

基于DQN的低轨卫星网络多目标智能路由算法

罗宗屹^{1,2}, 金世超^{1,2}, 董 涛^{1,2}, 殷 杰^{1,2}

(1. 航天恒星科技有限公司天地一体化信息技术国家重点实验室, 北京 100095;

2. 北京卫星信息工程研究所, 北京 100095)

摘 要: 近年来, 低轨卫星网络发展迅速, 被广泛应用于全球通信、互联网接入等领域。首先, 对当前低轨卫星网络发展现状与主流的低轨网络路由算法进行总结。然后, 针对星上路由算法设计面临的优化目标单一、收敛速度慢以及时延和丢包率等指标优化难度大等挑战, 提出基于深度Q网络(DQN)的多目标智能路由算法, 通过设计多目标奖励函数, 实现了时延、丢包率、负载均衡多目标的性能优化。在时延方面, 本算法优于最短通路优先算法; 在流量分布指数和丢包率方面, 本算法相较于最短通路优先算法与Dijkstra算法分别提高14%和10%以上。最后, 阐述所提智能路由算法在多种应用场景中的使用方法。

关键词: 低轨卫星网络; 路由算法; 深度Q网络

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-8930.2025003

Multi-Objective Intelligent Routing Algorithm for LEO Satellite Networks Based on DQN

LUO Zongyi^{1,2}, JIN Shichao^{1,2}, DONG Tao^{1,2}, YIN Jie^{1,2}

1. State Key Laboratory of Space-Ground Integrated Information Technology, Space Star Technology Co., Ltd., Beijing 100095, China

2. Beijing Institute of Satellite Information Engineering, Beijing 100095, China

Abstract: In recent years, LEO satellite networks have developed rapidly and are widely used in global communication and internet access. This paper first summarized the current development status of LEO networks and the main LEO network routing algorithms. At the same time, aimed at the challenges of single optimization objectives, slow convergence speed, and difficulty in optimizing indicators such as delay and packet loss, this paper proposed a multi-objective intelligent routing algorithm based on deep Q-network (DQN). With the designed DQN multi-objective reward function, it realized the optimization of delay, packet loss, load balancing. Regarding delay, the proposed algorithm is superior to the Shortest Path First (SPF) algorithm. In terms of traffic distribution index and packet loss rate, the performances of the proposed algorithm are improved by more than 14% and 10% compared with the SPF and Dijkstra algorithms respectively. In addition, this paper illustrated the application methods of the proposed intelligent routing algorithm in various scenarios.

Keywords: LEO satellite network, routing algorithm, DQN

0 低轨卫星网络发展现状

随着地面移动通信技术的迅速发展, 地面网络已经可以提供定制化、大带宽、低时延的服务。然而, 受限于地理位置和人口分布, 地面网络仅能覆盖约20%的陆

地范围, 占地表面积的6%^[1], 仅利用地面网络无法实现6G泛在通信的愿景。作为星地融合网络的重要组成部分, 卫星网络能够覆盖海洋、山川、森林、沙漠等难以依靠地面蜂窝基站提供网络服务的地区, 将在下一代无线通信系统中扮演重要的角色。其中, 根据卫星网络节点轨

收稿日期: 2024-11-11; 修回日期: 2025-02-10

基金项目: 国家重点研发计划资助项目 (No. 2022YFB2902504)

Foundation Item: National Key Research and Development Program of China (No. 2022YFB2902504)

道高度的不同，卫星网络可以分为低轨（Low Earth Orbit，LEO）卫星网络、中轨（Medium Earth Orbit，MEO）卫星网络和包含地球静止轨道（Geostationary Earth Orbit，GEO）卫星的高轨卫星网络。

近年来，低轨卫星网络发展迅速，与高轨卫星网络相比，具有低时延、低成本、弹性灵活等特点，在全球通信、互联网接入等方面具有广泛应用。20 世纪末，铱星、全球星等星座计划掀起了第一代低轨星座建设浪潮，但最终因卫星发射和运营成本过高，导致用户数量较少。近年来，随着卫星批量化制造、可重复回收发射技术的发展，星座发射、运营成本逐渐降低，以“星链（Starlink）”星座和“一网（OneWeb）”星座为代表的低轨卫星网络建设迎来了新的发展热潮。

其中，SpaceX 计划共计部署两代星座，采用多层多倾角架构。据 2021 年 8 月 SpaceX 向 FCC 提交的第二代星座修订申请，SpaceX 已放弃被批准的第一代星座中极低轨道（Very Low Earth Orbit，VLEO）星座计划。星链第一代星座参数见表 1。第二代星座将首选 29 988 颗卫星配置方案，具体见表 2，由轨道高度在 340~614 km 的多壳层卫星构成^[2]。

