Hand-Key: Leveraging Multiple Hand Biometrics for Attack-Resilient User Authentication Using COTS RFID

Jianwei Liu[1], Xiang Zou[1], Feng Lin[1,2], Jinsong Han[1,2], Xian Xu[1], and Kui Ren[1,2,3]

[1]School of Cyber Science and Technology, Zhejiang University, China
[2]Key Laboratory of Blockchain and Cyberspace Governance of Zhejiang Province, China
[3]Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies, China
{liujianwei,xiang_zou}@stu.xjtu.edu.cn,{flin,hanjinsong,xianxu,kuiren}@zju.edu.cn

***Abstract*—Biometrics have been widely used in user authentications. However, existing outer-body biometrics (e.g., fingerprint), collecting from body surface, are vulnerable to spoofing attacks. Although inner-body biometrics, such as the electrocardiogram, are hard to be forged, their complex acquisition methods and instability lead to unsatisfactory user experience. Therefore, achieving good user-friendliness and high security simultaneously in biometric-based authentication is challenging. In this paper, we propose *Hand-Key*, an attack-resilient and user-friendly user authentication system to address the above challenge. *Hand-Key* utilizes a low-cost radio frequency identification (RFID) tag array to simultaneously collect the inner-body composition and outer-body geometric features of human hand to identify users. Users are merely required to hold their hands in a 'handshaking' pose between a reader's antenna and a tag array during authentication. To further enhance the security, we tactfully leverage the inherent randomness of the anti-collision scheme in RFID systems to make *Hand-Key* immune against replay attacks. We built a prototype of *Hand-Key* and conducted extensive experiments with 30 volunteers. The results show that *Hand-Key* achieves an authentication success rate of 99%+.**

I. Introduction

User authentication is one of the most important security means. When authenticating a user, s/he is usually required to verify her/is authenticity by providing certain factors, including knowledge, possession, or inherence. Knowledge factors usually require users to remember certain information, such as patterns and PINs. To guarantee its security, such information should be sufficiently complex so that attackers cannot easily guess it out, which is quite inconvenient to legitimate users. Possession of certain tokens or devices, however, is also not secure considering the risks of theft or loss [1]. Indeed, anyone (including the attacker) who holds the token or device would be accepted as a valid user.

The inherence is commonly related to the user's biometric, such as fingerprint, voiceprint, facial feature, and the like. Since such information is inherent with the user, the user does not need to remember certain knowledge or possess tokens such as keys or ID cards. Therefore, biometric-based solutions [2]–[4] become popular in user authentication applications in recent years. Nevertheless, most commonly used biometrics (*e.g.*, fingerprint, facial feature, and voiceprint) are 'outer-body' ones, which are collected from the surface of human body. As a result, these outer-body biometrics are easy to be left by direct touch or captured remotely, and hence vulnerable to various spoofing attacks, including the duplication, counterfeiting, impersonation, and replay attacks [3]. For instance, a faked fingerprint duplicated by clay can pass the verification and allow the attacker to intrude the system [5]. Many traditional facial feature-based authentications can be easily spoofed by photos [6]. Although liveness detection has facilitated the facial feature-based authentication to resist the spoofing attack [2], the emerging adversarial attack [7] would easily invalidate the anti-spoofing countermeasure. Voice signals are utilized to extract voiceprint [8], yet are vulnerable to the replay attack [9].

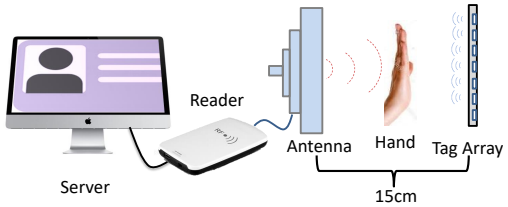


Fig. 1. The system illustration. Users only need to pose their hands between the reader's antenna and the tag array to achieve authentication.

Compared to the above biometrics, inner-body biometrics are relatively difficult to be intercepted or duplicated. However, their complex collection methods lead to poor user experience. For example, electrocardiography (ECG) [10] and photoplethysmography (PPG) [11] are considered as secure inner-body authentication credentials. But the collection of ECG and PPG requires body contact and complex body poses [10]. Besides, ECG and PPG might be severely influenced by human motion and emotion [11]. Therefore, a new biometric-based authentication that can guarantee the user-friendliness and security simultaneously is needed urgently.

Consider that RF signals are leveraged to capture human's biometric in a contact-free way in recent years [12]. Mean-

* Feng Lin is the corresponding author.

while, RFID tag array is able to sense fine-grained information of targets, e.g., tracking fingers [13]. In this paper, we attempt to utilize low-cost commercial radio frequency identification (RFID) tags to sense the inner composition and outer geometry feature of human hands to realize a new fusion biometric-based authentication system, which is secure yet easy to operate. We illustrate the system in Fig. 1, in which an RFID reader continuously emits RF signals and receives the ones backscattered from RFID tags. What a user needs to do for authentication is just holding her/is hand ('handshaking' gesture) between a reader's antenna and a tag array for a while. The RF signals would pass through and diffract around the hand to capture the hand's composition and geometry features. The system then extracts the hand biometrics from the signals and performs authentication.

However, realizing such a system is challenging due to the following reasons: 1) The composition feature is captured only when the RFID signals are capable to penetrate human hands. However, the proof of such penetrability is difficult because the penetration, if it happens, is nonvisible. 2) In the contact-free operation, the hand's position relative to the RFID reader in each authentication attempt is not identical, resulting in low authentication accuracy. 3) Wireless signal-based authentication systems are vulnerable to replay attacks [14].

In this paper, we propose a secure and contact-free user authentication system, namely *Hand-Key*. To overcome the first challenge, we theoretically prove that RF signals can penetrate human hands and activate tags afterwards by considering all potential power losses that RF signals meet during propagation. We also conduct validation experiments to validate the feasibility of sensing the desired hand features. To overcome the second challenge, we design a novel deep learning model, called 3D identity classifier, to effectively extract high-quality hand fusion biometrics from RF signals. Moreover, we propose a multi-group sampling scheme to increase the diversity of the hand position in model's training set to suppress the impact of the position variation. The sampling scheme would help 3D identity classifier improve its tolerance towards position variations. To address the last challenge, we design a novel anti-replay scheme that exploits the inherent randomness of the core anti-collision component in the RFID standard communication protocol, i.e., frame-based slot ALOHA (FSA) [15], to detect replay attacks. Furthermore, we employ a hash tree for this scheme to reduce its storage overhead and time cost.

Our contributions can be summarized as follows:

- We introduce a novel user authentication system using commodity RFID tags, in which we sense the inner-body composition and outer-body geometry of human hand in a contact-free way.
- We propose a multi-group sampling scheme to solve the problem caused by position variation. We also leverage the inherent randomness of FSA protocol to design an anti-replay scheme.
- Aiming to comprehensively evaluate the performance of *Hand-Key*, we invited 30 volunteers to collect experiment data within three months. The results show that *Hand-Key* can achieve an authentication success rate of $99\%+$.

II. Preliminary

Before we detail the feasibility and design of *Hand-Key*, it is necessary to briefly introduce preliminary elements of RFID systems, i.e., the FSA protocol and the special layout of tag array. The randomness of the FSA protocol is used to design our anti-replay scheme and the special layout can alleviate the coupling effect between tags.

