

6G语义通信中多无人机中继网络资源分配方法研究

肖力铭 官 铮* 刘 洁 俞纪宏 陈力源

(云南大学信息学院 昆明 650500)

摘 要: 针对6G多无人机中继网络在动态环境下面临的通信资源受限与传输效能瓶颈, 该文引入语义通信范式以降低系统对高带宽的依赖, 并提出一种基于两阶段混合强化学习的联合资源优化方法。首先, 综合考虑语义符号数量、无人机飞行轨迹、功率控制与信道分配, 构建了以最大化语义通信服务质量和长期用户公平性为目标的资源优化模型。其次, 设计了两阶段资源分配框架, 利用改进的容量感知K均值算法进行负载均衡预部署以规避无效探索, 并采用带参数共享的循环独立近端策略优化算法求解动态联合策略。仿真结果表明, 该文所提方法具有良好的收敛性, 能在保证用户公平性的同时有效提升语义传输效率并降低传输时延。

关键词: 语义通信; 多无人机网络; 资源分配; 轨迹优化; 多智能体强化学习

中图分类号: TN929.5; TP181

文献标识码: A

文章编号: 1009-5896(2026)00-0001-11

DOI: 10.11999/JEIT260520

CSTR: 32379.14.JEIT260520

1 引言

随着移动通信技术的快速发展, 第五代移动通信系统已实现商业部署, 但在广域无缝覆盖、频谱利用效率等方面仍面临诸多挑战。为突破传统网络的局限性, 构建空地一体化网络^[1]以实现全球无缝覆盖已成为6G发展的核心愿景与目标之一。在此背景下, 无人机(Unmanned Aerial Vehicle, UAV)通信系统凭借高机动性和灵活部署能力, 成为新一代信息基础设施中扩展地面网络覆盖、保障复杂场景可靠连接的关键载体^[2]。

由UAV集群构成的协同中继网络能够有效支持天地一体化网络中海量地面用户的接入。然而, 该架构在实际部署中面临着数据传输与多机协同的双重挑战。一方面, 有限的频谱资源难以满足实时高负载数据传输和环境感知信息交互需求^[3]; 另一方面, 高动态网络拓扑与多维资源耦合进一步增加了资源协同配置的难度^[4]。因此, 采用新型传输范式提升传输效率, 并结合多智能体分布式决策实现资源协同优化, 成为突破上述局限的关键^[5]。

语义通信通过在发送方提取语义特征并在接收方重建信息^[6-8], 能够有效降低网络传输业务量^[9], 为UAV在受限频谱资源下支持大规模用户接入提供了可行思路。然而, 语义通信特有的基于任务目

标的非线性质量度量特征使得传统基于比特的资源分配方法不再适用^[10]。为了准确量化语义通信的传输效率, 文献^[11,12]提出了语义谱效率(Semantic Spectral Efficiency, S-SE)等新型评价指标, 并设计了相应的语义驱动资源分配方法。但此类方法多基于节点固定与静态信道假设, 难以适应UAV高动态空地信道与多机同频干扰场景, 且未考虑UAV自身的通信资源限制。

目前, 针对UAV语义通信的研究^[13-15]主要集中于系统框架设计与特定任务导向的传输机制, 在多域资源联合分配层面的探索相对匮乏。文献^[16,17]虽利用强化学习实现了单UAV场景下的语义资源优化, 但在广域覆盖与高密度接入场景中, 单UAV系统面临着覆盖范围受限和服务能力瓶颈问题。为此, 采用多UAV组网提供协同中继服务, 成为解决该问题的关键^[18]。文献^[19]探索了感知任务下的多机协同机制, 并提出了一种多模态语义通信框架对用户体质量度和传输成本进行优化。然而, 该工作主要侧重于UAV间的中继场景, 并未解决多UAV作为空中基站服务地面用户的下行链路资源配置问题。在此背景下, 本文提出了一种面向6G的多UAV协同中继语义通信模型, 并构建了一套涵盖语义符号数量、UAV轨迹、频谱信道与发射功率等多维资源的联合优化求解框架, 以实现系统整体效能的提升。

本文的主要贡献为: (1)构建面向多UAV协同中继的语义通信网络联合优化模型。为解决传统指标难以准确刻画语义传输效能的问题, 定义了融合语义效率和传输时延的综合量化指标SC-QoS, 并将其与长期用户公平性结合, 建立联合优化目标函数; (2)提出一种两阶段资源分配优化求解框架。

收稿日期: 2026-04-27; 改回日期: 2026-07-15; 网络出版: 2026-07-25

*通信作者: 官 铮 guanzheng@ynu.edu.cn

基金项目: 云南省基础研究专项重点项目(202601AS070109), 云南大学专业学位研究生实践创新基金项目资助(ZC-252513639)

Foundation Items: Key Project of Yunnan Provincial Fundamental Research Program (202601AS070109), The Yunnan University Professional Degree Graduate Practice Innovation Fund (ZC-252513639)

第一阶段利用改进的容量感知K均值(Capacity-Aware K-means, CA-K-means)算法对UAV执行负载均衡的启发式预部署,第二阶段引入带参数共享的循环独立近端策略优化(Recurrent Independent Proximal Policy Optimization with Parameter Sharing, R-IPPO-PS)算法,实现对优化问题的高效求解;(3)通过仿真验证所提框架在密集接入场景下的性能优势。结果表明,该框架具有较好的收敛性能,能够在保障用户公平性的同时,提高系统语义传输效率并降低平均传输时延。

2 系统模型

本文研究如图1所示的多UAV中继语义通信网络。该网络包含 M 架UAV和 N 个地面用户,分别由集合 $\mathbf{U} = \{u_1, u_2, \dots, u_M\}$ 和 $\mathbf{N} = \{n_1, n_2, \dots, n_N\}$ 进行表示。地面控制中心负责UAV的远程控制与通信调度。UAV作为空中中继,基于正交频分多址技术为地面用户提供语义通信服务。为提升频谱利用率,所有UAV共享通信频谱^[20],因此需要考虑UAV间的同频干扰。在传输过程中,UAV对原始数据进行语义编码,并将提取的语义特征经空地链路发送给地面用户,由用户端完成语义解码与数据恢复。

考虑一个有限通信周期,总服务时间被划分为 T 个时长为 δ_t 的时隙,由集合 $\mathbf{T} = \{t_1, t_2, \dots, t_T\}$ 表示。在三维笛卡尔坐标系中,地面用户分布于半径为 R 的圆形区域内,UAV则部署于该区域上方空域。本文采用离散时间的移动悬停模型刻画UAV与用户的移动性:在每个时隙内,地面用户以较

低速度随机游走以模拟动态场景,UAV根据调度决策移动至目标位置并保持悬停,从而在当前空间拓扑下建立稳定的空地链路。此时,第 m 架UAV的三维坐标表示为 $\mathbf{w}_m[t] = (x_m[t], y_m[t], z_m[t])$;第 n 个用户的坐标为 $\mathbf{w}_n[t] = (x_n[t], y_n[t], 0)$ 。基于上述几何关系,UAV m 到用户 n 的欧氏距离表示为 $d_{m,n}[t] = \|\mathbf{w}_m[t] - \mathbf{w}_n[t]\|$ 。

