



专题：低空物联网与低空安全

面向高动态城市空中交通网络的多径智能传输策略

王煜婷, 冷甦鹏, 熊凯

(电子科技大学信息与通信工程学院, 四川 成都 611731)

摘 要: 飞行汽车作为低空经济的重要载体, 对于大力发展城市空中交通、构建高效的低空物联网具有十分重要的作用。然而, 灵活和广泛的飞行范围导致飞行汽车网络通信链路不稳定, 网络拓扑动态变化。针对这些问题, 提出了一种多径智能传输策略, 以实现高效的数据传输。首先提出了一种基于强化学习的多径智能路由算法, 该算法不仅能有效提升路由效率、降低路由时延, 还能够根据不同的传输需求动态划分数据包, 以适应物理环境的变化。为进一步提高路由算法对动态环境的适应性, 设计了一种环境检验机制来判断当前路由策略与动态网络的匹配程度, 实现了在路由过程中自适应地调整路由策略。仿真实验表明, 在不同情境下, 基于强化学习的多径智能传输策略在有效降低端到端传输时延的同时, 还提高了数据成功恢复的概率。

关键词: 飞行汽车; 多径智能路由算法; 动态网络; 强化学习; 城市空中交通

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025099

Multipath intelligent transmission strategy for UAM dynamic networks

WANG Yuting, LENG Supeng, XIONG Kai

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Abstract: As an important vehicle for the low-altitude economy, the flying cars play a crucial role in vigorously developing urban air mobility (UAM) and building an efficient low-altitude intelligent network. However, the high mobility of flying cars introduces instability in transmission links due to high-dynamic network changes. To address these challenges, a multipath intelligent transmission strategy was proposed to achieve efficient data transmission. Firstly, a multipath intelligent routing algorithm based on reinforcement learning was introduced, which was not only proven to effectively enhance routing efficiency and reduce routing latency but also capable of dynamically segmenting data packets according to different transmission requirements to adapt to changes in the physical environment. To further improve the adaptability of the routing algorithm to dynamic environments, an environmental validation mechanism

收稿日期: 2025-01-14; 修回日期: 2025-03-10

通信作者: 冷甦鹏, spleng@uestc.edu.cn

基金项目: 国家自然科学基金青年项目 (No.62201122); 四川省科技计划项目 (No.2024YFHZ0321)

Foundation Items: The National Natural Science Foundation of China Youth Project (No.62201122), Sichuan Science and Technology Program (No.2024YFHZ0321)

was designed to evaluate the compatibility between current routing strategies and dynamic network, enabling self-adaptive adjustment of routing strategies during the transmission process. Simulation experiments demonstrate that, under various scenarios, the multipath intelligent routing strategy based on reinforcement learning (MIRSRL) effectively reduces end-to-end transmission delay while also increasing the probability of successful data recovery.

Key words: flying car, multipath intelligent routing algorithm, dynamic network, reinforcement learning, urban air mobility

0 引言

大力发展低空经济, 构建低空智联网, 是近年来广受关注的话题^[1-3]。飞行汽车作为低空经济的关键载体, 对于构建新型城市空中交通网络, 提升城际运输效率具有重要的作用。为了实现高效交通运输、监控空中交通变化等复杂任务, 飞行汽车需要通过组建低空智联网^[4-5], 实现低时延、高可靠、适应动态环境的路由策略, 从而完成高效的协作和通信。然而, 飞行汽车的高机动性和广泛的飞行范围导致车间链路不稳定。同时, 动态环境和复杂的任务需求导致飞行汽车的飞行策略不断调整, 网络拓扑高度动态变化, 使传统的固定数据传输路径不再适用。可见, 飞行汽车网络的动态特性容易导致传统无线路由协议不断发起路由发现过程, 路由表难以收敛等问题。

多径技术正逐渐成为提升无线网络可靠性和路由效率的关键策略^[6]。利用多径技术, 业务数据被分割成多个数据包, 以不同路由路径传输到目的地。该技术通过在多条路径上分散和复制相同的流量, 显著提高了恢复丢失数据的可能性。多径传输不仅提高了链路可靠性, 还有助于实现链路聚合和负载均衡。然而, 向各个下一跳转发的数据量决策成为关键难题。转发较少的数据包能降低总体时延, 但会增加信息丢失的概率; 传输过多的数据包则会造成飞行汽车负载的浪费, 却有助于提升数据恢复的概率。强化学习 (reinforcement learning, RL) 作为新兴的计算和模拟技术, 为解决这些问题提供了一种新方向^[7]。目前已有研究采用基于多智能体策略梯度的算法实现无人机路由规划^[8], 有

效降低了负载与时延。为提升网络吞吐量并最小化传输时间, 还有研究设计了一种新型多智能体强化学习算法, 在无人机轨迹、频谱分配与下一跳选择间实现权衡优化^[9]。然而, 目前基于强化学习的路由策略研究往往只关注策略性能的优化, 忽略了现行转发策略对未来突发变化的拓扑和路由需求不适应的问题。

