DOI：10.13890/j.issn.1000-128X.2026.06.511

基于动态客流的城轨列车目标速度曲线优化

向超群[1]，杜京润[1]，杨　洋[2]，张清奇[2]，朱亨国[3]，范子寅[4]
（1.中南大学 交通运输工程学院，湖南 长沙 410075；
2.长沙市轨道交通集团有限公司，湖南 长沙 410027；
3.湖南中车时代通信信号有限公司，湖南 长沙 410100；
4.通号城市轨道交通技术有限公司，北京 100070）

摘　要：随着城市轨道交通系统能耗问题的日益凸显，一方面，既有目标速度曲线优化方案多基于静态客流量情景，缺乏对动态客流情景下的优化算法的探索；另一方面，多数节能研究未考虑牵引效率与牵引力的耦合关系，没有从提高牵引效率的角度来思考降低列车牵引损耗。针对上述问题，文章构建了融合动态客流与牵引效率优化的城轨列车节能控制框架，形成了从理论建模到算法设计的完整技术体系。提出了基于深度Q网络（deep Q-network, DQN）的目标速度曲线优化算法，解决了现有算法在动态客流场景下的适应性不足与缺乏对牵引效率因素考量的问题。首先，去除“巡航”驾驶动作，构建包含“高效率牵引”驾驶动作的新候选动作集；其次，通过多子区间密集划分扩大优化算法的寻优范围；最后，引入末端制动比例可调和理论最小剩余运行时长判定机制，降低寻优难度。算法将不同载荷状态下的车重信息设置为重要输入量，使寻优过程与客流变化相关联，并通过“高效率牵引”驾驶动作和DQN奖励环节的能耗计算将牵引效率引入优化环节，实现了对客流变化和牵引效率影响的综合考量。所提策略采用强化学习的框架，考虑了客流变化、牵引效率、线路条件等因素，通过多场景仿真与实车数据对比的方式，完成训练并验证了所提方法的节能效果。

关键词：城轨列车节能；城轨列车牵引效率；动态客流；目标速度曲线优化；深度Q网络；城市轨道交通

中图分类号：U231.6；U268.6　　**文献标志码：**A　　**文章编号：**1000-128X(2026)04-0014-10

0　引言

城市轨道交通具有运行速度快、运量大、准时安全等特点，是城市公共交通的重要形式之一[1-3]。官方统计[4]显示，列车牵引能耗约占运营总能耗的50%，具有很大的节能空间。通过优化列车运行时的目标速度曲线和牵引控制方式，可以提升对线路势能的利用率，提高牵引系统的工作效率，减小制动造成的动能损失，进而降低列车的牵引能耗，实现牵引节能。

根据最优解求解方法的不同，目标速度曲线优化算法大体上可以划分为3类：解析算法、数值算法和智能算法。解析算法通过为问题建立相应的数学关系式，并利用极值原理等解析手段计算最优解。1968年，文献[5]奠定了该领域的研究基础，开创性地将极值原理应用于列车节能运行领域，构建了基于车辆动力学方程的能耗模型，定义了包含最大牵引、匀速巡航、惰行和最大制动的四工况控制集。文献[6-10]系统研究了多约束条件下的控制策略，在长直轨道、连续变速约束、时变坡度及特殊工况场景中验证了庞特里亚金极大值原理（pontryagin's maximum principle, PMP）的适用性，并建立了离散/连续控制模式的等效映射关系。文献[11-12]引入协同状态变量建立工况切换的解析判定准则，为再生制动触发时机提供了存在唯一性证明。

基金项目：湖南省重点研发计划资助项目(2025JK2065)
通信作者：范子寅，男，硕士，研究员，主要从事列车速度曲线优化方面的研究；E-mail: 578362512@qq.com
引用格式：向超群, 杜京润, 杨洋, 等. 基于动态客流的城轨列车目标速度曲线优化[J]. 机车电传动, 2026(4): 14-23.DOI:10.13890/j.issn.1000-128X.2026.06.511.
Citation：XIANG Chaoqun, DU Jingrun, YANG Yang, et al. Optimization of target speed profiles for urban rail trains considering dynamic passenger flow[J]. Electric drive for locomotives,2026(4): 14-23.DOI:10.13890/j.issn.1000-128X.2026.06.511.

解析算法通常能高效地求解出列车驾驶策略的理论最优解，但需要对非线性问题做线性化处理，仅能得到理想条件下的最优。文献[13]提出基于连接函数的工况序列简化算法，量化加速-惰行-制动过程的衔接误差以降低计算复杂度。

数值算法在构建描述列车运行过程的动力学模型的基础上，结合速度、加速度等约束条件划定可行解域，并通过模拟自然进化或群体智能机制的启发式搜索策略在解空间内寻找最优解。文献[14]开发了一种基于随机增强遗传算法（genetic algorithm, GA）的最优驾驶策略求解方法，通过随机扰动机制增强全局搜索能力，有效避免陷入局部最优解。文献[15]采用收缩时域参数化预测控制（shrinking horizon parametrized predictive control, SPPC）方法进行了节能优化，提出了两阶段松弛策略，增强了算法对时变参数的适应性，拓展了算法的应用边界。文献[16]将列车运行轨迹离散为多个区段，利用时间逼近搜索方法进行求解，提高了算法的寻优效率。

随着运算设备算力的提升，智能算法成为了新的研究趋势。文献[17]将深度确定性策略梯度（deep deterministic policy gradient, DDPG）算法引入目标速度曲线优化，采用连续动作空间代替传统的离散动作集，极大地扩展了优化算法的解集，增强了寻优效果，但求解过程比较复杂。后续研究[18]进一步将专家知识和DDPG算法相融合，利用专家知识构建了节能驾驶特征推理方法，引导DDPG训练前期进化方向，约束了智能体的探索空间。为提高计算效率，文献[19]在DDPG框架中预先分配了各区段运行时间和运行模式，为DDPG训练前期提供了高质量的范本，有效缓解了连续集决策的计算压力。文献[20]提出一种两阶段优化框架：在特征学习阶段运用双向长短期记忆网络（bidirectional long-short-term memory, BLSTM）捕捉列车运行数据的时序特征，随后在策略优化阶段采用双重深度Q网络（double deep Q-network, DDQN）学习最优策略。列车运行过程中需面临多种受力叠加、多类因素耦合的复杂场景，智能算法无须依赖精确的列车模型或人为规则设定，自适应学习能力强大，但是优化过程的计算量庞大，存在训练数据依赖、可解释性不足等局限。

