

动态交通流下低空航路网的事件触发在线修正

汪一村^{1,2}

(1.北京城建勘测设计研究院有限责任公司, 北京 100101; 2.温州幸福轨道数智科技有限公司, 浙江 温州 325035)

摘要: 面向动态拥堵条件下低空航路网的战术级在线修正, 构建以同时占用架次为统一量纲的航段容量与动态边权模型, 提出“全局主路径—事件触发—有限子图局部规划—规则筛选—确定性回退”框架。局部 Dijkstra、锚点 Dijkstra、锚点 A* 与 DQN 在统一的有限子图、锚点和步数上限下接入, 触发式 D* Lite 作为采用相同事件触发器的增量全图基线。基于 1 063 个节点、4 044 条有向边及 30 组随机种子的控制仿真表明: 事件触发与局部化是减少在线计算范围的主要来源; Dijkstra-DQN 较纯 DQN 提高到达率, 但在局部突发和全网高峰场景中未优于同范围 A* 或触发式 D* Lite。因而本文不宣称 DQN 具有独立性能优势, 而将贡献限定为可复现、可审查、可回退的在线修正框架及其适用边界。

关键词: 低空航路网; 在线修正; 动态交通流; 事件触发; 深度 Q 网络

中图分类号: V279; U491

文献标志码: A

Event-Triggered Online Route Correction for Low-Altitude Air Route Networks under Dynamic Traffic Flow

Wang Yicun^{1,2}

(1. Beijing Urban Construction Survey and Design Institute Co., Ltd., Beijing 100101, China; 2. Wenzhou Xingfu Rail Transit Digital Intelligence Technology Co., Ltd., Wenzhou 325035, China)

Abstract: Tactical route correction in a low-altitude network must respond to congestion without transferring the entire planning task to a black-box policy. This paper defines segment capacity and estimated occupancy in the common unit of concurrent aircraft, builds non-negative dynamic edge weights, and proposes an event-triggered framework consisting of a global backbone route, bounded local replanning, rule-based action screening, anchor reconnection, and deterministic fallback. Local Dijkstra, anchor-based Dijkstra/A*, and DQN are evaluated under the same bounded subgraph, anchor, and step limit; trigger-based D* Lite serves as an incremental whole-graph baseline using the same event detector. A reproducible simulation with 1 063 nodes, 4 044 directed edges, and 30 paired random seeds shows that event triggering and localization, rather than the learning policy alone, are the main sources of reduced online search scope. The Dijkstra-DQN realization improves route-completion rate over pure DQN, but does not outperform same-range A* or trigger-based D* Lite in local-disruption and network-wide peak scenarios. The contribution is therefore limited to an auditable and recoverable online-correction architecture and an explicit account of its operating boundary; no claim of operational safety or independent DQN superiority is made.

Keywords: low-altitude air route network; online route correction; dynamic traffic flow; event triggering; deep Q-network

城市低空物流与公共航路网由起降点、分层走廊、容量约束和临时限制共同构成。本文讨论给定航路拓扑上的战术级在线修正: 输入起终点、当前位置及航段状态, 输出供上层交通管理系统复核的路径片段。该问题不同于飞控层传感器避障、多机冲突解脱和完整低空交通管理; 后者仍需处理最低间隔、通信、气象、运行许可和应急处置。

非负权图上的 Dijkstra、A* 和 LPA* 具有确定性、可解释和易复核等特点^[1-3]。低空航路网研究进一步考虑建筑物、风险、网络结构与拥堵定价^[4-7]。DQN 及多智能体强化学习可从交互中学习局部决策^[8-10], 但其训练分布依赖、约束处理和真实系统迁移仍是关键问题^[11-13]; 搜索与学习的混合框架能够兼顾全局骨架和局部适应性^[14-15]。