为解决上述问题, 本文针对动态路由需求和拓扑变化提出了一种多径智能传输策略。本文的主要贡献总结如下。

(1) 本文提出了一种多径智能传输策略。与现有忽略动态环境影响, 采用固定多路径的方法不同, 本文的策略将原始数据分割成多个数据包, 为数据包动态分配传输路径, 提高数据成功恢复的概率, 并降低全局时延。

(2) 本文提出了一种基于多智能体深度确定性策略梯度 (multi-agent deep deterministic policy gradient, MADDPG) 的多径路由算法。区别于将原始数据分割成固定数据包的传统方法, 该算法基于飞行汽车轨迹的预测结果, 不仅以优化路由性能为目标实现了动态划分数据包的初始设置, 同时有效降低了端到端传输时延, 提升了路由可靠性。

(3) 本文方案在转发过程中使用最大后验概率 (maximum a posteriori probability, MAP) 方法来判断现行路由策略是否匹配现行环境, 提高了路由策略对动态网络的适应性。同时, 为了在较低信息上传频率的情况下有效监控路由状态、降低传输成本, 本策略灵活定义了飞行汽车的拓扑变化周期, 计算出了全网数据传输状态的信息上传周期。



1 智能数据转发系统和模型

随着飞行汽车在各个领域的应用日益广泛，低空智能网对高性能数据转发策略的需求日益凸显。针对这一问题，本文提出了一种智能数据转发系统，其架构如图1所示。该系统基于包含多个飞行汽车节点的网络场景，通过中心节点与普通节点的协同工作实现高效数据传输。具体而言，中心飞行汽车节点负责从若干普通飞行汽车节点收集地理位置等信息，并将数据分割成具有动态数据量的数据包。在此基础上，中心节点为每辆飞行汽车生成智能路由策略。普通飞行汽车节点则负责数据包转发，并定期向中心飞行汽车节点传输状态消息。为确保系统数据转发性能，本文开发了一种基于中心节点的轨迹预测算法。考虑拓扑结构的动态变化和突发性数据传输需求，初始设计的策略可能无法适应实时网络环境。为此，本文引入了一种环境检验机制，通过实时监测和评估网络状态来动态调整策略。基于上述设计，该系统框架可解耦为3个主要部分：路由策略、消息上传与轨迹预测、环境检验。

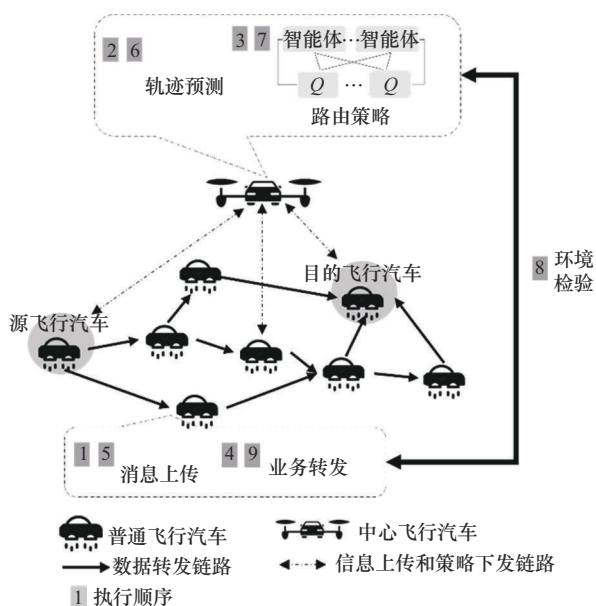


图1 智能数据转发系统架构

(1) 路由策略。在业务传输的初始阶段，以优化端到端业务性能为目标，动态计算数据包的数据量，同时为飞行汽车生成多径路由策略。在后续业务转发阶段，为每辆普通飞行汽车制订转发策略，帮助选择下一跳以及确定每个下一跳所要传输的数据。

(2) 消息上传与轨迹预测算法。普通飞行汽车节点向中心飞行汽车节点上传时空位置、一跳邻居表、所拥有的数据包等消息，为中心飞行汽车节点依据所上传的消息预测拓扑变化和生成路由策略提供帮助。

(3) 环境检验机制。中心飞行汽车节点根据目标性能和当前状态判断当前路由策略的有效性，并决定是否重新生成路由策略。

具体来说，在一个拥有若干飞行汽车的场景中，汽车组可以表示为集合 $\{N_1, \dots, N_i, \dots, N_n\}$ 。在该场景下，一组数据业务 F 在时刻 t^F 从源汽车 S 传输至目的汽车 D 。汽车 N_i 在时刻 t 的速度为 $v_i(t)$ ，所处的位置为 $I_i(t)$ 。可见，汽车 N_i 在时刻 t 的邻居集合 $U_i(t)$ 满足 $U_i(t) = \{N_j | \|I_j(t) - I_i(t)\| < r, \forall j\}$ ，其中 r 是汽车之间的通信半径， $\|\cdot\|$ 是计算向量二范数。规定该环境下丢包率为 ε 。每辆汽车的传输带宽被定义为 B 。假设带宽被均匀分配给每条传输链路，即带宽满足 $B_{i,j}(t) = B / |H_i(t)|$ 。其中， $B_{i,j}(t)$ 代表汽车 N_i 向汽车 N_j 进行数据转发时给该条链路分配的带宽， $H_i(t)$ 代表汽车 N_i 的下一跳汽车集合， $|\cdot|$ 代表计算集合元素个数。利用香农公式，2 辆汽车之间数据传输速率 $c_{i,j}(t)$ 可以表示为：

$$c_{i,j}(t) = \begin{cases} B_{i,j}(t) \log \left(1 + \frac{p_i(t) g_{i,j}(t)}{n_{i,j}} \right), & d_{i,j}(t) < r \\ 0, & d_{i,j}(t) \geq r \end{cases} \quad (1)$$

