

面向低空无人机集群的ILAC网络中联合 用户关联与资源优化方法

杨卓佳¹, 苏 伟^{1*}, 杨 斌², 马煜东³, 王文迪¹

(1. 北京交通大学电子信息工程学院, 北京 100044; 2. 滁州学院计算机与信息工程学院, 安徽滁州 239000;
3. 北京航空航天大学人工智能学院, 北京 100083)

摘 要: 通信定位一体化(Integrated Localization And Communication, ILAC)网络通过通信与定位的深度融合,能够在提供通信服务的同时实现高精度位置感知,成为未来无线网络发展的重要方向之一。然而,在应急救援场景下,传统依赖地面基站的ILAC网络易受到地面基础设施损毁影响,将面临通信覆盖中断、链路质量恶化和定位精度下降等问题。为此,本文构建了一种低空无人机集群辅助的ILAC网络,为受灾区域用户提供临时与分布式的通信与定位服务,并提出了一种基于分布式架构的无人机关联、频谱与功率联合优化方法,以保障高效的通信与定位服务,支持应急救援等关键任务的顺利执行。首先,本文建立基于双边双向测距的用户定位模型,描述无人机空中锚点与地面用户之间的测距关系及其对定位精度的影响。之后,构建以最大化最小用户吞吐量为目标,定位精度、无人机关联、频谱与功率等为约束的联合优化问题。针对此复杂优化问题,我们将其划分为无人机关联与功率-子信道优化两个子问题。针对第一个子问题,提出三维快速空中锚点选择方法,通过空间划分降低空中锚点搜索空间,实现用户通信接入点与定位锚点的快速关联;针对第二个子问题,提出基于多智能体近端策略优化的联合功率控制与子信道分配算法,利用分布式多智能体学习实现功率与子信道的协同决策。通过与多种基线算法仿真对比,结果表明本文所提出方法能够在保障定位精度约束的同时有效提升最小用户吞吐量,提高用户间通信公平性,并在相同网络场景设置下获得更优的定位性能,从而为应急场景下实现高效通信与可靠定位服务提供有力支持。

关键词: ILAC网络;无人机集群;无人机关联;资源优化;分布式强化学习;应急通信

基金项目: 雄安新区科技创新专项(No.2025XAGG0055);国家自然科学基金(No.62372076)

中图分类号: TN929.5

文献标识码: A

文章编号: 0372-2112(2026)06-2676-16

电子学报URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20260403

Joint User Association and Resource Optimization for Low-Altitude UAV Swarm-Based ILAC Networks

YANG Zhuojia¹, SU Wei^{1*}, YANG Bin², MA Yudong³, WANG Wendi¹

(1. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. School of Computer and Information Engineering, Chuzhou University, Chuzhou, Anhui 239000, China;

3. School of Artificial Intelligence, Beihang University, Beijing 100083, China)

Abstract: Integrated localization and communication (ILAC) networks deeply integrate communication and positioning functions, enabling efficient communication transmission services while providing reliable positioning services. Owing to their ability to support both information exchange and positioning perception within a unified wireless network, ILAC networks have emerged as a key direction for the evolution of future wireless systems. However, in emergency rescue scenarios, traditional ILAC networks that rely on terrestrial base stations may become unavailable or severely degraded due to damage to ground infrastructure. Such infrastructure failures can result in communication coverage interruption, link quality deterioration, insufficient localization anchors, and reduced positioning accuracy, making it difficult to continuously support mission-critical services in disaster-affected areas. To address this issue, this paper proposes an unmanned aerial vehicle (UAV)-swarm-enabled ILAC network, in which multiple UAVs are deployed as flexible aerial platforms to provide temporary, distributed, and reconfigurable communication and localization services for ground users in disaster-affected regions. Furthermore, a distributed framework is proposed for the joint optimization method of UAV association, spectrum allocation, and power control to ensure efficient communication and positioning services, thereby supporting the successful execution of critical tasks such as emergency rescue. First, a user positioning model based on double-sided two-way ranging is es-

established to describe the ranging relationship between aerial UAV anchors and ground users, as well as the impact of ranging measurements and anchor selection on positioning accuracy. On this basis, the coupling relationship among communication access, positioning anchor association, spectrum sharing, and transmit power allocation is further characterized. Then, a joint optimization problem is formulated with the objective of maximizing the minimum user throughput, while considering positioning accuracy requirements, UAV association constraints, sub-channel allocation constraints, and power constraints. Since the formulated problem involves both discrete association decisions and continuous resource allocation variables, and the communication and positioning performances are strongly coupled through shared UAV anchor nodes and wireless resources, the problem is highly non-convex and difficult to solve directly by conventional centralized optimization methods. To tackle this complexity, the original problem is decomposed into two subproblems: UAV association, power control, and sub-channel allocation. For the first subproblem, a three-dimensional fast aerial anchor selection method is proposed. By introducing spatial partitioning, the proposed method reduces the search space of candidate aerial anchors and enables rapid association between user communication access points and positioning anchors, thereby lowering the computational complexity of UAV association while preserving positioning performance. For the second subproblem, a multi-agent proximal policy optimization-based joint power control and sub-channel allocation algorithm is proposed. By leveraging distributed multi-agent learning, users can learn coordinated resource allocation policies under local network observations, so as to jointly determine spectrum and power configurations in dynamic multi-user scenarios. Simulation comparisons with multiple baseline algorithms demonstrate that the proposed method can effectively improve the minimum user throughput and enhance communication fairness among users while satisfying positioning accuracy constraints. Moreover, under the same network scenario settings, the proposed method achieves superior positioning performance, indicating its effectiveness in jointly balancing communication efficiency and positioning reliability. These results show that the proposed UAV swarm-enabled ILAC framework provides a feasible and effective solution for supporting efficient communication and reliable positioning services in emergency rescue scenarios.

Keywords: ILAC network; UAV swarm; UAV association; resource optimization; distributed reinforcement learning; emergency communications

Foundation Item(s): Science and Technology Innovation Program of Xiong'an New Area (No.2025XAGG0055); National Natural Science Foundation of China (No.62372076)

0 引言

通感一体化(Integrated Sensing And Communication, ISAC)网络通过在统一架构下深度融合通信与感知,成为未来无线通信的重要发展方向之一。定位作为感知的重要组成部分,通信定位一体化网络(Integrated Localization And Communication, ILAC)同样在6G等下一代通信网络发展中扮演重要角色,为智能交通、智慧城市、无人驾驶及应急救援等应用场景的快速发展提供关键技术支撑与基础设施保障^[1]。然而,以地面基站和全球导航卫星系统(Global Navigation Satellite System, GNSS)为主的传统定位方法,其可靠性容易受到地面设施与非视距(Non-Line-Of-Sight, NLOS)传播的影响。例如:受山脉和建筑物遮挡的偏远地区,环境遮挡常导致GNSS不可靠或拒止;或在灾害导致的地面网络瘫痪地区,基于基站(Base Station, BS)的定位和通信方法将受到严重限制,从而使得传统方法难以为地面用户提供稳定的通信与定位服务。尤其是在应急救援等关键场景中,高效的通信服务是保障应急信息传输、协同调度与指挥控制的基础,而可靠的定位能力则是实现人员搜救、

目标感知与任务精准执行的重要支撑,上述限制将直接影响任务执行的及时性与有效性。

随着低空经济的快速发展,无人机集群网络以其快速部署、空间灵活性和广域覆盖能力等特点,在保障灾害应急通信、增强复杂环境感知、支撑低空空域管理等方面展现出广阔前景^[2]。因此,基于无人机集群的ILAC网络为复杂应急场景下的高效通信与高精度定位提供了一种良好的解决方案。

目前,ILAC网络研究可以分为地面辅助的ILAC网络以及无人机辅助的ILAC网络。在地面辅助的ILAC网络中,研究主要通过优化波束赋形、可重构智能表面(Reconfigurable Intelligent Surface, RIS)相位控制、信道状态信息(Channel State Information, CSI)估计、子载波与功率分配等,提升系统定位精度、通信速率与可靠性。文献[3]推导了ILAC网络通信与定位性能上界,为资源分配提供了理论依据。文献[4]构建了ILAC的统一解析框架,分析了其在大规模网络中的覆盖、资源分配和通信定位性能。为了提高定位精度,文献[5]采用上行随机接入导频方法,同时完成用户检测、信道估计和多目标定位。文献[6]则联合优化波束赋形与RIS相位控制,通过传统数学方

法提高定位精度。更进一步地,文献[7]考虑现实场景中的NLOS、多用户干扰与不完美CSI影响,通过优化波束赋形,实现对多用户信号、信道与位置的同步估计,在不增加额外定位参考信号的前提下提升了定位精度。文献[8]联合优化了用户位置获取、通信功率分配与密集区域的资源调度,采用深度学习方法最大化感知效用函数。除了定位精度,通信性能也是网络关键性能之一。文献[9]通过联合优化上行ILAC中的CSI与到达角估计,同时提升通信可靠性与定位精度。文献[10]则是通过优化RIS配置、信道估计与最优编码器设计,提升了系统的整体性能。文献[11]引入了差分相位到达方法,联合优化调制、子载波和功率分配,以最大化通信速率与定位精度。类似地,文献[12]联合优化带宽与时间分配,通过迭代和启发式方法求解,提高了定位精度和通信数据速率。

然而,在应急场景下,受地面基础设施受损或失效的影响,地面辅助的ILAC网络往往难以满足救援任务对通信与定位的需求。为此,无人机辅助的ILAC网络为应急救援场景提供了一种新的解决方案。文献[13]研究了ILAC网络中无人机的部署问题,在满足地面用户的通信与感知需求下最小化无人机部署数量,并提出一种低复杂度的数学方法来求解该问题。文献[14]考虑NLOS传播对定位的影响,提出了一种基于城市物理环境地图的无人机部署优化方法,采用传统的数学方法来求解。文献[15]讨论了导频插入方式、无人机部署以及带宽与用户关联对通信与定位性能的影响,并分析了其可行的解决方案。然而,目前无人机辅助的ILAC网络研究较少,一些ISAC网络研究重点关注感知中的定位性能,从而为ILAC提供一定的参考。例如,文献[16]考虑了环境干扰的影响,通过混合到达时间差(Time Difference Of Arrival, TDOA)定位方法实现无人机集群的自定位与时钟同步,并推导了定位精度的理论下界。文献[17]通过优化无人机轨迹与设备间调度,基于差分进化方法,实现最大能耗的最小化,以辅助地面用户完成通信与定位任务。文献[18]优化了锚点选择,并分析了所提出方法在平均功率、最大功率以及定位性能方面的性能。文献[19]考虑车辆网场景下用户的高速移动,构建了基于多普勒效应的系统模型,结合TDOA与到达频率差实现车辆高精度定位,并通过差分进化算法优化无人机-车辆链路选择,在通信性能与感知精度间实现有效权衡。文献[20]讨论了资源有限场景下资源分配对通信与定位性能的影响,提出了一种基于分布式强化学习的方法为地面终端提供感知信息并最小化无人机能耗,联合优化无人机关联与功率分配。

