Feature Selection and Cross Validation for Physical-Layer RFID Counterfeit Tag Identification

Haifeng Wu, Wei Gao, Chongrong Pu, and Zeng Yu

***Abstract*— Since radio frequency identification (RFID) tags are easily cloned, tag anticounterfeit is an important issue for RFID security. Physical-layer identification is a feasible method for the anticounterfeit, and it introduces machine learning to train classification on tag physical-layer signals. However, there are still some problems further investigated for the method, such as features of tag and cross validation. For this, this article proposes a new learning method, which extracts time- and frequency-domain statistics not only from the raw tag responding signal but also from its expected, noise, and normalized signals and adopts feature selection to classify authentic and counterfeit tags. In addition, this article also proposes a new cross validation to objectively test the performance of the physical-layer method. This experiment uses a software-defined radio to collect data from 140 tags of seven classes from three manufacturers. The results show that the classification accuracy of this new method is 4%–5% higher than that of the traditional method. Besides, under the new cross validation, the classification accuracy of all the physical layer methods will drop by 8%–10%. From this, we get an important conclusion that the performance of the physical-layer methods will depend on whether a training set has attacking tag data.**

***Index Terms*— Cross validation, feature selection, radio frequency identification (RFID), security.**

I. Introduction

RADIO frequency identification (RFID) is an automatic identification technology [1], which reads information from electronic tags with unique identification (ID) through wireless communication, without any contact and manual intervention. This technology is considered to be one of the most promising technologies in market in the 21st century and has been applied in smartphones and 5G communication [2]. A simple RFID system consists of a reader, tags, and back-end databases [3], where the reader sends a command to tags through a wireless channel. After the tags receive the command, they will respond and send their ID information. The back-end database can retrieve the corresponding item information according to IDs. In theory, tags can respond to commands sent by different RFID readers, and any attackers with a reader can freely read the information from the tags lacking access control, such as electronic product code (EPC) [4]. Due to the low cost of RFID tags, the information read can easily be cloned into another tag. If the tag does not have anticloning, the attackers can counterfeit the tag and obtain illegal benefits [5]. For this, communication security has always been an important issue for RFID.

In the communication security issue, data encryption and authentication protocols [3] are common solutions. However, the low cost of RFID tags determines their simple structure and limited computing power. High-performance encryption and authentication protocols will increase the complexity of tags, while using lightweight encryption, attackers with sufficient resources can use brute-force search to steal the data. Moreover, for simple password methods, once the password is leaked, the data will be easily stolen, such as applications in an RFID supply chain [6]. For the problem of encryption, some researchers have proposed some methods to protect tags itself, rather than data, through some hardware to defend against counterfeit tags, such as inductive coupling [7], [8] or transmit antenna [9]. However, although the use of hardware improves the security, it increases the cost of RFID tags. In addition, if the hardware circuit is changed too much, the compatibility of the tags will not be strong, and they can only be used in a specific system with the hardware, which limits its applications.

In recent years, many studies have shown that the signal to which the electronic tag responds is unique at a physical layer [10] and has a physical unclonable function (PUF). Since the PUF of each tag is different, the PUF response will not be the same even if the input is the same, regardless of whether the EPC code of the tag or even the manufacturer is the same. More importantly, PUF is unpredictable and unrepeatable, so it can be applied to tag authentication and now has become a popular method for RFID security. The physical-layer identification methods are based on PUF and machine learning and extract features from the tag responding signal as a fingerprint [10], such as time- or frequency-domain statistics of the signal [11], [12], [13], [14]. The methods only need to load the corresponding algorithm on an RFID reader and do not need to change the circuit on tags. Therefore, they will not increase the cost of the tags, have good compatibility, and are suitable for low-cost tag authentication.

Manuscript received 19 April 2022; revised 11 July 2022; accepted 7 August 2022. Date of publication 19 August 2022; date of current version 2 September 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62161052. The Associate Editor coordinating the review process was Dr. Alice Buffi. *(Corresponding author: Haifeng Wu.)*

The authors are with the School of Electrical and Information Engineering, Yunnan Minzu University, Kunming 650504, China (e-mail: whf5469@gmail.com, 434357606@qq.com, puchongrong@gmail.com, yv.zeng@gmail.com).

Digital Object Identifier 10.1109/TIM.2022.3200437

However, there are still some problems in the physical-layer identification methods, which need to be further investigated. First, what features to extract from tag responding signals is not fully discussed. Some features have statistical significance in some types of tags but not necessarily in other types [15]. How to extract or select features with significance should be further investigated. Besides, the results from the existed evaluation for the physical-layer identification methods sometimes are not objective. The existed evaluation uses the classic cross validation [14] to test the classification accuracy of the authentic or attacking tags. In application, however, the information of attacking tag may not prior known and thus is difficult to be trained in advance. Therefore, the classification accuracy from the cross validation is not necessarily accurate.

For the above problems, this article proposes a new feature extraction to better identify counterfeited tags to improve the RFID security and redesign a cross validation to evaluate the security. In the feature extraction, we extract statistics not only directly from the responding signal of a tag but also from the processed EPC, the noise, and the normalized signal of the tag itself. To obtain effective features and remove redundant ones, moreover, we also use feature selection. In addition, the designed cross validation is closer to engineering applications, where the type of attacking tags in a testing set does not exist in a training set. In experiments, we use Universal Software Radio Peripheral (USRP) to test a total of 140 tags with seven types from three manufacturers. The experimental results show that the classification accuracy of the proposed method is 4–5 percentage points higher than that of the traditional method after the classic fivefold cross validation. In addition, the experiment uses the new validation to evaluate the security. The results show that under the new cross validation, both the traditional method and the proposed method have lower classification accuracy. Therefore, we can draw a conclusion that, if the information of counterfeited tags in a training set is incomplete, for the physical-layer identification methods in application, the anticounterfeiting performance will be affected.

II. Related Work

The RFID security is essentially a wireless communication security issue, where some common solutions are authentication protocols. In a popular standard, EPC C1 Gen2 [4], its protocol specifies that a password can be set to control a tag access, but the security is low. Once the password is leaked, the data can be easily stolen. A complete RFID authentication protocol should be able to prevent tags from being tracked, cloned, eavesdropped, and leaked [3]. Mature authentication protocols use symmetric and asymmetric encryption [16], [17], [18], [19], which can resist most common attacks. However, due to the high complexity of the encryption, applying them to RFID will inevitably increase the cost. For this, some lightweight authentication protocols [20], [21], [22], [23], [24] have been proposed, which support authentication algorithms for random numbers and one-way hash functions. To better adapt to the low cost and limited resources of RFID, in addition, some ultralightweight authentication protocols [25], [26], [27], [28], [29] have also appeared, which use simple bitwise, trigonometric, and adding operations. Whether mature or lightweight authentication protocols, there should be a compromise when applied to the RFID security. Complex authentication algorithms are more secure but increase the cost of RFID. On the contrary, simple algorithms are more suitable for low-cost tags, but with reduced security. In this regard, this article only adopts a technical route different from encryption to achieve anticounterfeiting because it does not require changing the structure on the tag end and thus has better adaptation.

If the encryption authentication is called a software security method, then there are some hardware ones. The so-called hardware methods use the hardware of the RFID system to resist attacks. One type of hardware methods is to design some new logic gates [30], [31], adders [32], and even baseband circuits [33] on a tag to adapt to some lightweight encryption. Another type of methods leverages some additional hardware without changing the circuit of the tag itself. In this type, various types of hardware can be added. For example, an antenna array is added on a tag [9] and noises are mixed into signals that a reader sends. Another example is some methods based on an inductive coupling effect [7], [8], [34], where an additional tag is placed next to the tag to be read and the effect will create a unique fingerprint. Regardless of the method of changing the tag circuit or additional hardware, the designer should consider the compatibility of the new system, whether the changed one can be applied to the original one, such as the system supported by EPC C1 Gen2. Even if the new system can be compatible with the original one, it is necessary to consider whether the new hardware brings higher costs and thus limits its promotion.

In addition to the methods of software and hardware, there are also physical-layer identification methods. The differences in manufacturing tag will bring the difference in the circuits of tags, and thus, they show the difference in their responding signals on the physical layer. Therefore, authentic and counterfeited tags can be identified through the features of the physical-layer signals. According to the features, the physical-layer identification methods can be classified into various categories. One is to directly use some physical quantities of the signal as features, such as the minimum power response [35], [36] and voltage [37]. Another type is to extract the statistics of the signal, such as the mean, variance, skewness, kurtosis, autocorrelation, and other statistics in the time-domain signals [14]. There are also methods for extracting the frequency-domain features [38] as well as time–frequency features [11], [12], [13], [14]. The performance of the physical-layer identification depends heavily on the extracted features, and the features with statistical significance will produce a better performance. However, different features have different performances on different types of tags, and it is difficult to find a unified feature that can distinguish all types of tags. In addition, how to objectively evaluate the performance of the physical layer identification is also a problem. As mentioned above, the classes of tags used by attackers may not be predicted in advance, and thus, there may be a deviation in the classification accuracy of traditional cross validation in machine learning.

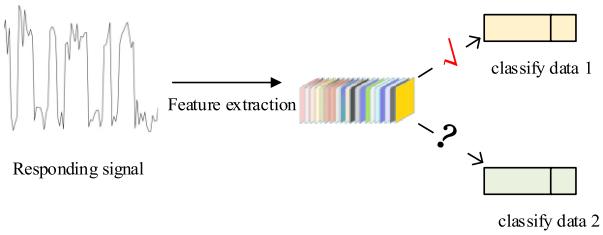


Fig. 1. Problem 1: feature extraction.

