

基于图强化学习的LEO卫星网络智能路由策略

白禄鑫¹, 胡宇翔^{1,2,3}, 马海龙^{1,2}, 江逸茗^{1,2}, 胡涛^{1,2}

(1. 信息工程大学信息技术研究所, 郑州 450002; 2. 网络空间安全教育部重点实验室, 郑州 450002;
3. 先进通信网全国重点实验室, 郑州 450002)

摘要: 新兴的低轨 (LEO, low earth orbit) 巨型星座在全球通信领域展示出巨大的潜力。然而, 庞大的网络规模和频繁的网络故障导致现有基于深度强化学习 (DRL, deep reinforcement learning) 的路由方案的扩展性和鲁棒性较差。为此, 提出一种基于图强化学习的鲁棒和可扩展的路由框架 GraphSR。首先, GraphSR 采用一种称为虚拟流量拓扑模型来屏蔽星座动态拓扑, 虚拟流量拓扑将网络故障引起的物理拓扑变化转换为流量需求的多样性, 从而实现鲁棒路由的计算; 其次, 为了实现路由可扩展性, GraphSR 将星座划分为多个软件定义网络 (SDN, software-defined networking) 控制域, 并设计了基于图神经网络 (GNN, graph neural network) 的 DRL 算法以摆脱对固定拓扑的输入依赖; 最后, 通过蒙特卡洛模拟快速评估关键故障场景并选择对网络性能影响较大的部分关键链路, GraphSR 专注于控制关键链路权重来降低 DRL 的动作空间维度。在 Starlink 巨型星座拓扑上的仿真表明, 与其他基线路由方案相比, GraphSR 在不同规模的链路失效场景下路径可靠性提升超过 7.04%, 同时端到端时延始终保持在可接受范围内。

关键词: 低轨巨型星座; 韧性路由; 软件定义网络; 图神经网络; 深度强化学习

中图分类号: TN915.08

文献标志码: A

doi: 10.11959/j.issn.2096-3750.WLW25023

Graph reinforcement learning-based intelligent routing strategy for LEO satellite networks

Bai Luxin¹, Hu Yuxiang^{1,2,3}, Ma Hailong^{1,2}, Jiang Yiming^{1,2}, Hu Tao^{1,2}

1. Institute of Information Technology, Information Engineering University, Zhengzhou 450002, China
2. Key Laboratory of Cyberspace Security, Ministry of Education, Zhengzhou 450002, China
3. National Key Laboratory of Advanced Communication Networks, Zhengzhou 450002, China

Abstract: Emerging low earth orbit (LEO) mega-constellations demonstrate significant potential in global communications. However, the vast network scale and frequent disruptions result in poor scalability and robustness of existing deep reinforcement learning (DRL)-based routing schemes. GraphSR, a robust and scalable routing framework based on graph reinforcement learning was proposed. First, GraphSR adopts a new topology model, called virtual traffic topology to shield dynamic topology disruptions. The physical topology changes caused by network failures are converted into the diversity of traffic demands to facilitate robust routing calculations. Second, to achieve routing scalability, GraphSR divides the constellation into multiple software-defined networking (SDN) control domains and designs a graph neural network (GNN)-based DRL algorithm to get rid of the input dependence on fixed topology. Lastly, by quickly evaluating critical failure scenarios and selecting some critical links that significantly impact network performance, GraphSR focuses on controlling the weights of critical links to reduce the action space dimension of DRL. Simulations on the Starlink mega-

收稿日期: 2025-02-18; 修回日期: 2026-05-15

通信作者: 胡涛, hutaondsc@163.com

基金项目: 雄安新区科技创新专项 (No. 2022XAGG0111)

Foundation Item: Xiong'an New Area Science and Technology Innovation Special Project (No. 2022XAGG0111)

constellation topology show that compared with other baseline routing schemes, GraphSR improves path reliability by more than 7.04% under various link failure scenarios, while the end-to-end delay remains within an acceptable range.

Key words: low earth orbit mega-constellation, resilient routing, software-defined networking, deep reinforcement learning, graph neural network

0 引言

随着通信技术的发展和卫星制造、发射成本的大幅降低,低轨(LEO, low earth orbit)卫星网络被视为未来全球通信和下一代互联网的关键基础设施^[1]。SpaceX、OneWeb等公司正积极推进部署基于LEO的巨型星座,以提供低时延、高带宽和广覆盖的网络接入服务。

为了实现上述目标,路由发挥着关键作用,直接影响着网络的整体性能和运行效率。然而,LEO卫星网络(LSN, LEO satellite network)的独特网络特性对路由算法的设计提出了诸多挑战^[2-4]。首先,LSN暴露在开放且恶劣的太空环境中^[5],面临着太空碎片和太阳风暴等自然威胁。同时,小型LEO卫星寿命较短,且存在多种安全漏洞和系统缺陷^[6],易成为网络攻击的潜在入口^[7]。最后,人口分布不均和经济发展差异等地理条件与轨道运动的动态特性影响全球流量模式,使得网络负载不均衡,即热点地区的链路带宽长期超载,而其他链路带宽闲置浪费。上述所有因素都可能导致LSN的网络故障。

针对这些挑战,结合软件定义网络(SDN, software-defined networking)的集中式控制架构^[8-9]和基于深度强化学习(DRL, deep reinforcement learning)的智能路由方法表现出对流量传输较好的控制能力,正在被广泛研究以优化LSN的流量分配^[10]。文献[11]引入SDN技术,通过改进的广度优先搜索算法识别多条最短路径,并利用Q-learning优化路由决策,以提升网络性能。考虑高部署成本和有限的星上资源,文献[12]提出了一种基于DRL的节能路由,在满足端到端时延约束的同时,避免星座间的电池能量出现不平衡。为应对高维连续动作空间的挑战,文献[13]采用信息素激励的多路径路由发现算法,快速识别和选择高效路径,并使用深度确定性策略梯度(DDPG, deep deterministic policy gradient)方法优化多路径流量调度策略。文献[14]结合多智能体强化学习(MARL, multi-agent

reinforcement learning)和拉格朗日方法,通过迭代更新卫星和地面站的路由策略,满足能耗和丢包率的约束。

然而,现有的智能路由方法应用于LSN仍面临两大问题^[15]:扩展性较差和鲁棒性较弱。基于DRL的路由通常需要控制网络中目标元素(链路或数据流)的所有单元^[16],而LSN中密集的链路元素导致状态、动作空间急剧增大,模型难以有效训练和收敛。此外,这些路由算法通常使用全连接等传统神经网络架构^[17],容易过拟合训练拓扑。一旦网络故障引起拓扑结构变化,路由性能会严重下降。