2.1 信道模型

与传统地面网络相比,UAV空地链路的传播状态主要受空间仰角影响。基于上述空地链路模型,采用概率路径损耗模型^[21]刻画信号传播特性。在时隙 t ,第 m 架UAV至第 n 个用户的自由空间路径损耗为:

$$PL_{m,n}^F[t] = 20 \lg \left(\frac{4\pi f_c d_{m,n}[t]}{c} \right) \quad (1)$$

考虑到实际环境中的障碍物遮挡,空地链路存在视距(Line of Sight, LoS)和非视距(Non Line of Sight, NLoS)两种传输状态,对应的路径损耗分别表示为 $PL_{m,n}^{LoS}[t] = PL_{m,n}^F[t] + \xi_{LoS}$ 和 $PL_{m,n}^{NLoS}[t] = PL_{m,n}^F[t] + \xi_{NLoS}$ 。其中, f_c 为信号载波频率, c 为光速, ξ_{LoS} 和 ξ_{NLoS} 分别是LoS和NLoS的附加损耗。根据空地传播特性,链路处于LoS状态的概率由UAV与地面用户间的仰角决定,定义为:

$$P_{m,n}^{LoS}[t] = \frac{1}{1 + \eta_a \exp \left(-\eta_b \left(\arcsin \left(\frac{z_m[t]}{d_{m,n}[t]} \right) - \eta_a \right) \right)} \quad (2)$$

其中, η_a 和 η_b 为环境相关参数,反映地形对LoS链路的影响。相应地,链路处于NLoS状态的概率为

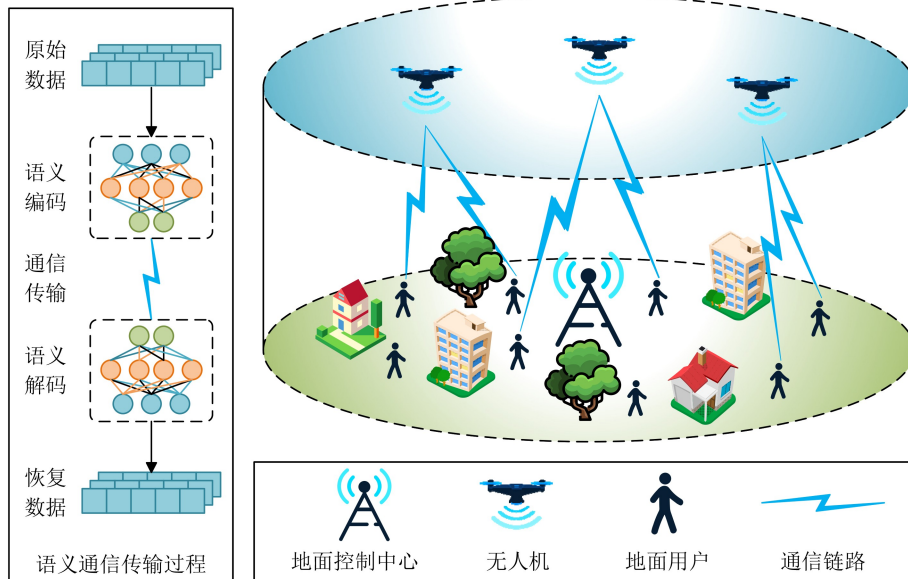


图1 多无人机语义通信场景

$P_{m,n}^{\text{NLoS}}[t] = 1 - P_{m,n}^{\text{LoS}}[t]$ 。综合上述两种状态, UAV m 到用户 n 的概率路径损耗表示为:

$$PL_{m,n}[t] = PL_{m,n}^{\text{F}}[t] + P_{m,n}^{\text{LoS}}[t]\xi_{\text{LoS}} + P_{m,n}^{\text{NLoS}}[t]\xi_{\text{NLoS}} \quad (3)$$

在下行传输中, 将每架 UAV 的可用带宽均等地划分为 S 个正交子信道, 由集合 $\mathbf{S} = \{1, 2, \dots, s, \dots, S\}$ 进行表示。由于采用同频复用, 用户在特定子信道上会受到其他 UAV 的同频干扰^[22]。因此, 用户 n 在子信道 s 上接收 UAV m 信号时的信干噪比(Signal to Interference plus Noise Ratio, SINR)定义为:

$$\text{SINR}_{m,n,s}[t] = \frac{P_{m,s}[t] \cdot 10^{-\frac{PL_{m,n}[t]}{10}}}{\sigma^2 + \sum_{i \in \mathbf{M}, i \neq m} P_{i,s}[t] \cdot 10^{-\frac{PL_{i,n}[t]}{10}}} \quad (4)$$

其中, $P_{m,s}[t]$ 表示 UAV m 在子信道 s 上的发射功率, 满足 $0 \leq \sum_{s=1}^S P_{m,s}[t] \leq P_m^{\text{max}}$, P_m^{max} 是 UAV m 的最大发射功率。 $\sigma^2 = n_0 \cdot B_{\text{sub}}$ 为加性高斯白噪声功率, n_0 为热噪声功率谱密度, B_{sub} 为每个子信道的带宽。

2.2 语义通信性能模型

由于系统吞吐量与误码率等传统基于比特的评估指标难以准确刻画语义传输效能, 本文构建了一套面向 UAV 集群协同网络的综合性能评估模型。该模型从语义传输效率与传输公平性两个维度出发, 旨在全面衡量系统在高动态环境下的资源分配质量与长期保障能力。

在语义传输效率方面, 为准确衡量单位语义符号资源所承载并被成功恢复的有效语义信息量, 本文首先定义广义语义谱效率(Generalized Semantic Spectral Efficiency, G-SSE) Φ_G 为:

$$\Phi_G = \frac{\tilde{H}}{K_s} \cdot \xi \quad (5)$$

其中, K_s 表示完成该传输任务所消耗的总语义符号数, $\xi \in [0, 1]$ 表示接收端的语义保真度, $\tilde{H}$ 为近似语义熵, 用于量化满足特定语义任务所需的最小语义信息量^[12]。

G-SSE 可用于统一刻画不同任务场景下的语义传输效率。对于包含 L_w 个单词的文本语义传输任务, 总语义符号数可表示为 $K_s = k \cdot L_w$, 其中 k 为每个单词编码后的平均语义符号数。基于语义密度 I/L 估算^[11], 得到近似语义熵为 $\tilde{H} \approx (I/L) \cdot L_w$ 。将其代入式(5), 可得:

$$\Phi_G = \frac{(I/L) \cdot L_w}{k \cdot L_w} \cdot \xi = \frac{I/L}{k} \cdot \xi = \Phi_{\text{S-SE}} \quad (6)$$

由上式可知, 在文本语义传输场景下, G-SSE

可退化为传统 S-SE 指标^[11], 表明 G-SSE 与已有语义效率评价方法兼容。

然而, 由于 G-SSE 缺乏对物理信道边界的显式约束, 直接对其优化难以收敛。为此, 定义语义量化效率(Semantic Quantization Efficiency, SQE) η_{se} , 对语义传输效率进行约束化表征:

$$\eta_{\text{se}} = \begin{cases} \frac{k_{\text{th}}}{k} \cdot \xi, & k \geq k_{\text{th}} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

其中, 语义保真度 $\xi = f(k, \text{SINR})$ 由离线查找表经双线性插值确定, 该查找表基于 DeepSC 模型^[23]对不同 SINR 和语义符号数条件下的 ξ 值进行离线实测构建。 k_{th} 为语义符号阈值, 与信道状态强相关, $k_{\text{th}} = \min\{k | f(k, \text{SINR}) \geq \xi_{\text{th}}\}$ 表示在当前信道条件下, 使语义保真度达到目标阈值 ξ_{th} 所需的最小语义符号数。因此, SQE 可视为在可靠性约束下对语义传输效率的归一化表征。

考虑到实际网络中低时延传输需求, 进一步将传输整个任务所需时间定义为传输时延 t_{tr} 。该时延由总的语义符号数 K_s 与符号速率 B_{sub} 决定:

$$t_{\text{tr}} = \frac{K_s}{B_{\text{sub}}} = \frac{k \cdot L_w}{B_{\text{sub}}} \quad (8)$$

