

面向数字孪生的混合业务确定性传输调度机制

王克文, 张维庭, 廖培希

引用本文

王克文, 张维庭, 廖培希. 面向数字孪生的混合业务确定性传输调度机制[J]. 计算机科学, 2024, 51(12): 37-45.

WANG Kewen, ZHANG Weiting, LIAO Peixi. [Deterministic Transmission Scheduling Mechanism for Mixed Traffic Flows Towards Digital Twin Networks](#) [J]. Computer Science, 2024, 51(12): 37-45.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于数字孪生的卫星网络移动边缘计算研究综述](#)

Review of Digital Twin Based Satellite Network Mobile Edge Computing
计算机科学, 2024, 51(12): 12-19. <https://doi.org/10.11896/jsjcx.240700046>

[元宇宙关键技术、研究进展与应用综述](#)

Review of Key Technologies, Research Progress and Applications of Metaverse
计算机科学, 2024, 51(12): 2-11. <https://doi.org/10.11896/jsjcx.240400166>

[基于强化学习考虑电池损耗的电动汽车充放电控制算法](#)

Reinforcement Learning Algorithm for Charging/Discharging Control of Electric Vehicles Considering Battery Loss
计算机科学, 2024, 51(11A): 231200147-7. <https://doi.org/10.11896/jsjcx.231200147>

[基于深度强化学习的云边协同任务迁移与资源再分配优化研究](#)

Cloud-Edge Collaborative Task Transfer and Resource Reallocation Optimization Based on Deep Reinforcement Learning
计算机科学, 2024, 51(11A): 231100170-10. <https://doi.org/10.11896/jsjcx.231100170>

[基于深度强化学习的无人机自主探索方法](#)

Autonomous Exploration Methods for Unmanned Aerial Vehicles Based on Deep Reinforcement Learning
计算机科学, 2024, 51(11A): 231100139-6. <https://doi.org/10.11896/jsjcx.231100139>

面向数字孪生的混合业务确定性传输调度机制

王克文^{1,2} 张维庭³ 廖培希³

1 北京交通大学电气工程学院 北京 100044

2 国能新朔铁路有限责任公司 内蒙古 鄂尔多斯 010300

3 北京交通大学电子信息工程学院 北京 100044

(81349177@qq.com)

摘要 针对铁路运维场景中混合业务流的端到端传输,提出了数字孪生架构下基于深度强化学习的确定性传输调度机制,即在线混合业务流端到端传输调度机制(End-to-End Transmission Scheduling Mechanism for Online Mixed-traffic, E2ETSM-OMT)。该机制基于差异化调度策略的思想,将业务流分为监控与数据采集流、控制与执行业务流和数据分析与业务优化流3类,通过确定性技术实现跨域端到端低时延传输。进一步地,通过模型映射和行为映射,将物理空间全方位、高精度地映射到虚拟空间,在数字孪生网络中构建混合业务的拓扑结构,预先分配数据传输路径和时隙资源,从而减少不同业务流之间的调度冲突和资源竞争。同时,通过深度强化学习(Deep Reinforcement Learning, DRL)智能体在线决策,兼顾效果与效率,对不同收益的业务流进行调度。与已有机制相比,数字孪生技术可以实现物理世界与虚拟世界的相互映射,实现非平稳通信环境下DRL的应用,避免在现实网络中探索造成的服务质量下降。仿真结果表明,所提出的面向数字孪生的确定性传输调度机制在保障成功调度混合业务流的同时,以较低的端到端整体时延实现了较高的传输收益。

关键词: 数字孪生; 确定性网络; 跨域传输; 差异化调度; 深度强化学习

中图分类号 TP393

Deterministic Transmission Scheduling Mechanism for Mixed Traffic Flows Towards Digital Twin Networks

WANG Kewen^{1,2}, ZHANG Weiting³ and LIAO Peixi³

1 School of Electronics and Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

2 Guoneng Xinshuo Railway Co., LTD., Ordos, Inner Mongolia 010300, China

3 School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

Abstract A deterministic transmission scheduling mechanism based on deep reinforcement learning under the digital twin architecture is proposed for the end-to-end transmission of mixed traffic flows in railway operation and maintenance scenarios, namely end-to-end transmission scheduling mechanism for online mixed-traffic (E2ETSM-OMT). Based on the idea of differentiated scheduling strategy, the proposed mechanism divides traffic flows into three categories: monitoring and data collection flow, control and execution traffic flow, and data analysis and business optimization flow, implementing cross domain end-to-end low latency transmission through deterministic technologies. Meanwhile, through model mapping and behavior mapping, the physical space is projected to the virtual space with high precision in all directions. In the digital twin network, after constructing the topology of mixed flows, deep reinforcement learning (DRL) agent makes pre-allocation decisions of transmission path and time slot resources, taking the effect and efficiency into account, so as to reduce scheduling conflicts and resource competition among different traffic flows. Compared with the existing mechanisms, digital twin technologies can realize the mutual mapping between the physical world and the virtual world, realize the application of DRL in non-stationary communication environment, and avoid the loss of service quality caused by exploration in the real network. Simulation results show that the digital twin-oriented deterministic transmission scheduling mechanism achieves high transmission benefits with low end-to-end overall delay while ensuring successful scheduling of mixed traffic flows.

Keywords Digital twin, Deterministic network, Cross-domain transmission, Differentiated scheduling, Deep reinforcement learning

到稿日期:2024-02-19 返修日期:2024-07-12

基金项目:国家自然科学基金(62201029);中国博士后科学基金(2022M710007, BX20220029)

This work was supported by the National Natural Science Foundation of China (62201029) and China Postdoctoral Science Foundation (2022M710007, BX20220029).

通信作者:张维庭(wtzhang@bjtu.edu.cn)

1 引言

近年来,大数据和人工智能等技术蓬勃发展,逐渐成为研究热点并应用于铁路智能运维等。这些场景中的数据往往比较精密,需要高效可靠的传输,而传统的“尽力而为”网络难以满足多种业务对低时延、高可靠、差异性服务的需求。在此背景下,确定性网络^[1]技术被提出,其技术关键在于实现确定性时延和抖动。

目前,已有大量研究工作聚焦于确定性传输调度。从局域网的时间敏感网络(Time Sensitive Network, TSN)^[2-10]到广域网的确定性网络(Deterministic Networking, Det-Net)^[11-14],再到跨域融合网络^[15-16],传输调度机制得到了广泛探索。文献[9]针对 TSN 中多种流量调度影响端到端时延确定性的问题,提出了一种在线混合流量分析的粒子群(Particle Swarm Optimization, PSO)改进算法;文献[10]提出了一种基于 SpaceFibre 星载网络的分类细粒度确定性调度算法,结合网络拓扑结构生成调度方案,在保证吞吐量的同时提升了时延性能和抗抖动性能;文献[15]提出了一种域间松耦合的端到端确定性传输和调度机制,通过拍卖模型联合考虑每条时间敏业务流的带宽分配和传输路径,实现了低复杂度的在线调度,并将传输时延保持在可控的范围内;文献[11]提出了使用循环指定的队列和转发(Cycle Specified Queuing and Forwarding, CSQF)求解大规模确定性网络中时间触发流的联合路由和调度问题的算法。在此基础上,文献[12]对 CSQF 机制的边缘流量整形方法进行了详解,并提出尽力而为(Best Effort, BE)流按较远路径传输的思想;文献[16]在大规模确定性网络^[13](Large-scale Deterministic Network, LDN)转发技术的基础上提出了循环队列转发(Cyclic Queuing and Forwarding, CQF)和确定性网际互连协议(Deterministic Internet Protocol, DIP)异构组网的确定性跨域传输并通过深度强化学习(DRL)进行流量调度,由于 DIP 技术中核心节点无逐流状态,设备之间仅需频率同步而不需要精准时间同步,因此可扩展性强。

