



XXXX

面向结构化空域的基站协同传输策略研究

熊凯¹, 刘瀚阳, 张瀚文¹, 冷甦鹏

(1. 电子科技大学 信息与通信工程学院, 四川省成都市 邮编 61173)

摘要: 随着低空经济快速发展, 构建面向飞行器协同出行的低空智能网已成为支撑低空活动安全运行的关键所在。然而现有地面基站的部署方案主要针对地面用户, 并未考虑到对结构化空域中的飞行器进行有效信号覆盖。此外, 低空飞行器的移动性远高于地面用户, 对基站的波束赋形以及发射功率调配提出了高实时性要求。本文聚焦于结构化空中航道场景, 提出了一种面向分层航道飞行器接入的地面基站部署与协同功率分配优化策略, 以最大化网络整体空-地通信效率。本文将优化设计为两个阶段: 首先采用启发式搜索算法对基站部署位置进行全局优化, 提升基站对结构化空域中分层定向航道的有效覆盖; 然后构建基于DeepSets架构的深度强化学习框架, 解决高动态场景下的基站功率分配问题, 实现具有预测能力的动态波束追踪。仿真结果表明, 采用该基站部署方案, 信道质量与基准方案相比带来了约9.78%的空地传输速率提升, 所采用的强化框架能够稳定快速完成波束追踪收敛, 为构建高效空-地通信与低空智能网提供了理论支撑。

关键词: 空中航道; 基站部署; 功率分配; 强化学习; 协同传输

中图分类号:

文献标志码:

doi: 10.11959/j.issn.1000-0801.

Cooperative Transmission Strategy of Base Stations for Structured Airspace Corridors

Xiong Kai¹, Liu Hanyang, Zhang Hanwen¹, Leng Supeng

1. School of Information and Communication Engineering, University of Electronic Science and Technology of China

Abstract: With the rapid development of the low-altitude economy, constructing a low-altitude intelligent network for aircraft has become a critical foundation for supporting the safe operation of low-altitude activities. However, existing deployment strategies for ground base stations are primarily designed for ground users, resulting in limited airspace coverage. Furthermore, the mobility of low-altitude aircraft is significantly higher than that of ground users, imposing stringent real-time requirements on beamforming and transmission power allocation of base stations. This paper focuses on structured air route scenarios and proposes a collaborative optimization strategy for base station deployment and power allocation to maximize the overall air-ground transmission efficiency of the network. The optimization is

收稿日期: XXXX-XX-XX; 修回日期: XXXX-XX-XX

基金项目: 国家自然科学基金青年项目 (批准号 62201122), 四川省自然科学基金项目 (批准号 2026NSFSC1431), 国家自然科学基金青年项目 (批准号 62501122), 国家自然科学基金重点项目 (批准号 62541103), 以及航空工业集团飞机智能决策与协同控制机理模型联合技术中心资助。



divided into two stages: first, a heuristic search algorithm is employed to globally optimize base station deployment positions, enhancing the geometric coverage performance of the network. Subsequently, a deep reinforcement learning framework based on the DeepSets architecture is constructed to address the issue of power allocation in highly dynamic scenarios, enabling three-dimensional dynamic tracking. Simulation results demonstrate that the proposed base station deployment scheme improves channel quality, achieving approximately 9.78% higher air-ground transmission rates compared to benchmark schemes. The reinforcement learning framework converges stably and quickly, providing effective three-dimensional tracking of aircraft users. This study offers theoretical support and practical guidance for building an efficient air-ground integrated low-altitude communication network.

Key words: Airspace corridor, base station deployment, power allocation, reinforcement learning, Coordinated Multi-Point.

1 引言

近年来,低空经济快速发展,物流、应急救援、农业等场景下的飞行器应用不断涌现。低空空域内网联飞行器的高密度、高频次运行对通信系统提出了前所未有的挑战。在此背景下,如何优化地面通信基础设施布局并实现动态资源分配,以满足结构化空域中多样化业务需求,成为当前学术界与工业界共同关注的核心问题[1]。

现代低空空域普遍采用分层航道的结构化设计,不同高度层分配不同航向,以确保飞行安全[2]。这种立体化空域结构与地面通信环境存在本质差异。首先空-地视距信道随飞行高度呈现非线性变化,且飞行器在预设航道内的高速移动导致用户空间分布高度动态,传统蜂窝架构所面临的小区边缘覆盖盲区、频繁切换引发的信令风暴问题更加严重。

协同传输体系下,地面基站(Base Station, BS)通过光纤前传链路连接,协同服务边缘处用户,显著提升了小区边界处的信号质量。相较于传统蜂窝系统,协同传输架构具有多方面优势。分布式BS的空间分集可有效对抗空地信道阴影衰落,提升覆盖均匀性,并且中央化资源调度实现全网干扰协调优化,避免了传统小区间干扰协调的高信令开销。文献[3]详细讨论了无蜂窝大规模MIMO系统中支持飞行器通信的性能,并比较

了多种功率分配策略。结果表明,无蜂窝及其低复杂度的以用户为中心架构在支持飞行器通信方面,优于传统的多小区大规模MIMO部署。一些其他学者也针对空地场景的通信进行了研究。文献[4]研究了将飞行器作为空中接入点的无蜂窝大规模MIMO系统,提出一种交替优化方法来优化飞行器的高度与水平位置以最大化总速率。文献[5]提出了一个联合优化接入点集合与飞行器传输功率的方案以最大化飞行器用户和地面用户的加权总速率。文献[6]提出了在智能反射表面辅助下、基于扩展卡尔曼滤波的飞行器毫米波MIMO波束跟踪方案,通过动态调整反射单元参数与线性化非线性系统,显著提升了复杂环境中波束的实时跟踪性能。