表 1 星链第一代星座参数

参数	星座 1	星座 2	星座 3	星座 4	星座 5	合计
卫星数/颗	1 584	720	348	1 584	172	4 408
轨道面数/个	72	36	6	72	4	190
每个轨道卫星数/颗	22	20	58	22	43	20~58
轨道高度/km	550	570	560	540	560	540~570
倾角	53°	70°	97.6°	53.2°	97.6°	53~97.2°

表 2 星链第二代星座参数

星座	轨道高度/km	倾角	轨道面数/个	每个轨道卫星数/颗	卫星总数/颗
星座 1	340	53°	48	110	5 280
星座 2	345	46°	48	110	5 280
星座 3	350	38°	48	110	5 280
星座 4	360	96.9°	30	120	3 600
星座 5	525	53°	28	120	3 360
星座 6	530	43°	28	120	3 360
星座 7	535	33°	28	120	3 360
星座 8	604	148°	12	12	144
星座 9	614	115.7°	18	18	324
合计	340~614	33°~148°	288	12~120	29 988

目前，星链用户数量超过了 300 万，截至 2024 年 6 月，卫星互联网服务已覆盖 100 个国家，预计将在 2025 年实现覆盖全球的语音、数据和卫星物联网服务的目标。截至 2024 年 8 月 10 日，Starlink 卫星升空总数达到 6 887 颗，其中 2024 年升空 1 227 颗。目前发射的 V2 Mini 卫星属于壳层 6 与壳层 7，连同先前发射的壳层 5，均属星链第二代星座。采用了轨道面内少量部署，进而加密面内卫星的分阶段部署方案^[3]。

OneWeb 第一代初始星座包括 648 颗卫星和 234 颗备份星，总数量达 882 颗。其中，648 颗卫星分布在轨道高度 1 200 km、倾角 87.9°的 18 个轨道面，每个轨道面部署约 36 颗卫星，工作在 Ku/Ka 频段。经过多次调整，按照 2021 年 1 月 OneWeb 公司向 FCC 提交的申请，目前的 OneWeb 星座布局应该包括 6 372 颗 LEO 卫星和 1 280 颗中轨卫星，具体情况见表 3^[4]。

表 3 OneWeb 星座参数

轨道高度/km	倾角	轨道面数/个	每个轨道卫星数/颗	各倾角卫星数/颗	卫星总数合计/颗
1 200	87.9°	36	49	1 764	6 372
1 200	55.0°	32	72	2 304	
1 200	40.0°	32	72	2 304	
8 500	—				1 280

OneWeb 已经完成第一代星座（全球宽带星座）的发射任务，在轨卫星总数 618 颗，其中 588 颗卫星用于全球覆盖，并计划增加额外的卫星以实现弹性和冗余。OneWeb 正逐步推广全球商业服务，在欧洲市场上，OneWeb 及其服务提供商 Gogo 商务航空，将开通 LEO 卫星通信服务。

亚马逊致力打造的全球性宽带网络 Kuiper 星座由 3 236 颗卫星组成，包括轨道高度为 590 km、倾角为 33°的 784 颗，轨道高度为 610 km、倾角为 42°的 1 296 颗，轨道高度为 630 km、倾角为 51.9°的 1 156 颗。其中，两颗原型卫星 Kuipersat-1 与 Kuipersat-2 于 2023 年上半年发射。目前，亚马逊已完成两颗原型卫星 Kuipersat-1 与 Kuipersat-2 的发射与多项测试，并将开始提供宽带服务。

综上，随着低轨卫星星座规模的快速增长和通信服务的迅速铺开，为确保低轨卫星星座网络效能的充分发挥，需综合考虑并提升低轨卫星网络在时延、丢包、负载均衡等方面的性能，在低轨卫星网络多项关键技术中，路由算法在提升服务质量方面发挥了重要的作用。

1 低轨卫星网络路由算法

低轨星座规模巨大、网络拓扑时变且受到卫星计算资源受限、全局路由计算评估开销大、用户分布不均匀等因素影响,低轨卫星网络路由算法设计面临优化目标单一、收敛速度慢和时延、丢包等指标优化难度大等诸多挑战。