A. Frame-Slot-based ALOHA Protocol

In a typical RFID system, a reader leverages an antenna to send RF signals to a tag. The tag is then identified by backscattering its electronic product code (EPC) to the reader. The reader can decode the backscattered signals to obtain the EPC as well as the signal indicators including received signal strength (RSS) and phase. In such identification procedures, collisions occur when multiple tags attempt to communicate with the reader simultaneously. In this case, the signals received by the reader are mixed by backscattered signals of multiple tags, leading the received signals to be un-decodable. To solve this problem, FSA protocol is proposed to control the read order of tags. In this protocol, the reader first sends a frame consisting of multiple slots to the tags. Each tag then selects a slot randomly and it will respond to the reader at the selected slot, despite other tags may respond at the same slot. If a collision happens, which indicates that two or more tags respond in a same slot, FSA will restart a new frame after the current frame and repeat the above procedure until all tags are read once. Since the tags are identified one after another in each frame and the identification order is random, we can define an EPC sequence that represents the read order of all tags. In this paper, we leverage the EPC sequence to design an anti-replay scheme to defend against replay attacks. The design of our anti-replay scheme is detailed in Section VII.

B. Special Layout of Tags

To collect hand biometrics by a tag array with suitable size, we build the tag array with 49 tags, the shape of which (a square with a side length of 19 centimeters) can cover most of the human hands. However, according to the research in [3], inductive coupling effect exists between two adjacent tags in the tag array. The coupling effect may weaken the received power of tags [16], making them unreadable. To deal with this issue, we adopt the layout of the tag array introduced in [13] to alleviate the coupling effect. In this layout, any two adjacent tags are arranged in a perpendicular way. The goal of this treatment is to guarantee that all tags are readable even if a hand is blocking the line-of-sight path between the reader's antenna and the tag array.

III. Feasibility Study

In this section, we prove that the composition and geometry biometrics of human hands can be captured by RF signals.

A. Capture of Human Hand Biometrics

For a human hand, its composition and geometry have been demonstrated distinguishable among persons [3], [17]. Song *et al.* [4] show that the hand geometry is distinguishable towards different persons and can be used as a critical biometric to verify person's identity. Zhao *et al.* [3] present that each person has an internal resistance, which varies in a range from 300 to 1,000 Ohms. The differences of the resistance among persons are derived from their composition differences. Inspired by these researches, we aim to leverage RF signals to capture the inner-body composition and outer-body geometry of human hand to achieve non-contact and device-free user authentication.

Intuitively, RF signals can capture the geometry feature of human hand, because different human hands produce different degrees of occlusion to the RF signals, which further cause different RSS and phase values of backscattered signals. However, to validate that the composition feature is also able to be captured by RF signals, we first should prove that RF signals can penetrate through human hand. To this end, we first show the penetrability of RF signals towards human hand theoretically, and then conduct a validation experiment to show the validity of our theoretical analysis.

We prove the penetrability through calculating the power loss during signal propagation. In *Hand-Key*, RF signals emitted by the reader first propagate in the air before reaching human hand. In this procedure, signal power would decay which is called free space path loss (FSPL) [18]. The FSPL can be calculated by:

$$FSPL(d_s, f, c) = 20\log_{10}\frac{4\pi d_s f}{c}, \tag{1}$$

where d_s, f, and c denote the traveling distance of the signal, the frequency of the transmitted signal, and the speed of light, respectively. Then, RF signals will meet the interface composed of skin and air during propagation. According to [19], reflection and incidence simultaneously happen at the interface. The reflection would cause some signal power loss (termed as reflection loss) because the reflected signals will not enter human hand. The ratio between the reflection power P_r and the incident power P_i can be formulated as:

$$\frac{P_r}{P_i} = \left|\frac{\sqrt{\epsilon_a} - \sqrt{\epsilon_s}}{\sqrt{\epsilon_a} + \sqrt{\epsilon_s}}\right|^2, \tag{2}$$

where the ϵ_a and ϵ_s are the electrical permittivity of air and skin, respectively. Next, when the incident signals propagate in the human hand, some signals would be absorbed by the composition of the hand, which is termed as absorption loss. Human hand is mainly composed of skin, sheath, muscle, and bones. The major part of RF signals will be absorbed by muscle, because the absorption ability is determined by the electrical permittivity and muscle has the largest electrical permittivity among these composition types [19]. According to [20], the absorption loss can be calculated by:

$$L_{abs}(\alpha, d_h) = 20\log_{10} e^{\alpha d} = 8.686\alpha d_h, \tag{3}$$

where α is the attenuation constant and d_h represents the propagating distance of the signal. To calculate the absorption loss, we consider an extreme case that a human hand is totally composed of muscle. We take five centimeters as the thickness of the hand because human hand is not totally composed of muscle and its thickness is normally less than five centimeters in our common sense. If the signal can penetrate such a hand with the extreme setting, it can penetrate human hands in most cases in the real-world. Given an emitted signal power of 32dbm (*i.e.*, the default distance in *Hand-Key*), the power of the signal at the tag will be approximately 6dBm after taking free space path loss, reflection loss L_{ref}, and absorption loss into consideration. This result indicates that the signal power harvested by the tag is sufficient to activate itself and is able to backscatter its EPC [15].

To show the validity of our theoretical analysis towards the penetrability, we conduct a validation experiment. in which we cover a tag with a cup wrapped by tinfoil paper. A volunteer is asked to cover the hole in the top of the cup with his hand. In this case, the reader can obtain the tag's EPC only when the signal can penetrate through the hand twice. As a result, the tag is readable when the distance between the reader's antenna and the tag is 15 centimeters (our default setting). Hence, in *Hand-Key*, RF signals indeed can penetrate human hands.

Under the proof of penetrability, we can further show that geometry and composition features can be captured by RF signals theoretically. If we denote the RSS of the transmitted signal as RSS_t, the RSS received by tag can be formulated as:

$$RSS_r = RSS_t - FSPL(d_s, f, c) - L_{ref} - L_{abs}(\alpha, d_h). \tag{4}$$

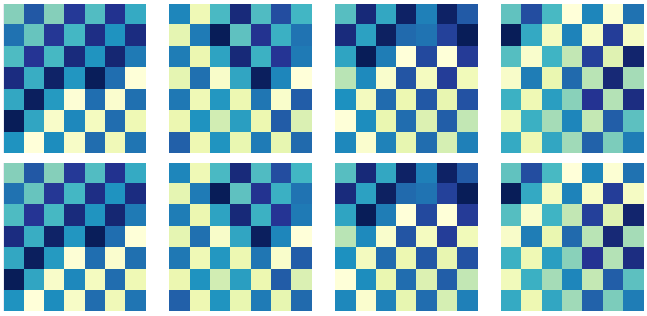
Since α is related to hand's composition, d_s and d_h are related to hand's geometry, RSS can indicate the geometry and composition biometrics. If we denote the phase of the transmitted signal as θ_t, the phase of the signal received by tag is:

$$\theta_r = (\frac{4\pi d_s}{\lambda_s} + \frac{4\pi d_h}{\lambda_h} + \theta_t) \mod 2\pi, \tag{5}$$

where λ_s and λ_h are the wavelengths of signals in air and hand respectively [21]. Likewise, λ_h is related to hand's composition, d_h and d_s are related to hand's geometry, demonstrating that phase can also quantify the geometry and composition of human hand.