尽管协同传输架构在飞行器通信领域展现出巨大潜力,现有研究仍存在明显局限。首先,多数工作聚焦于平面二维场景或自由空域模型,缺乏对三维分层航道结构的建模和针对该结构化空域地面基站部署的研究。其次,现有研究中动态功率分配问题多采用传统凸优化方式求解,并且对于低空用户的动态特性考虑不足。强化学习方法在该领域应用尚处于起步阶段。针对上述研究空白,本文聚焦于分层空中航道这一典型低空场景,采用两阶段优化的形式解决该场景下的通信问题。首先,我们采用启发式搜索算法优化对特定结构化空域下的协同传输基站部署进行优化。

然后运用多种强化学习策略对基站功率分配进行优化并且对比性能。本文的主要创新点如下：

1. 面向结构化空域的协同传输系统建模：本文针对结构化低空空域，构建了概率视距信道，以及波束成形的协同传输系统模型，为低空高密度飞行器通信提供了贴近实际场景的理论分析框架。

2. 启发式搜索与深度强化学习的联合优化框架：提出一种两阶段联合优化策略，首次针对结构化空域中的分层航道，提出启发式算法全局优化基站部署位置，提升接入网对结构化空域的有效覆盖；

3. 针对航道中飞行器高速移动特性，设计基于 DeepSets 架构的深度强化学习模型，实现在受限空域场景下对目标飞行器的预测性波束追踪，提升了对多飞行器同时波束追踪的收敛稳定性。

本文的组织结构如下。第 2 节介绍结构化空域和信道模型；第 3 节阐述优化算法；第 4 节比较不同模型的性能差异；第 5 节对上述研究进行了总结。

2 系统模型

本章构建地对空协同传输架构下系统的数学模型，旨在为后续的性能分析与算法设计奠定理

论基础。首先，本文给出了分层结构化空域的场景示例，并建立了该空域场景模型。在此基础上，本文详细推导下行链路用户可达速率的闭式表达式，并且在此过程中将信道估计对通信性能的影响纳入考虑。最后，以最大化系统下行总传输速率为目标，提出功率分配与位置部署的联合优化模型。

2.1 低空空间结构模型

鉴于飞行器飞行安全的需要，不同方向的飞行器必须安排在不同空间层次中，以确保飞行器不会发生碰撞。如图 1 所示，飞行器只能在航道中运动，在不同层的航道之间存在垂直通道，便于飞行器改变层级往不同方向移动。航道按照实际通行需求被设计为不同方向的圆柱体管道，并且在规划完成后具有较长周期的稳定性。此外，部署在地面的多个基站通过光纤连接至中央服务器。在该服务器的调度下，多个基站可以协同为同一个目标提供接入服务。

在实际航道中，出于安全性以及飞行效率的考虑，飞行器常常以编队的形式飞行。多个飞行器被划分为一个编队，并且编队中存在一个簇首 (Cluster Head, CH)，负责队内的飞行器管理避障等以及与地面基站通信。依据实际业务，基站需要对 CH 传递路径规划，飞行动力学控制等关

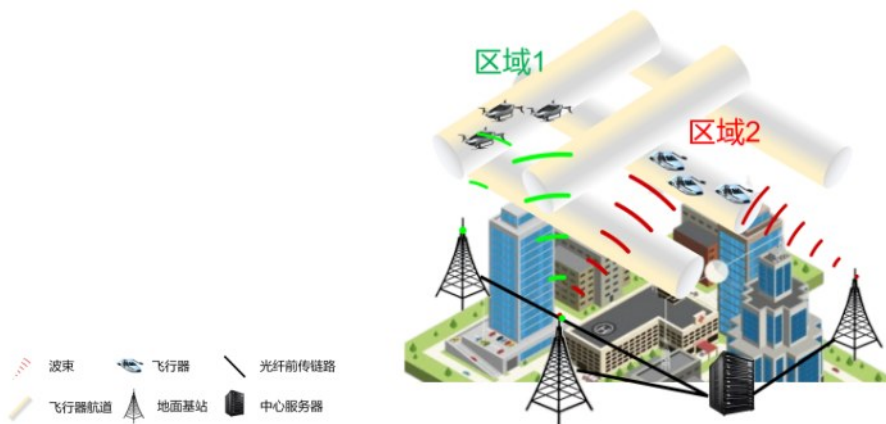


图 1 系统接入示意图



键信息。因此，CH需要于地面基站保持实时高质量通信。

2.2 信道模型

在本小节中我们阐述目标飞行器与基站通信系统的数学模型。我们将空地信道衰落、信道估计误差纳入考虑，给出在非完美信道状态信息下的下行链路数据传输闭式可达速率表达式，为后续优化问题构建提供理论基础。

2.2.1 信道传播模型

我们考虑CH与基站之间的空地通信链路。第 m 个基站与第 k 个CH之间的信道增益系数 g_{mk} 可以建模为大尺度衰落与小尺度衰落的乘积，即：

$$g_{mk} = \sqrt{\beta_{mk}} h_{mk} \quad \#(1)$$

其中， h_{mk} 表示第 m 个基站与第 k 个CH之间的小尺度衰落系数，假设其服从独立同分布的复高斯分布 $\mathcal{C}\mathcal{N}(0, 1)$ 。 β_{mk} 第 m 个基站与第 k 个CH之间的大尺度衰落系数。考虑到本场景的空地信道特性，大尺度衰落系数 β_{mk} 采用概率性视距 (Line of Sight, LoS) 与非视距 (Non Line of Sight, NLoS) 传播模型，该模型通过仰角相关的LoS概率统计描述城市环境中建筑遮挡对空地信道的宏观影响。 β_{mk} 的计算公式如下：

$$\beta_{mk} = P_{mk}^L D_{mk}^{-\alpha_1} + P_{mk}^N D_{mk}^{-\alpha_2} \quad \#(2)$$