针对现有城市轨道交通列车牵引节能研究主要从多车调度和单车控制两个维度展开，普遍存在影响因素考虑不全等问题，本文提出一种基于DQN的目标速度曲线优化策略，采用强化学习的框架，考虑了客流变化、牵引效率、线路条件等因素，以牵引能耗更低、运行时长偏差更小和到站位置偏差更小为优化目标，在满足限速、最大启动加速度、最大制动减速度等运行约束的前提下，求解最优目标速度曲线。依托线路实车运行数据开展对比仿真验证，结果表明本文所提优化方法具备可行性与有效性。

1　目标速度曲线优化解空间

1.1　运行区间离散化处理

研究团队对站间路程及驾驶过程进行离散化处理。如图1所示，设相邻两个站之间的路程为L，将这段路程平均划分为n个子区间，每个子区间长L/n。在每个子区间的起始处，列车选择驾驶动作，输出牵引力，并保持这一状态，直至子区间终止处。考虑寻优效果、寻优难度以及驾驶平顺性，取n=100。

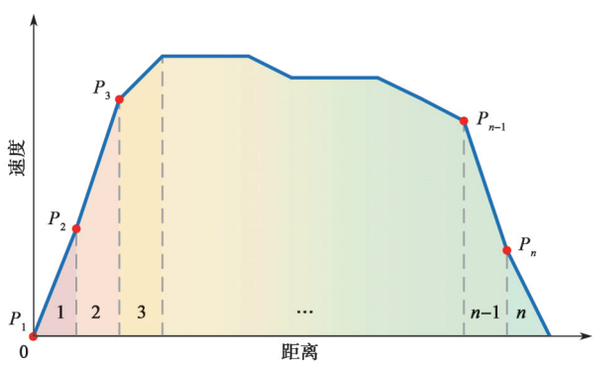


图1　站间运行离散优化示意图

Fig. 1　Schematic diagram of discrete optimization for inter-station operation

1.2　基于列车车重的候选动作集构建

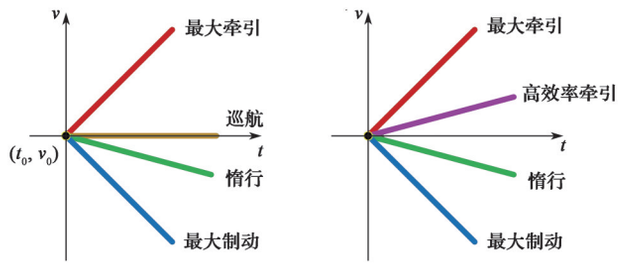
巡航工况通常出现在列车的中高速运行阶段，在此阶段，列车牵引效率主要受牵引力的影响。通过巡航工况来维持列车速度将损失较大能量，相反，在巡航工况的基础上略微增大牵引力输出值，既可保持速度稳定或小幅变化，也更利于节能。基于此，本文提出以“高效率牵引”驾驶动作替代“巡航”驾驶动作。如图2所示，为便于速度调节和限制列车速度的变化幅度，让“高效率牵引”落在“最大牵引”和t轴（对应“巡航”）之间这个区域，列车加速度不可超过0.2 m/s^2。图3和表1展示了在2种不同运行工况下执行“高效率牵引”驾驶动作的案例。另外，根据图3，当效率低于85%时，通过增大牵引力从而提升效率具有显著边际效益，而当效率高于85%后，增加较大范围的牵引力才能实现效率提升，因此，为使列车在较小的加速度下获得较高的牵引效率，列车牵引效率应尽可能接近85%。综上，“高效率牵引”牵引力公式为

$$F_{\mathrm{t}}=\underset{F_{\mathrm{t}}\leqslant F_{\mathrm{d}}+0.2M}{\arg\min}\left|\eta(v,F_{\mathrm{t}})-0.85\right| \tag{1}$$

式中：F_{t}表示牵引力；M表示列车当前车重；$\eta(v,F_{\mathrm{t}})$表示列车速度为v、牵引力为F_{t}时的牵引效率；F_{d}表示列车运行阻力。

由式(1)可知，“高效率牵引”时输出的牵引力与列车总重量以及牵引效率有关，研究团队同时考虑了列

车载客量和牵引效率，可以使列车在客流变化时仍能保持高效率的低功率牵引。



(a) 传统候选集　　(b) 所提候选集

图2　候选动作集

Fig. 2　Candidate action set

在图3工况1中，列车总质量为211 t，受到8.79 kN的运行阻力。若以0.2 m/s²的加速度运行，需输出50.99 kN的牵引力，对应的牵引效率为81.8%（如图3中p_1点所示）；若以85%的牵引效率运行，需输出70.96 kN的牵引力，对应的加速度为0.29 m/s²（如图3中p_2点所示）。为限制加速度不超过0.2 m/s²，列车最终输出的牵引力为50.99 kN。

在工况2中，列车重量为322.6 kN，受到12.55 kN的运行阻力。若以0.2 m/s²的加速度运行，需输出77.07 kN的牵引力，对应的牵引效率为88.2%（如图3中p_3点所示）；若以85%的牵引效率运行，需输出49.38 kN的牵引力，对应的加速度为0.11 m/s²（如图3中p_4点所示）。为使加速度尽可能接近巡航状态的特征，列车最终输出的牵引力为49.38 kN。

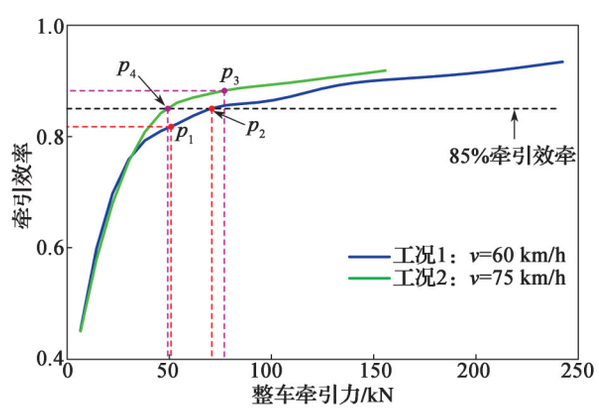


图3　不同运行工况下的“高效率牵引”驾驶动作案例分析

Fig. 3　Case analysis of "high-efficiency traction" driving action under different operating conditions

表1　不同运行工况下的“高效率牵引”驾驶动作案例分析

Table 1　Case analysis of "high-efficiency traction" driving action under different operating conditions