然而,目前已有的研究都着眼于实体网络的传输调度,一方面算法缺乏对实时变化的感知,事实上,实际网络中的通信环境是非平稳的,预先训练的调度策略无法确保服务质量(Quality of Service, QoS)的要求;另一方面铁路运维中的业务流对传输要求严苛^[17-18],直接在物理世界中应用传输调度机制可能会导致效率下降,甚至引起严重的网络震荡。而数字孪生技术作为一种高效的解决方案,通过物理世界与虚拟世界的相互映射,在虚拟空间观测对等体就可以高精度反映并预测现实空间实体的状态,被广泛应用于制造业^[18-20]。文献[20]利用数字孪生无人仓储模型的可视化技术,结合 DRL 方法,实现了无人仓储资源动态调度优化^[21]。文献[22]面向数字孪生场景下的智能制造系统提出了端-边-云协同架构,同时,在考虑任务调度时隙阈值和调度结果的情况下,相应地提出了一种端-边-云协同孪生任务混合重调度策略,满足大规模、并行式数字孪生任务的调度需求。此外,文献[23]基于数字孪生提出了面向卫星时敏网络中的确定通信框架,为

卫星通信的确定性提供了技术支撑。同样地,数字孪生技术也为铁路运维智能化发展提供了新的契机,在铁路站点等各层面广泛部署物联网感知设备,实现铁路运维要素数字化,通过物联网、通信网及承载网的数据传输与处理形成与实体铁路“孪生”的数字铁路,基于孪生体做出可靠决策并应用,得到 QoS 保证后再同步决策到实体,就能实现结构监测智能化及管理决策智能化。

据此,本文提出了一种基于 DRL 的数字孪生架构下的确定性传输调度机制,创新性地将确定性传输调度扩展到数字孪生空间,由可对等观测并预知物理实体空间的孪生空间 DRL 智能体进行传输调度决策。具体而言,首先介绍了虚实映射的网络场景,构建了“TSN+DIP”^[16]的网络架构,以实现源端到对端的跨域确定性传输;其次对传输调度问题进行了分析,构建了数字孪生架构下的确定性网络和业务流模型,通过在端域 TSN 中部署 CQF 机制,以及 CQF 与 CSQF 的映射,实现跨域的无缝确定性传输;然后设计了一种基于决斗双深度 Q 网络(Dueling Double Deep Q Network, D3QN)的时隙路径联合在线调度算法;最后通过实验验证了算法的收敛性和优势。实验结果表明,本文提出的数字孪生与 DRL 结合的确定性传输调度机制实现了端到端确定性传输,并且对不同重要性业务流进行了粗粒度分类,实现了差异化传输调度时资源分配的最优解。

本文第 2 章具体描述了本文提出的基于 DRL 的数字孪生架构下确定性混合业务传输调度算法;第 3 章描述了仿真实验的设置及实验结果;最后总结全文并对未来工作进行了展望。

2 数字孪生架构下混合业务流确定性传输调度

2.1 系统模型

2.1.1 网络场景

图 1 展示了数字孪生架构下铁路运维中产生的混合业务流端到端传输模式及优势。在物理空间中,混合业务流由源端 TSN1 域跨 DIP 域传输到对端 TSN2 域。TSN 域跳数较少、各传感器直连 TSN 交换机,运行 CQF 机制;而 DIP 域业务流汇聚、跳数较多,运行周期具化的 CSQF 机制以实现多队列缓存。其中,TSN 中时间同步,而 DIP 中频率同步,CQF 与 CSQF 机制均将时间划分为多个等长时隙。

为进行形式化描述,将整个网络抽象为有向连通图 $G = \{V, E\}$,其中 V 为网络节点集合, $E = \{e | e = (v_i, v_j), i \neq j\}$ 为由节点构成的链路集合,节点 v_i 和 v_j 分别为链路 e 的源端点和末端点,在链路 e 上的时延为 d_e (假定包含上游节点处的处理时延和排队时延), $\mathcal{F} = \{f_1, f_2, \dots, f_{|\mathcal{F}|}\}$ 为业务流集合。源站发出的每条业务流 f_i 由五元组 $\langle v_{f_i}, v_{f_i}, C_{f_i}, D_{f_i}, \xi_{f_i} \rangle$ 定义, v_{f_i} 和 v_{f_i} 分别为业务流的传输起点和终点, c_{f_i} 为该业务流包含的数据大小, D_{f_i} 为该业务流的最大端到端时延, ξ_{f_i} 为调度成功的收益。业务流 f_i 的传输路径为 $p_{f_i} = (v_0, v_1, \dots, v_{|p_{f_i}|})$,其中节点集合 (v_0, v_1, \dots, v_a) 位于 TSN1 域内, $(v_{a+1}, v_{a+2}, \dots, v_b)$ 位于 DIP 域内, $(v_{b+1}, v_{b+2}, \dots, v_{|p_{f_i}|})$ 位于 TSN2 域内。

为了跨域网关执行时隙映射与整流机制,定义超周期(Hyper-Period, HP)^[14,16],源端 TSN1 和对端 TSN2 的传输周期分别为 T_{TSN1} 和 T_{TSN2} ,DIP 的传输周期为 T_{DIP} ,映射关系为:

$$T_{HP} = N_{TSN1} T_{TSN1} = N_{DIP} T_{DIP} = N_{TSN2} T_{TSN2} \quad (1)$$

其中, $N_{TSN1}, N_{TSN2}, N_{DIP}$ 为一个超周期内不同域节点转发的周期数。

1) 跨域映射

我们定义链路 $e = (v_i, v_{i+1})$ 跨越不同的域 D_1 和 D_2 ,则上游节点 v_i 在周期 c 发出的业务流被映射至下游节点 v_{i+1} 的周期 $\Phi_e(c)$,即周期映射为:

$$\Phi_e(c) = \left[\frac{((c+1) \% N_{D1}) T_{D1} + d_e}{T_{D2}} - 1 \right] \% N_{D2} \quad (2)$$

2) 流量整形

为具象化理解,将两种机制的时隙/周期等抽象为节点

队列进行统一表述。被分配进入队列 0 的业务流将立即被转发,被分配进入队列 k 的业务流将等待 k 个周期后被转发 ($0 < k \leq N$),其中 N 为该节点预留的队列数。