其中， D_{mk} 表示第 m 个基站与第 k 个CH之间的距离， α_1 和 α_2 分别代表LoS和NLoS链路下的路径损耗指数。3GPP研究中对城市环境下该参数的取值给出了参考意见[14]。 P_{mk}^L 和 $P_{mk}^N = 1 - P_{mk}^L$ 分别表示第 m 个基站与第 k 个CH之间的出现LoS和NLoS链路的概率。假设基站位置用

$$\hat{y}_{mk} = \sqrt{\tau_p p_u} g_{mk} + \sqrt{\tau_p p_u} \sum_{k'=1, k' \neq k}^K g_{mk'} \phi_{k'}^H \phi_k + w_m \phi_k \quad \#(8)$$

基于最小均方误差 (MMSE) 准则，第 m 个基站与第 k 个CH之间的信道 g_{mk} 的估计值 \hat{g}_{mk} 为：

$\mathbf{q}_m = (X_m, Y_m)$ 表示，CH位置用 $u_k = (x_k, y_k, h_k)$ 表示，则第 m 个基站与第 k 个CH之间的三维距离 d_{mk} 由水平距离 r_{mk} 和CH所在高度 h_k 决定，可以表示为：

$$d_{mk} = \sqrt{r_{mk}^2 + h_k^2} \quad \#(3)$$

$$r_{mk} = \sqrt{(X_m - x_k)^2 + (Y_m - y_k)^2} \quad \#(4)$$

LoS概率 P_{mk}^L 主要取决于目标点相对于基站的仰角 θ_{mk} 以及所处环境的特性，在文献[7]中，作者给出了该概率的具体表达式：

$$P_{mk}^L = \frac{1}{1 + \zeta_1 e^{-\zeta_2 (\theta_{mk} - \zeta_1)}} \quad \#(5)$$

其中， ζ_1 和 ζ_2 是与环境相关的常数参数，该文献中，作者在不同环境下对此参数进行测量给出了参考取值。仰角 θ_{mk} 计算方法如下：

$$\theta_{mk} = \frac{180^\circ}{\pi} \arctan \left(\frac{h_k}{r_{mk}} \right) \quad \#(6)$$

2.2.2 上行链路导频训练与信道估计

在协同传输架构中，为了实现多个分布式基站对同一飞行器的相干联合传输，基站侧必须掌握准确的信道状态信息 (Channel State Information, CSI)。此阶段的信道估计效果将会影响实际的发射功率和干扰抑制效果，最终影响目标簇首的信号质量。假设所有CH同时发送长度为 τ_p 的导频序列 $\phi_k \in \mathcal{C}^{\tau_p \times 1}$ ，且满足 $\|\phi_k\|^2 = 1$ 。第 m 个基站接收到的导频信号 y_m 可以表示为：

$$y_m = \sqrt{\tau_p p_u} \sum_{k=1}^K g_{mk} \phi_k^H + w_m \quad \#(7)$$

其中， p_u 为CH上行发射功率， w_m 是服从 $\mathcal{CN}(0, 1)$ 的加性高斯白噪声。基站通过将接收信号投影到第 k 个CH的导频序列上来获取观测值：

$$\hat{g}_{mk} = \frac{\sqrt{\tau_p p_u} \beta_{mk}}{\tau_p p_u \beta_{mk} + 1} \hat{y}_{mk} \quad \#(9)$$

根据MMSE估计的性质，估计信道 \hat{g}_{mk} 与估

计误差 \tilde{g}_{mk} 是正交的。信道估计值的方差 $\gamma_{mk} \triangleq \mathbb{E}\{\hat{g}_{mk}^2\}$ 可表示为:

$$\gamma_{mk} = \frac{\tau_p p_u \beta_{mk}^2}{\tau_p p_u \beta_{mk} + 1} \quad \#(10)$$

2.2.3 下行链路数据传输

在下行数据传输阶段, 基站利用波束成形技术向簇首发送数据。用 K_m 表示基站 m 服务的区域集合, A_k 表示簇首 k 的服务基站集。 η_{mk} 是功率控制系数, 表示第 m 个基站给簇首 k 发送功率与其最大功率的比值。第 m 个基站发送的信号 x_m 为:

$$x_m = \sqrt{p_d} \sum_{k \in K_m} \eta_{mk}^{1/2} \hat{g}_{mk}^* q_k \quad \#(11)$$

其中, p 是归一化的最大发射功率, q_k 是发送给第 k 个 CH 的数据符号 (满足 $\mathbb{E}\{|q_k|^2\} = 1$), 为了满足基站的平均功率约束, 功率系数需满足:

$$\mathbb{E}\{\|x_m\|^2\} = p \sum_{k \in K_m} \eta_{mk} \gamma_{mk} \leq p \quad \#(12)$$

这等价于约束条件 $\sum_{k \in K_m} \eta_{mk} \gamma_{mk} \leq 1$ 。第 k

个 CH 接收到的信号 r_k 是来自其服务集 A_k 内所有基站信号的叠加:

$$r_k = \sqrt{p} \sum_{m \in A_k} \sum_{k' \in K_m} \eta_{mk'}^{1/2} g_{mk'} \hat{g}_{mk'}^* q_{k'} + z_k \quad \#(13)$$

其中, $z_k \sim \mathcal{CN}(0, 1)$ 为接收端的加性高斯白噪声。

2.2.4 下行链路可达速率分析

我们将接收信号分解为期望信号 (DS)、波束成形增益不确定性 (BU) 以及簇首间干扰 (UI)。第 k 个 CH 的下行链路可达速率 C_k 表达式为:

$$C_k = \log_2 \left(1 + \frac{|DS_k|^2}{\mathbb{E}\{BU_k^2\} + \sum_{k' \neq k} \mathbb{E}\{UI_{kk'}^2\} + \sigma^2} \right) \quad \#(14)$$

期望信号功率部分:

$$DS_k = \sqrt{p} \sum_{m \in A_k} \eta_{mk}^{1/2} \gamma_{mk} \quad \#(15)$$

波束成形不确定性引起的干扰功率:

$$BU_k \triangleq \sqrt{p} \left(\sum_{m \in A_k} \eta_{mk}^{1/2} g_{mk} \hat{g}_{mk}^* - \mathbb{E} \left\{ \sum_{m \in A_k} \eta_{mk}^{1/2} g_{mk} \hat{g}_{mk}^* \right\} \right) \quad (16)$$

