

引用:孔凌兴,望运武,顾家骅,等. 弹性光网络中面向提前预约服务请求的生存性感知路由频谱分配算法[J]. 聊城大学学报(自然科学版), 2026, 39(2): 205-216.

Citation: KONG Lingxing, WANG Yunwu, GU Jiahua, ZHU Min. A survivable routing and spectrum assignment algorithm for advanced reservation service requests in elastic optical networks[J]. Journal of Liaocheng University (Natural Science Edition), 2026, 39(2): 205-216.

文章编号 1672-6634(2026)02-0205-12

DOI 10.19728/j.issn1672-6634.2024110008

弹性光网络中面向提前预约服务请求的 生存性感知路由频谱分配算法

孔凌兴¹, 望运武², 顾家骅², 朱敏²

(1. 东南大学 集成电路学院, 江苏 南京 210000; 2. 东南大学 移动通信国家重点实验室, 江苏 南京 210096)

摘要 多路径保护能够有效提高光路请求的服务生存能力, 因其通过多条不相交的路径进行数据传输, 无论哪条路径发生故障也不会导致服务中断。这篇文章主要研究了基于深度强化学习(DRL)的弹性光网络(EON)中面向提前预约(AR)服务请求的生存性感知路由调制频谱分配(RMSA)问题, 其中 AR 请求不仅可以在服务请求到达时刻立即分配网络资源, 也可以选择在未来时刻中分配网络资源, 具有提前预约与等待的特点。为了有效地确定 AR 服务请求的多路径路由、调制格式、频谱分配以及调度时间, 本文提出了一种基于 DRL 的动态资源分配算法。在 14 节点网络拓扑上的仿真结果表明, 所提 DRL 算法在 300 Erlangs 的流量负载下, 减少了高达 31.9% 的阻塞概率。

关键词 多路径保护; 深度强化学习; 弹性光网络

中图分类号 TN929.1

文献标志码 A

开放科学(资源服务)标识码(OSID)



A survivable routing and spectrum assignment algorithm for advanced reservation service requests in elastic optical networks

KONG Lingxing¹, WANG Yunwu², GU Jiahua², ZHU Min²

(1. School of Integrated Circuits, Southeast University, Nanjing 210000, China; 2. State Key Laboratory of Mobile Communications, Southeast University, Nanjing 210096, China)

Abstract Multi-path protection can effectively enhance the service survivability of light-path requests by transmitting data through multiple disjoint paths. This approach ensures that even if one path fails, the service remains uninterrupted. This paper focuses on the survivability-aware routing, modulation, and spectrum assignment (RMSA) problem in elastic optical network (EON) based on deep reinforcement learning (DRL) for advanced reservation (AR) service requests, where AR requests not only allow for immediate allocation of network resources upon arrival but also enable scheduling at a future time, character-

收稿日期: 2024-11-18

基金项目: 国家自然科学基金项目(62271135)资助

通信作者: 朱敏, 男, 汉族, 博士, 教授, 江苏省“333”高层次人才计划(第三层次)省科技领军人才, 研究方向: 光网络与光通信,

E-mail: minzhu@seu.edu.cn.

ized by advanced reservation and waiting features. To efficiently determine the multi-path routing, modulation format, spectrum allocation, and scheduling time for AR service requests, this paper proposes a dynamic resource allocation algorithm based on DRL. Simulation results on a 14-node network topology demonstrate that the proposed DRL algorithm reduces the blocking probability by as much as 31.9% at 300 Erlangs traffic load.

Keywords multi-path protection; deep reinforcement learning; elastic optical networks

0 引言

随着超高清视频、云计算和数据中心等互联网流量的快速增长,运营商对网络基础设施的质量要求也不断提升,需同时确保基础设施的传输质量(QoT)、低延迟和无服务中断^[1]。传统的波分复用(WDM)光网络采用固定的均匀间隔,通常为 50 GHz 或 100 GHz,导致频谱利用效率较低。相比之下,弹性光网络(EON)提供了一种更具前景的解决方案,能够自适应地分配频谱资源,以满足下一代应用的动态和多样化需求^[2]。此外,EON 采用更细的频谱时隙(例如 6.25 GHz 或 12.5 GHz),灵活组合多个时隙以创建更宽的信道,从而满足光路请求^[3]。在 EON 中服务光路请求时,需满足频谱一致性、连续性和不重叠性的约束^[4],这些限制大大增加了 EON 中频谱资源分配的复杂性,给设计高效的光路请求分配方案带来了巨大挑战。

最近的研究主要集中在基于预定义传输距离(TR)限制的 EON 频谱资源分配上。每种调制格式都有具体的 TR 限制,大多数链路需要预留较大资源以保证 QoT,这种过度冗余配置^[5]往往导致频谱资源利用效率低下。为解决这一问题,有一些方案可供选择,例如通过精确考虑物理层损伤,或采用高斯噪声(GN)模型替代传统的基于 TR 的方法^[6]来提升频谱效率。最近,文献[7]考虑了一种非线性损伤(NLI)效应的传输模型。NLI 效应包括加性自发发射(ASE)、自通道干扰(SCI)和跨通道干扰(XCI)。其中,ASE 是光纤放大器产生的噪声,SCI 来自通道自身的非线性干扰,XCI 则源于两个通道间的非线性交互作用^[8]。

在 EON 中的生存性感知频谱资源分配面临两个关键挑战:(1)路由和频谱分配(RSA),需满足频谱一致性、连续性和非重叠性的约束条件^[9];(2)需考虑 NLI 效应的调制格式分配。这也是 EON 中生存性感知路由调制频谱分配算法(RMSA)的核心问题。文献[10]在多芯光纤 EON 中提出了一种物理层损伤感知的 RMSA 优化框架,有效提升了频谱碎片的细化水平,然而该研究主要针对静态光路请求,未考虑动态请求。文献[11]针对数据中心间 EON 中动态场景下的全光组播 RMSA 问题进行了研究,主要面向组播的分布式子树方案,提出了一种改进和优化该方案的算法,有效降低了网络中组播业务的总体阻塞率,然而,该研究未考虑 NLI 因素。文献[12]在 EON 支持的接入网(RAN)中,针对多播服务分布式单元(DU)和集中单元(CU)的 RMSA 问题进行了探讨,并提出了一种混合整数线性规划(MILP)模型,以最小化 EON 支持的 5G/B5G RAN 中多播服务的总能耗。为克服现有局限并进一步降低网络阻塞率,文献[13]在奖励函数中融入了资源占用情况和频谱邻接情况两个关键指标,提出了一种基于改进奖励设计的基于深度强化学习(DRL)的 RMSA 算法,然而,该算法未能考虑光网络的生存性问题。文献[14]围绕下一代灵活光网络中非线性损伤对网络容量的限制问题展开研究,重点探讨了非线性效应引起的非线性噪声损伤的建模、预测与补偿,以及级联可重构光分插复用器(ROADM)窄带滤波效应导致的滤波损伤预测,该研究提出了针对非线性噪声的建模模型和补偿方案,并建立了针对 ROADM 滤波损伤的预测模型。文献[15]则专注于基于相干光正交频分复用的 EON 中,路由频谱分配的时延带宽积、碎片率及链路拥塞的离线联合优化问题,并提出了一个最优的 MILP 模型,在考虑传输损伤的情况下,能够实现负载均衡,并最小化时延带宽积和碎片率。文献[16-17]研究了动态请求下的 NLI 感知 RMSA 算法,尽管其集中于动态请求,但这些请求均为即时预约服务请求,初始延迟容限为零,即请求到达时要么立即服务,要么被阻塞。除即时预约请求外,光网络中还存在提前预约(AR)请求^[18],此类请求在预定义的截止时刻前分配资源即可,因此在设置过程中允许一定的初始延迟。