在 TSN 域内,运行 CQF 机制同步入队和出队,预留两个队列。

在 DIP 域内,由于网络中汇聚的大量混合业务流共存,不同流的多个时隙重叠,导致可调度流数减少,因此我们采用 CSQF 机制进行宏观调度。在每个节点处引入周期偏移 $r_k \in \{0, 1, \dots, n-1\}$ 以便将业务流调度到特定的 CSQF 队列中,其中 N 为预留的队列数。当业务流在 DIP 域 v_m 节点的 c 周期内发出,则将在 DIP 域 v_{m+1} 节点的 $(c + r_k) \% N_{DIP}$ 周期内被转发。

综上可知,业务流端到端传输由 TSN1 域传输、TSN1-DIP 跨域映射、DIP 域传输、DIP-TSN2 跨域映射和 TSN2 域传输构成。

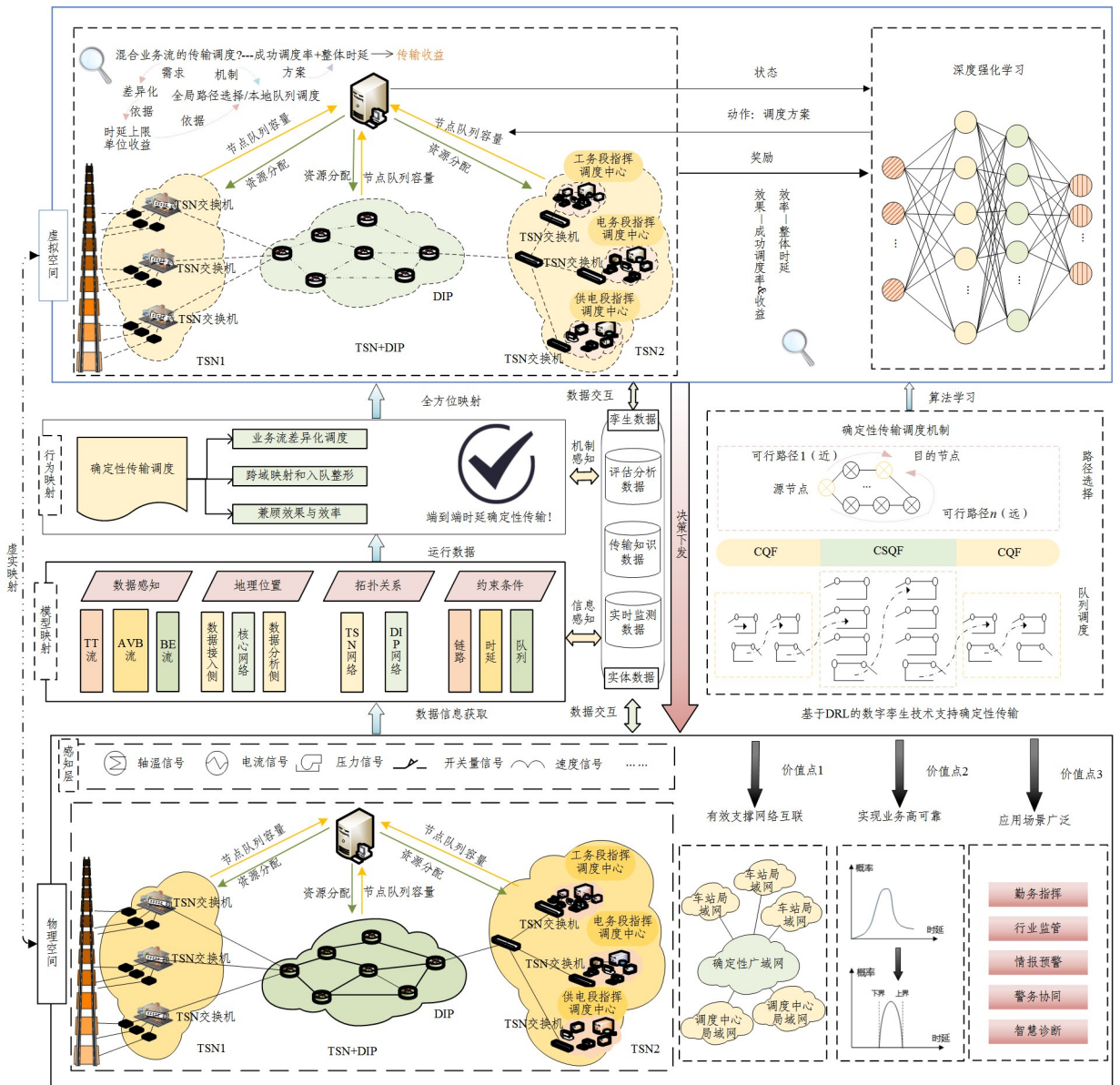


图 1 数字孪生架构下铁路运维业务流端到端传输场景

Fig. 1 End-to-end transmission scenario of railway operation and maintenance traffic flow under digital twin architecture

虚拟空间由物理空间高精度映射得到,通过从物理空间部署的各种传感器获取数据,构建刻画地理位置的数据接入侧、核心网络以及数据分析侧,确定 TSN 网络与 DIP 网络的拓扑结构,并对传感器、交换机、路由器等性能进行提取,得到虚拟空间的模型数据;同时,部署基于传输知识和评估知识的确定性端到端传输策略,以实现虚实空间的全方位高精度同步。在虚拟空间,采用在线混合业务流端到端传输调度算法(End-to-End Transmission Scheduling Mechanism for Online Mixed-traffic, E2ETSMT-OMT)在资源有限的情况下对确定性业务流进行集中式调度,在确保调度成功率的同时,最小化整体时延和最大化收益,最后将训练趋于稳定的策略同步到物理空间,以满足实际铁路运维对传输的有效性和可靠性要求。

2.1.2 业务流传输模型

根据铁路运维中传输业务流的特点,将其大体划分为对时延比较敏感的控制与执行业务流、对带宽要求相对较高的监控与数据采集流和用于数据分析与业务优化的业务流,对应于表 1 中 TSN 的时间触发(Time Triggered, TT)流、音视频桥接(Audio/Video Bridging, AVB)流和 BE 流。

TT 流多为非周期突发,数据量相对较小,但对于铁路运维却最重要;AVB 流周期采集,数据量大,重要性次之。TT 流和 AVB 流为时间敏感流;BE 流对时延的容忍度大,常用于对时延的高级分析,本文以传输成功的收益值来区别体现其重要性。由于混合流在有限节点资源同时路由,尤其是在 DIP 域进行汇聚,随着流的数量增多,路由决策的复杂度大大提升,算法不仅要考虑混合流量之间的冲突,而且要快速有效地找到优秀方案。本文提出的端到端传输机制如图 2 所示。