簇首间干扰功率:

$$UI_{kk'} \triangleq \sqrt{p} \sum_{m \in A_k} \eta_{mk'}^{1/2} g_{mk'} \hat{g}_{mk'}^* \quad \#(17)$$

将上述各项代入并化简, 最终得到第 k 个 CH 的下行链路闭式容量表达式为[8]:

$$C_k = \log_2 \left(1 + \frac{p \left(\sum_{m \in A_k} \eta_{mk}^{1/2} \gamma_{mk} \right)^2}{p \sum_{k' \neq k} \sum_{m \in A_k} \eta_{mk'} \cdot \gamma_{mk'} \beta_{mk'} + \sigma^2} \right) \quad \#(18)$$

此闭式表达式揭示了系统性能与大尺度衰落系数 β_{mk} 、信道估计质量 γ_{mk} 以及功率分配策略 η_{mk} 之间的关系。

2.3 优化模型

本文旨在合理分布基站位置, 并且在此基础上对基站合理分配发射功率, 使得在较短时间间隔内, 信号能够实际覆盖 CH 的当前位置以及

预期前进方向上的一段距离, 实现追踪对准的效果。

假设在某一较短的时间间隔内, 功率分配系数集合 $\eta = \{\eta_{mk}\}$ 和以及基站位置的坐标集合 $Q = \{q_m\}$ 均为固定值。在该时间间隔开始时第 k 个簇首所在位置为 $u_k = (x_k, y_k, h_k)$, 在该时间间隔内的预期速度为 $v_k = (v_{xk}, v_{yk}, v_{hk})$, 该预期速度在时间



间隔开始时由飞行器预测给出。若从该时间间隔开始到当前经过的时间为 t ，则飞行器实际位置为 $u_k' = u_k + v_k \cdot t$ 。公式18中给出的 C_k 是飞行器实际位置的函数，则 C_k 可以表示为 $C_k(t; \eta, Q)$ ，目标函数可以表示为：

$$\max_{\eta, Q} \sum_{k=1}^K \int_0^{t_m} C_k(t; \eta, Q) \cdot \varphi(t) \cdot dt \quad (19)$$

其中 t_m 表示时间间隔设定的长度， $\varphi(t)$ 是置信度函数，表示从0到 t_m 时预期速度 v_k 的置信度。求 C_k 与 $\varphi(t)$ 的积分将时间间隔内的飞行器预期移动方向上的信号覆盖纳入考虑。

该优化问题需满足四个约束条件：

$$C1: \sum_{k \in K_m} \eta_{mk} \gamma_{mk} \leq 1, \quad \forall m = 1, \dots, M \quad (20)$$

$$C2: \eta_{mk} \geq 0, \quad \forall m, k \quad (21)$$

$$C3: X_{\min} \leq X_m \leq X_{\max}, \quad \forall m = 1, \dots, M \quad (22)$$

$$C4: Y_{\min} \leq Y_m \leq Y_{\max}, \quad \forall m = 1, \dots, M \quad (23)$$

C1表示基站的功率分配约束。由于每个基站的归一化最大发射功率为 p_d ，分配给所有服务簇首的功率系数加权和不能超过1。C2表示功率分配系数必须是非负。C3、C4表示有效区域约束，基站部署簇首在 x 、 y 轴上的边界分别为 $[X_{\min}, X_{\max}]$ 、 $[Y_{\min}, Y_{\max}]$ 。

综上所述，原优化问题是一个关于功率控制与位置部署的联合优化问题。由于目标函数 C_k 关于变量 η 和 Q 是非凸的，且变量之间存在复杂的耦合关系，直接求解该问题具有较高的计算复杂度。

3 模型求解

该优化问题是一个非凸优化问题，涉及基站部署位置与功率分配系数的联合优化，并且飞行器网络的时变特性以及业务的突发不确定性传统的优化方法难以在满足实时性的情况下给出良好的解。

基站位置属于准静态部署变量，在网络规划

阶段确定后长期不变；而功率分配需根据目标簇首的动态变化进行实时调整，考虑到两个优化变量的不同时间尺度特性，如图所示，本文设计了一种两阶段分离优化框架：

阶段一：针对基站位置部署，在使用标准的Top-p功率分配策略时，采用差分进化（Differential Evolution, DE）算法进行全局启发式搜索，以获得最优的基站空间拓扑结构；

阶段二：在固定基站拓扑的基础上，将动态功率分配问题建模为马尔可夫决策过程，采用深度强化学习算法进行实时功率分配决策。

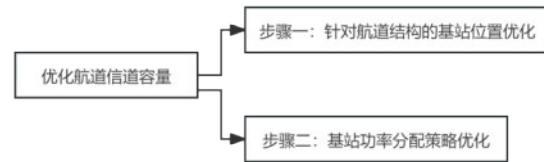


图2 两阶段优化

3.1小节中首先介绍步骤1的3.1基站位置部署优化；随后3.2小节中阐述阶段二的MDP建模，3.3小节中最后阐述基于DeepSets架构的深度强化学习框架。

3.1 基站位置部署优化

基站的拓扑布局对系统覆盖性能和平均速率具有决定性影响。依据实际需求，空中航道的实际结构各不相同，良好的基站分布是确保功率分配优化能够有效进行的前提。该优化问题可以描述为：当固定数量的CH在航道内随机均匀分布并采用固定功率分配策略时，优化地面基站的位置，使得飞行器簇首与地面基站之间的信道容量期望值最大。

鉴于基站位置优化问题具有高度非线性、非凸性的特点，传统的梯度下降法容易陷入局部最优。因此，本文采用一种启发式全局优化算法——差分进化算法来求解最优基站部署方案[13]。