其中， $p_i(t)$ 、 $g_{i,j}(t)$ 和 $n_{i,j}$ 分别为汽车的发射功率、信道增益和噪声功率。 $d_{i,j}(t) = \|I_j(t) - I_i(t)\|$ 为时刻 t 时汽车 N_i 与 N_j 之间的距离。

定义每辆汽车的缓冲区队列容量为 R 。 $R_i^F(t)$ 为汽车 N_i 在时刻 t 所拥有的业务 F 的负载。那么，汽车 N_i 的总负载可以表示为：

$$R_i(t) = \sum_F R_i^F(t) \quad (2)$$

数据业务 F 在转发传输时，可以被分成 k 个数据包，每个数据包的数据量为 m 。在带宽均分的假设下，汽车向下一跳进行业务转发的时间与所转发的数据量成正比。因此，汽车 N_i 的转发时间策略矩阵被表示为：

$$\Phi_i^F(t) = [\phi_{i,j}^F(t)]_n \quad (3)$$

其中， $\phi_{i,j}^F(t)$ 代表汽车 N_i 在 t 时刻向其下一跳 N_j 进行数据转发的实际时间。

对应地，汽车 N_i 的转发数据包策略矩阵被表示为：

$$K_i^F(t) = [k_{i,j}^F(t)]_n \quad (4)$$

其中， $k_{i,j}^F(t)$ 代表汽车 N_i 在 t 时刻向其下一跳 N_j 转发的数据包个数。

于是，依据数据转发的关系，传输递归模型

$$\min T^F \text{ s.t. } \begin{cases} \phi_{i,j}^F(t) \in [0, \phi_{\max,i,j}^F(t)] \\ \phi_{\max,i,j}^F(t) = \min \{\tau_1, \tau_2\}, \tau_1 = \frac{R - R_j(t)}{c_{i,j}(t)}, \tau_2 = \min \{\tau - t | d_{i,j}(\tau) \geq R\}, \tau > t \\ H_i^F(t) \subset U_i^t, \forall i, t \end{cases} \quad (8)$$

其中， $\phi_{\max,i,j}^F(t)$ 代表转发时间的上限。根据优化结果可知，业务 F 数据包的初始化数据量为 $m^F = \inf (\phi_{i,j}^F(t) c_{i,j}^F(t))$ ，数据包数量为 $k^F = \lceil F/m^F \rceil$ 。在后续转发过程中，汽车 N_i 向下一跳 N_j 转发的数据量为 $k_{i,j}^F(t) = \lceil \phi_{i,j}^F(t) c_{i,j}^F(t) / m^F \rceil$ 。

可见，多径传输的优化模型以降低端到端时延为目标，实现了数据包初始化的动态划分，并通过多径传输的方式确保了路由转发的可靠性。然而，由于网络拓扑不断动态变化，优化目标难以收敛，因此，我们根据强化学习^[10-14]，设计基于 MADDPG 和轨迹预测算法的多径智能传输策略，寻找表达式 (8) 的可行解。

可以表示为：

$$\begin{cases} R_i^F(t+\phi) = R_i^F(t) + \sum_j R_{j,i}^F(t+\phi) - \sum_j R_{i,j}^F(t+\phi) \\ E(R_{j,i}^F(t+\phi)) = \phi_{i,j}^F(t) \cdot c_{i,j}^F(t) \cdot (1-\varepsilon) \end{cases} \quad (5)$$

其中， $E(\cdot)$ 代表数学期望， ε 为丢包率， ϕ 表示未来的任意时间， $R_{j,i}^F(t+\phi)$ 代表汽车 N_i 在 $t+\phi$ 时刻向其下一跳 N_j 所转发的数据量。

考虑到汽车 N_i 在业务传输过程中，可能进行多次转发，则 ϕ 满足以下条件：

$$\begin{aligned} \phi &\in [\phi_{i,k}^F(t), \phi_{\text{next},i,j}^F(t)] \\ \phi_{\text{next},i,j}^F(t) &= \min \{\tau - t | R_{i,j}^F(\tau) \neq 0, \tau \geq t\} \end{aligned} \quad (6)$$

其中， $\phi_{\text{next},i,j}^F(t)$ 为汽车 N_i 进行下一次数据转发的时间。数据业务 F 的端到端传输时延 T^F 为：

$$T^F = \min \{\tau - t^F | R_D^F(\tau) \geq F\} \quad (7)$$

于是，以优化端到端传输时延为目标，在满足路由转发条件的情况下，建立数据业务多径传输的优化模型如下。

2 多径智能传输策略

在本节中，介绍基于强化学习的多径智能传输策略。为便于清晰分析，将复杂的转发方案解耦为 3 个算法：多径智能路由算法、消息上传和轨迹预测算法、环境检验机制。

2.1 多径智能路由算法

多径智能路由算法的目标是在任何时刻为每辆飞行汽车求解路由策略。MADDPG 是典型的多智能体强化学习算法，它能够处理具有连续变量特征的场景。为了获得有效的路由策略，即求解式 (8)。本文在轨迹预测结果的基础上，提出一种基于 MADDPG^[13] 的多径智能路由算法。该算法的关键设计如下。



