

基于有向图连通性的无监督辐射源个体识别方法

杨宁, 王桁, 张邦宁, 丁国如, 郭道省

(陆军工程大学通信工程学院, 南京 210007)

摘要: 辐射源个体识别 (Specific emitter identification, SEI) 是指利用接收电磁信号中的独特细微特征来区分发射设备的技术。深度学习由于其强大的特征提取能力, 逐渐成为实现辐射源个体识别的主要手段, 但在非合作场景中无法获取大量带标签的数据样本来训练神经网络, 且待识别的辐射源个数未知。为此, 本文提出了无需指定辐射源个数的基于有向图连通性的无监督辐射源个体识别方法。受层次聚类的启发, 首先基于局部密度将射频指纹特征空间划分为多个子簇, 并将特征向量之间的关系映射为有向图; 然后, 基于有向图的连通性, 将多个子簇进行合并, 得到最终的识别结果。实验结果表明, 在低信噪比条件下, 所提方法能准确进行辐射源个体识别, 识别性能较基准算法提高了 7.1%~53.1%。

关键词: 辐射源个体识别; 无监督; 射频指纹; 有向图; 层次聚类

中图分类号: TN911.7

文献标志码: A

Unsupervised Specific Emitter Identification Method Based on Directed Graph Connectivity

YANG Ning, WANG Heng, ZHANG Bangning, DING Guoru, GUO Daoxing

(College of Communications and Engineering, Army Engineering University of PLA, Nanjing 210007, China)

Abstract: Specific emitter identification (SEI) refers to the technique of distinguishing emitters by utilizing unique and subtle features in received electromagnetic signals. Due to its powerful feature extraction ability, deep learning has gradually become the main means of implementing SEI. However, in non-cooperative scenarios, labeled samples generally cannot be obtained to train the neural network, and the number of emitters to be identified is unknown. Therefore, this paper proposes an unsupervised SEI method based on directed graph connectivity without specifying the number of emitters. Drawing inspiration from the idea of hierarchical clustering, the radio frequency fingerprinting feature space is first divided into multiple sub-clusters based on local density, and the relationship between feature vectors is mapped to a directed graph. Then, based on the connectivity of the directed graph, the multiple subclusters are automatically merged to obtain the final identification result. Experimental results show that under low signal-to-noise ratio conditions, the proposed method can accurately identify individual emitters, and its identification performance is improved by 7.1%—53.1% compared to the benchmark algorithms.

Key words: specific emitter identification (SEI); unsupervised; radio frequency fingerprinting; directed graph; hierarchical clustering

引 言

随着现代通信技术的快速发展,电磁频谱成为支撑国家关键基础设施和军事信息化的基石。在军事领域,电磁频谱作为继陆、海、空、天、网之后的第6维作战空间,已成为大国战略博弈的制高点。在民用领域,物联网设备数量爆炸式增长,频谱供需矛盾加剧^[1]。但是非法设备网络入侵和非法用频行为时有发生,给电磁频谱管理带来了极大的挑战。准确识别和监控电磁频谱中各类无线设备变得至关重要。尤其是在军事作战和民用电磁频谱安全管控方面,确保电磁态势的清晰认知和快速响应已成为当务之急。在此背景下,辐射源个体识别技术因能破解电磁空间“目标身份难辨、行为意图难判”的难题,成为军民领域的共性需求。

辐射源个体识别是指利用接收的电磁信号包含的特征对辐射源进行个体身份识别。所有的电子器件都会有不同程度的缺陷。这些缺陷会对信号形成无意识调制,也被称为射频指纹^[2]。即使相同型号、同一批次的电子器件的射频指纹也互不相同。因此,每个设备的射频指纹都相当于它们唯一的“ID”,根据这个“ID”便可以实现辐射源的个体识别。辐射源个体识别技术现已被广泛用于频谱检测、安全认证和电子对抗等军民领域^[3]。

传统方法通常提取电磁信号的专家特征作为射频指纹来实现个体识别^[4]。根据信号段位置的不同,射频指纹分为基于暂态信号的射频指纹和基于稳态信号的射频指纹。信号的暂态部分是辐射源切换工作状态时产生的,与所发射的信号内容、调制方式等无关,仅与辐射源硬件结构和制造工艺相关,因此可从中提取出射频指纹。暂态信号的平均功率和幅度的比值,奇异频谱分析、双树复小波和多重分形等作为指纹特征均实现了有效的辐射源个体识别^[5]。然而,由于信号的暂态部分持续时间短,难以捕捉定位,后续研究者主要利用辐射源持续工作时发射的稳态信号中包含的指纹特征实现个体识别。Tang等^[6]设计了一种新的射频指纹特征表示,称为集中对数功率谱,以减少信道效应的影响,该特征具有良好的泛化性和稳定性。Huang等^[7]提出了一种新的指纹利用模式,即密度迹线图(Density trace plot, PlotDTP),以利用源自符号转换轨迹的射频指纹进行辐射源个体识别。文献[8]利用短时傅里叶变换(Short time Fourier transform, STFT)的时频信息实现了个体识别,并评估了STFT参数选择对所提方法识别性能的影响,在真实条件下捕获的数据集的认识准确率为98%,证明了所提方法的鲁棒性。但专家特征的提取需要极强的专业知识,且不同场景中能有效区分辐射源个体的专家特征各不相同。

