中国科学:信息科学 2026 年 第 56 卷 第 11 期: 2895–2916
SCIENTIA SINICA Informationis
Agentic AI 赋能的 6G 融合通信专题 • 论文

面向 Agentic AI-RAN 的无人机集群跨模态协同认知方法

舒慧桢[1], 姚启航[1], 徐辉[1], 李宇萌[1], 崔原豪[2*]

1. 北京航空航天大学电子信息工程学院, 北京 100191
2. 北京邮电大学信息与通信工程学院, 北京 100876
* 通信作者. E-mail: yuanhao.cui@bupt.edu.cn

收稿日期: 2026–03–31; 修回日期: 2026–08–13; 接受日期: 2026–09–22; 网络出版日期: 2026–11–09

国家自然科学基金 (批准号: 52302398) 资助项目

摘要 智能体化人工智能无线接入网 (Agentic AI-radio access network, Agentic AI-RAN) 作为 6G 时代的新兴网络架构概念, 其安全稳定面临着非合作无人机集群侵扰的严峻挑战, 关键在于如何提升对非合作集群行为状态的准确认知. 当前认知技术分别从物理域感知运动行为特征或从通信域获取交互行为特征, 但仅靠单一模态信息难以有效实现对非合作集群 “物理运动 – 通信交互” 行为的协同精准认知, 无法构建同一无人机下融合物理域与通信域特征的完整个体画像. 对此, 本文提出一种基于物理域和通信域的跨模态协同认知方法 (physical-communication cross-modal collaborative cognition, P-C CMCC). 在双域特征层, 分别提取物理域与通信域的异构特征, 建立双域模态子图实现非线性特征表示; 在双域节点层, 采用双模态异质图注意力网络挖掘各域内的节点结构性协同关系, 通过注意力机制将双域特征映射至公共交互空间, 并构建双域相似度矩阵实现双域节点关联; 在集群全局层, 通过软匹配和硬匹配两阶段求解, 获得全局最优的双域匹配结果, 进而构建双域协同认知图, 实现集群 “物理运动 – 通信交互” 行为的协同认知; 最后, 引入双域节点层与集群全局层的总损失函数, 前馈至双域节点层, 实现网络训练参数自智更新优化. 实验结果表明, 在不同通信信噪比、集群规模大小和集群规模突变场景下, 所提方法在双域匹配准确率和双域匹配均方根误差上优于基准和消融对比结果, 有效验证了本方法的性能.

关键词 非合作无人机集群, 跨模态协同认知, 物理运动 – 通信交互, 双域节点关联, 两阶段求解, 自智优化

1 引言

作为智能化自治网络架构, Agentic AI-RAN(Agentic AI-radio access network) 对终端行为状态的精准认知, 是实现其高效资源调度与稳定协同管控的重要前提 [1,2]. 对此, 当前 Agentic AI-RAN 研究应用主要面向协议透明、安全可信的合作终端, 分别从物理域依托雷达等设备探测运动轨迹与空间态势, 或从通信域通过信令交互实现状态识别与行为监测. 然而, 在真实复杂的部署场景中, 合作终端与非合作目标往往并存, 且后者常以对抗形

引用格式: 舒慧桢, 姚启航, 徐辉, 等. 面向 Agentic AI-RAN 的无人机集群跨模态协同认知方法. 中国科学: 信息科学, 2026, 56: 2895–2916, doi: 10.1360/SSI-2026-0122
Shu H Z, Yao Q H, Xu H, et al. Cross-modal collaborative cognition method for UAV swarm in Agentic AI-RAN. Sci Sin Inform, 2026, 56: 2895–2916, doi: 10.1360/SSI-2026-0122

© 2026《中国科学》杂志社

式侵扰网络. 此类非合作目标的恶意干扰极易造成智能体认知偏差与决策异常, 通过交互互联扩散蔓延, 引发网络协同失效、决策紊乱与全域可信性污染严重威胁 Agentic AI-RAN 的稳定运行. 因此, 为适配真实复杂环境下的网络稳定运行需求, Agentic AI-RAN 亟需强化面向非合作目标的行为状态认知水平, 提升网络抵御非合作目标侵扰的能力 [3~5].

非合作无人机集群作为典型恶意网络侵扰源之一, 具有通信协议未知和运动状态动态时变的特点, 使得现有 Agentic AI-RAN 中面向合作终端的物理域与通信域感知手段性能下降甚至失效. 具体而言, 通信域感知手段通过捕获无先验信令的非合作回波推算其通信交互行为状态, 存在静默目标探测盲区, 且难以将通信发射源与真实的无人机物理实体精准绑定; 物理域感知技术虽能探测目标的运动轨迹, 但无法获知集群内部的通信交互逻辑, 难以进一步支撑对集群关键节点识别、意图研判等深层认知 [6~9]. 因此, 仅依靠单一感知手段难以实现对非合作集群行为状态的精准认知, 亟需通过双域跨模态融合感知手段突破单域限制, 实现对非合作集群 “物理运动 – 通信交互” 行为的协同精准认知, 为 Agentic AI-RAN 抵御非合作目标侵扰提供关键支撑.

当前基于双域跨模态融合实现非合作集群 “物理运动 – 通信交互” 行为协同精准认知存在三大难点. 一是双域之间存在特征、节点和集群的三重尺度级联强耦合特性, 任一尺度的认知偏差将不断传播放大, 最终导致全局认知失效 [10, 11]; 现有方法多为特征与决策融合的两层级求解架构, 难以实现三重尺度下的深度建模与关系表征. 二是双域之间非线性关系映射复杂, 表现为双域行为特征异构、节点关联规则模糊和集群双域匹配认知冲突 [12]; 当前方法多面向关联规则明确的场景, 或难以在三层尺度下实现集群双域最优匹配认知, 易陷入局部最优. 三是非合作无人机集群通信协议未知且运动状态动态时变 [13], 无法依赖预设网络模板, 传统离线训练方法难以实现, 对网络在线校准的自智优化能力提出更高要求 [14].

综上, 为提升对非合作无人机集群的行为状态认知水平, 本文提出一种基于物理域和通信域的跨模态协同认知方法 (physical-communication cross-modal collaborative cognition, P-C CMCC). 在双域特征层, 分别提取物理域与通信域的异构特征, 并构建双域模态子图实现非线性特征表示; 在双域节点层, 采用双模态异质图注意力网络挖掘各域内的节点结构性协同关系, 通过注意力机制将双域异构特征映射至公共交互空间, 并据此构建双域相似度矩阵实现双域节点关联; 在集群全局层, 通过软匹配和硬匹配的两阶段求解, 获取全局最优的双域匹配结果, 进而构建双域协同认知图实现集群 “物理运动 – 通信交互” 行为的协同精准认知; 最后, 引入双域节点层与集群全局层的总损失函数, 前馈至双域节点层, 使其网络训练参数受到跨层级联合代价的约束进行自智更新优化. 本文主要贡献如下.

(1) 突破传统特征与决策融合的两层级求解架构局限, 基于双域三重尺度级联强耦合特性, 构建 “双域特征 – 双域节点 – 集群全局” 的跨模态协同认知求解框架, 从底层双域特征表示、中层双域节点关联到顶层集群双域匹配认知, 实现建模逐级深化.

(2) 针对双域非线性关系映射复杂问题, 在双域节点层利用双模态异质图注意力网络挖掘节点协同关系, 通过跨模态注意力机制将双域异构特征映射至公共语义空间并度量相似度, 在集群全局层引入软匹配和硬匹配的两阶段求解, 实现三层尺度下集群双域最优匹配认知, 避免陷入局部最优.

(3) 构建具备在线自智优化能力的端到端训练, 通过总损失函数使误差沿集群全局层向双域节点层反向传播, 从而实现跨层级的协同闭环优化.

(4) 实验结果表明, 在不同通信信噪比、集群规模大小和集群规模突变的验证场景中, 所提方法在双域匹配准确率和双域匹配均方根误差指标上优于基准和消融对比结果, 有效验证了本方法的性能.

本文后续结构安排如下: 第二节为相关工作综述; 第三节进行问题描述与方法概述; 第四节模型设置与算法设计; 第五节展示实验结果; 第六节对研究工作进行总结, 并展望未来研究方向.

2　相关工作

2.1　Agentic AI-RAN

Agentic AI-RAN 作为 6G 时代的新兴网络架构概念, 通过将基础设施节点转化为具备感知、通信、计算和控制等能力的智能体, 增强无线接入网络的自主性、智能化和自适应性. 在 Agentic AI-RAN 架构中, 智能体不再限于传统的无线接入和信号传输功能, 还具备数据采集、优化决策、协同控制、网络自愈和资源管理等能力, 从而使无线接入网络更加灵活、高效和可靠. 因此, 面向 Agentic AI-RAN 架构的智能化转型, 是未来 6G 网络时代的关键支撑 [15, 16].

Agentic AI-RAN 智慧内生特性赋能网络智能化升级的同时, 也使得其安全稳定性与运行可信性成为核心关键. 在真实复杂的部署场景中, 网络智能体具备独立感知与自主决策能力, 也使之易遭受恶意非合作目标入侵带来的感知欺骗、信号干扰等威胁, 由此引发的错误认知与决策会传导至网络资源调度、协同决策等环节, 引发网络可信性失准, 对 Agentic AI-RAN 的稳定运行带来安全隐患. Chergui 等人围绕 Agentic AI 网络智能体的认知偏差导致感知严重扭曲与执行问题, 系统提出偏差分类学、数学表达及缓解策略 [17]. Su 等人聚焦基于大模型的自主智能体在动态环境中面临的新型安全风险, 如记忆投毒、工具滥用、奖励黑客及涌现性错位等, 提出反思性风险感知智能体架构 [18]. 上述研究为保障 Agentic AI-RAN 稳定运行研究提供了重要参考, 但主要聚焦于网络内部的决策机制与认知偏差缓解, 而精准的外部威胁感知是实现有效安全响应的先决前提, 因此需从感知层面对非合作目标实现精准行为状态认知, 为后续精准防控与架构稳定运行奠定基础.

2.2　非合作无人机集群协同认知挑战

对非合作集群开展跨模态融合协同精准认知的研究, 当前仍存在多尺度级联强耦合、双域非线性关系映射复杂和网络自智优化能力不足的挑战.

针对特征、节点和集群的三重尺度耦合问题, 当前主要采用从特征到全局的两层架构实现跨模态融合. 例如 Liu 等人对可见光与红外图像进行特征级与决策级的分层融合实现无人机集群检测 [19]; Chen 等人提出多尺度自适应特征金字塔网络来解决车辆检测中的极端尺度变化问题 [20]; Liu 等人构造标签级和特征级图像 – 文本特征来提升跨模态融合准确性 [21]; Zhuang 等人分别从视觉的全局与区域特征、文本的句子和词汇特征进行多级跨模态融合 [22]. 上述方法在两层架构融合方面取得了显著进展, 但难以实现集群三重尺度问题的深度建模与关系表征.

针对双域行为特征异构、节点关联规则模糊和集群全局认知冲突的非线性关系映射复杂问题, 当前的主流方法之一是采用深度学习模型挖掘多模态间深层语义关联. Dabbabi 等人提出基于图神经网络的多模态数据融合方法来提升无人机集群的协同感知能力 [23]; Zahra 等人用动态图神经网络和 Bi-LSTM 对运动员骨架、速度和位置进行建模融合 [24]; Zhang 等人用图卷积网络和 Transformer 实现道路几何信息、历史信道状态和道路语义融合 [25]; Qin 等人利用异构图卷积网络融合图像和文本关系获取更丰富的跨模态表示 [26]. 上述工作为异构模态融合提供关键支撑, 但主要面向跨域融合规则明确场景. 基于此, 有研究采用深度学习方法与匹配决策算法结合的方法解决跨域融合规则模糊问题. Lu 等人通过时序图注意力网络捕捉多船舶关联下的局部交互和全局依赖, 采用匈牙利算法求解全局最优的一对一匹配 [27]; Hong 等人采用图注意力网络对语言和视觉模型进行融合, 通过匈牙利算法求解一对一的最优匹配 [28]; Zhu 等人利用图卷积网络融合语音信息和句法信息, 通过贪心策略求解最优匹配对应 [29]. 前述方法在解决跨域融合规则模糊问题上取得显著提升, 但主要针对局部特征与全局决策两层架构, 难以解决三层尺度下的集群双域最优匹配认知冲突, 易陷入局部最优.

