

面向 eVTOL 垂直起降场低空起降调度的拓扑博弈自博弈框架

Lanhaijian 大模型辅助

摘要

基于电动垂直起降飞行器 (eVTOL) 的城市空中交通 (UAM) 在垂直起降场面临核心调度瓶颈：传统集中式混合整数线性规划 (MILP) 求解器在高交通密度下存在计算量指数爆炸问题；固定先到先服务 (FCFS) 调度规则会引发严重航班延误、起降坪利用率偏低；传统单智能体强化学习缺少多步时序前瞻能力，决策逻辑可解释性差。针对上述缺陷，本文借鉴 AlphaZero 自博弈范式，提出一套拓扑博弈自博弈调度框架。

本文将标准化低空空域抽象为模块化空域拓扑基元，包含进港通道、离港通道、环形等待区与起降坪终端单元，可程序化生成垂直起降场博弈环境；将多 eVTOL 起降调度问题重构为定义在结构化 2.5 维空域拓扑上的时序时隙占用组合博弈。框架摒弃暴力轨迹拟合与人工枚举调度规则，采用融合拓扑感知混合专家模块 (TopoMoE) 的图神经网络 (GNN) 提取多尺度空域嵌入特征，并设计带时序折扣 UCT 的蒙特卡洛树搜索 (MCTS) 完成长时域时序序列推理；构建分层多目标奖励函数，兼顾飞行安全、起降坪吞吐量、航班延误与能耗约束。

通过仿真自博弈自主生成多智能体协同轨迹，并采用师生策略蒸馏得到轻量化机载调度决策模块。与现有文献中分设进港、离港调度模型不同，本文框架在同一图拓扑与 MCTS 搜索流水线下实现起降任务统一博弈建模。理论分析与分阶段仿真验证方案表明，所提拓扑博弈范式降低了对人工设计调度规则与全局集中求解器的依赖，具备五大突出优势：可解释多步时序调度、模块化可扩展空域拓扑生成、无中心控制分布式多智能体协同、自博弈涌现时隙礼让协同行为、起降一体化统一决策。本文为机载算力受限条件下高密度起降场运行提供全新自博弈调度技术路径。

关键词：城市空中交通；电动垂直起降飞行器；垂直起降场调度；拓扑博弈；自博弈强化学习；蒙特卡洛树搜索；图神经网络；多智能体协同

1 引言

1.1 现有起降场调度范式核心瓶颈

大规模商用城市 eVTOL 交通依赖高效、安全的起降场终端区调度。当前主流调度技术路线分为集中式数学优化、固定启发式规则、单智能体强化学习三类，均存在固有缺陷，制约高密度空中交通运行。

- 集中式 MILP/约束规划优化范式**混合整数线性规划与约束规划以安全间隔、起降坪容量、能耗约束构建全局优化模型，在小流量场景下可输出理论最优解。但求解耗时随 eVTOL 数量呈指数增长；高峰密集流量下，集中求解器无法在机载决策时间窗口内完成实时调度，且高度依赖全局完整交通信息，通信链路中断时无法支撑分布式独立决策。
- 固定启发式规则调度 (FCFS、EDF)** 先到先服务、最早截止时间优先因计算开销低被广泛部署于仿真平台。但刚性时序规则无法实现异质航班间主动时隙协同；高优先级医疗应急航班会被普通低优先级飞行器阻塞，造成整体延误过高、起降坪平均利用率低下。同时静态规则无法适配交通动态波动与临时空域关闭事件。

3. **无长时域规划的单/多智能体强化学习**现有基于多智能体强化学习（MARL）的起降场调度方法大多以瞬时奖励反馈作为优化目标，缺少多步时序前瞻推理能力；模型仅优化即时飞行收益，忽略未来时隙拥堵风险。同时多数强化学习框架将进港、离港建模分离，起降任务需两套独立网络，缺少空域资源统一拓扑表征。

本质上，现有调度方法忽略了起降场空域天然具备的博弈属性：低空通道、等待区、起降坪构成固定组合拓扑空间，航班资源为具备独占约束的离散时隙。传统方法将调度简化为纯轨迹优化或单步收益最大化，忽视时隙竞争的序贯博弈本质，导致多步前瞻能力不足、分布式协同性能差。

1.2 AlphaZero 自博弈范式迁移至起降场调度的可行性

AlphaZero 依托封闭博弈环境、离散序贯动作、策略-价值双网络与 MCTS 多步搜索实现规则驱动最优策略涌现，完全摆脱人工专家棋谱依赖。起降场调度与棋类博弈存在三点核心逻辑同源性：

