Hybrid RFF Identification for LTE Using Wavelet Coefficient Graph and Differential Spectrum

Linning Peng, *Member, IEEE,* Zhenni Wu, Junqing Zhang, *Member, IEEE,* Ming Liu, Hua Fu, and Aiqun Hu

***Abstract*—The growing popularity of 4G/5G mobile devices has led to an increase in demand for wireless security. Radio frequency fingerprint (RFF) technique is an emerging approach for device authentication using intrinsic and unique hardware impairments. In this paper, we propose an RFF-based method to identify rogue/unknown long term evolution (LTE) terminals. This is achieved by combining wavelet transform (WT) coefficient graphs and differential spectrum. The proposed method involves extracting 48 levels of wavelet coefficients from the transient power-off of the physical random access channel (PRACH) signal and representing them in a WT graph. The steady-state part of the PRACH signal after a frequency domain differential processing between the adjacent spectrum is extracted. To detect unknown attack devices, an identification scheme based on an autoencoder (AE) is designed. Two different AE network structures are designed based on the proposed features, and a hybrid identification structure is proposed. An experimental evaluation system is set up with seven mobile phones from three categories and one universal software radio peripheral (USRP) software-defined radio (SDR) platform. Training and testing datasets are collected under different conditions such as location, working times, and dates. Experimental results show that rogue devices can be identified with an accuracy up to 98.84% for different categories and 90.27% for different individuals.**

***Index Terms*—Radio frequency fingerprint, wavelet transform, spectrum, autoencoder, LTE.**

I. Introduction

Manuscript received xxx; revised xxx; accepted xxx. Date of publication xxx; date of current version xxx. This work was supported in part by the National Key Research and Development Program of China (2022YFB4300300), National Natural Science Foundation of China under Grant 62171120, 62001106, 61971029, and 62221001, Jiangsu Provincial Key Laboratory of Network and Information Security No. BM2003201, Guangdong Key Research and Development Program under Grant 2020B0303010001, and Purple Mountain Laboratories for Network and Communication Security. The work of J. Zhang was in part supported by the UK Engineering and Physical Sciences Research Council (EPSRC) New Investigator Award under grant ID EP/V027697/1. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Accepted Manuscript version arising. This paper was presented in part at the IEEE VTC-Fall 2022. The associate editor coordinating the review of this paper and approving it for publication was xx. (*Corresponding author: Linning Peng.*)

L. Peng, Z. Wu, and H. Fu are with the School of Cyber Science and Engineering, Southeast University, 210096 Nanjing, China. (e-mail: pengln@seu.edu.cn;220205055@seu.edu.cn;hfu@seu.edu.cn)

J. Zhang is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, United Kingdom. (email: junqing.zhang@liverpool.ac.uk)

M. Liu is with the Engineering Research Center of Network Management Technology for High Speed Railway of Ministry of Education, Collaborative Innovation Center of Railway Traffic Safety, Beijing Jiaotong University, 100044, Beijing, China. (email: mingliu@bjtu.edu.cn)

A. Hu is with the School of Information Science and Engineering, Southeast University, 210096, Nanjing, China. (email: aqhu@seu.edu.cn)

L. Peng, H. Fu, and A. Hu are also with the Purple Mountain Laboratories, Nanjing, 210096, China.

Digital Object Identifier xxx

THE proliferation of radio frequency communication devices has triggered massive connectivity of voice and data worldwide. This has increased wireless user exposure and driven the need for improved security measures [1], [2]. Traditional solutions usually use cryptographic mechanisms to provide protection for the integrity and confidentiality of communication data. These methods rely on security protocols and a common key between the transmitter and the receiver. Public key cryptography (PKC) is popular to share the cryptographic key but is threatened by quantum computers because PKC relies on computationally-expensive mathematical operations [3].

The global 4G long-term evolution (LTE) users have soared from 9 million to 4.7 billion from 2011 to 2021 [4]. The 5G NR technology is an extension of 4G LTE technology, with many common physical layer designs [5], [6]. The LTE standard specifies authentication mechanisms for users and base stations by using authentication and key agreement (AKA) protocols. Although LTE system has comprehensive protocols design for security [7], [8], there have still been attacks on LTE systems [9]–[12]. Fei and Wang revealed two vulnerabilities that could lead to denial of service (DoS) and leak a user's international mobile subscriber identity (IMSI) [9]. Lichtman *et al*. proposed a specific physical channel-oriented attack in [10], which requires less transmission power and is more efficient than the traditional method of jamming. In [11], Tan *et al*. proposed a data-plane signaling attack, which can cause DoS threats in cellular networks. Erni *et al*. proposed a man-in-the-middle (MITM) attack during the LTE radio resource control (RRC) connection process, which can cause either large-scale DoS or privacy leakage [12]. Most of the above-mentioned attacks were implemented at the upper layers. The main idea behind these attacks is to exploit vulnerabilities in the design of LTE protocols to carry out the attacks.

It is promising to design non-cryptographic-based solutions as complements in the physical layer. Radio frequency fingerprint (RFF) is such an approach for wireless device authentication. Due to the inevitable variation in the manufacturing process, there are innate differences in analog components of wireless transceiver chains, which are difficult to be imitated. These hardware imperfections are embedded in the transmitted signals, which can be extracted by a receiver [13]–[15]. RFF has good characteristics of tamper resistance, which can well provide a reliable mechanism for the authentication of radio devices. RFF identification has been widely studied with Wi-Fi [16]–[18], LoRa [19]–[22], ZigBee [23]–[26], Bluetooth [27], ADS-B [28], [29] etc.

The research on RFF in cellular communication systems

mainly focuses on identifying the RFF of base stations, which are referred to as evolved Node B (eNB) in LTE terminology. For example, Demers and St-Hilaire collected LTE signals of real base stations by a high-performance spectrum analyzer at different locations and extracted 13 signal characteristic coefficients. The base stations were identified by support vector machine (SVM) classifier [30]. The RFF-based base station identification could be primarily used for identifying fake base stations [31], [32]. However, there has been limited research on RFF for cellular user equipment (UE). A significant number of malicious network access attempts originated from UEs can result in DoS attacks on cellular network eNB. In the physical layer of 4G/5G LTE systems, the physical random access channel (PRACH) is the first message sent by a UE when connecting to the eNB [5]. The PRACH signal does not contain any identity information. Therefore, it might be possible that the malicious UEs flood the PRACH to eNB to launch PRACH spoofing attacks [10], [33]. In this case, physical layer DoS attacks become possible. A large number of invalid PRACH preambles will greatly consume the upper-layer resources of the eNB, resulting in network congestions that make the network unable to respond to the network access requests from legitimate UEs.

Therefore, RFF identification based on LTE PRACH signals will be of great significance. Li *et al*. used 10 different universal software radio peripheral (USRP) devices to emulate LTE terminals and achieved identification of 10 terminals based on PRACH signals [34]. It should be noted that there are significant differences in terms of hardware specifications between USRP-emulated LTE devices and real LTE terminals that make its applicability unclear. Moreover, in the PRACH accessing process, UE will compensate their carrier frequency offset based on the physical broadcast channel (PBCH) signal from the eNB, thereby reducing the differentiation of their RFF characteristics and complicating the RFF identification process. In [35], 6 real LTE terminal signals were collected, and a multi-channel convolutional neural network (MCCNN) was designed based on the extracted differential constellation trace figure (DCTF) features for LTE terminal identifications. Qiu *et al*. extracted residual transient segment (RTS) features from the varying PRACH preambles of real LTE eNB and classified LTE mobile phones from 5 different brands successfully [36]. In summary, most existing research on LTE terminal identification was applicable for fixed observation time and sampling location, and the LTE terminals considered for identification were chosen from different brands. LTE terminals used are also of various brands. The identification of LTE terminals of the same brands was not investigated yet in the literature. In reality, multiple terminals of the same brands may be associated with the same network eNB, which will increase the identification difficulty.

Most of the RFFI studies focus on the multi-class classification using convolutional neural network (CNN) [19], [22], [23], [37], long short-term memory (LSTM) [20], [22], [37], multilayer perceptron (MLP) [19], [21], [22], and tramsformer [37]. The work in [19] classified 22 LoRa devices by MLP and CNN. Rogue device identification is another set of RFFI works that can detect malicious/unknown devices [17], [18], [38]. Hanna *et al.* adopted a few methods to detect unknown devices using e.g., autoencoder (AE) and openMax [18]. And unsupervised learning is more applicable in these scenarios because it is not easy to involve rogue devices in the training stage. However, detecting rogue LTE terminals based on RFF has not been explored yet.

According to the segments of the signal, the main RF features can be categorized into transient parts and steady-state parts [16]. Transient parts are produced when the power of the transmission raises from zero to the specified working power or in the reverse process. Steady-state parts are present when the transmitter is in stable mode. The wavelet transform (WT) is a mathematical technique that provides a time-frequency representation of a signal, allowing for the adjustment of time-frequency resolution. Therefore, the WT can be regarded as a time-frequency analysis method with a flexible window, which offers an enhanced ability to analyze non-stationary signals. Hippenstiel and Payal transformed the transient signals of four transmitters into the wavelet domain, extracted the wavelet coefficients under three WT scales in turn, and used the peak values to obtain the Euclidean distance for classification [39]. Ho *et al*. extracted 4 WT scales of wavelet coefficients and applied the histogram to measure the frequency of different energy occurrences at each WT scale. They classified code-division multiple access (CDMA) signals and global system for mobile communications (GSM) signals according to a selected threshold. The accuracy can remain 100% even at low SNR [40]. Though showing a strong capability in signal classification, WT has not been considered in the RFF feature extraction for LTE devices.

In this paper, we proposed a hybrid RFF identification method using the transient and steady-state parts of the LTE PRACH signal. We conduct a comprehensive experimental evaluation under three environmental variables, namely sampling dates, sampling locations, and working hours. The main contributions of our work are as follows:

- A hybrid RFF feature extraction method that involves both transient and steady-state features is proposed for LTE devices. The transient features of PRACH signals are extracted by WT and the steady parts are processed by the differential fast Fourier transform (D-FFT).
- A hybrid AE detector is designed to detect rogue terminals based on the one-dimensional (1D) differential spectrum feature and two-dimensional (2D) WT feature. The AE can effectively fuse the features of different dimensions. The detector has environmental robustness and high discrimination.
- We have carried out extensive experiments by constructing a signal collection system and sampling real LTE signals under different times, locations, and working hours. The performance of the proposed method is intensively evaluated in the actual environment. The identification accuracies are up to 98.84% for identifying different categories and 90.27% for identifying different individuals.