工况编号	载荷状态	速度/(km·h⁻¹)	有无坡/弯	牵引力/kN	加速度/(m·s⁻²)	牵引效率
1	AW0	60	无/无	50.99	0.20	81.8%
2	AW2	75	无/无	49.38	0.11	85.0%

上述工况分别反映了列车轻载和重载时的情形，当客流变化时，用“高效率牵引”代替“巡航”总能提高列车中高速运行时的牵引效率，并降低牵引能耗。另外，相比图2(a)的传统候选集，图2(b)中所提候选集同样包含2个大幅增减车速和2个小幅调节车速的驾驶动作，可满足目标速度曲线的寻优需求。

1.3　优化过程边界条件构建

根据GB 50157—2013《地铁设计规范》，规定列车在第1子区间选择“牵引”驾驶动作，牵引力恒定为0.9M N；规定“最大制动”驾驶动作制动力恒定为0.9M N；若“高效率牵引”的牵引力超过了上限，则令其等于上限。列车不可超速行驶、不可在未到站时停车或倒行。规定当目标速度曲线终端和v=0轴的交点与目标终点的距离小于等于0.3 m时，目标速度曲线有效；为保证列车正点运行，同时预留充足的停站冗余时间，规定相邻两站之间的运行时长偏差不超过2%；为了降低解空间复杂度、实现精准停站和准时运行，列车在最后4个子区间（即第97~100子区间）选择“制动”，且可输出任意的制动力。

候选动作集A限制为式(2)。

$$A=\begin{cases}\{\text{"牵引"}\},\ \text{seg}=1\\ \{\text{"最大牵引"},\text{"高效率牵引"},\text{"惰行"},\text{"最大制动"}\},\\ \qquad\qquad 2\leqslant\text{seg}\leqslant 96\\ \{\text{"制动"}\},\ 97\leqslant\text{seg}\leqslant 100\end{cases}\tag{2}$$

式中：seg表示子区间编号。

对目标速度曲线的优化过程施加如下限制条件：

$$\begin{cases}0\leqslant v(x)\leqslant v_{\text{limit}}(x)\\ t_{\text{goal}}\times 98\%\leqslant t_{\text{run}}\leqslant t_{\text{goal}}\times 102\%\\ x_{\text{goal}}-0.3\leqslant x_{\text{stop}}\leqslant x_{\text{goal}}+0.3\\ \left|r_{\text{break}}(\text{seg}_1)|_{\text{seg}_1\in[97,100]}-r_{\text{break}}(\text{seg}_2)|_{\text{seg}_2\in[97,100]}\right|\leqslant 0.25\\ t_{\text{last_min}}(x)\leqslant t_{\text{goal}}-t(x)\\ v(x)\big|_{x=96L/100}\leqslant v_{\max}\end{cases}\tag{3}$$

式中：$v_{\text{limit}}(x)$为距离上一站x处的线路限速；t_{goal}为列车在站间的目标运行时长；t_{run}为列车运行时长；x_{goal}为列车目标停车位置；x_{stop}为列车停车位置；在最后4个子区间中，规定任意2个相邻子区间的制动比例相差不超过25%；r_{break}为4个子区间的制动比例，指输出制动力占最大制动力（限定为0.9M）的百分比；$t(x)$为列车在距离上一站x处已运行的时长；$t_{\text{last_min}}$为列车最小剩余运行时长。

列车理论最小剩余运行时长曲线如图4，当列车运行至k点位置，若列车立刻以“最大牵引”加速至最大ATO速度，并保持该速度，最后以最大制动减速并停止于终点位置，则列车在此过程中的运行时长为理论最小时长，记作$t_{\text{last_min}}(x_{\text{p}})$。

$v(x)\big|_{x=96L/100}$为列车在第97子区间起始处的最大速

度，若列车在第 97 子区间起始处开始持续“最大制动”，但仍会越过停车点才能完成停车，则列车已失去精准停站的可能性。列车在第97子区间起始处的最大速度如图5所示，如果列车在第97子区间起始处的速度小于蓝色实线标示的速度，则列车可以在终点前停车。如果彼时列车速度与蓝色实线重合，则列车通过持续地“最大制动”，可以刚好在终点处停车。现设定一个偏差裕量 x_r，若列车在终点后的 x_r 处可以停车，则都视为可接受。蓝色虚线表示停车精度处于可接受范围的速度上限，其在第97子区间起始处的纵坐标即为 v_{max}。

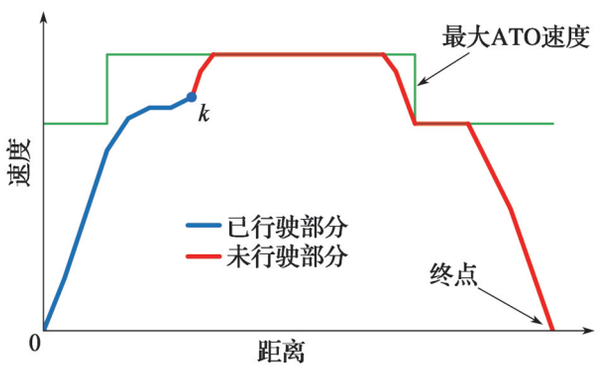


图4　列车理论最小剩余运行时长示意图

Fig. 4　Schematic diagram of theoretical minimum remaining running time

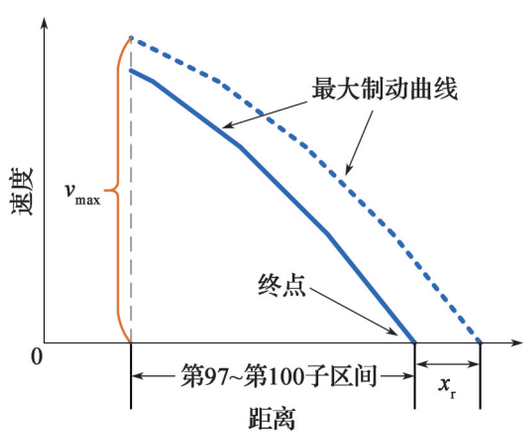


图5　列车在第97子区间起始处的最大速度示意图

Fig. 5　Schematic diagram of the maximum train speed at the beginning of the 97th sub-interval

