

基于软件定义车载网络的 QoS 路由规划研究*

崔峻玮¹ 翟亚红²

(1. 兰州信息科技学院计算机与人工智能学院, 兰州 730300; 2. 湖北汽车工业学院电气与信息工程学院, 十堰 442002)

【摘要】为了提高车载网络中的数据传输速率和保障业务的 QoS 需求, 结合软件定义车载网络(SDVN)技术, 设计了一种基于 SDVN 的深度强化学习 QoS 路由算法。该算法可以实现智能化控制和优化管理车载网络中的数据传输, 以保证车载网络流量的控制、分配和监控, 提高车载数据传输质量和效率。试验结果表明, 该路由算法能较好地降低车载网络的时延, 与传统路由算法相比, 该算法具有更好的优化性能。

主题词:软件定义车载网络 服务质量 路由规划 深度强化学习

中图分类号: TP393; U462.2 **文献标志码:** A **DOI:** 10.19620/j.cnki.1000-3703.20231160

Research on QoS Routing Planning Based on Software-Defined Vehicular Network

Cui Junwei¹, Zhai Yahong²

(1. School of Computer and Artificial Intelligence, Lanzhou College of Information Science and Technology, Lanzhou 730300; 2. School of Electrical & Information Engineering, Hubei University of Automotive Technology, Shiyan 442002)

【Abstract】With the increasing applications of new technologies such as smart driving, autonomous driving and Internet connectivity, traditional in-vehicle networks are difficult to meet the Quality of Service (QoS) demands of diverse applications. In order to improve the data transmission rate and guarantee the QoS demand of services in the in-vehicle network, a deep reinforcement learning QoS routing algorithm based on SDVN is designed in combination with Software-Defined Vehicular Network (SDVN) technology. The algorithm can realize intelligent control and optimized management of data transmission in the in-vehicle network to ensure the control, distribution and monitoring of in-vehicle network traffic and improve the quality and efficiency of in-vehicle data transmission. The experimental results show that the routing algorithm can better reduce the delay of the in-vehicle network and has better optimization performance compared to the traditional routing algorithm.

Key words: Software-defined vehicular network, Quality of service, Route planning, Deep reinforcement learning

【引用格式】崔峻玮, 翟亚红. 基于软件定义车载网络的 QoS 路由规划研究[J]. 汽车技术, 2025(1): 26-32.

CUI J W, ZHAI Y H. Research on QoS Routing Planning Based on Software-Defined Vehicular Network[J]. Automobile Technology, 2025(1): 26-32.

1 前言

目前大部分车辆网络系统采用 CAN 和 LIN 通信技术, 但面临信号负载高、延时增加等问题^[1]。随着车辆智能网联化不断发展, 需采取硬件最大化设计、软件迭代升级策略, 但可导致零部件内部数据量大、更新软件包大、现有 CAN 总线速率低、售后软件更新时间长等问

题。因此, 需要采用以太网等高速通信技术。软件定义车载网络 (Software-Defined Vehicular Network, SDVN) 是一种基于软件定义网络 (Software Defined Networking, SDN) 技术在汽车领域的应用^[2]。SDVN 旨在实现网络虚拟化和集中管理, 提高网络通信效率和用户体验^[3-4]。

目前, 各类智能算法性能突出, 车载网络路由规划结合智能算法是当前研究的热点问题之一。郭荣梅等

*基金项目: 湖北省教育厅科研计划重点项目 (D20211802); 湖北省科技厅计划项目 (2022BEC008)。

通信作者: 翟亚红 (1979—), 女, 教授, 研究方向为智能网联和人工智能技术, zhaiyh_dy@huat.edu.cn。

人^[5]研究了使用 K 条最短路径算法计算出最小时延和最大可用带宽的 K 条路径,但速度较慢;尹凤杰等人^[6]提出了改进的蚁群路由算法,可以保证网络稳定且利用率高,但存在收敛速度慢和时延问题;徐啸等人^[7]提出了基于强化学习的多路径路由算法,能实时选择最优路径,但时间复杂度高;Sun等人^[8]提出了基于深度强化学习的可扩展路由算法,但面对复杂网络时整体性能下降。多流量共存时的服务质量(Quality of Service, QoS)感知路由问题复杂度高,深度强化学习可应对此类问题。Pham等人^[9]使用深度强化学习(Deep Reinforcement Learning, DRL)智能体和卷积神经网络,在SDN下提高QoS感知路由的性能。Jalil等人^[10]提出了一种深度Q路由(Deep Q-Routing, DQR)算法,使用具有经验回放的深度Q网络计算出满足多个QoS指标的合理路径。

