

Received January 12, 2021, accepted January 28, 2021, date of publication February 10, 2021, date of current version February 17, 2021. *Digital Object Identifier 10.1109/ACCESS.2021.3058167*

Carrier Leakage Estimation Method for Cross-Receiver Specific Emitter Identification

MENGKA[I](https://orcid.org/0000-0002-7503-9570) SHI^{ID1,2}, YUANLING HUANG², AND GUILIANG WANG²
¹School of Information System Engineering, PLA Strategic Support Force Information Engineering University, Zhengzhou 450002, China

²National Key Laboratory of Science and Technology on Blind Signal Processing, Chengdu 610041, China Corresponding author: Mengkai Shi (shimengkai96@163.com)

ABSTRACT Specific emitter identification (SEI) technology uses fingerprint features originated from the emitter imperfections to identify the transmitting devices. However, the fingerprint features used in SEI are subtle and can be easily affected by the distortion of the receivers. This means that the features cannot be used universally among different receivers, which can bring many limits in identification but is rarely concerned in previous studies. To resolve this problem, the effects of receiver distortions on I/Q modulator fingerprints are discussed, and a carrier leakage estimation method for cross-receiver SEI is developed. In the proposed method, joint estimation of carrier leakage and filter distortions is included to eliminate the influence of the receiver. In the simulation and experiment, the performance of the proposed method and the traditional method is compared. The results validate that the proposed method can effectively reduce the feature offset caused by receiver distortions and the proposed method has the same effectiveness as the traditional one. The proposed method can be widely used in cross-receiver specific emitter identification.

INDEX TERMS Carrier leakage, radio frequency fingerprint, receiver distortions, specific emitter identification.

I. INTRODUCTION

Specific emitter identification (SEI) is a promising identification technique and can be potentially used in wireless security [1], [2], spectrum-management operations [3], cognitive radio [4] and self-organized networks [5]. In SEI, radio frequency fingerprints (RFF) are used to uniquely identify individual transmitters. These fingerprints are originated from the imperfection of hardware and are distinguishable in wireless networks with nonexclusive transmission medium [6]. However, RFF can be easily influenced by receiver distortions, which come from the imperfection of the analog components in receivers. This shortcoming makes it impossible to use SEI in some cross-receiver conditions such as recognition for the wide-area moving target and multi-platform collaborative recognition. Therefore, an effective cross-receiver SEI method is indispensable to expand the application of SEI.

Most existing SEI techniques focus on single-receiver conditions and the discussion about cross-receiver identification is limited. In general, existing SEI techniques fall into two categories [7]: traditional approaches and machine-learning approaches.

The associate editor coordinating the review of this manuscript and approving it for publication was Di Zhang¹[.](https://orcid.org/0000-0001-7190-1621)

In traditional approaches, the mechanism about the analog components of emitters is analyzed and a mechanism model or mathematical discrimination of signal is established to estimate the RFF. The mechanism modules of the components such as RF oscillator [8], power amplifier [9] and modulator [10] are commonly analyzed to get their fingerprints. The mathematic tools such as Hilbert-Huang transform [11], [12], discrete wavelet transform [13] and modulation constellation [14] are used to estimate the RFF from signals. However, similar analog components also exist in the receivers, which brings errors in estimation. For example, power amplifiers, converters, and modulators (or demodulators) exist in both emitters and receivers. So the RFFs about them will contain the influence from the receiver. Besides, some components such as receiver filters and oscillators may cause a complex effect on multiple RFFs as well.

As for machine-learning approaches, the methods from speech recognition [15] and image recognition [16] are used to design the identification system, and the algorithms in machine learning such as convolutional neural network (CNN) [16], [17] and deep residual network [18] also show effectiveness in RFF extraction. However, like the traditional approaches, these approaches cannot solve the influence from the receivers. The performance of the machine learning

algorism mainly consists of the used data. But the data of cross-receiver SEI is usually hard to get. Besides, because most RFF are subtle and the influence of receiver distortions may be complex, the network that can distinguish the distortions from emitters and receivers is hard to establish.

To sum up, most previous methods cannot meet the requirements of cross-receiver SEI. The existing discussion about the receiver distortion mainly focus on its influence in single-receiver identification [19] and there is almost no proven solution about cross-receiver SEI so far. In the existing study, receiver distortions are often considered as a kind of error, and the method to reduce its influence is rarely considered.

To solve this problem, we believe that a method to reduce the influence of the receivers is the key. Comparing two technical categories in SEI, we hold the view that the traditional approaches are more advantageous for cross-receiver identification in some ways. This is because the RFF obtained by the mechanism model is artificially designed and have better interpretability, which makes it more beneficial to analyze the impact of receiver distortion. Therefore, our discussion is based on the traditional approaches.

This article analyzes one existing mechanism model and discovers that the receiver distortion can be estimated and processed to reduce its influence on the carrier leakage estimation. According to this discovery, a carrier leakage estimation method for cross-receiver SEI is presented. Then the effectiveness and performance of the method are demonstrated in the experiment.

The main contributions are summarized as follows:

- 1) The effect of receiver distortions on carrier leakage is discussed and the exact mathematical representation of the effect is derived. To the best of our knowledge, this attempt is the first time that the relationship between carrier leakage and receiver distortions is mathematically discussed.
- 2) A mechanism model containing receiver distortions is presented. This method can jointly estimate I/Q modulator fingerprints and receiver filter characteristics. To the best of our knowledge, this method is the first method that considers a way to correct the fingerprint with estimated filter characteristics.
- 3) The way to evaluate the features in cross-receiver SEI is proposed. We further investigate the performance of our method under different SNR and different receivers in simulation and validate the effectiveness with the experimental data from different receiver devices in a practical scenario.

The rest of the paper is organized as follows. Section 2 details a modulation-error-based SEI method and discusses the influence of receiver distortions. In Section 3, a carrier leakage estimation method for cross-receiver SEI is presented. In Section 4, the comparison of the performance between the previous method and our proposed method is discussed with simulation. In Section 5, a testbed is built to demonstrate the effectiveness of the proposed method in a practical scenario. Finally, the conclusion of the paper is given in Section 6.

