

杨桂松, 陶挺, 何杏宇, 等. 卫星物联网中基于能量感知的自适应节能路由策略[J]. 智能计算机与应用, 2024, 14(6): 127-133. DOI:10.20169/j.issn.2095-2163.240617

卫星物联网中基于能量感知的自适应节能路由策略

杨桂松¹, 陶挺¹, 何杏宇^{1,2}, 杜平³

(1 上海理工大学 光电信息与计算机工程学院, 上海 200093; 2 上海理工大学 出版印刷与艺术设计学院, 上海 200093;

3 上海卫星互联网研究院有限公司/上海市卫星互联网重点实验室, 上海 201210)

摘要: 随着 5G 技术的发展和 6G 技术的研究, 低轨卫星网络在卫星物联网中的地位越发重要, 而作为网络核心技术的路由策略仍面临一些挑战。本文针对低轨卫星网络拓扑结构动态变化、链路的不连续性、卫星能量供应受限等问题, 为了及时感知星间链路状态和卫星的能量并选择正确的路由, 将卫星物联网中的路由选择问题转化为马尔可夫决策过程下的最优策略问题, 提出一种基于 Dueling DQN (Dueling Deep Q Network) 的自适应节能路由算法。该算法通过改进 DQN 中神经网络的架构, 大幅度地提升了学习的效果。仿真结果表明, 与传统的 DQN 算法相比, 该算法能有效降低系统能耗, 均衡网络负载, 提高网络吞吐量。

关键词: 卫星物联网; 路由算法; 能量感知; 深度强化学习

中图分类号: TP393

文献标志码: A

文章编号: 2095-2163(2024)06-0127-07

Adaptive energy-efficient routing strategy based on energy aware in satellite Internet of Things

YANG Guisong¹, TAO Ting¹, HE Xingyu^{1,2}, DU Ping³

(1 School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; 2 College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai 200093, China; 3 Shanghai Institute of Satellite Internet Engineering Co., Ltd./Shanghai Satellite Network Research Institute, Shanghai 201210, China)

Abstract: With the development of 5G technology and the research of 6G technology, LEO satellite networks have become increasingly important in the satellite Internet of Things. However, as the core technology of the network, routing strategies still face some challenges. In this paper, an intelligent sensing routing algorithm based on Dueling DQN (Dueling Deep Q Network) is proposed to address issues such as dynamic changes in the topology of low earth orbit satellite networks, link discontinuity, and limited satellite energy supply, which could sense the status of the inter-satellite link and the energy of the satellite in time and select the correct route, and the routing problem in the satellite Internet of Things is transformed into the optimal strategy problem in the Markov decision-making process. This algorithm greatly improves the learning effect by improving the architecture of neural networks in DQN. Simulation results show that compared with traditional DQN algorithm, this algorithm can effectively reduce system energy consumption, balance network load, and improve network throughput.

Key words: satellite Internet of Things; routing algorithm; energy aware; deep reinforcement learning

0 引言

卫星物联网 (Satellite Internet of Things, SIoT) [1] 将低轨 (Low Earth Orbit, LEO) 卫星与物联网相结合, 是在天地一体化信息网络架构下的一种泛在物

联网体系, 能够实现全球海洋、陆地和空中多层次的物联网数据互联互通。目前, 地面通信网络只能覆盖地球表面陆地约 20% 的地区 [2], 通过建立低轨卫星物联网, 使之成为地面物联网的补充和延伸, 则能够有效克服地面物联网在海洋、山区、沙漠等区域覆盖

基金项目: 南通市科技局社会民生计划项目 (MS12021060); 敏捷智能计算四川省重点实验室开放式基金资助项目; 浦东新区科技发展基金产学研专项 (PKX2021-D10)。

作者简介: 杨桂松 (1982-), 男, 博士, 副教授, 硕士生导师, 主要研究方向: 物联网, 普适计算; 陶挺 (1995-), 男, 硕士研究生, 主要研究方向: 卫星物联网; 杜平 (1978-), 男, 博士, 主要研究方向: 卫星互联网, 5G 移动通信, 计算机网络等。

通讯作者: 何杏宇 (1984-), 女, 博士, 副教授, 硕士生导师, 主要研究方向: 物联网, 群智计算大数据分析。Email: xy_he@usst.edu.cn

收稿日期: 2023-04-19

能力不足的问题,可实现全球网络的无缝覆盖^[3]。

随着 LEO 卫星网络的快速发展,卫星的成本逐渐降低,StarLink^[4]、OneWeb^[5]、TeleSat^[6] 等大规模星座相继出现。随着卫星规模的增大,星座星间路由技术的设计难度也显著增大。对于大规模 LEO 星座的网络路由技术存在以下挑战:LEO 卫星网络的拓扑变化复杂,很难实现路由算法的快速收敛,计算出稳定的星间路由^[7];LEO 卫星的电池容量是有限制的,在太阳光被遮挡时,过度地消耗电池能量,很容易导致星间链路的数据传输中断^[8];LEO 卫星受到计算资源和存储资源的限制,不能存储大量的路由信息和执行大规模的计算任务^[9]。

