

Research Article

On the Performance Supremum of CFO Based Physical Layer Identification

Shuiguang Zeng ^{1,2,3,4,5} Yin Chen,^{6,7} Xufei Li,^{1,2} Yulong Shen,^{1,2} Dongmei Zhao,^{3,4,5} Jinxiao Zhu ⁸ and Norio Shiratori⁹

¹School of Computer Science and Technology, Xidian University, Xi'an 710071, China

²Shaanxi Key Laboratory of Network and System Security, Xi'an 710071, China

³College of Computer and Cyber Security, Hebei Normal University, Shijiazhuang 050024, China

⁴Hebei Key Laboratory of Network and Information Security, Shijiazhuang 050024, China

⁵Hebei Provincial Engineering Research Center for Supply Chain Big Data Analytics & Data Security, Shijiazhuang 050024, China

⁶Graduate School of Media and Governance, Keio University, Fujisawa 252-0882, Japan

⁷Reitaku University, 2-1-1, Hikarigaoka, Kashiwa City, Chiba Prefecture 277-8686, Japan

⁸Faculty of Information Networking for Innovation and Design (INIAD), Toyo University, Tokyo 115-0053, Japan

⁹Research and Development Initiative, Chuo University, Tokyo 112-8551, Japan

Correspondence should be addressed to Jinxiao Zhu; zhu@iniad.org

Received 18 June 2022; Revised 1 July 2022; Accepted 20 July 2022; Published 8 August 2022

Academic Editor: A.H. Alamoodi

Copyright © 2022 Shuiguang Zeng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Physical layer identification is an emerging technique that exploits physical layer features to identify wireless devices. The identification accuracy and the device quantity that can be identified at most are significant for the identification scheme. Existing works primarily focus on the feature correlation analysis for multifeature selection without investigating the least upper bound (supremum) of the performance of a single feature. The supremum indicates the limit of the performance, which is another sight for evaluating the quality of features and improving the performance of the identification scheme. Therefore, this paper first investigates the supremum of the performance of the most commonly used physical layer feature, i.e., carrier frequency offset (CFO). Specifically, we offer a rigorous mathematical analysis and derive the closed-form expression of the supremum of identification accuracy based on the max-min distance analysis (MMDA) criterion. And then, the supremum of the number of distinguishable devices is also analyzed. Finally, we conducted a simulation study to verify the theoretical analysis result.

1. Introduction

Device identification plays a vital role in wireless networks, conventionally realized with pre-distributed information such as IP addresses, MAC addresses, and international mobile station equipment identity (IMEI) numbers. With this information, basic access control [1] and location tracking [2] can be implemented. However, the mentioned addresses and numbers can easily be spoofed, exposing wireless devices and infrastructures to security threats [3]. Furthermore, it is often restricted to collect identity information due to business, privacy, and legal reasons, while identity is necessary for some

applications. Therefore, there is an urgent need to find a more reliable or complementary way to identify devices.

Recently, the rich characteristics of the physical layer have been intensively investigated to implement device identification in wireless networks [4–6], also known as physical layer identification. Various physical layer features can be extracted and performed as the device's identity. These features stem from the small-scale hardware impairment in the transceivers or the location-specific characteristics of the wireless channel between the transmitter and receiver. According to the signal types collected for feature extraction, there are two categories of identification schemes, i.e.,

transient and steady-state signal-based ones [7]. Since the steady-state signal is easier to capture than the transient one and has attracted more attention from researchers, we concentrate on this type. Two approaches are reported in the literature for physical layer identification based on steady-state signals according to different classifier types.

The first approach, called shallow classifier-based device identification, implements the identification with hand-crafted features calculated from the received signals by carefully designed feature extraction algorithms. These features, usually relying on expert feature knowledge, will then be exploited with traditional shallow classifiers such as support vector machine (SVM) and K-nearest neighbor (KNN) or binary hypothesis testing to identify and authenticate transmitters [7–11].

The second approach takes advantage of the powerful learning ability of deep learning to identify wireless devices with the collected raw in-phase and quadrature (IQ) signal or its transformed information [6, 12–19]. Hence, it is known as deep learning-based physical layer identification. In this approach, hidden features can be automatically extracted from wireless frames with the aid of the representation learning ability of deep learning methods without using explicit feature calculation algorithms.

Both approaches are regarded as multiclass classification problems when using machine learning classifiers to discriminate multiple devices. It is intuitive that the identification accuracy will decrease as the quantity of devices increases. In other words, if we want to achieve the desired accuracy, the quantity of devices that can be identified will be limited. Recent work supports such a claim from the view of experiments, where the accuracies drop for both WiFi and ADS-B datasets using two deep learning models when the quantity of devices increases to 10,000 from 100 [19]. And there exist works focusing on the combination of multiple features to improve identification performance [8]. Their work is based on the view that a single feature leads to limited identification accuracy or limited number of distinguishable devices. User capacity of the physical layer identification system is investigated in [20, 21], where the authors consider the frequency characteristics from fast Fourier transform (FFT) as the radio frequency fingerprints.

Except for the mentioned related works, we still lack detailed analysis on the supremum of the identification accuracy and distinguishable devices for specific hand-crafted features, i.e., what is the highest identification accuracy and how many devices could identify at most under given conditions? This is important for investigating the performance and quality of a specific physical layer feature in an identification scheme and gives insight into finding approaches to improve the performance, such as identification with multiple features. Therefore, this paper focuses on issues not touched upon in existing works with the following contributions:

- (1) Firstly, with the max-min distance analysis (MMDA) criterion and other mathematical analyses, we derive the closed-form expression of the supremum of the identification accuracy of the shallow classifier-based scheme given specific conditions

- (2) Secondly, we also analyze the supremum of the number of distinguishable devices (device quantity supremum) of the shallow classifier-based scheme given the accuracy constraint
- (3) Finally, through comprehensive simulations, we confirm the theoretical analysis and provide some interesting insights. The results indicate that the feature range (the value range of the physical layer features such as CFO specified in the standard protocol) and the SNR are the main factors affecting the identification performance, consistent with the theoretical analysis. We compare the accuracy of the shallow classifier- and deep learning-based identification schemes with the accuracy supremum. We also investigate the device quantity supremum with simulation with different CFO ranges and accuracy constraints at various SNRs

The rest of this paper is organized as follows. Section 2 reviews related works and introduces basic knowledge. Section 3 describes the system and communication models. Section 4 focuses on the supremum analysis of the shallow classifier-based identification scheme. Section 5 presents the simulation settings and results. Finally, Section 6 concludes this paper.

2. Related Work and Background

In this section, we first present related works, and further introduce certain basic knowledge regarding machine learning.

2.1. Related Work. There is no difference between shallow classifier- and deep learning-based device identification in terms of essential processes, including feature extraction and device identification. However, steady-state radiometric features such as carrier frequency offset (CFO) [22] and in-phase and quadrature imbalance (IQI) [23] rely on the expert knowledge of signal processing. Therefore, the feature extraction is explicit and protocol-specific. On the other hand, deep learning methods [17, 24] can automatically extract implicit features rather than expert feature engineering based on the raw IQ samples of the signal. However, this process usually requires dedicated hardware, such as graphics processing units (GPUs), to accelerate the computation.

For the first type of identification approaches, the estimation method and the number of independent features are the key factors that affect the performance. It is acknowledged that employing multiple features can improve identification accuracy. Therefore, some existing works focus on exploring new features and integrating with other features [8]. Peng et al. smartly combine differential constellation trace figure, CFO, modulation offset, and IQI to identify 54 ZigBee devices and achieve classification error rates of 4.48% and 9.42% under the line of sight (LOS) and non-line of sight (NLOS) scenarios [25].

While recently, deep learning-based identification approaches have attracted considerable research attention, which apply various deep neural network models to implement the feature extraction and identification processes, raw IQ samples or their transformed information, such

as power spectrum and FFT sequence, can be used directly as the inputs of the models.

However, the upper bounds or the supremum of identification accuracy and the number of distinguishable devices for hand-crafted features are unclear for device identification. The supremum of a single feature in terms of accuracy and device number indicates the ultimate performance, with which we can design a better identification scheme and implement a more appropriate feature selection and combination. Although Wang et al. [20] [21] explore the user capacity of the physical identification system, they consider the frequency characteristics from FFT as the radio frequency fingerprints without analyzing hand-crafted features.

2.2. Machine Learning. From the view of model structure, machine learning can be categorized as shallow classifiers and deep learning. Shallow classifiers, which usually adopt statistical models with only a few layers of composition, are mainstream research before the breakthrough of deep learning. These classifiers include naive Bayes, support vector machine (SVM), AdaBoost, random forest, and KNN and are still adopted in many commercial classification systems. Deep learning technologies are neural networks with many layers of nonlinear information processing. In recent years, they have been widely studied in many fields such as computer vision, speech recognition, and cybersecurity.