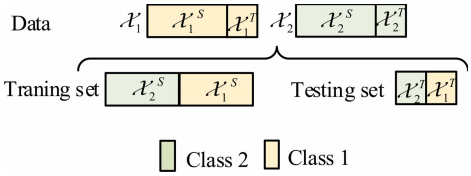


Fig. 2. Problem 2: traditional cross validation for classification of two classes.

III. Problem Description

From machine learning, the physical-layer identification is actually a classification problem. For the classification problem, no matter what kind of classifier is used, its performance depends much on the extracted features. In addition, which validation method is used is also very important for the performance evaluation of classification. Next, we will describe the two problems.

A. Feature Extraction

The tag circuits produced by different manufacturers will be different during the manufacturing tag. Even with the same manufacturer, the tag circuits will not be exactly the same due to the differences in the production time, batches, assembly lines, and so on. Thus, the signals that the circuits respond to are also different. The features from the responding signal on the physical layer embody the difference of the circuits, and thus, different types of tags can be distinguished by appropriate features. However, intertag differences tend to be diverse. As shown in Fig. 1, valid features used to classify data 1 are not necessarily available to data 2. In other words, the valid features for classification of the two data may be different. For example, the phase offset [15] differs in some tags but not in others. Of course, multifeature joint classification can be used. After the classifier is trained, more important features will be given larger weights [39], [40], while less important features will be given smaller weights. In this case, however, determining the number of features is still a problem to be solved. From the characteristics of the classification object, the more the number of features, the better, because the various characteristics of the object will be displayed as much as possible. However, when the number of features is large, redundancy will inevitably occur. Some features are not only redundant in classification but also reduce the classification performance. In this article, we extract features from both a tag responding signal and its noise and EPC signal, thus more features than that only from the responding signal. Since the EPCs of authentic and counterfeit tags are the same, the difference will be reflected more in their noise signals. Besides, the normalized EPC signals will also have differences, such as frequency drift [41]. In tests, we find that a classifier with a large number of features does not perform better than that with a small number of features. On the other hand, a classifier with fewer features will not be better either. Thus, there are actually some optimal values for the number of selected features. Too much will create redundancy, and too little will result in information loss. Therefore, the first problem to be addressed in this article is which valid features should be retained and which redundant features should be removed, regardless of any types of tags.

B. Classification Test

In general, classical cross validation can be used to evaluate the performance of the physical-layer identification methods. In the validation, data are randomly divided into two parts: one is used as a training set and the other is used as the testing set. Besides, the data classes in the testing set should also exist in the training one. Fig. 2 presents a fivefold cross validation, where classes 1 and 2 are authentic and attacking, respectively. From the figures, both the training and testing sets have the attacking class. However, the condition is not necessarily guaranteed in the application. Since it is impossible to predict which tag an attacker uses, although the training set can contain as many classes as possible, it is probable that there will be a lack of a class of tags from the attacker. If the validation model in Fig. 2 is used to evaluate the physical-layer identification methods, the result may not be accurate. Since the attacking class exists in the training set, the attacking class may also be identified in the testing set. From this, it will be falsely high for the classification accuracy to evaluate the physical-layer identification method by the traditional cross validation. Therefore, the second addressed problem in this article is how to adopt a validation model that can more fairly evaluate the physical-layer identification methods.

IV. Algorithm

A. Feature Extraction

The traditional method only extracts features from a tag responding signal, and however, the noise signal and the expected EPC signal of the tag also carry key information. Thus, the signal mentioned above should also be considered.

First, perform in-phase and quadrature (IQ) demodulation [38] on the tag signal received by a reader and compute the modulus of the IQ signal. Then, cut the EPC segment from the modulus signal as the final responding signals $a(n)$, where $n = 1, 2, \ldots, N$ are sampling points.

Next, cluster the modulus signal to obtain a cluster center vector $\mathbf{V} = [v_0, v_1]^{\mathrm{T}}$, which is expressed as

$$\mathbf{V} = \mathbf{clu}[a(n)] \tag{1}$$

where $\mathbf{clu}[\cdot]$ is a clustering function and v_0 and v_1 are clustering centers corresponding to symbols 0 and 1, respectively. Note that, which clustering center is 0 or 1 is unknown. From EPC C1 Gen 2, however, a tag will have a silent period before its RN16, where there is only a carrier leakage signal [4].

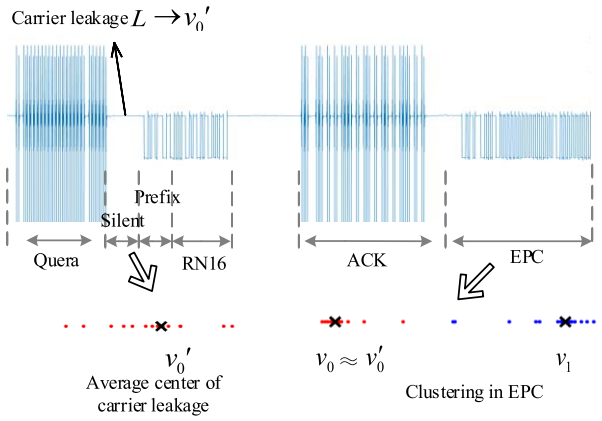


Fig. 3. Clustering in the modulus of a tag's IQ signals.

Thus, the average value of the carrier leakage will correspond to symbol 0. Fig. 3 shows the clustering process. First, v_0' is an average center during the silent period. Then, one center closer to v_0' should be v_0 and the other should be v_1. From the cluster center point, a Euclidean distance decision for $a(n)$ will get an expected EPC signal

$$a_e(n) = \text{dec}[a(n)] \tag{2}$$

where

$$\text{dec}(x) = \begin{cases} 0, & \text{if } |x - v_0| < |x - v_1| \\ 1, & \text{if } |x - v_1| < |x - v_0|. \end{cases} \tag{3}$$

Theoretically, the expected signal of the authentic tag and the counterfeit tag should be the same because both EPCs are the same. However, due to the frequency drift, their period or frequency will be different. Thus, the relevant features extracted from them will also be different.

Third, due to the difference in the transmission power, reading distance, and tag sensitivity, the amplitudes of the tag signals are different. To eliminate the difference, a responding signal can be normalized as

$$a_n(n) = \frac{a(n) - v_0}{v_1 - v_0}. \tag{4}$$

Equation 4 will normalize the received signal to 0–1, so as to reduce the difference between the signal amplitudes of each tag.

Finally, subtracting the normalized signal from the expected signal will produce the noise signal

$$a_\eta(n) = a_n(n) - a_e(n). \tag{5}$$

If the difference in frequency drift is ignored, the expected EPC signal of each tag signal should be the same. Thus, the difference between signals will be more reflected in the noise signal in (5).

After processing above, we will get the expected signal $a_e(n)$, the normalized signal $a_n(n)$, and the noise signal $a_\eta(n)$, together with the original responding signal $a(n)$, a total of four groups of signals. If the four ones are uniformly represented by $x(n)$, the mean u, variance σ^2, maximum autocorrelation R, Shannon entropy H, second-order center distance D, skewness S, and kurtosis K extracted from them are expressed as

$$u = \frac{1}{N}\sum_{n=1}^{N} x(n) \tag{6}$$

$$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N} [x(n) - u]^2 \tag{7}$$

$$R = \max_\tau \sum_{n=1}^{N} x(n)x(n+\tau) \tag{8}$$

$$H = -P(X_i) \sum_{i}^{\log_2 P(X_i)} \tag{9}$$

$$D = \frac{1}{N}\sum_{n=1}^{N} (x(n) - u)^2 \tag{10}$$

$$S = \frac{1}{\sigma^3 N}\sum_{n=1}^{N} [x(n) - u]^3 \tag{11}$$

$$K = \frac{1}{\sigma^4 N}\sum_{n=1}^{N} [x(n) - u]^4 \tag{12}$$

where X_i is the ith value of $x(n)$ after quantization. Also, we extract the signal spectral features such as centroid frequency, mean frequency, root-mean-square frequency, and frequency standard deviation, which can be expressed as

$$\text{FC} = \frac{\sum_{n=1}^{N} f_n P(n)}{\sum_{n=1}^{N} P(n)} \tag{13}$$

$$\text{RVF} = \sqrt{\frac{\sum_{n=1}^{N} (f_n - \text{FC})^2 P(n)}{\sum_{n=1}^{N} P(n)}} \tag{14}$$

$$\text{MF} = \frac{1}{N}\sum_{n=1}^{N} P(n) \tag{15}$$

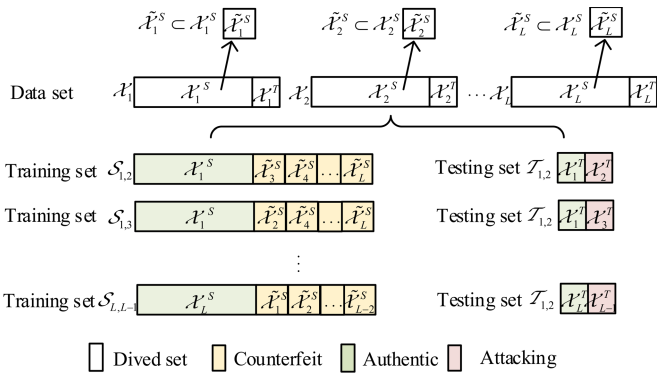
$$\text{PMSF} = \sqrt{\frac{\sum_{n=1}^{N} f_n^2 P(n)}{\sum_{n=1}^{N} P(n)}} \tag{16}$$

where $P(n)$ is the power spectrum of the nth sample point and f_n is the nth sample point's frequency. After feature extraction, each group of signals produces seven time-domain and four frequency-domain features. Thus, four groups of signals will have 44 features. We will use the features to classify tags later.

B. Cross Validation

Identifying authentic or counterfeit tags is a classification problem, and cross validation can be used to evaluate the classification accuracy. As mentioned above, traditional cross validations sometime are not objective. Here, we present a new K-fold cross validation, where the class of the attack tags does not exist in a training set.