综上,当前传统路由方法已无法为不同类型的网络流量分配合适的路由资源以满足其QoS需求。由于车载网络路由问题具有马尔可夫性^[11-12],本文提出基于SDVN的深度强化学习路由(Deep Reinforcement Learning Routing, DRLR)算法^[13-14],为车载网络提供更加稳定可靠的数据传输环境,同时保证车辆内部各种应用设备的良好运行效果,并通过与其他算法进行对比试验,验证算法的有效性。

2 基于SDVN的DRLR模型

2.1 软件定义车载网络

传统分布式网络由于通信带宽、计算能力不足和软件升级不便等瓶颈,已不能满足当前智能网联技术的发展需求。传统车载网络骨干正由CAN/LIN总线向以太网方向发展,以太网能够满足高速传输、高通量、低延迟等性能需求,同时也可以减少安装、测试成本。使用以太网作为主干网络,通过中央控制器计算平台可以实现多种信息的融合处理,实现整车协同功能,增强车载网络的交互性,实现软件的在线升级、操作系统可移植等功能。软件定义车载网络整体架构如图1所示。

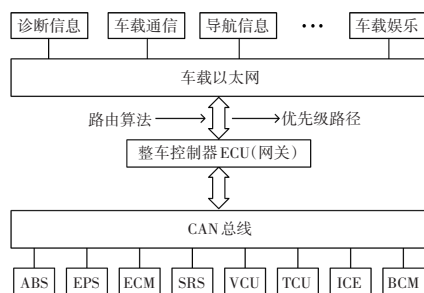


图1 软件定义车载网络整体架构

2.2 车载以太网QoS

不同类型的车载以太网流量在时延、带宽和丢包率等方向的QoS需求各有差异。车载以太网流量的优先级可以根据具体的应用和需求而有所不同,划分成4个QoS流量业务类别,即车辆诊断流量、车载通信流量、车载导航流量、车载多媒体流量。

a. 车辆诊断流量。这类流量用于将车辆诊断信息发送到网络上的服务中心,以进行远程诊断和维护。同时,服务中心也可以向车辆发送指令或更新,以改善车辆性能或解决问题,对快速诊断和故障排查具有重要作用,对时延、带宽、丢包率有较高的要求。

b. 车载通信流量。这类流量包括车内设备之间的通信和连接到互联网的数据传输,对时延、带宽、丢包率有一定要求,如车载通信实时语音、车载摄像头捕捉到的图像、互联网浏览等,此流量对网络传输的时延有较高的要求,需要确保良好的用户体验。

c. 车载导航流量。这类流量用于传输导航和地图数据,对于提供准确的导航指引和实时交通信息较为重要。车辆可以接收到来自网络的实时交通信息,以便选择最佳路线。同时,车辆的位置和速度等信息也可以上传到网络,从而为其他车辆提供实时交通状态。对时延、带宽、丢包率有一定要求。

d. 车载多媒体流量。这类流量包括音频、视频等娱乐内容的传输,对丢包率有较高的要求,对时延和带宽有一定要求,对于车辆的运行和安全性没有直接影响。

本文根据不同的车载流量类别进行网络QoS流量分类,如表1所示。

表1 车载网络流量类别

流量类型	应用类型
车辆诊断流量	车辆诊断信息等
车载通信流量	车载实时通信、互联网浏览等
车载导航流量	导航传输和地图数据等
车载多媒体流量	音频、视频数据等

2.3 马尔可夫决策过程

为了将深度强化学习引入网络路由决策中,需要将路由问题表示为一个马尔可夫决策过程(Markov Decision Process, MDP)^[15]。一个状态到另一个状态是通过执行动作策略产生的,这个过程可以用一组五元组 $\langle S, A, P, R, \gamma \rangle$ 表示。其中, S 为状态空间; A 为动作空间; P 为状态转移概率矩阵; R 为奖励函数; $\gamma \in (0,1)$ 为衰减因子,表示过去的奖励值在当前时刻的比例; γ 倾向于接近1,表示对未来的奖励有更大的偏好。

