Behavior Privacy Preserving in RF Sensing

Jianwei Liu, *Student Member, IEEE,* Chaowei Xiao, *Student Member, IEEE,* Kaiyan Cui, *Student Member, IEEE,* Jinsong Han, *Senior Member, IEEE,* Xian Xu, and Kui Ren, *Fellow, IEEE*

Abstract—Recent years have witnessed the booming development of RF sensing, which supports both identity authentication and behavior recognition by analysing the signal distortion caused by human body. In particular, RF-based identity authentication is more attractive to researchers, because it can capture the unique biological characteristics of users. However, the openness of wireless transmission raises privacy concerns since human behaviors could expose massive private information of users, which impedes the real-world implementation of RF-based user authentication applications. It is difficult to filter out the behavior information from the collected RF signals. In this paper, we propose a privacy-preserving deep neural network named *BPCloak* to erase the behavior information in RF signals while retaining the ability of user authentication. We conduct extensive experiments over mainstream RF signals collected from three real wireless systems, including the WiFi, radio frequency identification (RFID), and millimeter-wave (mmWave) systems. The experimental results show that *BPCloak* significantly reduces the behavior recognition accuracy, *i.e.*, 85%+, 75%+, and 65%+ reduction for WiFi, RFID, and mmWave systems respectively, merely with a slight penalty of accuracy decrease when using these three systems for user authentication, *i.e.*, 1%-, 3%-, and 5%-, respectively.

Index Terms—RF Sensing, Behavior Recognition, Privacy Preserving, Deep Learning.

1 Introduction

RF sensing technologies have gained widespread attention in recent years. They utilize RF signals to collect various object information. Compared with related solutions (*e.g.*, camera- and wearable-based solutions), RF-based ones have many advantages. For example, they can work under non-line-of-sight scenarios in a device-free manner. More importantly, RF sensing can mitigate visual privacy concerns since it would not reveal user's visual information. Therefore, RF signals are exploited to enable a variety of important human-computer interaction applications, such as user authentication [1], behavior recognition [2], and tracking [3].

Among these applications, user authentication is a particular interest of researchers in recent years [1], [4], [5], [6], [7]. This is because RF signals can penetrate human bodies and carry biological or physiological characteristics [4]. By analyzing the received RF signals, one can extract user-dependent and unique identity information to perform authentication. Besides, the identity information in RF signals is difficult to be stolen/duplicated/forged, compared with other kinds of identity information (*e.g.*, fingerprint and facial features) [8].

However, RF-based user authentication also introduces privacy concerns, because RF signals not only can record users' identity information, but also are able to sense users' behavior information. For example, while users may enjoy RF-based user authentication [4] provided by a service provider (SP) [9], it also allows the SP to steal their behavior privacy. A malicious SP could utilize the received signals to mine users' private behaviors, such as body movements [10], arm motions [2], and finger traces [11]. The SP can even speculate the users' personality [12], psychology [13], and some private passwords [14] based on the stolen behavior information.

Given these severe privacy and security consequences, one question is naturally raised: *can RF signal only record the identity information while avoiding capturing the behavior one?* To answer this question, we analyze the human influence on the RF signal. The analysis results show that as long as a person appears within the signal sensing range for identity authentication, his identity information and behavior information will be inevitably recorded at the same time.

In this paper, we aim to propose a 'once-deploy-forever-use' method to protect the behavior information in RF signals. This method enables users to 'disable the behavior recognition' by filtering their private behavior information out as much as possible from RF signals, yet still supporting RF-based user authentication with high accuracy.

However, achieving this goal is difficult due to the following challenges. First, the most intuitive way to protect behavior privacy is to separate identity information and behavior information, and then only retain the former. Nevertheless, directly separating the behavior and identity information is infeasible, because RF signals are not as intuitive as visual data (*e.g.*, image/video) and the exact function relationship between these two kinds of information is agnostic (detailed in Section 3.1). Second, it is unacceptable to degrade/ruin/erase the identity information too much when excluding the behavior information.

To overcome the first challenge, we transform the in-

- *Jianwei Liu is with the School of Cyber Science and Technology, Zhejiang University, China*
- *Chaowei Xiao is with Nvidia Research and Arizona State University, USA.*
- *Kaiyan Cui is with School of Software Engineering, Xi'an Jiaotong University, Xi'an, China, and the Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China.*
- *Jinsong Han (corresponding author) is with School of Cyber Science and Technology, Zhejiang University, China, and Key Laboratory of Blockchain and Cyberspace Governance of Zhejiang Province, China.*
- *Xian Xu is with Zhejiang University, China.*
- *Kui Ren is with School of Cyber Science and Technology, Zhejiang University, China, Key Laboratory of Blockchain and Cyberspace Governance of Zhejiang Province, China, and Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies, China.*

tractable separation problem into a simple similarity problem. We first dig out the reason why RF signals can be used for identity/behavior recognition. We find that signals are spatially distinguishable for both identity and behavior. Intrinsically, erasing the behavior information is equivalent to erase the distinguishability among the signals corresponding to different behaviors. One way to achieve this goal is to increase the similarity among the signals of different behaviors. Therefore, we design a *Siamese network*-based [15] deep neural network, named *Behavior Privacy Cloak* (*BPCloak*), to adjust the similarity to erase behavior information. To address the second challenge, we do not increase the similarity among the signals of different identities to maintain the distinguishability among different identities. Besides, we introduce an identity classifier in *BPCloak* to improve its identity feature extraction ability.

Specifically, in *BPCloak*, we treat the signals corresponding to different behaviors with the same identity as similar, and the signals corresponding to the same behavior with different identities as dissimilar. In this way, *BPCloak* can erase the behavior information without destroying the identity information. To further improve the identity feature extraction ability of *BPCloak*, we leverage an identity classifier to facilitate the optimization of *BPCloak*. For the real-world deployment, users first collect a batch of signal samples (each piece of data in RF signals is termed as a signal sample) and construct a training set. After training *BPCloak* using the training set as well as a reasonably-designed loss function, *BPCloak* can erase behavior privacy. *BPCloak* takes signal samples as input and outputs behavior-irrelevant identity feature vectors. Since such a feature vector does not contain behavior information but contains sufficient identity information, it can be directly made public or uploaded to the SP.

We conducted comprehensive experiments on different RF systems including a commercial off-the-shelf (COTS) WiFi system, a COTS RFID system, and a COTS mmWave system. These three systems represent omnidirectional RF technique, backscatter RF technique, and directional RF technique, respectively. The experiment results on the WiFi/RFID/mmWave system demonstrated that *BPCloak* can decrease the behavior recognition accuracy from 99%/95%/99% to 11%/18%/28%, while the accuracy of identity authentication only dropped 1%/3%/5%.

This paper makes the following contributions: ❶ We noticed the privacy issues in RF-based user authentication applications, which may lead to the leakage of user's behavior information. ❷ We propose *BPCloak* to erase behavior privacy in RF signals without hurting the performance of RF-based user authentication. ❸ We conduct comprehensive experiments on three representative RF systems. The experiment results demonstrate that *BPCloak* can effectively erase the behavior information in the signal while retaining high accuracy of identity authentication.

2 Privacy Concern in RF Signals

To answer the aforementioned naturally-raised question, we analyze the process of collecting RF signals and show the tight combination between the identity and behavior information in the collected signals. Then, we present a threat model and show how an adversary can compromise behavior privacy via *RF signals that have not been processed for behavior privacy preserving* (termed as vanilla RF signals).

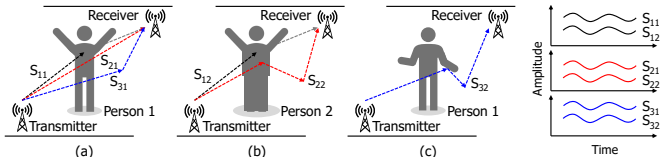


Fig. 1. Identity and behavior information are captured when the person is static.

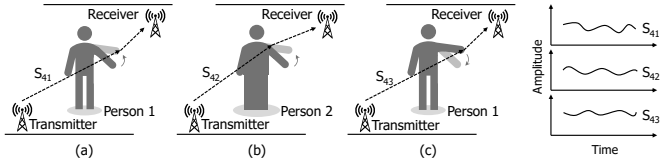


Fig. 2. Identity and behavior information are captured when the person is dynamic.