Let $\mathbf{X}_i = [x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(M)}]$ be a vector consisting of M features from the ith tag and form the vector and its classification label y_i into a cell $\chi_i = \langle \mathbf{X}_i, y_i \rangle$. If a class l has I tags, a cell set of the tags can be denoted as



Note: Both of the counterfeit and the attacking class have a false label, but the former is in the training set and the latter in the testing one.

Fig. 4. Proposed cross validation, where the class of attacking tags is not in a training set.

$\mathcal{X}_l = \{\chi_i | i = 1, 2, \ldots, I\}$. Then, divide the set $\mathcal{X}_l$ into two subsets $\mathcal{X}_l^S$ and $\mathcal{X}_l^T$ satisfying

$$\mathcal{X}_l^S \cup \mathcal{X}_l^T = \mathcal{X}_l \tag{17-a}$$

$$\mathcal{X}_l^S \cap \mathcal{X}_l^T = 0 \tag{17-b}$$

$$|\mathcal{X}_l^T| / |\mathcal{X}_l^S| = (K-1)/K \times 100\% \tag{17-c}$$

where $l = 1, 2, \ldots, L$. Suppose that a class k is used as the authentic class and a class j used as the attack one, where $k \neq j$ and $k, j \in \{1, 2, \ldots, L\}$. Then, a training and a testing set can be expressed as

$$\mathcal{S}_{k,j} = \tilde{\mathcal{X}}_1^S \cup \cdots \mathcal{X}_k^S \cup \cdots \tilde{\mathcal{X}}_{j-1}^S \cup \tilde{\mathcal{X}}_{j+1}^S \cdots \tilde{\mathcal{X}}_L^S \tag{18}$$

$$\mathcal{T}_{k,j} = \mathcal{X}_k^T \cup \mathcal{X}_j^T \tag{19}$$

respectively, where $\tilde{\mathcal{X}}_m^S$ is the trained counterfeit tag set and $m = 1, 2, \ldots, L$, $m \neq k$ and $m \neq j$. Note that, the elements of the set are randomly from $\mathcal{X}_m^S$, i.e.,

$$\tilde{\mathcal{X}}_m^S \subset \mathcal{X}_m^S \tag{20-a}$$

and its cardinal number satisfies

$$\left|\tilde{\mathcal{X}}_m^S\right| = \left|\tilde{\mathcal{X}}_n^S\right|, \quad \text{for any } m \neq n \tag{20-b}$$

$$\left|\cup_m \tilde{\mathcal{X}}_m^S\right| = \left|\mathcal{X}_k^T\right|. \tag{20-c}$$

Equation (20) ensures that the size of the total counterfeit training sets is equal to that of the authentic training set. For (19), due to $k \neq j$, there are $A_L^2 = L(L-1)$ combination in the training and the testing set, as shown in Fig. 4. From the figure, $\mathcal{X}_j^T$ as an attacking tag set (denoted by red), its class j does not appear in the training set. This means that the class of the tag by an attacker is not prior known, so it cannot be pretrained.

C. Feature Selection

Since there are many features extracted in Section IV-A, it is necessary to retain valid features and remove redundant features. A commonly used solution is feature selection. Feature selection is usually categorized into filtering, wrapping, and embedding. The latter two feature selection methods involve choosing a classifier and adding a validation set in addition to the training and test sets. In the proposed cross validation, the classification labels of the training set and the test set are not consistent, and thus, the wrapping and embedded feature selection may be prone to overfitting. Here, this article will use the filtering method. Next, we will describe how to select features in the new cross validation.

Let a cell be $\chi = \langle \mathbf{X}, y \rangle$, where $\mathbf{X} = [x^{(1)}, x^{(2)}, \ldots, x^{(M)}]$ is a tag's feature vector and y is its classification label. If the cell is in the training set $\mathcal{S}_{k,j}$, i.e., $\chi = \langle \mathbf{X}, y \rangle \in \mathcal{S}_{k,j}$, then we can sort the feature weights $\omega_{x^{(m)}}$, $m = 1, 2, \ldots, M$ via filtering feature selection and then select W features with the largest weight, which can be expressed as

$$\langle p_1, p_2, \ldots, p_W \rangle = \arg\max_m \omega_{x^{(m)}}. \tag{21}$$

From (21), a new cell $\chi^S = \langle \mathbf{X}^S, y \rangle$ can be obtained where $\mathbf{X}^S = [x^{(p_1)}, x^{(p_2)}, \ldots, x^{(p_W)}]$ is a vector of the selected features. All of the new cells will form into a new training set $\bar{\mathcal{S}}_{i,j}$, i.e.,

$$\chi^S = \langle \mathbf{X}^S, y \rangle \in \bar{\mathcal{S}}_{k,j}. \tag{22}$$

Similarly, the cell $\chi^T = \langle \mathbf{X}^T, y \rangle$ in the testing set can be obtained, and the new set $\bar{\mathcal{T}}_{i,j}$ can also be obtained, i.e.,

$$\chi^T = \langle \mathbf{X}^T, y \rangle \in \bar{\mathcal{T}}_{k,j} \tag{23}$$

where $\mathbf{X}^T$ is a vector of the largest weights features in the testing set.

After the feature selection, cross validation can be performed. If the weight w of the classifier $f_{\text{clas}}(\cdot)$ satisfies

$$y = f_{\text{clas}}(w, \mathbf{X}^S), \quad \langle \mathbf{X}^S, y \rangle \in \bar{\mathcal{S}}_{k,j} \tag{24}$$

the training will complete. Then, test results are obtained by

$$\hat{y} = f_{\text{clas}}(w, \mathbf{X}^T), \quad \langle \mathbf{X}^T, y \rangle \in \bar{\mathcal{T}}_{k,j}. \tag{25}$$

Comparing the test label $\hat{y}$ with the expected label y can produce the classification accuracy. The steps of the algorithm in this section can be seen in Table I.

In the above feature selection, the number of selected features W is an important parameter that affects the classification performance. If W is too large, the performance of selection will reduce. In an extreme case, W is equal to the original number of features, and it is equivalent to no feature selection. If W is too small, the key classification information is easily lost. A common method is to introduce a validation set. Try different W values in a test set, take the value of W with the highest classification accuracy in the test set as that in the validation set, and get the final classification accuracy. As mentioned above, however, the classification information of the attack label is unknown in advance, so it cannot be pretrained or pretested. Therefore, introducing the validation set will also lead to overfitting. In application, we can choose an intermediate value. The details of the parameter values can be discussed in Section V.

V. Experiment Setup

In this experiment, several different classes of tags will be selected as test ones. First, a tag writer writes the same EPC code to all tags to eliminate the influence of different EPC codes on the identification of authentic and fake tags. Then,

TABLE I
ALGORITHM STEPS

input:
Tag responding signal $a(n)$
output:
Classification label $\hat{y}$
known conditions:
Feature selection algorithm
The number of feature selections W
steps:
①Signal processing: get the expected signal $a_e(n)$, the normalized signal $a_n(n)$ and the noise signal $a_n(n)$ from (1-5).
② Feature Extraction: extract the feature vector $\mathbf{X}=[x^{(1)},x^{(2)},...x^{(M)}]$ of from (6-16).
③ Set division: divide data into a training set $S_{i.i}$ and a test set $T_{i.i}$ from (17-20), or from a *K*-fold cross validation like Fig. 4.
④Feature selection: get a new training set $S'_{i.i}$ and a new set $T'_{i.i}$ from (21-23).
⑤Train & Test: compute the label $\hat{y}$ from (24-25).

a reader is used to obtain the EPC signal of each tag, and some features are extracted from the signal as the input of a classifier. Finally, the experiment sets one class of tags as true and the rest as false, testing the classification performance after many times repetition.

A. Data Sources

The data in this experiment come from passive UHF RFID tags specified by EPC C1 Gen2, and 140 tags of seven classes commonly found in the market. The seven classes of tags are manufactured by three manufacturers, as shown in Table II. Before collecting data, write the same EPC code to all of the 140 tags. The writer used is a reader from Guangzhou Wang yuan Electronics and its system parameters are shown in Table III.

The data collection is completed by an ultrahigh-frequency RFID system and a USRP software radio [23]. The system follows the EPC C1 Gen2 standard, and the software is implemented by GNU Radio. For detailed parameters, please refer to Table IV, and the code download address is https://github.com/nkargas/Gen2-UHF-RFID-Reader.

During each data collection, only one tag is placed in front of antennas, and other tags are not within the magnetic field of the antennas to reduce the risk of tag collisions. All data collection is not performed in an isolated environment, which may include thermal noise, cell phone noise, wireless network, and radio frequency noise. The tags are randomly placed in a rectangular area formed by two antennas, as shown in Fig 5. Each tag records 10 s of data, randomly intercepts the signal containing a silent period and an EPC code, and stores it in a MATLAB format. Finally, 140 tags will produce 140 tag data.

The tags in Table II adopt the technical standards of Alien Company, which is the most trusted RFID tag supplier in the world (https://www. alientechnology.com/) and its tags have passed the EPC C1 GEN2 standard. The tags in Table II are the common Alien tag models on the market, which have good generalization. The parameters of the UHF100U reader in Table III also comply with ISO 18000-6C protocol and EPC C1 GEN2 standard and can read and write any tags meeting the above standards. In this experiment, its role is to write the same EPC code for each tag and make the tag as a counterfeit one.

