Hindawi
Security and Communication Networks
Volume 2021, Article ID 3530509, 16 pages
https://doi.org/10.1155/2021/3530509

Research Article

A LoRa-Based Lightweight Secure Access Enhancement System

Yu Jiang,[1,2,3] **Hua Fu**,[1,3] **Aiqun Hu**,[1,3] **and Wen Sun**[1]

[1]*School of Cyber Science and Engineering, Southeast University, Nanjing, China*
[2]*Key Laboratory of Computer Network Technology of Jiangsu Province, Nanjing, China*
[3]*Purple Mountain Laboratories for Network and Communication Security, Nanjing, China*

Correspondence should be addressed to Yu Jiang; jiangyu@seu.edu.cn

Received 9 June 2021; Revised 28 July 2021; Accepted 16 August 2021; Published 26 August 2021

Academic Editor: Jingyu Feng

The access control mechanism in LoRa has been proven to have high security risks. In order to improve the secure access ability of LoRa terminals, this paper presents a physical layer-based authentication system for security enhancement. Different from the security access technology of cryptography, a lightweight gateway architecture called LW-LoRaWAN is proposed to realize a data frame-based authentication with radio frequency fingerprint (RFF). A novel RFF feature of Cross Power Spectral Density (CPSD) is used to achieve a fast authentication with one single frame. Theoretical analysis and experimental results show that the proposed system not only reinforces the authentication security of LoRa network but also protects the LoRa terminals against the Sybil attacks. The LW-LoRaWAN provides new security approach from physical layer for LoRa network.

1. Introduction

The Internet of things (IoT) is related to distributed control, dynamic topology, and limited power of terminal nodes. These characteristics of IoT make the security threats different from those of the Internet [1–4]. Due to the openness of wireless communication, the IoT is more vulnerable to eavesdropping, counterfeiting, tampering, and denial-of-service attacks. When nodes are compromised by attackers, they can still access the network with their legitimate identities, making IoT networks exposed to both external and internal threats. Due to the frequent accessing and exiting of nodes and the dynamic changes of network topology, it is impossible to effectively prevent the attack of illegal nodes without a reliable secure access mechanism [5–7].

At present, the network structure of IoT system has not been clearly defined. It is generally considered that the network structure includes the perception layer, the transport layer, and the application layer [8]. There have been mature solutions for security management of the transport layer and application layer, but the research on the identity authentication technology of perception layer is not enough [9]. Common authentication techniques include authentication based on MAC address, security certificates, instructions, and so on. Unfortunately, all of these methods have certain risks in practical scenarios. MAC address-based authentication can easily be cracked by forging the MAC address, and the black and white list strategy is bypassed [10]. Certificate-based authentication is limited by the lack of resources such as memory or computing ability of IoT terminals [11]. Instruction-based authentication suffers from the instruction leakage and weak instructions, which makes the terminal unable to be effectively protected [12].

The security issues of IoT are different from those of the Internet, because complex security policies cannot be deployed with the limited resources of IoT devices. The problem of accessing security for the IoT has been studied for more than a decade, but it remains challenging to find an effective solution satisfying both lightweight and security.

In order to solve the identity authentication problem of the IoT, it is necessary to propose a complete secure access solution based on the limited resources of IoT devices. In fact, the signal emitted by wireless device carries the unique features on the waveform, which can be deployed as the identity of the terminal to distinguish the counterfeiters

[13–15]. Compared with traditional approaches based on MAC address or authentication instructions, the physical layer features of devices, also known as radio frequency fingerprint (RFF), are difficult to forge and can be applied for identity authentication [16–18].

LoRa is a low-power wireless communication technology, which uses Chirp Spread Spectrum (CSS) to serve remote communication [19]. According to the needs of different scenarios, the physical layer of LoRa is highly configurable, including spread factor (SF), coding rate (CR), bandwidth (BW), optional header, and other parameters. LoRaWAN is an open-source protocol of LoRa, which can realize effective communication and networking between LoRa terminals and gateway.

As the basic technology of Low-Power Wide-Area Network (LPWAN) communication, LoRa is expected to be widely used. Meanwhile, due to the use of unauthorized frequency bands and public protocols, LoRa network is vulnerable to attacks. At present, the analysis of access authentication of LoRa terminals is mostly based on LoRaWAN protocol [20, 21], and the research of access authentication based on the RFF of physical layer is still in the early stage. This paper firstly proposes a physical layer-based authentication system based on RFF features to improve the security of access for LoRa terminals. The main contributions of the proposed authentication solution for LoRa are listed below:

(1) A lightweight gateway architecture called LW-LoRaWAN is proposed to realize a data frame-based authentication with RFF

(2) A novel RFF feature of Cross Power Spectral Density (CPSD) is used to achieve a fast authentication within one data frame

(3) LW-LoRaWAN can protect the LoRa terminals against the Sybil attacks

(4) The proposed system for security enhancement only needs to upgrade the gateway; no change is needed for terminal devices, which is more feasible than the existing enhancement schemes

The remainder of this paper is organized as follows: In Section 2, the state of the art of the lightweight security enhancement techniques for IoT is presented. The lightweight secure access scheme for LoRa is proposed in Section 3. The novel RFF extraction method for single data frame is presented in Section 4. Experimental results and system analysis are presented in Section 5. Finally, the conclusion is drawn in Section 6.

2. Background and Related Work

Due to the limited resources of IoT devices, the research of the IoT security focuses on the lightweight access technologies, where the lightweight cryptographic algorithm and the lightweight security protocol are the two main research directions. Meanwhile, the RFF-based authentication has been widely studied, which provides a different approach from modern cryptography.

2.1. Lightweight Cryptographic Algorithm. Lightweight cryptographic algorithm [22] is an innovative algorithm for devices with limited computing resources. In recent years, with the increasing security requirements for IoT, the research on lightweight cryptographic algorithm has achieved good results. The International Organization for Standardization (ISO) has developed some standards for algorithms such as lightweight block ciphers and stream ciphers, most of which are symmetric.

Lightweight cipher mainly includes lightweight block cipher, lightweight Hash function, and lightweight stream cipher. Among the symmetric ciphers, the block cipher algorithm [23] has been studied for a long time and has many achievements. It has typical security structures, such as Feistel and Substitution Permutation Network (SPN). After the PRESENT algorithm was published, many lightweight Hash functions have been designed based on PRESENT algorithm, such as C-PRESENT, H-PRESENT, and DM-PRESENT [24]. In recent years, there has been a new trend to design Hash functions by using Sponge structure [25]. The design of stream ciphers is mainly based on the linear and nonlinear feedback shift registers. A variety of lightweight stream cipher algorithms have been proposed such as Espresso, Lizard, Grain-128a, Welch Gong 8 (WG-8), Sprout, P1antlet, and Fruit [26].

2.2. Lightweight Security Protocol. The perception layer of the IoT cannot meet the requirements of computing, storage, and communication overhead of traditional security protocols, so it is necessary to research and develop lightweight protocols. In general, lightweight security protocols, which mainly include lightweight authentication protocol, lightweight key agreement protocol, and lightweight key management protocol, are designed to reduce the amount of computation, information flow, and number of communication rounds by sacrificing certain reliability and even security [27].

The lightweight authentication protocols are mainly used in resource-constrained system to ensure the legitimacy of the identities. It includes one-way and two-way authentication and can be widely used in point-to-point and multihop communications. After the identity authentication is completed, the lightweight key negotiation protocol establishes a session key for the subsequent communication. It can be widely used in access control of RFID, IoT, and other systems [28]. The lightweight key management protocol is used to create, distribute, and maintain the key in the cryptographic mechanism of resource-constrained system and to realize the key management in the secure communication.

2.3. LoRa RFF Technology. As a new wireless communication technology, LoRa RFF begins to receive the attention of researchers. The radar signal has the same modulation scheme as the LoRa signal and the identification of radar equipment is called Specific Emitter Identification (SEI). The U.S. Naval Research Bureau has conducted research on SEI technology for more than decades [29]. The purpose of SEI

research is to extract weak and robust features from radar signals to achieve individual identification of emission sources. Due to the sensitivity of radar research, the information available for inquiry is very limited. In addition, radar is mainly used in the military field, so its production accuracy is better than the commercial LoRa devices. The identification of LoRa devices should be based on its practical scenarios and device characteristics.

In recent years, the research on the communication protocol of LoRa devices [30–32] and the synchronization of LoRa signals [33, 34] are in progress. The identification of LoRa devices based on RFF has also been carried out. In 2017, Eletreby et al. [35] proposed using the time, frequency, and phase offset of the signal to identify LoRa devices and applied them to the access authentication process of LoRaWAN network. Robyns et al. [36] proposed a supervised machine learning method to recognize LoRa devices. It takes the data after signal preprocessing as the recognition object for machine learning. In 2019, Jiang et al. [37] extracted the RFF features of LoRa device based on the differential constellation trace figure (DCTF).

3. Lightweight Secure Access Scheme for LoRa

The current secure access scheme of LoRaWAN protocol adopts modern cryptography technology. This section proposes a physical layer secure access control scheme based on RFF of LoRa terminals and establishes a lightweight access protection architecture, referred to as LW-LoRaWAN.

