applied sciences

MDPI

Article

Radio Frequency Fingerprinting for Frequency Hopping Emitter Identification

Jusung Kang [1], Younghak Shin [2], Hyunku Lee [3], Jintae Park [4] and Heungno Lee [1,*]

[1] School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea; k92492@gist.ac.kr
[2] Department of Computer Engineering, Mokpo National University, Muan-gun 58554, Korea; younghak@mokpo.ac.kr
[3] LIG Nex1 Company Ltd., Yongin 16911, Korea; hyunku.lee@lignex1.com
[4] Agency for Defense Development, Daejeon 34063, Korea; jtpark@add.re.kr
* Correspondence: heungno@gist.ac.kr; Tel.: +82-62-715-2237

Abstract: In a frequency hopping spread spectrum (FHSS) network, the hopping pattern plays an important role in user authentication at the physical layer. However, recently, it has been possible to trace the hopping pattern through a blind estimation method for frequency hopping (FH) signals. If the hopping pattern can be reproduced, the attacker can imitate the FH signal and send the fake data to the FHSS system. To prevent this situation, a non-replicable authentication system that targets the physical layer of an FHSS network is required. In this study, a radio frequency fingerprinting-based emitter identification method targeting FH signals was proposed. A signal fingerprint (SF) was extracted and transformed into a spectrogram representing the time–frequency behavior of the SF. This spectrogram was trained on a deep inception network-based classifier, and an ensemble approach utilizing the multimodality of the SFs was applied. A detection algorithm was applied to the output vectors of the ensemble classifier for attacker detection. The results showed that the SF spectrogram can be effectively utilized to identify the emitter with 97% accuracy, and the output vectors of the classifier can be effectively utilized to detect the attacker with an area under the receiver operating characteristic curve of 0.99.

Keywords: frequency hopping signals; radio frequency fingerprinting; emitter identification; outlier detection; physical layer security; inception block; deep learning classifier

Citation: Kang, J.; Shin, Y.; Lee, H.; Park, J.; Lee, H. Radio Frequency Fingerprinting for Frequency Hopping Emitter Identification. *Appl. Sci.* **2021**, *11*, 10812. https://doi.org/10.3390/app112210812

Academic Editor: Ernesto Limiti

Received: 8 October 2021
Accepted: 11 November 2021
Published: 16 November 2021

1. Introduction

The most important task in user authentication of a wireless communication system is to identify the emitter information of RF signals. A common way to confirm the emitter information, that is, the emitter ID, is to decode the address field of the medium access control (MAC) frame [1]. However, under this digitized information-based authentication process on a MAC layer, an attacker can possess the address information and imitate it as an authenticated user. To prevent this weakness, a physical layer authentication process, namely radio frequency (RF) fingerprinting, has been studied in recent years [2].

RF fingerprinting is an identification technique that utilizes a signal fingerprint (SF) to identify the unique emitter source of an RF signal. In the manufacturing of RF components inside an emitter, process tolerance is inevitable. These tolerances affect subtle differences in the features of the emitted RF signal. Because these process tolerances are not reproducible, an SF can act as the fingerprint of an emitter. It can also be utilized as a non-replicable authentication key to identify the authenticated user [3].

RF fingerprinting can be used to distinguish SFs in RF signals [4–8]. A conventional approach is to design a handcrafted feature from the SFs based on domain knowledge. In [4], the statistical moment and entropy were calculated from spectrograms of the transient signal to identify Bluetooth devices. In [5], statistical moments were calculated from

preamble signals to identify Bluetooth devices. In [6], the principle components of the transient and steady state signals using sparse representation were proposed to identify Walkie-Talkies. A recent approach is to train SFs directly using a deep learning-based classifier. In [7], the signal difference between the received signal and the ideally encoded signal was calculated as the SF. It was trained using a one-dimensional convolutional neural network (CNN)-based ensemble classifier to identify ZigBee devices. In [8], the Hilbert spectrogram of the SF was utilized to train a residual-based classifier.

The frequency hopping spread spectrum (FHSS) is a highly secure communication protocol frequently used in secure communication systems [9]. With an FHSS system, the frequency hopping (FH) signal rapidly hops from one frequency to another in a predefined pseudo-random fashion. This hopping pattern is known only to the transmitter–receiver pair. Thus, an attacker who does not possess the hopping pattern cannot pretend to be an authenticated user. In this case, the hopping pattern is a key for the authentication process on the physical layer of the FHSS network.

However, in recent days, attackers may have possessed the predefined hopping patterns. Especially for the FHSS network in the industry–science–medical (ISM) band as the hopping sequence is described in the IEEE 802.11 standard [10]. Scholars have also estimated the hopping pattern as a blind estimation condition [11–13]. In [11], the hopping sequence was extracted from the USRP device, and an attack model based on the extracted hopping sequence was discussed. In [12], a real-time hopping frequency tracking model based on the autoregressive moving average was proposed. In [13], the FH signals were sorted based on the power and hopping time information. From these studies, hopping patterns are expected to be traceable today and reproduced in the future.

Recently, a frequency hopping network based on non-orthogonal multiple access (NOMA) was proposed [14]. In the NOMA system, the probability that the attacker intercepts the FH signal can be reduced through the two-stage relay communication [15], the additive artificial noise method [16], and the optimization of the power allocation for the beamforming scheme [17]. However, this anti-interception capability is closely related to the outage probability of the NOMA users, closely related to the signal power. This means that if the attacker is closely located to the near-user side with a high SNR value, the attacker can intercept the FH signal and trace the hopping pattern.

Once a hopping pattern is reproducible, an attacker can generate FH signals similar to those of the authenticated user. The two hopping patterns become undiscernible and the attacker can pretend to be the user. In this case, the received signal can be demodulated to proceed to the MAC layer inspection step. The MAC layer authentication system should discern the attacker unless even the digital key is exposed to the attacker. That is, if the attacker knew the digital key of the network system, the attacker would be able to pretend to be the authenticated user, which is the case in deceptive jamming attacks [18] or man-in-the-middle attacks [19]. These attacks are not easily detectable and can flood fake data to mislead the network system [18]. To prevent such attacks, a non-replicable authentication system that can detect an attacker who even knows the digital key is required.

This study aims to propose an enhanced solution to the physical layer authentication problem in the case in which the attacker can reproduce the hopping pattern. The scenario of the problem is shown in Figure 1. It is assumed that the user, attacker, and receiver exist in the FHSS network. The goal of the attacker is to deceive the receiver by emitting the imitated FH signal based on the replicated hopping pattern. The primary goal of the receiver is to decide if the signal received came from the user or from the attacker.

The novel receiver algorithm we propose in this study is an RF fingerprinting-based emitter identification (RFEI) method that targets the physical layer of the FHSS network. By examining the emitter ID on the received FH signal, the receiver can decide if the current FH signal is emitting from one of the allowed users. If the emitter ID of the current FH signal is not included in the set of authenticated user IDs, the receiver can reject the current FH signal before it is passed to the MAC layer. The RFEI method can achieve system enhancement by being applied to the user authentication process. As the key of the RFEI

method, that is, the SF, is generated by the process tolerances during the manufacturing process, the attacker cannot reproduce it. By detecting these attackers based on the SFs, non-replicable authentication systems can be achieved wherein the receiver can reject FH signals even if an attacker knows the hopping pattern and the digital key.

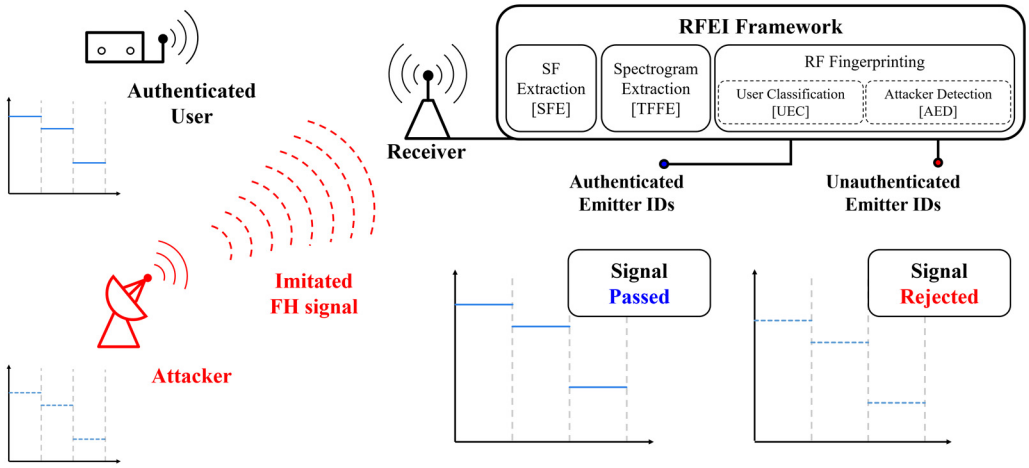


Figure 1. Non-replicable authentication scenario based on the RFEI method.

The RFEI method consists of four steps: SF extraction (SFE, Section 3.1), time–frequency feature extraction (TFFE, Section 3.2), user emitter classification (UEC, Section 3.3), and attacker emitter detection (AED, Section 3.4). As a preprocessing step, the target hop signal is down-converted to the baseband based on the hopping pattern known to the receiver. The baseband hop signal is passed to the SFE step to extract the analog SFs, i.e., rising transient (RT), steady state (SS), and falling transient (FT) signals are extracted. The SF is provided to the TFFE step to transform the SF into the time–frequency domain, i.e., the spectrogram. The spectrogram is provided to the UEC stage to train and test the spectrogram on a custom deep inception network (DIN)-based classifier. In addition, the ensemble approach is applied to exploit the multimodality of the analog SFs. Finally, the classifier output vector is provided to the AED step in which a detection algorithm is applied to detect the FH signal of the attacker. The novelties of this study are that (1) RF fingerprinting methods were evaluated targeting for FH signals, (2) the ensemble approach was applied to utilize the multimodality of SFs, and (3) the RFEI framework was employed to identify users and detect attackers simultaneously.

The RFEI algorithm was evaluated on a few SFs and ensemble-based approaches. The algorithm compares to well-designed baselines inspired by recent approaches described in the RF fingerprinting literature [4,5,7,8]. The experiments were performed using an actual FH dataset to evaluate the reliability of the algorithm. The results confirm that the proposed DIN classifier could improve the emitter ID identification accuracy by more than 1% compared to the baseline (Section 5.1). In addition, the multimode SF ensemble approach proved to be the most effective, achieving the best results with 97.0% identification accuracy for the seven FHSS emitters (Section 5.2). Regarding the detection performance, the classifier output vector of the outliers exhibited a much lower value than those of the training sample. By utilizing these differences, the detector based on the DIN-based ensemble classifier can improve the area under the receiver operating characteristic curve (AUROC) from 0.97 to 0.99 compared to the baseline. This result indicates that the classifier output vectors can effectively be used to detect the attacker signal input (Section 5.4).

The remainder of this study is organized as follows. The problem formulation is presented in Section 2. The details of the RFEI method are described in Section 3, and the baseline algorithms are explained in Section 4. The results, a discussion, and other details of the experiments are described in Section 5. The conclusion is presented in Section 6.

2. Problem Formulation

2.1. Frequency Hopping Signals of Frequency Hopping Spread Spectrum Network

In this study, we consider an FHSS network in which K FH signals are observed in a single receiver. To consider the ability of attackers to imitate FH signals similar to those of an authenticated user, we assume that the hth hopping times of the kth FH signals t_h^k have the same value, that is, the FH signals hop simultaneously. An example of an FHSS network with the two different FH signals is presented in Figure 2.