2.1 RF Signal Collection

When collecting RF signals, the influence of human body on the RF signals can be divided into static one and dynamic one according to whether the human body is static [4], [11] or dynamic [2], [16]. Both influences will inevitably embed identity information and behavior information into the signals. These two kinds of information lead to the distinguishability of signals in identity and behavior. Such distinguishability enables RF-based identity authentication and behavior recognition to be realized by means of feature vector extraction and machine learning classification [6], [17]. Particularly, such distinguishability can be reflected by the distinguishability of signal indicators, such as amplitude (higher amplitude means higher signal energy) and phase. Below, we take the distinguishability of amplitude as an example for illustration.

In the static case, *i.e.*, the user keeps still, the identity is distinguishable in terms of the body's biomaterial and shape. As shown in Fig. 1(a) and (b), when two different persons pose the same static gesture, the received signals are identity-distinguishable because 1) the biomaterials of different persons are distinguishable [1] and RF signals are sensitive to the material they pass through [18]; 2) the shapes of different persons are also distinguishable [19] and different shapes would cause different signal propagation paths by reflection [4], [20]. When the same signal experiences different persons (see S_{11} and S_{12} or S_{21} and S_{22} in Fig. 1) in different scenarios, its received amplitudes are distinguishable. As shown in Fig. 1(a) and (c), when the same person poses different gestures, the received signals (S_{31} and S_{32}) are gesture-distinguishable because different gestures also cause different amplitudes [11].

In the dynamic case, the user is performing certain activities. Besides the aforementioned biomaterial and shape, we involve behavioral patterns (which are referred to as behavioral biometrics [8], [21]). For two persons who perform the same activity, not only their biomaterials and body shapes will produce amplitude distinguishability, but their behavioral patterns of performing activities are also different from person to person [19], [22] (see S_{41} and S_{42} in Fig. 2). Similarly, when the same person performs two different activities (see S_{41} and S_{43}), different activities cause different signal

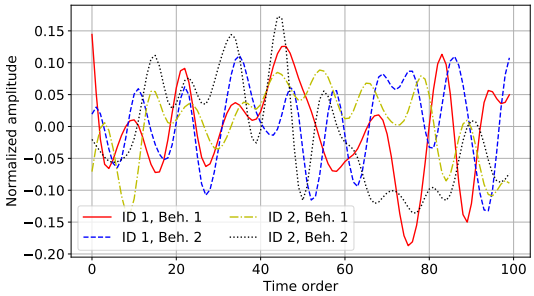


Fig. 3. Abstract and complex RF signals.

propagation paths [23] and further produce distinguishability on amplitude variations. Therefore, identity information and behavior information are also recorded by signals in the dynamic case. From the above analysis, we conclude that as long as a person appears in the sensing range of the signal, his identity and behavior information will inevitably and simultaneously be captured. As an application, WiHF [24] have confirmed the feasibility of using the same RF signal simultaneously for user identification and gesture recognition.

2.2 Threat Model and Attack Test

Threat model: The attack in this paper aims at mining behavior information from vanilla RF signals. Our threat model is a gray-box one. in which we assume that the adversary knows the feature extraction method of the user. This assumption is necessary and reasonable because: 1) Necessity. The attacker may only obtain the extracted features rather than the original signals. If the attacker trains a malicious classifier to achieve behavior inference, the category of the features used for classifier training (attacker's data) and testing (victim's data) should be consistent. 2) Rationality. The category (*e.g.*, statistical scalar or original signal values) of the features can be inferred according to the form of the features.

Attack test: After the adversary obtained the vanilla signals, he can leverage some techniques (*e.g.*, machine learning classifier) to speculate the behaviors corresponding to the signals. To show the hazard of such attacks, we conduct an attack experiment. We first ask a volunteer (victim) to performs six activities introduced in [25] in an environment to collect a batch of signal samples. Then, we invite another volunteer (adversary) to collect another batch of signal samples in another environment in the same way. After extracting frequency features [25] from the two batches, we train a logistic regression (LR) classifier [26] with the adversary's features and test the classification accuracy with the victim's features. As a result, the activity classification accuracy is 64%+. This result demonstrates the feasibility of using machine learning classifiers to mine the behavior privacy from vanilla signals. Moreover, Zheng *et al.* [2] show that when the adversary extracts domain-independent features from vanilla signals, the malicious behavior classification accuracy can reach 92.7%, even if the adversary's signals and the victim's signals are collected by different persons in different environments.

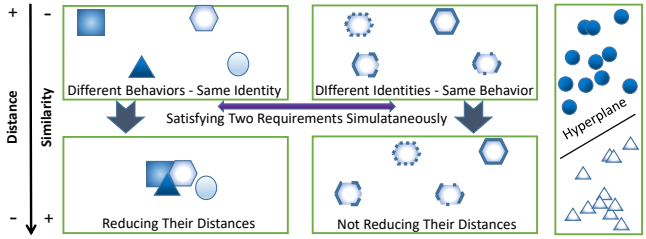


Fig. 4. Reducing the distances among the signals of different behaviors but the same identity while not reducing the distances among the signals of different identities but the same behavior.

Fig. 5. Illustration of SVM.

3 Paving The Way For Privacy Preserving

In this section, we first introduce the reason why we use the deep learning technique to achieve behavior privacy preserving. Then, we transform the distinguishability problem into a similarity problem, which enables us to leverage *Siamese network* to change the similarity among signal samples.

3.1 Why Deep Learning?

To protect behavior privacy in vanilla signals, intuitive methods are to directly separate them via visual observation or numerical analysis (*e.g.*, parameters estimation [27]). However, these methods are infeasible due to the following reasons: 1) RF signal is not as intuitive as image/video. Fig. 3 shows four amplitude traces of two persons (ID) and two gestures (Beh). Apparently, one cannot figure out which part of the signal represents the identity or behavior information. 2) To separate the two kinds of information through numerical analysis, we need to build a function relationship between the two kinds of information. According to [28], such function relationship can be formulated as:

$$H = \sum_{k=1}^{P} (a_k^i + a_k^b) e^{-j2\pi f(\tau_k^i + \tau_k^b)}, \quad (1)$$

where H is the channel state information that describes how signal propagates from transmitters to receivers. P is the number of multi-paths. a_k^i and τ_k^i are identity information-related amplitude and propagation delay, respectively. a_k^b and τ_k^b are behavior information-related amplitude and propagation delay, respectively. Since the signal is invisible and P cannot be measured by any devices, P is unknown in reality. Further, the exact function relationship between the two kinds of information is unknown, not mention to separate them through numerical analysis. Fortunately, we notice that a deep neural network can be trained to fit any complex function. Thus, it is possible to leverage the deep neural network to learn the complex function relationship between the two kinds of information to separate them.

3.2 From Distinguishability To Similarity

As mentioned in Section 2.1, existing identity/behavior recognition techniques dominantly leverage the distinguishability among signals, extracted feature vectors, and classifiers to achieve high accuracy. Intrinsically, removing the behavior information in RF signals is to make the feature vectors extracted from the signal samples associated with different behaviors no longer be distinguishable in behavior. To this end, as shown in Fig. 4, we can narrow

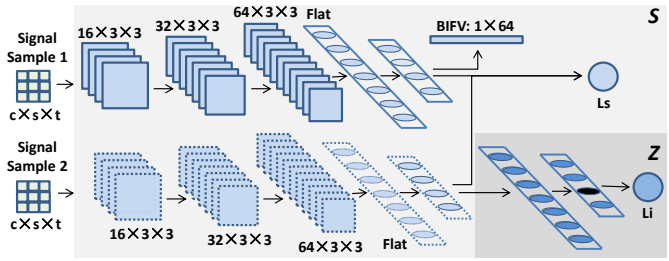


Fig. 6. Architecture of *BPCloak*.

the spatial distances among the feature vectors associated with different behaviors to make it inseparable, because classifiers rely on this spatial distance to classify [29]. For example, as shown in Fig. 5, support vector machine (SVM) will construct a hyperplane between different classes at a distance and then classify the samples according to which side of the hyperplane the samples are on. However, more importantly, the ability of the feature vector to be used for identity authentication cannot be excessively affected, *i.e.*, the distances among the feature vectors associated with different identities cannot be reduced (see the middle part of Fig. 4). From the perspective of similarity, distance can be characterized by similarity: small distance means high similarity while large distance means low similarity. Thus, it can be derived that eliminating behavior distinguishability is equivalent to improving the similarity among feature vectors associated with different behaviors; retaining identity distinguishability means not improving the similarity among feature vectors associated with different identities.

We notice that *Siamese network* has outstanding ability in terms of similarity control. Therefore, we design a *Siamese network*-based deep neural network named *BPCloak* to solve the privacy-preserving problem. In the next section, we will detail *BPCloak*'s design, training set construction method, and optimization process.