为了应对上述挑战,国内外已有很多学者从多个角度对 LEO 卫星动态路由算法进行了研究。文献[10]提出了一种动态自适应的路由算法,其策略是先找出多条传输时延小、切换次数少的备选路径,再从备选路径中选择一条负载最轻的路径进行传输。文献[11]提出了一种显式负载均衡算法,该算法在当卫星出现链路拥塞时,拥塞节点主动向邻居节点发送拥塞信息,使得节点在选择下一跳节点时规避拥塞的相邻节点,从而有效地提高了网络的负载均衡性。文献[12]提出了一种能量感知的路由方案去解决能量受限的卫星网络场景下不同业务的传输需求,提高了优先级高的数据的传输质量。文献[13]将虚拟拓扑法与虚拟节点法相结合,利用虚拟拓扑法处理星座层间拓扑动态性,而采用虚拟节点法处理层内拓扑动态性,从而简化了星间路由算法的设计。文献[14]建立了虚拟节点动态资源图,利用蚁群算法来建立多目标约束的模型并求解,缓解了 LEO 卫星网络拥塞,提高了网络吞吐量。文献[15]针对低轨卫星通信系统空间段动态漏斗型拥塞问题,根据卫星网络状态和源与目的之间距离计算转移概率,提出了一种基于区域流量绕路的多路径搜索负载均衡路由算法,实现全局最优的区域业务绕行。

现有卫星网络的路由方法,虽然大多采用动静结合的方式,可以利用星座网络的多路径特性,对数据进行分流,一定程度上提升网络的吞吐量,但仍有许多可以改进的地方。例如,针对低轨卫星网络高动态的拓扑结构,当低轨卫星网络的规模较大时,现有的路由算法很难直接输出满足复杂约束的路径,同时也会导致算法的处理时间较长。

本文针对 LEO 卫星网络拓扑结构动态变化、通信链路质量不稳定、卫星能量供应受限等问题,提出

了一种基于 Dueling DQN 的自适应节能路由算法。该算法利用深度强化学习对 LEO 卫星网络的路由进行优化,能够适应大规模网络的动态性变化,实时捕获网络状态信息,并根据网络能耗采取不同的措施,保证可靠的数据服务传输。与传统的算法相比,该算法有效降低了系统的能耗,提高了网络的吞吐量,实现了网络的负载均衡。

1 系统模型

1.1 任务模型

在低轨卫星物联网场景中,系统由 LEO 卫星网络、终端用户和云计算中心组成。由于低轨卫星不具备处理任务的能力,当终端用户产生任务时,需要接入 LEO 卫星网络将任务转发至云计算中心处理。根据终端用户接入 LEO 卫星网络方式不同,可以分为直连型终端和汇聚型终端。通常情况下,直连型终端具有较强的射频能力,能够直接与卫星进行通信完成数据的传输。而汇聚型终端由于设计相对简单,射频能力较弱,无法直接与卫星进行通信,需要通过具有接入 LEO 卫星网络能力的地面基站间接接入到 LEO 卫星网络。

在本文的研究场景中,针对的是直连型终端,终端用户可随时向 LEO 卫星发起接入请求。终端用户产生的任务直接上传至覆盖该终端用户的卫星上,如果云计算中心和终端用户同在一个卫星服务区域内,卫星直接选择将任务发送至云计算中心;否则,通过 LEO 卫星网络中的卫星作为中继,将任务发送至云计算中心。由于 LEO 卫星节点的资源受限,LEO 卫星可同时与多个卫星建立连接,因此,当 LEO 卫星对任务进行转发时,需要根据星间链路状态选择合适的卫星节点转发,以确保将任务发送至云计算中心的过程中消耗的能量最小^[16]。低轨卫星物联网架构如图 1 所示。

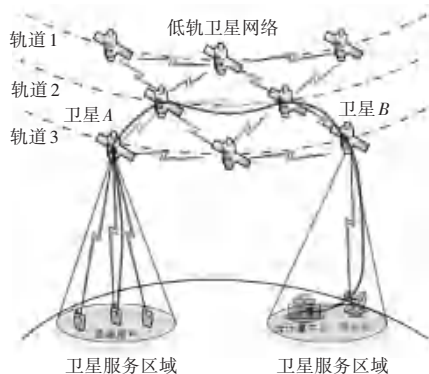


图 1 低轨卫星物联网架构图

Fig. 1 Architecture of LEO satellite Internet of Things

一般地,研究采用一个元组 (φ, A, B) 表示终端用户产生的任务。这里, φ 表示终端用户产生任务的数据大小; A 表示与任务产生的终端用户连接的卫星节点; B 表示与云计算中心连接的卫星节点。

