Reliable Multi-Factor User Authentication with One Single Finger Swipe

Jianwei Liu, *Student Member, IEEE,* Kaiyan Cui, *Student Member, IEEE,* Xiang Zou, *Student Member, IEEE,* Jinsong Han, *Senior Member, IEEE,* Feng Lin, *Senior Member, IEEE,* and Kui Ren, *Fellow, IEEE*

***Abstract*—Multi-factor user authentication becomes increasingly popular due to its superior security comparing with single-factor user authentication. However, existing multi-factor user authentication methods usually require multiple interactions between users and different authentication components when inputting the multiple factors, leading to extra overhead and bad user experience. In this paper, we propose a secure and user-friendly multi-factor user authentication system named *BioDraw*. It utilizes four categories of biometrics (impedance, geometry, behavior, and composition) of human hand plus the pattern-based password to identify and authenticate users. User only needs to draw a pattern on a radio frequency identification tag array, while four biometrics can be collected simultaneously. Specifically, we first design a gradient-based pattern recognition algorithm to precisely extract user's secret pattern. Then, a convolutional neural network- and long short-term memory-based classifier is utilized for user recognition. Furthermore, to guarantee the systemic security, an anti-replay method called *Binary ALOHA* is proposed to detect replayed signals. We conduct extensive experiments with 30 volunteers. The experiment results show that *BioDraw* can achieve high authentication accuracy (with a 2% – false reject rate) and is effective in defending against various attacks.**

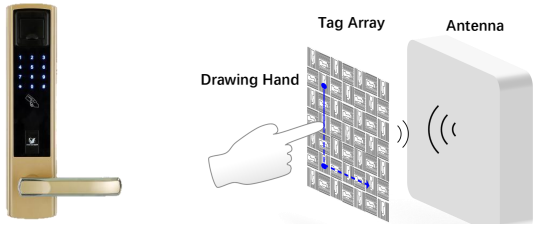
***Index Terms*—Radio Frequency Identification, Multi-factor User Authentication, Replay Defense.**

I. Introduction

USER authentication is pivotal to many security-critical applications, such as user identification, access control, and electronic transaction. Generally, user authentication utilizes specific equipment to collect user's secret or biometric information, including password, security token, fingerprint, *etc*. During this process, users mainly concern about two critical but usually contradictory points: security and user-friendliness.

Based on the number of authentication factors involved, existing user authentication approaches [1], [2], [3], [4], [5] can be divided into two main categories: single-factor user authentication and multi-factor user authentication (MFUA). Single-factor authentication approaches are regarded as user-friendly solutions because they only consider one factor and are less intrusive. However, conventional single-factor authentication approaches are vulnerable to attacks since the security of the whole system only depends on one factor [6]. For example, attackers can easily compromise a fingerprint verification system by forging legitimate user's fingerprint [7].

J. Liu, J. Han (corresponding author, email: hanjinsong@zju.edu.cn), and F. Lin are with Zhejiang University, China, and ZJU-Hangzhou Global Scientific and Technological Innovation Center, China. K. Cui is with Xi'an Jiaotong University, China, and the Department of Computing, Hong Kong Polytechnic University, China. X. Zou is with Xi'an Jiaotong University, China. K. Ren is with School of Cyber Science and Technology, Zhejiang University, China, and Zhejiang Provincial Key Laboratory of Blockchain and Cyberspace Governance, Hangzhou, China.



(a) PIN- and ID card-based lock. (b) Sketch of *BioDraw*.

Fig. 1. (a) shows a lock with multi-factor access control mechanism. (b) is the sketch of *BioDraw*.

To enhance the security of the single-factor approaches, MFUAs are proposed and they involve multiple factors to conduct joint authentication [6]. For example, the lock in Fig. 1(a) is a MFUA scheme applied in access control, which requires to verify user's PIN and ID card. RF-Mehndi [8] combines the physiological characteristics of the user's hand and the physical characteristics of RFID tag array to jointly verify user's identity. Nevertheless, most of the existing MFUA approaches simply 'stack' multiple factors to enhance the security, resulting in additional hardware overhead. More importantly, these approaches require frequent interactions of users, which reduces user-friendliness and increases the risk of information (e.g., biometrics) leakage [6]. Users often need to make trade-offs between the security and user-friendliness.

In this paper, we aim to address the aforementioned problems and enhance the security of user authentication by involving as many necessary factors as possible while simplifying the operation at a minimal level. In this way, we can guarantee high security and user-friendliness at the same time. For doing so, we still adopt the widely acceptable factor of "something they know": the lock/unlock pattern, as the basic factor in MFUA. Meanwhile, we attempt to collect four biometrics of user's hand, *i.e*., impedance, geometry, behavioral, and composition biometrics [8], [9], [10], as the other four factors of "something they have" and "something they are". Now, the problem becomes how to collect these four biometrics with only one single pattern input, instead of performing multiple interactive operations. Fortunately, the recent advance of sensing technologies inspires us to fuse the four biometric factors via wireless signals [8], [11], [12]. These technologies show that radio frequency (RF) signals have the ability of capturing human biometrics in both contact-based and non-

contact ways. Hence, as illustrated in Fig. 1(b), our goal is to allow users to draw a pattern on a radio frequency identification (RFID) tag array to capture the aforementioned four biometrics to achieve MFUA.

However, to achieve our goal is non-trivial due to the following challenges: 1) Severe coupling effect exists amongst adjacent tags in the tag array, deteriorating the accuracy of pattern recognition. 2) It is difficult to extract fine-grained biometrics from coarse-grained tags' signals. 3) RFID systems suffer from replay attacks [13], [8].

By addressing the above challenges, we propose a user-friendly RFID-based MFUA system, namely *BioDraw*. Specifically, we first adopt a perpendicular layout in the tag array to alleviate the impact of the coupling effect. Meanwhile, we design a gradient-based algorithm to accurately extract patterns from received signals. Then, we design a convolutional neural network and long short-term memory (CNN-LSTM) based classifier to extract fine-grained biometrics from coarse-grained signals to identify users. Finally, we design an anti-replay method called *Binary ALOHA*, which leverages the inherent randomness of the collision avoidance protocol (frame-based slot-ALOHA protocol [14], [15]) of RFID system to detect replayed signals without requiring extra hardware or modification on the existing RFID infrastructures. We conduct extensive experiments with 30 volunteers to evaluate *BioDraw*'s performance. The experiment results show that *BioDraw* is secure and user-friendly, with an average $2\%-$ false reject rate.

Our contributions can be summarized as follows:

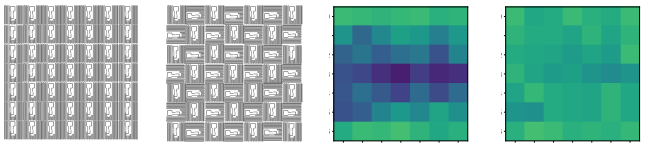
- We design a secure and user-friendly multi-factor user authentication system named *BioDraw*. Through one single pattern input, *BioDraw* captures the fusion biometric consisting of impedance, geometry, behavioral, and composition features of human hand as the authentication credential.
- We propose an easy-to-implement anti-replay method called *Binary ALOHA*. It exploits the inherent randomness of the frame-based slot-ALOHA protocol [14] to detect replay attacks.
- We conduct comprehensive real-world experiments with 30 volunteers. The experiment results demonstrate the high accuracy and security of *BioDraw*.

II. Preliminary

We start by introducing some background knowledge of RFID system, including the received signal strength (RSS) and phase information. Next, we describe the principle of the frame-based slot-ALOHA protocol [14] in RFID systems, which will be reused to design *Binary ALOHA*. At last, the layout of tags will be discussed to illustrate the impact of the coupling effect [8].

A. RFID System

A typical passive RFID system mainly consists of two parts: a reader equipped with an antenna and a passive tag. In the communication process, the reader first sends commands and continuous wave (CW) for supplying energy to tags. After receiving the commands, the tag will backscatter its electronic product code (EPC) to the reader. In addition to the EPC, the reader can also measures two RF parameters: RSS and phase. When the channel conditions (including the surrounding environment, tag's antenna state, and the distance between the reader antenna and tag) are static, the RSS and phase will remain stable. Once a human hand gets close to or touches the tag, the surrounding environment or the antenna state of the tag will change, affecting the tag's RSS and phase. Therefore, we can leverage RSS and phase to record the biological and behavioral characteristics of human hand.



(a) Parallel layout. (b) Perpendicular layout. (c) Detrimental RSS distribution. (d) Profitable RSS distribution.

Fig. 2. (a) is the parallel layout and (c) is its RSS distribution. (b) is the perpendicular layout and (d) is its RSS distribution. The deeper the color is, the lower the RSS is.

B. Randomness in Frame-Based Slot-ALOHA

Framed-based slot-ALOHA protocol [14] has been widely used for collision avoidance in industrial RFID systems. The core technique supporting this protocol is the random slot allocation. Specifically, in a multi-tag RFID session, each tag needs to select a slot in a frame (a frame contains multiple slots) to reply its EPC. A frame is broadcasted from the reader to all the readable tags at the beginning of each communication round. Afterwards, each tag reports its EPC to the reader via backscattered RF signals in its randomly-selected slot. In this way, most tags will be identified in their exclusive slots without collision. It can be found that randomness is the theme that helps the system avoid collision. In Section VII, we will present how we leverage this randomness to detect replay attacks.