2.2.1. Shallow Classifiers. Shallow classifiers always have a very efficient and effective performance on high-quality samples [26]. This paper adopts KNN as a shallow classifier for device identification based on hand-crafted features. KNN is a simple but efficient machine learning algorithm, usually used for classification and regression. Usually, the new sample/case will be assigned to the class that is most common among its K-nearest neighbors measured by a distance function, i.e., the majority voting of the new case's neighbors according to the distance such as Euclidean distance [27].

2.2.2. Deep Learning. There are different deep learning models, such as recurrent neural network (RNN), convolutional neural network (CNN), and generative adversarial network (GAN). This paper focuses on CNN since it has been investigated in much recent literature and has shown great potential in device identification. A general CNN comprises one or more convolutional layers, pooling layers, and fully connected (FC) layers [28]. The convolutional layers aim to promote important hidden features of the input data through the specially designed structures called "filters" having different dimensions, also known as feature extractors. Also, different types of CNN have been investigated for device identification. 1D and 2D CNNs with one/two-dimensional convolutional layers are exploited to identify wireless devices [19, 29]. Complex-valued neural networks are explored in [30] to improve the wireless identification performance.

3. System Model

In this section, we first describe the considered identification and communication model. Then, we formulate the concerning problem about the least upper bound analysis.

TABLE 1: List of variables and notations.

Notation	Definition
M	Class (transmitter) number
U_ϵ	CFO range
L	Number of multipath components
γ/η	Signal-to-noise ratio (SNR)
L_s	The length of training sequence
$(\cdot)^*$	Complex conjugate operator
$\Im(\cdot)$	Imaginary part of complex number
$\Re(\cdot)$	Real part of complex number
$Q(\cdot)$	Q-function
$\sup A$	Supremum of set A
$f'(\cdot)$	First-order derivative
$f''(\cdot)$	Second-order derivative
$f^{-1}(\cdot)$	Inverse function
$\lfloor \cdot \rfloor$	Floor function

Table 1 summarizes the main variables and notations used in this paper.

3.1. Identification Model. As shown in Figure 1, the model comprises a wireless receiver (RX) and M wireless transmitters (TX). The receiver attempts to identify each transmitter using the received wireless frames. The receiver first collects the raw IQ samples of the concerned field, e.g., baseband preamble, via frame detection and synchronization from the received wireless signals subject to the concerned channels. The receiver can calculate the hand-crafted features using the raw IQ samples for identification. Also, it can directly use the raw IQ samples to identify transmitters with deep learning. Then, the transmitter identification will be formulated as a multiclass classification problem based on hand-crafted features or raw IQ samples, depending on the adopted classifier.

3.2. Communication Model. At the receiver, the passband signal is down-converted to the baseband. Then, the received baseband signal is sampled by the analog-to-digital converter (ADC) to obtain the discrete complex-valued preamble signal, i.e., the raw IQ samples. The baseband signal with CFO is given as [31]:

$$r(n) = e^{j2\pi n\Delta f} s(n) + w(n), \quad (1)$$

where $\Delta f = \epsilon T_s$ is the normalized CFO with ϵ being the CFO in the corresponding range $[-U_\epsilon, U_\epsilon]$ parts per million (ppm) to the carrier frequency f_c . The range is usually specified in the communication standard concerned. Here, $T_s \triangleq 1/B_w$ is the sampling interval with B_w being the total communication bandwidth, and $w(n) \sim \mathcal{CN}(0, \sigma_n^2)$ is the circular symmetric additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2 . Let $a(n)$ denotes the long training symbols with length of $2 \times L_s$, which usually contains

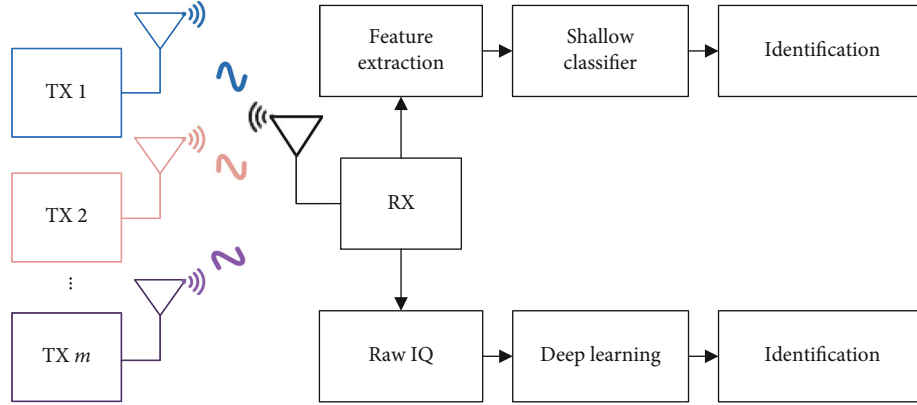


FIGURE 1: System model of typical physical layer identification.

two L_s -length repetitive sequences. And it is a classical preamble structure in many wireless protocols such as most specifications in the IEEE 802.11 family. When the baseband signal is only subject to the AWGN channel, $s(n)$ is the same as $a(n)$, and the received signal is denoted as (1). When the baseband signal is subject to multipath channel, $s(n)$ is given by

$$s(n) = \sum_{l=0}^{L-1} h(l)a_{n-l}, \quad (2)$$

where $h(l)0 \leq l \leq (L-1)$ represents the channel coefficients of the multipath fading channel. Notably, we assume the locations of transmitters and receiver are fixed. Hence, the Doppler offset is zero, and the channel profile is static during the operation time, which can be considered a quasi-static channel similar to [32]. This is reasonable for many wireless networks such as wireless sensor networks (WSN) and wireless local area networks (WLAN) where fixed sink nodes or routers create a static channel profile when the receiver location is also fixed. We then denote the concerned preamble containing the two repetitive long training sequences as $\mathbf{r} = [r(0), r(1), \dots, r(2L_s - 1)]$ after synchronization and denote the SNR of the received signal as γ with the unit of dB, which is calculated as follows:

$$\gamma = 10 \log(\eta) = 10 \log\left(\frac{1}{2L_s} \frac{\sum_{n=0}^{2L_s-1} |s(n)|^2}{\sigma_n^2}\right), \quad (3)$$

where $\eta = \sigma_s^2 / \sigma_n^2$ and σ_s^2 and σ_n^2 represent the power of signal and noise, respectively.

3.3. Problem Formulation. For the shallow classifier-based identification scheme with CFO, the CFO will be first estimated from raw IQ samples of the preamble field of the received signal. Then with the collected feature of each frame, we can train a shallow classifier for device identification.

We are interested in how the identification accuracy will vary with the range of a single feature and other conditions. And are there any supremum or upper bound of identification accuracy and the number of distinguishable devices with the considered feature? In a word, our primary aim is to investigate the performance limits of each specific feature adopted in physical layer identification by answering the above questions.

4. Identification Performance and the Supremum

At the receiver, the raw IQ samples of the long training sequence are adopted to implement hand-crafted feature estimation for identification. We then analyze the supremum of identification accuracy and the device quantity supremum considering CFO.

4.1. CFO Estimation. Similar to [33–35], when the length of the long training sequence is larger than the maximum channel delay L in (2), i.e., $L_s \geq L$, the CFO can be estimated by the two repetitive long training sequences. We first calculate the phase difference ϕ between the frequency responses of two identical and consecutive long training sequences as

$$\phi = \arctan \left\{ \frac{\Im \left(\sum_{n=0}^{L_s-1} r^*[n] r[n+L_s] \right)}{\Re \left(\sum_{n=0}^{L_s-1} r^*[n] r[n+L_s] \right)} \right\}, \quad (4)$$

where $r[n]$ is the complex I and Q samples of the frequency response of a training sequence and n is the time index in a window of $2L_s$ samples. Then, we achieve the estimated CFO as

$$\hat{\epsilon}_h = \frac{\phi}{2\pi L_s}, \quad (5)$$

and according to [35], the standard deviation of the CFO

estimation is

$$\sigma_{\varepsilon_h}^{\wedge} = \frac{1}{2\pi L_s \sqrt{L_s \eta}}, \quad (6)$$

which can be considered as the lower bound of the CFO estimation error in the receiver. Usually, there exists more than one method for estimating the same feature with different estimating errors, i.e., variance, which results in different performances for the shallow classifier-based identification.