目前传统的低轨卫星网络路由算法可分为面向虚拟拓扑的路由算法和面向虚拟节点的路由算法^[5]。其中,虚拟拓扑算法是将网络拓扑在时间上划分成一系列的时间片,将每个时间片内的网络拓扑视为静态,分别针对每个时间片计算路由。Werner 等^[6]首次提出了基于虚拟拓扑的路由方法。Chang 等^[7]将卫星网络建模为有限状态自动机,将卫星网络运行周期划分为若干个状态,每个状态作为固定的网络拓扑片。Tan 等^[8]在虚拟拓扑快照的基础上添加了动态检测机制,提高了算法的鲁棒性和适应性。但是随着卫星星座规模的不断增加,上述算法需要大量的计算与存储资源以维护庞大的路由表,同时无法较好地适应网络动态变化,实际部署难度较大。

面向虚拟节点的路由算法基于虚拟节点构造拓扑。虚拟节点的位置固定不变,虚拟节点与物理卫星的映射关系发生改变,实际卫星网络拓扑抽象为静态的拓扑以计算路径。DRA 算法^[9]是最早采用虚拟节点概念的路由算法,该算法将分组传输时延作为优化目标。Liu 等^[10]在虚拟节点的基础上,根据当前节点和目的节点的经纬度选择传播时延最小的下一跳节点,并考虑了交换队列。基于虚拟节点的路由算法减小了卫星网络拓扑动态变化对路由协议的影响,降低了路由计算和存储开销,但是当网络中的节点或链路失效时,路由表不能快速更新,网络易发生拥塞。

随着人工智能技术的快速发展,智能路由算法已成为当前低轨卫星星座路由计算的研究热点。内生智能是未来网络技术的核心特征,基于人工智能的路由算法能够显著提升网络管理效能。与传统路由算法相比,基于人工智能的路由算法能适配多种网络架构,适应复杂的低轨卫星网络环境,根据实时网络状况动态实现最优路径选择。因此,对于大规模低轨卫星星座的复杂网络场景,智能路由算法将能够发挥重要作用。

随着人工智能技术的快速发展,基于深度学习和强化学习的路由算法已成为当前研究的热点。深度学习对大量数据进行学习并进行拟合,例如,Na 等^[11]提出了一种基于极限学习机 (Extreme Learning Machine, ELM)

的分布式路由策略,采用 ELM 对卫星节点的流量进行预测,并以流量预测为基础,引入移动代理同时独立地搜索 LEO 卫星网络,确定路由信息。但深度学习的路由算法,严重受限于训练样本的数量和质量。强化学习的主要特点是通过智能体选择不同的行为来获取最大的奖励。例如,Wang 等^[12]设计了一种基于强化学习的路由算法,对 Q 学习算法进行了改进,通过限制算法的跳数和采用动态贪婪系数,改进了传统算法收敛时间过长的問題。但是,复杂的网络环境容易导致强化学习难以收敛或者不稳定。

近年来,深度强化学习^[13]作为一种新型的人工智能方法被广泛应用于网络领域。You 等^[14]提出了一种基于多智能体深度强化学习的数据包路由框架,利用长短期记忆网络从数据包和过去动作信息中提取路由特征,以不断逼近 Q 学习中的 Q 函数。Bhavanasi 等^[15]提出了一种基于深度 Q 网络 (Deep Q-Network, DQN)^[16]的多智能体路由算法,每个地面节点作为智能体,根据状态信息进行路由决策。但是,上述基于深度强化学习的智能路由算法没有综合考虑时延、丢包率、负载均衡,在训练时通常只对部分指标进行单独优化。同时,参考文献^[15]提出的多智能体路由算法重点关注地面网络,在训练时根据下一跳节点与目的节点间的距离制定奖励函数,容易陷入局部最优,难以适应时延敏感业务在卫星网络内端到端低时延、低丢包率的要求。

2 算法设计

针对上述问题,本文提出了一种基于 DQN 的低轨卫星网络多目标智能路由算法,对 DQN 的奖励函数进行综合优化设计,加入时延、丢包率、负载均衡等相关指标,同时对环路、拥塞丢包等情况进行单独惩罚处理,在降低时延和丢包率的同时,提升了负载均衡性能。

2.1 场景建模

为了简化仿真分析,本文构造了一个每条轨道升交点赤经均匀分布的 Walker 星座,共有 12 个轨道,每个轨道面 11 颗卫星。每颗卫星都有星上路由处理功能。此外,每颗卫星有 4 条星间链路,包括 2 条轨间链路和 2 条轨道内链路,分别连接 2 颗同轨道面卫星和 2 颗相邻轨道面卫星。星座中的每颗卫星周期性地向其邻居节点发送 HELLO 数据包,并根据接收到的 HELLO 数据包实时更新路由表,确定其相邻的卫星节点。