TABLE II
TAG MODEL AND MANUFACTURER

Class	Model	Manufacturer
1	Alien9640	Guangzhou Wang yuan Electronics
2	Alien9662	Shenzhen Qibao Technology
3	Alien9654	Shenzhen Qibao Technology
4	Alien9662	Guangzhou Wang yuan Electronics
5	Alien7017	Guangzhou Wang yuan Electronics
6	Alien9662	Nanjing Lejay Technology
7	Alien9654	Nanjing Lejay Technology

TABLE III
PARAMETERS OF THE READER WRITING EPC TO TAGS

Parameter	Description
Manufacturer	Guangzhou Wang yuan Electronics
Model	UHF 100U
Frequency	865~868MHz or 902~928MHz
Standard	EPC C1 Gen2
Distance	0-0.1m
Communication port	USB
Voltage	DC+5V
Maximum power	4W

TABLE IV
PARAMETERS OF USRP

Parameter	Description
Motherboard	USRP N200
Daughter board	RXF900
Antenna	
Quantity	2
Type	Circularly polarized
Gain	7dBic
Distance	0.5-1.5m
Link frequency	40kHz
Maximum queries	1000
Encoding	FM0
Transmission power	17.8dBm
Emission amplitude	0.1
Sampling frequency	1000kHz

B. Classification and Algorithms

This experiment uses the following two fivefold cross-validation methods to evaluate algorithms.

1) Fivefold Cross Validation I: There are L classes, where if the kth is authentic, then it will be used for the traditional fivefold cross validation of the binary classification with each of the jth classes, $j = 1, 2, \ldots, L$, $j \neq k$, as shown in Fig. 6. The classification accuracy of the kth label will be the average of the binary classification results for each test set $T'_{k,j}$, $j \neq k$.

2) Fivefold Cross Validation II: There are L classes, where if the kth is authentic, then it will be used for the traditional fivefold cross validation of the binary classification with each of the jth class, $j = 1, 2, \ldots, L$, $j \neq k$ respectively, shown

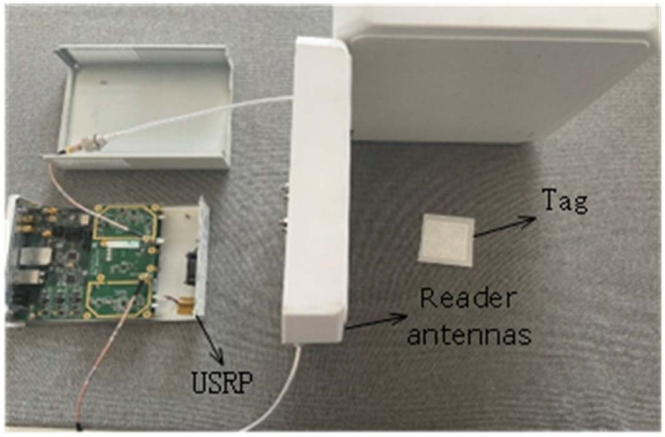


Fig. 5. Experiment setup, where a USRP reader collects tag data.

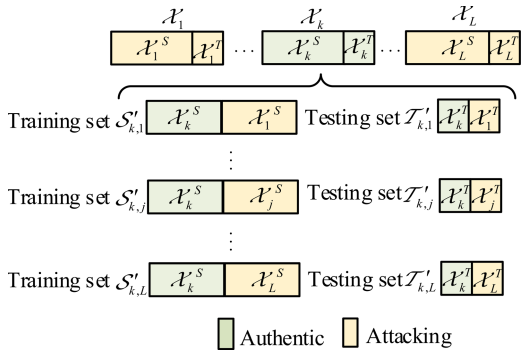


Fig. 6. Fivefold cross validation I.

in Fig. 4. Note that the jth class is an attacking class, not in the training sets. The classification accuracy of the kth label will be the average of the binary classification results for each test set $\mathcal{T}'_{k,j}$, $j \neq k$.

This experiment evaluates the performance of the following algorithms for different features or feature selection.

3) Seven Without FS: Using the method of literature [14], the seven statistics in (6–12) are extracted from the original responding signal, and the classification is performed directly without feature selection.

4) PSD Without FS: Using the method of literature [10], this method can identify genuine and fake RF signals by extracting the PSD fingerprint feature of the signal.

5) Twenty-Eight Without FS: The expected, normalized, and noise signals are obtained from (1–5), the method of [14] is used to extract seven statistics in (6–12) for the above signal and the original responding signal, a total of 28 features, and the classification is performed directly without feature selection.

6) Seven/Fourteen/Twenty-One With ReliefF: The algorithm in Table I is used for feature extraction and feature selection. The feature selection method adopts ReliefF algorithm [39], realized by a function *relieff* in Statistics and Machine Learning Toolbox of MATLAB 2021a. The number of feature selections is 7, 14, or 21, that is, $W = 7$, 14, or 21.

7) Seven/Fourteen/Twenty-One With Chi2: The algorithm in Table I is used for feature extraction and feature selection. The feature selection method adopts a chi-square test algorithm [40], realized by a function *fscchi2* in the Statistics and Machine Learning Toolbox of MATLAB 2021a. The number of feature selections is 7, 14, or 21, that is, $W = 7$, 14, or 21.

8) Forty-Four Without FS: Four frequency-domain features are extracted from four signals, a total of 16 features, and then combined with the 28 features proposed in this method, a total of 44 features $W = 11$, 22, or 33.

9) Eleven/Twenty-Two/Thirty-Three With ReliefF/Chi2: Four frequency-domain features are extracted from four signals, a total of 16 features, and then combined with the 28 features proposed in this method, a total of 44 features $W = 11$, 22, or 33.

10) Two/Four/Six With ReliefF/Chi2: The statistics in (6–12) are extracted from the original responding signal, and then, the feature selection is performed. The feature selection method adopts ReliefF or chi-square. The number of feature selections is 2, 4, or 6.

The classifier adopted in this experiment is a support vector machine (SVM), which is implemented by a function *fitcsvm* in MATLAB2021a Statistics and Machine Learning Toolbox. The classification accuracy is expressed as acc, which is defined as

$$\text{acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \tag{26}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

VI. Experimental Results

A. Preprocessing Results

The signal collected by the USRP software radio platform is the sampling point signal of the real part and the imaginary part after demodulation of the signal communicated by the reader and the tag. Compute the modulus of the real part and imaginary part of the collected signal, and then, intercept a complete communication signal, as shown in Fig. 7. In the figure, the complete signal includes query, RN16, ACK, and EPC signal, which are consistent with EPC C1 Gen2. The main signal used to classify the tags is the EPC segment, as well as the normalized, expected, and noise signals via (1–5), as shown in Fig 8. As can be seen from the figure, the normalized signal is mainly concentrated around 0 and 1, the expected signal is a binary signal, and the noise signal is the difference between the above two.

In addition, we present the spectrogram of the tag signal. Fig. 9 shows the spectrums of the two tags, which reach more than 20 dB only on the lower frequency band and are below -20 dB on the rest of the frequency bands. One reason is that no noise of other frequencies is in the test environment. Besides, the tag signal passes through a low-pass filter [41] before IQ demodulation, so the high-frequency noise is also filtered out.

To show the spatial distribution of the extracted features in this article, Fig. 10 shows the 3-D display after feature extraction and dimensionality reduction of two classes from

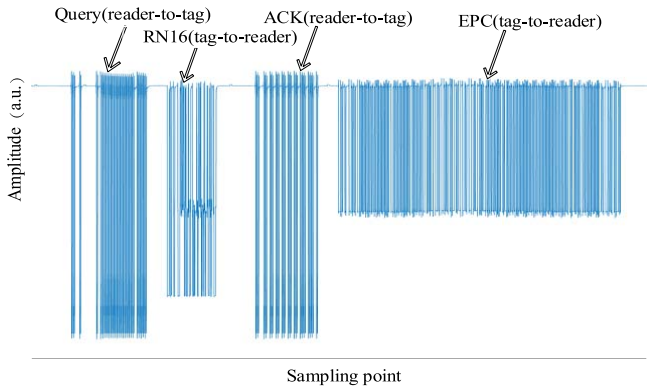


Fig. 7. USRP software radio platform collects a complete signal segment communicated by a reader and the tag, which is the modulus of the real part and the imaginary part after demodulation of the signal.

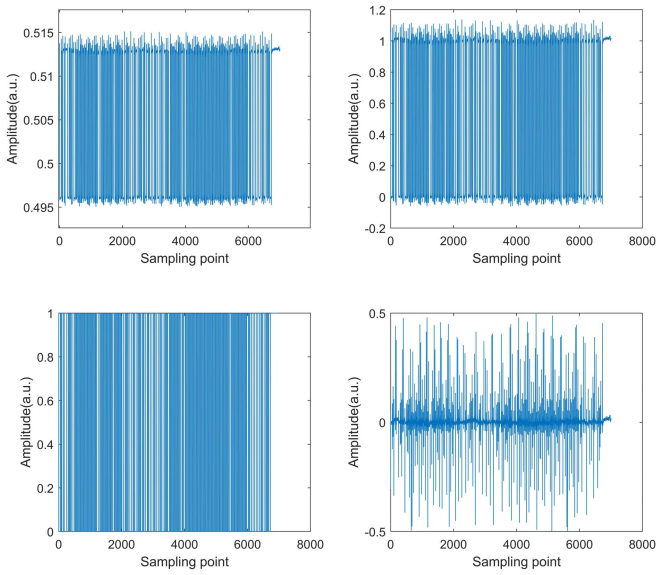


Fig. 8. Responding signal, normalized signal, expected signal, and noise signal from the EPC segment via (1–5).

TABLE V
EUCLIDEAN DISTANCE BETWEEN TAGS (A.U.)