差分进化算法是一种基于种群的随机搜索技

术, 通过变异、交叉和选择三个核心算子在解空间内进行并行搜索。在本文中, 优化变量为 M 个基站的二维平面坐标向量 $\mathbf{x}=[x_1, y_1, \dots, x_M, y_M]$ 。算法的流程包含四个步骤: 初始化、变异、交叉、选择。其中, 在变异步骤中, 为平衡全局探索与局部开发能力, 本文采用 best/1 策略生成变异向量 \mathbf{v}_i :

$$\mathbf{v}_i = \mathbf{x}_{best} + F \cdot (\mathbf{x}_{r1} - \mathbf{x}_{r2}) \quad (24)$$

\mathbf{x}_{best} 为当前种群中适应度最高的个体, \mathbf{x}_{r1} 和 \mathbf{x}_{r2} 为随机选取的不同个体。在交叉步骤中, 算法通过二项式交叉操作将目标向量 \mathbf{x}_i 与变异向量 \mathbf{v}_i 混合生成试验向量 \mathbf{u}_i 。在选择步骤中, 采用贪婪选择策略, 若试验向量 \mathbf{u}_i 对应的系统平均总速率优于原个体 \mathbf{x}_i , 则在下一代中保留 \mathbf{u}_i , 否则保留 \mathbf{x}_i 。为了表示在均匀分布时簇首的信道容量期望, 我们采取蒙特卡洛采样的方法进行多次采样, 适应度函数可以表示如下:

$$\max_{\mathbf{Q}} \frac{1}{N_s} \sum_{n=1}^{N_s} \sum_{k=1}^K C_k(\mathbf{Q}, \mathbf{u}_k^{(n)})$$

其中 N_s 为蒙特卡洛采样场景数, $\mathbf{u}_k^{(n)} \sim \mathcal{U}(\mathcal{C})$ 表示第 n 个采样场景中第 k 个簇首在航道区域 \mathcal{C} 内的均匀采样位置。基站功率分配方案采用稳定有效的 Top-p 方案, 即某飞行器按照信号强度从备选基站中选择, 使得选出的基站信号功率高于所有选择总和的 $p\%$ 。

3.2 功率分配MDP建模

针对功率分配部分的优化, 我们将其建模为一个马尔可夫决策过程 (Markov Decision Process, MDP)。在MDP模型中, 中央服务器通过与飞行器通信环境的交互来学习最优的功率分配策略, 由元组 $\langle \mathcal{S}, \mathcal{A}, \mathcal{R} \rangle$ 定义。其中, 状态空间 \mathcal{S} 定义为:

$$s = \{U, V\} \quad (25)$$

其中 $U = \{\mathbf{u}_k = (x_k, y_k, h_k)\}$ 表示当前时间间隔开始时所有 CH 的初始三维位置坐标。 $V = \{\mathbf{v}_k =$

$(v_{xk}, v_{yk}, v_{hk})\}$ 表示当前时间间隔开始时所有 CH 在该时间间隔内的预计三维位置速度信息。功率切换的实际间隔大概在 0.1s 左右, 该速度向量只需要预测段时间的移动方向。

动作空间 \mathcal{A} 定义为:

$$a = \{\eta_{mk} \mid m \in \mathcal{M}, k \in \mathcal{K}\} \quad (26)$$

其中 \mathcal{M} 表示地面基站集合, \mathcal{K} 表示全部 CH 集合。

奖励函数 \mathcal{R} 方面, 公式 19 中给出了优化的总目标。由于积分运算在训练过程中计算复杂度高, 我们将奖励函数其近似替换为:

$$r = (1 - \lambda) \sum_{k=1}^K C_k(t=0) + \lambda \sum_{k=1}^K C_k(t=t_m)$$

即只取首位两点的信道容量纳入考虑, 其中 $\lambda=0.5$ 。

3.3 功率分配强化学习设计

针对此高维非凸场景, 我们从强化学习框架、骨干神经网络设计、输出处理三个方面进行强化学习设计。

强化学习框架方面, 由于功率分配系数为连续变量, 我们分别实现了 DDPG[10]、TD3[11]和 SAC[12]三种深度强化学习框架进行对比。其中, DDPG 是基础经典的 Actor-Critic 算法; TD3 在此基础上添加了双 Critic 网络与延迟策略更新机制缓解 Q 值过估计问题; SAC 则引入最大熵机制, 其随机性策略在此问题的高维非凸场景下有更好的搜索能力。三种强化学习框架伪代码如下:

骨干神经网络设计方面, 传统的全连接神经网络 (Multilayer Perceptron, MLP) 要求输入维度固定且对输入顺序敏感, 这不适用于簇首位置输入无序的飞行器通信网络。为了提取系统中所有目标点与基站之间的复杂拓扑特征, 本文采用基于 DeepSets 的 Encoder-Decoder 架构作为 Actor 与 Critic 网络的主干。该架构通过将每个 CH 的特征独立编码后求和, 无论 CH 以何种顺序输入网络, 输出的功率分配策略均保持一致。这一特性



 强化学习训练框架

```

Initialize Actor(s), Critic(s), Replay_Buffer D
Loop (Training_Epochs):
  Loop (Collection_Steps):
    s = Env.observe()
    [DDPG, TD3]a = Clip(Actor(s) + exploration_noise)
    [SAC]a = Actor_Distribution(s).sample()
    r = Env.execute(a)
    D.store(s, a, r)
  Loop (Gradient_Steps):
    b_s, b_a, b_r = D.sample_batch()
    [DDPG] Loss_C = MSE(Critic1(b_s, b_a), b_r)
    [TD3, SAC]Loss_C = MSE(Critic1(b_s, b_a), b_r) + MSE(Critic2(b_s, b_a), b_r)
    Update_Critics(Minimize(Loss_C))
    [DDPG]run_actor_update = True
    [TD3, SAC]run_actor_update = (step % delay == 0)
    If run_actor_update:
      [DDPG, TD3]Loss_A = -Mean(Critic1(b_s, Actor(b_s)))
      [SAC]Loss_A = Mean(alpha * log_prob - Min_Critic(b_s, new_sampled_a))
      Update_Actor(Minimize(Loss_A))
      Soft_Update(Target_Networks)
  