4 Privacy Preserving Approach

4.1 *BPCloak* Design

The architecture of *BPCloak* is shown in Fig. 6. *BPCloak* is composed of *BPCloak-S* and *BPCloak-Z*. *BPCloak-S* is a feature extraction network, the basic structure of which is *Siamese network*. It is used to extract behavior-irrelevant identity feature vectors. *BPCloak-Z* is an auxiliary identity classifier. It is utilized to improve the identity feature extraction ability of *BPCloak*. Here, we introduce the architecture in detail through the data flow. The inputs of this model are twofold, *i.e.*, pairwise vanilla signal samples $\vec{x_1}$ and $\vec{x_2}$ are fed into *BPCloak* simultaneously. Then, $\vec{x_1}$ and $\vec{x_2}$ respectively pass through two five-layer convolutional branches CB_1 and CB_2. These two convolution branches form the whole structure of *BPCloak-S*. They have the same structure and share parameters. Taking branch CB_1 as an example, its first three layers are convolutional layers and each of them consists of three sub-functions, *i.e.*, convolution operation, batch normalization, and rectified linear unit (ReLU) [30]. Among them, the convolution operation is used to extract features from signal samples. The convolution kernel is two-dimensional, and the convolution operation will extract both the spatial feature between sub-channels and the temporal feature between adjacent sampling points. Batch normalization is used to avoid the offset of data distribution and away from the derivative saturation zone. ReLU is used to Increase the non-linearity of the network to help the network complete the complex filtering task. The remaining two layers of CB_1 are fully-connected ones. The output of the third convolutional layer will be flattened as a vector before being fed into the fully connected layer. The first fully-connected layer is followed by a Sigmoid activation function [30] to increase the non-linearity between fully connected layers. A dropout layer is arranged behind the Sigmoid function. The dropout layer will nullify some neurons in the fully-connected layer, preventing the overfitting on training set [31]. Without the dropout layer, the learning on the former training samples in the training set may be disproportionately, and the features that appear only in later training samples may not be learned [31]. Thus, the dropout layer helps improve the generalization ability of the trained network. Afterwards, the output of the first fully-connected layer is fed into the second fully-connected layer and becomes $\vec{v_1}$. The dimensionality of $\vec{v_1}$ is $1 \times N_F$, where N_F is the number of neurons in the second fully-connected layer. N_F is set as 64 by default.

While $\vec{x_1}$ has undergone five layers of operations, $\vec{x_2}$ has undergone the same operations and is transformed into $\vec{v_2}$. $\vec{v_1}$ and $\vec{v_2}$ are behavior-irrelevant feature vectors (BIFVs). BIFV can be only utilized to identify users and it contains no information about users' behaviors. Henceforth, $\vec{v_1}$ and $\vec{v_2}$ undergo different operations to calculate different losses. In the first operation, $\vec{v_1}$ and $\vec{v_2}$ are simultaneously used to calculate *similarity loss* (will be introduced in Section 4.3). In the second operation, $\vec{v_1}/\vec{v_2}$ will pass through *BPCloak-Z*, *i.e.*, two fully-connected layers. The output of this layer $\vec{v_p}$ is a probability vector, the elements of which are the probabilities that $\vec{x_1}/\vec{x_2}$ belongs to each user. $\vec{v_p}$ and the correct identity label of $\vec{x_1}/\vec{x_2}$ are then used to calculate *identity loss* (will be introduced in Section 4.3). It enables *BPCloak* to extract high-quality identity features from signals.

When user wants to extract the BIFV of a vanilla signal sample $\vec{x}$, he only needs to let $\vec{x}$ pass through CB_1 or CB_2. Since CB_1 and CB_2 have the same parameters, we have $CB_1(\vec{x}) = CB_1(\vec{x})$, and $CB_1(\vec{x})$ or $CB_2(\vec{x})$ can be used as BIFV. The source code of *BPCloak* is opened[1].

4.2 Training Set Construction

To empower *BPCloak* with the ability to erase behavior information,, we need to construct a training set that includes behavior information. As aforementioned, as long as users appear within the signal's sensing range, their identity information and behavior information will necessarily and concurrently be captured. The simplest and effective way to collect signals that meet the requirement is to let users do sensitive behaviors that require protection while they are within the sensing range. In this way, each collected signal sample has two labels: an identity label and a behavior label. Additionally, according to our analysis of 'distinguishability-to-similarity', *BPCloak* needs to change the similarity between two feature vectors, which requires

1. https://github.com/DragonflyCaptainL/RF_behavior_privacy_preserving.git

Fig. 7. Training set construction method.

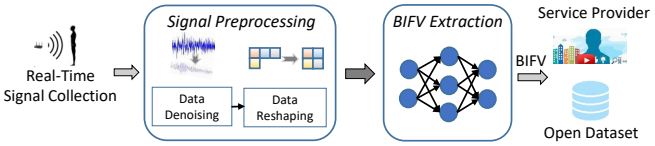


Fig. 8. Pipeline for *BPCloak* deployment.

transferring the behavior label and identity label into a similarity label.

As shown in Fig. 7, we are interested in two types of pairs (each pair is a training sample): 1) two signal samples with the same identity but different behaviors and 2) two signal samples with the same behavior but different identities. We set the similarity label of the first type of pair to '0' (similar) and set that of the second type of pair to '1' (dissimilar). In this way, *BPCloak* will learn to increase the similarity of the extracted feature vectors with different behaviors and the dissimilarity among extracted feature vectors of different identities will not be reduced. To further boost the extracted identity information contained in the extracted feature vectors, we also leverage the original identity labels of signal samples. Finally, each training sample in the training set has three labels: a similarity label and two identity labels.

4.3 Loss Function and Optimization Process

Given a training set containing n training samples, we form p batches and each batch contains $\frac{n}{p}$ training samples. Each batch is fed into *BPCloak* individually and the losses are calculated based on the outputs. We calculate losses with a unit of batch. For the *similarity loss*, it can be formulated as:

$$\begin{aligned}\mathcal{L}_s =&(1-Y_s)*(D_W(B^S(\vec{x_1}),B^S(\vec{x_2})))^2+\\ &Y_s*(max\left\{0,margin-D_W(B^S(\vec{x_1}),B^S(\vec{x_2}))\right\})^2.\end{aligned} \tag{2}$$

In this formula, $\vec{x_1}$ and $\vec{x_2}$ are two signal samples that belong to the same training sample. Y_s is the similarity label and $margin$ is a distance threshold. $B^S(\cdot)$ represents *BPCloak-S* and $D_W(B^S(\vec{x_1}),B^S(\vec{x_2}))$ is the *Euclidean distance* between the BIFVs $\vec{v_1}$ and $\vec{v_2}$, where $\vec{v_1}=B^S(\vec{x_1})$ and $\vec{v_2}=B^S(\vec{x_2})$. The $margin$ is set as 3 empirically. As for the *identity loss*, we opt to use *cross-entropy loss* [32]. This is because our user authentication is indeed a multi-class classification task, and *cross-entropy loss* is the most popular loss used in classification tasks [33]. It can be formulated as:

$$\mathcal{L}_i=-\sum_{c=1}^{M}y_c\log(P_c), \tag{3}$$

in which y_c is the indication variable, P_c is the probability that the signal sample belongs to identity label c and M is the maximum of the identity labels. Ultimately, the final loss consisting of the *similarity loss* and *identity loss* can be formulated as:

$$\mathcal{L}_f=\alpha*\mathcal{L}_s+\frac{(1-\alpha)}{2}*\mathcal{L}_i^1+\frac{(1-\alpha)}{2}*\mathcal{L}_i^2,\quad \alpha\in(0,1), \tag{4}$$

where $\mathcal{L}_i^1$ and $\mathcal{L}_i^2$ are the identity losses of the first and the second signal samples in the training sample, respectively.

