Received 12 May 2023; revised 18 June 2023; accepted 7 July 2023. Date of publication 14 July 2023; date of current version 4 August 2023.

Digital Object Identifier 10.1109/OJCOMS.2023.3295379

Deep Learning-Based RF Fingerprint Identification With Channel Effects Mitigation

HUA FU [1,2] (Member, IEEE), LINNING PENG [1,2] (Member, IEEE), MING LIU [3] (Member, IEEE), AND AIQUN HU [2,4] (Senior Member, IEEE)

[1] School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China

[2] Frontier Crossing Scientific Research Center, Purple Mountain Laboratories for Network and Communication Security, Nanjing 211111, China

[3] Engineering Research Center of Network Management Technology for High Speed Railway of Ministry of Education,
School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

[4] School of Information Science and Engineering, Southeast University, Nanjing 210096, China

CORRESPONDING AUTHOR: H. FU (e-mail: hfu@seu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB4300300; in part by the National Natural Science Foundation of China under Grant 62001106, Grant 62171120, and Grant 61971029; in part by the Natural Science Foundation of Jiangsu Province under Grant BK20200350; in part by the Guangdong Key Research and Development Program under Grant 2020B0303010001; in part by the Jiangsu Provincial Key Laboratory of Network and Information Security under Grant BM2003201; and in part by the Purple Mountain Laboratories for Network and Communication Security.

ABSTRACT The radio frequency fingerprint (RFF)-based device identification is a promising physical layer authentication technique. However, the wireless channel significantly affects the RFF features of the wideband wireless devices. In this paper, we extensively investigate the impact of channel variation on RFF identification using 20 MHz IEEE 802.11 signal. A time domain least mean square (LMS) equalization-based feature extraction method has been proposed. This method progressively restores the transmitted signal and preserves more details of RFF features than the classical frequency domain equalization (FDE) method. Moreover, a hybrid identifier is proposed to take advantage of both LMS-based and FDE-based methods. With the equalized samples, a four-layer convolutional neural network is designed for device identification. An experimental system has been set up to capture the waveform of 68 802.11 devices at different positions. The experimental results show that the LMS-based method outperforms others when the acquisition positions of the training dataset are the same as those of the testing dataset. On the other hand, the FDE-based method is shown to be more effective when the acquisition positions of the training dataset do not fully include those of the testing dataset. Moreover, the hybrid identifier achieves an improvement of 2% for overall identification accuracy.

INDEX TERMS Physical layer security, radio frequency identification, convolutional neural network, least mean square equalization, multipath fading.

I. INTRODUCTION

ACCESS authentication is a critical security concern for Internet of Things (IoT), particularly in the case of massive machine type communications. The classical access authentication techniques involve the utilization of media access control (MAC) address, pre-shared key or digital certificate. However, the MAC address is vulnerable to tampering, and updating the pre-shared key in IoT devices is often challenging. Additionally, cryptography-based certification mechanisms typically demand additional computational resources, making them unsuitable for energy-constrained IoT devices.

The radio frequency fingerprint (RFF)-based device identification is an emerging physical layer (PHY) authentication technique for wireless communication systems [1], [2], [3], [4]. Due to the minor hardware manufacturing variations among different transmission devices, the transmitted radio frequency (RF) waveform carries the inherent features

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

of the device. These unique and persistent features can be regarded as the "fingerprint" of a device [5], [6]. Many works in the literature have demonstrated the success of the RFF-based device identification for IoT devices, such as for Bluetooth [7], Wi-Fi [8], [9], [10], [11], [12], WiMax [13], ZigBee [14], [15] and LoRa [16], [17] specifications. In most of the works, the main concerned challenge of the RFF-based identification is the degradation of signal strength due to the propagation distance.

However, wireless channel effect is also a crucial challenge for RFF-based identification of wireless device, especially for wideband communication systems [9], [10], [12]. For instance, IEEE 802.11a system works at 5 GHz band with a signal bandwidth of 20 MHz [18]. As a result, the signal is susceptible to multipath fading in the indoor environment. Additionally, considering that the half-wavelength of the signal is about 3 cm, the wireless channel fading can be considered uncorrelated when the device is moved a short distance. Consequently, it becomes crucial to mitigate the impact of channel fading on RFF features.

Several works have demonstrated that the carrier frequency offset (CFO) can be used as an important feature for RFF identification [8], [9], [14], [16], [19]. However, as the CFO is a time-varying feature [20], [21], the stability of the CFO in real wireless communication systems remains an issue to be studied. A location-invariant RFF feature extraction approach, referred to as amplitude of quotient (AoQ), has been proposed in [19] for Wi-Fi devices, and the performance has been verified with different locations. The experimental results in [10] show that the wireless channel fading impacts the identification accuracy significantly from 85% to 9% in the experimental dataset. Moreover, equalizing the I/Q data can increase the identification accuracy to 23% [10]. An RFF extraction method using undercomplete demodulation has been proposed in [9], which consists in reapplying the frequency and sampling offsets after channel equalization. An accuracy of 80%∼95% can be achieved for Wi-Fi devices and universal software radio peripheral (USRP) radios in static environment. The authors in [9], [22] have demonstrated that adding active RFF feature can effectively combat the impact of wireless channel. However, the implementation of such method requires changing the constellation pattern or adding digital filter at the terminal device, which brings additional cost. In order to overcome the effects of time-varying channel on RFF features, a data augmentation method is used in [23], [24]. Many simulated channel variations are used to generate a large number of training samples, which improves the accuracy of identification. This method requires a priori knowledge about the channel model of the transmission environment, and the large number of samples also increases the complexity of model training. For multiple-input multiple-output (MIMO) system, the authors in [25] propose to use blind channel estimation to remove the channel effects. The classification accuracy has been tested with simulated RF waveforms and Rayleigh channel.

In fact, the research on the RFF extraction in presence of channel fading has great importance. The features that many traditional RFF methods extract contain both unique hardware impairments and channel effects and sometimes the channel effect may be dominant in the overall signal representation [9]. The extracted RFF will become outdated once the channel condition changes. In other words, the RFF is highly location-dependent. Therefore, it is crucial to have a reliable RFF extraction method that can overcome the impact of channel effect. This will greatly improve the practicality of the RFF identification technique and lead to a much improved network security condition. For instance, RFF identification can be used for access control of loT devices. Gateway devices can authenticate access devices by analyzing the physical layer signals and detect potential spoofing or distributed denial of service attacks using RFF. RFF identification can also be used for zero-trust authentication. By performing RFF identification for each frame of the received signal, a highly secure loT system can be constructed.

In this work, we focus on the RFF identification of IEEE 802.11 a/g/n devices under different wireless channel conditions. The main contributions of this paper are listed as follows:

- A time domain least mean square (LMS) equalization-based RFF extraction method has been proposed. Compared to the existing methods, the LMS-based method shows the best performance and reaches an accuracy of 99.34% when the target device is trained and tested at the same position, which means that this method preserves more detailed RFF features. However, when the training and testing datasets are collected from different positions, the frequency domain equalization (FDE)-based RFF extraction method [9] is shown to be more effective in eliminating the channel effects. Therefore, a hybrid identifier is proposed to combine the advantages of both the FDE and LMS approaches. In this identifier, two convolutional neural networks (CNNs) are trained with the FDE and LMS processed sequences, respectively, and the identification result is decided by a similarity test between the channel state information (CSI) of the testing frame and the CSI pre-stored in training process. Experimental results show that, using the proposed hybrid identifier, the overall identification accuracy can be enhanced by 2%.
- A four-layer deep learning CNN has been designed for RFF identification. The performance of the proposed CNN has been investigated with different RFF extraction methods. Compared to the existing CNNs, the proposed CNN shows the benefits on training speed and identification accuracy for the FDE and LMS methods, which means that the proposed CNN is useful in distinguishing the subtle features of equalized I/Q samples.
- An experimental system has been set up with different IEEE 802.11 devices and USRP N210 receiver for raw

I/Q sample capture. The I/Q samples are collected from 64 802.11 access point (AP) routers of 6 manufactures and 4 mobile phones, and the receiver has been placed at 4 different positions. The wireless channel fluctuates due to the people moving during the experiments. This real measured database can be useful for the study and research of the wireless channel influence on RFF identification.

The remainder of this paper is organized as follows. In Section II, the literature of RFF identification has been reviewed. In Section III, the 802.11 training symbol and the RFF model have been introduced. In Section IV, the RFF identification framework is presented, including signal preprocessing, channel equalization and CNN architecture. The experimental system and identification results are presented in Section V. Section VI concludes the paper.

II. RELATED WORK

The RFF identification technique typically comprises two aspects, feature extraction and training/identification. Regarding the waveform segment used, the RFF feature extraction can be categorized into transient feature extraction and steady-state feature extraction.

The transient features are extracted during the on-off transient period of waveform [5]. For some IoT devices, transient features also exist at the beginning of steady-state period, which can also be treated as region of interest for RFF identification [26]. The greatest advantage of transient feature extraction is that it does not require any prior information. Therefore, this feature extraction method is largely used in Radio [27], GSM [28], [29], Bluetooth [7], [30], Wi-Fi [8], [31] and ZigBee [32] device identification. However, the accuracy of signal acquisition poses a challenge for transient feature extraction, which requires high-cost receivers such as spectrum analyzer or digital storage oscilloscope [7], [28], [29], [30], [32].

The steady-state features are extracted from the transmitted symbols [5]. The available sample length of steady-state feature is longer than that of the transient feature. Therefore, the steady-state features are more stable. The steady-state feature extraction method is largely used in GSM and LTE mobile communications [33], [34], Bluetooth [35], Wi-Fi [6], [36], [37], [38], [39], WiMax [40], ZigBee [14], [15], [41], [42], LoRa [43], and RFID [44], [45] device identification. Multi-dimensional RFF features can be extracted from the steady-state signal, which enables the use of multi-dimensional classifier and leads to high identification accuracy. Hence, more and more works use steady-state signal for RFF feature extraction.