C. Perpendicular Layout of Tags

To enable users to draw patterns via tags, we build a tag array named TagPad. To make users smoothly and seamlessly draw patterns as well as save space, we place two adjacent tags to be close to each other with a three-millimetre gap between two adjacent tags. However, the inductive coupling effect exists between two adjacent tags according to [8]. As shown in Fig. 2(a), if we make the orientations of all tags consistent when building a 7×7 TagPad, the RSS distribution shown in Fig. 2(c) will be generated. The coupling effect will weaken the RSS of some tags in the central part of the TagPad, or even make them unreadable. To solve this problem without enlarging the distance between adjacent tags, we resort to the method proposed in [16]. As shown in Fig. 2(b), we arrange each pair of adjacent tags perpendicularly according to their antennas' orientations. In this manner, *BioDraw* effectively alleviates the impact induced by the coupling effect (shown in Fig. 2(d)). Finally, since the circuit's layout of a tag's antenna is very dense, direct contact by a fingertip may short-circuit

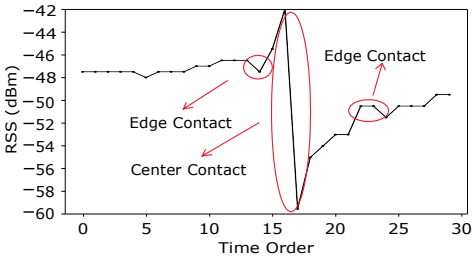


Fig. 3. Curve of the RSS variation when a fingertip swipes across a tag.

the tag and make it unreadable, we cover the TagPad with a layer of transparent plastic (normal adhesive tape) to avoid direct contact.

III. Biometrics in *BioDraw*

BioDraw utilizes four biometrics that can be obtained by TagPad for authentication: the impedance, geometry, behavioral, and composition biometrics of human hand. In this section, we detail how these biometrics are captured by RF signals and exploited as the factors in MFUA.

A. Biometrics Captured by RF Signals

In this part, we separately introduce these four biometrics and explain how they are quantified by the RSS and phase of RF signals.

Impedance biometric capture: The impedance of human body has been used for identifying individuals [8], so we can utilize it as a factor in MFUA. To explain how the impedance biometric is introduced by the contact of the fingertip, we first conduct a preliminary experiment, and then build a theoretical model to support the experiment result. In the experiment, a volunteer is asked to use his fingertip to swipe across a target tag in the TagPad. The RSS variation trace of the tag is shown in Fig. 3. It can be seen that the RSS values are stable before contact. When the fingertip starts to contact the edge of the tag, the RSS begins to increase. However, the RSS decreases rapidly when the fingertip contacts the central part of the tag. This 'decreasing phenomenon' can be theoretically explained. According to [8], a passive tag can be simplified as an equivalent circuit shown in the left part of Fig. 4. The impedance of the center chip and its surrounding antenna can be denoted by Z_A and Z_C respectively. If we denote the voltage produced by the harvested energy as V, the alternating current I in the tag's circuit can be formulated by:

$$I = \frac{V}{Z_A + Z_C}. \tag{1}$$

Recalling that our TagPad is covered by a layer of transparent plastic, such plastic can be regarded as a kind of dielectric. Particularly, when a fingertip contacts the tag's centre part, a capacitor will be constructed by three media: the tag's antenna, the transparent plastic, and the fingertip. This capacitor will influence the status of the original circuit, and a new equivalent circuit will be formed. The new equivalent circuit is displayed in the right part of Fig. 4, in which the capacitor C is formed by the aforementioned three media, and the impedance Z_F

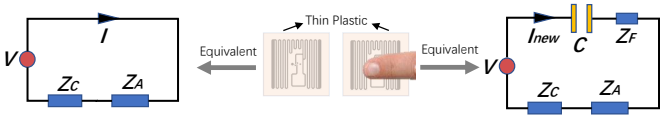


Fig. 4. Equivalent circuit of tag. The original equivalent circuit of the tag is shown in the left part. After touching the center part of the tag (shown in the right part), an extra capacitor and an extra impedance are introduced.

is introduced by the fingertip. In this case, the sum of all impedance in the new equivalent circuit is:

$$Z_S = Z_F + \frac{1}{j2\pi f C}, \tag{2}$$

where f is the frequency of the alternating current and j represents the square root of -1. Accordingly, the alternating current I_{new} in the new equivalent circuit can be formulated by:

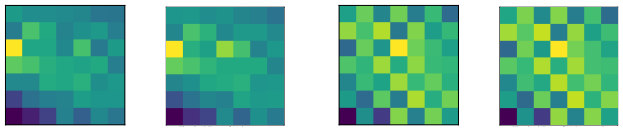
$$I_{new} = \frac{V}{Z_A + Z_C + Z_S}. \tag{3}$$

Due to the introduction of C and Z_S, I_{new} is smaller than I quantitatively. Therefore, the experiment results in Fig. 3 well demonstrate the validity of the right equivalent circuit in Fig. 4 and the corresponding theoretical model. It is noteworthy that the RSS and phase of backscattered signals are related to the alternating current in the tag's antenna [8]. Hence, for a specific tag, the contact of the fingertip enables the tag's RSS and phase to quantify the impedance biometric.

Geometry biometric capture: A human hand can be discerned by its geometry, which allows the geometry to be used as a biometric for identity authentication. For example, Song *et al.* [9] prove that the geometry of human hand is distinguishable for different people. In *BioDraw*, when drawing a pattern, the human hand acts as a reflector to reflect part of the in-air CW to tags, enabling the tags to harvest more energy. The extra energy will further influence the RSS and phase. This is also the reason why RSS increases when the fingertip swipes between the edge and the centre of the tag (in Fig. 3). Furthermore, different human hands, *i.e.*, different geometries, reflect varied extra energy. Therefore, RF signals backscattered by tags can capture the geometry biometric of human hand.

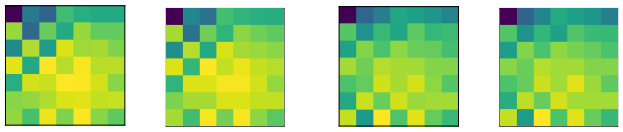
Behavioral biometric capture: In *BioDraw*, the behavioral biometric [17] of human hand stems from human's drawing habit. The research in [9] demonstrates that the drawing habit of human hand is person-distinguishable and can be utilized for user verification. When a user draws a pattern on the TagPad, his/er drawing speed and hand gesture are behavioral biometrics and can be recorded by the RSS and phase. This is because: 1) The drawing speed influences the RSS and phase variation speeds of the tags in the temporal domain. 2) Different hand gestures reflect different CW energy to the tags, resulting in different RSS and phase [18]. Thus, we can use RF signals to capture behavioural biometrics.

Composition biometric capture: The biological component of the surface of human hand is skin. Thus, in *BioDraw*, when CW reaches human hand, it will meet the interface between skin and air. At the interface, incident and reflection occur simultaneously. The power of the CW reflected to tags



(a) 1st distribution of volunteer 1. (b) 2nd distribution of volunteer 1. (c) 1st distribution of volunteer 2. (d) 2nd distribution of volunteer 2.

Fig. 5. RSS distributions caused by different impedance biometrics. Deeper color means larger RSS value.



(a) 1st distribution of volunteer 1. (b) 2nd distribution of volunteer 1. (c) 1st distribution of volunteer 2. (d) 2nd distribution of volunteer 2.

Fig. 6. RSS distributions of different geometry biometrics. Deeper color means larger RSS value.

(which is reflected by the RSS) is related to the skin's relative permittivity. Since skin's relative permittivity is different from other materials, skin can be regarded as a composition biometric. Specifically, according to [10], when RF signals meet the interface between skin and air, the ratio between the reflected power and incident power can be formulated by: $\frac{P_R}{P_I} = |\frac{\sigma_A - \sigma_S}{\sigma_A + \sigma_S}|^2$. P_R and P_I are the reflected power and the incident power, respectively. σ_A and σ_S are the relative permittivity of air and skin, respectively. When the power of the CW that reaches human hand is invariable, P_R only depends on σ_S. Hence, the signal power reflected by the skin is different from that reflected by other materials, which enables *BioDraw* to naturally distinguish between real human hands and counterfeited ones (*e.g.*, hand made by polyvinyl chloride), *i.e.*, to have the capability of liveness detection [19].

B. Validation Experiment

To validate that the four biometrics are indeed captured by the backscattered signals and can be utilized for authentication, we conduct four validation experiments.

Impedance biometric validation: In this experiment, we aim to validate that the impedance biometrics of different persons are distinguishable, while that of the same person is stable. We invite two volunteers and each volunteer is asked to contact the same tag in the TagPad twice, while the other tags are covered by tinfoil paper (avoiding the influence of geometry and behavioral biometrics). The RSS distributions of the TagPad are shown in Fig. 5, demonstrating that two distributions of the same volunteer are similar, but that of different volunteers are distinguishable. Thus, the impedance biometric is indeed embedded in the backscattered signals and can be used to identify individuals.

Geometry biometric validation: In this experiment, we try to visualize the geometry biometric also by RSS distribution. We ask two volunteers to place their right hands in the same position in front of the TagPad while collecting signals. As shown in Fig. 6, the RSS distributions of the same volunteer are similar, meanwhile, that of different volunteers show apparent distinguishability. Therefore, the geometry biometric is indeed captured by backscattered signals.

Behavioral biometric validation: In this experiment, we validate that the behavioral biometric, *i.e.*, the drawing speed and hand gesture, can be captured by the signals. Specifically, for the drawing speed, we employ a motor and a thin wire to conduct the experiment. One end of the thin wire is tied to the motor while the other end is tied to the fingertip. In this way, the drawing speed of the fingertip is determined by the rotation speed of the motor, which is further determined by the driving voltage. We use two different voltages to introduce two different drawing speeds V_1 and V_2. One volunteer is required to draw the pattern '1' respectively with the two speeds V_1 and V_2, where the only variable is the drawing speed, *i.e.*, the behavioral biometric. For each speed, we collect 50 signal samples (which are composed of RSS values and phase values). We then label the signal samples of two speeds by '0' and '1', respectively. By using 75% signal samples as the training set and 25% ones as the testing set, we obtain 100% classification accuracy on a neural network classifier, indicating that the drawing speed feature is contained in the backscattered signals. For the experiment towards the hand gesture, we ask one volunteer to draw the same pattern with two different gestures. Then, the same classification is performed and the accuracy, 100%, proves that the feature of hand gesture is also recorded by backscattered signals. Therefore, the behavioral biometric can also be used as an authentication factor in *BioDraw*.