由于卫星运动具有周期性和可预测性,本文将卫星星座动态拓扑结构离散划分为若干时间片,在时间片内

网络拓扑均为静态。每个时间片的网络拓扑可表示为一个无向图 $G = (V, E)$ ，其中 $V = \{v_1, v_2, \dots, v_n\}$ 表示网络中每个卫星节点， $E = \{(v_1, v_2), \dots, (v_i, v_j)\}$ 表示卫星节点之间的链路，并且 n, i, j 表示卫星的编号。低轨卫星网络场景示意如图1所示。

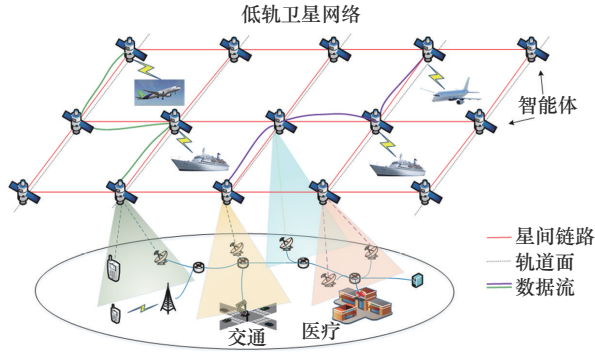


图1 低轨卫星网络场景示意

在网络时延模型中，本文考虑了传播时延和排队时延。对于排队时延，本文引入 $M/M/1$ ^[17] 排队模型。传播时延则由卫星节点间的距离计算，公式如下

$$\tau_{v_i, v_j} = \frac{d_{v_i, v_j}}{c}, \forall (v_i, v_j) \in E \quad (1)$$

其中， d_{v_i, v_j} 代表节点 v_i 与节点 v_j 之间的距离， c 代表链路上传输速度，即光速。

2.2 算法架构

本文提出的算法架构如图2所示，其主要包括目标网络 $Q(s, a; w^-)$ 、决策网络 $Q(s, a; w)$ 、经验回放缓冲区和

网络环境4部分。其中，目标网络与决策网络是两个参数不同但结构相同的神经网络；经验回放缓冲区负责存储在每次行动后获取的经验，一条经验由 (s_t, a_t, r_t, s_{t+1}) 来表示，其组成元素如下。

状态 s_t ： s_t 为状态信息，它是智能体决策的基础。本文中其包括相邻卫星的逻辑编号、相邻节点拥塞情况、源节点与目的节点编号。

动作 a_t ： a_t 为状态 s_t 下智能体要执行的动作，其根据决策网络的输出得出，即选择哪颗相邻卫星作为下一跳节点，并向其转发数据。

奖励 r_t ： r_t 为状态 s_t 下执行动作 a_t 得到的奖励，可通过奖励函数计算得到。

状态 s_{t+1} ：表示把动作 a_t 输入环境后获得的新状态。

本文将每颗卫星看作一个单独的智能体，每个智能体根据从环境中得到的状态信息进行分布式自主路由决策。在每个时刻，智能体与环境交互得到一个高维度的观察，对观察进行处理以得到具体的状态特征表示，然后通过决策网络以及相应策略将当前状态映射为相应的动作，并通过基于预期回报的奖励函数来评价该动作。环境对此动作做出反应，并得到下一个观察。通过不断循环以上过程，最终可以得到实现目标的最优策略。具体的，环境包括网络拓扑结构、拥塞情况、时延与排队模型、丢包、负载均衡等全方位的网络情况，每个卫星智能体仅从环境中获得相邻卫星的逻辑编号、相邻节点拥塞情况、源节点与目的节点编号状态3种状态信息，将3种状态信息处理后输入决策网络中，决策网络根据

