

基于深度强化学习的交叉口信号配时优化方法

李 琦, 冷荣梦

(上海电科智能系统股份有限公司, 上海市 200063)

摘 要: 随着城市化进程的加速, 交通拥堵成为影响城市效率和居民生活质量的主要问题之一。特别是在交叉口车流量高峰期, 不合理的信号配时方案设置常常导致严重的交通堵塞和排队溢流。为减少高峰期交叉口二次排队车辆数和最大排队长度, 提高交通流的通行效率, 通过深度强化学习中的 DQN (Deep Q Network, 深度 Q 网络) 算法来优化交叉口信号灯配时方案。首先构建交叉口仿真模型作为 DQN 算法的学习环境; 其次接受动作并反馈回报; 最后提取交通运行体征指标来评价 DQN 算法在交叉口信号控制中的优化效果。研究结果显示, 相比韦伯斯特信号配时方法, DQN 算法可以将最大排队长度降低 14%, 交叉口车均延误降低 16%, 证明 DQN 算法下的信号控制方案能更有效地适应交通流变化, 提高交叉口的车流运行效率。

关键词: 信号优化; 深度强化学习; DQN 算法; 排队溢出

中图分类号: U491.5⁺1

文献标志码: A

文章编号: 1009-7716(2025)12-0001-06

Optimization of Signal Timing at Intersections Based on Deep Intensive Learning

LI Qi, LENG Rongmeng

(Shanghai SEARI Intelligent System Co., Ltd., Shanghai 200063, China)

Abstract: With the acceleration of urbanization, traffic congestion has become one of the major issues impacting the urban efficiency and the life quality of the residents. Particularly during the peak traffic flow at intersections, the unreasonable signal timing schemes frequently result in severe traffic congestion and queue overflows. In order to reduce the number of vehicles experiencing secondary queuing and the maximum queue length during peak hours at the intersections, and enhance the efficiency of traffic flow, the DQN (Deep Q Network) algorithm in the deep intensive learning is used to optimize signal timing schemes at intersections. Firstly, the simulation model of intersection is established as the learning environment of DQN algorithm. Secondly, the actions are accepted and the feedback is given. Finally, the traffic operation indicators are extracted to assess the optimization effects of the DQN algorithm in intersection signal control. The research result shows that compared to the Webster signal timing method, the DQN algorithm can reduce the maximum queue length by 14% and the intersection vehicle delay by 16%, demonstrating that signal control schemes under the DQN algorithm can more effectively adapt to fluctuations in traffic flow, thus improving the operational efficiency of vehicle flow at intersections.

Keywords: signal optimization; deep intensive learning; DQN algorithm; queue overflow

0 引 言

随着城市化的快速发展, 交通拥堵已成为全球许多城市面临的主要问题之一。交叉口的拥堵直接影响着城市交通系统的运行效率, 为缓解该拥堵, 优

化交叉口信号方案成为了关键措施。传统信号控制系统多依赖于固定时序计划, 无法灵活应对交通流量的动态变化, 导致高峰期二次排队和长队列形成, 而低峰期绿灯时间浪费。深度强化学习(DRL)算法作为一种强大的决策优化工具, 因其无需人工介入即可从环境交互中学习, 已被证明在交通信号优化中具有显著优势。通过基于 DRL 的信号控制, 可以实时调整交通信号时序, 优化交通流量, 提高交叉口的通行能力。

交叉口信号优化是交通领域的重要研究方向, 已有大量研究聚焦于改进城市交通管理系统, 提高

收稿日期: 2025-02-13

基金项目: 国家科技部重点研发计划“自主式交通计算技术”(2023YFB4301900)

作者简介: 李琦(1985—), 男, 博士, 高级工程师, 从事智能交通系统研发工作。

通信作者: 冷荣梦(1999—), 女, 硕士, 工程师, 从事智能交通系统算法与研发工作。电子信箱: lengrongmeng@163.com

交叉口通行效率^[1],减少交通拥堵,改善交通安全性。现有研究主要关注交叉口的整体延误^[2]、车均延误^[3]、排队长度^[4]等指标,或结合多个优化指标建立多目标优化函数^[5],评估信号控制效果。此外,部分研究还探索了节能和排放优化,如油耗^[6]和碳排放^[7]。为了提高交叉口通行效率,信号优化通常通过调整关键参数如相序^[8]、周期时长^[9]和绿信比^[10]等实现。

