

卫星光网络中基于 DQN 的匹配度感知波长路由算法

艾琪洋, 何荣希*, 林润峰

(大连海事大学 信息科学技术学院, 辽宁 大连 116026)

摘要: 为了进一步降低卫星光网络的业务阻塞率, 提出了一种基于深度 Q 网络(DQN)的匹配度感知波长路由算法。该算法首先设计了路径波长匹配度因子, 用以反映当前路径与下一跳节点之间链路的共同空闲波长情况。随后, 算法综合考虑了路径波长匹配度、相邻链路时延以及下一跳节点到目的节点的最短跳数等多个因素, 以此为基础来构造奖励函数。在分配波长时, 为了后续业务请求在相邻链路上有更多的空闲波长, 设计了相邻链路波长匹配度指标, 用于描述该路径与其相邻链路之间的空闲波长对齐情况。仿真结果表明, 该算法可以降低网络阻塞率和时延, 提高波长利用率。

关键词: 卫星光网络; 路由波长分配; 深度强化学习; 波长匹配度

中图分类号: TN929.11 **文献标志码:** A **文章编号:** 1002-5561(2024)03-0030-08

DOI: 10.13921/j.cnki.issn1002-5561.2024.03.006

开放科学(资源服务)标识码(OSID):



Wavelength routing algorithm with matching degree awareness based on DQN in satellite optical networks

AI Qiyang, HE Rongxi*, LIN Runfeng

(College of Information Science and Technology, Dalian Maritime University, Dalian Liaoning 116026, China)

Abstract: In order to further reduce the traffic blocking rate of satellite optical networks, this paper proposes a wavelength routing algorithm based on deep Q-networks (DQN) and matching perception. This algorithm first designs a path wavelength matching factor to reflect the common free wavelengths between the current path and the link of the next hop node. Subsequently, the algorithm takes into account multiple factors such as path wavelength matching, adjacent link delay, and the shortest number of hops from the next hop node to the destination node, and constructs a reward function based on these factors. When allocating wavelengths, in order to have more free wavelengths on adjacent links for subsequent service requests, this paper designs an adjacent link wavelength matching index to describe the alignment of free wavelengths between the path and its adjacent links. The simulation results show that this algorithm can reduce network blocking rate and delay, and improve wavelength utilization.

Key words: satellite optical networks, routing and wavelength assignment, deep reinforcement learning, wavelength matching degree

0 引言

卫星光网络(OSN)已受到越来越多的关注^[1-2], 其星间采用激光链路传输数据, 相较于传统卫星网络的

收稿日期: 2024-03-12。

基金项目: 国家自然科学基金项目(61371091, 61801074, 62371085)资助; 大连市科技创新基金项目(2019J11CY015)资助。

作者简介: 艾琪洋(1999—), 男, 硕士研究生, 现就读于大连海事大学信息科学技术学院电子信息专业, 主要从事光通信网络研究工作, 曾荣获大连海事大学优秀学生奖学金、全国大学生数学建模大赛辽宁省三等奖等奖项。

***通信作者:** 何荣希(1971—), 男, 博士, 教授, 主要研究方向为光网络和无线网络技术。



微波传输, 具有不受频率限制、带宽大、传输速率高、保密性好等优势。OSN 采用波长路由技术, 以光信号为载体, 并利用波分复用(WDM)技术实现在同一链路上的多路数据流并行传输, 从而显著提高数据传输速率和系统容量, 满足未来不断增长的数据需求。

路由波长分配(RWA)是 OSN 中最具挑战性的任务之一, 已被证明是一个 NP(Non-deterministic Polynomial)难问题。该问题通常被拆解为路由选择和波长分配 2 个子问题来解决, 已有文献提出了多种启发式算法用以解决这些子问题。例如: 文献[3]不仅提出了一种虚拟节点思想, 用以屏蔽 OSN 拓扑的动态变化, 还设计了一种分布式路由算法, 该算法能够针对低轨

卫星网络的特点,基于每个数据包灵活做出路由决策,有效避免拥塞区域。文献[4]提出一种基于小窗策略的蚁群优化RWA算法,降低了OSN的阻塞率,并使收敛速度增加。文献[5]提出一种基于多业务的OSN蚁群优化路由算法,通过改进蚁群算法的启发函数,并考虑波长空闲率、时延等因素,使该算法在降低时延和提高波长利用率方面有着良好表现。文献[6]提出了一种创新的负载均衡算法,该算法基于拥塞避免原则,并通过人工蜂群算法优化链路代价。这一算法实现了流量的预先分配,从而最大限度地减少了拥塞对网络性能的负面影响。文献[7]针对低轨道OSN通信成功率下降的问题,提出一种基于蜂群适应度优化的RWA算法,改进了蜂群适应度函数,在保证业务实时传输的基础上,提高了卫星光通信业务的通信成功率。文献[8]针对大规模低轨OSN中存在的空间风险,利用一种管道风险评估模型来综合考虑碰撞概率和影响卫星正常运行的故障概率,并提出一种基于可靠性的路由算法,可降低通信失败率。文献[9]提出一种基于跨层设计的OSN RWA算法,考虑多普勒波长偏移、传输时延和波长一致性约束,建立跨层优化模型,并使用蚁群算法进行求解,该算法在通信成功率和传输时延方面有良好表现。文献[10]针对具有计算和存储功能的超级卫星节点OSN,提出一种RWA算法,可改善波长利用率和连通性等性能。但是,上述基于启发式的路由算法计算复杂度较大,而且容易陷入局部最优。