1.2 星座模型

本文所研究的 LEO 卫星网络是基于 Walker 星座^[17] 给出的,星座中所有卫星呈均匀对称分布,同一个轨道面内卫星分布均匀。每颗卫星都能与其它的一颗卫星相连接,其中包括 2 条轨间链路和 2 条轨内链路。

为了保护低轨卫星网络拓扑的动态性,本文利用虚拟拓扑^[18] 的思想,通过将动态的低轨卫星网络拓扑按一定的时间时隙划分为网络拓扑序列。在每个时隙的开始,低轨卫星网络系统构建相对稳定的拓扑结构;在每个时隙内无论卫星的移动性如何改变,每颗卫星与其它颗卫星的链接状态保持不变。假设在 t 时隙 LEO 卫星网络的拓扑结构,如图 2 所示。



图 2 低轨卫星网络拓扑结构图

Fig. 2 Illustration of LEO satellite networks

为了方便表示,用一个无向图 $M_t = (V, E)$ 表示 t 时隙 LEO 卫星网络拓扑结构中卫星之间的链接关系,将卫星作为节点,卫星之间的链接作为边,边的权重由卫星节点 i 与节点 j 之间的距离 $d^{(i,j)}$ 给定,其中, $V = \{v_1, v_2, \dots, v_N\}$ 表示所有卫星节点的集合,这里的 N 表示低轨卫星网络中卫星节点的数量; $E = \{(i, j) \mid i \neq j\}$ 表示所有边的集合。

卫星的无向图可以用大小为 $N \times N$ 的连接矩阵 M_t 表示:

$$M_t^{(i,j)} = \begin{cases} 0, & \text{卫星节点 } i \text{ 和 } j \text{ 之间没有链接} \\ d^{(i,j)}, & \text{卫星节点 } i \text{ 和 } j \text{ 之间有链接} \end{cases} \quad (1)$$

1.3 通信模型

一般通过定向天线发射电磁波的方式建立星地通信或者星间通信,能够聚集电磁波能量往某一个方向上进行辐射,并对发射功率进行放大,天线增益^[19] G 可表示为:

$$G = \frac{4\pi S}{\lambda^2} \eta = \frac{4\pi D^2}{\lambda^2} \eta \quad (2)$$

其中, S 表示天线的开口面积; D 表示天线的直径; η 表示天线的效率; λ 表示电磁波的波长。

根据式(2),定义卫星的发射天线增益为 G_T , 卫星的接收天线增益为 G_R 。通常把卫星的发射天线在波束中心轴向上辐射的功率称之为发送设备的有效全向辐射功率 (Effective Isotropic Radiated Power, EIRP), 即天线发射功率 $P_T^{(i,j)}$ 和发射天线增益 G_T 的乘积 (单位为 W), 这是表征卫星发射能力的一项重要技术指标,即:

$$EIRP^{(i,j)} = P_T^{(i,j)} G_T \quad (3)$$

星间链路间采用无线电波进行数据传输,根据香农公式定理^[20] 和式(3)得到,将卫星之间的传输速率 $R^{(i,j)}$ 表示为:

$$R^{(i,j)} = \omega^{(i,j)} \log_2 \left(1 + \frac{EIRP^{(i,j)}}{\sigma^2} \right) \quad (4)$$

其中, $\omega^{(i,j)}$ 为无线信道的带宽; σ^2 为噪声功率。

1.4 能耗模型

考虑到卫星通信链路的传输会出现损耗,其中最主要的是自由空间传播损耗^[21]。当电波在自由空间传播时,用增益为 G_R 的天线接收,则卫星的接收功率 $P_R^{(i,j)}$ 可表示为:

$$P_R^{(i,j)} = \frac{G_R G_T P_T^{(i,j)}}{L_p^{(i,j)}} \quad (5)$$

其中, $L_p^{(i,j)}$ 为自由空间传播损耗。根据电波的传播理论,卫星通信链路中的自由空间传播损耗 $L_p^{(i,j)}$ 由式(6)计算:

$$L_p^{(i,j)} = \frac{4\pi d^{(i,j)} f}{c} = \frac{4\pi d^{(i,j)} f^2}{c} \quad (6)$$

其中, $d^{(i,j)}$ 为卫星之间的距离; λ 为电磁波的波长; c 为光速; f 为工作频率。给定卫星的接收功率 $P_R^{(i,j)}$, 根据式(5)和式(6),则卫星的发射功率 $P_T^{(i,j)}$ 可以表示为:

$$P_T^{(i,j)} = \frac{16\pi^2 P_R^{(i,j)}}{G_T G_R \lambda^2} (d^{(i,j)})^2 \quad (7)$$