交通模式的多样化下,现代信号控制系统逐渐采用多模态交通管理模式,考虑不同交通参与者的需求,如自行车、行人^[11]以及公共和应急车辆^[12]的优先通行。此外,还需结合交通运行环境、恶劣天气^[13]等因素,以及智能网联车^[14]和道路干线协同^[15]等技术,以进一步提升交叉口效率。

随着信息和通信技术的发展,智能交通设备为交叉口信号优化提供了基础。研究者采用传感器、通信设备和大数据分析等技术,结合机器学习^[16]、启发式算法和深度学习^[17]等智能算法,实现动态信号调整。深度强化学习中的深度Q网络(DQN)算法在此领域应用广泛,通过设计状态空间^[18]、选择动作集^[19]和奖励函数^[20],学习最优信号控制策略。此外,神经网络^[21]方法也被用于基于历史数据预测未来流量,辅助信号时序调整。

本文结合交叉口的排队情况,重点优化二次排队车辆数和最大排队长度,采用DQN算法训练交叉口最优信号控制策略,同时通过SUMO仿真软件^[22]构建交叉口车流运行的仿真环境,验证DQN算法在交叉口信号优化中的有效性。

1 DQN 算法理论

深度Q网络算法是强化学习领域的一个重要突破,它将深度学习和强化学习相结合,用于解决具有高维状态空间的问题。

Q学习作为机器学习方法中的一种,智能体通过环境的互动来学习如何做出决策。在这个过程中,智能体尝试不同的行动,并从环境那里获得反馈。这个反馈包括新的环境状态和奖励,奖励告诉智能体它的行动是好是坏。在Q学习中,Q值表示在特定状态下采取某个行动可能得到的总奖励,智能体的目标是最大化这些Q值,来学习在某种状态下应该采取哪个行动来获得最大回报。Q值更新的表达式为:

$$Q(s_{i+1}, a_{i+1}) = Q(s_i, a_i) + \alpha(r_{i+1} + \gamma \max_a Q(s_{i+1}, a) - Q(s_i, a_i)) \quad (1)$$

式中: $Q(s_i, a_i)$ 为在*t*时刻当前状态*s*和动作*a*下的*Q*值; α 为学习率; r_{i+1} 为执行动作*a*后接收到的即时奖励; γ 为折扣因子,用于平衡即时奖励与未来奖励的重要性; s_i 为当前状态; a_i 是在当前状态下采取的动作。

Q学习在面对低维状态空间时效果显著,但难以处理高维状态空间。因此提出深度Q网络(DQN),通过应用深度神经网络来逾越这一障碍,利用网络的强大非线性拟合能力来近似*Q*值函数。其中网络参数 θ 被训练以最小化预测*Q*值与目标*Q*值之间的差异,即损失函数*L*(θ),其表达式为:

$$L(\theta) = E[(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta_x))^2] \quad (2)$$

式中: θ 为当前神经网络的参数; θ^- 为目标网络的参数,用于稳定学习过程,通常每几个步长更新一次;*E*表示期望;*r*表示回报。

DQN框架引入经验回放机制,将智能体的历史转换(*s, a, r, s'*)存储在经验回放池中,后续从中随机抽取小批量样本进行网络训练。这一机制减少了样本之间的时间相关性,提高了数据的利用率,并增强了学习过程的稳定性。为了在探索与利用之间取得平衡,DQN通常采用 ϵ -贪婪策略进行行动选择。在该策略下,智能体以 $1 - \epsilon$ 的概率选择当前*Q*值最大的行动以利用已知策略,以 ϵ 的概率随机选择一个行动以探索未知策略。随着智能体获得更多的经验, ϵ 会按照预定的策略逐步减少,从而使智能体逐渐转向更多地利用其学习到的策略,以实现长期累积奖励的最大化。

2 基于DQN算法的信号控制模型

本节针对交叉口交通运行情况构建基于DQN算法的交通信号控制模型,将交叉口交通运行过程建模为马尔可夫决策过程,设计DQN算法需要的交通运行状态空间和可供选择的信号相位时长动作集,根据动作实施后环境的反馈信息设计奖惩函数。

2.1 状态空间

作为DQN算法中的一个关键部分,状态空间是智能体在环境中观测到的信息,智能体基于状态信息选择动作。因此,状态空间的构建需尽可能接近交叉口的交通运行状况。为了准确描述交叉口的交通运行情况,本文基于交通流仿真环境提取交叉口运行状态来设置状态空间。