综上所述,在复杂动态场景中,无人机辅助的ILAC网络可以作为重要补充设施,提供高效的通信与高精度定位服务。目前,无人机辅助的ILAC网络相关研究较少,主要集中于无人机部署^[13-15]、锚点选择^[17-20]、轨迹优化^[17]以及功率分配^[15, 20]等,在一定程度上提升了系统的通信与定位性能。然而,仍存在以下问题亟待解决。首先,部分研究依赖于地面基站实现定位^[13-14, 17-18],难以有效应对地面基础设施损毁的应急救援场景,从而难以保障救援任务的可靠执行。其次,尽管一些研究引入了无人机集群以提供通信与定位服务^[15-16, 19-20],但其大部分采用传统的数学方法求解^[16, 18-19],或未给出具体解决方法^[15],难以有效应用于网络频繁变化的复杂低空环境。文献[20]提出了一种基于多智能体Q-learning的方法优化关联与功率分配,但ILAC网络中通信与定位资源高度耦合,其未充分考虑频谱资源对系统整体性能的影响,而通信与定位功率以及频谱分配共同决定了系统性能,若仅针对单一资源维度进行优化,难以同时满足通信性能与定位需求,从而限制系统整体性能的进一步提升。并且,上述研究采用TDOA等方法^[15-16, 19-20],而在实际场景中接收端设备难以实现高精度时间同步,可能出现定位性能下降问题。

针对以上问题,本文提出了一种复杂应急场景下无人机集群辅助的ILAC网络联合用户关联与资源优化方法,通过双边双向测距(Double-Sided Two-Way Ranging, DS-TWR)定位方法降低系统对严格时钟同步的依赖,并引入分布式架构,采用多智能体深度强化学习(Multi-Agent Deep Reinforcement Learning, MADRL)实现无人机关联与功率、频谱资源分配优化,为地面基础设施受损以及GNSS受限条件下的通信与定位服务提供支撑,从而保障目标搜寻与指挥决策等应急救援任务的成功执行。具体贡献有如下3点。(1)本文构建了一种基于无人机集群的ILAC网络架构,为缺乏基站与GNSS支撑的应急场景提供高效通信与高精度定位服务。在此基础上,本文建立基于双边双向测距的用户定位模型,并以最大化最小用户吞吐量为目标,在满足用户定位精度等约束的条件下,建立面向通信与定位协同优化的无人机关联、功率控制与频谱资源分配联合优化模型。(2)本文将上述联合优化问题划分为无人机关联与功率-频谱分配两个子问题。针对无人机关联问题,充分考虑无人机的三维空间分布特征,本文提出了一种基于三维空间的快速空中锚点选择方法(Three-Dimensional Fast Aerial Anchor Selection, 3D-FAS),以实现用户空中锚点的高效匹配,降低计算规模并提升关联效率。针对功率与频谱分配问题,本文采用分布式框架,通过引入多智

能体近端策略优化(Multi-Agent Proximal Policy Optimization, MAPPO)算法,实现了用户的通信与定位功率以及频谱资源的自主决策,从而保障通信性能的公平性以及定位性能的可靠性。(3)本文将所提算法与多个基线算法进行了对比,验证了所提算法在通信与定位性能上的有效性。结果表明,本文所提算法能够更好地提升用户间通信的公平性,同时有效保障了用户定位精度。

1 系统模型

1.1 网络模型

本文提出了一种应急救援场景下无人机集群辅助的ILAC网络,如图1所示。该网络由一个包含 M 个无人机的无人机集群与 N 个地面用户组成,分别表示为 $\mathcal{U}=\{U_1, U_2, \dots, U_M\}$, $\mathcal{C}=\{C_1, C_2, \dots, C_N\}$ 。其中,每个无人机 $U_m \in \mathcal{U}$ 的三维坐标表示为 $q_m=(x_m^u, y_m^u, z_m^u) \in L^U$,地面用户 C_n 的二维坐标表示为 $p_n=(x_n^c, y_n^c) \in L^C$ 。时间轴被划分为多个等长的时隙,每个时隙长度为 τ ,记为 $t_i \in \mathcal{T}$,其中索引 $i \in \{1, 2, \dots, T\}$ 。通信与定位信号采用独立的频段进行传输,其中,通信频段被划分为 K 个子信道,子信道带宽大小为 B ,表示为 $\mathcal{K}=\{k_i | i=1, 2, \dots, K\}$ 。在通信过程中,用户与无人机之间采用非正交方式,一个无人机最多占用一个子信道,一个子信道可被多个无人机占用。定位信号在独立定位频段内采用预设的频分正交与时分正交资源,以避免定位信号间干扰,具体过程见2.3节。本文仅对通信子信道进行分配优化。

在本文所考虑的场景中,受限与建筑物、地形或山脉森林等复杂环境的遮挡,GNSS无法为地面用户提供高精度的定位服务。同时,地面基站因灾害影响而失效,无法依赖传统的基站通信与辅助定位方法。无人机集群网络能够作为一种有效的替代方案,为地面用户提供持续、可靠的通信服务与高精度的定位支持。其工作流程为:在每个时隙中,地面用户首先向无人机发送定位信号与通信信号,无人机接收信号,在等待一段时间后向用户发送定位应答信号,用户接收并计算获得自身位置信息。在本文中,我们重点关注用户与无人机通信与定位过程中的关联与资源分配问题。

1.2 通信模型

本文考虑的通信模型如图2所示。本文采用概率信道模型对无人机与地面用户之间的通信进行建模。 U_m 与 C_n 之间信道为视距信道(Line-Of-Sight, LOS)的概率为^[21]

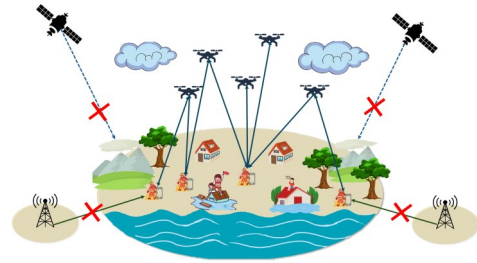


图1 系统模型

Figure 1 System model

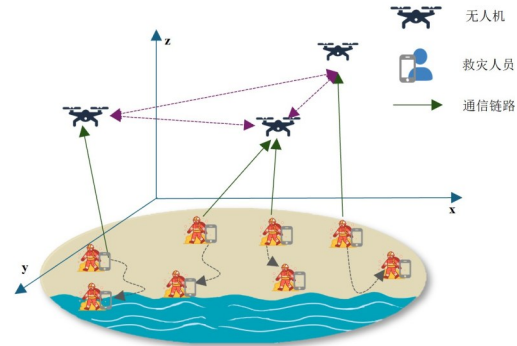


图2 通信模型

Figure 2 Communication model

$$P^{\text{LOS}} = \frac{1}{\left(1 + a \cdot \exp\left(-b \sin^{-1}\left(\frac{H}{d_{mn}(t)}\right)\right) - a\right)} \quad (1)$$

其中, a 和 b 是与环境相关的常数, $d_{mn}(t)=\|p_m - q_n\|_2$ 为 U_m 与 C_n 之间的距离, H 为无人机的高度。LOS信道与NLOS信道的概率满足: $P^{\text{NLOS}}=1-P^{\text{LOS}}$ 。

相应地, U_m 与 C_n 间的LOS和NLOS的路径损耗分别表示为

$$\text{PL}_{nm}^{\text{LOS}}(t) = L_{nm}^{\text{FS}}(t) + \eta^{\text{LOS}} \quad (2)$$

$$\text{PL}_{nm}^{\text{NLOS}}(t) = L_{nm}^{\text{FS}}(t) + \eta^{\text{NLOS}} \quad (3)$$

其中, $L_{nm}^{\text{FS}}(t)$ 为自由空间损耗模型,具体地:

$$L_{nm}^{\text{FS}}(t) = 20\lg d_{mn}(t) + 20\lg f + 20\lg\left(\frac{4\pi}{c}\right) \quad (4)$$

其中, f 为信号传输的中心频率, η^{LOS} 和 η^{NLOS} 分别为LOS和NLOS的平均额外损失。因此, U_m 与 C_n 之间的平均损耗为

$$\text{PL}_{nm}(t) = P^{\text{LOS}} \cdot \text{PL}_{nm}^{\text{LOS}} + P^{\text{NLOS}} \cdot \text{PL}_{nm}^{\text{NLOS}} \quad (5)$$

进而,得到 U_m 与 C_n 之间的信道增益为

$$g_n^m(t) = \frac{1}{\text{PL}_{nm}(t)} \quad (6)$$

本文采用离散化的形式对通信功率进行表示,假设将通信功率分成 J 份,则通信功率集合表示为 $P^c = \{P_1^c, P_2^c, \dots, P_J^c\}$, $j \in \{1, 2, \dots, J\} = \mathcal{J}$ 。用二进制变量

$p_{c_n}^j(t)$ 表示时隙 t 时 C_n 是否采用功率等级 j 进行通信。若采用, $p_{c_n}^j(t) = 1$, 否则, $p_{c_n}^j(t) = 0$, 其满足:

$$\sum_{j \in \mathcal{J}} p_{c_n}^j(t) = \{0, 1\} \quad (7)$$

其中, $p_{c_n}^j(t) \in \mathcal{P}_n^c(t) = \{p_{c_n}^1(t), p_{c_n}^2(t), \dots, p_{c_n}^J(t)\}$, 且 $\mathcal{P}_n^c(t) \subset \mathcal{P}^c(t)$ 。基于上述定义, 将 C_n 在时隙 t 的通信功率表示为

$$p_n^c(t) = \sum_{j=1}^J p_{c_n}^j(t) \cdot P_j^c \quad (8)$$

令 $\mu_{nm}^c(t) \in M_n^c(t)$ 表示 U_m 是否为 C_n 提供通信服务, 若是, 则 $\mu_{nm}^c(t) = 1$; 否则, $\mu_{nm}^c(t) = 0$, 其满足:

$$\sum_{m=1}^M \mu_{nm}^c(t) \in \{0, 1\} \quad (9)$$

令 $M_n^c(t) = \{\mu_{n1}^c(t), \mu_{n2}^c(t), \dots, \mu_{nM}^c(t)\}$ 。上式表示最多只有一个无人机为一个地面用户提供通信服务, $U_{m_c} \in \mathcal{U}$ 表示为 C_n 提供通信服务的无人机。令 $\theta_{c_n}^k(t) \in \Theta_{c_n}(t)$ 为 U_m 在时隙 t 是否占用子信道 k 来进行通信, 若是, 则 $\theta_{c_n}^k(t) = 1$; 否则, $\theta_{c_n}^k(t) = 0$, 其满足:

$$\sum_{k=1}^K \theta_{c_n}^k(t) \in \{0, 1\} \quad (10)$$

由此, U_{m_c} 在第 k 个子信道上的通信信干噪比 (Signal-to-Interference-plus-Noise Ratio, SINR) 可以表示为

$$r_{c_n}^k(t) = \frac{\sum_{m=1}^M g_n^m(t) \cdot \mu_{nm}^c(t) \cdot p_n^c(t)}{I_{c_n}^k(t) + \sigma^2} \quad (11)$$