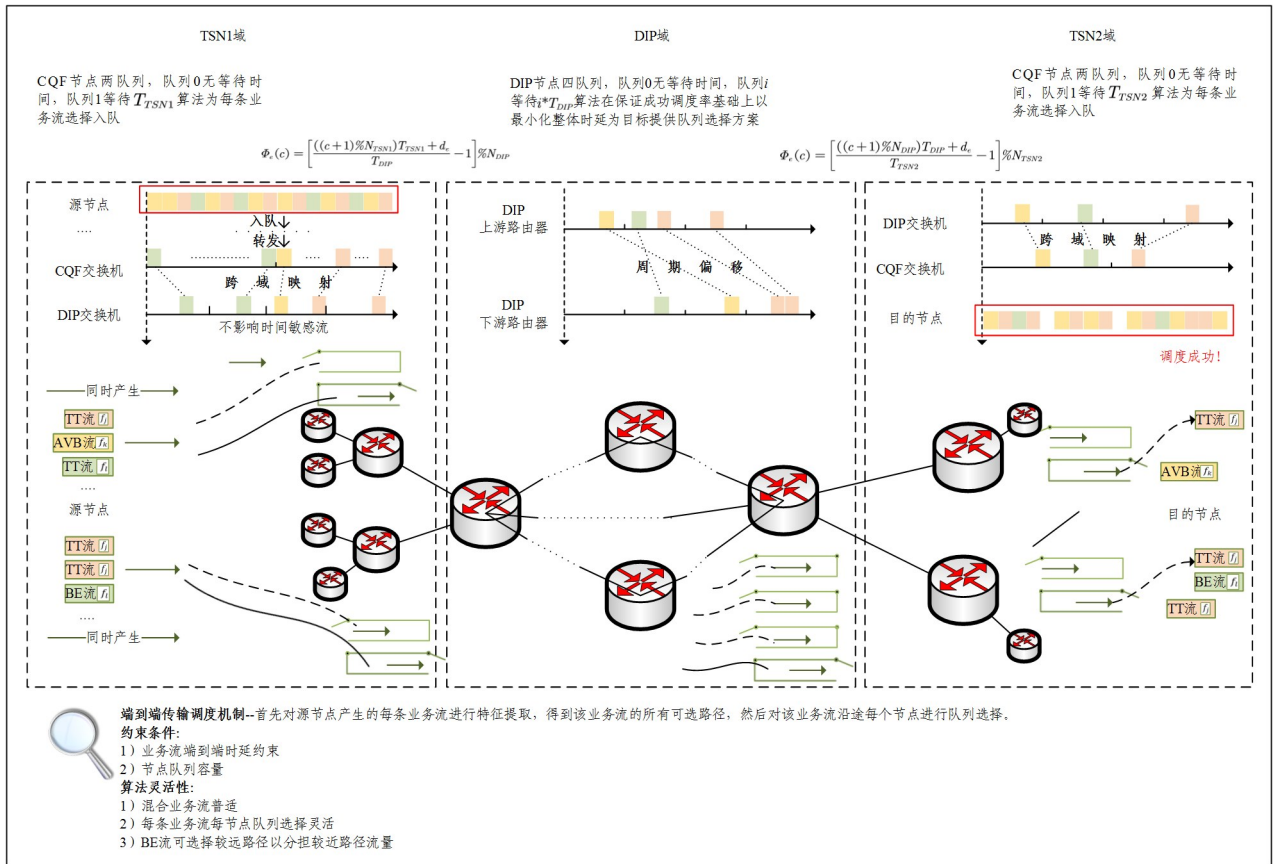


图 2 混合业务流端到端传输机制

Fig. 2 End-to-end transmission mechanism for mixed traffic flow

表 1 铁路运维中业务流粗粒度分类^[2]

Table 1 Coarse grained classification of traffic flows in railway operation and maintenance^[2]

铁路运维中的业务流	示例	QoS 需求			业务类别	调度成功收益	调度失败收益
		时延	带宽	可靠性			
控制与执行业务流	开关信号灯、道岔扳动...	高	较低	高	TT	$8 * \frac{D_{f_i}}{d_{f_i}}$	-10
监控与数据采集流	压力、温度、速度...	较高	高	较高	AVB	$4 * \frac{D_{f_i}}{d_{f_i}}$	-10
数据分析与业务优化流	轨道障碍物检测...	低	低	低	BE	$2 * \frac{D_{f_i}}{d_{f_i}}$	-10

2.2 问题定义

铁路运维场景中产生的混合业务流进行跨域端到端传输,每条业务流经由一条链路实现端到端传输。

2.2.1 约束条件

1) 时延约束

每条业务流跨域传输的总时延为:

$$d_{f_i} = d_{\text{TSN1}_i} + d_{\text{cross1}_i} + d_{\text{DIP}_i} + d_{\text{cross2}_i} + d_{\text{TSN2}_i} \quad (3)$$

其中, d_{TSN1} , d_{DIP} , d_{TSN2} 分别为 3 个域内传输时延, d_{cross1} 和 d_{cross2} 为跨域传输时延。

在 TSN1 域, 由于运行 CQF 机制, 在节点 v_m 的 c 周期发出的业务流, 会在节点 v_{m+1} 的 c 周期被接收, 然后在节点 v_{m+1} 的 $c+1$ 周期被转发。业务流 f_i 从到达 v_0 到离开 v_a 的时延为:

$$d_{\text{TSN1}_i} = (a+1)T_{\text{TSN1}} \quad (4)$$

TSN1-DIP 跨域传输时, 业务流 f_i 从离开 v_a 到离开 v_{a+1} 的时延为:

$$d_{\text{cross1}_i} = d_{(v_a, v_{a+1})} + \left\{ \left[\frac{((C_a + 1) \% N_{\text{TSN1}}) T_{\text{TSN1}} + d_{(v_a, v_{a+1})}}{T_{\text{DIP}}} - 1 \right] \% N_{\text{DIP}} \right\} T_{\text{DIP}} \quad (5)$$

在 DIP 域, 由于运行 CSQF 机制, 业务流 f_i 从离开 v_{a+1} 到离开 v_b 的时延为:

$$d_{\text{DIP}} = \sum_{n=a+1}^{b-1} [d_{(v_n, v_{n+1})} + r_{n+1} T_{\text{DIP}}] \quad (6)$$

DIP-TSN2 跨域传输时, 业务流 f_i 从离开 v_b 到离开 v_{b+1} 的时延为:

$$d_{\text{cross2}_i} = d_{(v_b, v_{b+1})} + \left\{ \left[\frac{((C_b + 1) \% N_{\text{DIP}}) T_{\text{DIP}} + d_{(v_b, v_{b+1})}}{T_{\text{TSN2}}} - 1 \right] \% N_{\text{TSN2}} \right\} T_{\text{TSN2}} \quad (7)$$

在 TSN2 域, 业务流 f_i 从离开 v_{b+1} 到离开 $v_{|p_{f_i}|}$ 的时延为:

$$d_{\text{TSN2}_i} = (|p_{f_i}| - b - 1) T_{\text{TSN2}} \quad (8)$$

对业务流 f_i 的端到端时延有:

$$d_{f_i} \leq D_{f_i} \quad (9)$$

2) 队列容量约束

混合业务流沿着确定的端到端链路传输, 一方面, 业务流只有在所分配的路径上每一跳分配的队列缓存资源都足够时才能传输成功。在实验中, 当所分配的资源小于业务流包大小时, 将宣告调度失败, 并转换为超时来量化处理, 这一点将在下文的算法设计中详细阐述; 另一方面, 节点队列的缓存容量决定了所能传输的业务流上限。

$$\sum C_{f_i, v_j} \leq C_{v_j}^m \quad (10)$$

其中, C_{f_i, v_j} 为同一周期被调度至节点 v_j 的业务流, $C_{v_j}^m$ 为节点 v_j 单个队列的缓存容量。当节点 $v_j \in \text{TSN}$ 域时, $m \in \{1, 2\}$; 当节点 $v_j \in \text{DIP}$ 域时, $m \in \{1, 2, 3, 4\}$ 。