对于车辆状态,从车道维度检测并记录每个进口车道上的车辆数、车辆平均速度、排队车辆数,此3个参量结合以反映进口车道*i*的交通运行状态 $s_{vehicle}^i$,其表达式为:

$$s_{vehicle}^i = [n_{vehicle}^i, v_{vehicle}^i, n_{queue}^i] \quad (3)$$

式中:*i*为车道序号; $n_{vehicle}^i$ 为车道*i*上的车辆数; $v_{vehicle}^i$ 为车道*i*上的车辆平均速度; n_{queue}^i 为车道*i*上的排队车辆数。

结合交叉口所有进口车道的状态参量,用于描述交叉口车辆状态 $S_{vehicle}$ 的表达式为:

$$S_{vehicle} = \{s_{vehicle}^i | i \in I\} \quad (4)$$

式中:*I*为交叉口进口车道编号集合, $I=\{1, 2, 3, \dots, i_{max}\}$,其中 i_{max} 为交叉口各进口车道总数量。

交叉口信号状态 S_{signal} 包含当前相位*p*和相位剩余时长 t_r 。因为动作的选择很大一部分取决于当前信号相位和相位剩余绿灯时长,因此将这2个参数纳入状态空间,其表达式为:

$$S_{signal} = [p, t_r] \quad (5)$$

将进口道车辆状态和信号状态参数合并,得到交叉口状态空间向量 S_{state} ,它代表整个交叉口的状态,其表达式为:

$$S_{state} = [S_{vehicle}, S_{signal}] \quad (6)$$

2.2 动作空间

智能体基于环境状态,从动作集中选择一个动作执行,动作执行后再反馈环境状态。在现有研究中,深度强化学习算法中的动作一般有切换相位和改变信号周期的绿信比这2种设置情况。

本文选择切换相位的方式设计动作选择集 A_{action} ,具体设置方式如下:

$$A_{action} = \{0, 1\} \quad (7)$$

式中:动作 action=0 表示不切换相位,动作 action=1 表示切换相位。(1)若当前相位未达到最小绿灯时长,智能体必须选择保持当前相位不变且不改变相位时长,即选择动作0。

(2)若当前相位为绿灯,相位绿灯时长达到最小绿灯时长且小于最大绿灯时长时,智能体根据贪婪策略选择动作。当选择切换相位,即选择动作1,信号立即执行3 s黄灯时间,然后切换到下一相位;若选择保持当前相位,即选择动作0,当前绿灯时长增加1 s。

(3)当前相位持续时间大于最大绿灯时长时,强制切换相位,智能体必须选择动作1,强制执行3 s黄灯,然后切换到下一相位。

2.3 奖励函数

为了减少交叉口的排队溢出情况,降低最大排队长度,设计了关于二次排队和最大排队长度的奖惩函数,作为奖励函数来指导智能体的学习。

2.3.1 二次排队奖惩函数 r_1

在最小绿灯时间结束时,根据进口车道车辆速度和位置信息,判断当前相位通行方向的排队车辆是否已经完全消散,再选择动作。如果排队已经完全消散,当选择动作0(保持当前相位不变,并为当前相位增加时长1 s),给予惩罚值-10;当选择动作0(立即切换下一相位),给予奖励值10。如果排队未完全消散,当选择动作0,给予奖励值未消散车辆排队长度;当选择动作0,给予惩罚值负未消散车辆排队长度。表1为二次排队奖惩函数 r_1 取值表。表1中: n_d 为当前相位通行方向已经产生过排队延误的车辆数; l_{car} 为排队车辆的车头间距。

表1 二次排队奖惩函数 r_1 取值表

排队情况	动作(action)	
	0	1
完全消散	-10	10
未完全消散	$l_{car} \cdot n_d$	$-l_{car} \cdot n_d$

2.3.2 最大排队长度奖励函数 r_2

如果只考虑交叉口二次排队情况,为了避免当前相位通行车辆的二次排队,智能体会不断选择保持当前相位,直到到达最大绿灯时长。这可能造成绿灯时间浪费,导致其他相位排队长度过大,降低交叉口的通行效率。因此将交叉口最大排队长度 q_{max} 纳入奖励函数的构建,以确保不同方向的车流通行权。最大排队长度奖励函数 r_2 的表达式为:

$$r_2 = \begin{cases} -q_{max}, & \text{action} = 0 \\ q_{max}, & \text{action} = 1 \end{cases} \quad (8)$$