1 动态航路网与时空边权

1.1 三维有向航路网

将低空航路网表示为三维有向动态加权图

$$G(t) = (V, E, \mathbf{w}(t)) \quad (1)$$

式中： $G(t)$ 为时刻 t 的三维有向动态加权图； V 为节点集合，节点包含水平坐标、高度层和运行属性； E 为有向航段集合； $\mathbf{w}(t)$ 为时变非负边权向量。相反方向的航段分别建弧，以表达单向走廊、分层运行和方向相关容量。建筑物、净空不足区和临时禁限飞区通过节点筛除与边扫描检查从可用拓扑中删除，而不是交由奖励函数“试错”。该处理仅保证路径满足本文编码的图层约束，不等同于多机无冲突或飞行安全保证。图 1 以分层节点、有向航段和换层航段直观说明上述拓扑关系。

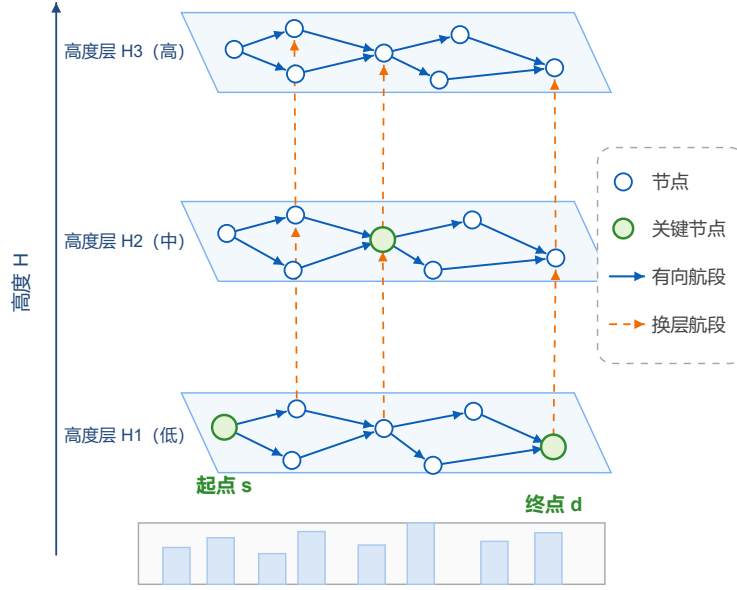


图 1 三维有向低空航路网示意

Fig. 1 Schematic of a three-dimensional directed low-altitude route network

1.2 航段容量、拥堵指数与动态边权

为消除量纲混用，本文把航段容量 K_{ij} 统一定义为同一观测时刻允许同时占用的飞行器架次，单位为“架”；估计占用量 $\hat{N}_{ij}(t)$ 亦以“架”计。其中， i 和 j 分别为有向航段的起、终节点索引。二者之比为无量纲拥堵指数

$$\rho_{ij}(t) = \frac{\hat{N}_{ij}(t)}{K_{ij}} \quad (2)$$

若工程系统只提供进入率 q_{ij} (架/min)，需先结合平均通过时间 τ_{ij} (min) 换算为占用量 $\hat{N}_{ij} \approx q_{ij}\tau_{ij}$ ，再代入式 (2)，不得将流率与同时占用容量直接相除。为抑制泊松抽样和监视噪声导致的路径摆动，采用指数平滑

$$\bar{\rho}_{ij}(t) = \eta \rho_{ij}(t) + (1 - \eta) \bar{\rho}_{ij}(t - \Delta t) \quad (3)$$

式中： $\eta \in (0, 1]$ 为指数平滑系数， Δt 为相邻状态更新的采样间隔。边权统一以分钟计，由自由流时间、拥堵惩罚、换层代价、风险代价和关闭指示组成

$$w_{ij}(t) = t_{ij}^0 \left[1 + \alpha \bar{\rho}_{ij}^\beta(t) \right] + \lambda_h I_{ij}^h + \lambda_r r_{ij}(t) + M z_{ij}(t) \quad (4)$$

式中： t_{ij}^0 为自由流通过时间； α 和 β 为无量纲拥堵参数； I_{ij}^h 为取 0 或 1 的换层指示； λ_h 为换层等效时间； $r_{ij}(t)$ 为归一化风险； λ_r 为风险等效时间； $z_{ij}(t)$ 为取 0 或 1 的禁限飞或容量失效指示。 M 取大于任一可行路径总代价上界的常数。所有项均非负，因此可使用 Dijkstra 类算法。主要符号及量纲见表 1。