4.2. The Supremum of the Identification Accuracy. Since the true CFOs of the transmitters are independent uniform random variables in the concerned range $[-U_\varepsilon, U_\varepsilon]$ [35] [36], we denote it as $\varepsilon_m \in [-U_\varepsilon, U_\varepsilon]$, $m = 1, 2, \dots, M$. And we define the spacing between two adjacent true CFOs as the distance of $d_m = \varepsilon_{m+1} - \varepsilon_m$, $m = 1, 2, \dots, M-1$; then, we have $\sum_{m=1}^{M-1} d_m \leq 2U_\varepsilon$. Assuming the mean distance of

the CFO $\bar{d} = \sum_{m=1}^{M-1} d_m / M - 1$, then we have

$$0 \leq \bar{d} \leq \frac{2U_\varepsilon}{M-1}. \quad (7)$$

The overall accuracy is widely used for multiclass classification problems, whose definition is as

$$ACC = \frac{TP_m}{S}, \quad (8)$$

where S is the total number of predictions and TP_m is the true positive predictions when considering the classification as a binary classification regarding the m -th class and other classes. We can also denote the classification error rate and accuracy as $p_e = 1 - p_a$ and p_a as (9) according to [37], where $Q(x) = 1/\sqrt{2\pi} \int_x^\infty e^{-t^2/2} dt$ is the Q-function:

$$p_a = \frac{\left((1 - Q(d_1/2\sigma_{\varepsilon_h}^{\wedge})) + \sum_{m=1}^{M-2} (1 - Q(d_m/2\sigma_{\varepsilon_h}^{\wedge}) - Q(d_{m+1}/2\sigma_{\varepsilon_h}^{\wedge})) + (1 - Q(d_{M-1}/2\sigma_{\varepsilon_h}^{\wedge})) \right)}{M} = 1 - \frac{2}{M} \sum_{m=1}^{M-1} Q\left(\frac{d_m}{2\sigma_{\varepsilon_h}^{\wedge}}\right). \quad (9)$$

As we assume that the variances of the estimated CFO of all devices are the same at a specific SNR, which means all the CFO samples are from homoscedastic Gaussians as the same standard deviation of (6). We can adopt the MMDA criterion to achieve the maximum separation of all devices concerning CFO [38]. According to this criterion, to achieve a maximum classification accuracy, we have to maximize the minimum distance of each class pair (two devices) to guarantee the separation as best as possible of any class pairs as

$$\max_{1 \leq m \leq M-1} \min d_m, \quad (10)$$

where the inner minimization chooses the minimum CFO distance $d' \in D'$ of all class pairs, while the outer maximization maximizes this minimum distance [38]. Here, D' is the set of minimum CFO distance.

Theorem 1. For M devices, the supremum of the minimum distance of the CFO $\sup D'$ in the range U_ε is $\bar{d} = 2U_\varepsilon/(M-1)$.

Proof. According to the definition of supremum, we first adopt proof by contradiction to prove that $2U_\varepsilon/(M-1)$ is an upper bound of the minimum distance set D' , i.e., $2U_\varepsilon/(M-1) \geq d'$, where $2U_\varepsilon/(M-1) = \max(\bar{d})$ as in (7). If $\max(\bar{d}) < d'$, then we have $2U_\varepsilon = (M-1) \max(\bar{d}) < (M-1)d' \leq d_1 + d_2 + \dots + d_{M-1} \leq 2U_\varepsilon$, which is a contradiction. Therefore, $2U_\varepsilon/(M-1) \geq d'$ holds. Second, we prove 2

$U_\varepsilon/(M-1)$ is the minimum of the upper bounds. $\forall 0 < \xi < 2U_\varepsilon/(M-1)$; we find $d'_0 = 1/2((2U_\varepsilon/(M-1)) - \xi) + (2U_\varepsilon/(M-1)) = (2U_\varepsilon/(M-1)) - (\xi/2)$, and $d'_0 \in D'$ fulfills $d'_0 > (2U_\varepsilon/(M-1)) - \xi$, which completes the proof. \square

Proposition 2. When the true CFOs of all devices are distributed with equal distance in the concerned range, i.e., the minimum distance of the CFO equals to its supremum, the separation of all devices will be the best. Thus, in this case, $d_1, d_2, \dots, d_{M-1} = \sup D' = 2U_\varepsilon/(M-1)$, we have the least upper bound, i.e., the supremum of accuracy p_a as

$$\begin{aligned} \bar{p}_a &= 1 - 2 \left(\frac{M-1}{M} \right) Q(Y) = 1 - 2 \left(\frac{M-1}{M} \right) Q\left(\frac{U_\varepsilon}{(M-1)\sigma_{\varepsilon_h}^{\wedge}} \right) \\ &= 1 - 2 \left(\frac{M-1}{M} \right) Q\left(\frac{U_\varepsilon}{2(M-1)\pi L_s \sqrt{L_s \eta}} \right), \end{aligned} \quad (11)$$

where $Y = \bar{d}/2\sigma_{\varepsilon_h}^{\wedge}$ and $\bar{d} = 2U_\varepsilon/(M-1)$.

Proof. First, we prove \bar{p}_a is an upper bound of the accuracy p_a . Since $Q(x) = 1 - \Phi(x)$, where $\Phi(x) = 1/\sqrt{2\pi} \int_{-\infty}^x e^{-1/2t^2} dt$ is the cumulative distribution function (CDF) for the standard Gaussian distribution. We have $Q'(x) = -\Phi'(x) = -P(x) < 0$, where $P(x) = 1/\sqrt{2\pi} e^{-1/2x^2}$ is the probability density function (PDF). And when $x > 0$, $Q''(x) = -P'(x) = x/(1$

$\sqrt{2\pi}e^{-1/2x^2} > 0$, which means the Q-function is a monotone decreasing convex function when $x > 0$. Using Jensen's inequality, we have

$$\frac{1}{M-1} \sum_{m=1}^{M-1} Q\left(\frac{d_m}{2\sigma_{\varepsilon_h}}\right) \geq Q\left(\frac{\sum_{m=1}^{M-1} d_m}{2\sigma_{\varepsilon_h}(M-1)}\right) = Q(Y), \quad (12)$$

$$1 - \frac{2}{M} \sum_{m=1}^{M-1} Q\left(\frac{d_m}{2\sigma_{\varepsilon_h}}\right) \leq 1 - 2\left(\frac{M-1}{M}\right)Q(Y). \quad (13)$$

Combining Equation (11) and the expressions of p_a and \bar{p}_a in Equations (9) and (12), we have $p_a \leq \bar{p}_a$, and thus, \bar{p}_a is an upper bound of p_a .

Second, we prove \bar{p}_a is the minimum of the upper bounds of p_a . With $\forall 0 < \xi < 2U_\varepsilon/(M-1)$, we construct a function as

$$f(\xi) = \frac{2}{M}(Q(Y-\xi) + Q(Y+\xi) - 2Q(Y)). \quad (14)$$

Given that the Q-function is a monotone decreasing convex function when $x > 0$, we have $f(\xi) > 0$, and when $\xi \rightarrow 0$, $f(\xi) \rightarrow 0$.

With (12) and (15), then we have

$$\bar{p}_a - f(\xi) = 1 - \frac{2}{M}((M-3)Q(Y) - Q(Y-\xi) - Q(Y+\xi)), \quad (15)$$

and we can find

$$p_0 = 1 - \frac{2}{M} \left((M-3)Q(Y) + Q\left(Y - \frac{\xi}{2}\right) + Q\left(Y + \frac{\xi}{2}\right) \right), \quad (16)$$

with $p_0 \in \mathbb{P}_a$ fulfills $p_0 > \bar{p}_a - f(\xi)$, where \mathbb{P}_a is the set of identification accuracy. Because combined with (16) and (17) and applying the Lagrange mean value theorem, we have $p_0 - (\bar{p}_a - f(\xi)) > 0$, i.e., $p_0 > \bar{p}_a - f(\xi)$ as in (17), where $Y - \xi < x_1 < Y - (\xi/2)$, $Y + (\xi/2) < x_2 < Y + \xi$, and $x_2 > x_1$. Finally, according to the definition of supremum, the proof completes.

$$\begin{aligned} p_0 - (\bar{p}_a - f(\xi)) &= \frac{2}{M} \left(\left(Q(Y-\xi) - Q\left(Y - \frac{\xi}{2}\right) \right) \right. \\ &\quad \left. - \left(Q\left(Y + \frac{\xi}{2}\right) - Q(Y+\xi) \right) \right) \\ &= \frac{\xi}{M} \left(Q'(x_2) - Q'(x_1) \right) > 0. \end{aligned} \quad (17)$$

□

We can observe from (11) that the accuracy supremum is determined by the feature range, the number of transmitters to be identified, and the precision of the feature estimation method (i.e., the standard deviation of the estimate).