2.2.2 目标函数

确定性传输的指标之一是时延, 通过对网络节点、链路资源的合理配置, 确保业务流在截止时间内到达目标节点。

混合业务流的整体时延为:

$$D = \sum_{f_i} d_{f_i} \quad (11)$$

由表 1 可知, 铁路运维中的 3 类混合业务流在实际场景中的意义和重要性各不相同。其中, 控制与执行业务流对时延要求最高; 监控与数据采集流产生快, 发送周期

更短; 数据分析与业务优化流产生慢, 但是其对铁路运维的效用最大。因此考虑不同业务流有不同的收益 ξ , 其中 $\xi_{\text{TT}} > \xi_{\text{AVB}} > \xi_{\text{BE}}$ 。

混合业务流的总体收益为:

$$\xi = \sum_{f_i} \xi_{f_i} \frac{D_{f_i}}{d_{f_i}} - 10 * m \quad (12)$$

其中, m 为调度失败的业务流条数。

本文确立了一种兼顾效果与效率的目标函数, 以成功传输时间敏感流为第一目标, 辅助目标为提高传输效率, 同时得到较高收益值。

$$\max F = \alpha * |F^*| + \beta * \frac{1}{D} \quad (13)$$

其中, α/β 为权重值, $\alpha + \beta = 1$; F^* 为调度成功的时间敏感流; D 为整体传输时延。

2.3 算法设计

为了解决铁路运维场景中混合业务流端到端确定性传输调度问题, 本文提出了一种基于 D3QN 的 E2ETSM-OMT 算法。

在虚拟空间, 智能体与孪生网络进行交互, 获取混合业务流信息和链路状态信息并引入状态空间, 根据输入状态对应地选择合适的动作输出到孪生网络中, 从而使业务流得以沿动作指定的路径和队列传输。不同业务流对传输时延有不同的要求, 传输成功的收益值也各不相同。智能体的决策目标是在保证调度成功率的同时, 最小化整体时延和最大化总体收益值。当训练趋于稳定时, 策略下发到物理空间, 以满足实际应用需要。

由以上分析可知, 网络各节点队列容量与上一条业务流调度结果有关, 因此具有马尔可夫性。将其建模为马尔可夫决策过程 (Markov Decision Process, MDP), 状态、动作和奖励设定如下。

2.3.1 马尔可夫决策过程建模

1) 状态空间 S

状态包含两方面: 各个网络节点队列的剩余缓存容量 S_i^{cap} 和业务流的传输要求 S_i^{req} , 其中 S_i^{cap} 可以表示为:

$$S_i^{\text{cap}} = \left\{ \begin{array}{cccc} C_0^{v_0} & C_0^{v_1} & \cdots & C_0^{v_m} \\ C_1^{v_0} & C_1^{v_1} & \cdots & C_1^{v_m} \\ \vdots & \vdots & \vdots & \vdots \\ C_{n-1}^{v_0} & C_{n-1}^{v_1} & \cdots & C_{n-1}^{v_m} \end{array} \right\} \quad (14)$$

其中, $C_j^{v_i}$ 表示节点 v_j 的队列 j 剩余的缓存容量。

2) 动作空间 A

动作空间由 (路径, 队列) 二维元组刻画, $A = \{a_i \mid a_i = (p_i, r_{v_i})\}$, 智能体以动作向量 a_i 指导业务流传输。在传输开始时刻, 提取业务流的传输起点和终点, 得到该业务流所有的可选传输路径, 这样可以避免冗余的探索过程。 $p_i \in \text{PATH}_i$ 为业务流 f_i 的所有可选传输路径, $r_{v_i} \in \{0, 1, \dots, N-1\}$ 为节点 v_i 的队列序号。

$$\mathbf{a}_t \begin{cases} \arg \max_{\mathbf{a}_t \in A} Q(\mathbf{s}_t, \mathbf{a}_t), & \text{if } x \geq \epsilon \\ \text{random } \mathbf{a}_t, & \text{otherwise} \end{cases} \quad (15)$$

其中, $x \in [0, 1)$ 为每步生成的随机变量。

3) 奖励函数 R

孪生网络中业务流以选定的调度方案进行对应传输, 以一定的奖励反馈智能体, 并将状态转到下一个状态。奖励用于即时评价动作的优劣, 以目标导向促使智能体探索得到最优调度方案。奖励如下:

$$r_t = \begin{cases} \xi_{f_t} \left(\frac{D_{f_t}}{d_{f_t}} \right), & \text{if success} \\ -10, & \text{if time_out} \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

其中, D_{f_t}/d_{f_t} 为传输时延比值。当业务流基于动作 a_t 的传输时延越小, 时延比值越大, 获得的奖励越多。

2.3.2 算法具体设计

基于队列缓存的路由算法考虑混合业务流端到端确定性传输, 具体如算法 1 所示。为保证确定性, 预先给每条业务流分配传输路径, 结合确定的传输机制, 以成功调度业务流为首要目标, 以最小化整体传输时延和最大化总体传输收益为次要目标。

算法 1 基于队列缓存的路由算法

输入: 网络信息 G , 业务流信息 \mathcal{F}

输出: 传输调度方案 s

1. for 业务流 $f_i \in \mathcal{F} = \{f_1, f_2, \dots, f_{|\mathcal{F}|}\}$ do
2. 进行特征提取, 按 D_i 对业务流进行粗粒度分类
3. 调用 Networkx 库中的 all_simple_paths 函数找所有可行路径, 删去不满足约束条件的路径并更新 $\rightarrow \text{PATH}_i$;
4. for 传输调度决策 \mathbf{a}_i do
5. 根据 \mathbf{a}_i 选择路径 $\text{path} \in \text{PATH}_i$
6. for path 中 TSN1 域中节点 do
7. 根据 \mathbf{a}_i 选择队列, 时延累加
8. end for
9. 跨域映射到 path 中 DIP 域首个节点的队列中, 时延累加
10. for path 中 DIP 域中节点 do
11. if $f_i \in \text{TT}$:
12. for j in range(4) do
13. 根据 $(\mathbf{a}_i + j)$ 选择队列, 时延累加
14. elif $f_i \in \text{AVB}$:
15. for j in range(4) do
16. 根据 $(\mathbf{a}_i + 1 + j)$ 选择队列, 时延累加
17. else:
18. for j in range(4) do
19. 根据 $(\mathbf{a}_i + 2 + j)$ 选择队列, 时延累加
20. end for
21. 跨域映射到 path 中 TSN2 域首个节点的队列中, 时延累加
22. for path 中 TSN2 域中节点 i do
23. 根据 \mathbf{a}_i 选择队列, 时延累加
24. end for
25. 第 6—24 步中若有节点所有队列均无法满足传输时, 时延