```

与协同传输场景高度契合——基站对飞行器的服务优先级仅取决于其实际位置与信道质量，而与飞行器ID无关。此外，DeepSets架构能够灵活处理变数量的CH输入，为未来扩展至更高密度飞行器场景提供了良好的可扩展性。

在输入特征设计上，本文选用 $\log_{10} \beta_{mk}$ 压缩大尺度衰落的动态范围、 $\sin \theta_{mk}$ 消除仰角周期性突变)、以及CH的三维坐标 (x_k, y_k, h_k) 。这些特征直接反映了结构化空域中分层航道带来的信道差异。不同高度层的视距概率与路径损耗呈非线性变化，而 $\sin \theta_{mk}$ 与 h_k 的组合能够有效表征这种高度依赖性。输出端采用Sigmoid与逐基站Softmax归一化，确保生成的功率系数满足非负性与基站总功率约束，使动作始终落在物理可行域内。

4 仿真结果与分析

在本节中，我们通过数值仿真验证了所提出

的联合优化框架的有效性。首先我们评估基站位置优化对系统容量的提升效果；其次，在优化后的拓扑结构下，对比SAC、TD3和DDPG三种深度强化学习算法在功率分配任务上的性能；最后，分析目标点所在高度对通信质量的影响。

4.1 1 仿真参数

本小节构建了一个覆盖范围为 $10 \times 10 \text{ km}^2$ 的场景。网络包含10个基站和5个簇首。基站的功率发射功率设定为50 W。信道建模采用了概率性的LoS与NLoS传播模型，其中LoS和NLoS路径损耗指数分别设置为 $\alpha_1=2$ 和 $\alpha_2=4$ 。视距概率参数设置为 $\xi_1=1.2$ ， $\xi_2=2.1$ 。

基站部署方面，我们采用一种常见的三层空中航道结构，一层的飞行器航向相同或相反，在上一层基础上顺时针偏移 60° ，三层飞行器航道高度分别为100m、200m、300m。航道区域长宽为10km，每个层内航道之间的间距为1500米。本文采用的算法对于任意航道形态都有效，此处

考虑本场景仅便于效果验证。算法缩放因子F设置在 [0.5,1.0] 区间内随机抖动，以增加搜索的多样性。交叉概率设定为 CR=0.7，这意味着新个体有大概率继承变异向量的特征。

在强化学习部分实验中，编码器与解码器的隐藏层节点数均设为 256。编码器采用 2 层全连接结构，层间使用层归一化与 ReLU 激活函数；解码器为 2 层 MLP，第一层同样配置 256 节点、层归一化与 ReLU，第二层输出层对于确定性策略 (DDPG/TD3) 采用 Tanh 激活将动作限制于 [-1,1]，对于随机性策略 (SAC) 则并列输出均值与标准差对数。

4.1 2 算法复杂度分析

为进一步分析不同方法的计算复杂度，本文在相同实验平台下对各方法的单次决策时间进行了对比。具体而言，测试时固定使用 30 个预先生成的状态样本，并仅统计算法在给定状态下输出动作所需的时间，而不计入环境重置、状态转移等额外开销，从而保证比较的公平性。实验结果表明，简单 Top-K 规则的平均单次决策时间仅为 0.0113 ms，具有最高的实时性；三种强化学习方法的推理开销处于同一量级，其中 DDPG、TD3 和 SAC 的平均单次决策时间分别为 1.5880 ms、1.6278 ms 和 2.0482 ms；相比之下，遗传算法由于需要进行迭代搜索，其平均单次决策时间达到 59.6833 ms，显著高于学习型方法。上述结果说明，所采用的强化学习模型在计算开销上虽然略高于极简规则法，但相较于基于搜索的遗传算法具有明显的效率优势，能够在保证较强决策能力的同时满足更高的实时推理需求。

实验平台方面，上述复杂度测试均在 CPU 条件下完成，硬件环境为 Intel Core Ultra 9 275HX 处理器，24 核，系统内存为 32 GB，操作系统为 Windows 11，软件环境为 Python 3.9.25 与 PyTorch 2.8.0。

4.1 3 鲁棒性

为评估各方法在不确定观测条件下的鲁棒性，本文在测试阶段对输入观测引入逐级增强的扰动，并采用性能保持率 (rate retention) 作为核心指标进行比较。结果如图 3，随着观测噪声提升，三者总信道容量的保持都均明显下降，当达到 0.2 时，Top-K、GA、DDPG、TD3 和 SAC 的保持率分别为 0.9948、0.9337、0.9615、0.9675 和 0.9822。可以看出，各方法在扰动环境下均表现出一定稳定性，但保持率存在明显差异：Top-K 与 SAC 的保持率相对更高，说明其对观测扰动的敏感性较低；TD3 与 DDPG 处于中间水平；GA 的保持率最低，表明其在当前扰动设置下性能退化更为明显。总体而言，学习型方法在鲁棒性方面具有可接受的稳定表现，其中 SAC 在保持率指标上优于另外两种强化学习基线。上述鲁棒性差异源于方法的内在特性：SAC 的随机策略通过动作采样对观测噪声形成隐式平均，DDPG/TD3 的确定性策略则直接放大扰动；Top-K 规则依赖排序而非精确数值，对噪声天然不敏感但性能上限低；GA 缺乏抗噪机制，适应度评估偏差导致搜索失效。这验证了 SAC 在高动态、非理想观测场景下的工程适用性。