(1) 观测值。观测每次决策后飞行汽车的负载 $R_i(t)$ ，以及基于轨迹预测的位置序列。

(2) 动作。执行一次转发时间策略矩阵 $\Phi_i^F(t)$ 。

(3) 奖励。飞行汽车之间的最大传输时延奖励为 $A_i(t)$ 。飞行汽车负载的均方误差奖励为 $A_i(R)$ 。步进距离奖励为 $A_i(L)$ 。成功传输奖励为 $A_i(S)$ 。因此，普通飞行汽车的总奖励为：

$$A_i(L) = \frac{1}{|Z'(F)|} \sum_i \sum_j [\phi_{i,j}(t) c_{i,j}(t) \sum_m \frac{\eta^m}{d_{j,D}(t+m\tau)}]$$

$$A_i = \frac{1}{F} (-\alpha(t) A_i(t) + \alpha(L) A_i(L)) - \alpha(R) A_i(R) + \alpha(S) A_i(S) \quad (9)$$

其中， α 代表若干个非负权重， η 代表折扣因子。

在本算法中，飞行汽车智能体的贪婪策略为 $\Pi_i(A_i|O_i)$ 。动作价值函数是 $Q_i(O_i, A_i) = E(R|O_i, A_i)$ 。因此，需要以最小化损失 $\nabla_{\theta_i} J(\mu_i) = E_{s \sim \mathcal{D}} [\nabla_{\theta_i} \mu_i(a_i|s) \cdot \nabla_{a_i} Q_i^{\mu}(s, a_1, \dots, a_N)]$ 训练该算法演员网络，以最小化损失 $L(\theta_i) = E_{(s, a, r, s') \sim \mathcal{D}} [(Q_i^{\mu}(s, \text{LSTM}(s, h), a_1, \dots, a_N) - y_i)^2]$ 训练该算法评论家网络。其中， μ_i 代表演员网络的策略函数， Q_i 代表评论家网络的目标价值网络输出， s 代表所有智能体的联合状态， a_i 代表每个智能体的动作， h 代表长短期记忆 (long short-term memory, LSTM) 网络中的隐藏状态， y_i 代表评论家网络的目标函数，该函数与每一步动作的奖励相关。

$$\mathcal{T}_i = \min \left\{ \tau \left| \begin{array}{l} (|U_i(t) - U_i(t+\tau)|/|U_i(t)| = \lambda \text{ 或 } (|U_i(t+\tau) - U_i(t)|/|U_i(t)| = \theta) \\ 0 \leq \lambda \leq 1, 0 \leq \theta \leq 1 \end{array} \right. \right\} \quad (10)$$

其中， λ 表示飞离 N_i 通信半径内的汽车所占的百分比， θ 表示飞入 N_i 通信半径的汽车所占的比例。

因此，全局拓扑变化的周期 $\mathcal{T}$ 可以被定义为所有汽车局部拓扑变化周期的平均值，即 $\mathcal{T} = E(\mathcal{T}_i)$ 。然而，这仅为理论值，为了在实际中应用该值，需要得到信息上传周期的计算值 $\hat{\mathcal{T}}$ 。在一

在实际场景中，该算法将搭载在中心飞行汽车上，基于各个普通飞行汽车所上传的观测值模拟数据转发和飞行轨迹，实现计算和优化。本路由策略有效提高了路由可靠性并减少了传输延迟。然而，为了确保本算法可以有效实施，需要对轨迹预测算法和普通飞行汽车消息上传的周期进行设计计算。

2.2 消息上传和轨迹预测算法

为了生成一种高性能且能适应动态环境变化的联合传输方案，模拟动态网络的移动情况是很有必要的。因此，需要中心飞行汽车节点收集来自其他飞行汽车的移动信息，帮助实现全网未来轨迹预测。在本节中，计算了信息上传的周期，并提出了一种轨迹预测算法。

信息上传的周期应与拓扑变化的周期保持一致，因此，需要对网络中的拓扑变化程度进行定义并确认其周期性。为解决这一难题，讨论了飞行汽车局部拓扑变化和全局拓扑变化之间的关系。将汽车的速度视作随时间变化的函数，那么汽车 N_i 与汽车 N_j 在时刻 t 之间的距离为 $d_{i,j}(t) = \|I_j(t) - I_i(t)\|$ ，在时刻 $t+\tau$ 之间的距离为 $d_{i,j}(t) = \left\| I_j(t) - I_i(t) + \int_t^{t+\tau} (v_i(t) - v_j(t)) dt \right\|$ 。那么，依据汽车间距离与一跳邻居之间的数学关系，定义汽车 N_i 的局部拓扑关系的最小变化时间 $\mathcal{T}_i$ 为：