表 1 主要符号、量纲与含义
Tab. 1 Main symbols, units and definitions

符号	单位	含义
V, E	—	节点集合、有向航段集合
$K_{ij}, \hat{N}_{ij}(t)$	架	航段同时占用容量、估计占用量
$q_{ij}, \tau_{ij}, \Delta t$	架/min; min; min	进入率、平均通过时间、采样间隔
$\rho_{ij}(t), \bar{\rho}_{ij}(t), \eta$	1	瞬时/平滑拥堵指数及平滑系数
$t_{ij}^0, w_{ij}(t)$	min	自由流时间、动态边权
α, β	1	拥堵增幅系数与指数
$I_{ij}^h, r_{ij}(t), z_{ij}(t)$	1	换层指示、归一化风险、关闭指示
λ_h, λ_r, M	min	换层/风险等效时间及关闭惩罚常数
$\theta_p, \theta_w, \epsilon_0$	1	拥堵/权重突变阈值及分母正则常数
$\mathcal{M}, t_p, R, L_{\max}, v_a$	—; min; 跳; 步; —	滚动窗口、上次规划时刻、局部半径、最大步数、稳定锚点

1.3 在线修正目标

给定起点 s 、终点 d 、时刻 t 和当前可用拓扑，可行路径集合记为 $\Omega_{s,d}(t)$ ； P 为候选路径， $P^*(t)$ 为当前快照的最优路径。路径层目标为

$$P^*(t) = \arg \min_{P \in \Omega_{s,d}(t)} \sum_{(i,j) \in P} w_{ij}(t) \quad (5)$$

状态更新后若仍有可行路径，则由上述目标函数确定当前快照下的最小代价路径；若临时关闭使 s 与 d 不连通，算法必须返回失败并请求上层采取等待、改派、容量重分配或返航等措施。本文不再使用“全局可达性保证”等无条件表述。

2 事件触发的全局—局部混合修正

2.1 “混合”的定义与事件触发

本文所称“混合”不是若干模型的简单拼接，而是四个职责明确的模块：Dijkstra 生成并保留可解释的全局主路径；事件检测器监测前方窗口；局部求解器仅在有限子图内接管；规则层完成动作屏蔽、锚点重接入和动态 Dijkstra 回退。图 2 给出完整流程。

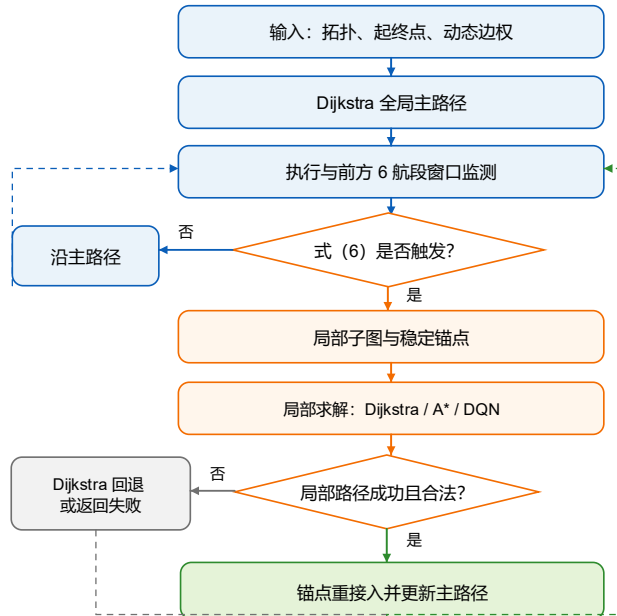


图 2 事件触发的全局—局部混合规划框架
Fig. 2 Event-triggered global-local hybrid planning framework

设 $\mathcal{W}_t$ 为当前主路径前方 6 条航段组成的滚动窗口, t_p 为上次规划时刻, e 为窗口内航段索引。当平滑拥堵指数或相对边权增量超过阈值时触发局部接管

$$\max_{e \in \mathcal{W}_t} \bar{\rho}_e(t) > \theta_\rho \quad \vee \quad \max_{e \in \mathcal{W}_t} \frac{|w_e(t) - w_e(t_p)|}{\max[w_e(t_p), \varepsilon_0]} > \theta_w \quad (6)$$