3.1. Overall Architecture. The proposed LW-LoRaWAN system architecture includes four parts: LoRa terminals, LoRa gateway, RFF database, and remote server. LoRa terminals and remote server continue to use the equipment in LoRaWAN, while LoRa gateway and RFF database are new equipment in LW-LoRaWAN. The system architecture is shown in Figure 1.

Maintaining the original functions of LoRaWAN, LW-LoRaWAN provides the functions listed in the following:

(1) LoRa modulation and demodulation
(2) LoRaWAN protocol support
(3) Bidirectional communication with remote server
(4) RFF extraction and identification of LoRa terminals
(5) Establishing the relevance between RFF and data frame of LoRa terminals
(6) Real-time illegal data blocking and abnormal terminal alarming

Among the above functions, Functions 1–3 are the original functions of LoRaWAN, while Functions 4–6 are the new functions of LW-LoRaWAN to enhance access protection. Function 4 requires the introduction of a new hardware platform, so the Universal Software Radio Peripheral (USRP) is used to receive RF signals from LoRa terminals and extract the RFF. Since the RFF is data-independent and the number of terminals in the LoRa network

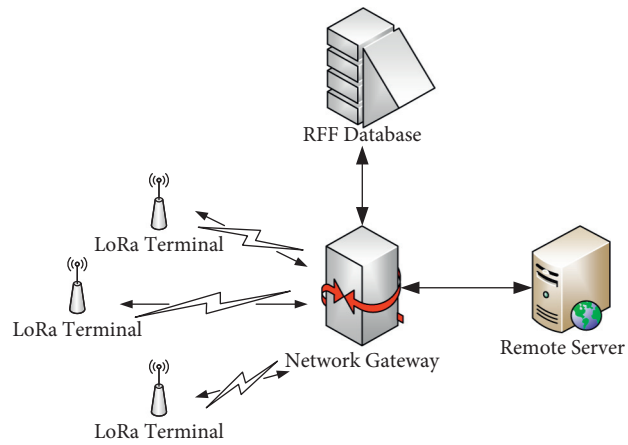

Figure 1: LW-LoRaWAN system architecture.

can be very large, Function 5 is deployed to establish the relevance between the RFF and the data frame of LoRa terminals for the implementation of secure policies. Function 6 is in charge of intercepting the data packets from illegal terminals based on the trained RFF database of legitimate terminals. When two or more terminals are found with the same ID but different RFFs, this function will be triggered to alert the remote server for the presence of counterfeit terminal.

3.2. The Design of New Functions in LW-LoRaWAN. The design of Functions 4–6 is described in detail in the following.

3.2.1. RFF Extraction and Identification of LoRa Terminals. The extraction and identification of LoRa RFF can be divided into three stages: signal acquisition and preprocessing, RFF extraction, and recognition and decision of RFF.

The first stage is acquisition and sampling the received LoRa signal. After the signal is collected by USRP, the signal frame is preprocessed including energy normalization and the coarse and fine synchronization.

The purpose of frame synchronization is to accurately locate and extract the signal frame from the received signal segment. Firstly, the coarse synchronization of the frame is conducted with the double sliding window method and the approximate position of the frame head is obtained. The coarse synchronization can meet the requirements of carrier frequency offset calculation, but more accurate synchronization is needed for RFF extraction. Therefore, the cross-correlation calculation is used to find the position of the maximum correlation peak, which corresponds to the frame starting point for the fine synchronization. According to the format of LoRa frame, the preamble part contains at least 6 identical up-chirp symbols. The cross correlation between the coarse-synchronized signal and the standard up-chirp signal is calculated. When the correlation peak is found, the position of the peak is the offset of the coarse-synchronized signal. Then the offset is compensated for the fine synchronization.

In the second stage, the RFFs of LoRa terminals are extracted and the flowchart is shown in Figure 2.

As shown in Figure 2, the RFF features of LoRa terminals are divided into transient features and steady features. The transient features can be found in the rising and falling edge of the signal frame. The steady features include the IQ offset and the carrier frequency offset [35].

In the third stage, the transient and steady features are extracted from each signal frame. The Euclidean distance of feature vectors between the devices is calculated and registered in the legitimate database. According to the size of the distance deviation, the legal and illegal terminals can be identified.

3.2.2. Establishing the Relevance between RFF and Data Frame. This function is the core idea of the proposed lightweight access scheme. LoRaWAN achieves the authentication of the terminals by presharing key, which means that the access control is based on the device ID. The proposed new function demonstrates a novel idea of using data as the object of access control rather than IDs. The data frame contains both the transmitted data and the RFF features of the terminal. When the ID information is included in the payload of the data frame, the binding relationship of "terminal ID-terminal data-RFF" can be established. Then, the access control of data packets based on the RFF can be realized.

The access security policy of the physical layer can achieve the access control of the terminal alone. From the perspective of the overall architecture, the gateway needs to implement the RFF extraction function, while the terminal requires no modification. There are a large number of terminals in the LoRa network, but the data throughput is limited. Therefore, it is possible to implement the RFF binding with the data frame without adding too much resources burden.

3.2.3. Real-Time Illegal Data Blocking and Abnormal Terminal Alarming. The above functions complete the RFF extraction of each packet and determine the attribution of the data packet only from the physical layer. This function performs real-time data processing, forwarding, or discarding according to the validity of the data packets. In addition, this function replaces the LoRaWAN communication function and it needs to be compatible with the data format, protocol, modulation, and demodulation of LoRaWAN.

Different from the existing security scheme, the greatest advantage of the proposed protection scheme is that it uses the uniqueness of RFF to achieve the identification of counterfeit terminals. When two or more different RFFs are found with the same terminal ID, it means that there exist counterfeit terminals in the network. Hence, this function performs real-time data blocking and abnormal terminal alarming for this terminal ID to prevent malicious data from being uploaded to the remote server. The subsequent data with this terminal ID then is blocked until the terminal returns to normal. Compared with some existing access control strategies based on the terminal behaviors, the proposed function can immediately detect the existence of counterfeit devices and effectively block the malicious data.

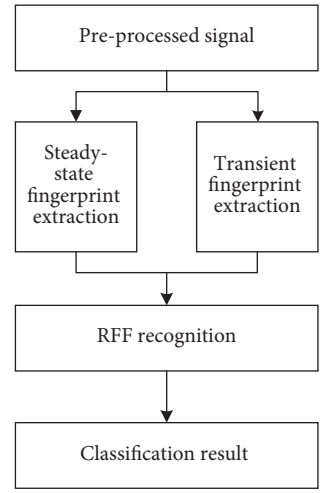

Figure 2: The flowchart of RFF extraction.

3.3. LW-LoRaWAN Workflow. The working procedure of LW-LoRaWAN is simpler than that of LoRaWAN, as shown in the following:

(1) The initialization phase: The RFF and corresponding ID of the legitimate terminals are stored in the RFF database. The RFF database can be an independent device as shown in Figure 1 or integrated with the gateway. When a new legitimate terminal joins the network, its ID and RFF can be added to the RFF database.

(2) Normal working phase: The gateway receives the RF signal and extracts the RFFs from LoRa terminals. Meanwhile, the RF signal is demodulated into the link layer packets with the terminal ID according to the LoRa physical layer protocol. The RFF is bonded to the corresponding ID and the terminal is marked as online.

(3) According to the current packet, the gateway matches the RFFs of the online terminal with that in the RFF database to determine whether the online terminal is legitimate. If the terminal is judged to be illegal, the working process goes to (4). If the terminal is judged to be legal, the gateway forwards the current packet to the remote server.

(4) The gateway blocks the packets of illegal terminals to prevent them from accessing the gateway. The gateway stores the alarm information of the illegal terminals and sends it to the remote server.

(5) For the server side, different servers transmit data to the terminals. The gateway receives the packets from the remote server and extracts the terminal ID.

(6) The gateway queries whether the terminal ID has an alarm message. If not, the terminal ID is judged to be

legal and the working process goes to (7). If the alarm message exists, the terminal is judged to be illegal and the gateway blocks its packets.

(7) The gateway forwards the legitimate packets of the remote server to the terminals through the physical layer protocol.

The workflow of LW-LoRaWAN is shown in Figure 3.

3.4. Comparative Analysis of LW-LoRaWAN. The proposed lightweight security enhancement scheme based on RFF and the modern cryptographic security enhancement scheme are designed to solve the existing LoRaWAN security issues. From a functional point of view, both of these schemes have promoted the secure access protection for LoRa terminals. Then, the implementation and performance of the two schemes are qualitatively compared and the analysis results are shown in Table 1.

As shown in Table 1, the modern cryptography enhancement scheme maintains the technical specifications of the original gateway and the compatibility of the original communication system due to the software upgrade of the original gateway. However, this implementation requires the software upgrade for all the existing terminals and gateways. Due to the small number of gateways in the LoRa network, the workload of gateway upgrade is limited, but, for the massive and widely deployed LoRa terminals, upgrade is almost impossible. Therefore, the enhancement scheme can only be implemented on the newly deployed devices. In addition, the conflict between modern cryptography and the limited resources of LoRa terminals still exists, which greatly reduces the battery life of LoRa terminals.

The proposed LoRa gateway security enhancement scheme replaces the original LoRaWAN gateway with the USRP and its RF performance depends on the USRP specifications. The advantages of the RFF scheme include the uniqueness of preventing counterfeit attacks; there is no need to upgrade a large number of terminals and the designing freedom for customized functions.