In wireless communication systems, the source of RFF features includes I/Q imbalance [1], CFO from crystal oscillator deviation [1], amplifier non-linearity [46], etc. The CFO feature can be estimated directly from the received signal. However, most of the steady-state features are mixed together in waveform and very hard to be separated. Therefore, the RFF features need to be extracted using different methodologies. The most commonly used RFF features include waveform feature, modulation feature and transform domain feature.

The waveform feature consists in using directly the received I/Q samples as the RFF feature. For ZigBee systems, the offset quadrature phase shift keying (O-QPSK) modulated samples can be used as the input of deep-learning CNN for RFF identification [42]. The I/Q samples are precisely synchronized in order to eliminate the CFO. The IEEE 802.11a/g/n pilot/training waveform is employed for RFF identification in [9], [10].

The modulation feature refers to the estimated modulation parameters extracted from the received waveform. Modulation features include I/Q imbalance [33], [39], [47], [48], [49], CFO [7], [8], [14], [33], [35], [38], [39], [41], [43], [50], [51], sampling offset [8], amplifier non-linearity [52], etc. Among these features, the CFO is the most important one for RFF identification, because it is relatively stable with serious noise and complex wireless channel conditions [14]. However, the most notable drawback of CFO is that its value can change with time or environmental variations [14], [41], [51].

The transform domain feature consists in transforming the captured time domain I/Q samples into frequency domain or other transform domain. A straightforward method is the spectrum-based RFF identification via fast Fourier transform (FFT) [31], [34], [46] or wavelet transform [6]. Recently, other transform domains such as Hilbert-Huang transform (HHT) [29], [34], [53], [54], [55], [56], [57] are applied for RFF identification. It is worth noting that, the authors of [46] have found that the spectrum-based RFF feature is very sensitive to the wireless channel variation. The RFF identification accuracy dramatically decreases when the location of the target device changes. The authors of [58], [59] use the tap coefficients of the LMS filter as the RFF feature for ZigBee devices. The performance has been tested with proximate line-of-sight channel.

Once the RFF features are extracted, the RFF identification problem transforms into a conventional machine learning classification problem. In the early studies of RFF identification, many machine learning algorithms have been used, including multiple discriminant analysis (MDA), k-nearest neighbor (KNN), random forest (RndF) [60], support vector machine (SVM) [43], [61], etc. However, the accuracy of identification using these algorithms is limited, particularly when the number of target devices increases. In recent studies, deep learning based CNN has been widely used for RFF identification [42], [62]. The waveform feature can be used as the input of CNN as 1-dimensional (1D) sequence [9], [10], [42]. Some modulation features and transform domain features can be used as the input of CNN as 2-dimensional images [63], [64]. The CNN-based classification methods show high accuracy even with a large number of target devices. However, training CNN to fully overcome the wireless channel variation still remains an unsolved problem.

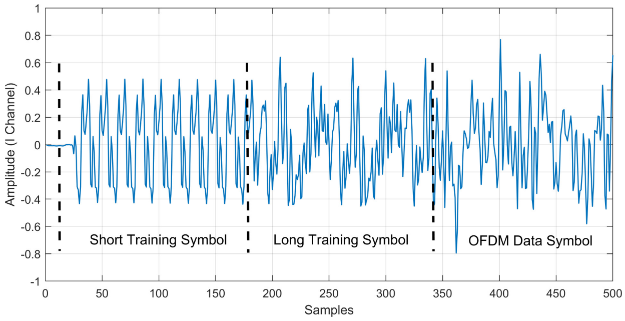


FIGURE 1. A sample of the preamble of IEEE 802.11 signal (in-phase component).

III. PRELIMINARY

A. IEEE 802.11 PACKET FORMAT

IEEE 802.11 specification adopts the orthogonal frequency division multiplexing (OFDM) modulation. The transmitted data bits are firstly mapped to phase shift keying (PSK) or quadrature amplitude modulation (QAM) symbols. These modulated symbols are then distributed across different frequency domain subcarriers and transformed into the time domain using the inverse fast Fourier transform (IFFT). The resulting time domain OFDM waveform is dependent on the combination of the frequency domain data symbols. However, due to the random nature of the information data, the signal waveform exhibits random variations, making it challenging to extract stable and data-independent RFF. In this paper, we will leverage the a priori information of the preamble of IEEE 802.11 system to extract RFF from the preamble signal part.

A captured preamble part of 802.11 signal is depicted in Fig. 1. A preamble consists of ten short training symbols (STSs) and two long training symbols (LTSs), resulting in a total duration of 16 μs [18]. The STS is mainly used for synchronization purpose and is generated from a frequency domain sequence

$$\begin{aligned}X_{S_{(0,63)}} = \sqrt{\frac{13}{6}} \times \{&0, 0, 0, 0, 0, 0, 0, 0, 1+j, 0, 0, 0, -1-j,\\ &0, 0, 0, 1+j, 0, 0, 0, -1-j, 0, 0, 0, -1-j, 0, 0,\\ &0, 1+j, 0, 0, 0, 0, 0, 0, 0, -1-j, 0, 0, 0, -1-j,\\ &0, 0, 0, 1+j, 0, 0, 0, 1+j, 0, 0, 0, 1+j, 0, 0, 0,\\ &1+j, 0, 0, 0, 0, 0, 0, 0\}. \end{aligned} \tag{1}$$

The time domain STS is obtained after a 64-point IFFT

$$x_S(t) = \mathrm{IFFT}_{64}\big(X_{S_{(0,63)}}\big). \tag{2}$$

In $X_{S_{(0,63)}}$ as defined in (1), one subcarrier of every four subcarriers carries non-zero symbol. This leads to a periodically changing time domain signal $x_S(t)$, cf. Fig. 1. For 20 MHz channel spacing, the sampling time is $T_s = 0.05$ μs. The corresponding IFFT period with 64 points is $T_{FFT} = 3.2$ μs. The length of one STS period is $T_{FFT}/4 = 0.8$ μs. The total length of ten STSs takes up 8 μs.

The LTS, on the other hand, is mainly used for fine CFO and channel estimation. Like STS, it is also generated from a frequency domain sequence

$$\begin{aligned}X_{L_{(0,63)}} = \{&0, 0, 0, 0, 0, 0, 1, 1, -1, -1, 1, 1, -1, 1, -1, 1, 1,\\ &1, 1, 1, 1, -1, -1, 1, 1, -1, 1, -1, 1, 1, 1, 1, 0, 1,\\ &-1, -1, 1, 1, -1, 1, -1, 1, -1, -1, -1, -1, -1,\\ &1, 1, -1, -1, 1, -1, 1, -1, 1, 1, 1, 1, 0, 0, 0, 0, 0\}\end{aligned} \tag{3}$$

where 53 subcarriers are loaded by non-zero symbols, while others being set to 0. The time domain LTS sequence is generated by a 64-point IFFT

$$x_L(t) = \mathrm{IFFT}_{64}\big(X_{L_{(0,63)}}\big). \tag{4}$$

Afterwards, two repetitions of $x_L(t)$ sequence, preceded by a guard interval (GI) form the overall LTS part. The duration of LTS part is $2 \times 3.2 + 1.6 = 8$ μs.

Overall, the preamble part of IEEE 802.11 specification can be expressed as

$$x = \big\{x_{S_{(32,63)}}, x_S, x_S, x_{L_{(32,63)}}, x_L, x_L\big\} \tag{5}$$

where $x_{S_{(32,63)}}$ and $x_{L_{(32,63)}}$ are the sequences of the 32th to 63th elements of x_S and x_L, respectively. As can be seen from the specification, the preamble consists of deterministic sequences and is known by the receiver. This a priori knowledge can be exploited to exclude the influence of data randomness in the RFF extraction.

B. RFF MODEL

The transmitted signal of an 802.11 device is affected by the hardware impairments of RF components, including local oscillator, mixer, filter and amplifier, which leave some traces of RFF in the signal. These RFF features can be identified in time and frequency domains. For instance, CFO often manifests as a continuous phase shift in the time domain signal. Similarly, the frequency distortion caused by the imperfect filter response can be observed in the frequency domain signal. Additionally, the band-limited and non-linearity effects of power amplifier exhibit memory effect and non-linear distortion in both time and frequency domain signal waveforms. In general, the 802.11 signal with RFF effects can be noted as

$$\hat{x}(t) = \tilde{x}(t) * h_{\mathrm{RFF}}(t) \cdot e^{-j\omega_{\mathrm{CFO}} t} \tag{6}$$

where $\tilde{x}(t)$ is the standard transmit signal $x(t)$ affected by the time domain RFF features including non-linear distortion, I/Q offset, etc. The operator $*$ denotes the convolution operation. $h_{\mathrm{RFF}}(t)$ denotes the response of comprehensive frequency domain RFF distortions. $e^{-j\omega_{\mathrm{CFO}} t}$ is the time domain phase rotation caused by CFO ω_{CFO}.

After passing the wireless channel, the received signal can be written as

$$y(t) = \hat{x}(t) * h_{\mathrm{Ch}}(t) + n \tag{7}$$

where $h_{\mathrm{Ch}}(t)$ is the wireless channel response, and n is the additive white Gaussian noise (AWGN) with variance σ_n^2.

