Fingerprinting Radio Frequency Devices Using the Instantaneous Frequency

Steven Sandoval, Matthew Bredin, Phillip L. De Leon, Susana Terrazas
Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, New Mexico 88003
Email: {spsandov,mbredin,pdeleon,selisath}@nmsu.edu

***Abstract*—One approach to enhancing security in a wireless network environment is by uniquely identifying particular network devices via physical, hardware-level traits. Such traits are not easily spoofed because they are typically tied to unique hardware variations and imperfections introduced in the manufacturing process. In this work, we classify individual radio frequency (RF) devices using a feature vector based on the instantaneous frequency (IF) of the power ON transient. To evaluate the performance of the proposed feature vector, we collected 1889 RF bursts from six Laird CL4790 ConnexLink RS232 RF modules in a high SNR environment. From these bursts we extracted the proposed feature vectors. With an ensemble-based classifier, we achieve better than 95% accuracy; we also include an evaluation of the feature in noisy conditions. From this work we find that a feature vector based on the IF during a transient event, appears to provide a unique RF device fingerprint which is easily extractable.**

I. Introduction

Providing security in a wireless network environment is a challenging problem because access to network resources is possible without being physically connected. One way to enhance security is through unique identifiers associated to particular network devices without considering easily forgeable identifiers such as MAC addresses and user credentials [1]. Prior research has demonstrated the ability to identify a device via physical, hardware-level traits due to variations/imperfections in the circuitry associated with manufacturing differences [2]. One way these variations/imperfections manifest is through unique transient events when an RF transmitter is activated or deactivated. Features extracted from these transient events may provide a unique “fingerprint” for a radio frequency (RF) device that is similar to a biometric for human identification [3]. Thus, RF device fingerprinting can offer an additional defensive layer of security beyond conventional (software) protocols for authentication [4], [5].

More specifically, hardware-based device-level fingerprinting offers additional security and safeguards against attacks such as device cloning, message replay, and spoofing attacks [6]. On the other hand, fingerprinting may also be used offensively to gain information about network operations or to gain information about specific network users [6]. A skilled attacker, for example, may be able to exploit a device fingerprint to violate sender anonymity, in order to associate a transmission to a specific sender. For example, FM transmitters have been shown in many cases, to have relatively short transient characteristics directly following when the transmitter is activated or “keyed” [7]. Potentially, these transient characteristics are similar enough that the device model may be identifiable, and in the best case are unique enough to allow identification of an individual device. Recently, the widespread availability/use of software defined radios (SDRs) has lowered the barrier to adversarial cloning and spoofing [8]. One potential defense against these attacks could be RF device fingerprinting.

Methods for RF device fingerprinting are broadly divided into two categories characterized by imperfections in timing or modulation [5]. Many of the studies to date have focused on identifying these imperfections in either the ON and/or OFF transients of the wireless transmission, or alternatively the “steady state” segments of the wireless transmission [3], [4], [6], [7], [9], [10]. We note that some authors refer to a more general problem known as specific emitter identification (SEI) which seeks to designate the unique transmitter of a given signal using only external feature measurements [11], [12]. However, the focus of SEI research is often the development of RF device fingerprints. Most common fingerprinting methods utilize statistical, time-frequency/time-scale, or parametric models for the fingerprint.

For example, in [6] features based on signal amplitude statistics are considered, in [13] higher order statistics for common digital modulation schemes are considered, in [14] a cycle-frequency domain profile (statistical) feature for orthogonal frequency division multiplexing (OFDM) signals is proposed and tested on an IEEE 802.11a/g WLAN device, and in [15] a normalized permutation entropy is used. Other work uses time-frequency and time-scale methods. For example features are considered based on wavelet coefficients extracted from the transients [16], empirical mode decomposition and Haar wavelet decompositions [17], Hilbert-Huang Transform [18], [19], ambiguity function and Wigner distribution [20], and the intrinsic time-scale decomposition [21]. Finally, other work seeks to use parametric model parameters as features. In [9], the authors identified a feature based on a proxy for fractal dimension which measures change in detail with respect to change in scale. In [22], the authors propose to model circuit-dependent nonlinear emitter characteristics using a complex power series while in [23] and [24], the authors propose to extract a set of intrapulse parameters modelling the AM and FM waveforms.