4.3. The Device Quantity Supremum. We define the supremum of the number of distinguishable devices or the device quantity supremum of an identification scheme as the device number under which the accuracy of the identification scheme will not exceed the given constraint. And then, with this supremum, we can evaluate the performance limit of the adopted feature for device identification. Intuitively, the supremum is related to the identification accuracy as shown in (11). However, it is difficult to deduce a closed-form expression of the inverse function of (11) concerning M and \bar{p}_a , then we denote (11) as $\bar{p}_a = f(M)$ for simplicity. And given the monotone decreasing property of $\bar{p}_a = f(M)$, we have the following proposition.

Proposition 3. For a specific feature range U_ε and feature estimation, given the target accuracy, the device quantity supremum is $\bar{M} = \lfloor f^{-1}(\bar{p}_a) \rfloor$, $\bar{M} \geq 2$, $\bar{M} \in \mathbb{N}^+$, where $\bar{p}_a = f(M)$ is as shown in (11).

Proof. We define two functions $g(m) = (m-1)/m$ and $h(m) = Q(U_\varepsilon/((m-1)\sigma_{\varepsilon_h}))$, then we can denote (11) as $\bar{p}_a = 1 - 2g(m)h(m)$. Since both $g(m)$ and $h(m)$ are strictly monotone decreasing with $0.5 \leq g(m) < 1$, $m \in \mathbb{R}$, $m \geq 2$ and $0 < h(m) < 0.5$, $-2g(m)h(m)$ will be strictly monotone increasing. Thus, $f(m)$ will be a strictly monotone decreasing function too. According to the properties of the inverse of strictly monotone function, $f^{-1}(\bar{p}_a)$ also will be a monotone decreasing function. The device number is an integer set, denoting as $\mathbb{M} = \{M \in \mathbb{N}^+ | M \leq \lfloor f^{-1}(\bar{p}_a) \rfloor\}$, and $\bar{M} = \lfloor f^{-1}(\bar{p}_a) \rfloor$ is the supremum of the number set \mathbb{M} . We prove it by contradiction as follows. Suppose that \bar{M} is not the supremum of \mathbb{M} , which means there is at least an integer $M' \in \mathbb{M}$ that fulfills $M' > \bar{M}$. Obviously, this is a contradiction because M' should be $\leq \bar{M}$ according to its definition, which means the premise cannot be true. Thus, the device number supremum is \bar{M} . □

It indicates that the maximum number of transmitters can be identified, i.e., device quantity supremum under the constraint of the desired accuracy is determined by the feature's range and the precision of the estimation method. Although it is difficult to give the closed-form expression of $f^{-1}(\bar{p}_a)$, we can depict the relationship between M and \bar{p}_a by simulation and observe the variation of the device quantity supremum.

5. Simulation Study

5.1. Simulation Settings. We simulated a typical wireless communication processing of 802.11a based on OFDM. Similar to [16], we also generated the beacon frames for transmitter identification where the (legacy) long training field (L-LTF) was adopted to estimate the CFO. We implemented data generation and processing, machine learning, and deep learning methods on a platform with MATLAB R2021a. The platform is a Dell Precision 3640 tower workstation (https://dl.dell.com/topicspdf/precision-3640-workstation_owners-manual2_en-

us.pdf) with an Intel(R) Core(TM) i9-10900K CPU and 32GB RAM running the Ubuntu 18.04 operating system. Further, we used an NVIDIA GeForce RTX 3080 GPU configured on the workstation to train and test the deep learning-based models. The main simulation parameters, including the communication system and the Rayleigh channel, are shown in Table 2.

As the system model shows in Section 3, the receiver collects signals from the transmitters and then uses L-LTF to extract the features and identify the devices. Following the specification, the transmitted L-LTF sequences are configured as the same for all transmitters, enabling the algorithm to avoid any data dependency. Since we assume the transmitters and receiver are static, the multipath channel profile and RF impairments do not vary in time.

After comparing several shallow classifiers in common use, we selected KNN for the shallow classifier-based identification. We tuned parameters of “100” as the number of neighbors, “Euclidean” as the distance metric, and “Equal” as the distance weight.

We adopted the same CNN architecture in [16] as the deep learning identification method. The detailed CNN architecture’s parameters, including convolutional layers (Conv2D), pooling layers (MaxPooling2D), and fully connected layers (FC, also known as the dense layer), is shown in Table 3.

To minimize the sampling bias (when selecting data from the dataset) and ensuring statistical confidence, we adopted a 5-fold cross-validation (CV) in the classification evaluation. We split the dataset into five blocks, ensuring that each block has 200 random frames from each device since we collected 1000 frames per transmitter. Then, we performed five rounds of training and testing for each shallow classifier- and deep learning-based model. One block was selected as the test dataset, and the rest were used for training in each round. We considered the averaged overall test accuracy of the five-round CV as the final metric to evaluate the identification performance, i.e., $\bar{ACC} = ACC_k/5$, where ACC_k is the overall test accuracy of the k -th round of CV.

5.2. Identification Accuracy and the Supremum

5.2.1. Identification Under AWGN Channel. As shown in Figure 2, in each CFO range (2.5CFO means $U_\epsilon = 2.5$ ppm), the identification accuracy of the shallow classifier-based scheme (i.e., with the classifier of KNN and hand-crafted feature CFO, denoted as HC) is always under the supremum (SUP). As the SNR increases, the accuracy first exceeds that of deep learning and then to the supremum. When the SNR is between -20 and 0 dB, the gap between the accuracy of the shallow classifier-based scheme and the supremum is small. When the SNR increases from 0 to 50 dB, the gap first widens and then closes. When SNR ≥ 45 dB, the accuracy reaches the supremum of 100% in the range of 20 ppm. It is reasonable because the error of the estimated CFO will be smaller at higher SNR, as discussed in Section 4. Moreover, the accuracy of the deep learning-based scheme (denoted as DP) in all CFO ranges converges to approximately 95% except for 20 ppm where the accuracy converges to 98%. It indicates that deep learning has limited discriminative capabilities for CFO at higher

TABLE 2: Simulation parameters of the communication system.

Parameter	Value
Transmitter number (M)	50,400
Frames per transmitter (F_d)	1000
Carrier frequency (f_c)	5.765 GHz
CFO ranges (U_ϵ)	2.5, 5, 10, 20 ppm
Bandwidth (B_w)	20 MHz
Sampling frequency (f_s)	20 MHz
Length of L-LTF (L_s)	64
Rayleigh path number	3
Discrete path delay in seconds	$[0, 1.8, 3.4]/f_s$ s
Average path gains	$[0, -2, -10]$
Maximum Doppler shift	0

TABLE 3: Architecture of the CNN Model [16].

Layer	Type	Kernel size	Stride	#kernel	Input size
L1	Input				$128 \times 2 \times 1$
L2	Conv2D	7×1		50	$128 \times 2 \times 1$
L3	MaxPool	$[2, 1]$	$[2, 1]$		$122 \times 2 \times 50$
L4	Conv2D	7×2		50	$61 \times 2 \times 50$
L5	MaxPool	$[2, 1]$	$[2, 1]$		$55 \times 1 \times 50$
L6	FC			256	$27 \times 1 \times 50$
L7	FC			80	$27 \times 1 \times 256$
L8	Out			M	$27 \times 1 \times 80$

SNR compared with hand-crafted feature estimation. On the other hand, the deep learning-based scheme can achieve better performance at lower SNR. It means that the influence of SNR on the deep learning scheme is not as significant as that of hand-crafted feature estimation since the former performs better under low SNR.

5.2.2. Identification on More Transmitters. To observe the supremum and identification accuracy with a larger device scale, we implemented the simulation with 400 transmitters. Figure 3 shows the results considering 400 transmitters with the same CFO ranges as in Figure 2. It also presents the identification accuracy of the two shallow classifier-based and deep learning-based schemes with the same CFO ranges and the same transmitter quantity under the AWGN channel. Comparing Figures 3(a)–3(d) with Figures 2(a)–2(d), it is evident that the accuracy of both schemes with 400 transmitters is lower than those with 50 transmitters, respectively. Also, the supremum decreases too. Combined with Figure 2, we find that at the SNR of about 45 dB, no matter how the quantity of devices changes, the shallow classifier-based scheme consistently exceeds the deep learning-based one.