与启发式算法相比,深度强化学习(DRL)可以通过自主学习的方式找到基于奖励的最优策略,更快得到接近全局最优解,并且无需预先建立底层网络数学模型,这使得其在实际应用中更具灵活性和适应性。因此,已有不少文献将DRL引入路由算法设计中。文献[11]首次将Q学习引入动态网络的路由算法,证明了在网络负载发生变化时也能有效选路。针对中、低轨双层OSN,文献[12]提出一种基于DRL的RWA算法,利用DRL动态感知网络当前状态和进行路由决策,能够降低时延和丢包率,不仅提高了网络的波长利用率,还降低了针对低时延、高优先级业务的阻塞率。为了提高规划过程的灵活性和敏捷性,文献[13]针对低轨卫星网络,考虑了节点的排队时延、存储空间、链路带宽以及传播时延等因素的影响,提出一种基于DRL的负载均衡路由算法,该算法可以有效降低网络丢包率和时延。但是,上述基于DRL的路由算法,往往侧重于解决卫星网络的选路问题,而很少考虑对波长

分配的影响。实际上,RWA中的选路是波长分配的前提和基础,波长分配成功与否与路由算法所选路径息息相关。不同的路由选择结果会导致不同的波长需求,直接影响波长分配成功的可能性。另外,所选路径上的波长使用情况也会影响后续波长资源的分配,决定业务连接能否成功建立。由此可见,OSN中选路与波长分配紧密相关,选路时也应关注所选链路的波长使用情况,鼓励选择具有更多共同空闲波长的链路构成路径,有利于提高在该路径建立业务连接的成功率。为此,本文提出基于深度Q网络(DQN)的匹配度感知波长路由算法(下文简称DQN-MD-RWA算法)。

1 系统模型

本文构建铱星星座的OSN模型,由 $N=N_0 \times N_s$ 颗卫星组成,其中 N_0 表示卫星轨道的数量, N_s 表示每个轨道上卫星的数量。卫星之间通过星间链路(ISL)连接。铱星星座作为典型的星座,所有的卫星在各自轨道上均匀分布,并且每个轨道上的卫星数相同,星座中所有卫星轨道的升交点也均匀分布在地球赤道面的 180° 内。 π 星座的升交点两侧卫星运行方向相反,会产生1条反向旋转接缝,该接缝不利于星间链路的建立。因此,除了位于反向旋转接缝上的卫星外,每颗卫星有4条相邻链路,具体包括2条轨道内相邻链路和2条轨道间相邻链路。而位于反向旋转接缝上的卫星只有3条相邻链路,它们与接缝另一侧的卫星之间无法建立星间链路。假设任意星间链路 $e(j,k)$ 都支持 m 个波长,其中 j 与 k 表示卫星节点,构成波长集 $W_{jk} = \{\lambda_{jk}^i\}, j, k \in N, i=1, 2, \dots, m$ 。

虽然OSN的节点随卫星的运动不断变化,但是由于卫星运动呈现周期性,因此OSN的拓扑结构可以预测。为了避免网络拓扑动态变化造成的影响,本文采用虚拟节点的思想^[4],假定整个地球被与星座相同数量的虚拟卫星覆盖,这些虚拟卫星不会移动。在任何给定时间,虚拟卫星代表距离其最近的物理卫星节点。当物理卫星相对于地球的位置发生变化,与虚拟卫星之间的距离变远时,其所存储的路由信息将被传输到另一颗距离虚拟卫星最近的物理卫星节点,从而达到因屏蔽卫星相对运动导致网络拓扑变化的目的。在OSN中,当业务请求到达时,可在虚拟拓扑上确定源和目的节点,然后为其寻找一条路径,并在该路径选择一条满足波长一致性约束的波长建立光路;如果光路建立不成功,则阻塞该请求。

艾琪洋, 何荣希, 林润峰: 卫星光网络中基于 DQN 的匹配度感知波长路由算法

2 DQN-MD-RWA 算法

本文提出的 DQN-MD-RWA 算法是一种智能 RWA 算法, 它结合了 DQN 的特点。该算法通过不断与环境交互来调整路由的选择, 通过最大化采取动作所获得的奖励, 从而得到最佳的选路决策。然后, 在所选路径上根据相邻链路匹配度, 选择合适的波长以建立业务连接。DQN-MD-RWA 算法包含两部分, 即基于 DQN 的路径波长匹配度感知选路 (DQN-PWMDR) 算法和基于相邻链路波长匹配度的波长分配 (ALWMD-WA) 算法。前者选路时通过 DRL 综合考虑网络拓扑和链路信息, 感知所选路径和下一跳节点之间链路的路径波长匹配度, 总是鼓励选择路径波长匹配度较大的相邻节点作为下一跳节点; 后者在所选路径为业务请求分配波长时, 首先根据预分配方案计算预分配路径与相邻链路的波长匹配度, 然后选择波长匹配度最大的相邻链路预分配方案来建立光路, 从而为后续业务请求留下更多满足波长一致性约束的波长资源。