为实现与 SQE 的统一度量, 采用极差法对传输时延进行归一化处理:

$$\tilde{t}_{\text{tr}} = \frac{t_{\text{tr}} - t_{\text{tr}}^{\min}}{t_{\text{tr}}^{\max} - t_{\text{tr}}^{\min}} \quad (9)$$

综合考虑上述语义效率和传输时延, 定义语义通信服务质量(Quality of Service for Semantic Communication, SC-QoS) Q 为 SQE 奖励与时延惩罚的加权和:

$$Q = w_1 \cdot \eta_{\text{se}} - w_2 \cdot \tilde{t}_{\text{tr}} \quad (10)$$

其中, w_1 和 w_2 分别是 SQE 和传输时延的权重系数。

在传输公平性方面, 为保障多用户间的长期服务均衡, 本文基于比例公平(Proportional Fairness, PF)调度窗口内的历史服务质量, 采用 Jain 公平性指数^[24]衡量各用户语义服务水平的均衡程度:

$$\mathcal{F}[t] = \frac{\left(\sum_{n=1}^N \bar{\eta}_{\text{se},n}[t]\right)^2}{N \cdot \sum_{n=1}^N (\bar{\eta}_{\text{se},n}[t])^2} \quad (11)$$

其中, $\bar{\eta}_{\text{se},n}[t]$ 表示用户 n 在时隙 t 的历史平均 SQE。该历史平均值采用指数加权移动平均方式更新:

$$\bar{\eta}_{\text{se},n}[t] = \left(1 - \frac{1}{T_c}\right) \bar{\eta}_{\text{se},n}[t-1] + \frac{1}{T_c} \cdot \eta_{\text{se},n}[t] \quad (12)$$

其中, T_c 是PF调度窗口大小, $\eta_{se,n}[t]$ 为用户 n 在时隙 t 获得的瞬时SQE。

2.3 问题建模

本文建立了一个多维资源优化模型, 旨在面向6G语义通信的多UAV中继网络, 在整个时隙周期 T 内, 最大化系统所有用户的平均SC-QoS和全局公平性的加权和。令 $\mathbf{A} = \{\mathbf{w}_m, \alpha_{m,n,s}, P_{m,s}, k\}$ 为所有UAV的联合决策变量集合, 包括UAV轨迹 $\mathbf{w}_m$ 、信道分配 $\alpha_{m,n,s}$ 、UAV m 在子信道 s 上的发射功率 $P_{m,s}$ 和平均每词语义符号数 k 。基于上述定义, 建立该问题的最大化优化目标函数如下:

$$P1: \max_{\mathbf{A}} \frac{1}{T} \sum_{t=1}^T \left(w_{sc} \cdot \left(\frac{1}{N} \sum_{n=1}^N Q_n[t] \right) + w_f \cdot \mathcal{F}[t] \right) \quad (13)$$

其中, w_{sc} 和 w_f 分别是SC-QoS和公平性的权重系数。 $Q_n[t]$ 表示 t 时刻用户 n 能够得到的SC-QoS总和, $\mathcal{F}[t]$ 是 t 时刻全局的公平性指数。该优化问题的约束条件表述为:

$$\begin{aligned} C1: & \alpha_{m,n,s}[t] \in \{0, 1\}, \forall m, n, s, t \\ C2: & \sum_{n \in \mathbf{N}} \sum_{s \in \mathbf{S}} \alpha_{m,n,s}[t] \leq S, \forall m, t \\ C3: & \sum_{n \in \mathbf{N}} \alpha_{m,n,s}[t] \leq 1, \forall m, s, t \\ C4: & 0 \leq \sum_{s=1}^S P_{m,s}[t] \leq P_m^{\max}, \forall m, t \\ C5: & \|\mathbf{w}_m[t+1] - \mathbf{w}_m[t]\| \leq V_{\max}^{\text{uav}} \cdot \delta_t, \\ & t = 1, 2, \dots, \forall m \in \mathbf{U} \\ C6: & \|\mathbf{w}_i[t] - \mathbf{w}_j[t]\| \geq d_{\min}^{\text{uav}}, t = 1, 2, \dots, \forall i \neq j, i, j \in \mathbf{U} \end{aligned} \quad (14)$$

其中, $\alpha_{m,n,s}[t]$ 表示UAV m 是否为地面用户 n 分配子信道 s 的指示符, $V_{\max}^{\text{uav}}$ 为UAV最大移动速度, $d_{\min}^{\text{uav}}$ 为UAV之间最小间隔距离。约束C1给定信道分配指示符的取值范围; 约束C2确保每个UAV分配的子信道数不超过最大可分配子信道数 S ; 约束C3确保每架UAV在同一子信道上最多服务一个用户; 约束C4是UAV的最大发射功率约束; 约束C5和C6分别是UAV的轨迹应满足的最大速度约束和避免碰撞约束。

3 算法设计

上述优化问题属于混合整数非线性规划(Mixed Integer Nonlinear Programming, MINLP)问题, 具有动作空间高维混合、环境动态时变及变量强耦合等特征, 传统优化方法难以实现面向长期规划的高效求解。针对此挑战, 提出一种两阶段混合强化学习(Two-Stage Hybrid Reinforcement Learning,

TS-HRL)资源分配框架。所提算法的整体框架如图2所示。

3.1 基于CA-K-means的UAV启发式预部署

为缓解UAV随机部署导致的算法收敛缓慢和易陷局部最优的问题, 首先将UAV的初始部署建模为兼顾空地通信距离与网络负载均衡的空间聚类问题。标准K-means算法以UAV数量作为聚类簇数, 能够在一定程度上缩短用户与簇中心之间的距离, 但由于缺乏对节点服务容量的感知, 容易引发网络局部过载。为此, 本文引入容量感知的用户重分配机制, 提出一种CA-K-means启发式预部署算法, 在空间距离和UAV剩余容量之间进行联合权衡。具体而言, 根据单架UAV的最大理论服务容量 $C_{\max}$, 每次迭代识别过载簇, 并依据最小加权距离准则, 将边缘用户 n 重新分配给未满载的候选簇 j 。加权距离 $D_w(n, j)$ 的公式表述如下:

$$D_w(n, j) = D_{n,j} - w_{\text{cap}} \cdot (C_{\max} - |\mathbf{N}_j|) \quad (15)$$

其中, $D_{n,j}$ 为用户 n 到候选簇中心 j 的物理距离, w_{cap} 为容量权重因子, $|\mathbf{N}_j|$ 为簇 j 当前已接入用户数。通过引入基于剩余容量的动态补偿项, 实现了通信距离与节点负载的联合度量。候选簇越空闲, 获得的距离补偿越大, 从而赋予低负载UAV更高的接入优先级。该机制能够有效引导过载区域的边缘用户重分配至空闲节点, 在兼顾空间邻近性的同时, 实现UAV负载均衡, 优化动态寻优起始状态, 有效提升算法的收敛速度。

3.2 Dec-POMDP建模

优化问题P1可视为一个分散式部分可观测马尔可夫决策过程(Decentralized Partially Observable Markov Decision Process, Dec-POMDP)^[25]。针对该优化问题的高维混合动作空间与多智能体协作特性, 本文对模型核心元素的设置如下。