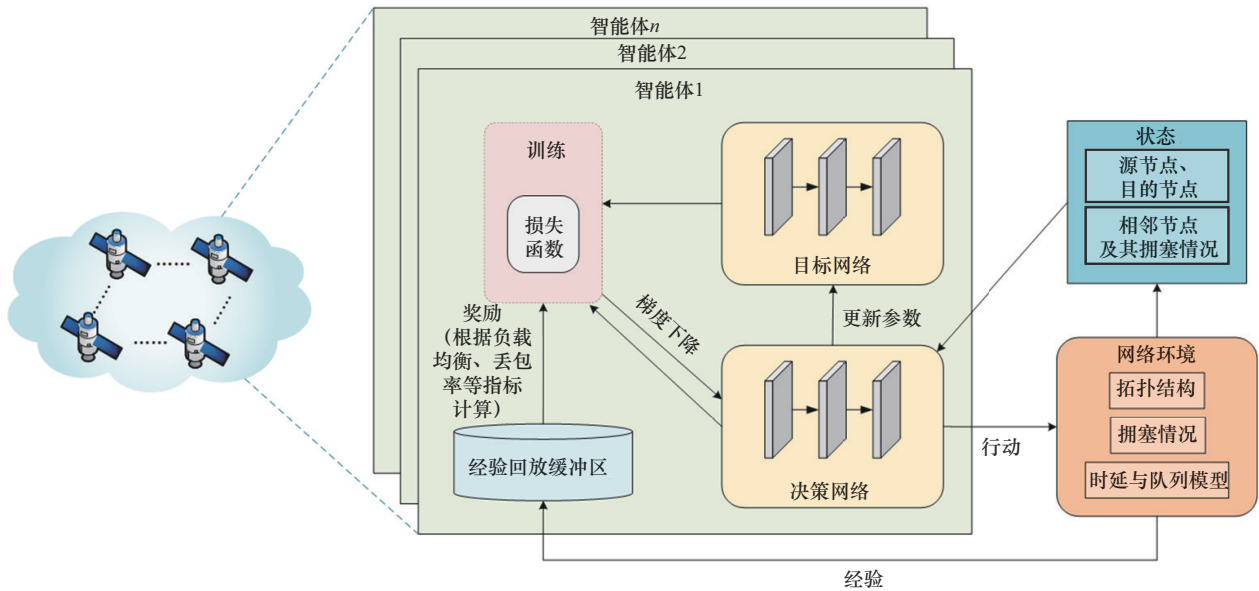


图2 算法架构

输出结果选择概率最大的动作，环境根据当前动作的选路情况与全局的网络信息通过奖励函数计算奖励，并在智能体执行动作向选择的节点转发数据后返回新的状态。

为了在训练过程中智能体能够综合考虑网络时延和负载均衡情况，更好地实现路由选择的多目标优化，本文将时延、丢包率、流量分布指数等相关指标加入 DQN 的奖励函数中。此外，本文在 DQN 训练过程中对出现环路和拥塞丢包的情况增加了额外的惩罚。奖励函数的定义如下

$$r_t = \begin{cases} \beta \cdot B + \theta, & l \notin S \text{ 且 } l \in L_{\text{right}} \\ T + \lambda, & l \notin S \text{ 且 } l \notin L_{\text{right}} \\ P + \eta, & l \in L_{\text{loss}} \\ -2, & l \in S \end{cases} \quad (2)$$

其中， B 表示流量分布指数， P 表示丢包率， T 表示当前卫星节点到目的节点的传输时延与下一跳卫星节点到目的节点传输时延的差值，在计算奖励函数前本文对上述 3 项指标进行归一化处理且对丢包率取负值。 β 、 θ 、 λ 、 η 为不同的权重， $\beta + \theta = 1$ 。 $l \in S$ 表示当前路径出现了环路， $l \in L_{\text{right}}$ 表示选择了通过 K 最短路径算法计算出的路径， $l \in L_{\text{loss}}$ 表示遇到拥塞发生了丢包。

在训练过程中，本文采用 K 最短路径算法进行路径计算。为了进一步考虑排队时延的影响，本文在使用 K 最短路径算法计算最短路径时，既要考虑节点间的传播时延，又要考虑下一跳节点的排队时延。

本文对 DQN 模型训练过程如下。

(1) 初始化网络环境、目标网络和决策网络。

(2) 随机选取源节点和目的节点。

(3) 将状态 s_t 输入决策网络中，决策网络根据 ϵ -贪心策略计算出动作 a_t ，目的是对探索和利用进行平衡，在训练初期让智能体尽可能多地探索可选路由，并且随着训练步数的增加逐渐减少探索率，增加利用已学习到的知识进行决策的可能性。 ϵ -贪心策略如式(3)所示

$$a_t = \begin{cases} \arg\max_a Q(s_t, a_t; w), & p \in [0, 1 - \epsilon] \\ \text{随机选择}, & p \in (1 - \epsilon, 1] \end{cases} \quad (3)$$

其中， p 是每次选择动作时生成的随机数，取值在 0 到 1 之间。参数 ϵ 的计算公式为

$$\epsilon = \epsilon_{\text{end}} + (\epsilon_{\text{start}} - \epsilon_{\text{end}}) \times e^{\frac{s}{\epsilon_{\text{decay}}}} \quad (4)$$

其中， ϵ_{start} 为 ϵ 的初值， ϵ_{end} 为 ϵ 的最小值， ϵ_{decay} 为衰减因子， s 为训练步数， ϵ 在实际训练中随着训练步数的增加逐渐递减。

