

基于时序-脉幅聚类的未知雷达辐射源 信号分选方法

林进健¹, 黄明军², 孙慧博¹, 林梓涵¹, 谢 恺^{1*}

(1. 中山大学电子与通信工程学院, 广东深圳 518107; 2. 中国电子科技集团公司第二十九研究所, 四川成都 610036)

摘 要: 在复杂电磁环境中, 多个未知雷达辐射源发射的脉冲在时域高度交织, 其射频、脉宽等参数彼此高度相似, 导致常规的分选方法性能下降。相比之下, 脉冲幅度(Pulse Amplitude, PA)受雷达天线方向图、波束扫描方式等物理机制影响, 特别是在机械扫描型雷达, 呈现出可识别的包络变化规律, 可为分选提供补充判别信息。基于此, 本文提出一种基于到达时间(Time Of Arrival, TOA)与脉冲幅度的雷达信号分选方法。本文首先分析PA在不同雷达工作模式下的时序变化规律, 借鉴密度聚类思想, 在TOA-PA 2维空间中, 结合邻域半径与局部斜率变化约束, 识别具有相似几何形态的脉冲子集, 生成初始聚类路径组。为解决因漏脉冲或噪声干扰导致的同源轨迹断裂问题, 提出了聚类路径融合方法。通过时间重叠率筛选候选路径对, 计算全局及局部斜率熵以评估PA趋势一致性, 并采用Hausdorff距离度量路径间空间相似性度量, 实现相似路径融合, 构建了具有物理可解释性的PA包络轨迹。最后, 对融合后的TOA序列构建一阶差分直方图, 结合关联脉冲对方法完成脉冲重复间隔(Pulse Repetition Interval, PRI)候选分组与参数统计。实验在4种仿真场景下进行, 涵盖10%~50%不同组合的漏脉冲率与噪声脉冲率。以纯度、F值、福尔克斯-马洛斯指数和调整兰德系数4项指标评估聚类性能, 并与7种主流聚类算法对比。结果表明, 所提方法在综合性能上显著优于对照组; 路径融合机制有效抑制“增批”问题, 提升聚类时序连续性与辐射源对应一致性; PRI估计平均相对误差不超过0.6%。本文研究了基于TOA与PA联合特征的初分选, 引入基于时空相似度的路径融合策略, 利用TOA一阶差分直方图进行PRI主分选, 完成对PRI的检测。该方法适用于非合作电子侦察中未知辐射源的分选问题。后续研究可聚焦于聚类超参数的自适应整定、多节拍分选结果的证据融合机制, 以及在实测脉冲描述字(Pulse Descriptive Word, PDW)数据集上的泛化性能验证。

关键词: 时序-脉幅; 聚类; 未知雷达辐射源; 信号分选; 融合

基金项目: 科技委基础加强重点基金(No.2022-JCJQ-JJ-0726); 国防预研项目(No.1002, No.917, No.1004); 中山大学科学研究重点建设项目(No.2024_76150_B25899); 电磁空间作战与应用重点实验室开放基金(No.CETCKL20230010)

中图分类号: TN957.51 **文献标识码:** A **文章编号:** 0372-2112(2025)09-3256-18

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250341

A Temporal-Pulse Amplitude Clustering-Based Unknown Radar Emitter Signal Sorting Method

LIN Jin-jian¹, HUANG Ming-jun², SUN Hui-bo¹, LIN Zi-han¹, XIE Kai^{1*}

(1. Department of Electronics and Communication Engineering, Sun Yat-sen University of China, Shenzhen, Guangdong 518107, China;

2. The 29th Research Institute of China Electronics Technology Group Corporation, Chengdu, Sichuan 610036, China)

Abstract: In complex electromagnetic environments, pulses emitted from multiple unknown radar emitters are highly interleaved in the time domain. Conventional deinterleaving methods suffer from performance degradation because key parameters, such as radio frequency (RF) and pulse width (PW), often exhibit high similarity to one another. In contrast, pulse amplitude (PA) is influenced by underlying physical mechanisms, including the antenna radiation pattern and the beam scanning mode. This is particularly evident in mechanically scanned radars, where PA presents recognizable envelope variation patterns that can provide supplementary discriminative information for deinterleaving. Based on this premise, this paper proposes a radar signal deinterleaving method founded on the joint use of time of arrival (TOA) and pulse amplitude (PA). The

proposed method first analyzes the temporal variation patterns of PA under different radar operational modes. Inspired by density-based clustering, it identifies pulse subsets with similar geometric morphologies in the 2D TOA-PA space by combining constraints on neighborhood radius and local slope variations, thereby generating an initial set of cluster paths. To address the fragmentation of co-source tracks caused by missing pulses or noise interference, a cluster path fusion method is introduced. It screens candidate path pairs through temporal overlap, calculates global and local slope entropy to assess PA trend consistency, and employs the Hausdorff distance to measure spatial similarity between paths. This process merges similar paths to reconstruct physically plausible PA envelope tracks. Finally, a first-order difference histogram is constructed from the TOA sequence of the fused tracks, and pulse repetition interval (PRI) candidate grouping and parameter statistics are completed through an associated pulse pair analysis. Experiments are conducted in four simulated scenarios, covering various combinations of missing pulse rates and noise pulse rates from 10% to 50%. Clustering performance was evaluated using four metrics—purity, *F*-score, Fowlkes-Mallows index (FMI), and adjusted rand index (ARI)—and benchmarked against seven mainstream clustering algorithms. The results demonstrate that the proposed method significantly outperforms the control group in overall performance. The path fusion mechanism effectively suppresses the generation of spurious emitters, enhances the temporal continuity of clusters, and improves their correspondence to true emitters. The average relative error for PRI estimation did not exceed 0.6%. In summary, this paper performs an initial sort using joint TOA-PA features, introduces a path fusion strategy based on spatio-temporal similarity, and conducts the main deinterleaving via a TOA first-order difference histogram to achieve robust PRI detection. The approach is well-suited for deinterleaving unknown emitters in non-cooperative electronic reconnaissance. Future research could focus on the adaptive tuning of clustering hyperparameters, an evidence fusion mechanism for multi-epoch deinterleaving results, and validation of generalization performance on measured, real-world pulse descriptive word (PDW) datasets.

Key words: temporal-pulse amplitude; clustering; unknown radar emitters; signal sorting; fusion

Foundation Item(s): Key Foundation Project of the Science and Technology Commission for Strengthening Basic Research (No.2022-JCJQ-JJ-0726); National Defense Pre-research Project (No.1002, No.917, No.1004); Key Scientific Research Construction Project of Sun Yat-sen University (No.2024_76150_B25899); Open Fund of the Key Laboratory of Electromagnetic Space Operations and Applications (No.CETCKL20230010)

1 引言

在现代电子频谱作战中,电子支援作为获取敌方雷达情报和实施电磁压制的重要手段,扮演着至关重要的角色.通过对战场中不断变化的电磁波信号进行实时监测和捕获,尤其是对复杂脉冲流的精准解析,获得非合作方的雷达参数信息,用于干扰非合作方的行动并引导我方向着有利方向发展.因此,雷达信号分选技术作为整个电子支援系统的核心环节,其任务在于从多个随机交叠的辐射源脉冲流中有效分离出各自的信号,并对其特征进行深入分析,从而既能够及时更新雷达数据库,显著提升目标识别的准确性与系统整体鲁棒性^[1].

在雷达信号侦察设备中,为了实现较高的信号截获概率,通常采用频域宽带接收策略.接收机经过预处理后提取出脉冲描述字(Pulse Descriptive Word, PDW),其中包含 TOA、载波频率(Radio Frequency, RF)、脉冲宽度(Pulse Width, PW)、PA 以及到达方向(Direction Of Arrival, DOA)等关键参数^[2].围绕着 PDW 的辐射源信号分选方法,学者和工程师进行了一系列研究,主要可以归结以下 2 个方面.

基于单参数的雷达信号分选方法.由于 PRI 在雷

达信号中通常表现出较高的稳定性和易于测量的特点,早期研究多采用基于 PRI 单参数的分选策略.经典方法包括统计直方图法^[3]、累计差直方图(Cumulative Difference histogram, CDIF)算法^[4]、序列差直方图(Sequential Difference histogram, SDIF)算法^[5]、PRI 变换算法^[6]以及平面变换法^[7].这些方法通常通过构造 TOA 差值直方图,结合合适的阈值对各辐射源的 PRI 进行估计,实现信号分离.随后,学者进一步提出了增强经典算法^[8,9]以及多种方法的集成策略^[10];文献[11]利用隐马尔可夫模型对交织脉冲流建模,识别了 5 种主要 PRI 调制类型;文献[12]将语义分割引入雷达脉冲分选过程,采用双向递归神经网络取得较好效果.然而,基于 PRI 分选的方法难以适应如今复杂电磁环境,处理性能较差^[13].为了构建鲁棒性更高的分选模型,基于多参数联合的雷达信号分选方法成为研究的热点,特别基于人工智能技术的多参数雷达信号分选研究成果不断涌现出来.文献[14]提出了一种基于 ResGCN-BiLSTM-SE 网络的分选模型,通过构建自适应加权邻接矩阵有效整合时空特征;文献[15]引入离散更新马氏链混合序列的分选方法,实现了对惩罚似然性得分的最大化;文献[16]则将 TOA、CF、PA 和 PW 映射到点云空间,将雷达脉冲分选问题转换为点云分割任务,基于图卷积的

多级图卷积网络在应对脉冲丢失和伪脉冲干扰方面展现出较强鲁棒性。尽管上述方法在不同应用场景中均取得一定的进展,但在非合作电子侦察环境下,截获的脉冲流往往缺乏充分的先验信息,而信号真实类别数也难以准确预知,这对现有分选算法提出了更高的要求和挑战。

此外,雷达辐射源全脉冲信号一般在时域、频域、空域上具备特定的变化规律,利用 TOA、RF、PW、DOA 等信息对雷达全脉冲序列进行稀疏主分选的脉冲流或者提取参数间相关性进行分选,达到较好分选效果。但是,当高脉冲密度和多个辐射源同时到达接收机时,其脉冲序列会按照脉冲的到达时间先后进行排序,而大多数信号在时域、频域、空域上有一定的交叠,这些具有捷变性参数,具有较高的相似性,给分选带来新的困难^[17,18]。因此,需要从新的视角探索脉冲信号分选方法。

由于 PA 的变化规律与雷达系统工作状态密切相关,这一特性为雷达信号的深入分析和脉冲分选提供了重要依据。近年来,一些研究人员开始关注到利用 TOA 和 PA 特征对辐射源进行脉冲信号分选。文献[18]认为天线扫描类型引入对雷达参数的分选处理是有益处的;文献[19]利用 PA 的变化特性,通过分段埃尔米特插值和幅度差算法设计 PA 包络提取方法,实现严重混叠情况下的雷达辐射源信号分选;文献[20]结合实例分割网络 SOLOv2 和天线扫描模式,提出一种端到端的信号分选方法;文献[21]利用双向选通递归单元处理脉冲流的到达时间差和 PA 数据,实现同优先级不同机械扫描雷达信号的分选方法。因此,本文将分析 PA 包络特性并利用 PA 特征完成脉冲信号分选。