II. A TRADITIONAL SEI METHOD AND RECEIVER DISTORTIONS

The discrepancies of emitter imperfections such as modulation error can provide distinguishability between emitters, whereas the discrepancies are affected by the receiver distortions. In this section, we introduce a typical traditional estimation method for modulation errors and then analyze the influence of receiver distortions on the estimation results.

A. INTRODUCTION OF A TRADITIONAL SEI METHOD

Many emitter distortions are used as RFF in the traditional SEI method and modulation errors are included in various methods. Modulation errors include I/Q gain imbalance, quadrature error, I/Q offset, and carrier leakage. The existing studies have established a proven solution about the modulation error in SEI [14], [20], [21]. In some studies, it is also called constellation errors.

The advantage of modulation errors is obvious. Modulation errors originate from the modulator and widely exist in the PSK signals, which means that it can be widely used as RFF. Besides, because the estimation of modulation errors is mainly based on system identification, the used mathematic tools are much easier, which means that the SEI system can have a simple structure and lower computational complexity.

More importantly, according to our discussion, the influence of receiver distortions on carrier leakage is easier to model than other RFFs, which makes it possible to find a method to reduce the influence of receivers.

The basic idea of a traditional SEI approach is as follows: Establish a reasonable signal model that can describe the distortion of the emitter, then use the received signal and demodulation results to estimate the distortion parameters in the model. This is the basic idea of system identification.

In [21], Wang introduced an SEI method based on this idea. The method can jointly obtain multiple modulation errors and its effectiveness has been proved by the actual SEI system. It is representative and our discussion is based on it. The details of the method are as follows:

As is mentioned above, a reasonable signal model that can describe the modulation error is basic. In this method, a complex baseband signal model containing variables about modulation error is established:

$$
z(t) = Ae^{j(2\pi f_0 t + \varphi)}(\mu_1 \rho(t) + \mu_2 \rho^*(t) + \xi)
$$
 (1)

where f_0 is the residual frequency offset, φ is the initial phase, ξ is the carrier leakage, *A* is the signal amplitude.

 $\rho(t)$ denotes the complex baseband signal and is defined as follows:

$$
\rho(t) = \sum_{n=-\infty}^{\infty} \left[\text{Re}\{c_n\} h(t - nT - \tau) + j \text{Im}\{c_n\} h(t - nT - \tau - \tau_d) \right]
$$
 (2)

where c_n is the transmission symbols, $h(t)$ denotes the transmitter filter, *T* is the duration of a symbol, τ is the delay of the filter, and τ_d denotes the delay between I path and Q path.

 μ_1 and μ_2 are the distortion parameters that describe the gain imbalance and quadrature error, denoted by:

$$
\mu_1 = 0.5(G_{I/Q} + 1)\cos(\frac{\zeta}{2}) + 0.5(G_{I/Q} - 1)\sin(\frac{\zeta}{2}) \tag{3}
$$
\n
$$
\mu_1 = 0.5(G - 1)\cos(\frac{\zeta}{2}) + 0.5(G - 1)\sin(\frac{\zeta}{2}) \tag{4}
$$

$$
\mu_2 = 0.5(G_{I/Q} - 1)\cos(\frac{5}{2}) + 0.5(G_{I/Q} + 1)\sin(\frac{5}{2})
$$
 (4)

where $G_{I/O}$ denotes the I/Q gain imbalance, ζ denotes the quadrature error.

The model in [\(1\)](#page-1-0) is in analog form and a step to transform it into a digital domain is necessary. Moreover, the model has to be simplified in this step to reduce the computation complexity. So Wang chooses to establish a digital signal model based on transmit symbols.

Define r_n as the received symbol (received signal at optimal sampling point). Then the model can be shown as:

$$
r_n = Ae^{j(2\pi f_0 nT + \varphi)}(\mu_1 c_n + \mu_2 c_n^* + \xi) + v_n \tag{5}
$$

where v_n is the zero-mean Gaussian white noise with variance σ_v^2 . The definition of other variables is the same as [\(1\)](#page-1-0).

In order to match the form of maximum likelihood estimation (MLE), convert [\(5\)](#page-2-0) into a vector form:

$$
\mathbf{U}(f_0)\mathbf{r} = \mathbf{G}\boldsymbol{\theta} + \mathbf{v} \tag{6}
$$

Here, $G = [c, c^*, 1], c = [c_0, \ldots, c_{N-1}]^T, 1 = [1, \ldots, 1]^T$. *N* is the number of samples used in estimation. θ = $Ae^{j\varphi}[\mu_1, \mu_2, \xi]^T$, it is the modulation distortion matrix. $U(f_0) = diag(1, e^{-j2\pi f_0}, \ldots, e^{-j2\pi f_0(N-1)})$, it is the vector that describes the effect of residual frequency offset. **r** is the vector form of *rⁿ* and it is a column vector.

From [\(6\)](#page-2-1), the MLE result of θ can be given by:

$$
\hat{\theta} = \left(\mathbf{G}^H \mathbf{G}\right)^{-1} \mathbf{G}^H (\mathbf{U}(f_0) \mathbf{r}) \tag{7}
$$

where $(\bullet)^H$ means conjugate transpose.

Here, the residual frequency offset $U(f_0)$ still remains unknown and an estimation is needed. Define **P** = **G**(**G**^{*H*}**G**)⁻¹**G**^{*H*}, **u**(*f*₀)=[1,e^{-*j*2 πf ₀},...,e^{-*j*2 πf ₀(*N*-1)]^{*T*},} $\mathbf{I}_N = diag(1, \ldots, 1)$ and $\mathbf{D}_r = diag(\mathbf{r})$, the MLE result of *f*₀ is given by:

$$
\hat{f}_0 = \arg\min_{f_0} \|(\mathbf{I}_N - \mathbf{P}) \mathbf{D}_r \mathbf{u}(f_0) \|^2 \tag{8}
$$

The calculation of [\(8\)](#page-2-2) can be done through FFT or CZT spectrum search. For details about this calculation, please refer to [21].

Substitute estimation of f_0 into [\(7\)](#page-2-3). And in an actual estimation mission, the demodulated symbols from demodulator $\hat{\mathbf{c}} = [\hat{c}_0, \dots, \hat{c}_{N-1}]^T$ can replace **c** to establish **G**. Now we can get the estimation value of θ .