1. 封闭结构化博弈边界：空域拓扑、安全间隔标准、起降场运行规则构成固定硬约束，等价于棋类规则；
2. 离散序贯决策空间：时隙申请、盘旋等待、改航、进近着陆均可抽象为离散高层博弈动作，对应棋步；
3. 可量化有限博弈收益：碰撞惩罚、准点着陆奖励、能耗损失构成多目标回报信号，存在清晰博弈终止状态（成功着陆/冲突碰撞）。

空中调度与棋类博弈的核心差异：起降场博弈为多智能体合作-竞争混合博弈，而非双人零和博弈。因此本文增加分层协同奖励项，引导自博弈生成互相礼让的协同行为，避免无限制抢占时隙等极端竞争纳什均衡。

基于上述同源性与针对性改进，本文将 AlphaZero 自博弈流水线迁移至起降场时序时隙调度：依托空域拓扑基元构建博弈棋盘，以时隙序列决策替代棋步，依托航空安全规则约束博弈空间，通过仿真自博弈生成分布式多智能体协同策略。

1.3 本文核心贡献

本文聚焦拓扑博弈系统设计与自博弈调度算法创新，而非提出全新基础 GNN/MCTS 模块，五项主要贡献如下：

1. **拓扑博弈调度范式创新**：将 eVTOL 起降场起降调度重构为定义在 2.5 维模块化空域拓扑上的时序时隙占用组合博弈，提供可替代集中式优化、固定启发式规则的分布式自博弈方案。
2. **模块化空域拓扑基元体系**：提出四类低空空域核心基元 + 起降坪终端基元，支持程序化生成任意规模单/多坪起降场终端博弈环境。
3. **时序折扣 MCTS-UCT 改进**：设计时序衰减 UCT 选择公式，提升近时时隙资源权重，匹配起降场时间敏感特性；结合 TopoMoE 实现稀疏远距离通道与密集起降坪区域统一编码。
4. **起降一体化统一自博弈流水线**：单套策略-价值网络同时处理着陆、起飞博弈任务，避免模型堆叠；融合拓扑感知分组相对策略优化（Topology-GRPO）稳定多智能体自博弈训练，抑制极端竞争策略。
5. **分阶段最小可行验证体系**：提出从单智能体单起降坪到多起降场网络自博弈的分步仿真路线，设计吞吐量、平均延误、冲突率等量化指标，与 FCFS、MILP 基线对比。

2 相关工作

2.1 起降场调度优化研究

现有起降场调度研究分为数学优化、启发规则、数据驱动学习三大技术分支。MILP、CP 等数学优化方法在小流量下可得到全局最优解，但求解复杂度随 eVTOL 数量激增，无法支撑机载实时决策；FCFS、EDF 等启发式算法计算开销低，但缺少主动多步协同能力，高峰时段起降坪吞吐量偏低。

数据驱动调度主要采用标准强化学习与多智能体强化学习框架。现有多数强化学习调度模型仅以瞬时延误、冲突作为奖励目标，未引入长时域多步规划机制；少量研究引入基础 MCTS 解决单机冲突，但未融合空域图拓扑表征与多智能体自博弈迭代。尚无现有工作将起降场资源竞争建模为拓扑组合博弈，并采用 AlphaZero 式全自博弈生成分布式协同策略。

2.2 低空空域拓扑感知图学习

几何深度学习（GNN）已广泛应用于空域轨迹预测、航线特征提取。现有空域图模型大多以静态飞行路径为输入完成感知预测，未构建可生成式拓扑博弈环境支撑序贯调度决策。TopoMoE 混合专家网络已被证实适用于异质空间特征提取，但尚未在城市空中交通场景统一稀疏远距离通道与密集起降终端区域。本文针对该空白定制 TopoMoE 骨干网络，完成双尺度空域拓扑编码。

2.3 序贯资源博弈下的 MCTS 与自博弈

AlphaZero 系列自博弈框架依托 MCTS 树搜索与策略迭代实现最优策略涌现。现有工业落地拓展集中于物流调度、机器人路径规划，极少应用于航空场景时序时隙资源博弈问题。传统 MCTS 采用固定 UCT 系数，未针对时间敏感资源竞争引入时序折扣。本文提出时序衰减 UCT 公式适配时隙调度特征，并结合 Topology-GRPO 解决空中交通场景多智能体自博弈收敛不稳定问题。