在网络中,当前的网络状态是过去路由状态的聚合。当前时刻的路由决策只与当前时刻的网络状态和流量信息有关,不受历史网络状态信息的影响。若当前的路由决策只依赖于当前的网络状态,而与之之前的状态无关,那么这个过程可视为具有马尔可夫性质。此外,路由决策问题可以表述为一个顺序决策问题。路由决策智能体与网络环境反复交互,在每个时间步骤的开始,它观察当前的网络状态,并从有限的状态集合中选择一个;然后,智能体移动至下一个新的状态并获得相应的奖励;循环地进行状态选择,得到一组最大化的奖励^[16-17]。

2.4 深度强化学习路由模型

本文将路由问题建模为一个MDP,并引入深度强化学习技术。首先,在一个时间间隔内,智能体需要观察网络拓扑中的环境 S_t ,并根据当前的策略做出路由决策;然后,在动作空间中选择相应的动作 A_t 执行,网络状态转为 S_{t+1} 并将奖励值反馈给智能体;其次,不断重复上述的步骤来反馈奖励值,使累积奖励值最大化;最后,通过不断改进策略来实现奖励值增大,从而获得最大的累积奖励,以输出符合QoS需求的最优路径。

QoS路由的本质是寻找满足QoS需求的端到端最优路径。为了描述这一问题,首先将整个网络建模为一个无向图 $G=(V,E)$ 。其中, V 为图中的节点集合, $V=\{v_1,v_2,\dots,v_n\}$; n 为节点数量, v_i 表示网络中交换机或路由器; E 为图中的边集,表示网络中的节点间的链路。从节点 v_i 出发的第 j 条边,其中 d_{ij} 、 b_{ij} 、 l_{ij} 表示QoS特征值,分别为时延、带宽和丢包率。智能体位于SDN控制平面,环境位于SDN数据平面。基于SDN的深度强化学习QoS路由模型如图2所示。

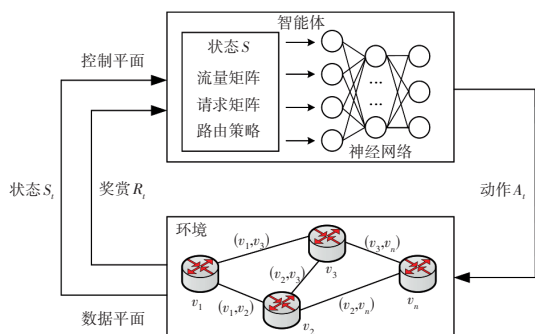


图2 深度强化学习路由算法模型

此模型中包括的概念描述如下:

a. 路由策略。一个路由策略是单个请求的一个合理路径。对于图中存在的每一对节点, $R(i,j)$ 定义为流量请求从节点 i 传输到节点 j 所经过的所有节点。例如,从节点1传输到节点4的流量,从节点1出发,途中经过

节点2和节点3,最终到达节点4,路由 $R(1,4)$ 生成的路径信息即为 $[1,2,3,4]$ 。其中,路径不能出现环路,意味着路径信息中不包括重复的节点。

b. 流量矩阵。在此模型中有 n 个网络节点,使用一个大小为 $n \times n$ 的矩阵 D 来表示网络中所有节点之间传输的数据量。 $D(i,j)$ 定义为从节点 i 传输到节点 j 的数据量,其中矩阵对角线的值被设置为-1。

c. 请求矩阵。在此模型中有 n 个网络节点,使用一个大小为 $n \times n$ 的矩阵 C 来表示网络中所有节点之间等待传输的数据量。 $C(i,j)$ 定义为从节点 i 等待传输到节点 j 的数据量,其中矩阵对角线的值被设置为-1。

d. 深度神经网络。深度神经网络(Deep Q-Network, DQN)通过输入的环境状态信息来输出当前状态下所有动作的 Q 值,即深度神经网络中用来评估特定状态下某个动作的预期回报,随后选择所有动作中 Q 值最大的动作执行。DQN中隐藏层数量和隐藏层中神经元的数量等均与输入和输出的数据规模相关。通常,输入数据的规模越大,维度越高,隐藏层的数量以及神经元的数量就越多,如果输入的数据为图片或者时序数据,则可能需要使用具有卷积层的卷积神经网络或具有循环结构的循环神经网络。在本文所设计的算法中,DQN中深度神经网络的结构如图3所示。