4. RFF Extraction Method for Single Data Frame

The primary requirement of the LoRa gateway security enhancement is not to affect the normal functions of the original network architecture. According to the analysis in the previous section, an access control mechanism for the south side of the gateway is introduced, which requires stable and effective extraction of packet information and the corresponding RFF from each data frame. At present, the LoRa RFF features include the frequency offset [35], the overall data [36], and the DCTF [37]. However, these features generally require the accumulation of a certain number of packets for statistical analysis. In order to solve the problem of extraction efficiency, this section proposes an RFF feature extraction method based on the Cross Power Spectral Density (CPSD) of LoRa signals, which can extract stable and unique RFF information from a single frame.

4.1. LoRa Signal Analysis. Though the RFF features are data-independent, they are generally weak relative to the modulation waveform of the signal. There are multiple identical preambles in RF communication signals, so it is easy to perceive the weak RFF information from the known preambles. Therefore, most of the existing algorithms extract the RFF features from the preambles. However, the preambles only occupy a small proportion of the entire data frame. When the payload behind the preamble is used, more raw data can be utilized for RFF extraction.

Compared with traditional CSS technology, LoRa modulation further improves the deployment of spectrum. LoRa modulation is essentially a circular shift of the standard chirp symbol to obtain the modulated signal and the information transmitted by each symbol is determined by the initial frequency offset. Therefore, all the LoRa symbols can be obtained theoretically by cyclic shift of any symbol, which means that all the symbols in a single data frame can be shifted into the same waveforms for RFF extraction.

The chirp signal is composed of sinusoidal signals and the frequence varies linearly with time. A time-domain waveform of duration T can be expressed as

$$c(t) = \mathrm{rect}\left(\frac{t}{T}\right) \cdot e^{j\varphi(t)}, \tag{1}$$

where $\mathrm{rect}(t/T)$ is a rectangular signal:

$$\mathrm{rect}\left(\frac{t}{T}\right) = \begin{cases} 1, & \left|\frac{t}{T}\right| \le \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

In equation (1), $\varphi(t)$ represents the phase of the chirp signal and the equation for calculating the instantaneous frequency $f(t)$ from the signal phase can be expressed as

$$f(t) = \frac{1}{2\pi} \cdot \frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}. \tag{3}$$

Due to the linear relationship between chirp signal frequency and time, there is

$$f(t) = f_c + \mu \cdot \frac{B}{T} \cdot t = f_c + \mu K t, \tag{4}$$

where f_c represents the carrier frequency, μ represents the instantaneous frequency changing slope of the chirp signal, B represents the bandwidth, and $K = \mathrm{TB}$ represents the frequency modulation slope. $\mu = 1$ means up-chirp and $\mu = -1$ means down-chirp. The IQ signals and instantaneous frequencies of the up-chirp and down-chirp are shown in Figures 4 and 5, respectively, and the signal frequency varies linearly within a bandwidth of 250 kHz.

LoRa modulation encodes the data by cyclic shifting the chirp signal by k bits, where $0 \le k \le 2^{SF} - 1$. After k-bit cyclic shifting on equation (4), the result can be expressed as

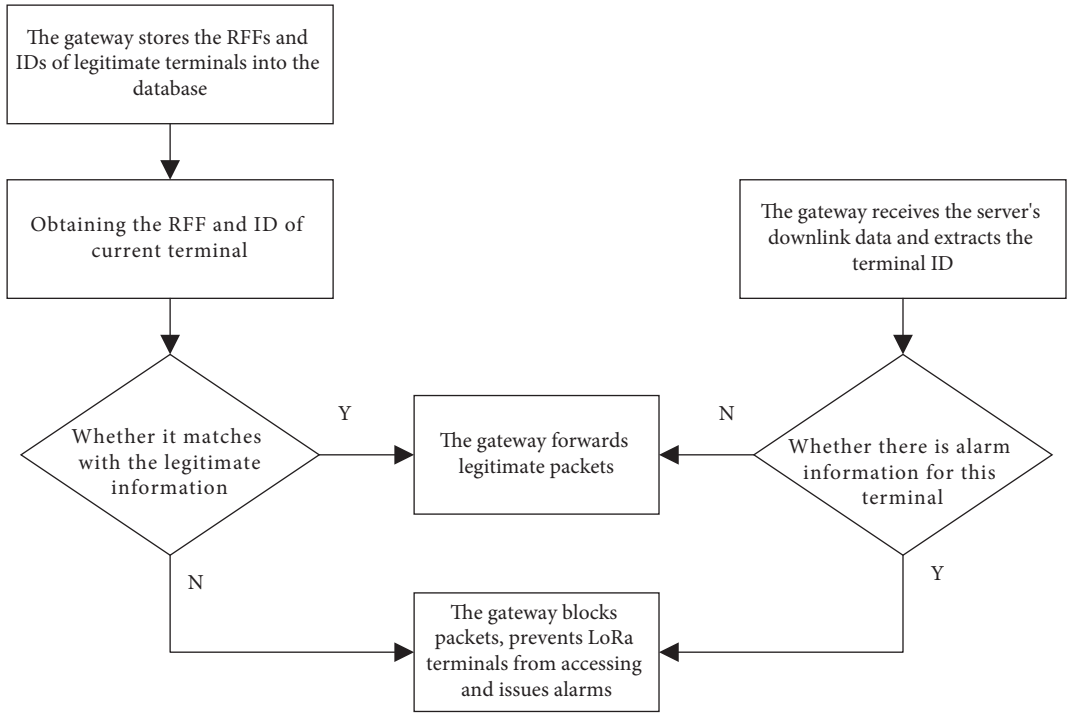

Figure 3: The workflow of LW-LoRaWAN.

Table 1: Comparison of two LoRa gateway security enhancement schemes.

Schemes	Modern cryptography	RFF technology
Modulation and demodulation	Using LoRaWAN	Self-realization
Communication distance	Far	Normal
Access algorithm freedom	No	High
Gateway upgrading content	Software upgrading	Hardware and software upgrading
Terminal upgrading content	Software upgrading	No
Protection mechanism	Increasing cracking workload	Uniqueness of physical features
Counterfeit attack protection	No	Yes

$$f(t) = \begin{cases} f_c + \mu K\left(t - \dfrac{k}{B}\right), & -\dfrac{T}{2} + \dfrac{k}{B} \le t \le \dfrac{T}{2}, \\ f_c + \mu K\left(t - \dfrac{k}{B}\right) + B, & -\dfrac{T}{2} \le t \le -\dfrac{T}{2} + \dfrac{k}{B}. \end{cases} \tag{5}$$

Taking Figure 4 as the reference, Figure 6 shows the waveform of chirp signal after cyclic shifting of 30 bits.

As shown from Figure 4 to Figure 6, the rule of cyclic shifting for LoRa modulation is obvious, which verifies the feasibility of obtaining the same waveform from the actual LoRa signals. The typical time-frequency diagram of LoRa data frame is shown in Figure 7. In this experiment, the data frame is in an explicit header mode and the data contain 10 up-chirp, 2.25 down-chirp, the explicit header, and the payload. As shown in Figure 7, the up-chirp and down-chirp remain unchanged in each data frame and the subsequent payload varies as the transmitted information changes. As long as each chirp symbol can be synchronized, the cyclic shifting of the payload into the same waveform can be realized.

4.2. LoRa Data Frame Composition. The LoRaWAN protocol mainly defines the technical details and specifications of the LoRa physical layer and MAC layer. The LoRa physical layer frame has two message formats: uplink and downlink. Uplink messages sent by the terminals reach the remote server through the gateway and the downlink is in the opposite direction. Both the uplink and downlink messages include the preamble and PHYPayload. In explicit mode, the message includes the physical layer header (PHDR) and its cyclic redundancy check (PHDR_CRC), which are not included in implicit mode. In terms of frame format, the only difference between uplink and downlink messages is that uplink messages have a cyclic redundancy check (CRC) to protect the integrity of the payload. The data frame format in explicit mode is shown in Figure 8.

The preamble is composed of n up-chirp symbols and 2.25 down-chirp symbols for data synchronization and the value of n is selected from 6 to 65536. The PHDR includes the payload data length, CR, and other values. When the above values are fixed, the implicit mode can be selected to shorten the transmission time. The length of the payload is variable, and its content includes data and MAC layer settings.

