

面向交互式视频业务的MEO-LEO分层协同卫星路由策略

王波¹, 卢锐¹, 万锦辉², 王吉平¹

(1. 桂林电子科技大学信息与通信学院, 广西 桂林 541004; 2. 广西科技信息网络中心, 广西 南宁 530012)

摘要: 针对低轨卫星网络中交互式视频业务的低时延、低抖动和会话连续性需求, 面向强时变拓扑与能量时空异构条件下路由决策复杂、链路切换频繁等问题, 提出MEO-LEO分层协同路由策略H-GMRL。采用GNN与Double-DQN进行多约束全局路径规划, 结合MARL与触发式平滑切换实现局部修复。仿真表明, 所提策略平均端到端时延最高降低12.1%, 时延抖动最高降低45.3%, 分组丢包率最高降低44.8%, 平均路径TTL最高提升25.9%, 可增强高负载场景下交互式视频业务的连续稳定承载能力。

关键词: 低轨卫星网络; 交互式视频; 分层协同路由; 强化学习

中图分类号: TN927.2

文献标志码: A

MEO-LEO hierarchical collaborative satellite routing strategy for interactive video services

Wang Bo¹, Lu Rui¹, Wan Jinhui², Wang Jiping¹

1. School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China

2. Guangxi Science and Technology Information Network Center, Nanning 530012, China

Abstract: To meet the requirements of interactive video services in low Earth orbit satellite networks for low latency, low jitter, and session continuity, a hierarchical collaborative routing strategy, termed H-GMRL, was proposed for medium Earth orbit (MEO)-low Earth orbit (LEO) satellite networks to address complex routing decisions and frequent link switching under highly dynamic topologies and spatiotemporally heterogeneous energy conditions. Graph neural networks (GNNs) and Double Deep Q-Network (Double-DQN) were adopted for multi-constraint global path planning, while multi-agent reinforcement learning (MARL) together with a trigger-based smooth switching mechanism was used for local repair. Simulation results show that the proposed strategy reduces the average end-to-end delay by up to 12.1%, decreases delay jitter by up to 45.3%, lowers the packet loss rate by up to 44.8%, and increases the average path time to live (TTL) by up to 25.9%. It provides more stable and continuous support for interactive video services under high-load scenarios.

Key words: Low earth orbit satellite networks, interactive video, hierarchical routing, reinforcement learning

0 引言

面向6G空天地一体化的泛在连接愿景, 低轨卫星网络(Low Earth Orbit Satellite Networks, LSNs)凭借覆盖范围广、传播时延低以及组网灵活等优

势, 被普遍认为是6G能力增强的重要承载形态之一^[1]。与此同时, 随着远程协作、在线会议和实时交互等应用场景的持续发展, 端到端视频通话、远程会议等交互式视频业务呈现快速增长趋势。与传

收稿日期: XXXX-XX-XX; 修回日期: XXXX-XX-XX

通信作者: 万锦辉, 邮箱: wlzx@kjt.gxzf.gov.cn

基金项目: 面向6G的LEO卫星时空能量预测与绿色调度优化研究(No. C25GAH30)

Foundation Items: Research on Space - Time Energy Prediction and Green Scheduling Optimization of LEO Satellites for 6G (No. C25GAH30)

统弹性业务不同,交互式视频不仅对端到端时延和丢包率较为敏感,而且对业务过程中的会话连续性和体验稳定性提出了更高要求^[2]。

然而,交互式视频业务在LSNs中的高质量承载并不能仅依赖接入覆盖能力的扩展来实现,其关键仍在于网络层路由机制能否适应卫星网络的动态特性并满足业务连续性需求。与地面网络相比,低轨卫星网络具有星座规模大、节点状态异构、链路可见性快速变化以及能量补给受阳照/地影交替显著影响等特点,使得面向交互式视频业务的路由决策面临更为复杂的约束条件。

与传统业务相比,交互式视频业务对端到端时延、时延抖动、分组丢失以及会话连续性更为敏感,其在LSNs中的高质量承载不仅依赖接入覆盖能力,更取决于网络层路由机制对动态网络环境的适应能力。此外,由于LSNs具有星座规模大、节点状态异构、链路可见性快速变化以及能量供给受阳照/地影交替显著影响等特点,面向交互式视频业务的路由决策需要同时应对两方面挑战:一方面,卫星节点状态、链路条件、网络负载与能量水平共同构成高维动态状态空间,路由选择需在多约束耦合条件下实现在线优化,对算法的实时性、可扩展性与决策效率提出了较高要求;另一方面,卫星高速运动导致网络拓扑强时变,星间链路频繁切换甚至发生可见性中断,容易引发路径失稳、时延抖动累积和会话连续性下降。由此可见,仅依赖单一目标或瞬时状态驱动的传统路由方法,已难以满足交互式视频业务在动态卫星网络中的连续稳定承载需求。

目前,针对大规模异构条件下的多约束路由决策,已有研究主要从综合代价构建、在线决策优化和图学习增强等多方面展开。一类研究将时延、拥塞、链路质量与能量等因素统一纳入路由代价函数或多约束优化框架,以在网络性能与能量利用之间实现权衡^[3-5];另一类研究面向大规模强时变拓扑场景,结合网络拓扑演化和节点状态构建在线路由决策机制。例如Xu等^[6]利用空间位置信息实现全分布式动态路由。近年来,随着图表示学习与强化学习方法的发展,也有研究尝试使用图神经网络(Graph Neural Network, GNN)对星间拓扑与节点状态进行结构化编码,并在此基础上结合强化学习学习动态路由策略,以提升复杂场景下的决策能

力^[7-9]。这些工作为LSNs中的多约束智能路由提供了有益参考。

针对强时变拓扑下的会话连续性与切换稳定性,现有研究主要从业务自适应控制和拓扑平滑设计两条思路展开。Zhao等^[10]通过联合优化切换策略、自适应码率控制以及视频传输过程中的拥塞调节机制,缓解链路切换对业务体验带来的不利影响。Li等^[11]通过构造动态离散拓扑或虚拟拓扑,削弱卫星网络拓扑剧烈演化对路径选择的直接冲击,以降低路由震荡并增强端到端路径稳定性。这类研究在提升业务传输平稳性、减弱切换扰动等方面取得了一定成效,也为交互式视频在时变卫星网络中的承载提供了有价值的思路。

综合上述研究可以看出,现有工作已分别围绕多约束路由优化与拓扑动态适应展开了大量探索,但对于交互式视频业务的长时连续稳定承载而言,仍存在以下不足:其一,对阳照/地影交替引起的能量时空异构特征刻画不足,难以兼顾短期传输性能与长期能量可持续性的协同优化;其二,缺乏对路径潜在失效风险和持续可用性的前瞻性建模,不利于长时会话业务的稳定保持;其三,全局路径规划与局部切换修复缺少统一协同设计,难以同时兼顾路径前瞻性、局部响应性与切换平滑性。

为满足交互式视频长时会话连续承载需求,本文构建了全局前瞻规划、局部受约束修复与平滑切换协同联动的统一路由机制,提出MEO-LEO分层协同路由策略H-GMRL。与单层GNN+DRL路由方法相比,本文通过中地球轨道(Medium Earth Orbit, MEO)层与低地球轨道(Low Earth Orbit, LEO)层协同,将长期全局路径规划与局部修复解耦,缓解了大规模动态拓扑下状态空间膨胀和决策非平稳性;与纯分布式多智能体强化学习(Multi-Agent Reinforcement Learning, MARL)路由方法相比,本文通过全局路径约束传递和局部受限动作选择,抑制局部贪婪调整导致的路径震荡;与面向动态拓扑设计的路由方法相比,本文进一步将链路寿命(Time To Live, TTL)、能量时空异构、拥塞状态以及平滑切

换机制纳入统一协同框架,更强调交互式视频业务连续性导向下的全局与局部联合优化。因此,本文的实质贡献在于提出一种面向交互式视频连续承载的分层协同路由机制,并非对已有方法的简单工程组合。本文的主要贡献如下:

