SPECIFIC MULTI-EMITTER IDENTIFICATION VIA MULTI-LABEL LEARNING

Yuhao Chen, Boxiang He, Shilian Wang, Jing Lei

College of Electronic Science and Technology, National University of Defense Technology, China

ABSTRACT

Specific emitter identification leverages hardware-induced impairments to uniquely determine a specific transmitter. However, existing approaches fail to address scenarios where signals from multiple emitters overlap. In this paper, we propose a specific multi-emitter identification (SMEI) method via multi-label learning to determine multiple transmitters. Specifically, the multi-emitter fingerprint extractor is designed to mitigate the mutual interference among overlapping signals. Then, the multi-emitter decision maker is proposed to assign the all emitter identification using the previous extracted fingerprint. Experimental results demonstrate that, compared with baseline approach, the proposed SMEI scheme achieves comparable identification accuracy under various overlapping conditions, while operating at significantly lower complexity. The significance of this paper is to identify multiple emitters from overlapped signal with a low complexity.

***Index Terms*—** Multi-label, overlapping signal, specific emitter identification

1. INTRODUCTION

Specific emitter identification (SEI) uses inherent hardware imperfections to identify devices at the physical layer [1]. Most SEI studies emphasize individual identification and overlook concurrent device transmissions [2, 3], despite multiple emitters operating simultaneously. For example, the industrial, scientific, and medical (ISM) band is often subject to simultaneous emissions from multiple devices. These emitters typically occupy adjacent sub-bands and transmit within closely spaced time intervals, leading to energy superposition that manifests as spectral overlap and multi-source signal mixtures [4]. In practical scenarios, the received signal comprises a mixture of emissions from multiple sources, motivating the extension of SEI to specific multi-emitter identification (SMEI), which aims to identify the entire set of active devices rather than a single emitter.

Enumerating all possible active combinations leads to exponential growth of the theoretical output space as the number of devices K increases, since the number of non-empty subsets equals $2^K - 1$. This exponential inflation not only escalates training and inference costs, but also introduces data-combination sparsity and degrades generalization to previously unseen combinations [5]. Furthermore, receiver noise can severely contaminate the subtle emitter-specific distortion features, making SMEI particularly challenging [6].

In multi-label classification, each input sample can simultaneously belong to multiple classes, unlike traditional single-label classification where each sample is associated with only one class. Multi-label classification can significantly reduce the number of model parameters by reducing the number of combinations to consider, thus decreasing the consumption of computational resources during training and inference [7]. This framework naturally models scenarios such as multi-emitter identification, where a received signal may contain contributions from several active emitters, requiring the system to assign multiple labels to a single observation. Here, we propose a SMEI method via multi-label learning. Specifically, we first model the SMEI problem. Then, we design a multi-emitter feature extractor to capture multi-emitter features, followed by a decision-making process that assigns emitter labels based on the extracted features. Finally, numerical results show that our scheme has the promising performance in terms of the identification accuracy and the complexity. The main contribution of this paper is the introduction of multi-label learning to address the challenging SMEI problem with extremely low complexity.

2. SYSTEM MODEL

We consider up to K potential emitters, where the received signal is denoted by

$$y \;=\; \sum_{m\in S} h_m\, f_m\!\big(x_m\big) \;+\; w, \tag{1}$$

where S is the set of active devices; x_m denotes the original transmitted of the m-th emitter; $f_m(\cdot)$ is the distortion function of the m-th emitter; h_m denotes the channel coefficient; w is Gaussian noise.

Following the literature [8], we model the distortion function $f_m(\cdot)$ using the I/Q imbalance, the spurious tone, the carrier leakage, and the power amplifier (PA) nonlinearity. Specifically, the distorted signal with the I/Q imbalance is expressed as

$$x'_m = \;\mu_m\, x_m \;+\; \nu_m\, x^*_m, \tag{2}$$

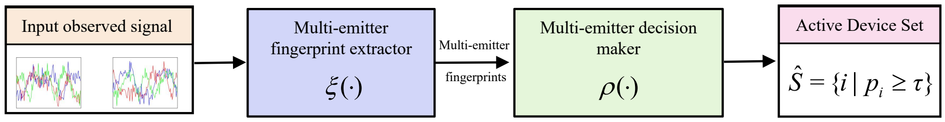

Fig. 1. SMEI scheme via multi-label learning.