(1) 状态与观测空间: 全局状态 $\mathbf{s}[t]$ 包含系统中所有节点的位置、速度与瞬时信道状态矩阵。考虑到动态网络拓扑下观测维度可能随用户数量变化, 将智能体 m 的局部观测 $\mathbf{o}_m[t]$ 设计为归一化定长向量:

$$\mathbf{o}_m[t] = [\mathbf{o}_{\text{self},m}[t], \mathbf{o}_{\text{uav},m}[t], \mathbf{o}_{\text{cand},m}[t], \mathbf{o}_{\text{other},m}[t], \mathbf{o}_{\text{pwr},m}[t-1]] \quad (16)$$

其中, $\mathbf{o}_{\text{self},m}[t]$ 为UAV自身状态分量, 由位置坐标和速度矢量组成; $\mathbf{o}_{\text{uav},m}[t]$ 为其他UAV信息分量, 包含平均相对位置、平均距离和最近两架UAV的相对位置; $\mathbf{o}_{\text{cand},m}[t]$ 为候选用户状态分量, 包含各候选用户相对位置与历史平均语义速率; $\mathbf{o}_{\text{other},m}[t]$ 为非候选用户聚合分量, 包含其数量占比、平均相

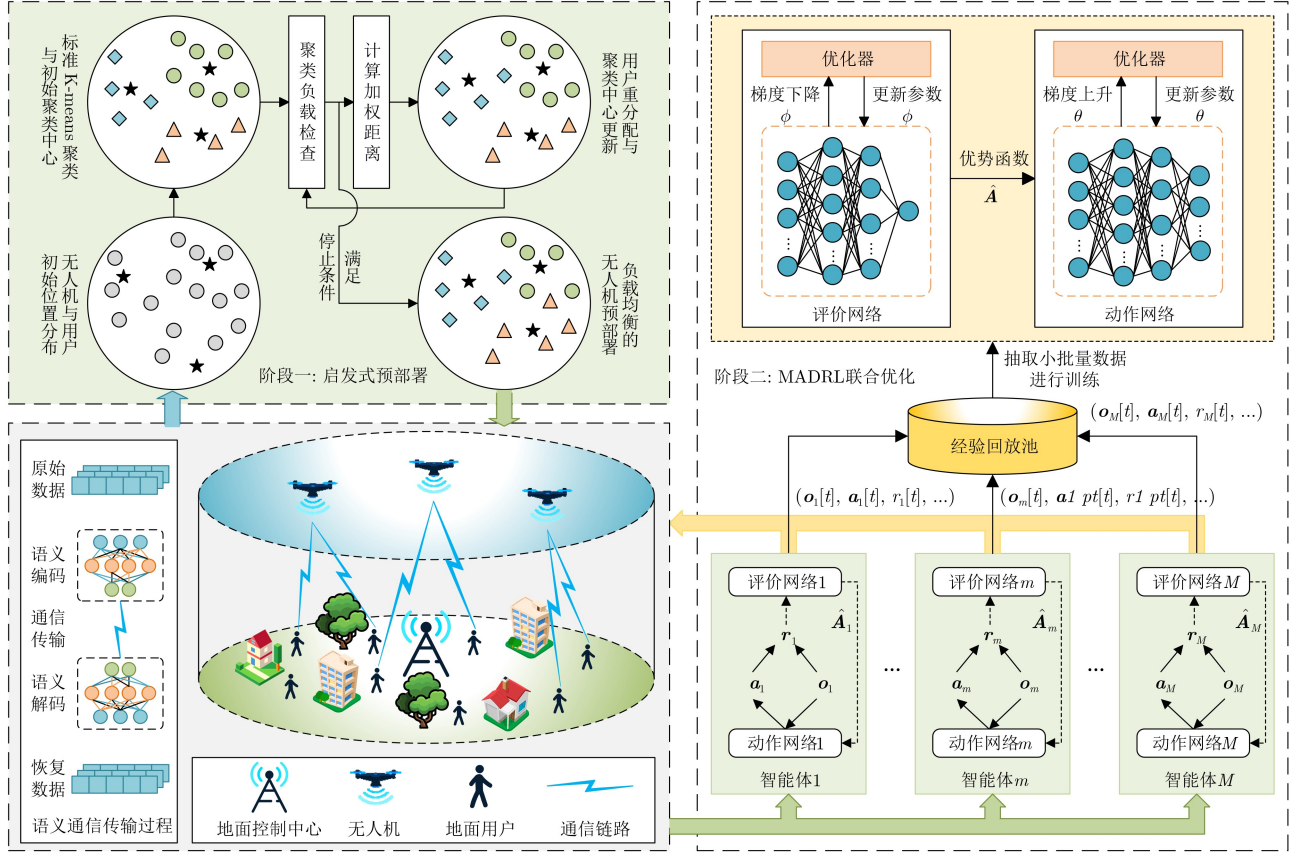


图2 两阶段算法框架

对位置和平均语义速率； $\mathbf{o}_{\text{pwr},m}[t-1]$ 为功率状态分量，由上一步各子信道发射功率组成。

(2)动作空间与偏好解耦：智能体的联合动作空间存在离散与连续变量强耦合的挑战。为此，本文将智能体 m 的动作 $\mathbf{a}_m$ 统一映射为定长的连续向量 $\mathbf{a}_m = [\mathbf{a}_{\text{vel}}, \mathbf{a}_{\text{pwr}}, \mathbf{a}_{\text{assoc}}, \mathbf{a}_k]$ 。该向量包含了速度控制 $\mathbf{a}_{\text{vel}}$ 、功率分配 $\mathbf{a}_{\text{pwr}}$ 、用户关联偏好 $\mathbf{a}_{\text{assoc}}$ 及语义符号数控制 $\mathbf{a}_k$ 四个连续分量。

在处理离散变量时，本文引入连续松弛与偏好解耦机制。对于离散的语义符号数 k ，通过线性映射与取整操作将 $\mathbf{a}_k$ 离散化，从而在一定程度上保留评价网络的梯度信息。对于高维信道分配问题，智能体仅需输出针对距离最近的 N_{cand} 个候选用户的连续关联偏好 $\mathbf{a}_{\text{assoc}}$ ，而非完整的高维离散信道分配矩阵。随后，中央控制器收集各UAV的关联偏好，并按最大偏好原则确定用户关联拓扑。各UAV针对其关联用户，通过PF调度器完成最终的信道分配。PF度量的优先权值 $\rho_{n,s}[t]$ 定义如下：

$$\rho_{n,s}[t] = \frac{\eta_{\text{se},n,s}[t]}{\bar{\eta}_{\text{se},n}[t]} \quad (17)$$

其中， $\eta_{\text{se},n,s}[t]$ 为信道 s 可为用户 n 提供的瞬时SQE， $\bar{\eta}_{\text{se},n}[t]$ 为用户 n 的历史平均SQE。该调度机

制优先保障用户基本覆盖，并将剩余信道资源分配给PF权值较高的链路，以兼顾传输效率与服务公平性。

(3)奖励函数：为引导多智能体在满足物理约束的同时优化系统效用与服务公平性，将智能体 m 的即时奖励 $r_m[t]$ 设计为局部效用、全局公平性与约束惩罚的加权和：

$$r_m[t] = Q_{m,\text{local}}[t] + r_{\text{fairness}}[t] - p_{\text{power},m}[t] - p_{\text{cluster},m}[t] \quad (18)$$

其中， $Q_{m,\text{local}}[t]$ 为智能体 m 的局部SC-QoS， $r_{\text{fairness}}[t]$ 为全局共享公平性奖励， $p_{\text{power},m}[t]$ 与 $p_{\text{cluster},m}[t]$ 分别为功率超限与碰撞惩罚，用于限制最大发射功率并避免碰撞。

3.3 基于R-IPPO-PS的算法求解

Dec-POMDP模型在刻画多UAV网络的分布式特征的同时，引入了局部观测受限等复合挑战。为此，本文提出一种基于R-IPPO-PS的联合资源分配算法。算法流程如[算法1](#)所示。