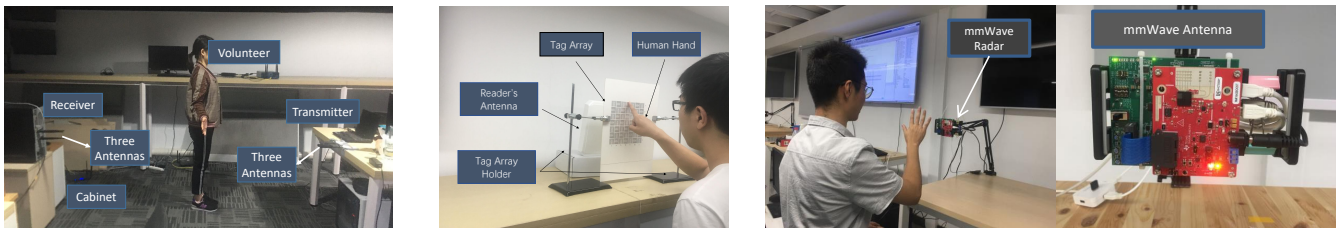
4.4 Privacy-Preserving Pipeline

As shown in Fig. 8, the privacy-preserving pipeline is composed of two modules: *signal preprocessing* module and *BIFV extraction* module. it is deployed at the receiver end of the identity authentication system. The received signal streams are first segmented to have N_t timestamps, and then preprocessed by the *signal preprocessing* module. The preprocessing includes data denoising and data reshaping. The former is used to remove high-frequency components (noise) from raw signals. For example, researchers usually utilize a low-pass filter to remove high-frequency noise in WiFi signals [10]. The latter is used to provide dimensionality-constant signal samples to the next module. The reshaped signal sample will have dimensionality of (N_i, N_s, N_t), where N_i and N_s represent the number of signal indicator types and the number of signal streams, respectively. For example, in an RFID system with N_{tag} tags, the dimensionality of a signal sample is $(2, N_{tag}, N_t)$, where 2 means received signal strength (RSS) plus phase. Then, the obtained signal sample is inputted into the *BIFV extraction* module. The output, i.e., the BIFV, can be uploaded to the server or made public. The challenge here is that the use of deep neural network may impact the real-time performance of the authentication system. To deal with this problem, users can use model compression technique [34] to greatly reduce the time consumption of computation.

5 Evaluation and Result

In this section, we conducted experiments on three real-world systems (*i.e.*, a COTS WiFi system, a COTS RFID system, and a COTS mmWave system). The condition for the adversary to achieve the best-case attack is that the adversary can use a part of the user's signal samples and corresponding behavior labels to train the malicious classifier. Therefore, to effectively evaluate our privacy-preserving method, all experiments were carried out under this condition.

Experiment setup: 1) WiFi system: we invited ten volunteers (two females and eight males) to perform ten gestures in the WiFi sensing range. The ages of volunteers varied from 22 to 35 and their heights varied from 160cm to 188cm. The experiment setup is shown in Fig. 9(a), the transmitter equipped with an *Atheros 9380* network interface card (NIC) was placed 2m away from the receiver (also equipped with an *Atheros 9380* NIC). Both the transmitter and the receiver were placed on wooden furniture, the top surfaces of which were 80cm off the ground. Ten gestures representing ten numbers from zero to nine are shown in Fig. 10. When posing gestures, the volunteer was standing between the transmitter and the receiver. We totally collected over 29000 signal samples in the WiFi system. 2) RFID system: we invited five volunteers (two females and three males) to perform ten activities. The ages of the volunteers varied from 21 to 31 and their heights varied from 165cm to 188cm. As shown in Fig. 9(b), the tag array was formed by 49 tags (*Alien-9629*), and the reader's antenna (*Impinj R420+Larid*



(a) WiFi. (b) RFID. (c) Millimeter wave.

Fig. 9. Experiment setups of three systems.

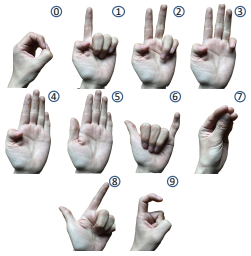


Fig. 10. Ten gestures: 0 to 9.

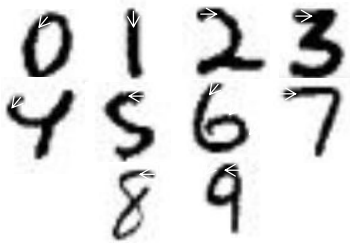


Fig. 11. Ten activities: 0 to 9.

A9028) is placed 10cm away from the tag array. The activities of zero-to-nine are shown in Fig. 11. The position of the white arrow is the starting position of writing and its direction is the starting direction of writing. Volunteers were asked to write ten numbers in front of the tag array. We totally collected over 4000 signal samples in the RFID system. 3) Millimeter wave system: we invited nine volunteers (seven males and two females) to pose ten gestures shown in Fig. 10. The ages of the volunteers varied from 22 to 24 and the heights of them varied from 160cm to 175cm. As shown in Fig. 9(c), We asked volunteers to sit in front of the mmWave radar (*IWR1642*) to pose gestures. The radar was connected to one transmission antenna and four receiving antennas. We totally collected 9000 signal samples in the mmWave system.

Data preprocessing: The received WiFi signals were CSI data, the elements of which were complex numbers. We first utilized a low-pass filter (with a cutoff frequency of 20Hz) to remove the high-frequency noise in the CSI data. Then, we calculated the absolute value of the CSI data to obtain the amplitude of the CSI. Afterwards, we segmented the time-series data of each sub-channel for each gesture so that each signal sample (which is associated with a specific gesture and a specific volunteer) had dimensionality of (1,504,10). In the RFID system, the received signals were also time-series data for each tag. We first selected the first 30 sampling points for each tag and then concatenated them together. In this way, each signal sample had dimensionality of (2,49,30). In the mmWave system, we also calculated the absolute values (*i.e.*, amplitudes) of collected signals. The dimensionality of each signal sample is (1,4,1024).

Metrics: We defined three metrics to quantify the attack effectiveness of the behavior privacy mining and the privacy-preserving performance of *BPCloak*: accuracy, defense rate (DR), and trade-off rate (TR). The accuracy is the probability that the identity/behavior label of any signal sample is correctly identified. It is termed as authentication accuracy (AA) in identity authentication, and recognition accuracy (RA) in behavior recognition. It can be formulated as: $accuracy = \frac{N_{cor}}{N_{all}}$, where N_{cor} is the number of correctly classified test signal samples and N_{all} is the number of all test samples. DR is the ratio of the RA that *BPCloak* reduces to the RA that the adversary can achieve. The larger the DR, the better the privacy-preserving performance of *BPCloak*. DR can be formulated as:

$$DR = \frac{RA_{att} - RA_{def}}{RA_{att}}, \tag{5}$$

where RA_{att} is the RA of vanilla signal samples and RA_{def} is the RA of extracted BIFVs. TR is the ratio of the AA that *BPCloak* loses during the privacy preserving to the AA of the vanilla signal samples. The smaller the TR, the better the performance of *BPCloak*. TR can be formulated as:

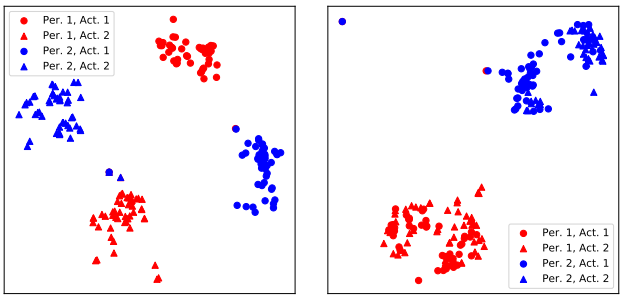
$$TR = \frac{AA_{ori} - AA_{def}}{AA_{ori}}. \tag{6}$$

In this formula, AA_{ori} is the AA of vanilla signal samples and AA_{def} is the AA of extracted BIFVs.

5.1 Overall Performance

For the sake of training *BPCloak*, we first constructed a training set with one thousand randomly-selected training samples for each system. Then, in order to ensure that the experiment results will not be affected by the randomness of the training set construction method, we train the *BPCloak* by using these training sets and their subsets in the following experiments. At the same time, to ensure that the signal samples involved in the *BPCloak* training and the signal samples used to extract BIFVs do not overlap, the data we used to extract BIFV consists of other signal samples that are not in the training set. Hereinafter, the vanilla signal samples used to extract BIFVs are called the data to be protected (DoP). The RA_{att} and AA_{ori} in Eq. 5 and Eq. 6 are the RA and AA of DoP.

t-SNE validation: To intuitively show that the distinguishability of behavior is eliminated while the distinguishability of identity is well retained, we try to visualize the DoP and BIFVs. To this end, we first use t-SNE [35] to reduce



(a) DoP. (b) BIFV.

Fig. 12. Distributions of WiFi DoP and BIFVs. Different colors represent different identities and different shapes represent different behaviors.

TABLE 1
Authentication accuracy (AA) and recognition accuracy (RA) of WiFi DoP and BIFV on different classifiers.