在光传输网络中,服务请求的生存性是至关重要的。文献[19]提出了一种启发式算法来设计生存性感知波分复用(WDM)网络。文献[20]针对基于正交频分复用(OFDM)的 EON 提出了一种生存性感知多路径配置方案,该方案可实现完全保护和部分保护,相较于传统的单路径配置方案效率更高。在该方案中,请求通过多条链路不相交的路径路由,并在这些路径上分配频隙(FS),以满足请求的带宽和保护需求。然而,

上述研究主要依赖启发式算法,因此难以实现最佳的网络性能,为此,可以利用机器学习技术设计更高级的算法,其中 DRL 效果最为显著^[21-24]。文献[25]提出了一种应用于单波段光网络设计的 DRL 算法,研究表明该算法在减少频谱碎片方面优于传统的启发式算法,同时网络阻塞率也有所降低。文献[26]在全光网络场景中,通过引入波长转换器对启发式算法和 DRL 算法进行了比较,结果显示无论是否使用波长转换, DRL 算法的阻塞率始终低于启发式方法。文献[27]则将 DRL 算法与遗传算法在虚拟网络中的性能进行了对比,发现 DRL 在服务功能链(SFC)拒绝率上表现出显著优势。文献[28]聚焦于灾害场景,研究表明 DRL 算法能够更及时、准确地执行路径规划。然而,上述 DRL 算法没有考虑在 EON 中面向 AR 请求的 RMSA 问题。

在本文中,旨在通过多路径保护来解决 EON 中生存性感知 RMSA 算法问题。为了有效地将时域请求调度与频谱域的多路径 RMSA 结合起来,本文提出了一种基于 DRL 的动态资源分配算法,并通过阻塞概率、平均频谱效率、平均初始延迟容限和零延迟请求比例这四个指标对其网络性能进行了仿真评估。本文的主要贡献有:(1) 提出了一种基于 DRL 的 EON 中面向 AR 服务请求的生存性感知 RMSA 算法,旨在通过优化调制格式分配、路由和请求调度,最小化请求阻塞概率;(2) 提出了一种多路径路由分解方案,便于 DRL 代理决策,该方案将每个请求的多条保护路径分解为多条不相交的单路径;(3) 提出了一种基于 DRL 的算法,通过引入时间冻结方案来缩小动作空间,从而降低请求阻塞概率。此外,该算法采用了带优先级队列的加权排序方法,进一步优化了 DRL 代理的决策过程;(4) 通过大量仿真实验验证了基于 DRL 算法的有效性。仿真结果表明,在不同到达率和滑动因子的 AR 流量模型下,基于 DRL 的算法能够有效降低阻塞概率,减少平均初始延迟容限,并提高频谱效率。

1 系统模型

本节首先介绍了网络和请求模型,随后概述了物理层的 NLI 模型,最后定义了几个关键的性能指标。

1.1 网络模型和请求模型

本文将 EON 采用无向图 $G(V, E, F)$ 表示,其中 V 和 E 分别表示节点集合和光纤链路集合, F 表示每条链路 $e(e \in E)$ 上 FS 的使用情况。一个 AR 光路请求可以表示为 $R_i(s_i, d_i, b_i, p_i, \tau_i, t_i^a, t_i^e)$,其中 s_i 为起点, d_i 为终点, b_i 为所需的传输速率, p_i 为所需的保护等级。若链路发生故障,首先需保证 $b_i \cdot p_i$ 的容量,因此,当 $Q \geq 1/(1-p_i)$ 时,每条路径分配 b_i/Q 的容量;否则,分配的容量为 $b_i \cdot p_i/(Q-1)$,其中 Q 为链路不相交路径的个数。最后, τ_i 表示持续时间, t_i^a 和 t_i^e 分别为请求的到达时间和结束时间。在请求 R_i 中,服务时长 τ_i 是时间窗口 $[t_i^a, t_i^e]$ 内的一个可预约的时间间隔。请求 R_i 的起始时间相对时间窗口 $[t_i^a, t_i^e]$ 的左边界 t_i^a 是可变的。若请求 R_i 在 t_i^a 开始后等待 t_i^w 个时间单位,则该请求将在时间区间 $[t_i^a + t_i^w, t_i^a + t_i^w + \tau_i]$ 期间保持激活,其中 $t_i^a + t_i^w + \tau_i \leq t_i^e$ 。空闲时间 $t_i^s(t_i^e - t_i^a - \tau_i)$ 表示请求 R_i 的最大等待时间,这一时间可以使服务推迟,以容纳更多的请求。AR 请求滑动因子定义为空闲时间与时间窗口长度之比,即 $s_f = t_i^s/(t_i^e - t_i^a)$ 。当 $t_i^s = 0$ 时,表示请求的起始时间和结束时间与时间窗口的左右边界重合,在这种情况下,请求必须在到达后立即处理,否则将过期并导致阻塞。在 AR 流量模型中,滑动因子(s_f)通常大于 0,这意味着请求的起始和结束时间可以根据当前 EON 的状态灵活调整。利用这种灵活性,可以在时间窗口内将请求的起始时间调整为更早或更晚的时刻,以避免资源竞争,减少请求阻塞。

在容纳请求 R_i 时,需要计算端到端路径 $p_{s,d}$,并选择合适的调制格式 m 以确保 QoT。在此基础上,根据传输速率 b 和调制格式 m ,为路径 $p_{s,d}$ 上的每条光纤链路分配一组频谱一致且连续的 FS。根据文献[29],分配给 R_i 的 FS 数量 n_i 可以通过公式(1)计算

$$n_i = \left\lceil \frac{b_i}{m_i \cdot B_{\text{PM-BPSK}}} \right\rceil, \quad (1)$$

其中 $B_{\text{PM-BPSK}}$ 表示采用偏振复用二进制相移键控(PM-

表 1 可用的调制格式及其线性阈值

Tab. 1 Available modulation formats and their linear scale thresholds

调制格式	$m/(\text{bit} \cdot \text{s}^{-1} \cdot \text{Hz}^{-1})$	S^{th}/dB
PM-BPSK	2	3.52
PM-QPSK	4	7.03
PM-8QAM	6	17.59
PM-16QAM	8	32.60