其中, $\sigma^2 = N_0 B$ 表示噪声功率, $I_{c_n}^k(t)$ 表示 C_n 在第 k 个子信道接收到的干扰功率大小, 表示为

$$I_{c_n}^k(t) = \sum_{l=1, l \neq n}^N \sum_{U_m \in \mathcal{U}} g_l^m(t) \theta_{cl}^k(t) p_n^c(t) \quad (12)$$

即在同一子信道上由其他复用链路产生的总接收干扰功率。之后, 得到由 C_n 到 U_{m_c} 的总 SINR 及吞吐量分别为

$$r_{c_n}(t) = \sum_{k=1}^K \frac{g_n^m(t) \cdot \mu_{nm}^c(t) \cdot \theta_{c_n}^k(t) \cdot p_n^c(t)}{I_{c_n}^k(t) + \sigma^2} \quad (13)$$

$$R_n = B \log_2(1 + r_{c_n}(t)) \quad (14)$$

1.3 定位模型

本文采用 DR-TWR 方法来对用户进行定位, 如图 3 所示。DR-TWR 是一种基于双向测距的定位方法, 能够降低系统对严格时钟同步的依赖^[22], 适用于本文无人机分布式协同提供通信与定位服务场景。DR-TWR 的定位流程如下: 在一个时隙内, 用户与其关联无人机进行双向测距交互。用户发送定位请求信号后, 无人机在预设应答间隔内返回携带接收与发

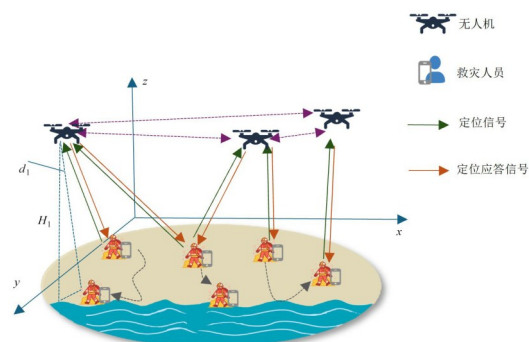


图3 定位模型

Figure 3 Positioning model

送时间戳的定位应答信号。用户根据本地时间戳、无人机反馈时间戳和已知应答间隔估计其与各关联无人机之间的伪距, 并进一步结合多个锚点的伪距测量结果实现定位。为避免定位信号间干扰, 定位上行和下行阶段, 通过正交频分多址技术区分不同的用户定位信号。在下行阶段, 用户通过时分多址技术区分不同的无人机发送的定位应答信号。假设与 C_n 关联的定位无人机为 $U_{m_p} \in U_n^p \subset \mathcal{U}$, 则 $|U_n^p| = \chi_A \geq 3$ 表示用户与 3 个以上无人机关联进行定位^[23]。由文献[24]可知, C_n 与 U_m 之间的伪距表示为

$$d_{nm} = d_{nm}^* + n_{nm} = \|p_n - q_m\| + n_{nm} \quad (15)$$

其中, d_{nm}^* 为 C_n 与 U_m 之间的真实距离, n_{nm} 为伪距测量方差。由于 DR-TWR 的测距涉及用户与无人机双向通信过程, 因此其伪距测量方差由上行和下行两阶段的到达时间 (Time Of Arrival, TOA) 测量方差共同决定。由文献[24]可知:

$$n_{nm} \sim \mathcal{N}\left(0, \frac{1}{4} \sigma_{n \rightarrow m}^2 + \frac{5}{4} \sigma_{m \rightarrow n}^2\right) \quad (16)$$

其中, $\sigma_{n \rightarrow m}^2$ 表示 U_m 接收到来自 C_n 的最小 TOA 测量方差, $\sigma_{m \rightarrow n}^2$ 表示 C_n 接收到来自 U_m 的最小 TOA 测量方差。具体地, U_m 接收到来自 C_n 的伪距测量方差为^[25]

$$\sigma_{n \rightarrow m}^2 (\text{m}^2) = \frac{c^2}{B^2 r_{n \rightarrow m}} \quad (17)$$

$r_{n \rightarrow m}$ 为 U_m 的定位信号接收信噪比 (Signal-to-Noise Ratio, SNR); $r_{m \rightarrow n}$ 同理。为了方便起见, 令 $(\hat{\sigma}_{nm})^2 = \frac{1}{4} \sigma_{n \rightarrow m}^2 + \frac{5}{4} \sigma_{m \rightarrow n}^2$ 。由此, C_n 的伪距测量标准差表示为 $\hat{\sigma}_{nm} = [\hat{\sigma}_{nm_1}, \hat{\sigma}_{nm_2}, \dots, \hat{\sigma}_{nm_{\chi_A}}]^T$ 。相应地, C_n 的定位均方根误差 (Root Mean Square Error, RMSE) 为

$$\epsilon_r = \text{RMSE} \geq \sqrt{H_n \cdot Q_n \cdot (H_n)^T} \quad (18)$$

其中,

$$H_n = \left[(G_n)^T G_n \right]^{-1} (G_n)^T \quad (19)$$

$$\text{且, } \mathbf{Q}_n = \begin{bmatrix} \hat{\sigma}_{nm_1^p}^2 & 0 & \dots & 0 \\ 0 & \hat{\sigma}_{nm_2^p}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \hat{\sigma}_{nm_{\lambda_A}^p}^2 \end{bmatrix} \quad (20)$$

$\mathbf{H}_n$ 表示与 C_n 关联的定位锚点几何位置相关的矩阵。 $\mathbf{Q}_n$ 为对角阵, 表示 C_n 与各关联定位锚点之间伪距测量误差的统计特性, 其每个对角元素对应一条用户-无人机测距链路的伪距测量方差。 $\mathbf{G}_n$ 是一个与待测点及其锚点几何形状相关的矩阵, 其每一行由对应锚点相对于用户的方向向量构成。具体地,

$$\mathbf{G}_n = \begin{bmatrix} l_x^{nm_1^p} & l_y^{nm_1^p} \\ l_x^{nm_2^p} & l_y^{nm_2^p} \\ \vdots & \vdots \\ l_x^{nm_{\lambda_A}^p} & l_y^{nm_{\lambda_A}^p} \end{bmatrix} \quad (21)$$

其中,

$$\begin{cases} l_x^{nm_i^p} = \frac{(x_n^p - x_{m_i^p}^p)}{d_{nm_i^p}} \\ l_y^{nm_i^p} = \frac{(y_n^p - y_{m_i^p}^p)}{d_{nm_i^p}} \end{cases} \quad (22)$$

其中, $l_x^{nm_i^p}$ 和 $l_y^{nm_i^p}$ 分别代表 C_n 与 $U_{m_i^p}$ 在 x 和 y 轴上的方向向量。在 95% 的置信度下, C_n 的定位精度为

$$\delta_n \approx 1.614 \epsilon_r \quad (23)$$

根据上述分析可知, 用户的定位精度由关联锚点的几何形状与其接收定位信号的伪距测量方差共同决定。

与通信功率类似, 令离散化的定位功率大小集合表示为 $P^p = \{P_1^p, P_2^p, \dots, P_J^p\}$, $j \in \{1, 2, \dots, J\} = \mathcal{J}$ 。用二进制变量 $p_{p_n}^j(t)$ 表示时隙 t 时, C_n 是否采用功率 j 进行定位。若采用, $p_{p_n}^j(t) = 1$, 否则, $p_{p_n}^j(t) = 0$, 其满足:

$$\sum_{j \in \mathcal{J}} p_{p_n}^j(t) = \{0, 1\} \quad (24)$$

$p_{p_n}^j(t) \in \mathcal{P}_n^p(t) = \{p_{p_n}^1(t), p_{p_n}^2(t), \dots, p_{p_n}^J(t)\}$, 且 $\mathcal{P}_n^p(t) \subset \mathcal{P}^p(t)$ 。基于上述定义, 将 t 时隙 C_n 的定位功率表示为

$$p_n^p(t) = \sum_{j=1}^J p_{p_n}^j(t) \cdot P_j^p \quad (25)$$

采用 $\lambda_{nm}^p(t) \in \mathcal{A}_n^p(t)$ 来表示 U_m 是否为 C_n 提供定位服务, 若是, 则 $\lambda_{nm}^p(t) = 1$; 否则, $\lambda_{nm}^p(t) = 0$, 其满足:

$$\sum_{m=1}^M \lambda_{nm}^p(t) = \chi_A \quad (26)$$

上式说明有且只有 N_{pos} 个无人机为一个地面用

户提供定位服务。由于用户采用正交频分多址技术向无人机发送定位信号, 因此 $U_{m_i^p}$ 接收到来自 C_n 的定位信号 SNR 为

$$r_{n \rightarrow m_i^p} = \sum_{m=1}^M \frac{g_n^{m_i^p}(t) \cdot \lambda_{nm_i^p}^p(t) \cdot p_n^p(t)}{\sigma^2} \quad (27)$$

定位下行阶段, 无人机采用正交频分多址与时分多址技术以避免信号间干扰, 故 C_n 接收到 $U_{m_i^p}$ 的定位应答信号的 SNR 为

$$r_{m_i^p \rightarrow n} = \sum_{m=1}^M \frac{g_n^{m_i^p}(t) \cdot \lambda_{nm_i^p}^p(t) \cdot p_u}{\sigma^2} \quad (28)$$

其中, p_u 为无人机定位应答信号的发送功率。为了保障用户端能接收并正确解调定位信号, 应满足:

$$r_{n \rightarrow m} \geq r_{\text{pos}}, r_{m \rightarrow n} \geq r_{\text{pos}} \quad (29)$$

由式 (16) (17) (27) 和 (28) 可得, 用户的定位精度与锚点的选择以及定位信号功率大小相关。用户与无人机之间的调度策略设计如下: 考虑到每个时隙的持续时间较短, 可合理假设在单个时隙内用户与无人机的位置与信道增益保持不变。同时, 为降低调度复杂度, 假设每个时隙内由策略得到的资源分配方案固定不变, 系统仅在每个时隙开始时根据当前网络状态选择新的分配方案, 或延续上一时隙的方案以实现调度的时序连续性与稳定性。这种策略为旨在优化无人机集群支持的 ILAC 最大化最小吞吐量的在线学习算法奠定了基础。通过利用实时反馈, 该算法可以适应动态网络条件并提高跨多个时隙的整体性能。

1.4 问题建模

本节构建无人机集群辅助 ILAC 网络优化问题, 联合优化无人机关联、通信与定位功率以及通信子信道选择, 降低通信信号间干扰, 在保障定位精度的前提下以最大化最小用户吞吐量。问题描述如下:

$$P: \max_{M_n^c(t), A_n^p(t), P_n^p(t), \theta_{c_n}(t)} \min_n R_n$$

$$n \in \{1, 2, \dots, N\}$$

$$\text{s.t.} \begin{cases} \text{C1: } \delta_n^2 \leq \delta_r^2 \\ \text{C2: } \sum_{m=1}^M \mu_{nm}^c(t) \in \{0, 1\} \\ \text{C3: } \sum_{m=1}^M \lambda_{nm}^p(t) = \chi_A \\ \text{C4: } \sum_{k=1}^K \theta_{c_n}^k(t) \in \{0, 1\} \\ \text{C5: } p_n^c(t) + p_n^p(t) \leq p^{\text{total}} \\ \text{C6: } r_{n \rightarrow m} \geq r_{\text{pos}}, r_{m \rightarrow n} \geq r_{\text{pos}} \\ \text{C7: } \chi_A \geq 3 \\ \text{C8: } \mu_{nm}^c(t), \lambda_{nm}^p(t), \theta_{c_n}^k(t), p_{c_n}^j(t), p_{p_n}^j(t) \in \{0, 1\} \end{cases} \quad (30)$$