2.4 城市空中交通分布式多智能体协同

集中式无人机交通管理（UTM）系统存在单点失效风险与高峰计算瓶颈；分布式多智能体强化学习协同方案逐步兴起，但大多依赖全局所有智能体状态通信，缺少可解释多步搜索逻辑。本文提出的拓扑博弈框架仅通过邻机飞行意图广播，各飞行器独立机载 MCTS 决策，为中央 UTM 通信中断时提供备用协同方案。

3 eVTOL 起降场调度拓扑博弈建模

3.1 空域拓扑基元：可生成式起降场博弈棋盘单元

本文定义四类核心低空通道基元与一类终端起降坪基元，所有城市起降场终端空域均可由上述模块程序化拼接、嵌套构成：

1. 线性进港通道单元：固定高度线性进近航道，供 eVTOL 下降进入起降场空域；
2. 线性离港通道单元：固定高度线性爬升航道，供 eVTOL 驶离终端区；
3. 环形等待区单元：闭合环形悬停空域，供延误航班等待时隙；
4. 汇流交汇单元：多条进港通道交叉口，存在多机时隙冲突风险；
5. 起降坪终端基元（核心微博弈单元）：由多组独立起降坪构成高密度约束微博弈空间，是时序时隙占用博弈的核心载体。

起降坪基元为本框架核心扩展模块：每一块起降坪为带固定时隙时长约束的独占时间资源节点；多架 eVTOL 竞争连续空闲时间窗口，形成典型长序列时序组合博弈，完美匹配 MCTS 多步前瞻搜索逻辑。TopoMoE 网络自动切换稀疏通道特征提取分支与密集起降坪分支，实现宏观空域与微观坪区统一编码。

3.2 AlphaZero 博弈逻辑至起降场调度映射关系

AlphaZero 棋类博弈概念	起降场拓扑博弈定义
固定棋盘	由通道/起降坪基元程序化拼接生成空域
离散博弈动作	时隙申请/盘旋等待/改航/执行着陆
博弈硬规则	最小安全间隔、起降坪时隙时长、高度约束
博弈收益	碰撞惩罚 + 准点着陆吞吐量 + 能耗损失
自博弈输出	仿真生成多智能体时隙协同序列
博弈拓扑	带起降坪时间资源的 2.5 维点边空域图

3.3 图结构博弈状态表征

全局起降场博弈状态定义为四元组图观测：

$$S = (G_{air}, Ego_{VTOL}, Traffic_{Intent}, SlotTable)$$

- G_{air} : 2.5 维多尺度空域拓扑图；节点 = 航路点/起降坪/等待点；边 = 带高度属性的飞行通道；
- Ego_{VTOL} : 本机 eVTOL 状态：三维位置、速度、剩余电量、任务优先级、目标起降坪；
- $Traffic_{Intent}$: 周边飞行器广播飞行意图（仿真添加噪声用于鲁棒性测试）；
- $SlotTable$: 全局起降坪时隙占用表，记录各起降坪占用/空闲时间窗口。

3.4 离散高层动作空间（时隙博弈动作）

所有动作均为抽象高层调度意图，非底层飞控指令；MCTS 输出最优动作后，轨迹生成器与飞控输出满足气动约束的平滑三维航路。经动作掩码过滤后包含三类合法动作：

1. 申请时隙 (REQUEST SLOT)：申请未来 N 个周期目标起降坪空闲时间窗口；
2. 盘旋等待 (HOLD)：进入指定环形等待区等候后续可用时隙，附带能耗惩罚；
3. 复飞改航 (GO AROUND)：放弃当前进近，转至备用起降场或长期延后。