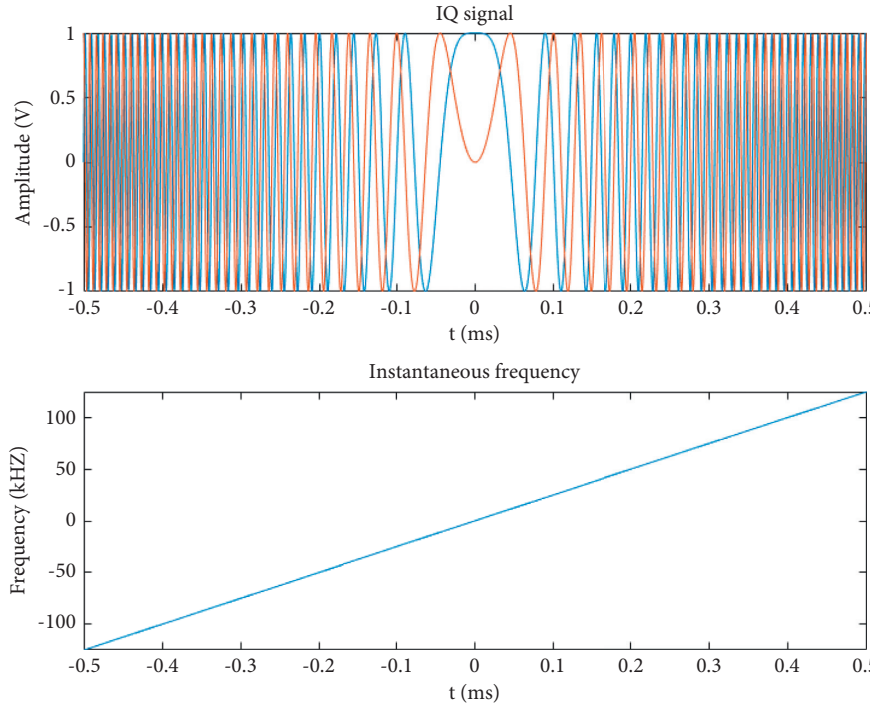

Figure 4: Parameters of up-chirp.

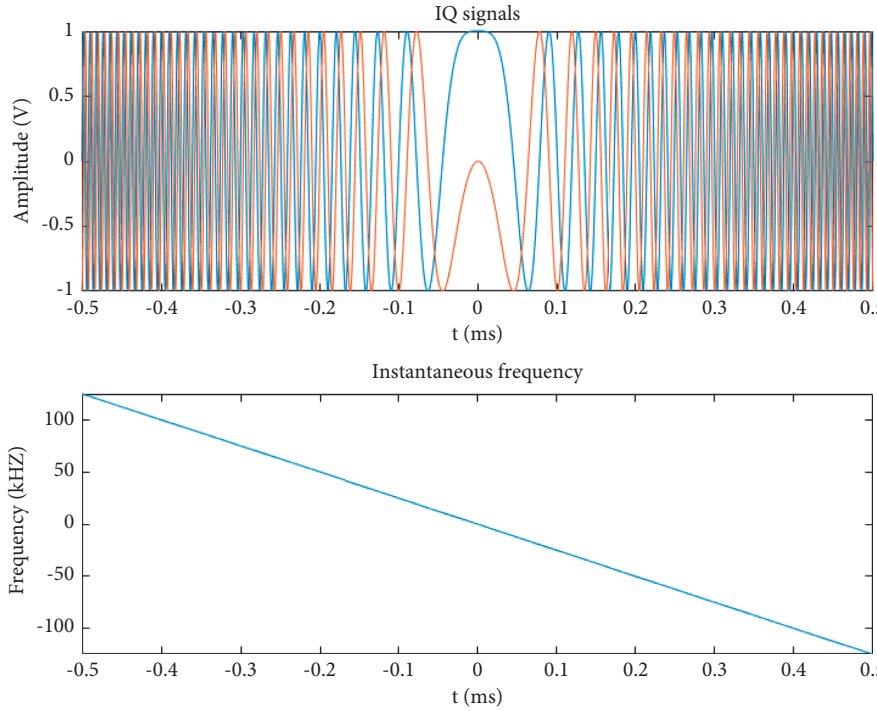

Figure 5: Parameters of down-chirp.

4.3. Cross Power Spectral Density (CPSD) Extraction. The CPSD can be used to describe the correlation between two random processes at each frequency point. The chirp symbol $x(n)$ in the LoRa signal is cyclically correlated with the up-chirp symbol $y(n)$ and then the CPSD features of the chirp symbol are obtained by the Fourier transform. CPSD reflects the energy features of the chirp symbols in the amplitude-frequency curve and the cyclic shifting features in the phase-frequency curve. Therefore, through the amplitude-frequency curve, the preamble and payload of the data frame can be effectively analyzed in the same dimension without considering the difference of the initial frequencies.

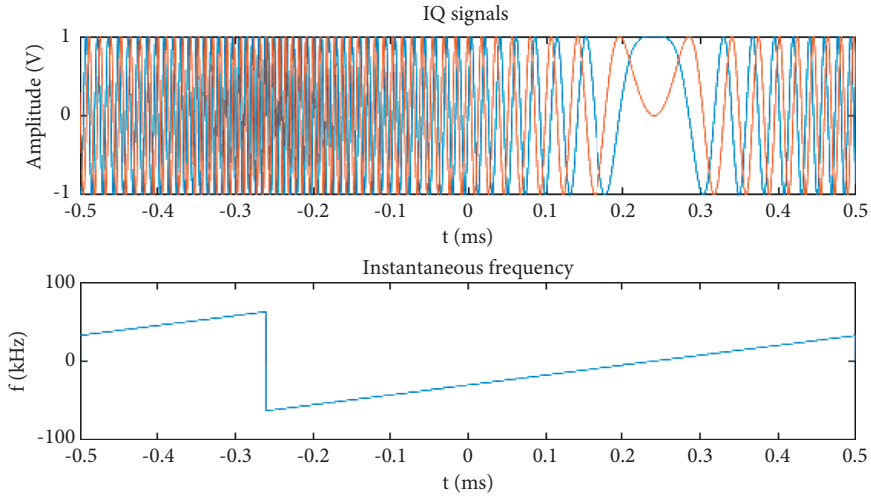

Figure 6: Chirp signal waveform after 30-bit shifting.

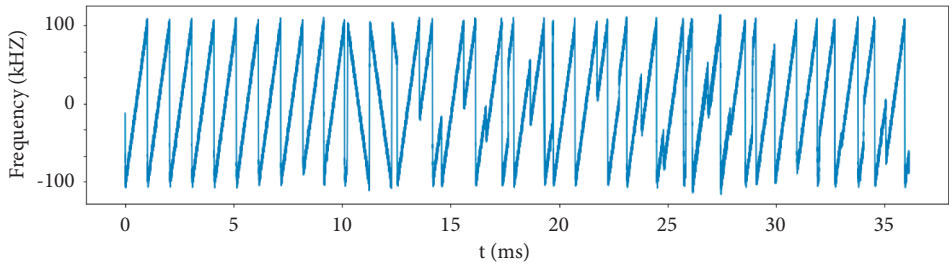

Figure 7: Time-frequency diagram of LoRa data frame.

(a) (b)

Figure 8: LoRa physical layer frame format (explicit mode). (a) Upstream frame format. (b) Downlink frame format.

Cyclic correlation is a kind of correlation operation for the cyclic shifting of the data sequences. Since both $x(n)$ and $y(n)$ are finite-length sequences of equal length, $y(n)$ can be selected for periodic extension. The period is the number of sampling points N of a chirp symbol and then $y(n)$ is shifted to the left by m bits after N points extension $R_N(n)$, which can be expressed as $y((n+m))_N R_N(n)$. The sequences can be shifted in one direction, because, after the periodic extension, shifting one position to the left is equivalent to shifting $N-1$ positions to the right. Finally, taking N values from $(0, N-1)$, the N sequence values are obtained after the cyclic shifting.

The N-point cyclic correlation of $x(n)$ and $y(n)$ can be expressed as

$$r_{xy}(m) = \sum_{n=0}^{N-1} x(n) y((n+m))_N R_N(n), \tag{6}$$

where $m = 0, 1, 2, ..., N$ and $r_{xy}(m)$ represents the correlation result of $x(n)$ and $y(n)$ after cyclic shifting by m points. The length of r_{xy} is also N points and the N-point CPSD of $x(n)$ and $y(n)$ can be expressed as

$$G_{xy}(k) = \frac{1}{N} \sum_{m=0}^{N-1} r_{xy}(m) e^{-j\frac{2\pi}{N}mk}, 0 \le k \le N-1. \tag{7}$$

There are multiple chirp symbols in a LoRa frame. In order to prevent the first and last symbols from possible power instability, the average CPSD of the rest of chirp symbols is taken as the RFF feature of the frame. The specific steps for calculating this feature are as follows:

(1) Extracting L chirp symbols $X_i, i \in (1, 2, \cdots, L)$ in the middle of the data frame.

(2) Calculating the cyclic correlation r_i between $X_i, i \in (1, 2, \cdots, L)$ and the up-chirp symbol $Y(n)$, respectively.

$$r_i(m) = \sum_{n=0}^{N-1} X_i(n) Y((n+m))_N \cdot R_N(n), m \in (0, 1, \cdots, N-1), \tag{8}$$

where $i \in (1, 2, \cdots, L)$.

(3) Calculating the CPSD vector G_i from the cyclic correlation vector r_i.

$$G_i(k) = \frac{1}{N}\sum_{m=0}^{N-1} r_i(m)e^{-j\frac{2\pi}{N}mk}, 0 \le k \le N-1, \quad (9)$$

where $i \in (1, 2, \cdots, L)$.

(4) Averaging the CPSD vectors to obtain the CPSD features of the frame.

$$\mathrm{CPSD}(n) = \frac{\sum_{i=1}^{L} G_i(n)}{L}, \quad 0 \le n \le N-1. \quad (10)$$

Equation (10) indicates that the CPSD feature of the signal is an N-dimensional vector. When $N = 2048$, the CPSD diagram of two terminals is depicted in Figure 9. The curves of the CPSD between different terminals are roughly the same. Then, the expanded view is shown in Figure 10 and there are intuitive differences between the two terminals.