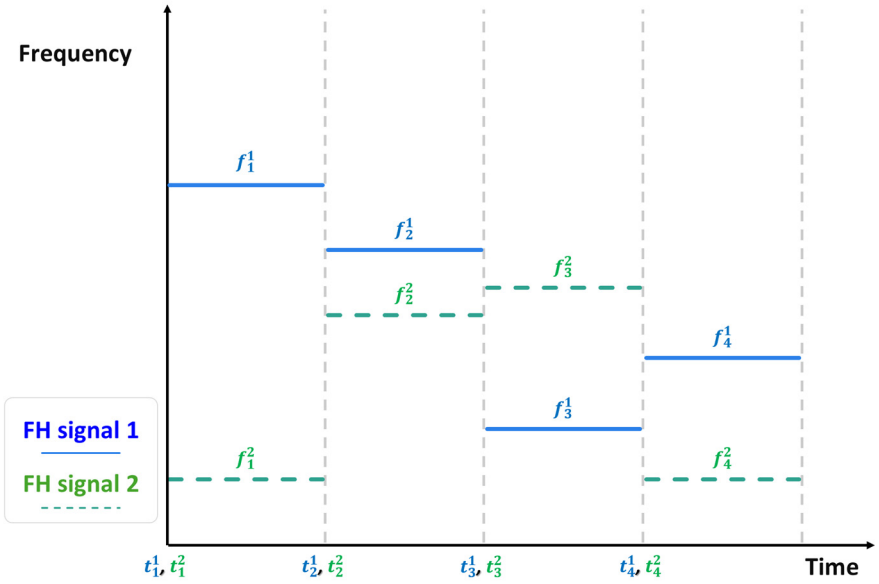


Figure 2. FH signals in two FHSS networks.

A single FH signal is defined as follows

$$x^k(t) = a^k e^{j2\pi(f^k(t)t+\varphi^k(t))} \tag{1}$$

where $x^k(t)$ is the FH signal emitted by the th emitter, a^k is the amplitude and is the hopping frequency of the kth FH signal $x^k(t)$, and $\varphi^k(t)$ is the phase difference modulated by the kth message signal $m^k(t)$. When the message signal is modulated with frequency modulation (FM), the phase difference is defined as follows

$$\varphi^k(t) = \int_{-\infty}^{t} m^k(\alpha)d\alpha \tag{2}$$

From Equation (1), all the K FH signals simultaneously observed by a single receiver can be defined as follows

$$y(t) = \sum_{k=1}^{K} x^k(t) + n(t) \tag{3}$$

where $y(t)$ is the observed RF signal and $n(t)$ is the additive white Gaussian noise (AWGN) present in the channel environment.

The FH signal is observed during the observation time T. During this time, a total of H hops are observed. Within a single hop duration of the hth hop signal, $t_h \leq t < t_{h+1}$, the hopping frequency $f^k(t)$ is held constant at f_h^k, denoting the hth hopping frequency of the kth FH signal. Thus, Equations (1) and (3) can, respectively, be reformulated as follows

$$x_h^k(t) = Ae^{j2\pi(f_h^k t+\varphi_h^k(t))}, \text{ for } t_h \leq t < t_{h+1} \tag{4}$$

$$y_h(t) = \sum_{k=1}^{K} x_h^k(t) + n(t), \text{ for } t_h \le t < t_{h+1} \tag{5}$$

where $x_h^k(t)$ is the hth hop signal of the kth FH signal and $y_h(t)$ is the observed RF signal during the hth hop duration, $t_h \le t < t_{h+1}$, where a total of K hop signals exist.

2.2. User Authentication in Frequency Hopping Spread Spectrum Networks

In an FHSS network, the core process for user authentication can be performed in two steps: (1) determining whether or not the appropriate hopping frequency is measured, and (2) determining whether or not the header information of the MAC frame is correct.

Because we assume that the attacker can reproduce the predefined hopping pattern $f^k = \left[f_1^k, f_2^k, ..., f_H^k\right]$, the imitated FH signal will display the same hopping frequency pattern. The imitated FH signal will be demodulated and passed through the MAC layer, that is, Step 1 is disabled. The process of inspecting the address field in the MAC header remains. However, because this address information has been digitized, the attacker can possess and imitate this address field. If an attacker sends an address field similar to an authenticated emitter, there is no way to detect and prevent it. Therefore, the emitter identification process based only on header information of the MAC frame is not sufficient to reject the imitated FH signal.

2.3. Emitter Identification Based User Authentication in Frequency Hopping Spread Spectrum Networks

We propose a non-replicable authentication system that operates on the physical layer of the FHSS network presented in Figure 3 and Algorithm 1. By adding the emitter identification framework within the authentication process, we can achieve an enhanced physical layer authentication system for the FHSS network by verifying (1) whether or not the appropriate hopping frequency is measured, (2) whether the emitter ID of the current FH signal is an authenticated user or attacker, and (3) whether or not the header information of the MAC frame is correct.

In this study, our target was to evaluate the RFEI framework for the FH signals corresponding to Step 2 of Algorithm 1. We intended to develop an algorithm to estimate the emitter ID from the baseband FH signal such that

$$s_h^k(t) = Ae^{j2\pi\varphi_h^k(t)}, \text{ for } t_h \le t < t_{h+1} \tag{6}$$

$$\widetilde{k} = \mathrm{F}_{\mathrm{RFEI}}\left(s_h^k(t)\right) \tag{7}$$

where $s_h^k(t)$ is the baseband hop signal down-converted from the hop signal $x_h^k(t)$ and $\widetilde{k}$ is the emitter ID estimated from the RFEI algorithm $\mathrm{F}_{\mathrm{RFEI}}$.

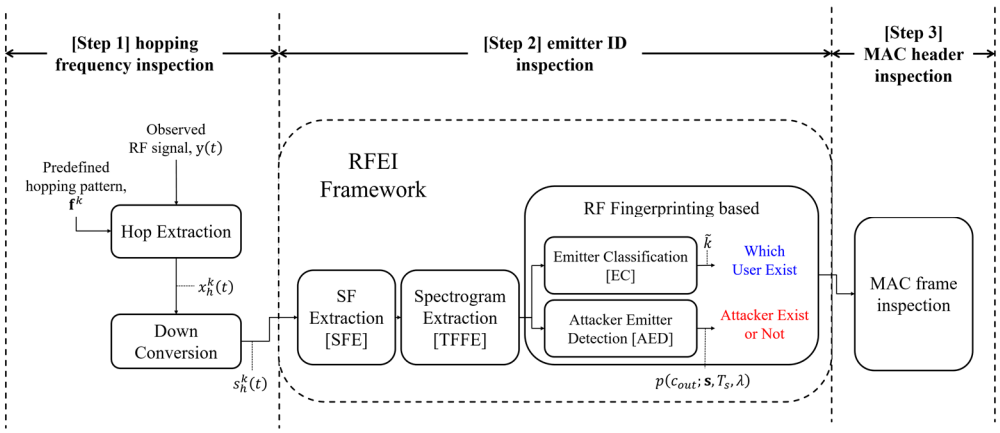


Figure 3. Block diagram of the RFEI-based non-replicable authentication system.

As the receiver knows the hopping frequency, f_h^k, the target hop signal, $x_h^k(t)$ can be extracted from the observed FH signal, $y_h(t)$. This approach is reasonable as the FH signal must be demodulated to an intermediate frequency (IF) or baseband and passed to the MAC layer to decode the digital data modulated by the message signal, $m^k(t)$. The SFs are non-replicable differences dependent on the manufacturing process of the emitter. Therefore, the SFs are independent of the hopping frequency and should be in the baseband of the hop signal, $s_h^k(t)$.

Algorithm 1. Non-replicable authentication system for the physical layer of the FHSS network.

Input: The observed RF signal $y(t)$
For each hop duration, $t_h \leq t < t_{h+1}$ **do:**
Step1: Extract and down-convert the target hop signal $x_h^k(t)$ to the baseband hop signal $s_h^k(t)$ from the observed signal $y_h(t)$ based on a predefined hopping pattern f_h^k.
If *RFEI* is activated **do:**
Step 2-1: Estimate the emitter ID based on the RFEI algorithm on (7)
Step 2-2: Pass the hop signal $x_h^k(t)$ when the emitter ID k is an authenticated emitter ID.
Step 2-3: Reject the hop signal $x_h^k(t)$ when the emitter ID k is an attacker's emitter ID.
Step 3: Send all passed baseband hop signals $s_h^k(t)$ to the next step, i.e., the MAC frame inspection.
Output: The authenticated baseband signal $x^k(t)$.

3. Proposed RF Fingerprinting-Based Emitter Identification Method

The RFEI algorithm is implemented as follows.

- SF extraction: An SF is an RF signal that contains feature information for emitter ID identification. It can be any signal involved in the demodulation process for communication. However, the SF used in this study focused on analog SF, i.e., RT, SS, and FT signals.
- Time–frequency feature extraction: A feature is a set of values containing physical measurements that can ensure robust classification. Any feature having a physical meaning can be applied from statistical moments to a raw preamble signal. In this study, a spectrogram of the SF was considered.
- User emitter classification: Classification is a decision process in which an emitter ID can be estimated from an input feature. A classifier was trained and tested on a large set of extracted features. Subsequently, the emitter ID was estimated from the classifier output vector. In this study, we consider a discriminative classifier model from a support vector machine (SVM) to a DIN-based ensemble classifier.
- Attacker emitter detection: This detection process enables the classifier to search whether the input feature has been trained for the classifier. The difference between the classifier output characteristics of the trained and outlier samples can be utilized. In this study, a simple but effective threshold based approach was applied.

The RFEI method can be formulated as a classification problem using the following expression

$$\mathbf{y} = \mathrm{F}_{\mathrm{RFEI}}(\mathbf{s}) \tag{8}$$

where $\mathbf{s} = [s(T_s), s(2T_s), ..., s(NT_s)] \in \mathbb{C}^{N\times 1}$ is a baseband hop signal sampled by the sampling period T_s. The vector representation of the signal is now used in this study for convenience. Further, N is the length of a complex-valued baseband hop signal, $\mathrm{F}_{\mathrm{RFEI}}$ is a mapping function from the signal space to the ID space referencing the RFEI algorithm, and $\mathbf{y} \in \mathbb{R}^{C\times 1}$ is the output vector of the algorithm containing the emitter ID information, where C is the number of emitters trained on the algorithm.

3.1. Signal Fingerprint Extraction

The SF can be defined as any subtle differences in the demodulation and decoding of the FH signal, which can uniquely identify the emitter ID. However, in this study, our objective was to identify the emitter ID before passing through the MAC layer. Thus, we targeted the analog SF that could pass the physical layer in the form of RT, SS, and FT signals. We represent them by

$$\mathbf{s}_{\mathrm{SF}} = \mathrm{g}_{\mathrm{SF}}(\mathbf{s}) \tag{9}$$

where g_{SF} is the extraction function of the SF, and $\mathbf{s}_{\mathrm{SF}} \in \mathbb{C}^{N_{\mathrm{SF}} \times 1}$ is the SF selected from a set of possible lists, that is, $\mathrm{SF} \in \{\mathrm{RT}, \mathrm{SS}, \mathrm{FT}\}$. Here, N_{SF} is the length of the SFs.

Based on the definition of the SF signal in [6], the RT signal is defined as an increasing RF signal that increases from the noise level to the designed level. The SS signal is defined as a region of the RF signal that contains a modulated signal with a designed energy level, and the FT signal is defined as an inverse case of the RT signal, decreasing the RF signal from the designed energy level to the noise level.

For accurate extraction, the extraction procedure is structured based on the energy variation of the SFs. For the windowed vector $\mathbf{s}_n = s[i + (n-1)/2 \times W_E : i + (n+1)/2 \times W_E]$ with the extraction window size W_E and its L_2 norm energy E_n, the detection rule for the transient signals can be expressed as follows

$$\begin{cases} E_n \geq (1+\delta) \times E_{n-1}; & T^{\mathrm{RT}} \leftarrow \begin{bmatrix} T^{\mathrm{RT}} & i \end{bmatrix} \\ E_n \leq (1-\delta) \times E_{n-1}; & T^{\mathrm{FT}} \leftarrow \begin{bmatrix} T^{\mathrm{FT}} & i \end{bmatrix} \end{cases} \tag{10}$$

where δ is the threshold value for detecting the energy variance and T^{RT} and T^{FT} are the detected time indices for the RT and FT signals, respectively.