结合二次排队奖惩函数 r_1 、最大排队长度奖励函数 r_2 ,得到本文的最终奖励函数 R 为:

$$R = r_1 + r_2 \quad (9)$$

3 案例分析

3.1 交叉口现状

本文选取上海市控江路与图们路交叉口进行数字孪生建模。该路口东侧由于铁路的影响,常常出现拥堵到上游军工路的情况,固定的信号配时难以适应随时变化的交通环境。该路口为十字形交叉口,4个进口道均为双向2车道,分别为左转与直右车道,每条车道宽度为3.25 m,路口范围内不存在展

宽段,路口各进口道交叉口车道渠化情况见图1。

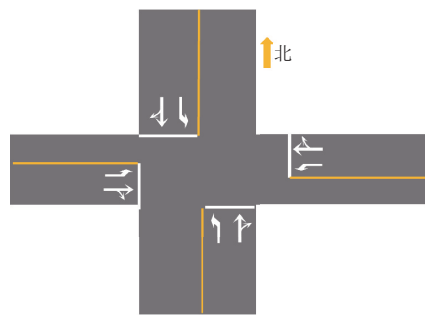


图1 交叉口车道渠化图

路口为4相位,直左分离放行,路口车道与相位关联关系见表2。

表2 路口车道与相位关联关系

进口道	车道	通行相位	当前相位时长/s
南	左转	B	18
	直右	A	28
北	左转	B	18
	直右	A	28
东	左转	D	20
	直右	C	24
西	左转	D	20
	直右	C	24

基于SUMO构建交叉口仿真模型,模拟交叉口交通运行状况,作为DQN算法的训练环境。选取某个周三(工作日)对路口进行交通数据调查,从卡口过车记录获取路口分车辆类型、分流向流量(见表3)。以此流量作为训练的中流量样本,将样本的0.7倍和1.3倍作为训练的小流量样本和大流量样本,对模型进行训练,获取最优信控方案。

表3 路口交通组成及流向流量

进口道	车辆类型比(小:中:大)	转向比(直:左:右)
南	90:3:7	285:244:71
北	95:3:2	256:212:64
东	91:4:5	241:269:60
西	91:4:5	277:230:69

按照车身长度将车辆类型分为小型车辆、中型车辆和大型车辆,车长小于6 m为小型,6~12 m为中型,大于12 m为大型车辆,路口大型车辆多为公交车。表中转向比同时表示路口进口道各流向流量。

该方法的实施对交叉口的交通检测设备要求较高,要求路口设备实时传输交通状况。通过路口进口道卡口分车道检测实时通过车辆信息,如车辆类型等;基于路口SCATS信号机实时传输路口信号状况,按秒传输当前相位、相位剩余时间等信息;车辆GPS系统实时传输车辆位置、速度等信息;另外结合路口进口道线圈实时信息,作为卡口数据的补充信息。

3.2 结果分析

将由韦伯斯特(Webster)方法求得的信号配时方案在该交叉口实施,与本文使用的DQN算法相对比,验证该方法的有效性。信号方案中每个相位转换之前执行3 s黄灯,信号周期保持为90 s。每次实验的前3 min不进行训练和数据提取,以避免仿真初始波动的影响,在此后6 min的车流运行中对比2种方法得到的信号控制方案对交叉口整体交通运行情况的控制效果。

在DQN算法实验中,智能体在数字孪生模型环境中共进行350次迭代实验,每次实验持续540个步长,其中前180个步长作为稳定交通运行的时间,在后续4个信号周期进行DQN算法学习,每个步长为1 s。

训练的累计回报值变化曲线见图2。图中横轴代表智能体与环境交互所经历的训练次数,纵轴是智能体在每次迭代中获得的累计奖励值。

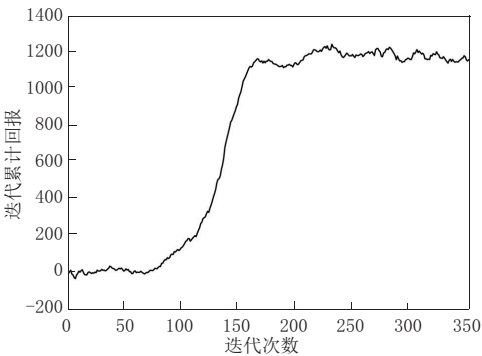


图2 DQN算法回报收敛图

由图2可知,最初智能体的累计回报随着训练次数增加而呈现出显著的上升趋势,即智能体在学习过程中逐步改善了其决策策略。在大约前100次训练中,智能体正在探索环境并尝试理解如何获得奖励,因此回报增长较为缓慢。当训练次数超过100之后,回报的增长速度加快,此时智能体开始利用其学习到的策略来获得更好的奖励。在约200次训练后,回报达到一个较高的水平并趋于稳定,表明智能体找到了一个有效的策略,并且在多次迭代中一直保持着较高的性能。