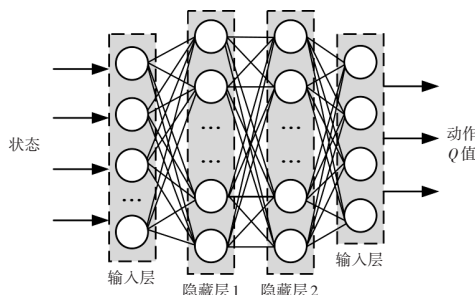


图3 DQN深度神经网络

该方案中神经网络的输入是网络状态信息经过预处理后形成的一维数组,因此省去了卷积层而只使用全连接层作为隐藏层。图3中神经网络中全连接层的数量和全连接层中神经元的数量并不是固定的,需要根据网络规模确定。根据大量试验结果,具有2个全连接层,每个全连接层中神经元数量为150个的神经网络结构在主机数量小于20台、交换机数量小于30台的小规模网络中表现较好。对于规模更大的网络,则需要适当增加全连接层或神经元的数量。对于DQN,最为直观的输入输出方式是将网络中的流量请求矩阵作为神经网络的输入,将所有支流的转发路径选择作为输出。

将每步迭代的总时间分为几个连续的时间段,定义

为 t 。在时间区间 t 的开始,从请求矩阵 $C(i, j)$ 读取节点 i 和节点 j 之间的流量请求信息,以覆盖流量矩阵 $D(i, j)$ 的值。然后,神经网络根据流量矩阵信息和所有历史时间区间 $(1, 2, \dots, t)$ 的路由策略来评估流量请求信息,得出从节点 i 到节点 j 的路由策略。路由策略 $R(i, j)$ 选定后,路由表也相应确定。若请求矩阵 $C(i, j)$ 中的所有值都被遍历,本次迭代结束,下次迭代开始,重复上述步骤。

3 深度强化学习路由算法设计

深度强化学习任务需要通过马尔可夫决策过程进行描述,对于用深度强化学习技术解决QoS路由分配问题,本文用5个部分来模拟该问题:智能体、环境、状态空间、动作空间和奖励函数。由于状态和动作维度较大,尝试使用深度强化学习算法来实现合理的路由策略。

a. 智能体。首先,智能体负责执行动作并获得环境的反馈,智能体计算路由表,与流量矩阵一起写入Mininet环境中;然后,从Mininet环境中获得时延、带宽、丢包率等QoS指标作为反馈;最后,智能体将状态推送到下一步。

b. 环境。本算法的环境主要建立在Mininet网络仿真环境中,其中主要包括路由器、交换机和Ryu控制器。

c. 奖励功能。本文利用QoS指标来评估网络的状态。采用常见并重要的指标作为反馈,即时延 d_{ij} 、带宽 b_{ij} 、丢包率 l_{ij} 。奖励函数为:

$$R_t = \beta d_{ij} + \theta b_{ij} + \varphi l_{ij} \quad (1)$$

式中: β, θ, φ 为QoS指标的权重值,不同的QoS需求通过设置权重来实现奖励值。

d. 动作空间。深度强化学习算法需要一个固定的状态空间,在每次迭代中都是一致的。因此,本文的状态空间采用无环路路径。由于状态空间的维度过大,通过部分训练方法来减少迭代次数。在每个时间间隔 t 开始,从请求矩阵 $C(i, j)$ 中读取节点对 (i, j) 。智能体需要为该节点分配一条路径,所选的动作必须满足源-目节点的要求,在 S_n 状态下参考 $Q(S_n, A_i)$ 执行的动作为 A_i 。因此,每次智能体选择动作时,控制程序只是选择满足要求的路径,而不是遍历整个路由表。这种机制使得每时刻可供选择的动作空间范围大幅缩小。

例如,当智能体处于状态 S_1 ,从请求矩阵 $C(i, j)$ 中读取的节点对是(2, 6),即需要从节点2到节点6寻找一条路径。假设在整个动作空间中只有动作2到动作10是对应的,这意味着有9个动作可选。在本次迭代中,只有 $Q(S_1, A_2)$ 到 $Q(S_1, A_{10})$ 的值被更新,这个状态下的其他