因此,为满足LEO巨型星座数据传输的高可用和低时延要求,需为空间段设计一种鲁棒、可扩展的路由算法,以应对频繁的网络故障和大规模网络拓扑变化。图神经网络(GNN, graph neural network)在处理具有复杂拓扑结构的网络问题时展现出显著优势^[18-19]。作为专为图结构数据而设计的神经网络,GNN能够动态适应拓扑变化。此外,分布式SDN控制器架构通过逻辑集中、物理分散的方式部署多个控制器节点,可有效管理整个大规模网络。

本文提出了一种用于LEO巨型星座的鲁棒、可扩展路由框架GraphSR,目标是最大化路径可靠性,并将其表述为时延有界约束下的整数非线性规划(INLP, integer nonlinear programming)问题。具体来说,GraphSR结合了两种关键技术来有效应对网络故障和巨型拓扑。首先,GraphSR采用一种新的拓扑模型,即虚拟流量拓扑(VTT, virtual traffic topology),将各种网络故障引起的拓扑变化转化为流量需求的多样性。在物理拓扑中某条链路的失效,可以认为在VTT中存在一个突发流量需求完全占据该链路带宽,进而在稳定的VTT上实现鲁棒路由计算。其次,受星座+Grid连接的拓扑特性启发,GraphSR在SDN范式下将整个网络划分为多个形状、规模一致的矩形,以保持网格特性。在每个控制域内,仅有少数关键链路的故障会对网络

性能产生较大影响。通过蒙特卡洛模拟面向关键故障场景的链路选择算法, GraphSR 采用基于 GNN 的 DRL 算法在 VTT 中控制部分关键链路来降低动作空间维度, 其中 GNN 用于摆脱对固定拓扑的输入依赖, 实现路由可扩展。

本文的贡献可以总结为以下几点。

1) 本文指出了目前 DRL 算法在大规模网络路由策略生成中普遍存在的可扩展性问题, 以及在网络故障下的鲁棒性问题, 并用数学建模了鲁棒空间段路由问题。

2) 本文提出了 GraphSR, 结合 VTT 和基于 GNN 的 DRL 路由方案, 在容易发生故障的 LEO 巨型星座中实现了鲁棒和可扩展的路由。

3) 通过在 Starlink 巨型星座拓扑上的广泛评估, 证明了 GraphSR 的优越性能。

1 系统模型与问题建模

1.1 场景描述

LEO 巨型星座的形式化描述可采用 Walker 编码 (T, P, F) , 其中 T 为卫星总数, P 为轨道数, F 为相位因子。具体来说, 星座由 P 个均匀分布的圆轨道面组成, 每个轨道面上均匀分布 T/P 颗卫星。星座组网采用网格连接 (+Grid) 方式^[20], 即每颗卫星配备 4 个激光终端^[21], 分别与同轨道面内和相邻轨道面的 4 颗卫星形成激光星间链路 (ISL, inter-satellite link)。

软件定义 LEO 巨型星座架构如图 1 所示。SDN 控制范式将网络分为逻辑上独立的数据平面和控制平面^[22], 通过控制器的集中控制能力实时监控网络状态, 并快速识别和响应网络故障。本文仅考虑空间段的星间路由, 而忽略分布在地球上的直连终端和地面站与 LEO 卫星之间的星地路由。因此, 数据平面由 LEO 卫星构成, 而控制器可以部署在中轨 (MEO, middle earth orbit) 卫星、地球静止轨道 (GEO, geostationary orbit) 卫星或地面。随着星上计算资源的增加, 控制功能甚至可以部署在 LEO 卫星内部, 实现带内控制, 但这超出了本文的讨论范围。

考虑控制器的处理能力有限, 本文采用多控制器的分布式部署策略, 根据特定的划分策略, 将整个 LEO 卫星网络划分为若干个独立的控制域, 每个域由一个从控制器管理。这些从控制器由一个

全局主控制器统一管理, 遵循主从关系协同工作。首先, 主控制器运行域间路由算法, 将跨域流量需求转换为多个连续的域内流量需求, 并生成各子域的流量需求矩阵。然后, 子控制器根据流量需求矩阵和链路状态生成特定链路权重, 并计算域内路由。最后, 当网络状态发生改变时, 控制器动态更新路由决策, 以保证网络的高效运行。

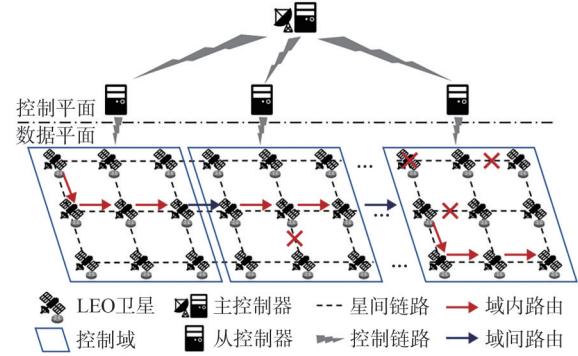


图1 软件定义LEO巨型星座架构

卫星星座的运行周期被划分为多个连续的拓扑快照, 将二维网格拓扑建模为时空图 $G = \{(V, E, T, N)\}$, 其中 V 表示所有 LEO 卫星, E 表示动态变化的 ISL 集合, T 表示观测的时间范围, N 表示划分时隙的数量。分布式 SDN 控制器架构采用主从控制器架构, 在水平和垂直方向将网络划分为 M 个不相交的域 $C = \{C_1, C_2, \dots, C_M\}$, 每个规模为 $m \times n$ 的域 C_i 的网络拓扑表示为 $G_i(V_i, E_i)$ 。