A sliding window method is applied to monitor the energy variation of the incoming signal, which is then used to detect the RT and FT signals. The RT signal is detected as a signal in which the L_2-norm energy of the window is increased by 10% or more. The FT signal is defined as a decreasing case. After detecting the RT and FT signals, the SS signal can be defined as the signal between the RT and FT signals using the following definitions:

$$\begin{aligned} \mathbf{s}_{\mathrm{RT}} &= s\left[T^{\mathrm{RT}}[1] : T^{\mathrm{RT}}[-1]\right] \\ \mathbf{s}_{\mathrm{FT}} &= s\left[T^{\mathrm{FT}}[1] : T^{\mathrm{FT}}[-1]\right] \\ \mathbf{s}_{\mathrm{SS}} &= s\left[T^{\mathrm{RT}}[-1] : T^{\mathrm{FT}}[1]\right] \end{aligned} \tag{11}$$

The extraction results for the SFs are presented in Figure 4.

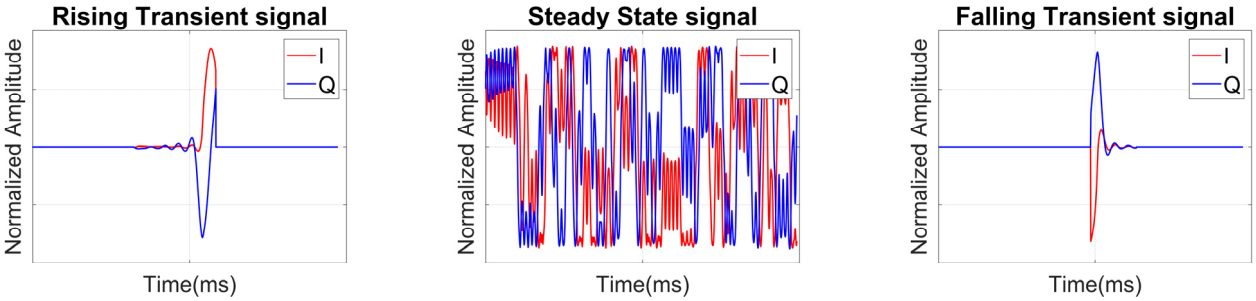


Figure 4. Examples of the SFs: (**a**) RT, (**b**) SS, and (**c**) FT signals.

3.2. Time–Frequency Feature Extraction

The next step is to design a feature from the SF. The purpose of this step is to transform the SF domain into a specific feature domain in which the physical measurements between different emitters could be well distinguished. In conventional approaches [4–6], the designed handcrafted features are calculated from signal characteristics of the SFs. In this

case, the goal is to obtain a feature domain that can ensure robust classification results. However, in more recent approaches [7,8], the purpose of this step is slightly modified. The SFs are transformed into domains that can express the signal characteristics of the SFs, and the identification of a feature domain that can ensure robust classification is entrusted to the classification step based on a deep learning-based classifier. The relevant procedure is expressed as follows

$$\mathbf{s}_{Feature} = \mathrm{q}_{\mathrm{SF}}(\mathbf{s}_{\mathrm{SF}}) \tag{12}$$

where q_{SF} is the transform function for the designed feature domain, $\mathbf{s}_{Feature} \in \mathbb{R}^{N^f_{\mathrm{SF}} \times N^t_{\mathrm{SF}}}$, where N^f_{SF} and N^t_{SF} are the sizes of the frequency and time indices, respectively, of the spectrogram transformed from the SF.

In this study, the time–frequency distribution of the FH signals, that is, the spectrogram, was analyzed. The spectrogram is a well-known time–frequency analysis method used to visualize the variation of the frequency components calculated from nonstationary signals [20]. The feature design strategy used in this study requires analysis of the power density behavior of the SFs in the time–frequency domain. The key idea of the FHSS system is that the carrier frequency of the FH signal hops within a predefined frequency range. Therefore, the signal characteristics must be implied in the distribution of the time–frequency domains.

A discrete-time short-time Fourier transform (STFT) is applied to compute the spectrogram of the SFs. With the sliding window $w[n]$ with a size of W_{STFT}, the STFT of the SFs can be calculated as follows

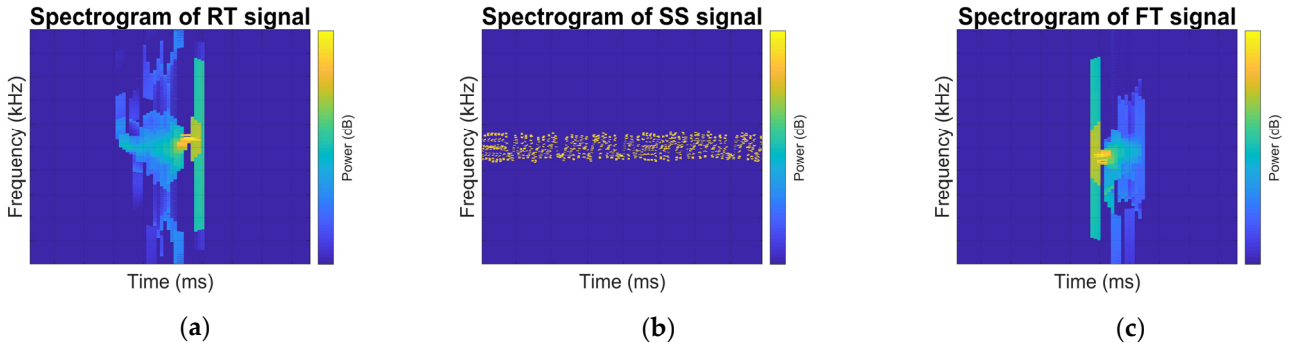
$$\mathrm{STFT}_{\mathbf{s}_{\mathrm{SF}}}[m,p] = \sum_{n=-N_{\mathrm{SF}}}^{N_{\mathrm{SF}}} s_{\mathrm{SF}}[n] w[n-m] e^{-j2\pi pm} \tag{13}$$

where $m = 1, 2, ..., K^t_{\mathrm{SF}}$ is the time sampling point along the time axis and $p = 1, 2, ..., K^f_{\mathrm{SF}}$ is the frequency sampling point along the frequency axis. We set N_{SF} as a sufficiently large value.

Next, the power density behavior of the spectrogram can be represented as the magnitude squared of the STFT such that

$$\mathrm{spectrogram}\{\mathbf{s}_{SF}\} = |\mathrm{STFT}_{\mathbf{s}_{SF}}[m,p]|^2. \tag{14}$$

The spectrogram results are presented in Figure 5.



(**a**) (**b**) (**c**)

Figure 5. Examples of the spectrograms: (**a**) RT, (**b**) SS, and (**c**) FT signals.

3.3. User Emitter Classification

The third step is to identify the emitter ID from the designed feature. The goal is to design a classification algorithm that can learn spectrograms for robust classification results. Owing to recent research in the field of deep learning, deep neural networks are well known for their abilities to extract spatial or temporal features with nonlinear computational

capabilities [21]. Thus, we aimed to construct a deep learning-based classifier to train the spectrogram of the SFs. The classification process can be obtained using

$$\mathbf{y} = \mathrm{f}_{Classifier}(\mathbf{s}_{Feature}) \tag{15}$$

where $\mathrm{f}_{Classifier}$ is the deep learning-based classification algorithm, and the output vector $\mathbf{y}$ implies the emitter ID information k.

3.3.1. Base Classifier: Deep Inception Network Classifier

There are two main blocks to construct the custom deep learning-based classifier: a residual block [22] and an inception block [23]. The residual block is designed to enable flexible training as the depth of the network increases. In the case of the inception block, the main purpose is to filter out input features with different receptive field sizes. Details of the architecture and design strategies of the main blocks are described in Appendix A.

The spectrogram consists of physical measurements calculated from the SF signals. It represents the power densities of the SFs along the time–frequency axes. Thus, the subtle differences exhibited by the SFs can be anywhere on the time–frequency axes of the spectrogram, and the size of the features can be varied. To train these SFs, we aimed to filter the spectrogram on multiple scales in the temporal and spatial domains by applying inception blocks to construct a custom deep learning classifier.

We utilized the inception-A and reduction-A blocks to construct the base classifier: the DIN classifier. The inception-A and reduction-A blocks are the basic blocks for constructing the Inception-v4 models [24]. The role of the inception-A block is to filter the input features with multiple receptive field sizes and concatenate them as the filter axis, thereby expanding its dimensions. The role of the reduction-A block is to downsize the feature map on the grid side, that is, the time–frequency axes of the spectrogram. It can effectively manage the number of weights inside the classifier, similar to the pooling layer.

We adopted the inception-A and reduction-A blocks, as shown in Figure 6. The structures of the blocks are the same as defined in [24]. However, the filter sizes N_F of the sublayers were set to 32 and 64, adjusted by the experiments. Batch normalization [25] and rectified linear unit activation units were applied immediately after every convolutional layer. The inception-A block was applied twice to expand the filter axis, and the reduction-A block was applied once to re-size the feature map on the grid axis. We applied these block sequences twice, adjusted by heuristic experiments. The total structure of the DIN classifier is provided in Table 1.

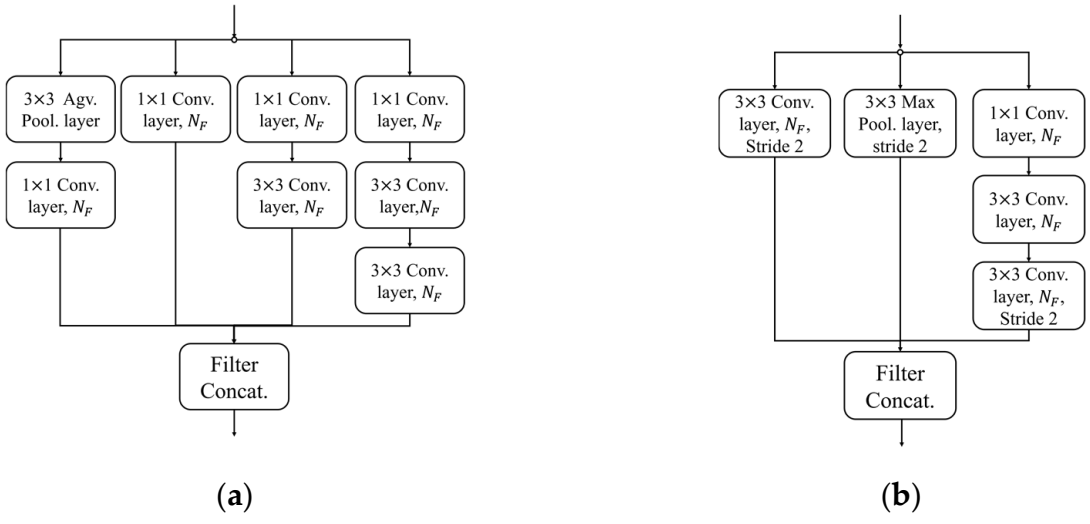


Figure 6. Basic block units used to construct the DIN: (**a**) the Inception-A block in [24] and (**b**) the Reduction-A block in [24].

Table 1. Structure of the base classifier: the DIN classifier.