Composition biometric validation: Different from the impedance, geometry, and behavioral biometrics that are used to distinguish identities, composition biometric is developed to distinguish skin from other materials. We thus validate that the signals reflected by the skin are distinguishable from that reflected by other materials in this experiment. One volunteer is asked to attach different materials (nothing, paper, and plastic glove) to his/er hand while drawing. For each material, we collect 50 signal samples. In this way, the only variable is the reflective material, *i.e.*, composition biometric. Then, we label the signal samples of skin as '0' and that of other materials as '1'. We also use a neural network to classify these signal samples. The 100% accuracy indicates that the composition biometric is indeed contained in the backscattered signals and can be utilized for liveness detection.

In *BioDraw*, we use the fusion biometric composed of the aforementioned four biometrics as the authentication credential, because if we verify only one biometric each time, the authentication process will bring multiple operations and weaken the user-friendliness of *BioDraw*.

IV. *BioDraw* Design

We design a reliable multi-factor user authentication system, *BioDraw*. This section first introduces the overview of *BioDraw*. Then, the modules in *BioDraw* are separately detailed.

A. BioDraw Overview

As illustrated in Fig. 7, the architecture of *BioDraw* can be divided into two phases: registration phase and authentication phase. The registration phase contains the *signal preprocessing* module and *identity recognition* module. The authentication

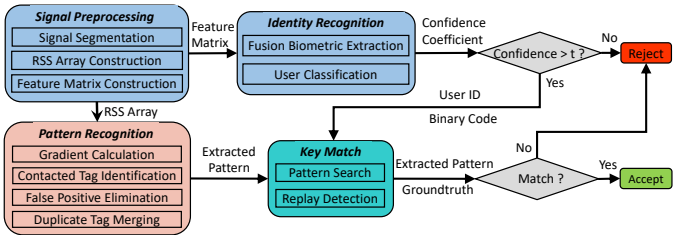


Fig. 7. Architecture of *BioDraw*.

phase uses not only the two modules in the registration phase, but also the *pattern recognition* module and *key match* module.

In the registration phase, users first need to save their patterns (termed as groundtruth patterns) that are selected according to their preferences in the database. Each pattern is stored in the form of EPC sequence (*e.g.*, $[EPC_a, EPC_b, \cdots, EPC_g]$). Then, each user needs to draw his/er pattern on the TagPad several times to collect a batch of data. This batch of data is first preprocessed by the *signal preprocessing* module and then used as a training set to train the classifier in the *identity recognition* module. Each individual input of the classifier is called a *feature matrix*. The registration phase is completed after the classifier is well trained using the *feature matrices* in the training set.

In the authentication phase, a user only needs to draw his/er pattern on the TagPad once to initiate an authentication request. Concurrently, the fusion biometric carried by backscattered signals is collected by the reader. The signals are first input into the *signal preprocessing* module to obtain a *feature matrix* and an *RSS array* (composed of the RSS values of all tags). After that, the *feature matrix* is input into the trained classifier in the *identity recognition* module. *BioDraw* can obtain a user ID (*i.e.*, a predicted class of the classifier [20]), a confidence coefficient, and a binary code in this module. As long as the confidence coefficient is smaller than a threshold, the authentication request will be rejected. Otherwise, the *pattern recognition* module leverages the *RSS array* to extract a pattern. *BioDraw* then tries to match the extracted pattern with the groundtruth pattern in the *key match* module. If it does not match, the authentication request will be rejected. Otherwise, *BioDraw* searches the database to confirm whether the binary code has been stored in the database or not. If the binary code is found to exist in the database, the authentication request will be regarded as a replayed one and rejected. Or, the authentication request will be accepted and the binary code will be stored in the database.

B. Module Introduction

Signal preprocessing: This module is responsible for preprocessing the received raw signals. It includes three steps: signal segmentation, constructing a *feature matrix* for the *identity recognition* module, and constructing an *RSS array* for the *pattern recognition* module. To segment the signal part that records the pattern drawing, we must first locate the start and end timestamps of the drawing process in the signal stream. We monitor the RSS values of each tag in the TagPad and use the first/last timestamp where the RSS drops drastically (the difference between two continuous values is larger than 5dbm) as the start/end timestamp of the drawing. Afterwards, We take the tag with the least number of reads (denoted as n) as the standard, and select n successive timestamps for each tag. Since each timestamp corresponds to an RSS value and a phase value, we obtain n RSS values and n phase values for each tag. To provide shape-consistent input to the CNN-LSTM-based classifier, we linearly interpolate the signals so that each tag has m (empirically set as 30) RSS values and m phase values. After aligning the RSS and phase values, we construct a *feature matrix* with dimensionality of $(2, 49, m)$ (a TagPad contains 49 tags in our default setting). This *feature matrix* is then sent to the *identity recognition* module to identify users. Next, *BioDraw* aligns the RSS values of each tag according to the reading order, and obtains an *RSS array* with dimensionality of $(49, n)$. The *RSS array* is then sent to the *pattern recognition* module to extract a pattern.

Identity recognition: In this module, a CNN-LSTM-based classifier is utilized to recognize users by taking as input the *feature matrix*. The outputs of this module are threefold: a user ID, the ID's confidence coefficient, and a binary code. If the confidence coefficient of the user ID is smaller than a threshold we set in advance, it means that the *feature matrix* belongs to an illegitimate user and the authentication request will be rejected. Only when the confidence coefficient is larger than the threshold, the user ID and binary code will be sent to the *key match* module for further verification. The details of our CNN-LSTM-based classifier are introduced in Section V.

Pattern recognition: This module aims to extract a pattern from the *RSS array*. *BioDraw* achieves this goal by our proposed gradient-based pattern recognition algorithm. This algorithm first constructs a *gradient array* using the *RSS array*. By comparing the gradients in the *gradient array* with a threshold, those tags that are contacted during drawing are identified. Then, the EPCs of these contacted tags form a specific pattern. The pattern is sent to the *key match* module to perform further matching. The details of the gradient-based pattern recognition algorithm are elaborated in Section VI.

Key match: This module is used to check if the extracted pattern is the same as the groundtruth pattern, and if the authentication request is a replayed one. It first uses the user ID obtained by the *identity recognition* module as a key to retrieve the groundtruth pattern from the database. Then, *BioDraw* compares the groundtruth pattern with the extracted one. If they are not identical, the authentication request is considered to be from an illegitimate user and rejected, or else *BioDraw* will compare the binary code with the ones stored in the database. Once we find that an identical binary code has been stored, the authentication request will be treated as a replayed one and rejected. Otherwise, the authentication request will be accepted and the binary code will be stored in the database. The method designed for binary code extraction, *i.e.*, *Binary ALOHA*, is detailed in Section VII.

V. CNN-LSTM-Based User Classifier

In this section, we leverage a CNN-LSTM-based classifier to effectively extract the fusion biometric from the *feature matrix* to achieve accurate user recognition. This is because:

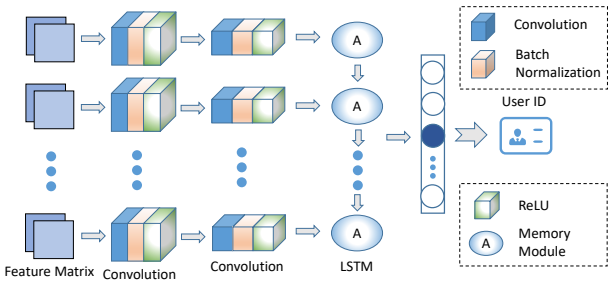
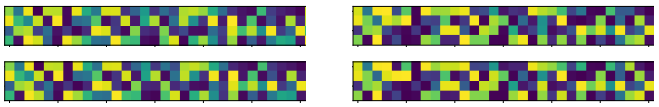


Fig. 8. Architecture of CNN-LSTM-based classifier.



(a) Two fusion biometrics of Volunteer 1.

(b) Two fusion biometrics of Volunteer 2.

Fig. 9. Fusion biometrics of two Volunteers 1 and 2.

1) The impedance, geometry, hand gesture, and composition biometrics are spatial biometrics, and CNN is effective in spatial feature extraction [21]. 2) The drawing speed is a temporal biometric, and LSTM is a qualified tool for temporal feature extraction. We will detail the spatial and temporal characteristics of the biometrics below.

Spatial and temporal biometrics: The TagPad used for biometric collection can be regarded as a two-dimensional plane with spatial structure. When a fingertip contacts a tag in the TagPad, although the contacted tag is mostly impacted, the slight coupling effect will make the entire TagPad impacted, turning the impedance into a spatial biometric. For the geometry, gesture, and composition, their impacts on the TagPad are like a hand being projected on the TagPad, so they are spatial biometrics as well. Therefore, using CNN to process the *feature matrix* can effectively extract the impedance, geometry, hand gesture, and composition biometrics. Second, in terms of the drawing speed feature extraction, *BioDraw* utilizes an LSTM component because it has been widely adopted to extract features in the temporal domain [21], [22], while the drawing speed is a temporal biometric intuitively.

Classifier architecture design: The architecture of our classifier is shown in Fig. 8. It mainly consists of two convolutional layers and an LSTM component. Each convolutional layer is followed by two functions: a batch normalization (BN) function and a rectified linear unit (ReLU). The BN is utilized to avoid entering the derivative saturation domain and prevent the data distribution from the offset. The ReLU is used to enhance the non-linearity between two layers of neurons and weaken the dependence between neurons. These two functions improve the complex task processing capability of our classifier and make the biometric extraction more effective. In respect to the convolution operation. we empirically set the size of each convolutional kernel as 2×2 with a sliding stride of 1×1. In the first convolutional layer, the channel number of input data changes from 2 (RSS+phase) to 16, and that of the second convolutional layer changes from 16 to 32. Behind the LSTM component, a fully connected layer is utilized to project extracted fusion biometric into the confidence coefficient of each user. In this process, the dimension of the output of the LSTM component changes from 128 to the number of registered users. The user ID with the largest confidence coefficient will be regarded as the user that initiates the authentication request. However, if the largest confidence coefficient is smaller than a threshold that is empirically set as 0.85, the authentication request will be considered to be initiated by an attacker (*e.g.*, an illegitimate user or a counterfeited hand) and rejected.