In our previous work [41], we identified different LTE devices based on wavelet coefficient graphs and AE network. Through experiments conducted in a fixed position, we have

found that the WT feature can achieve a higher identification accuracy compared to the D-FFT feature. This paper advances the research by combining WT and D-FFT features from different dimensions at different signal parts and building a dataset under different environmental variables. This approach improves the identification accuracy and proves the robustness of our method through comprehensive experiments in two typical scenarios.

The rest of this paper is organized as follows. Section II introduces the structure and properties of the LTE PRACH signal. Section III introduces the research problems and presents the proposed solutions. The signal preprocessing, RFF feature extraction, and hybrid AE design are presented in Sections IV, V, and VI, respectively. Section VII provides the experimental setup and analyzes the results in different experimental scenarios. Section VIII concludes the paper.

II. LTE PRACH PRIMER

In LTE systems, a PRACH signal is sent by a user equipment (UE) to an eNB for establishing RRC connection, obtaining uplink synchronization, and requesting resources for future interaction. The data transmission is performed after the radio resource control (RRC) connection. Random access occurs when UE starts uplink time synchronization or the handover between cells [42]. Malicious users can execute a flooding attack in this process [33]. Due to the limited spectrum resources, the massive number of devices causes an overload in the random access channel (RACH) that, in turn, decreases the overall performance of the LTE systems [43].

In LTE systems, the eNB broadcasts the parameters of the access signal through the physical broadcast channel (PBCH). The UE randomly selects one of the 64 PRACH signals for access based on the parameters from PBCH. Both eNB and UE can generate standard PRACH signals from the selected parameters. The LTE PRACH signal utilizes 6 resource blocks (RBs), occupying a bandwith of 1.08 MHz. The carrier frequency of the PRACH signal is also fixed from the PBCH parameters. The value of *rootSequenceIndex* (RSI) parameter ranges from 0 to 838, which determines the preamble format of the PRACH signal. Specifically, the PRACH preamble is generated from a Zaddoff-Chu (ZC) sequence

$$p_u(n) = e^{-j\frac{\pi u n(n+1)}{N_{\text{ZC}}}}, \quad 0 \le n \le N_{\text{ZC}} - 1, \tag{1}$$

where n is the time index of ZC sequence, and u is the physical root sequence index of RSI. The detailed mapping table of u and RSI is given in [5]. N_{ZC} is the length of the ZC sequence, whose value is 839 in preamble format 0 to 3 and 139 in preamble format 4.

RSI is used to generate a ZC root sequence as the base sequence. There are 64 different sequences generated by cyclically shifting the base sequence and a UE can randomly select one preamble based on the parameter sent from the eNB. The cyclic shift interval is referred to as C_v, and the definition of C_v differs between restricted and unrestricted sets. In this paper, we only consider the case of unrestricted sets, and C_v is defined as follows

$$C_v = vN_{cs},\ v = 0, 1, ..., \lfloor N_{\text{ZC}}/N_{\text{CS}} \rfloor - 1, \tag{2}$$

where N_{CS} is a fixed interval specified by the parameters of *zeroCorrelationZoneConfig* and *Highspeedflag* in [5], $\lfloor \cdot \rfloor$ is the floor operation, v is the cyclic shift index. The cyclically shifted ZC sequence $p_{u,v}(n)$ can be expressed as

$$p_{u,v}(n) = p_u((n + C_v)_{N_{\text{ZC}}}), \tag{3}$$

where the subscript $(\cdot)_{N_{\text{ZC}}}$ represents the modulo N_{ZC} operation. When the cyclic shift operation cannot generate 64 preambles by the same u-th root sequence, i.e. $63 \times N_{\text{CS}} \ge N_{\text{ZC}}$, the next *rootSequenceIndex* will be chosen to continue the sequence generation. In general, for a specific eNB, the RSI and the cyclic shift interval are unchanged over a period of time. In practice, all 64 PRACH preambles associated with an eNB can be generated by a single RSI using a small cyclic shift interval.

The time-discrete PRACH signal is represented as

$$\begin{aligned} s_{idx}(i) =& \beta_{\text{PRACH}} \sum_{k=0}^{N_{\text{ZC}}-1} \sum_{n=0}^{N_{\text{ZC}}-1} p_{u,v}(n) e^{-j\frac{2\pi nk}{N_{\text{ZC}}}} \\ & \cdot e^{j2\pi(k+\varphi+K(k_0+\frac{1}{2}))\Delta f_{\text{RA}}(i-i_{\text{CP}})}, \end{aligned} \tag{4}$$

where idx, ranging from 0 to 63, is the index of signals generated from 64 kinds of $p_{u,v}(n)$ whose content is determined by u and v; idx equals v when all the 64 PRACH preambles can be generated from the same u-th root ZC sequence. β_{PRACH} is an amplitude scaling factor, φ is a fixed offset in the frequency domain, Δf_{RA} is PRACH preamble subcarrier interval, $K = \frac{\Delta f}{\Delta f_{\text{RA}}}$, and $\Delta f = 15$ kHz, is the LTE uplink symbol subcarrier interval. k_0, φ, K together consist of the frequency offset factor. i_{CP} is the length of cyclic prefix (CP).

The diagram of the PRACH generation system is shown in Fig. 1. Taking the preamble format 0 as an example. The values of the above parameters are $N_{\text{ZC}} = 839$, $\varphi = 7$, $\Delta f_{\text{RA}} = 12.5$ kHz, $K = 12$, $k_0 = -36$. When RSI is set to 0, the corresponding basic ZC sequence value u is 129. The parameter *highSpeedFlag* is set to false, and *zeroCorrelationZoneConfig* is set to 1, which decides that the value of the fixed interval N_{cs} is 13. The maximum cyclic shift interval is $13 \times 63 = 819$, which is less than the ZC sequence length 839. Therefore, 64 PRACH symbols can be generated from the same u-th root ZC sequence. At first, the ZC sequence $p_{u,v}$ is generated by the aforementioned procedures. $p_{u,v}(n)$ is converted via the discrete Fourier transform (DFT) to

$$P_{u,v}(k) = \sum_{n=0}^{N_{\text{ZC}}-1} p_{u,v}(n) e^{-j\frac{2\pi nk}{N_{\text{ZC}}}}. \tag{5}$$

According to [5], the DFT size $N_{\text{DFT}} = \frac{1}{\Delta f_{\text{RA}} T_s} = 24,576$, where T_s is the sampling period $(1/(30.72\ \text{MHz}))$. The inverse discrete Fourier transform (IDFT) process is given as

$$s'_{idx}(i) \triangleq \sum_{n=0}^{N_{\text{ZC}}-1} P_{u,v}(k) e^{j\frac{2\pi ik}{N_{\text{IDFT}}}}. \tag{6}$$

Then $s'_{idx}(i)$ is multiplied by the frequency offset factor $e^{j2\pi(\varphi+K(k_0+1/2))i/N_{\text{IDFT}}}$ in the time domain. At last, the i_{CP} samples of the tail part of the sequence are copied to the beginning of the sequence, forming the CP.

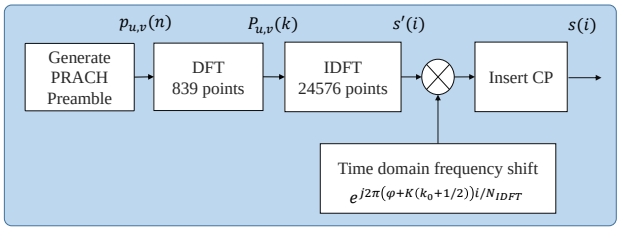
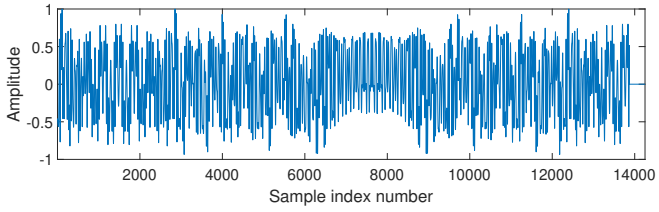
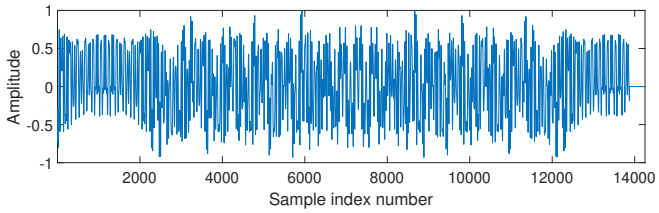


Fig. 1. PRACH baseband signal generation progress.



(a)



(b)

Fig. 2. Two LTE PRACH signals in the time domain from the same terminal. (a) Preamble index v at 0. (b) Preamble index v at 34.

We generated two PRACH preambles with different indexes using simulation and showed their time-domain waveforms in Fig. 2. The two waveforms are generated when the UE randomly selects v=0 and v=34 with the aforementioned settings. It can be observed that, although both time-domain waveforms are generated by the same u-th root ZC sequence, they exhibit significant differences after undergoing different cyclic shifts. The eNB is capable of detecting the PRACH signal randomly generated by the UE during competitive access process.

III. Problem Statement and System Overview

A. Problem Statement

As can be observed from (1) and (3), the preamble sequence is not fixed for a UE, which will result in various time domain signals. As exemplified in Fig. 2, the time domain waveforms are different when they are generated by a different preamble index. Therefore, using the time domain I/Q samples for RFF identification becomes difficult due to their temporal variations. It is important to extract other stable RFF features. In this paper, we assumed all the LTE terminals are connected to the same eNB, which means they have the same u-th root ZC sequence. However, in practical scenarios, LTE terminals will be connected to different eNBs when they are moving around. The different eNBs will have a different RSI, which results in different ZC root sequences u. Extracting RFF features under such circumstance has been studied in our previous work [36].

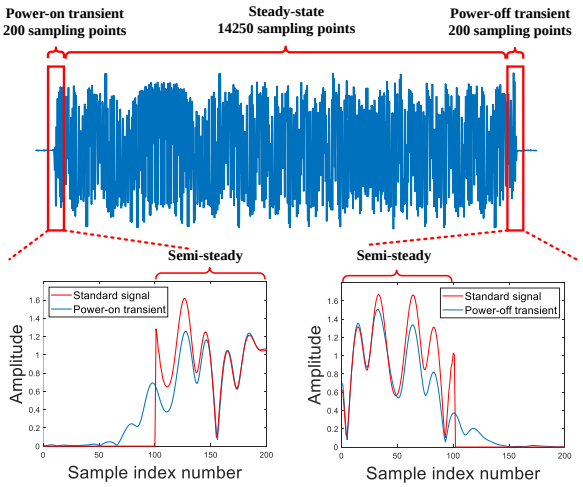


Fig. 3. The transient parts and steady-state part of PRACH preamble.