BPSK)技术的数据速率,参数 m 表示 M 中的调制格式等级,从集合 $M = [2, 4, 6, 8]$ 中选择,

其中分别对应 PM-BPSK、PM-正交相移键控(QPSK)、PM-8 正交调幅(QAM)和 PM-16 QAM。考虑到不同调制格式的频谱效率,本文使用 m_i 表示分配给请求 R_i 链接的相应调制格式。表 1^[7]列出了在特定前向纠错(FEC)比特误码率(本文中为 4×10^{-3})下,每种可用调制格式的 $m(m \in M)$ 所需的最小信噪比(SNR)阈值 S_m^{th} 。

1.2 物理层 NLI 模型

链接的 QoT 可以使用 GN 模型^[7]进行估计,GN 模型是用于计算 NLI 色散无补偿链路的解析模型。通过对不同类型的 NLI(包括 ASE 噪声、SCI 和 XCI)进行详细梳理,可以计算出请求 R_i 在每个链路上的 SNR:

$$S_i = \frac{G_i}{G_i^{\text{ASE}} + G_i^{\text{SCI}} + G_i^{\text{XCI}}}, \quad (2)$$

其中

$$G_i^{\text{ASE}} = (e^{aL} - 1)h\nu n^{\text{sp}} \cdot N_i^{\text{span}}, \quad (3)$$

$$G_i^{\text{SCI}} = \frac{3\gamma^2 G_i^3}{2\pi\alpha |\beta_2|} \ln\left(\frac{\pi^2 |\beta_2|}{\alpha} \Delta f_i^2\right) \cdot N_i^{\text{span}}, \quad (4)$$

$$G_i^{\text{XCI}} = \frac{3\gamma^2 G_i}{2\pi\alpha |\beta_2|} \sum_{j \neq i} G_j^2 \cdot \ln\left|\frac{\Delta f_{ij} + \Delta f_j/2}{\Delta f_{ij} - \Delta f_j/2}\right| \cdot N_{ij}^{\text{span}}. \quad (5)$$

相关参数定义如表 2 所示,有兴趣的读者可参考文献[7]了解更多细节。

表 2 SNR 计算的符号和含义

Tab. 2 Symbol and meaning for SNR calculation

符号	含义	符号	含义
G_i	信号功率谱密度(PSD)	n^{sp}	自发辐射因子
G_i^{ASE}	ASE 的 PSD	ν	光载波频率
G_i^{SCI}	SCI 的 PSD	L	每个跨度的长度
G_i^{XCI}	XCI 的 PSD	Δf_i	请求 R_i 使用的带宽
α	功率衰减	Δf_{ij}	R_i 和 R_j 占用的带宽之间的间隔
γ	光纤非线性系数	N_i^{span}	每个请求 R_i 的路径所经过的跨度
β_2	光纤色散	N_{ij}^{span}	两个请求 R_i 和 R_j 的两条路径所共同经过的跨度数
h	普朗克常数		

在满足请求 R_i 的调制格式 m 的 SNR 阈值条件(即: $S_i \geq S_m^{\text{th}}$)下,才能确保成功服务。在 EON 中的动态生存性感知多路径 RMSA 的目标是尽量降低请求阻塞概率。如公式(5)所示,请求 R_i 的光谱位置变化会导致显著的 XCI 非线性损伤(即 G_i^{XCI}),从而影响请求 R_i 的 SNR 并进一步影响其 QoT。公式(3~5)表明每个传输通道所面临的干扰受路由跨度数量以及已激活请求的影响。

在动态多路径 RMSA 的背景下,由于 AR 请求的到达和结束是动态的,所以在其到达后需要在空闲时间内及时为请求提供服务。因此,请求 R_i 不仅会受到已激活请求的 SCI 限制,还会受到未来请求可能产生的 SCI 影响,而在提供服务时,未来请求对 R_i 的具体影响尚无法预测。为此,本文假设:(1) 通过将平均到达率与当前 R_i 的服务时长相乘来估计未来请求的规模;(2) 未来请求的带宽等于所有请求的平均带宽;(3) 未来请求所需的 FS 数量基于最低调制格式计算;(4) 为未来请求分配的 FS 尽量贴近当前 R_i 最高 FS 附近的光纤链路。

1.3 优化目标

本文重点探讨了具备带宽资源的 EON 架构,并分析了其在动态业务环境中的应用。在该架构下,AR 服务请求会在不同时刻以动态的方式到达 EON 系统。每一个 AR 服务请求都包含了详细的时间参数,包括具体的到达时间、结束时间、持续时间以及空闲时间。此外,这些请求拥有确定的带宽和保护等级需求。结合 AR 业务模型的特点,本文旨在设计一种最优 AR 服务请求的多路径 RMSA 算法,在确保所有时延约

束得到满足的同时,尽可能地减少 AR 服务请求的阻塞概率。

1.4 定义性能指标

为了支持动态服务提供算法的设计,本文定义了四个性能评估指标。

1.4.1 阻塞概率。考虑到不同 AR 服务请求的带宽和时间需求不同,本文重新定义了请求阻塞概率^[30]

$$P_b = \frac{\sum_{i \in \mathbf{R}_{\text{block}}} b_i \cdot \tau_i}{\sum_{i \in \mathbf{R}} b_i \cdot \tau_i}, \quad (6)$$

其中 $\mathbf{R}$ 表示所有 AR 请求的集合, $\mathbf{R}_{\text{block}}$ 表示被阻塞请求的集合, b_i 和 τ_i 分别表示请求 $\mathbf{R}_i$ 所需的传输速率和持续时间。

1.4.2 平均频谱效率。考虑到不同请求选择了不同的路径和调制格式,平均频谱效率定义为

$$\eta = \frac{\sum_{i \in (\mathbf{R} - \mathbf{R}_{\text{block}})} (b_i \cdot \tau_i)}{\sum_{i \in (\mathbf{R} - \mathbf{R}_{\text{block}})} (n_i \cdot B_{\text{PM-BPSK}} \cdot \tau_i)}, \quad (7)$$

其中 n_i 表示请求 $\mathbf{R}_i$ 所需的 FS 数量, $B_{\text{PM-BPSK}}$ 表示单个 FS 的带宽。

1.4.3 平均初始延迟容限。在本文中,被成功服务的 AR 请求的平均初始延迟容限定义为

$$T_{\text{ID}} = \frac{\sum_{i \in (\mathbf{R} - \mathbf{R}_{\text{block}})} (t_i^w / (\tau_i^e - \tau_i^a))}{|\mathbf{R} - \mathbf{R}_{\text{block}}|}, \quad (8)$$

其中 t_i^w 表示请求 $\mathbf{R}_i$ 部署的等待时间, $(\tau_i^e - \tau_i^a)$ 表示请求 $\mathbf{R}_i$ 时间窗口的持续时间,符号 $|\cdot|$ 表示集合中元素的数量。

1.4.4 零延迟请求的比例。在 AR 流量模型中,本文使用被成功服务请求的总数作为分母来计算零延迟请求的比例,即

$$P_{\text{ZD}} = \frac{|\mathbf{R}_{\text{zero}}|}{|\mathbf{R} - \mathbf{R}_{\text{block}}|}, \quad (9)$$