针对网络自智优化问题, 非合作集群的通信协议未知且运动状态动态时变, 传统离线训练方法难以有效部署, 无需先验模型知识即可适应动态变化的在线方法优势显著. Kang 等人采用分布式有限记忆在线学习策略实现多无人机系统的状态实时辨识 [30]; Si 等人利用元学习框架实现优化模型的参数初始化, 实现对新动态频谱环境的快速适应 [31]; Yang 等人将自我监督学习集成到无线网络来提升可扩展性、适应性和泛化性 [32]; Yang 等人设计数据驱动的迁移学习方法实现非合作无人机轨迹预测与意图识别 [33]. 因此, 面向非合作集群的跨模态认知

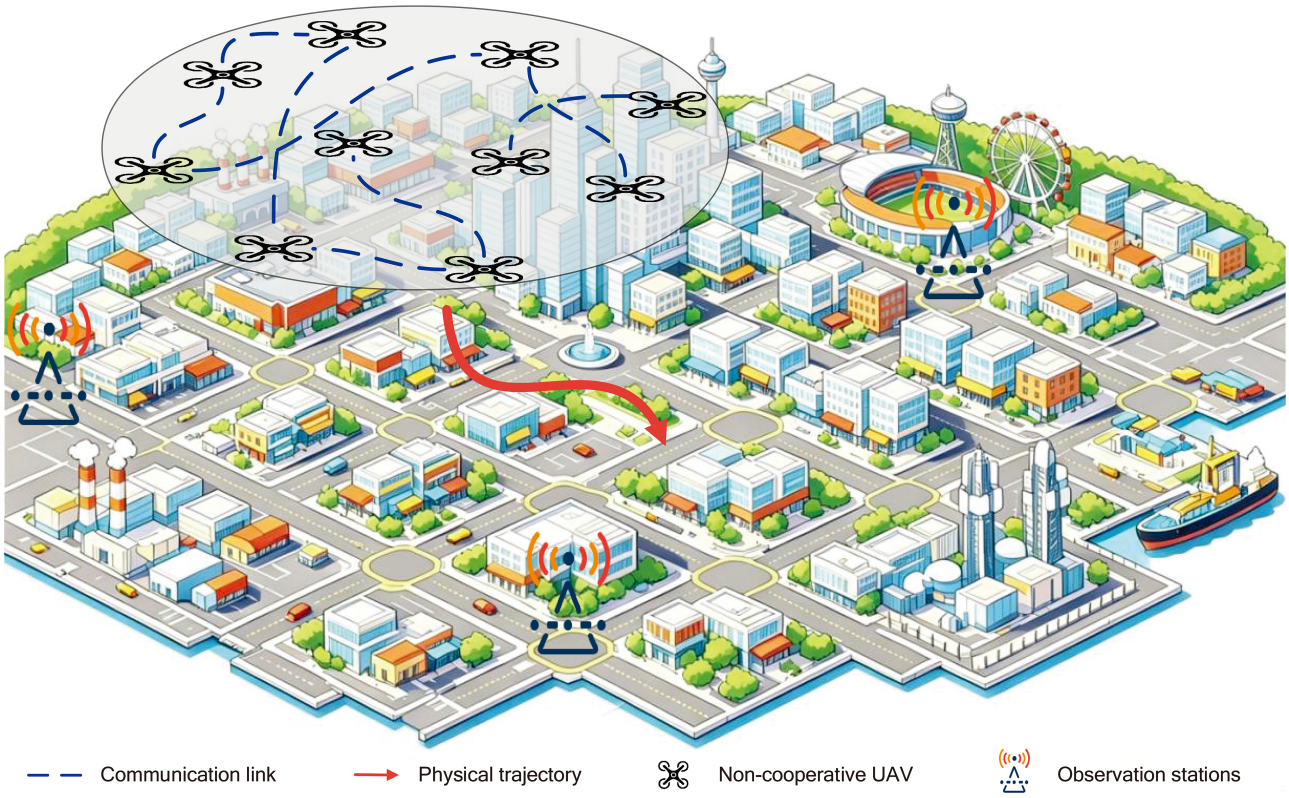

图 1　(网络版彩图) Agentic AI-RAN 中非合作无人机集群侵扰示意图.

Figure 1　(Color online) Schematic diagram of non-cooperative UAV swarm intrusion in Agentic AI-RAN.

系统需构建端到端的在线自智优化闭环, 使模型能够在无先验条件下实现跨层级参数的协同演进.

综上, 现有工作在跨模态协同认知方向积累了丰富研究成果. 但面向非合作无人机集群, 相关研究仍处于起步或碎片化阶段. 为此, 本文提出一种基于物理域和通信域的跨模态协同认知方法, 提升对非合作集群 “物理运动 – 通信交互” 行为的协同认知水平, 为 Agentic AI-RAN 抵御非合作目标侵扰、实现稳定运行提供保障.

3　问题描述与方法框架概述

3.1　问题描述

假设在 Agentic AI-RAN 中, 出现物理运动行为与通信交互行为未知的非合作 “低小慢” 无人机集群, 采用多观测站进行双域信号探测 [34,35]. 图 1 为 Agentic AI-RAN 中非合作无人机集群侵扰示意图.

3.2　P-C CMCC 框架概述

本文提出如图 2 所示的 P-C CMCC 框架, 构建 “双域特征 – 双域节点 – 集群全局” 的跨模态协同认知求解框架, 通过从底层双域特征表示、中层双域节点关联到顶层集群双域匹配认知的逐级深化建模, 进而实现非合作集群 “物理运动 – 通信交互” 行为的协同精准认知.

P-C CMCC 框架包括以下递进层级: 第一层双域特征层, 分别提取物理域与通信域的异构特征, 并构建双域模态子图实现非线性特征表示; 第二层双域节点层, 采用双模态异质图注意力网络挖掘各域内的节点结构化协同关系, 通过注意力机制将双域异构特征映射至公共交互空间, 并据此构建双域相似度矩阵实现双域节点关联; 第三层集群全局层, 通过软匹配和硬匹配的两阶段求解, 获取全局最优的双域匹配结果, 进而构建双域协同认知图实现集群 “物理运动 – 通信交互” 行为的协同精准认知; 最后, 引入双域节点层与集群全局层的总损失函数,

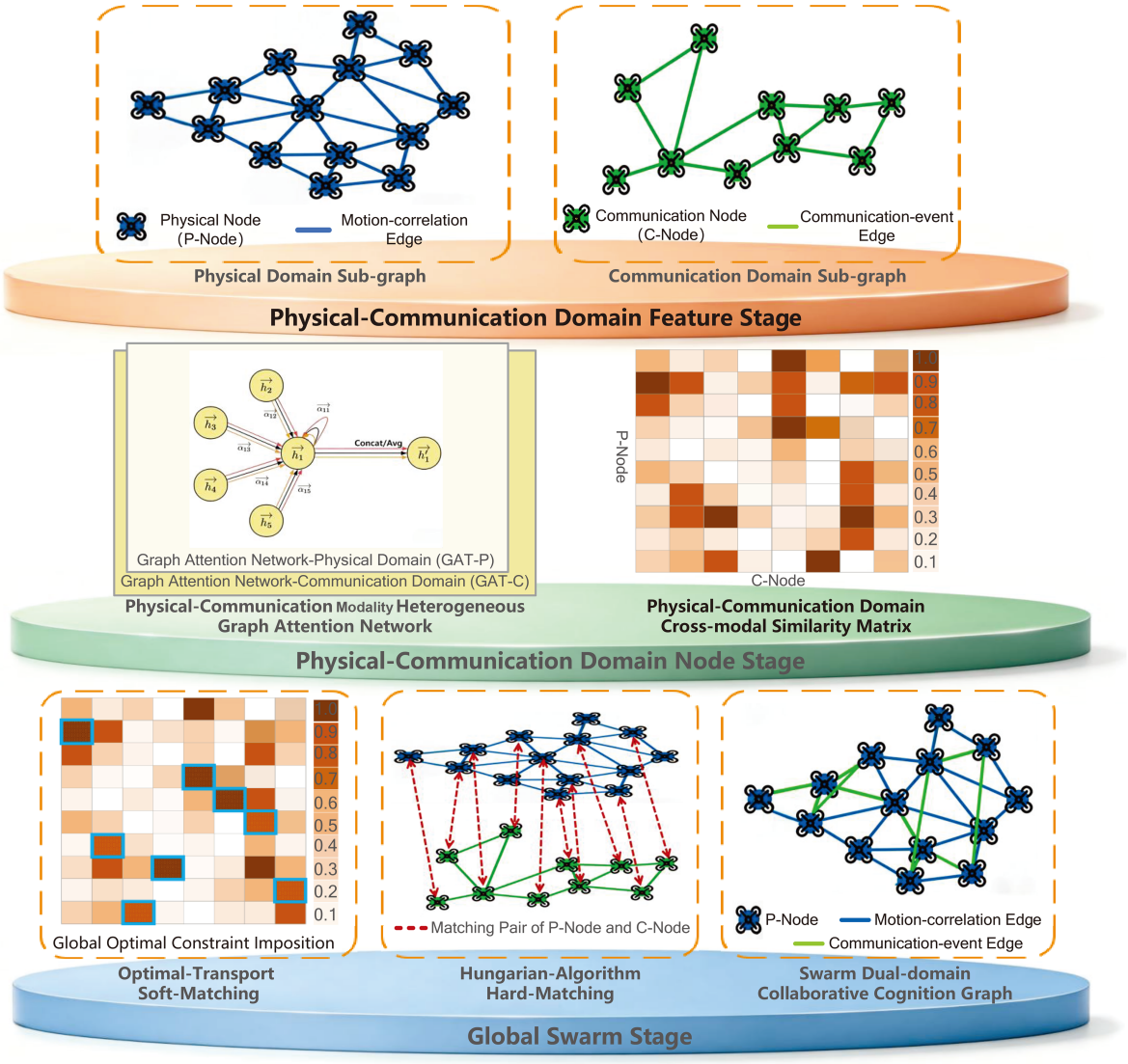

图 2　(网络版彩图) **P-C CMCC 框架示意图.**

Figure 2　(Color online) Schematic diagram of the P-C CMCC framework.

前馈至双域节点层, 使其网络训练参数受到跨层级联合代价的约束进行自智更新优化.

4　模型设置与算法设计

4.1　双域特征层: 双域特征提取与模态子图构建

4.1.1　集群物理运动行为建模与运动特征提取

定义非合作集群中第 i 架无人机 $(i=1,2,\ldots,N_{r,all})$ 在 t_r 时刻的物理状态 $\mathcal{X}_{r,i}(t_r)$ 为 [6,7]

$$\mathcal{X}_{r,i}(t_r)\equiv\left\{\mathbf{p}_{r,i}(t_r),\ \mathbf{v}_{r,i}(t_r)\right\},\tag{1}$$

其中, $\mathbf{p}_{r,i}(t_r)$ 为三维位置矢量, $\mathbf{v}_{r,i}(t_r)$ 为速度矢量, $N_{r,all}$ 为集群中的无人机总数.

设地面部署 N_s 个观测站, 定义第 s 个观测站 $(s=1,2,\ldots,N_s)$ 对第 i 架无人机的物理域观测结果 $\bar{\mathbf{z}}_{r,s,i}$ 为

$$\bar{\mathbf{z}}_{r,s,i}=h_s\left(\bar{\mathcal{X}}_{r,s,i}\right)+\mathbf{n}_{r,s,i},\tag{2}$$

其中 h_s 为观测函数, $\mathbf{n}_{r,s,i}$ 为噪声, $\bar{\mathcal{X}}_{r,s,i}$ 为无人机物理观测状态, 包括位置观测矢量 $\bar{\mathbf{p}}_{r,s,i}$ 和速度观测矢量 $\bar{\mathbf{v}}_{r,s,i}$.

地面观测站采用集中式处理架构, 通过融合中心对各观测站的物理域观测结果进行时间对齐与坐标统一, 通过高斯混合概率假设密度滤波器进行多目标跟踪 [36,37], 输出 t_r 观测时刻的集群物理运动状态估计序列

$$\mathcal{T}_r(t_r)=\{\bar{\mathbf{p}}_{r,i}(t_r),\ \bar{\mathbf{v}}_{r,i}(t_r),\ \mathbf{P}_{\mathrm{cov},r,i}(t_r),\ w_{r,i}(t_r),\ \ell_{r,i}\}_{i=1}^{N_r(t_r)}, \tag{3}$$

其中, $\bar{\mathbf{p}}_{r,i}(t_r)$ 和 $\bar{\mathbf{v}}_{r,i}(t_r)$ 分别为无人机 i 在 t_r 时刻的位置观测矢量与速度观测矢量, $\mathbf{P}_{\mathrm{cov},r,i}(t_r)$ 为估计协方差矩阵, $w_{r,i}(t_r)$ 为置信度, $\ell_{r,i}$ 为唯一航迹标识符, $N_r(t_r)$ 为该时刻确认的目标总数. 受雷达虚警、漏检等影响, 雷达探测的无人机总数 N_r 一般小于实际集群无人机总数 $N_{r,all}$.

4.1.2　物理域模态子图构建

基于集群物理运动状态估计序列, 构建物理域模态子图 $G_r=(V_r,E_r)$, 其中物理节点 $i\in V_r$ 表征无人机实体, 边 E_r 表征无人机间的时空关联关系. 具体实现步骤如下 [38,39].

(1) 构建物理节点的特征向量. 物理节点 i 在 t_r 时刻对应一个确认的无人机实体, 其特征向量 $\mathbf{h}_{r,i}(t_r)$ 为

$$\mathbf{h}_{r,i}(t_r)=[\bar{\mathbf{p}}_{r,i}(t_r),\ \bar{\mathbf{v}}_{r,i}(t_r),\ w_{r,i}(t_r)], \tag{4}$$

其中, $\bar{\mathbf{p}}_{r,i}(t_r)$、$\bar{\mathbf{v}}_{r,i}(t_r)$ 和 $w_{r,i}(t_r)$ 分别为物理节点 i 在 t_r 时刻的位置观测矢量、速度观测矢量和置信度.

(2) 时空关系建模. 通过双重准则约束, 针对物理域模态构建边集 E_r [40].

准则 1: 空间邻近性. 当物理节点 i_1 与 i_2 的空间位置满足欧氏距离约束 $\|\bar{\mathbf{p}}_{r,i_1}(t_r)-\bar{\mathbf{p}}_{r,i_2}(t_r)\|\leqslant d_{\mathrm{th}}$, 建立两个物理节点之间的空间关联边.

准则 2: 运动一致性. 当物理节点 i_1 与 i_2 的速度向量夹角满足 $\cos^{-1}\left(\frac{\bar{\mathbf{v}}_{i_1}\cdot\bar{\mathbf{v}}_{i_2}}{\|\bar{\mathbf{v}}_{i_1}\|\|\bar{\mathbf{v}}_{i_2}\|}\right)\leqslant\theta_{\mathrm{th}}$, 建立两个物理节点之间的运动关联边. 物理节点 i_1 与 i_2 之间边的权重 ρ_{r,i_1,i_2} 表示为

$$\rho_{r,i_1,i_2}=\exp\left(-\frac{\|\bar{\mathbf{p}}_{r,i_1}-\bar{\mathbf{p}}_{r,i_2}\|^2}{\sigma_{p_r}^2}-\frac{\|\bar{\mathbf{v}}_{r,i_1}-\bar{\mathbf{v}}_{r,i_2}\|^2}{\sigma_{v_r}^2}\right), \tag{5}$$

其中, σ_{p_r} 和 σ_{v_r} 为空间尺度参数和速度相似性尺度参数, $\bar{\mathbf{p}}_{r,i_1}$ 和 $\bar{\mathbf{v}}_{r,i_1}$ 为位置观测矢量和速度观测矢量.