	Calss1	**2**	**3**	**4**	**5**	**6**	**7**
Class1	1.73	6.26	5.90	5.61	5.81	4.68	8.77
2	6.26	1.67	1.68	2.50	2.75	6.62	10.46
3	5.90	1.68	1.49	2.31	2.66	6.36	10.26
4	5.61	2.50	2.31	2.52	2.86	5.42	9.38
5	5.81	2.75	2.66	2.86	2.97	5.30	9.25
6	4.68	6.62	6.36	5.42	5.30	1.43	5.46
7	8.77	10.46	10.26	9.38	9.25	5.46	8.80

the seven tag classes. From the figure, the Euclidean distances between most of the same class tags are close, while those between most different classes are far. In Section VI-B, we present the classification accuracy results after feature selection from the features. Table V gives seven classes of

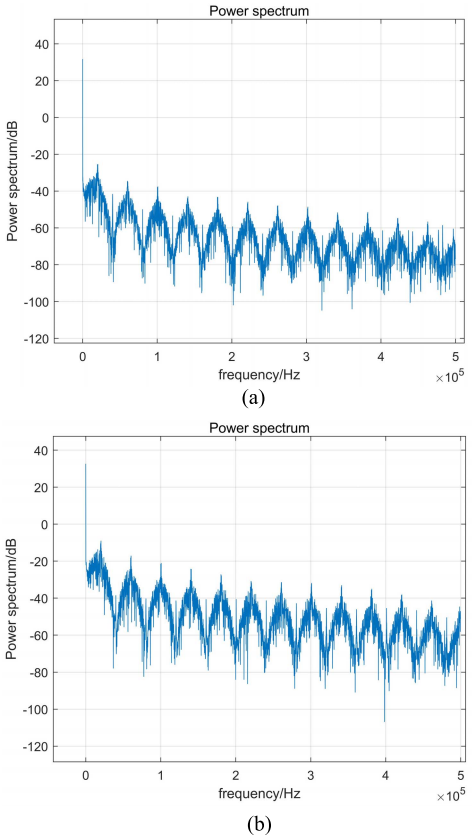


Fig. 9. Power spectral density. (a) One tag's power spectrum. (b) Other tag's power spectrum.

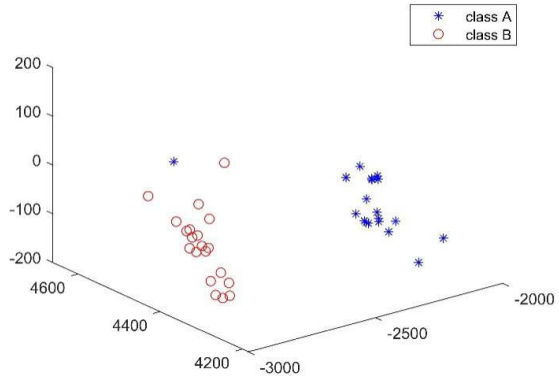


Fig. 10. 3-D display of the two classes of tags after feature extraction and dimension reduction, where the feature extraction adopts the 28 without FS method. Dimensionality reduction adopts PCA, which is reduced from 28 dimensions to three dimensions. Two classes of tags, A and B, are from the seven classes, that is, classes 6 and 7 in Table II, with a total of 40 tags.

intralabel distances as quantitative analysis, where the diagonal line is the intralabel distance and the nondiagonal line is the interlabel distance. As can be seen from the table, the intraclass distances of most labels are smaller than the interclass distances.

B. Results of Fivefold Cross Validation I

Figs. 11 and 12 show the average values of the 28 feature weight for the seven classes, after chi-square and ReliefF processing in cross validation I, where the weights are normalized. From the figure, the weight value distribution for each

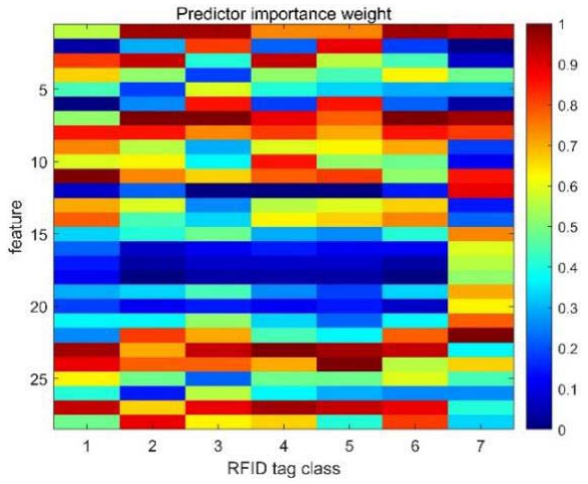


Fig. 11. Heatmap for the feature weights by ReliefF in different tag classes, where the weights are normalized and the tag classes are given in Table II. The 28 features are extracted from the responding, the noise, the expected, and the normalized signal, and their extraction methods are from (6–12).

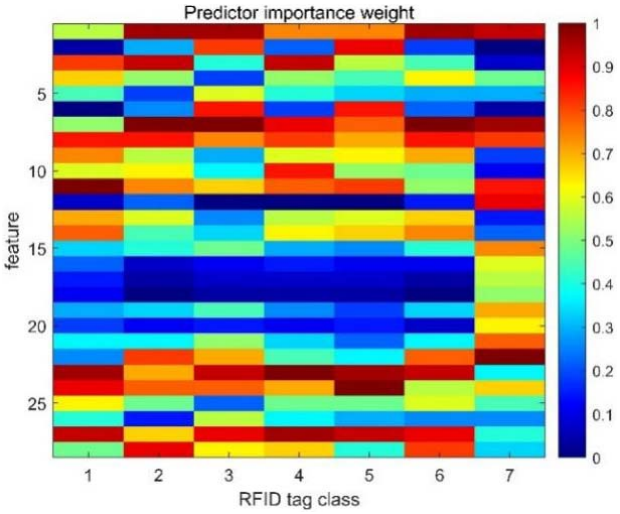


Fig. 12. Heatmap for the feature weights by chi-square in different tag classes, where the parameters are the same as in Fig. 11.

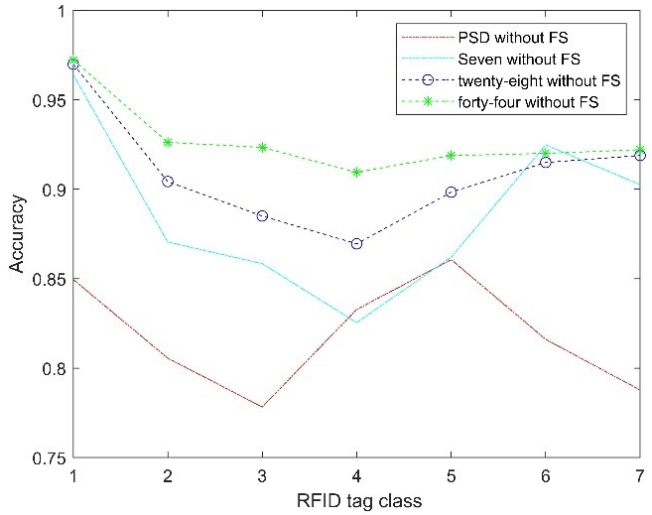


Fig. 13. Classification accuracy curves without feature selection in cross validation I, where the tag class is given in Table II.

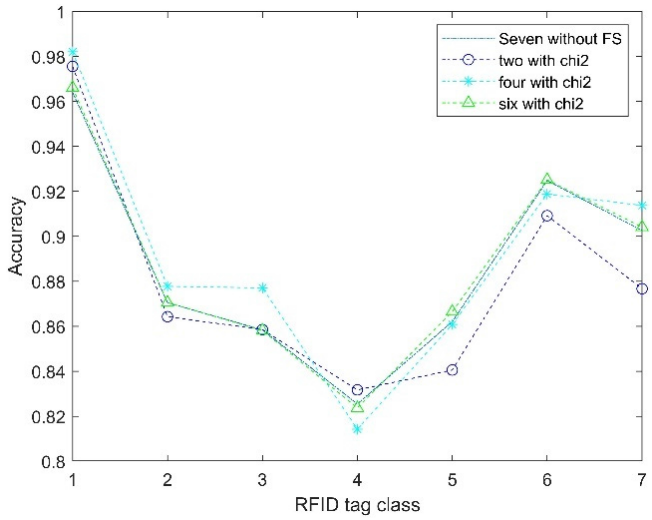


Fig. 14. Classification accuracy curves with feature selection by chi-square in cross validation I, where the features are extracted only from the tag responding signals and the tag class is given in Table II.

class is different. If the features are selected by the weights, features selected for different classes will be different.

Fig. 13 shows the classification accuracy curves in cross validation I using 44 features, 28 features, traditional seven features, and power spectral fingerprint features. From the figure, there is no feature selection, the 44 feature method is the highest, and the curve of 28 features is slightly higher than the curve of seven features, except for six types of RFID tags, but it is significantly higher than the power spectrum fingerprint feature. The method with 44 features without feature selection is about four percentage points higher than the method with seven features.

Fig. 14 shows the results of chi-square feature selection, where the seven features are extracted only from the tag responding signal. From the figure, no matter how many features are selected, the accuracy improvement is not obvious. Moreover, the difference between the methods is only about 0.5%. Therefore, without additional valid features, it is difficult to improve the accuracy even after feature selection. Fig. 15 shows the results of the ReliefF selection method, which is similar to Fig. 14.

Fig. 16 presents the classification accuracy results using the chi-square feature selection for 28 features in cross validation I. As can be seen from the figure, no matter how many features are selected from the 28 features, the classification accuracy is higher than that of the traditional seven-feature method, except for class 6. The average classification accuracy of the method using feature selection is higher than that of the traditional method, and the average accuracy is over 90%. The highest classification accuracy of the method of selecting seven features is about 4% higher than that of the traditional method. In addition, from the above results, the average difference of the methods for selecting three different numbers of features is only 1%–2%.