Type	Filter Size/Stride /Padding	Output Shape (for the SS Input)
Input signal	-	$205 \times 340 \times 1$
Conv_1	$3 \times 3/2/0$	$102 \times 169 \times 32$
Conv_2	$3 \times 3/1/0$	$100 \times 167 \times 32$
Conv_3	$3 \times 3/1/1$	$100 \times 167 \times 32$
Max. pool	$3 \times 3/2/0$	$49 \times 83 \times 32$
$2 \times$ inception	Inception-A [$N_F = 32$]	$49 \times 83 \times 128$
$1 \times$ reduction	Reduction-A [$N_F = 32$]	$24 \times 41 \times 192$
$2 \times$ inception	Inception-A [$N_F = 64$]	$24 \times 41 \times 256$
$1 \times$ reduction	Reduction-A [$N_F = 64$]	$20 \times 11 \times 384$
Avg. pool	Adaptive avg. pooling	$1 \times 1 \times 384$
Linear	Logits	$1 \times 1 \times 7$

Finally, we obtained the deep learning classification framework, as in Equation (15). From the M training samples in $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_M]$ and output samples $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_M]$, the cross-entropy loss, was applied such that loss function can be expressed as follows

$$\text{loss} = -(1/M)\sum_{i=1}^{M} \log\left(e^{\mathbf{y}_i[c_k]} / \sum_{j=1}^{C} \mathrm{e}^{\mathbf{y}_i[c_j]}\right) \tag{16}$$

where c_k is the true label of sample $\mathbf{y}_i$ with the kth emitter ID, $\mathbf{y}_i[c_j]$ is the jth element of output sample $\mathbf{y}_i$. Based on the cross-entropy losses, the Adam optimizer [26] is utilized to update the weights of our DIN classifier.

After finishing the training of the DIN classifier, the emitter ID of input sample $\mathbf{y}_i$ can be estimated as follows

$$\begin{aligned} p(c_l; \mathbf{s}_{\mathrm{SF}}) &= \mathrm{softmax}(\mathbf{y})_{c_l} \\ &= \frac{e^{\mathbf{y}[c_l]}}{\sum_{j=1}^{C} \mathrm{e}^{\mathbf{y}[c_j]}} \end{aligned} \tag{17}$$

$$\begin{aligned} \widetilde{k} &= \underset{c_j \in C}{\mathrm{argmax}}\left(p\left(c_j; \mathbf{s}_{\mathrm{SF}}\right)\right) \\ &= \underset{c_j \in C}{\mathrm{argmax}}\left(\mathrm{softmax}(\mathbf{y})_{c_j}\right) \end{aligned} \tag{18}$$

where $p(c_l; \mathbf{s}_{i,\mathrm{SF}})$ is the probability that the emitter ID of the input sample is c_l, which can be defined as the softmax output of sample $\mathbf{y}_i$. In this probability, the estimated emitter ID $\widetilde{k}$ is defined as the maximum probability that input samples will be included in a particular emitter ID c_j (see the Equation (18)).

3.3.2. Ensemble Approach for Multimodal Signal Fingerprints

The ensemble approach is a well-known method that ensures better generalization performance of classification models [27]. It combines the results of multiple base classifiers trained on the same training dataset and makes a final decision based on these results. Stacking is a combined method that uses the final model to combine the outputs of the base model [27]. It is useful when multimodal features are present in applications such as video signal processing where audio, video, and text segments exist simultaneously [28].

It was reported that multiple SFs, that is, the RT, SS, and FT signals, can be considered as multimodal features for an accurate RF fingerprinting model [6]. To utilize the multi-modality features of the SFs, we adapted the stacking ensemble approach to the DIN model as presented in Figure 7. The SFs $\mathbf{s}_{\mathrm{SF}}$ were extracted from hop signal $\mathbf{s}$ in Equation (10). These SFs can act as independent features for emitter identification. Thus, each of the SFs,

i.e., RT, SS, and FT, is assumed to be independent of the others. For the ensemble approach, the probability that the emitter ID is c_l can be defined as follows

$$p(c_l; \mathbf{s}) = \prod_{\mathrm{SF} \in \{\mathrm{RT,SS,FT}\}} p\left(c_j; \mathbf{s}_{\mathrm{SF}}\right). \tag{19}$$

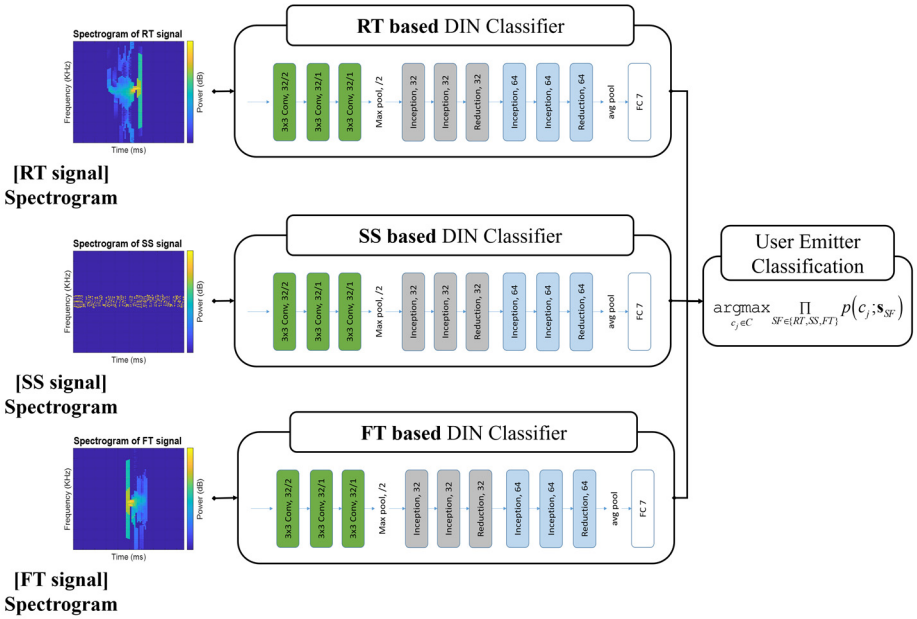


Figure 7. Stacking ensemble approach for the multimodal SF signals.

According to the DIN classifier trained on the RT, FT, and SS signals presented in Section 3.3.1, the final decision was performed by a linear combination of each base classifier (i.e., DIN classifier) such that

$$\begin{aligned} \widetilde{k} &= \underset{c_j \in C}{\mathrm{argmax}} p\left(c_j; \mathbf{s}\right) \\ &= \underset{c_j \in C}{\mathrm{argmax}} \prod_{\mathrm{SF} \in \{\mathrm{RT,SS,FT}\}} p\left(c_j; \mathbf{s}_{\mathrm{SF}}\right) \\ &= \underset{c_j \in C}{\mathrm{argmax}} \prod_{\mathrm{SF} \in \{\mathrm{RT,SS,FT}\}} \mathrm{softmax}\left(\mathbf{y}_{\mathrm{SF}}\right)_{c_j} \end{aligned} \tag{20}$$

3.4. Attacker Emitter Detection

The last step of the RFEI method is an outlier detection step implemented to detect the imitated FH signal. An outlier is a sample included in specific emitter IDs that is not considered during training. In this study, the imitated FH signal was the outlier. This step is aimed at detecting the differences in the classifier output characteristics between the outputs of the classifier when the trained and outlier samples are input. This objective can be achieved by comparing the classifier outputs [29–31], exposing the outliers during the training step to magnify the differences between the trained and outlier samples [32,33], and analyzing the likelihood of the inputs from a generative adversarial network [34,35].

The proposed outlier detection scheme is presented in Figure 8. We considered the outlier detection framework proposed in [30]. Temperature scaling [36] and the opposite application of an adversarial attack [37] have been reported to be effective in detecting outlier samples. After preprocessing the input sample, outliers can be detected when the maximum probability of the output vector is lower than the threshold. The key idea of this approach is that the output vector of the outlier represents a much smaller value than the output vector of the trained sample.

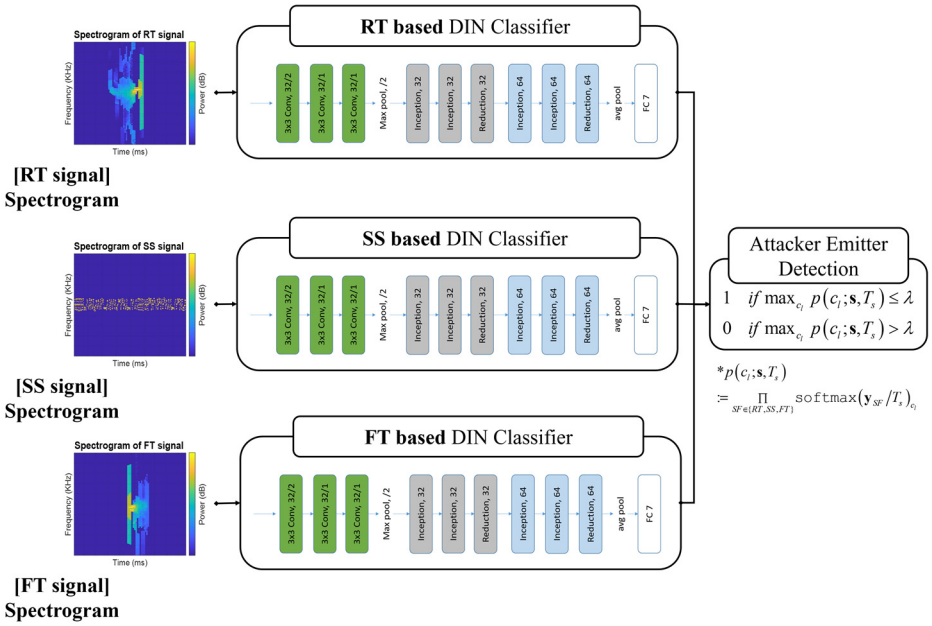


Figure 8. Attacker detection scheme based on stacking ensemble approach.

Utilizing this approach, we constructed the outlier detector to alert the signal input when the imitated FH signal was input by performing two steps: (1) calibration of the output vector of the classifier by a temporal scaling factor, T_s, and (2) comparison of the maximum probability of the output vector to the outlier detection threshold, λ. In this study, opposite application of the adversarial attack was not performed because a small perturbation of the input sample may affect the SFs, defined as subtle differences in the FH signal.

Mathematically, the temporal scaling process was applied to Equation (17) such that

$$\begin{aligned} p(c_l; \mathbf{s}_{\mathrm{SF}}, T_s) &= \mathrm{softmax}(\mathbf{y}/T_s) \\ &= \frac{\exp(\mathbf{y}[c_l]/T_s)}{\sum_{j=1}^{C} \exp(\mathbf{y}[c_j]/T_s))} \end{aligned} \tag{21}$$

In the case of the ensemble approach, the probability in Equation (19) was modified as the temporal scaled version as follows

$$\begin{aligned} p(c_l; \mathbf{s}, T_s) &= \prod_{\mathrm{SF} \in \{\mathrm{RT,SS,FT}\}} p(c_l; \mathbf{s}_{\mathrm{SF}}, T_s) \\ &= \prod_{\mathrm{SF} \in \{\mathrm{RT,SS,FT}\}} \mathrm{softmax}(\mathbf{y}_{\mathrm{SF}}/T_s)_{c_l} \end{aligned} \tag{22}$$

Based on the scaled output probability, the detection rule for the outlier sample can be defined as follows

$$p(c_{out}; \mathbf{s}, T_s, \lambda) := \begin{cases} 1 & if \max_{c_l} p(c_l; \mathbf{s}, T_s) \leq \lambda \\ 0 & if \max_{c_l} p(c_l; \mathbf{s}, T_s) > \lambda \end{cases} \tag{23}$$

where $p(c_{out}; \mathbf{s}, T_s, \lambda)$ is the probability that the current input sample is an outlier. This detection rule is a binary classifier with trained class c_{train} and outlier class c_{out}. Thus, parameters T_s and λ were optimized experimentally based on the minimum false positive rate (i.e., the part of the actual outliers that were misdetected as trained samples, FPR) when the true positive rate (i.e., the part of the actual trained samples that were detected as trained samples, TPR) was higher than 95%.

The final version of the algorithm used for our proposed RFEI process is presented in Algorithm 2.