 $\hat{\theta}$ contains the estimated distortion parameters $\hat{\mu}_1$, $\hat{\mu}_2$, $\hat{\xi}$. However, $\hat{\theta}$ also contains $Ae^{j\varphi}$, which means initial phase and amplitude. In order to eliminate the influence and get the correct estimation value of the distortions, a further step is needed.

VOLUME 9, 2021 26303

According to [\(3\)](#page-2-4) and [\(4\)](#page-2-4), the final result of gain imbalance and quadrature error can be given by:

$$
\hat{\zeta} = \arg\left(\frac{\hat{\theta}(1) + \hat{\theta}(2)}{\hat{\theta}(1) - \hat{\theta}(2)}\right)
$$
(9)

$$
\hat{G}_{I/Q} = \left| \frac{\hat{\theta}(1) + \hat{\theta}(2)}{\hat{\theta}(1) - \hat{\theta}(2)} \right| \tag{10}
$$

where $\hat{\theta}(p)$ denotes the *p*th element in the vector $\hat{\theta}$. $\hat{\zeta}$ denotes the estimation of quadrature error ζ . $\hat{G}_{I/Q}$ denotes the estimation of gain imbalance *GI*/*Q*.

Substitute [\(8\)](#page-2-2) and [\(9\)](#page-2-5) into [\(3\)](#page-2-4) and [\(4\)](#page-2-4) to obtain $\hat{\mu}_1$ and $\hat{\mu}_2$, then we can get the correct estimation of carrier leakage:

$$
\hat{\xi} = \hat{\theta}(3) \frac{|\hat{\mu}_1|^2 + |\hat{\mu}_2|^2}{\hat{\mu}_1^* \hat{\theta}(1) + \hat{\mu}_2^* \hat{\theta}(2)}
$$
(11)

Now we get the estimation results of carrier leakage and other modulation distortions. However, the influence of receiver distortions is out of consideration in this method and further discussion is needed. To simplify the expression, this method will be referred to as the ''traditional method'' in the following discussion.

B. THE EFFECT OF RECEIVER DISTORTIONS

The traditional method described in Section 1.A is influenced by receiver distortions. To eliminate the influence, the exact mathematical description of the influence is essential.

Receiver distortion can be divided into deterministic distortion and stochastic distortion. The former mainly refers to linear and nonlinear distortions caused by the imperfection of amplifiers and filters in the receiver. The latter mainly refers to various types of noise.

Both of the distortions can reduce the distinguishability of RFFs. Huang [19] has analyzed this influence and gives the upper bound and lower bound of RFF performance.

The offset caused by stochastic distortions originates from the random error in receivers, which means that this problem exists both in single-receiver identification and cross-receiver identification. Its randomness also means that it can hardly be eliminated by establishing a signal model. Hence, we hold the view that the key to eliminating stochastic distortions lies in the improvement of receiver design, which is not the main focus of this article.

Deterministic distortions can bring a deterministic offset to RFF and influences the identification performance, and they only exist in cross-receiver identification. Therefore, we hold the view that deterministic distortions are the main problem in cross-receiver identification and our following discussion is mainly about it.

The popularly used receivers can be divided into two kinds: super heterodyne receiver and homodyne receiver. Their components and distortions are different. Fig[.1](#page-3-0) shows their basic structure.

FIGURE 1. Basic structure of a super heterodyne receiver (a) and a homodyne receiver (b).

The main sources of deterministic distortion are as follows:

- 1) The distortion between I and Q path. This distortion originated from the subtle difference between components in the I path and Q path. This distortion is determined by the receiver's structure. It only exists in the homodyne receiver and can be avoided by using a receiver with other structures (e.g., super heterodyne receiver). To simplify the problem, our discussion will be based on the super heterodyne receiver and ignore I/Q distortion.
- 2) The nonlinear distortion of the power amplifiers. To meet the needs of the analog-to-digital converter (ADC) and improve the signal to noise ratio (SNR), multiple power amplifiers are used in receivers and their nonlinear distortion may cause interference. With the use of linearization technology, this distortion is not significant, which can also be demonstrated by our experiment' in Section 5. Thus we ignore this distortion.
- 3) The imperfect characteristics of filters on the receiving path. Because of the diversity in filter design and production, the subtle characteristics of filters cannot be completely the same. Some common indexes, such as the in-band jitter, can also bring errors in cross-receiver SEI, even when the filters are of the same model. This distortion is deterministic and widely exists in all kinds of receivers, which means that the discussion about the

method to reduce its influence on RFF is meaningful. Therefore, our discussion will focus on eliminating the effect of filter distortion.

4) Distortions from other components. Some other components, such as oscillators and ADC, may also bring error. These errors are mostly stochastic distortions, and their deterministic distortion such as DC bias and frequency offset can be easily removed by digital signal processing. Thus we ignore them in our discussion.

To sum up, our focus is to eliminate the effect of filter distortion and we believe that other distortions have little effect on this problem. The experiment with actual receivers and emitters will confirm our view in Section 5.

To analyze the influence of the receiver filter, a signal model that can describe the distortion of the filter is established.

Here, to simplify the discussion, assume that the channel is a multipath channel and model it as an FIR filter $h_c(t)$ with an order of *Mc*.