1.2 问题建模

路由旨在确定数据在网络中从源到目的的路径, 设 $f_{ij} = \{t_n, d_{ij}\}$ 表示在时隙 t_n 内, 从节点 i 到节点 j 的流量带宽需求为 d_{ij} , $F = \{f_{ij}|i, j \in V\}$ 定义为所有流量需求的集合。设 b_{uv} 为链路 (u, v) 在 $u \rightarrow v$ 方向上的容量, 定义二进制决策变量 $y_{ij}(u, v, t_n)$ 表示从节点 i 到节点 j 的流量是否通过链路 (u, v) 。为了确保路由策略 $r_{ij}^{t_n} = \{y_{ij}(u, v, t_n)\}$ 能够构成首尾相接的有效路径, 模型满足以下约束

$$\sum_{u \in V} y_{ij}(u, k, t_n) \cdot f_{ij}^{t_n} - \sum_{v \in V} y_{ij}(k, v, t_n) \cdot f_{ij}^{t_n} = \begin{cases} d_{ij}, & k = i \\ -d_{ij}, & k = j, \forall k \in V, \forall t_n \in T \\ 0, & \text{其他} \end{cases} \quad (1)$$

$$\sum_{(i,j) \in V \times V} y_{ij}(u, v, t_n) \cdot f_{ij}^{t_n} \leq b_{uv}, \forall (u, v) \in E, \forall t_n \in T \quad (2)$$

式(1)表示流守恒约束, 式(2)表示链路容量约

束，可确保每条链路 (u, v) 的负载不超过其最大容量 b_{uv} 。

此外，本文考虑以下两个与路由鲁棒性相关的指标。

1) 路由路径的可靠性。

人口分布不均衡和恶意的网络攻击均会增加链路负载，本文将链路故障的发生概率 $^{[23]}p_{fail}$ 建模为

$$p_{fail} = L_{uv} / \sum_{(u, v) \in E} L_{uv} \quad (3)$$

其中， L_{uv} 表示链路 (u, v) 的负载。在路由方案 r_{ij}^t 下，链路 (u, v) 的失效概率 p_{uv}^{fail} 表示为

$$p_{uv}^{fail} = \frac{\sum_{i, j \in V} d_{ij} \cdot y_{ij}(u, v, t_n)}{\sum_{(u, v) \in E} \sum_{i, j \in V} d_{ij} \cdot y_{ij}(u, v, t_n)} \quad (4)$$

本文定义路由路径的可靠性为数据成功传输的概率，即对于流 f_{ij} ，所有组成该路径的链路成功传输的联合概率，则所有路由路径的平均可靠性 P^t 表示为

$$P^t = \frac{1}{|V|^2} \sum_{(i, j) \in V \times V} \prod_{(u, v) \in r_{ij}^t} (1 - p_{uv}^{fail}) \quad (5)$$

2) 路由路径的时延要求。

链路 (u, v) 的时延表示为 $\delta_{uv} = l_{uv}/c + \delta_{proc}$ ，包含传播时延 l_{uv}/c 和处理时延 δ_{proc} 两部分，其中 l_{uv} 为链路的可视距离， c 为光速， δ_{proc} 为常数。引入时延约束 $D_{ij} = \alpha \cdot D_{ij}^{min}$ ，表示流 f_{ij} 的单向延迟应低于 D_{ij} ，其中 D_{ij}^{min} 是从节点 i 到节点 j 的最短路径时延，而系数 $\alpha \geq 1$ 用来容忍一定的时延增长。

综上所述，本文将鲁棒空间段路由问题建模为在有界的时延和流约束下，最大化路由可靠性。其数学模型表示为

$$\begin{cases} \max_{y_{ij}^t} P^t, \forall t_n \in T \\ \text{s.t.} \sum_{(u, v) \in r_{ij}^t} l_{uv} c + |r_{ij}^t| \delta_{proc} \leq D_{ij}, \forall f_{ij} \in F \end{cases} \quad (6)$$

在每个时隙 t_n ，需要求解上述INLP问题以得到由变量 $y_{ij}(u, v, t_n)$ 表示的最优路由策略。在大规模

模动态网络环境中，即使对于某个时隙下的简单静态情况，路由规划的解空间依然十分庞大。

2 GraphSR方法

2.1 域间和域内路由

鉴于LSN拓扑呈现高度规则的网格状，GraphSR采用A*算法（而不是Dijkstra）^[24]快速确定经过的域边界节点和域内链路，其中启发式函数选择曼哈顿距离。主控制器将跨域流量请求分解为多个域内流量请求，由相应的从控制器来确保流量在其域内可靠传输。

为应对网络故障导致的拓扑变化，GraphSR设计了VTT，将链路故障视为带宽耗尽。VTT将物理拓扑中各种故障引起的变化转换为流量矩阵 $\mathbf{TM}$ 的多样性。GraphSR方法在域内的路由计算流程如图2所示。在第一阶段，GraphSR通过SDN控制器整合跨域流量需求，构建VTT并生成新的 $\mathbf{TM}$ 。第二阶段，GraphSR在线选择关键链路，并通过基于GNN的DRL智能体调整VTT链路权重。在第三和第四阶段，利用控制域的网格结构，GraphSR通过A*算法计算在最小跳数区域内满足时延约束的 K 条路径。最终，路由决策在虚拟拓扑上完成，而流量转发基于物理拓扑实施。

2.2 关键链路选择

文献[25-26]提出根据网络拓扑关系（如度、介数中心性等指标）选取部分关键链路生成控制信号，以减少控制策略的输出动作维度。然而，卫星网络具有更规则的拓扑结构，+Grid的连接模式使网络元素趋于同质化。图3展示了不同网络中边介数中心性的比较，其中Starlink第一壳层和One-Web^[27]拓扑的边中心性分别分布在[0.003 4, 0.011]、[0.003, 0.028]，而地面拓扑Geant的边中心性分布在[0.002 5, 0.271]。鉴于人口分布不均衡导致的流量分布差异，链路在网络中的重要性评估更适合以流量为基础进行。