Classifier	RFo	LR	KNN	SVM	DT	NB	NN
AA-DoP	99.84%	99.96%	99.63%	99.56%	97.83%	98.39%	99.79%
RA-DoP	15.21%	19.09%	73.86%	83.80%	75.54%	73.93%	93.69%
AA-BIFV	99.45%	99.40%	99.35%	99.37%	99.24%	99.34%	97.69%
RA-BIFV	10.49%	10.87%	10.51%	10.66%	10.42%	10.95%	10.02%

TABLE 2
Authentication accuracy (AA) and recognition accuracy (RA) of RFID DoP and BIFV on different classifiers.

Classifier	RFo	LR	KNN	SVM	DT	NB	NN
AA-DoP	99.60%	95.44%	99.50%	93.35%	62.60%	63.89%	100.00%
RA-DoP	95.36%	53.10%	95.63%	94.64%	48.12%	27.38%	89.78%
AA-BIFV	97.63%	97.71%	97.68%	97.74%	97.00%	95.81%	76.20%
RA-BIFV	17.95%	12.45%	15.37%	12.27%	17.44%	11.46%	10.75%

TABLE 3
Authentication accuracy (AA) and recognition accuracy (RA) of mmWave DoP and BIFV on different classifiers.

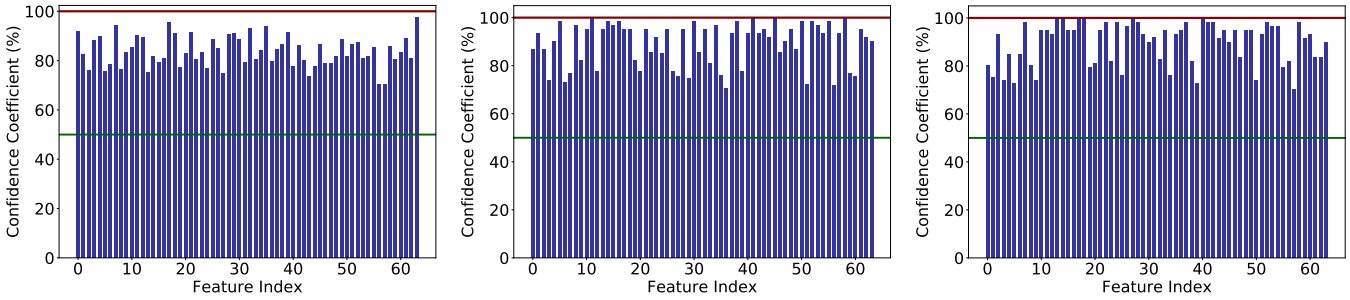
Classifier	RFo	LR	KNN	SVM	DT	NB	NN
AA-DoP	77.36%	99.57%	59.33%	99.97%	36.27%	78.24%	99.70%
RA-DoP	38.92%	86.41%	32.04%	93.15%	18.81%	38.68%	93.68%
AA-BIFV	96.05%	94.94%	94.57%	95.18%	93.62%	84.79%	94.71%
RA-BIFV	27.22%	9.62%	22.24%	19.47%	21.42%	10.52%	10.19%

the dimensionality of the DoP and BIFVs so that they only have two elements. Afterwards, we plot the dimensionality-reduced DoP and BIFVs of the WiFi system in Fig. 12(a) and (b), respectively. The t-SNE results of the RFID and mmWave systems are similar to those of the WiFi system, and are not redundantly shown here. It is apparent that both the behavior distinguishability and the identity distinguishability in DoP are high. In BIFVs, the distinguishability of identity can still be observed but the distinguishability of behavior has been greatly reduced. Such results demonstrate that *BPCloak* can effectively erase the behavior information while retaining the identity information.

RA and AA of DoP: We first calculated the RA and AA of DoP. It should be noted that in the WiFi system, RA is closely related to the person [2], *i.e.*, the mixed signal samples of multiple people cannot be used for multi-person behavior recognition with high accuracy. So, in the WiFi experiment, we calculated the RA of each person individually and then calculated the mean of these RA. We used mainstream machine learning classifiers that existing identity authentication and behavior recognition works have used and proved effectively: random forest (RFo), logistic regression (LR), k-nearest neighbors (KNN), SVM, decision tree (DT), naive Bayes (NB), and three-layer neural network (NN). We did not test on deep learning models containing convolutional layers because convolutional layer is used to extract features from input, but BIFVs are already the extracted identity features. In all experiments, we randomly selected 75% of the DoP/BIFVs to train the classifier and used the remaining 25% of the DoP/BIFVs to calculate the accuracy. This process was repeated ten times and we took the average of ten accuracy as the final result. The experimental results of the WiFi system, the RFID system, and the mmWave system are shown in the top two rows in Tab. 1, 2, and 3, respectively. The results of the WiFi system indicate that the best attack effectiveness (the highest RA) exceeds 99%. Meanwhile, an SP can use the DoP to provide an AA of 99%+. In both RFID and mmWave systems, an SP can use the DoP to provide high AA (99%+). The adversary's possible attack effectiveness of the RFID system and mmWave system exceed 94% and 99%, respectively. These experiment results show that although SP can use the DoP to provide users with accurate identity authentication, an adversary may use the DoP to monitor their behaviors.

RA and AA of BIFV: Then, we evaluated the RA and AA of the BIFVs extracted from the DoP. The results of the WiFi, the RFID, and the mmWave systems are shown in the bottom two rows in Tab. 1, 2, and 3, respectively. It can be observed that the AA of BIFVs are not much lower than that of the DoP, or even increase on some classifiers. But RA is greatly reduced on all classifiers. The increase of AA is rational because *BPCloak* helps extract clear identity information from DoP. The identity features in BIFVs could be more prominent than those in DoP. The RA in the WiFi system even approximates random guess (10%). Therefore, the identity information is well retained in the BIFV, so an SP can still provide user authentication with high accuracy. At the same time, the adversary cannot achieve decent attack effectiveness even under the best attack condition.

Chi-Square test: In order to verify that BIFVs are irrelevant to the behavior information from a statistical perspective, we performed a Chi-Square independence test towards each element position in BIFV (64 positions by default). Specifically, we removed every position of the element in BIFV alternatively and examined the effect of BIFV on RA after removing the element at that position. Our hypothesis to be validated is: the element at this position is independent of RA, *i.e.*, the element at this position is irrelevant to the behavior information. In each system, we randomly selected 500 BIFVs and counted the number of BIFVs that are correctly and incorrectly classified towards behavior before and after removing an element. After removing an element, we train and test the classifiers with the BIFVs containing 63 elements. With a degree of freedom of one, we show the confidence coefficients of the three systems to accept the hypothesis in Fig. 13(a), (b), and (c), respec-



(a) WiFi. (b) RFID. (c) Millimeter wave.

Fig. 13. Confidence coefficients for accepting the hypothesis that BIFV is irrelevant to the behavior information. The green horizontal line means 50% and the red horizontal line means 100.00%.

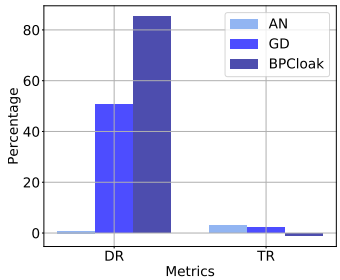


Fig. 14. DR and TR of the WiFi system.

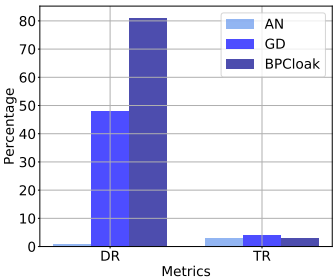


Fig. 15. DR and TR of the RFID system.

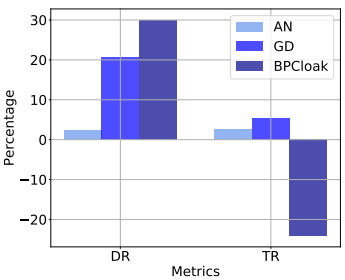


Fig. 16. DR and TR of the mmWave system.

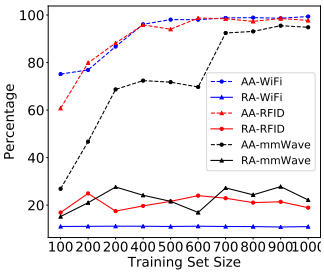


Fig. 17. Effect of training set size.