The filters in receiver can be modeled by a single FIR filter. We define $h_r(t)$ as the impulse response of this equivalent filter and define M_r as its order. Define $z_e(t)$ as the signal at the emitter antenna, $z_r(t)$ as the signal at the receiver antenna, $z_h(t)$ as the analog signal processed by the receiver. In order to facilitate the derivation, $z_e(t)$, $z_r(t)$ and $z_h(t)$ are transformed into baseband form and their relationship can be given by:

$$
z_h(t) = z_r(t) \otimes h_r(t)
$$

= $z_e(t) \otimes h_c(t) \otimes h_r(t)$ (12)

According to the model given by (1) , $z_e(t)$ is in the same form with $z(t)$, then the received signal can be represented as:

$$
z_h(t) = z_e(t) \otimes h_c(t) \otimes h_r(t)
$$

= $e^{j(2\pi f_0 t + \varphi)}(\mu_1 \rho(t) + \mu_2 \rho^*(t) + \xi) \otimes h_c(t) \otimes h_r(t)$ (13)

The definition of variables in [\(13\)](#page-3-1) are the same as [\(1\)](#page-1-0). Define $h_R(t) = h_c(t) \otimes h_r(t)$, then we can get:

$$
z_h(t) = e^{j(2\pi f_0 t + \varphi)} (\mu_1 \rho(t) \otimes h_R(t)
$$

$$
+ \mu_2 \rho^*(t) \otimes h_R(t) + \xi \otimes h_R(t)) \quad (14)
$$

This expression is in analog form and a step to transform it into digital form is needed. Define *r^h* as the received symbol (received digital signal at optimal sampling point). Then *r^h* can be represented as:

$$
r_h = \sum_{k=0}^{M} r_{n-k} h_R(kT)
$$

= $e^{j(2\pi f_0 nT + \varphi)} A[\mu_1 \sum_{k=0}^{M} c_{n-k} h_R(kT)$
+ $\mu_2 \sum_{k=0}^{M} c_{n-k}^* h_R(kT) + \xi \sum_{k=0}^{M} h_R(kT)] + v$ (15)

26304 VOLUME 9, 2021

Define $c_k = [c_{-k}, c_{1-k}, \dots, c_{N-k}]^T$, $G_k = [c_k, c_k^*, 1]$, $\theta = Ae^{j\varphi}[\mu_1, \mu_2, \xi]^T$ and $h_k = h_R(kT)$. Then [\(12\)](#page-3-2) can be further written in vector form as:

$$
\mathbf{r}_h = \sum_{k=0}^{M} h_k \mathbf{G}_k \boldsymbol{\theta} + \mathbf{v}
$$
 (16)

Based on the above signal model, the effect of filter distortion can be derived. The details of the derivation are given in the Appendix A, and the result is given as follows:

$$
\lim_{N \to \infty} E\begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\xi} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \xi \sum_{k=0}^{M} h_k \end{bmatrix}
$$
(17)

Equation [\(17\)](#page-4-0) shows that, considering the influence of the filter, μ_1 and μ_2 obtained by the traditional method is still a progressive unbiased estimate of its true value. However, the traditional method obtains a biased estimate of ξ . When *N* increases, the estimation of ξ tends to be closer to the product of carrier leakage and the amplitude-phase response of the equivalent filter $h_R(t)$ at zero frequency. Besides, the result also shows that the influence of the multipath channel can be discussed together with the filter distortions in the receiver.

Obviously, the estimation result of ξ is not steady with the change of the filter characteristics of the receiver and cannot meet the requirements in cross-receiver identification. But [\(17\)](#page-4-0) also offers an exact mathematical description of the relationship between estimation results and filter distortions, which makes it possible to eliminate the influence with an improved model.

III. CARRIER LEAKAGE ESTIMATION METHOD FOR CROSS-RECEIVER SEI

Equation [\(17\)](#page-4-0) indicates that the result of the traditional method is not steady in cross-receiver identification. To solve this problem, in this section, we propose a cross-receiver carrier leakage estimation method (cross-receiver method). In our method, we establish a model to jointly estimate the carrier leakage and the filter characteristics, then a cross-receiver identification scheme is given.

A. CROSS-RECEIVER SEI METHOD

A signal model that can describe the distortion characteristics of the filter needs to be established. To describe the filter characteristics accurately, it is necessary to increase the input sample density and introduce the parameters of the filter characteristics. Based on [\(6\)](#page-2-1), we construct an over-sampling baseband signal model with filter characteristic parameters:

Define *L* as the length of the filter in symbols and *P* as the over-sampling rate, then the length of filter in samples will be $M = L \cdot P$. Define N as the number of symbols used in estimation, then the number of used samples is *N* • *P*. Define $\mathbf{g} = [g_0, \dots, g_{LP}]^T$ as the filter coefficient vector. Then we can get a signal model:

$$
\mathbf{U}(f_0)\mathbf{y} = Ae^{\varphi}(\mu_1 \mathbf{Cg} + \mu_2 \mathbf{C}^* \mathbf{g} + \mathbf{\xi}_g) + \mathbf{v}
$$
 (18)

where **y** denotes the input signal vector with a length of *N* •*P* and it's a column vector. **C** denotes the over-sampling symbol matrix and its elements can be expressed as:

$$
c(i,j) = \begin{cases} c_{I-J+L}, & i\%P = j\%P\\ 0, & else \end{cases} \tag{19}
$$

where c_n denotes the transmission symbols, $0 \le i \le NP - 1$, $0 \le j \le LP$. Here, ξ_g denotes the carrier leakage with the influence of the filter and can be expressed as:

$$
\xi_g = [\xi \sum_{i=1}^{LP} g(i), \dots, \xi \sum_{i=1}^{LP} g(i)]_{NP}^T
$$

= $\xi_g [1, \dots, 1]_{NP}^T$ (20)

The definition of other variables in [\(18\)](#page-4-1) is the same as [\(5\)](#page-2-0).

Define $\mathbf{g}_{\mu,\xi} = Ae^{\phi}[\mathbf{g}\mu_1; \mathbf{g}\mu_2; \xi_g]$ and $\tilde{\mathbf{C}} = [\mathbf{C}, \mathbf{C}^*, \mathbf{1}],$ then [\(18\)](#page-4-1) can be rewritten as:

$$
\mathbf{U}(f_o)\mathbf{y} = \tilde{\mathbf{C}}\mathbf{g}_{\mu,\xi} + \mathbf{v} \tag{21}
$$

The maximum likelihood estimate of $\mathbf{g}_{\mu,\xi}$ can be obtained as:

$$
\hat{\mathbf{g}}_{\mu,\xi} = (\tilde{\mathbf{C}}^H \tilde{\mathbf{C}})^{-1} \tilde{\mathbf{C}}^H (\mathbf{U} (\hat{f}_o)^H \mathbf{y})
$$
(22)

 $\hat{\mathbf{g}}_{\mu,\xi}$ contains the offset caused by the filter distortion. In order to extract the features independent of the receiver, we define a new feature vector as:

$$
\hat{\theta}_C = \frac{\hat{\mathbf{g}}_{\mu,\xi}(3)}{\mathbf{1}^T \hat{\mathbf{g}}_{\mu,\xi}(1)} = \frac{A e^{\phi} \hat{\xi}_g}{A e^{\phi} \mu_1 \sum_{i=1}^{LP} \hat{g}(i)} = \frac{\hat{\xi}}{\mu_1}
$$
(23)

Equation [\(20\)](#page-4-2) shows that the new feature is related to carrier leakage and modulation distortion coefficient μ_1 . Theoretically, when using a super heterodyne receiver, μ_1 is not affected by receivers and the result will be the unbiased estimation of carrier leakage.