Because the bandwidth of LTE PRACH signals is 1.08 MHz, we used a sampling rate of 16 MSamples/s. Such oversampling will allow us to get good resolution signals for RFF. A real captured waveform of the PRACH preamble is shown in Fig. 3. There is a semi-steady part, including the power-on and power-off transients. The preamble signal received in the wireless propagation channel deviates from the standard one. In particular, with a sampling frequency of 16 MSamples/s, we take 100 samples before and after the synchronization start point to form a window of 200 samples to capture the power-on signal features. Similarly, we take 200 samples, with 100 samples before the synchronization endpoint and 100 samples after the synchronization endpoint, to capture the power-off signal features.[1] The relative position of transient parts and steady-state part of the PRACH signal is shown in Fig. 3.

Look more closely at the power-on and power-off transients as well as the steady-state parts. Fig. 4 depicts the transient amplitude and part of the steady-state amplitude of 70 PRACH preambles sent by LTE terminals. We constructed a testbed with 8 LTE terminals, which are from three models, as shown in Table I. Among the devices used in our study, LTE1-3 are from the same model and brand, while LTE5-7 are from a different brand but the same model. We also considered the potential for adversaries to use a software-defined radio (SDR) platform to launch a flooding attack on the PRACH signal in the case of denial of service (DoS) attacks. As a result, we also included the USRP B205 SDR. as an access terminal for experiments. These experimental devices cover a variety of types, brands, and models, allowing us to conduct a large number of experiments and collect a diverse dataset of PRACH signals in different scenarios. It can be seen from Fig. 4(a) and Fig. 4(b) that the transient parts of the same terminal

[1]Note that the length of the transient part should not exceed 20 μs according to the specification [5]. It means that our choice of 200-point observation window size can capture most of the transient part. In fact, the number of non-zero samples before the synchronization start point or after the synchronization endpoint is usually less than 100 in most of our practical measurements. This window size is sufficient for the observation of signal transient behaviour.

TABLE I
LTE DEVICE INFORMATION.

Device	Model	RF Transceiver	Baseband Chip	LTE Protocol Version
LTE1	Google Nexus 5	Qualcomm WTR1605L	Qualcomm MSM8974	LTE Cat.4
LTE2	Google Nexus 5	Qualcomm WTR1605L	Qualcomm MSM8974	LTE Cat.4
LTE3	Google Nexus 5	Qualcomm WTR1605L	Qualcomm MSM8974	LTE Cat.4
LTE4	Google Nexus 6P	Qualcomm WTR3925	Qualcomm MSM8944	LTE Cat.6
LTE5	Huawei P9	Hisilicon Hi6362	Kirin 955	LTE Cat.6
LTE6	Huawei P9	Hisilicon Hi6362	Kirin 955	LTE Cat.6
LTE7	Huawei P9	Hisilicon Hi6362	Kirin 955	LTE Cat.6
LTE8	USRP B205	AD9364	-	-

are relatively stable, which suggests that the signal features of the same terminal are stable. Moreover, the transient parts of PRACH preambles sent by the UEs of different models show significant differences. For instance, LTE1 and LTE6 are from different brands, and their power-on transient signals start at sample index 72 and 135, respectively, which is notably distinct. Furthermore, the peak-amplitude of the power-off transient signal at approximately sample index 85 exhibited significant differences and remained consistent across multiple experiments. In the meantime, according to Fig. 4(c) and Fig. 4(d), the features of the transient parts are somewhat similar for the terminals of the same model, which could lead to certain difficulties when distinguishing the terminals of the same model. Fortunately, as can be seen from Fig. 4(e) and Fig. 4(f), the steady-state parts of the preambles sent by the terminals of the same brand show identifiable differences. For instance, LTE1 shows a notable distinction from LTE2 and LTE3 of the same brand around sample index 10 and 125, respectively. This means that it is possible to build a distinguishable fingerprint if we take both transient and steady-state parts into account.

In summary, the time domain waveform of an LTE terminal will vary according to the preamble index, which makes them not ideal for RFF identification. A suitable signal representation should be designed. In addition, different parts of the waveform exhibit varied features. A hybrid approach to exploit these features independently is desirable but missing.

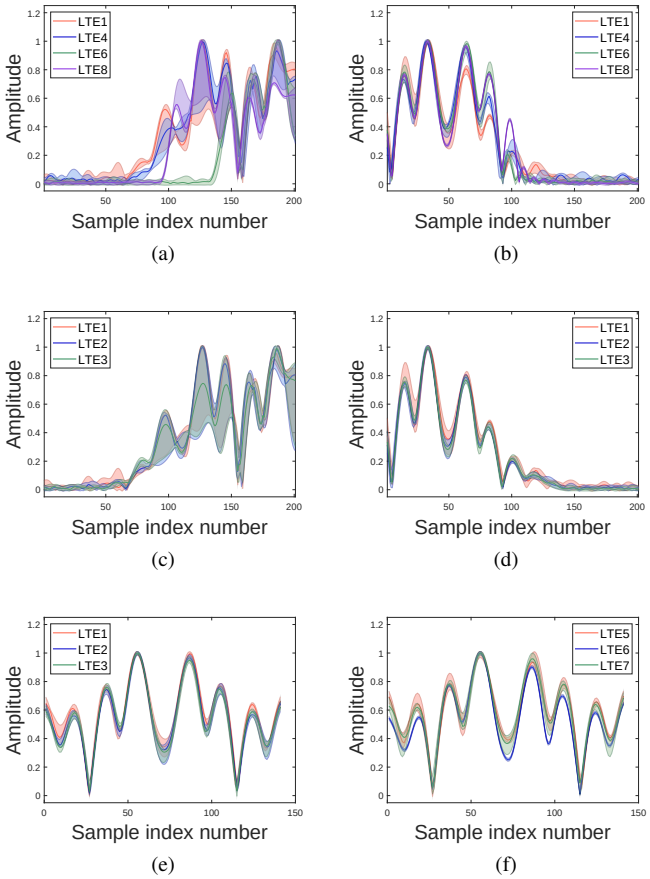


Fig. 4. Real collected LTE signals. (a) Power-on transient of different brands. (b) Power-off transient of different brands. (c) Power-on transient of the same brand (Google Nexus5). (d) Power-off transient of the same brand (Google Nexus5). (e) Part of the steady state of the same brand (Google Nexus5). (f) Part of the steady state of the same brand (Huaiwei P9).

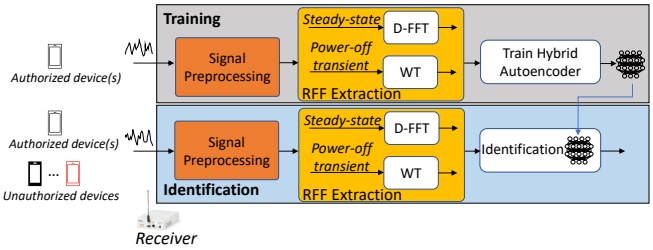


Fig. 5. Overall framework of the hybrid LTE identification system.

B. Proposed Solution

In this paper, we designed a hybrid identification approach for LTE terminals. While the majority of the RFF identification literature focuses on closed-set classification, this paper targets identifying rogue unknown devices by exploiting autoencoder (AE) architecture for open-set identifications. The system consists of two stages, namely training and identification, as shown in Fig. 5.

1) Training: In the training stage, the receiver will capture a large number of signals from the authorized LTE terminals. The signals will be first processed by signal preprocessing algorithms, including time synchronization and frequency offset compensation, introduced in Section IV.

RFF extraction is the essential step as it extracts the unique hardware impairments from the captured wireless signals. As indicated in Section II, the preamble of PRACH is generated by different root ZC sequences, which will result in various I/Q sequences. Therefore, directly using I/Q samples for identification is not viable. In addition, as shown in Fig. 3, there are both steady-state and transients in the received signal, which inspires us to design customized extraction algorithms for them. Based on the above motivations, we designed D-FFT for the steady-state part and WT coefficient graph for the transient-off part. The two methods will be elaborated in

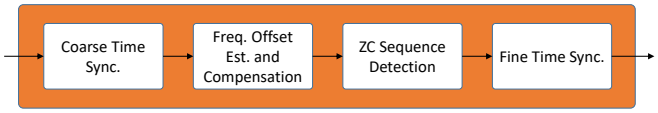


Fig. 6. LTE PRACH Signal preprocessing before RFF extraction.

Section V.

Finally, we designed two autoencoder models to process the D-FFT features and the WT graphs. We further integrate their MSE. Once the training is completed, a trained model will be produced, which will be used in the identification stage. The design details will be explained in Section VI.

2) Identification: In the identification, the receiver will capture a waveform, use the same signal preprocessing algorithms and RFF extraction methods in the training stage. Based on the extracted RFF features and the pre-trained model, the receiver will identify whether it is from the authorized device(s). We studied two cases, i.e., identifying authorized individual devices and identifying authorized sets (e.g., devices from the same brand). Different from closed-set classification which does not consider unknown or rogue devices, our paper will be able to detect them.

IV. Signal Preprocessing

In order to accurately obtain the start and end points of the transient RFF feature, the receiver must synchronize the PRACH signal accurately. In common usage scenarios, PRACH signal detection during UE access does not demand precise time and carrier frequency synchronization. We have designed a preprocessing method for synchronization based on the characteristics of the PRACH signal. The signal preprocessing is shown in Fig. 6, consisting of time synchronization and frequency offset estimation and compensation.

A. Coarse Time Synchronization

Coarse time synchronization is performed based on the CP of the ZC sequence. A coarse estimation of the starting point of the signal can be given by the peak of the correlation operation

$$D_{\text{coar}} = \underset{m}{\text{argmax}} \left(\sum_{i=0}^{L_{CP}-1} \hat{x}(i+m) \cdot \hat{x}^*(i+m+L_{\text{ZC}}) \right), \tag{7}$$

where $\hat{x}(i)$ is the collected I/Q samples, L_{ZC} and L_{CP} are the number of samples of the ZC sequence and the CP in the time domain, respectively. $(\cdot)^*$ denotes the conjugate of a complex number. The summation in the equation is the correlation of a sequence $\hat{x}$ and its delayed copy. The operator $\underset{m}{\text{argmax}}(\cdot)$ gives the value of m when the following expression reaches the maximum value.