Effectiveness of fusion biometrics: To show the effectiveness of the fusion biometric extracted by our classifier, *i.e.*, the distinguishability of the fusion biometrics between different persons and the stability of the fusion biometric of the same person, we collect four *feature matrices* of two volunteers and extract their fusion biometrics using a trained classifier. To avoid the impact of pattern difference, they are asked to draw the same pattern 'l' while collecting *feature matrices*. Since the output of the LSTM component is regarded as the fusion biometric, we show the heatmaps (deeper colour represents larger value) of the four outputs in Fig.9. It can be observed that the fusion biometrics of different volunteers are distinguishable, meanwhile, that of the same volunteers are similar. Thus, our CNN-LSTM-based classifier performs well in fusion biometric extraction.

VI. Pattern Extraction

Since pattern is one of the crucial shields to protect *BioDraw* from attacks, precise pattern extraction is of vital importance to the security and user-friendliness of *BioDraw*. In this section, we first observe the trends of RSS variation of contacted tag and non-contacted tag. Then, we design a gradient-based pattern recognition algorithm according to our observations to extract pattern from *RSS array*.

A. RSS Variation Caused by Fingertip

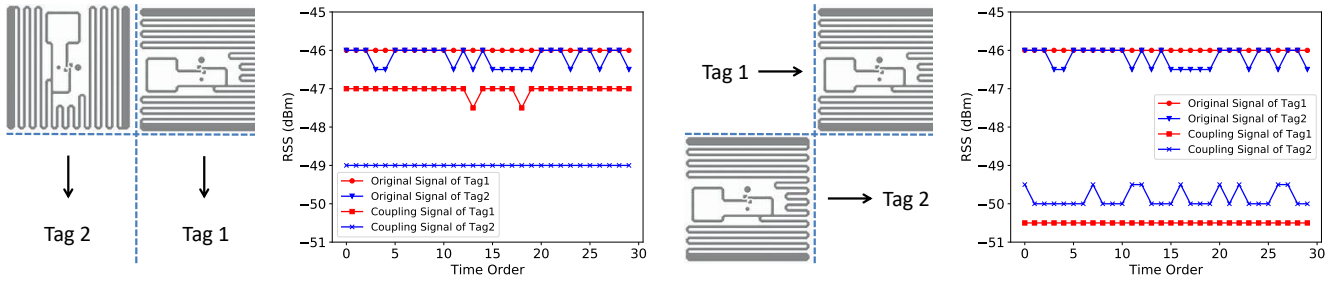
To extract pattern, two kinds of RSS relationships need to be figured out particularly: *R1*) The RSS relationship between contacted tag and non-contacted one; *R2*) The RSS relationship between prior contacted tag and posterior contacted tag. Leveraging *R1* and *R2*, we can distinguish contacted tags from non-contacted ones and determine the contact order of all contacted tags. A pattern is then uniquely determined by the EPCs of all contacted tags and their contact order.

***R1* analysis:** To analyze the RSS difference between contacted tag and non-contacted one, we conduct an experiment, in which a volunteer draws a pattern 'L' (shown in Fig. 10(a)) on the TagPad while recording the tags' RSS values. We take Tag 2 (contacted) and Tag 3 (non-contacted) as examples to explore *R1*. The RSS curves of these two tags are shown in Fig. 10(b). By comparing the two curves, we can get the following observations: 1) When the fingertip swipes across Tag 2, the RSS of both curves changes noticeably. 2) When the RSS decreases, the curve of the contacted tag is much steeper than that of the non-contacted tag. Therefore, we can define a metric that quantifies the rate of RSS decline to distinguish contacted tags from non-contacted ones.

***R2* analysis:** We use Tag 1 and Tag 2 in Fig. 10(a) as examples of prior contacted tag and posterior contacted one to analyze *R2*. The RSS curves of these two tags are shown in Fig. 10(c).

(a) Pattern 'L'. (b) RSS variation curves of Tag 2 and 3. (c) RSS variation curves of Tag 1 and 2.

Fig. 10. In (a), Tag 1 is the prior contacted tag, Tag 2 is the posterior contacted tag, and Tag 3 is the non-contacted tag. (b) and (c) show their RSS variation curves.



(a) Perpendicular layout. (b) RSS values of perpendicular tags. (c) Parallel layout. (d) RSS values of parallel tags.

Fig. 11. (a) and (c) are perpendicular layout and parallel layout, respectively. (b) and (d) are the RSS values corresponding to (a) and (c), respectively.

It can be observed that: 1) The drastic RSS decline of the prior contacted tag is similar to that of the posterior contacted tag. 2) The RSS of the prior contacted tag decreases earlier than that of the posterior contacted one.

Based on the two observations, we can sort the EPCs of all contacted tags according to the timestamp of RSS decline.

B. Gradient-based Pattern Recognition Algorithm

To precisely recognize the pattern, our algorithm needs to consider four consecutive operations: gradient calculation, contacted tag identification, false positive elimination, and duplicate tag merging.

Gradient calculation: Since *R1* demonstrates that the RSS decline rates of the contacted tag and the non-contacted one show a large difference, we define gradient G that can reflect the RSS decline rate to identify contacted tags. To be specific, We use three consecutive values of each tag in the *RSS array* to calculate a gradient, so that each tag will have $n-2$ gradients. The G can be calculated by:

$$G_i^k = |\frac{RSS_{i+2}^k - RSS_i^k}{t_{i+2}^k - t_i^k}|, \quad i \in [1, n-2], k \in [1, 49], \quad (4)$$

where G_i^k denotes the i_{th} gradient of the k_{th} tag in the TagPad, and $|t_{i+2}^k - t_i^k|$ is the corresponding normalized time interval. By using this formula, we calculate gradients for every tag in the TagPad and obtain a gradient array $([[G_1^1, G_1^2, \cdots, G_1^{49}], \cdots, [G_{n-2}^1, G_{n-2}^2, \cdots, G_{n-2}^{49}]])$ with dimensionality of $((n-2), 49)$.

Contacted tag identification: In this operation, we first find out all the tags that are likely contacted, which are termed as candidate tags. Then, all contacted tags are identified from candidate tags based on the timestamps. Specifically, we sequentially select those tags with gradients larger than a threshold GT (empirically set as 2) from each *gradient sub-array* $([G_i^1, G_i^2, \cdots, G_i^{49}])$ as candidate tags. If multiple tags in a *gradient sub-array* are selected as candidate tags, the tag with the earliest reading timestamp will be regarded as the contacted tag. In this way, we obtain no more than one contacted tag from each *gradient sub-array*.

False positive elimination: So far, it seems that we have obtained all indeed-contacted tags. However, due to the slight coupling effect, some tags that are not actually contacted are mistakenly treated as contacted ones, which are termed as false positives. To solve this problem, we dig the essence of the coupling effect. According to the theoretical analysis towards the coupling effect in [23], the coupling effect between parallel tags is much stronger than the one between perpendicular tags, and hence is easier to cause false positive. This is because two parallel tags have a much larger overlapped electromagnetic field than that of two perpendicular ones.

In order to verify the validity of our analysis, we conduct two validation experiments with two tags. In the first experiment, we individually collect the RSS from each tag. We then place them adjacently but with their orientations perpendicular to each other, as illustrated in Fig. 11(a). We plot the collected RSS of the two tags in Fig. 11(b). We find that in this case, the coupling effect is too weak to induce false positives. In the second experiment, as shown in Fig. 11(c), we place the two tags on the same diagonal but with parallel orientation. The collected RSS is shown in Fig. 11(d). It can be observed that the RSS difference induced by the coupling effect approximates 5dBm, which is highly possible to result in false positives. Inspired by these observations, we define a layout-based contacted tag selection principle to mitigate the impact of the coupling effect. Suppose that we have selected i contacted tags, according to this principle, *BioDraw* selects the $i+1_{th}$ tag from the horizontal/vertical direction of the i_{th} contacted tag. For example, in Fig. 11, if Tag 1 is the i_{th} contacted tag, then only Tag 2 will be selected as the $i+1_{th}$ contacted tag (instead of tag 3), even if the gradient of Tag

3 is larger than the GT. Thus, this principle can effectively minimize the probability of incurring false positives.

Duplicate tag merging: As aforementioned, our algorithm selects contacted tags from each *gradient sub-array* alternately. If the drawing speed is relatively slow, *i.e.*, the fingertip stays on a tag for a long time, the tag may be identified as contacted ones in multiple successive *gradient sub-arrays*, leading to the inclusion of some duplicate tags in the extracted pattern. To deal with this issue, *BioDraw* traverses the EPC sequence (*i.e.*, the extracted pattern) from the beginning and deletes the duplicate EPCs.

After the aforementioned four operations, we finally obtain a continuous EPC sequence that uniquely specifies a pattern. Note that in frame-based slot-ALOHA [14], the tags are read one by one, *i.e.*, the tag reading shows asynchronization. In this case, if the tags are read slowly in a frame, the EPC sequence of contacted tags may be disordered. Fortunately, as demonstrated by [24], *BioDraw* only takes ~ 0.25 seconds to read all tags in a frame. This time length is very short. Thus, such asynchronization hardly impacts our pattern extraction results.