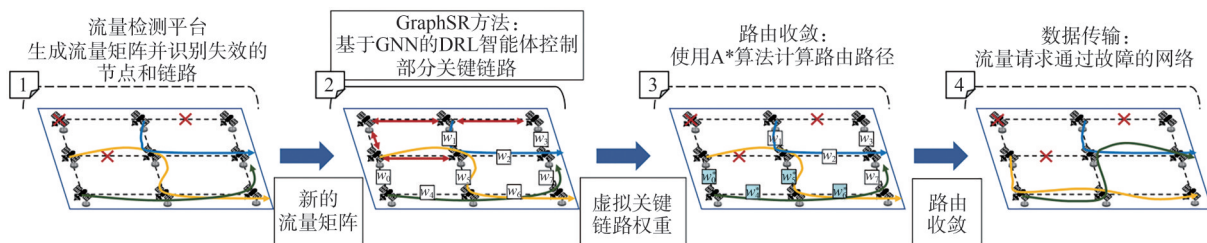


图2 GraphSR方法在域内的路由计算流程

GraphSR 采用了一种面向关键故障场景的链路选择方法, 对于特定的网络拓扑和流量需求, 只有少数故障场景是关键的。图4展示了在最短路径优先 (SPF, shortest path first) 路由下, 不同网络在两条链路故障下的最大链路利用率分布。结果表明, 无论是具有 23 个节点的地面拓扑 Geant, 还是 Starlink 中 4×4 、 6×6 等不同规模的集群网络拓扑, 在所有故障场景中分别仅有 1.39%、1.05% 和 1.04% 的 MLU 超过 80%, 这意味着对网络可用性造成显著影响的关键链路比例很小。

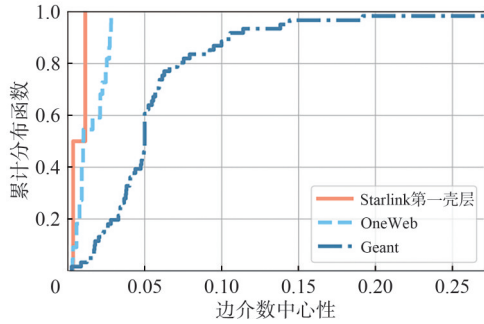


图3 不同网络中边介数中心性的比较

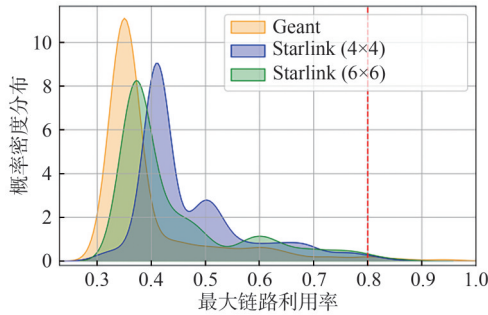


图4 不同网络在两条链路故障下的最大链路利用率分布

为此, 本文采用蒙特卡洛算法对故障场景的组合进行有限的 N_{mc} 次模拟, 以识别网络中的关键链路, 其中链路被选择的概率正比于链路在 SPF 路由方案下的初始负载。当 N_{mc} 为 1 时, 关键链路选择算法退化为选择网络中负载最大的前 λ 条链路。

算法1 关键链路选择算法

输入: 网络拓扑 $G_i(V_i, E_i)$ 、链路选择数 λ 、蒙特卡洛模拟数 N_{mc}

输出: 关键链路集合 $\hat{s}$

捕获流量需求并生成流量矩阵 $\mathbf{TM}$;

根据所有流量需求计算基于 SPF 路由下的链路负载;

基于链路负载计算链路抽样概率 P_{sample} ;

初始化 S, I 为 $\emptyset$, 用于存储故障场景和 MLU 指标;

while $|S| < N_{mc}$ **do**

 根据 P_{sample} 抽样得到 λ 条链路故障的场景 s ;

if $s \notin S$ **then**

 根据 SPF 路由计算场景 s 下所有流量需求的 MLU 指标 i_s

$S \leftarrow S \cup \{s\}, I \leftarrow I \cup i_s$

 选择 S 中具有最高 i_s 的链路集合 $\hat{s}$ 。

算法1详细描述了 GraphSR 中关键链路选择的过程。首先根据 SPF 路由和当前流量需求矩阵 $\mathbf{TM}$ 计算每条链路的负载, 得到链路抽样概率 P_{sample} 。其次在模拟次数少于 N_{mc} 时, 重复抽样出一个包含 λ 条链路的故障场景 s , 并检查这个场景是否已存在于集合 S 中。如果 s 为新场景, 则计算该场景下所有流量需求对应的 MLU, 并将场景 s 和 MLU 指标 i_s 分别添加到集合 S 和 I 中。最后从存储的故障场景中选择具有最高 MLU 指标的链路集合 $\hat{s}$ 作为关键链路。

2.3 基于 GNN 的 DRL 实现

GraphSR 将域内路由决策视为马尔可夫决策过程 (MDP, Markov decision process) 的状态转移, 通过基于 GNN 的 DRL 算法, 动态调整虚拟链路属性 (时延) 来实现路由计算。