2.1 DQN-PWMDR 算法

本文利用 DQN 算法对 RWA 的选路进行建模, 可表示为一个三元组 (s_t, a_t, r_t) 。其中, s_t 表示在 t 时刻的状态, a_t 表示智能体基于当前状态 s_t 做出的动作, r_t 表示奖励函数。建模具体描述如下:

1) 状态空间: 假设虚拟卫星节点 j 的相邻卫星节点集为 $N_j = \{N_{j,k}\}$, $k=1, 2, \dots, K_j$, 其中 K_j 为节点 j 的相邻节点数。考虑到链路上许多因素都会对智能体的动作选择有影响, 本文设定虚拟卫星节点 j 在时刻 t 的状态 $s_t^j = \{o, d, W_j, W_{P(o,j)}, T_j, H_j\}$ 。其中, o 和 d 分别表示到达业务请求的源和目的节点, 在起始状态时节点 j 即为源节点 o ; T_j 表示节点 j 与相邻节点 k 之间链路 $e(j, k)$ 的时延, $T_j = \{T_{j,k}\}$ 。由于星载计算能力较强, 本文仅考虑传播时延。 H_j 表示节点 j 的相邻节点 k 到目的节点的最短跳数, $H_j = H_{k,d}$; W_j 表示链路 $e(j, k)$ 上波长的使用情况 (状态), $W_j = \{W_{j,k}\}$ 。其中, $W_{j,k} = \{\lambda_{j,k}^i\}$, $i=1, 2, \dots, n$, $\lambda_{j,k}^i$ 表示链路 $e(j, k)$ 上第 i 个波长的状态, 定义为

$$\lambda_{j,k}^i = \begin{cases} 1, & \text{波长 } i \text{ 可用 (空闲)} \\ 0, & \text{波长 } i \text{ 不可用 (被占用)} \end{cases} \quad (1)$$

用 $P(o, j)$ 表示源节点 o 到节点 j 的当前路径, $W_{P(o,j)}$ 表示路径 $P(o, j)$ 上的波长状态, $W_{P(o,j)} = \{\lambda_{P(o,j)}^i\}$, $\lambda_{P(o,j)}^i$ 为该路径上第 i 个波长的状态。由于波长一致性限制, 当且仅当任意波长在路径的所有链路上都空闲

时, 才能用来建立光路, 因此 $\lambda_{P(o,j)}^i$ 定义为

$$\lambda_{P(o,j)}^i = \prod_{e(j,k) \in P(o,j)} \lambda_{e(j,k)}^i \quad (2)$$

其中, $\lambda_{e(j,k)}^i$ 表示链路 (j, k) 上第 i 个波长的状态。

2) 动作空间: 由于本节主要考虑选路, 即如何从当前节点的相邻节点选择下一跳节点, 因此设置动作空间 $D = \{n_{j,1}, n_{j,2}, \dots, n_{j,K_j}\}$, 其中 $n_{j,k} \in \{0, 1\}$ 。当 $n_{j,k}=1$ 时, 表示下一跳选择节点 k , 否则不选择。为了加快算法的运行时间, 如果在相邻链路上不存在与当前路径上相同的可用波长, 则停止选路, 并直接阻塞该业务请求。

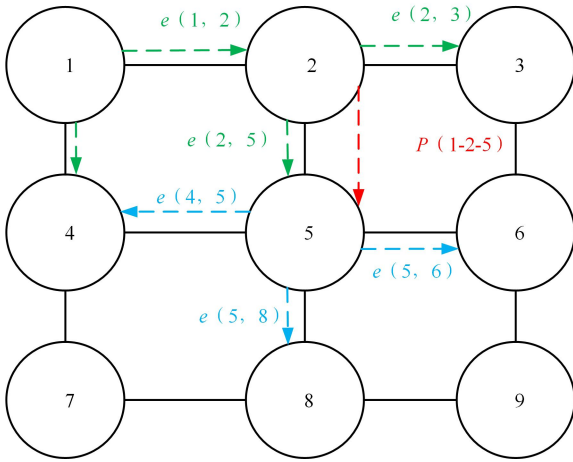
3) 瞬时奖励: 瞬时奖励代表智能体节点通过一次动作从环境中获得的奖励值, 其大小应与降低网络阻塞率和时延的算法优化目标紧密相关。

在 RWA 过程中, 选路后还将进行波长分配, 而所选路径的波长使用情况会影响后续的波长能否成功分配。为了提高在所选路径上波长分配的成功率, 本文提出采用路径波长匹配度因子来描述当前路径 (从源节点 o 到节点 j 的路径 $P(o, j)$) 与节点 j 和 k 之间的链路 $e(j, k)$ 上空闲波长的匹配情况。如果路径 $P(o, j)$ 与链路 $e(j, k)$ 之间有更多的共同空闲波长, 则有利于最终所选路径上存在更多满足一致性的可用波长, 从而提高光路成功建立概率, 降低业务阻塞率。基于此, 本文定义链路 $e(j, k)$ 与路径 $P(o, j)$ 的路径波长匹配度 $\eta_{e(j,k)}^P$ 为

$$\eta_{e(j,k)}^P = \sum_{i=1}^n (\lambda_{P(o,j)}^i \times \lambda_{e(j,k)}^i) \quad (3)$$