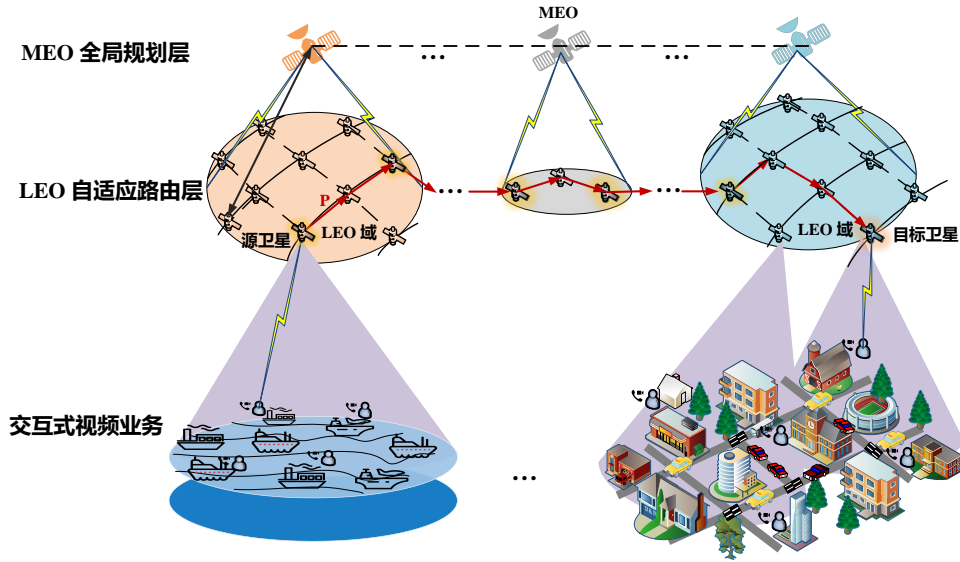


图1 分层网络架构图

(1) 提出一种面向交互式视频业务的分层协同路由框架, 实现MEO层全局规划与LEO层局部修复的协同联动。

(2) 构建面向能量、链路TTL与拥塞状态的联合多约束全局规划机制, 实现复杂动态环境下的全局路径决策。

(3) 设计触发式平滑切换机制, 通过受约束的局部修复与渐进式业务迁移减弱切换扰动, 提升交互式视频会话连续承载能力。

1 系统模型与问题建模

1.1 系统模型

本文考虑一种面向交互式视频业务的分层架构, 如图1所示。整个网络由MEO全局规划层和LEO自适应路由层构成: 其中, MEO层基于较强的全局状态感知能力执行跨域路径规划与控制, 并向LEO层下发路由决策及约束信息; LEO层则面向局部拓扑快速变化场景执行在线转发与自适应修复。交互式视频业务在LEO星间网络中沿端到端路径进行传输, 记从源卫星到目的卫星的跨域路径为 P 。

对于交互式视频业务, 仅优化端到端时延尚不足以保证会话体验, 还需进一步关注由路径中断、频繁切换和拥塞波动引起的时延抖动、分组丢失及会话连续性下降等问题。因此, 本节从通信特性、能量消耗、路径持续性以及会话连续性相关度量四个方面构建系统模型, 并在此基础上给出面向交互

式视频业务的多目标路由问题建模。

1.2 通信模型

定义相邻卫星 i 与卫星 j 之间链路为 (ij) , 在 t 时刻链路 (ij) 的瞬时链路容量为

$$C_{ij}(t) = B \log_2 \left(1 + \frac{P_i G_t G_r}{N_0 B L_{ij}(t)} \right) \quad (1)$$

其中 B 为系统带宽, P_i 为发射功率, G_t 、 G_r 分别为发射和接收增益, N_0 为噪声功率谱密度, $L_{ij}(t)$ 为路径损耗。该容量模型用于刻画链路容量随信噪比变化的基本趋势, 忽略调制与编码等物理层细节, 适用于网络层路由决策建模。

设沿路径 P 转发的第 k 个数据包 (长度为 M bit), 则该数据包在链路 (ij) 的时延为^[12]

$$d_{ij}^{(k)}(t) = [d_{ij}^{\text{prop}}(t) + d_{ij}^{\text{tx}}(t) + d_{ij}^{\text{queue}}(t)] \quad (2)$$

其中, 传播时延 $d_{ij}^{\text{prop}}(t) = r_{ij}(t)/c$, $r_{ij}(t)$ 为欧几里得距离, c 为光速; 传输时延 $d_{ij}^{\text{tx}}(t) = M/C_{ij}(t)$; $d_{ij}^{\text{queue}}(t)$ 为排队时延。

设节点 i 在时刻 t 的发送队列长度为 $Q_i(t)$ (单位: packet), 则链路 (ij) 的服务速率为 $\mu_{ij}(t) = C_{ij}(t)/M$, 相应的排队时延 $d_{ij}^{\text{queue}}(t)$ 估算为

$$d_{ij}^{\text{queue}}(t) = \frac{Q_i(t)}{\mu_{ij}(t)} = \frac{Q_i(t)M}{C_{ij}(t)} \quad (3)$$

因此, 数据包 k 沿路径 P 转发的端到端时延为

$$D_P(t) = \sum_{(ij) \in P} d_{ij}^{(k)}(t) \quad (4)$$

1.3 能耗模型

在 LSNs 中, 卫星转发业务时不仅消耗链路建立与维护能量, 还需持续消耗发送与接收过程中的通信能量。设数据包 k 在 t 时刻由链路 (ij) 转发, 则其通信能耗定义为

$$E_{ij}^{\text{com}}(t) = E_{ij}^{\text{set}}(t) + [P_i^{\text{tx}}(t) + P_j^{\text{rx}}(t)] \cdot d_{ij}^{\text{tx}}(t) \quad (5)$$

其中, $E_{ij}^{\text{set}}(t)$ 为链路 (ij) 建立或切换开销, 仅在新建链路或发生切换时计入; $P_i^{\text{tx}}(t)$ 和 $P_j^{\text{rx}}(t)$ 分别表示卫星 i 的发送功率与卫星 j 的接收功率。

考虑到卫星在轨运行过程中会受到阳照/地影交替影响, 节点可用能量状态具有明显的时空异构性。为反映不同能量状态下链路转发对节点资源的占用程度, 进一步引入卫星 i 在时刻 t 的动态能量权重系数:

$$\lambda_i(t) = f[\text{SoC}_i(t)] \times [1 + \eta_{\text{pen}}(t)], \quad \eta_{\text{pen}}(t) \in [0, 1] \quad (6)$$

其中, $\text{SoC}_i(t)$ 为卫星电池荷电状态, $f(\cdot)$ 为关于 $\text{SoC}_i(t)$ 的单调函数, 用于刻画电量紧张程度。本文取

$$f(\text{SoC}_i(t)) = 1 + \rho \cdot \frac{\text{SoC}_{\max} - \text{SoC}_i(t)}{\text{SoC}_{\max} - \text{SoC}_{\min}} \quad (7)$$

其中, $\rho > 0$ 为能量敏感系数, SoC_{\max} 与 SoC_{\min} 分别表示允许的最大和最小荷电状态阈值。当卫星 i 剩余电量降低时, $f(\text{SoC}_i(t))$ 单调增大, 表示低电量节点参与转发时应承担更高的能量代价。

$\eta_{\text{pen}}(t)$ 为地影能耗影响系数, 用于表征卫星处于地影环境时能量约束的增强效应, 定义为

$$\eta_{\text{pen},i}(t) = \begin{cases} 0, & \text{卫星 } i \text{ 处于阳照区} \\ \eta_e, & \text{卫星 } i \text{ 处于地影区} \end{cases}, \quad 0 < \eta_e \leq 1 \quad (8)$$

其中, η_e 为地影惩罚系数。当卫星 i 处于阳照区时, 能量可获得持续补给, 故不额外增加能耗惩罚; 当卫星 i 处于地影区时, 由于太阳能输入中断, 单位通信能耗对剩余能量的影响更为显著, 因此取 $\eta_{\text{pen},i}(t) = \eta_e$ 。结合 SoC 安全工作区间 $(0.2, 0.8)$ 及阳照/地影工况差异, 本文取 $\rho = 1$, $\eta_e = 0.5$ 。

因此, 路径能耗不仅与链路建立、发送和接收过程本身有关, 还受到节点实时荷电状态以及阳照/地影环境的共同影响, 数据包 k 沿路径 P 转发的加权能耗可表示为

$$E_P(t) = \sum_{(ij) \in P} \lambda_i(t) \cdot E_{ij}^{\text{com}}(t) \quad (9)$$

1.4 链路与路径 TTL 模型

对于交互式视频业务而言, 仅优化瞬时传输时延和能耗仍不足以保证会话连续性。由于 LSNs 拓扑随卫星运动持续演化, 链路可见窗口不断变化; 同时, 节点剩余能量的衰减也会影响链路持续转发能力。因此, 需要进一步刻画链路和路径在未来一段时间内的持续可用性。定义相邻卫星 i 与卫星 j 之间链路 (ij) 在 t 时刻的 TTL 为

$$T_{ij}(t) = \min \{ T_{ij}^{\text{vis}}(t), T_{ij}^{\text{enc}}(t) \} \quad (10)$$

其中, $T_{ij}^{\text{vis}}(t)$ 为链路窗口剩余时长, 用于反映由轨道运动和星间几何关系决定的链路可达性约束; $T_{ij}^{\text{enc}}(t)$ 为能量约束下的可持续转发剩余时长, 用于刻画由能量约束引起的链路潜在失效风险。