图3是2种信号配时方案下交叉口的累计最大排队长度对比。可以看到,DQN算法下一开始的累计最大排队长度明显较大,但随着算法的学习,DQN的性能有所提升,并在某一点后稳定在一个显著低于韦伯斯特法的排队长度。这表明DQN算法在经过足够的训练之后能够找到比传统韦伯斯特方法更优的信号控制策略,以降低交叉口的排队长度。

图4是2种方法得到的信号控制方案下交叉口

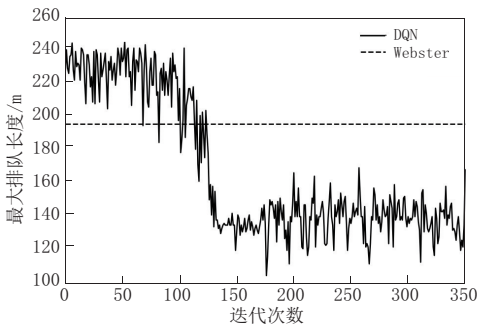


图 3 最大排队车辆数变化图

的累计二次排队车辆数对比。可以看到,随着迭代次数的增加,DQN 算法下的二次排队性能得到改善,即 DQN 智能体在逐步学习如何更有效地选择信号配时方案,以减少交叉口二次排队情况。在累计二次排队车辆数小于韦伯斯特法后维持在一个较低值范围内,说明 DQN 方法训练得到的信号控制方案可以明显改善交叉口二次排队情况。

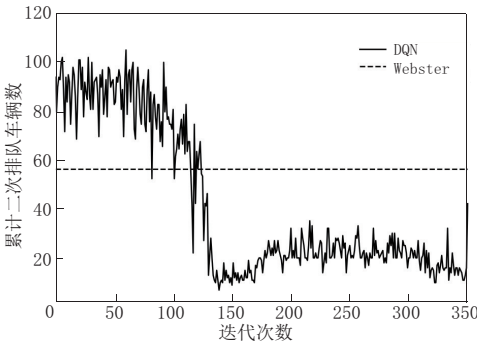


图 4 二次排队车辆数对比

图 5 比较了 DQN 算法与韦伯斯特法在一段时间内的交叉口信号方案。横坐标为仿真时间 180~540 s 内的信号控制方案,表示了不同信号相位的持续时长。

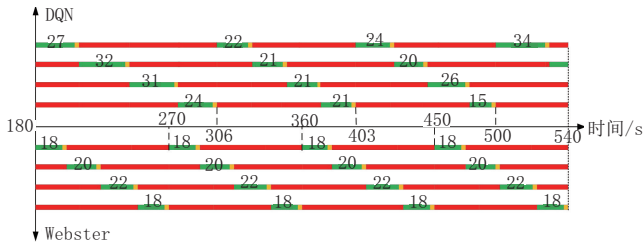


图 5 信号方案对比

由图 5 可知,DQN 算法根据实时交通状态数据和历史信息来动态调整信号相位的时长,提供了一个更加定制化的信号控制方案,而韦伯斯特法则提供了一个固定的基于平均交通流量计算得出的标准方案。

表 4 给出了 2 种方法下的交叉口评价指标,分别对比了 2 种方法下的交叉口最大排队长度、二次排队车辆数、车均延误、平均速度、平均行程时间、最大通过车辆数指标。由表 4 可见,DQN 方法下,车辆的平均等待时间更短,平均运行速度更大,车辆通过交

叉口所需的平均时间较短,交叉口每小时可以通行的车辆数量更多。DQN 方法在列出的 6 个评价指标上都优于韦伯斯特法,表明 DQN 在交通信号控制的应用中可能提供更高效的性能。