5.2.3. Identification Under Static Rayleigh Channel. In Figure 4, we also compare the identification accuracy under the static Rayleigh channel with the supremum. Comparing the performance of the deep learning-based scheme in

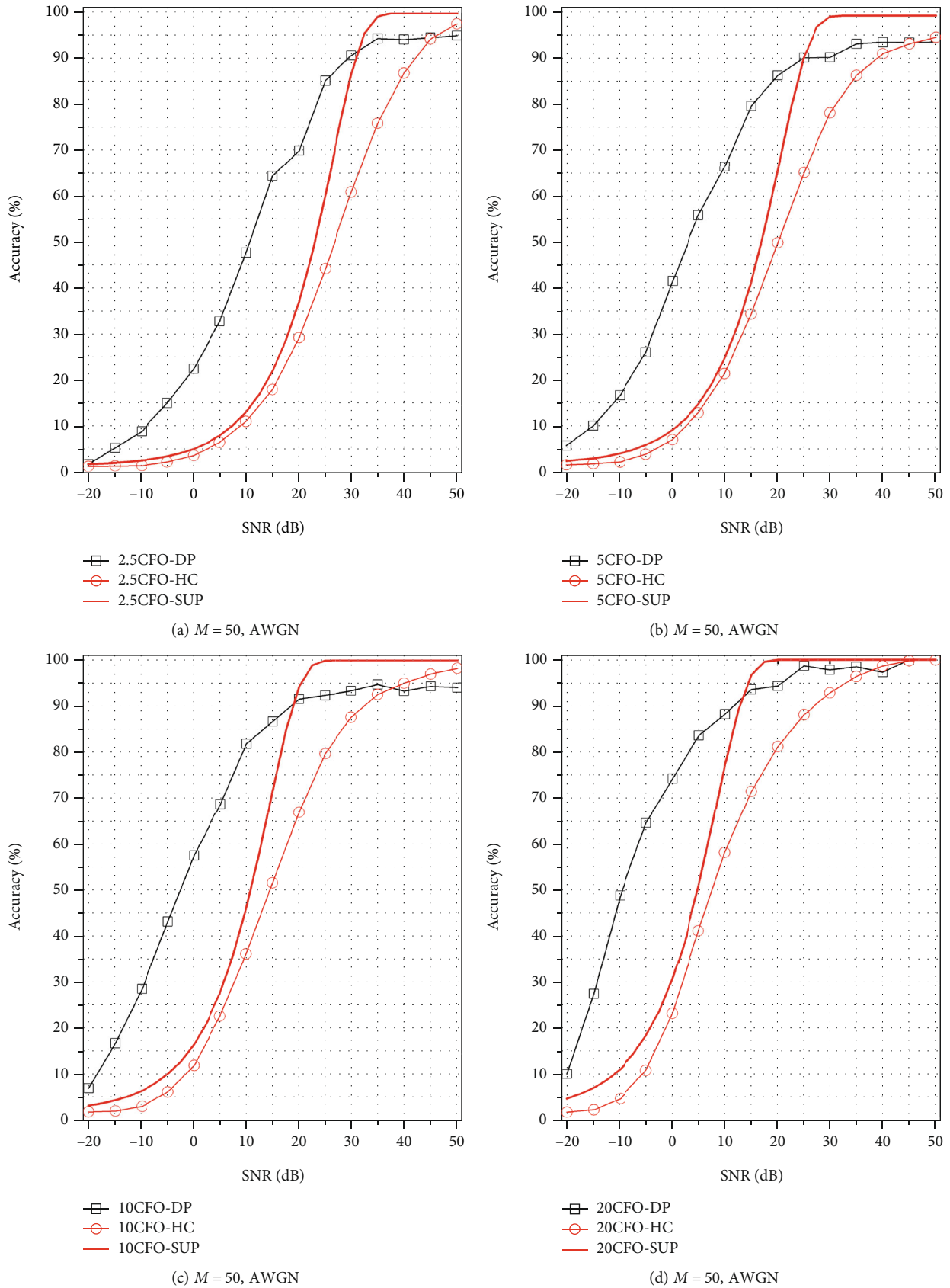
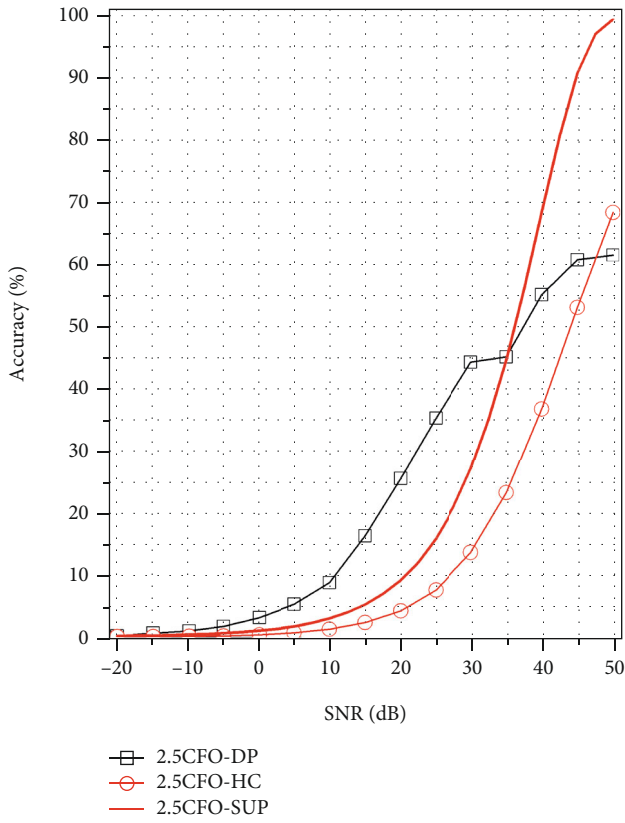
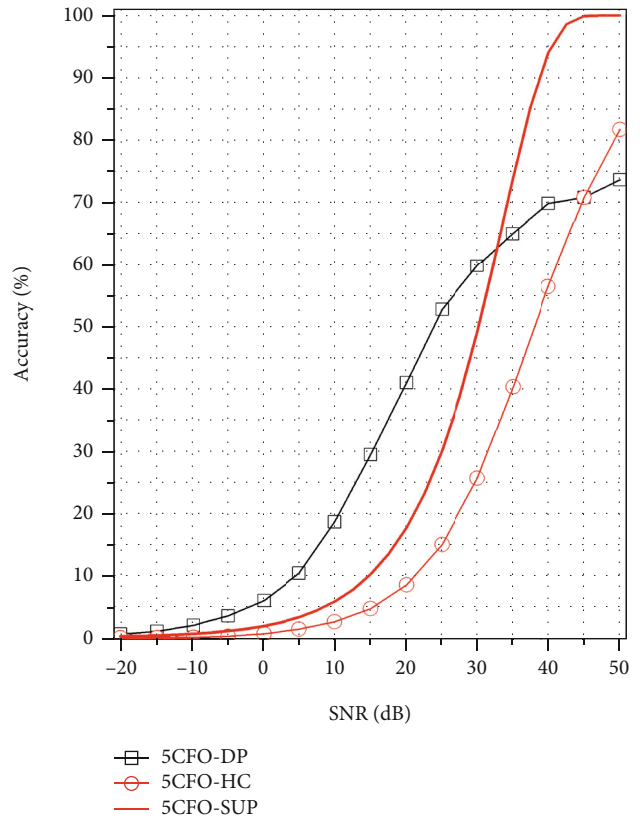


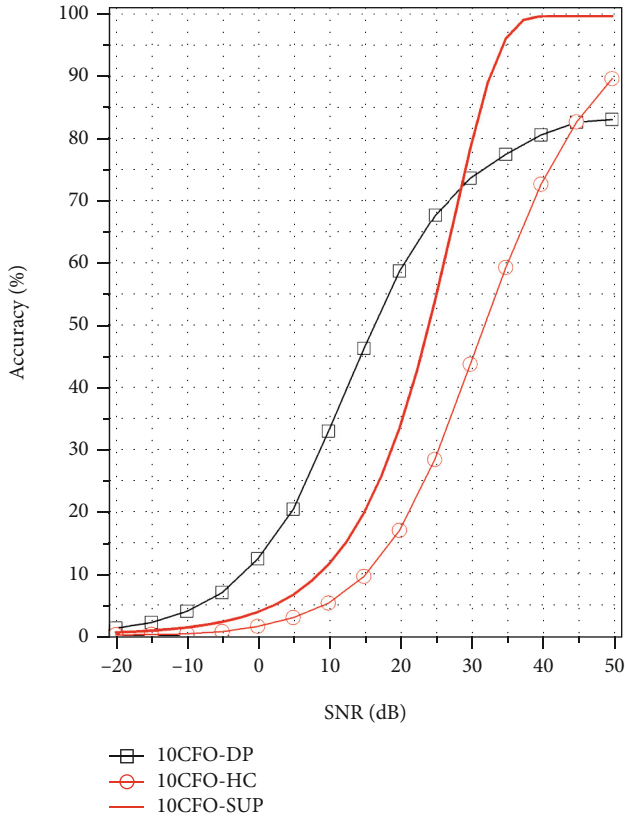
FIGURE 2: Identification accuracy of schemes using KNN with CFO (HC) and deep learning with raw IQ samples (DP) under different CFO ranges (e.g., 2.5CFO means $U_\epsilon = 2.5$ ppm) and AWGN channel vs. the accuracy supremum (SUP).