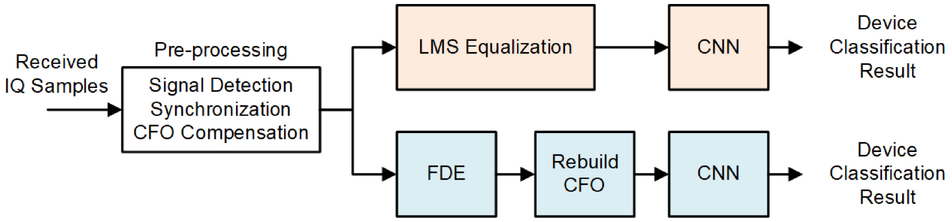


FIGURE 2. The RFF identification with LMS and FDE based channel fading compensation.

Obviously, the RFF features embedded in (6) is affected by the fading of wireless channel as shown in (7). This motivates us to investigate mechanism that can mitigate the influence of wireless channel when extracting the RFF from the received signal.

IV. RFF IDENTIFICATION FOR IEEE 802.11 DEVICES

This section proposes an RFF identification method that involves time and/or frequency domain approaches to mitigate the channel effect before extracting RFF. The global diagram of the RFF identification method is shown in Fig. 2. The received RF signal is first processed to detect the coarse start of frame, synchronize the time signal, and compensate the CFO. Then, the FDE or LMS based equalization is performed to compensate the channel fading effect in the received signal. Afterwards, the restored preamble signal is input into a 1D CNN for RFF extraction and device identification. Additionally, a hybrid identifier that aims to exploit the advantages of both FDE and LMS methods is proposed for further performance improvement. The details of the proposed method are presented in the following sections.

A. PREAMBLE PRE-PROCESSING

As mentioned in the previous section, in order to extract RFF, it is necessary to properly locate the preamble part in the received signal. We begin by detecting the coarse start of a signal frame from the captured raw data using a spectrum detection method. The received time domain signal within a 64-sample observation window starting from time t, is first converted to the frequency domain signal $Y^{(t)}$ by the 64-point FFT. Then, for each sampling time t, a metric is evaluated such that

$$D(t) = \frac{\sum_{m=1}^{12} Y_S^{(t)}(m)}{\sum_{n=1}^{52} Y_0^{(t)}(n)} \tag{8}$$

where the sequence $Y_S^{(t)}(m)$ corresponds to 12 elements of $Y^{(t)}$ with the indices that are same to the 12 non-zero elements in $X_{S_{(0,63)}}$. The sequence $Y_0^{(t)}(n)$ corresponds to 52 elements of $Y^{(t)}$ with the indices that are same to the 52 zero elements in $X_{S_{(0,63)}}$. When $D(t)$ is larger than a threshold γ, denoted as $D(t^\dagger) > \gamma$, then the signal $y(t)$, $t \in [t^\dagger - 30, t^\dagger + 420]$ is used for time synchronization, which contains the preamble part of the signal frame.

The time synchronization can be performed by computing the cross-correlation of the signal $y(t)$, $t \in [t^\dagger - 30, t^\dagger + 420]$, such that

$$t_{\mathrm{Syn}} = \arg\max_t \sum_{m=0}^{32} y(t+m)y^*(t+64+m) + y(t+64+m)y^*(t+128+m) \tag{9}$$

where the superscript a^* denotes the conjugate of a complex value a. With time synchronization, the preamble part of the signal frame is located, which corresponds to $y(t)$, $t \in [t_{\mathrm{Syn}}, t_{\mathrm{Syn}} + 319]$.

Moreover, the CFO can be estimated using STS. As the STS pattern repeats every 16 samples, the CFO can be derived from the phase difference between two samples that are 16 samples apart, such that

$$\hat{\Delta}_f = \frac{1}{2\pi} \frac{\sum_{m=0}^{143} \mathrm{angle}\big(y(t_{\mathrm{Syn}}+m)y^*(t_{\mathrm{Syn}}+m+16)\big)}{144 \times 16} \tag{10}$$

where angle$(\cdot)$ denotes the phase of a complex value. The phase difference is averaged over 144 samples to reduce the influence of noise.

As shown in (6) and (7), CFO causes a continuously changing phase in the received signal and therefore needs to be compensated before equalization. With the CFO $\hat{\Delta}_f$ estimated in (10), the phase shift in the received signal caused by CFO can be compensated as

$$\bar{y}(t) = y(t) \cdot e^{-j2\pi\hat{\Delta}_f t}, t \in \big[t_{\mathrm{Syn}}, t_{\mathrm{Syn}} + 319\big]. \tag{11}$$

After the aforementioned pre-processing, we obtain the raw IEEE 802.11 I/Q data samples containing the RFF features and the multipath channel fading effect. A straightforward RFF identification method has been proposed in [42] for ZigBee system, where the raw I/Q data samples were directly used for CNN training and identification. The learned features for identification included both RFF features and channel fading effects. However, wireless channel fading can seriously distort the wideband 802.11 signal and dominate the overall features of the received signal. An illustration of the STSs captured at four different locations is presented in Fig. 3 (a), and the CSIs measured using LTSs received at different locations are shown in Fig. 4 (a). It can be seen from Fig. 3 (a) that even when the same STSs are transmitted from the same device, the waveforms received at different locations are quite different from one another due to the distinct channel fading that the signal experienced as shown in Fig. 4 (a).

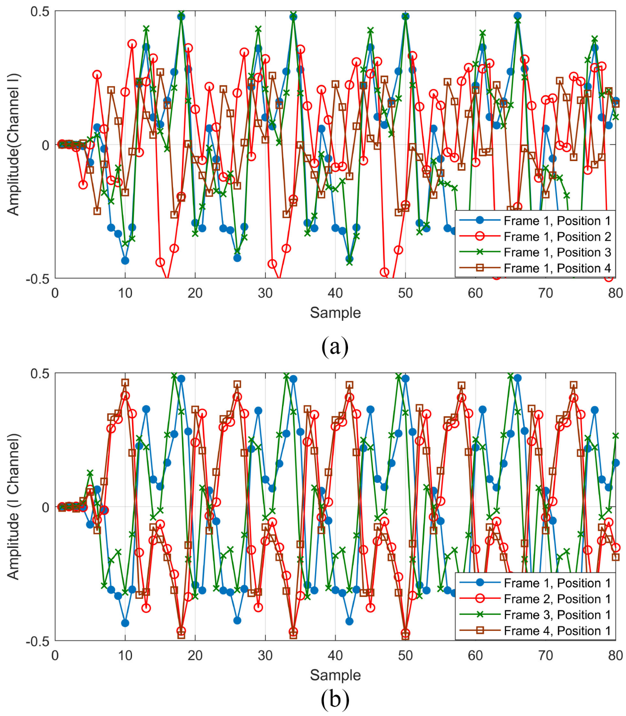


FIGURE 3. The first 80 STS samples received (a) at different locations and (b) at a fixed location.

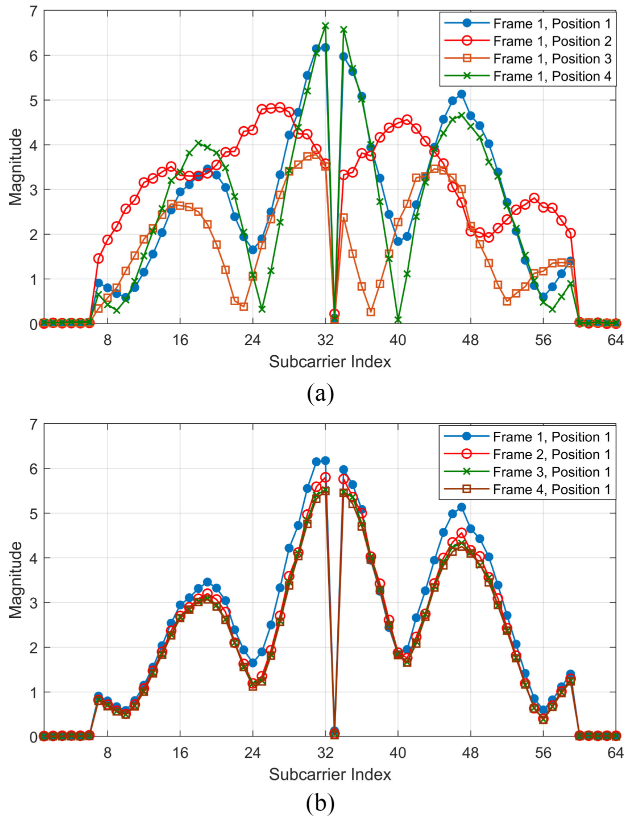


FIGURE 4. CSIs measured (a) at different positions and (b) at a fixed position.

Meanwhile, the received RF waveforms of different frames are also varying even observed at the same location. Fig. 3 (b) and Fig. 4 (b) illustrate four STS waveforms and estimated CSIs of one device, when the target device and the receiver do not change the location. It can be seen that, although the CSIs in Fig. 4 (b) are relatively stable, the STS waveforms in Fig. 3 (b) are still varying due to the residual synchronization error.

In summary, the raw IEEE 802.11 signal is not suitable for RFF identification due to channel fading and synchronization errors, even when the device is stationary. Therefore, we aim to explore effective channel mitigation method to ensure accurate RFF identification.

B. CLASSICAL FREQUENCY DOMAIN EQUALIZATION

The FDE is widely used in OFDM systems for the purpose of data recovery. However, according to (6) and (7), the received signal experiences the composite influence of RFF features and wireless channel. The FDE may cause the loss of some RFF features, such as the transmit filter response.