Then, an SEI system can be established. The diagram is shown in Fig[.2.](#page-4-3) The proposed system contains the same components as the traditional method. The cross-receiver function can be achieved by replacing the algorism used in RFF estimation.

B. TIME COMPLEXITY ANALYSIS

In the proposed method, the input sample density has been increased and the parameters to be estimated have become more than the traditional method, which will increase time complexity. Thus, we analyze the time complexity of the proposed method.

The main origination of time complexity is in the RFF estimation, which is shown in [\(7\)](#page-2-3) and [\(22\)](#page-4-4). The other steps such as the estimation of residual frequency offset f_0 or the calculation to normalize the RFFs will bring little difference in time complexity. So we ignore them in this analysis.

In [\(7\)](#page-2-3), **G** is a matrix with the size of $N \times 3$, r is a vector with length *N*. Then the time complexity of the traditional method can be derived:

$$
2O(9N) + O(3N)
$$
 (24)

In [\(22\)](#page-4-4), **C** is a matrix with the size of $NP \times LP$, *y* is a vector with length *NP*. Then the time complexity of our proposed method can be derived:

$$
2O((2LP + 1)^2 NP) + O((2LP + 1)NP)
$$
 (25)

The result can be further simplified into *O* (9*N*) and $O((2LP + 1)^2 NP)$.

RFF estimation in the proposed method requires much more calculation than the traditional one. The increase in complexity mainly comes from the over-sampling rate *P* and the length of the filter to be estimated *L*. If set $P = 1$ and $L = 1$, the proposed method will be the same as the traditional one. In this case, the function to eliminate the influence of receiver distortions is lost as well.

To sum up, our proposed method brings the cross-receiver ability with the cost of calculation speed. We hold the view that the cost is acceptable because the cost of the whole system is much lower than the limit of the existing device and there is no requirement for real-time performance for some of the SEI applications.

IV. SIMULATION AND ANALYSIS

In this section, a simulation is conducted to validate the feasibility and performance of our proposed method.

The simulation is organized into two parts. In the first part, the estimation results of the traditional method and the theoretical results from [\(17\)](#page-4-0) are compared. This is to validate our derivation in Section 2. In the second part, signals with different signal-to-noise ratios (SNR) and different filter distortions are simulated. And the RFF from the traditional method and the proposed method is analyzed to evaluate the performance of our method.

A. SIMULATION SETUP

1) RECEIVE FILTER GENERATION SETUP

In order to generate the imperfect filters, the module that can express the filter distortion is essential. In general, the filter distortions can be expressed by the following formula:

$$
G(f) = H(f)A(f)e^{j\phi(f)}
$$
\n(26)

TABLE 1. Distortion parameters of receive filter used in simulation.

Filter ID	α_k	β_k	a_k	b_k
А				
в	0.05	0.006	0.1	0.2
C	0.05	0.006	0.2	0.2
D	0.05	0.006	0.3	0.2
E	0.05	0.006	0.1	0.4
F	0.05	0.006	0.3	0.4
G	0.05	0.006	0.3	0.4

where $H(f)$ denotes the frequency response of a standard filter. $A(f)$ denotes the amplitude distortion and can be expanded into Fourier series as:

$$
A(f) = a_0 + a_k \cos(2\pi \alpha_k f) \tag{27}
$$

 $\phi(f)$ denotes phase distortion and can be expressed as:

$$
\phi(f) = 2\pi b_0 f + b_k \sin(2\pi \beta_k f) \tag{28}
$$

We can generate receive filters with different distortions by set the values of distortion parameters α_k , β_k , a_k and b_k .

We choose the band-pass FIR filter with a passband from 1.2 MHz to 2.8 MHz as the standard filter. The filter order is 800. Seven filters are simulated with different distortions according to the parameters in Table[.1.](#page-5-0)

2) SIGNALS GENERATION SETUP

The modulation mode of the simulation signal is QPSK, the transmitted symbol is a random sequence, the symbol rate is 1 MBaud/s, the carrier frequency is 2 MHz, the sampling rate is 8 MSps. SNR is defined as the ratio of signal energy to noise power spectral density. The shaping filter is a root-raised cosine filter with a roll-off coefficient of 0.35.

To reduce the contingency, the carrier leakage is set to random values that obey uniform distribution from 0+0i to 0.02+0.02i. (To show the result of feature extraction directly, carrier leakage is set to a fixed value as $\xi_{dc} = 0.01 + 0.01i$ when printing Fig. 2.).

In this way, 300 groups of signals are generated. Each group contains signals with the same carrier leakage value but different filters and signal-to-noise ratios. Different groups have different carrier leakage. One signal contains 10,000 symbols.

3) FEATURE EXTRACTION SETUP

Feature extraction is based on the algorithms in Section 2 and Section 3. It is necessary to set the signal length and filter length *M* used in the estimation.

To increase the accuracy of the correction as much as possible, the length of the estimated filter needs to be increased as long as possible. Generally speaking, a filter with a length of about 10 symbols can meet the demand. Therefore, we set a longer length in the experiment to increase the accuracy of estimation. We set the length of the filter to 18 symbols, and in the case of an oversampling rate of 8, the order of estimated

FIGURE 3. The feature distribution of carrier leakage from traditional method and cross-receiver method (SNR = 31dB).

filter *M* is 144. Carrier leakage and filter tap coefficients are estimated every 10,000 symbols.