由此, 通过构建物理域模态子图 G_r, 实现以图结构形式表征集群在每一观测时刻的物理位置与运动状态, 以及无人机间的空间邻近与运动一致性关联关系, 为后续双域节点关联提供物理域的结构化先验输入.

4.1.3　集群通信交互行为建模与通信特征提取

将无人机之间的通信交互定义为通信事件, 接收端在收到发射源数据后返回确认信号. 定义集群共有 $N_{c,all}$ 个发射源, 其中通信发射源 j_1 与接收端 j_2 在 t_c 时刻的通信事件 $\mathbf{E}_{c,j_1,j_2}(t_c)$ 为

$$\mathbf{E}_{c,j_1,j_2}(t_c)=\big\{j_1,\ j_2,\ f_{c,j_1}(t_c),\ B_{c,j_1}(t_c),\ \mathcal{M}_{c,j_1}(t_c),\ \mathbf{p}_{c,j_1}(t_c),\ \mathbf{p}_{c,j_2}(t_c)\big\}, \tag{6}$$

其中, $f_{c,j_1}(t_c)$、$B_{c,j_1}(t_c)$ 和 $\mathcal{M}_{c,j_1}(t_c)$ 分别为由发射源 j_1 决定的载波频率、信号带宽与调制方式, $\mathbf{p}_{c,j_1}(t_c)$ 和 $\mathbf{p}_{c,j_2}(t_c)$ 分别为发射源与接收端对应的空间位置.

地面观测站通过侦测集群内部通信事件形成通信域估计序列, 并利用跨站关联实现对同一通信事件的多站协同估计 [41,42]. 具体实现步骤如下.

(1) 求解第 s_1 和 s_2 个观测站之间通信域估计序列的相似度 r_{c,s_1,s_2}:

$$r_{c,s_1,s_2}=\mathbb{I}\big(\bar{\mathcal{M}}_{c,s_1}=\bar{\mathcal{M}}_{c,s_2}\big)\cdot\mathrm{e}^{-\frac{(t_{c,s_1}-t_{c,s_2})^2}{\sigma_{t_c}^2}-\frac{(\bar{f}_{c,s_1}-\bar{f}_{c,s_2})^2}{\sigma_{f_c}^2}-\frac{(\bar{B}_{c,s_1}-\bar{B}_{c,s_2})^2}{\sigma_{B_c}^2}-\frac{\left\|\bar{\mathbf{u}}_{c,s_1}-\bar{\mathbf{u}}_{c,s_2}\right\|^2}{\sigma_{u_c}^2}}, \tag{7}$$

其中, $\mathbb{I}(\cdot)$ 为指示函数, 括号内条件成立时取 1, 否则取 0. $\bar{\mathcal{M}}_{c,s_1}$ 和 $\bar{\mathcal{M}}_{c,s_2}$ 分别表示两个观测站估计的调制方式, t_{c,s_1}、$\bar{f}_{c,s_1}$、$\bar{B}_{c,s_1}$ 和 $\bar{\mathbf{u}}_{c,s_1}$ 分别为观测站侦测脉冲的到达时间、信号载频估计、信号带宽估计和信号到达方向单

位向量估计, σ_{t_c}、σ_{f_c}、σ_{B_c} 和 σ_{u_c} 分别为时间、频率、带宽和角度的一致性容忍参数. 为获得跨站脉冲间全局最优的关联关系, 采用匈牙利算法求解最大权关联, 关联成功的脉冲组成为跨站观测组.

(2) 几何定位与误差评估. 利用跨站观测组进行交叉定位, 建立加权最小二乘定位模型, 通过高斯 – 牛顿等迭代算法, 求解该跨站观测组的位置估计最优解 $\bar{\mathbf{p}}_c^*$ [43, 44]:

$$\bar{\mathbf{p}}_c^* \equiv \arg\min_{\bar{\mathbf{p}}_c} \sum_{s=1}^{N_s} w_{c,s} \left\| \left(\mathbf{I} - \bar{\mathbf{u}}_{c,s}\, \bar{\mathbf{u}}_{c,s}^{\top}\right)\left(\bar{\mathbf{p}}_c - \mathbf{p}_s\right) \right\|^2, \tag{8}$$

其中, $\bar{\mathbf{p}}_c$ 为待求解的位置估计变量, $\mathbf{p}_s$ 为第 s 个观测站的站址坐标, $\mathbf{I}$ 为单位矩阵, $w_{c,s}$ 为该观测站在通信域交叉定位中的置信度, N_s 为观测站总数, $\bar{\mathbf{u}}_{c,s}$ 为观测站测得的信号到达方向单位向量, $\bar{\mathbf{u}}_{c,s}\, \bar{\mathbf{u}}_{c,s}^{\top}$ 为到达方向向量的正交投影矩阵, 用于度量估计目标与第 s 个观测站测向射线之间的垂直偏差.

(3) 通信事件估计. 通信事件估计首先进行通信事件中的发射源与接收端配对. 同一通信事件包含两个在时序上成对出现的观测组, 发射源发射数据包后, 接收端发射确认包, 因此可将时空上成对出现的观测组归为同一通信事件. 对于发射源 j_1 与接收端 j_2 在 t_c 时刻的通信事件估计 $\bar{\mathbf{E}}_{c,j_1,j_2}(t_c)$ 可表示为

$$\bar{\mathbf{E}}_{c,j_1,j_2}(t_c) = \left\{\bar{f}_{c,j_1}(t_c),\ \bar{B}_{c,j_1}(t_c),\ \bar{\mathcal{M}}_{c,j_1}(t_c), \bar{\mathbf{p}}_{c,j_1}(t_c),\ \bar{\mathbf{p}}_{c,j_2}(t_c),\ w_{c,j_1,j_2}(t_c)\right\}, \tag{9}$$

其中, $\bar{f}_{c,j_1}(t_c)$、$\bar{B}_{c,j_1}(t_c)$ 和 $\bar{\mathcal{M}}_{c,j_1}(t_c)$ 分别为载波频率估计、信号带宽估计和调制方式估计, $\bar{\mathbf{p}}_{c,j_1}(t_c)$ 和 $\bar{\mathbf{p}}_{c,j_2}(t_c)$ 分别为该通信事件估计中的发射源和接收端位置估计, $w_{c,j_1,j_2}(t_c)$ 为通信事件估计置信度. 最终观测站侦测到的发射源总数为 N_c, 受环境条件等影响, 观测站侦测到的发射源总数 N_c 一般小于集群发射源总数 $N_{c,all}$.

4.1.4　通信域模态子图构建

基于通信事件估计结果构建通信域模态子图 $G_c = (V_c, E_c)$, 节点 $j_1 \in V_c$ 对应发射源无人机, 边 $e_{j_1,j_2} \in E_c$ 表征发射源 j_1 与接收端 j_2 之间有通信交互行为, t_c 时刻下 j_1 的通信特征向量 $\mathbf{h}_{c,j_1}(t_c)$ 为

$$\mathbf{h}_{c,j_1}(t_c) = \left[\bar{f}_{c,j_1}(t_c),\ \bar{B}_{c,j_1}(t_c),\ \bar{\mathcal{M}}_{c,j_1}(t_c), \bar{\mathbf{P}}_{c,j_1}(t_c),\ \bar{\mathbf{P}}_{c,j_2}(t_c),\ w_{c,j_1,j_2}(t_c)\right], \tag{10}$$

其中, $\bar{f}_{c,j_1}(t_c)$、$\bar{B}_{c,j_1}(t_c)$ 和 $\bar{\mathcal{M}}_{c,j_1}(t_c)$ 分别为载波频率估计、信号带宽估计与调制方式估计, $\bar{\mathbf{P}}_{c,j_1}(t_c)$ 和 $\bar{\mathbf{P}}_{c,j_2}(t_c)$ 为发射源 j_1 和接收端 j_2 的位置估计, $w_{c,j_1,j_2}(t_c)$ 为通信事件置信度. 发射源 j_1 和接收端 j_2 之间的通信事件边权重 ρ_{c,j_1,j_2} 为

$$\rho_{c,j_1,j_2} = \frac{N_{c,j_1,j_2}}{\sum_q N_{c,j_1,q}} \cdot \mathrm{e}^{-\sigma^2_{j_1,j_2}}, \tag{11}$$

其中 N_{c,j_1,j_2} 为 j_1 与 j_2 间通信事件总数, $\sum_q N_{c,j_1,q}$ 为 j_1 与其他接收端通信事件总数, $\sigma^2_{j_1,j_2}$ 为定位误差方差.

由此, 通过构建通信域模态子图 G_c 实现以图结构形式, 表征集群的通信事件参数特征、空间分布和活跃程度情况, 为后续双域节点关联提供通信域的结构化先验输入.

4.2　双域节点层: 模态子图结构化表征与双域节点关联

本小节基于各域内的节点协同关系, 采用双模态异质图注意力网络聚合邻居信息进行结构化表征, 通过注意力机制将双域异构特征映射至公共交互空间, 并据此构建双域相似度矩阵实现双域节点关联.

4.2.1　模态子图协同关系结构化表征

对于物理域模态子图 $G_r = (V_r, E_r)$, 物理节点 i_1 的特征向量 $\mathbf{h}_{r,i_1}$ 经多头图注意力网络聚合邻居信息后, 输出其高阶物理特征向量 $\mathbf{h}'_{r,i_1}$:

$$\mathbf{h}'_{r,i_1} = \Big\|_{k=1}^{K} \sigma\left(\sum_{i_2 \in \mathcal{N}_{r,i_1}} \alpha^k_{r,i_1,i_2} \mathbf{W}^k_r \mathbf{h}_{r,i_2}\right), \tag{12}$$

其中, K 表示注意力头数, σ 为非线性激活函数, $\mathbf{W}_r^k$ 对应第 k 个注意力机制的可学习权重矩阵, $\|$ 表示特征向量的拼接操作, $\mathcal{N}_{r,i_1}$ 为节点 i_1 的邻居节点总集合, $i_2 \in \mathcal{N}_{r,i_1}$ 为 i_1 的邻居节点, α_{r,i_1,i_2}^k 是邻居节点 i_2 对 i_1 在第 k 个注意力机制的归一化系数, 表征邻居节点 i_2 的贡献程度, 具体表示为

$$\alpha_{r,i_1,i_2}^k = \frac{\exp\left(\mathrm{LeakyReLU}\left((\boldsymbol{a}_r^k)^\top \left[\mathbf{W}_r^k \mathbf{h}_{r,i_1} \| \mathbf{W}_r^k \mathbf{h}_{r,i_2}\right]\right)\right)}{\sum\limits_{i_2' \in \mathcal{N}_{r,i_1}} \exp\left(\mathrm{LeakyReLU}\left((\boldsymbol{a}_r^k)^\top \left[\mathbf{W}_r^k \mathbf{h}_{r,i_1} \| \mathbf{W}_r^k \mathbf{h}_{i_2'}\right]\right)\right)}, \tag{13}$$

其中, $\boldsymbol{a}_r^k$ 是可学习注意力向量, $\mathbf{W}_r^k$ 为可学习权重矩阵, $\|$ 表示特征向量的拼接操作, LeakyReLU 为激活函数, $i_2' \in \mathcal{N}_{r,i_1}$ 为对节点 i_1 的所有邻居节点进行遍历求和, 实现归一化.

所有物理节点的高阶物理特征向量共同构成物理域模态子图的编码矩阵 $\mathbf{H}_r = \left[\mathbf{h}_{r,1}'; \mathbf{h}_{r,2}'; \cdots; \mathbf{h}_{r,N_r}'\right]^\top$, 其中 N_r 为物理节点总数. 物理域模态子图的编码矩阵通过聚合物理邻居信息, 编码集群物理域协同行为引发的节点间相对空间结构与运动协调性, 将孤立节点的原始量测转化为反映集群动态角色的结构化深度表征.

对于通信域模态子图 $G_c = (V_c, E_c)$, 发射源 j_1 的通信特征向量 $\mathbf{h}_{c,j_1}$ 经多头图注意力网络聚合邻居信息后, 输出高阶通信特征向量 $\mathbf{h}_{c,j_1}'$:

$$\mathbf{h}_{c,j_1}' = \Big\|_{k=1}^{K} \sigma\left(\sum_{j_2 \in \mathcal{N}_{c,j_1}} \alpha_{c,j_1,j_2}^k \mathbf{W}_c^k \mathbf{h}_{c,j_2}\right). \tag{14}$$

所有通信节点的高阶通信特征向量共同构成通信域模态子图的编码矩阵 $\mathbf{H}_c = \left[\mathbf{h}_{c,1}'; \mathbf{h}_{c,2}'; \cdots; \mathbf{h}_{c,N_c}'\right]^\top$, 其中 N_c 为通信发射源总数. 通信域模态子图编码矩阵 $\mathbf{H}_c$ 通过对通信邻居信息的聚合, 编码了节点间的时频协同关系与通信拓扑结构, 将孤立的通信事件转化为反映集群通信交互模式的结构化深度表征.

4.2.2　跨模态特征对齐与双域节点关联

由于物理模态编码矩阵 $\mathbf{H}_r$ 与通信域模态子图编码矩阵 $\mathbf{H}_c$ 分别源于两个异构的语义空间, 使用传统计算欧氏距离或余弦相似度的方法将面临语义鸿沟与度量失准的问题. 对此, 本文借助注意力机制, 将双域异构特征进行拼接, 而后将联合特征映射至公共交互空间进行对齐, 并据此构建双域相似度矩阵实现双域节点关联. 具体而言, 将物理编码矩阵 $\mathbf{H}_r$ 与通信编码矩阵 $\mathbf{H}_c$ 作为输入, 计算基于注意力机制的跨模态双域节点相似度矩阵 $\mathbf{S}_{\mathrm{cross}}$, 其元素 $S_{\mathrm{cross},i,j}$ 量化了物理节点 i 与通信发射源节点 j 属于同一物理实体的可能性

$$S_{\mathrm{cross},i,j} = \mathrm{Attn}_{\mathrm{cross}}\left(\mathbf{h}_{r,i}', \mathbf{h}_{c,j}'\right) = \mathbf{v}_a^\top \cdot \tanh\left(\mathbf{W}_a \left[\mathbf{h}_{r,i}' \| \mathbf{h}_{c,j}'\right] + b_a\right), \tag{15}$$

其中 $\mathbf{h}_{r,i}'$ 为高阶物理特征向量, $\mathbf{h}_{c,j}'$ 为高阶通信特征向量, $\mathbf{v}_a$ 是可学习的权重向量, $\|$ 表示向量拼接操作, $\mathbf{W}_a$ 为线性变换矩阵, b_a 为偏置项, $\tanh(\cdot)$ 为双曲正切激活函数. $S_{\mathrm{cross},i,j}$ 值越高表示物理节点 i 与通信发射源 j 匹配为同一物理实体的可能性越大.