The frequency domain channel response can be obtained from LTS using the least square (LS) estimation

$$\hat{H} = \frac{\text{FFT}_{64}(\bar{y}_L)}{X_{L_{(0,63)}}} \tag{12}$$

where $\bar{y}_L$ is the received LTS after CFO compensation, and $X_{L_{(0,63)}}$ is the ideal LTS. Since there are two LTS sequences in the preamble, the channel estimation can be performed twice. We choose the average of these two estimations as the CSI, denoted as $\bar{H}$. Then, the received STSs can be equalized using $\bar{H}$, such that

$$\bar{z}_{S_{\text{FDE}}} = \text{IFFT}_{64}\left(\frac{\text{FFT}_{64}(\bar{y}_S)}{\bar{H}}\right) \tag{13}$$

where $\bar{y}_S$ is the received STSs with CFO compensation. After removing the guard interval, the 128-point equalized STS $z_{S_{\text{FDE}}}$ can be obtained with two times FDE. It is worth noting that the FFT and IFFT operations should be performed without CFO. Otherwise, the orthogonality of OFDM modulation will be broken, which also leads to inter-carrier interference in the post-equalization signal.

Moreover, since CFO is an important RFF feature for device identification, as in the work of [9], the CFO is rebuilt after FDE operation

$$\ddot{z}_{S_{\text{FDE}}}(t) = z_{S_{\text{FDE}}}(t) \cdot e^{j2\pi\hat{\Delta}_f t}. \tag{14}$$

The I/Q signal after FDE and CFO restoration will be used for the device identification via a neural network, as will be elaborated in Section IV-D.

In summary, while the FDE may affect certain RFF features such as the transmit filter response, removing channel fading from the signal is still beneficial for the RFF identification of wideband wireless devices. This will be demonstrated in the experimental results.

C. PROPOSED TIME DOMAIN EQUALIZATION

Due to the fact that FDE can result in the loss of RFF features in the equalized I/Q samples, it is crucial to preserve more RFF features while eliminating channel effects. This motivates us to propose a designated channel elimination method for RFF identification.

In this section, we propose a time domain equalization approach for RFF extraction in OFDM systems. The time domain LMS adaptive equalization is widely used in wired and wireless communication systems with single-carrier modulations. Thanks to its low complexity and iterative process, the LMS filter progressively equalizes the received signal, which helps to preserve more detailed RFF features of signal waveform.

The STS sequence is used for LMS equalization, because the shape of STS waveform is more regular than that of LTS, cf. Fig. 1, which helps the LMS filter to converge. Moreover, applying LMS equalization requires long training sequence to capture the channel response with high precision. Whereas, STS sequence of 802.11 specification is relatively short, namely 160 samples (ten STSs) which are not sufficient for LMS equalizer. To address this problem, a region of interest repetition (ROIR) method is proposed for the LMS equalization. When the channel variation is not fast, the channel fading can be regarded as constant for adjacent signal sequences. The ten repeated STSs can be considered to be affected by the similar channel fading effect. Therefore, it is feasible to repeat the received STSs and build a long sequence for LMS equalization. Note that the multipath fading may cause inter-symbol interference. Hence, we discard the first two STSs, i.e., the first 32 samples of $\bar{y}_S$, and use the subsequent eight STSs of 128 samples as the basis to build the longer STS sequence. The average power of sequence of 128 samples is normalized to 1. Then, this sequence of 128 samples is repeated twice and concatenated with another fragment of its first 64 samples. The resulting long STS sequence of 320 samples is represented as

$$y_{\text{LMS}} = \left\{ \bar{y}_{S_{(32,159)}}, \bar{y}_{S_{(32,159)}}, \bar{y}_{S_{(32,95)}} \right\}. \tag{15}$$

The sequence y_{LMS} is input into LMS adaptive filter. Accordingly, the desired sequence x_{LMS} of the LMS adaptive filter is generated from perfect STS

$$x_{\text{LMS}} = \{x_S, x_S, x_S, x_S, x_S\}. \tag{16}$$

The LMS equalization process is written as

$$z_{\text{LMS}}(k) = \sum_{i=0}^{L-1} \omega_i(k) \cdot y_{\text{LMS}}(k - L + i) \tag{17}$$

where z_{LMS} is the output of the LMS equalizer, k is the sample index, L is the filter length, $[\omega_0, \ldots, \omega_{L-1}]$ are the LMS filter weights and are updated iteratively as

$$\begin{aligned} \left[\omega_0(k+1), \ldots, \omega_{L-1}(k+1)\right] &= \left[\omega_0(k), \ldots, \omega_{L-1}(k)\right] \\ &+ \Delta\epsilon(k)\left[y_{\text{LMS}}(k-L), \ldots, y_{\text{LMS}}(k-1)\right] \end{aligned} \tag{18}$$

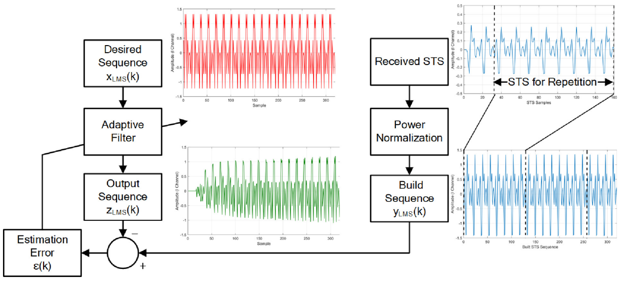


FIGURE 5. Time domain LMS equalization process.

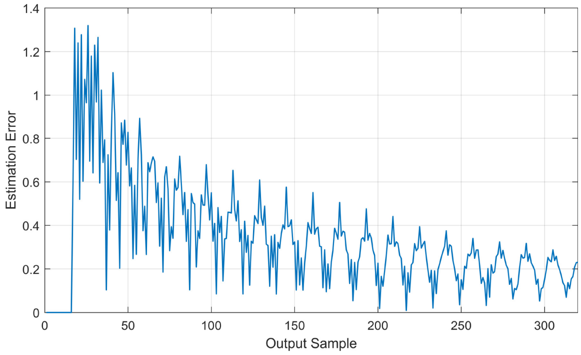


FIGURE 6. An illustration of LMS estimation error $\epsilon(k)$ for 320 samples.

where Δ is the step size of LMS filter, $\epsilon(k)$ is the estimation error

$$\epsilon(k) = x_{\text{LMS}}(k) - z_{\text{LMS}}(k). \tag{19}$$

Hence, the LMS algorithm involves $L + 1$ additions and $2L$ multiplications per iteration.

The process of time domain LMS equalization is presented in Fig. 5. An example of the estimation error is shown in Fig. 6. It can be seen from Fig. 6 that, the estimation error ϵ is high at the beginning stage and decreases gradually with the iteration process. A more detailed illustration is given in Fig. 7 where sequences of x_{LMS}, y_{LMS} and z_{LMS} are presented. After 320 points LMS time domain equalization, the channel effect has been removed and the final segment of z_{LMS} becomes similar to that of the desired signal x_{LMS}, cf. Fig. 7(b) in comparison with Fig. 7(a). This suggests that the quality of signal restoration is improving over time, and the influence of channel fading and RFF is removed gradually. We note that this progressive elimination of channel fading allow us to harvest the detailed signal features related to RFF. In other words, the signal z_{LMS} records the entire signal process results output from the LMS equalization, where the beginning part contains more raw I/Q signal information and therefore more RFF features, and the ending part achieves better signal restoration quality. The gradual changing of equalization quality provides us rich signal samples with a variety of channel effect mitigation levels. This diversity offers more information and flexibility for the RFF extraction and identification than the FDE-based method where the signal is uniformly equalized. In next section, a deep neural

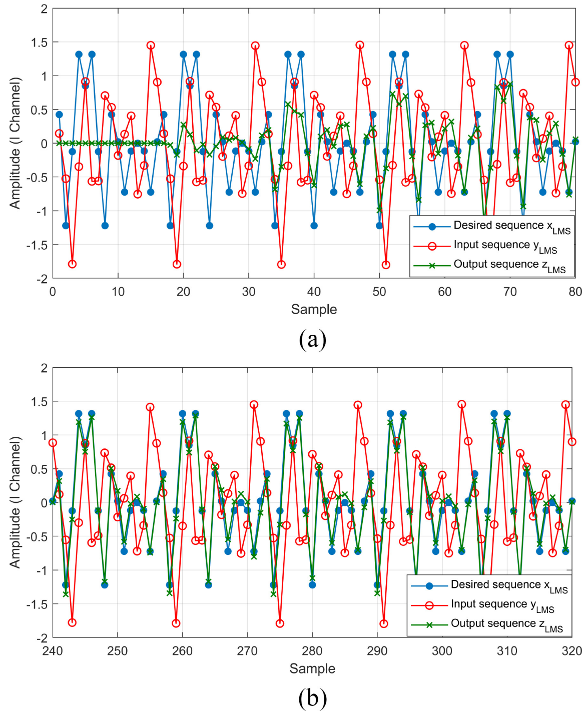


FIGURE 7. An illustration of (a) the first and (b) the last 80 samples of the desired perfect sequence x_{LMS}, the input sequence y_{LMS} and the output sequence z_{LMS}.

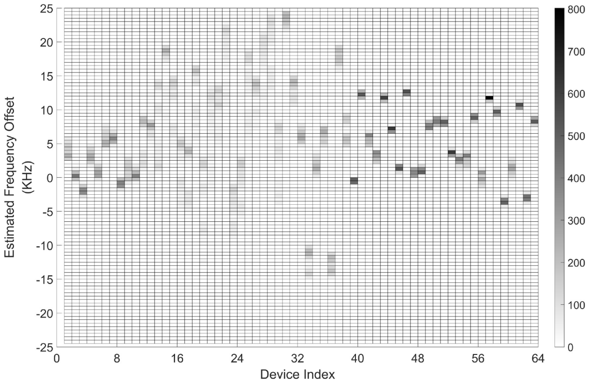


FIGURE 8. The measured CFO for the AP devices. The grayscale intensity of a grid represents the frequency of the CFO falls into an interval.

network based method will be proposed to implicitly extract the RFF features from the diversified signal samples and achieve device identification.