其中 $\mathbf{R}_{\text{zero}}$ 是在到达时立即被成功服务(没有初始延迟)的请求集合。

2 EON 中基于 DRL 的生存性感知多路径 RMSA 算法

本节首先介绍了 EON 中复杂的生存性感知多路径 RMSA 算法,包括状态表示、动作空间和奖励函数。随后详细阐述了所提 DRL 算法的训练过程。

2.1 DRL 建模

当有足够的频谱资源容纳新到达的 AR 请求时,它们可以立即被服务;否则,请求将进入 N 个请求槽中等待,超过 N 个请求的部分会被记录在积压队列中,如图 1 所示。

DRL 算法没有将请求槽和积压队列合并为一个队列,原因有两个:(1) 合并后会显著延长等待队列,增加挂起请求的缓冲;(2) 较大的请求槽可能扩大 DRL 代理的动作空间,从而增加迭代次数,降低学习效率。在请求槽中,每个请求都有一个权重,计算公式如

$$w_i = \frac{b_i \cdot \tau_i}{t_i^r + 1}, t_i^r \geq 0, \quad (10)$$

其中 b_i 表示所需的传输速率, τ_i 表示服务的持续时间, t_i^r 表示剩余的空闲时间。该权重综合了传输速率 b_i 、服务时长 τ_i 和空闲时间 t_i^r ,并考虑了较大请求的优先级($b_i \cdot t_i^r$)。这种加权方法不仅提高了资源利用率,还降低了请求阻塞概率。此外,算法优先考虑空闲时间较短的请求,因为空闲时间归零的请求若还未被服务,就会被阻塞,从而增加了阻塞概率。基于这种加权排序,算法按计算出的权重从高到低顺序处理请求槽队列中的请求。

为了有效地确定 AR 请求的多路径 RMSA 和调度时间,本文提出了一种基于 DRL 的 EON 动态资源分配算法。DRL 中的状态、动作和奖励定义如下。

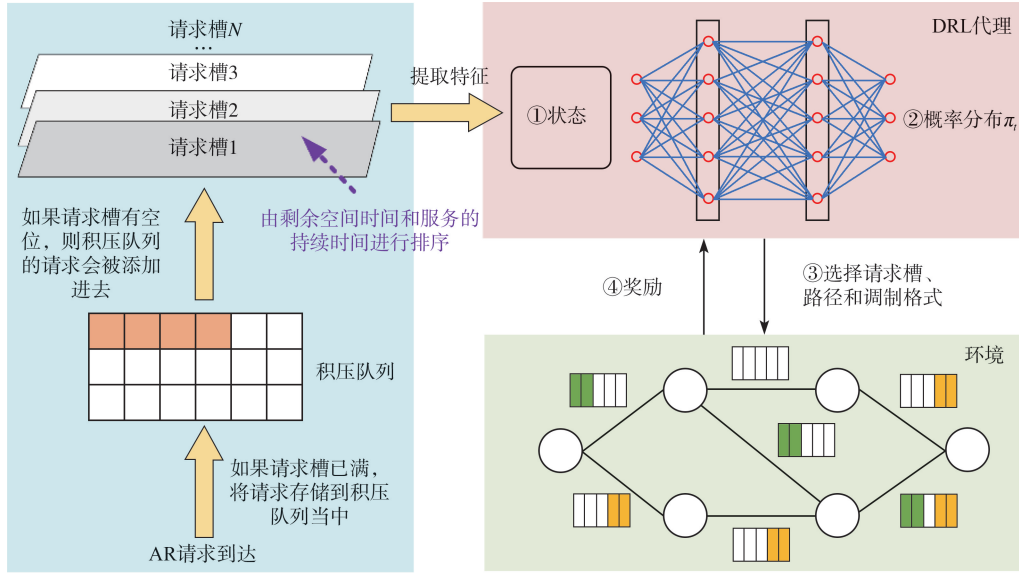


图1 基于DRL的NLI感知生存多路径RMSA算法的训练过程

Fig. 1 Training process of DRL-based NLI-aware survivable multipath RMSA algorithm

2.1.1 状态。状态可表示为一个 $1 \cdot [2 + 1 + (2 \cdot J + 3) \cdot |M| \cdot K + 2]$ 大小的数组。该数组包含选择的 AR 请求信息以及 K 条最短候选路径的频谱利用率信息。每条路径上有 J 个可用的 FS 块, 并支持 $|M|$ 种不同的调制格式。因此, 状态可以定义为以下形式:

$$S_t = \{[s_i, d_i, \tau_i, \{\{z_{k,m}^{1,j}, z_{k,m}^{2,j} \mid j \in [1, J], z_{k,m}^3, z_{k,m}^4, z_{k,m}^5 \mid m \in [1, M]\} \mid k \in [1, K]\}, c_s, c_b]\} \quad (11)$$

其中 s_i 和 d_i 分别表示起点和终点, τ_i 表示持续时间。对于路径 k ($k \in K$), 本文采用了 M 种不同的调制格式来计算 $z_{k,m}^3$, 其含义为请求 R_i 基于调制格式 m 所需的 FS 数。 $z_{k,m}^{1,j}, z_{k,m}^{2,j}, z_{k,m}^4, z_{k,m}^5$ 分别表示第 j 个可用 FS 块的大小、第 j 个可用 FS 块的索引、 J 个可用 FS 块的平均大小, 以及路径 k 上光纤链路的可用 FS 块总大小。另外, c_s 和 c_b 分别表示请求槽和积压队列中的请求数。

2.1.2 动作。为了便于 DRL 代理进行决策, 本文提出了一种多路径路由分解方案。具体而言, 每个请求的 Q 个保护路径被分解为 Q 条不相交的独立路径, 并通过 Q 个步骤为其提供服务。在本文中, 时间被离散化为多个时隙, 每个时隙内, DRL 代理会调度多个 AR 服务请求。如果需要调度的请求数量为“ N ”, 那么 DRL 代理的动作空间将达到 $(K \cdot M + 1)^N$ 的规模, 这会使 DRL 代理在学习过程中面临较大的挑战, 其中 K 表示可用候选路径的数量, M 表示候选调制格式的数量。为了解决这一问题, 算法通过允许 DRL 代理在单一时隙内执行多个动作来减小动作空间, 这一策略被称为“时间冻结步”^[31]。在一个冻结步中, 当一个请求被选中时, DRL 代理将从 K 条可用候选路径中选择一条路由路径, 接着从 M 种调制格式中选择一种, 最后在所选路径上从 J 个 FS 块中分配一个 FS 块。

在本文中, 动作空间表示为 $\{\emptyset, (1, 1), (1, 2), \dots, (1, |M|), (2, 1), \dots, (K, |M|)\}$, 其中 $a = (k, m)$ 表示将请求调度到第 k 条路径和第 m 种调制格式; $a = \emptyset$ 表示空动作, 即 DRL 代理不调度当前请求。一旦请求成功调度, 它将从请求槽中删除。若积压队列非空, 积压队列中的最前请求将填充至请求槽。若执行空动作或无效动作 (例如选择的路径和调制格式不符合 SNR 要求), DRL 代理将跳出当前冻结步。本文提出的基于 DRL 的算法仅用于路由和调制格式分配, 频谱槽资源采用经典的首次拟合 (FF) 算法进行分配。