Algorithm 2. Proposed RFEI algorithm.

Input: The target baseband hop signal $s_h^k(t)$
Initialize: $i = 1$, $T^{\mathrm{RT}} = T^{\mathrm{FT}} = \{\}$ for time periods, W_E and bandwidth of interest (BOI) BW_{BOI}.

Step 1: (Extract the target SF)
while do:

Detect the transient signal with Equation (10).
Extract the target SF $\mathbf{s}_{\mathrm{SF}}$ with Equation (11).
Set $i \leftarrow i + 0.5 \times W_E$

Step 2: (Calculate the spectrogram)

Calculate the spectrogram $\mathbf{s}_{Feature}$ of the SF with Equation (13) with respect to the BOI, BW_{BOI}.

Step 3: (Perform emitter classification) $i + W_{Ext.} < length(\mathbf{s})$

Estimate the emitter IDs from the decision rule using either the base classifier (18) or ensemble approaches in Equation (20).

Step 4: (Perform outlier detection)

Scale the output vector for temporal scaling factor T_s with Equation (22) and detect the outliers with Equation (23)

Output: Return the authenticated baseband hop signals $s_h^k(t)$

4. Baseline Algorithms for RF Fingerprinting Method

In this study, for performance comparison, three other baseline methods were carefully designed and implemented based on algorithms from the literature [4,5,7,8].

Before describing the details, we note that the signal preprocessing steps, such as preamble extraction [5] and signal difference calculation after signal decoding [7], are not covered in this study. The goal of this study was to identify the emitter ID in the physical layer of the FHSS network. Therefore, we focused on analog SFs that can be obtained from the physical layer of the system. To this end, all baseline SFs were set to RT, SS, and FT, and the feature extraction and classification processes were designed to reflect the approaches in the literature.

4.1. Baseline 1: Statistical Moments Based RF Fingerprinting

The first baseline aims to reflect the conventional RF fingerprinting approaches based on handcrafted features. It was designed for statistical moments of the SFs, similar to that in [4,5].

The SF extraction process was the same as that of the proposed method described in Section 3.1.

For feature extraction, the SFs were segmented using $N_{seg.}$. Because the RT and FT signals were too short to be segmented, segmentation was applied only to the SS signal.

$$\mathbf{s}_{\mathrm{SF}} = \left[\mathbf{s}_{\mathrm{SF}|1}, \mathbf{s}_{\mathrm{SF}|2}, \ldots, \mathbf{s}_{\mathrm{SF}|N_{seg}}\right] \tag{24}$$

where $\mathbf{s}_{\mathrm{SF}|n}$ is the nth segment of SF. For each segmented SF, a total of six sub-features were considered. The instantaneous amplitude, phase, and frequency, described in [5], were calculated as sub-features, and the time, frequency, and time–frequency axes of the spectrogram, identified as good features in [4], were applied as sub-features. Subsequently, the statistical moments (i.e., mean m, variance σ^2, skewness γ, and kurtosis κ) and entropy H were calculated for each sub-feature. Thus, a total of 30 features were calculated and arranged in a vector form such that

$$\mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|n}} = \left[\left(m, \sigma^2, \gamma, \kappa, H\right)_1, \left(m, \sigma^2, \gamma, \kappa, H\right)_2, \ldots, \left(m, \sigma^2, \gamma, \kappa, H\right)_6\right] \tag{25}$$

where $\mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|n}} \in \mathbb{R}^{1\times 30}$ is the vector form of the handcrafted features calculated from the nth segments of the SF. Finally, the composite handcrafted feature $\mathbf{s}_{Feature} \in \mathbb{R}^{N_{\mathrm{SF}}^{stats}\times 1}$ can be defined as follows

$$\mathbf{s}_{Feature} = [\mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|1}}, \mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|2}}, \ldots, \mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|N_{seg.}}}] \tag{26}$$

where N_{SF}^{stats} was the size of the statistic moments vector.

For classification, a linear SVM from [4] was applied. Random forest or multi-class AdaBoost from [5] and linear discriminant analysis from [4] were also investigated. We compared these algorithms when applied to our FH signal dataset, and the linear SVM showed the best classification results.

4.2. Baseline 2: Raw Signal-Based RF Fingerprinting

The second baseline aims to reflect the recent methods of RF fingerprinting based on raw signal processing. It was designed to train raw SF signals directly in the ensemble approaches of the deep learning classifiers described in [7].

As described at the beginning of Section 4, the SF extraction process was the same as that of the proposed method described in Section 3.1.

For feature extraction, the SFs were segmented using $N_{seg.}$ in Equation (24). The core idea of this approach was to train the raw signals in the ensemble classifiers, and the RT and FT were also segmented in this case. The feature vectors of each segment were set to a two-channel I/Q vector $\mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|n}} \in \mathbb{R}^{N_{\mathrm{SF}}^{raw}\times 2}$ such that

$$\mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|n}} = \begin{bmatrix} \mathrm{Re}\left(\mathbf{s}_{\mathrm{SF}|n}\right) \\ \mathrm{Im}\left(\mathbf{s}_{\mathrm{SF}|n}\right) \end{bmatrix} \tag{27}$$

where N_{SF}^{raw} is the size of each segment $\mathbf{s}_{\mathrm{SF}|n}$.

For the ensemble classification approach, the base classifier was set to a one-dimensional CNN as an identification network for outdoor data in [4]. After training each base classifier using segmented feature $\mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|n}}$, classification was performed using an ensemble approach, as in [7]

$$\widetilde{k} = \underset{c_j \in C}{\mathrm{argmax}} \prod_{n \in N_{seg.}} p\left(c_j; \mathbf{s}_{Feature|\mathbf{s}_{\mathrm{SF}|n}}\right) \tag{28}$$

4.3. Baseline 3: Spectrogram-Based RF Fingerprinting

The third baseline aims to reflect the recent approach in [8], which is based on the SF spectrogram. As described in [8], the author trained the Hilbert spectrum of the received hop signal in a residual unit-based deep learning classifier. To reflect this approach in baseline 3, the algorithm was designed to train an SF spectrogram directly in the residual-based deep learning classifier.

The SF extraction and feature extraction processes were the same as those of the proposed method described in Sections 3.1 and 3.2.

For classification, the classifier structure was set to the residual-based deep learning classifier described in [8]. After training the classifier, classification was performed using Equation (18).

5. Experimental Results and Discussion

This section describes the experimental investigation of the emitter identification performance of the proposed RF fingerprinting method. Before discussing the results, several experimental setups are discussed.

A custom DA system was set up for our experiments, as shown in Figure 9. The DA system consisted of a high-speed digitizer and a Raid-0 configuration with six SSD disk drives. The digitizer, PX14400, supports sampling rates of up to 400 MHz with a 14-bit

analog-to-digital converter resolution, resulting in a streaming rate of 0.7 GB/s for real-time data acquisition. With write speeds of up to 1.6 GB/s in our Raid-0 configuration, the DA system can acquire data in real-time streaming.

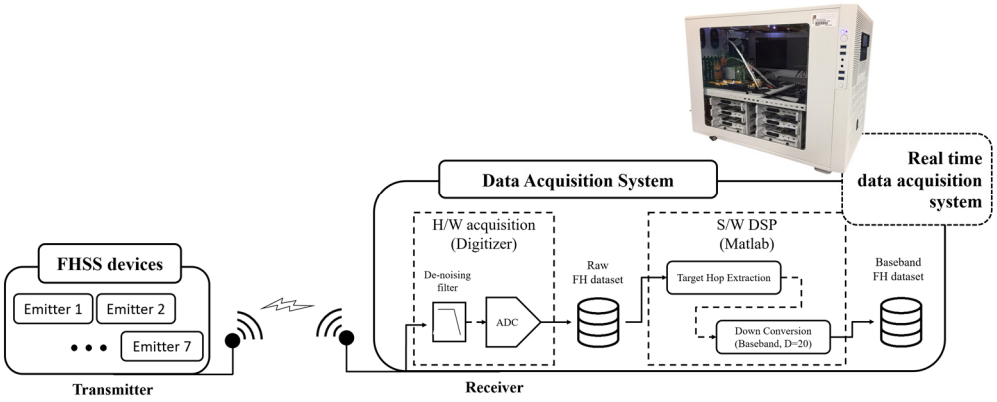


Figure 9. Custom-made data acquisition (DA) system.

We collected FH signals from a real experiment to determine the reliability of the algorithm. Seven FHSS devices were used to experiment. Each device utilized the same hopping rate for secure voice communication. The FH signal was frequency-modulated, and the carrier frequency was set to hops in the very high frequency range. The exact hopping rate and frequency range will not be disclosed owing to security issues. The FHSS device was connected under laboratory environmental conditions. The FH signal was acquired at a 400 MHz sampling rate and stored as raw FH data in the DA system.

Target hop extraction and down-conversion were performed on the stored raw training FH data. Because we assumed the predefined hopping pattern to be known, an energy detection approach was applied to the exact hopping frequency f_h^k and the target hop samples $\mathbf{x}_h^k$ were extracted from the observed RF signal $\mathbf{y}$. Subsequently, the hop sample was down-converted to the baseband using a decimation factor of 20, i.e., 20M sample rate baseband hop signals $\mathbf{s}_h^k$ were acquired. These were stored as baseband FH training data in the DA system. This down-conversion approach is reasonable because the FH signals were also demodulated to the IF or baseband to decode the digital data modulated by the message signal $m^k(t)$ as in Equation (2). As the SFs depend on the component characteristics of the emitter, the SFs also should exist in the baseband hop signal, $\mathbf{s}_h^k$.

Another set of FH signals was acquired to prepare an outlier dataset. Two more FHSS devices were recruited, and the FH signals were acquired on different dates compared with those of the training dataset. The emitter specifications were the same as those of the training emitter. However, in this experiment, the FH signal was down-converted to baseband and stored as outlier FH data with a sampling rate of 2.34 MHz. For fair comparison, the sampling rate of the signal was resampled using the Fourier-domain based sampling rate conversion method, which can improve the accuracy and computational cost compared to the time domain-based method [38]. These outlier data were considered only in the outlier detection experiment described in Section 5.5.

An average of 168 hop FH signals were obtained for each training emitter, and an average of 310 hop FH signals were obtained for each outlier emitter; a total of 1796 samples from nine emitters were obtained. The details are presented in Table 2.

The results were obtained using the experimental setup as follows. For the training and testing datasets, the FH dataset was partitioned according to a 7:3 ratio; a total of 823 samples were trained, and a total of 353 samples were tested from seven emitters. In the outlier detection experiment, the test dataset for training emitters and the outlier dataset for outlier emitters were considered; a total of 353 samples from seven training emitters were tested, and a total of 620 outlier samples from two outlier emitters were tested. All the results were tested 10 times, and the average performance was presented.

The experiments were conducted with an Intel i7-6850K CPU unit and an NVIDIA Titan RTX GPU unit. The dataset generation task in Figure 9 was performed using MATLAB 2018a, and all RF fingerprinting algorithms were implemented in Python 3.6 with PyTorch 1.6.0. The other implemented parameters of the experiments are described in Appendix B.

Table 2. Details of the FH dataset.

Dataset	Emitters	Emitter Type	Number of Acquisitions	Number of Samples
Training dataset	Emitter 1	Model 1	5 times	170
	Emitter 2	Model 1		168
	Emitter 3	Model 1		170
	Emitter 4	Model 1		171
	Emitter 5	Model 2		160
	Emitter 6	Model 2		169
	Emitter 7	Model 2		168
Outlier dataset	Emitter 8	Model 3	10 times	308
	Emitter 9	Model 3		312
Total emitters	9		Total samples	1796

5.1. Emitter Identification Accuracy

We firstly investigated the emitter identification performance of the proposed RFEI algorithm and the baselines. All algorithms were applied to all SFs, and the mean and standard deviation of the experimental values were investigated. The results are listed in Table 3.