4) RESULT EVALUATION

The performance of the estimation results can be evaluated with two indexes, the variance of the features and the feature offset. The variance can describe the dispersion of the features. The feature offset can describe the drift caused by filter distortions. The feature offset is defined as follows:

$$
\Pi(\Omega_i, \Omega_j) = \frac{1}{N_i N_j} \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} d(\mathbf{x}_k^{(i)}, \mathbf{x}_l^{(j)})
$$
(29)

where $\mathbf{x}_1^{(i)}$ $\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_{N_i}^{(i)}$ N_i are the samples in Ω_i , N_i denotes the number of samples in Ω_i , $d(\mathbf{x}_k^{(i)})$ $\mathbf{x}_l^{(i)}$, $\mathbf{x}_l^{(j)}$ $\binom{0}{l}$ denotes Euclidean distance of two samples. In our experiment, Ω_i and Ω_j should be the set of symbols with different filter distortions.

Furthermore, before calculating the offset, the samples should be normalized with the true value of carrier leakage or the average of the estimated values. The accuracy of identification is not included in the simulation. Because the main idea of this article is to evaluate the feature drift caused by receiver distortions, and feature offset can show the feature drift more directly. However, the accuracy is an effective index to evaluate the distinguishability of the feature, thus it will be used in the experiment on the testbed in Section 5.

B. RESULTS OF FEATURE EXTRACTION

The feature distribution of the traditional method and cross-receiver method is given in Fig[.3.](#page-6-0) Here, to show the result of feature extraction directly, carrier leakage is set to a fixed value as $\xi_{dc} = 0.01 + 0.01i$.

To validate the derivation in Section 1, the theoretical feature distribution is calculated according to [\(17\)](#page-4-0) and signed in Fig[.3\(](#page-6-0)a) as well. These theoretical values (''in theory'') are the carrier leakage influenced by the receive filters, thus there should be seven points. If our derivation is correct, these points should be close to the result of feature extraction. The SNR of the used signal is 31dB.

It is shown in Fig[.3\(](#page-6-0)a) that the estimation results of the traditional method are significantly affected by the filter, and the theoretical value given by [\(17\)](#page-4-0) are in good agreement with the simulation results. We can also see that the results of the cross-receiver method are independent with the filter in Fig. $3(b)$.

C. PERFORMANCE EVALUATION

We select four filters numbered A, B, C, and D from Table[.1](#page-5-0) to filter the signals, and the amplitudes of the estimation results versus SNR are given with boxplot in Fig[.4](#page-7-0) and the medians of the estimation results are connected with curves. The carrier leakage is set to random values that obey uniform distribution from 0+0i to 0.02+0.02i. The minimum value of the simulated SNR is set as 5 dB. Because the SNR lower than 5dB will result in an unacceptable error rate in demodulation.

Then the feature offset and variance of the extracted features are calculated. Feature offsets versus SNR are given in Fig[.5.](#page-7-1) The variances of the feature extraction results versus SNR are given in Fig[.6.](#page-8-0)

The following conclusions can be obtained from these figures:

1) The cross-receiver method can effectively eliminate the influence of the receiver filter. It is shown in Fig[.4](#page-7-0) and Fig[.5](#page-7-1) that the traditional method is significantly affected by the filter, while the results of

FIGURE 4. The distribution of estimation results from traditional method(a) and cross-receiver method(b) versus SNR.

FIGURE 5. Feature offset of results from traditional method and cross-receiver method versus SNR.

the cross-receiver method are similar no matter which filters or SNR is used. This can be directly seen from the feature offset. The feature offset of the traditional method will be largely influenced by the used filters, but that of the cross-receiver method is always at a very low level.

FIGURE 6. The variance of features from traditional method and cross-receiver method versus SNR.

- 2) The correlation between the performance of the cross-receiver model and the receive filter is low. It is shown in Fig[.5](#page-7-1) that the feature offset of the cross-receiver method is close to zero in all filters. In other words, the proposed method shows good applicability to different filters. This means that it is not necessary to design special filters in practical cross-receiver identification.
- 3) The dispersion of the estimation results of the cross-receiver model is similar to the traditional model. It is shown in Fig[.6](#page-8-0) that there is little difference in feature dispersion between the two methods. This means that the features obtained by the cross-receiver method have similar distinguishability as the traditional method.

V. EXPERIMENT ON TESTBED

In this section, to verify the effectiveness of the proposed model in a practical scenario, a testbed is built with universal software radio peripheral (USRP) and super heterodyne receivers. Here, the practical scenario means the condition in which the classifier is trained by the RFFs from one receiver but the identified targets are from another receiver.

A. EXPERIMENT SETUP

The testbed includes five USRP devices and two super heterodyne receivers. All the devices are shown in Fig[.7](#page-8-1) and the diagram of the experiment system is shown in Fig[.8.](#page-9-0)

FIGURE 7. Devices used in experiment.

Two X310s and three B210s are used as emitters (targets to be identified). The emitters transmit QPSK signals with a carrier frequency of 1 GHz and a symbol rate of 1 Baud/s. The transmitted symbol is a random sequence. Two super heterodyne receivers are used with a sampling rate of 80 MSps. The signals were received by two receivers at the same time with the help of a splitter. The signals are transmitted in a line-of-sight condition.

10⁶ symbols were transmitted for each emitter. Then every 10⁴ symbols are set as one sample to estimate their RFF. Thus in the input of the classifier, there will be 5 classes and each class contains 100 samples. The other settings in

TABLE 2. Variance, accuracy and offset of the features obtained with traditional method and cross-receiver method.

FIGURE 8. Diagram of the experiment system.

preprocessing and feature extraction are the same as the simulation setup.

The RFF used in identification is carrier leakage. The real part and imaginary part are used as two dimensions of the input of the classifier. The extracted features are classified by the SVDD open-set classification method. To meet the actual cross-receiver condition, the RFFs of the signals from Receiver A (RFF *A*) are used as train data and that from Receiver B (RFF *B*) are used as test data.

The feature variance, feature offset and the accuracy of identification is calculated as performance indexes. The identification of feature offset is shown in Section 4.