表 4 交叉口运行评价指标对比

评价指标	DQN	韦伯斯特 (Webster)	优化效果/%
最大排队长度/m	167.7	195	14.02
二次排队车辆数/辆	43	57	24.70
车均延误/s	66	79	16.46
平均速度/(km·h ⁻¹)	5.76	5.46	5.49
平均行程时间/s	76	89	14.61
最大通过车辆数/(辆·h ⁻¹)	1 620	1 440	12.50

图 6 对比了不同交通需求下的交叉口通行评价指标。以调查流量(路口流量为 2 278 辆/h)为路口中流量,分别取该流量的 0.7 倍和 1.3 倍作为路口的低流量和高流量进行仿真,对比交叉口运行指标。图中单位 C 表示信号周期。

选取了 8 个交通运行评价指标,包括车辆在交叉口范围内的平均运行时间、平均速度、通行流量(某一周期)、平均停车次数、平均延误、最大排队长度、二次排队车辆数(某一周期)及绿灯空放时长,来比较 DQN 算法和韦伯斯特方法下的信号方案对交叉口的控制效果。

由图 6 可知,与韦伯斯特法相比,DQN 算法下的车辆平均运行时间、平均停车次数、平均延误及绿灯空放时长在不同的交通需求下都较低。同时,DQN 算法下的平均速度、通行总流量都大于韦伯斯特法下的值。这些结果都证明 DQN 算法更能适应交叉口的交通变化,能给出更适合交通环境的信号控制方案,提高交叉口的运行效率。

针对不同流量下学习过程所用算力及耗时进行适用性分析,结果见表 5。另外,本方法的实施需要在交叉口配备一些智能交通设备,以获取实时的交通数据和信号状态信息,对其进行成本分析,结果见表 6。

4 结 语

针对交叉口的排队溢出问题,本文采用深度强化学习(DQN)算法,对交叉口信号配时方案进行了优化。实验结果表明,这种方法能显著降低交叉口的二次排队车辆数和最大排队长度,有效提高交通效率并减少拥堵。然而,尽管本文的研究取得了积极成果,但仍存在一些限制。例如,仿真环境基于仿真模型搭建,并不能完全模拟现实复杂多变的交通

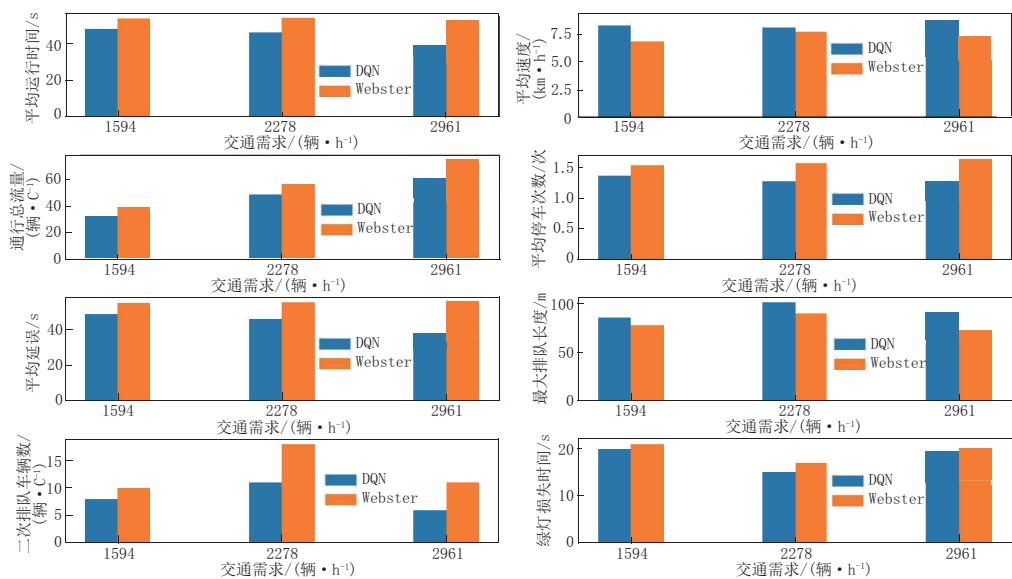


图6 不同交通需求下的评价指标

表5 适用性分析结果

处理器	机待RAM	流量	耗时/h
13th Gen Intel(R) Core(TM) i7-1360P 2.20 GHz	16 GB	低	7.3
	16 GB	中	8.5
	16 GB	高	9.2

表6 成本分析

设备	数量/台	成本/元
卡口设备	4	2 000×4
网络设备和服务器	1	4 000

环境,仍需在更多真实场景中验证算法的效果和适用性;此外,考虑到城市交通系统的复杂性,未来研究可以探索如何将该算法与其他交通管理措施相结合,以进一步优化交通效率和提升道路安全性。

参考文献:

[1] 童蔚苹,袁诗琳,刘菲菲,等.基于通行能力的借对向车道左转交叉口信号控制模型优化[J].公路交通科技,2021,38(3):87-95.