Since the feature dimension is 2048 in Figure 9, the amount of data is large. In order to reduce the computation, it is necessary to reduce the feature dimensions. For example, only the 925th to 1124th dimensions of the CPSD have larger amplitudes and more obvious differences, so this 200-dimensional vector can be used as the CPSD features of the terminal.

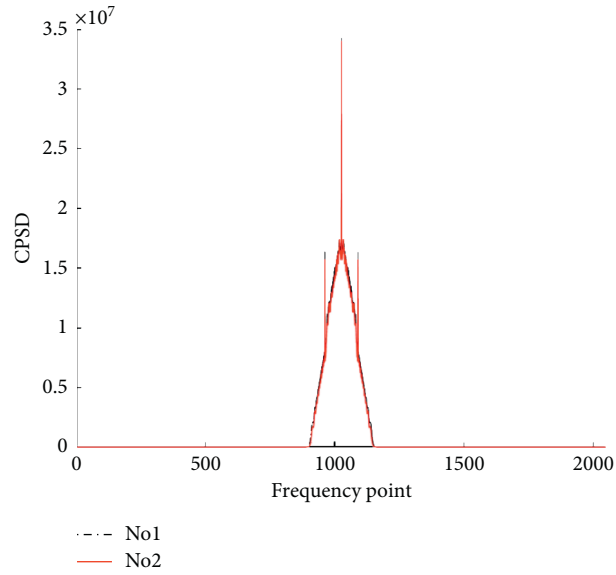

Figure 9: The CPSD of two terminals.

5. System Implementation and Testing

5.1. System Structure. The overall structure of the proposed system is shown in Figure 11. After the USRP completes the signal acquisition, downconversion, and analog-to-digital conversion, the RFF System executes the RFF feature extraction of the digital baseband signals and the training and recognition of the physical layer identity of different terminals. Then, the RFF System transmits the judgment results to the Management System for data management and integration, and then the results can be handled and displayed on the remote interface.

The physical hardware diagram of the system is shown in Figure 12. The sending device is a LoRa terminal powered by the USB interface and the receiving device is the NI USRP N210. The USRP and the computer deployed with the RFF System are connected through a gigabit network cable to exchange data. Meanwhile, in order to facilitate the development and demonstration of the program, the Management System and the display interface are deployed on the same computer. In practical applications, the display interface and part of the Management System can be deployed remotely and accessed through the Internet.

During the process, GNU Radio [38] is used to realize the sampling control and signal collection for USRP N210. GNU Radio is an open-source software toolkit for building and deploying software defined radio (SDR) systems. It can process wireless signals and control the parameters such as the sampling frequency, the spectrum range, and the gain. The system uses Python3 to support the signal demodulation and MATLAB to support the RFF feature extraction.

The LoRa terminals select the same batch of products from the same manufacturer, as shown in Figure 13. The product consistency makes the RFF features similar and poses a greater challenge for the classification and identification. The LoRa terminals set the carrier frequency to 433 MHz, the SF to 7, the BW to 125 kHz, and the number of up-chirp symbols in the preamble to 10. In the identification phase, the terminal ID is written to the payload to distinguish different terminals.

5.2. System Software Design. According to the system functions and workflow in Section 3, the system software can be divided into seven parts: signal acquisition module, data preprocessing module, RFF extraction module, device registration module, identification module, system management module, and display module.

The downconversion frequency is set at 433 MHz to realize zero IF acquisition, and the sampling frequency is set at 2 MHz. The file receiver module stores the acquired binary data locally, while the header module controls the amount of data collected at once.

GNU Radio Companion [39] is a visual interface supported by GNU Radio to achieve the signal acquisition. The configuration of the parameters is shown in Figure 14. For UHD, USRP Source is the parameter control module to collect signals from USRP. The downconversion frequency is set to 433 MHz to achieve the zero intermediate frequency acquisition and the sampling frequency is set to 2 MHz. The File Sink module stores the acquired binary data locally, while the Head module controls the amount of data collected for one time. The LoRa Receiver module has the same parameters as the LoRa terminal; it demodulates and decodes the LoRa signals. The Message Socket Sink module transmits the demodulated data to the local computer through the Socket and establishes the relevance between the terminal ID in the demodulated data and the baseband data output by the File Sink module.

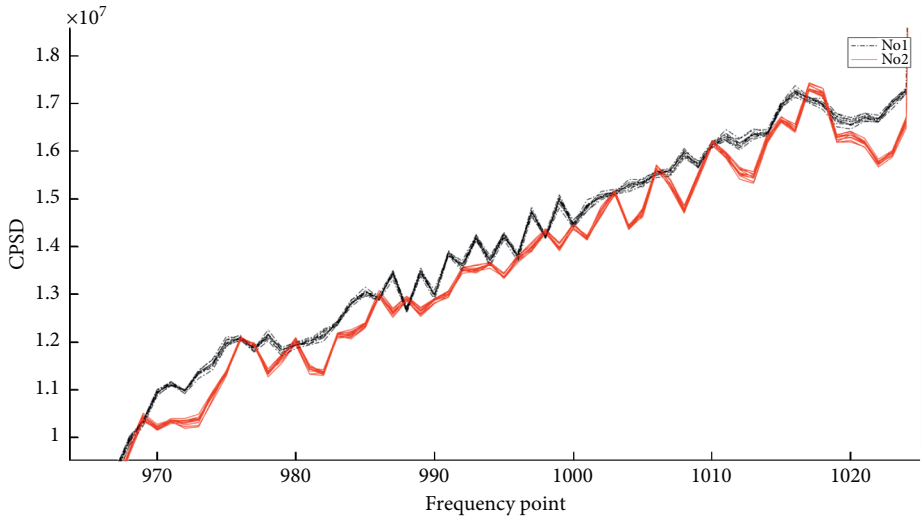

Figure 10: The expanded view of CPSD of the two terminals.

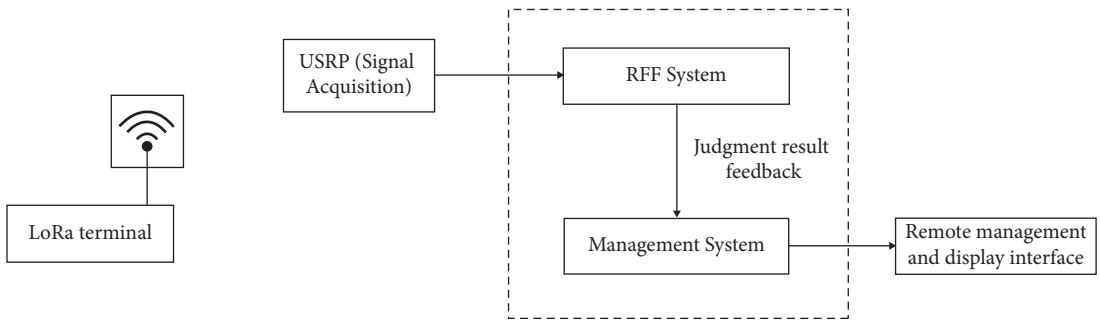

Figure 11: System structure diagram.

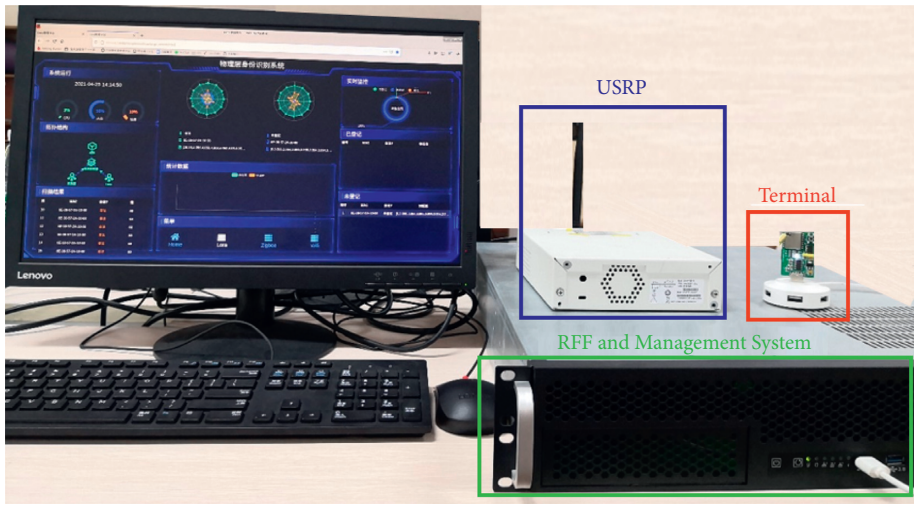

Figure 12: System hardware diagram.