假设星间通信和星地通信的工作波长相同,卫星的接收天线增益 G_T 与发射天线增益 G_R 也都相同。则当卫星的接收功率相同时,由式(7)可知,卫星所需的发射功率与 2 颗卫星之间的距离 $d^{(i,j)}$ 的平方相关。

由式(4)可得,将数据大小为 φ 的任务从卫星 i 转发至卫星 j 所需的时间 $\Gamma^{(i,j)}$ 可表示为:

$$\Gamma^{(i,j)} = \frac{\varphi}{R^{(i,j)}} \quad (8)$$

其中, $R^{(i,j)}$ 表示卫星 i 与卫星 j 之间的传输速率。因此,由式(7)和式(8)得出,卫星 i 将数据大小为 φ 的任务转发至卫星 j 的发射能耗 $e_T^{(i,j)}$ 为:

$$e_T^{(i,j)} = P_T^{(i,j)} \Gamma^{(i,j)} \quad (9)$$

卫星 j 接收数据大小为 φ 的任务所需的接收能耗 $e_R^{(i,j)}$ 可表示为:

$$e_R^{(i,j)} = P_R^{(i,j)} \Gamma^{(i,j)} \quad (10)$$

2 问题描述

在 LEO 卫星物联网场景下,为使得任务转发过程中系统能量消耗最低,本文设计了一种自适应节能路由策略。由式(9)和式(10)可以得到卫星节点 A 覆盖区内终端设备产生的任务通过低轨卫星网络发送到卫星节点 B 覆盖区内云计算中心的总能耗 $e_{A,B}$, 即:

$$e_{A,B} = \sum_{(i,j) \in L_{A,B}} (e_R^{(i,j)} + e_T^{(i,j)}) \quad (11)$$

其中, $L_{A,B}$ 表示从卫星节点 A 到 B 最优路由路径上的链路节点对。为了最小化低轨卫星网络路由过程中的总能耗 $e_{A,B}$, 最优路径 $L_{A,B}$ 可以通过以下优化问题 $P1$ 得到:

$$P1: \min_{L_{A,B}} e_{A,B} \quad (12)$$

$$\text{s.t. } C1: (A, j) \in L_{A,B}, 1 \leq j \leq N \quad (13)$$

$$C2: (B, \text{None}) \in L_{A,B} \quad (14)$$

$$C3: (i, j) \in E, 1 \leq i \leq N, 1 \leq j \leq N \quad (15)$$

约束 $C1$ 表示 $L_{A,B}$ 应包含 (A, j) , 以节点 A 为起始节点, 节点 j 为第一个链路的结束节点; 约束 $C2$ 定义了节点 B 是最后一个节点; 约束 $C3$ 表示链路限制在无向图 M_i 中的卫星链接边集合 E 。

3 基于 Dueling DQN 的自适应节能路由算法

本节中,首先将上述问题 $P1$ 重新表示为马尔可夫决策过程求解问题,并定义问题的状态空间、动作空间以及奖励函数,然后提出了一种基于 Dueling DQN 的自适应节能路由算法,通过改进传统 DQN 算法中神经网络的架构,以提取状态之间的特征关系,从而提高模型收敛速度和稳定性。

3.1 问题建模

(1) 状态空间: t 时隙的状态 $s_t \in S$ 定义为:

$$s_t = \{l(t), u(t), \tau(t)\} \quad (16)$$

其中, $l(t)$ 表示 t 时隙任务所处的卫星节点; $u(t)$ 表示 t 时隙任务所处卫星节点的链路连接情况,可由式(1)中的链接矩阵 M_i 得到; $\tau(t)$ 表示 t 时隙任务的数据量大小。

(2) 动作空间:为了使任务在转发过程中的能量消耗最小化,并尽可能利用低轨卫星通信资源,需要智能体根据每个时隙的状态 s_t 做出适当的决策。本文不仅要考虑卫星节点的选择,还要考虑带宽资源分配,因此,动作空间 a_t 可以定义为:

$$a_t = \{h(t), B(t)\} \quad (17)$$

其中, $h(t)$ 表示 t 时隙智能体进行下一次转发的卫星节点, $B(t)$ 表示 t 时隙为任务分配的带宽资源。

(3) 奖励函数:本文的优化目标是为最小化任务转发过程中的能量消耗,最大化强化学习模型中的长期奖励 R , 长期奖励表示为:

$$R = \sum_{i=1}^{|T|} r_i \quad (18)$$

在 t 时隙,当智能体执行动作 a_t 后会从环境中得到一个即时奖励 r_t , 用于评价当前决策的好坏,奖励函数 r_t 定义如下:

$$r_t = -(e_R^{(i,j)} + e_T^{(i,j)})_t \quad (19)$$

3.2 算法设计

近年来,许多基于价值的深度强化学习算法被应用于处理拓扑结构变化的 LEO 卫星网络环境,如 DQN^[22]、Double DQN^[23-24]。针对 LEO 卫星网络拓扑结构动态变化、通信链路质量不稳定、卫星能量供应受限等问题,提出了一种基于 Dueling DQN 的自适应节能路由算法。