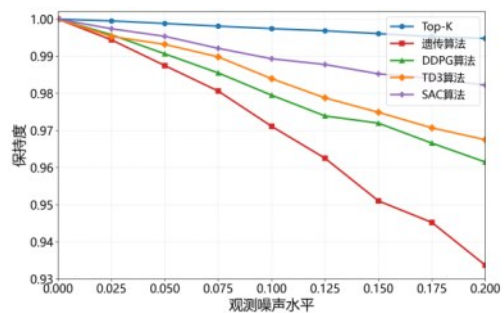


图 3 观测噪声-保持率分析

4.1 4 基站位置优化效果分析

优化前基站位置分布如图 4 所示，其中可以观察到分布散乱。为了量化基站部署策略的性能增益，我们采用差分进化算法对基站坐标进行了

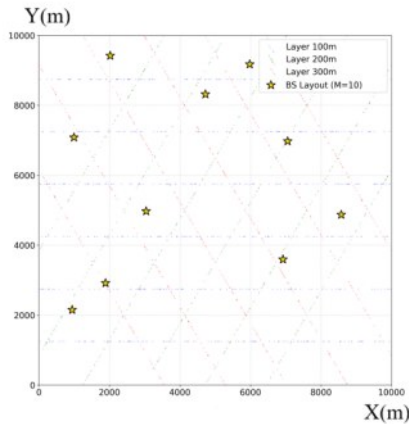


图4 优化前基站位置

全局搜索。

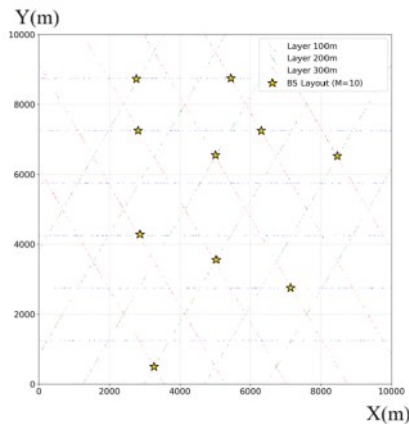


图5 优化后基站位置

优化后的基站位置如图5，在图5中，基站呈现出在三层航道水平投影的交叉点处集中的特征。在这些位置处部署基站能够更好的同时覆盖三个层级的飞行器，图6中体现了优化后的基站布局将网络的平均总速率从 2.1594 bps/Hz 提升至 2.3706 bps/Hz，实现了约 9.78% 的性能增益。这证明了通过启发式搜索改善空间拓扑结构是提升协同传输系统性能的有效预处理手段。优化后的基站更加趋于集中分布，因而可能导致优化后的方差略高于优化前的方差。

4.1 5 不同强化学习算法性能对比

基于优化后的基站位置，我们进一步引入深

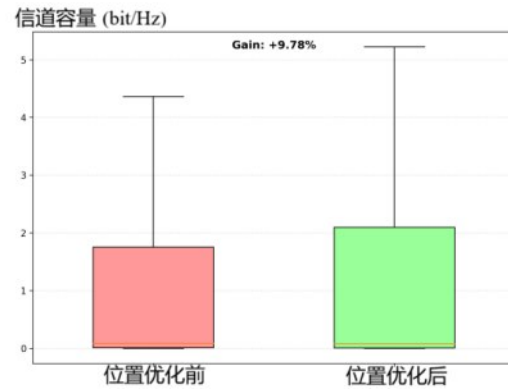


图6 优化前后基站部署方案对比

度强化学习算法进行动态功率分配。对比了 DDPG、TD3 和 SAC 三种主流框架。如图7所示，在训练初期，所有模型的 Loss 均快速下降，SAC 模型的下降速度快于其他两者。在训练 4000 轮次以后以后模型 Loss 趋于平稳，达到收敛状态。DDPG 和 TD3 的表现相近，而 SAC 模型的最终收敛奖励值效果优于前两者。从大致 0.2 下降到小于 0.02，而其他两者大致稳定在 0.05 左右。SAC 采用最大熵框架，其策略输出为动作分布而非确定性值，在训练初期通过熵正则化鼓励主动探索，从而更快遍历高维非凸的功率分配动作空间，加速损失函数下降；同时，随机策略对估计误差与干扰具有天然的平滑作用，使最终收敛奖励值更低。相比之下，DDPG 与 TD3 采用确定性策略，依赖外部噪声进行探索，效率较低且易陷入局部最优，因此收敛速度与最终性能均弱于 SAC。TD3 相比 DDPG 的改进主要体现在缓解 Q 值过估计与提升训练稳定性方面。然而，由于功率分配问题受基站功率约束限制，且信道增益随距离平滑变化，Q 值过估计程度本身不严重；同时两者均采用确定性策略加外部噪声的探索方式，在有限训练轮次内探索效率相近。因此，TD3 的额外机制未能带来显著的性能增益，导致其与 DDPG 表现相近。

4.1 6 目标点高度对通信性能的影响

为了探究飞行器三维位置对信道质量的影

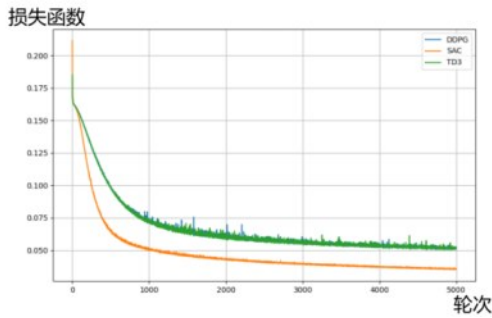


图7 训练 Loss 下降比较

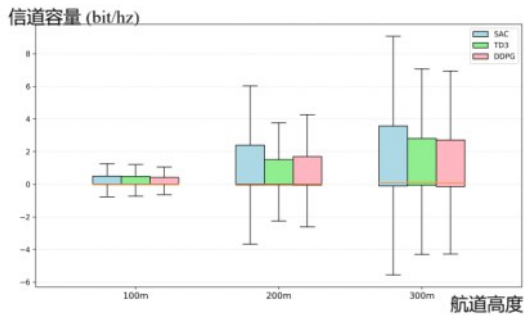


图8 高度以及强化学习算法对比

响，我们将测试集划分为三个高度层级：100m、200m 和 300m，如图8，从数据可以观察到高度-性能正相关趋势。这一现象与我们在信道建模中的预期一致。根据 LoS 概率模型公式，随着飞行高度 H 的增加，飞行器与基站之间的仰角 θ_{mk} 增大，从而显著提高了视距传播的概率 P_{mk}^L 。高概率的 LoS 链路大幅降低了路径损耗和阴影衰落，从而提升了接收端的信噪比和系统总容量。当簇首位置处于较高层级时，性能方差明显增大。这说明在较高层级中，由于视距概率提高，可达的最大速率提高，并且覆盖范围也更加广泛，而航道仍然存在部分区域无法覆盖，因而存在更高的方差。