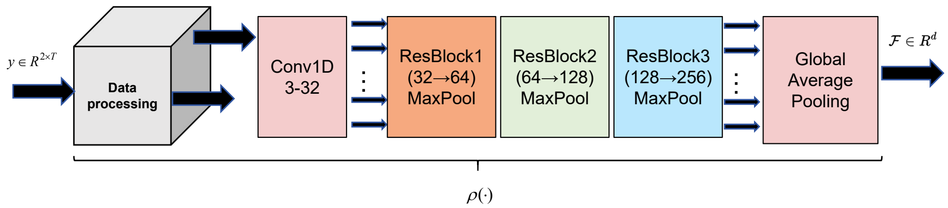

Fig. 2. Schematic diagram of the multi-emitter fingerprint extractor.

where μ_m and ν_m are the parameters used to describe the distortion of the modulator, which can be represented as

$$\mu_m = \tfrac{1}{2}(G_m + 1)\cos\left(\tfrac{\zeta_m}{2}\right) + j\,\tfrac{1}{2}(G_m - 1)\sin\left(\tfrac{\zeta_m}{2}\right), \quad (3)$$

$$\nu_m = \tfrac{1}{2}(G_m - 1)\cos\left(\tfrac{\zeta_m}{2}\right) + j\,\tfrac{1}{2}(G_m + 1)\sin\left(\tfrac{\zeta_m}{2}\right), \quad (4)$$

where ζ_m is the phase bias and G_m denotes the gain imbalance. With the spurious tone and carrier leakage, the distorted signal is further formulated by

$$x''_m = x'_m e^{j2\pi ft} + a_m^{ST}\, e^{j2\pi(f_m + f_m^{ST})} + \zeta_m\, e^{j2\pi ft}, \quad (5)$$

where a_m^{ST} and f_m^{ST} are the amplitude and the frequency of the spurious tone; $\zeta_m e^{j2\pi ft}$ is the carrier leakage.

Finally, the signal x''_m is fed into the PA with the nonlinear distortion, which can be expressed as

$$x'''_m = \sum_{l=1}^{L} b_{l,m}\, \left(x''_m\right)^l, \quad (6)$$

$$= f_m(x_m), \quad (7)$$

where L is the order and $b_{m,l}$ denotes the coefficient of the taylor polynomial. The goal of this paper is to determine the multiple emitters from the received overlapped signal y.

3. PROPOSED SMEI SCHEME VIA MULTI-LABEL LEARNING

The core idea of our SMEI scheme is to simultaneously identify multiple active emitters using a multi-label learning. Figure 1 presents the block diagram of the proposed SMEI scheme. The process starts with the overlapped signal y, followed by the multi-emitter fingerprint extractor $\xi(\cdot)$, which obtains multi-emitter fingerprints. The features are then input to the multi-emitter decision maker $\rho(\cdot)$, resulting in a set of active emitters S.

3.1. Multi-emitter fingerprint extractor

We design a multi-emitter fingerprint extractor $\xi(\cdot)$, which aims to extract multi-emitter fingerprints from received overlapped signals for subsequent multi-label classification tasks. The obtained multi-emitter fingerprints can be represented as

$$\mathcal{F} = \xi(y) : R^{2\times T} \to R^d, \quad (8)$$

where $y \in R^{2\times T}$ indicates the overlapped complex signal with a length of T samples; d defines the dimensionality of the extracted feature space. Figure 2 illustrates the architecture of the fingerprint feature extractor, which employs convolutional layers, residual blocks, and pooling layers to achieve multi-scale feature extraction.

Specifically, the overlapped signal y is converted into its real and imaginary parts after data preprocessing, resulting in an input of shape $\mathcal{R}^{2\times T}$. The input is then expanded to 32 dimensions through convolutional layers and progressively increased to 256 dimensions via multiple residual blocks. This process enhances feature representation through skip connections and mitigates the vanishing gradient problem. The max pooling layers incrementally reduce the size of the feature maps. Global average pooling compresses the feature space from $(B, 256, L/8)$ to a fixed-dimensional feature vector $(B, 256)$, where B is the batch size. Finally, $\mathcal{F}$ effectively captures the multi-emitter fingerprint information from y.