深度学习出现后,可以提取更加复杂内部的表示,从而实现更细粒度的识别,有效提高了辐射源个体识别的准确性。Robyns等^[9]最早将卷积神经网络和多层感知器用于辐射源个体识别。此后,残差网络^[10-11]、长短时记忆网络^[12]、递归神经网络和Transformer^[13]等网络均被用于辐射源个体识别并取得了良好的效果。Wang等^[14]提出了利用复值神经网络识别以更好地适配接收信号的复数形式。

但在合作场景中,获取带标签样本非常困难。文献[15]提出了无监督辐射源个体识别方法,将信号双谱的灰度直方图作为射频指纹特征,并利用无需带标签样本训练的信息最大化生成对抗网络来实现无监督识别。Hao等^[16]通过引入对比自监督聚类来应对在不可信场景或非合作环境中辐射源信号的无标签聚类。然而上述方法均需要类别数即辐射源个体数作为超参数,在非合作场景中,提取辐射源个数几乎是不可能的,因此上述方法并不完全适用非合作场景。文献[17]提出了在聚类中比距离测量更有效的数据相关性度量标准。Cheng等^[18]利用自动编码器和密度峰值聚类(Density peak clustering, DPC)实现无需指定个数的无监督聚类,通过提取信号的t分布-随机邻近嵌入(t-distributed stochastic neighbor embedding, t-SNE)E降维特征、均值、方差、峰度和偏斜度等特征并进行特征融合来提高无监督识别的准确率。密度峰值聚类可以很好地对单峰子簇(即只有一个密度峰的簇)的特征进行聚类^[19-20],但射频指纹特征分布不均匀,同一类的特征空间可能包含多个密度峰,密度峰值聚类寻找密度峰的启发式方法在处理多峰聚类时可能会错误地选择一些密度峰作为聚类中心,导致聚类结果

不佳^[21]。

分层聚类更适合于复杂形状的射频指纹特征空间。但传统层次聚类方法主要依赖于距离度量来进行样本分类,难以处理高维特征空间中数据分布不均匀的问题。本文将层次聚类与图论相结合,将射频指纹特征向量之间的关系映射到有向图中,采用图中的边权和节点连通性来判断和合并子簇。图论方法的引入,使得特征向量之间的复杂关系得以更精确的表示,尤其是在处理特征向量分布复杂、不均匀的情况下,比传统的基于距离的层次聚类方法更具优势。

为了解决非合作场景下辐射源数量未知、样本无标签的问题,本文提出了一种基于有向图连通性(Directed graph connectivity, DGC)的无监督辐射源个体识别方法。本文的主要贡献总结如下:

(1) 提出了一种基于有向图连通性的无监督辐射源个体识别方法,从图的角度确定合并子簇的依据,从而实现射频指纹的自动聚类,而无需指定辐射源的数量,为非合作场景中的无监督辐射源个体识别问题提供了解决方案。

(2) 为了提高无监督辐射源个体识别的准确率,将射频指纹特征向量之间的关系映射成有向图,提出了基于相似矩阵的连通性度量方法,通过最大代表性分布一致性综合评估子簇间的连通性,进一步揭示了射频指纹特征空间的分布特点。

(3) 在CC2530数据集和公开的LoRa数据集上验证了所提算法的性能。DGC的识别准确率为97.6%,与基准算法相比,识别性能提高了7.1%~53.1%。

1 问题定义

假设前期已通过信号分离技术将混合信号分离,信号样本仅包含单个辐射源的时域信号,用 $s_k(t)$ 表示第 k 个辐射源的发射信号,连续时间内接收信号可表示为

$$r_k(t) = s_k(t) * h_k(t) + n_k(t) \quad k = 1, 2, \dots, K \quad (1)$$

式中: $h_k(t)$ 为第 k 个辐射源和接收机之间的信道冲激响应, $n_k(t)$ 为加性高斯白噪声,“*”表示卷积运算, K 表示辐射源总数。对 $r_k(t)$ 进行采样后得到离散时间样本 $\mathbf{X} = [x_1, x_2, \dots, x_N]$,其中 N 表示样本总数。

无监督辐射源个体识别的目标就是找到相似性函数 s 和聚类函数 $f(\mathbf{X}, s)$,以获得信号样本到逻辑矩阵 \mathbf{M} 的映射,其中逻辑矩阵的元素 $m_{ij} = 1$ 表示信号样本 x_i 和 x_j 来自同一个辐射源。相似性函数 s 用于度量信号样本之间的相似度,通过减小类内距离、增大类间距离,将初始分布杂乱的时域信号映射到一个新的特征空间中。在这个特征空间中,信号的特征向量具有更好的可区分性,能够有效区分来自不同辐射源的样本。相似性函数 s 的作用可以看作是特征提取,因为它通过转换信号样本来提高样本的可区分性。除了传统的距离度量方法,神经网络也可以被视作一种相似性函数。通过一系列非线性变换,神经网络能够提取时域信号中的细微特征,识别来源于同一辐射源的样本之间的相似性,并增大来自不同辐射源的特征向量之间的距离。因此,聚类函数 $f(\mathbf{X}, s)$ 将经过相似性函数 s 转换后的特征向量划分为不同的簇,每个簇中的特征向量代表一组来自同一辐射源的信号样本。因此,聚类函数在辐射源个体识别中负责根据特征空间中的相似性将信号样本进行分类。