B. Distinguishability Validation

In order to experimentally confirm that the composition feature and geometry feature can be brought out by RF signals, we conduct three validation experiments. In the first experiment, we make a human hand using a glove filled with water and oil. In particular, we mix the water and oil with different ratios to simulate different compositions, while keeping the volume (geometry) of the mixed liquid unchanged. The reason behind is that according to [19], muscle and skin are water-based biomaterials and fat is oil-based biomaterial. After the signal collection, the RSS and phase differences



(a) Volunteer 1. (b) Volunteer 2. (c) Volunteer 3. (d) Volunteer 4.

Fig. 2. The RSS distributions of four volunteers. The deeper the color is, the lower the RSS value is.

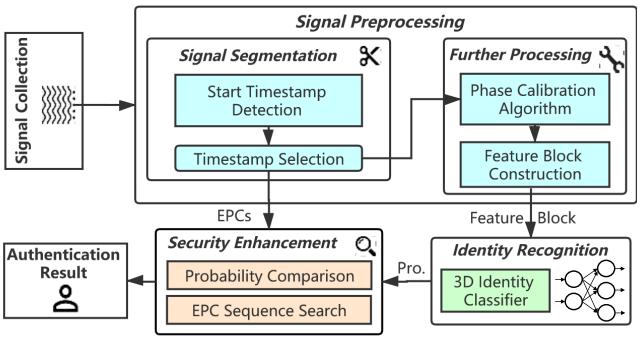


Fig. 3. The workflow of *Hand-Key*

between each pair of tags are treated as the training features. We adopt support vector machine (SVM) to identify different ratios. The recognition accuracy is 100%, which means that the composition feature is indeed carried out when RF signals pass through the hand. In the second experiment, we vary the volume of the mixed liquid while fixing the mix ratio between water and oil, simulating different hand geometries. We use the same training method as the one in the first experiment. We find that the accuracy is still 100%, which indicates that the geometry feature can also be captured by RF signals. In the last experiment, we invite four volunteers and collect two different signal samples for each volunteer. The RSS distributions of their signal samples are shown in Fig. 2. It can be found that the RSS distributions of the same volunteer are similar. Meanwhile, the RSS distributions of different volunteers are differentiable. In addition, we find that the phase distributions are distinguishable between any two users as well. Hence, it is feasible to utilize the feature fused by composition and geometry biometrics to achieve high-distinguishability authentication.

IV. Sysem Design

In this section, we first present the overview of *Hand-Key* and then introduce each module in Hang-Key respectively.

A. Overview

The architecture of *Hand-Key* is primarily composed of two phases: registration phase and authentication phase. As shown in Fig. 3, the registration phase contains the *data preprocessing* module and *identity recognition* module. The authentication phase not only uses the two modules contained in the registration phase, but also the *security enhancement* module.

In the registration phase, users put their hands between the reader's antenna and the tag array for several seconds to collect a batch of signal samples. These signal samples are then inputted into the *signal preprocessing* module to obtain a batch of feature blocks (composed of RSS and phase values). Afterwards, a deep neural network in the *identity recognition* module is trained by using these signal samples as a training set. The registration is complete after the training,

In the authentication phase, user initiates an authentication request by putting her/is hand between the reader's antenna and the tag array for a short time. A signal sample is collected in this process. The signal sample is then preprocessed by the first module to obtain a feature block. Then the *identity recognition* module utilizes the feature block and pre-trained deep neural network to calculate the probability that the signal sample should be identified as each registered user. The largest probability is then compared with an acceptance threshold in the *security enhancement* module. If the probability is smaller than the acceptance threshold, the authentication request will be considered to be initiated by an illegitimate user and thus rejected. Otherwise, the *security enhancement* module extracts an EPC sequence from the signal sample and compare the EPC sequence with the ones stored in the database. If the EPC sequence is found in the database, the authentication request will be considered as a replayed one and rejected. Only when the EPC sequence is different from every one in the database, the authentication request will be accepted and this EPC sequence will be stored in the database for later comparison in new authentication requests.

The deployment of Hand-Key also supports a more distributed implementation. Similar to the massive MIMO, many tags deployed on multiple users' devices would provide more plentiful information for more fine-grained authentication. Moreover, a pair of reader and a tag array can be regarded as a distributed node. Through wide deployment of these nodes, Hand-Key can provide authentication services for a large number of users at different locations at the same time.

B. Function of Module

Data preprocessing: The goal of this module is to denoise the raw signal sample and generate a feature block containing the fusion biometric of human hand. Specifically, this module first utilizes a phase calibration algorithm to replace the outliers in the raw phase values by normal ones. Afterwards, *Hand-Key* calculates the RSS difference and phase difference as feature values to suppress the impact of environment noise. The feature values are then constructed to be a feature block, which will be used as the input of our deep neural network. The details of this module are arranged in Section VI.

Identity recognition: In this module, *Hand-Key* identifies users by utilizing a well-designed deep neural network called 3D identity classifier. The input of this model is a three-dimensional feature block and it has two sub-blocks, RSS

sub-block and phase sub-block. The output of the network is a vector, each element of which is a probability that the input belongs to a specific user. The user that has the largest probability is regarded as the provider of the feature block. The design of the 3D identity classifier will be detailed in Section VI.

Security enhancement: This module is responsible for defending against potential attacks. First, the largest probability outputted by the *identity recognition* module is compared with an acceptance threshold in this module. If the acceptance is larger than the probability, the authentication request is likely to be initiated by an attacker. Otherwise, *Hand-Key* utilizes our anti-replay scheme to extract a new EPC sequence from the signal sample. The signal sample is treated as a replayed one once the new EPC sequence is found to exist in the database. Only when *Hand-Key* does not find an identical EPC sequence to the new one, the authentication request is considered as a benign one. The new EPC sequence will be stored in the database. We elaborate our anti-replay scheme in Section VII.

V. Data Preprocessing

In this section, we first show the signal segmentation method and then elaborate why and how we perform phase calibration to deal with abnormal phase values. Finally, we introduce our feature block construction method, in which RSS and phase difference values are calculated to suppress the impact of environment noise.

A. Signal Segmentation

To construct a feature block that contains hand features, we first need to detect the start timestamp of the authentication. Since the occlusion of the hand would greatly reduce the RSS of received signals, we detect the start timestamp by monitoring the RSS variation of the center tag. Specifically, we use a time window with k continuous RSS values of the center tag to calculate the sum S of the RSS values in the window. Once the S of a window is t larger than the sum of the previous window, we will set the timestamp of the third RSS value in this window as the start timestamp. We empirically set k as 6 and t as 5. Then, we segment the raw signals so that each tag has five timestamps after the start timestamp (each timestamp corresponds to an RSS value and a phase value). Next, we align the RSS and phase values of all tags in the tag array to construct two sub-blocks: RSS sub-block and phase sub-block. Each sub-block has a dimension of (5,7,7). Though the two sub-blocks contain hand features, we cannot directly use them to identify users because: 1) Abnormal phase values are mixed in the normal ones; 2) The sub-blocks contain not only hand features but also noise induced by ambient environment. The abnormal phase values and environment noise would impact the performance of *Hand-Key*. We thereby design a phase calibration algorithm and a feature block construction method to improve the robustness of *Hand-Key*.