B. Frequency Offset Estimation and Compensation

After the coarse time synchronization, a coarse estimation of the carrier frequency offset (CFO) is performed based on the offset of the signal spectrum

$$\Delta f_{\text{coar}} = \frac{(\hat{I}_L - I_L) + (\hat{I}_R - I_R)}{2N_{\text{FFT}}} \cdot F_s, \tag{8}$$

TABLE II
The CFO statistical results of LTE Devices

Device	Model	CFO Mean (Hz)	CFO Standard Deviation (Hz)
LTE1	Google Nexus 5	-1.3221	16.4308
LTE2	Google Nexus 5	1.1934	15.1441
LTE3	Google Nexus 5	2.8097	15.8251
LTE4	Google Nexus 6P	-18.5898	13.3214
LTE5	Huawei P9	7.2652	18.9580
LTE6	Huawei P9	2.7723	26.4250
LTE7	Huawei P9	6.0214	36.3126
LTE8	USRP B205	-2.1164	138.4374

where I_R and I_L are the right band index and the left band index of the standard signal on the spectrum, respectively. Similarly, $\hat{I}_R$ and $\hat{I}_L$ are the right and the left frequency band indexes of the received signal, respectively. N_{FFT} is the FFT size, the same as the signal length, and F_s is the sampling rate of the received signal. Next, we compensate frequency offset in the received signal

$$x_{\text{coar}}(i) = \hat{x}(i) \cdot e^{\frac{-j2\pi i \Delta f_{\text{coar}}}{F_s}}. \tag{9}$$

Then, the fine estimation of frequency offset is performed based on the CP

$$\Delta f_{\text{fine}} = \frac{\sum_{i=1}^{L_{CP}} \text{angle}\left(x_{\text{coar}}(i) \cdot x^*_{\text{coar}}(i+L_{\text{ZC}})\right)}{2\pi L_{CP} \cdot L_{\text{ZC}}} \cdot F_s, \tag{10}$$

where $\text{angle}(\cdot)$ represents the phase of a complex number. Then, the frequency offset is compensated as

$$x_{\text{fine}}(i) = x_{\text{coar}}(i) \cdot e^{\frac{-j2\pi i \Delta f_{\text{fine}}}{F_s}}. \tag{11}$$

With the help of aforementioned frequency offset estimation method, we can derive the distribution of the mean and standard variance of the CFO among different LTE terminals. The specific results are presented in Table II. As indicated by the table, the CFO mean of LTE devices predominantly centers around 0 Hz, with only a few devices exhibiting deviations. However, the standard deviation distribution of the CFO across different devices is considerably larger than the distribution of its mean value around 0 Hz. Consequently, the CFO characteristics of LTE devices are unable to effectively contribute to RFF. Therefore, we compensate the residual CFO of the LTE devices to eliminate the influence of residual CFO on RFF extraction.

C. ZC Sequence Detection

The signal detection method based on cross-correlation can effectively identify similar signal sequences within the received signal using local sequences. In actual eNB signal processing, signal cross-correlation can be conducted using the gate circuit implemented by the hardware. However, in our baseband algorithms based on software processing, extensive signal cross-correlation calculations consume significant computational resources. With 64 possibilities for the signals transmitted by PRACH, accurate PRACH signal detection requires cross-correlation based on the 64 generated PRACH local sequences at the receiving end. Leveraging the Fourier

transform property, the convolution operation of the time-domain signal can be rapidly achieved through the point-multiplication operation of the frequency-domain signal. We have designed a fast signal detection algorithm based on the point-multiplication of 64 generated standard ZC sequences in the frequency domain. This method expedites the cross-correlation calculation of 64 PRACH signals, facilitating the completion of ZC sequence detection.

According to the good autocorrelation of the ZC sequence, ZC sequence detection is performed. After removing the CP, the compensated signal $\overline{x}_{\text{fine}}(i)$ and the standard preamble signal $\overline{s}_{\text{idx}}(i)$ are converted via the fast Fourier transform (FFT)

$$\begin{cases} \overline{X}_{\text{fine}}(q) = \mathcal{F}(\overline{x}_{\text{fine}}(i)) \\ \overline{S}_{\text{idx}}(q) = \mathcal{F}(\overline{s}_{\text{idx}}(i)) \end{cases}, \tag{12}$$

where the $\mathcal{F}(\cdot)$ is the FFT operation. In our system, the received PRACH signal is sampled at 16 MSamples/s. Based on this, the length of the ZC signal after removing the CP is 12800 points, hence, indicating that the signal processing involves 12800 FFT points.

The correlation sequences of $\overline{x}_{\text{fine}}(i)$ and $\overline{s}_{\text{idx}}(i)$ can be calculated efficiently in the frequency domain by

$$\rho_{idx}(i) = \mathcal{F}^{-1}(\overline{X}_{\text{fine}}(q) \cdot \overline{S}^{*}_{\text{idx}}(q)), \tag{13}$$

where the $\mathcal{F}^{-1}(\cdot)$ is the inverse fast Fourier transform (IFFT) operation.

Then, the corresponding ZC serial number can be found by

$$I(idx) = \underset{i}{\text{argmax}} |\rho_{idx}(i)|, \tag{14}$$

where $I(idx)$ calculates the value of i when the correlation peak appears between the compensated signal and standard ZC sequences with different idx. The smaller the value of i, the smaller the offset of the two signals in the time domain. The smallest i and the corresponding idx can be given as

$$\widehat{idx} = \underset{idx}{\text{argmin}} \left(I(idx) \right). \tag{15}$$

The standard synchronization PRACH can be given as

$$x_{\text{standard}}(i) = s_{\widehat{idx}}(i). \tag{16}$$

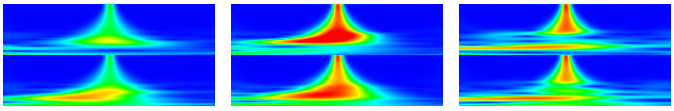
D. Fine Time Synchronization

Finally, we use the standard sequence selected by the previous step to perform cross-correlation with the compensated signal to find the position of the correlation peak and achieve fine time synchronization

$$D_{\text{fine}} = \underset{m}{\text{argmax}} \left(\sum_{i=0}^{L_{\text{CP}}-1} x^{*}_{\text{standard}}(i) \cdot x_{\text{fine}}(i+m) \right). \tag{17}$$

We take D_{fine} as the starting point of the fine synchronized signal and obtain $x(i)$ for the following RFF extraction.

From the transient and steady state waveform of various LTE devices depicted in the Fig. 4 using multi-frame signals, it is evident that we can precisely determine the start and end points of the transient RFF feature for each device through the time synchronization method proposed in this paper, and consistently maintain stability across different experiments.



LTE3 LTE4 LTE7

Fig. 7. WT coefficient graphs of the LTE PRACH transient part.

V. RFF Feature Extraction

A. Wavelet Transform for Transient Part

WT has superior time-frequency localization capability compared with the Fourier transform, which is a more suitable method for transient signal analysis [39]. It has a very short base function duration and supports the simultaneous location of time and frequency information. The WT is defined as

$$\text{WT}(\alpha, \gamma) = \int_{-\infty}^{\infty} s(t) \cdot w\Big(\frac{t-\gamma}{\alpha}\Big) \mathrm{d}t, \tag{18}$$

where $s(t)$ is the expression of signal in the time domain, $w(t)$ stands for the wavelet in the time domain, α is the scale of the wavelet and γ decides the translation of the wavelet. By stretching or compressing the wavelet, the length and frequency of the wavelet can be changed, and so does the corresponding time window length. At the high frequency, the wavelet is compressed and the time window is narrowed, which makes the time resolution higher. These changes at low frequencies are reversed. Hence, WT can express the frequency components of the signal and provide its specific position in the time domain.

We propose a method to represent multi-scale WT coefficients in a 2-D graph. The CWT function in the Matlab toolbox is employed to calculate the wavelet coefficient matrix, using Morse wavelets [44]. Discretization is performed with a specified number of sounds per octave of 10. The minimum and maximum scales are determined automatically based on the energy spread of the wavelet in frequency and time [44]. After WT processing, a 48×200 coefficient matrix is generated corresponding to the 200 transient points and 48 WT scales. We further combine the real and imaginary parts to form a new matrix of 96×200. Fig. 7 shows the WT coefficient graphs of some LTE terminals after the aforementioned feature extraction. The upper and lower parts of the images consist of the real and imaginary parts of wavelet coefficients, respectively. Significant differences can be observed among different LTE terminals.

B. D-FFT Spectra for Steady Part

1) D-FFT Processing: As each UE will randomly select one of the 64 preambles, there will be significant differences between their time domain waveforms, as shown in Fig. 2. Therefore, the features of the steady-state part in RFF training and testing may change when time-domain signals are used.

In this paper, we propose to use D-FFT to mitigate the variation of the PRACH preambles. Intuitively, though the signal waveform can change significantly in the time domain, the responses of adjacent frequencies remain similar, which will be shown later. Moreover, as the typical multipath channels are band-limited, the channel frequency responses are also similar

in adjacent frequencies [45]. Therefore, using a differential operation to calculate the difference of adjacent frequencies of the preamble can effectively mitigate the influence of signal variation on RFF features. Danev and Capkun initially adopted such operations to polish the frequency domain RFF features of CC2420 wireless sensor nodes [46].

The signal $x(i)$ is first converted to the frequency domain by FFT

$$X(q) = \mathcal{F}(x(i)). \tag{19}$$

Then, the D-FFT feature extraction is carried out by

$$\text{D-FFT}(q) = |X(q+1) - X(q)|, \tag{20}$$

where $|\cdot|$ is the absolute value of a complex number. D-FFT(q) actually provides the difference of the adjacent spectra of signal $x(i)$.

2) Effect of D-FFT Based PRACH Variation Mitigation: In this part, we will evaluate the preamble similarity via the cross-correlation of two sequences. A large peak value of the cross-correlation indicates a high similarity of the sequences. The correlation property of the PRACH preamble sequences will be derived and will demonstrate that the D-FFT operation is able to mitigate the variation of the PRACH signal and therefore reduce the influence on RFF extraction and recognition.

a) Correlation Property of the Original PRACH Signal: According to (4) and (5), the time domain PRACH preamble signal $s_{idx}(i)$ can be simplified as

$$s_{idx}(i) = \beta_{\text{PRACH}} \sum_{k=0}^{N_{\text{ZC}}-1} P_{u,v}(k) \cdot e^{j\theta_k(i-i_{\text{CP}})}, \tag{21}$$

where $\theta_k = 2\pi(k+\varphi+K(k_0+\frac{1}{2}))\Delta f_{\text{RA}}$, $P_{u,v}(k)$ is the result of the DFT of cyclically shifted ZC sequence $p_{u,v}(n)$.

The cross-correlation function of the two sequences can be expressed as

$$r_{s<a,b>}(m) = \frac{1}{N} \sum_{i=1}^{N} s_a(i) \cdot s_b^*(m+i), \tag{22}$$

where m represents the offset between $s_b(i)$ and $s_a(i)$, ranging from $-N$ to N where N is the number of samples.