由于端到端路径的持续可用性通常受最脆弱链路制约, 因此将路径 P 的 TTL 定义为路径上各链路 TTL 的最小值, 即

$$T_P(t) = \min_{(ij) \in P} T_{ij}(t) \quad (11)$$

对于长时交互式视频会话而言, 较大的 $T_P(t)$ 意味着当前路径在未来一段时间内更可能保持可用, 从而有助于降低频繁切换和会话中断风险。

1.5 会话连续性相关度量

对于交互式视频业务而言, 仅从端到端时延、路径能耗和路径 TTL 三个维度进行建模, 仍不足以完整刻画交互式会话过程中由频繁切换引发的传输波动、分组丢失与路径扰动。因此, 除上述通信、能耗与 TTL 相关模型外, 本文进一步引入端到端时延抖动、丢包率和路径变化幅度三个会话连续性相关度量, 用于从业务平稳性、传输可靠性和路径重构扰动三个方面补充描述交互式视频业务在动态卫星网络中的承载特征。

设会话期间成功接收的数据包数为 N_{rx} , 第 k 个成功接收数据包端到端时延为 D_k , 则端到端时延抖动定义为两个相邻成功接收数据包端到端时延差的平均绝对值, 即

$$J = \frac{1}{N_{rx} - 1} \sum_{k=1}^{N_{rx} - 1} |D_{k+1} - D_k| \quad (12)$$

进一步地, 设会话期间发送的总数据包数为 N_{tx} , 成功接收的数据包数为 N_{rx} , 则会话期间的丢包率定义为

$$PLR = \frac{N_{tx} - N_{rx}}{N_{tx}} = 1 - \frac{N_{rx}}{N_{tx}} \quad (13)$$

该指标用于衡量网络在链路失效、拥塞累积及路径切换场景下的可靠传输能力,其值越小,说明业务传输可靠性越高。

此外,为衡量路由切换前后路径重构对业务传输造成的扰动强弱,设会话期间共发生 N_{sw} 次路径切换,第 r 次切换前后的路径分别记为 P_r^{old} 和 P_r^{new} ,其对应的路径边集合分别记为 $\mathcal{L}(P_r^{old})$ 和 $\mathcal{L}(P_r^{new})$,则路径变化幅度定义为

$$A = \frac{1}{N_{sw}} \sum_{r=1}^{N_{sw}} \frac{|\mathcal{L}(P_r^{old}) \Delta \mathcal{L}(P_r^{new})|}{|\mathcal{L}(P_r^{old}) \cup \mathcal{L}(P_r^{new})|} \quad (14)$$

其中, Δ 表示对称差运算。该指标越小,说明每次路径调整越局部、切换越平滑,对交互式视频业务连续传输的扰动越弱。

上述三个度量分别从时延波动、传输可靠性和平滑切换扰动三个方面刻画交互式视频业务在动态网络中的传输稳定性,为后续实验评价与结果分析提供统一依据。

1.6 问题建模

交互式视频业务在 LSNs 中的路由选择需要同时兼顾传输实时性、能量可持续性和路径连续性。然而,端到端时延 $D_p(t)$ 、路径能耗 $E_p(t)$ 和路径寿命 $T_p(t)$ 在物理含义与数值尺度上存在显著差异,若直接进行线性组合,容易导致某一指标因量纲或取值范围不同而主导优化结果。为此,本文首先将三项指标归一化至 $[0,1]$ 区间,分别记为 $\tilde{D}_p(t)$ 、 $\tilde{E}_p(t)$ 与 $\tilde{T}_p(t)$,如下所示:

$$\begin{aligned} \tilde{D}_p(t) &= \frac{D_p(t) - D_{\min}}{D_{\max} - D_{\min}} \\ \tilde{E}_p(t) &= \frac{E_p(t) - E_{\min}}{E_{\max} - E_{\min}} \\ \tilde{T}_p(t) &= \frac{T_p(t) - T_{\min}}{T_{\max} - T_{\min}} \end{aligned} \quad (15)$$

其中, $D_{\min/\max}$ 、 $E_{\min/\max}$ 、 $T_{\min/\max}$ 分别为对应指标在系统运行区间内的上下界,由网络运行状态和任务场景共同确定。

此外,面向交互式视频业务,考虑到主观体验质量(QoE)难以直接嵌入网络层路由决策,而交互式视频中的帧级截止时间、关键帧丢失和码率自适应等特征通常会在网络层体现为时延、抖动、丢包、路径扰动和路径持续性变化,故 LSNs 的路由选择可进一步建模为一个以交互响应实时性、资源

可持续性和路径连续性为核心目标的多目标优化问题。因此,本文以端到端时延、路径能耗和路径 TTL 作为主优化变量,记端到端路径为 P ,则问题可表示为

$$\begin{aligned} \min_P & [\alpha \cdot \tilde{D}_p(t) + \beta \cdot \tilde{E}_p(t) - \gamma \cdot \tilde{T}_p(t)] \\ \text{s.t.} & \quad C1: D_p(t) \leq D_{\max} \\ & \quad C2: T_p(t) \geq T_{\min} \\ & \quad C3: E_p(t) \leq E_{\max} \\ & \quad C4: H_p \leq H_{\max} \\ & \quad C5: \alpha + \beta + \gamma = 1, \alpha, \beta, \gamma > 0 \end{aligned} \quad (16)$$

其中, H_p 为路径 P 的中继跳数, H_{\max} 为允许的最大跳数; α, β, γ 为无量纲化权重系数,用于调节时延、能耗与路径 TTL 三者在整个优化目标中的相对重要性。约束 C1 - C4 分别对应端到端时延、路径 TTL、路径能耗以及跳数等可行性限制, C5 用于保证权重归一化。

需要指出的是,1.5 小节中的时延抖动、丢包率和路径变化幅度不直接纳入主优化目标,而作为会话连续性评价指标,用于补充刻画传输平稳性、可靠性和切换扰动控制效果。由于该问题同时具有路径组合选择、多指标耦合和动态约束演化等特征,因此本质上属于动态多约束组合优化问题。对于大规模、强时变的低轨卫星网络,传统精确优化方法难以满足在线决策时效性要求。基于此,本文采用学习驱动的分层协同路由策略对其进行近似求解,以兼顾决策效率与业务承载性能。

2 MEO-LEO 分层协同路由策略 H-GMRL

2.1 分层协同决策过程

针对低轨卫星网络中大规模异构星座条件下全局路由决策复杂、以及强时变拓扑环境下交互式视频会话连续性难以保障的问题,本文提出一种 MEO-LEO 分层协同路由策略 H-GMRL。H-GMRL 路由决策流程图如图 2 所示。

其中, MEO 层面向相对稳定的全局网络视图,在控制周期尺度上执行跨域路径规划,并向 LEO 层下发全局路径及相应约束信息; LEO 层则面向局部拓扑快速变化场景,在全局规划路径的基础上执行逐跳转发、局部修复和平滑切换。

此外,考虑到单颗 MEO 卫星对 LEO 层的瞬时可见范围有限,本文采用全局感知机制。具体而言,LEO 层的几何位置、链路可见性及链路窗口等

可预测信息主要依据卫星星历与轨道参数获得；在此基础上，各 LEO 卫星仅周期性向当前可见的服务 MEO 卫星上报队列负载、剩余能量、链路 TTL 统计量和容量统计量等低维动态状态摘要。随后，各 MEO 控制卫星通过控制层链路交换区域级状态信息，从而在每个控制周期内形成面向 LEO 业务层的近实时全局状态表征，并据此完成跨域路径规划与控制信息下发。

H-GMRL 分层协同路由策略的伪代码如表 1 所示：

2.2 基于 GNN-Double DQN 的全局路径规划

考虑到 LSNs 拓扑规模大、链路连接关系动态变化且节点能量状态具有显著异构性，本文在 MEO 层采用 GNN 对动态星间拓扑进行结构化表征。记时刻 t 的动态拓扑图为 $G_t = (\mathcal{V}, \mathcal{E}_t)$ ，其中 \mathcal{V} 表示卫星节点集合， \mathcal{E}_t 表示时刻 t 的星间链路集合。对于卫星节点 i ，其输入特征向量定义为

$$x_i(t) = [E_i(t), \bar{T}_i(t), Q_i(t), \bar{B}_i(t)] \quad (17)$$

其中， $E_i(t)$ 表示节点剩余能量， $\bar{T}_i(t)$ 表示与节点相关的平均链路 TTL 统计量， $Q_i(t)$ 表示节点队列负载， $\bar{B}_i(t)$ 表示平均链路容量统计量。为兼顾大规模拓扑下的表示能力与计算复杂度，本文采用 GraphSAGE 作为图编码模型，设置 2 层消息传递结构，隐藏维度为 64。经邻域聚合后得到节点嵌入 $h_i(t)$ ，再通过池化操作得到全局状态表示 s_t^G ，用于后续的全局路径决策。