3.2. Multi-emitter decision maker

Multi-emitter decision maker $\rho(\cdot)$ is to map the feature space R^d to emitter state space, thereby effectively identifying up to K active emitters. Formally, we define the label vector $\boldsymbol{\lambda} = \begin{bmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_K\end{bmatrix}^\top$, where $\lambda_m \in \{0, 1\}$ indicates the active state of the m-th emitter.

The mapping relationship of the multi-emitter decision maker $\rho(\cdot)$ can be expressed as $\rho : R^d \to 2^{[1,K]}$, where the input is multi-emitter fingerprints $\mathcal{F} \in R^d$ and the output is the estimated active emitter set $\hat{S}$. Thus, the set of active emitters recognized by the SMEI can be expressed as

$$\hat{S} = \{m \mid \sigma(\rho(\mathcal{F}))_m > \theta, m \in [1, K]\}, \tag{9}$$

where σ denotes the sigmoid activation function. We use the binary cross-entropy loss function, which calculates the error between the predicted activation probabilities and the true labels for each emitter, i.e.

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{K}\sum_{m=1}^{K}\Big[\lambda_m \log(p_m) + (1-\lambda_m)\log(1-p_m)\Big], \tag{10}$$

where

$$p_m = \sigma(\rho(\mathcal{F}))_m, \tag{11}$$

where p_m is the predicted activation probability for the m-th emitter; $\sigma(\cdot)$ is the sigmoid activation function.

Remark 1 *The proposed SMEI scheme overcomes single-label limits by using multi-label learning and independent sigmoid modeling for each emitter, accurately capturing activation in signal-overlap scenarios while maintaining identification accuracy and reducing system complexity.*

3.3. Evaluation metric for SMEI

To evaluate the performance of the proposed SMIE sheme, we adopt two key accuracy metrics, subset accuracy [9] and hamming accuracy [10], which is expressed as

$$P_c^{\text{subset}} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{I}\left(\hat{\boldsymbol{\lambda}}^{(i)} = \boldsymbol{\lambda}^{(i)}\right), \tag{12}$$

$$P_c^{\text{Hamming}} = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{N}\sum_{i=1}^{N}\mathcal{I}\left(\hat{\boldsymbol{\lambda}}_k^{(i)} = \boldsymbol{\lambda}_k^{(i)}\right), \tag{13}$$

where N is the total number of samples; $\hat{\boldsymbol{\lambda}}^{(i)}$ is the predicted label vector for the i-th sample; $\boldsymbol{\lambda}^{(i)}$ denotes the ground truth label vector for the i-th sample; $\hat{\boldsymbol{\lambda}}_k^{(i)}$ is the predicted label for the k-th emitter of the i-th sample; $\boldsymbol{\lambda}_k^{(i)}$ denotes the ground truth label for the k-th emitter of the i-th sample; $\mathcal{I}(\cdot)$ is the indicator function that returns one if the condition is true and zero otherwise.

In addition, we introduce the macro F1 score as an auxiliary metric, which can be expressed

$$F1_{\text{macro}} = \frac{1}{K}\sum_{k=1}^{K}\frac{2\text{TP}_k}{2\text{TP}_k + \text{FP}_k + \text{FN}_k}, \tag{14}$$

where TP_k is the number of true positives; FP_k denotes the number of false positives; FN_k is the number of false negatives for the k-th class.

Table 1. System configuration parameters in multi-emitter identification experiments.

Category	Parameter	Value / Range
Signal parameters	Sampling rate	F_s = 120 MHz
	Symbol rate	R_s = 20 MHz
	Oversampling ratio	*SPS* = $F_s/R_s = 6$
	RRC roll-off	α = 0.3
	RRC span	*span* = 10· (symbols)
	Rician factor	K_R = 10 dB
Training hyperparameters	Batch size	64
	Epochs	100
	Learning rate	3×10^{-4}
	Step size	20
	Learning rate decay factor	0.5
Optimizer	Type	Adam
	Initial learning rate	3×10^{-4}
	Learning rate scheduler	StepLR

Here, we emphasize the characteristics of these three metrics. P_c^{subset} focuses on the accuracy of the overall multi-emitter combination identification, while P_c^{Hamming} provides a fine-grained analysis of the identification precision for individual emitters. However, since P_c^{subset} is extremely stringent for the proposed SMEI scheme, directly comparing P_c^{subset} of the two approaches is not fair. To achieve a more just and comprehensive performance comparison, this paper introduces $F1_{\text{macro}}$ as the auxiliary metric for evaluating the performance differences between the proposed SMEI scheme and prior multi-classification scheme.