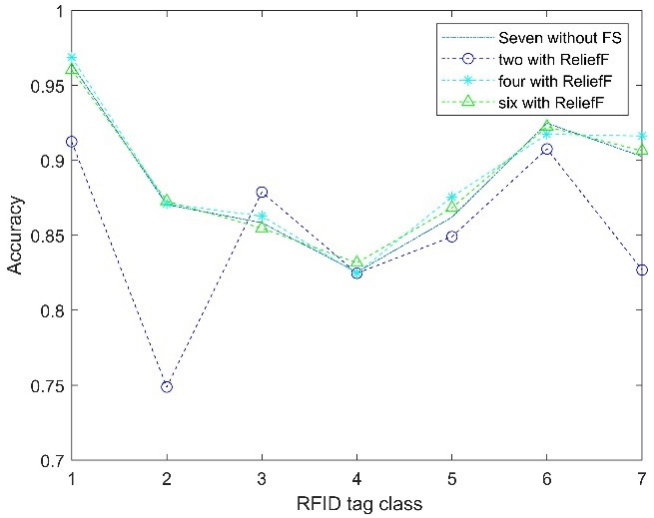


Fig. 15. Classification accuracy curves with feature selection by ReliefF in cross validation I, where the features are extracted only from the tag responding signals and the tag class is given in Table II.

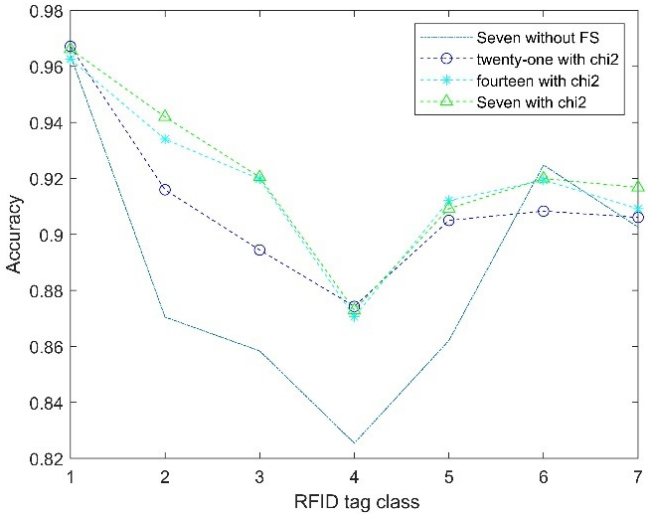


Fig. 16. Classification accuracy curves with feature selection by chi-square in cross validation I, where the features are extracted from the tag responding, the expected, the noise and the normalized signal.

Fig. 17 presents the classification accuracy results using ReliefF feature selection for 28 features in cross validation I. Similar to the results of chi-square, the methods with feature selection perform better than those without selection, no matter how many features are selected. The improvement of the classification accuracy is also about 4% and the highest accuracy occurs at selecting 21 features. Figs. 18 and 19 present the classification accuracy results of ReliefF and chi-square feature selection for 44 features in cross validation I. Similar to the results of 28-feature selection, the methods with feature selection perform better than those without selection, no matter how many features are selected. The improvement of the classification accuracy is also about 5%.

Fig. 20 shows the classification accuracy and time of the algorithms, whose values are average for many repetitions. Besides, the classification time is to train the model once.

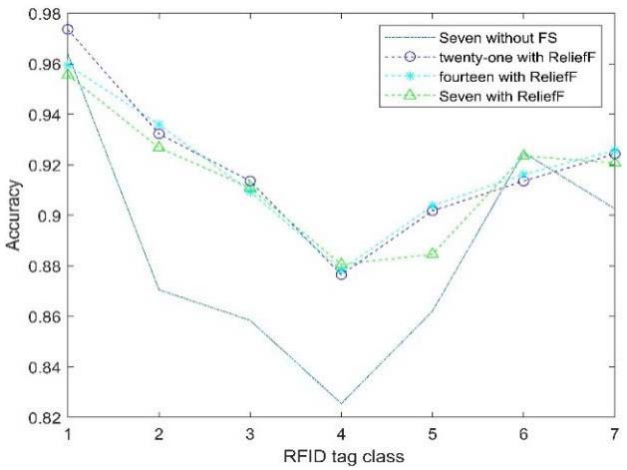


Fig. 17. Classification accuracy curves with feature selection by ReliefF in cross validation I, where the features are extracted from the tag responding, the expected, the noise, and the normalized signal.

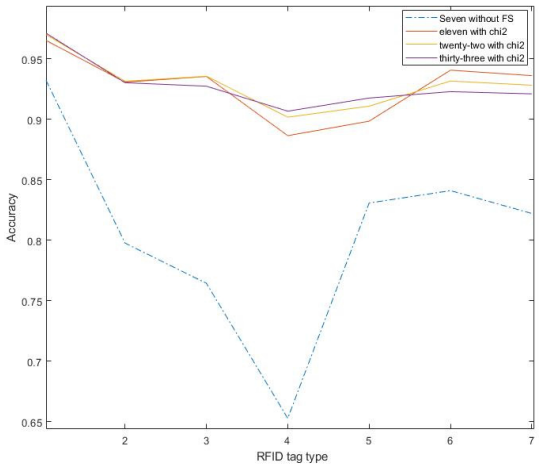


Fig. 18. Classification accuracy curves with feature selection by chi-square in cross validation I.

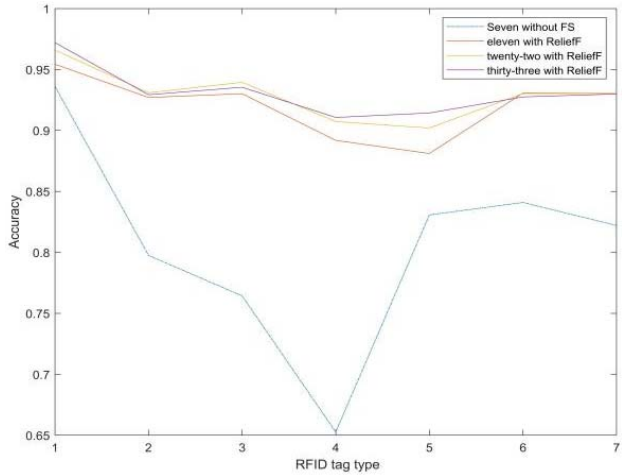


Fig. 19. Classification accuracy curves with feature selection by ReliefF in cross validation I.

From the figure, the proposed algorithm's accuracy is higher than the traditional ones, but its time is more.

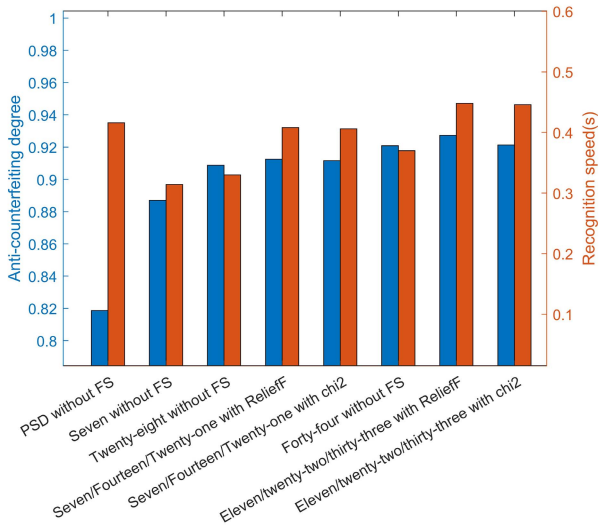


Fig. 20. Algorithm accuracy and recognition speed in cross validation I.

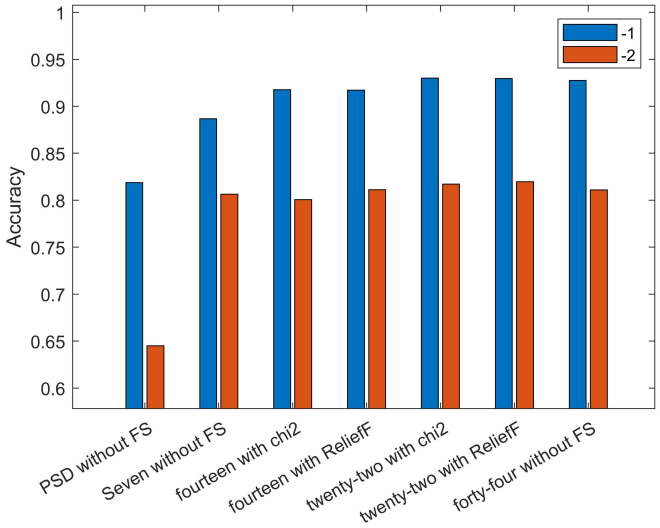


Fig. 21. Average classification accuracy histogram, where "−1" represents cross validation I and "−2" represents II.

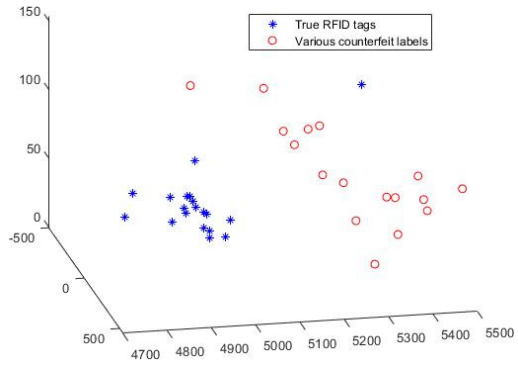


Fig. 22. PCA clustering diagram.

C. Comparison of Cross Validations I and II

Fig. 21 shows the classification accuracy histogram under cross validations I and II, where the number of features selected by chi-square is 14. Since the results for the number of features seven or 21 are similar, we give the results only for 14. Similarly, 44 features also select the middle number of features, 22 features. From the figure, the classification accuracy of the proposed and the traditional method in cross validation II is lower than that in cross validation I.