动作掩码硬约束：禁止申请已占用时隙、禁止低电量长时间盘旋、禁止应急高优先级航班改航。MCTS 在起降坪终端区自动扩大搜索深度，处理长序列时隙竞争决策。

3.5 分层多目标时隙奖励函数

按安全优先级分为四层加权奖励体系：

3.5.1 第一层：安全硬惩罚（最高权重）

空域间隔违规、起降坪时隙碰撞给予大额负奖励，直接终止当前博弈片段。

3.5.2 第二层：核心任务完成奖励

在申请时隙内准点着陆给予大额正奖励；晚到延误周期对应梯度负奖励。

3.5.3 第三层：资源利用与能耗成本

选择低利用率起降坪获得小额正奖励均衡负载；盘旋飞行持续扣除能耗损失负奖励。

3.5.4 第四层：协同激励奖励

主动盘旋礼让高优先级 eVTOL 获得正向奖励，引导自博弈生成非自私协同策略，规避极端抢占时隙纳什均衡。

4 五层整体框架核心算法设计

4.1 五层起降场拓扑博弈流水线

1. 空域拓扑生成层：拼接通道/起降坪基元，生成单/多坪起降场仿真环境；
2. TopoMoE 多尺度编码层：GNN+ 专家路由提取统一稀疏通道/密集起降坪拓扑嵌入 z_{air} ；
3. 时序 MCTS 博弈决策层：输入 $z_{air} \rightarrow$ 时序折扣 UCT 搜索 $\rightarrow$ 策略/价值双网络输出；
4. 师生蒸馏层：将 MCTS 高质量调度策略蒸馏为轻量化机载决策网络；
5. 飞行动态仿真迁移层：将离散时隙决策转化为满足气动约束的 eVTOL 平滑三维轨迹。

三层递进学习核心：1. 几何深度学习构建空域拓扑结构先验；2. TopoMoE 实现远距离通道与终端起降坪跨尺度统一编码；3. Topology-GRPO 稳定多智能体自博弈，抑制极端竞争时隙策略。

4.2 GNN+TopoMoE 空域拓扑编码

空域图定义 $G_{air} = (V, E, X_V, X_E)$ ：

- 节点集 V ：进近航路点、等待节点、起降坪终端节点；
- 边集 E ：带固定高度、最小间隔约束的有向进港/离港通道边；
- 节点特征 X_V ：节点类型、起降坪容量、当前时隙占用、盘旋时长上限；
- 边特征 X_E ：通道飞行时长、垂直/水平安全间隔要求。

图卷积分聚合公式：

$$h_v^{(l)} = \sigma \left(W^{(l)} \cdot \text{AGG} \left(h_v^{(l-1)}, \left\{ h_u^{(l-1)} | u \in \mathcal{N}(v) \right\} \right) \right)$$

TopoMoE 路由规则：远距离稀疏通道节点分配至远程专家分支；起降坪密集节点路由至微时隙专家分支，自适应特征提取，无需两套独立网络。

4.3 时序折扣 MCTS-UCT 核心算法

4.3.1 策略-价值网络输入输出

策略网络输入：空域嵌入 z_{air} + 本机 eVTOL 状态；输出经飞行约束掩码后的合法时隙动作概率分布。价值网络输入与策略网络一致；输出片段期望累积回报 $v(S) \in [-1, 1]$ 。

4.3.2 适配时隙调度的时序折扣 UCT 选择公式

$$a^* = \arg \max_a \left[\frac{Q(s, a)}{N(s, a)} + c_{puct} \cdot \pi(a|S) \cdot \frac{\sqrt{\sum_b N(s, b)}}{1 + N(s, a)} \cdot \gamma^{t_{ahead}} \right]$$

式中 $\gamma \in (0, 1)$ 为时序折扣因子, t_{ahead} 代表目标时隙距离当前时刻的未来周期数。近时时隙资源在搜索中获得更高权重, 匹配起降场调度时间敏感需求。MCTS 四步流程: 选择 $\rightarrow$ 扩展 $\rightarrow$ 评估 $\rightarrow$ 反向传播。

4.3.3 单轮多智能体自博弈片段算法

算法 1: 多 eVTOL 起降场时隙自博弈片段输入: 空域生成器、策略网络、价值网络、最大周期 T_{max}
输出: 多智能体调度轨迹数据集

1. 初始化起降场空域图、同构 eVTOL 智能体组、空轨迹列表;
2. **while** 当前周期 $< T_{max}$ 且无全局碰撞 **do**
 - (a) 通过 TopoMoE GNN 编码全局空域图得到 z_{air} ;
 - (b) 对每组 eVTOL 智能体循环: 基于局部观测执行完整时序折扣 MCTS-UCT 搜索, 从 MCTS 访问频次分布采样时隙动作, 将 $(S, \cdot)$ 存入智能体轨迹;
 - (c) 同步执行所有智能体时隙动作, 更新全局时隙表与空域交通状态, 计算各智能体单步奖励;
3. 通过反向回报传播, 用最终片段回报 v_{final} 标记全部轨迹样本, 输出完整多智能体轨迹数据集。