式中: θ_ρ 和 θ_w 分别为拥堵与相对边权突变阈值, ε_0 为防止分母为零的小正数。实验取 $\theta_\rho=0.70$ 、 $\theta_w=0.25$ 、 $\varepsilon_0=10^{-6}$; 恢复阈值取 0.58 以形成滞回。未触发时直接沿主路径执行, 从而避免每个时间片都进行全图重算。

2.2 局部子图、稳定锚点与公平接入

触发后, 以当前位置为中心, 按无向跳距半径 $R=9$ 截取诱导子图; 沿原主路径向前寻找至少 5 跳之外、连续局部航段 $\bar{\rho} \leq 0.58$ 且可在最大局部步数 $L_{\max}=16$ 内到达的首个节点, 并将其记为稳定锚点 v_a 。图 3 示意局部绕行和锚点重接入。

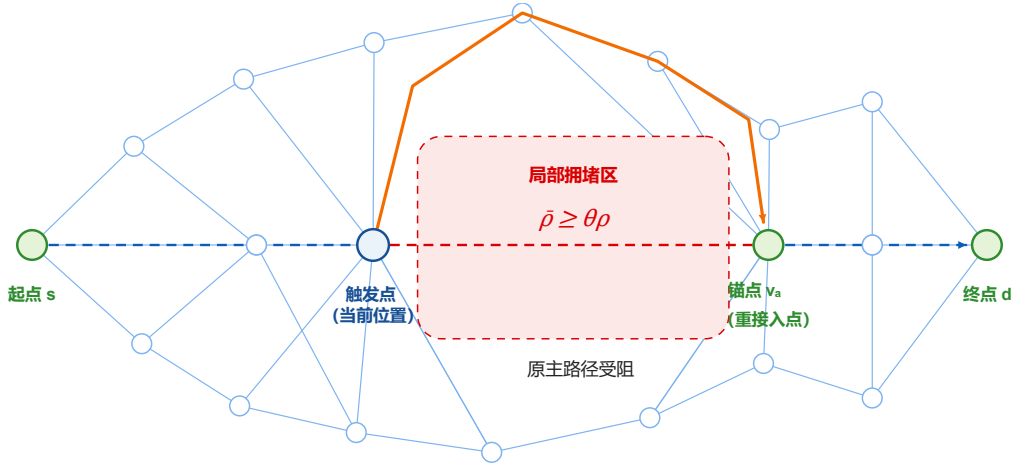


图 3 局部拥堵下的绕行与锚点重接入
Fig. 3 Local detour and anchor reconnection under congestion

局部模块采用统一接口: 事件触发式局部 Dijkstra 搜索局部边界; 锚点 Dijkstra 与锚点 A* 在相同子图、相同锚点和相同步数上限内求解; 滚动锚点 A* 每一步重新求局部最短路径; Dijkstra-DQN 在同一局部范围内用 DQN 逐步选边。触发式 D* Lite 使用相同事件触发器和动态边权进行增量全图搜索, 作为非局部基线。上述设置用于区分“有限计算范围”与“学习能力”。所有候选动作先经过禁限飞、关闭、换层和局部范围检查; 无合法动作或超出步数上限时, 调用当前快照下的全局动态 Dijkstra。

2.3 DQN 状态、奖励与训练目标

DQN 在时刻 t 的状态向量 $\mathbf{x}_t \in \mathbb{R}^{46}$: 六个候选方向各包含合法性、归一化边权、拥堵指数、拥堵趋势、朝向锚点的进展、历史访问标记和换层标记共七项, 再加三维归一化锚点偏移及剩余步数预算四项。网络为 128-128-64 三层全连接结构, ReLU 激活, 输出候选动作 a 的动作价值 $Q_\theta(\mathbf{x}_t, a)$; 非法动作在取最大值前置为不可选。

单步奖励定义为无量纲训练信号

$$u_t = -0.85c_t + 1.40\Delta\ell_t - 1.20I_t^{\text{vis}} + 30I_t^{\text{arr}} \quad (7)$$