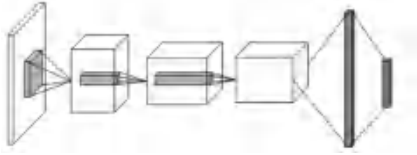
竞争网络架构 (Dueling DQN) 是针对 DQN 的一种改进,对 DQN 中神经网络的架构进行了简单的修改,大幅度地提升了学习的效果。该架构在评估某个状态下动作的价值 $Q(s, a)$ 的同时,也评估了与动作无关的状态的价值函数 $V(s)$ 和在该状态下各个动作的相对价值函数 $\Lambda(s, a)$ 的值。

图 3(a) 是传统 DQN 算法的网络图,图 3(b) 是 Dueling DQN 的网络图。从图 3 中可以看出, DQN 的神经网络直接输出 Q 函数的值,即某个状态下动作对应的价值,而其前一层是全连接层。Dueling DQN 针对 DQN 的改进主要就是全连接层,将全连接层改成 2 条流,其中一条输出关于状态的价值,另外一条输出关于动作的优势函数值,最终合并为 Q 函数。Dueling DQN 的 Q 函数是由状态的价值函数 $V(s)$ 加上每个动作的优势函数 $\Lambda(s, a)$ 得到,表示公式如下:

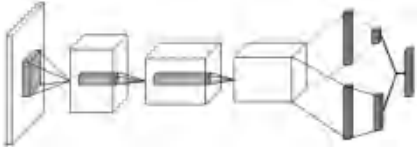
$$Q(s, a; \theta, \zeta, \eta) = V(s; \theta, \eta) + \Lambda(s, a; \theta, \zeta) \quad (20)$$

其中, θ 为公共特征提取层的参数; η 为价值函数 $V(s)$ 的参数; ζ 为优势函数 $\Lambda(s, a)$ 的参数; 价值函数 $V(s)$ 表示一个状态的好坏程度; 优势函数

$\Lambda(s, a)$ 表明在这个状态下某一个动作相对于其它的动作的好坏程度; Q 函数则表示这个状态下确定的某个动作的价值。



(a) 传统 DQN 算法的网络图



(b) Dueling DQN 的网络图

图3 传统 DQN 与 Dueling DQN 网络图对比

Fig. 3 The comparison diagram of traditional DQN and Dueling DQN network

然而,由式(20)来确定的 Q 值还会存在一个问题,由于 $V(s)$ 是一个变量,所以在神经网络中这个值不管偏左,还是偏右,对最后 Q 值都是没有影响的。也就是说,如果给定一个 Q 值,也无法得到唯一的 V 和 Λ 。为了解决这个问题,可以强制令所选择的贪心动作的优势函数为 0,即:

$$Q(s, a; \theta, \zeta, \eta) = V(s; \theta, \eta) + (\Lambda(s, a; \theta, \zeta) - \max_{a' \in A} \Lambda(s, a'; \theta, \zeta)) \quad (21)$$

这样就可以得到唯一的值函数:

$$a^* = \arg \max_{a \in A} Q(s, a; \theta, \zeta, \eta) = \arg \max_{a \in A} \Lambda(s, a; \theta, \zeta) \quad (22)$$

$$Q(s, a^*; \theta, \zeta, \eta) = V(s; \theta, \eta) \quad (23)$$

通过上述一系列的改进,相比较于传统的 DQN 算法,通过引入了 Dueling 网络架构来修改 Q 值的估计方式,对 DQN 中的神经网络进行了重新设计,使其能够大幅度地提升了学习的效果。因此,提出了一种基于 Dueling DQN 去求解卫星物联网中动态路由的规划问题,利用卫星物联网环境与智能体进行实时的交互,所产生的状态、动作以及奖励存在经验池中,通过随机采样经验池中的数据,构建目标函数与损失函数反复迭代训练网络模型参数,最后得到最佳的路由选择策略。

算法 1 基于 Dueling DQN 的自适应节能路由算法

1. 随机初始化价值网络 Q_θ 的参数 θ , 价值函数参数 ζ , 优势函数参数 η ;
2. 初始化目标网络参数 $\theta' \leftarrow \theta$;
3. 初始化经验复用池 $D = \emptyset$;

4. For each episode:

5. 获取环境初始的系统全局状态 s_0 ;

6. For each step:

7. 以一定的贪心率 ε 随机选择动作 a_t ,

8. 否则根据价值网络 Q_θ 选择动作 a_t : $a_t = \arg \max_{a \in A} Q(s_t, a; \theta, \zeta, \eta)$;

9. 执行动作 a_t , 从环境中获得奖励 r_t 和下一个状态 s_{t+1} ;