个汽车群中，所有汽车间的平均距离为 $E(d)$ ，平均速度为 $E(v)$ ，平均邻居数量为 $E(|U|)$ 。讨论一辆代表性的普通汽车 N_i ，即假设它的邻居数量和速度均为平均值，那么它所有的邻居 N_j 飞行情况可以分为3种类型。

(1) 汽车 N_j 飞入汽车 N_i 通信半径：汽车 N_i

和汽车 N_j 之间的距离满足 $[d_{i,j}(t) = \sqrt{d_x^2 + d_y^2} < r]$, $d_{i,j}(t + \hat{T}) = \sqrt{[d_x + (v_{i,x} - v_{k,x})\hat{T}]^2 + [d_y + (v_{i,y} - v_{k,y})\hat{T}]^2} \geq r$ 。此类汽车的极限飞行为在拓扑变化周期内飞入汽车 N_i 的一跳范围内, 则代表速度值为 $v_j^a(t) = (v_{i,x} - d_x/\hat{T}, v_{i,y} - d_y/\hat{T})$ 。依据大数定律, 根据频率估计此类汽车出现的概率为 $P^a = \theta E(U)/n$ 。

(2) 汽车 N_j 飞离汽车 N_i 通信半径: 汽车 N_i 和汽车 N_j 之间的距离满足 $d_{i,j}(t) \geq r$, $d_{i,j}(t + \hat{T}) < r$ 。考虑此类汽车的极限飞行, 以及实际场景中汽车群的速度存在范围 $[v_{\min}, v_{\max}]$, 则代表速度值为 $v_j^b(t) = (-v_{\max} v_{i,x} / \|v_{il}\|, -v_{\max} v_{i,y} / \|v_{il}\|)$ 。同时, 此类出现的概率为 $P^b = \lambda E(U)/n$ 。

(3) 汽车 N_j 与汽车 N_i 的相对邻居关系不变: 此类汽车的代表速度为 $v_j^c(t) = E(v)$, 出现的概率为 $P^c = [n - (\lambda + \theta)E(U)]/n$ 。

根据以上情况的典型概率和极限速度, 全概率公式 $E(v) = \sum_{i=a,b,c} P^i v_j^i(t)$, 可推导出信息上传的周期为:

$$\hat{T} = \theta E(d) / (\lambda E(v) + \lambda v_{\max}) \quad (11)$$

普通飞行汽车依据信息上传周期, 将实时位置和速度发送至中心飞行汽车。为了更好地分析端到端时延, 中心飞行汽车使用基于网络的轨迹预测算法^[15], 以最小化预测位置与实际位置之间的均方误差 (mean square error, MSE) 来为路由分析提供未来分析数据, 从而提高路由准确性。

2.3 环境检验机制

由于丢失数据包、环境剧烈变化等意外, 现有的策略可能无法与当下的环境条件相匹配。为了进一步提高转发策略对动态网络的适应性, 参考强化学习中的时间差分 (temporal difference, TD) 算法, 提出一种基于MAP准则的环境检验机制, 判断是否需要重新生成路由解决方案。

参考TD算法的思路, 定义上一次迭代中的多径智能路由算法所计算得到的目标端到端时延为

T^F , 可以将其视为理想估值。经过一段时间 T 后, 多径智能路由算法依据当下环境重新计算得到的目标端到端时延为 T^F , 则目前的实际时延可以表示为 $T^F = T + T^F$ 。于是, 本文定义了3种假设:

$$S(1): T^F = T^F$$

$$S(2): T^F = (1 + \eta) T^F$$

$$S(3): T^F = (1 - \eta) T^F \quad (12)$$

其中, η 表示偏差程度。假设 $S(1)$ 指当前的实际时延近似于策略最佳值, 这表明现行路由策略适合于当前的环境, 其他二者则代表不适合。考虑不确定的数据包丢失概率, 轨迹预测结果的不准确性等影响, 可以分别计算 $S(2)$ 和 $S(3)$ 的概率。于是, 根据贝叶斯定理, 判决条件可以被分为2类:

$$J_i: J_1 = \frac{P(S(1)|T^F)}{P(S(2)|T^F)} \geq 1 \text{ 和 } J_2 = \frac{P(S(1)|T^F)}{P(S(3)|T^F)} \geq 1$$

$$J_f: \bar{J}_1 = \frac{P(S(1)|T^F)}{P(S(2)|T^F)} < 1 \text{ 或 } \bar{J}_2 = \frac{P(S(1)|T^F)}{P(S(3)|T^F)} < 1 \quad (13)$$