由于跨模态双域节点相似度矩阵 $\mathbf{S}_{\mathrm{cross}}$ 是原始结果, 其数值范围未经约束. 为获得更稳定的训练过程, 将 $\mathbf{S}_{\mathrm{cross}}$ 归一化为双域节点关联概率分布 $\mathbf{P}_{\mathrm{cross}}$. 定义物理节点 i 在通信发射源节点集合上的双域节点关联概率 $\mathbf{P}_{\mathrm{cross},i} = \left[P_{\mathrm{cross},i,1}, P_{\mathrm{cross},i,2}, \ldots, P_{\mathrm{cross},i,N_c}\right]$, 则物理节点 i 与通信发射源 j 的双域节点关联概率 $P_{\mathrm{cross},i,j}$ 为

$$P_{\mathrm{cross},i,j} = \frac{\exp\left(S_{\mathrm{cross},i,j}/\tau\right)}{\sum_{j'=1}^{N_c} \exp\left(S_{\mathrm{cross},i,j'}/\tau\right)}, \tag{16}$$

其中, $\tau > 0$ 为温度系数, 用于调节概率分布的平滑程度, N_c 为通信发射源节点的总数, $j' \in \{1, 2, \ldots, N_c\}$ 为对物理节点 i 的所有通信发射源节点进行遍历求和, 实现归一化.

4.3　集群全局层: 两阶段双域匹配与双域协同认知图构建

双域节点层的双域节点关联概率 $\mathbf{P}_{\text{cross}}$ 仅为双域节点候选配对的关联概率, 在集群全局层还需输出确定的集群双域匹配关系. 可信的双域匹配结果需要满足每个物理目标只与一个通信发射源匹配, 且集群全局的双域匹配置信度最高. 传统直接求解方法容易陷入局部冲突出现次优解, 因此提出一种软匹配和硬匹配的两阶段双域匹配求解策略, 实现双域节点一一对应与集群全局最优的双域匹配结果. 基于集群双域匹配结果, 构建集群 “物理运动 – 通信交互” 行为的协同精准认知.

4.3.1　基于最优传输理论的软匹配

首先将双域节点关联问题形式化为最优传输 (optimal transport, OT) 问题, 利用 Sinkhorn 算法求解满足双随机约束的双域软匹配概率矩阵 [45]. 定义双域软匹配概率矩阵 $\mathbf{\Pi}$, 其元素 $\Pi_{i,j}$ 表示满足双随机约束下物理节点 i 与通信发射源 j 的双域匹配软概率, 其理论最优解 $\mathbf{\Pi}^*$ 可表述为

$$\mathbf{\Pi}^* = \arg\max_{\mathbf{\Pi}\in\mathcal{U}} \left(\langle \mathbf{\Pi}, \mathbf{S}_{\text{cross}}\rangle_F + \varepsilon \mathrm{H}(\mathbf{\Pi})\right), \tag{17}$$

其中, $\langle\cdot,\cdot\rangle_F$ 是 Frobenius 内积, $\langle \mathbf{\Pi}, \mathbf{S}_{\text{cross}}\rangle_F = \sum_{i=1}^{N_r}\sum_{j=1}^{N_c} \Pi_{i,j}\cdot S_{\text{cross},i,j}$ 用于在全局约束下将双域匹配总相似度最大化. 熵正则化项 $\varepsilon\mathrm{H}(\mathbf{\Pi})$ ($\varepsilon > 0$) 用于平滑优化问题, 防止退化为硬分配, 保持可微性便于端到端训练. $\mathrm{H}(\mathbf{\Pi}) = -\sum_{i,j}\Pi_{i,j}\log\Pi_{i,j}$ 是双域软匹配概率矩阵 $\mathbf{\Pi}$ 的熵. $\mathcal{U}$ 是满足双随机约束的集合, 要求双域软匹配概率矩阵 $\mathbf{\Pi}$ 的所有行和与所有列和均等于相同常数, 确保双域匹配关系在全局上是一一对应的, 消除双域匹配冲突的可能性, 为后续求解一一对应的双域匹配结果提供唯一性基础. 满足双随机约束的集合 $\mathcal{U}$ 表示为

$$\mathcal{U} = \left\{ \mathbf{\Pi} \in \mathbb{R}_+^{N_r\times N_c} \mid \mathbf{\Pi}\mathbf{1}_{N_c} = \frac{1}{N_r}\mathbf{1}_{N_r}, \mathbf{\Pi}^{\top}\mathbf{1}_{N_r} = \frac{1}{N_c}\mathbf{1}_{N_c} \right\}, \tag{18}$$

其中, N_r 为物理节点总数, N_c 为通信发射源总数, $\mathbf{1}_n$ 为 n 维全 1 向量.

在双域软匹配概率矩阵求解过程中, 熵正则化项 $\varepsilon\mathrm{H}(\mathbf{\Pi})$ 的引入使得原本可能退化为硬分配的线性规划问题转化为严格凸优化问题, 保证最优解的唯一性, 且保持可微性便于端到端训练. 通过对双域软匹配概率矩阵求解问题引入拉格朗日乘子 $\alpha\in\mathbb{R}^{N_r}$ 和 $\beta\in\mathbb{R}^{N_c}$, 并对拉格朗日函数关于 $\Pi_{i,j}$ 求导令其为零, 可得双域软匹配概率理论最优解的指数形式

$$\Pi_{ij}^* = \exp\left(\frac{S_{\text{cross},i,j} + \alpha_i + \beta_j}{\varepsilon}\right). \tag{19}$$

通过对 α 和 β 进行交替迭代更新来逼近软匹配概率理论最优解 Π_{ij}^*. 具体地, 定义核矩阵 $\mathbf{K} = \exp\left(\frac{\mathbf{S}_{\text{cross}}}{\varepsilon}\right)$, 通过交替缩放行与列进行迭代逼近双随机矩阵

$$\mathbf{U}^{(t+1)} = \frac{\mathbf{1}_{N_r}}{\mathbf{K}\mathbf{V}^{(t)} + \epsilon_{\text{num}}}, \quad \mathbf{V}^{(t+1)} = \frac{\mathbf{1}_{N_c}}{\mathbf{K}^{\top}\mathbf{U}^{(t+1)} + \epsilon_{\text{num}}}, \tag{20}$$

其中 ϵ_{num} 为数值稳定项, 避免归一化和指数运算中的数值不稳定. 经 T 次迭代后得到双域软匹配概率矩阵 $\mathbf{\Pi}$:

$$\mathbf{\Pi} = \mathrm{diag}\left(\mathbf{U}^{(T)}\right)\mathbf{K}\mathrm{diag}\left(\mathbf{V}^{(T)}\right). \tag{21}$$

双域软匹配概率矩阵 $\mathbf{\Pi}$ 是近似最优解, 能够消除集群双域匹配中的局部冲突, 且求解过程可微, 使得梯度能够从 $\mathbf{\Pi}$ 回传至双域节点层, 进而更新网络训练参数, 实现整个网络的端到端优化.

4.3.2　基于匈牙利算法的硬匹配

经过最优传输软匹配后, 获得满足双随机约束的双域软匹配概率矩阵 $\mathbf{\Pi}$, 其元素 $\Pi_{i,j}$ 表示满足双随机约束下物理节点 i 与通信发射源 j 的双域匹配软概率. 由于 $\mathbf{\Pi}$ 输出的是概率化的双域匹配关系, 而非明确的确定性

决策, 需进一步将 $\mathbf{\Pi}$ 转化为离散的硬匹配结果, 为此引入匈牙利算法进行离散化决策 [46]. 首先将双域软匹配概率矩阵 $\mathbf{\Pi}$ 转换为分配代价矩阵 $\mathbf{C}$:

$$\mathbf{C} = 1 - \mathbf{\Pi}. \tag{22}$$

其元素 $C_{i,j} = 1 - \Pi_{i,j}$ 表示将物理节点 i 分配给通信发射源 j 的双域匹配代价, 代价越小表示双域匹配置信度越高. 在此基础上, 采用匈牙利算法求解使全局分配代价最小且满足一一对应约束的最优硬匹配 $\mathbf{\Phi}^*$:

$$\mathbf{\Phi}^* = \arg\min_{\mathbf{\Phi}} \sum_{i=1}^{N_r} \sum_{j=1}^{N_c} C_{i,j} \cdot \mathbb{I}\big(\mathbf{\Phi}(i) = j\big), \tag{23}$$

满足

$$\begin{cases} \displaystyle\sum_{j=1}^{N_c} \mathbb{I}\big(\mathbf{\Phi}(i) = j\big) \leqslant 1, \quad \forall i \in \{1, \ldots, N_r\}, \\ \displaystyle\sum_{i=1}^{N_r} \mathbb{I}\big(\mathbf{\Phi}(i) = j\big) \leqslant 1, \quad \forall j \in \{1, \ldots, N_c\}, \end{cases} \tag{24}$$

其中, $\mathbb{I}(\cdot)$ 是指示函数, 当条件成立时取 1 否则取 0, $C_{i,j}$ 为代价矩阵元素, $\mathbf{\Phi}(\cdot)$ 为双域正确匹配映射函数, $\mathbf{\Phi}(i) = j$ 表示双域匹配正确. 当 $N_r \neq N_c$ 时, 通过引入虚拟节点扩展为方阵, 将虚拟节点对应的代价设为极大值, 使虚拟节点仅在必要时参与双域匹配, 从而获得近似最优解, 最终完成全局最优的双域匹配.

4.3.3　集群双域协同认知图构建

通过求解全局最优硬匹配映射 $\mathbf{\Phi}^*$, 建立双域之间节点一一对应且集群全局最优的映射关系, 进而构建集群双域协同认知拓扑图, 实现对集群 “物理运动 – 通信交互” 行为的协同关系认知.

定义集群双域协同认知拓扑图 $G_{\text{p-c}} = (V_{\text{p-c}}, E_{\text{p-c}}, W)$. $V_{\text{p-c}}$ 为满足双域最优匹配关系下的无人机节点对集合, 其元素 $v_{\text{p-c}} \in V_{\text{p-c}}$ 表征双域最优匹配 $(i, \mathbf{\Phi}^*(i))$ 下的物理节点与通信发射源节点对, $N_{\text{p-c}}$ 为节点对总数, $v_{\text{p-c}}$ 的特征向量 $\mathbf{h}_{v_{\text{p-c}}}$ 是双域模态特征的融合

$$\mathbf{h}_{v_{\text{p-c}}} = \text{MLP}_{\text{fuse}}\left(\left[\mathbf{h}'_{r,i} \,\middle\|\, \mathbf{h}'_{c,\mathbf{\Phi}^*(i)}\right]\right), \tag{25}$$

其中, $\mathbf{h}'_{r,i}$ 与 $\mathbf{h}'_{c,\mathbf{\Phi}^*(i)}$ 为物理域与通信域模态子图的高阶特征向量, $\|$ 表示向量拼接, MLP_{fuse} 是一个轻量级多层感知机 [47], 用于最优的非线性特征融合方式, 使双域节点对获得兼具 “物理运动 – 通信交互” 行为双域特性, 进而构建同一无人机下融合双域行为状态的完整个体画像.

集群双域协同认知图的边集合 $E_{\text{p-c}}$, 是以物理运动关联边为基础, 叠加通信交互边. 即: 保留无人机物理实体间基于运动关联的固有物理边, 同时当两物理实体进行通信交互时, 在固有物理边基础上新增通信交互边, 实现物理域和通信域边信息的融合表征. $\mathbf{W}_{\text{p-c}}$ 为边权矩阵, 可基于物理运动关联特性与通信活跃度动态赋值.

集群协同认知拓扑图的节点、边和权重可随时间窗滑动而更新, 从而刻画非合作集群 “物理运动 – 通信交互” 行为双域特性的动态演化过程, 深化对非合作无人机集群的认知水平, 为后续理解非合作集群意图、辨识关键节点与协作模式提供结构化的知识表示.

4.4　网络训练参数自智优化

为提升方法的自智优化能力, 使双域节点层的双模态异质图注意力网络与跨模态注意力模块根据集群全局层的双域匹配结果, 在跨层级联合代价的约束下学习得到判别性更强的训练参数, 引入总损失函数对模型进行端到端训练. 定义总损失函数 $\mathcal{L}$ 为

$$\mathcal{L} = \lambda \mathcal{L}_{\text{smooth}} + \mathcal{L}_{\text{match}} + \eta \|\Theta\|_2^2, \tag{26}$$

其中, Θ 表示模型全部可学习参数, 平滑正则项系数 λ 与 L2 正则项系数 η 通过交叉验证设定, $\mathcal{L}_{\text{smooth}}$ 为平滑正则项, $\mathcal{L}_{\text{match}}$ 为基于软匹配概率矩阵的损失函数.

平滑正则项 $\mathcal{L}_{\text{smooth}}$ 是为进一步提升匹配的物理空间合理性, 通过约束物理邻域内节点的匹配概率分布趋于一致, 使双域节点关联结果在空间上保持平滑, 计算公式为

$$\mathcal{L}_{\text{smooth}} = \frac{1}{|E_r|} \sum_{(i,i')\in E_r} \left\| \mathbf{P}_{\text{cross},i} - \mathbf{P}_{\text{cross},i'} \right\|_2^2, \tag{27}$$

其中, E_r 为物理域模态子图边集合, i' 为物理节点 i 的邻近物理节点, $\mathbf{P}_{\text{cross},i}$ 为 i 在其通信发射源节点集合上的双域节点关联概率, $\mathbf{P}_{\text{cross},i'}$ 为 i' 的双域节点关联概率. 平滑正则项 $\mathcal{L}_{\text{smooth}}$ 可使邻域内具有相似运动行为的物理节点呈现相近的双域节点关联结果分布, 从而提升整体双域匹配的平滑性.