路径波长匹配度越大, 表示当前路径与连接下一跳节点的链路上有更多的共同空闲波长, 有利于提高业务连接成功建立概率。

本文以图 1 为例来说明如何计算路径波长匹配度。图 1(a) 中, 当前已选路径 P 为 1-2-5, 此时对于节点 5, 下一跳会在节点 4、6、8 中选择, 其对应链路上的波长资源使用情况如图 1(b) 所示。图 1(b) 中, 数字 1 表示该波长资源处于空闲状态, 数字 0 表示该波长已被占用, 假设每条链路支持 10 个波长。此时, 路径 P 与链路 $e(4, 5)$ 的共同可用波长为 λ_2 , 与链路 $e(5, 6)$ 的共同可用波长为 λ_2, λ_5 和 λ_9 , 与链路 $e(5, 8)$ 的共同可用波长为 $\lambda_2, \lambda_5, \lambda_9$ 和 λ_{10} 。链路 $e(4, 5)$ 与路径 $P(1-2-5)$ 的路径波长匹配度为 1, 链路 $e(5, 6)$ 与路径 $P(1-2-5)$ 的路径波长匹配度为 3, 而链路 $e(5, 8)$ 与路径 $P(1-2-5)$ 的路径波长匹配度为 4。因此, 下一跳应选择



(a) 网络拓扑

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
$P(1-2-5)$	0	1	0	0	1	0	1	0	1	1
$e(1, 2)$	0	1	0	1	1	0	1	0	1	1
$e(2, 3)$	1	0	1	0	1	0	0	0	0	1
$e(2, 5)$	1	1	0	0	1	0	1	0	1	1
$e(1, 4)$	1	0	0	1	0	1	0	0	0	1
$e(4, 5)$	1	1	1	0	0	1	0	1	0	0
$e(5, 6)$	0	1	1	0	1	0	0	0	1	0
$e(5, 8)$	0	1	0	0	1	1	0	0	1	1
$e(7, 8)$	0	1	0	1	0	1	1	1	0	1
$e(3, 6)$	0	0	1	0	0	1	0	1	1	0
$e(4, 7)$	1	1	0	0	0	1	1	0	1	1
$e(6, 9)$	1	0	0	1	1	0	0	0	1	1
$e(8, 9)$	1	1	1	0	0	0	1	1	1	0

(b) 波长资源使用情况

图1 路径及链路波长资源使用情况

节点8, 可使路径上的可用波长更多, 有利于提高波长成功分配概率。

为了降低业务阻塞率, 选路时除了考虑路径波长匹配度以外, 还应考虑下一跳节点距离目的节点的最小跳数, 这样可以节省建立光路上所需波长数, 从而使后续业务请求能有更多波长可用, 降低其被阻塞的可能性。另外, 还需考虑时延因素。因此, 结合状态中的参数, 奖励函数 r_t 应与邻接点 k 距离目的节点的最小跳数 H_k 、链路 $e(j, k)$ 与路径的波长匹配度 $\eta_{e(j, k)}^P$ 、链路 $e(j, k)$ 的时延 T_j 有关。由于每个因子的数量级不一致, 对以上因子进行归一化处理, 即

$$\eta_{e(j, k)}^P = \frac{\eta_{e(j, k)}^P}{n} \quad (4)$$

$$T_j' = \frac{2 \arctan(T_j)}{\pi} \quad (5)$$

$$H_k' = \frac{2 \arctan(H_k)}{\pi} \quad (6)$$

综合上述因素, 可定义奖励函数 r_t 为

$$r_t = \mu_1 \eta_{e(j, k)}^P - \mu_2 T_j' - \mu_3 H_k' \quad (7)$$

由于选路时希望得到更大的路径波长匹配度、更小的跳数和更低的时延, 因此奖励函数将3种因素归一化并加权后做减法运算。其中, μ_1 、 μ_2 、 μ_3 分别表示路径波长匹配度、时延和到达目的节点的最小跳数在奖励中所占的权重, 并且 $\mu_1 + \mu_2 + \mu_3 = 1$, 通过调整权重, 可以改变奖励函数对不同因子的侧重。

智能体根据式(3)采取动作来选择下一跳节点, 并得到相应的奖励, 在不断与环境交互和学习的过程中得到奖励最大的路径。DQN-PWMDR 路由的 DQN 结构如图2所示。当业务到来时, 神经网络首先获取当前的网络状态 s_t 。随后, 智能体依据这一状态信息 s_t 选择下一步的动作 a_t , 并依据奖励函数获取相应的奖励。同时, 智能体会将当前状态、所采取的动作以及获得的奖励一并存入经验池中, 以便通过经验回放机制来更新神经网络。之后, 智能体执行动作 a_t , 网络状态随之更新为 s_{t+1} 。通过这一连串的精准确操作和不断优化, 智能体最终能够成功找到一条从源节点到目的节点的路径 P 。