(4) 在网络环境中执行动作 a_t ，计算奖励 r_t ，并返回新的状态 s_{t+1} ，将该条经验 (s_t, a_t, r_t, s_{t+1}) 存入经验回放缓冲区，从经验回放缓冲区中随机取出一批数据，通过目标网络与决策网络计算损失函数，通过梯度下降法对决策网络的参数 w 进行更新。

(5) 判断数据包是否到达目的节点或者发生了丢包；若数据包成功到达目的节点或者发生了丢包，将决策网络的参数 w 赋给目标网络。否则返回步骤 3，继续选择下一跳卫星节点。

(6) 重复步骤 2 到步骤 5，直至网络逐渐收敛。

3 仿真实验

本文选择最短通路优先 (Shortest Path First, SPF) 算法和 Dijkstra 算法作为基线算法与本文提出的算法进行比较。SPF 算法选择从源节点到目的节点跳数最少的路径作为最短路径，Dijkstra 算法基于全局信息计算源节点到目的节点的最短传播时延路径。本文在仿真实验时，将正常节点转发数据最大速率设置为 2 Mbit/s，网络负载设置为节点接收数据的速率，当网络负载超过节点转发数据的最大速率并逐渐增大时，节点易发生拥塞并丢包。本文从网络中随机选择源节点和目的节点，将数据包从源节点发出，直到数据包到达目的节点或发生丢包为止，并据此统计网络的各项评价指标。

本文从以下 3 个评价指标进行性能评估。

时延：每个数据包从源节点到目的节点的总时延的平均值，包含传播时延和排队时延。如果数据包发生丢包，则不统计该数据包所经过的路径的时延。

丢包率：未到达目的节点的数据包占总发送数据包的比例。

流量分布指数^[18]：表示网络的负载均衡状况，该指标范围从 0 到 1，指数越大网络的负载均衡性能越好，其定义如下

$$B = \frac{\left(\sum_{i=1}^q x_i^2 \right)}{q \sum_{i=1}^q x_i^2} \quad (5)$$

其中， x_i 为当前星间链路上实际通过数据包的个数， q 为星间链路的数量。

不同网络负载下的时延如图 3 所示，当网络负载较轻时，本算法在时延性能方面优于 SPF 算法，在网络负载为 1.4 Mbit/s 时提高了 4%。并且，在不考虑数据包丢包后进行重传的情况下，本算法与 Dijkstra 算法基本接近，时

延性能在一定程度上逼近了最优解,同时,与Dijkstra算法相比,本算法不需要全局拓扑信息,仅根据邻居节点的信息即可进行分布式自主决策。随着网络负载的不断增加,为了避免拥塞,本算法会选择其他路径,导致在网络负载超过2.2 Mbit/s时,时延略高于其他两种算法。时延分布函数如图4所示,时延在0~100 ms时,本算法时延分布情况与Dijkstra算法基本接近;时延在100~180 ms时,本算法时延分布略低于Dijkstra算法,但全面优于SPF算法。

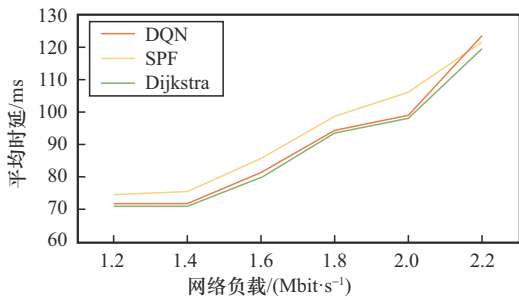


图3 不同网络负载下的时延

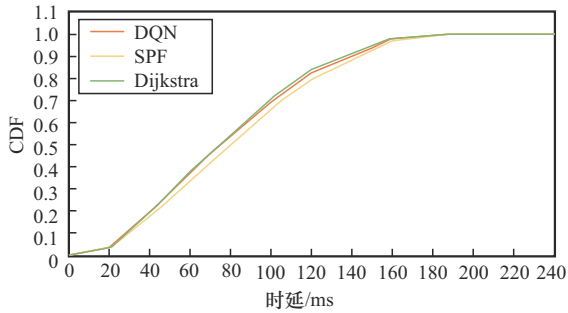


图4 时延分布函数

如图5所示,在丢包率方面,本算法明显低于其他两种算法。在网络负载为2.4 Mbit/s与2.6 Mbit/s时,本算法较其他两种算法至少提高46%与40%。实验结果表明,随着网络负载的增加,本算法能够动态适应网络的变化,根据相邻节点的状态信息选择未发生拥塞的下一跳节点。因此,同SPF算法和Dijkstra算法相比,本算法在避免拥塞、降低丢包率方面更加有效。