为克服现有方法受未知辐射源数量及信号捷变性参数的影响,本文提出了一种基于 TOA-PA 聚类的未知雷达辐射源分选方法。首先,根据全脉冲数据中 PA 特征的数学模型,分析了 PA 变化规律,提出一种基于 TOA-PA 的聚类算法,该算法无须预先知道辐射源的个数。然后,依据前一步得到的聚类路径组,联合时间和空间信息对多个聚类路径组进行融合,显著降低聚类路径数量,构建具有物理可解释性的聚类结果。最后,通过构建 TOA 一阶差分直方图,对融合后的聚类结果进行 PRI 主分选,实现了对复杂 PRI 的检测。实验结果表明,在漏脉冲率和信号干扰较高的复杂场景下,本方法在 4 项评价指标累计差异上均展现出与 7 种传统聚类算法相比更为优异的性能。

因此,本文提出的基于 TOA-PA 聚类的新方法,在复杂电磁环境下展现出优于 7 种传统算法的准确性。其核心创新主要体现在:

(1) 基于 TOA-PA 联合建模的路径聚类策略。区别于传统方法并未考虑脉冲的时序关系进行聚类,本文方法结合 TOA 和 PA 的时序变化关系,构建具有物理意义的聚类路径,提升对复杂交织信号的分选能力。

(2) 引入路径融合机制提升聚类一致性。针对 PDW 数据中常见的脉冲交错与混叠问题,本文设计了基于时空相似性的路径融合机制,有效缓解“增批”现象,增强了聚类结果的连续性和稳定性。

(3) 结合 PRI 检测增强工程实用性。在聚类基础上,进一步利用 TOA 一阶差分直方图完成 PRI 分选,作为对聚类结果的补充验证手段,显著提升了对未知雷达信号的解析能力和整体分选流程的完整性。

2 TOA-PA 联合建模与模式分析

2.1 TOA-PA 联合模型构建

在一些典型的电子侦察应用场景如机械扫描雷达环境中,雷达信号的 TOA 与 PA 之间呈现出特定的关联模式。例如,当雷达天线以恒定角速度沿固定轨迹进行扫描时,接收端所接收到的脉冲幅度会随时间呈现周期性的波动变化。这种变化反映了天线主瓣方向增益的最大值随着扫描角度的变化而动态移动,从而导致接收信号强度的非线性变化。

基于这一物理机制,本文提出了一种新的聚类方法,该方法充分利用 TOA 与 PA 之间的联合变化特性,通过构建路径扩展模型和相似性度量机制,在存在多个类型相同或相近的雷达信号交织的情况下,也能有效地区分不同辐射源,实现对复杂电磁环境下未知雷达信号的有效分选。

假设某一雷达辐射源发射的脉冲序列为 $S = \{(t_i, a_i)\} (i=1, 2, \dots, N)$, 其中 t_i 表示第 i 个脉冲的到达时间, a_i 表示对应的 PA 值。将 PA 视为关于时间的函数 $a(t)$, 并在此基础上建立 TOA-PA 的联合函数模型。

通常情况下,侦察接收系统接收到的雷达功率来衡量雷达脉幅特征,计算公式为^[22,23]

$$P_{\text{receive}}(t) = \frac{P_{\text{transmit}} G_{\text{receive}} \lambda^2}{(4\pi R)^2 L} G_{\text{transmit}}[\theta(t), \phi(t)] \quad (1)$$

其中, $P_{\text{receive}}(t)$ 是接收器接收到的脉冲功率, P_{transmit} 是发射功率, G_{receive} 是接收器天线增益, λ 是信号波长, R 是雷达与目标之间的距离, L 是传播损耗, $G_{\text{transmit}}[\theta(t), \phi(t)]$ 是发射天线在方向 $(\theta(t), \phi(t))$ 上的增益。

在雷达辐射源与侦察平台相对静止,或雷达天线扫描速率远大于两者之间的相对运动速度的情况下,脉冲信号的功率主要由发射天线的方向图函数 $G_{\text{transmit}}[\theta(t), \phi(t)]$ 决定。在电子战中大多属于这种情况,所以 $P_{\text{transmit}} G_{\text{receive}} / (4\pi R)^2 L$ 被认为是恒定的。考虑到实

际应用中常用对数形式表示功率水平,一般取功率密度的对数值作为PA值.由此可得侦测设备接收到的目标雷达脉冲幅度表达式为

$$PA(t) = 10 \lg \frac{P_{\text{transmit}} G_{\text{receive}} \lambda^2}{(4\pi R)^2 L} G_{\text{transmit}} [\theta(t), \phi(t)] \quad (2)$$

其中, $PA(t)$ 为接收到目标雷达的脉幅值, \lg 是取对数操作.

注意到, $\theta(t)$ 和 $\phi(t)$ 是随时间变化的方位角和仰角, 导致 $F(\theta(t), \phi(t))$ 也是时间的函数. 因此, PA 不是一个静态值, 而是随着时间演化的一个可观测量.

由前面分析可知, 在侦察机与目标雷达的距离一定时, 雷达脉幅值的变化取决于雷达天线的扫描方式, 天线方向图以及所处工作状态. 由于天线方向图函数的精确表达式通常是复杂的, 在后续的分析实验中, 为了便于计算分析, 机械扫描雷达(如圆形扫描)的天线增益方向图通常近似为 sinc^2 函数:

$$G(t) = G_0 \cdot \text{sinc}^2 \left(\frac{\omega t - \theta_0}{\theta_{\text{BW}}} \right) \quad (3)$$

其中, G_0 是最大天线增益, ω 是天线扫描角速度, θ_0 是初始相位角, θ_{BW} 是半功率波束宽度.

当主瓣指向接收机时 ($\theta(t) \rightarrow 0$), 通过泰勒展开将 sinc^2 函数近似为二次函数, 即 $G(t) \propto -a(t - t_0)^2$. 由于 $P_{\text{receive}}(t) \propto G(t)$, 脉冲幅度与 TOA 的关系可建模为

$$A(t) = a(t - t_0)^2 + b(t - t_0) + c \quad (4)$$

其中, $A(t)$ 表示脉冲幅度, t_0 是对应于主瓣指向接收平台的时间点.

2.2 雷达工作模式下的 TOA-PA 模式规律

为了进一步分析 TOA-PA 联合空间中数据点的模式规律, 可以从典型雷达工作模式出发, 分析不同模式下 PA 在时间轴上的变化趋势及其物理成因.

单目标跟踪模式 (Single Target Tracking, STT) 中, 雷达持续锁定单一目标, 天线主瓣始终指向目标方向. 此时, $F(\theta, \phi)$ 近似恒定, PA 值基本稳定. 由于目标距离缓慢变化, PA 可能呈现缓慢上升或下降的趋势. 所以, 在 TOA-PA 平面中, STT 模式下的 PA 曲线表现为一条近似水平或缓慢变化的直线. 边扫描边跟踪模式 (Track While Scan, TWS) 中, 雷达在搜索的同时保持对多个目标的跟踪. 波束周期性扫描空域, 当扫过某目标时, PA 显著增强; 远离时则迅速衰减. 所以, PA 值呈现周期性起伏, 其变化形态近似 sinc 函数. 在 TOA-PA 平面中, 这种模式下的 PA 曲线呈现出明显的周期性包络结构, 每个周期对应一次波束扫描过程. 范围搜索 (Range While Search, RWS) 模式主要用于探测目标的大致位置. 在此模式下, 雷达波束在较大范围内扫描, PA 值随波束指向的变化而波动, 呈现出类似正弦函数

的周期性变化. 所以, 在 TOA-PA 平面中, RWS 模式下的 PA 曲线表现为具有一定振幅和频率的波动曲线. 速度搜索 (Velocity Search, VS) 模式适用于远距离目标的速度估计. 尽管 PA 值受到多普勒效应等因素的影响, 但仍保留了扫描包络的 sinc 函数特性, 整体上仍体现出一定的周期性和包络结构^[18,24].

综上所述, 不同雷达工作模式下的 PA 变化具有不同的时域特征, 这些特征在 TOA-PA 平面中形成了具有辨识性的路径样式. 这为后续基于路径匹配和趋势一致性的聚类方法提供了坚实的物理依据.

3 本文提议的方法

在复杂电磁环境下, 多部雷达发射的脉冲信号往往互相交织. 基于各辐射源 PA 变化规律的差异, 本文提出了一种全新的 TOA-PA 聚类算法, 实现对脉冲数据的初聚类, 形成多条 PA 轨迹样式的聚类路径. 随后, 结合聚类路径组的时间和空间信息约束, 对聚类路径进行同类合并, 降低由脉冲数据混叠、交错引起的“增批”问题, 得到了具有物理可解释性的聚类结果. 最后, 对聚类合并结果进行 PRI 主分选与参数统计, 实现了不同雷达脉冲信号的分选.

3.1 TOA-PA 聚类算法

根据 PDW 数据中 PA 时序特征分析, 本方法旨在通过识别局部几何相似性较强的点来形成簇, 实现对复杂脉冲信号模式的分选. 借鉴了具有噪声的基于密度的聚类方法 (Density-Based Spatial Clustering of Applications with Noise, DBSCAN)^[25] 的思想, 本文提出了一种基于 PA 局部几何特征扩展的包络聚类算法——TOA-PA 聚类算法.

为了清楚起见, 下面借鉴 DBSCAN 中的概念, 并给出一些定义.

定义 1 邻域. 给定对象在半径 r 内的区域称为该对象的 r -邻域, 即为

$$N_r(q) = \{p \text{ 属于 } D \mid \text{dist}(p, q) \leq r\} \quad (5)$$

其中, $\text{dist}(\cdot)$ 表示计算距离函数, 如切比雪夫距离.

定义 2 直接几何可达. 如果对象 p 在对象 q 的 r -邻域内, 并且满足斜率变化小于阈值 θ , 则对象 p 从对象 q 出发是直接几何可达的. 即为

$$\text{DGR}(p, q) = \begin{cases} \text{dist}(p, q) \leq r \\ |\text{slope}(p, q)| \leq \theta \end{cases} \quad (6)$$

其中, $\text{slope}(p, q)$ 表示计算对象 p 与 q 斜率函数.

定义 3 几何可达. 如果存在一个对象链 p_1, p_2, \dots, p_n , 其中 $p_1 = q, p_n = p$, 并且每一对相邻点 p_i 和 p_{i+1} 都是直接几何可达, 则对象 p 是从对象 q 关于 r -邻域和 θ 几何可达.