2　目标速度曲线寻优算法

基于深度Q网络（DQN），建立目标速度曲线优化算法，考虑动态客流场景下的适应性问题与对牵引效率考量的问题。

DQN的框架[21-22]如图6所示，为了用具有参数 θ 的模型近似替代Q学习的价值函数 $Q(s,a;\theta)$，DQN引入了深度卷积神经网络。在每次时间步长 t 进行经验回放时，智能体的经验 $e_t=(s_t,a_t,r_t,s_t+1)$ 被存储到数据集 $D_t=\{e_1,\cdots,e_t\}$ 中。

研究团队采用迭代寻优的方式优化目标速度曲线，

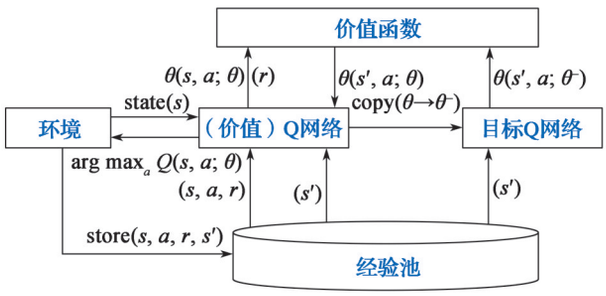


图6　深度Q网络算法框架[21-22]

Fig. 6　Deep Q-network algorithm framework

具体的迭代寻优规则如下：

① 从第1子区间开始，按顺序逐一决策每个子区间的驾驶动作。若其间违反了边界条件[式(3)]，则立刻中断本轮寻优。

② 在第1~第96子区间的每个子区间起始处，进行当前子区间的驾驶动作决策，并计算子区间结束时的列车速度、累计运行时长和累计牵引能耗。

③ 在第97子区间的起始处，以最后4个子区间内每个子区间应采取的制动比例为寻优对象，采用群优化算法进行求解。

求解最优制动比例时，适应度函数设置为

$$\text{fitness}=\lambda_1 f_{\text{time}}+\lambda_2 f_{\text{stop}} \tag{4}$$

式中：f_{time} 为“准时运行”适应度；f_{stop} 为“精准停站”适应度；λ_1、λ_2 分别为两部分适应度对应的权重系数，其取值是基于量纲平衡、工程优先级两个主要方面综合考量的结果。

f_{time} 用以考量列车实际运行时长与目标时长的偏差程度，表示为

$$f_{\text{time}}=\begin{cases}0, & \left|t_{\text{cast}}-t_{\text{last}}\right|>t_{\text{limit}}\\ \left(t_{\text{limit}}-\left|t_{\text{cast}}-t_{\text{last}}\right|\right)^2, & \left|t_{\text{cast}}-t_{\text{last}}\right|\leqslant t_{\text{limit}}\end{cases} \tag{5}$$

式中：t_{cast} 为列车在最后4个子区间的运行时长；t_{last} 为列车在第97子区间起始处剩余的运行时长；t_{limit} 为算法迭代时允许的运行时长偏差上限。

f_{stop} 用以考量列车实际停车位置与目标位置的偏差程度，写作

$$f_{\text{stop}}=\begin{cases}0, & \left|x_{\text{stop}}-x_{\text{goal}}\right|>x_{\text{limit}}\\ x_{\text{limit}}-\left|x_{\text{stop}}-x_{\text{goal}}\right|, & \left|x_{\text{stop}}-x_{\text{goal}}\right|\leqslant x_{\text{limit}}\end{cases} \tag{6}$$

式中：x_{limit} 为算法迭代时允许的停车位置偏差上限。

列车在最后4个子区间的制动范围如图7。若按照前96个子区间的寻优方式，列车在最后4个子区间只能选择“最大制动”驾驶动作，故当列车到达第97子区间的起始处时，其速度只有处于橙色区域内时，才有可能实现准时运行和精准停站。若采用所提寻优方式，列车的制动范围将扩大至青绿色区域内，这样不仅降低本轮寻优的解空间复杂度和求解难度，而且可以扩大前96个子区间的寻优范围，实现准时运行和精准停站。

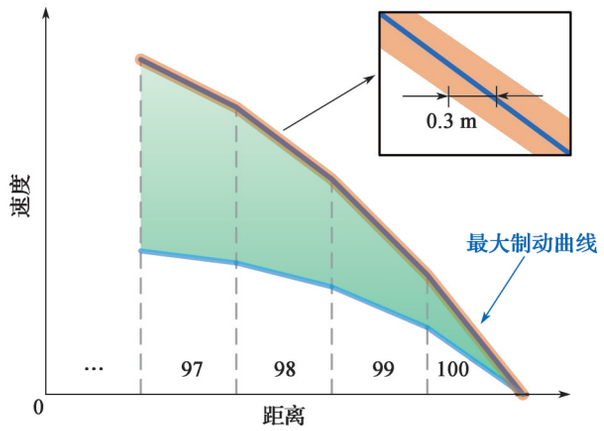


图 7　列车在最后 4 个子区间的制动范围

Fig. 7　Train braking range in the last four sub-intervals

在强化学习中，奖励是智能体从环境中获得的反馈信号，用于评估其行为的优劣。奖励引导智能体学习，通过不断尝试和调整策略，智能体旨在最大化长期累积奖励。合理的奖励设计能够激励智能体探索有效策略，加速学习过程，显著提升智能体的训练效率和策略优化能力。在目标速度曲线优化策略中，为获取符合列车运行规范的目标速度曲线，实现节省牵引能耗、准时运行和精准停车的目标，设计如下奖励获取机制：

① 每完成一个子区间的驾驶动作决策后，计算该子区间终止处的列车状态，并根据列车状态立即进行奖励结算；

② 按照“稀疏奖励”的原则，仅在每轮寻优结束后设置奖励点，其余时候奖励默认为 0；

③ 采用“轻奖重罚”的思想，对违反约束条件的情况给予绝对值较大的负奖励，对满足或接近满足应用需求的结果给予绝对值较小的正奖励；

④ 根据寻优结果的好坏，奖励值按一定的梯度进行设计，以引导智能体逐步向最大化长期累积奖励的方向收敛。

基于上述机制，具体的奖励函数设计如下：在当前子区间终止处，若本轮寻优既没有因违反约束条件而中断，也没有随着到达终点而结束，则当前子区间的奖励为 0，即 r=0；在第 n 子区间，若列车速度小于等于 0，则立即结束本轮寻优，获取奖励 r=−50；在第 n 子区间，在列车速度大于 0 的前提下，若列车速度大于线路限速，或者列车失去准时到站的可能性，则立即结束本轮寻优，获取奖励 r=−50+0.1n；在第 97 子区间起始处，若列车速度小于等于线路限速，但是大于精准停站的速度上限 $v_{\max}$，则立即结束本轮寻优，获取奖励 $r=-30-10(v(x)|_{x=96L/100}-v_{\max})/(v_{\text{limit}}(x)|_{x=96L/100}-v_{\max})$；完成最后 4 个子区间的制动比例寻优后，若寻优结果的“准时运行”适应度和“精准停站”适应度没有同时大于 0，则获取奖励 r=−10；完成全部子区间的驾驶动作决策后，若得到的目标速度曲线同时满足“准时运行”和“精准停站”要求，则获取奖励 $r=10+20(E_s-E)/E_s$，其中，E 为列车在站间运行的总牵引能耗，E_s 为列车节能奖励的计算标准值。除了上述提到的情形外，其他情况下奖励均按 0 进行结算。