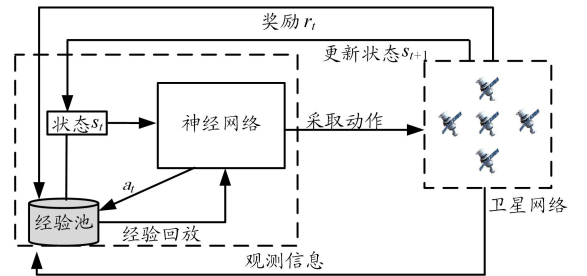


图2 DQN-PWM 路由的 DQN 结构

2.2 ALWMD-WA 算法

为了实现业务的传输, 不仅需要为到达业务请求找到一条源、目的节点之间的路径, 同时还需要在此路径上分配波长资源。对于 OSN 的波长分配, 需要满足波长一致性的约束, 即在一条路径的所有链路上选择的波长必须相同。在一条路径上, 空闲波长资源的对齐程度越高, 越容易选择合适的波长资源建立光路。因此, 在为当前业务请求选择波长资源时, 应尽可能使分配波长后路径上的链路与其相邻链路上的占用波长保持对齐, 从而为后续业务请求留下更多可用的空闲波长。

对于选定的业务请求路径, 存在其相邻链路集。如果在该路径上选择一个波长来建立光路, 能够使其相邻链路集中保留更多的共同空闲波长, 这将有利于

艾琪洋,何荣希,林润峰:卫星光网络中基于 DQN 的匹配度感知波长路由算法

后续业务请求成功建立光路,进而降低后续业务的阻塞率。因此,在所选路径上为当前业务请求分配波长时,所选波长会影响到该路径上相邻链路的可用波长资源的对齐情况,继而会影响到后续到达业务请求成功分配波长。为了更好地描述路径与其相邻链路之间的波长资源对齐情况,本文提出相邻链路波长匹配度指标,定义为

$$M_p = \sum_{a=1}^A \sqrt[B]{\prod_{b=1}^B \left(\sum_{i=1}^n \left(\lambda_{e_p^a}^i \times \lambda_{e_p^{a,b}}^i \right) \right)} \quad (8)$$

其中, M_p 表示路径 $P(o, d)$ 的相邻链路波长匹配度, e_p^a 表示路径 $P(o, d)$ 上的第 a 条链路, A 表示该路径包含的所有链路数, B 表示链路 e_p^a 的相邻链路数, $e_p^{a,b}$ 表示链路 e_p^a 的第 b 条相邻链路, $\lambda_{e_p^a}^i$ 与 $\lambda_{e_p^{a,b}}^i$ 分别表示链路 e_p^a 和 $e_p^{a,b}$ 上第 i 个波长的使用情况。

算法在选择波长建立光路的过程中,总是选择波长分配后 M_p 值最大的方案建立当前连接。本文以图 1 为例来说明路径 P 与其相邻链路波长匹配度的计算过程。路径 $P(1-2-5)$ 由链路 $e(1,2)$ 和 $e(2,5)$ 组成。链路 $e(1,2)$ 的相邻链路为 $e(2,3)$ 、 $e(2,5)$ 和 $e(1,4)$, 链路 $e(2,5)$ 的相邻链路为 $e(1,2)$ 、 $e(2,3)$ 、 $e(4,5)$ 、 $e(5,6)$ 和 $e(5,8)$ 。因此,路径 P 的相邻链路集 $E=\{e(1,2), e(2,3), e(2,5), e(1,4), e(4,5), e(5,6), e(5,8)\}$ 。当路径 $P(1-2-5)$ 选择波长 λ_2 时,链路 $e(1,2)$ 与相邻链路 $e(2,3)$ 、 $e(2,5)$ 和 $e(1,4)$ 具有相同空闲的波长数分别为 2、3、2,则路径 $P(1-2-5)$ 上第一条链路与其相邻链路波长匹配度为 $\sqrt[3]{2 \times 4 \times 2} = 2.52$; 同样地,第二条链路与其相邻链路的波长匹配度为 2.352。因此,路径 P 选择波长 λ_2 后的相邻链路波长匹配度 M_p 为两者之和,即 $M_p=2.52+2.352=4.872$ 。类似地,可求出选择其它空闲波长时的相邻链路波长匹配度。路径 $P(1-2-5)$ 上空闲波长的相邻链路波长匹配度如表 1 所示。在该路径所有可用波长中,选择 λ_7 后的相邻链路波长匹配度达到最大,因此应选择 λ_7 为该业务请求建立光路。

表 1 路径 $P(1-2-5)$ 上空闲波长的相邻链路波长匹配度

链路或 M_p	匹配度				
	λ_2	λ_5	λ_7	λ_9	λ_{10}
$e(1,2)$	2.520	2.00	2.520	2.520	1.587
$e(2,5)$	2.352	2.491	3.104	2.702	2.702
M_p	4.872	4.491	5.624	5.222	4.289