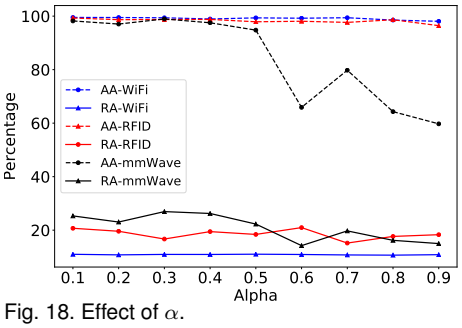


Fig. 18. Effect of α.

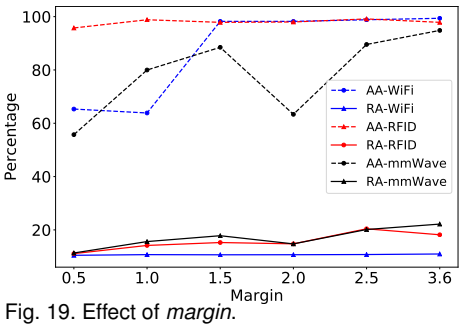


Fig. 19. Effect of *margin*.

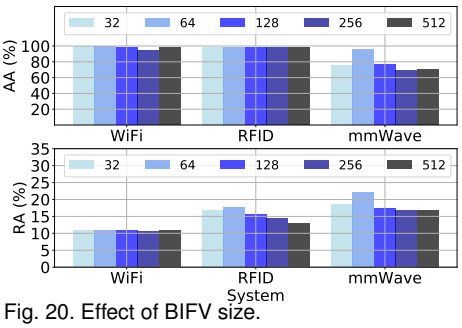


Fig. 20. Effect of BIFV size.

tively. The experiment results show that all the confidence coefficients are larger than 50% or even 70%. Some of the confidence coefficients even approximate 100%. This result demonstrates that we have a high degree of confidence to accept the hypothesis that the BIFV is irrelevant to the behavior information. Therefore, *BPCloak* has effective behavior privacy-preserving capabilities and it is extremely difficult for an adversary to recover the behavior privacy.

DR and TR: According to the experiment results in Table 1, 2, and 3, NB has the highest RA on WiFi BIFVs, and RFo has the highest RA on RFID and mmWave BIFVs. Therefore, we used NB and RFo to evaluate the DR and TR of *BPCloak*. We compared *BPCloak* with two other deep learning-based baselines. One is the adversarial network (AN) [10]. Its principle is to use a generator to extract BIFV, the other two discriminators are used to reduce the loss of identity authentication and increase the loss of behavior recognition. The other one is gradient descent (GD) [36]. We first randomly generate a perturbation, and then update the perturbation based on reducing the loss of identity recognition and increasing the loss of behavior recognition. The comparison results of the WiFi system, the RFID system, and the mmWave system are shown in Fig. 14, 15, and 16, respectively. The experiment results indicate that no matter it is in the WiFi system, RFID system, or mmWave system, the DR of *BPCloak* is higher than that of other two baselines. This demonstrates that *BPCloak* can effectively erase the behavior information in DoP and it outperforms the other two baselines. At the same time, in the three systems, the TRs of *BPCloak* are lower than that of other two baselines. In the WiFi and mmWave systems, the TRs of *BPCloak* are even negative, which means that *BPCloak* performs better than the other two baselines in identity information retention. More importantly, *BPCloak* can effectively retain the identity information in DoP and even make it more prominent.

5.2 Effect of Training Set Size

In order to explore the effect of the training set size (*i.e.*, the number of training samples in the training set) on the performance of *BPCloak*, we varied the training set size from 100 to 1000 with a stride of 100. The AA and RA of the BIFVs of the three systems are shown in Fig. 17. It can be seen from the experiment results that the variation trends of the AA and RA are basically the same in the three systems. RA basically remains stable. AA increases as the training set size increases. When the size of the training set is 700, the AA curves of the three systems become flat. Henceforth, the AA of the BIFVs of the three systems are all significantly high and the RA is low. Therefore, only a small training set is needed to train an effective *BPCloak*.

TABLE 4
Authentication accuracy (AA) and recognition accuracy (RA) of *Widar3.0* dataset. The value in () is TR.

Person	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Person 7	Average	AA
RA-DoP	70.00%	88.89%	83.33%	100.00%	87.50%	88.89%	89.30%	86.84%	86.63%
RA-BIFV	18.18%	17.27%	20.91%	15.45%	17.88%	17.27%	18.79%	17.88%	71.13%
DR	74.03%	80.57%	74.91%	84.55%	79.57%	80.57%	78.96%	79.02%	(12.12)%

TABLE 5
Authentication accuracy (AA) and recognition accuracy (RA) of *Wiar* dataset. The value in () is TR.

Person	Person 1	Person 2	Person 3	Person 4	Person 5	Average	Authentication Accuracy
RA-DoP	98.32%	96.22%	98.00%	95.14%	91.38%	95.88%	100%
RA-BIFV	5.43%	6.45%	6.59%	6.03%	6.20%	6.14%	100.00%
DR	94.48%	93.32%	93.28%	93.66%	93.27%	93.59%	(0.00%)

5.3 Transferability Study

In this experiment, we explored the transferability of *BPCloak* in terms of unseen identities and behaviors. To be specific, we first used the trained *BPCloak* to extract the BIFVs of the identities (unseen identities) not participating in the training and then classified these BIFVs. Then, we used the trained *BPCloak* to extract the BIFVs of the behaviors (unseen behaviors) not participating in the training and recognized them. In the identity experiment, we used two identities as unseen identities and used the signal samples of other identities to train *BPCloak*. The AA of unseen identities are 99.13%, 82.14%, and 100.00% in the WiFi system, RFID system, and mmWave system, respectively. The AA of the RFID system is relatively low because the training set used to train *BPCloak* has only three persons' data, which makes *BPCloak* not learn the expected capability for identity information retention. Nevertheless, the results still prove that *BPCloak* has outstanding transferability for new users. In the behavior experiment, we used the signal samples of the first five behaviors to train *BPCloak* and then extract the BIFVs of the signal samples of the other five behaviors. The experiment results show that the RA of unseen behaviors are 23.60%, 34.58%, and 69.99% in the WiFi, the RFID, and the mmwave systems, respectively. The results show that *BPCloak* still has a strong ability to erase the privacy of unseen behaviors in the WiFi system and RFID system. In the mmWave system, *BPCloak* has relatively weaker privacy protection capabilities compared with the other two systems. This result shows that *BPCloak* also has outstanding transferability for unseen behaviors.

5.4 Ablation Study

In this part, we explored the effect of hyperparameters α, *margin*, and feature vector size (the number of elements in BIFV) on the privacy-preserving capability of *BPCloak*.

Effect of α: In this experiment, we varied the value of α from 0.1 to 0.9 in steps of 0.1. The experiment results in Fig. 18 show that the variation of α has no obvious effect on the RA of BIFV. However, a smaller α can better retain identity information. In the mmWave system, a large α will make the AA of the BIFV unstable. Hence, 0.2 is a better choice for α.

Effect of *margin*: In this experiment, we varied the value of *margin* from 0.5 to 3.0 in steps of 0.5. The experiment results in Fig. 19 show that the increase of $margin$ will increase both the AA and the RA of BIFV. However, the $margin$ has a much larger effect on AA than RA. In order to achieve a significantly high AA, the $margin$ is best to take 3.0.

Effect of BIFV size: In this experiment, we selected five commonly used feature sizes: 32, 64, 128, 256, and 512. The AA and RA are shown in the upper part and the lower part of Fig. 20, respectively. It can be found that in the WiFi and RFID systems, the variation in feature size does not have an obvious effect on AA. In the mmWave system, as the feature vector size increases, the AA of BIFV first increases and then decreases. The maximum value is reached when the feature size is 64. As for the RA, it does not change markedly in the WiFi system. But it increases first and then decreases in both RFID and mmWave systems. It also reaches the maximum when the feature size is 64. However, even the maximum value is only close to 20%. Therefore, if in order to provide high-quality identity authentication service as much as possible while protecting behavior privacy acceptably, 64 is the best choice.

5.5 Evaluation on Open-Source Dataset

In addition to the dataset we collected, we also used two open-source datasets *Widar 3.0* [2] and *Wiar* [37] to evaluate *BPCloak*.