In this work, we propose to use the sequence of instantaneous frequencies (IFs) [25] during the input-on transient as the feature vector for RF device fingerprinting. As we will demonstrate, this feature is easily extractable and can be quite discriminating. The remainder of this paper is organized as follows. In Section II, we provide information about our data collection environment and feature extraction. In Section III, we describe our observations of the data using principal component analysis (PCA), then evaluate the proposed feature vectors using an ensemble-based classifier at various levels of SNR. We discuss the results and other observations in Section IV and conclude the article in Section V.

II. Methods

A. Data Collection

Our data collection environment consists of the following arrangement. We use a SDR model USRP B210 connected to a PC using

GNU Radio for radio device control and data acquisition. The SDR independently captures the communications signals from each of the two radio devices, simultaneously. The devices under test consist of six Laird CL4790 ConnexLink RS232 RF modules operating at 900 (902-928) MHz with 1 watt output power. We designated the six radios as CL4790$_1$ through CL4790$_6$. While the devices do not have individual serial numbers, radios CL4790$_1$ through CL4790$_4$ had consecutive MAC addresses and radios CL4790$_5$ and CL4790$_6$ had different but consecutive MAC addresses. We assume that consecutive MAC addresses indicate a common manufacturing lot.

During the data acquisition, the radios operate using a frequency-hopping spread spectrum wireless communication protocol in a half-duplex mode. The resulting communications were sampled using a rate of $f_\mathrm{s} = 30$ MHz centered at $f_\mathrm{c} = 915$ MHz and 16-bit interlaced in-phase and quadrature (I/Q) data were saved as `.sc16` files. We note that while the radios are connected to the acquisition system using a minimum of 30 dB attenuation, the data collection environment is relatively free from noise and other RF disturbances due to the wired connections from the devices under test to the acquisition system. Thus we consider the recordings as noise-free. We evaluate classifier performance in various SNR environments in Section III-B.

B. *Feature Extraction*

Data collection results in a continuous recording of the intermediate frequency signal sampled at 30 MHz, approximately 40 seconds in length. The STFT magnitude for the beginning of a recording is shown in Fig. 2(a). Prior to feature extraction, GNU radio segments the continuous recording into individual RF bursts by shifting the frequency band of interest to DC, lowpass filtering, and downsampling to 6 MHz. The resulting signal bursts, summarized in Table I, are approximately 40 ms in duration and are roughly aligned such that the RF burst begins at approximately $t = 125$ μs. The STFT magnitude for the first 1/8th of a single RF burst is shown in Fig. 2(b) with the transient event highlighted in the grey box.

In this work, we propose to directly use the IF sequence within the transient event as the feature vector. Shown in Fig. 1 is the block diagram for the signal processing involved in feature extraction. Because of the temporal nature of the proposed feature vector, a second fine-grained time-alignment is necessary prior to classification. Time-alignment consists of determining a time-shift parameter Δt using a cross-correlation of a "template" with the instantaneous power-weighted (IA squared) lowpass filtered IF. This step occurs within the "estimate alignment parameters" block in Fig. 1. The template is pre-built by choosing a random subset of RF bursts from each class and averaging the instantaneous power-weighted IFs. Because of the frequency-hopping nature of the communication protocol used by the radios [see Fig. 2(a)], a second fine-grained frequency alignment is also necessary. Frequency-alignment centers the IF about zero using a frequency-shift parameter ω_s. The frequency-shift parameter is determined by averaging the IF after completion of the transient, i.e. in the steady state. This step also occurs within the "estimate alignment parameters" block in Fig. 1.

Using the two alignment parameters, the IF sequence is both time-aligned to the template and frequency-centered at zero. Once aligned, we segment the transient to approximately 40 μs. This is depicted in the alignment/segment block in Fig. 1. The aligned and segmented transient is demodulated and lowpass filtered resulting in a length 251 feature vector, $\omega_\mathrm{LP}(t)$. This length encapsulates approximately 20 μs before and after the IF peak and captures the majority of the transient event. Fig. 2(c) shows STFT analysis for a single RF transient with the lowpass filtered IF estimate (black line) overlaid.

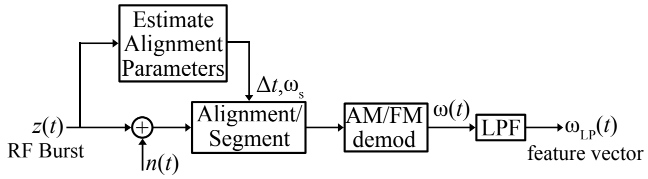

Figure 1. Block diagram for the signal processing involved in extracting a feature vector ω_LP from the RF burst $z(t)$ consists of estimation of signal alignment parameters $\{\Delta t, \omega_\mathrm{s}\}$, alignment, segmentation, IF estimation $\omega(t)$, and lowpass filtering. For evaluation of robustness, additive white Gaussian noise $n(t)$ is included.