We note that, being different from the FDE-based method, no CFO restoration is performed for the LMS-equalized signal to avoid the influence of CFO variation over time. In IEEE 802.11 specification, the maximum transmit frequency tolerance is 20 ppm, which corresponds to a maximum of 48 KHz CFO for 2.4 GHz carrier frequency and 116 KHz CFO for 5.8 GHz carrier frequency, making CFO a potential source of RFF feature. The measured CFO of 64 AP devices is depicted in Fig. 8. From the experiments, it can be observed that the CFO varies within a range not exceeding the maximum transmit frequency tolerance. Whereas, for some devices, the CFO fluctuates during a long observation time. For example, the average CFO deviation of the devices no. 11 to 38 is around 5 KHz, which is sufficiently large to affect the accuracy of the CFO-based RFF identification. Therefore, in the proposed LMS-based method, we choose not to rebuild the CFO for equalized sequence z_{LMS}.

D. PROPOSED CNN-BASED RFF EXTRACTION AND IDENTIFICATION

After the equalization, the channel effect in the signal is mitigated and has less impact on the signal waveform. The next challenge is how to efficiently synthesize the RFF features from equalized signal. The waveform features that are useful for the RFF identification are difficult to be measured and expressed in a quantitative manner. Alternatively, the deep neural network can be employed to extract the useful RFF features and implicitly synthesize them to form high-level features that directly serve for device identification.

In this paper, we use 1D CNN to process the equalized time-domain I/Q samples. The 1D CNN can be immediately applied to I/Q data and performs well in identifying features regardless the location of such features within the data segment [9], [10]. The 1D CNN works like a filter that goes through the I/Q data sequence and extracts the "low-level" features that reflect the signal waveform and variation in different dimensions. The obtained features are consequently processed by several 1D convolution layers in a cascading manner to progressively yield more sophisticated and implicit features. Finally, a variety of high-level features output from convolution layers will be synthesized by fully connected layers to compute the possibilities that an input belongs to different classes, i.e., different devices.

The detailed architecture of the proposed 1D CNN is presented in Fig. 10. It consists of four convolution layers and two fully connected layers. Four 1D convolution layers are composed of 16, 32, 48 and 64 convolution kernels, respectively, with the filter size of 1×2. After the first three convolution layers, a rectified linear units (ReLU) activation function layer and a 1×2 maximum pooling (MaxPool) layer are used to reduce complexity and parameter number. In the last convolution layer, a ReLU activation layer and a 1×2 average pooling layer is used. After four convolution layers, the output is fed to a fully connected layer of the length 256, followed by a dropout layer with 0.5 dropout rate. Then, a fully connected layer of the size that is the same as the number of target devices and a softmax layer are used to calculate the possibilities that the input sample belongs to all the classes. Finally, the class with highest possibility is output as the identification result. During the training stage, we adopt the stochastic gradient descent with momentum (SGDM) optimizer, with a momentum value of 0.9. The training process consists of a maximum of 30 epochs with a mini-batch size of 256. The learning rate is set to 0.02 and the factor for dropping the learning rate is 0.1 with a number of epochs for dropping the learning rate set to 9.

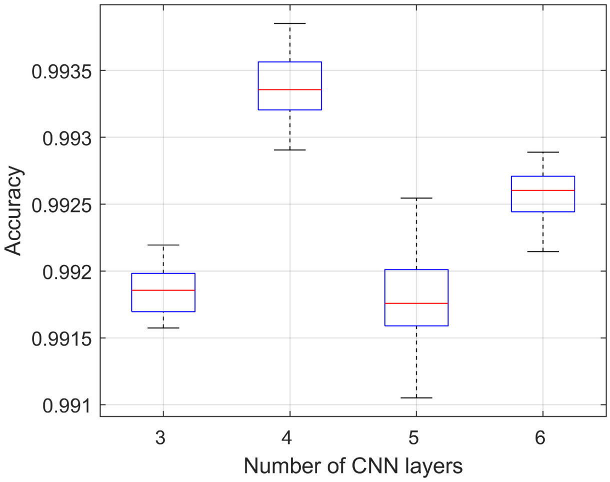


FIGURE 9. Classification performance of different numbers of CNN layers for LMS method.

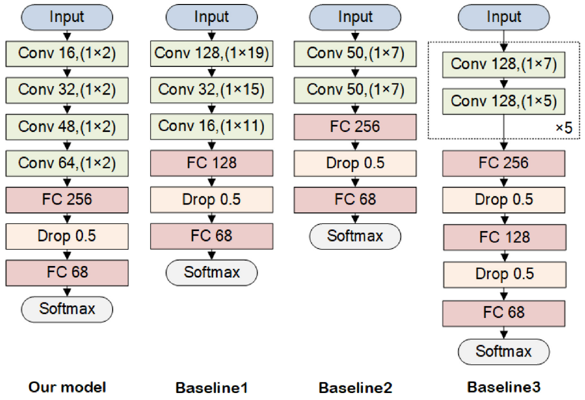


FIGURE 10. The CNNs for RFF identification.

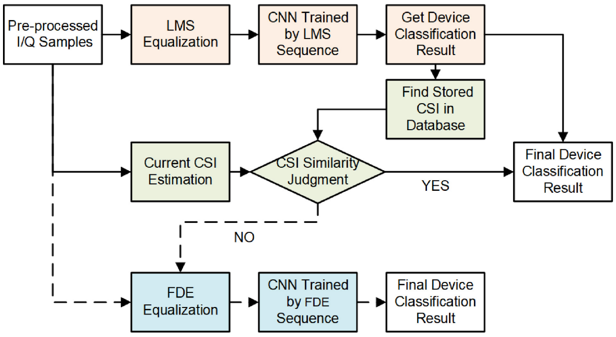

Pre-processed I/Q Samples
LMS Equalization
CNN Trained by LMS Sequence
Get Device Classification Result
Find Stored CSI in Database
Current CSI Estimation
CSI Similarity Judgment
YES
Final Device Classification Result
NO
FDE Equalization
CNN Trained by FDE Sequence
Final Device Classification Result

FIGURE 11. An illustration of the proposed hybrid identifier.

For the proposed CNN network, the influence of different numbers of network layers has been analyzed. The classification performance of 3-layer, 4-layer, 5-layer, and 6-layer CNN networks has been tested. The results are presented in Fig. 9 using box plot with 20 repetitions. It is worth noting that as the number of network layers increases, the training time also increases. The results indicate that designing a 4-layer CNN network achieves an effective balance between performance and complexity.

To evaluate the performance of the proposed CNN-based RFF identification method, a few CNN structures proposed in the literature are also implemented as baseline. *Baseline1* model [42] is a first CNN that exploits raw I/Q samples for RFF identification. It is composed of three 1D convolution layers with the kernel size of 1 × 19, 1 × 15 and 1 × 11, respectively. The number of kernels is 128, 32 and 16 for three convolution layers, respectively. *Baseline2* model proposed in [10] is a simpler CNN with one less convolution layer and fewer convolution kernels. The kernel size is 1 × 7 and the number of kernel is 50 for both convolution layers. *Baseline3* model [10] adopts a more complex network architecture. The block of two 1D convolution layers (consisting of a layer of 128 1 × 7 convolution kernels and a layer of 128 1 × 5 convolution kernels) is repeated five times, which yields a feature extraction module of ten convolution layers. Two dropout layers with dropout rate of 0.5 are employed to reduce overfitting.

E. HYBRID IDENTIFIER

In previous sections, we have introduced the FDE and LMS methods to equalize the received signal, which can offer different levels of channel effect elimination and RFF feature preservation. As will be demonstrated later in Section V-B, these two methods have their own favored application scenarios. When the training and testing signal samples experience similar CSI conditions, the LMS method shows higher accuracy in identification. On the other hand, when the training and testing signal samples experience different CSI conditions, the FDE method is more effective to remove the channel effects and yields a more accurate identification performance, although a considerable part of RFF features is eliminated. This observation inspires us to design a hybrid identifier that combines these two approaches, aiming to achieve good performance in all scenarios. The block diagram of the hybrid identifier is presented in Fig. 11.

The basic idea is to adaptively select the equalization method according to the CSI condition. Specifically, two CNN-based identifiers are trained during the training stage using the I/Q samples obtained from the FDE and LMS equalizers, respectively. Moreover, a database is established to store a subset of estimated CSIs along with the corresponding device identities for future reference. During the testing stage, the received signal samples are first processed by the LMS equalizer. Then, the equalized signal is input into the pre-trained CNN based identifier. The identification result is then used to retrieve the corresponding CSIs from the database. These retrieved CSIs are then compared with the CSI of the current signal. Once the current CSI is found to be similar to any of the CSIs in the database, it can be concluded that the current signal has experienced a similar CSI as the signal samples used for training. In such case, the LMS-based method is expected to provide better RFF extraction and identification performance. Therefore, the identification result given by the LMS-based method will be output as the final result. Otherwise, if the current CSI is determined as different from any of the CSIs stored in the database, the received I/Q signal will be processed using the FDE-based method, and the identification result provided by the FDE method will be output as the final result. By combining the LMS and FDE methods in this manner, a hybrid identifier can be constructed, allowing us to leverage the advantages of both approaches and achieve an improved overall RFF identification performance. The complexity of the hybrid model mainly comes from the computation of LMS algorithm and CNN. However, when the processing of the model is implemented in hardware, such as field programmable gate array (FPGA), the processing speed can be significantly improved.