加上 D_i , 将资源约束转化为时延约束, 便于量化和收敛

26. end for

算法 2 基于 D3QN 的在线混合业务流端到端传输调度算法 (E2ETSA-OMT)

输入: 网络信息 G , 业务流信息 \mathcal{F} , 迭代轮次 M , 经验回放池 D 大小 N ,

$m, L^-,$ 探索率 $\epsilon_{\text{ini}}, \epsilon_{\text{fin}}$

输出: 传输调度方案 s

1. 初始化当前 Q 网络参数 θ 和目标 Q' 网络参数 $\theta', \theta \rightarrow \theta', \epsilon_{\text{ini}} \rightarrow \epsilon, s = \{\}$
2. for episode $e \in \{1, 2, \dots, M\}$ do
3. 初始化环境, 获取状态 $S = S_0, R = 0, \text{done} = \text{False}$;
4. for 业务流 $f_i \in \mathcal{F} = \{f_1, f_2, \dots, f_{|\mathcal{F}|}\}$ do
5. 做出传输调度决策: $\mathbf{a}_t \leftarrow Q(\mathbf{S}_t, \mathbf{A}; \theta) / \epsilon$ — 贪婪算法;
6. 根据 \mathbf{a}_t 选择路径和节点队列
7. 状态转移: $\mathbf{A}_t = \mathbf{a}_t$, 新状态 $\mathbf{s}_{t+1} = \mathbf{s}'$, 奖励 r_t, d_t
8. 将 $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1}, d_t)$ 五元组存入经验回放池 D
9. while D 缓存 $\text{Buffer}_{\text{size}} \geq N$
10. 替换掉旧元组
11. end while
12. 从 D 中随机采样 m 个样本 $(S_j, A_j, R_j, S'_j, \text{done}_j), j = 1, 2, \dots, m$, 计算当前 Q 网络的 y_j 和 Loss
13.
$$y_j = \begin{cases} R, & \text{if done} \\ R_j + \gamma Q' * \mathbf{s}', \arg \max_{\mathbf{a}' \in A} Q(S'_j, \mathbf{a}'; \theta'), \theta^- \end{cases}, \text{ else}$$
14. $\text{Loss} = \frac{1}{M} \sum_{j=1}^m (y_j - Q(S_j, A_j; \theta))^2$ 并更新 θ
15. 每 L^- 步同步参数,
16. end for
17. if $\epsilon \geq \epsilon_{\text{fin}}$ then
18. $\epsilon^- = 0.8 * \frac{\epsilon^- - \epsilon_{\text{fin}}}{1000}$
19. end if
20. end for

3 实验

本章将演示模拟的设置环境, 并验证本文提出的算法的性能。

3.1 实验设计

3.1.1 网络拓扑

本文提出的跨域融合网络由两个端域 TSN1、TSN2 和一个广域 DIP 组成, DIP 域内部署 7 个转发节点, 广域网传播时延均为 $150 \mu\text{s}$, 链路带宽为 1 Gbps。各域传输周期分别为: $T_{\text{TSN1}} = T_{\text{TSN2}} = 20 \mu\text{s}, T_{\text{DIP}} = 10 \mu\text{s}$ 。

构造需要从源节点 0 和 1 传输到目的节点 9 的数量比为 3:5:2 的 TT/AVB/BE 这 3 类流, 包大小分别为 {1600, 2600, 400} bit, 时延上界分别为 {0.82, 1.1, 1.26} ms。

3.1.2 参数设置

以 4×10 矩阵表示 10 个网络节点的队列容量, 业务流信息从预定义的列表流表中导出。为加快学习速度, E2ETSA-OMT 采用了两个完全相同的双神经网络, 每 4 步对其参数进行同步。神经网络包括卷积神经网络 (CNN) 和两个完全连接的层。CNN 的第一个池化层有 32 个通道, 第二个 64 个

通道在“压平”操作后连接到完全连接的层。

3.2 实验结果分析

当业务流传输同时满足时延约束和资源约束时,即在时延上限允许范围内到达目的地,同时每节点所预留的缓存资源都足够,业务流被成功调度。在仿真实验中,我们以调度成功率、整体传输时延和总体传输收益 3 个指标作为评估确定性传输调度性能的依据。其中调度成功率反映传输调度效果;整体传输时延反映传输调度效率,整体传输时延越小,效率越高。最后,本文从服务质量 (Quality of Service, QoS)

角度入手,由于不同种类业务流有着不同的单位传输收益,不同业务流在传输时延上的优化也会带来 QoS 不同程度的提升,因此制定了总体传输收益指标,其中单条业务流传输调度收益如表 1 所列。

1) 算法收敛情况

图 3 展示了 E2ETSM-OMT 算法在不同学习率下的收敛情况。可以看出,最高调度成功率、最小整体时延和最大总体收益均在 $lr=10^{-5}$ 时得到。实验表明,较高的学习率不一定能带来更好的性能,而较低的学习率往往会导导致算法收敛缓慢。

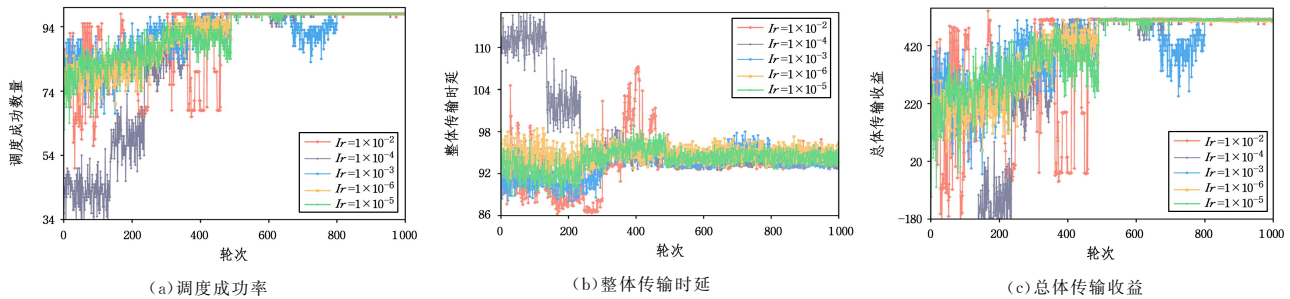


图 3 算法在不同学习率下的性能及收敛情况

Fig. 3 Performance and convergence of E2ETSM-OMT algorithm at different learning rates