III. Simulations and Results

A. *Data Exploration*

To visually assess the separability of the classes, we project the feature vectors onto three dimensions, using PCA. The first three principal components (corresponding to the three directions of maximum variance) of the data are shown in Fig. 3, and the color associations for the devices are given in Table I. We note that radios CL4790$_1$ through CL4790$_4$ are tightly clustered in PCA space and as noted earlier, are possibly from the same manufacturing lot.

B. *Classification*

Next we use the proposed feature vector as a "device fingerprint" for classification. For the classifier, we use an ensemble of 1000 tree classifiers with bagging, implemented using the `fitcensemble` function in Matlab. The data in Table I was randomly partitioned into subsets: 80% for training and 20% for testing. To simulate the effects of noise, we mixed complex additive white Gaussian noise (AWGN) $n(t)$ to the signal $z(t)$ as shown in Fig. 1 at prescribed SNR values. This allows the evaluation of robustness of the proposed device fingerprint in non-ideal conditions. To isolate the deleterious effects of noise on the feature vector from the effects on the alignment procedure, we use the same alignment parameters for all SNR levels as illustrated in Fig. 1. We justify this choice because the relatively simple alignment procedure used in this work could be improved or replaced with a more robust methodology. Alternately, alignment could be eliminated by utilizing a classifier which natively models time-dependent phenomena such as a recurrent neural network. For each SNR level we perform 30 trials and report the average percent accuracy. Finally, we consider the feature extraction as described above, both with and without the lowpass filtering. The results of these analyses are summarized in Table II and Fig. 4.

IV. Discussion

In prior research, methods for RF device fingerprinting are characterized by imperfections in either timing or modulation. In this paper, we proposed the use of the IF as a feature for RF device fingerprinting. The instantaneous nature of the IF encapsulates both timing and modulation imperfections. Furthermore, extraction of the proposed feature vector is both conceptually and computationally simple, potentially allowing real-time implementation.

Classification results show that in high SNR environments, the proposed feature provides excellent classification accuracy. At lower SNRs, lowpass filtering improves accuracy, e.g. at 5 dB the improvement is 11.2%. We find that classifier accuracy smoothly degrades over the range of SNRs from 10 dB to -10 dB. However,

Device	CL4790$_1$	CL4790$_2$	CL4790$_3$	CL4790$_4$	CL4790$_5$	CL4790$_6$	**Total**
# RF Bursts	323	291	321	305	342	307	**1889**

Table I
THE NUMBER OF RF BURSTS FOR THE DEVICES UNDER CONSIDERATION. THE TEXT COLORING CORRESPONDS TO MARKER COLOR IN FIG. 3.