1) 状态空间

在传统 DRL 路由模型中, 以流量矩阵或链路容量矩阵的形式表示网络状态, 无法显式地学习网络拓扑信息, 也不符合 GNN 的输入格式。基于 GNN 的 DRL 架构如图5所示。本文将第 t 步的网络状态表示为 $s_t = [\mathbf{D}, \mathbf{H}, \text{edge_index}]$, 即包括节点信息 $\mathbf{D}$ 、链路信息 $\mathbf{H}$ 以及表示网络连接关系的边索引信息 edge_index 。以集群 C_i 为例, 节点信息 $\mathbf{D}$ 是一个维度为 $2|V_i| \times L$ 的矩阵, 包含网络流量需求信息

$$\mathbf{D} = [d_1, d_2, \dots, d_{|V_i|}]^T \quad (7)$$

其中, $|V_i|$ 是网络节点数, L 是每个节点的特征向量长度。 d_j 由流入和流出该节点的流量向量组成, 并通过 0 填充扩展到长度 L 。 d_j 表示如下

$$d_j = \begin{bmatrix} d_{j1}, d_{j2}, \dots, d_{j|V_i|}, \dots, d_{jL} \\ d_{1j}, d_{2j}, \dots, d_{|V_i|j}, \dots, d_{Lj} \end{bmatrix}_{2 \times L} \quad (8)$$

其中, 向量 $(d_{j1}, d_{j2}, \dots, d_{j|V_i|}) = \mathbf{TM}[j, :]$, 向量 $(d_{1j},$

$d_{2j} \dots, d_{|V_i|j}) = \mathbf{TM}[:, j]^T$ 。链路信息 $\mathbf{H}$ 是维度为 $|E_i| \times 3$ 的矩阵, $|E_i|$ 是网络链路数, 3维向量 (x_1, x_2, x_3) 包括链路时延、故障链路标志以及关键链路标志。边索引信息 $\mathbf{edge_index}$ 是维度为 $2 \times |E_i|$ 的矩阵, 每列中的两个元素表示网络中存在由这两个节点组成的链路, 表示如下

$$\mathbf{edge_index} = \begin{bmatrix} u_1, u_2, \dots, u_E \\ v_1, v_2, \dots, v_E \end{bmatrix} \quad (9)$$

其中, u_j, v_j 分别表示第 j 条边的起始和终止节点, $j=1, 2, \dots, E$ 。

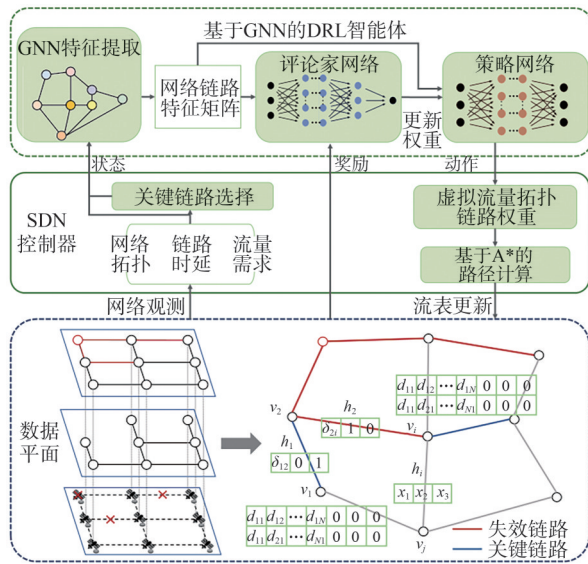


图5 基于GNN的DRL架构

2) 动作空间

在每个时间步, 根据 $a_t = \{a_1, a_2, \dots, a_\lambda\}$ 调整每个关键链路的虚拟权重, 在SDN控制器中形成用于路由计算的权重。其中, a_i 是第 i 个关键链路的虚拟调整比例, λ 是关键链路的数量。

3) 奖励函数

奖励函数包括3个方面, 分别为路由可靠性奖励 $R_{\text{reliability}}$ 、时延满足奖励 R_{delay} 以及链路容量溢出惩罚 R_{capacity} 的求和, 计算公式如下

$$R_{\text{reliability}}^t = \sum_{(i,j) \in V \times V} \prod_{(u,v) \in r_{ij}^t} (1 - p_{uv}^{\text{fail}}) \quad (10)$$

$$R_{\text{delay}}^t = - \sum_{(i,j) \in V \times V} \max \{0, D_{ij}^t - D_{ij}\} \quad (11)$$

$$R_{\text{capacity}}^t = - \sum_{(u,v) \in E} \max \left\{ 0, \sum_{(i,j) \in V \times V} y_{ij}(u, v, t_n) f_{ij}^{t_n} - b_{uv} \right\} \quad (12)$$

在基于GNN的DRL智能体中, GraphSR采用图卷积神经网络 (GCN, graph convolutional net-

works) 作为近端策略优化 (PPO, proximal policy optimization) 算法^[28]中策略网络和价值网络的特征提取器, 包括初始化、消息传递和读出3个阶段。在初始化阶段, 输入的 s_t 转换为特征向量 $\mathbf{X}^{(0)} = \phi(s_t)$, 然后通过聚合和更新的机制实现消息传递, 利用边特征嵌入更新节点特征, 直至各节点稳定。在迭代 l 次后, 消息表示为

$$a_v^{(l)} = \sum_{u \in \mathcal{N}(v)} M_t(\mathbf{W}^{(l)} \mathbf{h}_u^{(l-1)} + \mathbf{B}^{(l)} \mathbf{e}_{uv}^{(l-1)}) \quad (13)$$

$$\mathbf{h}_v^{(l)} = \sigma(\mathbf{h}_v^{(l-1)} + a_v^{(l)}) \quad (14)$$

其中, $a_v^{(l)}$ 为第 l 层节点 v 聚合得到的邻域消息特征, $\mathcal{N}(v)$ 为节点 v 的邻居节点集合, $\mathbf{W}^{(l)}$ 、 $\mathbf{B}^{(l)}$ 分别为第 l 层用于节点特征、边特征变换的可学习权重矩阵; $\mathbf{h}_u^{(l-1)}$ 、 $\mathbf{e}_{uv}^{(l-1)}$ 分别为上一层邻居节点 u 的特征、节点 u 与 v 间的边特征; $\mathbf{h}_v^{(l)}$ 、 $\mathbf{h}_v^{(l-1)}$ 分别为第 l 层与上一层节点 v 的自身特征, $\sigma(\cdot)$ 表示非线性激活函数。在读出阶段, 将图中所有节点的信息汇总成一个全图级别的表示, 通常采用最后一层节点特征的加权和, 表示为 $\mathbf{h}_G = \text{softmax}(\sum_{v \in V} \mathbf{h}_v^{(l)})$ 。