2 基于有向图连通性的无监督辐射源个体识别方法

本文提出的无监督辐射源个体识别框架如图1所示,可分为3个步骤:信号预处理、特征提取和聚类。信号预处理包含小波降噪、数据增强和归一化。小波降噪的目的是提高接收信号的信噪比,以降低噪声对识别性能的影响。本文采取的数据增强方法包括相位旋转、循环移位和随机加噪^[22-23],目的是生成同一信号样本的增强视图,以便后续利用对比学习进行特征提取。特征提取指的是利用对比学习将接收到的时域信号映射为更具区分性的射频指纹特征。对比学习通过构建正负样本对,让相似的样

本在特征空间中更加接近,而不相似的样本更加远离。最后一步为聚类,首先根据局部密度将射频指纹特征划分为多个子簇,然后根据子簇之间的连通性进行合并,得到最终的识别结果。

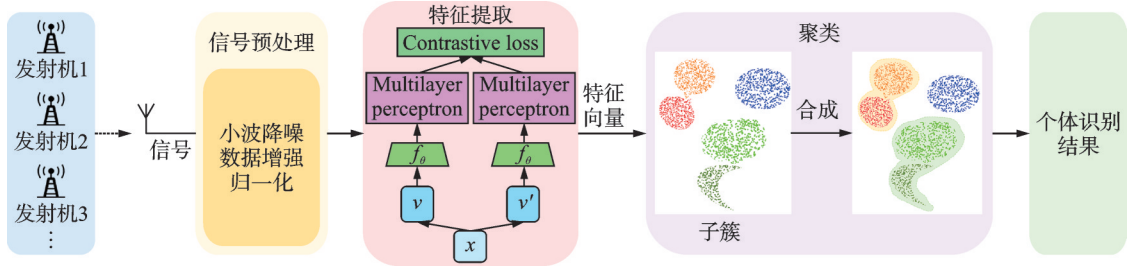


图1 基于有向图连通性的无监督辐射源个体识别流程图

Fig.1 Framework of unsupervised specific emitter identification based on directed graph connectivity

2.1 信号预处理

2.1.1 小波降噪

噪声通常表现为高频信号,而发射的信号通常为低频信号或者一些较为平稳的信号,因此通过小波分解将接收信号分解为多个子频带,并对高频带信号进行滤除,而后重构信号,重构后的信号信噪比提高,从而降低了噪声对识别的影响。

图2和图3为小波降噪前后的CC2530的时域信号对比图。CC2530是常用的基于 Zigbee 的物联网节点,采用 BB60C 在室内采集其发射的无线信号,并进行小波降噪。如图2和图3所示,可以看到降噪后的信号毛刺减少,更有利于神经网络从信号中提取与辐射源相关的特征,避免信道特征的影响。通过实验测试,本文选取的小波基函数为“db4”,阈值估计方法采用固定式阈值,分解层数为4,小波降噪可将0 dB 信号信噪比提高到8 dB。

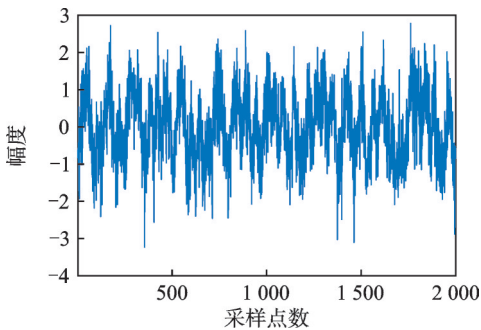


图2 小波降噪前时域信号

Fig.2 Time-domain signal before wavelet denoising

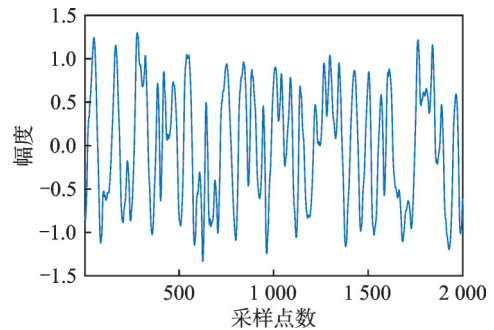


图3 小波降噪后时域信号

Fig.3 Time-domain signal after wavelet denoising

2.1.2 数据增强

为了便于后续利用对比学习进行特征提取,本文采用相位旋转、循环移位和随机加噪对信号进行数据增强以生成样本相关视图。

(1) 相位旋转:如图4所示,使用正弦和余弦函数将接收信号 I/Q 两路分别旋转一定角度 α 。旋转后的信号可表示为

$$\begin{bmatrix} I' \\ Q' \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} I \\ Q \end{bmatrix} \quad (2)$$

式中 $\alpha \in \{0, \pi/2, \pi, 3\pi/2\}$, 表示旋转角度。

(2) 循环移位: 将 I/Q 两路信号向后进行 n 个样本点的循环移位, 即

$$\begin{bmatrix} I' \\ Q' \end{bmatrix} = \begin{bmatrix} i_{N_I-n+1}, i_{N_I-n}, \dots, i_{N_I}, i_1, i_2, \dots, i_{N_I-n} \\ q_{N_I-n+1}, q_{N_I-n}, \dots, q_{N_I}, q_1, q_2, \dots, q_{N_I-n} \end{bmatrix} \quad (3)$$