损失函数 $\mathcal{L}_{\text{match}}$ 是基于软匹配阶段的双域软匹配概率矩阵 $\mathbf{\Pi}$ 构建的, 通过最小化该损失, 引导模型将更高关联概率分配给双域最优匹配 $(i, \mathbf{\Phi}^*(i))$ 下的物理节点与通信发射源对, 驱动网络学习更具判别力的参数表示. 损失函数 $\mathcal{L}_{\text{match}}$ 为

$$\mathcal{L}_{\text{match}} = -\frac{1}{N_{\text{p-c}}} \sum_{i=1}^{N_{\text{p-c}}} \log\left(\Pi_{i,\mathbf{\Phi}^*(i)} + \delta\right), \tag{28}$$

其中, $N_{\text{p-c}}$ 为双域匹配下的节点对总数, $\mathbf{\Phi}^*(i)$ 是双域最优匹配下物理节点 i 对应的通信发射源. $\Pi_{i,\mathbf{\Phi}^*(i)}$ 为双域节点关联概率, δ 为极小常数以避免数值不稳定.

由此, 通过最小化总损失函数 $\mathcal{L}$, 并利用随机梯度下降类优化器对模型训练参数进行更新

$$\Theta - \mu \frac{\partial \mathcal{L}}{\partial \Theta} \to \Theta, \tag{29}$$

其中, μ 为学习率. 随着训练参数的优化, 双域节点层的双模态异质图注意力网络与跨模态注意力模块在集群全局级联合代价的约束下在线自智优化, 实现在无先验条件下的跨层级参数协同演进.

5　实验与分析

5.1　实验设置

5.1.1　实验场景与参数设置

本小节以非合作无人机集群侵扰低空 Agentic AI-RAN 网络为实验场景, 开展方法验证工作. 本文所提方法均在 MATLAB 与 Python3.8.19 环境下实现, 所有实验在搭载 Intel Core i5 CPU 与 NVIDIA RTX 3000 系列 GPU 平台进行. 在双域特征层, 基于集群典型物理运动和通信交互参数进行系统建模 [6~8,13,36,37], 具体的物理运动参数和通信交互参数设置分别如表 1 和 2 所示. 本实验采用 3 个观测站对非合作集群进行协同探测, 物理域与通信域观测站共址部署, 但射频通道与信号处理参数独立配置, 观测站参数设置见表 3 [31,33,34,36]. 双域模态子图中节点与边构建规则的阈值参数见表 4 [6,7,39]. 在双域节点层, 利用双模态异质图注意力网络对物理域模态子图与通信域模态子图进行独立分支编码, 双模态异质图注意力网络参数设置见表 5 所示 [22~24,26,27,29,38,39]. 在集群全局层, 设置基于最优传输软匹配环节的正则系数 $\varepsilon = 0.10$, 迭代次数 $T = 20$, 数值稳定项 $\epsilon_{\text{num}} = 1.0\times 10^{-6}$ [45]. 设置网络总损失函数中平滑正则项系数 λ 为 0.1, L2 正则项系数 η 为 1.0×10^{-4}, 数值稳定项 δ 为 1.0×10^{-8}. 网络已包含 L2 正则系数与早停机制, 设置训练超参数 Dropout 率为 0, 训练方式为 mini-batch, 优化器为 adam, 学习率 μ 为 2.0×10^{-4}, 批大小为 8 图/批, 训练轮数耐心值为 10 轮 [7,23,24,39].

5.1.2　数据集构建

非合作无人机集群的双域仿真数据集共包含 2300 个相互独立的场景样本, 覆盖 5~50 架无人机数量、集群航向、高度分布、飞行速度和通信调制方式的不同组合. 数据集按 6:2:2 的比例随机划分为训练集、验证集和测

表 1　非合作无人机集群物理运动参数设置.

Table 1　Physical motion parameter setting for non-cooperative UAV swarm.

Parameter name	Symbol	Value/range	Unit
Total number of UAVs in swarm	$N_{r,all}$	$[5, 50]$	–
Initial position of swarm center	P_{c0}	$x \in [-1400, 1400], y \in [0, 1500], z \in [90, 180]$	m
Minimum distance between UAVs	$D_{\min}$	40	m
Altitude and fluctuation of UAVs	h	$[40, 240]$	m
Initial velocity direction of UAVs	$\mathbf{V}_0$	$[0, 1, 0]^\top$	–
Horizontal velocity magnitude of UAVs	V_{xy}	$[10, 30]$	m/s
Vertical velocity magnitude of swarm	V_z	0	m/s
UAV RCS	σ_{RCS}	0.1	m^2

表 2　非合作无人机集群通信交互参数设置.

Table 2　Parameter setting of the communication interaction for non-cooperative UAV swarm.

Parameter name	Symbol	Value/range	Unit
Carrier frequency band	f_c	$\{2.4, 5.2\}$	GHz
Signal bandwidth	B_c	5.0	MHz
Transmit antenna gain	G_t	3	dBi
Command frame duration	T_{cmd}	50	ms
Acknowledgment frame duration	T_{ack}	0.2	ms
Mean guard interval	T_{guard}	50	μs
Standard deviation of guard interval	$\sigma_{T_{\text{guard}}}$	10	μs
Transmission scheme and modulation	$\mathcal{M}_c$	{BPSK, 8PSK, 64QAM, FHSS, OFDM}	–

表 3　非合作集群物理域与通信域信号观测站参数设置.

Table 3　Parameter setting of observation stations in physical domain and communication domain for non-cooperative swarm.

Parameter name	Symbol	Physical domain	Communication domain	Unit
Number of observation stations	N_s	3	3	–
Operating frequency	f_o	10	$\{2.4, 5.2\}$	GHz
Signal bandwidth	B	120	5	MHz
Sampling frequency	f_s	300	15	MHz
Transmit power for physical observation	$P_{s,t}$	10	–	W
Receive gain	$G_{s,r}$	30	12	dBi
Pulse repetition interval	PRI	100	–	μs
Number of pulses within coherent processing interval	N_{pulse}	384	–	–
Scanning range (azimuth)	θ_{az}	±70	Omni	°
Scanning range (elevation)	θ_{el}	±10	Omni	°
Maximum detection range	$R_{\max}$	4500	–	m
Range resolution	ΔR	1.25	–	m
Velocity resolution	Δv	0.39	–	m/s
Direction-finding accuracy	σ_{DOA}	–	1	°
Time synchronization accuracy	σ_t	–	10	ns
Frequency estimation accuracy	σ_f	–	1	kHz

表 4　双域子图构建阈值参数设置.

Table 4　Parameter setting of relevant thresholds for dual-domain subgraph construction.

Parameter name	Symbol	Value	Unit
Euclidean distance constraint of spatial position in physical domain	d_{th}	250	m
Velocity vector angle constraint in physical domain	θ_{th}	35	°
Spatial scale parameter of physical domain	σ_{p_r}	180	m
Velocity similarity scale parameter of physical domain	σ_{v_r}	12	m/s
Temporal consistency tolerance in communication domain	σ_{t_c}	5	μs
Frequency consistency tolerance in communication domain	σ_{f_c}	3	kHz
Bandwidth consistency tolerance in communication domain	σ_{B_c}	0.1	MHz
Angle consistency tolerance in communication domain	σ_{u_c}	2	–

表 5　图注意力网络参数设置.

Table 5　Parameter setting of the graph attention network.

Parameter name	Symbol	Value	Unit
Number of attention heads	K	8	–
Hidden dimension per head	d_h	16	–
Output dimension of graph encoder	d_{out}	128	–
Number of layers	L	2	layer
Activation function of hidden layers	$\sigma_h(\cdot)$	LeakyReLU (α=0.10)	–
Activation function for output layer	$\sigma_e(\cdot)$	ELU	–
Multi-head aggregation strategy	–	Concat	–

试集 [13,33,38]. 其中, 训练集用于 P-C CMCC 方法的模型参数优化, 验证集用于 P-C CMCC 方法的超参数选择、早停决策及模型结构筛选, 测试集用于最后的 P-C CMCC 方法、基线对比和消融实验的性能评估.

5.1.3　评估指标

为全面评估 P-C CMCC、基线方法与消融实验的匹配认知性能, 定义以下核心指标.

(1) 双域匹配准确率 (dual-domain matching accuracy, DMA): 双域匹配正确的节点对 $N_{\text{p-c}}$ 占双域有效节点总数的比例, 衡量双域匹配的正确程度 [21,27]

$$\text{DMA} = \frac{1}{\min(N_r, N_c)} \sum_{i=1}^{N_{\text{p-c}}} \mathbb{I}\big(\boldsymbol{\Phi}^*(i) = j\big), \tag{30}$$

其中, $\mathbb{I}(\cdot)$ 为指示函数, $\boldsymbol{\Phi}^*(i)$ 为最优匹配映射, N_r 和 N_c 分别为观测站探测到的物理节点和通信辐射源节点的总数, 通常 $N_r \neq N_c$, 以 $\min(N_r, N_c)$ 作为有效样本量, 仅对双域同时观测到的节点评估均方根误差, 避免单域缺失引入非对称偏差.

(2) 双域匹配均方根误差 (dual-domain matching root-mean-square error, DRMSE): 双域匹配下物理节点与通信辐射源节点之间的空间位置偏差 [8,13,34]

$$\text{DRMSE} = \sqrt{\frac{1}{\min(N_r, N_c)} \sum_{\min(N_r, N_c)} \left\| \bar{\mathbf{p}}_r - \bar{\mathbf{p}}_c \right\|_2^2}, \tag{31}$$

其中, $\bar{\mathbf{p}}_r$ 为物理节点的空间位置估计, $\bar{\mathbf{p}}_c$ 为通信辐射源节点的空间位置估计.

(3) 处理耗时 (runtime): 算法处理单帧数据的平均时间, 用于评估实时性与计算开销 [23,24,38].

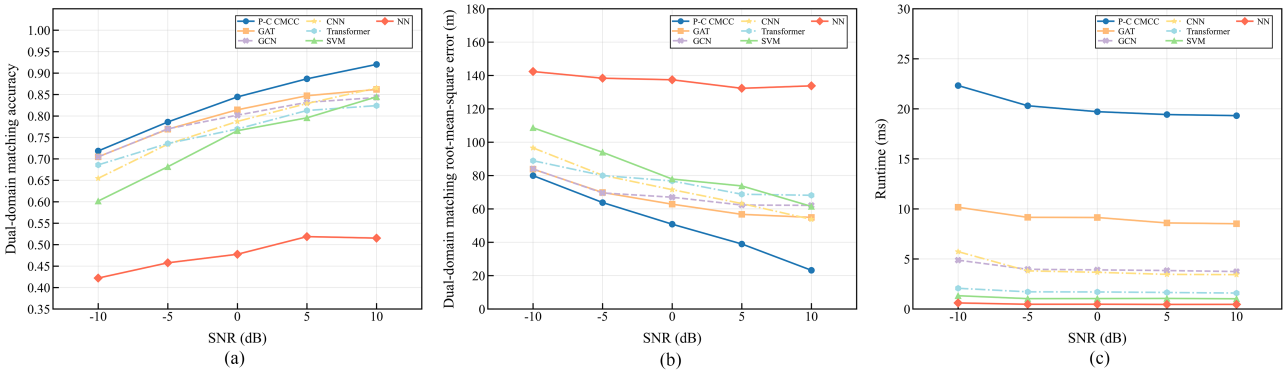

图 3　(网络版彩图) 不同接收信噪比下的算法性能对比. (a) 双域匹配准确率; (b) 双域匹配均方根误差; (c) 处理耗时.

Figure 3　(Color online) Performance comparison of algorithms under different received SNR conditions. (a) Dual-domain matching accuracy; (b) dual-domain matching root-mean-square error; (c) runtime.

5.2　实验结果与分析

5.2.1　实验结果

为验证本文 P-C CMCC 方法对非合作集群双域匹配的有效性, 选取以下基准方法进行比较: 传统最近邻匹配 (NN) [6~8,21]、支持向量机 (SVM) [48]、卷积神经网络 (CNN) [19,20,31]、Transformer [14,41]、图卷积神经网络 (GCN) [27,29]、图注意力网络 (GAT) [22,38].

(1) 不同接收信噪比下的算法鲁棒性. 为评估 P-C CMCC 方法在实际环境中面对通信链路质量波动时的鲁棒性, 本小节以 20 架的集群规模作为实验对象, 测试不同接收信噪比条件下的方法效果. SNR 梯度设置为 −10、−5、0、5 和 10 dB 以模拟典型信道条件. 为模拟真实环境, 每个接收机信噪比场景均随机生成 50 个独立的测试样本, 结果为实验统计平均值. P-C CMCC 方法与基线对比算法在不同接收信噪比条件下的评估指标结果如图 3 所示.

(a) 双域匹配准确率 (DMA). 随着 SNR 提升, 所有算法的 DMA 均呈上升趋势. 原因是 SNR 越高, 地面观测站侦测的无人机信号质量越高, 有益于双域观测样本之间的匹配. 此外, P-C CMCC 算法在全部 SNR 下均优于其他基线算法, 10 dB 时 DMA 为 0.92, 在 −10 dB 下仍能保持 0.72. 最近邻方法在 −10 和 10 dB 下 DMA 仅为 0.42 和 0.51, 主要原因是最近邻方法依赖特征空间中的欧氏距离进行模板匹配, 缺乏对特征深化表征提取能力. SVM 的 DMA 高于最近邻结果但也低于其他方法, 原因是 SVM 通过最大间隔和核函数提升了非线性分类能力, 使其抗噪性优于最近邻, 但 SVM 本质是浅层模型, 难以在耦合复杂和噪声干扰中提取深层结构. Transformer 与 CNN 凭借深度架构带来的强非线性表征与拟合能力, 其性能优于传统浅层模型, 然而二者对图拓扑关系的适应性不足, 因此 DMA 虽有提升但增益有限. GAT 和 GCN 在处理图结构问题具备天然优势, 但仅依靠网络的深层表征能力有限, 缺乏对双域特性的针对性建模, 因此其 DMA 低于 P-C CMCC.