为避免在全图上直接枚举所有端到端路径导致动作空间爆炸，MEO 首先基于当前时刻的 LEO 层拓扑，对源接入 LEO 与目的接入 LEO 之间执行 $K = 15$ 候选路径搜索，得到候选路径集合 \mathcal{P}_t 。随后，对每条候选路径分别计算其端到端时延、路径 TTL、路径能耗与跳数，并剔除违反式 (16) 中约束 $C1 \sim C4$ 的不可行动作，最终得到全局可行动作集为

$$A_t^G \triangleq \left\{ P \in \mathcal{P}_t \mid D_P(t) \leq D_{\max}, T_P(t) \geq T_{\min}, E_P(t) \leq E_{\max}, H_P \leq H_{\max} \right\} \quad (18)$$

表 1

H-GMRL 分层协同路由策略伪代码

输入： 动态拓扑图 $G_t = (\mathcal{V}, \mathcal{E}_t)$ ；控制周期 T_c ；触发阈值集合 $\{\theta_T, \theta_q, \theta_E\}$ ；平滑切换窗口 ΔT ；已训练网络：GNN 编码器 ϕ 、MEO 层 Double-DQN 网络 Q_ψ 、LEO 层 MARL 局部策略网络 Q_ϕ 。

输出： 全局路径 P_t 及逐跳转发决策。

MEO 层：全局路径规划

- 1: 根据当前拓扑 G_t 与节点特征 $\{x_i(t)\}$ ，利用 GNN 编码得到节点嵌入 $\{h_i(t)\}$ ，并生成全局状态表示 s_t^G 。
- 2: 基于当前拓扑与资源状态构造候选路径集合，并筛除不满足时延、路径 TTL、能量与跳数约束的非法动作，形成全局可行动作集 A_t^G 。
- 3: 在 A_t^G 上采用 ϵ -greedy 策略，根据 $Q_\psi(s_t^G, a)$ 选择全局动作 a_t^G 。
- 4: 将动作 a_t^G 解码为端到端路径 P_t ，并将 $\{P_t, \theta_T, \Delta T\}$ 下发至 LEO 层。

LEO 层：局部修复与平滑切换

- 5: 当前转发节点 i 观测局部状态 $s_i(t)$ 及邻居集合 $\mathcal{N}_i(t)$ 。
- 6: 若满足触发条件（链路 TTL 低于阈值、下一跳队列超过阈值或剩余能量低于安全阈值），则启动局部修复。
- 7: 对邻居集合 $\mathcal{N}_i(t)$ 进行约束筛选，构造局部可行邻居集合 $\mathcal{N}_i^{feas}(t)$ 。
- 8: 若 $\mathcal{N}_i^{feas}(t) \neq \emptyset$ ，则选择新的下一跳 $n^* = \arg \max_{j \in \mathcal{N}_i^{feas}(t)} Q_\phi(s_i(t), j)$ 。
- 9: 在时刻 t_0 触发切换，并计算平滑切换权重 $C(t) = \text{clip}[\frac{(t - t_0)}{\Delta T}, 0, 1]$ 。
- 10: 在平滑窗口 $[t_0, t_0 + \Delta T]$ 内，按比例 $C(t)$ 将业务逐步迁移至新下一跳 n^* ，同时按比例 $1 - C(t)$ 保持旧下一跳转发。
- 11: 当 $t \geq t_0 + \Delta T$ ，或旧链路 TTL 耗尽、旧队列清空时，停止旧路径转发，完成局部切换。
- 12: 若 $\mathcal{N}_i^{feas}(t) = \emptyset$ 或在有界窗口内转发无推进，则保持保守转发并向 MEO 层请求重规划。
- 13: 若未触发局部修复条件，则继续沿当前路径 P_t 转发。
- 14: 若距离上一次全局规划时间超过 T_c ，则返回步骤 1，由 MEO 层重新执行全局路径规划。

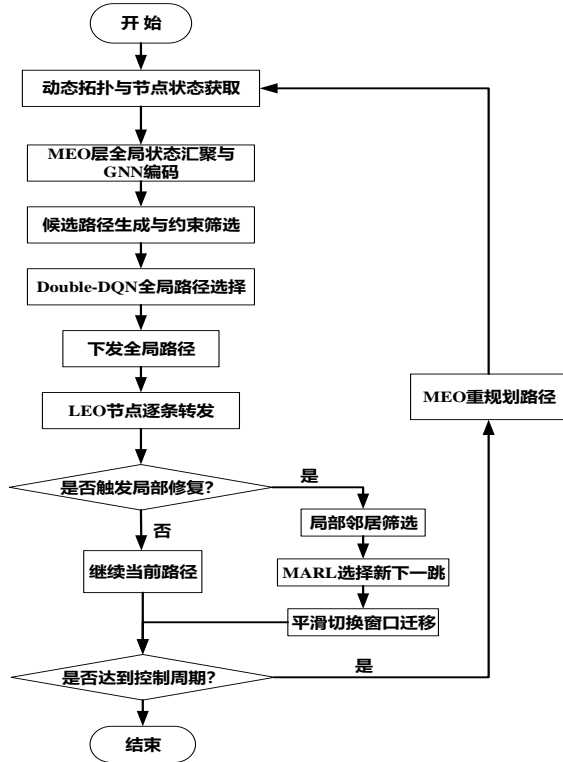


图2 H-GMRL路由决策流程图

其中 D_{\max} 、 T_{\min} 、 E_{\max} 、 H_{\max} 分别对应式(16)中的时延约束、路径TTL约束、能耗约束与跳数约束。通过在动作选择前进行约束筛选,可从机制上保证全局规划结果满足基本可行性要求,避免强化学习在大量非法动作上进行无效探索,从而提高训练稳定性与在线决策可行性。在动作选择阶段,本文采用双深度Q网络(Double Deep Q-Networks, Double-DQN)在 A_t^G 上执行 ϵ -greedy 选择,其中 ϵ 从 1.0 逐步衰减至 0.1,折扣因子和学习率分别设为 0.95 和 0.001。为与式(16)的问题建模保持一致,定义MEO层全局规划奖励为

$$r_t^G = -\alpha \tilde{D}_t - \beta \tilde{E}_t + \gamma \tilde{T}_t \quad (19)$$

其中, \tilde{D}_t 、 \tilde{E}_t 和 \tilde{T}_t 分别表示当前候选路径在端到端时延、路径能耗和路径TTL维度上的归一化评价价值, α 、 β 和 γ 为对应的权重系数。

综上, MEO层通过 GraphSAGE 对动态图结构与节点异构状态进行统一编码,再结合候选路径筛选与 Double-DQN 决策,实现面向交互式视频业务的全局路径规划。与直接在全图路径空间中进行搜索或学习相比,该机制能够在控制动作空间规模的同时增强全局规划的前瞻性与稳定性,更适用于大规模、强时变低轨卫星网络中的在线路由决策。

2.3 基于MARL的平滑切换机制

尽管MEO层能够在控制周期尺度上生成全局可行路径,但由于LEO层拓扑变化更快,在全局规划周期内仍可能出现链路TTL快速衰减、局部拥塞加剧或节点资源紧张等情况,从而导致当前路径难以持续稳定承载交互式视频业务。为此,本文在LEO层设计基于值函数的多智能体强化学习框架,使节点在继承全局路径约束的基础上具备快速局部修复能力,并结合平滑切换机制减弱链路波动对会话连续性的影响。

将每颗卫星视为一个智能体,对于当前转发节点 i ,其在时刻 t 的局部状态向量定义为

$$s_i(t) = [E_i(t), Q_i(t), h_i^{rem}(t), d_{ij}(t), \tilde{t}_{ij}(t), \tilde{b}_{ij}(t)]_{j \in \mathcal{N}_i} \quad (20)$$

其中, $E_i(t)$ 、 $Q_i(t)$ 、 $h_i^{rem}(t)$ 分别表示当前节点 i 的剩余能量、队列长度和剩余跳数,用于表征节点自身状态; $d_{ij}(t)$ 、 $\tilde{t}_{ij}(t)$ 、 $\tilde{b}_{ij}(t)$ 分别表示当前节点到邻居节点的链路时延、归一化TTL和归一化带宽,用于刻画邻域链路质量。

为避免局部修复引入明显不可行的下一跳,LEO层在触发切换后首先对邻居集合 $\mathcal{N}_i(t)$ 进行约束筛选,构造局部可行邻居集合

$$\mathcal{N}_i^{feas}(t) = \{j \in \mathcal{N}_i(t) | d_{ij}(t) \leq d_{\max}, TTL_{ij}(t) \geq TTL_{\min}, E_j(t) \geq E_{\min}, Q_j(t) \leq Q_{\max}, b_{ij}(t) \geq b_{\min}\} \quad (21)$$