针对局部观测 $\mathbf{o}_m[t]$ 缺乏全局马尔可夫性的问题，本文设计了一个包含多层感知机(Multilayer Perceptron, MLP)与长短期记忆(Long Short-Term Memory, LSTM)网络的特征处理前端。观测数据

算法 1 基于R-IPPO-PS的联合资源分配算法

输入：最大回合数 N_{eps} ，回合内时隙数 T ，每次参数更新迭代数 N_{epoch}

输出：优化后的共享网络参数 θ, ϕ

- (1) 初始化参数 θ, ϕ ，同步旧策略参数 $\theta_{\text{old}} \leftarrow \theta$ ，初始化经验回放池 $\mathcal{D}$
- (2) 对于每个回合 $e \in \{1, 2, \dots, N_{\text{eps}}\}$:
- (3) 获取初始观测 $\mathbf{o}_m[0]$ ，初始化LSTM的隐藏状态 $\mathbf{h}_m[0]$ 与细胞状态 $\mathbf{c}_m[0]$
- (4) 对于每个时隙 $t \in \{1, 2, \dots, T\}$:
- (5) 通过MLP与LSTM更新隐藏状态 $(\mathbf{h}_m[t], \mathbf{c}_m[t])$ ，采样独立动作 $\mathbf{a}_m[t]$
- (6) 执行联合动作，获取奖励 $r_m[t]$ 与下一时刻观测 $\mathbf{o}_m[t+1]$
- (7) 结束循环
- (8) 获取本回合的交互轨迹 τ_m ，采用GAE计算优势函数 $\hat{A}[t]$ ，并将完整交互数据存入 $\mathcal{D}$
- (9) 对于每次训练迭代 $j \in \{1, 2, \dots, N_{\text{epoch}}\}$:
- (10) 从 $\mathcal{D}$ 中抽取小批量数据，最大化总目标函数 $J(\theta, \phi)$ ，联合更新参数 θ 与 ϕ
- (11) 结束循环
- (12) 更新旧策略参数 $\theta_{\text{old}} \leftarrow \theta$ ，并清空经验回放池 $\mathcal{D}$
- (13) 结束循环

经MLP提取特征后，结合前一时刻动作 $\mathbf{a}_m[t-1]$ 与奖励 $r_m[t-1]$ 构成联合输入 $\mathbf{x}_m[t]$ ，利用LSTM内部细胞状态 $\mathbf{c}_m[t]$ 累积历史轨迹：

$$(\mathbf{h}_m[t], \mathbf{c}_m[t]) = \text{LSTM}(\mathbf{x}_m[t], \mathbf{h}_m[t-1], \mathbf{c}_m[t-1]; \theta_{\text{rnn}}) \quad (19)$$

输出的隐藏状态 $\mathbf{h}_m[t]$ 作为环境的时序紧凑表征送入Actor-Critic架构，分别生成策略分布 $\pi(\cdot|\mathbf{h}_m[t])$ 与状态价值 $V_\phi(\mathbf{h}_m[t])$ ，从而赋予智能体在信息缺失下的长期规划能力。

为克服传统多智能体算法随UAV数量增加而引发的维度灾难，算法采用多智能体参数共享机制。在训练时所有UAV共享动作网络参数 θ 与评价网络参数 ϕ ，交互产生的经验轨迹 τ_m 汇入全局回放池 $\mathcal{D}$ 集中更新。该机制解耦了模型复杂度与智能体数量，有效提升了样本利用率与训练效率。在执行时，各UAV依赖实时的局部观测 $\mathbf{o}_m[t]$ 与独立维护的隐藏状态 $\mathbf{h}_m[t]$ 实现差异化的协同行为。

在策略迭代环节，算法依托PPO裁剪机制限制策略更新幅度以克服并发探索带来的环境非平稳性。动作网络目标函数 $\mathcal{L}^{\text{CLIP}}(\theta)$ 设计为：

$$\begin{aligned} \mathcal{L}^{\text{CLIP}}(\theta) &= \mathbb{E} \left[\min \left(r[t](\theta) \hat{A}[t], \text{clip}(r[t](\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}[t] \right) \right] \end{aligned} \quad (20)$$

其中， $r[t](\theta)$ 为新旧策略概率比， $\hat{A}[t]$ 为利用广义

优势估计 (Generalized Advantage Estimation, GAE) 计算的优势函数， ϵ 为裁剪阈值。同时，评价网络通过最小化均方误差 $\mathcal{L}^{\text{VF}}(\phi)$ 逼近真实回报 $R[t]$ ：

$$\mathcal{L}^{\text{VF}}(\phi) = \mathbb{E}[(V_\phi(\mathbf{h}_m[t]) - R[t])^2] \quad (21)$$

结合用于鼓励探索的策略熵 $S[\pi_\theta]$ ，端到端的联合优化总目标函数 $J(\theta, \phi)$ 表示为：

$$J(\theta, \phi) = \mathcal{L}^{\text{CLIP}}(\theta) - c_1 \mathcal{L}^{\text{VF}}(\phi) + c_2 S[\pi_\theta] \quad (22)$$

其中， c_1, c_2 为权重系数。

4 仿真分析

仿真场景设定为一个半径为200 m的圆形区域，区域内默认随机分布有20个地面用户，以小于5 m/s的速度进行随机游走。在固定200 m的飞行高度部署了3架UAV进行水平机动，为地面用户提供中继服务。每架UAV的最大飞行速度为20 m/s，最大发射功率为0.5 W，分配有5个带宽为5 MHz的正交子信道。仿真关键参数如表1所示，训练参数经调优后设置如表2所示。

本文将所提方案与以下基准方案进行性能对比：(1)初始位置随机部署方案。该方案将UAV的初始位置随机部署，随后使用R-IPPO-PS算法进行

表 1 仿真参数设置

参数	参数值
载波频率	$f_c = 2.4 \text{ GHz}$
热噪声功率谱密度	$n_0 = 1.27 \times 10^{-20} \text{ W/Hz}$
LoS路径损耗指数	$\eta_a = 9.61$
NLoS路径损耗指数	$\eta_b = 0.16$
语义保真度阈值	$\xi_{\text{th}} = 0.8$
每回合最大步数	$T=600$
时隙持续时间	$\delta_t = 0.1 \text{ s}$

表 2 训练参数设置

参数	参数值
MLP隐藏层	[256, 256, 128]
MLP激活函数	ReLU
LSTM隐藏单元数	256
LSTM序列长度	40
学习率	$\alpha = 1.5 \times 10^{-5}$
折扣因子	$\gamma = 0.85$
GAE参数	$\lambda = 0.95$
PPO裁剪阈值	$\epsilon = 0.09$
SC-QoS权重系数	$w_1 = 1.0, w_2 = 1.0$
优化目标权重系数	$w_{\text{sc}} = 1.0, w_{\text{f}} = 1.0$

优化。(2)无记忆资源分配方案。该方案移除了TS-HRL架构中的LSTM时序层,智能体仅依赖当前时刻的局部观测信息进行决策。(3)基于EI-SAC的资源分配方案。该方案使用文献[26]中提出的增强型独立SAC算法进行优化。(4)传统启发式优化方案。该方案采用传统的分离式优化思想,利用K-means聚类引导UAV的初始部署[27]与动态追踪[28],结合等功率分配策略[27,29]控制发射功率,并基于贪心算法[28,29]完成信道与语义符号的匹配,以局部最大化系统效用。(5)基于TS-HRL的传统通信方案。该方案使用TS-HRL算法对传统比特通信进行优化,而不进行语义特征提取。