The data preprocessing module performs the signal processing on the baseband signal, including reading valid signals, dividing data segments, normalization, frequency, and phase offset estimation and compensation. The RFF extraction module operates on the preprocessed data and establishes the relationship among the terminal ID, the timestamp, the signal strength, the signal-to-noise ratio (SNR), the carrier frequency offset, the CPSD, and other values extracted from each signal frame to form the unique RFF. The device registration module, identification module,

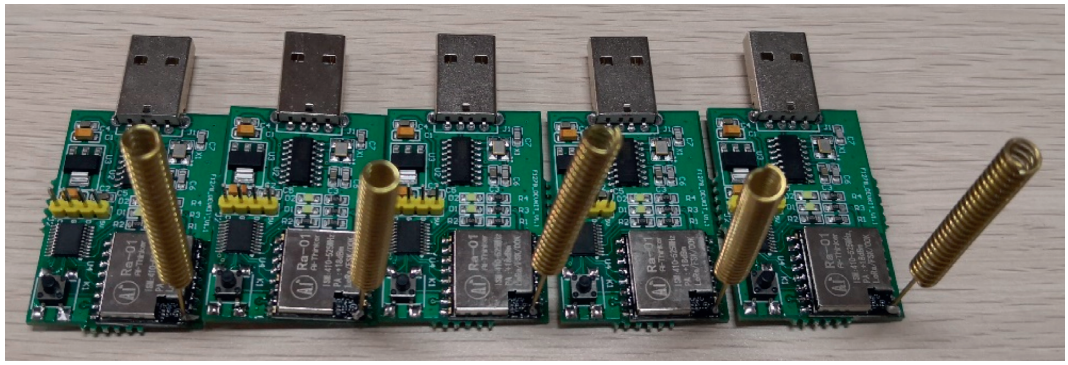

Figure 13: LoRa terminals.

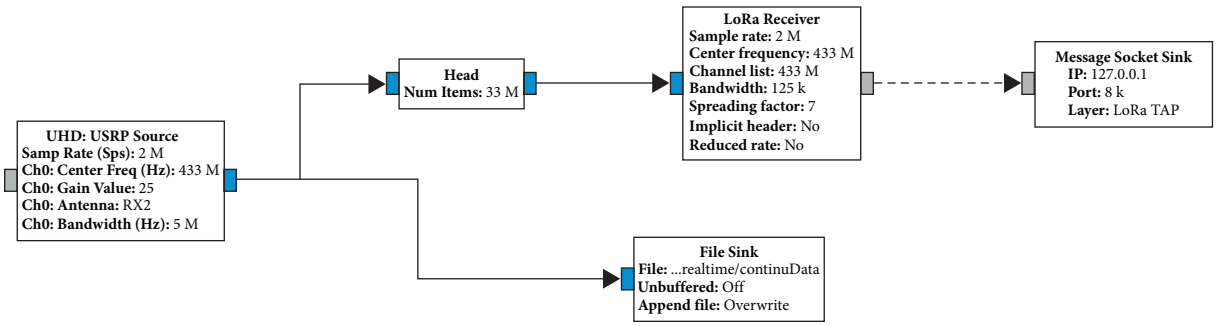

Figure 14: Signal acquisition module diagram.

and system management module operate according to the workflow in Figure 3. The display module provides the interface for parameter configuration and information display.

5.3. System Test

5.3.1. Test Scenario. After establishing the recognition system based on LoRa RFF features, the performance tests are carried out to ensure that the system can operate normally in different environments. This section focuses on the analysis of the performance tests.

System performance refers to the classification and recognition accuracy of legitimate terminals and the recognition rate of counterfeit terminals in a certain environment. Three experimental scenarios are tested. In each scenario, 8 LoRa terminals and one USRP are used. 50 sets of single-frame data are firstly collected for each terminal to generate a training model and then 100 sets are used to test the classification and recognition accuracy of the system.

In the first scenario, the LoRa terminals and the USRP communicate in a line-of-sight (LOS) indoor environment; hence, less interference comes from the surrounding and the RFF features are stable. The second scenario is the non-line-of-sight (NLOS) indoor environment and the received signal is greatly affected by multipath channel. The third scenario is a NLOS partition wall environment which leads to long distances, complex channels, and low SNR. By artificially adding white Gaussian noise to change the SNR values, the classification performance with different SNR has been evaluated.

5.3.2. Experiment 1: Comparison for Different RFF Features. The first experiment is carried out in Scenario 1 to verify the recognition efficiency of CPSD features by comparing the recognition accuracy of the carrier frequency offset, the IQ offset, and the CPSD.

Firstly, the performance of the carrier frequency offset features is analyzed, as shown in Figure 15. The linear discriminant analysis (LDA), linear kernel Support Vector Machine (SVM), and Gaussian kernel SVM are used for terminal recognition with multiple SNRs. Even at high SNR, the accuracy of carrier frequency offset features is only about 92%. This feature provides similar results with the three classification algorithms, but unfortunately some terminals with similar frequency offsets are difficult to distinguish.

Secondly, the IQ offset features are tested under multiple SNRs and the results are shown in Figure 16. The test results show that this feature requires high SNR and the recognition accuracy is less than 60% when the SNR is lower than 20 dB. Therefore, the accuracy of this feature is not sufficient, and it can only be used as an auxiliary feature for the device recognition. The analysis of CPSD features is drawn in Experiment 2.

5.3.3. Experiment 2: Comparison of CPSD for Different Scenarios. After dimension reduction for the CPSD features, the recognition accuracies under multiple SNRs for the three scenarios are shown in Figure 17. Compared with the results of Experiment 1 in Scenario 1, the recognition accuracy based on the CPSD feature has been significantly improved and exceeded 99% when SNR is 30 dB. In Scenario 2, when

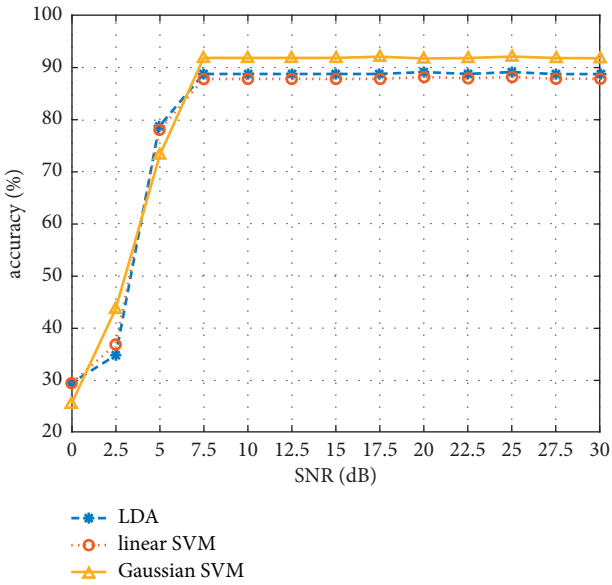

Figure 15: Test results of carrier frequency offset features.

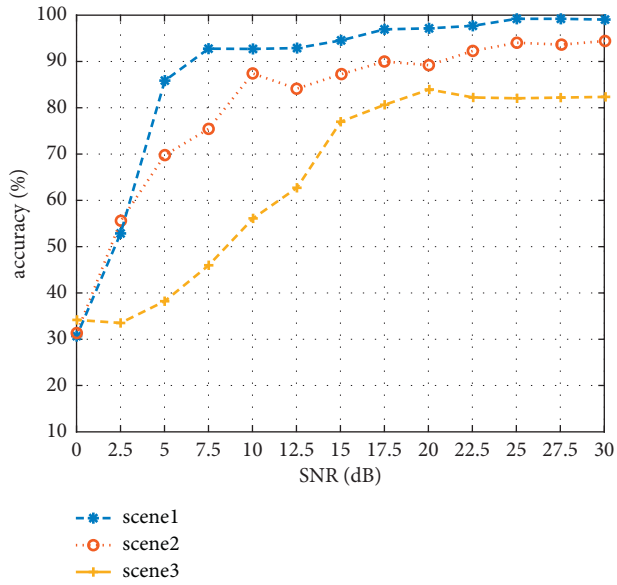

Figure 17: Test results of CPSD feature in three scenarios.

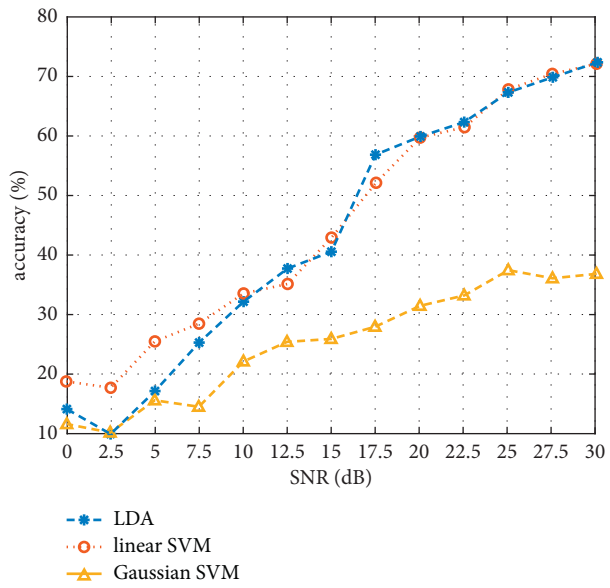

Figure 16: Test results of IQ offset feature.

the SNR is greater than 10 dB, the recognition accuracy exceeds 80%. In Scenario 3, under the condition of low SNR, the recognition accuracy is degraded, and the maximum recognition accuracy is about 83%.

The experimental results show that the CPSD feature using the payload information has a better recognition accuracy compared to the other two features. Therefore, the CPSD feature is more feasible in the proposed system.