式中 N_I 表示样本长度。

(3) 随机加噪: 对信号加上随机高斯白噪声, 使其在幅度上产生细微差别。

2.2 基于对比学习的特征提取

能否提取具有独特性和可区分性的射频指纹特征直接影响后续聚类的性能。对比学习作为一种无监督学习方法, 不依赖于数据标签, 通过比较不同样本之间的相似性和差异性来学习有用的特征表示。它利用构建正样本对(相似或相关的样本对)和负样本对(不相似或不相关的样本对), 来训练模型以拉近正样本对之间的距离, 同时推开负样本对之间的距离。本文采用的模型结构如图5所示。

(1) 数据增强: 对于样本 x_i , 随机选取两种数据增强方式以获得正样本对 v_i 和 v'_i 。对于大小为 B 的批次, 共得到 $2B$ 个样本对, 其中 $\{v_i, v'_i\}$ 为正样本对, 其余 $2B-2$ 为负样本对。

(2) 利用特征提取器 f_θ 对 v_i 和 v'_i 进行特征提取, 得到每个样本的特征表示 $rf_i = f_\theta(v_i)$ 和 $rf'_i = f_\theta(v'_i)$ 。其中, 特征提取器对于不同的样本共享权重, 本文采用 1D-Resnet18 的结构作为特征提取器。

(3) 利用多层感知机 (Multilayer perceptron, MLP) 将特征向量投影到 128 维特征空间中, 得到投影向量 z_i 和 z'_i , 以减轻由对比损失引起的信息损失^[16]。MLP 的结构为: 线性层、修正线性单元 (Rectified linear unit, ReLU) 层、线性层。

(4) 利用对比损失函数减小正样本对的投影向量之间的距离, 增大负样本对的投影向量之间的距离。对比损失函数的表达式为

$$L_i = -\lg \frac{e^{z_i^T z'_i / \tau}}{\sum_{j=0}^{2B} e^{z_i^T z_j / \tau}} \quad (4)$$

式中: “T” 表示转置运算, τ 是控制灵敏度的温度超参数。

2.3 基于有向图连通性的无监督辐射源个体识别

2.3.1 基于局部密度的单峰子簇识别

根据对比学习得到的特征向量集合 $R_F = \{r_{f_1}, r_{f_2}, \dots, r_{f_n} | r_{f_i} \in \mathbf{R}^d\}$, 对于每一个特征向量 r_{f_i} , 选取距离其最近的 k 个特征向量为近邻点, 表示为 $N_k(r_{f_i})$, 那么它的局部密度 ρ_i 可表示为^[17]

$$\rho_i = \frac{1}{\frac{1}{k} \sum_{r_{f_j} \in N_k(r_{f_i})} d_{ij}} \quad (5)$$

式中 d_{ij} 表示特征向量 r_{f_i} 和 r_{f_j} 之间的距离。如果 $\rho_i > \max_{r_{f_j} \in N_k(r_{f_i})} (\rho_j)$, 那么特征向量 r_{f_i} 可视为一个子簇的中心点。根据式(5)计算得到所有的子簇中心点, 从而将特征向量集合划分为 m 个子簇。

式(5)的关键在于近邻点个数 k 的选择, 根据文献[24], k 值的计算公式为

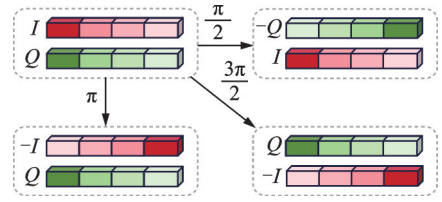


图4 相位旋转

Fig.4 Phase rotation

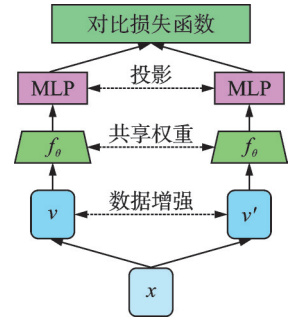


图5 对比学习框架图

Fig.5 Framework of contrastive learning

$$\begin{aligned} & \text{maximize } \text{range}(T'_k) \\ & \text{s.t. } \forall k_1, k_2 \in T'_k \subseteq T_k = \left[1, \lceil \sqrt{n} \rceil\right] \quad C^{k_1} = C^{k_2} \end{aligned} \quad (6)$$

式中: T_k 表示 k 的取值区间, T'_k 表示 k 的近似取值区间, C^{k_1} 表示当 $k = k_1$ 时的子簇数量, 符号 $\lceil \cdot \rceil$ 表示向上取整运算, 将 T'_k 的均值作为 k 的最终取值。

2.3.2 边缘特征向量代表性

结合子簇的识别结果, 特征向量之间的关系可映射为有向图的邻接矩阵, 其中节点表示特征向量, 边的取值定义为

$$A[i][j] = \begin{cases} 1 & r_{f_i} \text{ 指向其邻近的更高局部密度的特征向量 } r_{f_j} \\ 0 & \text{其他} \end{cases} \quad (7)$$

图6展示了子簇映射为有向图的示意图。在有向图中, 指向该节点的边数称为入度; 反之, 则称为出度。出度为0的点就是子簇的中心点, 入度为0的点对应子簇中的边缘点。