[2] 林丽,蒋帅捷.基于Ring-Barrier相位的交叉口信号配时优化模型[J].重庆交通大学学报(自然科学版),2021,40(1):12-16.

[3] 吴场建,任刚,曹奇,等.基于预信号的快速路出口匝道衔接交叉口控制优化方法[J].中国公路学报,2023,36(10):269-280.

[4] 魏吉敏.基于网联车辆信息的路口信号配时优化方法[J].交通工程,2022,22(2):7-11.

[5] YAN J, HAN X L. Improved genetic particle swarm algorithm based single intersection signal timing optimization[J]. Proceedings of SPIE, 2023(7): 127080.

[6] 张鹏,顾云翔,孙超,等.基于最小油耗的货车车速与信号配时协同优化模型[J].交通运输系统工程与信息,2023,23(2):84-91.

[7] KANG L L, HUANG H, LU W K, et al. A dueling deep q-network method for low-carbon traffic signal control[J]. Applied Soft Computing, 2023, 141: 16.

[8] ZHANG W B, YAN C, LI X F, et al. Distributed signal control of arterial corridors using multi-agent deep reinforcement learning[J]. Ieee Transactions on Intelligent Transportation Systems, 2023, 24(1):178-190.

[9] YU Z Y, XU G H, GAYAH V V, et al. Incorporating phase rotation into

a person-based signal timing optimization algorithm[J]. IEEE Transactions on Intelligent Transportation Systems,2022,23(1):513-521.

[10] HAN Y H, LEE H Y S, KIM Y. Extensible prototype learning for real-time traffic signal control[J]. Computer-Aided Civil and Infrastructure Engineering, 2023, 38(9): 1181-1198.

[11] 张旭,杨晓光,戴佩家,等.基于老年人行为特性的行人过街信号优化[J].重庆交通大学学报(自然科学版),2022,41(7):39-45.

[12] WANG Z W, YANG K K, LI L, et al. Traffic signal priority control based on shared experience multi-agent deep reinforcement learning [J]. Iet Intelligent Transport Systems, 2023, 17(7): 1363-1379.

[13] 李岩,南斯睿,马静,等.降雨天气单点交叉口交通信号控制优化方法[J].交通运输工程学报,2018,18(5):185-194.

[14] SONG L, FAN W. Performance of state-shared multiagent deep reinforcement learning controlled signal corridor with platooning-based CAVs[J]. Journal of Transportation Engineering Part a-Systems , 2023, 149(8): 10.

[15] LI Y K, NIU W J, TIAN Y Z, et al. Multiagent reinforcement learning-based signal planning for resisting congestion attack in green transportation[J]. IEEE Transactions on Green Communications and Networking, 2022, 6(3): 1448-1458.

[16] 马东方,陈曦,吴晓东,等.基于强化学习的干线信号混合协同优化方法[J].交通运输系统工程与信息,2022,22(2):145-153.

[17] LEE H, HAN Y, KIM Y, et al. Effects analysis of reward functions on reinforcement learning for traffic signal control[J]. Plos One, 2022, 17(11): 18.

[18] 陈元洁.基于DQN算法的单点交叉口信号控制[J].交通与运输, 2023,36(增刊1):89-94.

[19] 张尊栋,刘雨珂,刘小明.深度强化学习在交通信号控制中的应用[J].自动化博览,2022,39(12):30-37.

[20] 宋太龙,贺玉龙,刘钦.基于深度强化学习的大型活动关键交叉口信号控制[J].科学技术与工程,2023,23(22):9694-9701.

[21] QIN H, ZHANG W S. Short-term traffic flow prediction and signal timing optimization based on deep learning[J]. Wireless Communications & Mobile Computing, 2022: 11.

[22] GUO J, HUANG X, LI R, et al. Based on MOPSO algorithm of real-time traffic signal optimization control for intelligent transportation intersections[J]. Journal of Physics: Conference Series, 2023(9): 012085.