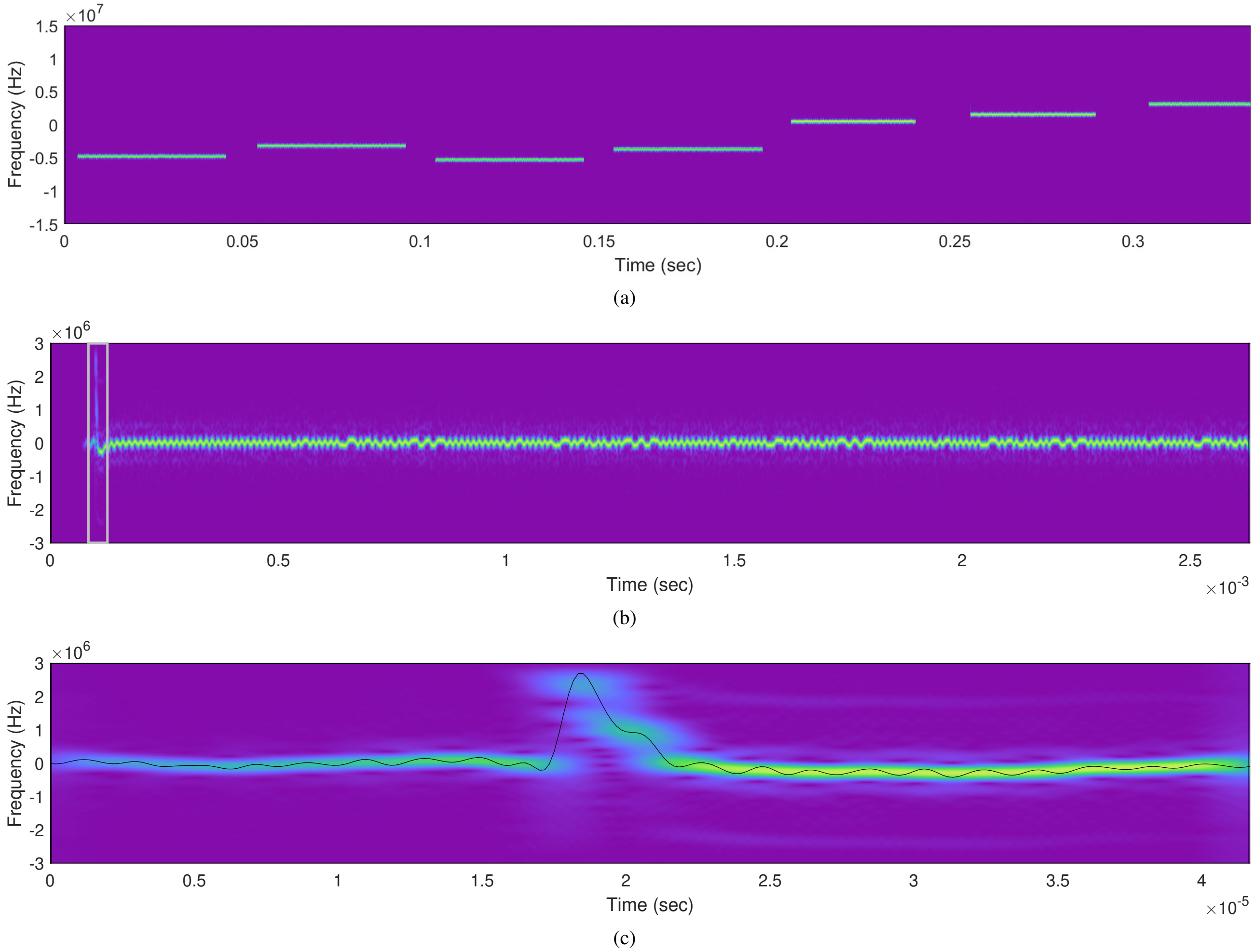

Figure 2. Time-frequency (STFT) visualization of the data at various time scales. (a) A segment of the continuous RF recording, from the data acquisition system, showing the presence of multiple RF bursts. (b) Start of a single RF burst with the transient event highlighted in the boxed region. (c) The transient event corresponding to the boxed region of the RF burst in (b), with the filtered IF estimate overlaid (black line). For this example, at about $t = 1.75 \times 10^{-5}$, we observe a rapid change in IF during the input-on transient. This transient is unique enough to discriminate between individual radio devices.

	SNR (dB)													
Feature Vector	**60**	**30**	**20**	**15**	**10**	**8**	**5**	**3**	**0**	$\mathbf{-3}$	$\mathbf{-5}$	$\mathbf{-8}$	$\mathbf{-10}$	$\mathbf{-20}$
$\omega(t)$	95.2	94.1	91.2	86.1	77.5	70.4	57.8	52.9	41.4	33.7	28.2	21.2	18.8	17.2
$\omega_{\text{LP}}(t)$	97.4	97.4	96.6	95.5	91.0	87.7	79.0	72.1	58.0	43.5	34.8	26.7	21.7	17.1

Table II
PERCENT ACCURACY FOR DEVICE CLASSIFICATION USING THE UNFILTERED $\omega(t)$ AND FILTERED $\omega_{\text{LP}}(t)$ IF ESTIMATE AS A FUNCTION OF SNR IN DECIBELS.

more sophisticated noise reduction techniques could be employed. Additionally, in a more realistic networked environment of wireless devices, source separation may have to be considered in order to isolate individual transients prior to feature extraction.