B. Phase Calibration

In the phase collection in an RFID system, phases change periodically and a phase jumping happens when the increased phase value crosses the boundary of $[0, 2\pi]$. In *Hand-Key*, the phase jumping comes from two components, i.e., hardware offset and hand jittering. First, the hardware offset would add extra phase values on the normal ones and thus induce phase jumping, resulting in unexpected errors to the final authentication. Second, user's hand is slightly jittering during the signal collection in *Hand-Key*, leading to phase jumping as well. Therefore, the collected phase values are not temporally stable. We formulate a tag's phase received by reader as follows:

$$\theta = (\frac{4\pi d}{\lambda} + \theta_{\mathrm{r}} + \theta_{\mathrm{t}}) \quad mod \quad 2\pi, \tag{6}$$

where λ, d, θ_r, and θ_t are the signal wavelength, the distance between the reader's antenna and the tag, the additional phase offset induced by the reader, and the offset induced by the tag's circuits, respectively. Here, θ_r and θ_t are the hardware offset. Since the jumped phase values should be treated as outliers, we design a calibration algorithm to replace these jumped values.

Our calibration algorithm is based on the observation that in a time window with 10 continuous phase values of a specific tag, the majority of these values are larger than π and the rest of them are smaller than π, or vice versa. Meanwhile, the ratio between the number of values in the majority and that in the minority is about 4:1. Obviously, those values included in the minority group are anomalies. Hence, the intuition behind our calibration algorithm is that 'the minority is subordinate to the majority'. For each window, the algorithm replaces each phase value in the minority group with the mean of the phase values in the majority group. For validating the effectiveness of this algorithm, we separately use calibrated phase values and uncorrected phase values to train the deep neural network in the *identity recognition* module. The results show that the calibration significantly improves the authentication accuracy, i.e., 15%.

C. Feature Block Construction

During signal collection in *Hand-Key*, ambient environment (*e.g.*, pedestrian) may influence the RSS and phase values of signals due to the multi-path effect [19]. To solve this problem. we calculate the RSS difference and phase difference as feature values to suppress the impacts of environment noise. This is because the distance between a tag and the ambient environment is relatively longer than that between any two adjacent tags, any two adjacent tags are affected by approximately the same environment noise. Thus, calculating the difference between adjacent tags can make the same environment noise cancel each other out [22]. Specifically, we calculate the difference value for both RSS and phase sub-blocks respectively. Taking RSS sub-block as an example, we calculate the RSS difference between any two adjacent rows:

$$R_{dif} = \begin{bmatrix} R_2^1 - R_1^1 & R_2^2 - R_1^2 & \cdots & R_2^c - R_1^c \\ R_3^1 - R_2^1 & R_3^2 - R_2^2 & \cdots & R_3^c - R_2^c \\ \vdots & \vdots & \cdots & \vdots \\ R_r^1 - R_{r-1}^1 & R_r^2 - R_{r-1}^2 & \cdots & R_r^c - R_{r-1}^c \end{bmatrix} \tag{7}$$

where R_r^c is the RSS value at the r_{th} row and c_{th} column. In this way, we obtain two new sub-blocks with the dimension of each (5,6,7). The two sub-blocks are then concatenated together as a feature block with a dimension of (2, 5, 6, 7), which is the input of our 3D identity classifier.

VI. Identity Recognition

In this section, we design a deep learning model, namely 3D identity classifier, to identify users. To improve the robustness of our 3D identity classifier, we propose a multi-group sampling scheme to combat the impact of hand position variation.

A. Modeling Considerations

Why a deep learning model? To identity users, an intuitive solution is to leverage a traditional classifier to perform identity classification. We thus invite ten volunteers to collect a batch of feature blocks to perform classification on five classic learning models, including SVM, decision tree (DT), naive Bayes classifier (NB), K-nearest neighbors (KNN) and three-layer neural network (NN). For each of them, we use 75% and 25% of feature blocks as the training and testing set respectively. We then calculate the ratio (*i.e.*, classification accuracy) between the number of correctly classified testing feature blocs and the number of all testing feature blocks. The results show that NN can achieve the highest accuracy of 86%. However, this accuracy is insufficient for a usable authentication system. Therefore, we need to design a better classifier that can extract more effective hand features from feature blocks to achieve accurate identity classification. Recently, deep convolutional neural network (DCNN) [23] shows powerful deep feature extraction ability. We are thereby encouraged to adopt DCNN as the basis of our identity classifier.

3D identity classifier: To extract hand biometrics by a DCNN, an intuitive method is directly feeding the feature block into convolutional layers. However, RSS value and phase value are not at the same numeric interval, i.e., they have different signs ('-' for RSS and '+' for phase) and different orders of magnitude (10^1 for RSS and 10^0 for phase). Therefore, using two convolutional branches to separately extract deep features from RSS sub-block and phase sub-block is more reasonable. Besides, to extract temporal features, we should perform convolution on the first dimension of the sub-blocks. To extract spatial features, it is necessary to perform 2D convolution on the second and the third dimensions. To this end, we design a two-branch 3D DCNN to effectively extract temporal-spatial features from feature blocks. As illustrated in Fig. 4, our identity classifier has two convolutional branches, and each of them has three 3D convolutional layers and two fully-connected layers. Each convolutional layer is followed by a batch normalization function and a rectified linear unit (ReLU). Batch normalization can help the network avoid data distribution offset. ReLU is used to increase the nonlinearity of the network and reduce the interneuronal dependency. The size of each 3D convolutional kernel and the convolutional stride are empirically set as (3,3,3) and (1,1,1), respectively. The outputs of the two fully-connected layers are first multiplied by two weights respectively and then added to form a probability vector. The weights for the RSS and phase branches are 0.8 and 0.2 respectively, because RSS is more stable compared with phase. The user that has the largest probability is regarded as the provider of the input feature block.

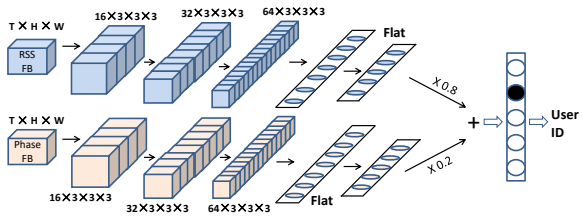


Fig. 4. The architecture of our 3D identity classifier.

B. Training Scheme

Impact of position variation: In a real-world authentication scenario, a user cannot guarantee that at each authentication her/is hand is placed at the position identical to the one in the registration phase. The hand position differences between the feature blocks in training set and the feature blocks collected during authentication may impact the authentication performance. To verify our conjecture, we conduct a validation experiment. We first collect four groups of feature blocks with ten volunteers in four periods. and then perform two types of identity classification. In the first type, we train and test our 3D identity classifier with the feature blocks in the same group, the classification accuracy can reach 99.8%. However, when we train our classifier with one group and test it with another group, the classification accuracy drops to 85%. The results demonstrate that the hand position difference between two groups indeed lowers the authentication performance.

Multi-group sampling scheme: To deal with the position variation issue, we try to improve the diversity of the hand positions in the training set. Specifically, we randomly sample feature blocks from multiple groups to form the training set, as illustrated in Fig. 5. In this way, we can improve the classifier's robustness to the position variation, because the 3D identity classifier can 'see' more positions during training. To show the effectiveness of our sampling scheme, we conduct a classification experiment, in which we sample feature blocks from different numbers of groups and test with different groups. To avoid the impact of training set size, we fix the number of feature blocks in the training set. The experiment results indicate that the classification accuracy increases to 99%+ by constructing the training set with three groups.