其中, $M_n^c(t)$ 和 $A_n^p(t)$ 表示用户与无人机的关联决策, $\mathcal{P}_n^c(t)$ 和 $\mathcal{P}_n^p(t)$ 表示用户通信与定位的发射功率等级选择, $\theta_{c_n}(t)$ 表示通信子信道选择。在优化目标中, 本文最大化最小用户吞吐量, 相比于系统总吞吐量与平均吞吐量, 该指标更加关注于应急场景中较弱用户性能, 更体现了多用户场景下通信的高效性与公平性, 尽可能保障所有用户的基本通信需求。在约束条件中: C1 表示用户在 95% 置信度下的定位精度需满足预设阈值; C2 表示每个用户最多只能选择一个无人机进行通信; C3 表示每个用户有且只能选择 χ_A 个无人机进行定位; C4 表示每个用户最多只能选择一个子信道进行通信; C5 表示用户在单个时隙的通信与定位功率之和小于总功率; C6 表示无人机的接收定位 SINR 大于一个阈值以能在接收端正确解调; C7 表示二维定位锚点数量需大于等于 3; C8 表示二进制变量约束。

需要指出的是, 本文中通信性能与定位性能并非相互独立。由上述优化方程可知, 通信功率分配、定位功率分配以及通信子信道选择共同影响用户吞吐量与干扰水平。通信功率与子信道选择主要影响通信链路干扰水平, 定位性能则主要由定位锚点选择与定位功率共同决定。因此, 本文在满足定位精度约束的前提下, 对上述变量进行联合优化。

2 算法描述

由于非凸约束以及二进制变量 $\mu_{nm}^c(t)$, $\lambda_{nm}^p(t)$, $p_{c_n}^j(t)$, $p_{p_n}^j(t)$, $\theta_{c_n}^k(t)$, 问题 P 为一个混合整数非线性非凸优化问题, 难以用传统的数学方法来求解。为了获得满意的次优解, 将上述问题分解成两个子问题: 无人机关联问题(子问题 1)以及功率控制与子信道分配问题(子问题 2)。首先, 针对第一个子问题, 本文提出了 3D-FAS 方法, 该方法充分考虑无人机的不同高度和空间覆盖特性, 缩小搜索空间选择较优通信接入无人机与定位无人机锚点组合, 以保障用户的定位和通信服务质量, 并将其输出的无人机关联方案作为后续算法的初始条件。其次, 针对第二个子问题, 本文引入 MAPPO 算法, 将每个用户视为一个智能体, 并通过与环境交互实现通信功率、定位功率与通信子信道分配的动态优化。基于此, 提出基于 MAPPO 的联合优化(MAPPO-based Joint Optimization, MAPPO-JO)方法, 将 3D-FAS 输出的无人机关联方案与 MAPPO 资源分配策略相结合, 在满足定位精度约束的前提下最大化最小用户吞吐量, 从而提升通信公平性与定位服务可靠性。

2.1 基于三维空间的空中锚点快速选择方法(3D-FAS)

本节提出了 3D-FAS 算法, 为地面用户选择合适

的通信无人机与定位无人机。在无人机集群辅助的 ILAC 网络中, 通信与定位无人机锚点直接影响系统的通信与定位性能。因此, 本文考虑无人机的三维空间分布特性, 为后续资源分配提供高效的初始解。

3D-FAS 算法包括通信接入点选择与定位锚点选择两阶段。在通信接入点选择阶段, 考虑到用户与无人机之间的信道增益对通信性能具有重要影响, 同时为降低整体搜索复杂度, 本文采用最近距离关联准则, 将与用户直线距离最近的无人机选为其通信接入点, 以获得低复杂度且具有较好初始性能的通信接入方案。在定位锚点选择阶段, 由于本文主要关注地面用户的二维定位性能, 故采用水平精度衰减因子(Horizontal Dilution Of Precision, HDOP)评估无人机锚点的几何形状, 较小的 HDOP 值意味着选择的锚点分布更均匀, 有助于提高定位精度^[26]。需要说明的是, HDOP 主要刻画几何分布对定位误差的影响, 而链路质量差异则由后续方法中的功率优化进一步补偿。HDOP 表示为

$$\text{HDOP} = \sqrt{\sigma_x^2 + \sigma_y^2} \quad (31)$$

其中, σ_x^2 与 σ_y^2 分别为矩阵 $(\mathbf{H}^T \mathbf{H})^{-1}$ 的前两个对角线元素。在选择定位无人机锚点时, 我们的目标是从所有的候选集合 $\mathcal{O}$ 中选择最小 HDOP 的锚点组合 $\mathcal{O}^*$, 表示为

$$\mathcal{O}^* = \min(\text{HDOP}(\mathcal{O})) \quad (32)$$

其中, 候选集合 $\mathcal{O}$ 的大小与无人机数和定位锚点数相关, 即 $|\mathcal{O}| = C_M^{\chi_A}$ 。值得注意的是, 候选集合的规模随着无人机数量而增加, 而用户终端与无人机平台受限于计算能力和能量供给, 特别是在大规模网络中, 难以承受穷举搜索带来的高计算复杂度与能耗开销。为此, 本文提出了一种空中定位锚点的快速选择方法, 根据无人机的位置分布特征采用 K -means 聚类对无人机进行分类以降低搜索空间, 其原理图如图 4 所示。

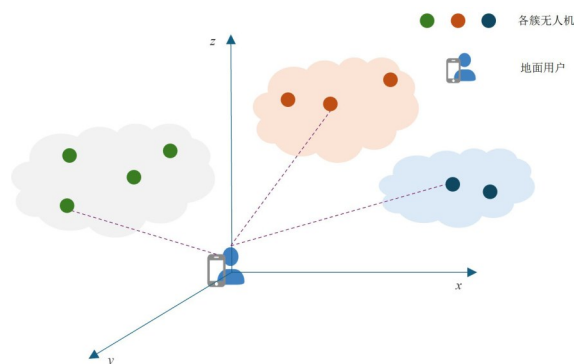


图4 基于3D-FAS的空间划分示意图

Figure 4 Spatial partitioning of 3D-FAS

K -means 是一种基于距离度量的无监督聚类算法,本文利用 K -means 对无人机的位置分布进行聚类,从而缩小定位锚点的搜索范围,提高候选锚点选择效率。对于每个用户,其优化目标为最小化各无人机相对位置到所属聚类中心的距离平方和,即

$$J = \sum_{k=1}^K \sum_{C_n \in \mathcal{C}^k} \|\hat{q}_m^n - \mu_k\|^2 \quad (33)$$

其中, $\mathcal{C}^k$ 表示第 k 个无人机簇, μ_k 表示对应的聚类中心, $\hat{q}_m^n$ 表示 U_m 与 C_n 的相对位置向量。在每次迭代中, K -means 首先按照最小距离准则将样本分配到最近的聚类中心,其分类规则为

$$c_i = \arg \min_{k \in \{1, 2, \dots, K\}} \|\hat{q}_m^n - \mu_k\|^2 \quad (34)$$

之后根据当前簇内所有样本重新计算聚类中心,即

$$\mu_k = \frac{1}{|\mathcal{C}^k|} \sum_{C_n \in \mathcal{C}^k} \hat{q}_m^n \quad (35)$$

3D-FAS 算法的主要流程如下:首先,对于每个用户,选择与其距离最近的无人机作为通信接入无人机。然后,遍历所有用户,以每个用户位置为中心,计算无人机的相对位置,执行 K -means 聚类算法,将无人机划分为 χ_A 类。之后,分别从 χ_A 类无人机中选择一个无人机进行组合,形成筛选后的候选解集合 $\mathcal{O}'$ 。对于每个候选解,计算其 HDOP 值,选择最小的 HDOP 对应的组合作为最优定位锚点。最终得到通信与定位无人机选择方案。3D-FAS 算法的伪代码如算法 1 所示。

对于 3D-FAS 算法,其时间复杂度为 $O(N(M + IM_{\chi_A} + |O_p|f(\chi_A)))$,其中 I 为 K -means 迭代次

算法 1 3D-FAS

输入: 无人机数量 M 与位置 L^p , 定位锚点数 χ_A

输出: 通信与定位锚点选择方案 $\mathcal{O}^*$

1. 初始化通信与定位锚点方案 $\mathcal{O}_c^* \leftarrow \emptyset$, $\mathcal{O}_p^* \leftarrow \emptyset$
2. For $C_n \in \mathcal{C}$:
3. $U_n^c \leftarrow \min_{m \in \mathcal{M}} (\|q_m - p_n\|_2)$
4. $\mathcal{O}_c^* \leftarrow \mathcal{O}_c^* \cup \{U_n^c\}$
5. 得到通信无人机选择方案 $\mathcal{O}_c^*$
6. For $C_n \in \mathcal{C}$:
7. 以 $(x_n^c, y_n^c, 0)$ 为中心, 计算所有无人机的相对位置, 基于相对位置执行 K -means 聚类方法, 将所有无人机划分为 χ_A 类
8. 分别从 χ_A 类无人机中选择一个无人机进行组合, 得到定位锚点候选集合 $\mathcal{O}_p$
9. For $o_p^i \in \mathcal{O}_p$:
10. 计算 $\mathcal{O}_p$ 内所有组合对应的 HDOP 值 $\text{HDOP}(o_p^i)$
11. 将 $\min(\text{HDOP}(o_p^i))$ 添加进 $\mathcal{O}_p^*$
12. $\mathcal{O}^* \leftarrow \mathcal{O}_c^* \cup \mathcal{O}_p^*$

数, $|O_p| = \prod_{k=1}^{\chi_A} m_k$ 为候选定位锚点组合数, m_k 为第 k 个簇内无人机个数, $f(\chi_A)$ 为 HDOP 计算复杂度, 其复杂度可近似为 $O\left(N\left(IM_{\chi_A} + \prod_{k=1}^{\chi_A} m_k\right)\right) \leq O\left(N\left(IM_{\chi_A} + \left(\frac{M}{\chi_A}\right)^{\chi_A}\right)\right)$ 。

穷举方法的复杂度近似为 $O\left(N\left(\frac{M}{\chi_A}\right)\right)$ 。对于定位锚点选择, 3D-FAS 通过 K -means 预筛选显著缩小了定位锚点搜索空间, 使得评估的候选组合数由 $\left(\frac{M}{\chi_A}\right)$ 减少为

$\prod_{k=1}^{\chi_A} m_k$, 有效减少了实际搜索开销, 从而提高定位锚点选择效率。

2.2 基于多智能体近端策略优化的联合优化方法 (MAPPO-JO)