2.1.3 奖励。在 DRL 代理采取行动后, 会立即获得奖励 r_t 。本文的主要目标是尽量减少阻塞请求的数量, 因此, 在完成一次训练回合后, 奖励 r_t 定义为所有成功服务的请求的总时间, 即

$$r_t = \frac{\sum_{i \in \{\mathbf{R} - \mathbf{R}_{\text{block}}\}} b_i \cdot \tau_i}{\sum_{i \in \mathbf{R}} b_i \cdot \tau_i}, \quad (12)$$

若训练回合未完成, 则 $r_t = 0$ 。其中, $\mathbf{R}$ 表示所有 AR 请求的集合, $\mathbf{R}_{\text{block}}$ 表示被阻塞的请求集合。奖励在每

个训练回合结束时计算,并通过策略梯度(PG)训练传递给训练过程中的每个动作,从而确保全局最优性。

2.2 训练过程

图1展示了EON中基于DRL的生存性感知多路径RMSA算法的训练过程,其中DRL代理被表示为深度神经网络(DNN),也称为策略网络。当请求 R_i 在时刻 t 到达时,策略网络将 R_i 的状态和当前网络状态合并为状态 s_t 作为输入(步骤1);随后,策略网络在动作空间 A 上生成概率分布 $\pi_t(A|s_t, \theta)$,其中 θ 表示DNN的策略参数(步骤2);接着,DRL代理根据生成的概率分布 π_t 在动作空间 A 中选择一个动作 $a_t \in A$,并尽力为请求 R_i 建立光路(步骤3);最后,代理根据奖励 r_t 对策略进行反馈更新(步骤4)。因此,三元组 (s_t, a_t, r_t) 被用作训练DRL代理的输入。

本文提出的基于DRL的生存性感知多路径RMSA算法在EON中的训练过程如算法1所示。第1行初始化了策略网络;为确保充分训练,算法通过随机生成多个请求集合,对每组进行 E 个训练回合,生成了不同的环境代理交互样本集合,之后初始化缓冲区集合、存放数据的索引和算法状态(第2~6行);为了便于算法的表达,本文将算法的状态划分为“待分配状态”、“更新状态”和“空闲状态”三种类型。如果当前算法状态为“待分配状态”时,算法将优先保持此状态,若为“空闲状态”,则根据DRL智能体与请求候选区和积压区的反馈,决定下一个状态是转为“待分配状态”还是“更新状态”(第7~13行)。在“待分配状态”下,算法会为当前请求分配路径,并维持该状态直到满足生存性要求;而在“更新状态”下,算法会将当前请求更新为下一个请求,并推进至下一个时间步,同时进入“空闲状态”。处理完这两个状态后,算法会跳转回第8行,进行下一个状态的判断。由于所有决策可以在同一时间步内完成,这一过程被称为“时间冻结步”^[31](第14~20行)。环境和代理的所有交互 (s, a, r) 被存储在缓冲区 Δ_{ep} 中,以用于后续神经网络参数 θ 的更新(第21行);当所有请求分配完毕,满足回合终止条件时,当前训练回合的数据收集结束(第7行);在数据收集结束后,开始对折扣值奖励进行计算,其中 r_t 由公式(12)获得,其他时刻都为0(第23行);随后,算法使用经典的PG算法^[31]通过存储的 (s, a, r) 数据集迭代计算返回值 v 、基准 b ,并更新策略网络参数 $\Delta\theta$,最终更新 θ ,调整策略网络,使其输出的动作概率分布倾向于选择能够最大化长期回报折扣值奖励的动作(第24~32行)。

算法1 基于DRL的生存性感知多路径RMSA算法在EON中的训练过程

1	初始化策略网络参数 θ
2	for 每一次迭代, do
3	网络更新参数 $\Delta\theta$ 赋值为0
4	for 每一个请求集合, do
5	运行每一个回合 $e = 1, 2, \dots, E$
6	初始化缓冲区集合 $\Delta_e = \{\}$, 存储训练数据的索引 $l_e = 1$, 当前算法状态为“空闲状态”(一共有“更新状态”, “待分配状态”, “空闲状态”这3个状态)
7	while 请求数量没有达到目标数量或且请求候选区和请求积压区不为空, do
8	if 当前算法状态为“待分配状态”, then
9	维持“待分配状态”
10	else if 选择了一个空动作或无效的动作, 或者请求候选区和请求积压区为空, then
11	当前算法状态跳转到“更新状态”
12	else
13	当前算法状态跳转到“待分配状态”
14	if 当前算法状态为“待分配状态”, then
15	if DRL智能体选择的路径和调制格式合理, then
16	当前请求的可用路径加1, 更新路径上的频谱, 维持“待分配状态”
17	if 可用路径的数量达到要求, then
18	当前请求更新到下一个请求, 当前算法状态跳转到“空闲状态”

```

19     else
20     更新至下一个时间冻结步,产生请求,当前算法状态跳转到“空闲状态”
    记录当前时间冻结步中,DRL 智能体与 EON 环境的交互样本  $(s, a, r)$ ,以  $l_e$  为索引,将  $(s, a, r)$  赋值给
21  $(s_{l_e}^e, a_{l_e}^e, r_{l_e}^e)$  并存储至缓冲区  $\Delta_e$ 。然后,索引  $l_e$  加 1
22     end
23     计算折扣值奖励:  $v_l^e = \sum_{k=l}^{l_e} \gamma^{k-l} r_k$ , 当  $k$  为  $l_e$  时  $r_k = \frac{\sum_{i \in \{\mathbf{R}-\mathbf{R}_{\text{block}}\}} b_i \cdot \tau_i}{\sum_{i \in \mathbf{R}} b_i \cdot \tau_i}$ , 当  $k$  为其他值时,  $r_k = 0$ 
24     for 每一个索引  $l = 1$  到  $\max\{l_1, l_2, \dots, l_E\}$ , do
25         计算基准:  $b_l = \frac{1}{E} \sum_{e=1}^E v_l^e$ 
26         for 每一个  $e$  到  $E$ , do
27              $\Delta\theta = \Delta\theta + \alpha \cdot \nabla_{\theta} \log \pi_{\theta}(s_l^e, a_l^e)(v_l^e - b_l)$ 
28         end
29     end
30 end
31  $\theta = \theta + \Delta\theta$ 
32 end