式中: u_t 为时刻 t 的单步奖励; $c_t = w_{e_t}(t)/t_{\text{ref}}$ 为无量纲归一化航段代价, 其中 e_t 为第 t 步选择的航段, $t_{\text{ref}}=0.40 \text{ min}$ 为参考时间, 保留“少耗时、少拥堵、少风险”的物理对应; $\Delta\ell_t = \ell(n_t, v_a) - \ell(n_{t+1}, v_a)$ 为锚点图跳距离的减少量, n_t 为第 t 步所在节点, $\ell(\cdot, \cdot)$ 为图跳距离; $I_t^{\text{vis}}=1$ 表示重复访问, 否则为 0; $I_t^{\text{arr}}=1$ 表示完成锚点重接入, 否则为 0。奖励函数不参与前述全局路径优化, 因而不存在把同一代价重复计入的问题。目标值与 Huber 损失分别为

$$y_t = u_t + \gamma(1 - I_t^{\text{done}}) \max_{a \in \mathcal{A}_{t+1}} Q_{\theta^-}(\mathbf{x}_{t+1}, a) \quad (8a)$$

$$L(\theta) = \mathbb{E}[\text{Huber}(y_t - Q_{\theta}(\mathbf{x}_t, a_t))] \quad (8b)$$

式中： y_t 为目标值， γ 为折扣因子， I_t^{done} 为回合终止指示， $\mathcal{A}_{t+1}$ 为下一时刻合法动作集合， Q_{θ} 为在线网络动作价值函数， θ 为在线网络参数， θ^- 为目标网络参数， $L(\theta)$ 为 Huber 损失的期望。

训练采用 $\gamma=0.95$ 、经验池 50 000、批量 64、Adam 学习率 10^{-4} 、3 500 回合、目标网络每 500 步更新；采用 ϵ -greedy 探索策略，探索率由 1.00 线性降至 0.05。训练阶段随机化起终点与起始时刻，场景采样概率比为 S1 : S2 : S3=1 : 3 : 1。

2.4 保护逻辑与退化边界

规则层只对本文已建模的路径层约束进行条件筛选，不能替代冲突探测、间隔保持和飞控校验。局部求解成功后把绕行段与主路径尾段拼接；失败时从当前位置全局重算。若全网长期处于高负荷，事件会连续触发，局部策略可能频繁回退，此时混合框架的计算开销可接近或超过全局重算，且单机绕行不能解决系统性容量不足。

3 可复现实验与公平对比

3.1 仿真系统与随机过程

实验定位为控制仿真而非真实城市运行验证。固定拓扑覆盖 5 km×5 km，采用 17×17 水平网格和 60、90、120、150 m 四个高度层；建筑物遮挡后得到 1 063 个节点、4 044 条有向边。水平航段自由流速度 60 km/h，同时占用容量为 8 架；换层航段容量为 7 架、等效通过时间 0.22 min。

对每条航段 e 和离散时间步 t ，先由基础负荷、空间高斯扰动与正弦波动构成期望占用比 $\mu_e(t)$ ；再从参数为 $K_e \mu_e(t)$ 的泊松分布中抽样 $\hat{N}_e(t)$ ，其中 K_e 为航段 e 的同时占用容量。将实际占用比 $\rho_e(t) = \hat{N}_e(t)/K_e$ 截断于 [0,1.5] 并按 $\eta=0.55$ 平滑。S1 为自由流；S2 在中心走廊施加局部突发负荷，并于第 17~22 步关闭部分低层航段；S3 叠加两处移动高负荷区和同样的临时关闭。每个场景使用 30 个配对随机种子。可复现配置见表 2；完整生成公式、种子和代码随附件提供。