(b) 双域匹配均方根误差 (DRMSE). 随着 SNR 提升, 所有算法的 DRMSE 均呈下降趋势, 原因是信号质量提升降低了双域匹配偏差. P-C CMCC 方法始终保持最低 DRMSE, 在 −10 dB 下 DRMSE 为 79.96 m, 在 10 dB 时为 23.21 m, 验证了本方法的鲁棒性. 对比而言, 最近邻方法的 DRMSE 始终最高, 在 −10 dB 下为 142.35 m, 在 10 dB 下为 137.78 m. SVM 的 DRMSE 低于最近邻方法, 但浅层模型难以应对复杂非线性耦合噪声, RMSE 仍处于较高水平. Transformer、CNN 以及 GAT、GCN 凭借深度架构的非线性表征能力, 使 DRMSE 低于最近邻与 SVM, 然而未能针对非合作集群的双域特性进行协同建模, 因而其误差抑制能力仍劣于 P-C CMCC.

(c) 处理耗时. 各算法的处理耗时受 SNR 变化的影响极小, 原因是处理耗时主要取决于模型自身的架构复杂度与计算规模. P-C CMCC 方法的处理耗时最长, 平均为 20.22 ms, 其计算复杂度主要来源于三个核心环节: 跨模态异质图注意力网络、跨模态注意力计算、软匹配和硬匹配两阶段求解. 相比之下, 各基线算法在处理耗时上占据优势. 最近邻无需参数训练耗时最低, 平均为 0.49 ms. SVM、CNN、Transformer 及 GCN 的计算时

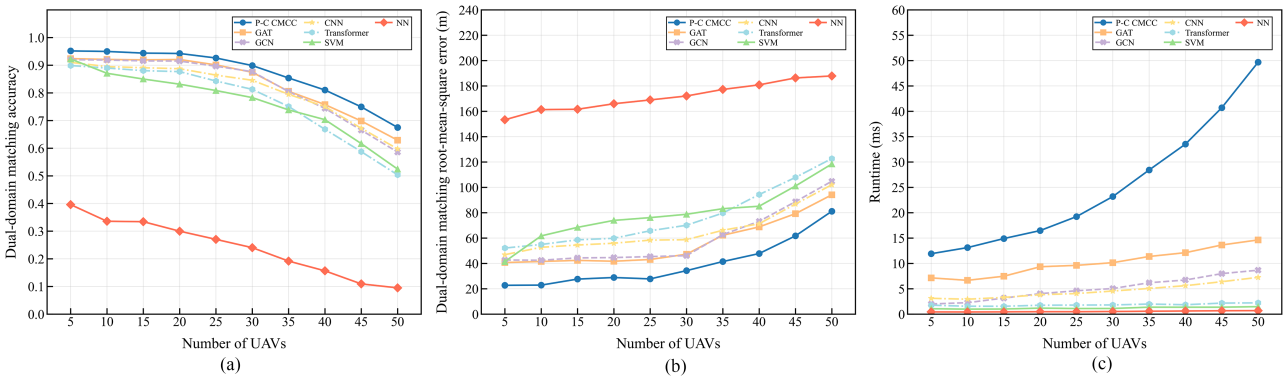

图 4　(网络版彩图) 不同集群规模下的算法性能对比. (a) 双域匹配准确率; (b) 双域匹配均方根误差; (c) 处理耗时

Figure 4　(Color online) Performance comparison of algorithms under different swarm scales. (a) Dual-domain matching accuracy; (b) dual-domain matching root-mean-square error; (c) runtime.

间处于中低水平, GAT 由于采用注意力机制对邻域节点特征进行自适应加权聚合, 致使计算开销略高, 平均为 9.12 ms, 为 P-C CMCC 方法的 45.10%. 综合来看, P-C CMCC 在处理耗时上不占优势, 以计算资源和时间成本换取匹配正确率, 适用于低 SNR 下对匹配精度要求高且计算资源相对充足的场景.

(2) 不同集群规模下的算法泛化性. 为评估 P-C CMCC 方法在不同集群规模下的泛化能力, 本小节在 SNR 为 5 dB 下, 设置无人机集群规模从 5 至 50 架变化, 涵盖小型到中大规模的典型非合作集群场景. 每个规模下随机生成 50 个独立测试样本, 实验结果取统计平均值. P-C CMCC 方法与基线对比算法在不同集群规模下的评估指标结果如图 4 所示.

(a) 双域匹配准确率 (DMA). 随着集群规模从 5 架增加至 50 架, 不同无人机之间的特征混叠与互干扰现象显著加剧, 使得特征空间中的样本分布愈发密集, 同时候选匹配的数量呈指数级增长, 导致 P-C CMCC 方法与其他基线算法的 DMA 均大幅下降. P-C CMCC 算法在所有集群规模下的 DMA 均为最优, 在集群规模由 5 架增长至 30 架的过程中, DMA 在 0.89 以上; 当集群规模由 35 架扩展至 50 架的中大型集群时, DMA 仍能保持在 0.67 之上, 验证了算法的泛化性能. 相比之下, 最近邻方法的 DMA 随集群规模增大迅速衰减至 0.10 以下, 说明仅依赖原始特征间的欧氏距离进行硬判决, 在目标密集分布时极易发生错配. SVM 性能优于最近邻方法, 但同样受限于浅层模型能力, DMA 随目标数增多而大幅下滑. Transformer 与 CNN 虽借助深度架构提升了特征表征能力, 但缺乏应对非欧空间问题的适应性, DMA 下降明显. GAT 和 GCN 在利用邻域聚合挖掘图结构关系上具备优势, 然而在面对大规模集群时, 邻域信息的过度平滑及固定邻接矩阵的适应性不足, 导致其 DMA 低于 P-C CMCC. 总体而言, 大规模集群带来的混叠问题与候选匹配的指数级膨胀, 仍是当前研究领域内难以逾越的性能瓶颈. 尽管 P-C CMCC 通过双域深度特征协同取得了领先优势, 但其性能根本上受限于固有的物理分辨极限与计算复杂度约束, 方法架构创新难以突破由规模效应引发的内在瓶颈.

(b) 双域匹配均方根误差 (DRMSE). 随着集群规模增大, 目标间维度与结构复杂度的增加使得双域匹配偏差放大, P-C CMCC 方法与其他基线算法的 DRMSE 均显著上升. P-C CMCC 方法的 DRMSE 始终保持在最低水平, 在 50 架次规模下 DRMSE 为 81.13 m, 展现了本方法的泛化性. 最近邻方法的 DRMSE 始终处于极高区间, 且几乎不受规模改变的影响, 原因是其极低的特征区分度导致方案失效. SVM 的 DRMSE 居中, 但同样难以抑制目标增多带来的累积偏差. Transformer 和 CNN 等深度模型在 5 架至 30 架规模时误差抑制较好, 但 30 架以上后 DRMSE 陡升. GAT 与 GCN 借助图结构关联, 其 DRMSE 在中小规模下较低, 但面对 40 架以上的大规模集群, 其注意力机制和邻域聚合在拥挤场景下易受到无关节点的干扰, DRMSE 亦快速攀升.

(c) 处理耗时. P-C CMCC 方法与其他基线算法的处理耗时普遍随集群规模增大而增长, 但增长速率存在显著差异. P-C CMCC 算法的处理耗时与集群规模呈现显著的强正相关增长, 在集群规模为 5 架时耗时为 11.92 ms, 50 架时处理耗时为 49.68 ms, 分析原因是随着集群规模的增大, P-C CMCC 算法以承担较高的计算时间开销为代价, 换取大规模集群下的高 DMA 和低 DRMSE. 最近邻、SVM、Transformer、GAT 及 GCN 等基线

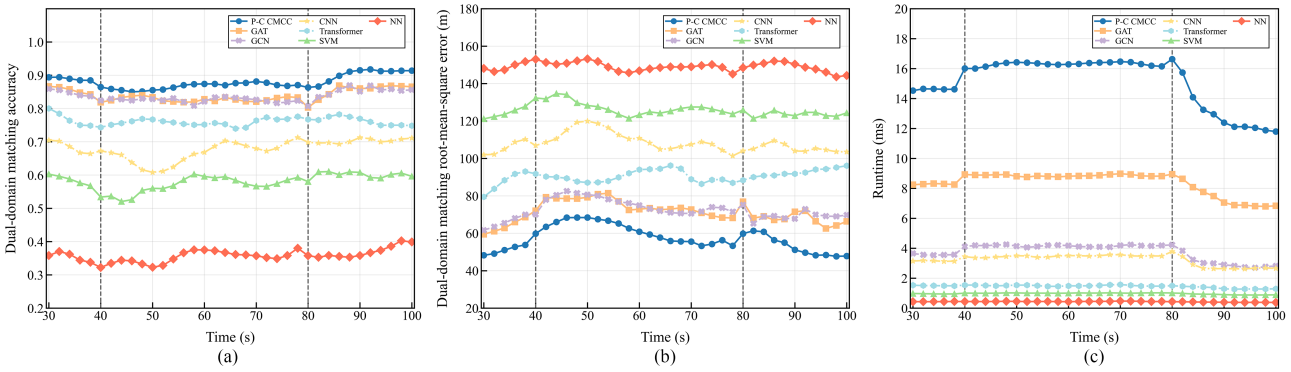

图 5　(网络版彩图) 集群规模动态变化下的算法性能对比. (a) 双域匹配准确率; (b) 双域匹配均方根误差; (c) 处理耗时.

Figure 5　(Color online) Performance comparison of algorithms under dynamic swarm scale variation. (a) Dual-domain matching accuracy; (b) dual-domain matching root-mean-square error; (c) runtime.

算法的处理耗时在整个规模变化区间内始终维持在低水平, 且几乎不受目标数量增加的影响. GAT 虽为深度模型, 但其运行时间随规模扩展呈现温和上升, 远低于 P-C CMCC.

(3) 集群规模动态变化下的算法适应性. 为验证 P-C CMCC 方法在真实动态环境下的适应性, 本小节在 SNR 为 5 dB 下, 设置集群规模动态变化场景, 评估本文方法针对集群时变特性的性能表现. 本小节设置无人机集群规模由初始 20 架, 在 40 s 时刻增加至 25 架, 在 80 s 时刻减少至 15 架. 每个规模下随机生成 50 个独立测试样本, 实验结果取统计平均值. P-C CMCC 方法与基线对比算法在集群规模动态变化下的评估指标结果如图 5 所示.

(a) 双域匹配准确率 (DMA). 受集群规模在 40 s 进入 (由 20 架增加至 25 架) 与 80 s 离开 (由 25 架减少至 15 架) 的影响, 各算法的匹配准确率均产生相应波动. 本文所提 P-C CMCC 方法在动态变化的全程始终维持最高 DMA, 且在集群出现规模突变后能迅速收敛至新稳态, 验证了算法的动态适应能力. 最近邻方法的 DMA 始终处于极低, 无法难以满足时变场景需求. SVM 的 DMA 虽有提升, 但本质是浅层模型, 难以有效应对由规模突变带来的复杂结构耦合问题, 导致 DMA 提升有限. Transformer 与 CNN 的双域匹配准确率优于 NN 和 SVM, 然而二者难以适配非欧拓扑结构问题, 在集群规模突变的情况下适应性不足. GCN 与 GAT 在集群规模突变时波动幅度较大, 主要原因是仅采用图结构模型, 在应对非合作集群的动态演化时存在结构敏感性, 无法有效解决集群的三重尺度问题及非线性映射问题.

(b) 双域匹配均方根误差 (DRMSE). DRMSE 在集群规模突变时各算法均出现同步波动, 原因是集群规模的动态调整改变了目标空间分布, 节点加入或离开引起的特征偏移. P-C CMCC 方法在集群规模突变过程中始终保持最低的 DRMSE. 在 40 和 80 s 的规模瞬变期, 其 DRMSE 最高为 68.41 m, 经过短暂震荡后迅速恢复平稳, 尤其是集群规模减小的情况下恢复更快速, 充分证明了 P-C CMCC 方法的适应性. 最近邻方法几乎未随集群规模变化产生波动, 且 DRMSE 一直处于最高水平, 主要原因是该算法仅依靠原始特征空间的欧氏距离进行硬判决, 缺乏对集群规模变化引起特征偏差的有效补偿. SVM 的 DRMSE 低于 NN, 原因是最大间隔和核函数虽提升了抗噪性, 但浅层模型在面对集群规模变化问题时, 误差拟合能力十分有限. Transformer 与 CNN 因缺乏对动态图拓扑关系的先验适应性, 在面对集群规模突变时 DRMSE 曲线依然存在明显震荡. GAT 和 GCN 方法在 40 和 80 s 集群规模变化时, 其 DRMSE 曲线呈现较明显的波动, 原因是它们主要依靠网络性能挖掘双域特征, 无法有效捕捉双域特征之间的时变耦合关系, 因而其 DRMSE 仍高于 P-C CMCC.