10. 将 (s_t, a_t, r_t, s_{t+1}) 存储经验复用池 D 中;

11. 从 D 中随机采样含有 N 个经验数据 (s_i, a_i, r_i, s_{i+1}) 的小批量;

12. 获得目标值:

$$y_i = \begin{cases} r_i, & s_{i+1} \text{ is terminates} \\ r_i + \gamma \arg \max_{a' \in A} Q(s_{i+1}, a'; \theta, \zeta, \eta), & \text{otherwise} \end{cases};$$

13. 通过最小化损失函数更新价值网络 Q_θ :

$$\theta \leftarrow \arg \min \frac{1}{N} \sum (y_i - Q(s_i, a_i; \theta', \zeta, \eta))^2;$$

14. 更新目标网络的参数: $\theta = \tau \theta + (1 - \tau) \theta'$;

15. end for

16. end for

4 仿真与结果分析

4.1 仿真环境与参数设置

本文假设在低轨卫星物联网场景中,系统由 16 颗低轨卫星节点、1 个终端用户和 1 个云计算中心组成。终端用户产生的任务数据大小在 100 ~ 400 kb 之间;星间链路通信使用 Ka 频段,频率为 23.18 GHz,卫星间的距离都在 2 000 ~ 4 000 km 之间;卫星的发射天线增益为 15.4 dB,接收天线增益为 16.7 dB,有效全向辐射功率在 40 ~ 50 dBW 之间。实验中的默认参数见表 1。

表 1 实验参数

Table 1 Experimental parameters

名称	值
星间链路工作频率 f / GHz	23.18
卫星接收天线增益 G_R / dB	16.7
卫星发射天线增益 G_T / dB	15.4
卫星之间距离 $d^{(i,j)}$ / km	2 000 ~ 4 000
有效全向辐射功率 $EIRP^{(i,j)}$ / dBW	40 ~ 50
任务数据大小 φ / kb	200
通信带宽 $\omega^{(i,j)}$ / MHz	10

对于所提出的 Dueling DQN 算法,使用全连接神经网络来训练模型。Dueling DQN 的神经网络由 PyTorch 框架实现,每个神经网络包含 4 层,包括输入层、输出层和 2 个隐藏层,隐藏层的神经元数量为

256。每次训练采用数据大小为 64 的小批处理,并使用大小为 10 000 的经验复用池,目标网络软更新的 τ 设置为 0.02。

4.2 算法收敛性分析

为了评估超参数对所提算法性能的影响,本文验证了 Dueling DQN 算法的学习率对算法收敛性和稳定性的影响。在实验中,将学习率分别被设定为 $1e^{-3}$ 、 $1e^{-4}$ 和 $1e^{-5}$ 。不同学习率对算法收敛性的影响如图 4 所示。从图 4 可以看出,学习率为 $1e^{-5}$ 时,曲线在 100 个 episode 后收敛到最优值,并在达到收敛后保持稳定状态。然而,当学习率为 $1e^{-4}$ 或 $1e^{-5}$ 时,需要花费大于 50 个 episode 才能达到收敛状态。由此可以得出结论,曲线的最优值与学习率的大小不成正比,当算法中优化器的学习率太小时,算法需要更多的训练回合来达到收敛状态;当学习率过大时,曲线不一定能收敛到一个较好的值,甚至会导致训练不稳定。

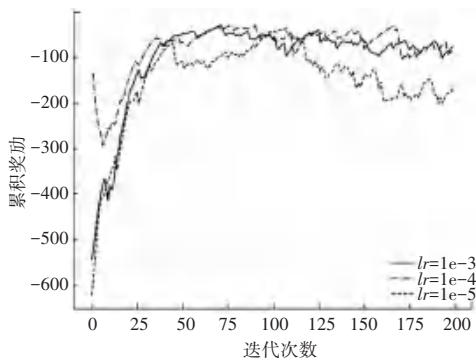


图 4 不同学习率对算法收敛性的影响

Fig. 4 Impact of learning rate on convergence

不同算法收敛性的对比如图 5 所示,展示了传统 DQN 算法与基于 DQN 改进的 Dueling DQN 算法的收敛性能对比。研究提出 Dueling DQN 大约在学习迭代 50 次后智能体的累积奖励曲线波动逐渐趋于平缓,相比较于 DQN 的方法,Dueling DQN 前期具有更高的累积奖励,这主要是 DQN 中目标网络对目标值的过估计,就使得当训练的迭代次数较少,策略不太稳定。在智能体前期学习过程中,智能体对环境认知较少,仍然处于探索阶段,导致了累积奖励较小,波动性较大。逐渐地,智能体对环境的认知慢慢地增加,逐渐学习到了更好的动作策略,最终使得累计奖励趋于稳定。