在同一高度下横向对比三种算法在不同的平均奖励，SAC 在所有测试高度下均取得了最高的平均总速率，TD3 和 DDPG 呈现出相似的性能水平。由于 SAC 的动作生成的随机采样机制，从标准差来看，SAC 的方差略高于 TD3 和 DDPG。

5 结束语

本文针对低空智能空中结构化航道场景下的通信覆盖与资源分配问题，提出了一种两阶段协同优化策略，融合了基站部署的全局优化与功率分配的动态优化管理。首先，通过差分进化算法对基站位置进行启发式搜索，显著提升了系统对三维结构化空域的覆盖能力；然后基于 DeepSets 架构构建了深度强化学习模型，实现了对高动态飞行器的实时功率分配与预测性波束追踪。仿真结果表明，所提部署策略在空-地通信平均总速率上有明显提升，而基于 SAC 的强化学习框架在收敛速度与系统性能上均优于 DDPG 与 TD3，展现了良好的鲁棒性与适应性。此外，本文揭示了飞行高度对通信质量的正向影响，为未来分层空域通信系统设计提供了理论依据。在未来工作中，我们考虑通过多飞行器协同航迹优化进一步降低空-地通信干扰。



熊凯 (1991-)，男，电子科技大学副教授，主要研究方向为先进空中交通、无人集群组网与控制。

参考文献 (Reference):

- [1] 殷哲. 低空通信网络优化研究[J/OL]. 电信科学, 2025, 41(12): 1-15. (2025-12-09) [2026-01-20]. <https://link.cnki.net/urlid/11.2103.TN.20251209.1252.002>.
- [2] 王尔申, 张圣隆, 郭婧, 等. 面向低空经济应用场景的飞行器路径规划技术研究综述[J]. 电信科学, 2025, 41(12): 27-43.
- [3] MEI W, ZHANG R. Cooperative downlink interference transmission and cancellation for cellular-connected UAV: a divide-and-conquer approach[J]. IEEE Transactions on Communica-



- tions, 2020, 68(2): 1297-1311. DOI: 10.1109/TCOMM.2019.2955953.
- [4] D'ANDREA C, GARCIA-RODRIGUEZ A, GERACI G, et al. Analysis of UAV communications in cell-free massive MIMO systems[J]. IEEE Open Journal of the Communications Society, 2020, 1: 133-147. DOI:10.1109/OJCOMS.2020.2964983.
- [5] MEI W, ZHANG R. Aerial-ground interference mitigation for cellular-connected UAV[J]. IEEE Wireless Communications, 2021, 28(1): 167-173. DOI:10.1109/MWC.001.2000173.
- [6] 吴万民. 飞行器通信MIMO信道估计及波束跟踪研究[D]. 南京: 南京邮电大学, 2024. DOI: 10.27251/d.cnki.gn-jdc.2024.000113.
- [7] AL-HOURANI A, KANDEEPAN S, LARDNER S. Optimal LBS altitude for maximum coverage[J]. IEEE Wireless Communications Letters, 2014, 3(6): 569-572. DOI: 10.1109/LWC.2014.2342736.
- [8] WANG L, ZHANG Q. Cell-free massive MIMO with UAV access points: UAV location optimization[C]// *2022 IEEE/CIC International Conference on Communications in China (ICCC) *. Sanshui, Foshan, China: IEEE, 2022: 262-267. DOI:10.1109/ICCC55456.2022.9880734.
- [9] ZAHEER M, KOTTUR S, RAVANBAKSH S, et al. Deep sets [EB/OL]. (2018-04-14) [2024-12-01]. <https://arxiv.org/abs/1703.06114>. DOI:10.48550/arXiv.1703.06114.
- [10] LILLICRBS T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning[EB/OL]. (2019-07-05) [2024-12-01]. <https://arxiv.org/abs/1509.02971>. DOI: 10.48550/arXiv.1509.02971.
- [11] FUJIMOTO S, VAN HOOF H, MEGER D. Addressing function approximation error in actor-critic methods[EB/OL]. (2018-10-22) [2024-12-01]. <https://arxiv.org/abs/1802.09477>. DOI:10.48550/arXiv.1802.09477.
- [12] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[EB/OL]. (2018-08-08) [2024-12-01]. <https://arxiv.org/abs/1801.01290>. DOI:10.48550/arXiv.1801.01290.
- [13] BILAL, PANT M, ZAHEER H, et al. Differential evolution: a review of more than two decades of research[J]. Engineering Applications of Artificial Intelligence, 2020, 90: 103479. DOI: 10.1016/j.engappai.2020.103479.
- [14] 3 GPP. Study on channel model for frequencies from 0.5 to 100 GHz: 3GPP TR 38.901 V16.1.0[R/OL]. Valbonne: 3GPP, 2019[2023-10-27]. https://www.3gpp.org/ftp/Specs/archive/38_series/38.901/.

[作者简介]



刘瀚阳 (2003-), 男, 电子科技大学硕士研究生在读。主要研究方向为无人编队多模态感知识别及组网协议设计。

张瀚文 (1995-), 女, 电子科技大学博士后, 主要研究方向低空无线通信, 太赫兹通信与组网,



语义通信。

冷甦鹏 (1973-), 男, 电子科技大学教授、博士生导师, 主要研究方向为先进空中交通、低空智联网、物联网、车联网、智能交通信息系统的资源管理等。