5.3.4. Experiment 3: Analysis of Counterfeit Attack Protection. This experiment tests the system's ability to recognize counterfeit attacks. Since the ID of the LoRa terminal is written into the payload of the data frame, it can be tampered by changing the payload to achieve the counterfeit attack for the specified ID.

In this experiment, all terminals are numbered from 1 to 8 in order, and terminal 8 is used as the attacker to impersonate terminals 1 to 7, respectively. In Scenario 1, each terminal is tested for 100 counterfeit attacks and the results are shown in Table 2 for SNR of 30 dB. The detection success rate of counterfeit terminals refers to the percentage of counterfeit attacks recognized by the system. The false alarm rate refers to the probability that a legitimate terminal is identified as an illegal terminal.

The detection success rate of counterfeit terminals indicates the system's ability to recognize counterfeit attacks. The detection success rate reflects the system's ability to resist attacks from counterfeit terminals. The false alarm rate indicates the system's ability to recognize the legitimate terminals. A lower false alarm rate leads to a higher ability to recognize legitimate terminals.

As shown in Table 2, the proposed system has a high success rate of counterfeit attack detection. Meanwhile, the false alarm rate of legitimate terminals remains to be acceptable. Therefore, the proposed system can resist the counterfeit attack, which is unable to be achieved by the existing IoT security protection system.

5.4. Comparative Experiments. The proposed CPSD method is compared with the existing methods [35–37] mentioned in Section 2. The RFF features used in [35] are the time plus frequency offset (TFO). The supervised machine learning methods in [36] include MLP, convolutional neural network, and SVM, where MLP achieves the highest recognition accuracy. The classification method in [37] analyzed the features of the DCTF with the image recognition algorithm. Then, the performances of the four types of methods are evaluated with different fingerprinting experiments.

Table 2: Test results of counterfeit attacks.

Counterfeited terminal number	Detection success rate of counterfeit terminal (%)	False alarm rate (%)
1	95	3
2	96	4
3	98	2
4	97	2
5	95	3
6	94	5
7	96	3

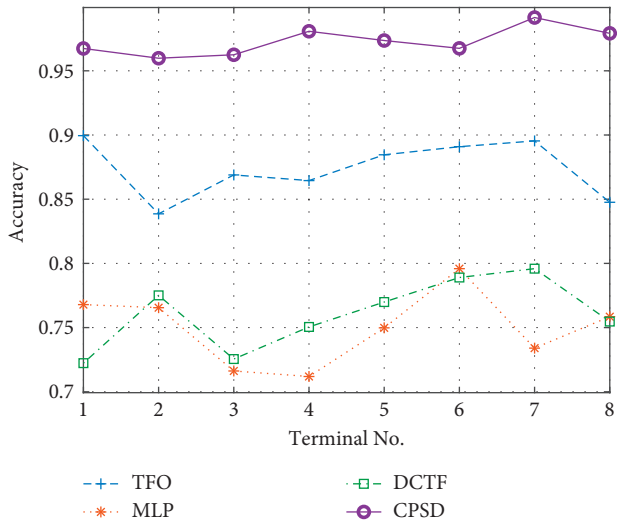

Figure 18: Accuracy comparison for 8 terminals.

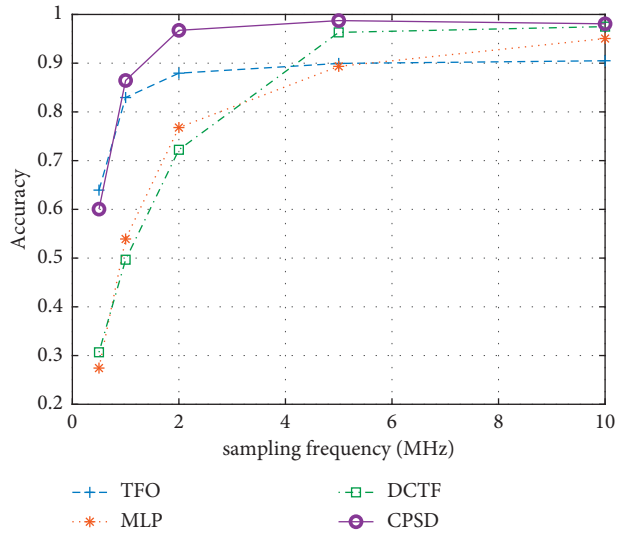

Figure 19: Accuracy comparison for different sampling frequency.

The first experiment tests the recognition accuracy of 8 terminals in Scenario 1 with a fixed SNR of 30 dB and a sampling frequency of 2 MHz. The comparison results are shown in Figure 18. The MLP and DCTF rely on the signal details at high sampling frequency, so the identification accuracy is poor. The TFO and CPSD are less affected by the sampling frequency, and the accuracy is above 80%.

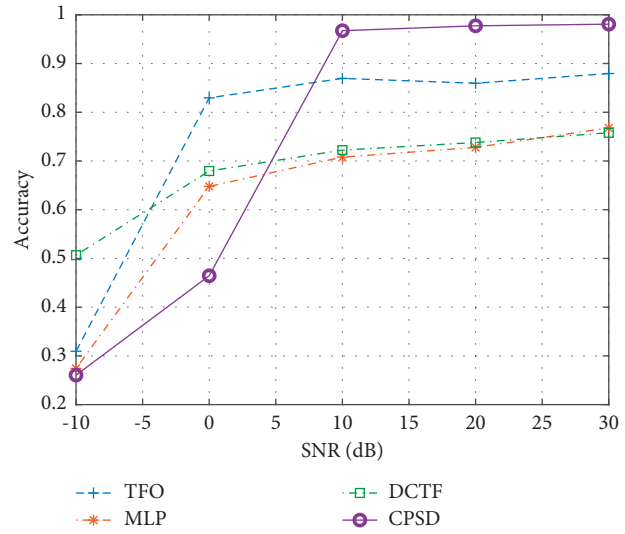

Figure 20: Accuracy comparison for different SNR.

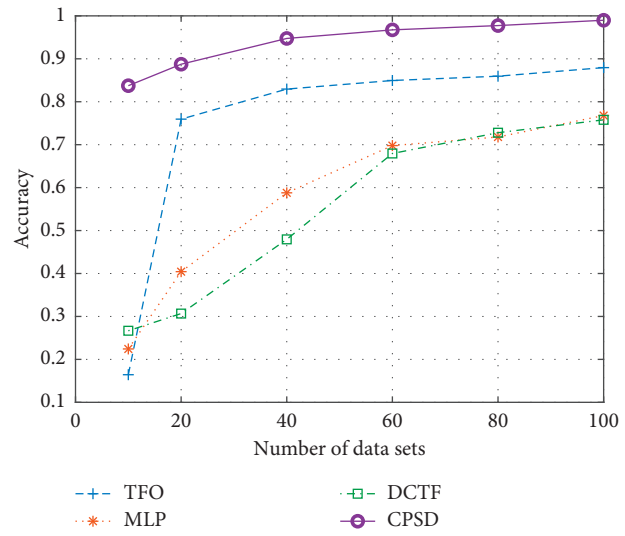

Figure 21: Accuracy comparison for different number of data sets.

The second experiment tests the effect of sampling frequency in Scenario 1, where each terminal has a fixed SNR of 30 dB. The average results are shown in Figure 19. The accuracies of TFO and CPSD are significantly higher than those of MLP and DCTF at low sampling frequencies. With the increase of sampling frequency, the accuracy of all the

algorithms is improved. When the sampling frequency exceeds 5 MHz, all the algorithms achieve better accuracy.

The third experiment tests the effect of SNR when the sampling frequency is 2 MHz in Scenario 1, and the average results are shown in Figure 20. The RFF reflects the subtle features of the signal, so the SNR has a great influence on the RFF. With the decline of SNR, the accuracy of all algorithms decreases obviously. Compared with TFO and DCTF, MLP and CPSD are more sensitive to the change of SNR.

The fourth experiment tests the effect of the number of data sets in Scenario 1 with a fixed SNR of 30 dB and a sampling frequency of 2 MHz. The average results are shown in Figure 21. The number of data sets determines the ability of the algorithm to quickly extract stable RFFs. The smaller the amount of data required to extract stable RFFs, the more practical the algorithm is. Unfortunately, most of the current research on RFF algorithms does not consider this problem. Since the CPSD algorithm makes effective use of the data part of the signal, higher accuracy can be obtained in the case of a small amount of data, as shown in Figure 21.

Through the comparative analysis of the above 4 experiments, the following conclusions can be drawn. Compared with the other three algorithms, the proposed CPSD algorithm achieves the best performance under the conditions of low sampling frequency and high SNR and can obtain high stable RFF with the least amount of data.

6. Conclusions and Future Work

Different from the secure access technology of modern cryptography, a lightweight gateway architecture called LW-LoRaWAN is proposed to achieve a data frame-based authentication with RFF. Compared to the two kinds of lightweight access techniques presented in Section 2, the RFF-based access method uses a different security policy. The main advantages and differences are listed as follows:

(1) LW-LoRaWAN is proposed to achieve a data frame-based authentication by establishing the relevance between RFF and data frame of LoRa terminals, which conforms to the concept of zero trust.

(2) The current RFF extraction methods for LoRa terminals cannot provide a stable RFF within a small number of packets, so a novel RFF feature of CPSD is proposed to achieve a fast authentication within one data frame.