定义 4 簇. 通过局部几何特征扩展形成的路径也

是最大几何相连的对象集合。路径扩展过程中,所有符合条件的点被依次加入路径,直到无法找到新的符合条件的点为止。形成的路径即为一个簇。

定义 5 噪声点。未被任何路径访问过的点被视为噪声点。这些点不包含在任何簇中,被认为是异常值或孤立点。

在明确了上述关键概念后,接下来详细介绍 TOA-PA 聚类算法的具体操作步骤,包括路径扩展、簇形成与终止条件以及完整的算法 1 流程。

算法 1 路径扩展

输入:脉冲数据集 data, 搜索半径 radius, 斜率变化阈值 theta

输出:当前路径 current_path

1. 初始化

数据预处理:

data_sorted ← 对 data 按 TOA 列排序

初始化变量:

将所有对象标记为未访问

全局访问数组标记 ← FALSE

2. 遍历脉冲数据点查找邻域对象

FOR 每个未访问的数据点 p DO

初始化新路径:

current_path ← { p }; #将 p 作为路径起点

current_slope ← NaN; #初始化当前斜率

WHILE TRUE DO

获取候选对象集:

未访问点 q 且 $q(1)$ 大于 $p(1)$

IF 候选对象集为空

BREAK

END

FOR q DO #检查是否可以几何可达

q 在 p 的邻域内 \leq radius

计算 q 与 p 的斜率变化 ← temp_slope(p, q)

IF temp_slope \leq theta

current_path ← current_path \cup { q }; #将 q 添加到路径中

更新当前斜率 current_slope

标记 q 为已访问

BREAK 内层 FOR 循环; #继续寻找下一个候选点

END

END

IF 未找到新的候选点

BREAK;

END

IF length(current_path) $>$ 3

clusters ← current_path

3.1.1 路径扩展

路径扩展是 TOA-PA 聚类算法的核心过程,它通过逐步添加邻近对象来构建路径,进而形成簇。

3.1.2 簇形成与终止条件

完成路径扩展后,需要确定哪些路径应被视为簇,同时识别出噪声点。具体步骤如下。

步骤 1: 路径长度检查。检查当前路径的长度,如果路径长度大于或等于最小长度 3, 则将该路径视为一个簇。

步骤 2: 簇存储。将形成的簇存储到簇列表中,同时更新全局访问状态数组,标记路径中的所有对象为已访问。

步骤 3: 噪声点识别。从未被任何路径访问过的对象被视为噪声点。

3.1.3 算法步骤

综上所述, TOA-PA 聚类算法的整体步骤见算法 2。

算法 2 TOA-PA 聚类算法

输入:脉冲数据 data, 搜索半径 radius, 斜率变化阈值 theta

输出:簇列表 clusters, 噪声点列表 noise

1. 对输入数据进行归一化处理。

[data_normalized, min_max_params] ← NORMALIZE_DATA(data)

2. 初始化必要的变量

n ← SIZE(data_normalized, 1)

clusters ← \emptyset

3. 路径扩展过程, 形成符合条件的簇。

算法 1 路径扩展

4. 识别未被任何路径访问过的噪声点。

noise ← data_normalized[全局访问标记数组 == FALSE, :]

5. 清理不包含任何点的空簇。

clusters ← REMOVE_EMPTY(clusters)

3.2 聚类路径融合算法

尽管 TOA-PA 聚类算法能够较好地提取出 PA 轨迹,但在复杂环境中, PDW 数据存在严重的漏脉冲、噪声脉冲以及多个未知辐射源的脉冲信号的 PA 包络严重混叠和交错的复杂情况,同一条 PA 包络轨迹往往会被分裂为 2 个或更多聚类,带来“增批”的分选问题。

为了应对这一挑战,采用融合方法合并相似聚类。一种简单的方法使用相似度量函数计算每个簇的距离,当小于某个阈值可以认为是同一个簇。目前已有基于欧几里得距离^[26]和基于脉冲组的时序关系的聚类合并算法^[27]。虽然基于欧几里得距离的算法较为直观,但要求待对比 2 个向量的维度一致。而利用脉冲组时序关系的算法假设聚类结果以连续脉冲组形式存在。

在本节中,介绍了一种新的聚类路径合并算法。在 TOA-PA 聚类算法中,由于利用了搜索半径和斜率变化对接收到脉冲序列进行了约束,输出的聚类组将以线条状路径的形式呈现。因此,通过联合时序和空间关系分析这些线条状聚类组之间的相似度,并对满足条件的聚类路径结果进行合并处理,从而保证了聚类结果

的整体连续性和一致性。

3.2.1 合并原理

设 P 为所有聚类路径组的集合,每次迭代中选择一个基准路径 P_{base} ,并寻找与之相似的路径集合 S 。基于时空相似度量方法对潜在相似的路径分配相同的标签,合并后的路径 P_{merged} 为

$$P_{\text{merged}} = \bigcup_{P \in S} P \quad (7)$$

通过不断迭代,直到所有聚类组都被合并或无法再找到相似的聚类组为止。

3.2.2 时空相似度量

在评估 2 条聚类路径之间的相似性时,本文提出了一种基于时间交集、时序相关性和空间相似性的多阶段分析方法。

首先,计算时间范围分别为 $[x_{\min_1}, x_{\max_1}]$ 和 $[x_{\min_2}, x_{\max_2}]$, 2 条聚类路径 P_1 和 P_2 的重叠比例(Overlap Ratio)。若重叠比例大于等于 50%,或者当一条路径完全包含另一条路径且重叠比例至少达到 30% 时,则认为这 2 条聚类路径在时间上具有显著的交集。具体公式如下:

$$\text{overlap_ratio} = \frac{\max(0, \min(x_{\max_1}, x_{\max_2}) - \max(x_{\min_1}, x_{\min_2}))}{\max(x_{\max_1}, x_{\max_2}) - \min(x_{\min_1}, x_{\min_2})} \quad (8)$$

(1) 时序相似性计算。根据时间交集的结果,选择不同的时序相似性度量方法。全局斜率熵是用于衡量路径在整体范围内的趋势一致性。如果 2 条路径的全局斜率熵差异较小,则可以认为它们在整体趋势上具有较高的相似性。对于全局斜率熵的计算公式如下:

$$H = - \sum_{i=1}^m p_i \log p_i \quad (9)$$

其中, m 是划分区间数, p_i 是斜率出现的概率。

通过比较合并前后熵值的变化来判断 2 条路径是否具有相似的整体趋势。合并前的总熵值定义为

$$H_{\text{before}} = H_1 + H_2 \quad (10)$$

其中, H_1 是聚类路径 1 全局斜率熵, H_2 是聚类路径 2 全局斜率熵。

将 2 条路径的数据点合并,按时间轴排序,计算合并后路径的全局斜率熵为 H_{after} 。如果合并后的熵值满足以下条件,则认为 2 条路径在整体趋势上具有较高的相似性:

$$H_{\text{after}} \leq H_{\text{before}} + \varepsilon_{\text{global}} \quad (11)$$

其中, $\varepsilon_{\text{global}}$ 是一个可调阈值。

当路径没有明显的时间交集时,计算其在局部区域内的斜率熵,以评估局部趋势的一致性。局部斜率熵与全局斜率熵类似,但仅考虑路径在重叠部分的数据点。

(2) 空间相似性计算。仅当时序相关性分析结果表明 2 条路径相似时,才会进一步计算空间相似性。评估 2 条聚类路径的空间距离相似性,使用 Hausdorff 函数度量 2 条距离路径的空间距离,当 2 条聚类路径的 Hausdorff 距离小于等于 $\varepsilon_{\text{distance}}$ 时,认为该 2 条聚类路径相似,即为

$$\begin{cases} d(A, B) \leq \varepsilon_{\text{distance}}, & \text{聚类路径相似} \\ d(A, B) > \varepsilon_{\text{distance}}, & \text{聚类路径不相似} \end{cases} \quad (12)$$

其中, $\varepsilon_{\text{distance}}$ 是一个可调阈值, $d(A, B) = \max\left(\sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|a - b\|\right)$ 。

3.2.3 相似聚类路径合并过程

对于满足相似性条件的路径集合,将其合并为一个新的聚类组。合并过程中,确保新聚类组的时间顺序性和空间连续性,同时排除可能存在的离群点。具体步骤见算法 3。

4 实验与分析

4.1 实验环境、仿真数据与预处理

4.1.1 实验软硬件配置

本文实验使用硬件环境为 12th Gen Intel(R) Core(TM) i7-12700 CPU 2.10 GHz,操作系统是 Windows 10,软件是 Matlab 2021b 和 Python 3.7。

4.1.2 仿真数据集设置

本文仿真的雷达辐射源分选数据集包含了 7 部雷达辐射源。在构建仿真数据集时,仿真过程参考了文献[20]中的思路,考虑了影响排序性能的关键参数,包括调制类型、天线扫描模式和增益模式。通过参考现有先进信号分选方法^[28,29],确定了调制参数范围,如 PRI、RF 和组变 PRI 情况。由于天线扫描参数影响接收到的信号功率,在仿真中考虑影响脉冲幅度的关键参数,如扫描周期和波束宽度。仿真信号样本包含常规 PRI、组变 PRI、常规频率、频率捷变和频率分集等脉冲调制类型,对脉冲数据进行脉冲丢失和脉冲干扰处理,增加数据的复杂性。具体的每部雷达源参数如表 1 所示。丢失率设置为从 10%~50% 变化,具体实现时,会根据该丢失率随机选择一部分脉冲作为漏脉冲并从交错序列中移除,然后重新计算剩余脉冲的时间差。而脉冲干扰处理过程在脉冲参数仿真过程中,针对 RF、PW、PRI 和 DOA 等参数添加高斯噪声,各参数对应的高斯噪声标准差设定如下:对于 RF 调制,标准差为 2;对于 PRI 调制,标准差为 1;对于 PW 调制,标准差为 0.1;对于 DOA 设定,标准差为 1;对于机械扫描的 PA 计算,标准差为 0.05。

此外,还设置了噪声脉冲,交错脉冲序列中的噪声脉冲率可设定为从 10%~50% 变化,表示脉冲序列中有相应比例的脉冲为随机噪声脉冲。其中,噪声脉冲的

算法3 聚类路径合并

输入:待处理路径集合 paths,全局时序相似度阈值 $\varepsilon_{\text{global}}$,局部时序相似度阈值 $\varepsilon_{\text{local}}$,空间相似度阈值 $\varepsilon_{\text{distance}}$

输出:合并后的路径集合 merged_paths

```

merged_paths ← 空列表
WHILE paths 非空 DO:
    reference_path ← 提取 paths 中的第一条路径
    similar_paths ← {reference_path} // 初始化相似路径集合
    FOR i FROM 1 TO len(paths) - 1 DO:
        candidate_path ← paths[i]
        // 判断时间重叠是否满足条件
        is_overlap_ratio ← calculate_overlap(reference_path, candidate_path)
        is_similar ← FALSE
        IF is_overlap_ratio THEN: // 计算全局时序相似度
            temporal_similarity ← 全局时序相似性(reference_path, candidate_path)
            IF temporal_similarity <  $\varepsilon_{\text{global}}$  THEN:
                spatial_similarity ← 空间相似性(reference_path, candidate_path)
                IF spatial_similarity <  $\varepsilon_{\text{distance}}$  THEN:
                    is_similar ← TRUE
                END IF
            ELSE: // 计算局部时序相似度
                temporal_similarity ← 局部时序相似性(reference_path, candidate_path)
                IF temporal_similarity <  $\varepsilon_{\text{local}}$  THEN:
                    spatial_similarity ← 空间相似性(reference_path, candidate_path)
                    IF spatial_similarity <  $\varepsilon_{\text{distance}}$  THEN:
                        is_similar ← TRUE
                    END IF
                END IF
            IF is_similar THEN:
                similar_paths ← similar_paths ∪ {candidate_path}
                DELETE paths[i]
            END IF
        END FOR // 合并路径
        IF len(similar_paths) > 1 THEN:
            merged_path ← concatenate_paths(similar_paths)
        ELSE:
            merged_path ← reference_path
        END IF
        merged_paths ← merged_paths ∪ {merged_path}
        DELETE paths[0]
    END WHILE

```

RF、PW、DOA 的范围及分布特性如下:RF 遵循均匀分布,范围为[1 000, 1 2000] MHz;PW 遵循均匀分布,范围为[1, 10] μs ;DOA 遵循均匀分布,范围为[0°, 360°];

PA 遵循均匀分布,范围为[-90, 0] dBm. 因此,在生成交错脉冲序列后,再处理漏脉冲和增加噪声脉冲,根据设定的丢失率随机移除脉冲,根据噪声脉冲率生成随机噪声脉冲并与真实脉冲序列合并,最后重新计算合并序列的时间差.为了统一表述,下面使用 (a, b) 分别表示漏脉冲率和噪声脉冲率参数设置,如(0.1, 0.2)表示漏脉冲率为10%,噪声脉冲率为20%.