B. RESULTS AND ANALYSIS

The accuracy and feature offset of the traditional method and cross-receiver method are given in Table[.2.](#page-9-1)

The following conclusions can be obtained from the result:

- 1) The feature variance of traditional method and cross-receiver method are similar. This result is in good agreement with the simulation result in Fig[.6.](#page-8-0) This phenomenon verifies our analysis in the simulation and shows that the features obtained by cross-receiver method have similar distinguishability as the traditional method.
- 2) The features obtained by our proposed method have better stability in cross-receiver identification. The feature offset of the cross-receiver method is reduced to lower than 0.4%, while the value is about 6% in the traditional method.

Besides, it is noteworthy that the feature offset of the cross-receiver method is very low but not zero, this is because the non-linear distortions in receivers can also affect the features, but we can see from the value that this influence is very limited.

3) The identification accuracy of the cross-receiver method is better than the traditional method. This phenomenon verifies the effectiveness of our proposed method. However, it is noteworthy that the improvement is not obvious but only 8%-10%. This is because the used two receivers have the same structure and feature drift caused by filter distortions is not very significant. When using receivers in a different structure, the improvement will be much more obvious.

VI. CONCLUSION

This article analyzes the influence of the receiver filter distortions on the existing carrier leakage estimation method and a cross-receiver carrier leakage estimation method is proposed. The effectiveness of the new method is illustrated in the experiment, which shows that the feature offset can be reduced to a very little value and the identification accuracy is improved by up to 10%.

The experimental results show that the proposed method can effectively reduce the drift of carrier leakage caused by receiver filter distortions in cross-receiver identification, and the feature from the new method has similar distinguishability and reliability as the traditional method. In cross-receiver SEI, our method can be used to extract the precise carrier leakage feature which is independent of receivers without any extra design requirement on receivers. The feature can be jointly used with modulation distortions to establish a cross-receiver feature library to identify the target signals from different receivers.

The receiver distortions can influence almost all the features used in the previous SEI method, investigation of the influence on other features will be our future work.

APPENDIX A

THE DERIVATION IN SECTION 2

In this appendix, we will show the details of the derivation about [\(17\)](#page-4-0).

Substitute [\(16\)](#page-4-5) into [\(7\)](#page-2-3), we can arrive at:

$$
\hat{\theta} = (\mathbf{G}_0^H \mathbf{G}_0)^{-1} \mathbf{G}_0^H (\sum_{k=0}^M h_k \mathbf{G}_k \theta + \mathbf{v})
$$

\n
$$
= \sum_{k=0}^M h_k (\mathbf{G}_0^H \mathbf{G}_0)^{-1} \mathbf{G}_0^H \mathbf{G}_k \theta + (\mathbf{G}_0^H \mathbf{G}_0)^{-1} \mathbf{G}_0^H \mathbf{v}
$$

\n
$$
= \theta + \sum_{k=1}^M h_k (\mathbf{G}_0^H \mathbf{G}_0)^{-1} \mathbf{G}_0^H \mathbf{G}_k \theta + (\mathbf{G}_0^H \mathbf{G}_0)^{-1} \mathbf{G}_0^H \mathbf{v} \quad (30)
$$

$$
\begin{bmatrix}\n\hat{\mu}_{1} \\
\hat{\mu}_{2} \\
\hat{\xi}\n\end{bmatrix} = \begin{bmatrix}\n\mu_{1} \\
\mu_{2} \\
\mu_{3}\n\end{bmatrix} + \left(\mathbf{G}_{0}^{H}\mathbf{G}_{0}\right)^{-1}\mathbf{G}_{0}^{H}\mathbf{v} + \frac{\mu_{1}\sum_{k=1}^{M}h_{k}\begin{bmatrix}\nN^{2}R - (P_{0}^{*})^{2}R + P_{0}P_{0}^{*}S^{*} - NQ^{*}S^{*} + P_{0}^{*}Q^{*}P_{k} - NP_{0}P_{k} \\
P_{0}^{*}QR - NQ_{0}R + N^{2}S^{*} - P_{0}^{2}S^{*} + P_{0}QP_{k} - NP_{0}^{*}P_{k} \\
P_{0}^{*}QR - NP_{0}R + P_{0}Q^{*}S^{*} - NP_{0}^{*}S^{*} + N^{2}P_{k} - QQ^{*}P_{k}\end{bmatrix}}{\nN^{3} + P_{0}P_{0}^{*}(Q + Q^{*}) - N(P_{0}^{2} + QQ^{*} + (P_{0}^{*})^{2})}\n+ \frac{\mu_{2}\sum_{k=1}^{M}h_{k}\begin{bmatrix}\nN^{2}S - (P_{0}^{*})^{2}S + P_{0}P_{0}^{*}R^{*} - NQ^{*}R^{*} + P_{0}^{*}Q^{*}P_{k}^{*} - NP_{0}P_{k}^{*} \\
P_{0}^{*}QS - NP_{0}S + N^{2}R^{*} - P_{0}^{2}R^{*} + P_{0}Q^{*}P_{k}^{*} - NP_{0}^{*}P_{k}^{*} \\
P_{0}^{*}QS - NP_{0}S + P_{0}Q^{*}R^{*} - NP_{0}^{*}R^{*} + N^{2}P_{k}^{*} - QQ^{*}P_{k}^{*} \\
N^{3} + P_{0}P_{0}^{*}(Q + Q^{*}) - N(P_{0}^{2} + QQ^{*} + (P_{0}^{*})^{2})\n\end{bmatrix} + \frac{\sum_{k=1}^{M}h_{k}\begin{bmatrix}\nN^{2}P_{0} - (P_{0}^{*})^{2}S + P_{0}P_{0}R^{*} + NP_{0}Q^{*} - NP_{0}P_{k}^{*} \\
P_{0}P
$$

Define $P_k = \sum_{k=1}^{N} P_k$ $\sum_{m=1}^{N} c_k(m), Q = \mathbf{c}_0^H \mathbf{c}_0^*, R = \mathbf{c}_0^H \mathbf{c}_k, S = \mathbf{c}_0^H \mathbf{c}_k^*,$ $\Theta = \left(\mathbf{G}_0^H \mathbf{G}_0\right)^{-1} \mathbf{G}_0^H \mathbf{G}_k \boldsymbol{\theta}$. Then we can get:

$$
\Theta = \begin{pmatrix} N & Q & P_0 \\ Q^* & N & P_0^* \\ P_0 & P_0^* & N \end{pmatrix}^{-1} \begin{pmatrix} R & S & P_0 \\ S^* & R^* & P_0^* \\ P_k & P_k^* & N \end{pmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \xi \end{bmatrix}
$$
 (31)

Substitute [\(31\)](#page-10-0) into [\(30\)](#page-9-2), we can arrive at [\(32\)](#page-10-1), as shown at the top of the page.