图3展示了训练期间UAV平均累积奖励的收敛趋势。所提TS-HRL算法具有较高的初始回报,且收敛曲线较为平滑,能够在较少的迭代次数内稳定收敛至较高的性能水平。相比之下,随机部署方案由于缺乏预部署引导,前期探索效率较低;无记忆方案因缺乏时序建模引发后期震荡;EI-SAC则受离线更新局限陷入局部最优。上述结果表明,CA-K-means预部署与R-IPPO-PS时序决策机制有助于提升算法的收敛效率和训练稳定性。

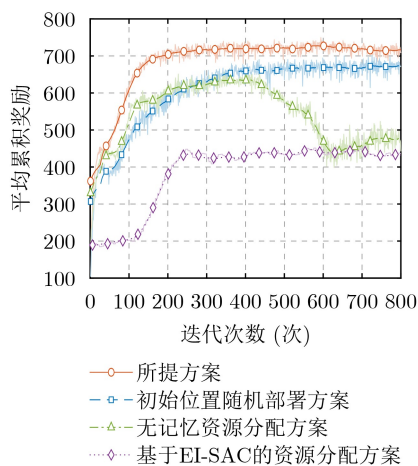


图3 各方案在训练中获得奖励

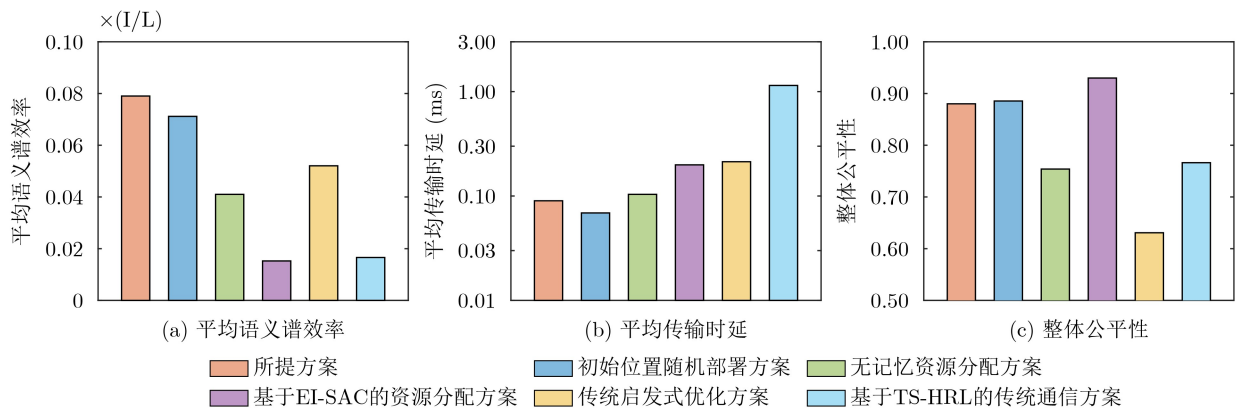


图4 性能评估结果

为实现语义通信与传统通信的等效评估,本文选取了S-SE^[11]、传输时延和系统公平性作为主要的性能评估指标。评估时,仅针对成功传输的任务计算时延,传输失败的代价则反映为S-SE性能的下降。

图4对比了各方案在上述三个评价指标上的性能表现。与传统通信方案相比,本文所提方案通过语义特征提取减少冗余传输,使S-SE提升约3.6倍,传输时延降低92.1%,整体公平性提升14.3%。在语义通信基准方案中,所提方案相较无记忆方案与传统启发式方案,S-SE分别提升92.7%和51.9%,时延分别降低13.5%和57.7%,公平性分别提升17.3%和39.7%。与随机部署方案相比,所提方案在公平性仅小幅下降1.1%的情况下,实现了11.3%的S-SE提升,体现出更优的效率与公平性权衡。EI-SAC算法虽然获得了0.93的公平性指标,但其S-SE相对较低;相比之下,本方案在保持较高公平性的同时,S-SE约为EI-SAC的4.3倍。

为进一步验证系统在不同网络负载下的可扩展性,图5给出了用户数量从10增加至40时的性能表现。随着网络负载加重,多数基准方案的S-SE与公平性均显著下降。相比之下,所提方案采用自适应调度与全局公平性奖励机制,有效缓解了性能衰减,且时延始终稳定在0.1 ms以内。该结果表明,所提方案在不同用户规模下均能保持较好的通信性能表现,具有一定的负载适应能力。

5 结束语

本文针对6G多UAV中继语义通信场景中的资源受限与动态协同优化问题,构建了以系统SC-QoS与长期公平性为目标的资源联合优化模型。针对该MINLP问题的高维耦合特性,提出了一种TS-HRL资源分配框架,通过CA-K-means预部署与R-IPPO-PS动态调度算法的协同,实现了多UAV轨迹、发射功率、信道分配和语义符号数量的联合优化。仿

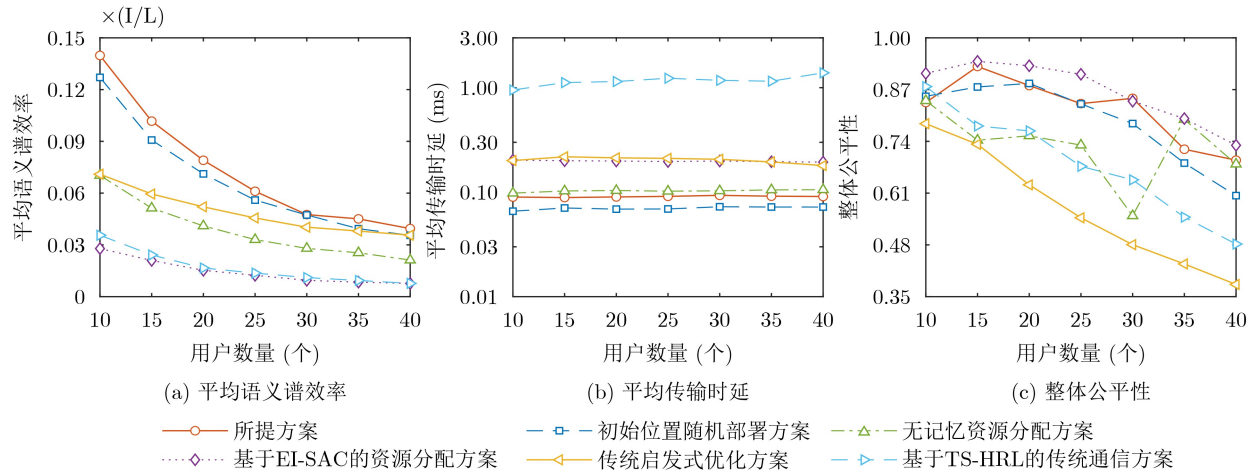


图5 不同用户数量性能表现

真结果表明, 所提方法具有良好的收敛性能, 能够在保障用户公平性的同时提高系统S-SE并降低传输时延。该研究为6G空天地一体化网络中UAV辅助语义通信的智能资源分配提供了一种可行思路。