仿真实验构建了4个场景由表2给出,每个场景的所有雷达辐射源同时开机,侦察接收机并非实时对准非合作方雷达辐射源主瓣进行数据接收,所以接收的脉冲流中会有空档期,按到达时间 TOA 顺序编程实现脉冲的交错处理.

4.1.3 算法参数设置

本文的 TOA-PA 聚类算法中涉及5个阈值.尽管这些阈值需要人为设定,并对测试效果直接产生影响,但是其具有明确的物理意义,可以根据测试场景和环境进行预设,如表3所示.

4.1.4 预处理工作

为了确保每种算法的输入数据一致,对雷达脉冲数据进行了统一的限幅滤波操作和最小—最大归一化.如图1展示了在场景4下脉冲数据滤波前后的3维可视化效果,图1(a)是原始数据分布,图1(b)是经过限幅滤波后的数据分布,其中红色脉冲数据是噪声脉冲.最小—最大归一化的等式在等式(13)中给出:

$$x' = \frac{x_{ij} - \min x_j}{\max x_j - \min x_j} \quad (13)$$

4.2 对比算法与评价指标**4.2.1 对比算法**

为了验证所提方法的有效性,选取了7种常用聚类算法进行性能对比,包括了K-means^[30]、DBSCAN、MS (MeanShift)^[31]、AP (Affinity Propagation)^[32]、GMM (Gaussian Mixture Model)^[33]、SC (Spectral Clustering)^[34]以及HDBSCAN^[35].在参数配置方面,为了保证不同算法在最优状态下运行,K-means、GMM 和 SC 使用了实际类别数作为参数输入,而其他算法则通过网格搜索调参以获得最佳效果.鉴于这些算法存在一定随机性,实验中每种方法均进行了50次随机试验,并采用平均结果作为最终评估值.

在预分选阶段,通常选取 DOA、RF、PW 3个维度作为聚类特征输入.图2展示了基于 DOA、RF、PW 3维特征的7种对比聚类算法的可视化结果;而图3则展示了以 TOA、PA、DOA、RF 和 PW 5维特征作为输入时的聚类效果.直观上,在场景4中可以看出7部雷达的 TOA-PA 曲线具有明显的包络特征.但从图2与图3的对比结果可知,相较于 TOA、PA、DOA、RF 和 PW 5维特征输入,利用 DOA、RF、PW 3维特征能更好地分选出不同的

表 1 仿真雷达辐射源参数设置

雷达编号	雷达 1	雷达 2	雷达 3	雷达 4	雷达 5	雷达 6	雷达 7
RF 类型	分集	固定	固定	分集	分集	分集	捷变
RF/MHz	6 017,6 517,6 717,6 917	6 517	6 910	6 316,6 416,6 916	7 120,7 320,7 420,7 320	7 120,7 320	6 750~7 050
PW 类型	抖动	抖动	抖动	抖动	抖动	抖动	抖动
PW/ μ s	3	1	3	3	2	3	2
PRI 类型	组变	组变	抖动	组变	组变	固定	组变
PRI/ μ s	400,400,300,300,200,200	99.4,699.4,399.4	699.4	700,300,900	899.4,699.4,399.4	600	300,300,700,700
DOA/ $^{\circ}$	71	68	70	78	80	75	73
PA 范围/dBm	[-56,-12]	[-63,-26]	[-58,-21]	[-42,-17]	[-43,-23]	[-46,-29]	[-63,-54]

表 2 分选仿真场景示意图

编号	场景 1	场景 2	场景 3	场景 4
雷达 1	√	√	√	√
雷达 2	√	√	√	√
雷达 3	√	√	√	√
雷达 4	√	√	√	√
雷达 5		√	√	√
雷达 6			√	√
雷达 7				√

表 3 TOA-PA 聚类算法的阈值

序号	阈值名称	阈值符号	设置依据	参考范围
1	搜索半径	radius	用于确定 2 个点是否足够接近形成路径的一部分	[0.16,1.38]
2	斜率变化阈值	theta	用于控制路径斜率的变化程度	[0.091,0.41]
3	全局斜率熵阈值	ϵ_{global}	用于控制全局范围内序列相似性	[0.05,1.95]
4	局部斜率熵阈值	ϵ_{local}	用于控制局部范围内序列相似性	[0.05,1.35]
5	空间距离阈值	$\epsilon_{\text{distance}}$	用于确定 2 个序列空间上的偏差	[0.25,4.81]

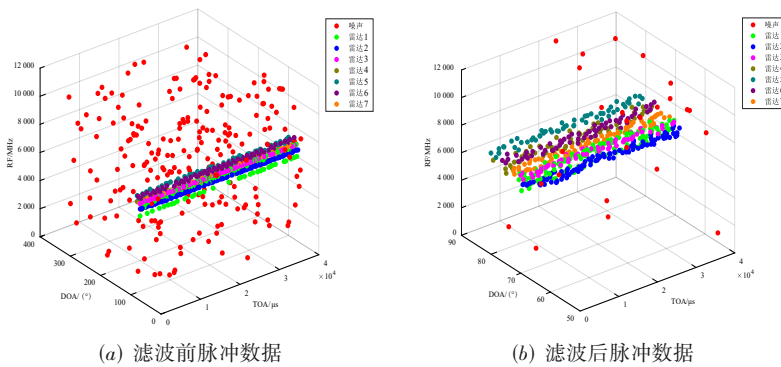


图 1 脉冲数据滤波前后效果

辐射源信号. 因此,本文在后续对比实验中,7种对比算法均采用 DOA、RF 和 PW 作为输入特征,本文算法使用 TOA、PA、DOA、RF 和 PW 5 维特征作为输入.

4.2.2 评价指标

在聚类分析过程中,评价指标起到了衡量各方法聚类结果质量的重要作用. 本文采用了以下 4 项指标: 纯度 (Purity)^[36]、F 值 (F-score)^[37]、福尔克斯-马洛斯基

数 (Fowlkes-Mallows Index, FMI)^[38] 和调整兰德系数 (Adjusted Rand Index, ARI)^[39]. 它们用来衡量聚类结果与真实类别 (Ground Truth) 标签之间的一致性. 以下是这几个指标的严格数学定义.

在给出指标前,先明确定义以下 4 个基本概念,在聚类任务中用于刻画聚类结果与真实类别的匹配情况,有以下 4 种情况:(1)真阳性 (True Positive, TP) 是指被聚

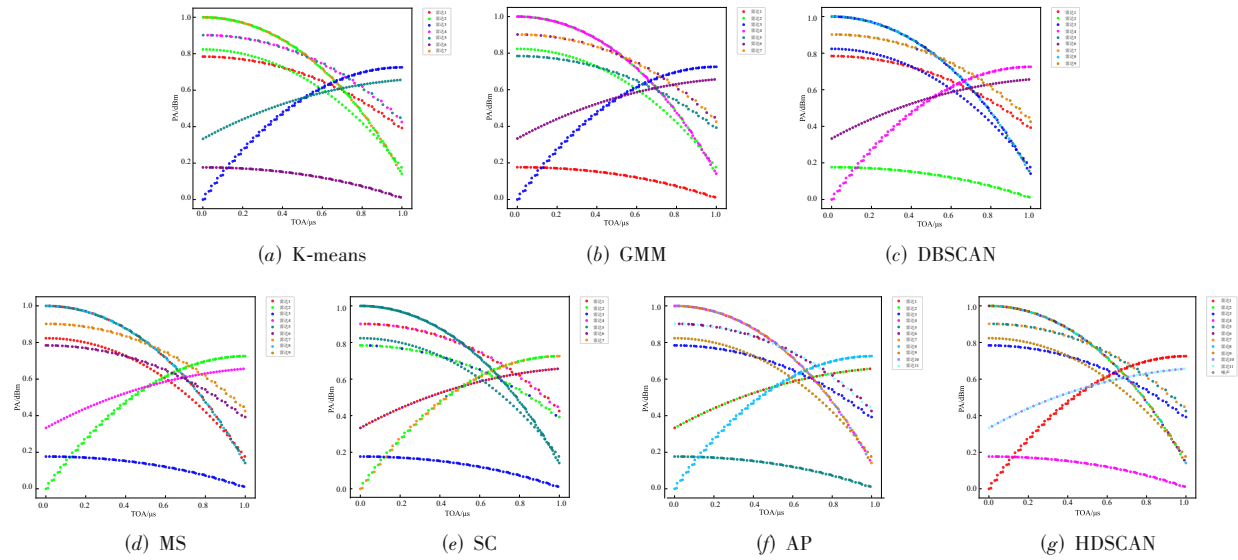


图2 场景4下3维特征输入的7种聚类算法结果

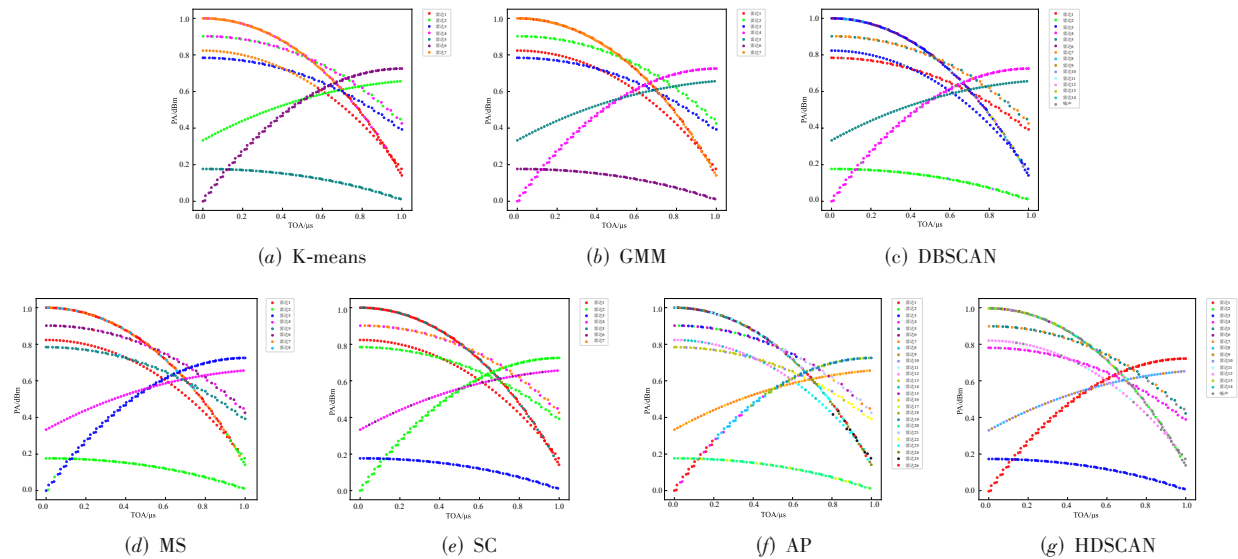


图3 场景4下5维特征输入的7种聚类算法结果

类算法认为属于同一类且在真实标签中也确实属于同一类的样本对数量;(2)假阳性(False Positive, FP)是指被聚类算法认为属于同一类、但在真实标签中不属于同一类的样本对数量;(3)真阴性(True Negative, TN)是指在真实标签中属于同一类、但被聚类算法错误地分到不同类的样本对数量;(4)假阴性(False Negative, FN)是指在真实标签中不属于同一类且被聚类算法正确地分到不同类的样本对数量。

(1) 纯度

设 $C = \{C_1, C_2, \dots, C_K\}$ 为聚类的结果, $T = \{T_1, T_2, \dots, T_J\}$ 为真实的类别标签集合。记 $|C_i \cap T_j|$ 为第 i 个聚类中属于第 j 个真实类别的样本数。纯度定义为

$$\text{Purity} = \frac{1}{N} \sum_{i=1}^K \max_j |C_i \cap T_j| \quad (14)$$

其中, N 是样本总数, K 是聚类的数量, C_i 是第 i 个聚类, T_j 是第 j 个真实类别。

(2) F 值

F 值是综合考虑查准率(Precision)和查全率(Recall)的评估指标,常用于分类和聚类评估。 F 值通过调和平均来平衡这2个指标,尤其适用于不平衡数据集。 F 值定义为

$$F_\beta = \frac{(1 + \beta^2) \cdot \text{Precision} \cdot \text{Recall}}{\beta^2 \cdot \text{Precision} + \text{Recall}} \quad (15)$$

当 $\beta=1$ 时,得到的是标准的 F_1 值。其中, Preci

sion=TP/(TP+FP), Recall=TP/(TP+FN).