4.3.4 网络损失函数

$$\mathcal{L}_{policy} = - \sum_a \rho_{mcts}(a) \log \pi_{net}(a) \quad (1)$$

$$\mathcal{L}_{value} = (v_{net}(S) - v_{final})^2 \quad (2)$$

$$\mathcal{L}_{total} = \mathcal{L}_{policy} + \lambda \cdot \mathcal{L}_{value} \quad (3)$$

4.3.5 机载轻量化策略蒸馏算法

将 MCTS 搜索动作分布作为教师标签, 训练小型多层感知机学生网络。蒸馏完成后, 学生网络无需在线运行高开销 MCTS 即可输出实时时隙调度决策, 满足机载低算力实时约束。

4.4 动态 MCTS 搜索深度策略

1. 远距离进/离港通道: 浅搜索深度, 侧重粗粒度时隙预留规划;
2. 起降坪终端区: 深搜索深度, 多周期时隙序列前瞻, 解决多机竞争冲突。

5 相较传统调度方法的差异化优势

1. **无中心求解器分布式决策:** 每架 eVTOL 依托局部空域观测独立运行 MCTS 调度, 不依赖全局 UTM 集中优化, 通信中断场景鲁棒性强;
2. **无需专家数据, 自博弈生成调度轨迹:** 不需要人工标注历史航班时刻表; 仅依靠规则约束仿真自博弈涌现协同策略, 降低数据依赖;

3. **可解释多步时序规划**：全部时隙决策源自 MCTS 搜索树，冲突、延误成因可追溯，满足航空安全审计要求；
4. **起降一体化统一建模**：单套网络同时处理进港、离港时隙博弈，避免现有研究起降模型分离堆叠；
5. **高峰流量性能优于启发式规则**：自博弈涌现主动礼让协同行为，相较 FCFS 固定规则降低平均航班延误、提升起降坪吞吐量。

与 MILP 优化核心区别：MILP 在完整信息下追求全局单步最优解；本文框架通过多步博弈搜索生成长期协同时隙序列，适配动态实时交通与分布式机载执行场景。

6 四项核心科学假设分阶段验证方案

6.1 假设 1：离散时隙动作可表征全部合法起降场调度策略

核心风险：离散周期划分粒度过粗导致时隙匹配次优、航班延误增加。验证方案：设置多组周期时长消融实验组（5s/10s/15s），对比吞吐量、平均延误指标选取最优时间粒度；若性能下降则采用参数化时隙窗口动作实现细粒度调整。

6.2 假设 2：TopoMoE 图神经网络可泛化未见过的拼接空域拓扑

核心风险：模型对固定起降场布局过拟合，无法适配全新多坪空域组合。验证方案：在单坪起降场完成训练，在未参与训练的双坪/三坪拼接空域测试，量化冲突率、准点率验证拓扑泛化能力。

6.3 假设 3：多智能体自博弈可涌现理性协同时隙礼让行为

核心风险：无约束自博弈收敛至贪婪抢占时隙纳什均衡，引发终端严重拥堵。验证方案：增加优先级协同奖励项；自博弈训练初期引入规则背景飞行器作为训练锚点，约束极端竞争策略。