```

3 仿真分析

3.1 仿真环境设置

本文所提 DRL 算法的网络性能基于 14 节点的 NSFNET 拓扑进行了评估,如图 2 所示^[32]。物理损伤的仿真参数设定为: $\alpha = 0.22 \text{ dB/km}$, $\gamma = 1.3 (\text{W} \cdot \text{km})^{-1}$, $\beta_2 = -21.7 \text{ ps}^2/\text{km}$, $n^s = 1.58$, $v = 193.55 \text{ THz}$, $L = 100 \text{ km}$ 。所有请求链接的 PSD 均假定为 15 mw/THz (即 $G = 15 \text{ mw/THz}$)。底层 EON 中的每条光纤链路可容纳 358 个 FS^[33],每个 FS 的带宽为 12.5 GHz 。到达的 AR 请求有平均速率 λ 和平均服务时长 u 这两个特性,两者分别服从泊松分布和指数分布^[32]。EON 系统的流量负荷用 Erlangs 表示为 $\rho = \lambda/u$,每个请求的带宽均匀分布在 $80 \sim 320 \text{ Gb/s}$ 范围内。对于 DRL 训练,学习率设置为 0.001,折扣率为 0.99。为便于网络性能的对比,本文使用了四种经典的启发式算法,即 K-最短路径-最先匹配算法(KSP-FF)、K-最短路径-最优匹配算法(KSP-BF)、最大频谱槽指数路径-最先匹配算法(MFSI-FF)和最大频谱槽指数路径-最先匹配算法(MFSI-BF),其中 K 的值为 3。仿真请求数量设定为 1 000,滑动因子设定为 0.2,请求槽的个数取值为 5,保护等级设为 0.8。本文其它参数的设置如表 3 所示。

表 3 仿真参数数值

Tab. 3 Values of simulation parameter

仿真参数	数值	仿真参数	数值
α	0.22 dB/km	λ	5、7.5、10、12.5、15、17.5、20
γ	1.3 W/km	学习率	0.001
β_2	-21.7 ps ² /km	折扣率	0.99
n^s	1.58	候选路径数量	3
v	193.55 THz	候选调制格式数量	4
L	100 km	仿真请求数量	1 000
PSD	15 mw/THz	滑动因子	0.2
光路上频谱槽数量	358	请求槽的个数	5
每个 FS 的带宽	12.5 GHz	保护等级	0.8
u	20		

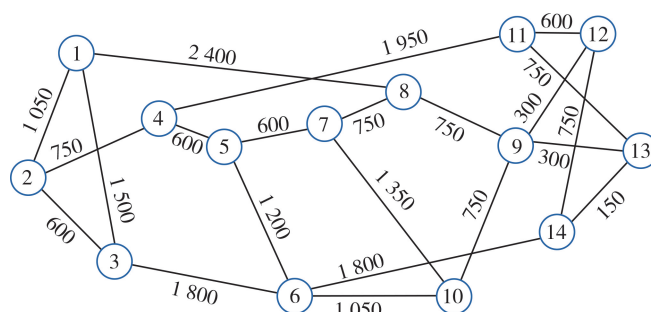
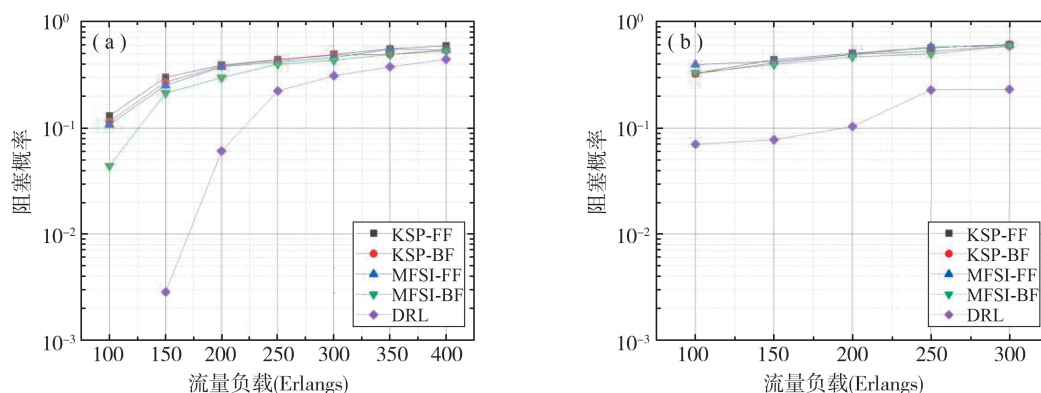


图2 14节点 NSFNET 拓扑

Fig. 2 14-node NSFNET topology

3.2 仿真分析

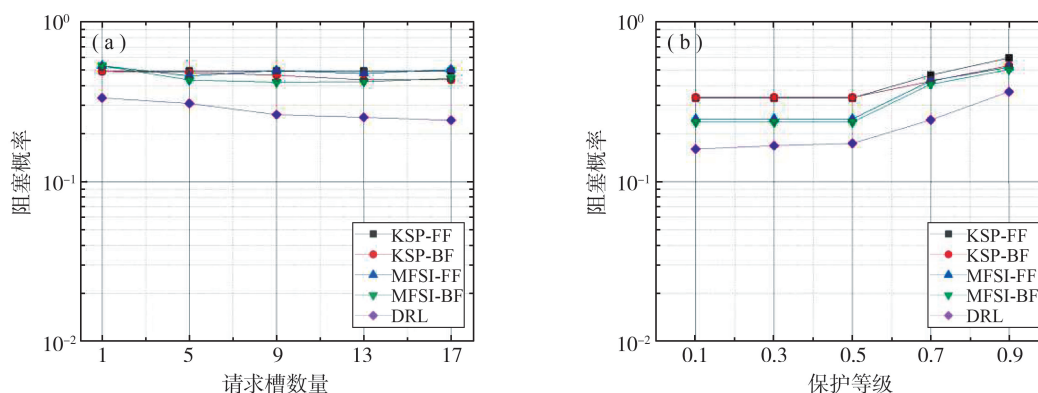
在本小节中,流量负载从 100 Erlangs 增加到 300 Erlangs,增量为 50 Erlangs。其中流量负载为平均到达率与请求的平均服务率之比。



注:(a) NSFNET 网络拓扑;(b) 28 节点 US 骨干网络拓扑。

图3 不同网络拓扑下的阻塞率

Fig. 3 Blocking probability under different network topologies



注:(a) 不同请求槽数量下的阻塞率;(b) 不同保护等级下的阻塞率。

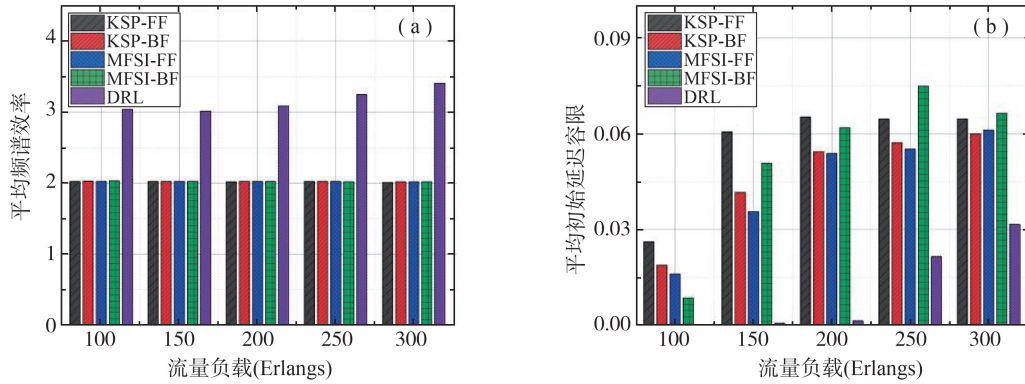
图4 NSFNET 网络拓扑的网络性能

Fig. 4 Network performance of the NSFNET topology

图3清晰展示了在不同网络下的阻塞率性能。阻塞概率均随流量负载增加而上升,这是因为流量负载的增加会直接导致频谱需求的增长。在高负载流量下,更多的请求会占用资源更长的服务时长,从而阻碍了其它未服务的请求,这种资源的长时间占用不仅使光通道资源愈加紧张,还导致部分请求需绕路传输,加剧非线性损伤并进一步降低 SNR,从而导致更多的 AR 服务请求被阻塞。如图3所示,所提出的基于 DRL 的算法在两个网络中,性能上明显优于基准启发式算法。这是由于基准启发式算法仅仅使用固定的路由策略,然而 DRL 算法根据不同的网络状态,灵活地选取不同的光路,从而有效减少请求的阻塞概率。与基准启发