(3) 福尔克斯-马洛斯指数

FMI 是用来评估 2 个聚类结果相似度的指标,它通过计算真阳性、假阳性和假阴性来衡量聚类的精度与召回率的平衡. FMI 值定义为

$$FMI = \frac{TP}{\sqrt{(TP+FP)(TP+FN)}} \quad (16)$$

(4) 调整兰德系数

调整兰德系数是通过标准化的 Rand 指数来衡量聚类与真实标签之间的一致性. ARI 定义为

$$ARI = \frac{RI - E(RI)}{\max(RI) - E(RI)} \quad (17)$$

其中, RI 是 Rand 指数, 表示样本对的一致性. $E(RI)$ 是随机聚类情况下的期望值. $\max(RI)$ 是 Rand 指数的最大值.

Purity、 F -score 和 FMI 取值范围为 $[0, 1]$, 值越大说明聚类结果与实际标签越相近; ARI 取值范围为 $[-1, 1]$, 数值越大说明聚类结果与实际标签一致性越高.

4.3 预分选效果与分析

本节将从聚类可视化效果和评价指标 2 方面, 对不同聚类算法的性能进行分析. 根据分选仿真场景漏脉冲

率和噪声脉冲率不同组合, 抽取了仿真场景 4 漏脉冲率和噪声脉冲率分别是 $(0, 0)$ 和 $(0.4, 0.4)$ 的效果进行分析.

图 4 展示了在无脉冲丢失和无噪声干扰条件下, 8 种聚类算法可视化结果, 而各算法在 4 项指标上的评估结果详见表 4. 从表 4 中可观察到, 在全部评价指标的平均值中, GMM 算法表现最佳, 其平均值较本文方法高出 4.86%. 从图 4 的直观效果可以看出, 尽管 GMM 在局部区域中存在少量脉冲误聚类现象, 但大部分脉冲仍被正确划分至对应的 PA 包络曲线上. 这主要归因于 GMM 能够通过多个高斯分布的线性组合灵活地拟合数据, 适合描述任意椭球状、非对称或重叠的簇结构. 与之相比, 尽管本文方法存在错误地将一些脉冲识别为噪声点, 但同样能够将大部分脉冲数据按 PA 包络曲线正确分类, 有效处理了包络曲线混叠和脉冲交错的问题. 对于 DBSCAN 和 HDBSCAN 来说, 其 4 项指标平均值分别为 75.76% 和 78.41%. 2 种算法在处理高密度区域时较为出色, 通常能将这些区域划分为单一簇, 但在脉冲数据密度分布不均或混叠较为严重时, 会出现部分脉冲归入错误簇的情况.

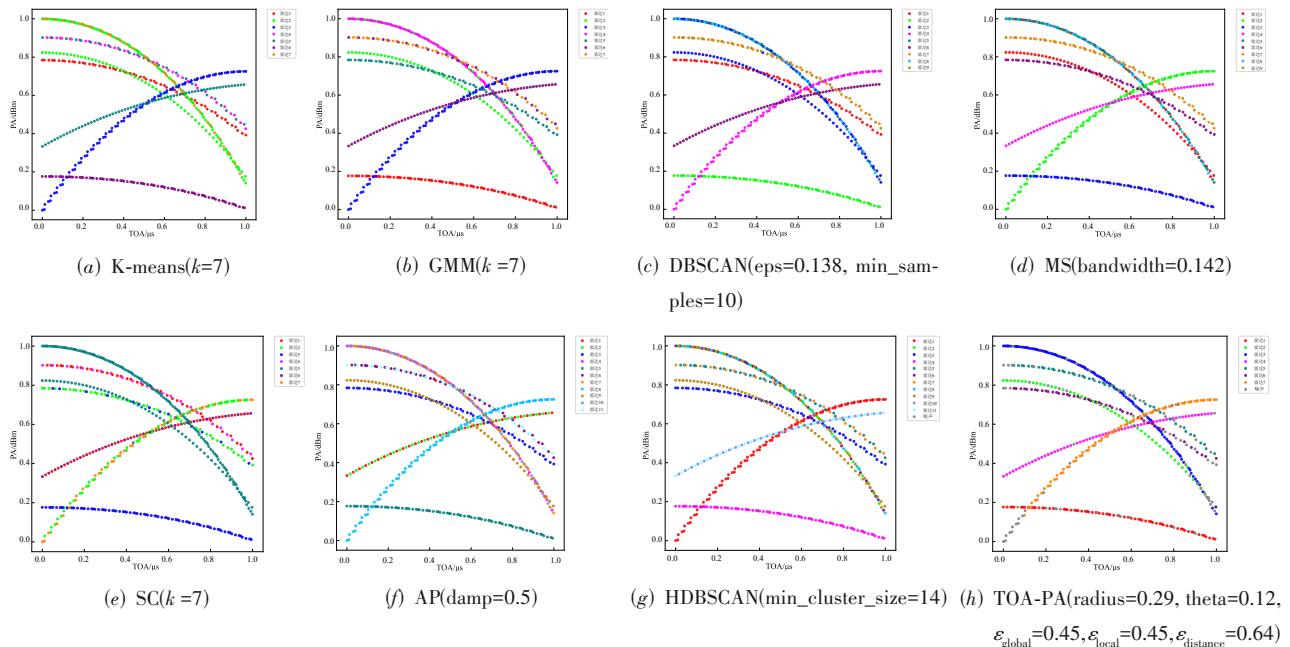


图 4 8 种聚类算法在场景 4(0, 0) 上的可视化结果

当漏脉冲率和噪声脉冲率均提高至 0.4 时, 观察到 7 种对比算法的 4 项指标均出现了大幅下降. 然而, 本文所提方法依然表现稳健, 性能相对改善了 13.02%. 从图 5 可见, 本文方法依然能够将绝大多数脉冲数据正确划分到对应的 PA 包络曲线内, 说明该方法对脉冲丢失和噪声干扰的鲁棒性较好. 这主要归因于 TOA-PA 聚

类算法在局部几何特征相似性扩展生成路径簇方面的优势, 即 PA 包络主骨干越清晰, 有利于聚类算法依据局部斜率和搜索半径扩展聚类路径簇.

为了深入探究漏脉冲率和噪声脉冲率在场景 1~4 下, 对 8 种聚类算法性能的具体影响, 开展了相关实验. 如图 6 所示, 每一行代表了相同漏脉冲率不同场景下,

表 4 不同聚类算法在场景 4 下性能

脉冲率	评价指标	K-means	GMM	DBSCAN	MS	SC	AP	HDBSCAN	本文方法
(0,0)	Purity	0.853 1	0.899 1	0.853 1	0.853 1	0.592 9	0.791 2	0.936 3	0.869 0
	F-score	0.773 6	0.841 3	0.742 4	0.773 6	0.540 1	0.672 1	0.742 5	0.788 4
	FMI	0.782 5	0.841 4	0.745 4	0.774 6	0.552 6	0.673 8	0.753 3	0.790 2
	ARI	0.715 1	0.812 5	0.689 7	0.729 1	0.433 1	0.606 5	0.704 1	0.752 4
	Average	0.781 1	0.848 6	0.757 6	0.782 6	0.529 7	0.685 9	0.784 1	0.800 0
(0.4,0.4)	Purity	0.631 7	0.784 7	0.917 8	0.832 9	0.634 6	0.858 4	0.818 7	0.949 0
	F-score	0.500 2	0.784 7	0.658 9	0.636 8	0.516 3	0.563 9	0.595 2	0.928 1
	FMI	0.636 3	0.802 6	0.680 0	0.641 7	0.517 2	0.583 6	0.597 1	0.928 4
	ARI	0.550 1	0.735 1	0.614 0	0.579 9	0.421 8	0.507 4	0.528 1	0.915 3
	Average	0.579 6	0.776 8	0.717 7	0.672 8	0.522 5	0.628 3	0.634 8	0.930 2

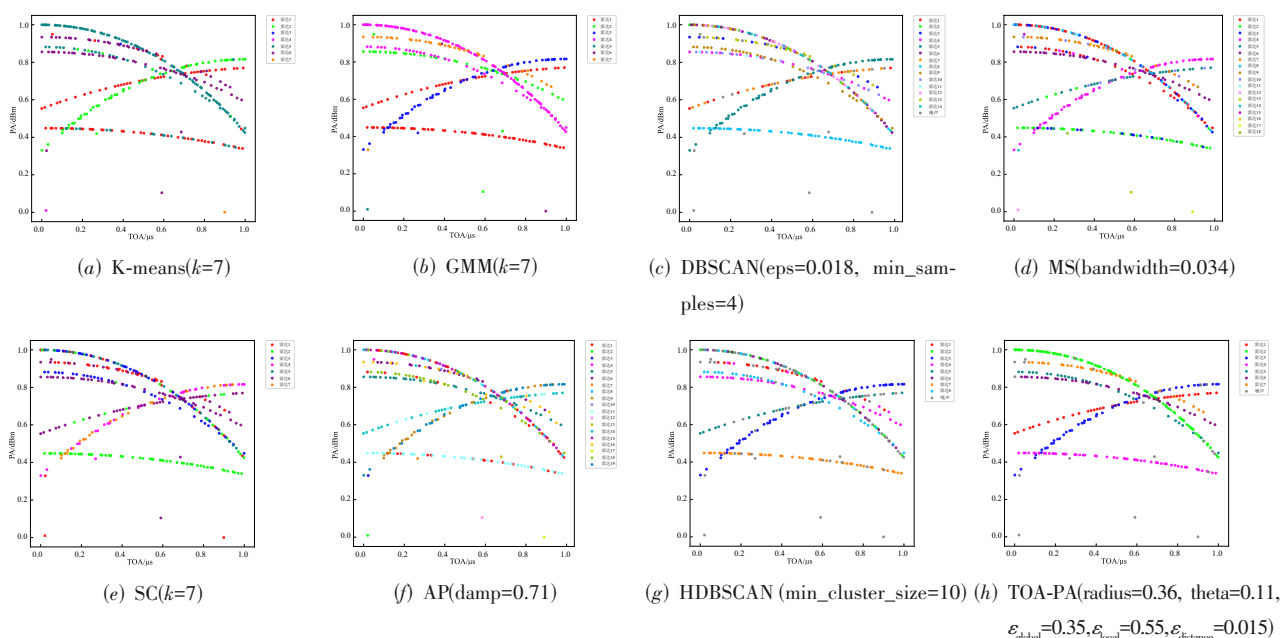


图 5 8 种聚类算法在场景 4 (0.4,0.4) 上的可视化结果

随着噪声脉冲率从 10% 增加到 50%, 各聚类算法的 FMI 值的变化情况; 而每一列则展示了在相同噪声脉冲率变化的不同场景中, 漏脉冲率分别为 10%、20%、30%、40% 和 50% 时, 各聚类算法的 FMI 表现。