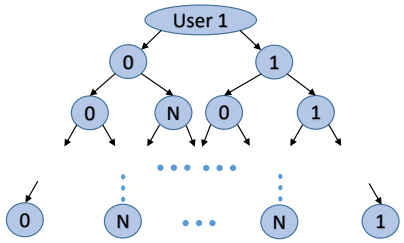


Fig. 12. Structure of bit tree.

VII. *Binary ALOHA*

According to the background knowledge introduced in Section II-B, off-the-shelf randomness is contained in the frame-based slot-ALOHA protocol [14]. In this section, we leverage this randomness to design an anti-replay method called *Binary ALOHA*.

Binary encoding scheme: In frame-based slot-ALOHA protocol [14], each tag randomly selects a slot in a given frame. In case a round of communication is finished, *i.e.*, each tag has been read once, the reader will reallocate a new frame. In the next round, each tag will re-select a new slot. Therefore, in each round, the order that the tags response to the reader is random as well. This randomness of response order allows us to devise a binary encoding scheme.

Specifically, we arbitrarily select two tags T_a (with EPC_a) and T_b (with EPC_b) in our TagPad as the encoding targets. In each round, the response order of T_a and T_b is uncertain and can be represented by their EPC sequence. An optional encoding scheme is that the order $[EPC_a, EPC_b]$ is encoded as '0' and the order $[EPC_b, EPC_a]$ is encoded as '1'. By adopting this scheme, a '0' or a '1' can be randomly obtained as a bit of a binary code in each round. Although only one bit can be obtained in a round, the encoding space, *i.e.*, the number of bits, will be enlarged infinitely when the randomness accumulates with the increase of the communication rounds,

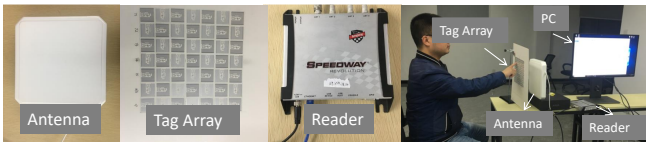


Fig. 13. Experiment setup of *BioDraw*.

In our default setting, we use 30 rounds to construct a binary code with 30 bits.

Theoretical feasibility: Since the length of each binary code is 30 bits, the encoding space of our encoding scheme is 2^{30} (larger than 10^9). Statistically, if the encoding space and the number of authentication requests are denoted as S and T, the probability that two identical binary codes will not appear in T authentication requests is:

$$P_{(S,T)} = \frac{(S-1)(S-2)\cdots(S-T+1)}{S^{T-1}}. \tag{5}$$

When we respectively set S and T as 10^9 and 10^4, the probability $P_{(10^9,10^4)}$ is larger than 95%. It means that in 10^4 authentication requests, the probability that each extracted binary code is unique exceeds 95%. Once we extract a repeated binary code in an authentication request, there is a 95%+ confidence that the authentication request is initiated by a replay attacker. Therefore, we regard all the authentication requests corresponding to repeated binary codes as replay attacks. *Binary ALOHA* is able to defend against replay attacks theoretically. Further, *Binary ALOHA* can be easily extended so as to have larger encoding space. For example, we can involve three tags in the extended encoding scheme. Then, our replay detection method can be called *Ternary ALOHA*. From *Binary ALOHA* to *Ternary ALOHA*, the encoding space is enlarged from $2^{30}{>}10^9$ to $3^{30}{>}10^{14}$. Apparently, *Ternary ALOHA* can protect more authentication requests from replay attacks than *Binary ALOHA*. However, the computational overhead will increase accordingly. Thus, users need to make trade-offs between security and computing-friendliness.

Reducing overhead: Since *Binary ALOHA* needs to store each user's used binary codes for comparing with the binary codes extracted from later authentication requests, *Binary ALOHA* would cause some storage overhead and time cost. To reduce the overhead and cost, we design a bit tree (the basic structure of which is a binomial tree) to store binary codes for each user. Specifically, each user has a bit tree to store his/er used binary codes. As shown in Fig. 12, this bit tree belongs to 'user 1' The 'N' in this tree means 'null', *i.e.*, this binary code has not been used by 'user 1'. When we use the bit tree to store binary codes, the storage overhead will be greatly reduced, because the same segment in different binary codes only needs to be stored once. For instance, to store the binary codes '$1011\cdots00$' and '$1010\cdots10$' without adopting the bit tree, the same segment '101' needs to be stored twice. After adopting the bit tree, '101' only needs to be stored once. Moreover, the time complexity (*i.e.*, time cost) will decrease from $\mathcal{O}(p)$ to $\mathcal{O}(q)$, where p and q are the numbers of binary codes and the length of the binary code, respectively. Apparently, p will increase constantly with the use of *BioDraw*, yet q is invariable and small.

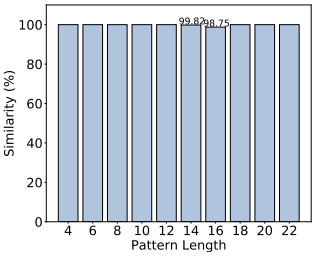


Fig. 14. Similarity between extracted pattern and groundtruth.

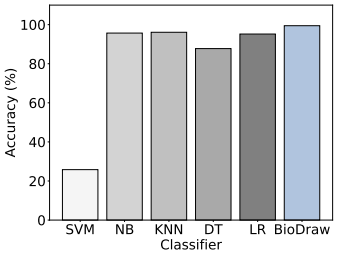


Fig. 15. Recognition accuracy of six classifiers.

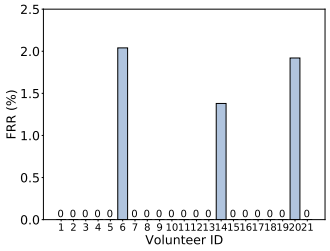


Fig. 16. FRR of 21 volunteers.

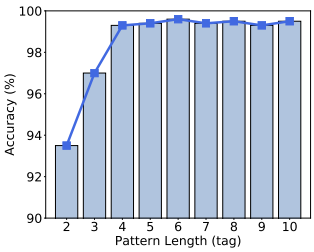


Fig. 17. Recognition accuracy when users use the same pattern.

VIII. Evaluation

We built a prototype of *BioDraw* and conducted comprehensive experiments to evaluate *BioDraw* quantitatively.

Implementation: Our experiments were conducted in two environments: a typical laboratory environment and an office environment. The default experiment setup (in laboratory) is illustrated in Fig. 13. The hardware of *BioDraw* is mainly composed of four parts: a personal computer with 2.8GHz CPU, an *Impinj R420* reader, a *Larid A9028* one-dimensional antenna, and a tag array containing 49 *Alien-9629* tags. The one-dimensional antenna was connected to the reader and placed 15cm away from the tag array in parallel. Users can draw patterns on the surface of the tag array comfortably. We adopted the EPC C1G2 standard [14] and frame-based slot ALOHA protocol [14] for reading process controlling and collision avoidance, respectively. All the experiments were conducted by adhering to the approval of our university's Institutional Review Board (IRB).

Data collection: We invited 30 volunteers (10 females and 20 males) aged from 8 to 44 to participate in the experiments. The weights of them varied from 36kg to 90kg and the heights of them varied from 143cm to 188cm. In the laboratory/office environment, 21/15 volunteers were asked to input their patterns and at least 50 times for each volunteer. In the default setting, we used 75% feature matrices for classifier training and the rest 25% for testing.

Metrics: We define three metrics to characterize the performance of *BioDraw*: *similarity*, *accuracy*, and *false reject rate (FRR)*. The *similarity* measures the performance of the gradient-based pattern recognition algorithm. It can be calculated by: $Similarity = \frac{L_O}{L_W}$, where L_O is the length (the number) of the tags of the overlapped fraction of the groundtruth pattern and the extracted pattern, and L_W is the length of the longest pattern between these two patterns. The *accuracy* is the probability that the user is correctly recognized. It can be calculated by: $Accuracy = \frac{N_{cor}}{N_{all}}$, where N_{cor} is the number of correctly classified *feature matrices* and N_{all} is the number of all testing *feature matrices*. The FRR is the probability that a legitimate user is falsely rejected. The FRR^i, *i.e.*, the *FRR* of user i can be calculated by: $FRR^i = \frac{N_M^i}{N_A^i}$, where N_M^i represents the number of the mistakenly rejected authentication requests of user i and N_A^i is the number of all authentication requests of user i.

A. Overall Performance

Performance of the gradient-based pattern recognition algorithm: In this experiment, we asked the volunteers to draw 10 patterns with different lengths from 4 tags to 22 tags. For each length, each volunteer was required to draw it 20 times. The final *similarity* for a specific length is calculated by averaging the *similarity* of all volunteers. The results are shown in Fig. 14. It can be found that in most of the lengths, the *similarities* are 100%. When the length is 16 tags, the *similarity* is 98.75%, indicating that *BioDraw* extracted several wrong patterns. This 'imperfect' *similarity* may be caused by accidental and incorrect drawing operations. Therefore, the overall similarity of 99.86% indicates that our gradient-based pattern recognition algorithm can accurately extract patterns.

Performance of the CNN-LSTM-based classifier: In this experiment, we first let each volunteer select a pattern according to his/er preference, and then each volunteer drew his/er pattern in the laboratory environment. As a result, the overall recognition accuracy is larger than 99%, which demonstrates that our classifier can effectively extract the fusion biometrics of volunteers and recognizes volunteers accurately. To explore if electromagnetic interference would impact the classification performance, we carried out an experiment in the office environment with WiFi signals because WiFi signals are one of the most common electromagnetic signals in our daily life. In this experiment, a WiFi router was placed 30cm away from the TagPad while a smartphone connected to the router was playing online videos. The classification results are shown in Fig. 15. We compare the CNN-LSTM used by *BioDraw* with other five classic classifiers including support vector machine (SVM), naive Bayes (NB), K-nearest neighbours, decision tree, and logistic regression (LR). It can be found that the accuracy of *BioDraw* is larger than 99% and it outperforms other five classifiers. Thus, *BioDraw* is robust against WiFi signal interference. Furthermore, we show the FRR^i of each volunteer caused by low confidence coefficient in Fig. 16. It can be seen that all the $FRRs^i$ are smaller than 2.5%, and the majority of them are 0.0%. The low $FRRs^i$ indicate that *BioDraw* can provide users with good experiences.