***Widar3.0* dataset** is an open WiFi dataset. It was collected by using COTS WiFi devices equipped with *Intel 5300* NIC. Since it is a cross-domain dataset, we only used 261 signal samples of six gestures of seven volunteers that belong to the same domain. We utilized the classifiers used in the previous experiments to calculate the AA and RA of the DoP and BIFVs, and show the maximum values of each person in Tab. 4. It can be seen from the table that the average value of all DRs approximate 80%, and the TR value is only 12%. This result indicates that *BPCloak* protects behavior privacy significantly well, while retaining the ability of *Widar3.0* to be used for identity authentication. The TR here is not as high as the ones in the previous experiments because the training data is limited and *BPCloak* is not fully trained. But even so, the AA of the BIFV is only 15.50% smaller than that of the DoP. Therefore, *BPCloak* performs well on the *Widar3.0* dataset.

***Wiar* dataset** contains the signal samples of ten persons and 16 activities. It was also collected by using *Intel 5300* NIC. Since the RA of five persons' signal samples in the dataset are low, we only used the 2601 signal samples of the other five persons to evaluate *BPCloak*. The experiment method here is consistent with *Widar3.0*'s experiment method. The experiment results are shown in Tab. 5. It can be found that every person's DR exceeds 90%, and the TR is 0.00%. This indicates that the extracted BIFV retains the identity information significantly completely, and at the same time completely erases the behavior information. It is

TABLE 6
Overhead and time cost of BPCloak.

System	WiFi	RFID	mmWave	MN2
FLOPs (M)	56.97	3.23	22.09	∅
Para. (M)	10.37	0.51	3.26	3.47
Time (ms)	9.5	2.6	4.6	33.3−

reasonable that the results of this experiment are better than those of *Widar3.0*, because the number of signal samples used to construct the training set of *Wiar* is almost ten times that of *Widar3.0*. In short, *BPCloak* also performs well on *Wiar* dataset.

5.6 Storage Overhead and Time Cost

In this part, we evaluated the practicability and real-time performance of our privacy-preserving pipeline from two aspects: computational overhead and storage overhead. Since the other parts of the pipeline except *BPCloak* have negligible computational and storage overhead, we mainly evaluated the overhead of *BPCloak*. We counted the number of floating-point operations (FLOPs), the numbers of parameters, and the average time required to extract a BIFV of a signal sample in the three systems by using a 2.8GHz *i7* CPU. The smaller the number of FLOPs, the smaller the computational overhead. The smaller the number of parameters, the smaller the storage overhead. The smaller the average time, the better the real-time performance. Meanwhile, we compare *BPCloak* with *MobileNet V2* [38]. *MobileNet V2* is a deep neural network that can run smoothly on mobile devices (*e.g.*, iPhone6s). The comparison results are shown in Tab. 6 ('MN2' is *MobileNet V2*). It can be observed from the results that *BPCloak*'s overhead on the WiFi system is the largest, and the number of its parameters is 2.99 times that of *MobileNet V2*. But its FLOPs and parameter number are still acceptable. For the RFID and mmWave systems, their numbers of parameters are less than that of *MobileNet V2*. From a time-consuming perspective, *BPCloak* can achieve excellent real-time performance, because *MobileNet V2* has good real-time performance, and the time to extract a BIFV from each system is much less than the time in *MobileNet V2*. Therefore, *BPCloak* is computational-friendly and has outstanding real-time performance.

5.7 Scalability Study

In this part, we evaluated *BPCloak*'s performance (i.e., scalability) towards behavior privacy preserving in WiFi-based human localization. We first asked a volunteer to do ten activities representing the numbers from zero to nine at five different locations. We collected 50 signal samples for each activity per location. Then, we constructed a training set with 1000 training samples. To quantify the localization performance, we define localization accuracy (LA), which represents the probability that the location of a signal sample is correctly identified. The LA and RA of DoP and BIFV are shown in Tab. 7. It can be observed that the average RA of DoP is as high as 85.0%. When we extract BIFVs from DoP, the average RA decreases to 10.3%. The DR is higher than 87%, demonstrating that *BPCloak* can effectively protect users' behavior privacy in WiFi signals used for localization. Meanwhile, we can find that the LA only decreases from 96.5% to 76% after *BPCloak* processing, and the TR is only 21.1%. Thus, *BPCloak* can effectively retain the location information in the signals, so that BIFV can be used for accurate localization.

6 Related Work

This work is mainly related to two kinds of techniques: RF signal-based identity authentication techniques and privacy-preserving techniques.

RF signal-based identity authentication: Existing authentication technologies based on RF signals generally follow the same authentication mode, that is, first extracting feature vectors, and then using learning-based classifiers to distinguish feature vectors. For example, RF-Rhythm [39] extracts the user's tapping rules on the tag array as a feature vector, and uses three classifiers (*i.e.*, NN, SVM, and CNN) for identity recognition. WiPIN [4] extracts the human body features in the WiFi signal as a feature vector after the signal passes through the human body, and finally uses SVM for identification. MU-ID [40] leverages mmWave radar to capture the limb motions and gaits of multiple users as features. Then, it uses CNN to achieve multi-user identification. Different from previous works, *BPCloak* not only pursues high-quality identity authentication but also ensures that behavior privacy is not leaked.

Privacy-preserving technique: Such techniques usually leverage perturbation or machine learning to achieve the privacy-preserving goal. For instance, in [41], Rezaei *et al.* utilize a generative adversarial net to generate perturbation. The perturbation is then added to the privacy-sufficient data to protect users' private individual attributes. To protect database privacy, Dwork *et al.* [42] add certain noise to the data. Bayerl *et al.* [43] realize privacy-preserving automatic speaker verification by outsourced secure two-party computation. To prevent mobile devices from re-identification attacks while retaining the utility of activity, Jourdan *et al.* [44] propose a machine learning-based framework to reduce the re-identification accuracy. However, there has been very little work to protect the privacy of RF signals. We thereby proposed a promising deep learning-based method to erase the behavior privacy contained in RF signals.

7 Discussion and Future Work

Effect of participant number. Although our experiment on each RF system has no more than ten participants, we prove that *BPCloak* can also be used for a larger population. Specifically, we experimented with different numbers of participants from two to ten persons. The results of the curves fitting the BIFV's AA and the loss of AA ($AA_{ori} - AA_{def}$) are shown in Fig. 21. The experiment results demonstrate that as the number of people participating in authentication increases, *BPCloak*'s ability to retain identity information becomes stronger, and the loss of AA is also getting lower. Therefore, *BPCloak* can ensure that sufficient identity information is provided even with a large set of users.

Explainability and vulnerability. As RF signals are very abstract and the relationship between identity information and behavior information is extremely complex, we leverage deep neural networks to achieve the privacy-preserving goal. Despite the outstanding performance of *BPCloak*, the

TABLE 7
Localization accuracy (LA) and recognition accuracy (RA) of WiFi signal. The value in () is TR.

Location	Location 1	Location 2	Location 3	Location 4	Location 5	Average	Localization Accuracy
RA-DoP	88.0%	96.0%	80.0%	84.8%	76.0%	85.0%	96.5%
RA-BIFV	11.8%	8.3%	10.9%	10.4%	10.1%	10.3%	76.0%
DR	86.6%	91.4%	86.4%	87.7%	86.7%	87.9%	(21.2%)

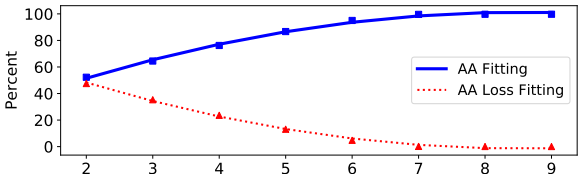


Fig. 21. Number of participants v.s. AA and AA loss.

extracted BIFVs lack explainability. In future work, we will study the parameters learned by *BPCloak* to explore why *BPCloak* works well. Besides, deep neural network is reported vulnerable to adversarial samples [45]. In future work, we will explore if there is a backdoor for the adversary to generate adversarial perturbation against *BPCloak*.

8 Conclusion

In this paper, we first noticed that there is a risk of behavior privacy leakage in RF-based user authentication. To erase the behavior privacy in RF signals, we designed a novel deep neural network named *BPCloak*. *BPCloak* can effectively filter behavior information out from RF signals without impairing the ability of signals to be used for accurate identity authentication. We conducted comprehensive experiments on real-world WiFi, RFID, and mmWave systems. The experimental results show that *BPCloak* can effectively hide behavior information while retaining sufficient identity information. The results of experiments on two open-source datasets also demonstrate the effectiveness of *BPCloak*. AT last, we show that *BPCloak* can be easily extended to other RF-based applications.

Acknowledgements

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, 61972348, and 52178175, Research Institute of Cyberspace Governance in Zhejiang University, Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Grant No. 2018R01005), Zhejiang Key R&D Plan (Grant No. 2019C03133), Ant Group Funding No.Z51202000234, and Alibaba-Zhejiang University Joint Institute of Frontier Technologies.