3　最优目标速度曲线求解过程

如表 2 所示，研究团队求解最优目标速度曲线时，首先在 DQN 自学习架构中，将目标速度曲线优化策略

表 2　基于 DQN 的目标速度曲线优化求解过程伪代码

Table 2　Pseudocode for the DQN-based target speed profile optimization solution process

算法：基于 DQN 的目标速度曲线优化求解
输入：
列车总质量 M（反映当前客流量）
列车站间目标运行时长 t_{goal}
线路条件及限速信息
初始化：
初始化记录子区间驾驶动作决策历史的经验回放池 D
初始化评价驾驶动作优劣的价值函数 Q，其权重 θ 为随机值
初始化目标动作价值函数 $\hat{Q}$，其权重 θ^- 设置为与 θ^- 相同
执行：
1：**for** episode=1 to N **do**
2：初始化序列 $s_1=\{x_1\}$ 和预处理序列 $\phi_1=\phi(s_1)$；
3：**for** t=1 to T **do**
4：找出 Q 值最大的驾驶动作 $a_{\text{best}}=\arg\max_a Q(\phi(s_t),a;\theta)$；
5：以概率 ε 或者 $1-\varepsilon$ 选择一个随机驾驶动作 a_t/Q 值最大的驾驶动作 $a_t=a_{\text{best}}$；
6：执行驾驶动作 a_t，并计算奖励 r_t 和下一个子区间的列车状态 x_{t+1}；
7：令 $s_{t+1}=s_t,a_t,x_{t+1}$，并预处理 $\phi_{t+1}=\phi(s_{t+1})$；
8：将驾驶动作决策的经验 $(\phi_t,a_t,r_t,\phi_{t+1})$ 存储到回放记忆 D 中；
9：**if** 时间步长达到迭代要求 **do**
10：从 D 中采样一组随机的小批量经验 $(\phi_j,a_j,r_j,\phi_{j+1})$；
11：**if** episode 在时间步长 $j+1$ 处结束 **do**
12：$y_j=r_j$；
13：**else**
14：$y_j=r_j+\gamma\max_{a'}\hat{Q}(\phi_{j+1},a';\theta^-)$；
15：**end if**
16：对损失函数 $(y_j-Q(\phi_j,a_j;\theta))^2$ 执行关于参数 θ 的梯度下降步骤；
17：**end if**
18：每隔 C 个时间步长，重置 $\hat{Q}=Q$，即更新 $\theta^-=\theta$；
19：**end for**
20：**end for**
21：基于目标动作价值函数 $\hat{Q}$ 计算各子区间的最优驾驶动作序列
22：根据最优驾驶动作序列生成目标速度曲线
输出：
列车站间最优目标速度曲线

抽象为采用基于深度神经网络构建的动作价值函数Q（即Q网络），通过对Q网络进行M轮次的迭代训练，得到具有节能特性的目标速度曲线优化策略；再基于目标速度曲线优化策略，计算各子区间的最优驾驶动作序列；最后将最优驾驶动作序列代入列车数学模型，解出列车在站间运行的最优目标速度曲线。选取的运行线路段为某城轨线的跨河线路段，沿线地形变化丰富，能确保DQN算法在目标速度曲线优化策略训练过程中的稳定性和鲁棒性。

4　仿真与实际案例验证

4.1　目标速度曲线优化仿真对比验证

研究团队在MATLAB上搭建仿真环境，基于某城轨线路和某A型列车特性参数，对不同载客状态和运行时长下的列车能耗优化过程进行仿真，并与传统算法以及实车运行数据进行对比。

深度Q网络训练过程奖励变化如图8。由图8可看出，算法的训练效率很高，在2 000轮训练后进一步提升了寻优能力，并在4 000轮之后性能逐渐稳定。算法很好地平衡了寻优效果与收敛难度。智能体多次跳脱出局部最优，展现了出色的全局寻优能力。

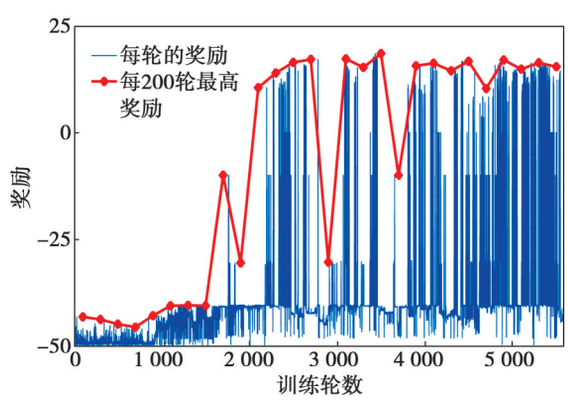


图8　深度Q网络训练过程奖励变化

Fig. 8　Reward changes during deep Q-network training process

另外，研究团队将所提算法与其他两种算法进行了对比：传统算法采用传统候选动作集，采用与所提算法相同的DQN架构和训练方法，但不考虑牵引效率；考虑牵引效率因素的算法同样采用候选动作集，将牵引效率引入节能奖励计算中，具体分析如下：

①固定载荷状态和运行时长下的优化结果分析。

列车载荷状态为AW2，目标运行时长为158 s，此时3种算法的优化结果如图9所示。图9中3条目标速度曲线均在首段和末段进行了持续的大幅加速和制动，但在中间部分有所不同。传统算法在中间段保持相对较低的速度，速度波动较小，在较晚的时机开始惰行降速和制动减速；考虑牵引效率因素的算法和本文所提算法则采取“先快后慢”的策略，先将速度提升至接近限速，维持一段路程后再开始惰行降速，充分利用列车存储的动能和线路势能向前行进。从运行安全的角度来看，传统算法通过降低中间段靠前部分的最大速度来保持更多的安全冗余，并通过延后降速时机的方式来确保准时到站；另外两种算法虽然减少了安全冗余，但得益于DQN训练时“轻奖重罚”的奖励制度，确保了线路限速的高优先级，因此依然有2 km/h的速度裕量，保证了运行安全。