V. EXPERIMENTAL RESULTS

A. DATA COLLECTION

A software defined radio (SDR) based hardware system is designed and implemented to capture the IEEE 802.11 signal frame. A USRP N210 with CBX daughterboard is employed as the signal receiver [65]. We use 64 Wi-Fi AP routers with five brands and six models, including Huawei$^{\text{TM}}$ WS5100 (2.4 GHz/5.8 GHz dual-band), Huawei$^{\text{TM}}$ WS5200 (2.4 GHz/5.8 GHz dual-band), Xiaomi$^{\text{TM}}$ R4A (2.4 GHz/5.8 GHz dual-band), TP-LINK$^{\text{TM}}$ TL-WDR4310 (2.4 GHz/5.8 GHz dual-band), MERCURY$^{\text{TM}}$ MW305R (2.4 GHz band only) and Dlink$^{\text{TM}}$ DWL-2000AP+A (2.4 GHz band only). Some AP models only work at 2.4 GHz band, and other dual-band AP models also work at 5.8 GHz band. It is worth noting that the IEEE 802.11 standard defines that both beacon and data frames using OFDM modulation at 5.8 GHz band. However, the beacon frame at 2.4 GHz band use direct sequence spread spectrum (DSSS) modulation with forward compatibility considerations. To capture the OFDM data frames at 2.4 GHz band, we use four additional mobile phones, including Huawei$^{\text{TM}}$ Honor 6, Honor V10, Honor 20 and Mate 20, to connect to the AP routers. A simple packet Internet groper (PING) hypertext transfer protocol (HTTP) application has been designed for mobile phones to trigger the transmission of data frames. Because the signal of mobile phones are also captured by the USRP, the number of target devices for RFF identification comes to 68. The RF signal is captured at a sampling rate of 20 MSamples/s. In order to reduce the size of the dataset, we use the first 200 frames of each signal capture, i.e., each 0.3 second raw I/Q samples. We also limit the maximum number of frames for each device by 10,000. The frames are stored based on the time of reception. For each device, we take one frame out of every five frames to build the training and validation set. Hence, there are 20% frames are used for training and validation, 80% frames are used for testing.

To obtain the IEEE 802.11 data frames with different channel conditions, we put the USRP N210 receiver on a trolley cart and move it in a room of 8.8 m × 8.8 m. The received signal suffers from serious multipath fading in the indoor environment. In each experiment, 10 AP routers are fixed on the edges around the room. The receiver has been placed at four different positions. After collecting the data, we place another set of 10 AP routers and repeat the process until all the APs are collected. All data is collected on the same day. An illustration of the experimental setup is shown in Fig. 12.

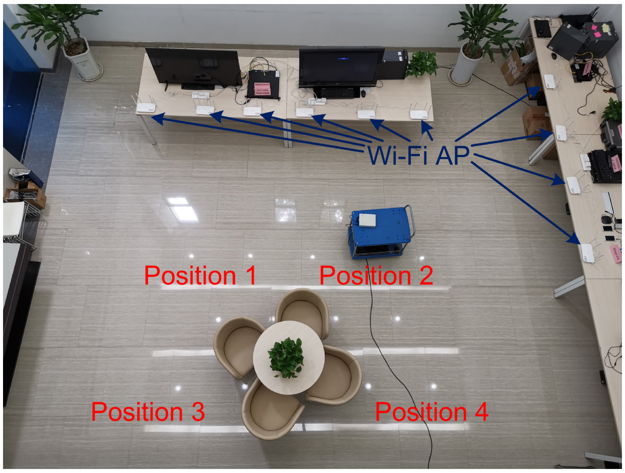


FIGURE 12. An illustration of the experimental setup with four receiver locations.

B. EXPERIMENTAL SETUP

To evaluate the performance of the proposed RFF identification scheme, two experimental scenarios have been considered. In the first scenario, the RFF identifier is trained and tested with the data samples acquired from the same positions. This scenario represents the typical IoT applications with fixed positions, such as sensor networks and smart appliances. In the second scenario, the data samples used to train the RFF identifier are obtained from locations that are different from the data samples used to test the identifier. It indicates that the identifier will be tested with the data samples from unknown position. This scenario represents the communications with indoor mobility. In the sequel, the two scenarios will be referred to as "*known position*" and "*unknown position*", respectively.

The proposed RFF identification based on the LMS equalizer is evaluated with a LMS filter length of $L = 24$. It is referred to as "*LMS*" in the sequel. The RFF identification methods based on the raw I/Q sample with only CFO compensation $\bar{y}_S$ (referred to as "*Raw*" [42]), the quotient of two LTSs [19] (referred to as "*AoQ*"), and the output of FDE with rebuild CFO $\ddot{z}_{S_{\text{FDE}}}$ (referred to as "*FDE*" [9]) are also evaluated for comparison. These features are extracted from the training dataset and stored separately to train their respective CNN models.

The CNN models are trained using MATLAB Deep Learning Toolbox with NVIDA GTX1070 GPU. The training time are presented in Table 1. It can be seen that the proposed CNN model requires a training time of less than 5 minutes, which is much faster than that of the *Baseline3*

TABLE 1. Training time for different CNN models using *LMS* equalized signal.

Model	*Baseline1*	*Baseline2*	*Baseline3*	*Our model*
Time	18'17"	6'52"	87'30"	**4'52"**

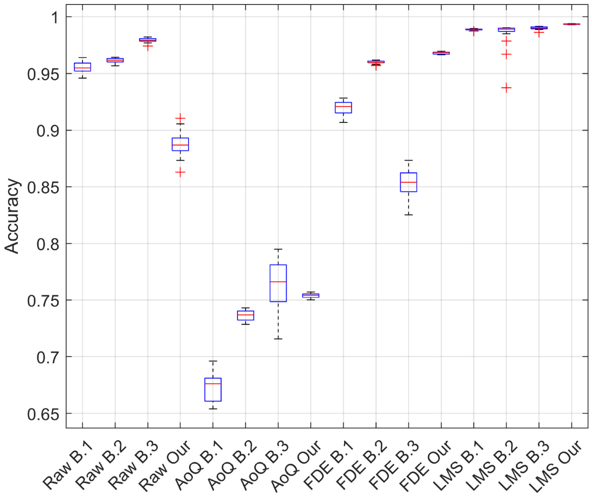


FIGURE 13. Classification performance of different RFF extraction methods and CNN models for *known position* scenario.

model. This is due to the fact that the proposed CNN model has less parameters than other baseline models. The simpler network structure also means lower implementation costs.

C. RFF IDENTIFICATION PERFORMANCE

We first investigate the identification performance in the *known position* scenario, where the IEEE 802.11 data samples received at all four positions are used for both training and testing. The performance of the RFF identification is illustrated in Fig. 13 using box plot with 20 repetitions. It can be seen that the LMS-based equalization provides the highest identification accuracy compared with other methods. It can achieve over 98.88% the median of accuracy with all CNN models considered in this work. In particular, the median of accuracy is as high as 99.34% with the proposed CNN model. Moreover, the proposed CNN model with *FDE* sequence also provides a considerably high accuracy, with a median of accuracy of 96.83%. We also notice that the *Baseline3* model also shows good performance with input of *Raw* and *AoQ*. This is attributed to a relatively deeper network of *Baseline3* model with 10 layers, which permits the network extract more details from raw I/Q samples.

To evaluate the performance for the *unknown position* scenario, we trained the CNN model using IEEE 802.11 signal samples captured from different locations. Specifically, the data samples collected from 1 position, 2 positions and 3 positions are mixed to form the training dataset, which is denoted as "1 position", "2 positions" and "3 positions" in the sequel. The testing dataset are collected from all four positions.[1] Only the *FDE* and *LMS* methods have been evaluated using the proposed CNN model, because these two methods are more effective than the *Raw* and *AoQ* methods in the *known position* scenario.

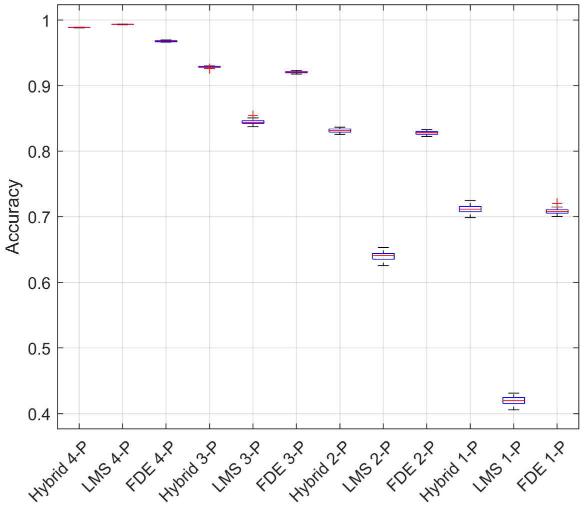


FIGURE 14. Classification performance of different RFF extraction methods for *unknown position* scenario.

The performance of the RFF identification is presented in Fig. 14 using box plot with 20 repetitions. It can be seen that, when the testing dataset contains samples collected from unknown positions than the training dataset, *FDE* data processing provides better performance than other methods. For instance, the median of accuracy can reach 70.76% when the training dataset contains data samples captured from 1 position and testing dataset contains data samples captured from all four positions (i.e., "1 position" case). In contrast, the median of accuracy of LMS method is 41.98% for this case. This is because "1 position" represents a case where the data samples in the testing dataset contain more unknown channel response, which affects the signal waveform. Hence, the FDE-based method can better mitigate the channel effects and achieve better RFF identification performance. With data samples from more positions available in the training dataset, higher identification accuracy can be achieved for both RFF extraction methods. This is because more diversified training data samples are available, which helps to improve the performance of the identifier network. In particular, with *LMS* and *FDE* methods, the median of accuracy is higher than 84.42% for "3 positions" case.

The hybrid identifier proposed in Section IV-E combines the *FDE* and *LMS* methods with CSI similarity judgment. We analyze the CSI similarity of the collected signal frames by calculating the Pearson correlation coefficient of the magnitude values of power-normalized CSIs. We first calculate the correlation coefficient for the CSIs of the four positions of each device. The resulting correlation coefficients

[1] The case where data samples of all four positions are mixed together for both training and testing coincides with the previous "known position" scenario.