(3) Since the RFF is unique and unclonable, LW-LoRaWAN can protect the LoRa terminals against the Sybil attacks.

(4) The proposed security enhancement system only needs to upgrade the gateway, without any change to the large number of terminals, which is more feasible than the existing enhancement schemes in practical applications.

From the results of our work, we can arrive at a conclusion that the proposed security policy could be a promising approach for LoRa terminal authentication. The theoretical analysis and experimental results show that the proposed system not only improves the authentication security of LoRa network but also protects the LoRa terminals against the counterfeit attacks. The LW-LoRaWAN provides new ideas from the physical layer for the security of LoRa devices.

In future work, we plan to test the system performance and terminal recognition rate for more communication scenarios. In addition, the RF performance of LW-LoRaWan depends on the capacity of the USRP, and we will try other RF platforms to improve the gateway performance.

Data Availability

The data supporting this study are available within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported in part by Jiangsu Key R&D Plan BE2019109; the National Natural Science Foundation of China under Grants 61601114, 61602113, 61801115, and 62001106; Natural Science Foundation of Jiangsu Province under Grants BK20160692, BK20200350, and BK20200352; Jiangsu Provincial Key Laboratory of Network and Information (Security no. BM2003201); and the Purple Mountain Laboratories for Network and Communication Security.

References

[1] M. Stoyanova, Y. Nikoloudakis, S. Panagiotakis, E. Pallis, and E. K. Markakis, "A survey on the internet of things (IoT) forensics: challenges, approaches, and open issues," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1191–1221, 2020.

[2] S. H. Haji and S. Y. Ameen, "Attack and anomaly detection in IoT networks using machine learning techniques: a review," *Asian Journal of Research in Computer Science*, vol. 9, no. 2, pp. 30–46, 2021.

[3] H. Mrabet, S. Belguith, A. Alhomoud, and A. Jemai, "A survey of IoT security based on a layered architecture of sensing and data analysis," *Sensors*, vol. 20, no. 13, p. 3625, 2020.

[4] W. Iqbal, H. Abbas, M. Daneshmand, B. Rauf, and Y. A. Bangash, "An in-depth analysis of IoT security requirements, challenges, and their countermeasures via software-defined security," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10250–10276, 2020.

[5] M. M. Ogonji, G. Okeyo, and J. M. Wafula, "A survey on privacy and security of Internet of Things," *Computer Science Review*, vol. 38, p. 100312, 2020.

[6] K. Lounis and M. Zulkernine, "Attacks and defenses in short-range wireless technologies for IoT," *IEEE Access*, vol. 8, pp. 88892–88932, 2020.

[7] Y. Jia, F. Zhong, A. Alrawais, B. Gong, and X. Cheng, "FlowGuard: an intelligent edge defense mechanism against IoT DDoS attacks," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9552–9562, 2020.

[8] I. Yaqoob, E. Ahmed, I. A. T. Hashem et al., "Internet of things architecture: recent advances, taxonomy, requirements, and

open challenges," *IEEE wireless communications*, vol. 24, no. 3, pp. 10–16, 2017.

[9] F. Olivier, G. Carlos, and N. Florent, "New security architecture for IoT network," *Procedia Computer Science*, vol. 52, pp. 1028–1033, 2015.

[10] N. Pimple, T. Salunke, U. Pawar et al., "Wireless security—an approach towards secured wi-fi connectivity," in *Proceedings of 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 872–876, IEEE, Coimbatore, Tamil Nadu, February 2020.

[11] Y. Wen and T. Liu, "WIFI security certification through device information," in *Proceedings of 2018 International Conference on Sensor Networks and Signal Processing (SNSP)*, pp. 302–305, IEEE, Xi'an, China, October 2018.

[12] E. Baray and N. K. Ojha, "WLAN security protocols and WPA3 security approach measurement through aircrack-ng technique," in *Proceedings of 2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 23–30, IEEE, Erode, India, April 2021.

[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] J. Sakhnini, H. Karimipour, A. Dehghantanha, and R. M. Parizi, "Physical layer attack identification and localization in cyber-physical grid: an ensemble deep learning based approach," *Physical Communication*, vol. 47, Article ID 101394, 2021.

[15] L. Bai, L. Zhu, J. Liu et al., "Physical layer authentication in wireless communication networks: a survey," *Journal of Communications and Information Networks*, vol. 5, no. 3, pp. 237–264, 2020.

[16] Z. Li, W. Xu, R. Miller et al., "Securing wireless systems via lower layer enforcements," in *Proceedings of the 5th ACM workshop on Wireless security*, pp. 33–42, Los Angeles California, September 2006.

[17] R. Zhang, L. Song, Z. Han et al., "Physical Layer Security for Two Way Relay Communications with Friendly Jammers," in *Proceedings of 2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pp. 1–6, IEEE, Miami, Florida, USA, December 2010.

[18] J. Chen, R. Zhang, L. Song et al., "Joint relay and jammer selection for secure two-way relay networks," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 310–320, 2011.

[19] P. Neumann, J. Montavont, and T. Noel, "Indoor deployment of low-power wide area networks (LPWAN): a LoRaWAN case study," in *Proceedings of 2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 1–8, IEEE, New York, NY, USA, October 2016.

[20] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, "Understanding the limits of LoRaWAN," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 34–40, 2017.

[21] M. Bor, J. E. Vidler, and U. Roedig, "LoRa for the Internet of Things," in *Proceedings of International Conference on Embedded Wireless Systems and Networks (EWSN) 2016*, pp. 1–7, Graz, Austria, 2016.

[22] A. Shah and M. Engineer, "A survey of lightweight cryptographic algorithms for iot-based applications," *Smart innovations in communication and computational sciences*, Springer, Singapore, pp. 283–293, 2019.

[23] A. Biswas, A. Majumdar, S. Nath et al., "LRBC: a lightweight block cipher design for resource constrained IoT devices," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 1–15, 2020.

[24] S. Aruna, G. Usha, P. Madhavan, and M. V. R. Kumar, "Lightweight cryptography algorithms for IoT resource-starving devices," *Role of Edge Analytics in Sustainable Smart City Development: Challenges and Solutions*, pp. 139–169, 2020.

[25] B. Seok, J. Park, and J. H. Park, "A lightweight hash-based blockchain architecture for industrial IoT," *Applied Sciences*, vol. 9, no. 18, p. 3740, 2019.

[26] M. A. Philip, "A survey on lightweight ciphers for IoT devices," in *Proceedings of 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy)*, pp. 1–4, IEEE, Kollam, India, December 2017.

[27] X. W. Wu, E. H. Yang, and J. Wang, "Lightweight security protocols for the internet of things," in *Proceedings of IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–7, IEEE, Montreal, QC, Canada, October 2017.

[28] A. S. Sani, D. Yuan, P. L. Yeoh et al., "A lightweight security and privacy-enhancing key establishment for internet of things applications," in *Proceedings of 2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, Kansas City, MO, USA, May 2018.

[29] K. I. Talbot, P. R. Duley, and M. H. Hyatt, "Specific emitter identification and verification," *Technology Review*, p. 113, 2003.

[30] J. Tapparel, O. Afisiadis, P. Mayoraz et al., "An open-source LoRa physical layer prototype on GNU radio," in *Proceedings of 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, IEEE, Atlanta, Georgia, USA, May 2020.

[31] M. O. Farooq, "Multi-hop communication protocol for LoRa with software-defined networking extension," *Internet of Things*, vol. 14, Article ID 100379, 2021.

[32] J. Souifi, Y. Bouslimani, M. Ghribi et al., "Smart home architecture based on LoRa wireless connectivity and LoRaWAN® networking protocol," in *Proceedings of 2020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP)*, pp. 95–99, IEEE, El-Oued, Algeria, March 2020.

[33] C. Bernier, F. Dehmas, and N. Deparis, "Low complexity LoRa frame synchronization for ultra-low power software-defined radios," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3140–3152, 2020.

[34] L. Tessaro, C. Raffaldi, M. Rossi et al., "Lightweight synchronization algorithm with self-calibration for industrial LORA sensor networks," in *Proceedings of 2018 Workshop on Metrology for Industry 4.0 and IoT*, pp. 259–263, IEEE, Brescia, Italy, April 2018.

[35] R. Eletreby, D. Zhang, S. Kumar et al., "Empowering low-power wide area networks in urban settings," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, pp. 309–321, Los Angeles CA USA, August 2017.

[36] P. Robyns, E. Marin, W. Lamotte et al., "Physical-layer fingerprinting of LoRa devices using supervised and zero-shot learning," in *Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 58–63, Boston Massachusetts, July 2017.

[37] Y. Jiang, L. Peng, A. Hu et al., "Physical layer identification of LoRa devices using constellation trace figure," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–11, 2019.

[38] T. W. Mathumo, T. G. Swart, and R. W. Focke, "Implementation of a GNU Radio and python FMCW Radar Toolkit," in *Proceedings of IEEE AFRICON*, pp. 585–590, IEEE, Cape Town, September 2017.

[39] M. Chino, H. Miyashiro, and A. J. Luis, "Implementation of SNR estimation algorithms, using LabVIEW communications and GNU radio companion," in *Proceedings of 2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, pp. 1–4, IEEE, Lima, Peru, August 2018.