Without loss of generality, we normalize the correlation function by dividing the correlation coefficient at offset 0

$$R_{s<a,b>}(m) = \frac{r_{s<a,b>}(m)}{r_{s<a,b>}(0)}. \tag{23}$$

The peak of the cross-correlation coefficients is

$$\text{Peak}_{s<a,b>} = \max\left(R_{s<a,b>}(m)\right). \tag{24}$$

Therefore, the mean value of $C_{64}^2 = 2016$ correlation peaks between 64 signal sequences can be inferred as

$$\text{Peak}_{\text{signal}} = \frac{1}{C_{64}^2} \sum_{a=1}^{64} \sum_{b=a+1}^{64} \text{Peak}_{s<a,b>}. \tag{25}$$

b) Correlation Property of the Differential Spectrum: According to (21), the N-point FFT of PRACH is given by

$$\begin{aligned} F_{idx}(q) &= \beta_{\text{PRACH}} \sum_{i=0}^{N} s_{idx}(i) e^{j\frac{2\pi iq}{N}} \\ &= \beta_{\text{PRACH}} \sum_{i=0}^{N} \sum_{k=0}^{N_{\text{ZC}}-1} P_{u,v}(k) e^{j\theta_k(i-i_{\text{CP}})} e^{j\frac{2\pi iq}{N}}, \end{aligned} \tag{26}$$

where $F_{idx}(q)$ is the frequency response of the idx-th PRACH signal at the qth subcarrier. Accordingly, the differences between adjacent FFT spectra can be inferred as

$$\begin{aligned} F_{idx}^{'}(q) &= F_{idx}(q+1) - F_{idx}(q) \\ &= \beta_{\text{PRACH}} \sum_{i=0}^{N-1} \sum_{k=0}^{N_{\text{ZC}}-1} P_{u,v}(k) e^{j\theta_k(i-i_{\text{CP}})} e^{j\frac{2\pi i(q+1)}{N}} \\ &- \beta_{\text{PRACH}} \sum_{i=0}^{N-1} \sum_{k=0}^{N_{\text{ZC}}-1} P_{u,v}(k) e^{j\theta_k(i-i_{\text{CP}})} e^{j\frac{2\pi iq}{N}}. \end{aligned} \tag{27}$$

Then, the cross-correlation function of two sequences can be calculated as

$$r_{f<a,b>}(m) = \frac{1}{N-1} \sum_{q=0}^{N-1} F_a^{'}(q) \cdot F_b^{'*}(m+q), \tag{28}$$

where m represents the relative offset of the two sequences, ranging from $-(N-1)$ to $+(N-1)$. Similarly, the normalization is performed as follows

$$R_{f<a,b>}(m) = \frac{r_{f<a,b>}(m)}{r_{f<a,b>}(0)}. \tag{29}$$

Peak cross-correlation between two differential spectral sequences can be obtained

$$\text{Peak}_{f<a,b>} = \max\{R_{f<a,b>}(m)\}. \tag{30}$$

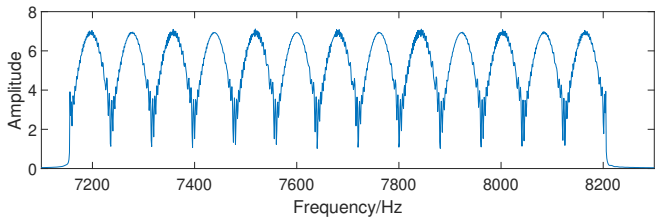
The mean values of $C_{64}^2 = 2016$ correlation peaks between 64 sequences can be expressed as

$$\text{Peak}_{\text{D-FFT}} = \frac{1}{C_{64}^2} \sum_{a=1}^{64} \sum_{b=a+1}^{64} \text{Peak}_{f<a,b>}. \tag{31}$$

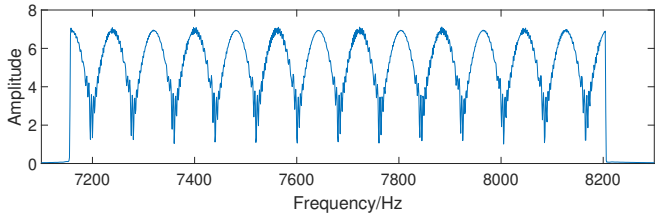
c) Improved Correlation Property by D-FFT Processing: At the sampling rate of 16 MSamples/s, it can be calculated that $\text{Peak}_{\text{signal}}$ is 0.7747 and $\text{Peak}_{\text{D-FFT}}$ is 0.9130. The overall correlation of 64 sequences is highly improved after D-FFT processing. Fig. 8 shows the D-FFT of two signals with different preamble indexes. Compared with the time domain signals shown in Fig. 2, the cross-correlation degree of the two sequences with a large difference in the time domain has been greatly improved after the aforementioned processing.

VI. Hybrid Autoencoder Design

In the actual environment, the eNB needs to determine whether it is a legitimate device or an unknown illegal device based on the PRACH signal transmitted from UE. We are also concerned about the identification of anonymous terminals. Identifying rogue/unknown terminals should be carried out without any information about the unknown terminal. Therefore, we employ AE to address this problem.



(a)



(b)

Fig. 8. The D-FFT of two LTE PRACH signals from the same terminal. (a) Preamble index = 0. (b) Preamble index = 34.

A. The Design of Autoencoder

AE is an unsupervised neural network model designed to learn how to represent the input information by taking it as the learning object [47]. The input data is regarded as the supervision to guide the neural network to learn a mapping relationship, so as to obtain a reconstructed output. In our work, the stacked autoencoder (SAE) is used, which consists of several autoencoders stacked in series. In the detection scenario, when the difference between the output reconstructed by the network and the original input exceeds a certain threshold, it is considered an abnormal one.

After the aforementioned signal processing and RFF feature extraction, the ith input feature of AE is denoted as G_i. It is worth noting that G_i could be multi-dimensional due to the format of the input. For example, if the input is a segment, G_i is 1D. If the input is a graph, G_i is 2D. In order to make the output of the subsequent decoder between 0 and 1, and eliminate differences between data categories, min-max normalization should be carried out first

$$G_i^{'} = \frac{G_i - \min(G_i)}{\max(G_i) - \min(G_i)}, G_i^{'} \in [0,1]^{l \times d}, \tag{32}$$

where l is the length of the extracted feature vector, and d is the dimension of one feature vector. When the feature is D-FFT, l equals the length of D-FFT sequence and d equals 1. When the feature is a WT graph, l equals the length of the signal and d equals the amount of WT scales.

Assume the function learned by the encoder is $\phi_{w,z}(\cdot)$, where w and z are the network parameters. The input after encoding can be expressed as

$$T_i = \phi_{w,z}(G_i'), \quad T_i \in R^{l^{'} \times 1}. \tag{33}$$

where $l^{'}$ represents the length of the encoder output, which is fixed at 20 in our system.

Similarly, the reconstructed output after decoding can be described as

$$Y_i = \varphi_{w,z}(T_i), \quad Y_i \in R^{l \times d}, \tag{34}$$

where $\varphi_{w,z}(\cdot)$ represents the operations at decoding layer.

Then, the mean square error (MSE) is applied to evaluate the reconstruction error

$$\text{MSE}(G^{'}, Y) = \frac{1}{N_s} \sum_{i=1}^{N_s} \left| Y_i - G_i^{'} \right|^2, \tag{35}$$

where N_s stands for the number of training samples.

B. Autoencoders for WT-Based 2D Input

The wavelet coefficient map that we extract is a 2D graph. Therefore, we design an AE for 2D input. As shown in Fig. 9, the $96 \times 200 \times 3$ graph is used as input. Three convolution layers, two max-pooling layers, and four fully connected layers are added to the encoding layer. The feature map size of each layer is as shown in the figure, and the kernel size is chosen as $[6 \times 6]$, $[5 \times 5]$, $[4 \times 4]$ for convolution, and $[2 \times 2]$ for max-pooling. Then, a completely symmetrical network structure forms the decoder, which is composed of four fully connected layers, two up-pooling layers, and three deconvolution layers. The parameters of the encoder layers are as follows

- The first convolutional layer: The number of trainable parameters is $(6^2) \times 3 \times 6 = 648$. After a $[2 \times 2]$ max-pooling layer, the graph size is $[45 \times 97]$.
- The second convolutional layer: The number of trainable parameters is $(5^2) \times 6 \times 16 = 2,400$. After a $[2 \times 2]$ max-pooling layer, the graph size is $[20 \times 46]$.
- The third convolutional layer: The number of trainable parameters is $(4^2) \times 16 \times 32 = 8192$. After shape change, the data size is $[23392 \times 1]$.
- The fourth fully connected layer: The number of trainable parameters is $1000 \times 23392 + 1000 = 23,393,000$.
- The fifth fully connected layer: The number of trainable parameters is $120 \times 1000 + 120 = 120,120$.
- The sixth fully connected layer: The number of trainable parameters is $84 \times 120 + 84 = 10,164$.
- The seventh fully connected layer: The number of trainable parameters is $20 \times 84 + 20 = 1,700$.

The number of encoder parameters of the AE network for 2D input equals 23,536,224, nearly half of the amount of the whole network due to symmetry. In this paper, we used a GeForce RTX 2080 graphics card to train the network. In the experiment of selecting 100 frames of signals for training with 100 iterations, the network took 18.292 seconds. In the experiment of testing the signal with 1,687 frames, the time consumption was 1.885 seconds.

In RFF identification, the AE trained by legal devices will demonstrate the learned mapping relationships to every input device (including the illegal devices), thereby exhibiting the characteristics of a legal device in the reconstruction result. Due to environmental variations, deviations may occur among legal devices. Our goal is to minimize the MSE of legal devices after AE reconstruction, while maximizing the MSE of illegal devices after AE reconstruction. A ratio R_{MSE} is defined to evaluate the discrimination effect for AE as follows:

$$R_{\text{MSE}} = \frac{\text{MSE}_{\text{legal}}}{\text{MSE}_{\text{illegal}}}, \tag{36}$$

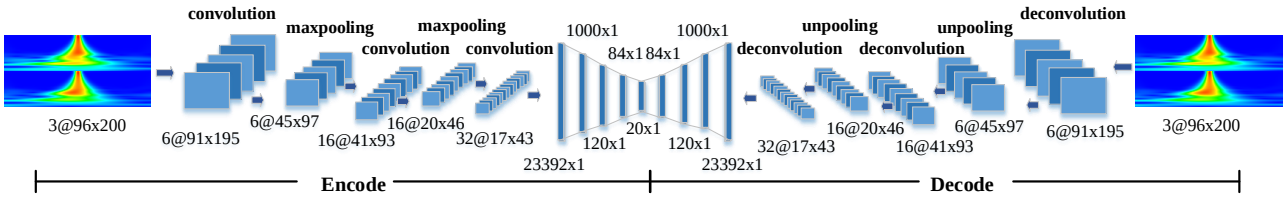


Fig. 9. The structure of the autoencoder for features extracted from the WT coefficient graph.