如 2.1 节所述, 我们通过 3D-FAS 算法解决了用户的通信与定位无人机关联问题。在此基础上, 本节引入 MAPPO 以优化功率与子信道分配, 并结合 3D-FAS, 提出 MAPPO-JO, 通过联合优化锚点选择、功率与子信道分配, 在保障定位精度约束的前提下, 最大化最小用户吞吐量, 从而提升整体资源分配的公平性与服务质量。

本文将功率控制和子信道分配问题转化为 N 个用户完全协作的多智能体任务, 表示为 $\langle \mathcal{S}, \{\mathcal{O}_n\}_{n \in \mathcal{N}}, \{\mathcal{A}_n\}_{n \in \mathcal{N}}, R, \gamma \rangle$ 。其中, 将每个用户 C_n 看成一个智能体, $\mathcal{S}$ 表示全局网络的环境状态, $\mathcal{O}_n$ 表示每个智能体的局部观测状态, $\mathcal{A}_n$ 为每个智能体的动作空间, R 和 $\gamma \in [0, 1]$ 分别代表奖励和折扣因子。在每个时隙, 用户观测得到局部环境状态 $O_n(t) \in \mathcal{O}_n$, 根据策略网络选取一个动作 $A_n(t) \in \mathcal{A}_n$, 并得到所有智能体的联合动作 $A(t) = A_1(t) \times A_2(t) \times \dots \times A_N(t)$ 。所有智能体在执行动作之后, 环境返回一个全局奖励 $R(t)$, 并转移到下一时刻的状态 $S(t)$ 。我们分别对状态空间、观测空间、动作空间以及奖励进行定义。MAPPO 的原理如图 5 所示。

(1) 状态空间

在时隙 t , 状态空间包括所有用户的通信与定位的功率控制、通信子信道占用以及所有子信道接收功率大小信息, 表示为

$$S(t) = \{\mathcal{P}^c(t), \mathcal{P}^p(t), \Theta_c(t), \mathcal{Q}(t)\} \quad (36)$$

其中, $\mathcal{P}^c(t)$ 和 $\mathcal{P}^p(t)$ 分别表示 N 个用户在时隙 t 的通信与定位功率等级选择, $\Theta_c(t) = \{\theta_{c_1}(t), \theta_{c_2}(t), \dots, \theta_{c_N}(t)\}$ 为 N 个地面用户的通信子信道占用状态, $\mathcal{Q}(t) =$

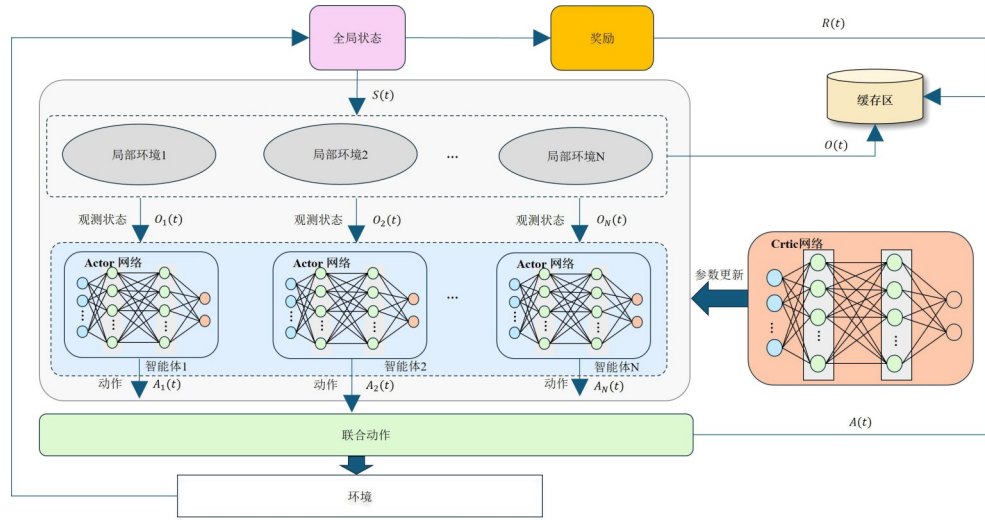


图5 MAPPO原理图

Figure 5 MAPPO schematic diagram

$\{Q_1(t), Q_2(t), \dots, Q_N(t)\}$ 则表示 N 个用户在所有子信道上接收到的其他用户的功率大小, 且 $|Q_n(t)| = K$ 。

(2) 局部观测状态

每个智能体观测到部分状态, 包括当前用户在时隙 t 的通信与定位信号的功率等级、通信子信道占用状态, 以及该用户在所有子信道上接收到其他用户的功率大小, 表示为

$$O_n(t) = \{P_n^c(t), P_n^p(t), \Theta_{c_n}(t), Q_n(t)\} \quad (37)$$

(3) 动作空间

每个智能体的动作空间表示为

$$\mathcal{A}_n = \{A_1, A_2, \dots, A_I\} \quad (38)$$

其中, $A_i = [a_{i1}, a_{i2}, a_{i3}]$, a_{i1}, a_{i2}, a_{i3} 分别表示通信与定位功率等级调整量, 以及通信子信道索引调整量, 且 $a_{i1}, a_{i2}, a_{i3} \in \{0, 1, -1\}$ 。并满足以下条件:

- (i) 对于任意两个动作 $A_i, A_j \in \mathcal{A}_n, A_i \neq A_j$;
- (ii) $a_{i1} + a_{i2} = 0$ 。

(4) 奖励函数

为了有效地将本文优化目标转化为奖励, 考虑以下因素: (i) 本文的主要优化目标是最大化最小用户吞吐量以保障网络的公平性, 因此, 奖励函数应与用户的最小吞吐量呈正相关关系; (ii) 保障定位精度为本文的强约束条件, 故对所有不满足定位精度的情况应给予负奖励, 以引导智能体学习, 保障用户定位精度需求; (iii) 由于智能体的动作空间由离散的功率等级与子信道索引组合构成, 为防止非法动作的出现, 对于超界状态施加惩罚, 促使策略神经网络仅学习在可行域内进行有效探索; (iv) 为了加速智能体的收敛速度, 应该将奖励设置在一个范围内。

综上所述, 我们设计了全局奖励如下:

$$R(t) = \begin{cases} \min_n \lg R_n(t), & Y_{\text{npos}}(t) = 0 \\ -\lg(50 \times Y_{\text{npos}}(t)), & Y_{\text{npos}}(t) \neq 0 \end{cases} \quad (39)$$

其中, $Y_{\text{npos}}(t)$ 表示时隙 t 不满足定位精度的用户个数, 具体为

$$Y_{\text{npos}}(t) = \sum_{n=1}^N \mathbf{1}(\delta_n(t) > \delta_r) \quad (40)$$

值得注意的是, 本文通过奖励函数对不满足精度要求的情况施加惩罚的“软约束”方式来引导智能体学习, 从而在满足精度约束的同时优化整体性能。

近端策略优化算法 (Proximal Policy Optimization, PPO) 是一种稳定且高效的基于策略的在线强化学习方法^[27-28], 在每次策略更新时限制新旧策略之间的变动幅度, 从而避免传统策略梯度方法在更新过程中易出现策略剧烈变化导致的性能下降问题。PPO 通过引入了概率比值函数以控制新旧策略的变化程度, 从而实现稳定的策略改进。策略目标函数为

$$L^{\text{CLIP}}(\theta) = \mathbb{E}_t \left[\min \left(r_t(\theta) \hat{A}(t), \text{clip} \left(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon \right) \hat{A}(t) \right) \right] \quad (41)$$

其中: $r_t(\theta) = \pi_\theta(A(t)|S(t)) / \pi_{\theta_{\text{old}}}(A(t)|S(t))$ 为新旧策略在同一状态下执行某一动作的概率比; ε 为裁剪阈值; $\text{clip}(\cdot)$ 是裁剪函数, 以将 $r_t(\theta)$ 限制在 $[1 - \varepsilon, 1 + \varepsilon]$ 内; $\hat{A}(t)$ 为优势函数, 表示为

$$\hat{A}(t) = \hat{Q}(S(t), A(t)) - V_\phi(S(t)) \quad (42)$$

其中, $\hat{Q}(S(t), A(t))$ 是基于样本估计的动作-价值函数。为了提高估计的准确性并降低方差, 通常采用广义优势估计 (Generalized Advantage Estimation, GAE) 来计算 $\hat{A}(t)$, 通过对 n 步引导进行线性组合, 从而在降低偏

差与减小方差之间实现权衡:

$$\hat{Q}(S(t), A(t)) = \hat{A}(t) + V_{\phi}(S(t)) = \sum_{l=t}^{\infty} (\gamma \lambda)^{l-t} \delta_l + V_{\phi}(S(t)) \quad (43)$$

其中, γ 为折扣因子, λ 为 GAE 平衡系数, $V_{\phi}(s_t)$ 是当前状态下的值函数估计, δ_l 是时间差分误差, 表示为

$$\delta_l = R(l) + \gamma V_{\phi}(S(l+1)) - V_{\phi}(S(l)) \quad (44)$$

进一步地,

$$\hat{R}(t) = \hat{A}(t) + V_{\phi}(S(t)) \quad (45)$$

表示用于训练值函数的目标回报。状态值函数 $V_{\phi}(S(t))$ 用于最小化估计值与奖励之间差异, 表示为

$$L^{\text{VF}}(\phi) = \mathbb{E}_t \left[\left(V_{\phi}(S(t)) - \hat{R}(t) \right)^2 \right] \quad (46)$$

在上述基础上, 我们提出了 MAPPO-JO, 首先根据 3D-FAS 为用户选择合适的通信与定位锚点, 之后通过 MAPPO 为每个用户选择合适的功率等级与子信道, 从而在满足定位精度约束的前提下提升最小用户吞吐量, 实现通信与定位资源的协同优化。算法伪代码如算法 2 所示。

算法 2 MAPPO-JO

输入: 用户通信与定位无人机关联方案 $\mathcal{O}^*$, 定位阈值 δ_r , 折扣因子 γ , GAE 系数 λ , 裁剪阈值 ε , 训练轮数 MAX_EP, 步数 MAX_STEP, 批次大小 N_{batch} , 缓存区 D

输出: 用户通信与定位功率等级, 通信子信道分配策略

1. 初始化策略网络参数 θ 与价值网络参数 ϕ
2. For $\text{ep} = 1, 2, \dots, \text{MAX_EP}$
3. 重置环境, 得到初始状态
4. 初始化缓存区 D
5. For $\text{step} = 1, 2, \dots, \text{MAX_STEP}$
6. 执行 3D-FAS 算法(算法 1), 得到当前时隙无人机关联方案 $\mathcal{O}^*$
7. 每个智能体观测状态, 得到各自的局部观测 $O_n(t)$
8. 根据观测状态与策略网络采样动作 $A_n(t)$, 执行动作组合 $\{A_n(t)\}_{n \in \mathcal{N}}$
9. 根据式(39)计算奖励, 得到 $R(t), S(t+1)$
10. 将经验样本存入缓存区 D
11. For $\text{epoch} = 1, 2, \dots, \text{EPOCH}$
12. For $i = 0, 1, \dots, D/N_{\text{batch}}$
13. 取 N_{batch} 份样本, 计算新的策略概率比值, 得到 $r_i(\theta)$
14. 计算优势函数 $\hat{A}(t)$
15. 计算策略目标函数和价值损失函数 $L^{\text{CLIP}}(\theta)$ 与 $L^{\text{VF}}(\theta)$
16. 利用梯度下降更新策略网络参数 θ 和价值网络参数 ϕ