Apparently, if **c** is a random sequence, P_k , Q , R , S are bounded when $N \to \infty$. Then the expectation of θ can be derived as [\(33\)](#page-10-1), as shown at the top of the page.

Here, we can get the relationship between the estimated features $\theta = [\hat{\mu}_1, \hat{\mu}_2, \hat{\xi}]$ and the original feature $\theta = [\mu_1, \mu_2, \xi].$

REFERENCES

- [1] J. Yu, A. Hu, G. Li, and L. Peng, ''A robust RF fingerprinting approach using multisampling convolutional neural network,'' *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6786–6799, Aug. 2019.
- [2] O. Ureten and N. Serinken, ''Wireless security through RF fingerprinting,'' *Can. J. Electr. Comput. Eng.*, vol. 32, no. 1, pp. 27–33, 2007.
- [3] R. G. Wiley, *ELINT: The Interception and Analysis of Radar Signals*. London, U.K.: Artech House, 2006.
- [4] K. Kim, C. M. Spooner, I. Akbar, and J. H. Reed, "Specific emitter identification for cognitive radio with application to IEEE 802.11,'' in *Proc. IEEE Global Telecommun. Conf.*, New Orleans, LA, USA, Nov. 2008, pp. 1–5.
- [5] Z. Zhang, K. Long, and J. Wang, "Self-organization paradigms and optimization approaches for cognitive radio technologies: A survey,'' *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 36–42, Apr. 2013.
- [6] K. I. Talbot, P. R. Duley, and M. H. Hyatt, ''Specific emitter identification and verification,'' *Technol. Rev. J.*, pp. 113–133, Jan. 2003.
- [7] B. Danev, D. Zanetti, and S. Capkun, ''On physical-layer identification of wireless devices,'' *ACM Comput. Surv.*, vol. 45, no. 1, pp. 1–29, 2012.
- [8] A. C. Polak and D. L. Goeckel, ''Wireless device identification based on RF oscillator imperfections,'' *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2492–2501, Dec. 2015.
- [9] M.-W. Liu and J. F. Doherty, ''Specific emitter identification using nonlinear device estimation,'' in *Proc. IEEE Sarnoff Symp.*, Princeton, NJ, USA, Apr. 2008, pp. 1–5.
- [10] F. Zhuo, Y. Huang, and J. Chen, ''Radio frequency fingerprint extraction of radio emitter based on I/Q imbalance,'' *Procedia Comput. Sci.*, vol. 107, pp. 472–477, 2017.
- [11] 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, Sep. 2014.
- [12] J. Zhang, F. Wang, Z. Zhong, and O. Dobre, ''Novel Hilbert spectrumbased specific emitter identification for single-hop and relaying scenarios,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.
- [13] C. Bertoncini, K. Rudd, B. Nousain, and M. Hinders, ''Wavelet fingerprinting of radio-frequency identification (RFID) tags,'' *IEEE Trans. Ind. Electron.*, vol. 59, no. 12, pp. 4843–4850, Dec. 2012.
- [14] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless device identification with radiometric signatures,'' in *Proc. 14th ACM Int. Conf. Mobile Comput. Netw.*, 2008, pp. 116–127.
- [15] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication networks,'' *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 160–167, Feb. 2018.
- [16] L. Ding, S. Wang, F. Wang, and W. Zhang, ''Specific emitter identification via convolutional neural networks,'' *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2591–2594, Dec. 2018.
- [17] M. Kulin, T. Kazaz, I. Moerman, and E. De Poorter, "End-to-end learning from spectrum data: A deep learning approach for wireless signal identification in spectrum monitoring applications,'' *IEEE Access*, vol. 6, pp. 18484–18501, 2018.
- [18] Y. Pan, S. Yang, H. Peng, T. Li, and W. Wang, ''Specific emitter identification based on deep residual networks,'' *IEEE Access*, vol. 7, pp. 54425–54434, 2019.
- [19] Y. Huang and H. Zheng, ''Theoretical performance analysis of radio frequency fingerprinting under receiver distortions,'' *Wireless Commun. Mobile Comput.*, vol. 15, no. 5, pp. 823–833, Apr. 2015.
- [20] Y. Huang and H. Zheng, ''Radio frequency fingerprinting based on the constellation errors,'' in *Proc. 18th Asia–Pacific Conf. Commun.*, Oct. 2012, pp. 900–905.
- [21] G. Wang and Y. Huang, ''A novel special emitter identification method based on improved subclass discriminant analysis,'' in *Proc. IEEE 2nd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Chongqing, China, Mar. 2017, pp. 114–118.

MENGKAI SHI was born in Weihai, China, in 1996. He received the degree in communications engineering in 2018. He is currently pursuing the master's degree with Strategic Support Force Information Engineering University. His main research interest includes specific emitter identification.

GUILIANG WANG was born in Weifang, China, in 1987. He received the Ph.D. degree in communication and information systems from the National Key Laboratory of Science and Technology on Blind Signal Processing, in 2017. He is currently an Assistant Researcher with the National Key Laboratory of Science and Technology on Blind Signal Processing. His main research interests include specific emitter identification and signal processing.

 $\ddot{\bullet}$ $\ddot{\bullet}$ $\ddot{\bullet}$

YUANLING HUANG was born in Hunan, China, in 1981. He received the Ph.D. degree in communication and information systems from the National Key Laboratory of Science and Technology on Blind Signal Processing, in 2013. He is currently an Associate Researcher with the National Key Laboratory of Science and Technology on Blind Signal Processing. His main research interests include specific emitter identification and signal processing.