在同一漏脉冲率水平下, 本文方法在场景 1~4 中相对于其他 7 种聚类算法展现了更高的 FMI 值。观察发现, DBSCAN 和 HDBSCAN 算法表现出一定的鲁棒性, 但相较于 GMM 算法略显不足。而在同一场景但漏脉冲率不同的条件下, 本文方法在各个漏脉冲率条件下, 均展现出优于或至少不劣于其他算法的性能表现。综合上述, 在绝大多数场景和漏脉冲率设置中, 相较于其他聚类算法, 本文方法的 FMI 值都位居前列, 在聚类性能上展现出显著优势。

为了系统地评估 8 种聚类算法的性能差异, 采用定量分析方法。在某一特定场景的不同漏脉冲率和噪声

脉冲率组合下, 将每种聚类算法的指标数值与该条件下表现最优的算法指标值进行比对, 计算其差值, 并将每种算法的差值进行累加, 得到在各个场景下对应指标的 8 种算法累计差异, 结果如图 7 所示。

图 7 展示了不同聚类算法在 4 种场景中, 基于 Purity、F-score、FMI、ARI 这 4 项指标的累计差异情况。累计差异越低表明算法性能越好, 反映了各算法在不同场景下与最佳表现的偏离程度。K-means 算法在所有指标上的累计差异均处于较高水平; AP 算法除了在 Purity 指标下有不错的表现外, 在其余 3 项指标上的累计差异均为最高; DBSCAN 算法在 4 项指标上的累计差异表现较为稳定; GMM 算法整体表现良好, 在 3 项指标中的累计差异仅次于本文方法。而本文方法在所有指标上的累计差异均最低, 且各场景贡献小。综上所述, 在 4 种场景及 4 项聚类评价指标下, 本文方法的累计差异值均低于其他 7 种对比算法。这充分表

明本文方法在不同场景下,在这些指标上都能更稳定地

接近最佳性能,具有更好的聚类性能和场景适应性.

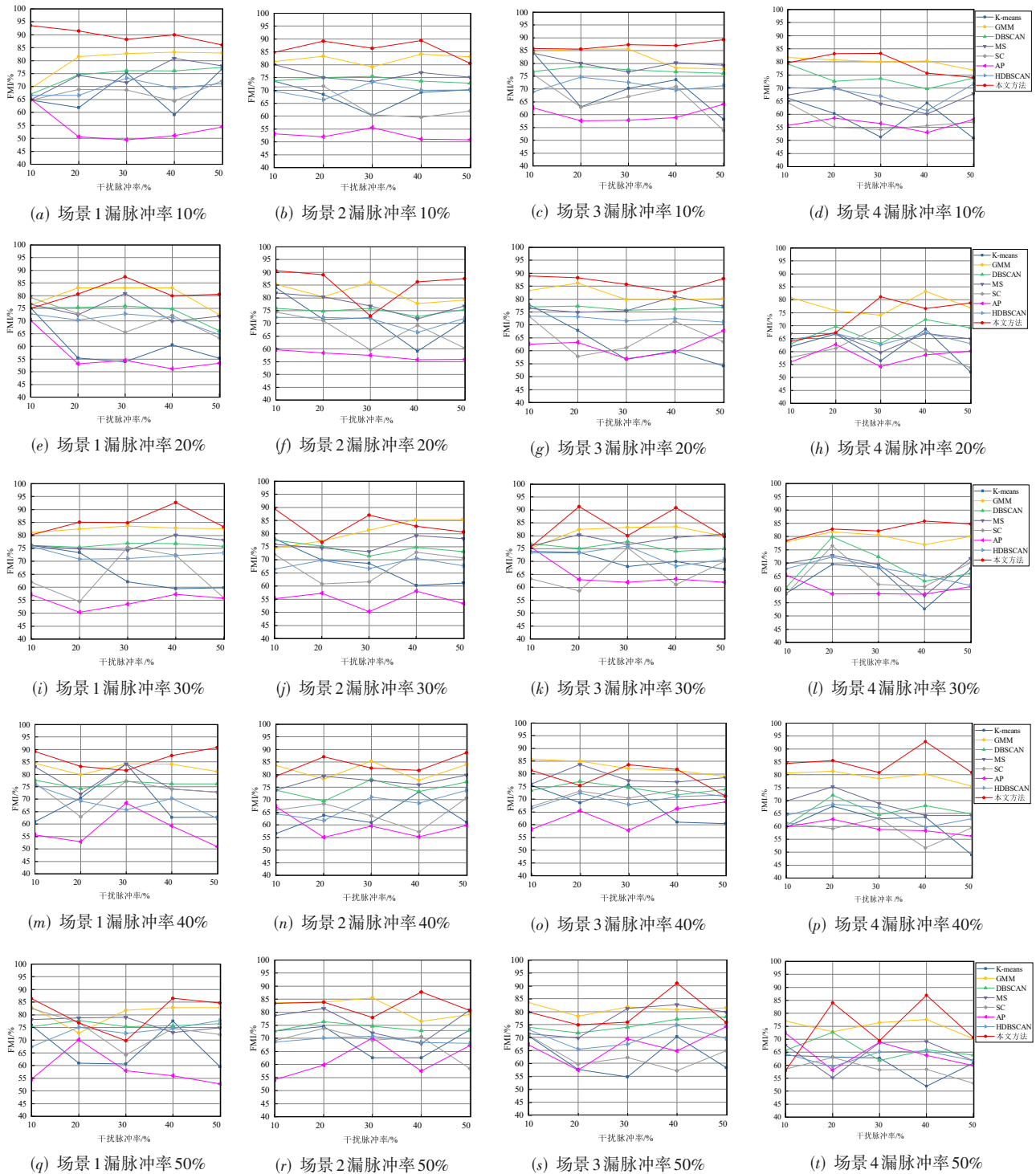


图6 4种场景下不同漏脉冲率和噪声干扰脉冲率的8种聚类算法性能

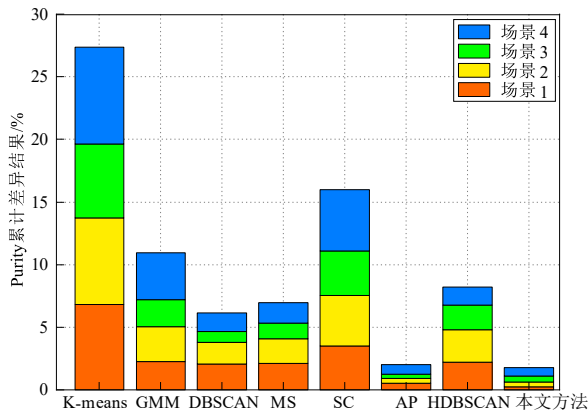
进一步地,在一般漏脉冲环境(20%~50%)下,本文算法的性能相对7种对比算法性能更优.主要原因在于7种对比聚类算法的设计未考虑到脉冲数据的时序关系,而是依据空间距离、空间分布特性或者数据密度等方式完

成聚类,性能变化相对平缓.而本文算法利用局部几何特征相似性完成聚类过程,一般漏脉冲环境下PA包络的主骨干更加清晰,有助于聚类算法更准确地捕捉PA包络特征.同时,本文方法设计了路径匹配算法的容错机

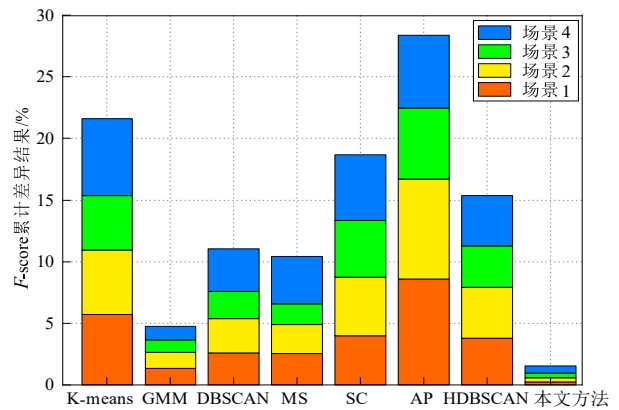
制,能够在一定程度上弥补少量漏脉冲带来的影响。

下面,给出8种算法的时间复杂度分析.K-means的时间复杂度为 $O(tkn)$;DBSCAN在最坏情况下的时间复杂度为 $O(n^2)$,但在使用索引结构时可以降低至 $O(n\log n)$;MS的时间复杂度为 $O(tn^2)$;AP的时间复杂度

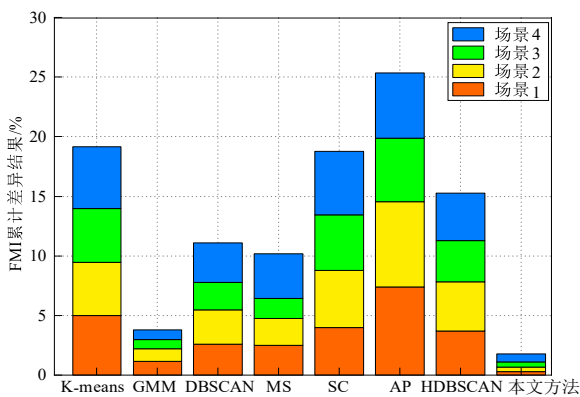
为 $O(n^2)$;GMM的时间复杂度为 $O(tkn)$;SC的时间复杂度为 $O(n^3)$;HDBSCAN的时间复杂度为 $O(n^2)$;本文算法在最坏情况下的时间复杂度为 $O(n^2)$,但在使用索引结构时可以降低至 $O(n\log n)$.其中, t 是迭代次数, n 是样本数, k 是簇的数量。



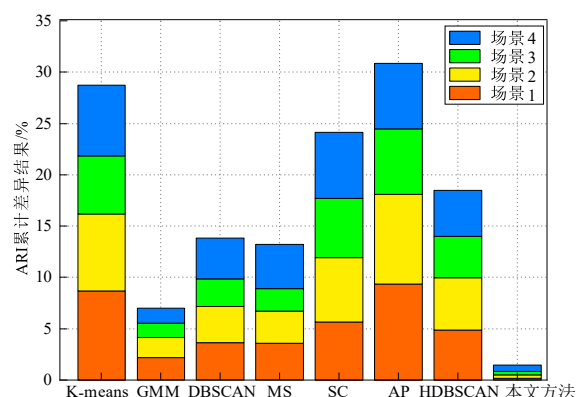
(a) 4个场景Purity指标的累计差异



(b) 4个场景F-score指标的累计差异



(c) 4个场景FMI指标的累计差异



(d) 4个场景ARI指标的累计差异

图7 不同聚类算法在4个场景4项性能指标的累计差异

4.4 主分选

在完成聚类结果融合后进行TOA的主分选.具体而言,通过构建到达时间差值(Difference of Time Of Arrival, DTOA)的一阶差分直方图,借鉴了关联脉冲对的方法对潜在PRI合批^[40],完成对复杂类型的PRI检测。

图8展示了第4个场景(0,0)下TOA一阶差分直方图,直观反映了各潜在PRI的分布情况.通过关联脉冲对方法实现了对PRI候选脉冲的合批分组,合批的结果如表5所示.表5表明了测量得到的PRI结果基本接近原始的PRI,平均相对误差不超过0.6%.在雷达4、5和7分选过程中,由于侦察系统中PDW脉冲参数的干扰不可避免,出现了一定程度的漏脉冲现象.从表6中可以看出,随漏脉冲率和噪声脉冲率的增加,分选过程中