其中, J_i 表示策略仍然有效, J_f 则表示相反情况。 J_1 和 J_2 分别代表决策中的关键变量。通过计算和化简, 可以得到:

$$J_1: T^F \in B_1, \rho_{12} = \sqrt{2(\sigma_1^2 - \sigma_2^2) \ln \frac{P(S(2))\sigma_1}{P(S(1))\sigma_2} + (b_1 + b_2)^2}$$

$$B_1 = [(b_2\sigma_1^2 - b_1\sigma_2^2 + \sigma_1\sigma_2\rho_{12})/(\sigma_1^2 - \sigma_2^2), (b_2\sigma_1^2 - b_1\sigma_2^2 - \sigma_1\sigma_2\rho_{12})/(\sigma_1^2 - \sigma_2^2)]$$

$$J_2: T^F \notin B_2, \rho_{13} = \sqrt{2(\sigma_1^2 - \sigma_3^2) \ln \frac{P(S(3))\sigma_1}{P(S(1))\sigma_3} + (b_1 + b_3)^2}$$

$$B_2 = [(b_3\sigma_1^2 - b_1\sigma_3^2 + \sigma_1\sigma_3\rho_{13})/(\sigma_1^2 - \sigma_3^2), (b_3\sigma_1^2 - b_1\sigma_3^2 - \sigma_1\sigma_3\rho_{13})/(\sigma_1^2 - \sigma_3^2)] \quad (14)$$

根据高斯分布的3- σ 原则, 式(15)中的变量分别取值 $b_1 = T^F$, $\sigma_1 = T^F/3$, $b_2 = (1 + \eta)T^F$,



$\sigma_2 = (1+\eta)T''^F/3$, $b_3 = (1-\eta)T''^F$, $\sigma_3 = (1-\eta)T''^F/3$ 。最终, 当前策略有效的判决 J_i 和当前策略无效的判决 J_f 可以分别表示为:

$$\begin{aligned} J_i: T^F \in B_1 \text{ 和 } T^F \notin B_2 \\ J_f: T^F \notin B_1 \text{ 或 } T^F \in B_2 \end{aligned} \quad (15)$$

3 仿真实验和分析

在本节中, 基于 Python 3.9 和 Anaconda 3 实现基于强化学习的多径智能传输策略 (multipath intelligent routing strategy based on reinforcement learning, MIRSRL) 仿真, 并基于仿真结果给出分析。

本次仿真实验中, 普通飞行汽车的初始位置是随机分布的, 并在每个时刻以平均速度附近值随机移动。中心飞行汽车始终位于普通汽车群的中心位置, 且具有与普通汽车群相同的平均速度。随机选择 2 辆普通飞行汽车作为源汽车和目标汽车。其他基本场景参数是: 通信半径为 30 m, 带宽为 2 kHz, 路由结束时钟为 50 s, 初始位置范围为 50 m×50 m。为了评估性能, 将 MIRSRL 与 3 种基准路由方案进行比较, 分别是洪泛 (Flood)^[16]、贪婪周边无状态路由 (greedy perimeter stateless routing, GPSR) 算法^[17]、增强型贪婪周边无状态路由 (enhanced greedy perimeter stateless routing, EGPSR) 算法。其中, EGPSR 算法是在传统 GPSR 算法基础上进行改进, 考虑动态网络中节点状态不稳定的特点, 在选择下一跳时优先选择距离目的节点较近、相对位置较为稳定、速度方向指向目的节点附近、负载较小的邻居节点。

为了评估业务流量对不同策略的平均端到端时延和路由成功率的影响, 本文建立一个包含 15 辆飞行汽车, 飞行汽车平均速度为 10 m/s, 且丢包率为 30% 的场景。业务数据量对策略效果的影响如图 2 所示。在高业务数据量需求下, 由于可靠的多径传输, MIRSRL 的表现明显优于

其他 3 种基准算法。即使在 1 600 bit 数据量的业务需求下, MIRSRL 也能达到路由成功率 85.5%, 平均时延 18.2 s, 路由可靠性和效率显著提升。

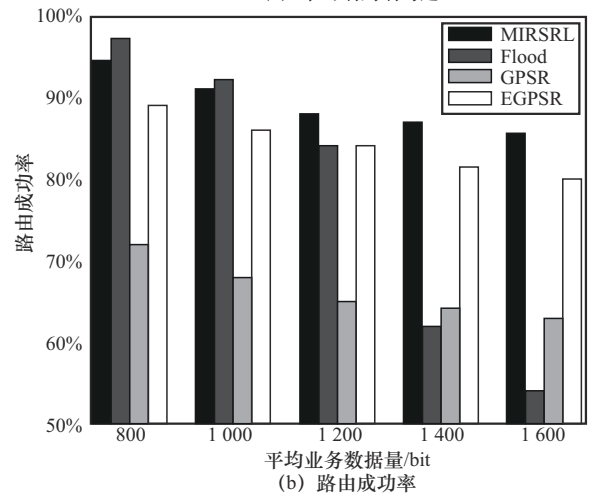
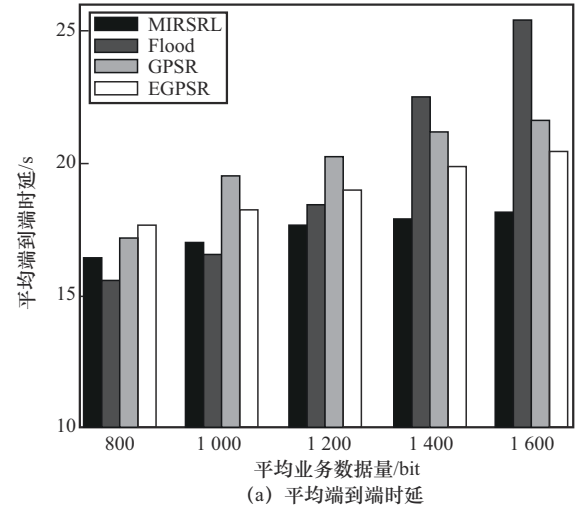