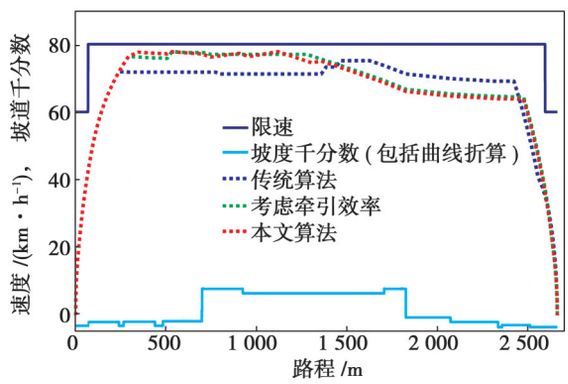


图9　AW2载荷状态-158 s运行时长下优化结果

Fig. 9　Optimization results of 158 s operation duration under AW2 load condition

表3汇总了前文3种算法的性能参数。从表3可以看出，通过引入末端制动比例可调机制，3种算法均将运行时长和停站位置的偏差分别控制在±1 s和±0.3 m以内，实现了准时运行和精准停站。基于传统算法优化得到的目标速度曲线具有较高的牵引能耗，而在考虑了牵引效率因素后能耗降低了9.64%，采用本文所提算法则进一步达到了12.79%的节能比例。

表3　3种算法优化后的目标速度曲线性能指标

Table 3　Performance metrics of optimized target speed profiles by three algorithms

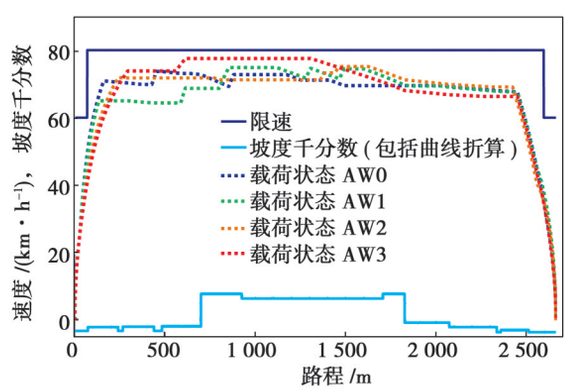
优化算法	运行时长偏差/s	停站位置偏差/m	牵引能耗/kWh	节能比例
传统算法	≈0	≈0	34.350	
考虑牵引效率	0.869	≈0	31.037	9.64%
本文算法	≈0	≈0	29.958	12.79%

②固定运行时长、载荷状态变化下的优化结果分析。

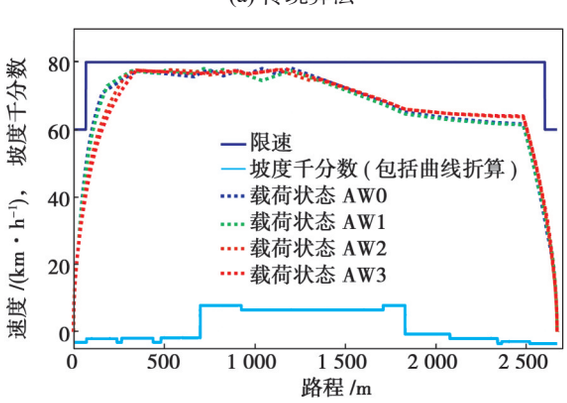
为检验客流变化场景下采用本文所提算法的节能效果，现设置列车的目标运行时长为158 s，载荷状态在AW0~AW3范围内变化，通过不同的载荷状态体现不同客流量的运行场景，对比传统算法和本文所提算法的目标速度曲线优化结果。

图10(a)和图10(b)分别展示了传统算法和本文所提算法的目标速度曲线优化结果。由于列车车重较大时运行阻力也比较大，故列车在重载状态下巡航工况的能耗较高。随着列车载荷的增加，为避免因巡航工况

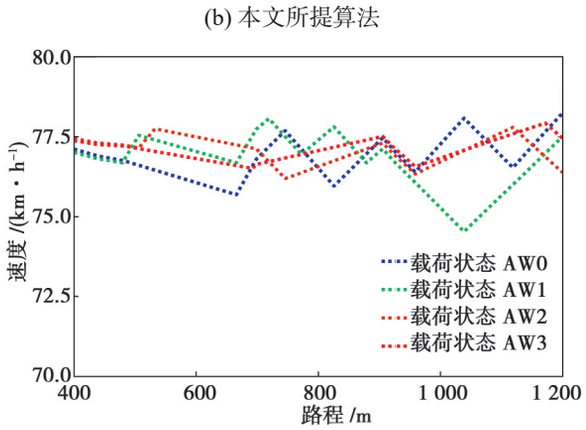
而产生过多的能耗，传统算法采用减少“巡航”动作频率和增加“惰行”动作频率的方式来优化目标速度曲线。例如，在AW3载荷状态下传统算法采用了类似本文所提算法的“先快后慢”的策略，通过先高速巡航，再惰行降速的方式减少中间段的牵引能耗。本文所提算法则在不同的载荷状态下采用相似的驾驶策略，其优化出的目标速度曲线整体上比较相近，但细节处仍有差异。如图10(c)所示，在400~1 200 m区段，不同载荷状态下算法采用了不同的速度控制方式。当列车轻载时，“高效率牵引”动作输出的牵引力较小，与“惰行”动作的牵引力差值显著，导致算法需频繁切换两种动作以维持速度，速度波动明显；当列车重载时，“高效率牵引”牵引力与运行阻力接近，使得算法可采用长距离单一动作策略，速度曲线趋于平缓。上述对比表明，若不将车重作为状态输入（如传统算法），则无法自适应调整驾驶动作切换频率，这解释了传统算法在轻载时节能比例较低（AW0下仅14.69%）的原因，故在优化目标速度曲线时考虑客流变化因素是非常有必要的。



(a) 传统算法



(b) 本文所提算法



(c) 本文所提算法优化后的目标速度曲线的中间段前部

图10 158 s运行时长-不同载荷状态下的优化结果

Fig. 10 Optimization results of 158 s operation duration under different load conditions