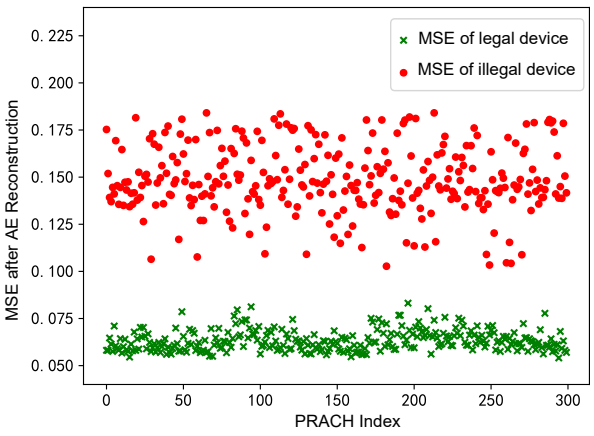


Fig. 10. An illustration of MSE after AE reconstruction.

where $\text{MSE}_{\text{legal}}$ and $\text{MSE}_{\text{illegal}}$ represent the average MSE of legal and illegal devices, respectively. A lower R_{MSE} indicates a larger reconstructed MSE value of illegal devices after the AE reconstruction, making it easier to detect the illegal device. We extract wavelet transform features from the power-on transient, steady state, and power-off transient parts of signals respectively, and calculated R_{MSE} for each part. The ratio of the power-on transient is 0.6935, that of the power-off transient is 0.5896, and that for the steady state is 0.9811. Fig. 10 illustrates the distribution of MSE between legal and illegal devices after AE reconstruction using power-off transient feature. As shown in the figure, it is evident that legal and illegal devices can be easily distinguished based on the MSE after AE reconstruction.

C. Autoencoders for D-FFT Based 1D Input

The adjacent spectrum we extracted is a sequence of 1D. As shown in Fig. 11, 1200×1 data is used as input, and three fully-connected layers are used to form the encoder and decoder separately. The parameters of the encoder layers are as follows

- The first fully connected layer: The number of trainable parameters is $120 \times 1200 + 120 = 144{,}120$.
- The second fully connected layer: The number of trainable parameters is $84 \times 120 + 84 = 10{,}164$.
- The third fully connected layer: The number of trainable parameters is $20 \times 84 + 20 = 1{,}700$.

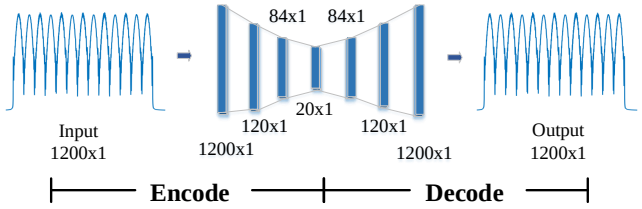


Fig. 11. The structure of the autoencoder for features extracted from D-FFT sequence.

Similarly, the number of encoder parameters of AE network for 1D input is 155,984, nearly half of the amount of the whole network. In the experiment of selecting 100 frames of signals for training with 100 iterations, the network took 5.462 seconds. In the experiment of testing the signal with 1,687 frames, the time consumption was 0.780 seconds.

D. Integration of Two MSE Outputs

We extract the WT-based 2D RFF feature and D-FFT based 1D RFF feature and design AE respectively for unknown device identification. The fusion authentication of the transmitter with multiple RFFs is a prospective direction for RF fingerprinting [48]. Based on the extracted two-dimensional graph features from WT coefficients and one-dimensional sequence features from D-FFT, we designed a fusion autoencoder that combines different dimensions of features. The fusion is achieved by utilizing the MSE of the previously designed autoencoder's output as a new autoencoder input, which also includes the fusion of transient and steady-state RFF features. The structure of the autoencoder for feature fusion from different dimensions is illustrated in Fig. 12. MSE1 and MSE2 are the output MSE of WT feature and D-FFT feature, MSE3 is the output MSE of the above two MSE after passing through the AE network.

We use the data from each authorized terminal to train the network model. Then, we test the data set containing the data from all the terminals. When the combined MSE exceeds a certain threshold, it is considered as a rogue terminal. Otherwise, it is recognized as an authorized one.

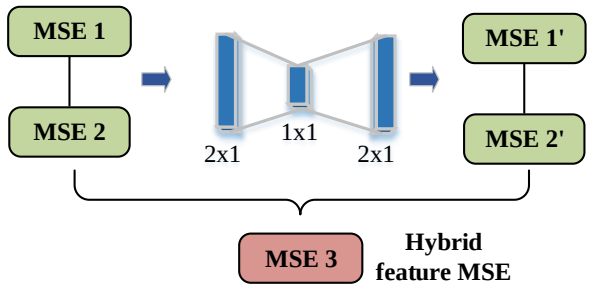


Fig. 12. The structure of the fusion autoencoder combing WT and D-FFT features.

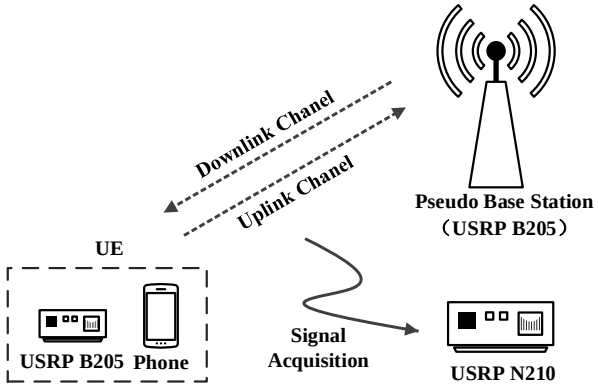


Fig. 13. Experimental testbed.

VII. Experimental Evaluation

A. Experimental Setups

1) Devices: We built an experimental testbed as shown in Fig. 13, including UEs, a pseudo base station and a receiver. The reference clock of the pseudo base station and the receiver is connected via a 10 MHz signal generation source.

We adopted a USRP B205 SDR platform to build a pseudo-base station with the open-source SDR LTE software suite [49]. In order to prevent experimental LTE device from being interfered with by commercial eNB or other LTE signals collected by the receiver, the uplink frequency selected for the experiment is 2565 MHz. This frequency falls within a free frequency band that is not used by operators, as determined by detecting LTE PBCH signals in the environment at the experimental site of the Purple Mountain Laboratory, Nanjing, China. The RB number of the LTE communication system is set to 50, and the PRACH frequency offset is set to 2. In accordance with the PRACH time-frequency structure introduced in Section II, the PRACH occupies 6 consecutive RBs, with each RB having a bandwidth of 180 KHz, the center frequency f_c transmitted by the PRACH can be calculated as follows:

$$f_c = 2565\text{MHz} - (\frac{50}{2} - \frac{6}{2} - 2) \times 180\text{KHz} = 2561.4\text{MHz}. \tag{37}$$

RSI is set to 0, and the cyclic shift interval index is set to 1, corresponding to the CP interval of 13.

We used 8 UE for evaluation, including 7 LTE mobile phones from 3 brands and 1 USRP B205 SDR platform. Since USRP devices can be set up as LTE terminals to access the pseudo-base station, we consider that USRP devices can be potential attackers in LTE networks and therefore also capture the USRP signals in our study. The parameters of each terminal information are shown in Table I. We studied both individual devices and category identification.

- Individual Identifications: One LTE terminal is selected as the authorized terminal (positive class) while the remaining seven terminals as rogue terminals (negative class).
- Category Identifications: The LTE terminals are first categorized into different brands. One brand of the LTE terminals is selected as the authorized class (positive class) while the remaining brands as rogue ones (negative class).

We employed a USRP N210 platform for signal acquisition. The captured signals were transferred to a server for further processing. The server configuration for training the AE network in our experiment is as follows: a GeForce RTX 2080 graphics card, four Intel(R) Xeon(R) E5-2678 V3 @ 2.50 GHz CPU with 12 cores. The version of CUDA is 11.0.

2) Signal Collection: The signal collection was carried out in an office environment, as shown in Fig. 14. The variation of the environment will affect the stability of RFF. We considered the effects of location, acquisition date, and working time, and constructed the following three groups of datasets:

- Acquisition Date, D1, D2, and D3: For each LTE terminal, we collected PRACH frames at one fixed location on three different dates when the terminal was just switched on, forming datasets D1, D2, and D3, each with 1677, 1834, and 1340 frames in total, respectively. Please note the dates for different devices in the same dataset might be different.
- Acquisition Location, L1, L2, and L3: For each LTE terminal, we collected PRACH frames at three different locations on the same day, forming datasets L1, L2, and L3, each with 1565, 1488, and 1460 frames, respectively. There were eight locations, including six LOS locations and two NLOS locations. The terminal was randomly put in one location.
- Working Time, T1, T2, and T3: For each terminal, we collected PRACH frames at the same location and after the device has been turned on or worked for 6 hours, 12 hours, and 24 hours, forming datasets T1, T2, and T3, each with 1615, 1490, and 1687 frames, respectively.

A total of 14,156 PRACH segments were collected, including 4513 frames at different locations, 4851 frames at different dates, and 4792 frames at different working hours. There were 1300 to 2500 frames collected from each terminal. Regarding LTE8, i.e., USRP B205, the signals were sampled from multiple positions, which may be different from the ones marked in Fig. 14. Table III summarizes the detailed collection information of the data sets.

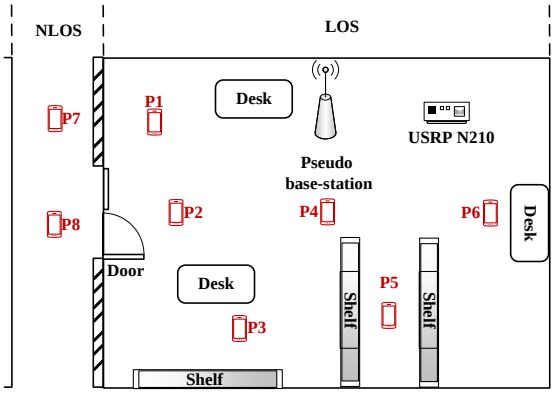


Fig. 14. The layout and location distribution of experimental environment.

TABLE III
DATASET INFORMATION UNDER DIFFERENT ENVIRONMENTAL FACTORS.