3 仿真结果

3.1 实验设置

本文使用卫星工具包 (STK, satellite tool kit)^[29]导出Starlink第一壳层星座的两行轨道数据 (TLE, two-line element set), 并采用StarPerf^[30]加载以构建网络拓扑。每个激光ISL的容量设置为20 Gbit/s。结合Starlink可用性地图和基于人口的重力模型生成流量需求。具体来说, 任意两颗卫星之间的通信需求根据其覆盖区域内的人口分布^[31]确定。这些需求基于可用性地图生成, 并且流量需求大小与覆盖区域内的人口规模成正比

$$\mathbf{TM}_{ij} = k \cdot \frac{p_i p_j}{l_{ij}^\phi} \quad (15)$$

其中, $\mathbf{TM}_{ij}$ 表示卫星节点 i, j 之间的业务流量需求大小, p_i, p_j 分别表示卫星节点 i, j 覆盖的人口数, l_{ij} 表示节点间物理距离, ϕ 表示距离衰减因子, k 表示调整系数, 用于控制流量平均水平。

为了反映现实流量的时间动态性, 本文对流量引入时间调制, 以模拟真实网络环境中的流量波动

$$\mathbf{TM}_{ij}(t) = \mathbf{TM}_{ij} \cdot (1 + \sin(2\pi f t)) \quad (16)$$

其中, f 为星座运行周期和地球自转周期的最小公倍数。

最终, 仿真得到 1 000 个流量需求矩阵用于模型的训练和测试。此外, 本文使用 PyTorch Geometric 实现 GCN 模型, 并在 OpenAI Gym 中构建相关的环境以训练和评估智能体。GCN 包含 2 个图卷积层和 1 个全连接层, 图卷积层输出通道分别为 64 和 128, 并集成了一个基于边特征的多层感知器 (MLP, multilayer perceptron), 通过全局平均池化聚合得到图级别特征。PPO 模型采用学习率为 3×10^{-4} 的 Adam 优化器进行参数更新, 输出层使用 Tanh 激活函数, 并设置裁减系数 ε 为 0.2, 用于控制策略更新的幅度。折扣因子 γ 设置为 0.98, 用于衡量未来奖励。将 GraphSR 与以下几种基线算法进行对比, 得出以下结论。

1) SPF 路由: 以链路时延作为权重, 通过计算网络中任意两点间的最短路径来确定数据包的传输路径。

2) 等价多路径 (ECMP, equal-cost multi-path): 将流量均匀分布在多个具有相同成本的等价路径上来分担网络负载, 提高吞吐量。

3) DRL-TE^[32]: 根据链路的剩余带宽和时延计算出多条预定义路径, 并使用 DRL 动态控制流量在这些候选路径之间的分配比例。

4) ScaleDRL^[26]: 根据拓扑中流的最短路径计算边中心性, 并识别出关键链路。采用 DRL 动态调整关键链路的权重, 并根据加权最短路径算法计算流量转发路径。

3.2 性能分析

GraphSR 将网络划分为多个 6×6 的域, 并针对链路故障率进行路由仿真实验, 使用标准化可靠性能比 (PR, performance ratio) $PR = P/P_{INLP}$ 作为评估指标^[33], 其中 P_{INLP} 为 INLP 计算的最优路由方案的路径可靠性。图 6 展示了在不同故障场景下各种路由方法的 PR 指标及标准差。由图 6 可知, 在 0%、5%、10% 和 15% 的链路失效情况下, GraphSR 的平均 PR 指标分别为 0.864、0.851、0.846 以及 0.841, 比基线方案中最好的结果分别高出 7.86%、7.04%、7.08% 和 10.94%, 能够保持较好的鲁棒性。ECMP 的整体性能优于 SPF, 通过多条路径分担流量, 在一定程度上降低了链路负载和失效概率。传统 DRL 方案 (DRL-TE 和 ScaleDRL) 的 PR 指标方差随着链路故障规模的增大而增大, 难以保持较好的路径可靠性。相反, GraphSR 通过构建虚拟流量

拓扑, 并使用 GNN 处理网络拓扑, 解决了模型输入限制的问题。同时, 面向故障场景的关键链路选择使得模型能够动态调整虚拟链路权重, 保持稳定的高路径可靠性。图 7 展示了在不同故障场景下各种路由方法的时延性能。与具有最短时延的 SPF 方法相比, GraphSR 在不同场景下的时延分别仅增加了 12%、22.58%、31.42% 和 42.5%。图 7 最后一项 α *SPF 展示了最大路径时延容忍度, 其在鲁棒空间段路由问题中被定义为最短路径时延的 1.5 倍。由此可知, 与其他算法相比, GraphSR 在保持高路径可靠性的同时, 实现了接近最优的时延表现, 具有可接受的时延。

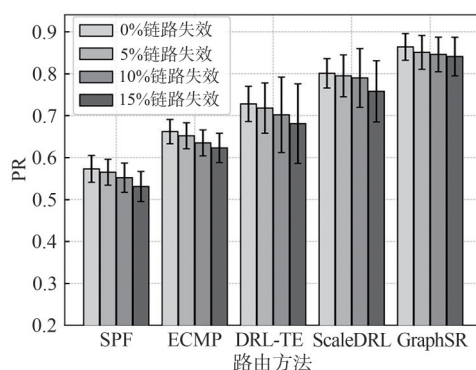


图6 在不同故障场景下各种路由方法的PR及标准差

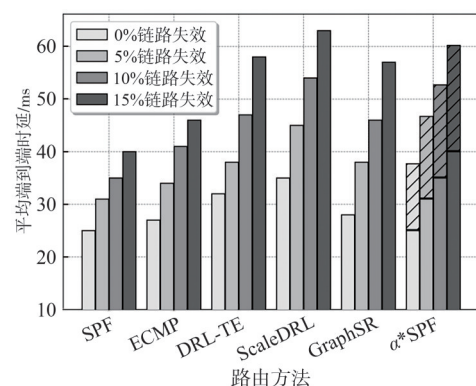


图7 在不同故障场景下各种路由方法的时延性能

图 8 展示了在不同规模拓扑下路由方法的 PR 指标。由图 8 可知, 当链路故障比例为 10% 时, DRL-TE 和 ScaleDRL 的 PR 指标随着网络规模的变化出现明显的下降, 分别下降了 8.13% 和 6.88%。相反, GraphSR 仅下降了 2.1%, 表明它具有动态适应网络变化的能力, 无须重新训练模型。