3 仿真与分析

在本节中, 我们对本文所提出的无人机关联与资源分配优化方法进行仿真验证。首先给出了仿真实

验相关参数, 然后将所提出的优化方法与基线算法进行对比与分析。

3.1 实验参数

本文在 Debian 12 系统环境下完成仿真实验, 实验平台配置包括 Intel i9@3.20 GHz 处理器、64 GB 内存及配备 24 GB 显存的 NVIDIA GeForce RTX 4090D 显卡。仿真所用主要参数如表 1 所示。用户与无人机在 $300 \text{ m} \times 300 \text{ m}$ 范围内随机分布, 用户数量为 6, 无人机数量为 8。我们将定位精度需求阈值设置为 $0.4 \sim 2.6 \text{ m}$, 间隔为 0.2 m , 无特殊说明默认为 1.6 m 。为保证系统模型的合理性, 本文在仿真中设置了如下边界条件。首先, 由于二维定位至少需要 3 个锚点, 因此每个用户至少需要与 3 架无人机关联, 并采用独立定位频段中的预设正交资源实现定位。其次, 定位精度必须满足预设阈值, 以确保定位服务的可靠性。此外, 用户移动被建模为马尔科夫过程。当用户位置超出预设空间范围时, 采用反向反弹机制将其重新映射至有效区域。

为了验证所提出算法的性能, 本文的基线算法如下: 采用 MADDPG 进行功率与子信道分配(记为

表 1 仿真实验参数

Table 1 Simulation parameters

参数	数值
无人机与用户分布范围	$300 \text{ m} \times 300 \text{ m}$
无人机高度范围	$30 \sim 150 \text{ m}$
M/N	8/6
K	[3-7]
f	2 GHz
B	100 kHz
τ	1 s
p^{total}	$0.3 \sim 1.2 \text{ W}$, 步长 0.1 W
p_u	0.5
χ_A	3
a, b	$9.61^{[29]}, 0.16^{[29]}$
$\eta^{\text{LOS}}, \eta^{\text{NLOS}}$	$1, 20^{[29]}$
ε, δ_r	$1/1.6^{[29]}$
B_{fe}	200 kHz
最大训练回合与步数	500, 128
更新周期 EPOCH	600
训练批次大小 B_{batch}	256
裁剪系数 ε	0.2
折扣因子 γ	0.99
学习率	5×10^{-4}
隐藏层大小	[256, 128, 64]
GAE 平衡系数 λ	0.95
K-means 迭代次数	100

MADDPG-JO);采用 MAPPO 进行固定通信与定位功率比例下的子信道分配(记为 MAPPO-FP),其中通信与定位功率比例为 1:2;固定通信与定位功率(比例为 1:2),随机选择通信子信道方法(记为 FP-RS)。基线算法所使用的无人机关联方法均为 3D-FAS。

为了验证所提出方法在联合优化上的性能,本文分别从收敛性能、通信性能以及定位性能三个方面进行比较分析。其中,收敛性能主要用于评估各算法的训练稳定性;在通信性能中将最小用户吞吐量作为评价指标,以刻画系统中性能最弱用户的通信性能,从而反映出所提出方法在保障定位精度前提下通信性能与用户间公平性的提升;定位性能中则以不满足定位精度需求的用户比例作为评价指标,用于分析所提方法在不同参数设置下对定位约束的满足能力及其对系统参数变化的敏感性。

3.2 收敛性能

我们首先验证了本文所提出的基于 MAPPO 的联合优化方法(记为 MAPPO-JO)与两种基线算法(MADDPG-JO 与 MAPPO-FP)的收敛性能,如图 6 所示。整体来看,三种算法在训练过程中,其奖励从负值逐渐增加至正值直到收敛,说明本文所设计的奖励函数能够很好地引导智能体学习,从而满足用户的定位精度约束。并且,本文所提出的基于 MAPPO 的联合优化方法可以在较少训练回合中快速收敛,并在收敛后保持较高的稳定性和较小的波动幅度。在基线算法中,MAPPO-FP 相对较快收敛,但其稳定后的性能低于 MAPPO-JO,而 MADDPG-JO 则收敛速度明显滞后且波动较大。总的来说,MAPPO-JO 的奖励值始终显著高于两种基线算法,表明本文所提方法在联合无人机关联与资源分配中具备更优的收敛速度、更强的鲁棒性以及更高的长期性能。

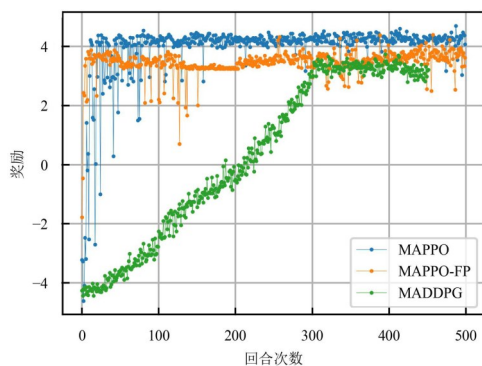


图6 三种算法收敛性能

Figure 6 Convergence performance of the three algorithms

本文采用离线训练与在线执行方式,表2列出了 MAPPO-JO 在不同子信道设置下的训练时间与推理时间。可以看出,随着子信道数量的增加,训练时间

和推理时间均呈现轻微上升趋势。这表明子信道数量的增加扩大了状态空间,从而在一定程度上提升了算法的计算开销,但整体增幅较小,说明所提方法具有较好的可扩展性。除此之外,由图6可以看出,训练阶段智能体在约80个回合附近趋于收敛,对应训练时间约为1500s。在线执行阶段,智能体单次策略执行时间处于几十毫秒量级,相较于所设置的1s时隙间隔,其能够满足在时隙内做出决策的实时性要求。由于本文采用集中式离线训练与在线执行相结合的策略,整体来看,所提出方法的训练与执行时间处于可行且合理的范围内,具有一定的实际部署能力。

表2 MAPPO-JO 训练时间与推理时间 单位:s
Table 2 Training and inference time of MAPPO-JO unit: s

时间	$K=4$	$K=5$	$K=6$	$K=7$
训练时间 (每step)	19.74	19.80	19.82	19.87
推理时间 (每step)	0.052 9	0.054 1	0.054 7	0.055 1

3.3 通信性能

图7给出了 MAPPO-JO 与其他三种基线算法在不同子信道下最小用户吞吐量随着其发射功率变化情况。值得注意的是,定位精度约束为本文的强约束条件,故当定位精度不满足时,吞吐量被设置为0。可以看出,在所有的子信道数设置下,本文所提出的基于 MAPPO 的联合优化方法(蓝色交叉)始终取得最高的最小用户吞吐量,显著优于三种基线,即 MADDPG-JO(红色点)、MAPPO-FP(黄色斜线)、FP-RS(绿色竖线),说明所提联合优化策略能更有效地兼顾用户间公平性。在固定子信道数下,最小用户吞吐量随着功率增加而增加。除此以外,在小功率设置下,基线算法 MAPPO-FP 与 FP-RS 无法满足用户定位精度需求,说明 MAPPO-JO 和 MADDPG-JO 能很好地通过调整通信与定位功率比例,在保证定位精度的同时提升通信公平性。随着功率增加,固定功率分配方法(MAPPO-FP 与 FP-RS)中的定位功率逐渐增大,从而能够满足定位精度要求,但由于其通信功率缺乏灵活调整能力,最小用户吞吐量性能仍显著低于 MAPPO-JO。总的来说,本文所提算法相对于 MADDPG-JO 与 MAPPO-FP,在用户最小吞吐量方面分别提高了16.28%、20.68% ($K=4$), 34.16%、41.62% ($K=5$), 46.20%、40.90% ($K=6$), 42.10%、66.80% ($K=7$),平均提高了34.69%与42.50%。

图8给出了 MAPPO-JO 算法在相同功率下,随着不同子信道变化其最小吞吐量的性能表现。从图8可以看出,在相同功率设置下,最小用户吞吐量随着

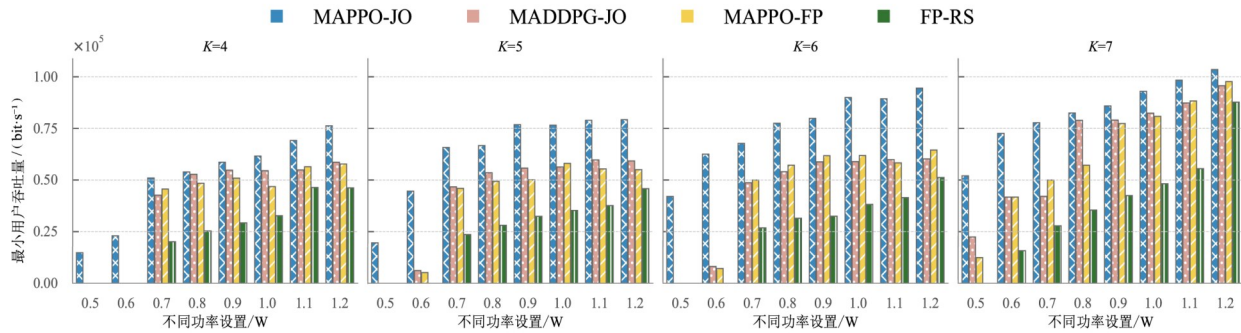


图7 不同子信道数下 MAPPO-JO 与三种基线算法最小用户吞吐量随功率变化情况

Figure 7 Communication performance of MAPPO-JO and baselines versus transmit power under different sub-channel settings

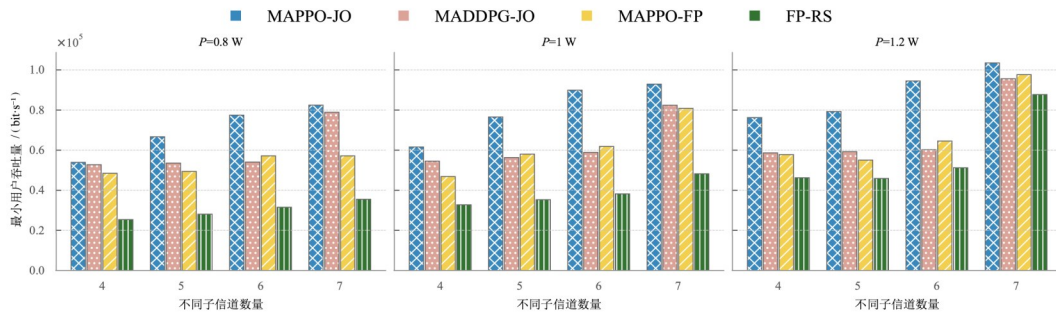


图8 不同功率设置下 MAPPO-JO 与三种基线算法最小用户吞吐量随子信道数变化情况

Figure 8 Communication performance of MAPPO-JO and baselines versus sub-channel settings under different transmit power