Brand	Device	Date	Location	Hours worked
Brand1	LTE1	9-Jul-21, D1	P8, L1	6H, T1
		19-Jul-21, D2	P5, L2	12H, T2
		31-Aug-21, D3	P3, L3	24H, T3
Brand1	LTE2	6-Jul-21, D1	P1, L1	6H, T1
		16-Jul-21, D2	P3, L2	12H, T2
		1-Sep-21, D3	P5, L3	24H, T3
Brand1	LTE3	6-Jul-21, D1	P6, L1	6H, T1
		16-Jul-21, D2	P5, L2	12H, T2
		30-Aug-21, D3	P7, L3	24H, T3
Brand2	LTE4	8-Jul-21, D1	P6, L1	6H, T1
		18-Jul-21, D2	P5, L2	12H, T2
		1-Sep-21, D3	P3, L3	24H, T3
Brand3	LTE5	4-Jul-21, D1	P3, L1	6H, T1
		14-Jul-21, D2	P1, L2	12H, T2
		23-Jul-21, D3	P5, L3	24H, T3
Brand3	LTE6	2-Jul-21, D1	P6, L1	6H, T1
		14-Jul-21, D2	P4, L2	12H, T2
		22-Jul-21, D3	P2, L3	24H, T3
Brand3	LTE7	2-Jul-21, D1	P5, L1	6H, T1
		10-Jul-21, D2	P1, L2	12H, T2
		22-Jul-21, D3	P3, L3	24H, T3

3) Scenario: We set two experimental scenarios to verify the robustness of the proposed method against the environment variables (location, date, or working hours).

- Scenario 1: The training and testing share the same dataset. For example, both training and test datasets are D1.
- Scenario 2: The training and testing have different datasets. We use D1 (L1/T1) as the training set and the test datasets consist of D1-D3 (L1-L3/T1-T3), when studying the effect of acquisition dates (location/working time).

4) Benchmark Approaches: The hybrid method we proposed is referred to as W-D. We also studied the following three benchmark approaches for comparison.

- *CWT* uses a single WT feature. The AE structure for 2D input designed in Section VI-B is adopted.
- *HWT* [40] extracts the wavelet coefficients at 4 scales, draws the histogram of the four groups of coefficients and calculates the number of wavelet coefficients within each numerical interval as the input feature. The AE structure for 1D input designed in Section VI-C is used.
- *D-FFT* [16] employs the adjacent spectral differences of frequency domain steady-state signal as the input feature. The AE structure for 1D input designed in Section VI-C is used.

B. Metrics

Receiver operating characteristic (ROC) describes the true positive rate (TPR) and the false positive rate (FPR) with different thresholds.

We can also select the threshold value when TPR and FPR have the maximum difference. The TPR with this threshold is defined as the identification accuracy of the current LTE terminal/category:

$$\begin{aligned} Threshold &= \underset{K_{Th}}{\operatorname{argmax}}(P_{\text{TPR}} - P_{\text{FPR}}), \\ Accuracy &= \underset{P_{\text{TPR}}}{\operatorname{argmax}}(P_{\text{TPR}} - P_{\text{FPR}}), \end{aligned} \tag{38}$$

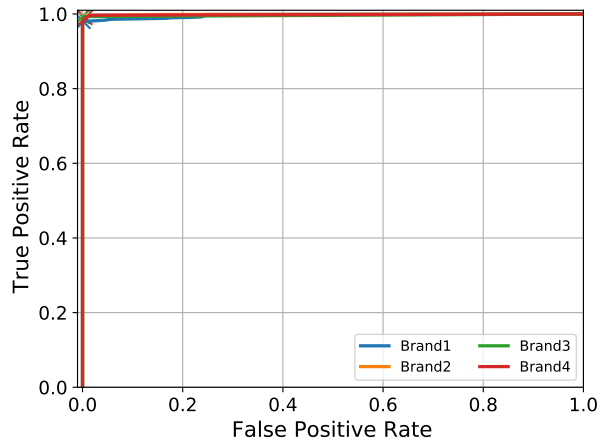
where K_{Th} is a list of potential thresholds, while P_{TPR} and P_{FPR} are the lists calculated from the corresponding values of K_{Th}.

C. Results and Analysis

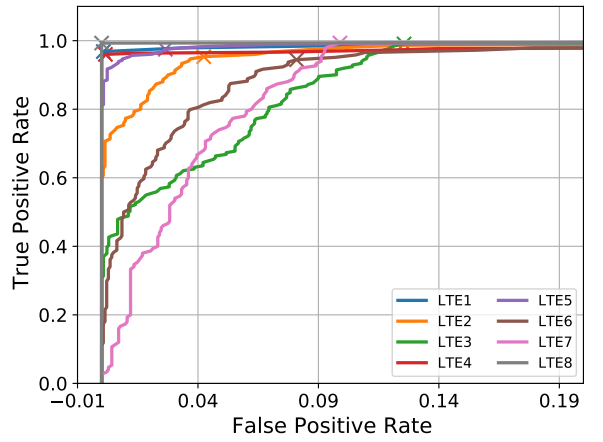
1) ROC Results: Fig. 15 illustrates the ROC curve of W-D obtained from the dataset of L1 in scenario 2, the $\times$ markers at each curve are the accuracy we defined in (38).

Fig. 15(a) shows the results of category identification. The curve of each category is close to a right angle, which reveals that each category has a good degree of discrimination.

Fig. 15(b) shows the results of individual identification. LTE2 and LTE3, LTE5, LTE6, and LTE7 terminals may have a certain degree of confusion respectively. According to Table I, LTE2 and LTE3 are both manufactured by Google with the same model, and LTE6 and LTE7 are both manufactured by Huawei with the same model. Terminals of the same model have the same RF design, which caused high similarity in the RFF features. Besides, the serial sequence of LTE6 and LTE7 are closer than LTE5 and the similar production batches reduced differences between analog circuits greatly. While compared to LTE5 with the same model, the serial sequence may reflect that LTE5 may not be produced in the same batch as the LTE6 and LTE7. This phenomenon is also observed in LTE1. Therefore, despite being the same model of LTE terminals, there will be certain variations in the RFF of different product batches. It is also worth noting that LTE8 performs well both in the category and individual identification, which is due to the inherent obvious differences between LTE mobile phones and USRP platforms.

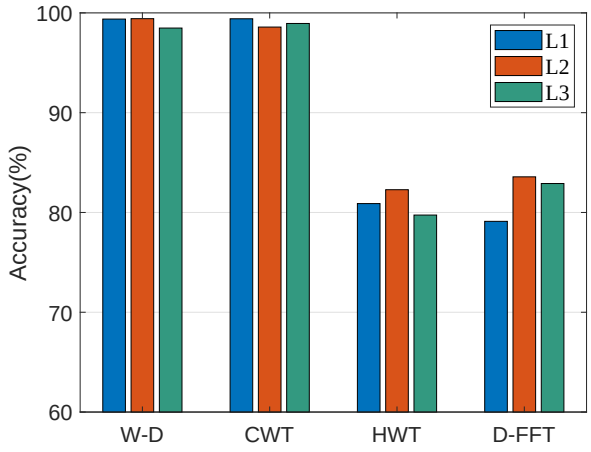


(a)

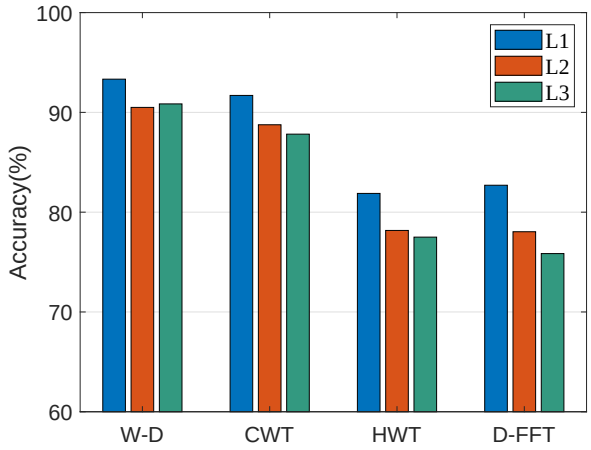


(b)

Fig. 15. ROC curve in Scenario II of L1. (a) ROC for the category identification. (b) ROC for individual identification.



(a)



(b)

Fig. 16. Accuracy in Scenario 1. (a) Category Identification. (b) Individual Identification.

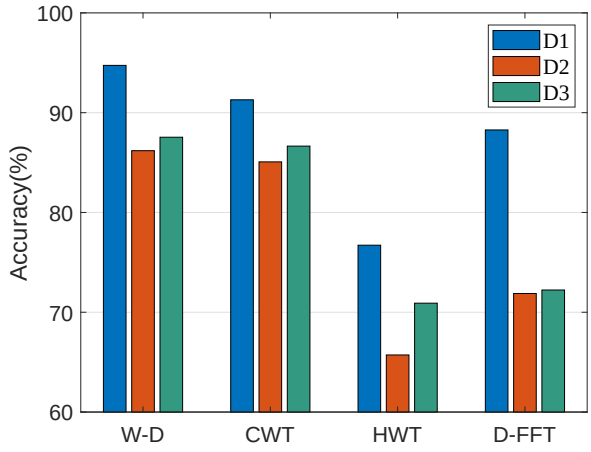
TABLE IV
THE AVERAGE IDENTIFICATION ACCURACY.

Method	Scenario 1		Scenario 2	
	Category	Individual	Category	Individual
W-D	98.80%	90.47%	98.84%	90.27%
CWT	98.41%	88.17%	98.40%	87.32%
HWT	80.43%	79.17%	82.73%	75.85%
D-FFT	79.92%	80.31%	77.99%	78.51%

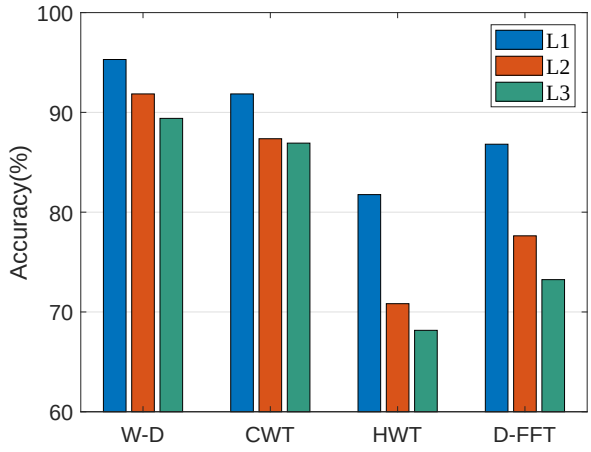
2) Identification Accuracy: Table IV shows the identification accuracy of the four methods to identify the category and individual for LTE terminals in both scenarios, which is the average result against the above three environmental variables. The identification accuracy of W-D and CWT can achieve over 87% for individual identification and over 98% for category identification, which is much higher than HWT and D-FFT. In both scenarios, the W-D method improves the accuracy both in the category and individual identification compared to CWT.

Fig. 16 exemplifies the identification results of the four methods at different locations in Scenario 1. It can be seen that the accuracies of the two methods, W-D and CWT, are significantly higher than the other two methods. Moreover, the training results under different sets of data are also relatively stable. In experiments at L1, L2, and L3, the fluctuation of recognition accuracy of W-D and CWT is less than 3%, while reaches 7% with D-FFT. The results for acquisition date and working time are similar trends, which are not included here for simplification.