令边缘点 r_{f_i} 周围 k_l 个近邻点的集合为 N_{k_l} , 其中 $k_l \ll k$, 如果边缘点 $r_{f_i} \in C_1$, 且边缘点 $r_{f_j} \in C_2$, 其中 C_1 和 C_2 表示两个不同的子簇, 且 $r_{f_i} \in N_{k_l}, r_{f_j} \in N_{k_l}$, 则认为子簇 C_1 和 C_2 可以合并。为了提高合并的准确性, 需要计算边缘点在子簇中的代表性。将有向图中点 r_{f_i} 指向点 r_{f_j} 的边的权重定义为

$$\theta(r_{f_i}, r_{f_j}) = \frac{\rho_i}{\rho_j} \quad A[i][j] = 1 \quad (8)$$

边缘点的代表性定义为

$$\eta_i = \theta(r_{f_i}, c) = \prod_{r_{f_p}, r_{f_q} \in \Delta r_{f_i}} \theta(r_{f_p}, r_{f_q}) \quad A[p][q] = 1 \quad (9)$$

式中 Δr_{f_i} 表示点 r_{f_i} 到中心点 c 的路径上所有的点。以图6中的点5为例, $\eta_5 = \theta(r_{f_5}, 11) \times \theta(r_{f_5}, 7) \times \theta(r_{f_5}, 1)$ 。边缘点离中心点越远, η 越小, 代表性越低。一对跨子簇相邻边缘点的代表性 $\alpha_i = \{x_i, \beta_i\}$ 用它们的均值表示为

$$\chi_{\alpha_i} = \frac{\eta_{x_i} + \eta_{\beta_i}}{2} \quad (10)$$

2.3.3 无监督辐射源个体识别

如果两个子簇具有高度连通性, 那么它们会有多个边缘点相邻。从中选取 n_s 个代表性最高的相邻边缘点对, 组成相似矩阵 $Z_{C_1 C_2} = [z_1, z_2, \dots, z_{n_s}]$ 。两个子簇连通性的计算公式为

$$g_{C_1 C_2} = \max(Z_{C_1 C_2}) \times \Gamma(Z_{C_1 C_2}) \quad (11)$$

$$\Gamma(Z) = 1 - \frac{\frac{1}{n_s} \sum_{i=1}^{n_s} |Z(i) - \max(Z)|}{\max(Z)} \quad (12)$$

式中 n_s 的选择表示为

$$n_s = \lceil \beta \times \min(n_e(C_1), n_e(C_2)) \rceil \quad (13)$$

式中: $\beta \in [0, 1]$ 是一个比值参数, $n_e(C_1)$ 表示子簇 C_1 中边缘点的总数。

不同子簇间的连通性构成连通矩阵 G 。根据矩阵 G , 将连通性强的子簇合并为一类, 表示其来自同一辐射源。

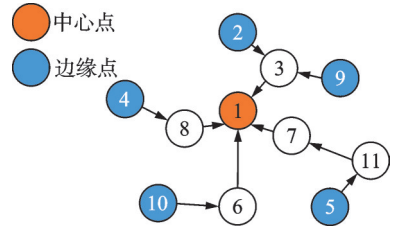


图6 子簇映射为有向图示意图
Fig.6 Schematic diagram of sub-cluster mapping to directed graph

3 实验结果

3.1 数据集

(1) CC2530数据集:使用BB60C采集了8台CC2530发射的信号。CC2530是常用的使用Zigbee的物联网节点,其物理层协议为IEEE 802.15.4。所有辐射源处于相同的采集环境中,信号中心频率为2.405 GHz。BB60C的采样率设置为40 MHz,为了保留更多的细微特征,采用了过采样方式。数据集共包含8个辐射源,每个辐射源有800个信号样本,每个样本包含1 024个采样点的I/Q两路时域信号数据。

(2) LoRa数据集^[25]:该数据集包含利用USRP N210软件无线电平台采集的60个LoRa设备的时域信号数据,其载波频率为868.1 MHz,传输间隔设置为0.3 s。接收机采样率设置为1 MHz。选取其中8个相同型号的设备作为待识别辐射源,来验证本文所提算法的相关性能。

3.2 评价指标

(1) 调整兰德系数(Adjusted Rand index, ARI):兰德指数用来衡量两个分布的吻合程度,在本文中用ARI衡量识别结果和真实标签之间分布的吻合程度。它的取值范围是 $[-1, 1]$,数值越接近于1越好,其计算公式定义为

$$ARI = \frac{RI - E[RI]}{\max(RI) - E[RI]} \quad (14)$$

$$RI = \frac{a + b}{\binom{n}{2}} \quad (15)$$

式中:RI表示兰德系数, n 表示样本总数, $\binom{n}{2} = C_n^2 = \frac{n(n-1)}{2}$ 。令 U 表示实际标签信息, V 表示无监督识别结果, a 表示在 U 与 V 中都是同类别的元素对数, b 表示在 U 与 V 中都是不同类别的元素对数, $E[\cdot]$ 表示求算数平均函数。

(2) 调整互信息(Adjusted mutual information, AMI):互信息可衡量两个分布的吻合程度,取值范围 $[-1, 1]$,值越大,识别结果与真实情况越吻合。