在此基础上,局部策略网络对候选邻居输出动作价值,采用参数共享机制以提升样本利用效率与泛化能力。训练过程中进一步结合经验回放与目标网络机制,以减弱样本相关性并提高多智能体训练稳定性。训练阶段采用 ϵ -greedy 机制进行探索,评估阶段采用贪婪选择,新的下一跳表示为

$$n^* = \arg \max_{j \in \mathcal{N}_i^{feas}(t)} Q_\phi(s_i(t), j) \quad (22)$$

其中, Q_ϕ 表示经局部约束修正后的动作价值函数。若可行邻居集合 $\mathcal{N}_i^{feas}(t)$ 为空,则节点不执行激进切换,保持当前保守转发状态并向MEO层请求重规划。

为使局部策略与全局优化目标保持一致,定义节点 i 在时刻 t 的局部奖励为

$$r_i^L(t) = -\lambda_d \tilde{d}_{ij}(t) - \lambda_e \tilde{e}_{ij}(t) + \lambda_\tau \tilde{t}_{ij}(t) \quad (23)$$

其中, $\tilde{d}_{ij}(t)$ 、 $\tilde{e}_{ij}(t)$ 和 $\tilde{t}_{ij}(t)$ 分别表示候选局部链路

在时延、能耗和 TTL 维度上的归一化评价值, λ_d , λ_e, λ_r 为对应权重系数。该奖励函数鼓励智能体优先选择时延较低、能耗较小且持续可用时间更长的邻居节点, 从而在局部尺度上与全局规划目标保持一致。

对于交互式视频业务而言, 局部修复并不等价于对当前下一跳的瞬时替换。若在链路风险出现时直接切换, 容易引发缓存状态不一致、队列重构和突发时延抖动。为减弱上述不利影响, 本文进一步设计触发式平滑切换机制。

当当前转发节点 i 检测到与当前下一跳节点 n 之间的链路 TTL 不足, 或者候选节点资源状态恶化时, 触发局部切换。触发条件定义为

$$\tilde{T}_{i,n} < \theta_T \vee Q_n(t) > \theta_q \vee E_n(t) < \theta_E \quad (24)$$

其中, $\tilde{T}_{i,n}$ 表示链路 (i,n) 的归一化 TTL, $Q_n(t)$ 和 $E_n(t)$ 分别表示节点 n 的队列负载和剩余能量, $\theta_T, \theta_q, \theta_E$ 分别为链路 TTL 预警阈值、拥塞阈值和能量安全阈值。

触发后, 节点 i 并不立即中断旧路径, 而是在时刻 t_0 启动平滑切换窗口 ΔT , 同时维护旧下一跳 n 与新下一跳 n^* 两条转发项, 并采用线性权重逐步完成业务迁移。线性权重如下所示

$$C(t) = \text{clip}\left[\frac{(t-t_0)}{\Delta T}, 0, 1\right] \quad (25)$$

于是, 在区间 $[t_0, t_0 + \Delta T]$ 内, 新下一跳承担比例 $C(t)$ 的业务流量, 旧下一跳承担比例 $1 - C(t)$ 的业务流量; 当 $t \geq t_0 + \Delta T$ 时, 即 $C(t) = 1$, 或旧链路 TTL 已耗尽、旧队列已清空时, 停止旧路径转发, 完成局部平滑切换。本文取 $\Delta T = 10 \text{ s}$ 。

综上, LEO 层的 MARL 机制并非独立于全局规划的纯局部贪婪决策, 而是在继承 MEO 层全局路径约束的基础上, 通过局部状态感知、约束筛选、动作价值评估与平滑迁移相结合的方式, 实现对局部链路风险的快速自适应响应。与直接切换或单纯逐跳重选相比, 该机制更适合交互式视频业务对低抖动和会话连续性的传输需求。

3 实验仿真

3.1 仿真参数设置

为评估 H-GMRL 在交互式视频业务中的路由性能, 本文使用系统工具包 (Systems Tool Kit, STK) 与 PyCharm 仿真程序进行实验, STK 用于生成卫星

轨道并实时更新动态拓扑, PyCharm 用于实现路由决策与算法执行。此外, 本文构建了一个由 MEO 全局规划层和 LEO 业务转发层组成的分层异构卫星网络。其中, MEO 层设置 12 颗控制卫星, LEO 层采用三壳层低轨卫星星座, 总计 1440 颗卫星。需要说明的是, 本文将三壳层 LEO 星座统一抽象为单一业务转发层, 只要满足可见性、链路距离、容量与 TTL 约束的同壳层或跨壳层 LEO 卫星对, 均允许建立星间链路并参与候选路径构造。

实验场景设置为地面用户对之间的交互式视频通信。业务流首先由源用户接入当前可见的源 LEO 卫星, 经 LEO 星间网络多跳转发至目标用户所接入的 LEO 卫星后, 再下传至目标用户。本文主要关注星间网络层的路由决策, 后续路径优化与性能统计仅针对源接入 LEO 与目标接入 LEO 之间的星间转发段展开, 用户接入端则采用固定接入开销处理, 不参与路由选择。每轮实验中, 随机生成源宿用户对, 并根据业务发起时刻的星地可见性关系确定对应的源接入与目标接入 LEO 卫星。

其余关键仿真参数如表 2 所示。链路 TTL 相关参数依据 STK 链路可见窗口统计与切换提前量需求设定; 荷电状态 (State of Charge, SoC) 安全范围参考电池感知能量优化与避免深度放电的安全边界思想设定^[13]; 阳照/地影净能量变化率用于近似表征昼夜工况下的净能量演化差异^[14]。其中, 三壳层 LEO 异构星座及其动态链路可见性共同构成了强时变拓扑环境; SoC 约束与阳照/地影净功率差异共同刻画了能量时空异构的形成条件, 而各卫星节点在不同轨道位置和不同时刻呈现出的剩余能量差异, 则体现了该异构性在网络中的动态分布特征。

3.2 基准策略

为全面评估 H-GMRL 的有效性, 本文选取经典非学习型参考与代表性学习型路由策略作为基准进行对比。所有方法均在相同的拓扑环境下运行, 并采用一致的链路延迟和 TTL 计算规则, 以确保对比的公正性。具体基准策略如下:

1) 最小时延最短路 (CR): 以端到端时延最小为目标, 采用 Dijkstra 算法求解最小时延路径, 用于表征传统确定性低时延路由方法。

2) 能量感知最短路 (DEB-SP): 在最短路代价中引入能量项, 代表传统能量约束路由。

表2 关键仿真参数表

参数名称	参数取值
MEO 卫星数量	12 颗
LEO 卫星数量	1440 颗
载波频率	60 GHz
带宽	2 GHz
星间链路业务速率	35 Mb/s
天线增益	55 dBi
端到端时延预算	200 ms
发送/接收功率	5 W / 3 W
链路 TTL 预警阈值	25 s
系统损耗	3 dB
噪声系数	3 dB
链路建立能耗	5 J
电池容量/SoC 范围	100 Wh / (0.8, 0.2)
ϵ -greedy 策略参数 (初值 \rightarrow 末值)	1.0 \rightarrow 0.1
Double-DQN 参数(γ, η)	(0.95, 0.001)
日照净充电功率/地影净耗功率	25 W / -2 W
会话时间	7200 s
视频业务负载	8 Mb/s
100% 网络负载	35 Mb/s
控制周期 T_c	480 s
平滑切换窗口 ΔT	10 s

3) 纯分布式 MARL (Flat-MARL): 不引入 MEO 层全局规划与平滑切换机制, 仅依赖 LEO 层局部可观测状态进行逐跳决策。

4) 时序 GNN+Double-DQN (Temporal-GNN-DRL): 采用时序 GNN 获取全局状态表征, 并结合 Double-DQN 执行端到端路径选择, 不引入 LEO 自适应层与平滑切换机制。

3.3 性能指标

为全面评估 H-GMRL 在交互式视频业务场景下的传输效率、会话稳定性与切换控制能力, 本文选取以下五项指标进行性能评价:

1) 端到端平均时延: 数据包从源卫星传输至目标卫星的平均端到端时延, 用于衡量业务传输的实时性。

2) 端到端时延抖动: 会话期间端到端时延的波动程度, 用于衡量业务传输的平稳性与会话连续性。

3) 丢包率: 会话期间丢失数据包数与发送总数据包数之比, 用于衡量网络在链路失效、拥塞及切换场景下的可靠传输能力。

4) 路径变化幅度: 会话期间路径切换前后新旧传输路径的变化程度, 用于衡量路径调整对业务连续性的扰动程度。

5) 路径 TTL: 会话期间当前传输路径的可持续有效时间, 用于衡量路由策略对路径稳定维持能力的保障水平。

上述五项指标刻画了交互式视频业务在动态卫星网络中的承载效果, 也能够反映强时变拓扑和能量时空异构对业务传输的影响。其中, 强时变拓扑主要体现为链路可见性变化和路径频繁重构, 对时延抖动、路径变化幅度和路径 TTL 影响最为直接, 并对端到端时延和丢包率产生间接影响; 能量时空异构则体现为节点剩余能量和持续转发能力的时空差异, 主要影响路径 TTL 和丢包率, 并通过局部修复触发和路径绕行影响端到端时延。