6.4 假设 4：时序折扣 MCTS 提升高峰调度效率

核心风险：时序衰减系数 γ 设置不当削弱长期时隙规划前瞻能力。验证方案：设置有无时序折扣 UCT 消融对照实验，在高交通密度下对比起降坪吞吐量、平均延误。

7 五阶段最小可行仿真路线图

7.1 阶段 1：单智能体单起降坪最小可行验证（1 周）

简化环境：单进港通道 + 单起降坪，规则背景 eVTOL 周期性占用时隙。验证目标：智能体自主学习盘旋等待空闲时隙、无碰撞。关键指标：冲突率、着陆准点率。

7.2 阶段 2：双智能体单起降坪多机博弈（2 周）

两架同构 eVTOL 竞争单块起降坪，设置差异化任务优先级。预期行为：低优先级飞行器主动盘旋礼让高优先级航班。指标：平均延误、高优先级航班准点率。

7.3 阶段 3：多智能体多坪空域泛化（3 周）

在单坪场景训练，在未见过的双坪组合空域测试，验证 TopoMoE 跨拓扑泛化能力。指标：起降坪平均利用率、系统总吞吐量。

7.4 阶段 4：无背景规则飞行器完整自博弈（3 周以上）

取消固定规则流量，全部 eVTOL 采用同一套自博弈策略，观测自发协同行为。指标：多智能体冲突率、延误分布。

7.5 阶段 5：基线对比实验（1 周）

在低/中/高峰三类交通密度下，将本文拓扑博弈框架与 FCFS 启发式、小规模 MILP 求解器对比。核心对比指标：起降场小时吞吐量、航班平均延误、冲突发生频次。

8 学术价值与工程应用前景

8.1 学术价值

本文构建面向 eVTOL 起降场时序调度的拓扑组合博弈理论体系，打破城市空中交通调度领域依赖集中式优化或单步强化学习的研究倾向；创新性将 AlphaZero 自博弈范式迁移至时隙资源竞争航空场景，融合 GNN 多尺度拓扑编码、时序折扣 MCTS、Topology-GRPO 稳定自博弈优化构建统一调度流水线，填补分布式博弈式起降场协同研究空白。

8.2 工程应用场景

1. 分布式机载调度备用模块：部署于 eVTOL 机载计算机，中央 UTM 通信中断时提供独立时隙协同决策；
2. 起降场预仿真评估平台：无需真实试飞，批量测试全新空域拓扑高峰交通吞吐量；
3. 高密度城市终端协同工具：解决市中心紧凑型起降场坪数量有限下的多机冲突、延误问题。

8.3 未来研究方向

1. 在 MCTS 推演中引入风扰、广播意图噪声，提升调度对不确定环境信息的鲁棒性；
2. 将框架拓展至多起降场跨空域时隙协同博弈；
3. 优化参数化连续时隙动作，降低离散周期粒度限制带来的延误；
4. 开发开源起降场拓扑博弈仿真库，支撑城市空中交通算法研究。

9 结论

本文提出一套面向 eVTOL 起降场起降的拓扑博弈自博弈调度框架。通过抽象模块化低空空域拓扑基元，将起降场时隙竞争问题建模为定义在结构化 2.5 维空域上的时序占用组合博弈；结合 TopoMoE 多尺度图编码与时序折扣改进 MCTS-UCT 自博弈搜索，构建起降任务统一分布式调度流水线，仅依靠仿真自博弈自主生成多智能体协同策略，无需人工航班时刻表专家数据。

区别于集中式数学优化与固定启发式调度方案，本框架具备分布式机载决策、可解释多步时序前瞻、原生起降一体化建模特征。文中提出分阶段仿真验证路线与四项核心科学假设，指导后续消融实验与基线对比。本文为高密度城市起降场运行提供全新自博弈博弈技术路径，同时为低空空域交通其他时序资源调度任务提供通用拓扑博弈范式参考。

参考文献

- [1] Bronstein M M, Bruna J, LeCun Y, et al. Geometric deep learning: Going beyond Euclidean data[J]. IEEE Signal Processing Magazine, 2017, 34(4): 18-42.
- [2] Varnousfaderani, et al. Mixed integer programming for eVTOL vertiport dispatching[C]//AIAA Aviation Conference, 2026.
- [3] Shao T, et al. Group relative policy optimization[EB/OL]. arXiv preprint arXiv:2406.09279, 2024.
- [4] Speed-based vertiport conflict resolution via CP-SAT[J]. Aerospace, 2026, 13(6).
- [5] eVTOL[R]. 2026.
- [6] Heterogeneous vertiport selection deep RL framework[EB/OL]. arXiv:2601.21316, 2026.
- [7] Guerreiro, et al. FCFS-based vertiport capacity evaluation[J]. Transportation Research, 2025.
- [8] Wei P. MCTS multi-UAM deconfliction at terminal airspace[C]//Aviation Proceedings, 2019.
- [9] Silver D, et al. Mastering the game of Go without human knowledge[J]. Nature, 2017, 550(7676): 354-359.
- [10] Silver D, et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play[J]. Science, 2018, 362(6419): 1140-1144.
- [11] Murthy, et al. MCTS UAM aperiodic scheduling[C]//FMAS Workshop, 2022.
- [12] Elbanhawi M, Simic M. Sampling-based robot motion planning: A review[J]. IEEE Access, 2014, 2: 56-77.
- [13] Vertiport multi-direction surface scheduling[EB/OL]. arXiv:2408.01152, 2026.
- [14] Topology-aware multi-UAV potential game coordination[EB/OL]. arXiv:2603.07456, 2026.
- [15] TAG-MAPPO 3D UAV topology coordination[EB/OL]. arXiv:2602.10029, 2026.
- [16] Bandara W C. Aerial graph road extraction[C]//ICRA, 2022.