图2 业务数据量对策略效果的影响

为分析在不同丢包率下, 不同方案在平均端到端时延和路由成功率方面的表现, 在一个包含 15 辆飞行汽车、飞行汽车平均速度为 10 m/s、平均业务数据量为 1 000 bit 的场景中, 即丢包率对策略效果的影响如图 3 所示。其中, MIRSRL 的平均路由时延始终接近于洪泛算法, 路由成功率保持在 85% 以上。这得益于其适应动态环境的模型设计, MIRSRL 避免了洪泛广播风暴的风险。相比于 GPSR 和 EGPSR, MIRSRL 提高了路由成

功率, 降低了数据丢失的概率, 并提升了整体路由效率。

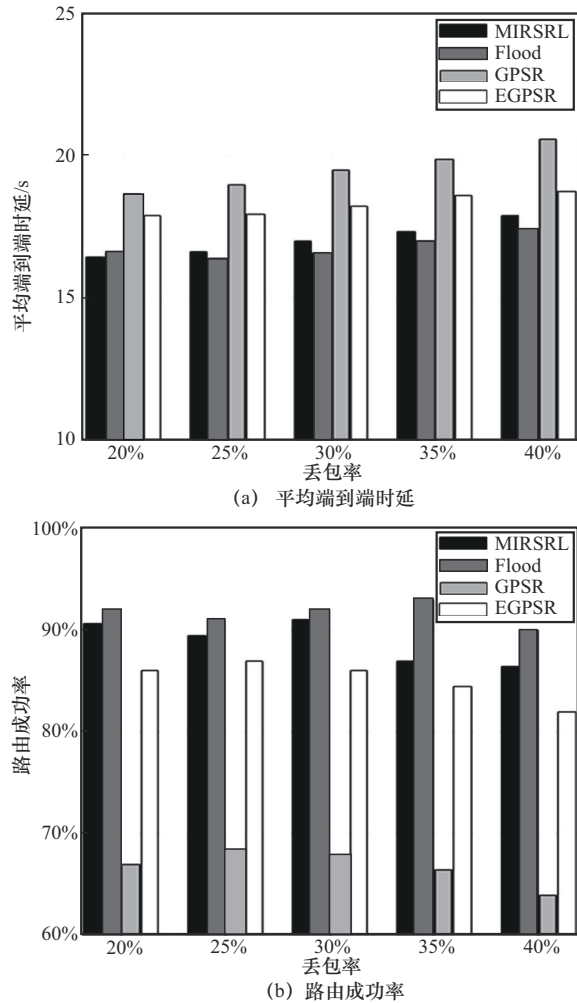


图3 丢包率对策略效果的影响

为探究在不同的平均速度下, 不同策略在平均端到端时延和路由成功率方面的表现, 在一个包含 15 辆飞行汽车、丢包率为 30%、平均业务数据量为 1 000 bit 的场景中, 平均速度对策略效果的影响如图 4 所示。其中, MIRSRL 在多种飞行汽车速度情况下表现均优于 GPSR 和 EGPSR。同时, MIRSRL 在有限冗余传输的情况下, 避免了洪泛广播风暴的风险, 达到了近似但略高于洪泛的低时延 17 s, 近似但略低于洪泛的高成功率 91%。