2.3 DQN-MD-RWA 算法描述和复杂度分析

DQN-MD-RWA 算法旨在利用 DRL 的决策能力,让智能体在不断试错中学习,从而找到奖励最大的方案,最终完成路由决策,得到一条奖励最大的路径。同时,在波长分配时,该算法会选择使路径与相邻链路波长匹配度最大的分配方案,以便为后续的业务请求留下更多的连续空闲波长资源。DQN-MD-RWA 算法主要包含 2 个核心子阶段:路由选择和波长分配。在路由选择阶段,本文采用了 DQN-PWMDR 算法。当有业务请求时,算法会首先获取当前网络的状态 s_t ,然后智能体根据这个状态 s_t 选择一个路由动作 a_t 。根据式(7),智能体会获得一个相应的奖励 r_t ,并据此更新网络状态至 s_{t+1} 。最终,算法会输出一条从源节点到目的节点、奖励最大的路径 P ,详细的执行过程可参见算法 1 的伪代码描述。而在波长分配阶段,本文则运用了 ALWMD-WA 算法。对于已选定的路径 P ,算法会根据式(8)计算预分配方案后相邻链路波长匹配度 M_p ,并选择 M_p 值最大的方案作为最终的波长分配方案,具体的执行过程可参见算法 2 的伪代码描述。

算法 1 DQN-PWMDR 算法

输入 状态 s_t (伪代码中用 s 表示),动作 a_t (伪代码中用 a 表示),折扣因子 γ ,学习率 α ,源节点和目的节点,循环次数 y ,贪婪概率 ε , ε 的衰减速度 $\varepsilon_{\text{decay}}$, ε 的最小值 $\varepsilon_{\min}$,学习阈值 Γ

输出 路由路径 P

- 1) 重置环境和经验回放池 Ω ,主网络 $Q(s,a|\theta)$ and 目标网络 $Q(s,a|\theta')$ 并赋予随机权重 $\theta=\theta'$,总奖励 $R=0$, $y=1$
- 2) while $\varepsilon > \varepsilon_{\min}$ do
- 3) $\varepsilon \leftarrow \varepsilon \varepsilon_{\text{decay}}$
- 4) for $y=1,2,\dots,Y$ do
- 5) 生成一个从 0 到 1 的随机值 p ;
选择动作并根据式(4)计算 η ;
- 6) if η is 0 then
阻塞该请求, break;
- 7) 采取动作 a 并将状态 s 改变为 s' ,将 a 存入动作列表 P 中并计算奖励 r
- 8) 将 (s,a,r,s') 存入经验回放池 Ω 中;
 $R \leftarrow R+r$
- 9) if $y > \Gamma$ then
- 10) 从经验回放池 Ω 中随机抽样一组经验值

训练网络,更新权重 θ 和 $y \leftarrow y+1$

if 动作是目的节点 then

break

11) end for

12) end while

算法2 ALWMD-WA 算法

输入 DQN-PWMDR 算法选出的路由路径 P 和网络中每条星间链路的波长占用信息

输出 波长分配方案

1) if 存在候选波长 then

2) 找到路径 P 及其相邻链路的所有空闲波长

3) for 每个空闲波长 λ do

4) 根据式(8)计算 M_p

5) end for

6) 按使 M_p 最大的方案分配波长

7) else

8) 阻塞该请求

DQN-PWMDR 算法的复杂度主要取决于实现 DQN 的神经网络结构。假设 DQN 的全连接层数为 G , v_g 表示第 g 层网络的输入尺寸大小,每层网络都有一个 v_g 和一个表示数量积的矩阵 $v_g v_{g+1}$,因此 DQN-PWMDR 的

算法复杂度为 $\sum_{g=0}^{G-1} v_g v_{g+1} + \omega v_g = O\left(\sum_{g=0}^{G-1} v_g v_{g+1}\right)$ 。其中, ω 表

示对应网络层使用的激活函数的相关系数。而对于 ALWMD-WA 波长分配算法,假设 OSN 的节点数为 N ,每条链路支持的波长数为 n ,所选路径的候选链路数最大不会超过 $(N-1)$,因此其复杂度近似为 $O(Nn)$ 。因此,DQN-MD-RWA 算法的复杂度为二者相加,其大小

为 $O\left(\sum_{g=0}^{G-1} v_g v_{g+1} + Nn\right)$ 。

3 仿真分析

本文对铱星星座进行仿真,先使用 STK 仿真软件生成相应的网络拓扑和连接关系,在此基础上再使用 Python 语言对所提算法性能进行验证,具体的仿真参数如表 2 所示。

在仿真过程中,每种业务的到来服从泊松分布,所有卫星节点中,业务请求的源节点和目的节点对是随机选取的。每个卫星节点没有排队等待机制,当业务请求无法建立光路时,立即被阻塞。

为了验证所提出 DQN-MD-RWA 算法的有效性,本文将其与 DRL-RWA 算法^[12]、K 最短距离路径算法+首次命中 (KSP-FF) 算法、DQN-PWMDR-FF 算法、KSP-ALWMD-WA 算法进行比较,性能指标包括业务

表 2 仿真参数设置

参数	数值
拓扑结构	π 型
卫星数/个	66
轨道数/个	6
每个轨道上的卫星数/个	11
轨道高度/km	780
轨道倾角/(°)	86
运行周期/min	114
轨内星间链路数/条	60
轨间星间链路数/条	55
星间链路支持波长数/个	10