令 U 表示实际类别信息,共有 u_c 个不同的取值,即共有来自 u_c 个辐射源的样本,那么其信息熵可以被计算为

$$I(U) = \sum_{i=1}^{u_c} p_i \log_2 p_i \quad (16)$$

式中 p_i 为归属于第 i 类的数据个数占数据总量的比例,其计算公式为

$$p_i = \frac{|U_i|}{n} \quad (17)$$

令 $V \in \mathbf{R}^n$ 表示无监督识别结果,共有 v_c 个不同的取值,互信息的计算公式为

$$MI(U, V) = \sum_{i=1}^{u_c} \sum_{j=1}^{v_c} p_{ij} \log_2 \frac{p_{ij}}{p_i \times p_j} \quad (18)$$

$$p_{ij} = \frac{|U_i \cap V_j|}{n} \quad (19)$$

U 和 V 的调整互信息可表示为

$$AMI(U, V) = \frac{MI(U, V) - E\{MI(U, V)\}}{\max(I(U), I(V)) - E\{MI(U, V)\}} \quad (20)$$

(3) Fowlkes-Mallows 系数 (Fowlkes-Mallows index, FMI): FMI 是对识别结果和真实值之间召回率和精确率进行几何平均的结果,取值范围为 $[0, 1]$,越接近 1 越好。

$$FMI = \frac{X_{TP}}{\sqrt{(X_{TP} + X_{FP})(X_{TP} + X_{FN})}} \quad (21)$$

式中: X_{TP} 表示在实际类别分类 U 中是同一类,同时在识别结果 V 中也是同一类的一对特征向量; X_{FP} 表示在 V 中属于同一类,但在 U 中不是同一类的一对特征向量; X_{FN} 表示在 U 中属于同一类,但在 V 中不是同一类的一对特征向量。

3.3 性能对比

文献[18]提出的基于 DPC 的无监督辐射源个体识别算法和经典的多模态子空间聚类 (Multimodal subspace clustering, MSC)^[26]、具有噪声的基于密度的聚类方法 (Density-based spatial clustering of applications with noise, DBSCAN)^[27] 聚类算法均无需指定类别数,可自动实现无监督识别,将它们与本文提出的基于 DGC 的无监督辐射源个体识别算法相比,识别结果如图 7 和表 1 所示。随着待识别辐射源数量的增加,特征向量在特征空间中重叠的概率增加,导致识别准确率下降,但 DGC 算法的性能总是优于其他算法。

如表 1 所示,只有本文提出的 DGC 算法和 MSC 才能识别出正确的辐射源数量。虽然 DPC 算法可以很好地识别单峰数据,但它主要关注全局信息,可能会对多峰数据产生误识别。如图 8 和 9 所示,同一个辐射源的特征向量在特征空间中存在多个密度峰值点,使得 DPC 算法可能会形成错误的聚类。本文只使用 DPC 算法来生成子簇,并且只关注局部信息。全局子簇合并是基于有向图连通性测量的结果。

为了检验算法的鲁棒性,验证了算法在不同信噪比

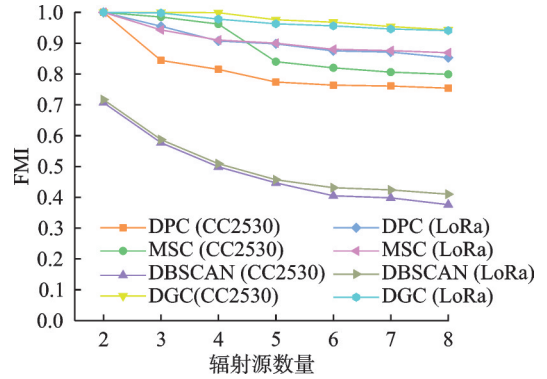


图 7 辐射源数量变化时不同算法识别性对比
Fig.7 Identification performance comparison of different algorithms under varying numbers of emitters

表 1 不同方法识别辐射源数量结果

Table 1 Results of number of emitters identified by different methods

方法	真实辐射源数量	识别结果
DPC	8	5
MSC	8	8
DBSCAN	8	1
DGC	8	8

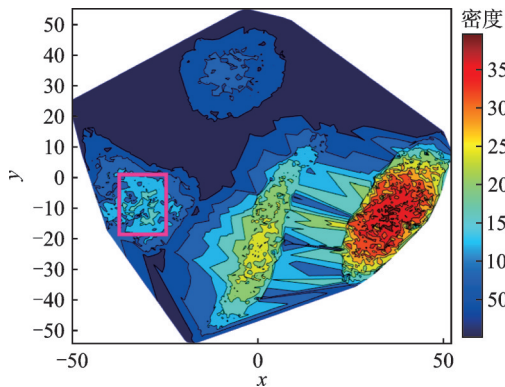


图 8 特征向量分布密度
Fig.8 Density distribution of feature vectors

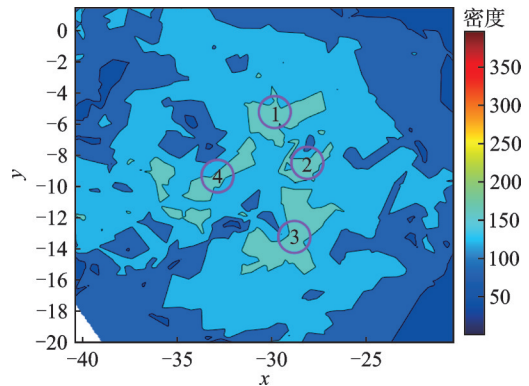


图 9 单一辐射源特征向量分布密度局部放大图
Fig.9 Zoomed-in view of feature vector density distribution for a single emitter