图 4 展示了 E2ETSM-OMT 算法在不同折扣因子下的收敛情况。可以看出,图中 $\gamma=0.99$ 的曲线收敛较为平缓。

这是由于 γ 越大智能体往前考虑的步数越多,但训练难度也越高,更难收敛。

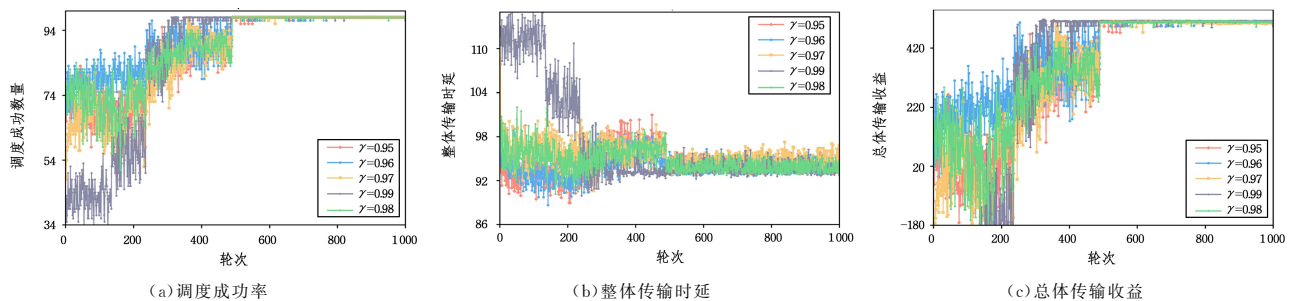


图 4 算法在不同折扣因子下的性能及收敛情况

Fig. 4 Performance and convergence of E2ETSM-OMT algorithm with different discount factors

2) 算法性能

我们以调度成功率作为算法的第一性能评估指标,同时考虑整体传输时延和总体传输收益。图 5—图 7 给出了不同算法在 $lr=10^{-5}$, $\gamma=0.98$ 时的训练情况。可以看出,Random 算法波动较大且无法收敛,而 E2ETSM-OMT 算法及其他 DRL 算法会随着训练的进行而逐渐收敛。

其中, Nature DQN 算法在 800 回合左右达到收敛, DDQN 在 600 回合左右达到收敛, E2ETSM-OMT 则在 500 回合左右到达收敛。相比 Nature DQN, DDQN 在不增加计算复杂度的情况下,通过两个 Q 网络来分别选择动作和评估 Q 值,从而减少过估计的问题,提高了解的准确性和稳定性;而基于 D3QN 的 E2ETSM-OMT 算法结合了 Dueling DQN 和 DDQN 两种算法的优点,可以在不增加模型复杂度的情况下,通过同时考虑状态值函数和优势函数来更好地估计 Q 值,从而提升算法的性能。

(1) 调度成功率

图 5 给出了不同算法的调度成功率随训练过程的变化。可以看出,所有 DRL 算法最终都达到了 98% 的调度

成功率,而本文提出的基于 D3QN 的 E2ETSM-OMT 算法最为稳定。

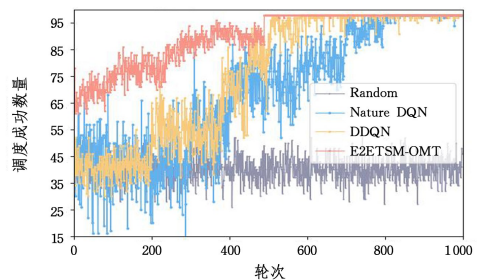


图 5 不同算法的调度成功率随训练轮次的变化

Fig. 5 Variation of successful scheduling rates of different algorithms with training rounds

(2) 整体传输时延

图 6 给出了不同算法的整体传输时延随训练过程的变化。随着训练回合增加, Nature DQN 算法整体传输时延稳定在 96.04ms, DDQN 稳定在 95.97 ms, E2ETSM-OMT 则降至 93.82 ms 并收敛。从这一角度考虑, E2ETSM-OMT 的

性能相应地提升了 2.31% 和 2.24%。

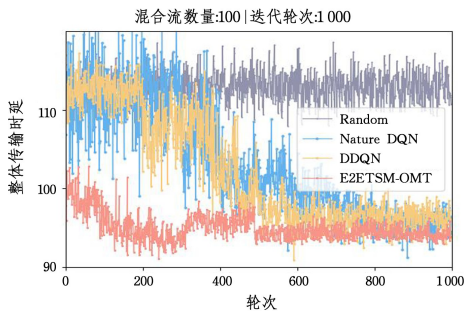


图 6 不同算法整体传输时延随训练轮次的变化

Fig. 6 Variation of overall transmission delay of different algorithms with training rounds

(3) 总体传输收益

图 7 给出了不同算法的总体传输收益随训练过程的变化。随着训练回合增加, Nature DQN 算法总体传输收益提升至 505.07, DDQN 提升至 506.28, E2ETSM-OMT 则高达 511.56。从这一角度考虑, E2ETSM-OMT 性能相应地提升了 1.29% 和 1.04%。

以上实验结果表明, 基于 D3QN 的 E2ETSM-OMT 算法收敛快、稳定性好, 兼顾效果与效率, 在保证高调度成功率的同时实现了较低的整体时延和较高的总体收益, 性能良好, 满足铁路运维等严苛场景对有效性和可靠性的要求。同时, 在

训练开始时, 就能实现较低整体传输时延, 但调度成功率和总体传输收益较低, 随着训练的成熟, 均得到了较好的结果。这也印证了本文算法是以成功调度混合业务流为第一目标, 以最小化整体时延和最大化收益为辅助目标, 符合效果第一效率第二这一设计理念。

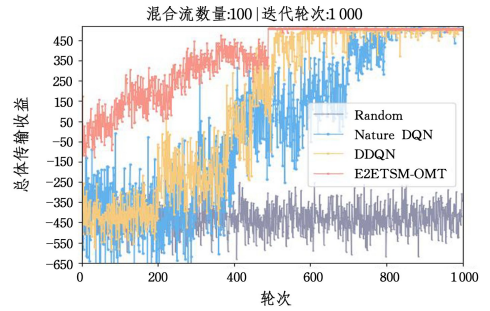


图 7 不同算法总体传输收益随训练轮次的变化

Fig. 7 Variation of overall transmission benefits of different algorithms with training rounds

此外, 上述实验均在资源有限的设定下完成, 算法收敛后仅有小幅抖动, 可以预见的是, 当资源充足时, 性能更佳。图 8 给出了节点缓存资源充足时, 混合业务流端到端传输结果的 3 个评估指标。可以看出, 调度成功率很快收敛至 100%, 而整体传输时延和总体传输收益在抖动中趋于最优。