It is noteworthy that under the worst condition, all the users will choose the same pattern. Therefore, the recognition accuracy under this condition should be measured. We then designed additional nine groups of experiments that all the volunteers shared the same pattern in the same group. The lengths of patterns in different groups are different. In group one, the length is two tags and the length of each latter

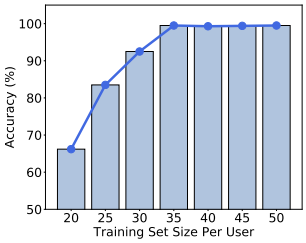


Fig. 18. Effect of training set size.

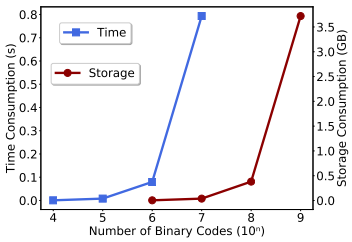


Fig. 19. Storage and time consumption of *Binary ALOHA*.

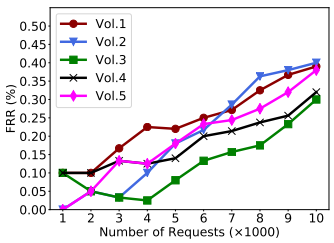


Fig. 20. FRRs of five volunteers caused by *Binary ALOHA*.

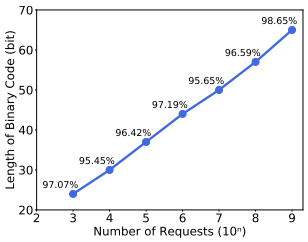


Fig. 21. Required number of bits for maintaining high security (probability).

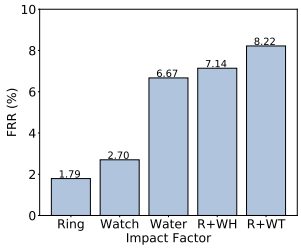


Fig. 22. Impacts of ring, water, and watch.

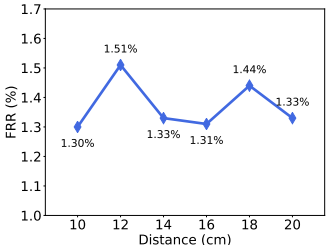


Fig. 23. Effect of the distance between antenna and TagPad.

group is one tag longer than that of its previous group. The experiment results are shown in Fig. 17, which indicate that a recognition accuracy as high as 90% can be achieved by using only two tags. When the length is larger or equals four tags, the recognition accuracy increases to 99%+. Thus, users are recommended to use patterns having at least four tags,

To explore the impact of the size of the training set on the recognition accuracy, we collected extra *feature matrices* and varied the size of the training set from 20 to 50 in steps of 5. The experiment results are shown in Fig. 17. It can be seen that by training the classifier with only 35 *feature matrices* per user, the recognition accuracy can exceed 99%.

Since the storage overhead is related to the user-friendliness of *BioDraw*, we also assess the space required for the storage of our CNN-LSTM-based classifier. We find that saving the whole classifier as *.pt* file [25] consumes about 7.67MB space. The small storage overhead indicates that our classifier is user-friendly in terms of storage.

Performance of *Binary ALOHA*: We evaluated *Binary ALOHA* from the perspectives of time cost, storage overhead, *FRR*, and scalability. We first assess the time cost and storage overhead without adopting the bit tree. To assess the time cost, we varied the number of stored binary codes from 10^4 to 10^7 and recorded the time cost of comparison between a new binary code and the stored ones. The experiment results are shown in Fig. 19. When the number of stored binary codes is 10^4, the time cost approximates 0.01 seconds, which is extremely low. When the number of stored binary codes increases to 10^7, the time cost is smaller than one second. The short time cost demonstrates that *Binary ALOHA* has outstanding real-time performance. We also plot the storage overhead in Fig. 19. The results indicate that only 3.72GB is needed for storing 10^9 binary codes. When we adopted the bit tree to store binary codes, the time cost approximates 0.01s and it will not increase. Meanwhile, the storage overhead

TABLE I
COMPARING *BioDraw* WITH OTHER THREE AUTHENTICATION SYSTEMS. D.F., R.A.R, AND M.F. REPRESENT DEVICE-FREE, REPLAY ATTACK-RESILIENT, AND MULTI-FACTOR, RESPECTIVELY.

System	Basics	FRR $\leq$ 5%	D.F.	R.A.R.	M.f.
BioDraw	**RFID**	✓	✓	✓	✓
RF-Mehndi	**RFID**	✓	×	×	✓
WiFi-ID	**WiFi**	×	✓	×	×
WiWho	**WiFi**	×	✓	×	×

approximately drops to one-twelfth of the original one (*i.e.*, without adopting the bit tree) when we store 10^9 binary codes. Thus, our bit tree can effectively reduce the time cost and storage overhead.

According to our theoretical analysis of *Binary ALOHA*, the *FRR* caused by *Binary ALOHA* is very low. To verify this conclusion, we invited five volunteers and asked each volunteer to collect 10^4 binary codes. Thereafter, we calculated the FRR^i of each volunteer and show it in Fig. 20. It can be observed that when the length of the binary code is 30 bits, the FRR^i is basically increasing linearly with the increase of the number of binary codes. Meanwhile, the maximal FRR in 10^4 requests is 0.4%, which is significantly low. Therefore, the FRR caused by *Binary ALOHA* is acceptable.

To assess the scalability of *Binary ALOHA*, We calculated the number of bits that can make the probability in Eq. 5 greater than 95% when the number of authentication requests is fixed. The experiment results shown in Fig. 21 indicate that the number of bits is increasing linearly with the increase of the order of magnitude of the request number. The number of bits is as small as 65 when the number of requests is 10^9. Hence, *Binary ALOHA* has outstanding scalability.

B. Related Factors

Impact of the ring, water and watch: In this part, we first consider that some daily accessories (*i.e.*, ring and watch) may impact *BioDraw*'s performance. Besides, since the water on the fingertip may also cause extra *FRR*, we also explore the impact of water. The experiment results are shown in Fig. 22, in which $R+WH$ means the ring and watch, and $R+WT$ means the ring and water. The results demonstrate that ring, watch, and water indeed induces extra *FRRs*. We find that the ring causes the smallest *FRR* (1.79%) and $R+WT$ causes the largest *FRR* (8.22%). The largest FRR is still small.

Effect of distance: We explored the impact of the distance between the reader's antenna and TagPad. We first varied the distance from 10 centimeters to 20 centimeters with a stride

of two centimeters. Then we show the experiment results in Fig. 23. It can be found that the FRR induced by different distances are all less than 2%, *i.e.*., the distance (within 20cm) has negligible effect on *BioDraw*. This is reasonable due to that the tag can be read even several meters away from the reader's antenna [14]. So, when the distance is within 20 centimeters, the RF signals still can effectively capture biometrics. Users can take advantage of this result in real-world deployment, *i.e.*, adjusting such distance flexibly to meet their needs.

Effect of tag array volume: We consider that the volume of the TagPad may affect the performance of *BioDraw*. We varied the volume from 7×7 to 3×3. The experiment results shown in Fig. 24 indicate that the *FRR* of 3×3 is only 6.32%. This result demonstrates that *BioDraw* is flexible on the TagPad volume selection.

Latency: The latency is the time cost for processing an authentication request. The time cost of *BioDraw* mainly comes from three components: drawing a pattern, user recognition with the classifier, and replay detection. Drawing a pattern may take several seconds. Using the CNN-LSTM-based classifier to process a *feature matrix* consumes 0.0003 seconds on average. To detect replay attacks, *Binary ALOHA* needs to consume 0.01 seconds. It can be found that the main time cost comes from pattern input. BioDraw is able to process an authentication request in a timely manner.

C. Comparison with Existing Works

To show the advantages of *BioDraw*, we compare *BioDraw* with three existing RF signal-based authentication systems: RF-Mehndi [8], WiFi-ID [26], and WiWho [11]. The comparison results are shown in Tab. 1. In terms of the *FRR*, *BioDraw* and RF-Mehndi outperform WiFi-ID and WiWho. However, RF-Mehndi is device-need while the other three systems are device-free. For the security, *BioDraw* can defend against replay attacks. while the other systems suffer from replay attacks. As for the involved factors, *BioDraw* and RF-Mehndi are MFUAs yet WiFi-ID and WiWho only utilize one factor. Therefore, *BioDraw* outperforms the other three authentication systems.

IX. Security Analysis

We specifically evaluate the security of *BioDraw* in this section. The metric used for security assessment is *the defence success rate* (DSR). If we let N_{def} denote the number of the attacks that be rejected by *BioDraw* and N_{att} denote the number of all attacks, DSR can be represented by: $DSR = \frac{N_{def}}{N_{att}}$.

A. Threat Model

In addition to the replay attack model, we also considered four other threat models: zero-effort attack, pattern-behavior-aware attack, pattern-geometry-behavior-composition-aware (PGBC) attack, and pattern-impedance-geometry-behavior-aware (PIGB) attack.

Zero-effort attack: In this attack, we assume that the attacker neither knows the principle of *BioDraw* nor the victim's pattern. The attacker randomly selects a pattern to attempt to be recognized as the victim by *BioDraw*. While drawing, the attacker uses his/er own natural drawing habit.

Pattern-behavior-aware attack: We assume that the attacker has observed the victim's authentication process. Thus, the attacker can mimic the victim's drawing habit and input the correct pattern through the TagPad.

PGBC attack: This attack is a multi-biometric-aware attack. We assume that the attacker knows the correct pattern and can counterfeit correct geometry, behavioral, and composition biometrics of the victim to trick *BioDraw*.