(c) 处理耗时. 集群规模动态变化下各算法的处理耗时呈现出分化特征. P-C CMCC 的处理耗时与集群规模呈现显著的强正相关性, 在集群规模为 20 架时, 其处理耗时平均为 14.61 ms; 在规模突增至 25 架时, 处理耗时跃升至平均 16.31 ms; 在规模骤减至 15 架后, 耗时回落至平均 12.19 ms, 表明 P-C CMCC 在获取最优 DMA 和最低 DRMSE 的同时, 其计算代价随规模而增长. 最近邻方法, SVM, Transformer 以及 GCN 的基线算法, 其处理耗时始终处于较低水平. 分析原因是基线方法依赖较浅的网络结构或简单的度量公式, 其计算时间在整个

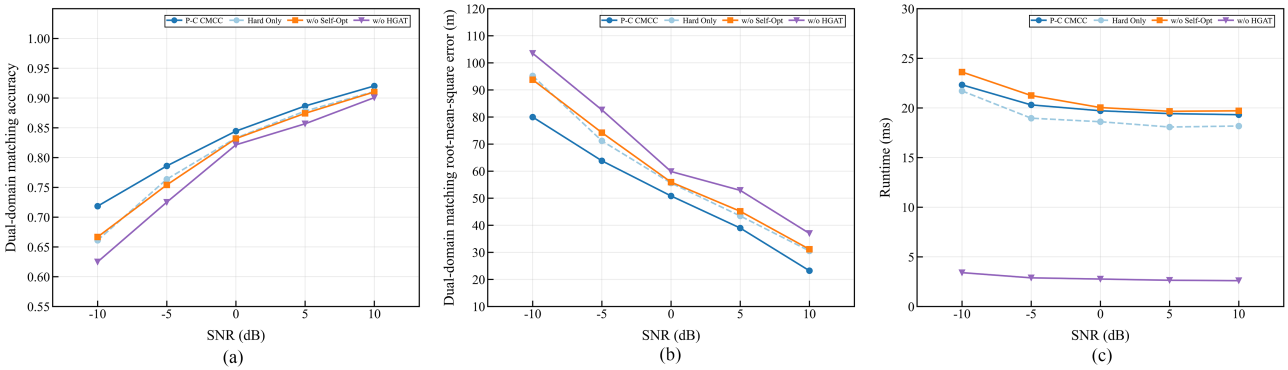

图 6　(网络版彩图) 不同接收信噪比下的消融实验性能对比. (a) 双域匹配准确率; (b) 双域匹配均方根误差; (c) 处理耗时.

Figure 6　(Color online) Performance comparison of ablation experiments under different received SNR conditions. (a) Dual-domain matching accuracy; (b) dual-domain matching root-mean-square error; (c) runtime.

动态过程中基本一致, 集群规模增减的影响较小. GAT 因注意力机制的引入, 其计算耗时处于中低水平, 并随规模变化呈现出一定程度的波动, 但整体耗时及波动程度显著低于 P-C CMCC.

5.2.2　消融实验

为定量评估 P-C CMCC 算法中跨模态异质图注意力网络、软匹配与硬匹配两阶段求解策略和网络训练参数自智优化等核心组件对方法的贡献程度, 设计如下四组消融实验: 完整方法 (P-C CMCC)、移除跨模态异质图注意力网络替换为特征拼接 (w/o HGAT)、移除软硬匹配两阶段求解策略仅采用匈牙利算法的硬匹配 (Hard Only)、移除网络训练参数自智优化将超参数固定为静态经验值 (w/o Self-Opt). 为充分模拟实际测试的随机性, 每个场景随机生成 50 个独立测试样本, 最终实验结果为所有测试样本的统计平均值.

(1) 不同通信接收信噪比下的消融实验对比. 设置在 20 架集群规模下, 选取典型的接收信噪比梯度, 分别为 −10、−5、0、5 和 10 dB, 以涵盖从极恶劣到优良的信道条件. P-C CMCC 方法与消融对比在不同接收信噪比条件下的评估指标结果如图 6 所示.

(a) 双域匹配准确率 (DMA). 各消融模型随着 SNR 的提升, DMA 均呈明显的上升趋势, 原因是信道条件的改善使通信域信号质量更高. 与此同时, 在所有 SNR 下, 完整 P-C CMCC 方法始终保持着最高的 DMA, 体现了各核心模块联合作用的综合优势. w/o HGAT 模型的性能衰减最为显著, 在 SNR 为 −10 dB 的恶劣环境下 DMA 仅为 0.62, 原因是将异质图注意力网络替换为简单的特征拼接与等权聚合后, 模型无法有效聚合关键邻域特征, 导致大量误匹配发生. Hard Only 模型的 DMA 低于完整方法, 尤其在低 SNR 下差异较大, 主要原因是去除了软匹配阶段的概率过渡直接进行硬判决, 容易在特征模糊时陷入局部极值, 从而引发双域匹配错误. w/o Self-Opt 模型在 −10 dB 下性能衰退较大, 但随着 SNR 提升至 0 dB 以上, 其 DMA 迅速回升至接近完整方法, 说明参数自智优化在应对极端恶劣信道环境时起着关键的泛化调节作用.

(b) 双域匹配均方根误差 (DRMSE). 各消融模型随着 SNR 的提升, DRMSE 呈明显的下降趋势. 完整 P-C CMCC 方法在所有 SNR 下均保持最低的 DRMSE, 展现出优异的抗噪稳定性. w/o HGAT 的 DRMSE 在各 SNR 下均为最大, 在 −10 dB 时误差为 103.50 m, 在 10 dB 时仍保持在 37.00 m, 说明异质图注意力机制在跨域邻域特征聚合去噪与误差抑制中具有重要作用. Hard Only 缺少软匹配的排序与筛选环节, 误匹配产生的偏差无法在求解阶段得到纠偏, 因此 DRMSE 曲线始终位于完整方法之上. w/o Self-Opt 由于固定了网络训练参数, 在低信噪比下无法对噪声分布做出自适应响应, 因此其 DRMSE 在各 SNR 下均高于 P-C CMCC, 且在低 SNR 情况下更为显著. 三个消融实验共同证明了完整模型的误差抑制是由各模块协同发挥作用.

(c) 处理耗时. 各消融模型在处理耗时上呈现显著差异化特征. 由于算法的前向推理结构与 SNR 无关, 各耗时曲线在 −10~10 dB 内均保持水平直线, 其差异主要来源于网络架构的复杂度变化. P-C CMCC 完整模型的计算复杂度依次为跨模态异质图注意力网络、跨模态注意力计算、软匹配和硬匹配两阶段求解. P-C CMCC 完整

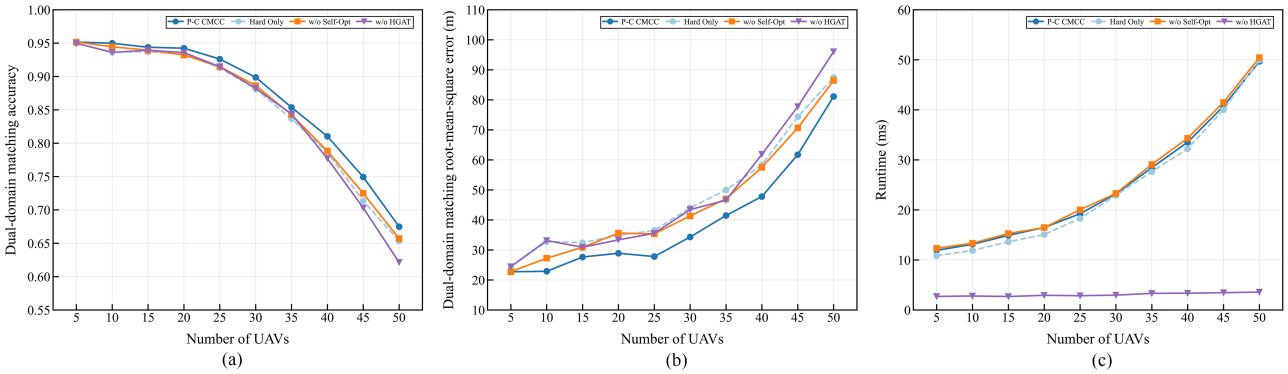

图 7　(网络版彩图) 不同集群规模下的消融实验性能对比. (a) 双域匹配准确率; (b) 双域匹配均方根误差; (c) 处理耗时.

Figure 7　(Color online) Performance comparison of ablation experiments under different swarm scales. (a) Dual-domain matching accuracy; (b) dual-domain matching root-mean-square error; (c) runtime.

模型处理耗时最高, 平均为 20.22 ms. w/o Self-Opt 由于前向推理结构与完整方法基本一致, 耗时与 P-C CMCC 持平. Hard Only 因去除软匹配阶段的计算开销, 耗时略有降低, 平均为 19.11 ms. w/o HGAT 的耗时急剧下降, 平均仅为 2.86 ms, 说明跨模态异质图注意力网络虽然大幅提升了算法抗噪鲁棒性, 但也为此承担了主要的计算代价. 在计算资源受限的实时系统中, 可进行精度与效率的折中部署.

(2) 不同集群规模下的消融实验对比. 本小节设置接收信噪比为 5 dB, 分析无人机集群规模梯度从 5 至 50 架, 以涵盖从微型到中大型的典型非合作目标数量变化. 不同集群规模下的消融实验评估结果如图 7 所示.

(a) 双域匹配准确率 (DMA). 随着无人机集群规模增加, 双域特征空间中的样本密度与特征混叠现象不断加剧, 候选匹配数呈指数级增长, 各消融模型的准确率均呈现明显的下降趋势. 完整方法 P-C CMCC 始终保持最高 DMA, 在 50 架规模时 DMA 仍稳定在 0.67, 展现了多模块协同带来的泛化增益. w/o HGAT 的性能衰减最为剧烈, 在 50 架规模时 DMA 已降至 0.62, 原因是将跨模态异质图注意力网络替换为简单特征拼接后, 模型丧失了对高密度邻域节点的有效聚合与特征挖掘能力. Hard Only 的 DMA 略低于完整方法 P-C CMCC, 且在大规模时 DMA 差异扩大, 表明在面对模糊交织的相似度空间时, 直接进行硬匹配缺乏有效容错过渡. w/o Self-Opt 的性能在小规模下与完整方法 P-C CMCC 接近, 但随着规模扩大至 30 架以上, 静态经验超参数无法自适应调整, 其 DMA 明显下滑, 验证了参数自智优化在大规模集群场景中提升泛化性的关键作用.

(b) 双域匹配均方根误差 (DRMSE). DRMSE 随集群规模增加呈上扬态势, 完整方法 P-C CMC 始终保持最低的 DRMSE, 在 50 架时仍控制在 81.13 m, 体现了多模块协同下的误差抑制能力. w/o HGAT 的 DRMSE 抬升最为显著, 在 50 架时, 其 DRMSE 已突破 96.04 m, 证明异质图注意力网络对关键邻域特征的噪声剔除与误差抑制具有关键作用. Hard Only 因缺少软匹配阶段的概率排序与容错平滑, 其 DRMSE 随目标数量呈线性累积, 在 50 架时 DRMSE 约为 87.42 m. w/o Self-Opt 受限于静态固定参数, 无法根据高维大规模特征空间下的梯度分布进行动态调整, DRMSE 曲线在规模扩大后逐步偏离完整方法, 在 50 架时 DRMSE 为 86.43 m.

(c) 处理耗时. 各消融模型在处理耗时上出现显著分化, 完整方法 P-C CMCC、Hard Only 以及 w/o Self-Opt 三者的耗时曲线高度重合, 说明软匹配阶段与自智优化环节的额外计算开销极小, 主要来自跨模态异质图注意力网络, 因此 w/o HGAT 的耗时基本脱离随规模变化的轨迹, 但牺牲了关键的匹配精度与误差抑制能力.

(3) 集群规模动态变化下的消融实验对比. 设置接收 SNR 为 5 dB, 集群初始规模为 20 架, 在 40 s 时突增为 25 架, 在 80 s 内骤减为 15 架, 以模拟真实非合作环境中节点动态进入与离开引起的时变扰动. 集群规模动态变化下的消融实验评估结果如图 8 所示.

(a) 双域匹配准确率 (DMA). 各消融模型的 DMA 与集群动态变化呈现显著的时序联动性. 完整方法 P-C CMCC 始终维持最高 DMA, 在 20 架突增至 25 架时, 其 DMA 波动下降 4.71% 后保持稳定; 25 架骤减至 15 架后, 完整方法 P-C CMCC 的 DMA 波动上升 7.06% 后保持稳定. Hard Only 在瞬变期的 DMA 表现出最明显下降, 降幅达 11.21%, 且恢复速度慢于其他模型, 缺乏软匹配的概率容错机制, 面对特征分布偏移时直接硬判决易

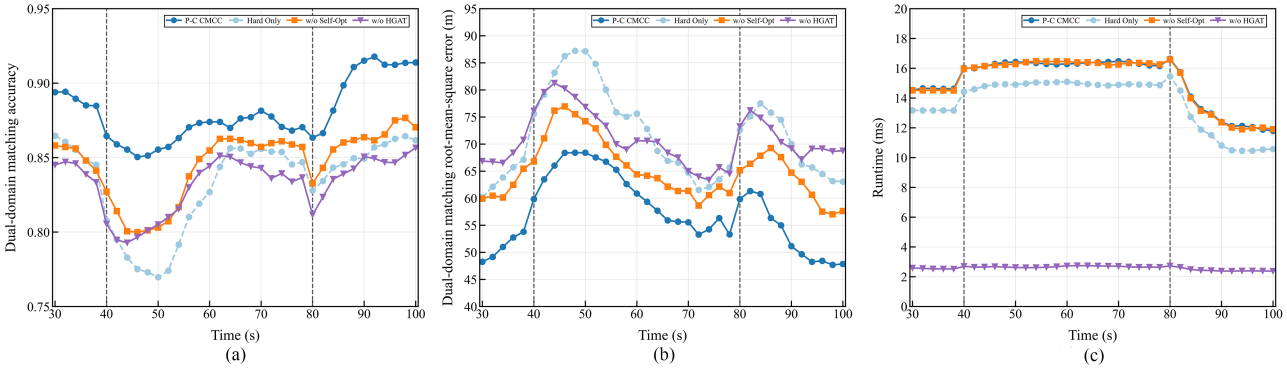

图 8　(网络版彩图) 集群规模动态变化下的消融实验性能对比. (a) 双域匹配准确率; (b) 双域匹配均方根误差; (c) 处理耗时.

Figure 8　(Color online) Performance comparison of ablation experiments under dynamic swarm scale variation. (a) Dual-domain matching accuracy; (b) dual-domain matching root-mean-square error; (c) runtime.

引发关联混乱. w/o HGAT 的 DMA 波动也较为剧烈, 在无人机加入后 DMA 波动下降 6.15%, 简单特征拼接无法在拓扑突变的动态邻域中提取关键特征, 导致邻域错误聚合. w/o Self-Opt 在规模突变时 DMA 波动也较为剧烈, 说明静态超参数在拓扑扰动下难以自适应调整, 影响了模型的快速收敛与稳定.