3.4 结果分析

基于上述指标, 本文从训练收敛性、传输时效性、会话平稳性、可靠性以及路径稳定性五个方面, 对所提 H-GMRL 策略与各基准方法进行比较分析。同时, 为减弱随机源宿选取、策略初始化与探索过程带来的偶然影响, 本文在每个负载点下采用 5 个不同随机种子进行独立重复实验, 并对结果取平均。对关键结论进一步采用统计显著性检验进行验证, 当 $p < 0.05$ 时认为差异具有统计显著性。

图 3 给出了 H-GMRL、Flat-MARL 与 Temporal-GNN-DRL 三种学习型策略的平均奖励收敛曲线, 三种方法的平均奖励均随训练轮次增加而逐步提升并最终趋于稳定。相比另外两种学习型基线, H-GMRL 在训练早期具有更快的奖励增长速度, 并能在较少训练轮次内进入稳定区间, 且收敛后的平均奖励整体更高、波动更小, 表明所提分层协同机制能够有效缓解大规模动态拓扑下学习的非平稳性, 提高策略的长期决策收益与训练稳定性。需要说明的是, CR 与 DEB-SP 为经典非学习型策略, 其决策过程不涉及参数更新与训练迭代, 不存在可比的奖励收敛过程。故图 3 仅展示 H-GMRL 与两种学习型基准策略的训练收敛结果。

为比较不同路由策略在高负载条件下的性能差异, 本文以网络负载为横轴进行对比分析。网络负

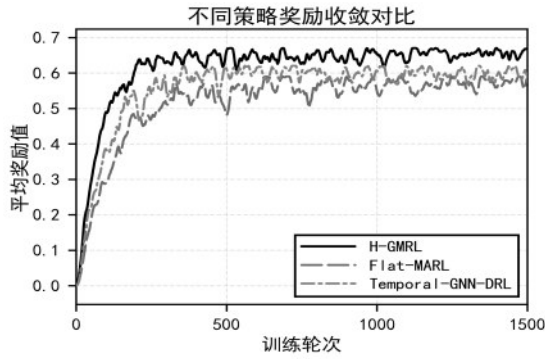


图3 不同学习型路由策略的奖励收敛对比图

载用于表征网络资源占用程度与拥塞水平，通过调节在网业务会话数及其聚合注入速率构造不同负载水平，使业务注入速率逐步逼近并超过局部星间链路容量上限，从而触发拥塞与链路切换，以考察各策略在高负载场景下的稳定性与鲁棒性。不同负载条件下各策略的性能指标对比如图4所示。

图4(a)和图4(b)分别给出了不同网络负载下各策略的平均端到端时延和平均时延抖动。总体来看，随着网络负载升高，各策略的时延与抖动均呈上升趋势，说明高负载条件下链路竞争、排队累积和切换扰动进一步加剧。相比各基线，H-GMRL在全负载范围内始终保持最低时延，且在中高负载区

间优势更为明显。以100%负载为例，H-GMRL的平均端到端时延约为196 ms，较CR、DEB-SP、Flat-MARL和Temporal-GNN-DRL分别降低约7.5%、12.1%、4.4%和7.1%。在时延抖动方面，H-GMRL的优势更为突出。以40%负载为例，其平均时延抖动约为0.41 ms，较四种基线分别降低约43.8%、45.3%、38.8%和42.3%。这表明，在强时变拓扑导致链路可见性快速变化和频繁切换的条件下，H-GMRL能够通过MEO层前瞻性路径规划与LEO层平滑局部修复协同，降低链路重构和队列扰动对业务实时性与会话平稳性的影响。

图4(c)和图4(d)分别给出了各策略的分组丢包率和平均路径变化幅度。可以看出，随着网络负载增加，各策略的丢包率和路径变化幅度整体上升，反映出高负载条件下链路拥塞、缓存溢出和路径失效风险进一步增强。在丢包率方面，H-GMRL在全负载区间内始终最低。以80%负载为例，其分组丢

包率较CR、DEB-SP、Flat-MARL和Temporal-GNN-DRL分别降低约44.8%、41.1%、33.8%和31.3%；在100%负载下仍分别降低约39.8%、35.5%、21.1%和32.0%。在平均路径变化幅度方

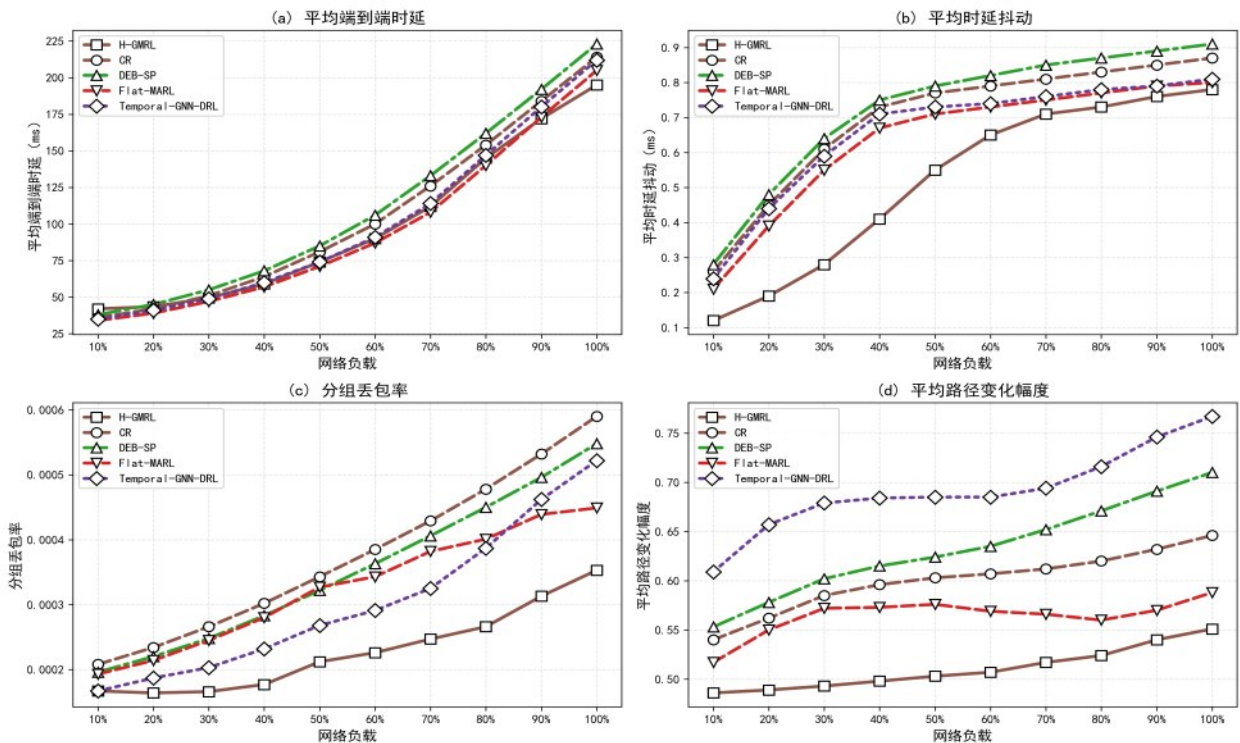


图4 四项性能指标综合对比图

面, H-GMRL 同样表现最优。以 100% 负载为例, 其路径变化幅度较 CR、DEB-SP、Flat-MARL 和 Temporal-GNN-DRL 分别降低约 14.9%、22.5%、6.5% 和 28.3%。这说明, 在拓扑重构和资源状态波动共同作用下, H-GMRL 能够将路径调整控制在较小范围, 并减少高风险链路或低能量节点引发的转发中断, 从而提升业务传输可靠性和切换稳定性。

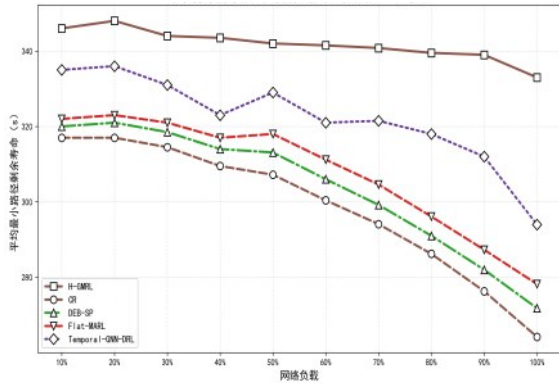


图5 平均路径TTL对比图

图5给出了不同网络负载下各策略的平均路径TTL对比图。相比之下, H-GMRL在全负载区间内始终保持最高路径TTL, 且下降幅度最缓, 体现出更强的路径持续承载能力。以100%负载为例, H-