关键链路选择的消融实验如图 9 所示。设置链路选择比例 λ 为 15%, 图 9(a) 展示了不同链路选择方法下 PR 指标的箱型图。由图 9(a) 可知,

GraphSR 实现了最优的结果, 采用基于链路负载的选择概率, 构建关键故障场景, 合理利用了网络拓扑和流量信息。而随机选择、基于链路中心性以及基于节点度的选择方法受链路故障影响较大, 平均性能较低同时含有少量离群点。图9(b)展示了在不同 λ 下奖励随训练步数变化的曲线。由图9(b)可知, 较大的 λ 表现出较大的波动幅度, 说明较高的动作输出维度使得模型难以有效学习和收敛。

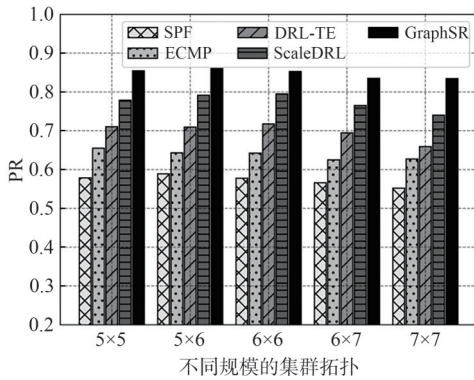
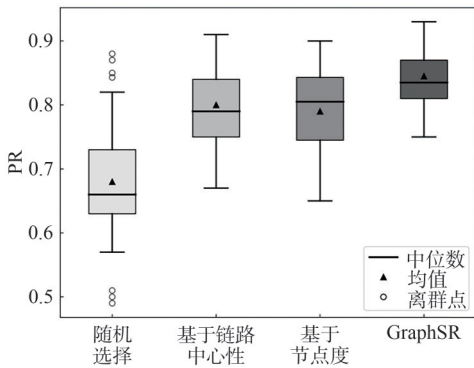
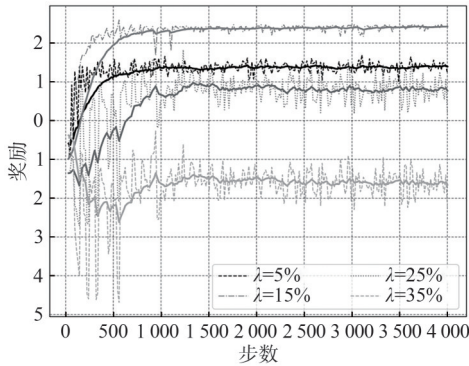


图8 不同规模拓扑下路由方法的PR指标



(a) 不同链路选择方法



(b) 不同链路选择比例

图9 关键链路选择的消融实验

4 结束语

本文围绕LEO巨型星座, 提出了一种基于图强化学习的鲁棒、可扩展路由框架GraphSR, 以应对大规模网络拓扑和各种网络故障, 将路由指定为INLP问题, 并尝试以分布式方式简化路由控制。具体来说, GraphSR将巨型星座划分为多个矩形控制域, 构建虚拟流量拓扑以屏蔽网络故障, 并设计了基于GNN的DRL算法用于控制网络关键链路。在Starlink星座拓扑上的仿真结果表明, GraphSR在不同链路故障场景中保持较高的路径可靠性, 同时具有可接受的时延。

参考文献:

- [1] Liu J J, Shi Y P, Fadlullah Z M, et al. Space-air-ground integrated network: a survey[J]. IEEE Communications Surveys & Tutorials, 2018, 20(4): 2714-2741.
- [2] Yue P Y, An J P, Zhang J K, et al. Low earth orbit satellite security and reliability: issues, solutions, and the road ahead[J]. IEEE Communications Surveys & Tutorials, 2023, 25(3): 1604-1652.
- [3] Yan Y J, Han G J, Xu H H. A survey on secure routing protocols for satellite network[J]. Journal of Network and Computer Applications, 2019, 145: 102415.
- [4] 张更新, 揭晓, 曲至诚. 低轨卫星物联网的发展现状及面临的挑战[J]. 物联网学报, 2017, 1(3): 6-9.
Zhang G X, Jie X, Qu Z C. Development status and challenges of LEO IoT[J]. Chinese Journal on Internet of Things, 2017, 1(3): 6-9.
- [5] Liu M Y, Quan W, Liu Z R, et al. Combating eavesdropping with resilient multipath transmission for space/aerial-assisted IoT[C]// Proceedings of the ICC 2022-IEEE International Conference on Communications. Piscataway: IEEE Press, 2022: 2230-2235.
- [6] Tedeschi P, Sciancalepore S, Di Pietro R. Satellite-based communications security: a survey of threats, solutions, and research challenges[J]. Computer Networks, 2022, 216: 109246.
- [7] Scharnowski T, Buchmann F, Wörner S, et al. A case study on fuzzing satellite firmware[C]// Proceedings of the 2023 Workshop on Security of Space and Satellite Systems. Internet Society, 2023.
- [8] Papa A, de Cola T, Vizarreta P, et al. Design and evaluation of reconfigurable SDN LEO constellations[J]. IEEE Transactions on Network and Service Management, 2020, 17(3): 1432-1445.
- [9] Shi Y P, Cao Y R, Liu J J, et al. A cross-domain SDN architecture for multi-layered space-terrestrial integrated networks[J]. IEEE Network, 2019, 33(1): 29-35.
- [10] Jiang W W. Software defined satellite networks: a survey[J]. Digital Communications and Networks, 2023, 9(6): 1243-1264.