PIGB attack: This attack is a multi-biometric-aware attack as well. The attacker not only knows the correct pattern, but also can exhibit correct impedance, geometry, and behavioral biometrics of the victim to *BioDraw* when launching attacks.

Replay attack: In this attack, we assume that the attacker first eavesdrops the physical signals of a legitimate authentication process in a digital form when the victim is using *BioDraw*. Then, the attacker replays the exactly same signals towards the reader to fool *BioDraw*.

B. Defensive Ability Analysis and Experiment

Defense analysis: In *BioDraw*, the pattern is one of the important shields to protect the system from attacks. If the attacker attempts to attack *BioDraw* by imitating the victim, he must guess the victim's pattern. However, due to the large selection space of the pattern, it is difficult for an attacker to guess the correct pattern. Moreover, the fusion biometric is composed of four different biometrics, and all of them are simultaneously utilized to specify a person. Although behavioral biometric is easy to be mimicked, the geometry and impedance biometrics are hard to be filched. Meanwhile, the composition biometric would help *BioDraw* reject counterfeited hands. Even if more than three factors are forged by the attacker, the remaining factor would prohibit the attack. Furthermore, *Binary ALOHA* can protect *BioDraw* from replay attacks. Therefore, *BioDraw* is able to defend against all the aforementioned attacks.

Case study: Before doing security experiments, we surveyed users' view points on different authentication methods in terms of security. We totally surveyed 43 volunteers including the 21 volunteers who participate in the experiments in Section VIII. We asked them to answer three questions to complete the case study. In the first question, we investigate how they are content about five authentication methods (PIN, ID card, fingerprint, facial feature, and wireless signal-based biometric). We set a score bar from 1 to 10 to let them estimate the security quantitatively. As a result, the ID card based authentication and the wireless signal based authentication get the smallest and the largest mean scores, respectively. Hence, from the volunteers' points of views, wireless signal based authentication is most secure. In the second question, we let volunteers make a most-secure choice among five authentication methods including PIN, ID card, fingerprint, facial feature, and *BioDraw*. The results show that more than half of the volunteers choose *BioDraw* as the most secure one. These results demonstrate the high security of *BioDraw* from users' view points.

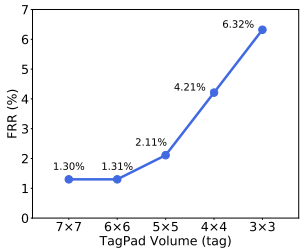


Fig. 24. Effect of TagPad's volume.

Fig. 25. DSR of 10×45 times of zero-effort attacks.

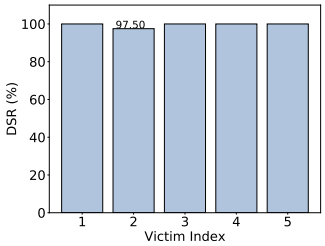


Fig. 26. DSR of 5×40 times of pattern-behavior-aware attacks.

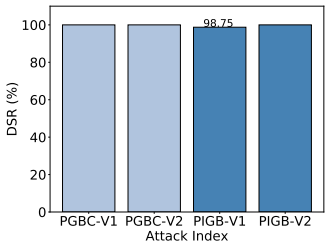


Fig. 27. DSR of PGBC attacks and PIGB attacks.

Zero-effort attack experiment: In this experiment, we first invited 10 volunteers to select 10 patterns according to their preferences. Then, nine volunteers (attackers) were asked to guess the pattern of the remaining volunteer (victim). Each attacker had five chances. The experiment results in Fig. 25 show that no attacker can correctly speculate the victim's pattern because the pattern selection space is significantly large.

Pattern-behavior-aware attack experiment: To launch this attack, we invited five pairs of volunteers as (attacker, victim). In each pair, the attacker first learnt the pattern and the drawing habit of the victim. Then, the attacker attempted to authenticate with *BioDraw* 40 times. The experiment results are shown in Fig. 26. It can be found that *BioDraw* is capable to defend against this attack effectively.

PGBC attack experiment: To counterfeit the geometry, behavioral, and composition biometrics, we asked two legitimate volunteers to attach paper to their fingertips to draw the correct pattern 40 times. The results in Fig. 27 indicate that impedance biometric can prevent *BioDraw* from PGBC attacks.

PIGB attack experiment: In this attack, we asked one out of two legitimate volunteers to wear a thin plastic glove while drawing. The fingertip part of the glove was broken to make the impedance biometric captured. The experiment results in Fig. 27 show that *BioDraw* is capable to prohibit PIGB attacks.

Replay attack experiment: We invited 10 volunteers to replay their previously used signals to *BioDraw*. As a result, all replayed signals are rejected.

X. Related Work

Single-factor user authentication: This kind of authentication only considers one factor during authentication, such as PIN [1], fingerprint [7], and iris [27]. This kind of authentication is convenient and user-friendly. The most commonly used authentication factors in the early days were knowledge and possessions. For example, PIN and pattern [3], [1] are knowledge factors, and ID card [2] is one of the ubiquitous authentication possessions. With the advances of smart sensors, biometrics are regarded as key authentication factors because they can reflect the unique internal characteristics of humans. For example, Li *et al.* [28] propose a fuzzy extractor to facilitate fingerprint-based authentication. Xie *et al.* [29] utilize the finger vein plus deep learning for user authentication. Ye *et al.* [30] utilize electrocardiogram to accurately classify the heartbeat. Although the simple operation of single-factor authentication in factor collection makes the authentication process user-friendly, the easy acquisition and forgery of single-factor also make its security relatively fragile.

Multi-factor user authentication: Compared with single-factor user authentication, multi-factor user authentication is more secure because it considers a combination of multiple factors. For instance, a finger vein and iris image can be used for authentication to improve system security [31]. An advanced design for multiple factors acquisition is collecting all the required certificates through only one operation. Several works have been proposed to approximate such a goal. Zhao *et al.* [8] propose to fuse two factors, *i.e.*, user's impedance and tag' physical feature, to achieve user identification. Chen *et al.* [32] leverage the geometry and palm-print of human hand to set up a bi-model for authentication. Song *et al.* [9] use the geometry biometric and behavioral biometric of user's hand to secure smartphone login. Nevertheless, existing MFUA has the following drawbacks. First, the number of involved factors is limited, usually only two types of factors. Second, the multiple factors acquisition relies on extra overhead, either from factor-collecting hardware or from human operations for inputting the factors. Therefore, the motivation of *BioDraw* is to involve as many factors as possible to improve the authentication security, while not introducing extra hardware or interaction.

Interaction in MFUA: To input multiple factors into the authentication system, a primitive manner is to let user exhibit one factor in each interaction and complete the authentication through multiple interactions. For example, in Fig. 1, user first needs to show his/er ID card, and then enter the PIN. In [27], the authentication system requires to collect user's fingerprint, iris, and voice features via multiple sensors such as camera and microphone. To reduce the interaction frequency for factor collection, RF-Mehndi [8] only requires user to hold his/er ID card in front of the reader's antenna, while both impedance feature and hardware feature are collected at the same time. Song *et al.* [9] propose to simultaneously collect geometry and behavioral biometrics of hand when user's hand is moving on the phone's screen. Different from previous works, *BioDraw* only asks user to draw a pattern on the TagPad. Four biometrics and the secret pattern are simultaneously captured in this simple interaction.

Replay attack resilience technique: Replay attack is one of the most intractable threats to wireless systems. Hu-Fu [13] is a replay-attack-resilient tag authentication technique achieved by introducing random noise into raw signals. Oo *et al.* [33] utilize phase features to prohibit replay attacks in speaker recognition/verification systems. Ye *et al.* [34] design a stochastic coding scheme to protect their system from replay attacks. The survey in [35] investigates some replay attack detection methods adopted by automatic speaker verification

systems. Different from the above works, *Binary ALOHA* can protect RFID systems from replay attacks without introducing extra randomness. This would avoid the destruction of biometrics caused by introducing random noise (*e.g.*, Hu-Fu [13]).

XI. Limitation and Future Work

Possible backdoor of *Binary ALOHA*: As aforementioned, we achieve replay detection by comparing binary codes extracted from tags' reading order. In frame-based slot ALOHA [14], each tag is read in an individual slot and multiple slots are received by the reader sequentially. So, if an attacker is able to change the order of slots in eavesdropped signals, the tags' reading order can be changed. Further, the binary code can be manipulated and the altered eavesdropped signal may bypass *BioDraw*'s replay detection. However, it is difficult for the attacker to attack successfully, because the attacker must segment the signal stream into slots precisely. Otherwise, the signal stream after segmentation would be un-decodable, causing the attack to fail. On the other hand, altering the order of the tags may cause the extracted pattern to be different from the groundtruth, which also renders failed attack. Therefore, it is possible to attack *BioDraw* in this way, but it is extremely hard to attack successfully. We will study this backdoor in future work.

Overhead caused by binary code: In *Binary ALOHA*, we need to perform binary code comparison/search and store all previously-used binary codes. This would cause a large amount of time and storage overhead. We thereby design bit tree to reduce such overhead. In fact, there is a potential to reduce the overhead at a minimal level. The potential solution is, instead of leveraging the randomness of the frame-based slot-ALOHA [14], to specify the reading order (pre-set reading order) of tags when sending reading commands, meanwhile, randomly generating a reading order in each authentication request. In this way, the reading order of a legitimate user is consistent with the pre-set one. But the reading order of the replayed signal (eavesdropped in previous authentication request) differs from the pre-set one, because the pre-set one has been changed. More importantly, we do not need to store previously-used reading orders. We will explore the feasibility of this method in future work.