GMRL的平均路径TTL约为333s, 而CR、DEB-SP、Flat-MARL和Temporal-GNN-DRL分别约为265s、272s、279s和294s, 所提策略分别提升约25.9%、22.4%、19.6%和13.3%。这表明, H-GMRL能够在全局规划阶段优先规避可见窗口较短或持续转发能力较弱的候选路径, 并在局部状态恶化时及时完成修复, 因此在强时变拓扑与能量时空异构共同作用下仍能保持更高的路径持续可用性。

为进一步验证H-GMRL中平滑切换、MEO全局规划和LEO局部自适应修复等关键模块的独立贡献, 本文在主结果对比基础上开展消融实验, 分别移除上述模块并比较各策略在不同负载条件下的性能变化, 结果如图6所示。

总体来看, 完整H-GMRL在平均时延抖动、分组丢包率、平均路径变化幅度和平均最小路径TTL四项指标上均优于各消融策略, 说明平滑切换、MEO全局规划和LEO局部自适应修复三类机制对提升整体性能均具有重要作用。

从时延抖动和路径变化幅度看, 移除平滑切换机制后, 会话平稳性明显下降。以40%负载为例, H-GMRL的平均时延抖动约为0.41ms, 而w/o Smooth Switching约为0.51ms, 降低约19.6%; 在100%负载下, H-GMRL的平均路径变化幅度约为

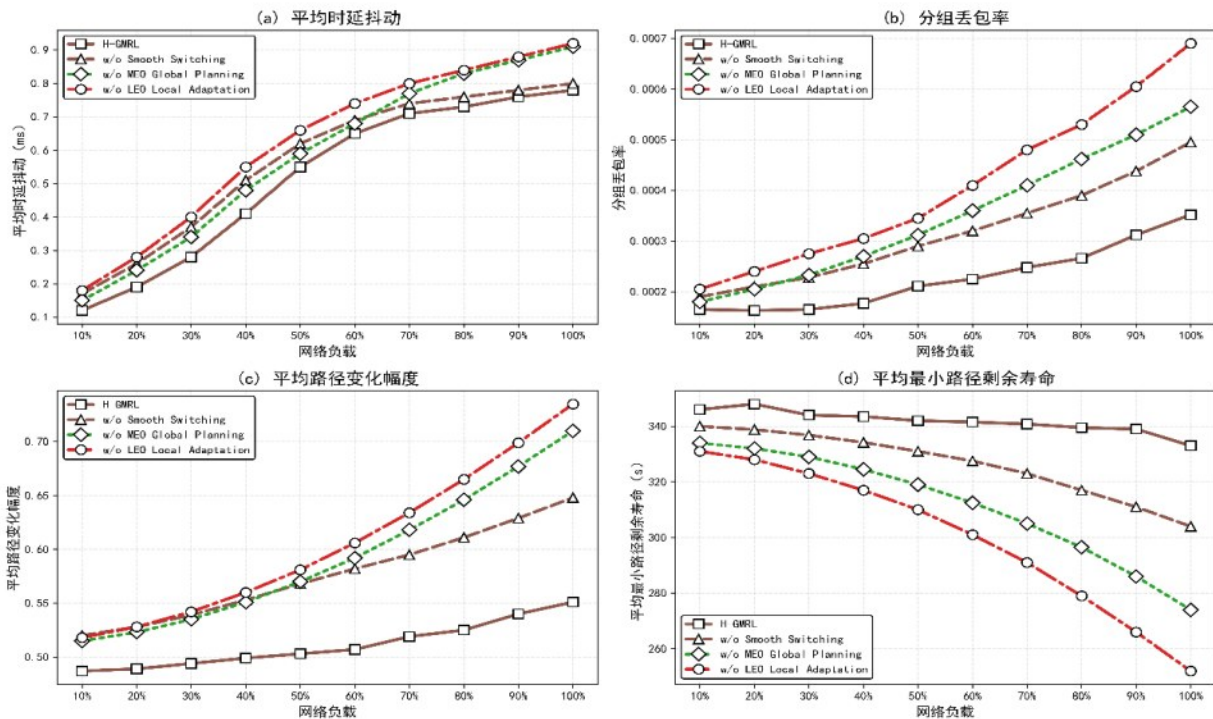


图6 不同负载条件下各消融策略性能指标对比图

0.55, 而 w/o Smooth Switching 约为 0.65, 降低约 15.0%。这表明平滑切换能够有效减弱路径迁移过程中由队列重构和转发重定向引起的瞬时扰动。

从丢包率和路径 TTL 看, 移除 MEO 全局规划或 LEO 局部自适应修复后, 性能均明显退化, 且后者影响更为显著。以 100% 负载为例, H-GMRL 的分组丢包率约为 3.55×10^{-4} , 较 w/o MEO Global Planning 和 w/o LEO Local Adaptation 分别降低约 37.2% 和 48.6%; 对应平均最小路径 TTL 约为 333 s, 较两种消融策略分别提升约 21.5% 和 32.1%。这说明 MEO 全局规划对提升路径前瞻性和持续可用性具有重要作用, 而 LEO 局部自适应修复则是保障高负载条件下可靠传输与路径稳定性的关键环节。

为验证关键控制参数设置的合理性, 本文进一步开展 H-GMRL 关键参数敏感性分析。针对控制周期 T_c 、平滑切换窗口 ΔT 和链路 TTL 预警阈值 θ_T , 采用单因素分析方法逐一调整参数, 其余设置保持默认值不变。相关结果均在 80% 网络负载下统计得到, 如表 3 所示。选择该负载点的原因在于, 此时网络已表现出较明显的拥塞与切换特征, 但尚未完全饱和, 能够较好反映参数变化对策略性能的影响。

由表 3 可见, 对于控制周期 T_c , 当 $T_c = 480$ s 时, 平均时延最低且路径 TTL 最高, 综合性能最优, 说明该取值能够在全局视图新鲜度与控制开销之间取得较好平衡。对于平滑切换窗口 ΔT , 当 $\Delta T = 10$ s 时, 时延抖动和路径变化幅度均最小, 表明该取值在切换平滑性与响应速度之间具有最佳折中。对于链路 TTL 预警阈值 θ_T , 较小阈值有利于维持更高的路径持续性, 而较大阈值有助于降低丢包率。综合比较可知, $\theta_T = 25$ s 时在可靠性与路

表 3 关键参数敏感性分析表(80% 网络负载)

参数	取值	指标 1	指标 2
T_c	360 s	时延=144.72 ms	路径 TTL=339.19 s
T_c	480 s	时延=144.18 ms	路径 TTL=342.05 s
T_c	600 s	时延=145.44 ms	路径 TTL=340.93 s
ΔT	5 s	抖动=0.752 ms	路径变化幅度=0.61
ΔT	10 s	抖动=0.712 ms	路径变化幅度=0.52
ΔT	15 s	抖动=0.739 ms	路径变化幅度=0.55
θ_T	20 s	丢包率= 2.49×10^{-4}	路径 TTL=342.64 s
θ_T	25 s	丢包率= 2.45×10^{-4}	路径 TTL=341.16 s
θ_T	30 s	丢包率= 2.43×10^{-4}	路径 TTL=330.79 s

径连续性之间取得了更合理的折中, 因此本文选取 $\theta_T = 25$ s 作为默认阈值。

3.5 复杂度与方法代价分析

为进一步评估所提 H-GMRL 策略在在线部署中的可实现性及其相较于基准策略的方法代价, 本文从理论复杂度和可量化在线开销两个层面进行分析。

设 $|V|$ 和 $|E|$ 分别表示网络中的卫星节点数和星间链路数, L_g 表示 GNN 的消息传递层数, d_g 表示 GNN 节点嵌入维度, d_l 表示局部状态特征维度, $|A_t|$ 表示时刻 t 的全局可行动作集规模, \bar{k} 表示局部可行邻居集合的平均规模, W 表示平滑切换窗口内的子窗口数。

对于所提 H-GMRL 策略, 其时间复杂度由 MEO 层全局规划与 LEO 层局部修复两部分构成。

对于 MEO 层, 其时间复杂度主要由网络状态汇聚与动态图更新、GNN 状态编码、候选动作约束筛选以及 Double-DQN 前向决策构成, 因此单个控制周期内的时间复杂度可表示为 $O(|V| + |E| +$