漏脉冲现象显著上升,而平均相对误差略微增加。

4.5 消融实验

为验证本文聚类路径融合方法的有效性,本文进行了消融实验,对比了聚类路径合并前后的聚类数目结果变化.表7~表10展示了在场景1~场景4中,不同漏脉冲率和噪声脉冲率组合下,聚类路径数量的变化情况.其中,符号“—”表示合并后聚类数量与合并前无显著变化,“ $\uparrow(x)$ ”表示合并后聚类数量比合并前增加了 x 个,“ $\downarrow(x)$ ”表示聚类数量减少了 x 个。

从表7~表10的数据可以看出,在各个场景下执行聚类路径融合后,聚类数量均呈现出减少,表明本方法能够有效合并同类聚类结果.图9展示了在场景4(0.4,0.4)条件下,合并前后的聚类可视化对比效果.可

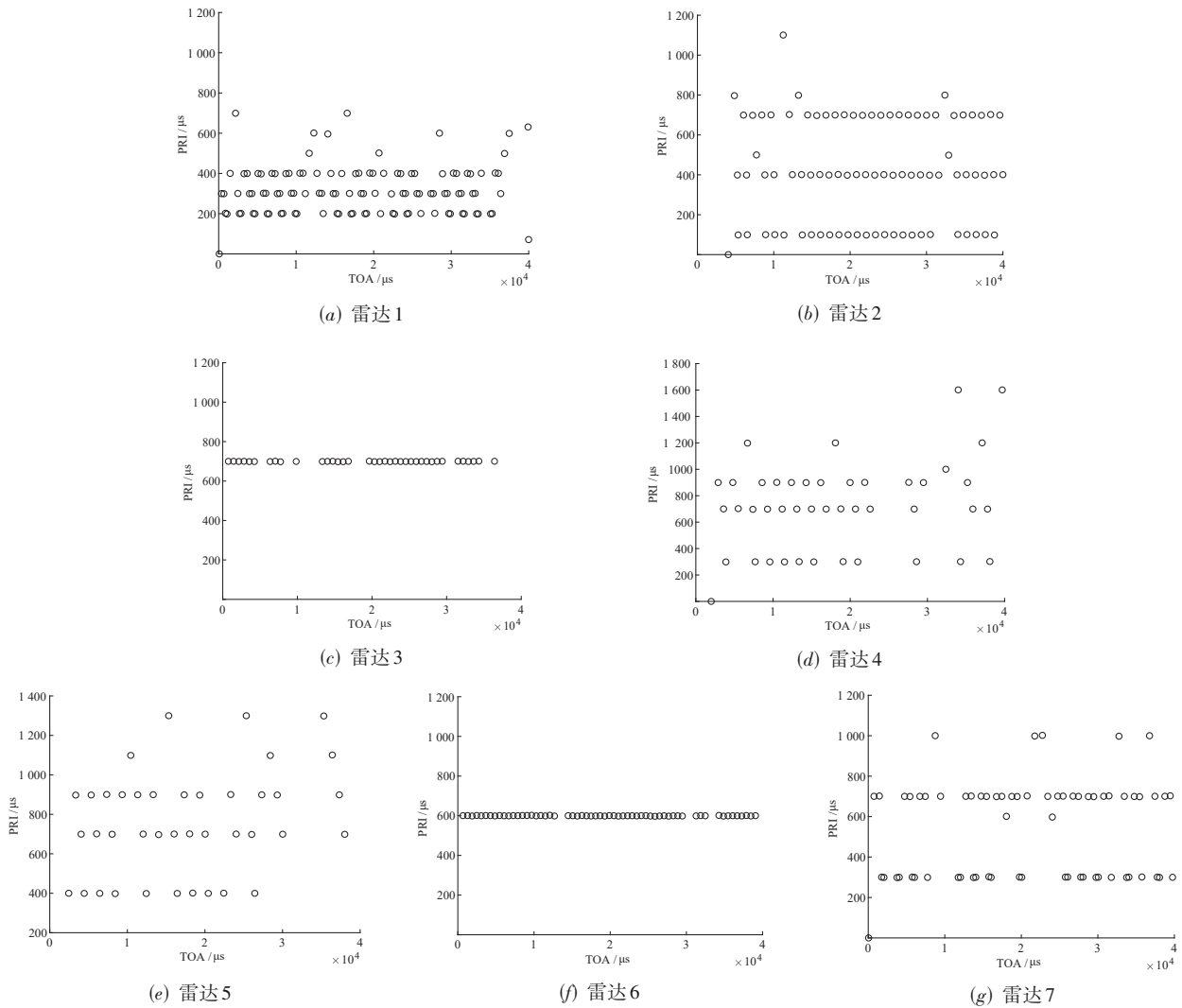


图8 PRI统计结果

表5 第4个场景(0,0)的PRI分选结果

编号	PRI	平均相对误差/%
雷达1	201.5, 301.5, 401	0.50
雷达2	99.5, 400.5, 700	0.15
雷达3	700.5	0.16
雷达4	302.5, 701.5, 902, 1 201	0.42
雷达5	401.5, 702, 901.5, 1 101	0.38
雷达6	601	0.17
雷达7	301.5, 701.5, 1 000	0.36

表6 第4个场景(0.4,0.4)的PRI分选结果

编号	PRI	平均相对误差/%
雷达1	201.5,300.5,401,699.5,802	0.39
雷达2	101.5,401.5,701,1 101.5,1 997.5	0.96
雷达3	703.5,1 402.5	0.59
雷达4	302.5,705,905	0.70
雷达5	401,701.5,901.5,1 301,1 598.5	0.31
雷达6	602,1 200,1 800.5	0.33
雷达7	301,700.5,1 001.5,1 399.5	0.20

以观察到,原本分散的聚类路径3、8、9、10最终合并为聚类路径3,聚类路径7、11、12、14也成功融合为聚类路径7.因此,本文提出的聚类路径融合方法在消除“增批”现象、减少聚类数量方面效果显著.

4.6 讨论

PA在传统雷达信号分选任务中被认为是一个易受干扰且稳定性较差的参数,其变化受到传播路径损耗、

目标 RCS起伏、多径效应等多种因素影响,但在特定作战场景和雷达工作模式下,PA的变化趋势具有较强的物理可解释性和可观测性.如在空对空交战环境中,接收平台通常位于高空、远距离探测目标,空间环境相对开阔,电磁波传播路径较为直接,缺少地面或建筑物引起的反射和散射.所以,在此类环境下,PA的变化主要由雷达发射器本身的特性决定,包括天线方向图、扫描

表 7 场景 1 合并前后的聚类数目增减情况

漏脉冲率和噪声脉冲率	0.1	0.2	0.3	0.4	0.5
0.1	↓(2)	↓(3)	↓(2)	↓(9)	↓(10)
0.2	↓(8)	↓(6)	↓(6)	↓(10)	↓(5)
0.3	↓(10)	↓(4)	↓(2)	↓(3)	↓(7)
0.4	↓(6)	↓(5)	↓(4)	↓(7)	↓(6)
0.5	↓(4)	↓(3)	↓(5)	↓(6)	↓(5)

表 8 场景 2 合并前后的聚类数目增减情况

漏脉冲率和噪声脉冲率	0.1	0.2	0.3	0.4	0.5
0.1	↓(8)	↓(4)	↓(11)	↓(6)	↓(5)
0.2	↓(7)	↓(6)	↓(6)	↓(10)	↓(13)
0.3	↓(5)	↓(3)	↓(6)	↓(9)	↓(5)
0.4	↓(4)	↓(4)	↓(7)	↓(5)	↓(5)
0.5	↓(5)	↓(5)	↓(2)	↓(7)	—

表 9 场景 3 合并前后的聚类数目增减情况

漏脉冲率和噪声脉冲率	0.1	0.2	0.3	0.4	0.5
0.1	↓(3)	↓(6)	↓(2)	↓(10)	↓(11)
0.2	↓(3)	↓(5)	↓(7)	↓(3)	↓(3)
0.3	↓(6)	↓(7)	↓(7)	↓(6)	↓(7)
0.4	↓(8)	↓(7)	↓(5)	↓(5)	↓(9)
0.5	↓(5)	↓(4)	↓(4)	↓(4)	↓(7)

表 10 场景 4 合并前后的聚类数目增减情况

漏脉冲率和噪声脉冲率	0.1	0.2	0.3	0.4	0.5
0.1	↓(11)	↓(9)	↓(12)	↓(9)	↓(5)
0.2	↓(9)	↓(11)	↓(10)	↓(15)	↓(11)
0.3	↓(6)	↓(8)	↓(6)	↓(7)	↓(7)
0.4	↓(6)	↓(8)	↓(5)	↓(7)	↓(6)
0.5	↓(3)	↓(3)	↓(1)	↓(3)	↓(3)

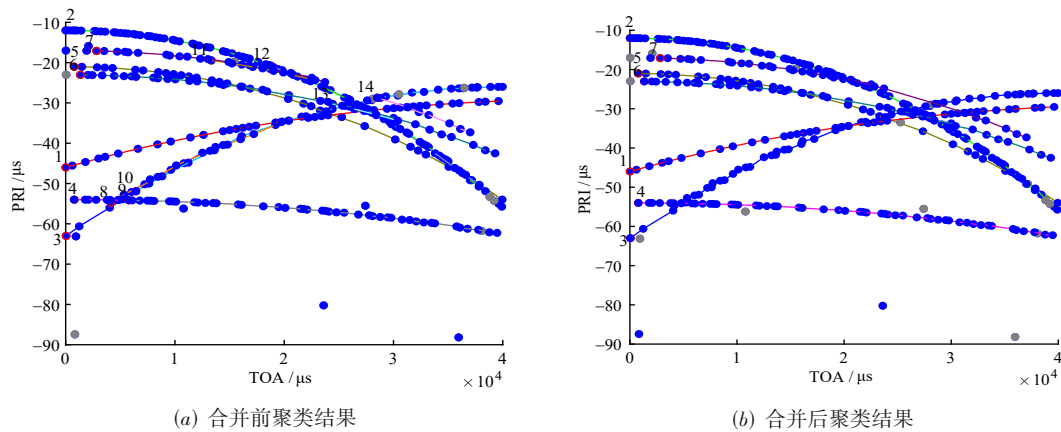


图 9 合并前后聚类结果

方式、波束指向等^[41]. 以及,在机械扫描雷达或TWS模式下,PA随时间的变化呈现出周期性或准周期性的规律,这种变化与雷达主瓣增益随扫描角度变化的过程

密切相关,具备一定的重复性和可建模性^[42]. 此外,现代电子侦察系统普遍配备自动增益控制机制,能够在一定程度上抑制PA测量中的异常波动. 通过对接收信

号进行归一化处理、滑动窗口滤波等预处理手段,也能进一步提升PA数据的可用性与稳定性。

4.7 局限性

(1)聚类算法超参数优化问题.当前方法中所采用的聚类算法仍存在较多超参数,这在一定程度上增加了模型调优的复杂度.未来应考虑进一步优化聚类算法,减少过多超参数的依赖,或探索自适应超参数调整技术,以实现参数的自动调节和更高的鲁棒性。

(2)提升雷达辐射源的参数挖掘与分析.在实际场景中,各个场景单个节拍的漏脉冲、干扰噪声和脉冲数量少时,严重影响了辐射源主分选环节,可以进一步通过相关融合算法来融合各个节拍结果,如基于DS证据理论的分选节拍融合,从而提升整体分选的准确性和稳定性。