Table 3. Emitter identification accuracy.

	RT	SS	FT
	Mean Accuracy (%) ± Standard Deviation		
Statistical moments *	61.8 ± 0.0	92.6 ± 0.0	66.4 ± 0.0
Raw signal **	17.7 ± 1.3	89.5 ± 0.7	20.4 ± 2.1
Spectrogram—residual ***	83.7 ± 2.1	93.7 ± 1.2	93.9 ± 1.2
Spectrogram—DIN †	84.6 ± 1.5	95.3 ± 1.2	92.8 ± 1.1
Ensembles †		97.0 ± 0.6	

*: (Baseline 1) statistical moments approach in [4,5]. **: (Baseline 2) raw signal approach in [7]. ***: (Baseline 3) spectrogram and residual block-based approach in [8]. †: (Proposed) spectrogram, DIN classifier, and ensemble-based approach in the proposed method.

Table 3 demonstrates the efficiency of the proposed RFEI algorithm showing that the proposed algorithm for identifying the emitter ID based on the SS spectrogram and DIN base classifier performs with an accuracy of 95.3%, which is better than other baseline algorithms. In addition, the ensemble approach of RT, FT, and SS based on the proposed algorithm yielded an accuracy of 97.0%, demonstrating its efficiency with a higher identification accuracy than other baseline algorithms.

In terms of the SF efficiency, the results show that the SS signal is the most effective SF, as it is more accurate than the RT and FT signal-based results. In addition, in terms of the efficiencies of the feature extraction and classification approaches, the spectrogram feature is effective for representing the differences in the SF for each emitter. The most effective means of identifying the emitter ID in the FH signals is to ensemble the multimodal SFs, i.e., the RT, FT, and SS, trained by a DIN.

The emitter identification performance at SNRs is shown in Figure 10. The AWGN signal $n(t)$ can be artificially added to the received hop signal $\mathbf{s}$ as follows

$$\mathrm{SNR} = 10\log_{10}\left(\frac{||\mathbf{s}||_2^2}{N\sigma_n^2}\right) \tag{29}$$

where N and σ_n^2 are the length and variance of the noise signal $n(t)$, respectively.

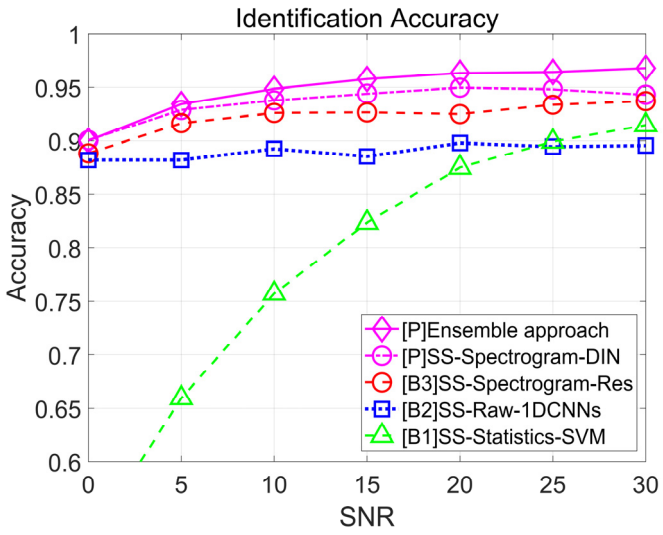


Figure 10. Emitter identification accuracy at different signal-to-noise ratios (SNRs).

We found that the classification accuracy obtained by applying the proposed method to the SS signal was nearly 3% above baselines 1 and 2 and at least 1% above baseline 3 over the entire range of SNRs. In addition, the ensemble approach of the proposed algorithm can improve the accuracy by more than 1% compared to the proposed method. In particular, applying the proposed method to the SS signal at 20 dB SNR, which is the typical operating SNR of the FHSS network [39], yielded an accuracy of more than 95.0%. For the ensemble approach, the identification accuracy was measured to be greater than 96.4%, making it the most effective algorithm. These accuracies are higher than those of baseline 1 (87.5%), baseline 2 (89.8%), and baseline 3 (92.5%).

The validity of the proposed algorithm was verified again. At low SNR, the accuracy of baseline 1 decreases dramatically, whereas the other algorithms maintain their accuracies. Baseline 2, baseline 3, and the proposed method work well when applied to the SS signal, even at low SNRs. However, the proposed method outperforms the baselines, and the ensemble approaches outperform the other algorithms at all SNRs. These findings imply that a deep learning-based classifier at baselines 2 and 3 can learn the differences in the SFs for RF fingerprinting, but our proposed algorithm (i.e., using the spectrogram and DIN classifier) with the ensemble approach is more effective than the baselines.

The confusion matrix of the ensemble approach based on the proposed method is presented in Table 4. The confusion matrix is a specific metric for a classifier that can represent the relationship of each emitter. This matrix can be obtained by simply counting the results of the test samples with their true label information. The rows of the matrix indicate the true emitter IDs, and the columns indicate the predicted emitter IDs. The diagonal terms in the confusion matrix represent the correct classification result cases, and the off-diagonal terms represent the incorrect classification result cases. Thus, Table 4 shows that our ensemble approach based on the proposed method can identify the FH emitters with more than 94.6% accuracy without confusion between emitters.

5.2. Efficiency of the Inception Blocks

We constructed the DIN classifier based on the inception blocks. To confirm the efficiency of the inception blocks, the identification accuracy of the proposed method was compared with that of baseline 3. The difference between the proposed method and baseline 3 lies in the classifier. As in baseline 3, the classifier was set to the residual-based classifier described in [8]. Two experiments were performed for comparison. One was conducted to identify the emitter ID from the received hop signal **s** without the SF extraction, and the

other was performed to identify the emitter ID from the ensemble approach of the SFs. The results are presented in Table 5 and Figure 11.

Table 4. Averaged confusion matrix of the ensemble approach based proposed method.

		Predicted Emitter (%)						
		1	**2**	**3**	**4**	**5**	**6**	**7**
Actual Emitter (%)	1	100.0	0	0	0	0	0	0
	2	0.2	98.6	0	0.2	0.4	0	0.6
	3	0	0	98.0	0.2	0	1.8	0
	4	0	1.6	0.6	95.5	0.6	0.4	1.4
	5	0	0.2	1.9	0.4	96.0	1.0	0.4
	6	0	0	2.6	0	1.0	95.8	0.6
	7	0.6	1.0	0.4	2.8	0.6	0	94.6

Table 5. Identification accuracies of the residual and inception blocks.

	Hop Signal without SF Extraction	**Ensemble Approach with SF Extraction**
	Mean Accuracy (%) ± Standard Deviation	
Spectrogram—Residual ***	94.4 ± 1.1	96.4 ± 0.7
Spectrogram—DIN †	95.1 ± 1.0	97.0 ± 0.6

***: (Baseline 3) spectrogram approaches in [8]. †: (Proposed) spectrogram approach of SF.

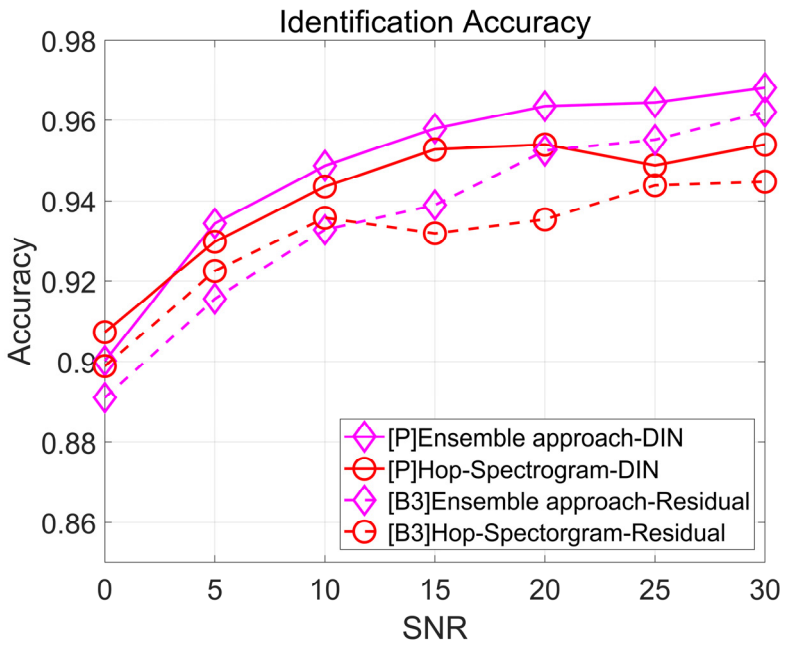


Figure 11. Identification accuracies of the residual and inception blocks at different SNRs.

Table 5 presents the identification accuracies of the proposed algorithm and baseline 3. The identification accuracy results at different SNRs are presented in Figure 11. Both sets of results demonstrate the efficiency of the inception blocks. Table 5 reveals that the DIN-based approach can produce higher accuracies than the residual-based approach. This result is also shown in Figure 11. As the SNR changes, the accuracy of the DIN-based approach is superior to that of the residual block-based approach, except when the ensemble approach of the residual-based method overcomes the hop and DIN-based method in environments with SNRs of 20 dB or more. However, if we focused on the classifier structure, i.e., compared the performance between hops approaches or ensemble approaches, the performance of the residual network could not overcome the performance

of the inception blocks. As described in Section 3.3.1, this result may stem from the fact that filtering features with different receptive field sizes can help train SFs within deep learning architectures.

5.3. Class Activation Map (CAM) Analysis of the DIN Classifier

We investigated the feature map of the DIN classifier to understand why the DIN-based model works well. To this end, we applied a gradient-weighted CAM (GCAM) to visualize the feature map. The GCAM is a well-known feature visualization that identifies parts of the input signal that positively influence the class decision [40]. This can be achieved by back-propagating the gradient of the inference to the input layer and highlighting the input parts using positive gradient values. The details of the GCAM are described in Appendix C.

The average GCAM (AGCAM) results are presented in Figure 12. Interestingly, for each emitter classification, we found that the activated region of the AGCAM is the location at which the head and tail of the signal are located. The GCAM of the positive sample with an inference score of 0.99 or higher is shown in Figure 12b. These results show that when the classifier model correctly identifies the emitter ID, the filter maps of the model are activated similarly to the AGCAM of the target emitter. In other words, the intensity of the activated region differs from that of the AGCAM, but the shape and location of the activated region are similar to those in the AGCAM results. Conversely, the GCAM of the negative sample with an inference score of 0.30 or less is shown in Figure 12c. The results demonstrate that when the model misidentifies the emitter ID, the activated region of the filter maps is completely different from those of the target emitter and other emitters.

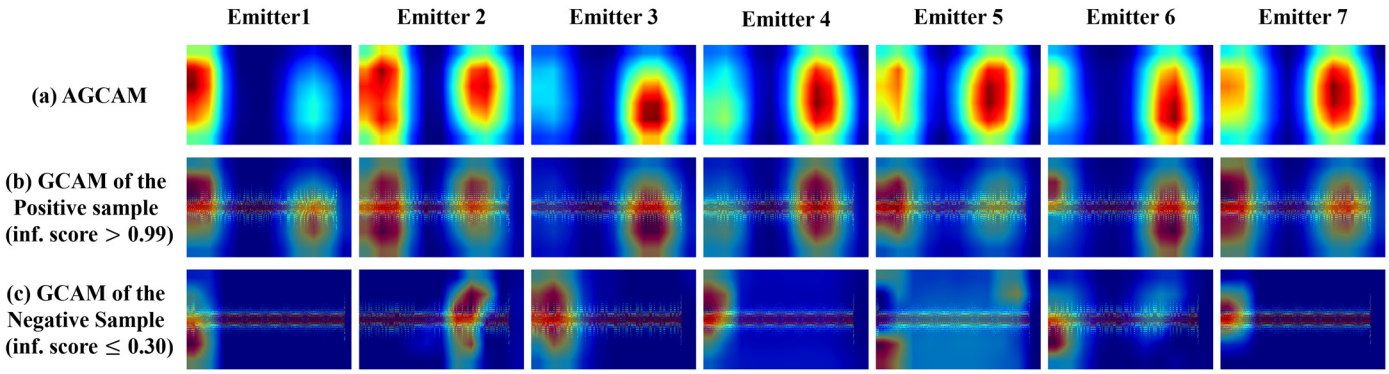


Figure 12. Examples of GCAM of the DIN classifier: (**a**) AGCAM for target emitters, (**b**) positive sample with an inference score greater than 0.99, and (**c**) negative sample with an inference score less than 0.30.