We assumed that consecutive MAC addresses indicate a common manufacturing lot. If true, the plot in Fig. 3 suggests that devices manufactured within a common lot have features which are more similar than those from other lots. This could be exploited to identify devices within a common manufacturing lot and thus be used to detect MAC address spoofing. As with any classifier, performance may degrade as the number of classes (radio devices) increases assuming the feature vectors are similar. However, our observation noted above suggests that devices manufactured in different lots may not have similar feature vectors. Further investigation is necessary to ensure

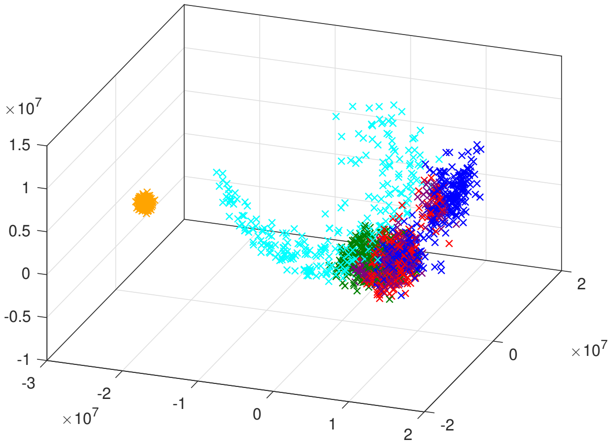

Figure 3. The first three principle components of the time-aligned lowpass filtered IFs where color denotes devices 1 (×), 2 (×), 3 (×), 4 (×), 5 (×), and 6 (×). All devices are distinctly clustered in PCA space, i.e. not distributed randomly. The level of separability in the three dimensions shown, varies among the devices. The ×s are well separated from the other devices. The ×s are the most spread out in a crescent moon shape. The other 4 are more tightly arranged.

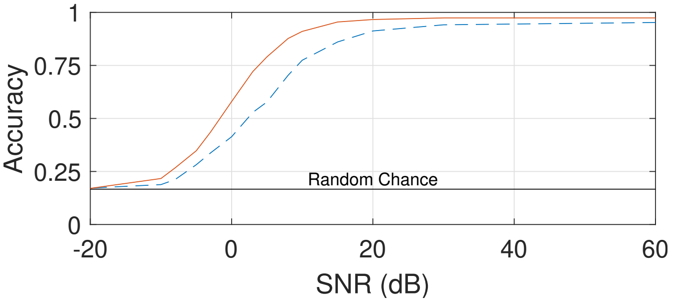

Figure 4. Classifier accuracy as a function of SNR for the unfiltered $\omega(t)$ (---) and filtered IF ω_{LP} (—) estimates. Below -10 dB the accuracy approaches random chance. Above 10 dB the accuracy is above 75%. Filtering improves the classifier accuracy regardless of SNR.

that the proposed method scales appropriately when considering a large number of radio devices. Finally, while some authors have noted that a device's RF transient does in fact change over time, there is limited investigation reported in the literature on the consistency of a device's RF transient over time and how it changes as the device ages. In addition, other effects on the IF due to Doppler, multi-path, fading, temperature variation, and battery condition many also need to be considered.

V. Conclusion

In this paper, we have proposed the viability of directly using the IF of the power ON transient as a feature for RF device fingerprinting. Previous research has shown that wireless model (e.g. model number) identification is an easier problem than individual device (e.g. serial number) identification. We considered individual RF device identification using six wireless devices from the same model family. With an ensemble-based classifier, we achieved better than 95% accuracy in a high SNR environment. In our evaluation of the classifier in a noisy environment, we found that accuracy degrades smoothly over a 20 dB SNR range. Based on this analysis, our proposed feature vector which based directly on the IF of the power ON transient event can provide a unique device fingerprint which is easily extractable.

References

[1] Y. Zhang, W. Lee, and Y.-A. Huang, "Intrusion Detection Techniques for Mobile Wireless Networks," *Wireless Networks*, vol. 9, no. 5, pp. 545–556, Sep. 2003.

[2] S. Banerjee and V. Brik, *Encyclopedia of Cryptography and Security*. Boston, MA: Springer US, 2011, pp. 1388–1390.

[3] J. Hall, M. Barbeau, and E. Kranakis, "Detection of Transient in Radio Frequency Fingerprinting using Signal Phase," *Wireless and Optical Commun.*, pp. 13–18, 2003.

[4] B. Danev, D. Zanetti, and S. Capkun, "On Physical-Layer Identification of Wireless Devices," *ACM Computing Surveys*, vol. 45, no. 1, p. 6, 2012.

[5] Sieka and Bartlomiej, "Using Radio Device Fingerprinting for the Detection of Impersonation and Sybil Attacks in Wireless Networks," *Lecture Notes in Computer Science*, vol. 4357, p. 179, 2006.