References

[1] C. Zhao, Z. Li, T. Liu, H. Ding, J. Han, W. Xi, and R. Gui, "Rf-mehndi: A fingertip profiled RF identifier," in *IEEE Conference on Computer Communications, INFOCOM*, 2019.
[2] 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 *International Conference on Mobile Systems, Applications, and Services, MobiSys*, 2019.
[3] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2.0: Passive human tracking with a single wi-fi link," in *International Conference on Mobile Systems, Applications, and Services, MobiSys*, 2018.
[4] F. Wang, J. Han, F. Lin, and K. Ren, "Wipin: Operation-free person identification using wifi signals," in *IEEE Global Communications Conference, GLOBECOM*, 2019.
[5] T. Xin, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "Freesense: Indoor human identification with wi-fi signals," in *IEEE Global Communications Conference, GLOBECOM*, 2016.
[6] 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.
[7] 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 *IEEE Conference on Computer Communications, INFOCOM 2021*, 2021.
[8] 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.
[9] X. Ma, J. Qu, J. Li, J. C. S. Lui, Z. Li, and X. Guan, "Pinpointing hidden iot devices via spatial-temporal traffic fingerprinting," in *IEEE Conference on Computer Communications, INFOCOM*, 2020.
[10] W. Jiang, C. Miao, F. Ma, S. Yao, Y. Wang, Y. Yuan, H. Xue, C. Song, X. Ma, D. Koutsonikolas, W. Xu, and L. Su, "Towards environment independent device free human activity recognition," in *International Conference on Mobile Computing and Networking, MobiCom*, 2018.
[11] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "Signfi: Sign language recognition using wifi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, IMWUT*, vol. 2, no. 1, pp. 23:1–23:21, 2018.
[12] P. O. POSITIVITY, "What do your daily habits reveal about your personality?" https://www.powerofpositivity.com/personality-habits/, 2020.
[13] M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," in *International Conference on Mobile Computing and Networking, MobiCom*, 2016.
[14] M. Li, Y. Meng, J. Liu, H. Zhu, X. Liang, Y. Liu, and N. Ruan, "When csi meets public wifi: Inferring your mobile phone password via wifi signals," in *ACM SIGSAC Conference on Computer and Communications Security, CCS*, 2016.
[15] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*, 2005.
[16] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016.
[17] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing, TMC*, vol. 16, no. 2, pp. 581–594, 2017.
[18] 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.
[19] 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.
[20] L. Fan, T. Li, R. Fang, R. Hristov, Y. Yuan, and D. Katabi, "Learning longterm representations for person re-identification using radio signals," in *IEEE Conference on Computer Vision and Pattern Recognition,CVPR*, 2020.
[21] M. Abuhamad, A. Abusnaina, D. Nyang, and D. A. Mohaisen, "Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 65–84, 2021.
[22] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *International Conference on Distributed Computing in Sensor Systems, DCOSS*, 2016.
[23] Y. Wang and Y. Zheng, "Modeling RFID signal reflection for contact-free activity recognition," *Proceedings of the ACM on Inter-*

active, Mobile, Wearable and Ubiquitous Technologies, IMWUT, vol. 2, no. 4, pp. 193:1–193:22, 2018.
[24] C. Li, M. Liu, and Z. Cao, "Wihf: Enable user identified gesture recognition with wifi," in *IEEE Conference on Computer Communications, INFOCOM*, 2020.
[25] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using wifi channel state information," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 98–104, 2017.
[26] R. M. C. R. de Souza, D. C. F. Queiroz, and F. J. de A. Cysneiros, "Logistic regression-based pattern classifiers for symbolic interval data," *PATTERN ANALYSIS AND APPLICATIONS*, vol. 14, no. 3, pp. 273–282, 2011.
[27] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "802.11 with multiple antennas for dummies," *Computer Communication Review*, vol. 40, no. 1, pp. 19–25, 2010.
[28] F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang, "Person-in-wifi: Fine-grained person perception using wifi," in *IEEE/CVF International Conference on Computer Vision, ICCV*, 2019.
[29] B. Z. H. Zhao, H. J. Asghar, and M. A. Kaafar, "On the resilience of biometric authentication systems against random inputs," in *Network and Distributed System Security Symposium, NDSS*, 2020.
[30] A. M. Pretorius, E. Barnard, and M. H. Davel, "Relu and sigmoidal activation functions," in *South African Forum for Artificial Intelligence Research*, 2019.
[31] Baeldung, "How relu and dropout layers work in cnns," https://www.baeldung.com/cs/ml-relu-dropout-layers#:~:text=The%20Dropout%20layer%20is%20a%20mask%20that%20nullifies,in%20which%20case%20it%20nullifies%20some%20hidden%20neurons., 2020.
[32] K. Nar, O. Ocal, S. S. Sastry, and K. Ramchandran, "Cross-entropy loss and low-rank features have responsibility for adversarial examples," *CoRR*, vol. abs/1901.08360, 2019.
[33] T. . Team, "What is cross-entropy loss?" https://365datascience.com/tutorials/machine-learning-tutorials/cross-entropy-loss/, 2021.
[34] G. E. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *CoRR*, vol. abs/1503.02531, 2015. [Online]. Available: http://arxiv.org/abs/1503.02531
[35] P. E. Rauber, A. X. Falcão, and A. C. Telea, "Visualizing time-dependent data using dynamic t-sne," in *Eurographics Conference on Visualization, EuroVis*, 2016.
[36] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," in *International Conference on Learning Representations, ICLR*, 2018.
[37] L. Guo, S. Guo, L. Wang, C. Lin, J. Liu, B. Lu, J. Fang, Z. Liu, Z. Shan, and J. Yang, "Wiar: A public dataset for wifi-based activity recognition," *IEEE Access*, vol. 7, pp. 154 935–154 945, 2019.
[38] M. Sandler, A. G. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2018.
[39] J. Li, C. Wang, A. Li, D. Han, Y. Zhang, J. Zuo, R. Zhang, L. Xie, and Y. Zhang, "Rf-rhythm: Secure and usable two-factor RFID authentication," in *IEEE Conference on Computer Communications, INFOCOM*, 2020.
[40] X. Yang, J. Liu, Y. Chen, X. Guo, and Y. Xie, "MU-ID: multi-user identification through gaits using millimeter wave radios," in *IEEE Conference on Computer Communications, INFOCOM*, 2020.
[41] A. Rezaei, C. Xiao, J. Gao, and B. Li, "Protecting sensitive attributes via generative adversarial networks," *CoRR*, vol. abs/1812.10193, 2018.
[42] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *International Conference on the Theory and Applications of Cryptographic Techniques*, 2006.
[43] Sebastian P Bayerl, Ferdinand Brasser, Christoph Busch, Tommaso Frassetto, Patrick Jauernig, Jascha Kolberg, Andreas Nautsch, Korbinian Riedhammer, Ahmad-Reza Sadeghi, Thomas Schneider, Emmanuel Stapf, Amos Treiber, and Christian Weinert, "Privacy-preserving speech processing via STPC and TEEs," 11 2019.
[44] T. Jourdan, A. Boutet, and C. Frindel, "Toward privacy in iot mobile devices for activity recognition," in *EAI International Conference on Mobile and Ubiquitous Systems, EAI MobiQuitous*, 2018.
[45] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao, "Adversarial sensor attack on lidar-based perception in autonomous driving," in *ACM SIGSAC Conference on Computer and Communications Security, CCS*, 2019.

Jianwei Liu received the BS degree from Northwestern Polytechnical University in 2018. He received his Master degree from Xi'an Jiaotong University. 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.

Chaowei Xiao received his Ph.D. degree in computer science and engineering from University of Michigan, Ann Arbor in 2020. He is now research scientist at Nvidia Research and an assistant professor at the Arizona State University. His research interests focus on Trustworthy Machine Learning.

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 The Hong Kong Polytechnic University. Her research interests include RFID, mobile computing, and smart sensing. She is student member of the IEEE and ACM.

Jinsong Han received his Ph.D. degree in computer science from Hong Kong University of Science and Technology in 2007. He is now a professor at the School of Cyber Science and Technology, Zhejiang University. He is a senior member of the ACM and IEEE. His research interests focus on IoT security, smart sensing, wireless and mobile computing.

Xian Xu is a professor at the College of Civil Engineering and Architecture, Zhejiang University. His research interests include smart structural health monitoring.

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.