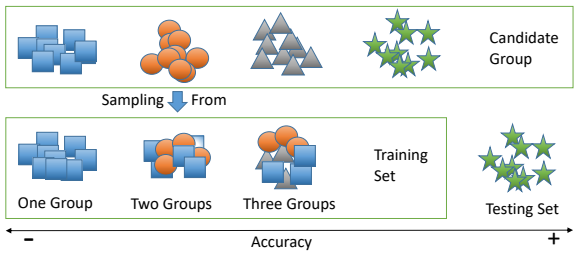


Fig. 5. The effectiveness of our multigroup sampling scheme.

VII. Anti-Replay Scheme

Replay attack is one of the most intractable attack types in RFID systems. In this section, we propose a novel anti-replay scheme to defend against replay attacks. We also design a structure called hash tree to reduce the overhead caused by the anti-replay scheme.

Fig. 6. Two EPC sequences in two read rounds.

A. EPC Sequence Construction

Existing most effective solution to achieve replay defense is to introduce random noise into the transmitted signals [14]. However, adding random noise in the signal in *Hand-Key* would destroy normal hand features. Therefore, we can only introduce randomness in the authentication procedure without changing the signal itself.

As aforementioned in Section II-A, mainstream passive RFID systems adopt FSA protocol to avoid collisions. At the beginning of each frame, each tag randomly selects a slot to respond. Thus, in each identification round (each tag is read once in each round), the read order of tags is random. Based on the above observation, we are motivated to fully utilize the inherent FSA protocol to yield sufficient randomness for the authentication procedure. Specifically, the off-the-shelf randomness in FSA protocol can be embodied by EPC sequence (*i.e.*, read order). As illustrated in Fig. 6, 'A', 'B', 'C', and 'D' are four EPCs of four tags. Intuitively, each of them could be read in any slot of the round due to the randomness of FSA protocol. Hence, the EPC sequence of these tags may be different in different rounds. In this example, the EPC sequences for 'Round X' and 'Round Y' are 'ACBD' and 'CBDA', respectively.

B. Theoretic Feasibility

To validate the feasibility of leveraging EPC sequence to defend against replay attacks, we first calculate the space of possible EPC sequence. Suppose that the tag array is formed by l tags (e.g., 49 in our default setting), the number of different read orders is $l!$, which means the EPC sequence space is $l!$. Theoretically, in m authentication requests, the

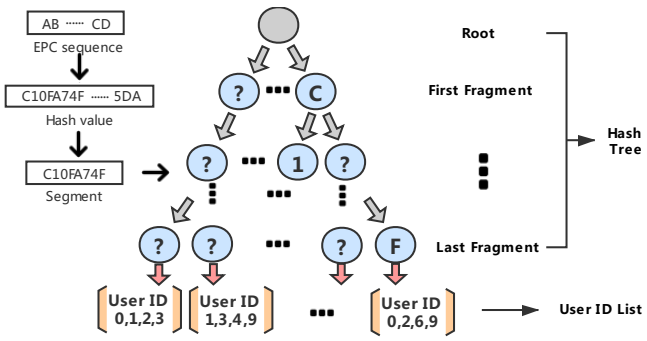


Fig. 7. The truncation method and structure of hash tree.

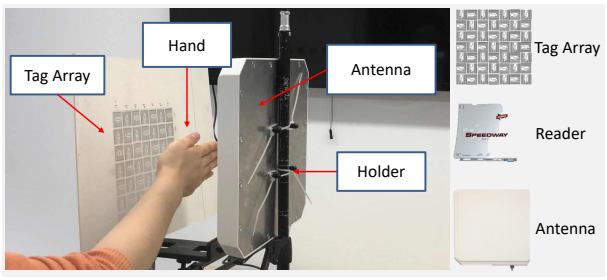


Fig. 8. The experiment setup.

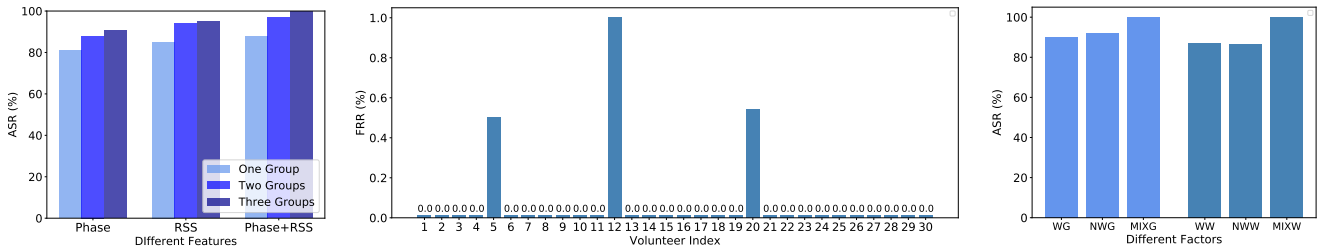
probability that m corresponding EPC sequences are different from each other can be formulated as:

$$P(n,m) = \frac{(n-1)!}{(n-m)!n^{(m-1)}}, \tag{8}$$

where n is the EPC sequence space. In *Hand-Key*, n can be set as 10^{60}, which is less than 49!. The m can be set as 10^5 because 10^5 is a considerable number of authentication requests for a user. Given n and m, the calculated non-duplication probability P is larger than 99%, which means that a specific EPC sequence is nearly impossible to appear twice even in 10^5 authentication requests. In this case, if in a new authentication request, *Hand-Key* extracts an EPC sequence that ever been extracted, we have 99%+ confidence that the new authentication request is a replayed one.

C. Hash Tree Design

In order to compare a new EPC sequence with previously-used ones, we need to store an EPC sequence after each successful authentication, which would cause some storage overhead. Moreover, the comparison between the new EPC sequence and stored ones would also cause some time cost. To reduce the storage overhead and time cost, we employ a hash function and tree-based search technique to transform and store EPC sequences. To be specific, we first hash each EPC sequence with MD5 Hash algorithm [24] and make an identical-length truncation for each hash value. Since an MD5 hash value contains a series of hexadecimal numbers, we term each hexadecimal number (four bits) as a fragment. For fast EPC sequence comparison (*i.e.*, search), we design a hash tree which has a hierarchical structure. As shown in the right part of Fig. 7, the tree has ten layers when including the root. The internal nodes of the tree are hash value fragments and the leaves are the lists of user IDs. A whole transformation and storing process is shown in Fig. 7, where the original EPC sequence is "AB$\cdots$CD". Then *Hand-Key* calculates the MD5 hash value and obtains "C10FA74F$\cdots$5DA". Afterwards, *Hand-Key* makes an eight-fragment truncation and adds the first eight fragments in the hash tree. In this example, "C" is the first fragment, so it is added in the first layer. Users 0, 2, 6, and 9 cannot use "C10FA74F" for a second time because they have used it in previous authentication requests. By using this tree, the time complexity (*i.e.*, time cost) decreases from $\mathcal{O}(M)$ to $\mathcal{O}(N)$, where M and N are the number of stored EPC sequences and the number of hash tree layers, respectively. M will increase over time but N is invariable.



(a) Training with different groups. (b) The FRR of 30 volunteers. (c) The impacts of related factors.

Fig. 9. The overall performance of *Hand-Key*: (a) Training with more groups can achieve better performance. (b) The FRR of most volunteers are 0.0%. (c) Glove and watch impact the ASR of *Hand-Key*.