表4记录了不同载荷状态下两种算法的牵引能耗，并计算了节能比例。在4个级别的载荷状态下，本文所提算法均展现出了更好的节能性能。随着载荷的增加，传统算法的驾驶策略与本文所提算法越来越接近，故本文所提算法在轻载时的优势更明显。

表4 不同载荷状态下两种算法优化后的目标速度曲线能耗

Table 4 Energy consumption of optimized target speed profiles by two algorithms under different load conditions

载荷状态	优化算法	牵引能耗/kWh	本文算法节能比例
AW0	传统算法	24.481	14.69%
	本文算法	20.884	
AW1	传统算法	25.399	13.56%
	本文算法	21.956	
AW2	传统算法	34.350	12.79%
	本文算法	29.958	
AW3	传统算法	36.093	7.49%
	本文算法	33.391	

4.2 基于线路数据的运行案例对比分析

为进一步检验所提目标速度曲线优化算法的节能性能，研究团队从实车车载ERM中提取了列车在前文所提路段运行的数据，验证其实用价值。根据ERM提取得到的实车数据，在某段区间列车载荷状态接近AW2，站间运行时长为158 s。基于此工况完成目标速度曲线优化，以下将进行详细分析。

图11对比了实车速度曲线与采用本文所提算法优化后的目标速度曲线。从图11可以看出，实车在中间段采取了长距离的近似匀速行驶的策略，通过巡航、小幅度加速和惰行保持稳定的运行速度，并在比较靠前的位置开始制动减速。实车在中间段保持匀速行驶，部分原因是驾驶员为避免延误而保留安全冗余。本文所提算法量化了列车在每个位置的理论最小剩余时长，使智能体能够在保证正点的前提下，将1 100~2 500 m区段设置为惰行（见图11），更加充分地利用了线路势能；实车在约2 200 m处开始制动减速，较早地进入制动阶段。相比之下，优化曲线将制动起始点从2 200 m延后至约2 500 m处，减少了制动过程中的动能损失。图12和表5对比了实车运行曲线与目标速度曲线优化情形下的牵引能耗。图12所示的累计牵引能耗曲线在2 500 m后趋于平缓，验证了奖励函数设计的有效性。

相比实车运行曲线，本文所提的目标速度曲线优化实现了节能减耗，从而验证了本文所提节能方案的有效性和实用性。通过目标速度曲线优化，牵引能耗由 35.037 kWh 降至 29.958 kWh，实现了 14.5% 的牵引节能。

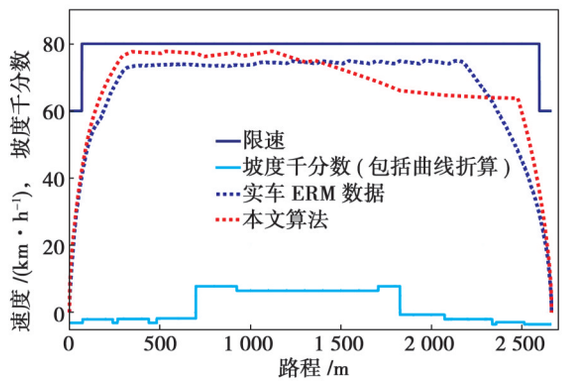


图 11　实车速度曲线与优化后的目标速度曲线对比

Fig. 11　Comparison between real train speed profile and optimized target speed profile

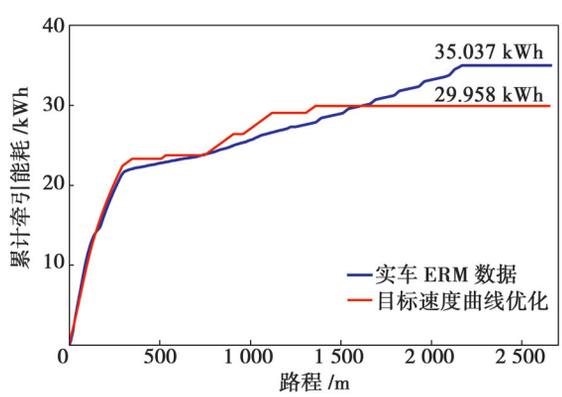


图 12　两阶段节能优化前后的累计牵引能耗变化

Fig. 12　Variations in cumulative traction energy consumption before and after two-stage energy-saving optimization

表 5　不同节能策略的能耗对比

Table 5　Comparison of energy consumption for different energy-saving strategies

节能策略	牵引能耗/kWh	节能比例
实车运行曲线	35.037	
进行目标速度曲线优化	29.958	14.50%

5　结论

本文基于某城轨线路和某列车特性参数，采用深度强化学习（DQN）框架，引入“高效率牵引”动作替代传统“巡航”动作，显著降低低效牵引工况占比。在 AW2 载荷、158 s 运行时长的场景下，牵引能耗由 34.350 kWh 降至 29.958 kWh，节能比例达 12.79%。算法通过调整惰行起始时机适应不同运行时长（158~168 s），在 168 s 时节能比例提升至 14.03%。轻载（AW0）与重载（AW3）场景下节能比例分别为 14.69% 与 7.49%，验证了算法对客流变化的适应性。

参考文献：

[1] 李蔚, 刘高峰, 赵思哲, 等. 基于不同速度控制模式的列车驾驶策略优化研究[J]. 铁道科学与工程学报, 2022, 19(8): 2169-2181.
Li Wei, Liu Gaofeng, Zhao Sizhe, et al. Research on the optimization of train driving strategy based on different speed control modes[J]. Journal of Railway Science and Engineering, 2022, 19(8): 2169-2181.

[2] 李茜, 李蔚, 曹悦, 等. 基于DDDQN的城轨列车节能运行控制方法研究[J]. 铁道科学与工程学报, 2024, 21(12): 4960-4970.
Li Qian, Li Wei, Cao Yue, et al. Research on energy-saving operation control method of urban rail train based on DDDQN[J]. Journal of Railway Science and Engineering, 2024, 21(12): 4960-4970.

[3] 李蔚, 曹悦, 娄永梅, 等. 基于FSMPC的列车精准停车阶段动态优化研究[J]. 铁道科学与工程学报, 2025, 22(10): 4424-4435.
Li Wei, Cao Yue, Lou Yongmei, et al. Study on optimization of train precise stop stage adjustment based on finite set model predictive control[J]. Journal of Railway Science and Engineering, 2025, 22(10): 4424-4435.

[4] 中国城市轨道交通协会. 城市轨道交通2024年度统计和分析报告[J]. 现代城市轨道交通, 2025(4): 113.
China Association of Metros. Urban rail transit 2024 annual data statistics and analysis[J]. Modern Urban Transit, 2025 (4): 113.