式算法相比,基于 DRL 的算法能够适应网络状态的变化。此外,KSP-BF(MFSI-BF)优先处理带宽大、持续时间长、空闲时间短的请求,因此略优于 KSP-FF(MFSI-FF)。

在图 4(a)中,随着请求槽数量的增加,KSP-FF 的阻塞概率保持不变,因为该算法按照请求到达的顺序进行处理,而不进行排序。但基于 DRL 的算法和 KSP-BF 算法与之不同,AR 请求的阻塞概率最初会随着请求槽数量的增加略微下降,随后趋于稳定。这是因为在多个请求槽中对请求进行排序能够带来优化效果,但由于因动态到达的请求可能无法完全占用所有可用的请求槽,因此阻塞概率不会随着请求槽数量的增加持续降低。从图 4(b)可以明显看出,在流量负载设置为 300 Erlangs 时,阻塞概率随着保护等级的增加而上升。这是因为随着保护等级的提高,资源需求增加,同时请求的 SNR 会降低。如图 4 所示,本文提 DRL 算法优于基准启发式算法。这是因为基准启发式算法只能采用固定策略来调度 AR 请求,而本文提出的基于 DRL 的算法能够根据网络状态的变化进行动态调整。

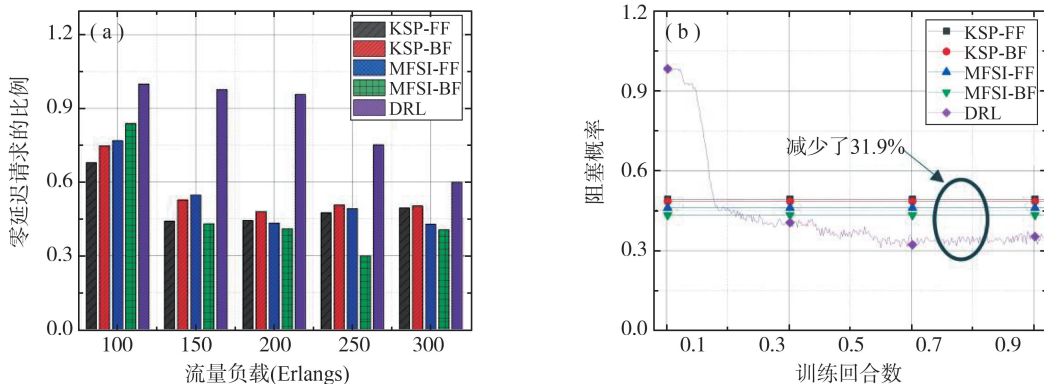


注:(a) 平均频谱效率与流量负载;(b) 平均初始延迟容忍与流量负载。

图 5 NSFNET 网络拓扑的网络性能

Fig. 5 Network performance of the NSFNET topology

从图 5(a)可以看出,基于 DRL 的算法在平均频谱效率上优于基准启发式算法,这是因为 DRL 算法能够根据当前网络状态灵活地选择最合适的调制格式,这种灵活性使得 DRL 算法在避免传输中的非线性损伤的同时,优先选择具有更高频谱效率的高阶调制格式。在图 5(b)中,随着流量负载的增加,DRL 的平均初始延迟容忍均呈现上升趋势,这是因为在流量负载较低时,EON 中有充足的频谱资源,可以在不延长等待时间的情况下快速部署请求,所以平均初始延迟较小。然而随着流量负载的增大,对频谱资源的需求也随之增加,因此新的请求需要等待已占用的频谱资源释放,从而导致等待时间占总时间窗口的比例增加,即平均初始延迟容忍上升。图 5(b)展示了系统平均初始延迟容忍的仿真结果,基于 DRL 的算法在平均初始延迟容忍的表现上仍优于基准启发式算法,这是因为基于 DRL 的算法对频谱资源进行了有效分配并减少了请求的等待时间。



注:(a) 零延迟请求的比例;(b) DRL 算法的训练曲线。

图 6 NSFNET 网络拓扑的网络性能

Fig. 6 Network performance of the NSFNET topology

图 6(a)展示了零延迟请求比例的仿真结果。随着流量负载的增加,DRL 算法的零延迟请求比例呈下降趋势,这是因为在 DRL 算法中,流量负载的增加会导致 EON 中的可用带宽资源减少,从而无法立即处理所有到达的请求。与基准启发式算法相比,DRL 算法能够有效利用频谱资源,大幅减少时域阻塞,从而使到达的请求能够及时部署,所以基于 DRL 的算法的性能优于基准启发式算法。图 6(b)展示了基于 DRL 的算法的阻塞概率随训练回合数的变化趋势,流量负载设置为 300 Erlangs。经过 100 个训练回合后,基于 DRL 的算法的阻塞概率显著低于基准启发式算法,且经过 200 回合训练后,阻塞概率趋于稳定,约为 0.33。造成这一显著差异的原因在于,基准启发式算法仅采用首次拟合算法,并未充分利用选择不同路径和调制格式的灵活性。相比之下,基于 DRL 的算法通过迭代优化获得了最优调度策略。与 KSP-FF 和 MFSI-FF 算法优先处理最先到达请求的策略不同,基于 DRL 的算法会根据服务时长和剩余空闲时间来调度请求。此外,基于 DRL 的算法在带宽消耗和空闲时间之间取得了较好的平衡。综合以上优点,DRL 算法有效地将阻塞概率降低了 31.9%。

4 总结

本文提出了一种基于 DRL 的生存性感知多路径 RMSA 算法,用于解决 EON 中动态 AR 请求的分配问题。采用 NLI 模型对传输信道的物理层非线性损伤进行估计,并引入 AR 流量模型来描述在特定时间间隔内对容量有需求的光路请求特征。所提出的 DRL 算法能够为到来的请求提供起始时间、路由和调制格式的决策。通过在不同流量负载、请求槽数量、保护等级条件下进行大量性能评估,仿真结果表明基于 DRL 的算法显著降低了请求的阻塞概率和平均初始延迟容限。在 14 节点网络拓扑上的仿真结果表明,所提 DRL 算法在 300 Erlangs 的流量负载下,减少了高达 31.9%的阻塞概率。