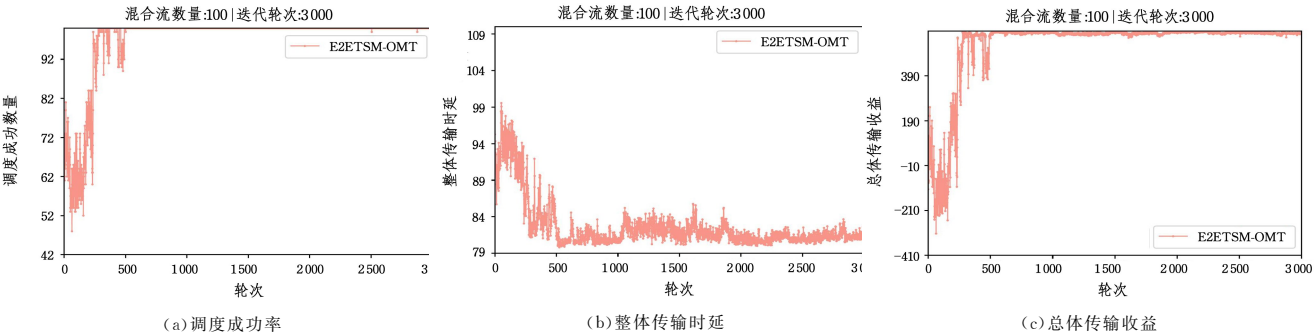


图 8 资源充足时算法效用

Fig. 8 Algorithm utility when resources are sufficient

结束语 针对铁路运维场景下混合业务流的跨域端到端传输调度, 本文创新性地引入数字孪生技术, 并在孪生空间中部署深度强化学习, 设计了数字孪生架构下混合业务流的端到端传输模型, 并进一步设计了基于队列缓存的 CQF 和 CSQF 机制相结合的路由算法。基于所提出的数字孪生技术与 DRL 技术结合的模型, 应用确定性传输调度机制, 提出了基于深度强化学习 D3QN 的多目标在线决策算法 E2ETSM-OMT。实验结果表明, 所提算法能有效应用于数字孪生架构下的严苛场景, 实现了较高的调度成功率和较低的整体传输时延, 并由于其支持差异化的调度而实现了较高的总体传输收益, 为端到端通信提供了差异化 QoS。同时, 本文提出的数字孪生架构能够有效支持网络互联且可移植性好。下一步工作是搭建此数字孪生应用的确定性传输调度系统, 并以虚实互通来验证方案的可行性。

参考文献

- [1] DUAN X D, LIU P, LU L, et al. Review of Deterministic Network Technology [J]. Telecommunications Science, 2023, 39(11): 1-12.
- [2] YAO J, WANG L, ZHOU Z, et al. Burst-Aware Mixed Flow Scheduling in Time-Sensitive Networks for Power Business [C]// 2023 IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC). 2023.
- [3] TAN W, WU B. Long-distance Deterministic Transmission among TSN Networks: Converging CQF and DIP [J]. arXiv: 2111.03246, 2021.
- [4] CHEN C L, ZHANG Y Z, XU Q M, et al. A Time Sensitive Network Gating Mechanism Traffic Shaping and Routing Planning Scheduling Method; CN111740924A [P]. 2020-10-02 [2024-02-07]. <https://wenku.baidu.com/view/293fd38b4ad7c1e708>

- a1284ac850ad02df800719?fr=xueshu_top&_wks_ = 1727075609594&-needWelcomeRecommand=1.
- [5] YANG S J, ZHUANG L, SONG Y, et al. Intelligent scheduling mechanism for time sensitive network modes in multimodal networks [J]. Journal of Communications, 2022, 43(5): 82-91.
- [6] SHEN Z, WANG H, WEI M, et al. Research on Global Deterministic Direct Forwarding and Scheduling of Mixed Flow Based on Time-Sensitive Network in Substation[J]. Electronics, 2023, 12(19): 4101.
- [7] LI Y H, ZHANG P, ZHOU Y C, et al. A Data Forwarding Mechanism based on Deep Reinforcement Learning for Deterministic Networks[C]// IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Institute of Electrical and Electronics Engineers, 2020: 285-290.
- [8] ZHANG Y, ZHANG H, LIU X, et al. Digital twin-enabled deep reinforcement learning for joint scheduling of ultra-reliable low latency communication and enhanced mobile broad band: A reliability-guaranteed approach[J]. Transactions on Emerging Telecommunications Technologies, 2023, 34(3): e4705.
- [9] WANG J X, YANG S J, ZHUANG L, et al. Multi objective online mixed traffic scheduling algorithm in time sensitive networks [J]. Computer Science, 2023, 50(7): 285-292.
- [10] ZHENG J Y, AN J S. SpaceFiber Low Delay Deterministic Scheduling Algorithm for Spaceborne Data Networks [J]. Journal of National University of Defense Technology, 2023, 45(2): 179-187.
- [11] QIANG L, LIU B Y, YU D L, et al. Large scale deterministic network forwarding technology [J]. Telecommunications Science, 2019(9): 8.
- [12] LEGUAY J, KROLIKOWSKI J, MEDAGLIANI P, et al. Joint routing and scheduling for large-scale deterministic IP networks [J]. Computer Communications, 2021, 165: 33-42.
- [13] PENG G, WANG S, HUANG Y, et al. Traffic Shaping at the Edge: Enabling Bounded Latency for Large-scale Deterministic Networks[C]// 2021 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2021.
- [14] KROLIKOWSKI J, MARTIN S, MEDAGLIANI P, et al. Joint Routing and Scheduling for Large-Scale Deterministic IP Networks[J]. arXiv: 2004. 02717, 2020.
- [15] WU B W, WANG S, TAN W Q. An end-to-end deterministic network transmission scheduling mechanism with cross domain decoupling [J]. Journal of Beijing University of Posts and Telecommunications, 2023, 46(5): 47-52.
- [16] TAN W Q, WU B W, WANG S. Deterministic Network Cross Domain Transmission Architecture and DRL Traffic Scheduling Algorithm [J]. Journal of Beijing University of Posts and Telecommunications, 2023, 46(3): 37-42.
- [17] LIU J H, SU Q, CHEN J J, et al. Research on Railway Operation and Maintenance Visualization Based on WebGL Extended Model [J]. Journal of Railway Science and Engineering, 2022, 19(4): 892-900.
- [18] XU S, XU F M, ZHAO C L. Research on self optimization technology of computing power networks based on digital twins [J]. ZTE Technology, 2023, 29(3): 46-50.
- [19] WU W, XU O. Digital twins and digital operations [J]. Intelligent Building Electrical Technology, 2023, 17(3): 115-119.
- [20] YE X Q, WANG D, LIU Y Y, et al. Optimization of unmanned warehouse scheduling based on digital twins and deep reinforcement learning [J]. Logistics Technology, 2023, 42(9): 97-100, 121.
- [21] QI B, LIU S F, LI B, et al. Optimization Algorithm for Power Demand Response Business Scheduling Based on Business Priority and SRLG [J]. Grid Technology, 2019, 43(7): 10.
- [22] WANG Y F, WANG C, XU Y T, et al. Digital Twin Task Scheduling Method for Intelligent Manufacturing Unit Operations under Edge Cloud Collaboration [J]. Journal of Mechanical Engineering, 2024, 60(6): 137-152.
- [23] LU Y, ZHAO G, XU C, et al. A Framework for Digital Twin-Based Deterministic Communication in Satellite Time Sensitive Networks[C]// IEEE International Conference on Communication (ICC 2023). 2023: 6301-6306.



WANG Kewen, born in 1980, master, senior engineer. His main research interests include railway information, communication, and automation.



ZHANG Weiting, born in 1992, Ph.D, associate professor. His main research interests include industrial Internet of Things, deterministic networks, edge intelligence, and machine learning for network optimization.

(责任编辑:杨雪敏)