[6] K. B. Rasmussen and S. Capkun, "Implications of Radio Fingerprinting on the Security of Sensor Networks," in *Proc. Int. Conf. Security and Privacy in Commun. Networks*, 2007, pp. 331–340.

[7] K. Ellis and N. Serinken, "Characteristics of Radio Transmitter Fingerprints," *Radio Science*, vol. 36, no. 4, pp. 585–597, 2001.

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

[9] O. Tekbas, O. Ureten, and N. Serinken, "Improvement of Transmitter Identification System for Low SNR Transients," *Electronics Letters*, vol. 40, no. 3, pp. 182–183, 2004.

[10] I. O. Kennedy, P. Scanlon, F. J. Mullany, M. M. Buddhikot, K. E. Nolan, and T. W. Rondeau, "Radio transmitter fingerprinting: A steady state frequency domain approach," in *IEEE Vehicular Tech. Conf.*, 2008, pp. 1–5.

[11] K. I. Talbot, P. R. Duley, and M. H. Hyatt, "Specific Emitter Identification and Verification," *Technology Review*, vol. 113, 2003.

[12] G. Huang, Y. Yuan, X. Wang, and Z. Huang, "Specific Emitter Identification for Communications Transmitter using Multi-Measurements," *Wireless Personal Commun.*, vol. 94, no. 3, pp. 1523–1542, 2017.

[13] Y.-J. Yuan, Z. Huang, and Z.-C. Sha, "Specific Emitter Identification based on Transient Energy Trajectory," *Progress in Electromagnetics Research*, vol. 44, pp. 67–82, 2013.

[14] K. Kim, C. M. Spooner, I. Akbar, and J. H. Reed, "Specific Emitter Identification for Cognitive Radio with Application to IEEE 802.11," in *IEEE Global Telecommun. Conf.*, 2008, pp. 1–5.

[15] G. Huang, Y. Yuan, X. Wang, and Z. Huang, "Specific Emitter Identification Based on Nonlinear Dynamical Characteristics," *Canadian J. Elect. and Comp. Eng.*, vol. 39, no. 1, pp. 34–41, 2016.

[16] R. Hippenstiel and Y. Payal, "Wavelet Based Transmitter Identification," in *IEEE Sym. Sig. Process. and its Applications*, vol. 2, 1996, pp. 740–742.

[17] C. Song, J. Xu, and Y. Zhan, "A Method for Specific Emitter Identification based on Empirical Mode Decomposition," in *IEEE Int. Conf. Wireless Commun., Networking and Info. Security*, 2010, pp. 54–57.

[18] Y. Yuan, Z. Huang, H. Wu, and X. Wang, "Specific Emitter Identification based on Hilbert-Huang Transform-Based Time-Frequency-Energy Distribution Features," *IET Commun.*, vol. 8, no. 13, pp. 2404–2412, 2014.

[19] J. Zhang, F. Wang, O. A. Dobre, and Z. Zhong, "Specific Emitter Identification via Hilbert-Huang Transform in Single-Hop and Relaying Scenarios," *IEEE Trans. Info. Forensics and Security*, vol. 11, no. 6, pp. 1192–1205, 2016.

[20] L. Li, H.-B. Ji, and L. Jiang, "Quadratic Time–Frequency Analysis and Sequential Recognition for Specific Emitter Identification," *IET Sig. Process.*, vol. 5, no. 6, pp. 568–574, 2011.

[21] C. Song, Y. Zhan, and L. Guo, "Specific Emitter Identification Based on Intrinsic Time-Scale Decomposition," in *IEEE Int. Conf. Wireless Commun. Networking and Mobile Computing*, 2010, pp. 1–4.

[22] M.-W. Liu and J. F. Doherty, "Specific Emitter Identification using Nonlinear Device Estimation," in *IEEE Sarnoff Sym.*, 2008, pp. 1–5.

[23] A. Kawalec and R. Owczarek, "Specific Emitter Identification using Intrapulse Data," in *European Radar Conf.*, 2004, pp. 249–252.

[24] H. Ye, Z. Liu, and W. Jiang, "Comparison of Unintentional Frequency and Phase Modulation Features for Specific Emitter Identification," *Electronics Letters*, vol. 48, no. 14, pp. 875–877, 2012.

[25] S. Sandoval and P. L. D. Leon, "The Instantaneous Spectrum: A General Framework for Time-Frequency Analysis," *IEEE Trans. Sig. Process.*, vol. 66, pp. 5679–5693, Nov. 2018.