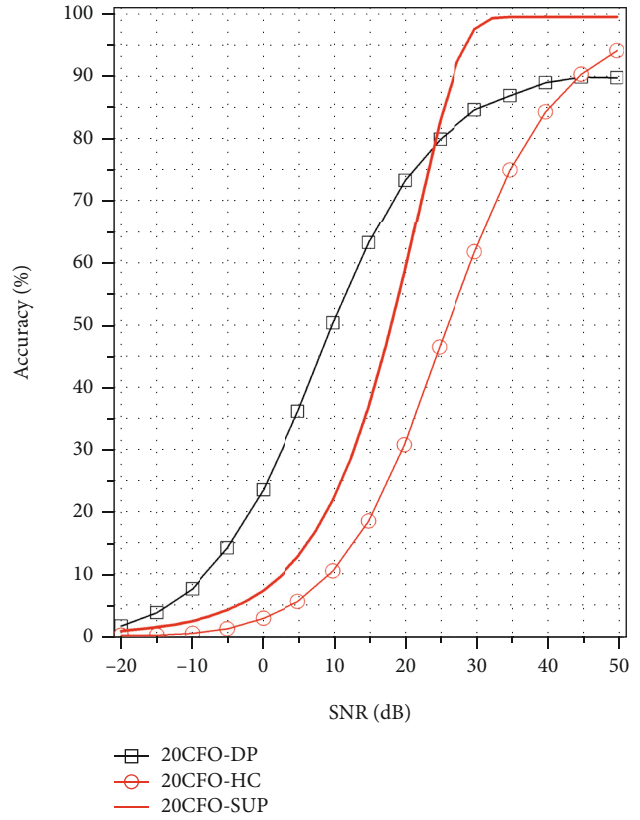
(a) $M = 400$, AWGN



(b) $M = 400$, AWGN

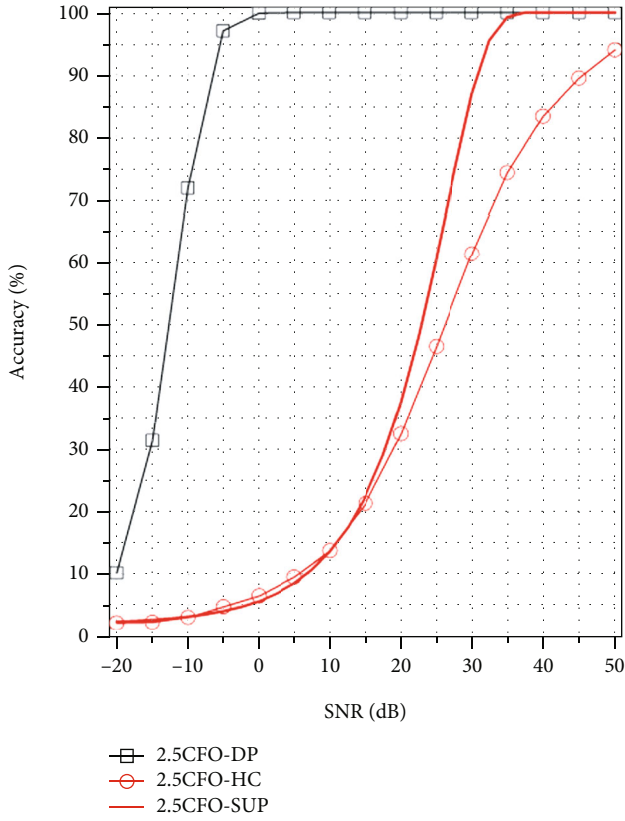


(c) $M = 400$, AWGN

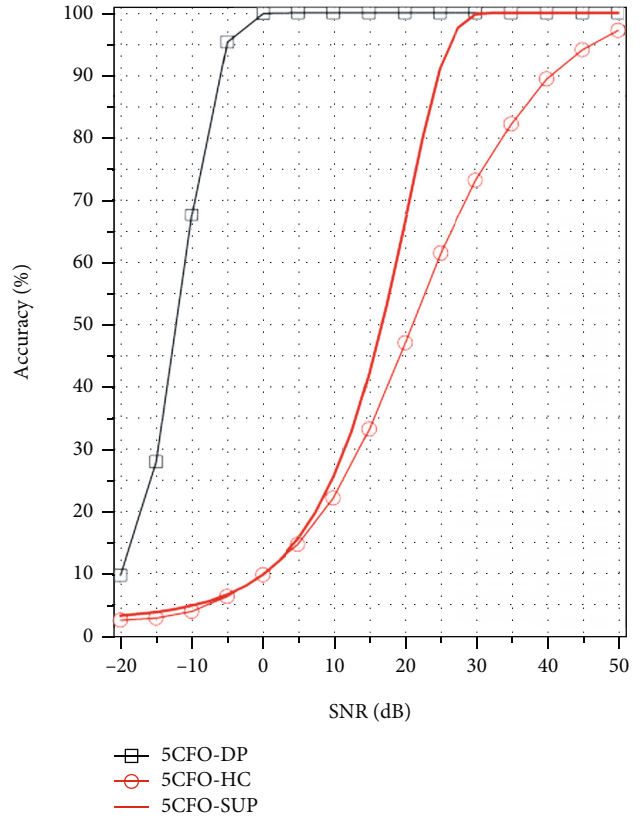


(d) $M = 400$, AWGN

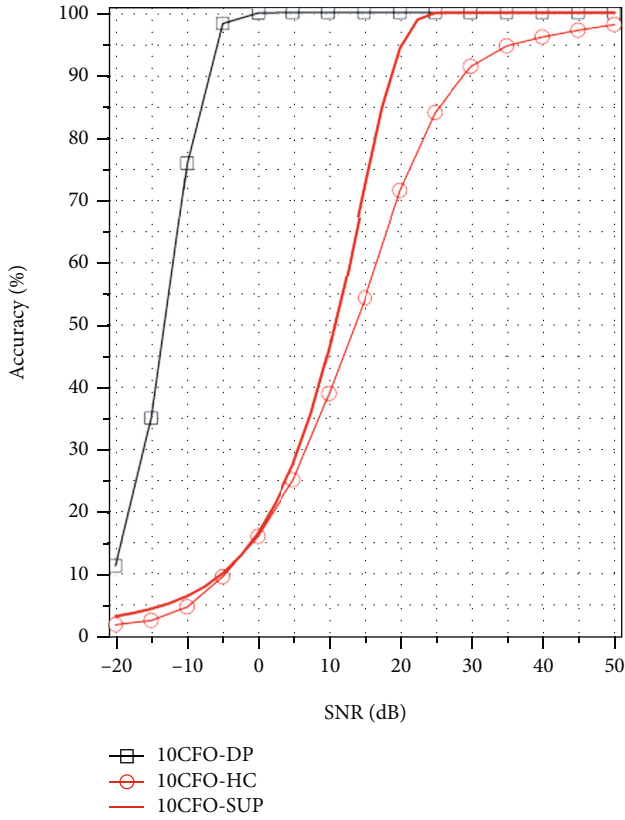
FIGURE 3: Identification accuracy under AWGN channel vs. the accuracy supremum.