To verify the meaning of the activated region in Figure 12, the physical layer convergence protocol (PLCP) frame format for the FHSS network as defined in the 802.11 standard [10] is presented in Figure 13. It can be verified that the preamble field is located at the head part of the frame, and the frame body is located at the tail part of the frame.

The preamble is a sequential signal for synchronization between the transmitter and receiver. Therefore, duplicated sync sequences must be transmitted repeatedly. When the data sequence contained in the frame is identical for each emitter, only the differences in SFs remain, which is helpful in RF fingerprinting. Consequently, many researchers have applied additional preprocessing steps to extract preamble signals [5,6]. However, the proposed method based on the DIN classifier can automatically learn the preamble field without additional preprocessing steps.

In the cases of the frame body, the GCAM is activated in this region because the FH signal dataset is collected in a laboratory environment; hence, the data sequences contained in the frame body are similar to each other. This similarity of the data sequences can help identify the differences in the SFs of the emitters. Again, the proposed method can automatically learn the fields in which the emitter IDs can be identified.

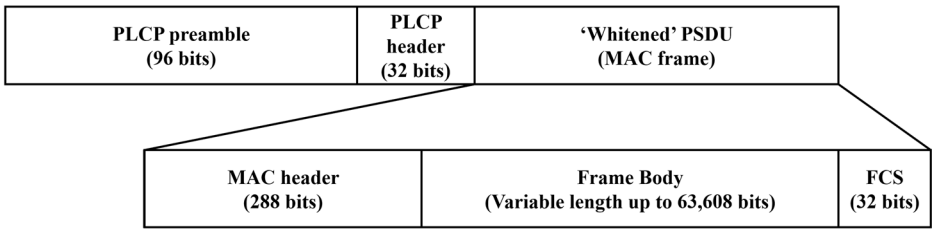


Figure 13. PLCP frame format for FHSS networks in the 802.11 standard [10].

5.4. Outlier Detection Performance

We evaluated the outlier detection performance of the proposed algorithm and baseline 3. The experimental dataset was prepared using the test dataset for trained emitters and the outlier dataset for outlier emitters. Before executing the experiment, the detector-related parameters, that is, the temporal scaling factor T_s and detection threshold λ, were optimized. For $T_s \in [1, 2, 3, 4, 5, 10, 15, 20]$ and $0 \leq \lambda \leq 0.3$, the parameters were set to $T_s = 2$ and $\lambda = 0.07$ for the proposed algorithm and $T_s = 1$ and $\lambda = 0.05$ for proposed baseline 3. These values were selected by finding the minimum FPR when the TRP was higher than 95%.

As discussed in Section 3.4, the key idea of the outlier detector is based on the fact that the maximum probability of the output vectors from the outlier samples has a smaller value than the maximum probability of the output vectors from the trained samples. To verify this idea, we plotted a histogram in Figure 14 showing the maximum probabilities of the output vectors obtained from the proposed method, i.e., the output vectors of the ensemble approach in the DIN. Evidently, the maximum probability values of the outlier samples occur at positions < 0.1. Conversely, the values of trained samples mostly exist at position ≥ 0.2. These results demonstrate that the differences between the characteristics of the outliers and trained samples are easily identified and can be utilized to detect the outlier samples.

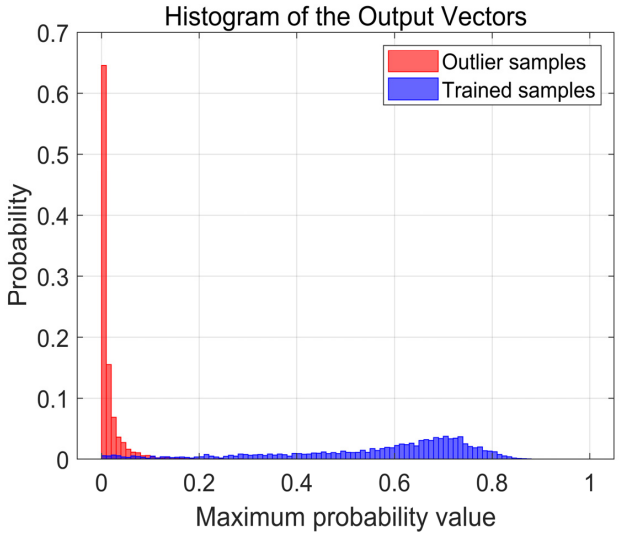


Figure 14. Histogram of the output vectors.

We present the confusion matrices of the outlier detectors based on the proposed method and baseline 3 in Tables 6 and 7. As we optimized our parameters based on the FPR values when the TPR was higher than 95.0%, both TPRs yielded similar rates in the detection of the actual trained samples. However, in the case of the true negative ratio, which represents the actual outlier sample detection ability, the proposed method can achieve a rate of 95.6%, which is 6.6% higher than that of baseline 3 (89.0%). In other words, the proposed method can reduce the FPR from 11.0% to 4.4%. These results indicate that

the DIN classifier-based approach is useful for training SF features in FH signals and can effectively detect outlier samples by using these trained features.

Table 6. Averaged confusion matrix of the outlier detectors based on the proposed method.

		Predicted Emitter (%)	
		Learned Classes	**Outlier Classes**
Actual emitter (%)	Learned classes	96.6	3.4
	Outlier classes	4.4	95.6

Table 7. Averaged confusion matrix of the outlier detectors based on baseline 3.

		Predicted Emitter (%)	
		Learned Classes	**Outlier Classes**
Actual emitter (%)	Learned classes	96.8	3.2
	Outlier classes	11.0	89.0

Figure 15 plots the ROC curve and compares the AUROCs. As was done for the previously presented results in Section 5, the values were averaged over 10 experiments. The ROC metric describes the relationship between the probability of detection (i.e., TPR) and the probability of a false alarm (i.e., FPR). This result can be achieved by plotting the FPR together with the TPR at different detector thresholds λ. Additionally, this ROC metric is known as the cost–benefit relationship in decision theory. Thus, when a high benefit is obtained at a low cost, i.e., when the probability of false alarms is low, high detection rates should be obtained. In other words, if the curve moves toward the upper left with a high AUROC, the model possesses strong detection ability. The results confirm that the proposed method can clearly improve the ROC curve compared with baseline 3. The AUROC also improves from 0.97 to 0.99. These results provide clear evidence that the proposed DIN-based ensemble method is more effective than the residual block-based method.

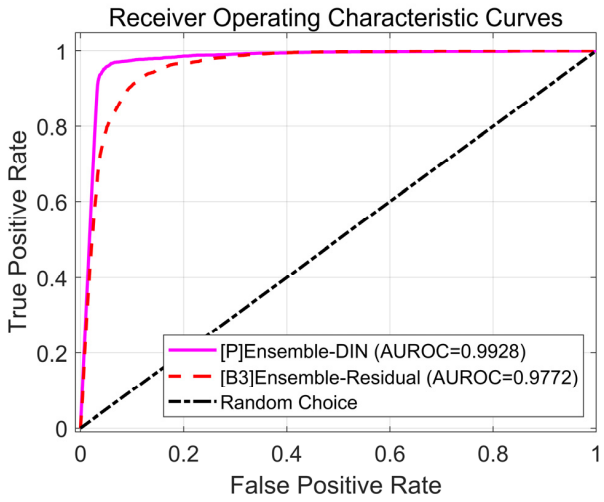


Figure 15. Receiver operating characteristic (ROC) curves.

6. Conclusions

In this study, an RFEI method that targets the physical layers of FHSS networks was proposed with the objective of directly identifying emitter IDs from received FH signals. An analog SF extraction process, SF spectrogram features, a DIN-based classifier for emitter classification, and an outlier detector algorithm for attacker detection were proposed and applied to the target hop signals. In addition, the ensemble approach that

utilized multimodality SFs was evaluated for robust classification. The results showed that the SF spectrogram extracted from the received FH signal can be effectively analyzed using the DIN-based classifier, and the classification accuracy was improved by at least 1.00% compared with those of other baselines. In addition, the multimodal SF ensemble approach, that is, the use of RT, FT, and SS, achieved the best results with a classification accuracy of 97.0% for the seven real FHSS emitters. In addition, the inception block-based approach was more effective than the residual block-based approach owing to its filtering ability at different receptive field sizes. From the analysis of the GCAM for each FH emitter, we found that the classifier model can train the region wherein the differences in the SFs can be maximized. In addition, the outlier detection performance of the proposed method was evaluated. We found that the output characteristics of the outliers differed from those of the training samples, and this property can be used by the detector to identify attacker signals with an AUROC of 0.99.

These results support that the proposed RFEI method can identify emitter IDs of the FH signals emitted by authenticated users and can detect the existence of the FH signals emitted by attackers. Because the SFs cannot be reproduced, it is possible to configure non-replicable authentication systems in the physical layer of the FHSS network. This study focused on evaluating the RFEI method, one of the components of the overall authentication system. Our future study will consider system improvement by utilizing the GCAM to detect misclassification cases.

As another future study, we will consider the property of the outliers in the RFEI system. We believe that further distinctions of the outliers, namely the detection of multi-labeled outliers, may be possible. We expect that this future consideration will help prevent the malicious application of the RFEI system, such as when eavesdroppers utilize the RFEI system. If the eavesdropper can successfully prepare the target FH sample, it can be used as a signal tracking method to decode the actual FH signal transmission. Our future study will consider the ways to prevent this malicious scenario by generating artificial outliers that can imitate authentication users.

Author Contributions: Conceptualization, J.K. and H.L. (Heungno Lee); methodology, J.K.; software, J.K.; validation, J.K. and Y.S.; formal analysis, J.K. and H.L. (Heungno Lee); data collection, J.K., H.L. (Hyunku Lee) and J.P.; writing—original draft preparation, J.K., Y.S. and H.L. (Heungno Lee); writing—review and editing, J.K., Y.S. and H.L. (Heungno Lee); visualization, J.K.; supervision, H.L. (Heungno Lee); project administration, H.L. (Hyunku Lee) and J.P.; funding acquisition, J.P. All authors have read and agreed to the published version of the manuscript.

Funding: The authors gratefully acknowledge the support from the LIG Nex1 which was contracted with the Agency for Defense Development (ADD), South Korea (Grant No. LIGNEX1-2019-0132).

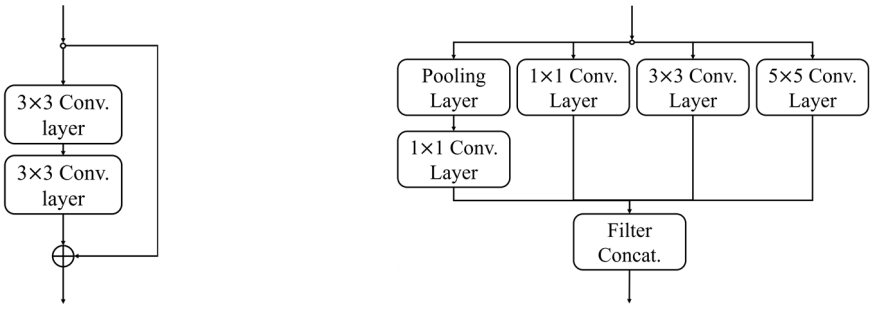
Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable. Due to security issues, the FHSS datasets are not disclosed.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, the writing of the manuscript, or the decision to publish the results. However, the funders helped prepare the FHSS emitters for data collection, analysis, and interpretation.