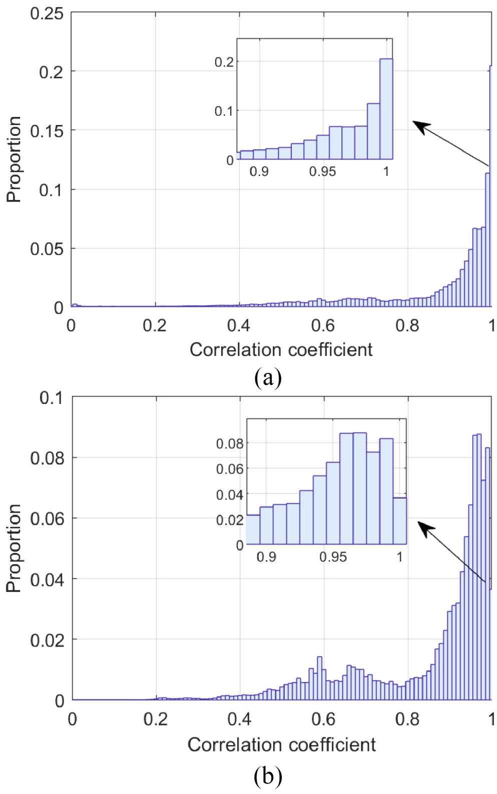


FIGURE 15. **Correlation coefficients between (a) the CSIs of the 4 positions (b) the CSIs of position 1 and the CSIs of the other 3 positions.**

are presented in Fig. 15 (a). It can be seen that 20.47% of the correlation coefficients have values of 1, and 11.34% have values of 0.99. This demonstrates that the measured CSI is relatively stable at different moments when the device location is fixed. To test the CSI similarity between different positions, we calculate the correlation coefficient between the CSIs of position 1 and the CSIs of the other 3 positions, the correlation coefficients are presented in Fig. 15 (b). It is shown that there are 3.65% correlation coefficients have values of 1 and about 8.32% have values of 0.99. Hence, the CSI similarity is possible to occur between channels of different positions.

With this basis, in the training process of hybrid identifier, we store 10 CSIs for each device, these CSIs are estimated from the randomly selected frames of training dataset. The similarity judgment of CSI consists in calculating the correlation coefficient between the stored CSIs and the CSI of testing frame. When the correlation coefficient is larger than 0.99, the similarity judgment of CSI is considered to be passed. The performance of the hybrid identifier is also presented in Fig. 14. It can be seen that, compared to the *FDE* method, the hybrid identifier can improve the median of accuracy by 2%. This proves that, with the stored CSIs and CSI similarity judgement, the proposed hybrid identifier can exploit the merits of both *FDE* and *LMS* methods. However, for the training set of all positions, the performance of hybrid identifier is slightly inferior to that of the *LMS* method. Moreover, we note that since the CSI is estimated using the LTSs, the resulted CSI may not fully cover the fluctuation of channel in the STS part. This leads to the case where the CSI similarity judgment is passed, but the classification result is incorrect. When the CSI correlation coefficient is larger than 0.99, the classification accuracy of LMS-based method is 99.8%.

VI. CONCLUSION

In this paper, the performance of RFF identification using IEEE 802.11 a/g/n devices has been investigated under various channel conditions. To mitigate multipath fading, a time domain LMS equalization based RFF extraction method is proposed. Compared to the FDE-based method, the equalization process of the LMS-based method is progressive, which results in higher identification accuracy when the training and testing datasets are collected from the same positions. Furthermore, a hybrid identifier combining the LMS-based and FDE-based methods is proposed to improve identification accuracy when the datasets are collected from different positions. A four-layer deep learning CNN is designed for RFF identification, demonstrating improved accuracy and training speed for equalized sequences. An experimental system with 64 AP routers and 4 mobile phones has been set up, and the 802.11 frames are captured from 4 different positions. The experimental results show that the proposed hybrid identifier can enhance overall identification accuracy by 2%.

REFERENCES

[1] X. Wang, P. Hao, and L. Hanzo, "Physical-layer authentication for wireless security enhancement: Current challenges and future developments," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 152–158, Jun. 2016.

[2] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 282–310, 1st Quart., 2021.

[3] L. Mucchi et al., "Physical-layer security in 6G networks," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1901–1914, 2021.

[4] Y. Liu, J. Wang, J. Li, S. Niu, and H. Song, "Machine learning for the detection and identification of Internet of Things devices: A survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 298–320, Jan. 2022.

[5] B. Danev, D. Zanetti, and S. Capkun, "On physical-layer identification of wireless devices," *ACM Comput. Surveys*, vol. 45, no. 1, pp. 1–29, Nov. 2012.

[6] D. R. Reising, M. A. Temple, and J. A. Jackson, "Authorized and rogue device discrimination using dimensionally reduced RF-DNA fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 10, pp. 1180–1192, 2015.

[7] E. Uzundurukan, A. M. Ali, and A. Kara, "Design of low-cost modular RF front end for RF fingerprinting of Bluetooth signals," in *Proc. 25th Signal Process. Commun. Appl. Conf. (SIU)*, Antalya, Turkey, 2017, pp. 1–4.

[8] T. D. Vo-Huu, T. D. Vo-Huu, and G. Noubir, "Fingerprinting Wi-Fi devices using software defined radios," in *Proc. ACM Conf. Security Privacy Wireless Mobile Netw. (WiSec)*, Darmstadt, Germany, 2016, pp. 3–14.

[9] K. Sankhe et al., "No radio left behind: Radio fingerprinting through deep learning of physical-layer hardware impairments," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 1, pp. 165–178, Mar. 2020.

[10] A. Al-Shawabka et al., "Exposing the fingerprint: Dissecting the impact of the wireless channel on radio fingerprinting," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Toronto, ON, Canada, 2020, pp. 646–655.

[11] S. Hanna, S. Karunaratne, and D. Cabric, "Open set wireless transmitter authorization: Deep learning approaches and dataset considerations," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 1, pp. 59–72, Mar. 2021.

[12] N. Wang, W. Li, L. Jiao, A. Alipour-Fanid, T. Xiang, and K. Zeng, "Orientation and channel-independent RF fingerprinting for 5G IEEE 802.11ad devices," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 9036–9048, Jun. 2022.

[13] D. Reising, J. Cancelleri, T. D. Loveless, F. Kandah, and A. Skjellum, "Radio identity verification-based IoT security using RF-DNA fingerprints and SVM," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8356–8371, May 2021.

[14] L. Peng, A. Hu, J. Zhang, Y. Jiang, J. Yu, and Y. Yan, "Design of a hybrid RF fingerprint extraction and device classification scheme," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 349–360, Feb. 2019.

[15] J. Bassey, D. Adesina, X. Li, L. Qian, A. Aved, and T. Kroecker, "Intrusion detection for IoT devices based on RF fingerprinting using deep learning," in *Proc. 4th Int. Conf. Fog Mobile Edge Comput. (FMEC)*, Rome, Italy, 2019, pp. 98–104.

[16] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2604–2616, Aug. 2021.

[17] S. Rajendran and Z. Sun, "RF impairment model-based IoT physical-layer identification for enhanced domain generalization," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 1285–1299, 2022.

[18] *Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, ANSI/IEEE Standard 802.11, 1998.

[19] G. Li, J. Yu, Y. Xing, and A. Hu, "Location-invariant physical layer identification approach for WiFi devices," *IEEE Access*, vol. 7, pp. 106974–106986, Aug. 2019.

[20] W. Hou, X. Wang, J.-Y. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1658–1667, May 2014.

[21] J. Zhang, R. Woods, M. Sandell, M. Valkama, A. Marshall, and J. Cavallaro, "Radio frequency fingerprint identification for narrowband systems, modelling and classification," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 3974–3987, 2021.

[22] F. Restuccia, S. D'Oro, A. Al-Shawabka, B. C. Rendon, S. Ioannidis, and T. Melodia, "DeepFIR: Channel-robust physical-layer deep learning through adaptive waveform filtering," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8054–8066, Dec. 2021.

[23] N. Soltani, K. Sankhe, J. Dy, S. Ioannidis, and K. Chowdhury, "More is better: Data augmentation for channel-resilient RF fingerprinting," *IEEE Commun. Mag.*, vol. 58, no. 10, pp. 66–72, Oct. 2020.

[24] G. Shen, J. Zhang, A. Marshall, and J. R. Cavallaro, "Towards scalable and channel-robust radio frequency fingerprint identification for LoRa," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 774–787, 2022.

[25] N. Basha, B. Hamdaoui, K. Sivanesan, and M. Guizani, "Channel-resilient deep-learning-driven device fingerprinting through multiple data streams," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 118–133, 2023.

[26] J. Yu, A. Hu, G. Li, and L. Peng, "A robust RF fingerprinting approach using multisampling convolutional neural network," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6786–6799, Aug. 2019.

[27] T. Iwamoto, "Radiometric identification of emitters in the automatic identification system," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, 2013, pp. 447–452.

[28] D. Zanetti, V. Lenders, and S. Capkun, "Exploring the physical-layer identification of GSM devices," Dept. Comput. Sci., ETH Zürich, Zürich, Switzerland, Rep. 763, 2012.

[29] Y. Yuan, Z. Huang, H. Wu, and X. Wang, "Specific emitter identification based on Hilbert-Huang transform-based time-frequency-energy distribution features," *IET Commun.*, vol. 8, no. 13, pp. 2404–2412, 2014.

[30] S. Ur Rehman, K. Sowerby, and C. Coghill, "RF fingerprint extraction from the energy envelope of an instantaneous transient signal," in *Proc. Aust. Commun. Theory Workshop (AusCTW)*, 2012, pp. 90–95.