表 2 仿真与训练的可复现配置
Tab. 2 Reproducible simulation and training settings

项目	设置
拓扑	17×17×4；1 063 节点；4 044 条有向边
空间与高度	5 km×5 km；60/90/120/150 m
速度与容量	60 km/h；水平 8 架、换层 7 架
动态边权	$\alpha=1.30$ ， $\beta=4.0$ ， $\eta=0.55$ ， $\lambda_b=0.12$ min， $\lambda_r=0.10$ min
触发与局部范围	窗口 6； $\theta_p=0.70$ ； $\theta_w=0.25$ ； $R=9$ 跳； $L_{\max}=16$ 步
场景	S1 自由流；S2 局部突发；S3 全网高峰
随机样本	每场景 30 个配对种子；主种子 20260701
DQN	46-128-128-64-6； $\gamma=0.95$ ；3500 回合；批量 64；Adam 10^{-4}
软件	Python 3.13.5；NetworkX 3.6.1；PyTorch 2.10.0，CPU

3.2 基线、指标与统计口径

为分离事件触发、局部化与学习策略的贡献，设置表 3 所列 9 种方法，并统一动态边权、事件触发器、局部半径和步数上限等条件。

表 3 对比方法与公平性控制

Tab. 3 Compared methods and fairness controls

方法	在线规划范围	公平性控制
静态 Dijkstra	初始主路径	不触发；自由流权重
全局 A*	全图	相同触发器；相同动态边权
事件触发局部 Dijkstra	局部边界	相同触发器和 R
锚点 Dijkstra	局部子图	相同锚点、 R 、 $L_{\max}$
锚点 A*	局部子图	相同锚点、 R 、 $L_{\max}$
滚动锚点 A*	局部子图	每步重算；其余相同
触发式 D* Lite	增量全图	相同触发器和动态边权
纯 DQN	逐步全程	相同网络、状态、奖励和动作屏蔽
Dijkstra-DQN	触发后局部	相同锚点、 R 、 $L_{\max}$ ；失败回退

评价指标包括到达率、累计动态边权对应的飞行时间、拥堵暴露率和执行期平均在线规划时间；到达率为在 120 个仿真步内到达终点的样本比例。拥堵暴露率 χ 定义为 $\bar{\rho} \geq \theta_\rho$ 的航段飞行时间占比

$$\chi = \frac{\sum_t \tau_t \mathbf{1}[\bar{\rho}_{e_t}(t) \geq \theta_\rho]}{\sum_t \tau_t} \quad (9)$$

式中： τ_t 为第 t 个执行航段的飞行时间， e_t 为对应航段， $\mathbf{1}[\cdot]$ 为示性函数。在线耗时只统计起飞后发生的搜索或策略推理，不计所有方法共有的初始主路径构造；以总在线规划耗时除以实际执行步数

$$\bar{T}_{\text{on}} = \frac{\sum_{k=1}^{N_{\text{call}}} T_k}{N_{\text{step}}} \quad (10)$$

式中： $\bar{T}_{\text{on}}$ 为执行期平均在线规划时间， T_k 为第 k 次在线规划或推理耗时， N_{call} 为在线规划调用次数， N_{step} 为实际执行步数。同一 DQN 网络在同一 CPU 上重复 5 000 次单样本前向推理，均值为 0.027 ms、中位数为 0.027 ms。因此，纯 DQN 与 Dijkstra-DQN 的单次网络推理开销相同；二者每执行步的累计耗时差异仅来自调用频率与回退搜索。飞行时间比较采用相同随机种子的双侧配对 t 检验。

3.3 结果与机理解释

S2 局部突发场景的配对统计结果见表 4；表中 p 值均以 Dijkstra-DQN 为参照。

表 4 S2 局部突发场景的公平对比结果

Tab. 4 Fair comparison results in the S2 local-disruption scenario

方法	到达率/%	飞行时间/min	拥堵暴露率	在线耗时/(ms·步 ⁻¹)	p 值
静态 Dijkstra	100.0	8.98±1.73	0.324±0.181	0.000	0.013
全局 A*	100.0	8.19±0.84	0.087±0.088	0.073	<0.001
局部 Dijkstra	100.0	8.77±2.10	0.206±0.149	0.767	0.046
锚点 Dijkstra	100.0	8.82±1.71	0.271±0.180	0.087	0.007
锚点 A*	100.0	8.82±1.71	0.271±0.180	0.021	0.007
滚动锚点 A*	100.0	8.85±1.69	0.224±0.170	0.184	0.017
触发式 D* Lite	100.0	8.15±0.82	0.095±0.092	0.906	<0.001
纯 DQN	86.7	9.92±2.18	0.389±0.189	0.118	0.072
Dijkstra-DQN	100.0	9.27±1.82	0.358±0.177	0.064	—