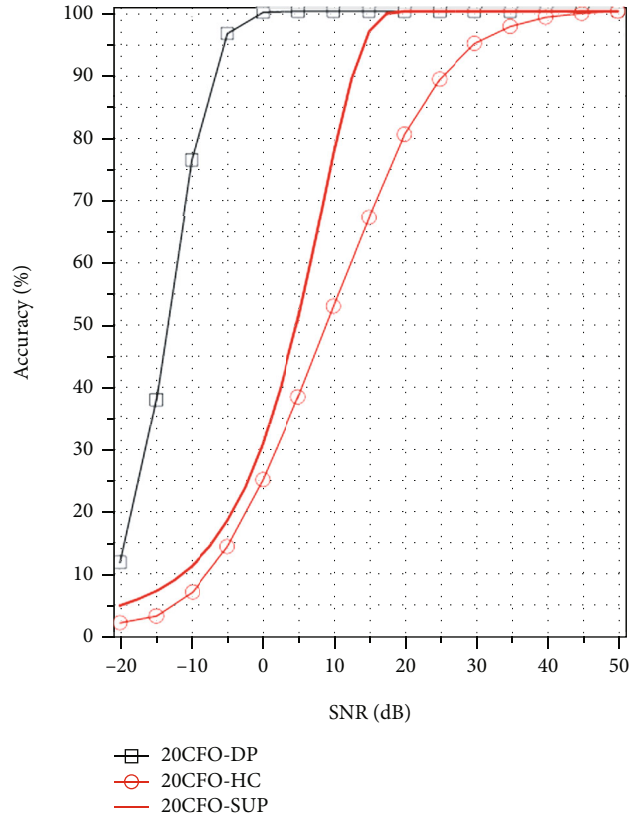
(a) $M = 50$, Rayleigh



(b) $M = 50$, Rayleigh

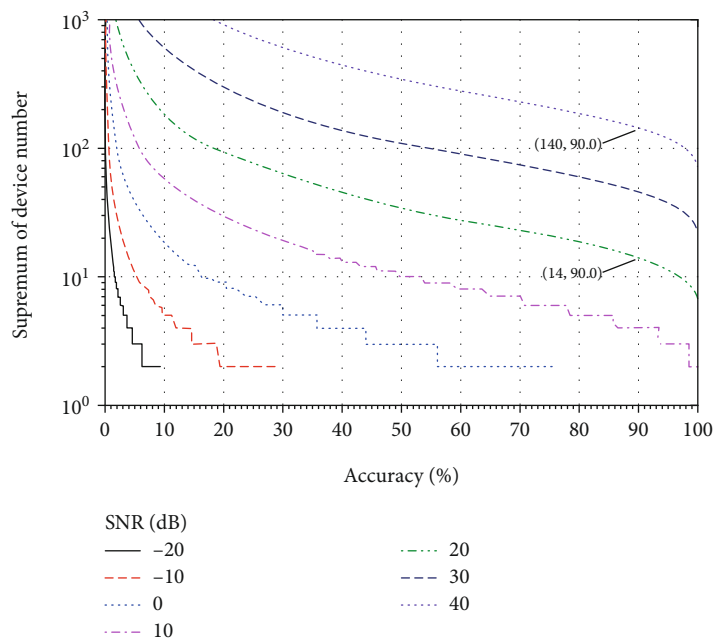


(c) $M = 50$, Rayleigh

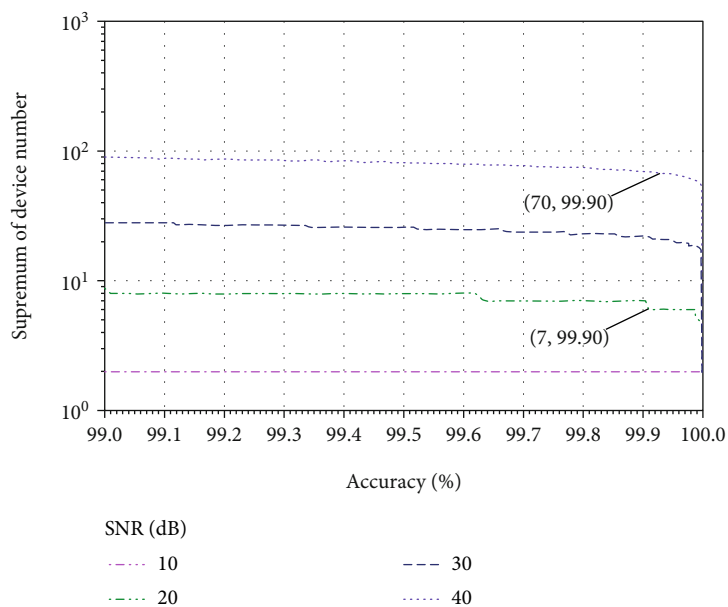


(d) $M = 50$, Rayleigh

FIGURE 4: Identification accuracy under Rayleigh channel vs. the accuracy supremum.



(a) $U_\epsilon = 2.5$ ppm



(b) $U_\epsilon = 2.5$ ppm

FIGURE 5: Continued.

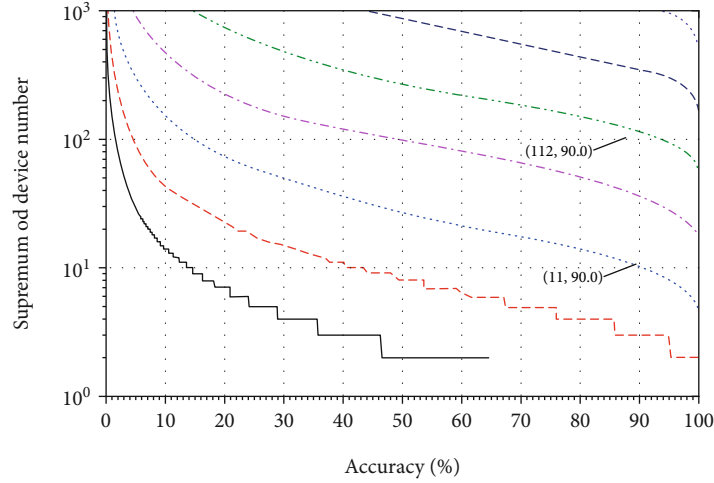
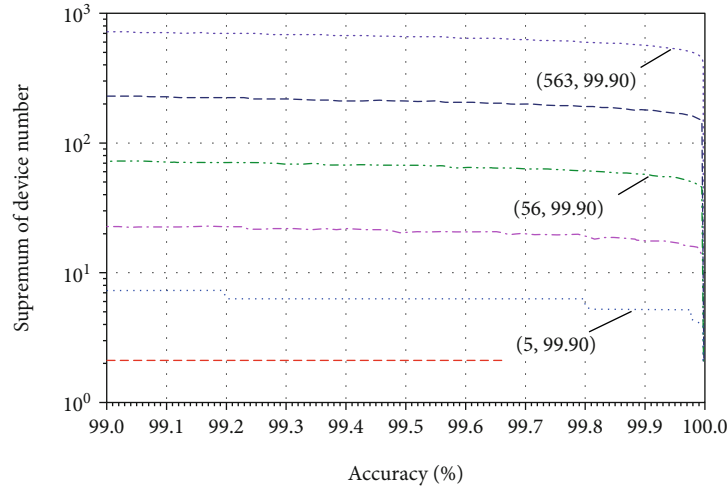
(c) $U_\epsilon = 20$ ppm(d) $U_\epsilon = 20$ ppm

FIGURE 5: The device quantity supremum using KNN with hand-crafted feature (i.e., CFO) under different identification accuracy and SNR with different CFO range.

Figure 2 with that of Figure 4, we can observe that the accuracy is improved obviously under the static Rayleigh channel in each CFO range. However, when looking into the three subfigures in Figure 4, we can find that expanding the CFO range has little effect on the accuracy improvement for the deep learning-based scheme. This result validates that deep learning can learn more information from the multipath channel than from the device's feature, which is consistent with the study of [32]. In other words, the identification accuracy of the deep learning-based scheme will be affected more by the channel than device features. On the other

hand, the shallow classifier-based scheme is less affected by the channel than the deep learning one since the accuracy under the Rayleigh channel is only improved slightly in each CFO range. Especially at low SNRs, the gap between the accuracy of the shallow classifier-based scheme and the supremum under the Rayleigh channel is slightly smaller than that of the AWGN channel.

5.3. *The Device Quantity Supremum.* Figure 5 presents the device quantity supremum under different SNRs considering CFO as the physical layer feature. Figures 5(b) and 5(d)

show the details of Figures 5(a) and 5(c), respectively, when the accuracy constraint $\geq 99\%$. As shown in Figures 5(a) and 5(c), the supremum will decrease with the increasing of the desired identification accuracy while also increases with the SNR. With the feature range of 2.5 ppm, the results in Figures 5(a) and 5(b) show that when the SNR < 20 dB, the receiver can identify no more than 15 transmitters under the required accuracy of 90%. Even at a high SNR of 40 dB, if a 90% accuracy is required, there should be no more than 140 transmitters. Moreover, when higher accuracy is required, the supremum will be smaller. For example, with the accuracy requirement reaching 99.9%, the device quantity supremum even decreased to 7 and 70 at the SNR of 20 and 40 dB, respectively. However, a small CFO range is quite common for some off-the-shelf wireless devices operating under rigorous synchronization requirements, such as mobile phones and high-end laptops [39].

In Figures 5(c) and 5(d), losing the feature range to 20 ppm, then we can see the supremum increases too. For the same accuracy requirement of 90%, the device quantity supremum can reach 112 at the SNR of 20 dB, more than seven times that in the CFO range of 2.5 ppm. But the supremum is only 11 when the SNR is 0 dB. With the high accuracy requirement of 99.9%, the supremum is 5, 56, and 563 at the SNRs of 0, 20, and 40 dB, respectively. Thus, it is evident that only considering one feature such as CFO, the device identification capability, i.e., the device number supremum, is small, especially when the range is small. Combined with the analysis of the supremum and the simulation results, we can find that a larger feature range, more precise feature estimation, and higher SNR of the signal can improve the accuracy and the device quantity supremum.