5 结论

本文聚焦于复杂电磁环境下未知雷达辐射源信号的分选问题,提出一种基于TOA-PA聚类的新方法.首先,本文分析了雷达不同工作模式下PA特征的时序变化规律,并据此设计了TOA-PA聚类算法,该算法通过识别局部几何相似性显著的脉冲点,实现了对复杂脉冲信号模式的有效分离.接着,为了应对复杂环境中PDW数据严重混叠、交错的问题,引入了聚类路径融合算法,通过时空相似度量将相似的聚类路径进行合并,显著减少了“增批”现象并提升了整体聚类结果的一致性和连续性,从而构建了具有物理可解释性的聚类路径.最后,利用TOA一阶差分直方图进行了PRI主分选,完成了对PRI的检测.实验结果显示,在10%~50%不同组合的漏脉冲率和噪声脉冲率的情况下,本文方法在Purity、F-score、FMI和ARI的4项评价指标累计差异上均优于7种聚类算法,体现了其出色的鲁棒性和准确性,有效解决了传统聚类方法在复杂脉冲模式分离中的局限性.因此,本文通过PA特征及其变化规律的深入挖掘,可以有效提高对交织脉冲流的解析能力,为复杂条件下未知雷达信号的精准分选提供了新思路。

参考文献

- [1] GUO Q, HUANG S, LIU L C, et al. A radar signal deinterleaving method based on enhanced sparse subspace clustering[J]. IEEE Transactions on Aerospace and Electronic Systems, 2025, 61(2): 2956-2972.
- [2] MOTTIER M, CHARDON G, PASCAL F. Deinterleaving RADAR emitters with optimal transport distances[J]. IEEE Transactions on Aerospace and Electronic Systems, 2024, 60(3): 3639-3651.
- [3] ROGERS J A V. ESM processor system for high pulse density radar environment[J]. IEE Proceedings F (Communications, Radar and Signal Processing), 1985, 132(7): 621-625.
- [4] MARDIA H K. New techniques for the deinterleaving of repetitive sequences[J]. IEE Proceedings F (Radar and Signal Processing), 1989, 136(4): 149.
- [5] MILOJEVIĆ D J, POPOVIĆ B M. Improved algorithm for the deinterleaving of radar pulses[J]. IEE Proceedings F (Radar and Signal Processing), 1992, 139(1): 98.
- [6] NISHIGUCHI K, KOBAYASHI M. Improved algorithm for estimating pulse repetition intervals[J]. IEEE Transactions on Aerospace and Electronic Systems, 2000, 36(2): 407-421.
- [7] 赵仁健,熊平,倪明.大脉冲重复周期调幅信号的压缩平面变换技术[J].四川大学学报(自然科学版),1997,34(2): 172-176.
ZHAO R J, XIONG P, NI M. The compression plane transformation technique of amplitude modulation signal with large pulse repeat period[J]. Journal of Sichuan University (Natural Science Edition), 1997, 34(2): 172-176. (in Chinese)
- [8] ZHANG C J, LIU Y C, SI W J. Synthetic algorithm for deinterleaving radar signals in a complex environment[J]. IET Radar, Sonar & Navigation, 2020, 14(12): 1918-1928.
- [9] 谢敏,赵闯,胡德秀,等.联合DTOA一阶差分曲线和相关分析的参差信号分选方法[J].电子学报,2023,51(11): 3271-3281.
XIE M, ZHAO C, HU D X, et al. Signal sorting method for jagged PRI radar based on first-order difference curve of DTOA and correlation analysis[J]. Acta Electronica Sinica, 2023, 51(11): 3271-3281. (in Chinese)
- [10] CHENG W H, ZHANG Q Y, DONG J M, et al. An enhanced algorithm for deinterleaving mixed radar signals[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(6): 3927-3940.
- [11] XIE M, HUANG J, ZHAO C, et al. Radar signal deinterleaving based on hidden Markov chains and residual fence networks[J]. IEEE Transactions on Aerospace and Electronic Systems, 2025, 61(3): 6011-6025.
- [12] WANG C, SUN L T, LIU Z M, et al. A radar signal deinterleaving method based on semantic segmentation with neural network[J]. IEEE Transactions on Signal Processing, 2022, 70: 5806-5821.
- [13] 隋金坪,刘振,刘丽,等.雷达辐射源信号分选研究进展[J].雷达学报,2022,11(3): 418-433.
SUI J P, LIU Z, LIU L, et al. Progress in radar emitter signal deinterleaving[J]. Journal of Radars, 2022, 11(3): 418-

433. (in Chinese)
- [14] QI L G, CHEN H Z, GUO Q, et al. GLS: A hybrid deep learning model for radar emitter signal sorting[J]. *Digital Signal Processing*, 2025, 161: 105117.
- [15] PINSOLLE J, GOUDET O, ENDERLI C, et al. Deinterleaving of discrete renewal process mixtures with application to electronic support measures[J]. *IEEE Transactions on Signal Processing*, 2024, 72: 4983-4991.
- [16] CHEN T, QIU B C, LI J X, et al. An end-to-end radar pulse deinterleaving structure based on point cloud mapping[J]. *Digital Signal Processing*, 2024, 155: 104773.
- [17] DUTT R, BALORIA A, PRASAD V R C, et al. Discrete wavelet transform based unsupervised underdetermined blind source separation methodology for radar pulse deinterleaving using antenna scan pattern[J]. *IET Radar, Sonar & Navigation*, 2019, 13(8): 1350-1358.
- [18] AYAZGOK S, ERDEM C, OZTURK M T, et al. Automatic antenna scan type classification for next-generation electronic warfare receivers[J]. *IET Radar, Sonar & Navigation*, 2018, 12(4): 466-474.
- [19] YANG W, LI C, XU Z, et al. Envelope extraction algorithm for amplitude sorting of radar emitter signals[C]//2019 IEEE 2nd International Conference on Computer and Communication Engineering Technology (CCET). Piscataway: IEEE, 2019: 223-228.
- [20] CHEN T, GUO X Q, LI J X. Radar signal sorting method with mimetic image mapping based on antenna scan pattern via SOLOv2 network[J]. *Remote Sensing*, 2024, 16(24): 4639.
- [21] WANG C, WANG Y, LI X Q, et al. A deinterleaving method for mechanical-scanning radar signals based on deep learning[C]//2022 7th International Conference on Intelligent Computing and Signal Processing. Piscataway: IEEE, 2022: 138-143.
- [22] BARSHAN B, ERAVCI B. Automatic radar antenna scan type recognition in electronic warfare[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2012, 48(4): 2908-2931.
- [23] GUO Q, TENG L, WU X L, et al. Deinterleaving of radar pulse based on implicit feature[J]. *Journal of Systems Engineering and Electronics*, 2023, 34(6): 1537-1549.
- [24] ZHANG Y J, HUO W B, ZHANG C, et al. Radar operation mode recognition via multifeature residual-and-shrinkage ConvNet[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, 16: 6073-6084.
- [25] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]//Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. New York: ACM, 1996: 226-231.
- [26] GALDINO S, DIAS J. Interval-valued data ward's hierarchical agglomerative clustering method: Comparison of three representative merge points[C]//2021 International Conference on Engineering and Emerging Technologies. Piscataway: IEEE, 2022: 1-6.
- [27] GUO Q, TENG L, QI L G, et al. A novel radar signals sorting method-based trajectory features[J]. *IEEE Access*, 2019, 7: 171235-171245.
- [28] ZHENG H P, XIE K, ZHU Y S, et al. An reconstruction bidirectional recurrent neural network-based deinterleaving method for known radar signals in open-set scenarios[J]. *IET Radar, Sonar & Navigation*, 2024, 18(6): 965-981.
- [29] FUYUE L I U, JINXIN L I. End-to-end radar signal sorting based on deep segmentation[J]. *Systems Engineering and Electronics*, 2023, 45(5): 1351-1358.
- [30] ARTHUR D, VASSILVITSKII S. How slow is the k-means method?[C]//Proceedings of the 22th Annual Symposium on Computational Geometry. Sedona, Arizona: ACM, 2006: 144-153.
- [31] 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.
- [32] FREY B J, DUECK D. Clustering by passing messages between data points[J]. *Science*, 2007, 315(5814): 972-976.
- [33] DEMPSTER A P, LAIRD N M, RUBIN D B. Maximum likelihood from incomplete data via the EM algorithm[J]. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1977, 39(1): 1-38.
- [34] SHI J B, MALIK J. Normalized cuts and image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905.
- [35] CAMPELLO R J G B, MOULAVI D, SANDER J. Density-based clustering based on hierarchical density estimates[C]//Advances in Knowledge Discovery and Data Mining. Berlin: Springer, 2013: 160-172.
- [36] KARYPIS G, KUMAR V, STEINBACH M. A comparison of document clustering techniques[C]//TextMining Workshop at KDD2000. New York: ACM, 2000: 428-439.
- [37] ZHAO Y, KARYPIS G. Evaluation of hierarchical clustering algorithms for document datasets[C]//Proceedings

of the 11th International Conference on Information and Knowledge Management. New York: ACM, 2002: 515-524.

- [38] FOWLKES E B, MALLOWS C L. A method for comparing two hierarchical clusterings[J]. Journal of the American Statistical Association, 1983, 78(383): 553-569.
- [39] HUBERT L, ARABIE P. Comparing partitions[J]. Journal of Classification, 1985, 2(1): 193-218.
- [40] 左峰, 曹兰英, 杨健. 一种基于关联脉冲对的动态直方图分选算法[J]. 现代防御技术, 2023, 51(1): 59-66.

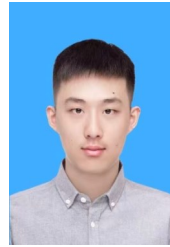
ZUO F, CAO L Y, YANG J. A dynamic histogram deinterleaving algorithm based on correlated pulse pairs[J]. Modern Defence Technology, 2023, 51(1): 59-66. (in Chinese)

- [41] GENÇOL K, KARA A, AT N. Improvements on deinterleaving of radar pulses in dynamically varying signal environments[J]. Digital Signal Processing, 2017, 69: 86-93.
- [42] KOCAMIŞ M B, ORDUYİLMAZ A, TAŞCIÖĞLU S. Object detection based deinterleaving of radar signals using deep learning for cognitive EW[J]. Signal, Image and Video Processing, 2024, 18(11): 7789-7800.

作者简介



林进健 男, 1993年1月生, 广东阳江人. 中山大学博士研究生. 主要研究方向为雷达信号处理、不确定信息融合.
E-mail: linjj77@mail2.sysu.edu.cn



林梓涵 男, 2003年1月生, 安徽合肥人. 中山大学硕士研究生. 主要研究方向为电子对抗、雷达信号分选.
E-mail: linzh266@mail2.sysu.edu.cn



黄明军 男, 1990年10月生, 安徽舒城人. 中国电子科技集团公司第二十九研究所高级工程师. 主要研究方向为作战支持、雷达信号与信息处理等.
E-mail: onion_hmj@163.com



谢恺 男, 1980年3月生, 安徽巢湖人. 中山大学教授、博士生导师. 主要研究方向为雷达信号处理、电子对抗信号处理等.
E-mail: xiek8@mail.sysu.edu.cn



孙慧博 男, 2002年4月生, 河南新乡人. 中山大学硕士研究生. 主要研究方向为电子对抗、雷达信号分选.
E-mail: sunhb6@mail2.sysu.edu.cn