如图6所示,该算法在网络拥塞程度增加时也能较好地提升负载均衡性能,能够灵活地自主选择其他路径,自主调节网络负载。特别是在高负载的情况下,智能体会为了避免拥塞而选择其他路径,进一步优化了负载均衡性能,例如流量分布指数在网络负载为2.4 Mbit/s与2.6 Mbit/s时本算法较其他两种算法至少提高22%与30%。

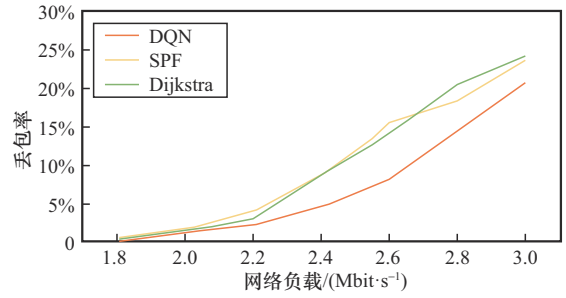


图5 不同网络负载下的丢包率

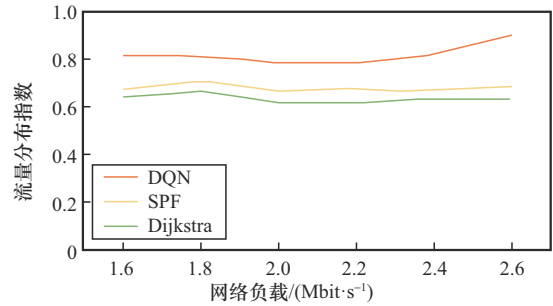


图6 不同网络负载下的流量分布指数

4 应用场景

除了大规模低轨卫星星座,基于DQN的多目标智能路由算法还能够高通量卫星网络及高低轨混合网络中应用,具体应用场景和应用方法如下。

(1) 高通量卫星网络。相比传统卫星,高通量卫星一般采用Ka、Ku等高频段,结合多点波束、频率时隙复用等技术,实现广阔区域连续覆盖。高通量卫星网络主要由高通量卫星、信关站、地面管控中心、各类型机动/固定终端及附属网络组成。可在地面管控中心及终端设备中部署智能体,同时利用网络拓扑信息搭建上文所述网络环境。网络环境部署地面管控中心及终端设备,并与设备中的智能体进行交互。当网络拓扑发生变化时,网络环境自主进行更新,并对智能体进行持续训练,使智能体能够通过在线学习的方式不断更新模型以保证路由选择的准确性。

(2) 高低轨混合网络。当前,传统高轨卫星通信运营商面临网络时延大、两极存在盲区等多种挑战,低轨卫星网络则面临系统容量整体利用效率较低、组网复杂等问题,单一的网络目前已无法同时满足用户对高可靠、大带宽、高速率、低费用、全域海量接入的需求。高低轨混合组网、多星座互补并存的新业态正在逐步形成,并成为未来的重要发展趋势,智能路由算法同样能够提

升高低轨混合网络效能。在高低轨混合网络中,智能体可部署在高轨卫星、低轨卫星、信关站、地面管控中心等核心节点上,并在每个核心节点上搭建网络环境。控制平面的各个控制器节点负责感知所在控制域的网络环境,并通过不断交互实时更新全网网络环境。在网络环境更新后,各个控制器节点将环境信息下发至控制域内其他节点,在网络环境发生改变时,高低轨混合网络中各个节点均能够通过在线学习的方式对智能体不断训练,保证智能体模型实现快速更新收敛,减少网络变化对路由选择带来的影响。

5 结束语

随着低轨卫星网络的建设与通信技术的不断发展,高效灵活智能的低轨卫星网络路由技术已成为当前研究的热点。针对现有智能路由算法在低轨卫星网络路由计算中的时延、丢包、负载均衡多目标优化问题,本文提出了基于 DQN 的低轨卫星网络多目标智能路由算法,通过设计基于 DQN 的多目标奖励函数,实现了针对时延、丢包、负载均衡等多目标的性能优化。在时延方面,本算法性能优于 SPF 算法,接近最短时延算法;在负载均衡和拥塞避免方面,本算法性能相较于 SPF 算法与 Dijkstra 算法分别提高 14% 和 10% 以上。此外,在低轨卫星网络基础上,本文针对高通量卫星网络和高低轨混合网络两种不同应用场景,分析并阐述了智能路由算法在各个场景中的使用方法。