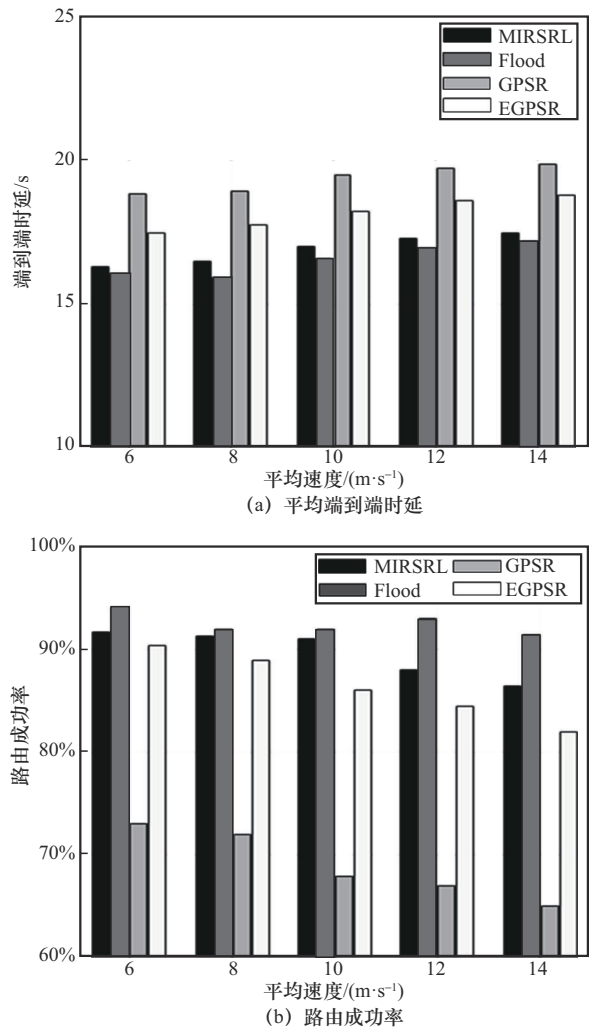


图4 平均速度对策略效果的影响

4 结束语

本文提出了一种面向高动态飞行汽车网络的多径智能传输策略。该策略能够动态调整路由转发, 动态划分数据包, 并实现动态寻路, 提高了路由可靠性并有效降低了传输时延。在未来的工作中, 希望考虑灵活数量的飞行汽车群, 增强模型在不同环境中的泛化能力。此外, 为了减轻中心飞行汽车的计算负担, 希望能够构建数字孪生分层架构, 提供本地和全局的转发策略计算平台, 提升即时决策和联合决策的综合性能。最后, 希望将复杂地形等实际环境因素对传输的影响引入模型设计中, 进一步提升路由策略的泛化性。



参考文献：

- [1] COHEN A P, SHAHEEN S A, FARRAR E M. Urban air mobility: history, ecosystem, market potential, and challenges[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(9): 6074-6087.
- [2] HUANG H L, SU J C, WANG F Y. The potential of low-altitude airspace: the future of urban air transportation[J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(8): 5250-5254.
- [3] WHITWORTH H, AL-RUBAYE S, TSOURDOS A. Urban air mobility link budget analysis in 5G communication systems[C]//Proceedings of the 2023 IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM). Piscataway: IEEE Press, 2023: 400-406.
- [4] XIONG K, LENG S P, CHEN L Y, et al. RIS-aided trajectory optimization in layered urban air mobility[J]. IEEE Internet of Things Journal, 2025, 12(2): 2188-2202.
- [5] XIONG K, YU H Q, LENG S P, et al. Multi-hop RIS-aided learning model sharing for urban air mobility[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(3): 3947-3959.
- [6] MACHUMILANE A, GOTTA A, CASSARÁ P, et al. Actor-critic scheduling for path-aware air-to-ground multipath multimedia delivery[C]//Proceedings of the 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring). Piscataway: IEEE Press, 2022: 1-5.
- [7] MAZYAVKINA N, SVIRIDOV S, IVANOV S, et al. Reinforcement learning for combinatorial optimization: a survey[J]. Computers & Operations Research, 2021, 134: 105400.
- [8] WANG Z L, YAO H P, MAI T L, et al. Cooperative reinforcement learning aided dynamic routing in UAV swarm networks[C]//Proceedings of the ICC 2022 - IEEE International Conference on Communications. Piscataway: IEEE Press, 2022: 1-6.
- [9] VERES M, MOUSSA M. Deep learning for intelligent transportation systems: a survey of emerging trends[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 21(8): 3152-3168.
- [10] ZHOU L Y, LENG S P, WANG Q, et al. Integrated sensing and communication in UAV swarms for cooperative multiple targets tracking[J]. IEEE Transactions on Mobile Computing, 2022, 22(11): 6526-6542.
- [11] LIU C Y, XU M W, YANG Y, et al. DRL-OR: deep reinforcement learning-based online routing for multi-type service requirements[C]//Proceedings of the IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2021: 1-10.
- [12] LI F, SONG X Y, CHEN H J, et al. Hierarchical routing for vehicular ad hoc networks via reinforcement learning[J]. IEEE Transactions on Vehicular Technology, 2021, 70(2): 1852-1865.
- [13] MAI T L, YAO H P, XIONG Z H, et al. Multi-agent actor-critic reinforcement learning based in-network load balance[C]//Proceedings of the GLOBECOM 2020 - 2020 IEEE Global Communications Conference. Piscataway: IEEE Press, 2020: 1-6.
- [14] LOWE R, WU Y, TAMAR A, et al. Multi-agent actor-critic for mixed cooperative-competitive environments[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 6382-6393.
- [15] YANG Z F, LIU D, MA L. Vehicle trajectory prediction based on LSTM network[C]//Proceedings of the 2022 International Conference on Artificial Intelligence and Computer Information Technology (AICIT). Piscataway: IEEE Press, 2022: 1-4.
- [16] MARINA M K, DAS S R. Ad hoc on-demand multipath distance vector routing[J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2002, 6(3): 92-93.
- [17] BUJARIA, PALAZZI C E, RONZANI D. A comparison of stateless position-based packet routing algorithms for FANETs[J]. IEEE Transactions on Mobile Computing, 2018, 17(11): 2468-2482.

[作者简介]



王煜婷 (2000-), 女, 电子科技大学信息与通信工程学院博士生, 主要研究方向为无线通信网络、无人机自组网、多智能体强化学习、数字孪生等。



冷甦鹏 (1973-), 男, 电子科技大学信息与通信工程学院教授、博士生导师, 主要研究方向为物联网、车联网、新一代宽带无线网络、无线自组织网络、智能交通信息系统的资源管理、介质访问控制、路由、组网与互联、智能算法理论及技术应用等。



熊凯 (1991-), 男, 电子科技大学信息与通信工程学院副研究员、在站博士后, 主要研究方向为无人机编队资源分配、移动边缘计算和机器学习。