表 4 各策略的时间复杂度与空间复杂度对比表

策略	时间复杂度	空间复杂度
CR	$O(E + V \log V)$	$O(V + E)$
DEB-SP	$O(E + V \log V)$	$O(V + E)$
Flat-MARL	$O(\bar{k})$	$O(\bar{k} d_l) + O(params_{MARL})$
Temporal-GNN-DRL	$O(V + E + L_g E d_g + A_t)$	$O(E) + O(V d_g) + O(params_{GNN-DRL})$
H-GMRL (MEO 层)	$O(V + E + L_g E d_g + A_t)$	$O(E) + O(V d_g) + O(params_{MEO})$
H-GMRL (LEO 层)	$O(\bar{k} + W)$	$O(\bar{k} d_l) + O(params_{LEO})$

$L_g|E|d_g + |A_t| + f_{\text{DDQN}}(|A_t|, d)$ 。当嵌入维度及网络前向规模视为常数时, 其主导项可近似写为 $O(|V| + |E| + L_g|E|d_g + |A_t|)$ 。

对于 LEO 层, 局部修复仅在触发条件满足时执行, 其时间复杂度主要由局部状态观测、邻居约束筛选、局部策略网络评估和平滑切换状态维护构成, 因此单次触发平滑切换的时间复杂度可表示为 $O(\bar{k} + f_{\text{MARL}}(\bar{k}) + W)$ 。当局部策略网络规模固定时, LEO 层单次触发平滑切换的计算开销约为 $O(\bar{k} + W)$ 。

对于 H-GMRL 的空间复杂度方面, MEO 层主要开销由动态图存储、节点嵌入缓存以及 GNN 与 Double-DQN 模型参数构成, 其空间复杂度可表示为 $O(|E|) + O(|V|d_g) + O(\text{params}_{\text{MEO}})$; 对于 LEO 层, 每个节点仅需维护局部邻域状态、当前转发项以及固定的局部策略网络参数, 其空间复杂度可表示为 $O(\bar{k}d_t) + O(\text{params}_{\text{LEO}})$ 。

总体来看, H-GMRL 通过分层协同将高频在线响应限制在局部邻域内完成, 而将全局路径规划置于控制周期尺度上执行, 从而在计算复杂度、控制代价与在线响应能力之间取得了较为合理的折中。

为比较所提方法与各基准策略在在线决策代价上的差异, 表 4 给出了不同路由策略的时间复杂度与空间复杂度对比。相较于 CR 和 DEB-SP, 其在在线决策由集中式最短路搜索 $O(|E| + |V|\log|V|)$ 扩展为在 MEO 层引入 $O(|V| + |E| + L_g|E|d_g + |A_t|)$ 全局规划开销, 并在 LEO 层增加 $O(\bar{k} + W)$ 局部维护开销; 相较于 Flat-MARL, H-GMRL 由局部逐跳决策 $O(\bar{k})$ 增加为全局规划与局部修复的分层协同; 相较于 Temporal-GNN-DRL, H-GMRL 的 MEO 层全局规划复杂度处于同一量级, 新增代价主要体现在 $O(\bar{k} + W)$ 局部修复与平滑切换维护上。

结合第 3.4 节结果可见, H-GMRL 虽引入了一定的复杂度增量, 但换来了更优的性能表现。相较于 CR 和 DEB-SP, 在 80% 负载下丢包率分别降低约 44.8% 和 41.1%, 在 100% 负载下平均路径 TTL 分别提升约 25.9% 和 22.4%; 相较于 Flat-MARL, 在 80% 负载下丢包率降低约 33.8%, 在 100% 负载下平均路径 TTL 提升约 19.6%; 相较于 Temporal-

GNN-DRL, 在 40% 负载下时延抖动降低约 42.3%, 在 100% 负载下路径变化幅度降低约 28.3%。总体来看, H-GMRL 在引入有限复杂度增量的基础上, 有效提升了可靠性、路径持续性与会话平稳性, 在计算复杂度、控制代价与在线响应能力之间取得了较为合理的折中。

除理论复杂度外, 本文进一步从控制消息规模、状态交换规模和切换维护开销三个方面分析 H-GMRL 的在线代价。其中, CR、DEB-SP、Temporal-GNN-DRL 及 H-GMRL 的全局规划部分均属于集中式全局决策过程, 因此都需要承担状态收集、全局汇聚和结果下发等基础控制开销; Flat-MARL 采用纯分布式决策, 不涉及上述集中式控制平面交互, 故不纳入该部分比较。按本文仿真设置, 一个控制周期内 LEO 层的状态摘要上报规模至多为 1440 条, MEO 间区域级状态交换规模不超过 66 条。在此基础上, H-GMRL 新增代价主要体现在 LEO 层局部修复与平滑切换维护上, 即切换期间仅额外维护 1 条新的转发项, 使并行维护总数由 1 条增至 2 条, 持续时间为 10 s。总体来看, H-GMRL 在继承集中式策略基础控制开销的同时, 仅引入了规模有限的增量维护代价, 换来更优的性能表现。

4 总结

面向低轨卫星网络中交互式视频业务的低时延、低抖动和会话连续性需求, 提出了一种 MEO-LEO 分层协同路由策略 H-GMRL。该策略通过 MEO 层多约束全局路径规划、LEO 层局部自适应修复以及触发式平滑切换, 实现了全局路径决策与局部快速响应的协同。研究结果表明, 所提策略能够适应强时变的低轨卫星网络环境, 对交互式视频业务的连续稳定承载具有较好的支撑能力。后续将进一步研究结合业务预测与状态前瞻估计的预测增强型分层协同路由机制。

参考文献:

- [1] Zhu X, Jiang C. Creating efficient integrated satellite-terrestrial networks in the 6G era[J]. IEEE Wireless Communications, 2022, 29(4): 154-160.
- [2] Wang X, Yu R, Yang D, et al. Space booking: enabling performance-critical applications in broadband satellite networks[C]//Proceedings of the IEEE International Conference on Distributed Computing Systems

(ICDCS). Glasgow: IEEE, 2025: 989-999.

[3] Zhao N, Long X, Wang J. A multi-constraint optimal routing algorithm in LEO satellite networks[J]. Wireless Networks, 2021, 27(4): 1107-1117.

[4] Wang C, Liu Y, Zhang X. Energy-efficient routing in LEO satellite networks for extending satellites' battery lifetime[J]. Journal of Communications and Networks, 2022, 24(3): 45-56.

[5] Li Y, Zhang H, Qi Y, et al. Available energy routing algorithm considering QoS requirements for LEO satellite networks[J]. Computer Communications, 2024, 188: 120-132.

[6] Xu G, Zhao Y, Ran Y, et al. Spatial location aided fully-distributed dynamic routing for large-scale LEO satellite networks[J]. IEEE Communications Letters, 2022, 26(12): 3034-3038.

[7] Xiang J, He X, Zhao Y, et al. Distributed dynamic routing for LEO satellite networks with temporal graph convolutions and imitation acceleration[J]. IEEE Communications Letters, 2025, 29(11): 2521-2525.

[8] Shi Y, Wang W, Zhu X, et al. Low Earth orbit satellite network routing algorithm based on graph neural networks and deep Q-network[J]. Applied Sciences, 2024, 14(9): 3840.

[9] Ran Y, Ding Y, Chen S, et al. Fully-distributed dynamic packet routing for LEO satellite networks: a GNN-enhanced multi-agent reinforcement learning approach[J]. IEEE Transactions on Vehicular Technology, 2025, 74(3): 5229-5234.

[10] Zhao J, Pan J. Low-latency live video streaming over a low-Earth-orbit satellite network[C]//Proceedings of the 15th ACM Multimedia Systems Conference (MMSys'24). New York: ACM, 2024: 100-108.

[11] Li S, Wu Q, Wang R. Dynamic discrete topology design and routing for satellite-terrestrial integrated networks[J]. IEEE/ACM Transactions on Networking, 2024, 32(5): 3840-3853.

[12] Zhang S, Liu A, Han C, et al. Federated graph reinforcement learning-driven adaptive routing strategy for mega LEO satellite networks[J]. IEEE Wireless Communications Letters, 2025, 14(10): 3344-3348.

[13] Li Q, Wang S, Ma X, et al. Battery-aware energy optimization for satellite edge computing[J]. IEEE Transactions on Services Computing, 2024, 17(2): 437-451.

[14] Baccari S, Mostacciolo E, Tipaldi M, et al. A deep reinforcement learning approach for energy management in low Earth orbit satellite electrical power systems[J]. Electronics, 2025, 14: 3110.

万锦辉 (1978-)，女，广西玉林人，硕士，广西科技信息网络中心副高级工程师，主要研究方向为项目管理、人工智能。



王吉平 (1977-)，男，四川威远人，硕士，桂林电子科技大学高级实验师，主要研究方向为移动通信、通信网络和水下图像处理。



王波 (1977-)，男，陕西西安人，博士，桂林电子科技大学副教授，主要研究方向为 6G、移动边缘计算和绿色通信。



卢锐 (2001-)，男，广西贵港人，桂林电子科技大学研究生，主要研究方向为低轨卫星网络和卫星路由。