(b) 双域匹配均方根误差 (DRMSE). 完整方法 P-C CMCC 始终保持最低 DRMSE, 且面对集群规模突变引起的特征分布偏移, 具备更快的收敛与抑制能力, DRMSE 整体波动最小. Hard Only 与 w/o HGAT 的 DRMSE 抬升最为显著. Hard Only 因移除了软匹配阶段对相似度空间的容错与概率筛选机制, 硬判决直接放大了误匹配偏差, 使其 DRMSE 大幅高于其他模型, 尤其是在规模突增时 DRMSE 增幅 34.62%. w/o HGAT 将异质图注意力网络替换为特征拼接, 丧失了对动态邻域中有效信息的筛选能力; 拓扑突变时大量无关邻域特征被错误聚合, 尤其在规模突增后 DRMSE 增幅 23.08%. w/o Self-Opt 由于依赖静态经验参数, 无法自适应调整, 其 DRMSE 较完整 P-C CMCC 高.

(c) 处理耗时. 计算处理耗时主要取决于无人机集群规模, 除 w/o HGAT 外各模型的耗时曲线趋势基本一致. 完整方法 P-C CMCC 在 20 架稳态时耗时平均为 14.61 ms, 规模突增至 25 架时上升至平均 16.31 ms, 缩减至 15 架后回落到 11.99 ms. w/o Self-Opt 耗时与完整方法完全一致, 说明自智优化环节在整体方法的计算耗时中占比基本可忽略. Hard Only 由于移除了软匹配阶段的计算开销, 其耗时曲线有较为明显的下降, 降幅为完整方法 P-C CMCC 的 7.52%. w/o HGAT 则因移除跨模态异质图注意力网络这一核心计算模块, 其处理耗时最低, 平均为 2.89 ms, 且几乎不随集群规模突变而变化, 说明该模块承担了主要计算代价.

6　结论

本文针对 Agentic AI-RAN 架构在真实部署中面临的非合作无人机集群侵扰问题, 为提升对其的精准态势感知和行为状态认知能力, 提出了一种基于物理域和通信域的跨模态协同认知方法 (P-C CMCC).

对比实验表明, 在不同接收信噪比、集群规模大小和集群规模突变场景下, 所提 P-C CMCC 方法在双域匹配准确率和双域匹配均方根误差上优于基线方法, 尤其在低 SNR、大规模集群与集群规模突变情况下优势更为突出, 验证了所提 P-C CMCC 方法的鲁棒性、泛化性和适应性. 在处理耗时方面, P-C CMCC 高于其他基线方法, 以计算开销换取性能提升. 消融实验进一步验证了所提 P-C CMCC 方法各核心模块的关键贡献, 移除跨模态异质图注意力网络在低 SNR 与大规模集群场景下性能急剧下降, 说明该模块在邻域信息聚合与深层特征提取中的重要作用; 移除软匹配仅保留硬匹配在集群规模动态变化场景下性能急剧下降, 表明软匹配阶段在相似度空间容错与误匹配抑制方面不可或缺; 移除网络训练参数自智优化在低 SNR 与集群规模动态变化场景下性能急剧下降, 揭示了自适应参数优化对提升模型泛化性与收敛稳定性的重要作用.

综上所述, P-C CMCC 方法在低 SNR, 大规模集群规模和集群规模动态时变下均展现出性能优势, 为 Agentic

AI-RAN 应对非合作集群侵扰提供了可靠的感知认知基础. 未来工作将进一步拓展至物理行为机动多样性、通信信号失真干扰的复杂场景, 同时考虑模型轻量化与边缘计算部署等.

参考文献

1 Cui Q M, You X H, Wei N, et al. Overview of AI and communication for 6G network: fundamentals, challenges, and future research opportunities. Sci China Inf Sci, 2025, 68: 171301
2 You X H, Huang Y M, Zhang C, et al. When AI meets sustainable 6G. Sci China Inf Sci, 2025, 68: 110301
3 Navidan H, Cheraghinia M, Fontaine J, et al. Toward autonomous O-RAN: a multi-scale agentic AI framework for real-time network control and management. ArXiv:2602.14117
4 He Z, Luo Y, Liu X, et al. Agentic AI-RAN: enabling intent-driven, explainable and self-evolving open RAN intelligence. ArXiv:2602.24115
5 Li Y, Wang T, Quek T Q. Empowering UAVs with large models: prospects and challenges. Chin J Aeronaut, 2025, 39: 103860
6 Kumari N, Lee K, Ranaweera C. Visually extracting the network topology of drone swarms. Rob Auton Syst, 2026, 198: 105313
7 Liang J, Qi M, Gu K, et al. The structure inference of flocking systems based on the trajectories. Chaos, 2022, 32: 101103
8 Zhang X, Hu W, Cui K, et al. A network-group target state and network topology estimation method based on signals containing delay, doppler and address. Remote Sens, 2024, 16: 1275
9 Liu Z, Wang W, Ding G, et al. Topology sensing of non-collaborative wireless networks with conditional granger causality. IEEE Trans Netw Sci Eng, 2022, 9: 1501–1515
10 Dong G, Sun N, Yan M, et al. Robustness of coupled networks with multiple support from functional components at different scales. Chaos, 2024, 34: 043122
11 Cui Y, Liang Y, Luo Q, et al. Resilient consensus control of heterogeneous multi-UAV systems with leader of unknown input against byzantine attacks. IEEE Trans Automat Sci Eng, 2025, 22: 5388–5399
12 Lim R, Gupta A, Ong Y S, et al. Non-linear domain adaptation in transfer evolutionary optimization. Cogn Comput, 2021, 13: 290–307
13 Mu L, Duan T, Wu J, et al. Motion behaviour based communication range estimation of adversarial drone swarms. IEEE Trans Netw Sci Eng, 2025, 12: 1953–1966
14 Yang Z, Cui Y, Li Y. Transformer-based cooperative UAV encirclement policies under uncertainty in low-altitude wireless networks. IEEE Trans Cogn Commun Netw, 2026, 12: 3525–3537
15 Chatzistefanidis I, Nikaein N, Leone A, et al. AGORAN: an agentic open marketplace for 6G RAN automation. Comput Netw, 2026, 275: 111927
16 Kaleem Z, Orakzai F A, Ishaq W, et al. Emerging trends in UAVs: from placement, semantic communications to generative AI for mission-critical networks. IEEE Trans Consum Electron, 2025, 71: 7412–7438
17 Chergui H, Rezazadeh F, Debbah M, et al. A tutorial on cognitive biases in agentic AI-driven 6G autonomous networks. IEEE Open J Commun Soc, 2025, 7: 5214–5240
18 Su H, Luo J, Liu C, et al. A survey on autonomy-induced security risks in large model-based agents. IEEE Trans Pattern Anal Mach Intell, 2026, 48: 11382–11401
19 Liu Z, Zou Y, Hu Z, et al. Research on multi-modal fusion detection method for low-slow-small uavs based on deep learning. Drones, 2025, 9: 852
20 Chen J, Huang J, Zhang Z, et al. ATM-NET: a lightweight multimodal fusion network for real-time UAV-based object detection. Drones, 2026, 10: 67
21 Liu H, Chen Z, Du Z, et al. 3D multi-object tracking driven by multi-level association and intelligent filtering. IEEE Trans Intell Transp Syst, 2026, 27: 1442–1457
22 Zhuang S, Li G, Wu Q, et al. CGATracker: correlation-aware graph alignment for referring multi-object tracking. IEEE Trans Circuits Syst Video Technol, 2025, 35: 11337–11349
23 Dabbabi K, Delleji T. Graph neural network-tracker: a graph neural network-based multi-sensor fusion framework for robust unmanned aerial vehicle tracking. Vis Comput Ind Biomed Art, 2025, 8: 18
24 Zahra I, Wu Y, Alhasson H F, et al. Dynamic graph neural networks for UAV-based group activity recognition in structured team sports. Front Neurorobot, 2025, 19: 1631998
25 Zhang S Y, Zhang S Y, Yuan W J, et al. Personalizing rate-splitting in vehicular communication via large multi-modal model. Sci China Inf Sci, 2025, 68: 170305
26 Qin X, Li L, Pang G, et al. Heterogeneous graph fusion network for cross-modal image-text retrieval. Expert Syst Appl, 2024, 249: 123842
27 Lu Y, Yang K, Yang D, et al. Graph learning-driven multi-vessel association: fusing multimodal data for maritime intelligence.

IEEE Trans Intell Transp Syst, 2026, 27: 5739–5754
28 Hong R, Yang J, Cai Z, et al. Cross-modal contrastive learning with multi-hierarchical tracklet clustering for multi object tracking. Pattern Recogn Lett, 2025, 192: 1–7
29 Zhu W, Liu S, Liu C. Learning multimodal word representation with graph convolutional networks. Inf Process Manage, 2021, 58: 102709
30 Kang H H, Ahn C K. Distributed finite memory online learning strategy for multi-UAV systems with neural networks. IEEE Trans Ind Electron, 2025, 72: 919–927
31 Si M, Zhang W, Qi B, et al. Meta-CT: a meta-learning-based spectrum sensing scheme using CNN-transformer for low-altitude intelligent networks. IEEE Trans Veh Technol, 2026, 75: 20602–20616
32 Yang Z, Du H, Niyato D, et al. Revolutionizing wireless networks with self-supervised learning: a pathway to intelligent communications. IEEE Wireless Commun, 2025, 32: 108–115
33 Yang Z, Li N, Cai L, et al. Transfer learning of non-cooperative drones trajectory prediction and intent recognition for unknown threats. Adv Eng Inf, 2026, 76: 105070
34 Liu G Y, Xi R Y, Jiang T, et al. Feasibility study of cooperative sensing: radar cross section, synchronization, cooperative cluster, performance and prototype. Sci China Inf Sci, 2025, 68: 150302
35 Zhang D, Cui Y, Cao X, et al. Integrated sensing and communications over the years: an evolution perspective. IEEE Commun Surv Tutorials, 2026, 28: 5014–5048
36 Hong T, Zhou C, Kadoch M, et al. Improvement of UAV tracking technology in future 6G complex environment based on GM-PHD filter. Electronics, 2022, 11: 4140
37 Li Y, Guo L, Huang X. Research and application of multi-target tracking based on GM-PHD filter. OPJ, 2020, 10: 125–133
38 Yang Z, Du W, Zhang X, et al. A spatiotemporal graph reasoning approach for pursuit-evasion game with communication limits. IEEE Trans Veh Technol, 2026, 75: 2069–2085
39 Bao Y J, Zhu X K, Yan W P, et al. Deep Learning on Graphs. Beijing: Tsinghua University Press, 2022 [包勇军, 朱小坤, 颜伟鹏, 等. 图深度学习从理论到实践. 北京: 清华大学出版社, 2022]
40 Deng H, Zhang Z, Li C, et al. Spatiotemporal information complementary modeling and group relationship reasoning for group activity recognition. J Supercomput, 2024, 80: 21790–21810
41 Cui Y, Cao X, Zhu G, et al. Edge perception: intelligent wireless sensing at network edge. IEEE Commun Mag, 2025, 63: 166–173
42 Zhao Y, Lv Q, Guan Y L, et al. Free space spot-beamforming for IoT multi-user near-orthogonal overlay communications enhanced by OAM waves and reconfigurable meta-surface. EMScience, 2025, 3: 0110022
43 Zou Y, Wu W, Fan J, et al. Constrained weighted least-squares algorithms for 3-D AOA-based hybrid localization. IEEE Open J Signal Process, 2024, 5: 436–448
44 Xia M, Wang J, Shan Y, et al. HPPS: a head-mounted pedestrian positioning system integrating fisheye camera/RTK/PDR with factor graph optimization. IEEE Internet Things J, 2025, 12: 50708–50721
45 Yang Z, Shi C, Guan Y. An optimal-transport-based multimodal big data clustering. Electronics, 2025, 14: 666
46 Sun R, Huang G, Wang X, et al. Text-augmented multi-modality contrastive learning for unsupervised visible-infrared person re-identification. Image Vis Comput, 2024, 152: 105310
47 Wang J, Luo M, Yue Z, et al. Fine-grained indoor-outdoor scene recognition: leveraging spatiotemporal characteristics from multimodal sensors via CNN-ViT-Tabtransformer fusion. IEEE Sens J, 2026, 26: 7362–7378
48 Zhao J, Zhang J, Li D, et al. Vision-based anti-UAV detection and tracking. IEEE Trans Intell Transp Syst, 2022, 23: 25323–25334

Cross-modal collaborative cognition method for UAV swarm in Agentic AI-RAN

Huizhen SHU[1], Qihang YAO[1], Hui XU[1], Yumeng LI[1] & Yuanhao CUI[2*]

1. *School of Electronic and Information Engineering, Beihang University, Beijing 100191, China*
2. *School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China*
* Corresponding author. E-mail: yuanhao.cui@bupt.edu.cn

Abstract The Agentic AI-RAN faces grave security risks from intrusive non-cooperative UAV swarms. The core difficulty lies in accurately perceiving their behavioral states. Existing solutions adopt unimodal sensing data acquired from either the physical or communication domain, which is inadequate to achieve joint cross-modal cognition of coupled physical-motion and communication-interaction behaviours. This paper presents a physical-communication cross-modal collaborative cognition (P-C CMCC) approach. It builds dual-domain subgraphs for feature representation at the bimodal feature stage. At the bimodal node stage, a bimodal heterogeneous graph attention network captures intra-domain spatial synergies among nodes, while the attention mechanism projects the bimodal features into a shared interaction space, where a cross-domain similarity matrix is constructed for node association. At the global swarm stage, a two-stage soft-and-hard matching scheme is adopted to derive the globally optimal bimodal correspondence across domains, thereby constructing a bimodal collaborative cognition graph that achieves joint precise cognition of the coupled physical-motion and communication-interaction behaviours of the swarm. A total loss function is formulated to enable end-to-end self-optimizing training. Experimental results show that across evaluation scenarios featuring varying communication SNRs, swarm sizes and abrupt variations in swarm scale, the proposed method surpasses the baseline and ablation models in dual-domain matching accuracy and root-mean-square error, thereby effectively verifying its superiority.

Keywords non-cooperative UAV swarm, cross-modal collaborative cognition, coupled physical-motion and communication-interaction, bimodal nodes association, two-stage soft-and-hard matching, self-optimizing training