4. EXPERIMENTS AND ANALYSIS

4.1. Experimental setup

This section outlines the key experimental configuration used for synthesizing the dataset, including signal modulation, spectrum allocation, device hardware parameters, and channel/noise models. The data synthesis targets the drone communication band from 2.40 to 2.48 GHz, utilizing quadrature phase shift keying modulation. To mitigate the boundary transients introduced by pulse shaping, the model input uniformly selects steady-state time intervals in the middle of each sequence. Table 1 lists the configuration parameters for the SMEI system, encompassing signal parameters, training hyperparameters and optimizer settings. Regarding spectrum allocation and digital intermediate frequency configurations, the experiment includes two levels of spectrum overlap: 50% and 100%. With 50% overlap, neighboring carrier frequencies are separated by 13 MHz, which is half the symbol bandwidth. Furthermore, emitter distortion parameters are detailed in [8] and unless otherwise specified the channel used by default is the rician channel.

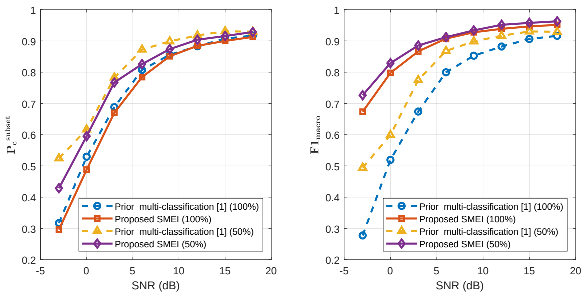

Fig. 3. Comparison of P_c^{subset} and $F1_{\text{macro}}$ between the proposed SMEI and prior multi-classification scheme under different overlap levels when $K = 3$.

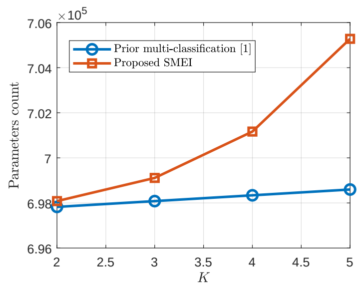

Fig. 4. Comparative analysis of parameters count in the proposed SMEI and prior multi-classification scheme across varying K.

4.2. Experimental results

As shown in Figure 3, We evaluate the performance of the proposed SMEI scheme in comparison with prior multi-classification scheme with $K = 3$. Under 100% overlap conditions, as the signal-to-noise ratio (SNR) increases from -3 dB to 18 dB, in the prior multi-classification scheme, P_c^{subset} increases from 0.3171 to 0.9171, while $F1_{\text{macro}}$ improves from 0.2775 to 0.9163. In our SMEI scheme, P_c^{subset} rises from 0.2962 to 0.9129, approaching the performance of the multi-classification scheme, while $F1_{\text{macro}}$ quickly jumps from 0.6735 to 0.9514, showcasing outstanding macro-level identification performance. Under 50% overlap conditions, the performance is further enhanced. Moreover, SMEI's identification accuracy significantly exceeds random guessing at a low SNR. This indicates strong feature extraction capabilities for effective multi-emitter identification, and as the SNR increases, identification performance improves substantially.

Figure 4 illustrates parameters count of the SMEI and prior multi-classification scheme across varying K values. The SMEI scheme shows a linear increase in parameters, growing from 697,826 at $K = 2$ to 698,597 at $K = 5$, with an increase of just 771 parameters. This indicates high parameter efficiency and low sensitivity to emitter count. In contrast, the multi-classification scheme experiences significant growth, jumping from 698,083 to 705,279 parameters (a 7,196 increase) as K rises from 2 to 5, suggestive of exponential growth. As K scales up, the multi-classification scheme risks catastrophic parameter inflation, leading to exponential increases in computational and storage requirements. Therefore, we can conclude that the advantage in lower parameter count makes SMEI easier to deploy.