VIII. Implementation and Evaluation

We built a prototype of *Hand-Key* and evaluated the performance by designing a series of experiments. We invited 30 volunteers (18 males and 12 females) aged from 20 to 35 to participate in our experiments. The heights of them ranged from 160cm to 188cm and their weights ranged from 45kg to 160kg.

Experiment setup: We used a commercial-off-the-shelf RFID reader, i.e., *Impinj R420*, and *Alien-9629* tags to build the prototype of *Hand-Key*. We conducted fundamental experiments in a laboratory environment and extended experiment in a crowded office environment with strong electromagnetic interference. The experiment setup is illustrated in Fig. 8, where a volunteer was asked to place her right hand between the tag array and the reader's antenna. The default distance between any two adjacent tags is three millimeters and the side length of the tag array is 19 centimeters. The default distance between the tag array and the reader is 15 centimeters. We collected 12432 feature blocks with the 30 volunteers.

Metrics: We defined four metrics to quantify the performance of *Hand-Key*: false reject rate (FRR), false accept rate (FAR), authentication success rate (ASR) and equal error rate (EER). FRR is the probability that a legitimate user is falsely rejected. It can be calculated by $FRR = \frac{N_{inc}}{N_{all}}$, where N_{rej} and N_{all} are the number of falsely rejected feature blocks and the number of all legitimate feature blocks, respectively. FAR is the probability that a feature block of an illegitimate user is accepted by *Hand-Key*. It can be calculated by $FAR = \frac{N_{acc}}{N_{ill}}$, where N_{acc} and N_{ill} are the number of accepted illegitimate feature blocks and the number of all illegitimate feature blocks, respectively. ASR is the probability that a legitimate user is successfully accepted by *Hand-Key*, which equals to $1-FRR$. EER is the FAR or FRR when FRR equals to FAR. It can be obtained by adjusting the acceptance threshold.

A. Overall Performance

In the laboratory environment, to show the effectiveness of our multi-group sampling scheme, we trained our 3D identity classifier with the feature blocks sampled from one, two and three groups respectively. Then, we tested our classifier with another group, in which no feature block was included in the training set. The experiment results are shown in Fig. 9(a), in which we can find that the ASR increases with the increase of the number of sampling groups. The ASR of using both RSS and phase is larger than that of using one of them. When the training set is composed of feature blocks from three groups, the ASR can achieve $99\%+$. Therefore, our multi-group sampling scheme can effectively improve *Hand-Key*'s robustness to the position variation. Moreover, the ASR in the office environment is 99%+ as well, which demonstrates that *Hand-Key* can effectively remove environment noise to achieve accurate authentication under harsh environments.

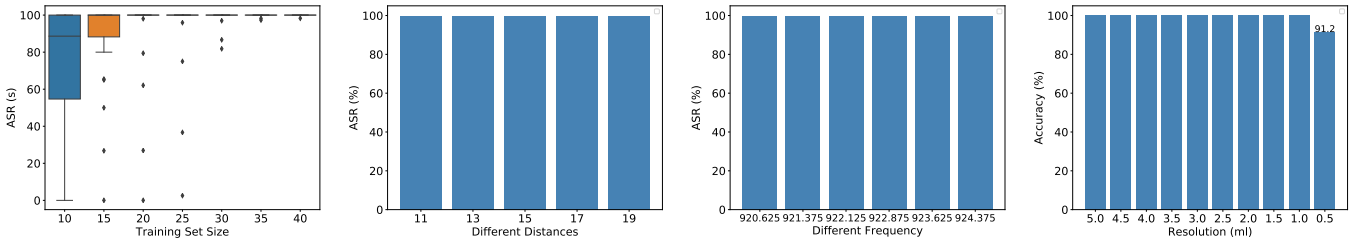
We then calculated the FRR of 30 legitimate volunteers and show the results in Fig. 9(b). It can be observed that the maximum of FRR is less than 1% and the mean FRR is less than 0.3%. Such a low FRR indicates that *Hand-Key* is significantly user-friendly. To calculate EER, we regarded 25 volunteers as legitimate users and the rest five volunteers as illegitimate users. The 75% feature blocks of legitimate users were used to train our 3D identity classifier. Afterwards, we varied the acceptance threshold from 0.7 to 0.9 with a stride of 0.01 and calculated FRR and FAR. The experiment results show that when the acceptance threshold is 0.8, we have $FRR = FAR$ and the EER is less than 1%. Thus, *Hand-Key* is usable and secure.

B. Impacts of Accessories

In this part, we considered the impacts of accessories, including glove and watch, that are commonly worn around human hand in people's daily life.

Impact of Glove: Consider that users may not wear gloves in the registration phase but wear gloves in the authentication phase, or vice versa. The impact of glove on the ASR should be explored. In the experiment, we collected a batch of feature blocks when volunteers were wearing normal winter gloves. Then we calculated the ASR under three conditions: 1) training with glove-on feature blocks and testing with glove-off ones (marked as WG); 2) Training with glove-off feature blocks and testing with glove-on ones (marked as NWG); 3) Training with both glove-on and -off feature blocks and testing with both of them as well (marked as MIXG). The experiment results, shown in Fig. 9(c), indicate that glove indeed impacts the performance of *Hand-Key*. However, using both glove-on and glove-off feature blocks can eliminate such impact. Besides, users can easily take on or off gloves to eliminate the impacts.

Impact of Watch: We used a metal watch to conduct the experiment because metal has larger impact on RF signals than other materials. The experiment method for watch is the same as that of glove. The experiment results are shown in



(a) The effect of training set size. (b) The effect of distance. (c) The effect of frequency. (d) Resolution experiment result.

Fig. 10. (a), (b), and (c) are the effects of different system settings. (d) is resolution classification result.

Fig. 9(c). It can be found that watch indeed impacts the ASR of *Hand-Key*. However, the impacts can also be suppressed by introducing both watch-on and -off feature blocks in the training set.

C. Effect of System Settings

In this part, we assess the effects of different system settings, including the training set size (*i.e.*, the number of feature blocks of each user in the training set), the distance between the reader's antenna and the tag array, and the frequency of transmitted signals.

Effect of training set size: We varied the training set size from 10 to 40 with a stride of 5 and calculated the ASR. As shown in Fig. 10(a), the ASR increases with the increase of the training set size. When the number of feature blocks per user is 20, the mean ASR can already exceed 90%. When the training set size is 40, the ASR is higher than 99%, which means collecting 40 feature blocks for each user in the registration phase is sufficient to train the 3D identity classifier. The collection of 40 feature blocks costs less than one minute. making *Hand-Key* significantly user-friendly.

Effect of distance: In this experiment, we varied the distance between the tag array and the antenna from 11 centimeters to 19 centimeters with a stride of two centimeters. As shown in Fig. 10(b), the distance has a negligible effect on the performance of *Hand-Key* when it is within 19 centimeters. Therefore, *Hand-Key* is flexible in distance selection.

Effect of frequency: It is worth noting that *Hand-Key* has 16 optional frequency from 920.62MHz to 924.375MHz. Thus, we varied the frequency from 920.62MHz to 924.375MHz with a stride of 0.75MHz and calculated the ASR. The result shown in Fig. 10(c), indicates that the variation of frequency does not affect the performance of *Hand-Key*. This result means that *Hand-Key* is also flexible in frequency selection.