[5] 侯秀芳, 冯晨, 燕汉民, 等. 2024年中国内地城市轨道交通运营线路概况[J]. 都市快轨交通, 2025, 38(1): 13-19.
Hou Xiufang, Feng Chen, Yan Hanmin, et al. Overview of urban rail transit operations in China's mainland in 2024[J]. Urban Rapid Rail Transit, 2025, 38(1): 13-19.

[6] Howlett P. An optimal strategy for the control of a train[J]. The ANZIAM Journal, 1990, 31(4): 454-471.

[7] Howlett P G, Milroy I P, Pudney P J. Energy-efficient train control[J]. Control Engineering Practice, 1993, 26(2, Part 4): 1081-1088.

[8] Howlett P. Optimal strategies for the control of a train[J]. Automatica, 1996, 32(4): 519-532.

[9] Howlett P. The optimal control of a train[J]. Annals of Operations Research, 2000, 98(1): 65-87.

[10] Howlett P G, Pudney P J, Vu X. Local energy minimization in optimal train control[J]. Automatica, 2009, 45(11): 2692-2698.

[11] Albrecht A, Howlett P, Pudney P, et al. The key principles of optimal train control—part 1: formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points[J]. Transportation Research Part B-

Methodological, 2016, 94: 482-508.

[12] Albrecht A, Howlett P, Pudney P, et al. The key principles of optimal train control—Part 2: existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques[J]. Transportation Research Part B-Methodological, 2016, 94: 509-538.

[13] 曲健伟, 王青元, 孙鹏飞. 基于极大值原理的地铁列车节能驾驶简化算法[J]. 铁道科学与工程学报, 2019, 16(6): 1577-1586.
Qu Jianwei, Wang Qingyuan, Sun Pengfei. A simplified algorithm of energy-efficient driving for metro trains based on the maximum principle[J]. Journal of Railway Science and Engineering, 2019, 16(6): 1577-1586.

[14] Cao Y, Wang Z C, Liu F, et al. Bio-inspired speed curve optimization and sliding mode tracking control for subway trains[J]. IEEE Transactions on Vehicular Technology, 2019, 68(7): 6331-6342.

[15] Farooqi H, Fagiano L, Colaneri P, et al. Shrinking horizon parametrized predictive control with application to energy-efficient train operation[J]. Automatica, 2020, 112: 108635.

[16] 厉高, 林建辉, 庄哲, 等. 基于再生制动的城市轨道列车节能控制研究[J]. 铁道运输与经济, 2019, 41(3): 121-126.
Li Gao, Lin Jianhui, Zhuang Zhe, et al. A research on energy-saving train control of urban mass transit based on regenerative brake[J]. Railway Transport and Economy, 2019, 41(3): 121-126.

[17] Zhou R, Song S J. Optimal automatic train operation via deep reinforcement learning[C]//2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), 2018: 103-108.

[18] Zhou K C, Song S J, Xue A K, et al. Smart train operation algorithms based on expert knowledge and reinforcement learning[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(2): 716-727.

[19] Ning L B, Zhou M, Hou Z P, et al. Deep deterministic policy gradient for high-speed train trajectory optimization [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 11562-11574.

[20] Huang J, Zhang E D, Zhang J R, et al. Deep reinforcement learning based train driving optimization[C]//2019 Chinese Automation Congress (CAC), 2019: 2375-2381.

[21] 郭啸, 孟建军, 陈晓强, 等. 城轨列车深度强化学习节能优化控制方法[J]. 铁道标准设计, 2024, 68(7): 185-191.
Guo Xiao, Meng Jianjun, Chen Xiaoqiang, et al. Energy-saving optimization control method for reinforced learning of urban rail train[J]. Railway Standard Design, 2024, 68 (7): 185-191.

[22] 张淼. 基于强化学习的列车自动驾驶方法研究[D]. 北京: 中国铁道科学研究院, 2020.
Zhang Miao. Research on automatic train operation based on reinforcement learning[D]. Beijing: China Academy of Railway Sciences, 2020.

Optimization of target speed profiles for urban rail trains considering dynamic passenger flow

XIANG Chaoqun[1], DU Jingrun[1], YANG Yang[2], ZHANG Qingqi[2], ZHU Hengguo[3], FAN Ziyin[4]

(1. School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan 410075, China;
2. Changsha Metro Group Co., Ltd., Changsha, Hunan 410027, China;
3. Hunan CRRC Times Signal & Communication Co., Ltd., Changsha, Hunan 410100, China;
4. CRSC Urban Rail Transit Technology Co., Ltd., Beijing 100070, China)

Abstract: Urban rail transit systems face growing energy consumption challenges. Existing target speed profile optimization methods typically assume static passenger flow, leaving a gap in exploring optimization algorithms that adapt to dynamic passenger flow scenarios. Moreover, most previous energy-saving studies overlook the coupling relationship between traction efficiency and traction force and thus fail to achieve reductions in train traction loss by improving traction efficiency. To address these gaps, this paper constructs an energy-saving control framework for urban rail trains that integrates dynamic passenger flow and traction efficiency optimization, forming a complete technical system from theoretical modeling to algorithm design. A target speed profile optimization algorithm based on a deep Q-network (DQN) is proposed, which improves adaptability to dynamic passenger flow scenarios and explicitly accounts for traction efficiency factors. Firstly, the "cruise" driving action is removed, and a new set of candidate driving actions including "high-efficiency traction" is constructed. Secondly, the optimization range of the algorithm is expanded through dense partitioning into multiple sub-intervals. Finally, a mechanism for adjustable terminal braking ratio and determination of the theoretical minimum remaining running time is introduced to simplify optimization. Train weight information under different load conditions is used as an important algorithm input, linking the optimization process with passenger flow variations. Moreover, traction efficiency is incorporated into the optimization process via the "high-efficiency traction" driving action and DQN reward function based on energy consumption calculations. These features enable the algorithm to account simultaneously for passenger flow variations, traction efficiency, and line conditions. The resulting strategy is trained within a reinforcement learning framework. Multi-scenario simulations and comparisons with real train data verify the energy-saving effect of the proposed method.

Keywords: energy conservation of urban rail train; traction efficiency of urban rail train; dynamic passenger flow; optimization of target speed profile; deep Q-network; urban rail transit

收稿日期：2026-03-18；修回日期：2026-05-11
责任编辑：钟　锋