子信道的增加而增加,说明在频谱资源更充分的情况下所有算法均能实现更好的性能,保障用户通信的公平性。在低功率条件下,本文所提优化算法相较于三种基线方法表现出明显优势,表明其能够在有限资源下更合理地进行通信功率、定位功率以及子信道分配,从而有效提升系统性能。然而,在高功率条件下,MAPPO-JO 与基线算法之间的性能差距逐渐缩小,表明当系统资源较为充裕时,基线方法也能在一定程度上满足通信与定位需求,但 MAPPO-JO 仍能在综合性能和稳定性上保持优势。总体来看,在 $P=0.8\text{ W}$ 时,本文所提算法相对于 MADDPG-JO 与 MAPPO-FP 分别提高了 18.61% 与 31.48%;在 $P=1\text{ W}$ 时,本文所提算法相对于 MADDPG-JO 与 MAPPO-FP 分别提高了 28.59% 与 30.96%;在 $P=1.2\text{ W}$ 时,本文所提算法相对于 MADDPG-JO 与 MAPPO-FP 分别提高了 32.23% 与 32.11%,相对于 MADDPG-JO 与 MAPPO-FP 平均提高了 26.48% 以及 31.52%。

为了验证所提出方法在不同定位需求场景下的适应性,图9展示了所提出方法与基线方法在不同定位需求($\delta_r = 1.6, 1.8, 2.0, 2.2, 2.4, 2.6\text{ m}$)以及不同子信道数量下的最小用户吞吐量的变化情况。从图9可以看出,在不同子信道设置($K=4, 5, 6$)以及定位精度需求下,所提算法 MAPPO-JO 相较于基线算法始终能实

现更高的最小用户吞吐量,说明 MAPPO-JO 算法在应急救援等实际应用场景中具有较强的适应性,能够适应不同定位精度需求场景。

图10展示了本文所提出的基于 MAPPO-JO 算法在不同的子信道下随着功率变化,其最小用户吞吐量的变化情况。从图10可以看出,随着功率与子信道数的增加,最小用户吞吐量均显著提升,这表明频谱资源和功率资源的联合优化能够显著提升系统在不同场景下的通信性能,用户间的资源分配更加灵活,显著改善系统整体性能,提升最小用户吞吐量,从而保证通信公平性。

图11和图12分别展示了四种用户边界分布场景及其对应的最小用户吞吐量大小,其中,四种边界场景分别为1个、2个、3个与4个用户位于场景边缘。可以看出,在不同数量的边界用户场景下,MAPPO-JO 通过合理的锚点选择以及优化的子信道与功率分配,始终能够满足用户的定位需求,并提升用户的最小吞吐量。因此,本文所提出的方法展现了较强的适应性和鲁棒性,能够有效应对不同网络场景。

3.4 定位性能

为了验证本文所提出方法的定位性能,图13展示了 MAPPO-JO 与基线方法在不同功率设置下用户不满足定位精度需求比例的变化情况。从图13可以看出,在网络资源较少的场景下,尽管算法无法满足

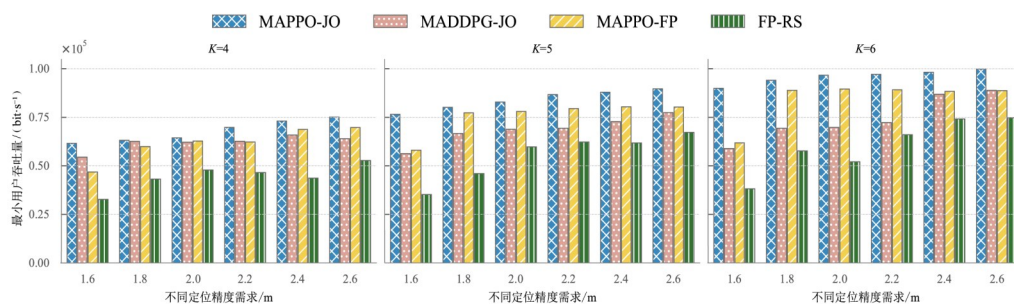


图9 不同子信道下 MAPPO-JO 与三种基线算法最小用户吞吐量随定位精度需求变化情况

Figure 9 Communication performance of MAPPO-JO and baselines versus positioning accuracy requirement

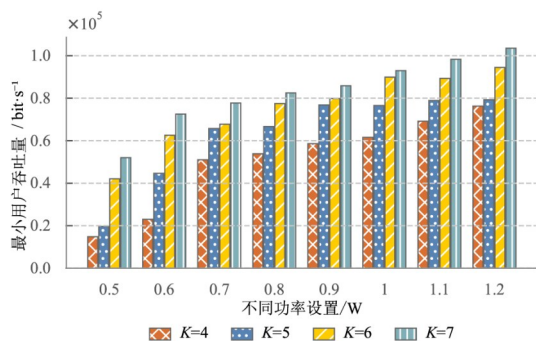


图10 MAPPO-JO 随子信道与功率的变化情况

Figure 10 MAPPO-JO performance versus sub-channels and transmit power

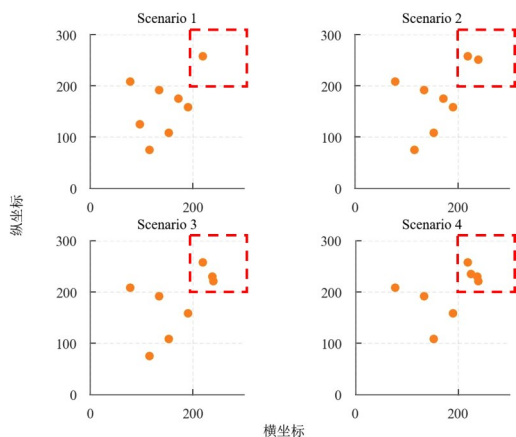
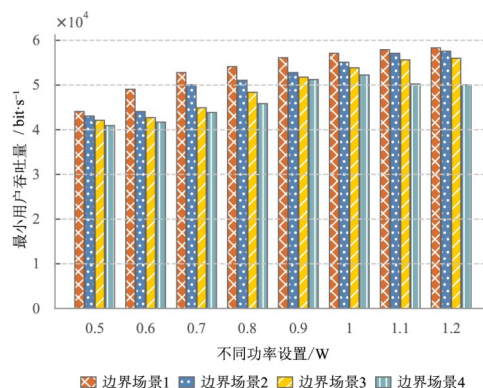


图11 用户边界情况分布

Figure 11 Distribution of user boundary scenarios

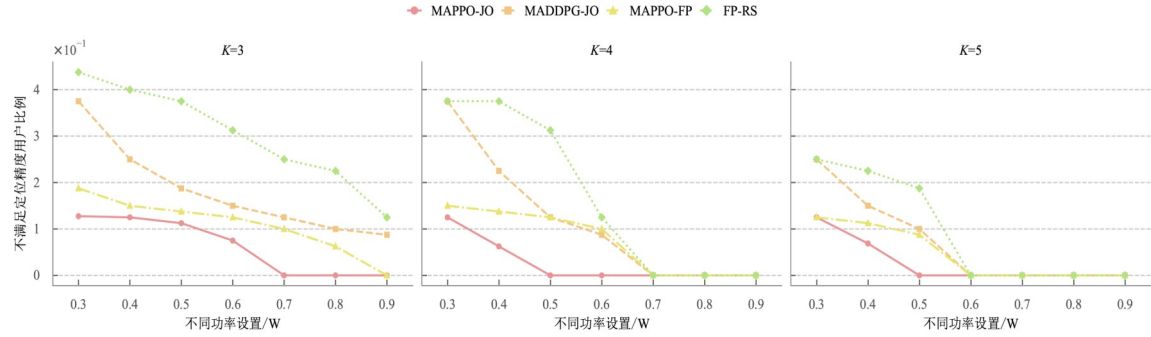
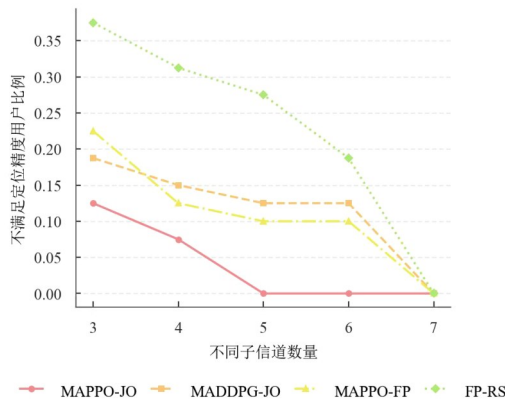
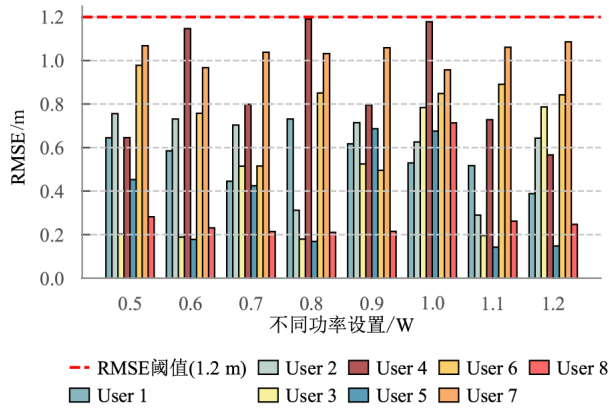
所有用户定位需求,但 MAPPO-JO (粉色实线) 在所有设置 ($K=4, 5, 6$) 下的定位性能仍优于基线算法,说明所提出方法在资源受限和约束严格条件下仍具有更强的鲁棒性与适应性。具体地,随着功率的增加,不满足定位精度需求的用户比例逐渐下降。这表明在更高功率设置下,定位所能利用的资源增加,信号质量得到改善,从而提升了定位精度。当 $P \geq 0.7$ W 时, MAPPO-JO 能满足所有的定位需求。

图 14 展示了 MAPPO-JO 与基线算法在不同子信

图12 不同用户边界分布场景下最小用户吞吐量性能变化 ($K=6, \delta_r=1.6$ m)Figure 12 MAPPO-JO performance under different user boundary distributions ($K=6, \delta_r=1.6$ m)

道数量下定位性能的变化情况,随着子信道数量增加,所有算法不满足定位精度需求的用户比例逐渐下降,说明算法能够被分配到更多的频谱资源来保障定位信号质量。除此以外,所提出方法的定位性能在所有情况下均优于基线算法,说明其能够更好地根据网络状态分配子信道与功率资源,从而保障应急场景下用户的定位需求。