Q 值保持不变。因此,对于 Q 值表中的每个状态,只有小部分的 Q 值会被训练,如表2所示。

表2 动作空间的 Q 值表

状态	A_1	A_2	\dots	A_{10}	\dots	A_n
S_1	$Q(S_1, A_1)$	$Q(S_1, A_2)$	\dots	$Q(S_1, A_{10})$	\dots	$Q(S_1, A_n)$
S_2	$Q(S_2, A_1)$	$Q(S_2, A_2)$	\dots	$Q(S_2, A_{10})$	\dots	$Q(S_2, A_n)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
S_n	$Q(S_n, A_1)$	$Q(S_n, A_2)$	\dots	$Q(S_n, A_{10})$	\dots	$Q(S_n, A_n)$

e. 状态空间。与Q-Learning和状态-动作-奖励-状态-动作(State-Action-Reward-State-Action, SARSA)算法相比,深度强化学习算法可以处理一个非常大的状态空间。因此,该算法使用流量矩阵 $D(i, j)$ 作为状态空间。此状态空间在每次迭代时只更新所有数值中的一个值。

基于上述描述的DRLR算法模型,设计了DRLR路由规划算法,DRLR算法描述如表3所示。

表3 DRLR算法

输入:节点对 (v_i, v_j)
输出:路由策略路径 $R(i, j)$
初始化经验回放和训练集;
初始化动作价值函数 Q 和网络参数 θ ;
设置 N 为迭代次数; t 为时间间隔区间; T 为整个时间区间;
for (episode=1; episode< N ; episode++)
初始化状态空间 S_1 和流量矩阵 D
step=0;
for ($t=1$; $t<T$; $t++$)
用概率 ε 选择一个随机动作 A ;
否则选择最大 $Q(S_t, A)$ 值对应的动作 A ;
在Mininet中执行动作 A ;
得到奖励 R 和下一个状态序列 S_{t+1} ;
在记忆池中储存临时数据 (S_t, A, R, S_{t+1}) ;
设置下一个状态序列 $S_t = S_{t+1}$;
if (step>200) & (step%5=0) then
从记忆池中随机抽取小批次的临时数据 (S_t, A, R, S_{t+1}) 进行训练;
for 所有样本
设置目标网络更新公式 $Y = R + \gamma \max Q(S_{t+1}, A)$;
设置损失函数 $L = (Y - Q(S_t, A))^2$;
通过计算目标网络 Y 值和 L 值用来训练神经网络;
step + 1;
end for
end if
end for
end for

首先,DRLR算法初始化后,控制平面读取本次迭代的输入,即节点对 (v_i, v_j) 和流量矩阵 D ,算法立即启动。根据 ε -greedy策略选择一个动作,并由智能体在Mininet网络仿真环境中执行;其次,智能体从Mininet网络仿真环境中获得环境反馈QoS指标时延、带宽和丢包率;然后,将当前状态、下一状态、所选动作和环境反馈

存储在内存卡记忆池中,当步骤数超过200后,每5步进行一次迭代。迭代过程是从记忆池中随机选择样本训练集,根据深度强化学习算法得到目标网络Y值,并通过梯度下降法进行优化计算损失函数L值,反向传回神经网络。

4 试验仿真与分析

4.1 试验仿真环境

为了对DRLR算法的有效性进行验证,本文在Ubuntu 20.04系统上使用Ryu、Mininet、OpenvSwitch搭建SDN网络仿真环境。试验所用的网络拓扑是经典的NSFNet拓扑,拓扑的链路时延设置为20ms,链路最大带宽设置为10M,并模拟发送车辆诊断流量、车载通信流量、车载导航流量、车载多媒体流量4种类型流量,持续时间为10min,时间间隔为5s,输出5条符合QoS需求的路径,合理分配路径给4种流量进行数据传输,保障其传输质量。试验拓扑如图4所示。

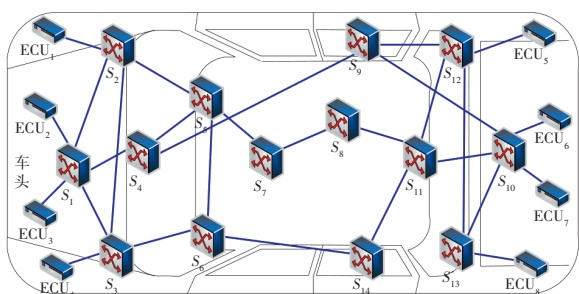


图4 DRLR路由算法拓扑

算法参数设置如表4所示。

表4 DRLR算法参数

参数	数值
学习率 α	0.6
奖赏衰减值 γ	0.9
探索率 ϵ	0.1
训练总轮数/次	800
每批次样本数/个	64
记忆池容量/个	5 000

4.2 试验对比与分析

以传输车载通信流量为例,因为车载通信流量对时延要求很高,设置奖励值参数时延权重 β 为0.8、带宽权重 θ 为0.1、丢包率权重 φ 为0.1。并采用Dijkstra算法、蚁群算法和DRLR算法在同等条件下进行了对比分析,设置每进行5个步骤对传输时延进行一次采样。