参 考 文 献

- [1] LEI Y, CHEN Q, JIANG Y F, et al. Crosstalk-aware routing, core, and spectrum assignment with core switching in SDM-EONs[C]// International Conference on Optical Communications and Networks, 2019: 1-3.
- [2] LIU S Q, LU W and ZHU Z Q. On the cross-layer orchestration to address IP router outages with cost-efficient multilayer restoration in IP-over-EONs[J]. IEEE/OSA Journal of Optical Communications and Networking, 2018, 10(1): 122-132.
- [3] ZHAO J Z, WYMEERSCH H and AGRELL E. Nonlinear impairment-aware static resource allocation in elastic optical networks[J]. Journal of Lightwave Technology, 2015, 33(22): 4554-4564.
- [4] LI Y J, ZHAO Y L, LI B, et al. Joint balancing of IT and spectrum resources for selecting virtualized network function in inter-data-center elastic optical networks[J]. Optics Express, 2019, 27(11): 15116-15128.
- [5] YAN L, AGRELL E, WYMEERSCH H, et al. Resource allocation for flexible-grid optical networks with nonlinear channel model[J]. IEEE/OSA Journal of Optical Communications and Networking, 2015, 7(11): B101-B108.
- [6] WANG R, BIDKAR S, MENG F C, et al. Load-aware nonlinearity estimation for elastic optical network resource optimization and management[J]. IEEE/OSA Journal of Optical Communications and Networking, 2019, 11(5): 164-178.
- [7] YAN L, AGRELL E, DHARMAWEERA M N, et al. Joint assignment of power, routing, and spectrum in static flexible-grid networks[J]. Journal of Lightwave Technology, 2017, 35(10): 1766-1774.
- [8] PENG S P, NEJABATI R, AZODOLMOLKY S, et al. An impairment-aware virtual optical network composition mechanism for future Internet[J]. Optics Express, 2011, 19(26): B251-B259.
- [9] LI X, ZHANG L, TANG Y, et al. On-demand routing, modulation level and spectrum allocation for multicast service aggregation in elastic optical networks[J]. Optics Express, 2018, 26(19): 24506-24530.
- [10] BEHERA S, SAVVA G, MANOUSAKIS K, et al. Impairment-aware routing, modulation, spectrum, and core allocation with bit loading in spectrally-spatially flexible optical networks[J]. IEEE/OSA Journal of Optical Communications and Networking, 2023, 15(6): 318-332.
- [11] 郭梦飞. 数据中心弹性光网络中的组播 RMSA 研究[D]. 郑州: 郑州轻工业大学, 2023.
- [12] ZHANG C, WANG X, WANG Y, et al. ENCG-DRL for multicast service oriented energy-efficient DU-CU deployment and RMSA in EON-enabled RAN[J]. Optical Fiber Technology, 2024, 88: 103958-103958.
- [13] 侯临风, 何荣希, 吴梓敬. 弹性光网络中基于 DRL 的 RMSA 算法[J]. 光通信技术, 2024, 48(3): 57-63.
- [14] 张博. 下一代灵活光网络中非线性损伤建模及补偿关键技术研究[D]. 北京: 北京邮电大学, 2019.

- [15] BEHERA S, GEORGE J, DAS G. Effect of transmission impairments in CO-OFDM based elastic optical network design[J]. *Computer Networks*, 2018, 144: 242-253.
- [16] DHARMAWEERA N, YAN L, KARLSSON M, et al. An impairment-aware resource allocation scheme for dynamic elastic optical networks[C]// *Optical Fiber Communications Conference and Exhibition*, Los Angeles, CA, USA, 2017: 1-30.
- [17] SHI C L, ZHU M, GU J H, et al. Deep-reinforced impairment-aware dynamic resource allocation in nonlinear elastic optical networks[C]// *Opto-Electronics and Communications Conference*, Hong Kong, 2021: 1-30.
- [18] LI S R, LU W, LIU X H, et al. Fragmentation-aware service provisioning for advance reservation multicast in SD-EONs[J]. *Optics Express*, 2015, 23(20): 25804-25813.
- [19] PATEL A N, JI P N, JUE J P, et al. Survivable transparent flexible optical WDM networks[C]// *Optical Fiber Communications Conference and Exhibition (OFC/NFOEC)*, 2011: 1-30.
- [20] RUAN L, XIAO N. Survivable multipath routing and spectrum allocation in OFDM-based flexible optical networks[J]. *IEEE/OSA Journal of Optical Communications and Networking*, 2013, 5(3): 172-182.
- [21] 陈斌, 顾家骅, 朱敏, 等. 基于深度强化学习的 OFDMA-PON 三维资源分配研究与性能分析[J]. *聊城大学学报(自然科学版)*, 2020, 33(6): 40-46.
- [22] 蔡晓峰, 望运武, 顾家骅, 等. 基于深度强化学习的细粒度 5G RAN 切片功能复用映射与路由[J]. *聊城大学学报(自然科学版)*, 2023, 36(5): 29-38.
- [23] ZHU M, CHEN Q, GU J H, et al. Deep reinforcement learning for provisioning virtualized network function in inter-datacenter elastic optical networks[J]. *IEEE Transactions on Network and Service Management*, 2022, 19(3): 3341-3351.
- [24] ZHU M, GU J H, SHEN T Y, et al. Energy-efficient and QoS guaranteed BBU aggregation in CRAN based on heuristic-assisted deep reinforcement learning[J]. *Journal of Lightwave Technology*, 2022, 40(3): 575-587.
- [25] 赵立刚. 基于机器学习的路由与资源分配算法的研究[D]. 北京: 北京邮电大学, 2022.
- [26] 刘兆洋, 潘必韬. 基于强化学习的光网络自适应高效率 RWA 算法[J]. *光通信研究*, 2024, 5: 66-73.
- [27] LIU Z W, SHU Z H, CHEN S W, et al. Low-latency virtual network function scheduling algorithm based on deep reinforcement learning[J]. *Computer Networks*, 2024, 246: 110418.
- [28] DY, ZH ZH, BAHAREH A, et al. A reinforcement learning-based routing algorithm for large street networks[J]. *International Journal of Geographical Information Science*, 2024, 38(2): 183-215.
- [29] CHEN X L, LI B J, PROIETTI R, et al. DeepRMSA: a deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks[J]. *Journal of Lightwave Technology*, 2019, 37(16): 4155-4163.
- [30] LU W AND ZHU Z Q. Dynamic service provisioning of advance reservation requests in elastic optical networks[J]. *Journal of Lightwave Technology*, 2013, 31(10): 1621-1627.
- [31] MAO H Z, ALIZADEH M, MENACHE I, et al. Resource management with deep reinforcement learning[C]// *Proceedings of the 15th ACM Workshop on Hot Topics in Networks*, 2016: 50-56.
- [32] WEI W T, GU H X, WANG K, et al. Improving cloud-based IoT services through virtual network embedding in elastic optical inter-DC networks[J]. *IEEE Internet of Things Journal*, 2018, 6(1): 986-996.
- [33] FANG W J, ZENG M L, LIU X H, et al. Joint spectrum and IT resource allocation for efficient VNF service chaining in inter-datacenter elastic optical networks[J]. *IEEE Communications Letters*, 2016, 20(8): 1539-1542.

(责任编辑:赵海军)