下的识别性能,并与MSC算法进行了比较,识别结果如表2所示。可以看出,即使在低信噪比的情况下,DGC仍然可以正确识别类别数量,实现辐射源个体识别,但MSC算法在信噪比低于0 dB时无法进行无监督辐射源个体识别。MSC是一种基于密度的聚类方法,它需要依赖带宽参数,即邻域的大小来控制聚类的粒度。带宽的选择非常敏感,错误的带宽值可能导致结果不准确,甚至无法有效地进行聚类。MSC特别适用于密度较为均匀的簇,它依赖于局部密度的变化来确定聚类结构。因此,对于存在明显密度差异的数据,均值漂移可能会出现异常,比如低密度区域的样本可能被错误地聚集到一处,而高密度区域可能被分割成多个簇。如图8所示,特征向量的密度分布并不均匀。低信噪比条件下,密度差异加大,这也是MSC无法正确识别的原因之一。而本文所提方法首先根据样本间的距离划分子簇,再根据子簇间的连通性进行合并,适应不同的簇形状。

表2 不同信噪比时无监督辐射源个体识别算法性能

Table 2 Performance of unsupervised specific emitter identification algorithms under different SNR conditions

信噪比/ dB	-10		-5		0		5		10		15	
	MSC	DGC	MSC	DGC	MSC	DGC	MSC	DGC	MSC	DGC	MSC	DGC
ARI	—	0.503	—	0.599	—	0.945	0.953	0.962	0.953	0.967	0.965	0.971
AMI	—	0.425	—	0.512	—	0.958	0.962	0.969	0.961	0.972	0.971	0.975
FMI	—	0.586	—	0.684	—	0.964	0.967	0.972	0.966	0.974	0.973	0.976

为了进一步验证模型鲁棒性,仿真了多径信道环境,以验证多径信道环境中模型的识别性能,信道参数设置如表3所示。识别结果如表4所示,相较于高斯噪声环境,多径信道环境中辐射源个体识别性能有一定的下滑,这是因为复杂的信道环境可能会淹没细微的射频指纹特征,增大了特征提取难度。

表3 多径信道参数

Table 3 Multipath channel parameters

路径	1	2	3
传播时延/ μs	0	0.1	0.2
平均信道增益	0	-4	-8

表4 多径信道环境中不同信噪比时无监督辐射源个体识别算法性能

Table 4 Performance of unsupervised specific emitter identification algorithms under multipath channels with varying SNR conditions

信噪比/ dB	-10		-5		0		5		10		15	
	MSC	DGC	MSC	DGC	MSC	DGC	MSC	DGC	MSC	DGC	MSC	DGC
ARI	—	0.485	—	0.559	—	0.884	0.884	0.905	0.886	0.913	0.894	0.926
AMI	—	0.386	—	0.459	—	0.916	0.894	0.907	0.894	0.921	0.906	0.934
FMI	—	0.559	—	0.608	—	0.916	0.899	0.926	0.896	0.934	0.906	0.937

4 结束语

本文提出了一种基于DGC的无监督辐射源个体识别算法。在不指定辐射源数量的情况下,DGC可以快速准确地实现非线性射频指纹特征聚类,从而实现无监督辐射源个体识别。本文创造性地将射频指纹特征聚类与有向图联系起来,利用DPC计算生成子聚类,保留足够的局部信息,然后通过代表性的边点评价计算有向图之间的连通性。根据连通性测量结果,可以合并任意形状的子簇,从而实现辐射源的有效识别。CC2530和公共LoRa数据集上的实验结果验证了该方法的性能。本文的特征提取器结构借鉴了1D-ResNet18,并未深入研究神经网络结构的影响,在未来的工作中,将针对神经网络结构优化进行深入研究,以进一步提高无监督辐射源个体识别性能。

参考文献:

- [1] WANG C, FU X, WANG Y, et al. Interpolative metric learning for few-shot specific emitter identification[J]. *IEEE Transactions on Vehicular Technology*, 2023, 72(12): 16854-16855.
- [2] 肖易寒, 王博煜, 于祥祯, 等. 基于双路射频指纹卷积神经网络与特征融合的雷达辐射源个体识别[J]. *电子与信息学报*, 2024, 46(8): 3238-3245.
XIAO Yihan, WANG Boyu, YU Xiangzhen, et al. Radar emitter identification based on dual radio frequency fingerprint convolutional neural network and feature fusion[J]. *Journal of Electronics & Information Technology*, 2024, 46(8): 3238-3245.
- [3] 刘志文, 陈旗, 郑恒权, 等. 基于特征融合的通信辐射源个体识别[J]. *数据采集与处理*, 2022, 37(6): 1280-1287.
LIU Zhiwen, CHEN Qi, ZHENG Hengquan, et al. Specific identification of communication emitter based on feature fusion[J]. *Journal of Data Acquisition and Processing*, 2022, 37(6): 1280-1287.
- [4] 刘高辉, 席宏恩. 改进变分模态分解与多特征的通信辐射源个体识别方法[J]. *电子与信息学报*, 2024, 46(10): 4044-4052.
LIU Gaohui, XI Hongen. Individual identification method for communication emitters based on improved variational modal decomposition and multiple features[J]. *Journal of Electronics & Information Technology*, 2024, 46(10): 4044-4052.
- [5] HAN X, CHEN S, CHEN M, et al. Radar specific emitter identification based on open-selective kernel residual network[J]. *Digital Signal Processing*, 2023, 134: 103913.
- [6] TANG P, DING G, XU Y, et al. Causal learning for robust specific emitter identification over unknown channel statistics[J]. *IEEE Transactions on Information Forensics and Security*, 2024, 19: 5316-5329.
- [7] HUANG D, AL-HOURANI A, SITHAMPARANATHAN K, et al. Deep learning methods for IoT device authentication using symbols density trace plot[J]. *IEEE Internet of Things Journal*, 2024, 11(10): 18167-18179.
- [8] LI B, CETIN E. A graphical deep learning approach to RF fingerprinting in the time-frequency domain[J]. *IEEE Sensors Journal*, 2024, 24(10): 16984-16990.
- [9] ROBYNS P, MARIN E, LAMOTTE W, et al. Physical-layer fingerprinting of LoRa devices using supervised and zero-shot learning[C]//*Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. New York, USA: ACM, 2017: 58-63.
- [10] SHEN G, ZHANG J, MARSHALL A, et al. Towards receiver-agnostic and collaborative radio frequency fingerprint identification[J]. *IEEE Transactions on Mobile Computing*, 2024, 23(7): 7618-7634.
- [11] SHEN G, ZHANG J, MARSHALL A, et al. Towards scalable and channel-robust radio frequency fingerprint identification for LoRa[J]. *IEEE Transactions on Information Forensics and Security*, 2022, 17: 774-787.
- [12] SHEN G, ZHANG J, MARSHALL A, et al. Toward length-versatile and noise-robust radio frequency fingerprint identification[J]. *IEEE Transactions on Information Forensics and Security*, 2023, 18: 2355-2367.
- [13] SHEN G, ZHANG J, MARSHALL A, et al. Radio frequency fingerprint identification for LoRa using deep learning[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(8): 2604-2616.
- [14] WANG Y, GUI G, GACANIN H, et al. An efficient specific emitter identification method based on complex-valued neural networks and network compression[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(8): 2305-2317.
- [15] GONG J, XU X, LEI Y. Unsupervised specific emitter identification method using radio-frequency fingerprint embedded InfoGAN[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 2898-2913.
- [16] HAO X, FENG Z, LIU R, et al. Contrastive self-supervised clustering for specific emitter identification[J]. *IEEE Internet of Things Journal*, 2023, 10(23): 20803-20818.
- [17] TING K M, ZHU Y, CARMAN M, et al. Overcoming key weaknesses of distance-based neighbourhood methods using a data dependent dissimilarity measure[C]//*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, USA: ACM, 2016: 1205-1214.
- [18] CHENG D, HUANG J, ZHANG S, et al. A novel approximate spectral clustering algorithm with dense cores and density peaks[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, 52(4): 2348-2360.
- [19] QIU T, LI Y J. Fast LDP-MST: An efficient density-peak-based clustering method for large-size datasets[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(5): 4767-4780.

- [20] TOBIN J, ZHANG M. A theoretical analysis of density peaks clustering and the component-wise peak-finding algorithm[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(2): 1109-1120.
- [21] GUAN J, LI S, HE X, et al. Fast hierarchical clustering of local density peaks via an association degree transfer method[J]. *Neurocomputing*, 2021, 455: 401-418.
- [22] GUAN J, LI S, HE X, et al. SMMP: A stable-membership-based auto-tuning multi-peak clustering algorithm[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(5): 6307-6319.
- [23] SWAMI A, SADLER B M. Hierarchical digital modulation classification using cumulants[J]. *IEEE Transactions on Communications*, 2000, 48(3): 416-429.
- [24] ZHANG X, LIN M, TIAN Y, et al. Data enhancement aided protocol-agnostic transmitter recognition for open-set in IoT[J]. *IEEE Internet of Things Journal*, 2023, 10(10): 8630-8644.
- [25] SHEN G, ZHANG J, MARSHALL A, et al. Towards scalable and channel-robust radio frequency fingerprint identification for LoRa[J]. *IEEE Transactions on Information Forensics and Security*, 2022, 17: 774-787.
- [26] COMANICIU D, MEER P. Mean shift: A robust approach toward feature space analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 603-619.
- [27] DIGNE F, BAUSSARD A, KHENCHAF A, et al. Classification of radar pulses in a naval warfare context using Bézier curve modeling of the instantaneous frequency law[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2017, 53(3): 1469-1480.

作者简介:



杨宁(1996-),女,博士研究生,研究方向:辐射源个体识别, E-mail: corrine_yang@126.com。



王裕(1983-),女,副教授,研究方向:智能信号识别。



张邦宁(1963-),男,教授,博士生导师,研究方向:卫星通信、通信抗干扰技术、辐射源个体识别等。



丁国如(1986-),男,教授,研究方向:电磁频谱深度感知及高效利用理论、方法与应用。



郭道省(1973-),通信作者,男,教授,博士生导师,研究方向:辐射源个体识别、卫星通信、通信抗干扰技术等, E-mail: xyzgfg@sina.com。

(编辑:陈琚)