Dijkstra算法在时延表现方面并不稳定,有较为严重的时延抖动,因为此算法无法保障不同类型流量差异

化的服务质量。Dijkstra算法时延曲线如图5所示。

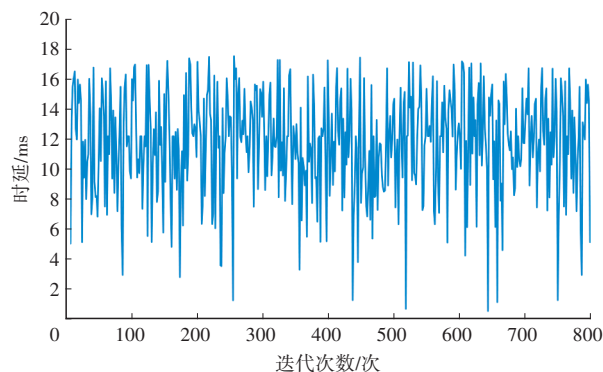


图5 Dijkstra算法时延曲线

蚁群算法相比传统路由算法,时延表现上较为稳定,抖动发生在8ms以内,可以较好地保障不同类型数据传输质量,但存在一定的抖动。蚁群算法时延曲线如图6所示。

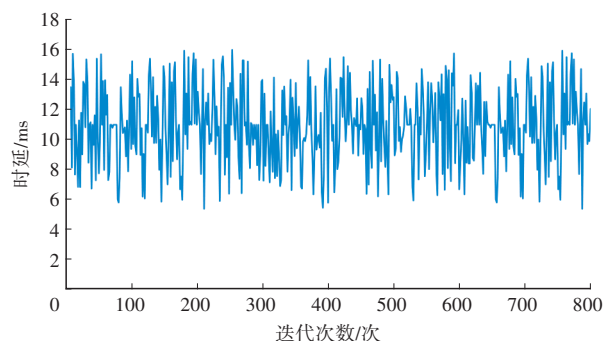


图6 蚁群算法时延曲线

DRLR算法与前2种算法进行对比,可以明显看到时延收敛更为稳定,抖动现象控制在4ms以内。DRLR算法时延曲线如图7所示。

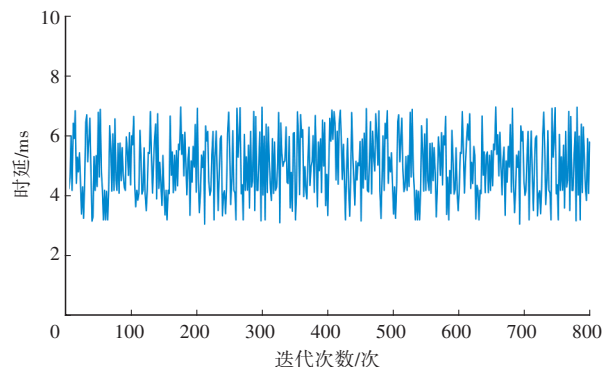


图7 DRLR算法时延曲线

本文将车辆诊断流量、车载通信流量、车载导航流量、车载多媒体流量通过车载以太网进行流量传输,减轻CAN等总线的负载,使得CAN总线可以更好地传输车辆控制、状态监测等方面的信息,必须确保车辆可靠、无误地为驾驶员提供实时指示,以便驾驶员做出快速反应和决策,车辆CAN总线报文如图8所示,车辆控制信息等报文传输更稳定。