参考文献

- [1] ZHANG Ruichen, DU Hongyang, NIYATO D, *et al.* Generative AI for space-air-ground integrated networks[J]. *IEEE Wireless Communications*, 2024, 31(6): 10–20. doi: [10.1109/MWC.016.2300547](#).
- [2] FENG Zikai, WU Di, HUANG Mengxing, *et al.* Graph-attention-based reinforcement learning for trajectory design and resource assignment in multi-UAV-assisted communication[J]. *IEEE Internet of Things Journal*, 2024, 11(16): 27421–27434. doi: [10.1109/JIOT.2024.3397823](#).
- [3] XIE Ping, CHEN Qian, WU Jingyan, *et al.* NOMA-based intelligent resource allocation and trajectory optimization for multi-UAVs assisted semantic communication networks[J]. *Ad Hoc Networks*, 2025, 171: 103762. doi: [10.1016/j.adhoc.2025.103762](#).
- [4] 张广驰, 顾泽霖, 崔苗. 空地协同通信感知一体化系统的轨迹与资源分配联合优化[J]. *电子与信息学报*, 2024, 46(6): 2382–2390. doi: [10.11999/JEIT230716](#).
ZHANG Guangchi, GU Zelin, and CUI Miao. Joint trajectory and resource allocation optimization for air-ground collaborative integrated sensing and communication systems[J]. *Journal of Electronics & Information Technology*, 2024, 46(6): 2382–2390. doi: [10.11999/JEIT230716](#).
- [5] 冯斯梦, 张云弈, 刘凯, 等. 低空混合障碍下无人机协同多智能体轨迹规划[J]. *电子与信息学报*, 2025, 47(5): 1291–1300. doi: [10.11999/JEIT250012](#).
FENG Simeng, ZHANG Yunyi, LIU Kai, *et al.* Collaborative multi-agent trajectory optimization for unmanned aerial vehicles under low-altitude mixed-obstacle airspace[J]. *Journal of Electronics & Information Technology*, 2025, 47(5): 1291–1300. doi: [10.11999/JEIT250012](#).
- [6] 张平, 牛凯, 姚圣时, 等. 面向未来的语义通信: 基本原理与实现方法[J]. *通信学报*, 2023, 44(5): 1–14. doi: [10.11959/j.issn.1000-436x.2023079](#).
ZHANG Ping, NIU Kai, YAO Shengshi, *et al.* Semantic communications for future: Basic principle and implementation methodology[J]. *Journal on Communications*, 2023, 44(5): 1–14. doi: [10.11959/j.issn.1000-436x.2023079](#).
- [7] QIN Zhijin, GAO Feifei, LIN Bo, *et al.* A generalized semantic communication system: From sources to channels[J]. *IEEE Wireless Communications*, 2023, 30(3): 18–26. doi: [10.1109/MWC.013.2200553](#).
- [8] FU Yuzhou, CHENG Wenchu, ZHANG Wei, *et al.* Scalable extraction based semantic communication for 6G wireless networks[J]. *IEEE Communications Magazine*, 2024, 62(7): 96–102. doi: [10.1109/MCOM.021.2300269](#).
- [9] DING Yahao, YANG Zhaohui, PHAM Q V, *et al.* Distributed machine learning for UAV swarms: Computing, sensing, and semantics[J]. *IEEE Internet of Things Journal*, 2024, 11(5): 7447–7473. doi: [10.1109/JIOT.2023.3341307](#).
- [10] WANG Lingyi, WU Wei, ZHOU Fuhui, *et al.* Adaptive resource allocation for semantic communication networks[J]. *IEEE Transactions on Communications*, 2024, 72(11): 6900–6916. doi: [10.1109/TCOMM.2024.3405355](#).
- [11] YAN Lei, QIN Zhijin, ZHANG Rui, *et al.* Resource allocation for text semantic communications[J]. *IEEE Wireless Communications Letters*, 2022, 11(7): 1394–1398. doi: [10.1109/LWC.2022.3170849](#).
- [12] YAN Lei, QIN Zhijin, LI Chunfeng, *et al.* QoE-based semantic-aware resource allocation for multi-task networks[J]. *IEEE Transactions on Wireless Communications*, 2024, 23(9): 11958–11971. doi: [10.1109/](#)

- TWC.2024.3386807.
- [13] XU Jiaqi, YAO Haipeng, ZHANG Ru, *et al.* Federated learning powered semantic communication for UAV swarm cooperation[J]. *IEEE Wireless Communications*, 2024, 31(4): 140–146. doi: [10.1109/MWC.005.2300147](#).
- [14] HOA N T, HAI C T T, LE HUNG H, *et al.* Joint edge computing and semantic communication in UAV-enabled networks[J]. *IEEE Communications Letters*, 2025, 29(1): 80–84. doi: [10.1109/LCOMM.2024.3496534](#).
- [15] DENG Yundi, LENG Supeng, ZHANG Ke, *et al.* Semantic communication assisted cooperative sensing in UAV networks[C]. Proceedings of 2025 10th International Conference on Computer and Communication System (ICCCS), Chengdu, China, 2025: 862–867. doi: [10.1109/ICCCS65393.2025.11069451](#).
- [16] SI Peiyuan, ZHAO Jun, LAM K Y, *et al.* UAV-assisted semantic communication with hybrid action reinforcement learning[C]. Proceedings of 2023 IEEE Global Communications Conference, Kuala Lumpur, Malaysia, 2023: 3801–3806. doi: [10.1109/GLOBECOM54140.2023.10437643](#).
- [17] XIE Wancheng, YANG Helin, and XIONG Zehui. Resource allocation for UAV-assisted anti-jamming semantic D2D networks: A graph reinforcement learning approach[J]. *Computer Networks*, 2025, 269: 111463. doi: [10.1016/j.comnet.2025.111463](#).
- [18] 时统志, 李博, 杨洪娟, 等. 基于传输公平性的多无人机通感一体化空间部署与波束成形设计[J]. *电子与信息学报*, 2025, 47(1): 57–65. doi: [10.11999/JEIT240590](#).
SHI Tongzhi, LI Bo, YANG Hongjuan, *et al.* Spatial deployment and beamforming for design multi-unmanned aerial vehicle-integrated sensing and communication based on transmission fairness[J]. *Journal of Electronics & Information Technology*, 2025, 47(1): 57–65. doi: [10.11999/JEIT240590](#).
- [19] HU Han, ZHU Xingwu, ZHOU Fuhui, *et al.* Resource allocation for multi-modal semantic communication in UAV collaborative networks[J]. *IEEE Transactions on Communications*, 2025, 73(9): 7599–7616. doi: [10.1109/TCOMM.2025.3552303](#).
- [20] YANG Helin, ZHAO Jun, NIE Jiangtian, *et al.* UAV-assisted 5G/6G networks: Joint scheduling and resource allocation based on asynchronous reinforcement learning[C]. Proceedings of IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Vancouver, Canada, 2021: 1–6. doi: [10.1109/INFOCOMWKSHPS51825.2021.9484604](#).
- [21] CHEN Ying, YANG Yaozong, WU Yuan, *et al.* Joint trajectory optimization and resource allocation in UAV-MEC systems: A lyapunov-assisted DRL approach[J]. *IEEE Transactions on Services Computing*, 2025, 18(2): 854–867. doi: [10.1109/TSC.2025.3544124](#).
- [22] TUSHA A and ARSLAN H. Interference burden in wireless communications: A comprehensive survey from PHY layer perspective[J]. *IEEE Communications Surveys & Tutorials*, 2025, 27(4): 2204–2246. doi: [10.1109/COMST.2024.3487068](#).
- [23] XIE Huiqiang, QIN Zhijin, LI G Y, *et al.* Deep learning enabled semantic communication systems[J]. *IEEE Transactions on Signal Processing*, 2021, 69: 2663–2675. doi: [10.1109/TSP.2021.3071210](#).
- [24] CHRYSOLOGOU A P, CHATZIDIAMANTIS N D, and KARAGIANNIDIS G K. Ergodic rate and fairness analysis of uplink RSMA with randomly deployed users[J]. *IEEE Wireless Communications Letters*, 2025, 14(11): 3764–3768. doi: [10.1109/LWC.2025.3603208](#).
- [25] TANG Qinqin, XIE Renchao, YU F R, *et al.* Decentralized computation offloading in IoT fog computing system with energy harvesting: A Dec-POMDP approach[J]. *IEEE Internet of Things Journal*, 2020, 7(6): 4898–4911. doi: [10.1109/JIOT.2020.2971323](#).
- [26] LI Rongyu, LU Wei, and CHU Jinghui. Multi-objective path planning for UAV swarm-assisted wireless information and power transfer system based on decentralized MADRL[J]. *IEEE Open Journal of the Communications Society*, 2025, 6: 8539–8555. doi: [10.1109/OJCOMS.2025.3615969](#).
- [27] CAI Xuli, LOHAN P, and KANTARCI B. Multi-agent deep reinforcement learning for optimized multi-UAV coverage and power-efficient UE connectivity[C]. Proceedings of 2025 IEEE 36th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Istanbul, Turkey, 2025: 1–6. doi: [10.1109/PIMRC62392.2025.11274644](#).
- [28] GHOSH S, MUKHERJEE P, and GHOSH S C. Energy-efficient UAV movement and user-UAV association in multi-UAV networks[C]. Proceedings of 2025 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Nice, France, 2025: 1–6. doi: [10.1109/MeditCom64437.2025.11104463](#).
- [29] LIN Xinyi, LI Peizheng, and ALJAZ A. RL-driven semantic compression model selection and resource allocation in semantic communication systems[C]. Proceedings of 2025 IEEE 36th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Istanbul, Turkey, 2025: 1–6. doi: [10.1109/PIMRC62392.2025.11274758](#).