6. Conclusion

This paper analyzed the accuracy supremum and the device quantity supremum of the shallow classifier-based physical layer identification scheme based on the hand-crafted feature. Specifically, we mathematically analyzed and deduced the closed-form expression of the accuracy supremum of the identification scheme based on CFO. We also investigated the device quantity supremum, i.e., the supremum of the number of distinguishable devices. The simulation results are consistent with the theoretical analysis. According to the analysis and simulation, there is insufficient fingerprinting space only considering one feature, such as CFO. Thus, if we want to identify more devices, a larger feature range or higher SNR of the signal can help.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported in part by the National Key R&D Program of China (Grant No. 2018YFE0207600), the Natural Science Foundation of China (NSFC) under Grant 61972308, the Natural Science Basic Research Program of Shaanxi (Program No. 2019JC-17), the Hebei Science Supported Planning Projects Under Grant 20310701D, the JSPS A3 Foresight Program (Grant No. JPJSA3F20200001), and the Grants-in-Aid for Scientific Research (JSPS KAKENHI, Grant Number 20K19801).

References

- [1] A. C. V. Gummalla and J. O. Limb, "Wireless medium access control protocols," *IEEE Communications Surveys & Tutorials*, vol. 3, no. 2, pp. 2–15, 2000.
- [2] D. J. Leith and S. Farrell, "Contact tracing app privacy: What data is shared by Europe's GAEN contact tracing apps," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, pp. 1–10, Vancouver, BC, Canada, 2021.
- [3] Y. Mirsky and M. Guri, "DDoS attacks on 9-1-1 emergency services," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 6, pp. 2767–2786, 2020.
- [4] P. Zhang, J. Liu, Y. Shen, and X. Jiang, "Exploiting channel gain and phase noise for PHY layer authentication in massive MIMO systems," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4265–4279, 2021.
- [5] P. Zhang, Y. Shen, X. Jiang, and W. Bin, "Physical layer authentication jointly utilizing channel and phase noise in MIMO systems," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2446–2458, 2020.
- [6] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2604–2616, 2021.
- [7] Y. Lin, J. Jia, S. Wang, B. Ge, and S. Mao, "Wireless device identification based on radio frequency fingerprint features," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, Dublin, Ireland, 2020.
- [8] S. Zeng, Y. Chen, X. Li, J. Zhu, Y. Shen, and N. Shiratori, "Visibility graph entropy based radiometric feature for physical layer identification," *Ad Hoc Networks*, vol. 127, p. 102780, 2022.
- [9] J. Chang, Y. Xiao, and Z. Zhang, "Wireless physical-layer identification assisted 5g network security," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–5, Paris, France, 2019.
- [10] Y. Liu, J. Wang, J. Li, S. Niu, and H. Song, "Machine learning for the detection and identification of internet of things (IOT) devices: A survey," *IEEE Internet Things J*, vol. 9, 2021.
- [11] S. Chen, Z. Pang, H. Wen, K. Yu, T. Zhang, and Y. Lu, "Automated labeling and learning for physical layer authentication against clone node and Sybil attacks in industrial wireless edge networks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2041–2051, 2020.
- [12] B. Li and E. Cetin, "Design and evaluation of a graphical deep learning approach for RF fingerprinting," *IEEE Sensors Journal*, vol. 21, no. 17, pp. 19462–19468, 2021.
- [13] M. Ramasubramanian, C. Banerjee, D. Roy, E. Pasilião, and T. Mukherjee, "Exploiting spatio-temporal properties of I/Q

- signal data using 3d convolution for RF transmitter identification.” *IEEE Journal of Radio Frequency Identification*, vol. 5, no. 2, pp. 113–127, 2021.
- [14] J. M. McGinthy, L. J. Wong, and A. J. Michaels, “Groundwork for neural network-based specific emitter identification authentication for iot,” *IEEE Internet Things Journal*, vol. 6, no. 4, pp. 6429–6440, 2019.
- [15] T. Jian, B. C. Rendon, E. Ojuba et al., “Deep learning for rf fingerprinting: a massive experimental study,” *IEEE Internet Things Magazine*, vol. 3, no. 1, pp. 50–57, 2020.
- [16] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, “Oracle: optimized radio classification through convolutional neural networks,” in *Proceedings of The 38th IEEE International Conference on Computer Communications (INFOCOM 2019)*, pp. 370–378, Paris, France, 2019.
- [17] K. Merchant, S. Revay, G. Stantchev, and B. Noursain, “Deep learning for RF device fingerprinting in cognitive communication networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.
- [18] S. Balakrishnan, S. Gupta, A. Bhuyan, P. Wang, D. Koutsonikolas, and Z. Sun, “Physical layer identification based on spatial-temporal beam features for millimeter-wave wireless networks,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1831–1845, 2019.
- [19] A. Al-Shawabka, F. Restuccia, S. D’Oro et al., “Exposing the fingerprint: dissecting the impact of the wireless channel on radio fingerprinting,” in *Proceedings of The 39th IEEE International Conference on Computer Communications (INFOCOM 2020)*, pp. 646–655, Toronto, ON, Canada, 2020.
- [20] Z. S. WenhaoWang, K. Ren, and B. Zhu, “User capacity of wireless physical-layer identification: an information-theoretic perspective,” in *2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, Kuala Lumpur, Malaysia, 2016.
- [21] W. Wang, Z. Sun, K. Ren, and B. Zhu, “User capacity of wireless physical-layer identification,” *IEEE Access*, vol. 5, pp. 3353–3368, 2017.
- [22] W. Hou, X. Wang, J.-Y. Chouinard, and A. Refaey, “Physical layer authentication for mobile systems with time-varying carrier frequency offsets,” *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1658–1667, 2014.
- [23] P. Hao, X. Wang, and A. Behnad, “Performance enhancement of I/Q imbalance based wireless device authentication through collaboration of multiple receivers,” in *2014 IEEE International Conference on Communications (ICC)*, pp. 939–944, Sydney, Australia, 2014.
- [24] F. Restuccia, S. D’Oro, A. Al-Shawabka et al., “Deepradioid: real-time channel resilient optimization of deep learning-based radio fingerprinting algorithms,” in *Proceedings of the 20th International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pp. 51–60, Hawaii, USA, 2019.
- [25] 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 Journal*, vol. 6, no. 1, pp. 349–360, 2019.
- [26] X.-C. Yin, C. Yang, W.-Y. Pei, and H.-W. Hao, “Shallow classification or deep learning: an experimental study,” in *2014 22nd International Conference on Pattern Recognition*, pp. 1904–1909, Stockholm, Sweden, 2014.
- [27] M. Kantardzic, *Data mining: concepts, models, methods, and algorithms*, John Wiley & Sons, 2011.
- [28] K. O’Shea and R. Nash, “An introduction to convolutional neural networks,” 2015, <https://arxiv.org/abs/1511.08458>.
- [29] Q. Wu, C. Feres, D. Kuzmenko et al., “Deep learning based rf fingerprinting for device identification and wireless security,” *Electronics Letters*, vol. 54, no. 24, pp. 1405–1407, 2018.
- [30] 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 Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2305–2317, 2021.
- [31] M. Morelli and U. Mengali, “Carrier-frequency estimation for transmissions over selective channels,” *IEEE Transactions on Communications*, vol. 48, no. 9, pp. 1580–1589, 2000.
- [32] K. Sankhe, M. Belgiovine, F. Zhou et al., “No radio left behind: radio fingerprinting through deep learning of physical-layer hardware impairments,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 165–178, 2019.
- [33] J. Li, G. Liu, and G. B. Giannakis, “Carrier frequency offset estimation for OFDM-based WLANs,” *IEEE Signal Processing Letters*, vol. 8, no. 3, pp. 80–82, 2001.
- [34] D. Huang and K. B. Letaief, “Carrier frequency offset estimation for OFDM systems using null subcarriers,” *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 813–823, 2006.
- [35] W. Hou, X. Wang, and J.-Y. Chouinard, “Physical layer authentication in OFDM systems based on hypothesis testing of CFO estimates,” in *2012 IEEE International Conference on Communications (ICC)*, pp. 3559–3563, Ottawa, Canada, 2012.
- [36] M. M. U. Rahman, A. Yasmeen, and J. Gross, “Phy layer authentication via drifting oscillators,” in *2014 IEEE Global Communications Conference*, pp. 716–721, Sydney, Australia, 2014.
- [37] L. Yang, S. Song, Y. Gong, H. Gao, and W. Cheng, “Nonparametric dimension reduction via maximizing pairwise separation probability,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 10, pp. 3205–3210, 2019.
- [38] W. Bian and D. Tao, “Max-min distance analysis by using sequential sdp relaxation for dimension reduction,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 1037–1050, 2010.
- [39] J.-C. Lin, “Synchronization requirements for 5g: an overview of standards and specifications for cellular networks,” *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, pp. 91–99, 2018.