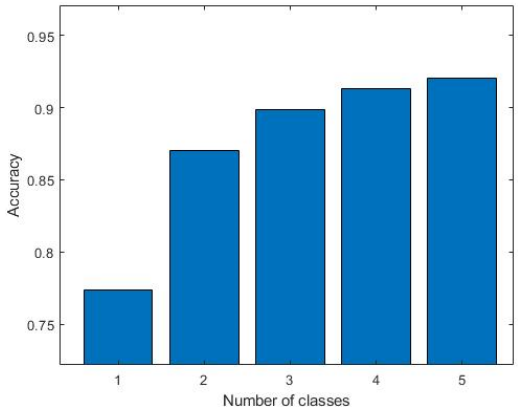


Fig. 23. Number of training tags and classification accuracy.

VII. Conclusion and Discussion

In the RFID communication security, the physical-layer identification is considered a feasible method to solve tag anticounterfeiting, and the performance of the method depends on how to extract features from tag signals. At present, most of the existing physical-layer tag identification focuses on which features perform better, such as time–frequency statistics and the physical quantities of the signal itself. However, no matter what kind of features will show differences in performance in classifying different types of tags, it is difficult to classify all tags through some fixed features. In this article, we increase the number of features, not only from the received signal but also from the noise signal, the desired signal, and the normalized signal, so that the number of features becomes 28 and 44. At the same time, a feature selection method is used to remove redundant features, retain effective features, and further improve the classification accuracy.

It is well known that in machine learning, K-fold cross validation is usually used to test the classification performance. In the validation, the class in a testing set also appears in a training one, and the feature selection method often has a better result in this case. However, in the application, the class of an attacking tag is not always predictable, so the class of the attacking tag may not exist in the training set. From that, this article designs a cross validation where the attacking tag is not in the training set. In this new validation, it will be necessary to reevaluate the method of feature selection and even the classification without feature selection. If the classification accuracy declines, it means that the performance of the methods may be falsely high. The algorithm proposed in this article is similar to semisupervised learning. It is trained by knowing the true tag label and various fake attack lag, and the corresponding machine learning model is obtained. Finally, it is tested by attack tags with unknown label. Fig. 22 shows a tag clustering

diagram in a 3-D space. From the figure and Table V, the intraclass distance of most of the tags is smaller than the interclass distance, which indicates that the use of Euclidean distance will theoretically distinguish different classes of tags. Of course, the classification accuracy of unsupervised learning is usually lower than that of supervised learning, and the experimental results also show that the accuracy of cross validation II (unsupervised) is 8%–10% lower than that of cross validation I (supervised). However, if more fake tags are trained, the recognition rate of unknown tags by our model will gradually increase. To this end, we give the result about the number of training tags classes and classification accuracy. As shown in Fig. 23, when the number of fake tags' categories in the training set increases, the classification accuracy will also increase. Therefore, we can infer that when there are enough kinds of tags in the training database, the model for recognizing tags with unknown labels will also be robust enough.

In the implementation, we first extracted seven commonly used features from the raw RFID tag responding signals, including first-order statistics, second-order statistics, and even entropy. However, in order to be able to find probable valid features or to expand the range of features that can be selected, we process the raw tag signals to obtain the expected, the noise signal, and the normalized signal, and extract features from the signals. Since both authentic and counterfeit tags have the same EPC, theoretically, their expected EPC signals should be the same, and the difference is mainly shown in the noise signal. Therefore, it would be reasonable to extract features from the noise. In addition, since each tag may have different frequency drift, the period of their expected EPC signal will also have some differences. Besides, the amplitude of the response signal of each tag will also be different. For example, the same tag will have different response amplitudes at different reading distances. Thus, the signals need to be normalized to 0 and 1 to eliminate the influence of amplitude on classification. Finally, we extracted a total of 28 features from the above signals and then performed feature selection for classification.

In the experiment, we first find that the method of extracting 28 features does not significantly improve the classification accuracy compared with the method of seven features, and the former is only about 2% higher than the latter. However, if feature selection is performed on 28 features, the highest classification accuracy reaches 92%, which is about four percentage points higher than the seven-feature method. The results show that there are redundant features in the 28 features, and removing some redundant features can improve the classification accuracy. In other words, using only seven features may have information loss, and adding valid features can improve the classification performance. The conclusion is also confirmed by another result that implementing feature selection from seven features does not significantly improve the classification performance, only about 0.5%. The reason is that although the probable redundant features are removed, the valid features are not increased.

When we select features for different tag classes, furthermore, we find that the weight of the selected features will change with the tag classes. The heatmaps drawn by the feature weights in different tag classes are not the same. For example, for Shenzhen Qibao Technology Alien9662 tag, the feature with the largest weight is the maximum autocorrelation of the received signal, while for Nanjing Lejay Technology Alien9654 tag, the feature with the largest weight becomes the mean. Therefore, it is difficult to try to find a few fixed features that can identify all tag classes. On the other hand, the feature selection does not select fixed features but relies on the training data to select features and the feature selection results vary with the training data. As long as the training data are suitable, the selected features may be available.

In addition, the experiment adopts two feature selection methods, chi-square and ReliefF, to test the performance. The main purpose is to verify whether the method proposed in this article must depend on a specific feature selection method. The experimental results show that as long as feature selection is performed, the classification accuracy can be improved to 92%. Of course, there are still differences between the two methods. The biggest difference is that the number of features selected will affect the final classification accuracy. In the chi-square test, the highest accuracy is when the number of the selected features is 7, while the number of the features in ReliefF is 21. How to determine the number is a problem that needs to be investigated. One popular method is to use embedded methods, where we add another validation set, find an optimal number in a testing set, and then verify it in the validation set. However, there is also a problem that different data tend to get different feature selection results, shown in the feature distribution heatmaps. The validation set method may fall into overfitting if the class validated is not in the testing set. Carefully observe the experimental results and find that when the number of features is selected as 7, 14, and 21, the difference between the three is not very big, about 1%–2%, but they all exceed the traditional seven-feature method. Therefore, as long as feature selection is performed, the classification accuracy can always be improved, and the number of the selected features does not necessarily have to take an optimal value because the fluctuation of performance is not very large. We can take an intermediate value, such as 14. Note that, we also try 16 features in the frequency domain, and let 28 features change into 44 features. Experimental results show that, without feature selection, it is 2% higher than 28 features, and with feature selection, it is 1% higher than 28 features. Therefore, the frequency-domain features will be helpful to classification performance.

Furthermore, it can be speculated that the number of selected features correlates with classification accuracy. The current number of selected features is 44, which improves the classification accuracy by 5%. In future work, if we can extract some common third-order or even fourth-order statistics to increase the number of features, the classification accuracy may be further improved. Of course, some experimental results in this article have some uncertainties. Due to the limitation of experimental conditions, the transmitting power of the USRP device we adopted is limited, and the tag can only be read in the small magnetic field range of the reader. Therefore, the amplitude changes of the responding signals are not large. Moreover, since the tag is closer to the reading antenna, the noise is small, and the signal-to-noise ratio is

between 16 and 19 dB. However, the actual environment of attacking tag may be more complex. Not only the amplitude of tag responding signal may change, but also the signal-to-noise ratio may be smaller. Therefore, we also hope that in future work, tag classification will be carried out under a more complex and changeable condition, such as some power-enhancing hardware circuits or antennas to increase the range of tag reading.

The algorithm code in this article has been uploaded to GitHub. Its download address is https://github.com/monk5469/counterfeit-identification.

REFERENCES

[1] K. Finkenzelle, "RFID Handbook (fundamentals and applications in contactless smart cards, radio frequency identification and near-field communication) ‖ physical principles of RFID systems," in *RFID Handbook*. Hoboken, NJ, USA: Wiley, 2010, pp. 361–418.

[2] B. Gianmarco, S. Gary, D. Franc, G. Raimondo, and K. J. S. Roman, "Experimental identification of smartphones using fingerprints of built-in micro-electro mechanical systems (MEMS)," *Sensors*, vol. 16, no. 6, p. 818, 2016, doi: 10.3390/s16060818.

[3] A. Ibrahim and G. Dalkilic, "Review of different classes of RFID authentication protocols," *Wireless Netw.*, vol. 25, no. 3, pp. 961–974, Apr. 2019, doi: 10.1007/s11276-017-1638-3.

[4] *EPC Radio Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz-960 MHz*, Version 2.0. 1, G.S. EPCglobal, Brussels, Belgium, 2015.

[5] S. A. Ahson and M. Ilyas, "RFID handbook: Applications, technology, security, and privacy," in *RFID Handbook*. Boca Raton, FL, USA: CRC Press, Dec. 2017.

[6] I.-C. Lin, H.-H. Hsu, and C.-Y. Cheng, "A cloud-based authentication protocol for RFID supply chain systems," *J. Netw. Syst. Manage.*, vol. 23, no. 4, pp. 978–997, Oct. 2015, doi: 10.1007/s10922-014-9329-1.

[7] H. Xu, X. Yin, F. Zhu, and P. Li, "An enhanced secure authentication scheme with one more tag for RFID systems," *IEEE Sensors J.*, vol. 21, no. 15, pp. 17189–17199, 2021, doi: 10.1109/JSEN.2021.3077575.

[8] G. Wang *et al.*, "Hu-Fu: Replay-resilient RFID authentication," *IEEE/ACM Trans. Netw.*, vol. 28, no. 2, pp. 547–560, Apr. 2020, doi: 10.1109/TNET.2020.2964290.

[9] G. Essam, H. Shehata, T. Khattab, K. Abualsaud, and M. Guizani, "Novel hybrid physical layer security technique in RFID systems," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Tangier, Morocco, Jun. 2019, pp. 1299–1304.

[10] S. U. Rehman, K. W. Sowerby, and C. Coghill, "Analysis of impersonation attacks on systems using RF fingerprinting and low-end receivers," *J. Comput. Syst. Sci.*, vol. 80, no. 3, pp. 591–601, 2014, doi: 10.1016/j.jcss.2013.06.013.