注：飞行时间、拥堵暴露率和在线耗时均为 30 组配对随机种子的统计结果； p 值为各方法与 Dijkstra-DQN 飞行时间的双侧配对 t 检验结果，“—”表示参照方法。

S2 中，触发式 D* Lite 和全局 A* 的平均飞行时间分别为 8.15 min 和 8.19 min，显著低于 Dijkstra-DQN 的 9.27 min ($p < 0.001$)；相同子图与锚点下，锚点 A* 为 8.82 min 且在线耗时仅 0.021 ms/步。Dijkstra-DQN 较纯 DQN 把到达率由 86.7% 提高到 100%，但飞行时间差异未达 0.05 显著水平 ($p = 0.072$)，且拥堵暴露率并未改善。该结果说明，当前训练配置下的 DQN 没有获得独立于局部化和回退机制的性能优势。S1 和 S3 的代表性结果见表 5。

表 5 S1 与 S3 场景的代表性结果
Tab. 5 Representative results in the S1 and S3 scenarios

场景	方法	到达率/%	飞行时间/min	拥堵暴露率	在线耗时/(ms·步 ⁻¹)
S1	静态 Dijkstra	100.0	6.65±0.67	0.003	0.000
S1	全局 A*	100.0	6.66±0.67	0.003	0.007
S1	锚点 A*	100.0	6.65±0.67	0.003	0.002
S1	触发式 D* Lite	100.0	6.67±0.66	0.003	0.030
S1	纯 DQN	83.3	6.94±1.10	0.002	0.112
S1	Dijkstra-DQN	100.0	6.65±0.67	0.003	0.016
S3	静态 Dijkstra	100.0	13.41±1.93	0.942	0.000
S3	全局 A*	100.0	12.89±1.79	0.841	0.452
S3	锚点 A*	100.0	14.30±1.99	0.942	0.064
S3	触发式 D* Lite	100.0	12.89±1.79	0.841	5.807
S3	纯 DQN	93.3	14.57±2.47	0.951	0.118
S3	Dijkstra-DQN	100.0	14.62±3.22	0.957	0.114

注：每个场景均使用 30 组配对随机种子；飞行时间为均值±标准差，拥堵暴露率与在线耗时的定义分别见式（9）和式（10）。

S1 自由流中，各搜索或混合方法的平均飞行时间均约 6.65 min，说明滞回触发没有引入明显绕行；纯 DQN 到达率仅 83.3%。S3 全网高峰中，全局 A*和触发式 D* Lite 为 12.89 min，Dijkstra-DQN 为 14.62 min；事件触发局部 Dijkstra 还出现 90%到达率和 40.05 min 平均时间，表明局部边界策略在系统性拥堵下可能反复绕行。

从公平基线可得两点。第一，在线计算范围的缩小主要来自事件触发和有限子图，而非 DQN 前向推理本身；锚点 A*在 S2 中同时取得更短路径和更低耗时。第二，主路径、动作屏蔽和回退能提高学习策略的完成率，但不能据此推出更优旅行代价。故摘要和结论均删除原稿中的“保证安全”“保证可达”和 DQN 显著优于既有算法等主张。

4 工程边界与局限

本文输出的是路径层候选航段序列。实际执行前仍需由上层系统确认航段预约、其他飞行器占用、通信链路、气象、最低间隔、飞控性能和运行许可。动作屏蔽只在已知拓扑和观测状态下排除不合规边；对非合作目标、监视漏检、突发通信失效和模型外风险不提供理论安全保证。

当前实验使用合成建筑、合成 OD 和合成动态占用，不能外推为真实城市运行有效性。DQN 在公平基线下未显示独立优势，后续应优先用真实轨迹和容量记录校准 K_{ij} 、 α 、 β 和风险等效时间，并在数字孪生、半实物和限定航线实飞中逐级验证；若继续采用学习策略，应增加不确定性估计、离线策略评估和约束强化学习，而不是仅通过扩大网络规模改善结果。