阻塞率、波长利用率、时延。其中 DRL-RWA 算法也是针对 OSN 设计的一种基于 DRL 的 RWA 算法,它将时延、波长利用率等因素作为奖励函数的因子进行选路,在分配波长时按照业务等级划分进行分配。由于本文不考虑对业务进行分级,因此所有业务为相同等级。

不同算法的业务阻塞率随网络负载的变化如图 3 所示。可以看出,所有算法的阻塞率都随着网络负载的增大而增大。在网络负载较小时,所提算法与对比算法的阻塞率相差无几。然而,随着网络负载的增大,KSP-FF 算法的阻塞率增长速度最快。这是因为 KSP-FF 算法仅考虑路径长度进行选路,导致所选路径逐渐拥塞,进而阻塞率迅速上升。而 KSP-ALWMD-WA 算法的阻塞率增长率较小。这是因为虽然两者都使用 KSP 选路,但 KSP-ALWMD-WA 算法采用了基于相邻链路波长匹配度感知的波长分配算法。这种算法使得分配波长后所选路径的链路和相邻链路的波长占用情况更加整齐,从而为后续业务留出更多满足一致性约束的波长。因此,当有业务经过时,KSP-ALWMD-WA 算法更有可能成功建立光路,阻塞率相对较低。

此外,DRL-RWA 算法的阻塞率增长相较于 KSP-

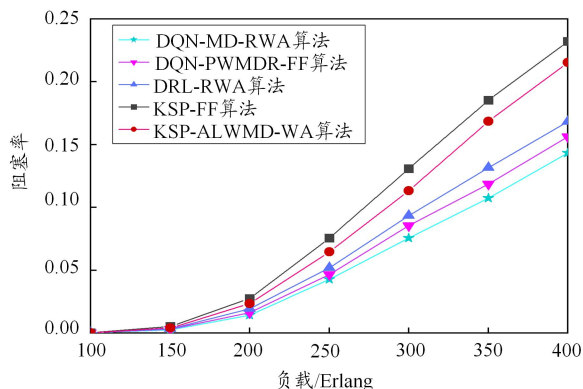


图 3 不同算法的业务阻塞率随网络负载的变化情况

艾琪洋,何荣希,林润峰:卫星光网络中基于DQN的匹配度感知波长路由算法

ALWMD-WA 算法更为缓慢。这是因为在选路过程中,该算法的智能体考虑了时延、波长利用率等因素,并通过与环境进行交互学习来选择奖励最大的路径。由于波长分配是在选路完成后进行的,因此选出的路径质量对于光路能否成功建立具有至关重要的作用。实际上,选路算法的影响往往超过了波长分配算法。

最后,相较于 DRL-RWA 算法,DQN-PWM-DR-FF 算法在路径选择上更加细致地考虑了波长资源的连续性。在选路阶段,该算法就注重优化波长资源的分配效率,从而有效降低了阻塞率。而 DQN-MD-RWA 算法则采用了基于相邻链路波长匹配度的波长分配方式,在分配波长时为后续到来业务留下更多满足一致性约束的波长资源,与基于 FF 算法的 DQN-PWM-DR-FF 算法相比,其在降低网络阻塞率方面的表现更加出色。

不同算法的波长利用率随网络负载的变化情况如图 4 所示。可以看出,在负载较小时,每种算法的波长利用率差别不大,所有算法的波长利用率都会随着网络负载的增大而增大。这是由于业务的密度增大,所需建立的光路也会增多,因此链路上使用的波长数也会增大。当负载较大时,由于 KSP-FF 算法、KSP-ALWMD 算法会出现网络拥塞,因此波长利用率增长缓慢。DQN-MD-RWA 算法由于在选路时综合考虑了路径波长匹配度和相邻链路波长匹配度,有利于建立更多光路,使得其在波长利用率方面较 DRL-RWA 算法和 DQN-PWM-DR-FF 算法有更好的表现。

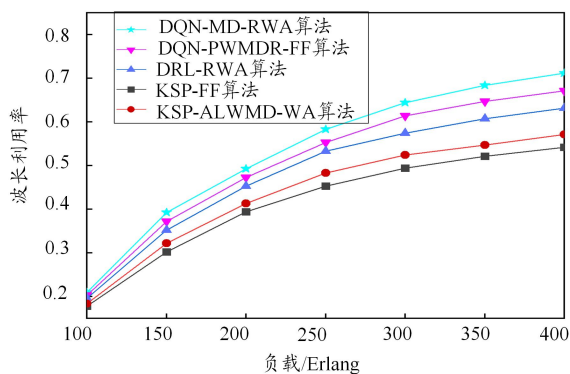


图4 不同算法的波长利用率随网络负载的变化情况

不同算法的时延随网络负载的变化情况如图 5 所示。可以看出,5 种算法的平均传输时延都随着网络负载的增加而呈上升趋势。这是因为当网络负载增大、业务越来越密集时,网络中的最短路径资源会优先被占用。为了成功建立光路,这些算法需要通过绕行部分卫星节点和星间链路的方式来选择路径。