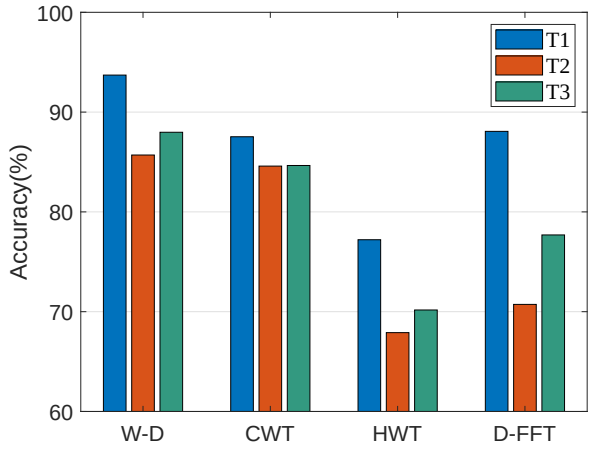
Fig. 17 shows the identification results of the four methods under three variables for individuals in Scenario 2. Taking Fig. 17(a) as an example. As the training dataset is D1, the accuracies of D2 and D3 test datasets are reduced compared to D1 test set. HWT and D-FFT have severe performance degradation. Similar conclusions can be drawn from the other two cases in Fig. 17(b) and Fig. 17(c). In contrast, W-D kept a relatively good performance over different environmental factors, which is desirable. In practice, the test date, location, and working hours will indeed be different from the training settings. This means that the proposed method is more robust to the changes in environmental factors, showing advantages over traditional methods in practical application scenarios.



(a)



(b)



(c)

Fig. 17. Accuracy in Scenario 2 for individual identification. (a) Different dates. (b) Different locations. (c) Different working hours.

VIII. Conclusions

In this paper, a hybrid AE-based LTE RFF identification scheme is proposed to identify rogue LTE terminals. The 2D wavelet coefficients graph from the transient-off part of the LTE PRACH signal and 1D differential spectrum from the steady part of the LTE PRACH signal are extracted as RFF features. These two RFF features can be combined to improve the robustness of the AE identification system effectively under different environmental factors. Besides, datasets containing different times, locations, and working hours are established to verify the stability of RFF features. The identification accuracy is up to 98.84% for different categories and 90.27% for different individuals, which significantly outperforms existing methods.

References

[1] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device fingerprinting in wireless networks: Challenges and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 94–104, First Quarter 2016.

[2] L. Xie, L. Peng, J. Zhang, and A. Hu, "Radio frequency fingerprint identification for internet of things: A survey," *Security and Safety*, vol. 3, 2024.

[3] C. Cheng, R. Lu, A. Petzoldt, and T. Takagi, "Securing the Internet of Things in a quantum world," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 116–120, Feb. 2017.

[4] Global Mobile Suppliers Association (GSA), "Ericsson mobility report," November 2021. [Online]. Available: https://gsacom.com/reports/.

[5] "3GPP TS 36.211 evolved universal terrestrial radio access (e-utra); physical channels and modulation, version 14.2.0 release 14." [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/

[6] "3GPP TS 38.211 5g; nr; physical channels and modulation, version 16.7.0 release 16." [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/

[7] "3GPP TS 33.401 3gpp system architecture evolution (sae); security architecture, version 17.4.0 release 17." [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/

[8] "3GPP TS 33.501 security architecture and procedures for 5g system, version 18.2.0 release 18." [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/

[9] T. Fei and W. Wang, "LTE is vulnerable: Implementing identity spoofing and denial-of-service attacks in LTE networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2019.

[10] M. Lichtman, R. P. Jover, M. Labib, R. Rao, V. Marojevic, and J. H. Reed, "LTE/LTE-A jamming, spoofing, and sniffing: Threat assessment and mitigation," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 54–61, 2016.

[11] T. Zhaowei, D. Boyan, Z. Jinghao, G. Yunqi, and L. Songwu, "Data-plane signaling in cellular IoT: Attacks and defense," in *Proc. 27th Annual International Conference on Mobile Computing and Networking*, Oct. 2021, p. 465–477.

[12] S. Erni, M. Kotuliak, P. Leu, M. Rschlin, and S. Apkun, "Adaptover: Adaptive overshadowing attacks in cellular networks," in *Proc. 28th Annual International Conference on Mobile Computing And Networking*, Oct. 2022, p. 743–755.

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

[14] 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, 2021.

[15] 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.

[16] 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, no. 6, pp. 1180–1192, 2015.

[17] A. Gritsenko, Z. Wang, T. Jian, J. Dy, K. Chowdhury, and S. Ioannidis, "Finding a 'new' needle in the haystack: Unseen radio detection in large populations using deep learning," in *Proc. IEEE Int. Symposium Dynamic Spectr. Access Netw. (DySPAN)*, Newark, NJ, USA, 2019, pp. 1–10.

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

[19] 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. 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec'17)*, 2017, p. 58–63.

[20] A. Al-Shawabka, P. Pietraski, S. B Pattar, F. Restuccia, and T. Melodia, "DeepLoRa: Fingerprinting LoRa devices at scale through deep learning and data augmentation," in *Proc. 32th International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, Jul. 2021, p. 251–260.

[21] 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.

[22] 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, 2021.

[23] 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, 2018.

[24] 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 of Things Journal*, vol. 6, no. 1, pp. 349–360, 2018.

[25] 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, 2020.

[26] J. Shi, L. Peng, H. Fu, and A. Hu, "Robust RF fingerprint extraction based on cyclic shift characteristic," *IEEE Internet of Things Journal*, pp. 1–1 Early Access, 2023.

[27] M. Liu, X. Han, N. Liu, and L. Peng, "Bidirectional IoT device identification based on radio frequency fingerprint reciprocity," in *Proc. IEEE International Conference on Communications (ICC)*, 2021, pp. 1–6.

[28] Y. Zhang, Y. Peng, J. Sun, G. Gui, Y. Lin, and S. Mao, "GPU-free specific emitter identification using signal feature embedded broad learning," *IEEE Internet Things J.*, vol. 10, no. 14, pp. 13 028–13 039, 2023.

[29] X. Fu, Y. Peng, Y. Liu, Y. Lin, G. Gui, H. Gacanin, and F. Adachi, "Semi-supervised specific emitter identification method using metric-adversarial training," *IEEE Internet Things J.*, vol. 10, no. 12, pp. 10 778–10 789, 2023.

[30] F. Demers and M. St-Hilaire, "Radiometric identification of lte transmitters," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2013.

[31] Z. Zhuang, X. Ji, T. Zhang, J. Zhang, and Y. Liu, "Fbsleuth: Fake base station forensics via radio frequency fingerprinting," in *Proc. Asia Conference on Computer and Communications Security*, 2018, p. 261–272.

[32] A. Ali and G. Fischer, "Enabling fake base station detection through sample-based higher order noise statistics," in *Proc. 42nd International Conference on Telecommunications and Signal Processing (TSP)*, 2019, pp. 695–700.

[33] M. Lichtman, R. Rao, V. Marojevic, J. Reed, and R. P. Jover, "5G NR jamming, spoofing, and sniffing: Threat assessment and mitigation," in *Proc. IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018, pp. 1–6.

[34] D. Li, X. Yang, A. Hu, F. Zhou, and O. A. Dobre, "LTE device radio frequency fingerprints blind extraction based on temporal-frequency domain PRACH signals," *IEEE Trans. Veh. Technol.*, pp. 1–1 Early Access, 2023.

[35] P. Yin, L. Peng, J. Zhang, M. Liu, H. Fu, and A. Hu, "LTE device identification based on RF fingerprint with multi-channel convolutional neural network," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[36] Y. Qiu, L. Peng, J. Zhang, M. Liu, H. Fu, and A. Hu, "Signal-independent RFF identification for LTE mobile devices via ensemble deep learning," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2022, pp. 37–42.

[37] G. Shen, J. Zhang, A. Marshall, M. Valkama, and J. R. Cavallaro, "Towards length-versatile and noise-robust radio frequency fingerprint identification," *IEEE Trans. Inf. Forensics Security*, 2023.

[38] D. Roy, T. Mukherjee, M. Chatterjee, E. Blasch, and E. Pasiliao, "RFAL: Adversarial learning for RF transmitter identification and classification," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 6, no. 2, pp. 783–801, 2019.

[39] R. Hippenstiel and Y. Payal, "Wavelet based transmitter identification," in *Proc. 4th International Symposium on Signal Processing and Its Applications*, vol. 2, 1996, pp. 740–742.

[40] K. C. Ho, H. Liu, and L. Hong, "On improving the accuracy of a wavelet based identifier to classify CDMA signal and GSM signal," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS'99)*, 1999.

[41] Z. Wu, L. Peng, J. Zhang, M. Liu, H. Fu, and A. Hu, "Authorized and rogue LTE terminal identification using wavelet coefficient graph with auto-encoder," in *Proc. IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, 2022, pp. 1–5.

[42] Y. He, J. Wang, Y. Su, E. Dutkiewicz, X. Huang, and J. Shi, "An efficient implementation of PRACH generator in LTE UE transmitters," in *Proc. 7th International Wireless Communications and Mobile Computing Conference*, 2011, pp. 2226–2230.

[43] H. Althumali and M. Othman, "A survey of random access control techniques for machine-to-machine communications in LTE/LTE-A networks," *IEEE Access*, vol. 6, pp. 74 961–74 983, 2018.

[44] "Continuous wavelet transforms." [Online]. Available: https://uk.mathworks.com/help/wavelet/continuous-wavelet-transforms.html

[45] M. Liu, M. Crussière, and J.-F. Hélard, "A novel data-aided channel estimation with reduced complexity for TDS-OFDM systems," *IEEE Trans. Broadcast.*, vol. 58, no. 2, pp. 247–260, 2012.

[46] B. Danev and S. Capkun, "Transient-based identification of wireless sensor nodes," in *Proc. International Conference on Information Processing in Sensor Networks*, 2009, pp. 25–36.

[47] M. Nowicki and J. Wietrzykowski, "Low-effort place recognition with WiFi fingerprints using deep learning," *arXiv:1611.02049*, 2016.

[48] H. Yuan, Y. Yan, Z. Bao, C. Xu, J. Gu, and J. Wang, "Multipath canceled RF fingerprinting for wireless OFDM devices based on hammerstein system parameter separation," *IEEE Canadian Journal of Electrical and Computer Engineering*, vol. 45, no. 4, pp. 401–408, 2022.

[49] "Open-source LTE software radio suite." [Online]. Available: https://www.srslte.com/