4.3 对比实验

为了验证提出算法的优越性和可靠性,本文将提出的 Dueling DQN 算法与 2 种基线算法进行比较:

(1) 最短跳数优先算法 (Shortest Hop First,

SHF):任务从源传送到目的地要经过卫星节点的转发次数最少;

(2) 最短路径优先算法 (Shortest Path First, SPF):任务从源传送到目的地要经过路径长度最短。

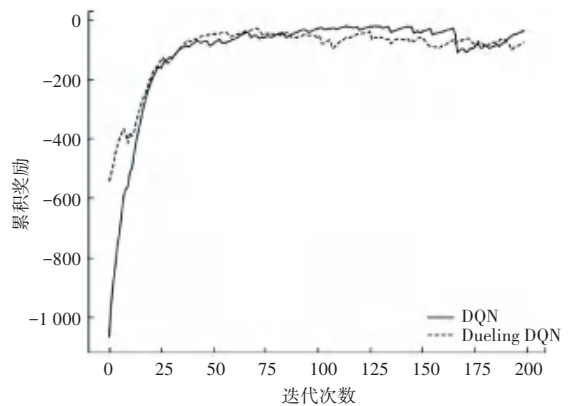


图 5 不同算法收敛性的对比

Fig. 5 Comparison of convergence of different algorithms

为验证任务数据量大小对系统能耗影响,本文设置了任务数据量从 100 增长到 400 kb,增量为 50 kb。任务的数据大小对系统能耗的影响如图 6 所示。由图 6 可以看出,随着任务数据量的增加,所有算法的系统能耗都在增加。这是因为任务的转发需要利用低轨卫星网络系统中的各种资源,任务数据量的增加,必然导致卫星能量消耗的加剧。通过比较不同的算法,能够得出 Dueling DQN 具有更低的系统能耗,这说明了在路由的决策过程中,Dueling DQN 比 SHF 和 SPH 拥有更好的效果。

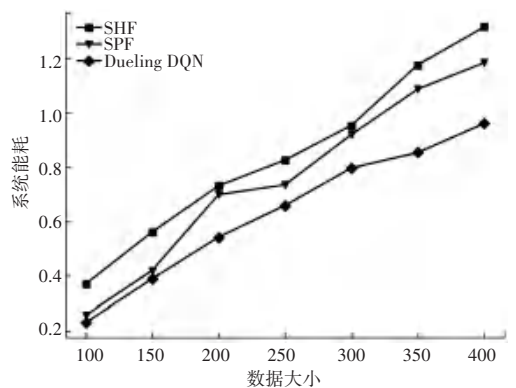


图 6 任务的数据大小对系统能耗的影响

Fig. 6 The impact of task data size on system energy consumption

任务的数据大小对平均吞吐量的影响如图 7 所示。由图 7 可知,网络平均吞吐量的值越大,说明 LEO 卫星网络的数据传输能力越强,低轨卫星网络中的资源越能够得到充分的利用。从图 7 中可以很明显地看出,相比较于 SPF 和 SHF 算法,Dueling DQN 分别平均提升了大约 85% 和 176% 的性能。

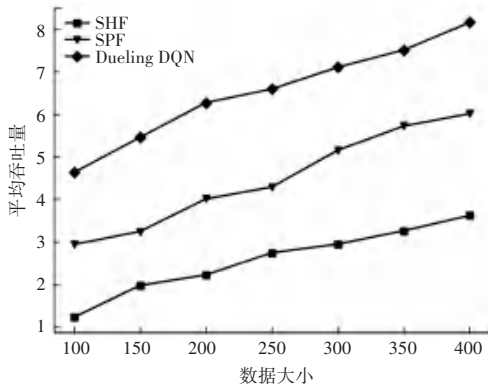


图7 任务的数据大小对平均吞吐量的影响

Fig. 7 The impact of task data size on average throughput

5 结束语

本文针对低轨卫星网络拓扑结构动态变化、通信链路质量不稳定、卫星能量供应受限等问题,提出了一种基于 Dueling DQN 的自适应节能路由算法。该算法实时捕获 LEO 卫星物联网的状态信息,利用深度强化学习对网络的路由进行学习,使之能够适应大规模网络的动态性变化,并根据网络能耗采取不同的措施,保证可靠的数据服务传输。仿真结果表明,与传统的 DQN、SPF 和 SHF 算法相比,本文提出的算法能够有效降低系统的能量消耗,均衡网络的负载压力,提高网络的吞吐量。