官 铮: 女, 博士, 教授, 研究方向为多源信息融合、语义通信、
无线网络资源配置。

陈力源: 男, 硕士生, 研究方向为机器学习、智能信息处理等。

刘 洁: 男, 博士生, 研究方向为机器学习、智能信息处理等。

责任编辑: 余 蓉

俞纪宏: 男, 硕士生, 研究方向为通信资源调度、无人机路径规划等。

Resource Allocation for Multi-UAV Relay Networks in 6G Semantic Communication

XIAO Liming GUAN Zheng LIU Jie YU Jihong CHEN Liyuan

(School of Information Science and Engineering, Yunnan University, Kunming 650500, China)

Abstract:

Objective Sixth-generation (6G) mobile networks aim to achieve global seamless coverage through space-air-ground integrated architectures. In this context, Unmanned Aerial Vehicles (UAVs) can serve as mobile aerial relay nodes to support massive ground user access in complex environments. However, traditional data-oriented communication paradigms incur high bandwidth overhead, making them difficult to apply under the limited spectrum resources of UAV networks. In addition, traditional centralized resource allocation methods struggle with real-time implementation due to highly dynamic network topologies and the complex coupling of multi-dimensional resources. To address these challenges, semantic communication has emerged as a new paradigm that extracts the essential meaning of information at the transmitter and reconstructs it at the receiver, thereby reducing redundant data traffic. Although semantic communication provides a feasible solution, existing semantic-driven resource allocation methods primarily focus on static ground networks or single-UAV scenarios, often overlooking the coverage limitations and co-channel interference in multi-UAV collaborative networks. Therefore, an intelligent joint resource optimization model and a distributed resource allocation framework are needed for multi-UAV relay networks. By jointly optimizing multi-dimensional resources, such a framework can improve semantic transmission efficiency and ensure long-term user fairness, thereby supporting intelligent resource scheduling in 6G integrated networks.

Methods This study formulates a joint resource optimization model for multi-UAV relay networks under the semantic communication paradigm (Fig.1), coupling the number of semantic symbols, UAV trajectories, transmission power, and channel allocation. To evaluate semantic transmission quality, the proposed model introduces a Semantic Communication Quality of Service (SC-QoS) metric, which integrates Semantic Quantization Efficiency with normalized transmission delay. The optimization objective is formulated as a Mixed Integer Nonlinear Programming problem that maximizes the weighted sum of system-wide SC-QoS and long-term user fairness evaluated by Jain's fairness index. To solve this problem, a Two-Stage Hybrid Reinforcement Learning (TS-HRL) framework is proposed (Fig.2). The first stage employs a Capacity-Aware K-means (CA-K-means) algorithm for heuristic UAV pre-deployment. By introducing a dynamic distance compensation term based on residual capacity, this algorithm guides edge users toward low-load UAVs, achieving load balancing among UAV clusters while preserving spatial proximity. The second stage models the dynamic scheduling problem as a Decentralized Partially Observable Markov Decision Process and solves it using a Recurrent Independent Proximal Policy Optimization with Parameter Sharing (R-IPPO-PS) algorithm (Algorithm 1). This algorithm leverages Long Short-Term Memory (LSTM) networks to aggregate historical trajectories and observations, thereby capturing hidden environmental states. Furthermore, the multi-agent parameter-sharing mechanism reduces model complexity with respect to the agent scale, while the preference decoupling strategy converts discrete channel allocation decisions into continuous preference variables to facilitate model training.

Results and Discussions The proposed TS-HRL framework is evaluated in a dynamic environment with randomly moving ground users. The convergence analysis shows that the proposed method achieves a higher initial reward and reaches a stable state with fewer iterations than the random deployment, memoryless

allocation, and Enhanced Independent Soft Actor-Critic (EI-SAC) baselines (Fig.3). By leveraging LSTM-based historical information and CA-K-means topology initialization, the proposed method reduces invalid exploration and improves convergence stability. Compared with traditional bit-based communication, the semantic communication framework increases Semantic Spectral Efficiency (S-SE) by 3.6 times, reduces transmission delay by 92.1%, and improves fairness by 14.3% (Fig.4). Within the semantic communication paradigm, compared with the memoryless and heuristic schemes, the proposed method improves S-SE by 92.7% and 51.9%, reduces transmission delay by 13.5% and 57.7%, and improves fairness by 17.3% and 39.7%, respectively. Although the EI-SAC baseline achieves a fairness index of 0.93, its S-SE remains relatively low. In contrast, the proposed TS-HRL framework maintains a high fairness level while achieving an S-SE approximately 4.3 times higher than that of EI-SAC. Compared with random deployment, the proposed method improves S-SE by 11.3% with only a marginal 1.1% decrease in fairness, demonstrating a better trade-off between transmission efficiency and fairness. As the number of users increases from 10 to 40, most baselines show decreases in S-SE and fairness due to intensified co-channel interference and spectrum constraints (Fig.5). In contrast, the adaptive scheduling and global fairness reward mechanisms of the proposed method mitigate performance degradation and maintain the average transmission delay below 0.1 ms. These results indicate that the proposed method improves overall system performance while ensuring long-term user fairness.

Conclusions This paper investigates joint resource allocation and trajectory planning in dynamic 6G multi-UAV relay networks by integrating the semantic communication paradigm. A joint optimization model coupling the number of semantic symbols, UAV trajectory, power control, and channel allocation is formulated to maximize system SC-QoS and long-term user fairness. To address the high-dimensional coupling of the formulated problem, a TS-HRL framework is proposed. This framework combines CA-K-means pre-deployment with the R-IPPO-PS algorithm to handle multi-dimensional resource scheduling under partial observability. Simulation results verify the convergence, stability, and scalability of the proposed method under different user densities. Through distributed cooperation among UAVs, the proposed method enhances semantic transmission efficiency and reduces transmission delay while ensuring long-term service fairness among ground users. These findings provide a viable approach for intelligent resource allocation in UAV-assisted semantic communication for future space-air-ground integrated networks.

Key words: Semantic communication; Multi-UAV network; Resource allocation; Trajectory optimization; Multi-agent reinforcement learning