为了进一步验证所提出方法在满足定位精度约束基础上的各用户的定位性能表现,图 15 展示了 MAPPO-JO 在 $K=6$, 定位 RMSE 阈值为 1.2 m 时,所有用户的 RMSE 随着功率变化情况。值得注意的是,图 15 直接展示了 RMSE 大小,以更直观地反映各用户定位均方根误差的分布及其与阈值之间的关系。从图 15 可以看出,在所有功率设置下,所有用户的 RMSE 均满足需求,表明本文设计的奖励函数能够很好地引导智能体朝向正奖励方向学习,通过功率与子信道优化,确保所有用户满足定位需求。除此以外,不同功率设置下,每个用户的 RMSE 并未呈现出明显的变化规律。这是因为本文仅将是否满足阈值作为约束条件,而未进一步考虑定位 RMSE 的具体数值。本文将

图 13 不同子信道数下 MAPPO-JO 与基线算法定位性能随功率变化情况($\delta_r = 1.6$ m)Figure 13 Positioning performance of MAPPO-JO and baselines versus different transmit power ($\delta_r = 1.6$ m)图 14 MAPPO-JO 与基线算法在不同子信道下定位性能变化情况 ($P = 0.4$ W, $\delta_r = 1.6$ m)Figure 14 Positioning performance of MAPPO-JO and baselines versus different sub-channel settings ($P = 0.4$ W, $\delta_r = 1.6$ m)图 15 所有用户在不同功率下 RMSE 分布情况($K = 6$)Figure 15 RMSE distribution of all users versus transmit power ($K = 6$)

该问题作为未来工作方向,进一步优化每一个用户的具体 RMSE 大小。

4 结论

本文针对应急场景中传统基于地面基站与 GNSS 的定位方法不可靠、通信服务难以保障的问题,提出

了一种无人机集群辅助的 ILAC 网络联合关联与资源优化方法,提升了用户间通信公平性并保障了用户定位性能。针对原始问题耦合度高、求解复杂的特点,本文将优化问题分解为无人机关联与资源分配两个子问题,并分别提出了 3D-FAS 和基于 MAPPO 的联合功率与子信道分配方法进行求解。仿真结果表明,所提方法在通信与定位性能方面均优于现有基线算法。在固定子信道条件下,本文算法相比于 MADDPG-JO 和 MAPPO-FP 算法,最小用户吞吐量分别提高了 34.69% 与 42.50%;在功率固定的条件下,相较于 MADDPG-JO 和 MAPPO-FP,其最小用户吞吐量分别提升了 26.48% 与 31.52%,验证了本文方法在无人机关联与资源分配联合优化方面的有效性。同时,所提方法在不同参数设置下均能够有效降低不满足定位精度需求的用户比例,能更好地保障用户定位需求。综上所述,本文所提出方法不仅能够提升用户通信的公平性,还能有效保障定位性能,并且在部分极端场景展示了一定的鲁棒性,说明该联合优化方法能够更好地协调通信与定位之间的资源,从而为应急场景下提供高效通信与可靠定位服务提供支撑。

除此以外,本文所提方法能够为基础设施不完善或突发中断场景下的通信与定位协同保障提供参考,从而在灾害救援、临时通信覆盖、偏远区域应急接入等场景中均具有一定的推广价值。未来将围绕更复杂的网络规模、更动态的用户分布以及更真实的环境约束开展进一步研究,例如联合考虑无人机部署与轨迹优化、用户差异化定位需求、信道不确定性以及能量受限等因素,以进一步提升系统在复杂实际场景中的适应能力与鲁棒性。

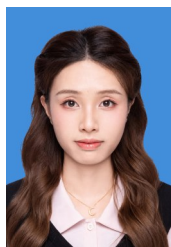
参考文献

- [1] Xiao Zhiqiang, Zeng Yong. An overview on integrated localization and communication towards 6G[J]. Science China Information Sciences, 2021, 65(3): 131301.
- [2] Meng Kaitao, Wu Qingqing, Xu Jie, et al. UAV-enabled in-

- egrated sensing and communication: Opportunities and challenges[J]. *IEEE Wireless Communications*, 2024, 31(2): 97-104.
- [3] Gao Yuan, Hu Haonan, Zhang Jiliang, et al. On the performance of an integrated communication and localization system: An analytical framework[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(7): 10845-10849.
- [4] Gao Yuan, Du Haoyu, Jiang Zhenwei, et al. A stochastic-geometry-based analytical framework for integrated localization and communication systems[J]. *IEEE Internet of Things Journal*, 2025, 12(18): 38710-38722.
- [5] Gao Junyuan, Zhu Weifeng, Zhang Shuowen, et al. Integrated massive communication and target localization in 6G cell-free networks[PP/OL]. V1. arXiv (2025-10-16) [2026-04-16]. <https://doi.org/10.48550/arXiv.2510.14281>.
- [6] Eskandari M, Savkin A V, Deghat M. RIS-enabled energy efficient integrated localization and communication for autonomous vehicles: Complementary to visual SLAM[J]. *IEEE Transactions on Green Communications and Networking*, 2026, 10: 332-343.
- [7] Peng Xingyu, Hu Xiaoling, Gao Jiabao, et al. Integrated localization and communication for IRS-assisted multi-user mmWave MIMO systems[J]. *IEEE Transactions on Communications*, 2024, 72(8): 4725-4740.
- [8] Adhikary A, Munir M S, Raha A D, et al. Integrated sensing, localization, and communication in holographic MIMO-enabled wireless network: A deep learning approach[J]. *IEEE Transactions on Network and Service Management*, 2024, 21(1): 789-809.
- [9] Chen Xu, He Xinxin, Feng Zhiyong, et al. Joint localization and communication enhancement in uplink integrated sensing and communications system with clock asynchronism[J]. *IEEE Journal on Selected Areas in Communications*, 2024, 42(10): 2659-2673.
- [10] Jiang Fan, Abrardo A, Keykhosravi K, et al. Two-timescale transmission design and RIS optimization for integrated localization and communications[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(12): 8587-8602.
- [11] Yang Helin, Zhang Sheng, Alphones A, et al. An advanced integrated visible light communication and localization system[J]. *IEEE Transactions on Communications*, 2023, 71(12): 7149-7162.
- [12] Li Biwei, Wang Xianbin, Xin Yan, et al. Value of service maximization in integrated localization and communication system through joint resource allocation[J]. *IEEE Transactions on Communications*, 2023, 71(8): 4957-4971.
- [13] Yang Zheyuan, Bi Suzhi, Zhang Y A. Deployment optimization of dual-functional UAVs for integrated localization and communication[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(12): 9672-9687.
- [14] Bi Suzhi, Zhuo Zhenpeng, Lin Xiaohui, et al. Physical-environment-map-aided 3-D deployment optimization for UAV-assisted integrated localization and communication in urban areas[J]. *IEEE Internet of Things Journal*, 2024, 11(9): 15490-15503.
- [15] Wang Li, Li Ruoguang, Xu Lianming, et al. Aerial-ground cooperative vehicular networks for emergency integrated localization and communication[J]. *IEEE Network*, 2023, 37(4): 323-330.
- [16] Wang Zijie, Liu Rongke, Liu Qirui, et al. Feasibility study of UAV-assisted anti-jamming positioning[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(8): 7718-7733.
- [17] Wang Zijie, Liu Rongke, Liu Qirui, et al. Energy-efficient data collection and device positioning in UAV-assisted IoT[J]. *IEEE Internet of Things Journal*, 2020, 7(2): 1122-1139.
- [18] Yang Jimin, Lee J, Lim J. Joint placement and communication optimization of uav base stations in GPS-denied environments[J]. *Journal of Communications and Networks*, 2024, 26(5): 490-501.
- [19] Zhu Qian, Liu Rongke, Wang Zijie, et al. Sensing-communication co-design for UAV swarm-assisted vehicular network in perspective of Doppler[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(2): 2578-2592.
- [20] Zhu Qian, Liu Rongke, Liu Qirui, et al. Resource allocation for UAV swarm-assisted green ISAC networks via multi-agent RL[J]. *IEEE Transactions on Green Communications and Networking*, 2025, 9(3): 1354-1367.
- [21] 高思华, 李军辉, 李建伏, 等. 面向公平性数据采集和能量补充的无人机路径规划算法研究[J]. *电子学报*, 2024, 52(11): 3699-3710.
- Gao Sihua, Li Junhui, Li Jianfu, et al. Research on UAV path planning algorithm for fairness data collection and energy supplement[J]. *Acta Electronica Sinica*, 2024, 52(11): 3699-3710. (in Chinese)
- [22] Frattasi S, Della Rosa F. Mobile positioning and tracking: From conventional to cooperative techniques[M]. New York: John Wiley & Sons, 2017.
- [23] Zekavat S A R, Buehrer R M. Handbook of position loca-

- tion: Theory, practice, and advances[M]. 2nd ed. Hoboken: Wiley, 2018.
- [24] Dong Fuwang, Liu Fan, Cui Yuanhao, et al. Sensing as a service in 6G perceptive networks: A unified framework for ISAC resource allocation[J]. IEEE Transactions on Wireless Communications, 2023, 22(5): 3522-3536.
- [25] Kaune R, Hörst J, Koch W. Accuracy analysis for TDOA localization in sensor networks[C]//14th International Conference on Information Fusion. Piscataway: IEEE, 2011: 1-8.
- [26] Del Peral-Rosado J A, I Castillo R E, Miguez-Sanchez J, et al. Performance analysis of hybrid GNSS and LTE Localization in urban scenarios[C]//2016 8th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing. Piscataway: IEEE, 2016: 1-8.
- [27] 陈平平, 张旭, 谢肇鹏, 等. 基于多智能体近端策略优化的多信道动态频谱接入[J]. 电子学报, 2024, 52(6): 1824-1831.
- Chen Pingping, Zhang Xu, Xie Zhaopeng, et al. Multi-channel dynamic spectrum access based on multi-agent proximal policy optimization[J]. Acta Electronica Sinica, 2024, 52(6): 1824-1831. (in Chinese)
- [28] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[PP/OL]. V2. arXiv (2017-08-28)[2026-04-16]. <https://doi.org/10.48550/arXiv.1707.06347>.
- [29] Cui Jingjing, Liu Yuanwei, Nallanathan A. Multi-agent reinforcement learning-based resource allocation for UAV networks[J]. IEEE Transactions on Wireless Communications, 2020, 19(2): 729-743.

作者简介



杨卓佳 女。现为北京交通大学移动专网网络国家工程研究中心博士研究生。主要研究方向为低空智联网、无人机通信与资源分配等。中国电子学会会员编号:E190202617A。
E-mail: 23111050@bjtu.edu.cn



苏伟 男。现为北京交通大学电子与信息工程学院教授,移动专网网络国家工程研究中心副主任。主要研究方向为新一代信息网络理论与关键技术等。
E-mail: wsu@bjtu.edu.cn



杨斌 男。现为滁州学院计算机与信息工程学院副院长、教授。主要研究方向为网络安全、边缘计算、物联网和无人机网络等。
E-mail: ybcup@chzu.edu.cn



马煜东 男。现为北京航空航天大学人工智能学院博士研究生。主要研究方向为基于L波段数字航空通信系统的导航与监视技术。
E-mail: yudongma1998@buaa.edu.cn



王文迪 男。现为北京交通大学移动专网网络国家工程研究中心硕士研究生。主要研究方向为无人机通感一体化、低空智联网。
E-mail: 24125060@bjtu.edu.cn