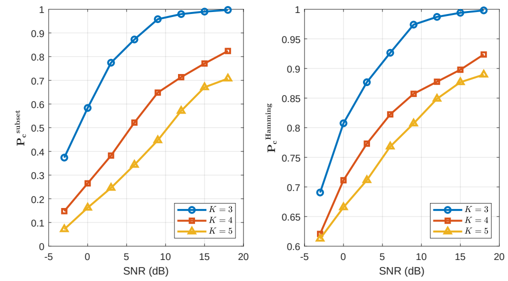

Fig. 5. P_c^{subset} and P_c^{Hamming} of proposed SMEI for different values of K in an AWGN channel with 100% overlap.

Figure 5 presents the P_c^{subset} and P_c^{Hamming} of the proposed SMEI scheme for different values of K in an additive white gaussian noise (AWGN) channel with 100% overlap. Specially, when $K = 3$, P_c^{subset} under the AWGN channel is significantly higher than the corresponding value shown in Figure 3 for the rician channel. This difference can be attributed to the inherent characteristics of the channels, where the AWGN channel typically provides a more stable and favorable environment for signal processing compared to the multipath effects present in the rician channel. Moreover as K increases the impact on P_c^{subset} becomes more significant. When $K = 3$, the identification accuracy nearly reaches complete identification at an SNR of 18 dB. However, for $K = 5$, P_c^{subset} is approximately 0.7074, indicating that as K increases, subset-level identification becomes progressively more challenging. In contrast, the bit-wise accuracy based on hamming indicates that most identification errors occur at the combinatorial level rather than at the individual label level. In practical engineering applications, if the goal is to identify certain key transmitters or determine the presence of a single transmitter, the evaluation metrics can focus on P_c^{Hamming}, allowing for more robust and easily implementable identification performance.

5. CONCLUSION

This paper proposed a SMEI scheme via multi-label learning to effectively identify multiple transmitting devices in overlapping signals. By designing a multi-emitter fingerprint extractor and a classification decision maker, the scheme significantly reduces the complexity while maintaining high accuracy. Experiment results show that the proposed SMEI scheme exhibits excellent robustness and scalability under various SNR and overlap conditions, making it suitable for multi-device concurrent communication scenarios.

6. REFERENCES

[1] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "Oracle: Optimized radio classification through convolutional neural networks," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*. IEEE, 2019, pp. 370–378.

[2] Huazhi Zheng, "Research on intelligent identification methods of target individuals in multi-drone signal overlapping scenarios," 2023.

[3] Mengfan Wei, "Research on individual identification of communication radiation sources in overlapping scenarios," 2024.

[4] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on uav communications for 5g and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, December 2019.

[5] H. Chen and S.-J. Kim, "Robust RF mixture signal recognition using discriminative dictionary learning," *IEEE Access*, vol. 9, pp. 1–1, 2021.

[6] Y. Jin, M. Wei, and Q. Li, "An RF fingerprint extraction method based on time-frequency domain feature fusion," in *2022 International Conference on Algorithms, Network and Computer Technology (ICANCT 2022)*. IEEE, 2023.

[7] J. Li, L. Yu, and Y. Wei, "A compound jamming signals recognition method based on one-dimensional multi-label convolutional neural network," in *2023 IEEE International Radar Conference (RADAR)*. IEEE, 2023.

[8] Boxiang He and Fanggang Wang, "Cooperative specific emitter identification via multiple distorted receivers," *IEEE Transactions on Information Forensics and Security*, pp. 3791–3806, 2020.

[9] W. de Silva and A. Fernando, "Explainable video topics for content taxonomy: A multimodal retrieval approach to industry-compliant contextual advertising," *IEEE Access*, vol. 13, pp. 30597–30612, 2025.

[10] A. Poonia, M. Kishor, and K. P. R. Ayyagari, "Designing of high entropy alloys with high hardness: A metaheuristic approach," *Scientific Reports*, vol. 14, no. 1, pp. 7692, 2024.