参考文献

- [1] KIM T, KWAK J, CHOI J P. Satellite edge computing architecture and network slice scheduling for IoT support[J]. IEEE Internet of Things Journal, 2021, 9(16): 14938-14951.
- [2] 苟亮,左云鹏,万扬洋,等. 低轨卫星物联网综述[J]. 信息化研究, 2022, 48(5): 1-9.
- [3] LATVA-AHO M, LEPPÄNEN K, CLAZZER F, et al. Key drivers and research challenges for 6G ubiquitous wireless intelligence[EB/OL]. [2019-09-09]. <http://urn.fi/urn:isbn:9789526223544>.
- [4] BOLEY A C, WRIGHT E, LAWLER S, et al. Plaskett 1.8 m observations of starlink satellites[J]. arXiv preprint arXiv:2109.12494, 2021.
- [5] HENRI Y. The OneWeb satellite system [M]// PELTON J. Handbook of Small Satellites: Technology, Design, Manufacture, Applications, Economics and Regulation. Cham: Springer, 2020: 1-10.
- [6] ZONG P, KOHANI S. Design of LEO constellations with inter-satellite connects based on the performance evaluation of the three constellations SpaceX, OneWeb and Telesat[J]. Korean Journal of Remote Sensing, 2021, 37(1): 23-40.
- [7] GAO X, LIU R, KAUSHIK A, et al. Dynamic resource allocation for virtual network function placement in satellite edge clouds[J]. IEEE Transactions on Network Science and Engineering, 2022, 9(4): 2252-2265.

- [8] TAKAHASHI M, KAWAMOTO Y, KATO N, et al. Adaptive power resource allocation with multi-beam directivity control in high-throughput satellite communication systems [J]. IEEE Wireless Communications Letters, 2019, 8(4): 1248-1251.
- [9] WANG Lina, WU Yanan, ZHANG Haijun, et al. Resource allocation for NOMA based space-terrestrial satellite networks[J]. IEEE Transactions on Wireless Communications, 2020, 20(2): 1065-1075.
- [10] CHEN Jing, ABBAS J. An adaptive path routing scheme for satellite IP networks[J]. International Journal of Communication Systems, 2003, 16(1): 5-21.
- [11] TALEB T, MASHIMO D, JAMALIPOUR A, et al. Explicit load balancing technique for N GEO satellite IP networks with on-board processing capabilities [J]. IEEE/ACM Transactions on Networking, 2008, 17(1): 281-293.
- [12] HAO Linchun, REN Pinyi, DU Qinghe. Satellite QoS routing algorithm based on energy aware and load balancing [C]//2020 International Conference on Wireless Communications and Signal Processing (WCSP). Nanjing, China: IEEE, 2020: 685-690.
- [13] LIU Zhiguo, ZHU Jie, ZHANG Jiangmei et al. Routing algorithm design of satellite network architecture based on SDN and ICN[J]. International Journal of Satellite Communications and Networking, 2020, 38(1): 1-15.
- [14] 邢川, 陈二虎, 韩笑冬. 基于动静结合方法的卫星网络路由方法研究[J]. 空间控制技术与应用, 2020, 46(3): 55-59.
- [15] ZHOU Ya, XIE Zhuochen, LIU Peilong, et al. Inter-satellite load balancing routing algorithm for LEO satellite constellation based on regional-traffic-detour [J]. Journal of University of Chinese Academy of Sciences, 2021, 38(5): 687-695.
- [16] WAQAR N, HASSAN S A, MAHMOOD A, et al. Computation offloading and resource allocation in MEC-enabled integrated aerial-terrestrial vehicular networks: A reinforcement learning approach [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 21478-21491.
- [17] GUAN Meiqian, XU Tianhe, GAO Fan, et al. Optimal walker constellation design of LEO-based global navigation and augmentation system[J]. Remote Sensing, 2020, 12(11): 1845.
- [18] WANG Y, CHE J, WANG N, et al. Load-balancing method for LEO satellite edge-computing networks based on the maximum flow of virtual links [J]. IEEE Access, 2022, 10: 100584-100593.
- [19] BENNER E, SESAY A B. Effects of antenna height, antenna gain, and pattern downtilting for cellular mobile radio [J]. IEEE Transactions on Vehicular Technology, 1996, 45(2): 217-224.
- [20] WYNER A. Recent results in the Shannon theory [J]. IEEE Transactions on Information Theory, 1974, 20(1): 2-10.
- [21] 李迎春, 朱诗兵. 星间链路自由空间传播衰减的分析[J]. 科学技术与工程, 2009, 9(13): 3840-3843.
- [22] LV Liangheng, ZHANG Sunjie, DING Derui, et al. Path planning via an improved DQN-based learning policy [J]. IEEE Access, 2019, 7: 67319-67330.
- [23] LEE D, SUN Y G, SIM I, et al. Neural episodic control-based adaptive modulation and coding scheme for inter-satellite communication link [J]. IEEE Access, 2021, 9: 159175-159186.
- [24] SEWAK M. Deep Q Network (DQN), Double DQN, and Dueling DQN: A step towards general artificial intelligence [M]// Deep Reinforcement Learning: Frontiers of Artificial Intelligence. Singapore: Springer, 2019: 95-108.