D. Resolution Study

For an authentication system, it is easy to think of the resolution, i.e., if it can identify a large number of people. In order to figure out how fine-grained *Hand-Key* is, we carried a resolution experiment out by adding different amounts of water into a water-filled rubber glove to simulate different hands. We use water as the substitute for biomaterial because muscle and fat are all water-based biomaterials [19]. We totally collected ten groups of feature blocks. In different groups, we added different amount of water, i.e., from 5ml to 0.5ml with a stride of 0.5ml. Then we performed classification on these feature blocks. The results in Fig. 10 show that even when the resolution is 1ml, the classification accuracy is 100%. When the resolution is 0.5ml, the classification accuracy drops 9% approximately. However, the accuracy, 91%, is still high. Hence, we believe that *Hand-Key* is capable to identify a large number of users.

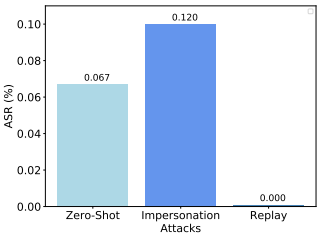


Fig. 11. The defense ability of *Hand-Key*.

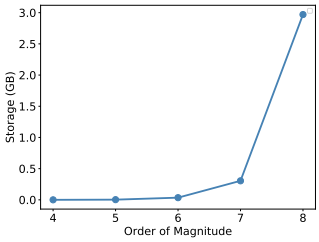


Fig. 12. The storage overhead caused by EPC sequence.

E. Security Analysis

To assess the security of *Hand-Key*, we considered three types of mainstream and potential attacks, including zero-effort attack, impersonation attack, and replay attack [3], and conducted corresponding defense experiments.

Zero-effort attack: In this attack, an attacker attempts to be accepted by Hand-Key by placing her/is hand between the reader's antenna and the tag array. To assess the defense ability of Hand-Key against this attack, we treated five volunteers as attackers to attack Hand-Key. The result is shown in Fig. 11, The mean ASR for attackers is only 0.067%, which indicate that Hand-Key is able to defend against zero-effort attacks effectively.

Impersonation attack: In this attack, we assume that the attacker is able to find an illegitimate user whose hand is similar to a legitimate user. The attacker tries to be accepted by Hand-Key by letting the illegitimate user exhibit her/is hand to Hand-Key. In the defense experiment, we selected three pairs of volunteer and the two volunteers' hands in every pair are similar. In each pair, we treated one volunteer as attacker and another one as victim. The experiment result (Fig. 11) shows that the ASR for impersonation attacker is only 0.120$. Therefore, Hand-Key can also defend against impersonation attacks.

Replay attack: In this attack, the attacker first eavesdrops RF signals of legitimate authentication processes, and then replay these signals to deceive Hand-Key. To simulate replay attacks, we input ten volunteers' signals into the *signal prepro-*

TABLE I
COMPARING *Hand-Key* WITH RF-MEHNDI AND WIPIN.

System	ASR$\geq$ 99%	Device-Free	Anti-Replay
Hand-Key	✓	✓	✓
RF-Mehndi	✓	×	×
WiPIN	×	✓	×

cessing module of Hand-Key. As a result, all replayed signals are rejected (as shown in Fig. 11). Thus, Hand-Key is resistant to replay attacks.

F. Overhead and Latency

Storage overhead: The storage consumption of Hand-Key mainly comes from two aspects: 3D identity classifier storage and EPC sequence storage. First, Hand-Key needs about 6MB to store the parameters of the 3D identity classifier. In terms of the EPC sequence storage, we varied the number of EPC sequences from 10^4 to 10^8. The experiment results are shown in Fig. 12, in which we can find that the storage consumption increases exponentially with the increase of the order of magnitude of EPC sequences' number. However, even if the number of EPC sequences reaches 10^8, the storage consumption is still less than 3GB. It means that if each user uses Hand-Key for 10^5 times, Hand-Key can protect 10^3 users from replay attacks simultaneously. Thus, the storage overhead of Hand-Key is acceptable.

Latency: The time cost of Hand-Key primarily comes from three components: signal collection, identity classification, and EPC sequence comparison. Firstly, to collect a feature block with a dimension of (5, 2, 6, 7), the process costs approximately 1 second. Secondly, with a 2.8GHZ CPU, the 3D identity classifier can process a feature block within 0.02 seconds. Thirdly, since we leverage hash tree to reduce the search complexity, the EPC sequence comparison can be finished within 0.01 seconds even if Hand-Key needs to compare 10^5 pairs of EPC sequences. Thus, it takes Hand-Key $2-$ seconds to process an authentication request. Hand-Key has outstanding real-time performance.

G. Comparison with Existing Works

We compare Hand-Key with a state-of-the-art RFID-based authentication system RF-Mehndi [3] and a state-of-the-art WiFi-based authentication system WiPIN [12]. The comparison results are shown in Table 1. Firstly, both Hand-Key and RF-Mehndi can achieve 99%+ ASR, yet the ASR of WiPIN is only 92% approximately. In terms of the user-friendliness, both Hand-Key and WiPIN are device-free, while RF-Mehndi requires users to carry an RFID token for identification. For the security, Hand-Key is capable to defend against replay attacks, but RF-Mehndi and WiPIN do not have such ability. Therefore, Hand-Key outperforms these two state-of-the-art authentication systems.

IX. RELATED WORK

Hand-based authentication: A large number of researches were proposed by using biometric features from human hands [3], [4], [25]–[29]. For instance, palm vein patterns can be extracted to achieve person recognition [26], [27] and finger vein patterns can also be captured for user identification [28]. Arora *et al.* [29] present a robust authentication method which extracts dorsal hand vein patterns to finish authentication. Song *et al.* [4] extract geometry and behavioral features of hands using smartphone sensors to identify persons. Zhao *et al.* [3] use RFID tag array to sense user's conductivity features to execute authentication. Hand-Key is different from the above works because Hand-Key is the first work to extract stable hand composition features to authenticate users.

Wireless authentication: Wireless authentication provides users good experience because it is non-contact. A lot of non-contact authentication methods have been realized via wireless devices [30]–[42]. For example, CardiacScan [43] uses a continuous-wave radar to capture the cardiac motion feature remotely. WiWho [30] employs a WiFi system to capture human biometrics to authenticate users. Similarly, WiFi-ID [44] also utilizes WiFi signals to extract human features. WiPIN [12] is a WiFi-based operation-free user identification system. RF-Mehndi [3] leverages RFID signals to capture human biometrics. Moreover, voice is also used to conduct user authentication [45]. Compared with the above works, Hand-Key shows superiority because Hand-Key is not only non-contact but also replay attack-resilient.

X. CONCLUSION

In this paper, we propose Hand-Key, an attack-resilient user authentication system, which captures inner-body composition feature and outer-body geometry feature to identify users. Hand-Key leverages a two-branch 3D identity classifier to identify users. The output of the 3D identity classifier, i.e., the probability, is used to detect illegitimate authentication requests. To defend against replay attacks, we design a novel anti-replay method, which reuse the inherent randomness of the FSA protocol. The experiment results with 30 persons show that Hand-Key can achieve an authentication accuracy as high as $99\%+$.

XI. ACKNOWLEDGEMENT

We sincerely thank our shepherd Prof. Kaishun Wu and the anonymous reviewers for their valuable comments and feedback. This work is supported in part by the National Key Research and Development Project (Grants No. 2017YFC0806100 and 2020AAA0107700), the National Natural Science Foundation of China (Grants No. 61872285, 62032021, 61772236, and 61972348), Zhejiang Key R&D Plan (Grant No. 2019C03133), Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Grant No. 2018R01005), major project of the National Social Science Foundation (Grant No. 20ZDA062), Alibaba-Zhejiang University Joint Institute of Frontier Technologies, and Research Institute of Cyberspace Governance in Zhejiang University.