[11] J. Han *et al.*, "GenePrint: Generic and accurate physical-layer identification for UHF RFID tags," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 846–858, Apr. 2016, doi: 10.1109/TNET.2015.2391300.

[12] D. Zanetti, B. Danev, and S. Apkun, "Physical-layer identification of UHF RFID tags," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, Chicago, IL, USA, 2010, pp. 353–364.

[13] A. Mehmood, W. Aman, M. M. U. Rahman, M. A. Imran, and Q. H. Abbasi, "Preventing identity attacks in RFID backscatter communication systems: A physical-layer approach," in *Proc. Int. Conf. U.K.-China Emerg. Technol. (UCET)*, Scotland, U.K., Aug. 2020, pp. 1–5.

[14] C. Bertoncini, K. Rudd, B. Nousain, and M. Hinders, "Wavelet fingerprinting of radio-frequency identification (RFID) tags," *IEEE Trans. Ind. Electron.*, vol. 59, no. 12, pp. 4843–4850, Dec. 2012, doi: 10.1109/TIE.2011.2179276.

[15] A. Candore, O. Kocabas, and F. Koushanfar, "Robust stable radiometric fingerprinting for wireless devices," in *Proc. IEEE Int. Workshop Hardw.-Oriented Secur. Trust*, Washington, DC, USA, Jul. 2009, pp. 43–49.

[16] J. Wei and N. Li, "Privacy-preserving and undeniable authentication for mobile RFID tags," in *Proc. IEEE 20th Int. Symp. 'World Wireless, Mobile Multimedia Netw.' (WoWMoM)*, Jun. 2019, pp. 1–6.

[17] M. S. Farash, O. Nawaz, K. Mahmood, S. A. Chaudhry, and M. K. Khan, "A provably secure RFID authentication protocol based on elliptic curve for healthcare environments," *J. Med. Syst.*, vol. 40, no. 7, pp. 1–7, Jul. 2016, doi: 10.1007/s10916-016-0521-6.

[18] A. Ibrahim and G. Dalkilic, "An advanced encryption standard powered mutual authentication protocol based on elliptic curve cryptography for RFID, proven on WISP," *J. Sensors*, vol. 2017, Aug. 2017, Art. no. 2367312, doi: 10.1155/2017/2367312.

[19] J. P. Kaps, "Chai-tea, cryptographic hardware implementations of xtea," in *Proc. Int. Conf. Cryptol. India*, 2008, pp. 363–375.

[20] M. Hosseinzadeh *et al.*, "An enhanced authentication protocol for RFID systems," *IEEE Access*, vol. 8, pp. 126977–126987, 2020, doi: 10.1109/ACCESS.2020.3008230.

[21] Y. J. Huang, C. C. Yuan, M. K. Chen, W. C. Lin, and H. C. Teng, "Hardware implementation of RFID mutual authentication protocol," *IEEE Trans. Ind. Electron.*, vol. 57, no. 5, pp. 1573–1582, May 2010, doi: 10.1109/TIE.2009.2037098.

[22] H. Gilbert and M. Robshaw, "Active attack against HB^{+}: A provably secure lightweight authentication protocol," *Electron. Lett.*, vol. 41, no. 21, pp. 1169–1170, 2005, doi: 10.1049/el:20052622.

[23] T.-C. Yeh, Y.-J. Wang, T.-C. Kuo, and S.-S. Wang, "Securing RFID systems conforming to EPC class 1 generation 2 standard," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 7678–7683, Dec. 2010, doi: 10.1016/j.eswa.2010.04.074.

[24] S. Wang, S. Liu, and D. Chen, "Security analysis and improvement on two RFID authentication protocols," *Wireless Pers. Commun.*, vol. 82, no. 1, pp. 21–33, May 2015, doi: 10.1007/s11277-014-2189-x.

[25] N. Pisharoty, "PICO: An ultra lightweight and low power encryption design for ubiquitous computing," *Defence Sci. J.*, vol. 66, no. 3, pp. 259–265, 2016, doi: 10.14429/dsj.66.9276.

[26] P. Peris-Lopez, J. C. Estevez-Tapiador, J. M. Estevez-Tapiador, and A. Ribagorda, "M^2AP: A minimalist mutual-authentication protocol for low-cost rfid tags," in *Ubiquitous Intelligence and Computing*. Berlin, Germany: Springer, 2006, pp. 912–923.

[27] P. Peris-Lopez, P. Hernandez-Castro, and J. C. Estevez-Tapiador, "LMAP: A real lightweight mutual authentication protocol for low-cost RFID tags," in *Ubiquitous Intelligence and Computing*. Berlin, Germany: Springer, 2006, pp. 12–14.

[28] P. Peris-Lopez, P. Hernandez-Castro, and J. C. Estevez-Tapiador, "EMAP: An efficient mutual-authentication protocol for low-cost RFID tags," in *Proc. Move Meaningful Internet Syst., OTM Workshops*, 2006, pp. 352–361.

[29] H.-Y. Chien, "SASI: A new ultralightweight RFID authentication protocol providing strong authentication and strong integrity," *IEEE Trans. Dependable Secure Comput.*, vol. 4, no. 4, pp. 337–340, Oct./Dec. 2007, doi: 10.1109/TDSC.2007.70226.

[30] S. D. Kumar, H. Thapliyal, and A. Mohammad, "EE-SPFAL: A novel energy-efficient secure positive feedback adiabatic logic for DPA resistant RFID and smart card," *IEEE Trans. Emerg. Topics Comput.*, vol. 7, no. 2, pp. 281–293, Apr./Jun. 2019, doi: 10.1109/TETC.2016.2645128.

[31] M. Todd, W. Burleson, and R. Tessier, "The design and assessment of a secure passive RFID sensor system," in *Proc. IEEE 9th Int. New Circuits Syst. Conf.*, Quebec, QC, Canada, Jun. 2011, pp. 26–29, doi: 10.1109/NEWCAS.2011.5981327.

[32] P. Israsena, "Design and implementation of low power hardware encryption for low cost secure RFID using TEA," in *Proc. 5th Int. Conf. Inf. Commun. Signal Process.*, Bangkok, Thailand, 2005, pp. 6–9, doi: 10.1109/ICICS.2005.1689288.

[33] J. Wang, H. Li, and F. Yu, "Design of secure and low-cost RFID tag baseband," in *Proc. Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Shanghai, China, Sep. 2007, pp. 21–25, doi: 10.1109/WICOM.2007.516.

[34] H. Cai, G. Wang, X. Shi, J. Xie, and C. Qian, "When tags 'read' each other: Enabling low-cost and convenient tag mutual identification," in *Proc. IEEE 27th Int. Conf. Netw. Protocols (ICNP)*, Chicago, IL, USA, Oct. 2019, pp. 1–11.

[35] S. C. G. Periaswamy, D. R. Thompson, and J. Di, "Fingerprinting RFID tags," *IEEE Trans. Depend. Sec. Comput.*, vol. 8, no. 6, pp. 938–943, Nov./Dec. 2011, doi: 10.1109/TDSC.2010.56.

[36] A. Gouissem, K. Abualsaud, E. Yaacoub, T. Khattab, and M. Guizani, "Hybrid physical layer security for passive RFID communication," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Xiamen, China, Dec. 2020, pp. 150–156, doi: 10.1109/CSCI51800.2020.00033.

[37] B. A. Alsaify, D. R. Thompson, and J. Di, "Exploiting hidden Markov models in identifying passive UHF RFID tags," in *Proc. IEEE Radio Wireless Symp. (RWS)*, Newport Beach, CA, USA, Jan. 2014, pp. 259–261.

[38] W. C. Suski, M. A. Temple, M. J. Mendenhall, and R. F. Mills, "Using spectral fingerprints to improve wireless network security," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, New York, NY, USA, Nov. 2008, pp. 1–5.

[39] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Mach. Learn.*, vol. 53, nos. 1–2, pp. 23–69, Oct. 2003.

[40] A. Satorra and P. M. Bentler, "A scaled difference chi-square test statistic for moment structure analysis," *Psychometrika*, vol. 66, no. 4, pp. 507–514, 2001, doi: 10.1007/BF02296192.

[41] H. Wu, X. Wu, Y. Li, and Y. Zeng, "Collision resolution with FM0 signal separation for short-range random multi-access wireless network," *IEEE Trans. Signal Inf. Process. over Netw.*, vol. 7, pp. 438–450, 2021, doi: 10.1109/TSIPN.2021.3093000.

Haifeng Wu was born in Kunming, Yunnan, China, in 1977. He received the M.S. degree in electrical engineering from Yunnan University, Kunming, China, in 2004, and the Ph.D. degree in electrical engineering from Sun Yat-Sen University, Guangzhou, China, in 2007.

He is currently a Professor with the Department of Information Engineering, Yunnan Minzu University, Kunming. His research interests include neural signal processing, machine learning, and mobile communications.

Wei Gao was born in Heze, Shandong, China, in 1998. He received the B.E. degree in communication engineering from Inner Mongolia University of Science and Technology, Baotou, China, in 2016. He is currently pursuing the M.S. degree with the Department of Information Engineering, Yunnan Minzu University, Kunming, China.

His research interests include mobile communications and wireless near-field communication.

Chongrong Pu was born in Chuxiong, Yunnan, China, in 1996. He received the B.E. degree in electronic information engineering from Yunnan Minzu University, Kunming, Yunnan, China, in 2020, where he is currently pursuing the M.S. degree with the Department of Information Engineering.

His research interests include UHF radio frequency identification (RFID) communication.

Zeng Yu was born in Kunming, Yunnan, China, in 1981. She received the M.S. degree in electrical engineering from Yunnan University, Kunming, China, in 2006.

She is currently an Assistant Professor with the Department of Information Engineering, Yunnan Minzu University, Kunming. Her research interests include wireless networks and mobile communications.