- [11] Shi Y J, Yuan Z W, Zhu X R, et al. An adaptive routing algorithm for inter-satellite networks based on the combination of multipath transmission and Q-learning[J]. *Processes*, 2023, 11(1): 167.
- [12] Liu J H, Zhao B K, Xin Q, et al. DRL-ER: an intelligent energy-aware routing protocol with guaranteed delay bounds in satellite mega-constellations[J]. *IEEE Transactions on Network Science and Engineering*, 2021, 8(4): 2872-2884.
- [13] Huang Y H, Jiang X F, Chen S W, et al. Pheromone incentivized intelligent multipath traffic scheduling approach for LEO satellite networks[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(8): 5889-5902.
- [14] Lyu Y F, Hu H, Fan R F, et al. Dynamic routing for integrated satellite-terrestrial networks: a constrained multi-agent reinforcement learning approach[J]. *IEEE Journal on Selected Areas in Communications*, 2024, 42(5): 1204-1218.
- [15] Valadarsky A, Schapira M, Shahaf D, et al. Learning to route[C]// *Proceedings of the 16th ACM Workshop on Hot Topics in Networks*. New York: ACM Press, 2017: 185-191.
- [16] Pei X L, Sun P H, Hu Y X, et al. Enabling efficient routing for traffic engineering in SDN with deep reinforcement learning[J]. *Computer Networks*, 2024, 241: 110220.
- [17] Almasan P, Suárez-Varela J, Rusek K, et al. Deep reinforcement learning meets graph neural networks: exploring a routing optimization use case[J]. *Computer Communications*, 2022, 196: 184-194.
- [18] Huang Y X, Yang D, Feng B H, et al. A GNN-enabled multipath routing algorithm for spatial-temporal varying LEO satellite networks[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(4): 5454-5468.
- [19] Shi Y J, Wang W A, Zhu X R, 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.
- [20] Bhattacharjee D, Singla A. Network topology design at 27 000 km/hour[C]// *Proceedings of the 15th International Conference on Emerging Networking Experiments and Technologies*. New York: ACM Press, 2019: 341-354.
- [21] O'Dwyer R. Musk confirms laser crosslinks for Starlink satellite network[EB]. (2021-01-27) [2024-09-01].
- [22] 滕颖蕾, 李鑫, 王剑, 等. 基于SDN的铁路系统空天地融合网络架构[J]. *物联网学报*, 2020, 4(3): 30-41.
Teng Y L, Li X, Wang J, et al. SDN enabled space-terrestrial integrated network architecture of railway system[J]. *Chinese Journal on Internet of Things*, 2020, 4(3): 30-41.
- [23] Zhang L, Du Y. Cascading failure model and resilience enhancement scheme of space information networks[J]. *Reliability Engineering & System Safety*, 2023, 237: 109379.
- [24] Candra A, Budiman M A, Hartanto K. Dijkstra's and A-star in finding the shortest path: a tutorial[C]// *Proceedings of the 2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA)*. Piscataway: IEEE Press, 2020: 28-32.
- [25] Ye M H, Hu Y, Zhang J J, et al. Mitigating routing update overhead for traffic engineering by combining destination-based routing with reinforcement learning[J]. *IEEE Journal on Selected Areas in Communications*, 2022, 40(9): 2662-2677.
- [26] Sun P H, Guo Z H, Lan J L, et al. ScaleDRL: a scalable deep reinforcement learning approach for traffic engineering in SDN with pinning control[J]. *Computer Networks*, 2021, 190: 107891.
- [27] Henri Y. The OneWeb satellite system[M]// Pelton J N, Madry S. *Handbook of Small Satellites: Technology, Design, Manufacture, Applications, Economics and Regulation*. Cham: Springer International Publishing, 2020: 1091-1100.
- [28] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[PP]. V2. arXiv(2017-08-28) [2025-02-22]. arXiv: 1707.06347.
- [29] 谢健骊, 陈龙, 张泽鹏, 等. 基于位置预测模型的空天地一体化网络切换算法[J]. *通信学报*, 2024, 45(12): 162-178.
Xie J L, Chen L, Zhang Z P, et al. Handover algorithm for space-air-ground integrated network based on location prediction model[J]. *Journal on Communications*, 2024, 45(12): 162-178.
- [30] Lai Z Q, Li H W, Li J H. StarPerf: characterizing network performance for emerging mega-constellations[C]// *Proceedings of the 2020 IEEE 28th International Conference on Network Protocols (ICNP)*. Piscataway: IEEE Press, 2020: 1-11.
- [31] Center for International Earth Science Information Network (CIESIN), Columbia University. Gridded population of the world, version 4 (GPWv4): population count, revision 11[DS]. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC), 2018. (2018-12-31)[2024-09-17].
- [32] Xu Z Y, Tang J, Meng J S, et al. Experience-driven networking: a deep reinforcement learning based approach[C]// *Proceedings of the IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. Piscataway: IEEE Press, 2018: 1871-1879.
- [33] Ye M H, Zhang J J, Guo Z H, et al. FlexDATE: flexible and disturbance-aware traffic engineering with reinforcement learning in software-defined networks[J]. *IEEE/ACM Transactions on Networking*, 2023, 31(4): 1433-1448.

[作者简介]



白禄鑫(2000—), 男, 信息工程大学信息技术研究所硕士生, 主要研究方向为卫星互联网、软件定义网络和网络安全。



胡宇翔(1982-), 男, 博士, 信息工程大学教授, 主要研究方向为新一代互联网体系结构、网络通信安全等。



江逸茗(1984-), 男, 博士, 信息工程大学信息技术研究所副研究员, 主要研究方向为卫星互联网、虚拟化。



马海龙(1980-), 男, 博士, 信息工程大学信息技术研究所教授, 主要研究方向为网络路由技术、网络主动防御。



胡涛(1993-), 男, 博士, 信息工程大学信息技术研究所助理研究员, 主要研究方向为新型网络体系、网络主动防御。