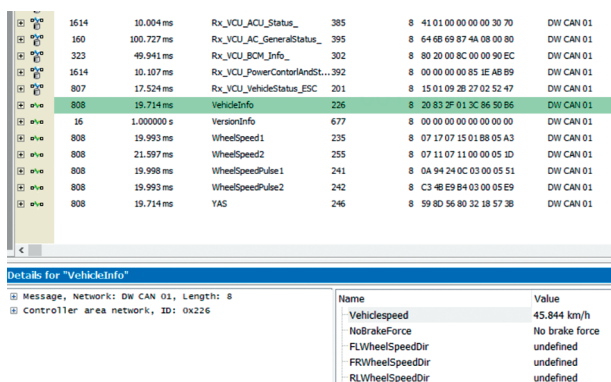


图8 车辆CAN总线报文

在整个训练中,在大约300轮训练后趋于稳定,奖励值有稳定的上升趋势,这表明DRLR算法有良好的收敛性。DRLR归一化奖励变化值曲线如图9所示。

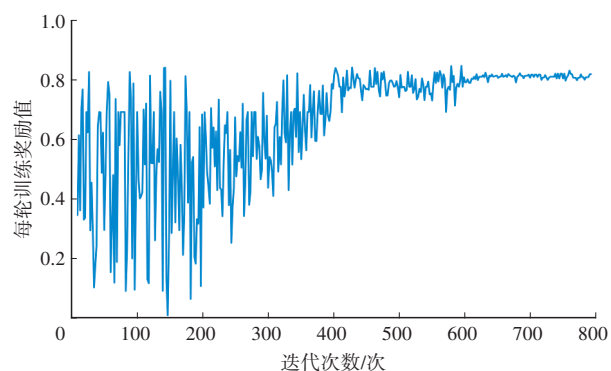


图9 DRLR归一化奖励变化值曲线

本节测试以传输车载通信流量为例,可以看到,在训练第300个循环左右时,可以趋于稳定输出路径。按时延性能“由强到弱”生成5条最优路径,并且按照网络流量优先级,将第一条最优路径分配给车载通信流量进行转发,从而对不同类型业务的QoS需求进行保障。测试结果充分证明,所提算法具备有效性和稳定性。DRLR车载通信流量路由规划如图10所示。

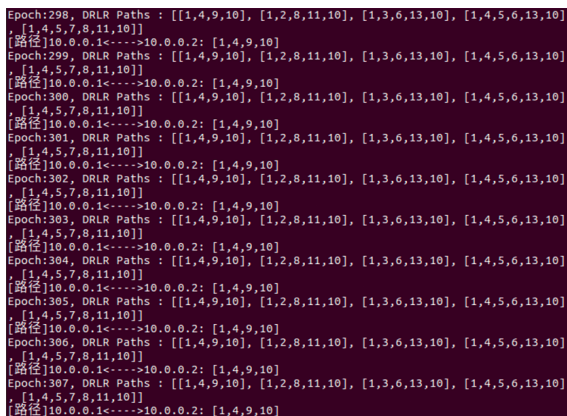


图10 DRLR车载通信流量路由规划

5 结束语

本文提出了一种基于SDN的DRLR路由算法,旨在
2025年 第1期

解决CAN总线负载过高、无法满足车载通信流量等QoS需求的问题。该算法将时延、带宽和丢包率QoS指标加入奖励函数中,以保障不同类型网络流量业务的QoS需求,并动态地实时分析网络状态进行路由规划。试验结果表明,相比其他传统路由算法,本文所提的DRLR路由算法传输时延降低至4ms左右,具有更快更稳定的传输速率,能够有效针对不同类型的车载流量进行模块之间的路由通信。本文还需进一步探索如何解决在车载网络环境中网络拥塞、时延等问题,并继续提高网络的性能和可靠性。

参考文献

[1] 谷晓会,章国安. SDN在车载网中的应用综述[J]. 计算机科学, 2020, 47(1): 237-244.
GU X H, ZHANG G A. A Review of SDN Applications in Vehicular Networks[J]. Computer Science, 2020, 47(1): 237-244.
[2] 翟亚红,崔峻玮. 软件定义网络安全研究进展[J]. 科技导报, 2023, 41(13): 76-88.
ZHAI Y H, CUI J W. Advances in Software-Defined Network Security Research[J]. Science and Technology Herald, 2023, 41(13): 76-88.
[3] 崔峻玮,翟亚红. 基于SDN的多媒体流QoS队列调度机制研究[J]. 湖北大学学报(自然科学版), 2023, 45(3): 404-413.
CUI J W, ZHAI Y H. Research on QoS Queue Scheduling Mechanism for Multimedia Streams Based on SDN[J]. Journal of Hubei University (Natural Science Edition), 2023, 45(3): 404-413.
[4] 崔峻玮,翟亚红. 基于自适应遗传蚁群算法的SDN路由规划研究[J]. 天津理工大学学报, 2023, 39(3): 39-48.
CUI J W, ZHAI Y H. Research on SDN Route Planning Based on Adaptive Genetic Ant Colony Algorithm[J]. Journal of Tianjin Polytechnic University, 2023, 39(3): 39-48.
[5] 郭荣梅,胡小兵. 求解时间窗口网络中前k条最短路径的方法[J]. 电子学报, 2020, 48(7): 1387-1395.
GUO R M, HU X B. A Method for Solving the First k Shortest Paths in Time Window Networks[J]. Electronics Letters, 2020, 48(7): 1387-1395.
[6] 尹凤杰,褚群森. 基于改进蚁群算法的QoS路由研究[J]. 辽宁大学学报(自然科学版), 2020, 47(4): 312-318.
YIN F J, CHU Q S. Research on QoS Routing Based on Improved ant Colony Algorithm[J]. Journal of Liaoning University (Natural Science Edition), 2020, 47(4): 312-318.
[7] 徐啸,顾玲丽,陈建平. 一种智能多路径路由及子流分配协同算法[J]. 计算机工程, 2021, 47(9): 136-144+152.
XU X, GU L L, CHEN J P. An Intelligent Multipath Routing and Subflow Assignment Cooperative Algorithm[J].