Appendix A. Architecture and Design Strategies of the Main Blocks



(**a**) (**b**)

Figure A1. Basic block for constructing the deep learning classifier used in this study: (**a**) the residual block [22] and (**b**) the inception block [23].

The custom deep learning-based classifier utilized in our study consists of two main blocks: a residual block [22] and an inception block [23]. The architecture of these blocks is shown in Figure A1.

The design strategy of the residual block is to handle the degradation problem as the network goes deeper [22]. The residual block contains skip connections between adjacent convolutional layers and helps mitigate the vanishing gradient problem. The goal of the residual network is to allow flexible training of the features as the network depth increases.

The design strategy of the inception block involves calculating features with different filter sizes in the same layer [23]. The inception block contains parallel convolutional layers with different filter sizes. The results for each layer are concatenated in the filter axis and pass through the next layer. These parallel connections can extract features with multiple receptive field sizes, which are useful when the features vary in location and size.

The spectrogram contains the physical measurements of the SF signals. It represents the power densities of the SF signals along the time–frequency axes. To train these two-dimensional density behaviors of the SF signals, we aimed to filter the spectrogram on multiple filter scales in the temporal and spatial domains by applying inception blocks.

Appendix B. Implemented Parameter Settings in Experiments

The implemented parameters of the RF fingerprinting algorithms performed at our experiments are described in Table A1.

Table A1. Implemented parameter settings.

Algorithm	Parameters	Values
Proposed algorithm	Number of FH signals, K	7
	Number of emitters trained on the classifier, C	7
	Length of the FH signal, N	194,475
	Length of the SFs, N_{SF}	38,895 for RT and FT 175,027 for SS
	Extraction window size, W_E	1945
	Energy variance detection threshold, δ	0.1
	Length of the frequency axis in the spectrogram, $N_{\mathrm{SF}}^{f} N_{\mathrm{SF}}^{stats}$	205 for all SFs
	Length of the time axis in the spectrogram, N_{SF}^{t}	74 for RT and FT 340 for SS
	STFT window size, W_{STFT}	1024

Table A1. *Cont.*

Algorithm	Parameters	Values
Baseline 1 algorithm	Number of segmented SFs, $N_{seg.}$ Length of the handcrafted feature vector, N_{SF}^{stats}	10 30 for RT and FT 300 for SS
Baseline 2 algorithm	Length of the raw vector segmented by the FH signal,	3889 for RT and FT 17,502 for SS

Appendix C. Gradient-Weighted Class Activation Map

The GCAM is a feature visualization method that identifies parts of the input signal that positively influence the class decision [40]. It can be obtained by performing the following steps. (1) Firstly, the gradient of the inference score from the target class c_j, that is, the jth element of the model output $\mathbf{y}$, is back-propagated to the last convolutional layer of the model, which is the last reduction-A block of the DIN. (2) Secondly, global average pooling of the back-propagated values on the grid axis, that is, the time and frequency axes of the feature map, is performed. This value serves as a weight to infer the importance of the current filter result. (3) With the linear combination of the entire filter map, the Grad-CAM for the input sample $\mathbf{s}$ and decision class c_j is obtained. Specifically, it follows

$$a_f^{c_j} = \frac{1}{P}\sum_z\sum_k \frac{\partial \mathbf{y}[c_j]}{\partial A_{zk}^f} \tag{A1}$$

$$\mathrm{GCAM}(\mathbf{s}, c_j) = \mathrm{ReLu}\left(\sum_f a_f^{c_j} A^f\right) \tag{A2}$$

where A_{zk}^f is the grid point (z,k) of the fth filter map existing on the last convolutional layer of the classifier model, P is the size of the fth filter map, and A^f and $a_f^{c_j}$ are the neuron importance weights of the fth filter map when the target class c_j is decided.

Finally, the GCAM is averaged for the positive samples that record the correct identification results. The positive sample dataset $\mathbf{S}_{True} = [\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_{M_{True}}]$ is collected when the classification result of the input sample $\mathbf{s}_j$ in Equation (15) is true. For the positive sample $\mathbf{s}_j$ and its true decision class c_j, the GCAM can be averaged as follows

$$\mathrm{AGCAM}(c_j) = \frac{1}{M_{True}} \sum_{\mathbf{x}_i \in \mathbf{S}_{True}} \mathrm{GCAM}(\mathbf{s}_j, c_j) \tag{A3}$$

References

1. Standard for Information Technology—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications IEEE Std. No. 802.11-2020. February 2021. Available online: https://ieeexplore.ieee.org/document/9363693 (accessed on 15 November 2021).
2. Soltanieh, N.; Norouzi, Y.; Yang, Y.; Karmakar, N.C. A review of radio frequency fingerprinting techniques. *IEEE J. Radio Freq. Identif.* **2020**, *4*, 222–233. [CrossRef]
3. Kennedy, I.O.; Scanlon, P.; Mullany, F.J.; Buddhikot, M.M.; Nolan, K.E.; Rondeau, T.W. Radio transmitter fingerprinting: A steady state frequency domain approach. In Proceedings of the IEEE 68th Vehicular Technology Conference, Calgary, AB, Canada, 21–24 September 2008; pp. 1–5.
4. Ali, A.M.; Uzundurukan, E.; Kara, A. Assessment of features and classifiers for Bluetooth RF fingerprinting. *IEEE Access* **2019**, *7*, 50524–50535. [CrossRef]
5. Patel, H.J.; Temple, M.A.; Baldwin, R.O. Improving ZigBee device network authentication using ensemble decision tree classifiers with radio frequency distinct native attribute fingerprinting. *IEEE Trans. Reliab.* **2015**, *64*, 221–233. [CrossRef]
6. Yang, K.; Kang, J.; Jang, J.; Lee, H.-N. Multimodal sparse representation-based classification scheme for RF fingerprinting. *IEEE Commun. Lett.* **2019**, *23*, 867–870. [CrossRef]

7. Merchant, K.; Revay, S.; Stantchev, G.; Nousain, B. Deep learning for RF device fingerprinting in cognitive communication networks. *IEEE J. Sel. Top. Signal Process.* **2018**, *12*, 160–167. [CrossRef]
8. Pan, Y.; Yang, S.; Peng, H.; Li, T.; Wang, W. Specific emitter identification based on deep residual networks. *IEEE Access* **2019**, *7*, 54425–54434. [CrossRef]
9. Stremler, F.G. *Introduction to Communication Systems*; Addison–Wesley: Reading, MA, USA, 1990; p. 658.
10. Standard for Information Technology—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications IEEE Std. No. 802.11-2012. March 2012. Available online: https://ieeexplore.ieee.org/document/6178212 (accessed on 15 November 2021).
11. Shin, H.; Choi, K.; Park, Y.; Choi, J.; Kim, Y. Security Analysis of FHSS-Type Drone Controller. In *International Workshop on Information Security Applications*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 9503, p. 240.
12. Liu, Z.; Huang, Z.; Zhou, Y. Hopping instants detection and frequency tracking of frequency hopping signals with single or multiple channels. *IET Commun.* **2012**, *6*, 84–89. [CrossRef]
13. Wang, Z.; Zhang, B.; Zhu, Z.; Wang, Z.; Gong, K. Signal Sorting Algorithm of Hybrid Frequency Hopping Network Station Based on Neural Network. *IEEE Access* **2021**, *9*, 35924–35931. [CrossRef]
14. Li, S.; Nie, H.; Wu, H. Performance Analysis of Frequency Hopping Ad Hoc Communication System With Non-Orthogonal Multiple Access. *IEEE Access* **2019**, *7*, 113171–113181. [CrossRef]
15. Feng, Y.; Yan, S.; Liu, C.; Yang, Z.; Yang, N. Two-stage relay selection for enhancing physical layer security in non-orthogonal multiple access. *IEEE Trans. Inf. Forensics Secur.* **2018**, *14*, 1670–1683. [CrossRef]
16. Liu, Y.; Qin, Z.; Elkashlan, M.; Gao, Y.; Hanzo, L. Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 1656–1672. [CrossRef]
17. Ghous, M.; Abbas, Z.H.; Hassan, A.K.; Abbas, G.; Baker, T.; AI-Jumeily, D. Performance Analysis and Beamforming Design of a Secure Cooperative MISO-NOMA Network. *Sensors* **2021**, *21*, 4180. [CrossRef] [PubMed]
18. Mpitziopoulos, A.; Gavalas, D.; Konstantopoulos, C.; Pantziou, G. A survey on jamming attacks and countermeasures in WSNs. *IEEE Commun. Surv. Tutor.* **2009**, *11*, 42–56. [CrossRef]
19. Conti, M.; Dragoni, N.; Lesyk, V. A survey of man in the middle attacks. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 2027–2051. [CrossRef]
20. Oppenheim, A.V.; Ronald, W.S.; John, R.B. *Discrete-Time Signal Processing*; Prentice Hall: Hoboken, NJ, USA, 1999.
21. Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* **2020**, *53*, 5455–5516. [CrossRef]
22. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
23. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.
24. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, Inception-ResNet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
25. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning (ICML), Lille, France, 6–11 July 2015; pp. 448–456.
26. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
27. Ganaie, M.A.; Hu, M.; Tanveer, M.; Suganthan, P.N. Ensemble deep learning: A review. *arXiv* **2021**, arXiv:2104.02395.
28. Guo, J.; Nie, X.; Yin, Y. Mutual Complementarity: Multi-modal enhancement semantic learning for micro-video scene recognition. *IEEE Access* **2020**, *8*, 29518–29524. [CrossRef]
29. Hendrycks, D.; Gimpel, K. A baseline for detecting misclassified and out-of-distribution examples in neural networks. In Proceedings of the International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April 2017; pp. 1–12.
30. Liang, S.; Li, Y.; Srikant, R. Enhancing the reliability of out-of-distribution image detection in neural networks. In Proceedings of the International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April 2017; pp. 1–27.
31. Lee, K.; Lee, K.; Lee, H.; Shin, J. A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In Proceedings of the Neural Information Processing Systems (NIPS), Montreal, QC, Canada, 3–8 December 2018; pp. 7167–7177.
32. Lee, K.; Lee, H.; Lee, K.; Shin, J. Training confidence-calibrated classifiers for detecting out-of-distribution samples. In Proceedings of the International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018; pp. 1–16.
33. Hendrycks, D.; Mazeika, M.; Dietterich, T. Deep anomaly detection with outlier exposure. In Proceedings of the 7th International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019; pp. 1–18.
34. Choi, H.; Jang, E.; Alemi, A.A. WAIC, but why? Generative ensembles for robust anomaly detection. *arXiv* **2018**, arXiv:1810.01392.
35. Serrà, J.; Álvarez, D.; Gómez, V.; Slizovskaia, O.; Núñez, J.F.; Luque, J. Input complexity and out-of-distribution detection with likelihood-based generative models. In Proceedings of the International Conference on Learning Representations (ICLR), Virtual Conference, 26 April–1 May 2020; pp. 1–15.
36. Hinton, G.; Vinyals, O.; Dean, J. Distilling the knowledge in a neural network. *arXiv* **2015**, arXiv:1503.02531.
37. Goodfellow, I.J.; Shlens, J.; Szegedy, C. Explaining and harnessing adversarial examples. In Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015; pp. 1–11.

38. Bi, G.; Mitra, S.K. FFT–based sampling rate conversion. In Proceedings of the IEEE Conference on Industrial Electronics and Applications (ICIEA), Singapore, 18–20 July 2012; pp. 428–431.
39. Sklar, B. *Digital Communications*; Prentice Hall: Upper Saddle River, NJ, USA, 2001; Volume 2, pp. 773–774.
40. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.