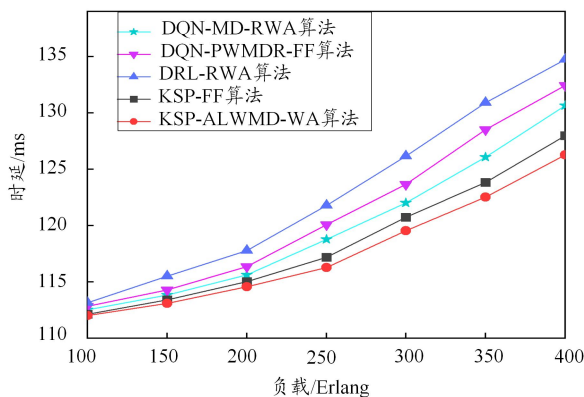


图5 不同算法的时延随网络负载的变化情况

当负载较小时,所有算法的传输时延相近,此时每种算法选择的路径都接近于最短路径。其中,KSP-FF 算法和 KSP-ALWMD-WA 算法的时延始终最小。这是因为 KSP 算法在选路时只考虑选择最短的路径。同时,KSP-ALWMD-WA 算法的时延始终小于 KSP-FF 算法的时延。这是因为使用 ALWMD-WA 算法可以为后续业务预留更多满足一致性约束的波长,从而有更高概率在候选路径中使用距离最短的路径建立光路。另外 DQN-MD-RWA 算法和 DQN-PWM-DR-FF 算法在选路时会避开负载较大的链路,并优先选择时延较小的链路。特别是 DQN-MD-RWA 算法,在波长分配时还考虑了相邻链路波长匹配度,对网络进行了负载均衡。因此,当负载增大时,DQN-MD-RWA 算法的时延表现更为优秀。

4 结束语

为了解决 OSN 中复杂的路由与波长分配问题,本文提出了一种 DQN-MD-RWA 算法。该算法综合考虑了网络拓扑和负载信息,通过不断与环境交互来调整路由的选择,获得采取动作的最大化奖励,从而得到最佳的选路决策。随后,在所选路径上根据相邻链路匹配度,选择合适的波长以建立业务连接。仿真结果表明,与 DRL-RWA 算法、KSP-FF 算法、DQN-PWM-DR-FF 算法、KSP-ALWMD-WA 算法相比,本文所提算法在降低网络阻塞率和时延,以及提高波长资源利用率方面的性能最优。

参考文献:

- [1] JIANG W. Software defined satellite networks: a survey [J]. Digital Communications and Networks, 2023, 9(6): 1243-1264.
- [2] 张家铭. 卫星激光通信网络研究进展 [J]. 光通信技术, 2023, 47(5): 37-44.

- [3] 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, 2001, 9(2): 137–147.
- [4] DONG Y, ZHAO S, DAN RAN H, et al. Routing and wavelength assignment in a satellite optical network based on ant colony optimization with the small window strategy[J]. Journal of Optical Communications and Networking, 2015, 7(10): 995–1000.
- [5] 石晓东,李勇军,赵尚弘,等. 软件定义卫星光网络蚁群优化波长路由技术[J]. 红外与激光工程, 2020, 49(10): 211–218.
- [6] LIU Y, ZHANG Q, XIN X, et al. Congestion avoidance based on artificial bee colony algorithm for routing and wavelength assignment in optical satellite networks[J]. Optical Engineering, 2021, 60(3): 036109-1–036109-14.
- [7] HE F, CHEN D, ZHANG Q, et al. Fitness optimization RWA algorithm based on bee colony in optical satellite network [C]//IEEE. Proceedings of 2022 20th International Conference on Optical Communications and Networks (ICOON). Shenzhen: IEEE, 2022: 1–3.
- [8] WEI Y, ZHU R, ZHANG Y, et al. Routing based on dynamic reliability in massive LEO satellite optical networks [C]//IEEE. Proceedings of 2023 Opto-Electronics and Communications Conference(OECC). Shanghai: IEEE, 2023: 1–3.
- [9] WEN G, ZHANG Q, WANG H, et al. An ant colony algorithm based on cross-layer design for routing and wavelength assignment in optical satellite networks[J]. China Communications, 2017, 14(8): 63–75.
- [10] SUN X, CAO S. A routing and wavelength assignment algorithm based on two types of LEO constellations in optical satellite networks [J]. Journal of Lightwave Technology, 2020, 38(8): 2106–2113.
- [11] BOYAN J A, LITTMAN M L. Packet routing in dynamically changing networks: a reinforcement learning approach [C]//Margan Kaufmann. Proceedings of the 6th International Conference on Neural Information Processing Systems. Denver Colorado: Margan Kaufmann, 1993: 671–678.
- [12] 李信,李勇军,赵尚弘. 基于深度强化学习的卫星光网络波长路由算法[J]. 系统工程与电子技术, 2023, 45(1): 264–270.
- [13] ZUO P, WANG C, WEI Z, et al. Deep reinforcement learning based load balancing routing for LEO satellite network [C]//IEEE. Proceedings of 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring). Helsinki: IEEE, 2022: 1–6.