5 结论

- 1) 统一以同时占用架次描述航段容量和占用量，构建量纲一致、非负且可校准的时空动态边权。
- 2) 提出事件触发的全局—局部修正框架，明确“混合”为全局主路径、受限局部求解、规则筛选、锚点重接入和确定性回退；框架允许多种局部求解器在同一条件下公平接入。
- 3) 30 组配对随机种子的控制仿真表明，Dijkstra-DQN 提高了纯 DQN 的到达率，但未优于同范围 A*或触发式 D* Lite；事件触发与局部化才是降低在线搜索范围的主要来源。本文据此限定贡献和适用边界，不作真实运行安全与 DQN 性能优势的外推。

参考文献：

- [1] Dijkstra E W. A note on two problems in connexion with graphs [J]. Numerische Mathematik, 1959, 1 (1): 269-271.

- [2] Hart P E, Nilsson N J, Raphael B. A formal basis for the heuristic determination of minimum cost paths [J]. IEEE Transactions on Systems Science and Cybernetics, 1968, 4 (2): 100-107.
- [3] Koenig S, Likhachev M, Furcy D. Lifelong planning A [J]. Artificial Intelligence, 2004, 155 (1/2): 93-146.
- [4] Zhang Honghai, Tian Tian, Feng Ouge, et al. Research on public air route network planning of urban low-altitude logistics unmanned aerial vehicles [J]. Sustainability, 2023, 15 (15): 12021.
- [5] Li Jie, Shen Di, Yu Fuping, et al. A method for air route network planning of urban air mobility [J]. Aerospace, 2024, 11 (7): 584.
- [6] Cummings C, Mahmassani H. Airspace Congestion, flow Relations, and 4-D fundamental Diagrams for advanced urban air mobility [J]. Transportation Research Part C: Emerging Technologies, 2024, 159: 104467.
- [7] He Xinyu, Li Lishuai, Mo Yanfang, et al. A distributed route network planning method with congestion pricing for drone delivery services in cities [J]. Transportation Research Part C: Emerging Technologies, 2024, 160: 104536.
- [8] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning [J]. Nature, 2015, 518 (7540): 529-533.
- [9] Boulares M, Fehri A, Jemni M. UAV path planning algorithm based on Deep Q-Learning to search for a floating lost target in the ocean [J]. Robotics and Autonomous Systems, 2024, 179: 104730.
- [10] Deniz S, Wu Yufei, Shi Yang, et al. A reinforcement learning approach to vehicle coordination for structured advanced air mobility [J]. Green Energy and Intelligent Transportation, 2024, 3 (2): 100157.
- [11] Garc{a} J, Fernández F. A comprehensive survey on safe reinforcement learning [J]. Journal of Machine Learning Research, 2015, 16: 1437-1480.
- [12] Dulac-Arnold G, Levine N, Mankowitz D J, et al. Challenges of real-world reinforcement learning: definitions, benchmarks and analysis [J]. Machine Learning, 2021, 110 (9): 2419-2468.
- [13] Chen Bolei, Zhong Ping, Cui Yongzheng, et al. EMExplorer: an episodic memory enhanced autonomous exploration strategy with Voronoi domain conversion and invalid action masking [J]. Complex & Intelligent Systems, 2023, 9 (6): 7365-7379.
- [14] Liu Ning, Shen Sen, Kong Xiangrui, et al. Cooperative hybrid multi-agent pathfinding based on shared exploration maps [J]. Journal of Intelligent & Robotic Systems, 2025, 111 (3): 97.
- [15] 赵星, 陈娴. 基于多尺度动态时空卷积网络的交通流速度预测 [J]. 华东交通大学学报, 2026, 43 (3): 13-21.
Zhao Xing, Chen Xian. Traffic speed prediction based on a multi-scale dynamic spatio-temporal convolutional network [J]. Journal of East China Jiaotong University, 2026, 43 (3): 13-21.



通信作者: 汪一村 (1980—), 男, 高级工程师, 研究方向为低空技术与工程。