REFERENCES

[1] C. Bessette, "How serious a crime is credit card theft and fraud," Technique Report. https://www.nerdwallet.com/blog/credit-cards/credit-card-theft-fraud-serious-crime-penalty/, 2018.

[2] T. Edmunds and A. Caplier, "Motion-based countermeasure against photo and video spoofing attacks in face recognition," *Journal of Visual Communication and Image Representation*, vol. 50, pp. 314–332, 2018.

[3] C. Zhao, Z. Li, T. Liu, H. Ding, J. Han, W. Xi, and R. Gui, "RF-Mehndi: A fingertip profiled RF identifier," in *IEEE International Conference on Computer Communications, INFOCOM*, 2019.

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

[5] TECHWORM, "iphone fingerprint sensor hacked with a finger made of clay at mwc," Report. https://www.techworm.net/2016/02/iphone-fingerprint-sensor-hacked-finger-made-clay-mwc-2016.html, 2016.

[6] N. M. Duc and B. Q. Minh, "Your face is not your password," Technique Report. https://www.blackhat.com/presentations/bh-dc-09/Nguyen/BlackHat-DC-09-Nguyen-Face-not-your-password.pdf.

[7] S. Komkov and A. Petiushko, "Advhat: Real-world adversarial attack on arcface face ID system," *CoRR*, vol. abs/1908.08705, 2019.

[8] S. Pradhan, W. Sun, G. Baig, and L. Qiu, "Combating replay attacks against voice assistants," *IMWUT*, vol. 3, no. 3, pp. 100:1–100:26, 2019.

[9] K. N. R. K. R. Alluri and A. K. Vuppala, "IIIT-H spoofing countermeasures for automatic speaker verification spoofing and countermeasures challenge 2019," in *Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2019.

[10] Y. Gao, W. Wang, V. V. Phoha, W. Sun, and Z. Jin, "Earecho: Using ear canal echo for wearable authentication," *Journal of Interactive, Mobile, Wearable and Ubiquitous Technologies, IMWUT*, vol. 3, no. 3, pp. 81:1–81:24, 2019.

[11] Y. Cao, Q. Zhang, F. Li, S. Yang, and Y. Wang, "PPGPass: Nonintrusive and secure mobile two-factor authentication via wearables," in *IEEE International Conference on Computer Communications, INFOCOM*, 2020.

[12] F. Wang, J. Han, F. Lin, and K. Ren, "Wipin: Operation-free passive person identification using wi-fi signals," in *IEEE Global Communications Conference, GLOBECOM*, 2019.

[13] 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 *IEEE International Conference on Computer Communications, INFOCOM*, 2018.

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

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

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

[17] A. Ross, "A prototype hand geometry-based verification system," in *Conference on Audio and Video Based Biometric Person Authentication*, 1999.

[18] Y. Luh and Y. Liu, "Measurement of effective reading distance of uhf rfid passive tags," *Modern Mechanical Engineering*, vol. 03, no. 03, pp. 115–120, 2013.

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

[20] M. J. Bentum, F. B. J. Leferink, A. Meijerink, and R. J. A. Wezel, "Analysis of radio propagation inside the humanbody for in-body localization purposes," 2014.

[21] C. Feng, J. Xiong, L. Chang, J. Wang, X. Chen, D. Fang, and Z. Tang, "Wimi: Target material identification with commodity wi-fi devices," in *39th IEEE International Conference on Distributed Computing Systems, ICDCS*, pp. 700–710.

[22] G. Wang, J. Han, C. Qian, W. Xi, H. Ding, Z. Jiang, and J. Zhao, "Verifiable smart packaging with passive RFID," *IEEE Transactions on Mobile Computing, TMC*, vol. 18, no. 5, pp. 1217–1230, 2019.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[24] X. Wang, D. Feng, X. Lai, and H. Yu, "Collisions for hash functions md4, md5, HAVAL-128 and RIPEMD," *IACR Cryptology ePrint Archive*, p. 199, 2004.

[25] J. Liu, X. Zou, J. Han, F. Lin, and K. Ren, "BioDraw: Reliable multi-factor user authentication with one single finger swipe," in *IEEE/ACM International Symposium on Quality of Service, IWQoS*, 2020.

[26] L. Mirmohamadsadeghi and A. Drygajlo, "Palm vein recognition with local binary patterns and local derivative patterns," in *IEEE International Joint Conference on Biometrics, IJCB*, 2011.

[27] M. I. Obayya, M. El-Ghandour, and F. Alrowais, "Contactless palm vein authentication using deep learning with bayesian optimization," *IEEE Access*, vol. 9, pp. 1940–1957, 2021.

[28] N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger vein patterns based on iterative line tracking and its application to personal identification," *Machine Vision and Applications*, vol. 35, no. 7, pp. 61–71, 2004.

[29] P. Arora, S. Srivastava, M. Hanmandlu, and S. Bhargava, "Robust authentication using dorsal hand vein images," *IEEE Intelligent Systems*, vol. 34, no. 2, pp. 25–35, 2019.

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

[31] R. Bharti and P. Bansal, "Real time speaker recognition system using mfcc and vector quantization technique," *International Journal of Computer Applications*, vol. 117, no. 1, 2015.

[32] E. Chandra and K. M. M. Kalaivani, "A study on speaker recognition system and pattern classification techniques," 2014.

[33] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2008.

[34] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?" in *IEEE International Conference on Computer Vision, ICCV*, 2011.

[35] W. Deng, J. Hu, and J. Guo, "Extended SRC: Undersampled face recognition via intraclass variant dictionary," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1864–1870, 2012.

[36] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *Transactions on Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.

[37] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[38] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition," in *IEEE International Conference on Computer Vision, ICCV*, 2005.

[39] B. Zhou, J. Lohokare, R. Gao, and F. Ye, "Echoprint: Two-factor authentication using acoustics and vision on smartphones," in *ACM International Conference on Mobile Computing and Networking, MobiCom*, 2018.

[40] H. Chen, W. Wang, J. Zhang, and Q. Zhang, "Echoface: Acoustic sensor-based media attack detection for face authentication," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 2152–2159, 2019.

[41] D. Tang, Z. Zhou, Y. Zhang, and K. Zhang, "Face flashing: a secure liveness detection protocol based on light reflections," in *Network and Distributed System Security Symposium, NDSS*, 2018.

[42] W. Xu, J. Liu, S. Zhao, Y. Zheng, F. Lin, J. Han, F. Xiao, and K. Ren, "RFace: anti-spoofing facial authentication using cots rfid," in *IEEE International Conference on Computer Communications, INFOCOM*, 2021.

[43] F. Lin, C. Song, Y. Zhuang, W. Xu, C. Li, and K. Ren, "Cardiac scan: A non-contact and continuous heart-based user authentication system," in *ACM International Conference on Mobile Computing and Networking, MobiCom*, 2017.

[44] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *IEEE International Conference on Distributed Computing in Sensor Systems, DCOSS*, 2016.

[45] H. Feng, K. Fawaz, and K. G. Shin, "Continuous authentication for voice assistants," in *ACM International Conference on Mobile Computing and Networking, MobiCom*, 2017.