[31] M. Köse, S. Tascioglu, and Z. Telatar, "RF fingerprinting of IoT devices based on transient energy spectrum," *IEEE Access*, vol. 7, pp. 18715–18726, 2019.

[32] B. Danev and S. Capkun, "Transient-based identification of wireless sensor nodes," in *Proc. Int. Conf. Inf. Process. Sens. Netw. (IPSN)*, San Francisco, CA, USA, 2009, pp. 25–36.

[33] F. Demers and M. St-Hilaire, "Radiometric identification of LTE transmitters," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2013, pp. 4116–4121.

[34] G. Baldini, C. Gentile, R. Giuliani, and G. Steri, "Comparison of techniques for radiometric identification based on deep convolutional neural networks," *Electron. Lett.*, vol. 55, no. 2, pp. 90–92, Jan. 2019.

[35] J. Huang, W. Albazrqaoe, and G. Xing, "BlueID: A practical system for Bluetooth device identification," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Toronto, ON, Canada, 2014, pp. 2849–2857.

[36] S. U. Rehman, K. W. Sowerby, and C. Coghill, "Radio-frequency fingerprinting for mitigating primary user emulation attack in low-end cognitive radios," *IET Commun.*, vol. 8, no. 8, pp. 1274–1284, May 2014.

[37] A. C. Polak and D. L. Goeckel, "Wireless device identification based on RF oscillator imperfections," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Florence, Italy, 2014, pp. 2679–2683.

[38] C. G. Wheeler and D. R. Reising, "Assessment of the impact of CFO on RF-DNA fingerprint classification performance," in *Proc. Int. Conf. Comput. Netw. Commun. (ICNC)*, Santa Clare, CA, USA, 2017, pp. 1–5.

[39] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless device identification with radiometric signatures," in *Proc. ACM Int. Conf. Mobile Comput. Netw. (MOBICOM)*, San Francisco, CA, USA, 2008, pp. 116–127.

[40] D. R. Reising and M. A. Temple, "WiMAX mobile subscriber verification using Gabor-based RF-DNA fingerprints," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Ottawa, ON, Canada, 2012, pp. 1005–1010.

[41] M. Pospíšil, R. Marsalek, and J. Pomenkova, "Wireless device authentication through transmitter imperfections—Measurement and classification," in *Proc. IEEE 24th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, London, U.K., 2013, pp. 497–501.

[42] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication networks," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 160–167, Feb. 2018.

[43] P. Robyns, E. Marin, W. Lamotte, P. Quax, D. Singelée, and B. Preneel, "Physical-layer fingerprinting of LoRa devices using supervised and zero-shot learning," in *Proc. ACM Conf. Security Privacy Wireless Mobile Netw. (WiSec)*, Boston, MA. USA, 2017, pp. 58–63.

[44] G. Zhang, L. Xia, S. Jia, and Y. Ji, "Identification of cloned HF RFID proximity cards based on RF fingerprinting," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, Tianjin, China, 2016, pp. 292–300.

[45] Q. Li, H. Fan, W. Sun, J. Li, L. Chen, and Z. Liu, "Fingerprints in the air: Unique identification of wireless devices using RF RSS fingerprints," *IEEE Sensors J.*, vol. 17, no. 11, pp. 3568–3579, Jun. 2017.

[46] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: Modeling and validation," *IEEE Trans. Inf. Forensics Security*, vol. 11, pp. 2091–2106, 2016.

[47] J. M. McGinthy, L. J. Wong, and A. J. Michaels, "Groundwork for neural network-based specific emitter identification authentication for IoT," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6429–6440, Aug. 2019.

[48] L. J. Wong, W. C. Headley, and A. J. Michaels, "Specific emitter identification using convolutional neural network-based IQ imbalance estimators," *IEEE Access*, vol. 7, pp. 33544–33555, 2019.

[49] G. Baldini, R. Giuliani, and C. Gentile, "An assessment of the impact of IQ imbalances on the physical layer authentication of IoT wireless devices," in *Proc. Global IoT Summit (GIoTS)*, Aarhus, Denmark, 2019, pp. 1–6.

[50] W. Wang, Y. Chen, and Q. Zhang, "Privacy-preserving location authentication in Wi-Fi networks using fine-grained physical layer signatures," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1218–1225, Feb. 2016.

[51] N. Nguyen, G. Zheng, Z. Han, and R. Zheng, "Device fingerprinting to enhance wireless security using nonparametric Bayesian method," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Shanghai, China, 2011, pp. 1404–1412.

[52] S. S. Hanna and D. Cabric, "Deep learning based transmitter identification using power amplifier nonlinearity," in *Proc. Int. Conf. Comput. Netw. Commun. (ICNC)*, Honolulu, HI, USA, 2019, pp. 674–680.

[53] Y. Pan, S. Yang, H. Peng, T. Li, and W. Wang, "Specific emitter identification based on deep residual networks," *IEEE Access*, vol. 7, pp. 54425–54434, 2019.

[54] Y. Lin, W. Li, J. Sun, and Q. Wu, "Improving wireless devices identification using gray relationship classifier to enhance wireless network security," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Honolulu, HI, USA, 2018, pp. 421–425.

[55] Q. Tian et al., "New security mechanisms of high-reliability IoT communication based on radio frequency fingerprint," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7980–7987, Oct. 2019.

[56] A. Aghnaiya, A. M. Ali, and A. Kara, "Variational mode decomposition-based radio frequency fingerprinting of Bluetooth devices," *IEEE Access*, vol. 7, pp. 144054–144058, 2019.

[57] U. Satija, N. Trivedi, G. Biswal, and B. Ramkumar, "Specific emitter identification based on variational mode decomposition and spectral features in single hop and relaying scenarios," *IEEE Trans. Inf. Forensics Security*, vol. 14, pp. 581–591, 2019.

[58] Y. Yang, A. Hu, Y. Xing, J. Yu, and Z. Zhang, "A data-independent radio frequency fingerprint extraction scheme," *IEEE Wireless Commun. Lett.*, vol. 10, no. 11, pp. 2524–2527, Nov. 2021.

[59] Z. Chen, L. Peng, and H. Fu, "Isolated forest-based ZigBee device identification using adaptive filter coefficients," in *Proc. 7th Int. Conf. Comput. Commun. Syst. (ICCCS)*, 2022, pp. 715–720.

[60] H. J. Patel, M. A. Temple, and R. O. Baldwin, "Improving ZigBee device network authentication using ensemble decision tree classifiers with radio frequency distinct native attribute fingerprinting," *IEEE Trans. Rel.*, vol. 64, no. 1, pp. 221–233, Mar. 2015.

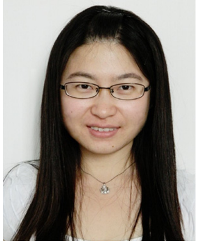
[61] M. Abdrabou and T. A. Gulliver, "Physical layer authentication for satellite communication systems using machine learning," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2380–2389, 2022.

[62] Y. Wang, G. Gui, H. Gacanin, T. Ohtsuki, O. A. Dobre, and H. V. Poor, "An efficient specific emitter identification method based on complex-valued neural networks and network compression," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2305–2317, Aug. 2021.

[63] L. Peng, J. Zhang, M. Liu, and A. Hu, "Deep learning based RF fingerprint identification using differential constellation trace figure," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1091–1095, Jan. 2020.

[64] L. Ding, S. Wang, F. Wang, and Z. Wei, "Specific emitter identification via convolutional neural networks," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2591–2594, Dec. 2018.

[65] "Networked software defined radio (SDR)." Apr. 2018. Accessed: 2022. [Online]. Available: https://www.ettus.com/product-categories/usrp-networked-series/



HUA FU (Member, IEEE) received the M.Eng. degree in electrical engineering from the École nationale supérieure d'électronique, informatique, télécommunications, mathématique et mécanique de Bordeaux, France, in 2011, and the Ph.D. degree in electrical engineering from the National Institute of Applied Sciences, Rennes, France, in 2015. She was with the Electrical and Computer Engineering Department, Université de Sherbrooke, QC, Canada, as a Postdoctoral Fellow. She is currently with the School of Cyber Science and Engineering, Southeast University, Nanjing, China, as a Lecturer. She also works with the Purple Mountain Laboratories for Network and Communication Security, Nanjing. Her research interests include multiple-input multiple-output systems with large antenna arrays and physical layer security for wireless communications.

LINNING PENG (Member, IEEE) received the Ph.D. degree from the Electronics and Telecommunications Institute of Rennes Laboratory, National Institute of Applied Sciences, Rennes, France, in 2014. He is an Associate Professor with Southeast University and also works with the Purple Mountain Laboratories for Network and Communication Security, Nanjing, China. His research interests include Internet of Things and physical layer security in wired and wireless communications.

MING LIU (Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from Xi'an Jiaotong University, China, in 2004 and 2007, respectively, and the Ph.D. degree in electrical engineering from the National Institute of Applied Sciences, Rennes, France, in 2011. He was with the Institute of Electronics and Telecommunications of Rennes as a Postdoctoral Researcher from 2011 to 2015. He is currently with Beijing Jiaotong University, China, as an Associate Professor. His main research interests include beyond 5G/6G, PHY security, and AI for wireless communications. He is a co-recipient of the Science and Technology Award of China Railway Society in 2020. He is a member of China Computer Federation and Young Computer Scientists and Engineers Forum.

AIQUN HU (Senior Member, IEEE) received the Ph.D. degree from Southeast University, Nanjing, China, in 1993, where he is a Full Professor. He also works with the Purple Mountain Laboratories for Network and Communication Security, Nanjing. He has published many papers on high-quality transactions and possessed many Chinese patents in wireless technology. His research interests are in wireless network technology and physical-layer security of wireless communications.