- Computer Engineering, 2021, 47(9): 136-144+152.
- [8] SUN P, GUO Z, LI J. Enabling Scalable Routing in Software-Defined Networks with Deep Reinforcement Learning on Critical Nodes[J]. IEEE/ACM Transactions on Networking, 2021, 30(2): 629-640.
- [9] PHAM T A Q, HADJADJ-AOUL Y, OUTTAGARTS A. Deep Reinforcement Learning Based QoS-aware Routing in Knowledge-Defined Networking[C]// 14th EAI International Conference. Ho Chi Minh City, Vietnam: Springer Publisher, 2019: 14-26.
- [10] JALIL S Q, REHMANI M H, CHALUP S. DQR: Deep Q-Routing in Software Defined Networks[C]// 2020 International Joint Conference on Neural Networks (IJCNN). Padua, Italy: IEEE, 2020: 1-8.
- [11] 廖岑卉珊, 陈俊彦, 梁观平, 等. 基于深度强化学习的SDN服务质量智能优化算法[J]. 物联网学报, 2023, 7(1): 73-82.
- LIAO C H S, CHEN J Y, LIANG G P, et al. Intelligent Optimization Algorithm for SDN Quality of Service Based on Deep Reinforcement Learning[J]. Journal of Internet of Things, 2023, 7(1): 73-82.
- [12] 黄万伟, 郑向雨, 张超钦. 基于深度强化学习的智能路由技术研究[J]. 郑州大学学报(工学版), 2023, 44(1): 44-51.
- HUANG W W, ZHEN X Y, ZHANG C Q. Research on Intelligent Routing Technology Based on Deep Reinforcement Learning[J]. Journal of Zhengzhou University (Engineering Edition), 2023, 44(1): 44-51.
- [13] 曹景祥, 刘其成. 基于深度强化学习的路径规划算法研究[J]. 计算机应用与软件, 2022, 39(11): 231-237.
- CAO J X, LIU Q C. Research on Path Planning Algorithm Based on Deep Reinforcement Learning[J]. Computer Applications and Software, 2022, 39(11): 231-237.
- [14] 黄鑫陈, 陈光祖, 郑敏. 基于Q-learning的飞行自组织网络QoS路由方法[J]. 中国科学院大学学报, 2022, 39(1): 134-143.
- HUANG X C, CHEN G Z, ZHEN M. Q-learning Based QoS Routing Method for Flying Self-Organizing Networks[J]. Journal of Chinese Academy of Sciences University, 2022, 39(1): 134-143.
- [15] 龙浩, 张书奎, 张力. 基于马尔科夫决策过程的高能效任务调度算法[J]. 计算机应用与软件, 2020, 37(1): 242-246+291.
- LONG H, ZHANG S K, ZHANG L. Energy Efficient Task Scheduling Algorithm Based on Markov Decision Process[J]. Computer Applications and Software, 2020, 37(1): 242-246+291.
- [16] TSAI K C, ZHUANG Z, LENT R. Tensor-Based Reinforcement Learning for Network Routing[J]. IEEE Journal of Selected Topics in Signal Processing, 2021, 15(3): 617-629.
- [17] CASAS-VELASCO D M, RENDON O M C, DA FONSECA N L S. Intelligent Routing Based on Reinforcement Learning for Software-Defined Networking[J]. IEEE Transactions on Network and Service Management, 2020, 18(1): 870-881.

(责任编辑 王 一)

修改稿收到日期为2024年4月6日。