XII. Conclusion

In this work, we propose a multi-factor user authentication system named *BioDraw*. Our main innovation points are twofold. First, we design *BioDraw* to collect as many authentication factors as possible in a simple interaction. Particular, five factors including pattern, impedance, geometry, behavioral, and composition biometrics are verified simultaneously. By our means, the security and user-friendliness are well integrated. On the other hand, we propose *Binary ALOHA* to overcome the dilemma that the existing anti-replay methods for RFID cannot be applied to sensing tasks. *Binary ALOHA* can be easily implemented on any multi-tag RFID systems by exploiting the off-the-shelf randomness of the collision avoidance protocol without any modification on software or hardware. Our real-world experiment results demonstrate the outstanding authentication performance and high security of *BioDraw*.

Acknowledgement

This work is supported in part by National Key R&D Program of China (2021QY0703), National Natural Science Foundation of China under grant U21A20462, 61872285, 62032021, 61772236, and 61972348, Research Institute of Cyberspace Governance in Zhejiang University, Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Grant No. 2018R01005), Ant Group Funding No.Z51202000234, and Alibaba-Zhejiang University Joint Institute of Frontier Technologies.

References

[1] X. Bultel, J. Dreier, M. Giraud, M. Izaute, T. Kheyrkhah, P. Lafourcade, D. Lakhzoum, V. Marlin, and L. Moták, "Security analysis and psychological study of authentication methods with PIN codes," in *Proceedings of the IEEE Conference on Research Challenges in Information Science (RCIS)*, 2018.

[2] T. H. Nam and V. D. H. Quan, "Multi-dimensional analysis of perceived risk on credit card adoption," in *Proceedings of the Econometric Conference of Vietnam (ECONVN)*, 2019.

[3] P. Andriotis, G. C. Oikonomou, A. Mylonas, and T. Tryfonas, "A study on usability and security features of the android pattern lock screen," *International Journal of Information and Computer Security*, vol. 24, no. 1, pp. 53–72, 2016.

[4] O. Ogbanufe and D. J. Kim, "Comparing fingerprint-based biometrics authentication versus traditional authentication methods for e-payment," *Decision Support Systems*, vol. 106, pp. 1–14, 2018.

[5] M. Komatsu and T. Akakura, "A facial authentication method robust to postural changes in e-testing," in *Human Interface and the Management of Information (HIMI)*, 2019.

[6] S. Das, B. Wang, and L. J. Camp, "MFA is a waste of time! understanding negative connotation towards MFA applications via user generated content," *CoRR*, vol. abs/1908.05902, 2019.

[7] "Iphone fingerprint sensor hacked with a finger made of clay at mwc," http://www.techworm.net/2016/02/, 2016.

[8] C. Zhao, Z. Li, T. Liu, H. Ding, J. Han, W. Xi, and R. Gui, "Rf-mehndi: A fingertip profiled RF identifier," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, 2019.

[9] Y. Song, Z. Cai, and Z. Zhang, "Multi-touch authentication using hand geometry and behavioral information," in *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2017.

[10] D. Vasisht, G. Zhang, O. Abari, H. Lu, J. Flanz, and D. Katabi, "In-body backscatter communication and localization," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, 2018.

[11] Y. Zeng, P. H. Pathak, and P. Mohapatra, "Wiwho: Wifi-based person identification in smart spaces," in *Proceedings of the ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2016.

[12] T. Xin, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "Freesense: Indoor human identification with wi-fi signals," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, 2016.

[13] G. Wang, H. Cai, C. Qian, J. Han, X. Li, H. Ding, and J. Zhao, "Towards replay-resilient RFID authentication," in *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2018.

[14] X. Lei and L. Sanglu, *Principle, Protocol and System Design of RFID*. Science Press, 2016.

[15] Z. Huang, R. Xu, C. Chu, Z. Li, Y. Qiu, J. Li, Y. Ma, and G. Wen, "A novel cross layer anti-collision algorithm for slotted aloha-based UHF RFID systems," *IEEE Access*, vol. 7, pp. 36 207–36 217, 2019.

[16] C. Wang, J. Liu, Y. Chen, H. Liu, L. Xie, W. Wang, B. He, and S. Lu, "Multi - touch in the air: Device-free finger tracking and gesture recognition via COTS RFID," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, 2018.

[17] OneSpan, "What is behavioral biometrics?" https://www.onespan.com/topics/behavioral-biometrics#:~:text=What%20is%20behavioral%20biometrics%3F%20Behavioral%20biometrics%20analyze%20the,of%20behavior%20that%20is%20unique%20to%20a%20person.

[18] H. Ding, L. Guo, C. Zhao, X. Li, W. Shi, and J. Zhao, "Device-free gesture recognition using time series rfid signals," in *Proceedings of the EAI International Conference on Broadband Communications, Networks, and Systems (BROADNETS)*, 2019.

[19] H. Farrukh, R. Mohamed, S. Cao, and H. Wang, "FaceRevelio: A face liveness detection system for smartphones with a single front camera," in *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2020.

[20] Z. Yang, J. Zhang, E. Chang, and Z. Liang, "Neural network inversion in adversarial setting via background knowledge alignment," in *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2019.

[21] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Zero-effort cross-domain gesture recognition with wi-fi," in *Proceedings of the ACM Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2019.

[22] M. T. Islam and S. Nirjon, "Wi-fringe: Leveraging text semantics in wifi csi-based device-free named gesture recognition," 2019.

[23] J. Guo, T. Wang, Y. He, M. Jin, C. Jiang, and Y. Liu, "Twinleak: Rfid-based liquid leakage detection in industrial environments," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, 2019.

[24] W. Xu, J. Liu, S. Zhang, Y. Zheng, F. Lin, J. Han, F. Xiao, and K. Ren, "Rface: Anti-spoofing facial authentication using COTS RFID," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, 2021.

[25] PyTorch, "Saving and loading models," https://pytorch.org/tutorials/beginner/saving_loading_models.html.

[26] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *Proceedings of the Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2016.

[27] M. S. El-Tokhy, "Robust multimodal biometric authentication algorithms using fingerprint, iris and voice features fusion," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 1, pp. 647–672, 2021.

[28] L. Li, S. Zhou, and H. Tu, "Fingerprint authentication based on fuzzy extractor in the mobile device," *International Journal of Electronic Security and Digital Forensics (IJESDF)*, vol. 11, no. 3, pp. 321–337, 2019.

[29] C. Xie and A. Kumar, "Finger vein identification using convolutional neural network and supervised discrete hashing," *Pattern Recognition Letters*, vol. 119, pp. 148–156, 2019.

[30] C. Ye, B. V. K. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2930–2941, 2012.

[31] S. Ilankumaran and D. Chelliah, "Multi-biometric authentication system using finger vein and iris in cloud computing," *Cluster Computing*, vol. 22, no. Suppl 1, pp. 103–117, 2019.

[32] W. Chen and W. Wang, "Fusion of hand-shape and palm-print traits using morphology for bi-modal biometric authentication," *International Journal of Biological Macromolecules (IJBM)*, vol. 10, no. 4, pp. 368–390, 2018.

[33] Z. Oo, L. Wang, K. Phapatanaburi, M. Liu, S. Nakagawa, M. Iwahashi, and J. Dang, "Replay attack detection with auditory filter-based relative phase features," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2019, p. 8, 2019.

[34] D. Ye, T. Zhang, and G. Guo, "Stochastic coding detection scheme in cyber-physical systems against replay attack," *Journal of Information Sciences*, vol. 481, pp. 432–444, 2019.

[35] H. A. Patil and M. R. Kamble, "A survey on replay attack detection for automatic speaker verification ASV system," in *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2018.

Jianwei Liu received the BS degree from Northwestern Polytechnical University in 2018. He received his Master degree from Xi'an Jiaotong University in 2021. He is working toward the Ph.D. degree at Zhejiang University. His research interests include RFID, mobile computing, and smart sensing. He is student member of the IEEE.

Kaiyan Cui received the BS degree from Taiyuan University of Technology in 2016. She is working toward the Ph.D. degree at Xi'an Jiaotong University and Hong Kong Polytechnic University. Her research interests include RFID, mobile computing, and smart sensing. She is student member of the IEEE and ACM.

Xiang Zou received the BS degree from Xi'an University of Posts and Telecommunications in 2014. He received the Master degree from Chang'an University in 2018. He is working toward the Ph.D. degree at Xi'an Jiaotong University. He research interests include RFID, mobile computing and IoT security.

Jinsong Han received his Ph.D. degree from Hong Kong University of Science and Technology in 2007. He is currently a professor of the College of Computer Science and Technology, Zhejiang University. His research interests focus on IoT security, smart sensing, wireless and mobile computing.

Feng Lin received the Ph.D. degree from the Department of Electrical and Computer Engineering, Tennessee Technological University, USA, in 2015. He is currently a Professor with the School of Cyber Science and Technology, College of Computer Science and Technology, Zhejiang University, China. He was an Assistant Professor with the University of Colorado Denver, USA, a Research Scientist with the State University of New York (SUNY) at Buffalo, USA, and an Engineer with Alcatel-Lucent (currently, Nokia). His current research interests include mobile sensing, wireless sensing, Internet of Things security, biometrics, and AI security. Dr. Lin was a recipient of the ACM SIGSAC China Rising Star Award, the Best Paper Awards from ACM MobiSys'20, IEEE Globecom'19, IEEE BHI'17, the Best Demo Award from ACM HotMobile'18, and the Best Paper Award Nomination from Infocom'21. He serves as an editor for IEEE Network and IEEE Access.

Kui Ren received the Ph.D. degree from the Worcester Polytechnic Institute, Worcester, MA, USA. He is currently a Professor of computer science and technology and the Director of the Institute of Cyberspace Research, Zhejiang University, Hangzhou, Zhejiang, China. His current research interests include cloud and outsourcing security, wireless and wearable system security, and artificial intelligence security. Dr. Ren is also a Distinguished Scientist and Fellow of the ACM. He was a recipient of the IEEE CISTC Technical Recognition Award 2017 and the NSF CAREER Award in 2011.