参考文献:

- [1] CHEN S Z, SUN S H, KANG S L. System integration of terrestrial mobile communication and satellite communication: the trends, challenges and key technologies in B5G and 6G[J]. China Communications, 2020, 17(12): 156-171.
- [2] 刘帅军, 徐帆江, 刘立祥, 等. Starlink 星座卫星入轨预测与仿真分析[EB]. 2023.
LIU S J, XU F J, LIU L X, et al. Orbit prediction and simulation analysis of Starlink constellation satellites[EB]. 2023.
- [3] 刘帅军, 徐帆江, 刘立祥, 等. Starlink 二代 GEN2 系统混合组网与分阶段部署探讨[EB]. 2023.
LIU S J, XU F J, LIU L X, et al. Discussion on hybrid networking and phased deployment of Starlink second generation GEN2 system[EB]. 2023.
- [4] 王学宇, 武坦然. OneWeb 低轨道卫星系统及其军事应用分析[J]. 航天电子对抗, 2022, 38(4): 59-64.
WANG X Y, WU T R. Analysis on OneWeb LEO satellite system and its military application[J]. Aerospace Electronic Warfare, 2022, 38(4): 59-64.
- [5] QI X G, MA J L, WU D, et al. A survey of routing techniques for satellite networks[J]. Journal of Communications and Information Networks, 2016, 1(4): 66-85.
- [6] WERNER M, DELUCCHI C, VOGEL H J, et al. ATM-based routing in LEO/MEO satellite networks with intersatellite links[J]. IEEE Journal on Selected Areas in Communications, 1997, 15(1): 69-82.
- [7] CHANG H S, KIM B W, LEE C G, et al. FSA-based link assignment and routing in low-earth orbit satellite networks[J]. IEEE Transactions on Vehicular Technology, 1998, 47(3): 1037-1048.
- [8] TAN H C, ZHU L D. A novel routing algorithm based on virtual topology snapshot in LEO satellite networks[C]//Proceedings of the 2014 IEEE 17th International Conference on Computational Science and Engineering. Piscataway: IEEE Press, 2014: 357-361.
- [9] EKICI E, AKYILDIZ I F, BENDER M D. A distributed routing algorithm for datagram traffic in LEO satellite networks[J]. IEEE/ACM Transactions on Networking, 2002, 9(2): 137-147.
- [10] LIU X M, JIANG Z Q, LIU C H, et al. A low-complexity probabilistic routing algorithm for polar orbits satellite constellation networks[C]//Proceedings of the 2015 IEEE/CIC International Conference on Communications in China (ICCC). Piscataway: IEEE Press, 2015: 1-5.
- [11] NA Z Y, PAN Z, LIU X, et al. Distributed routing strategy based on machine learning for LEO satellite network[J]. Wireless Communications and Mobile Computing, 2018(1): 3026405.
- [12] WANG X T, DAI Z Q, XU Z. LEO satellite network routing algorithm based on reinforcement learning[C]//Proceedings of the 2021 IEEE 4th International Conference on Electronics Technology (ICET). Piscataway: IEEE Press, 2021: 1105-1109.
- [13] MNIH V. Playing atari with deep reinforcement learning[J]. arxiv preprint arxiv, 2013.
- [14] YOU X Y, LI X J, XU Y D, et al. Toward packet routing with fully distributed multiagent deep reinforcement learning[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 52(2): 855-868.
- [15] BHAVANASI S S, PAPPONE L, ESPOSITO F. Routing with graph convolutional networks and multi-agent deep reinforcement learning[C]//Proceedings of the 2022 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN). Piscataway: IEEE Press, 2022: 72-77.
- [16] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.
- [17] SHORTLE J F, THOMPSON J M, GROSS D, et al. Fundamentals of queueing theory[M]. Hoboken: John Wiley & Sons, 2018.
- [18] TALEB T, MASHIMO D, JAMALIPOUR A, et al. Explicit load balancing technique for NGEOS satellite IP networks with on-board processing capabilities[J]. IEEE/ACM Transactions on Networking, 2009, 17(1): 281-293.

[作者简介]



罗宗屹（1998-），男，硕士生，主要研究方向为卫星网络。



董涛（1975-），男，博士，航天恒星科技有限公司研究员，主要研究方向为卫星通信和网络。



金世超（1980-），男，博士，航天恒星科技有限公司业务副总经理，研究员，主要研究方向为卫星通信系统与装备。



殷杰（1983-），男，博士，航天恒星科技有限公司研究员，主要研究方向为卫星通信网络。