

密 级： 公开
加密论文编号： _____

论文题目： 速度优化的学习型矿用无人驾驶车队调度研究

学 号： M202110447
作 者： 裴树伟
专 业 名 称： 机械工程

2024 年 05 月 27 日

速度优化的学习型矿用无人驾驶车队调度研究

**Learning-Based Dispatch System with Trajectory
Optimization for Mining Autonomous Fleet**

研究生：裴树伟

指导教师：杨珏

北京科技大学机械工程学院

北京 100083，中国

Master Degree Candidate: Shuwei Pei

Supervisor: Jue Yang

School of Mechanical Engineering

University of Science and Technology Beijing

30 Xueyuan Road, Haidian District

Beijing 100083, P.R.CHINA

中图分类号： U492.6+4
U D C: 621

学校代码： 10008
密 级： 公开

北京科技大学硕士学位论文

论文题目： 速度优化的学习型矿用无人驾驶车队调度研究

作者： 裴树伟

指导教师： 杨珏 单位： 机械工程学院 职称： 教授

小组成员： 单位： 职称：

 单位： 职称：

论文提交日期： 2024 年 05 月 27 日

学位授予单位： 北 京 科 技 大 学

摘 要

由于传统矿山运输业安全风险较高，并且人员老龄化严重，同时伴随着能源消耗和环境污染等问题，智能矿山在国内外引起了广泛关注。在矿山交通运输中，一方面，能源消耗占据了成本的一半；另一方面，装卸载排队的时间是卡车运输效率低下的主要原因。为了解决这些问题，我们针对经济型行驶策略和优化排队时间提出了速度优化的学习型矿用无人驾驶车队调度系统。

首先，我们利用动态规划算法将道路进行分段求解，计算得到车型的最佳行驶速度策略。对特定道路和特定车型进行针对性优化，得到最符合矿山运输成本的行驶速度策略。本文提出了一种考虑行驶时间、能量消耗和电池寿命的多目标优化方法，优化后的速度策略是一种帕累托前沿平衡。

其次在优化后的行驶策略上，本文采用多智能体 Deep Q-learning Network 强化学习方法训练调度策略。结合真实矿山参数创建了一个高度适应性的采矿模拟器，将每个车当作智能体，将深度强化学习与速度优化相结合，训练出基于人工智能强化学习的智能车队调度系统，经过近 10 万次的模拟训练，该系统可以生成高效的动态调度策略。

最后，总结归纳在运输效率、成本控制、单车利用率和运输灵活性的评价指标，结合工程经验制定矿山运输评价体系，对优化后的学习型矿山运输系统进行全面的评价。经过试验，结果表明速度优化后的学习型调度系统比传统的调度模式和匀速模式拥有更好的运输效率和运输灵活性，能够在一些特殊场景下表现出鲁棒性。

关键词：车队管理；能耗优化；运输调度；学习型调度；

Learning-Based Dispatch System with Trajectory Optimization for Mining Autonomous Fleet

ABSTRACT

The safety risks and aging workforce in traditional mining transportation industries, coupled with concerns over energy consumption and environmental pollution, have sparked widespread attention to intelligent mining practices both domestically and internationally. In the realm of mining transportation, energy consumption constitutes half of the operational costs, while inefficient queueing during loading and unloading significantly impedes trucking efficiency. To address these issues, we propose a speed-optimized learning-based dispatch system for autonomous mining fleets, focusing on economic driving strategies and queue time optimization.

Initially, we employ dynamic programming algorithms to segment roads for solving and calculate the optimal driving speed strategy for each vehicle type. Tailored optimizations for specific roads and vehicle types yield driving speed strategies that best align with mining transportation costs. This paper presents a multi-objective optimization method considering driving time, energy consumption, and battery lifespan, resulting in a Pareto-front balanced vehicle speed.

Subsequently, leveraging the optimized driving strategies, we employ multi-agent Deep Q-learning Network reinforcement learning to train dispatch policies. By integrating real mining parameters into a highly adaptable mining simulator, each vehicle is treated as an intelligent agent, combining deep reinforcement learning with velocity optimization to develop an AI-based dispatch system for autonomous fleets.

Finally, we summarize evaluation metrics for transportation efficiency, cost control, individual vehicle utilization, and transport flexibility. Drawing upon engineering expertise, we devise a comprehensive evaluation framework for mining transportation systems, facilitating the assessment of the optimized learning-based mining transportation system.

Key Words: Fleet management; Energy consumption; Dispatching System; Learning-based system;

目 录

摘 要.....	I
ABSTRACT.....	II
1 绪论.....	1
1.1 研究背景和意义.....	1
1.2 经济型速度策略研究现状.....	2
1.3 矿山车队智能调度研究现状.....	4
1.4 本文主要研究内容.....	6
2 矿用无人驾驶车辆多目标行驶速度策略优化方法.....	8
2.1 引言.....	8
2.2 速度轨迹优化问题设计.....	9
2.2.1 矿山工作场景简介.....	9
2.2.2 车辆经济型行驶策略的优化方法.....	11
2.2.3 优化算法：动态规划.....	12
2.3 车辆模型概述.....	13
2.3.1 车辆动力学系统建模方法.....	13
2.3.2 动力系统设置.....	15
2.3.3 车辆纵向动力学模型.....	16
2.3.4 发动机模型.....	17
2.3.5 电机模型.....	17
2.4 优化算法设计与分析.....	18
2.4.1 单一目标优化.....	19
2.4.2 多目标优化.....	24
2.5 仿真结果与分析.....	30
2.6 本章小结.....	37
3 基于深度强化学习的矿山无人驾驶车队调度系统.....	38
3.1 引言.....	38
3.2 车队调度的强化学习方法.....	39
3.2.1 深度强化学习求解.....	41
3.2.2 多智能体深度强化学习.....	45
3.3 矿山无人驾驶车队调度系统设计.....	47

3.3.1 基于强化学习的车队管理系统.....	48
3.3.2 动态规划速度策略优化方法.....	51
3.4 仿真器设计.....	53
3.4.1 网络参数设置.....	55
3.4.2 训练路网环境.....	55
3.4.3 车队参数设置.....	55
3.5 结果验证.....	59
3.6 小结.....	61
4 矿山无人驾驶车队运输评价体系.....	62
4.1 引言.....	62
4.2 矿山无人驾驶车队运输调度评价体系概述.....	62
4.2.1 运输效率指标.....	63
4.2.2 成本控制指标.....	64
4.2.3 单车利用率指标.....	64
4.2.4 运输均匀度指标.....	65
4.2.5 安全稳定性指标.....	65
4.3 学习型矿山无人驾驶车队调度系统评价.....	67
4.3.1 参考基准模型.....	67
4.3.2 仿真环境设置.....	67
4.3.3 指标分析对比.....	69
4.4 本章小结.....	81
5 结论与展望.....	82
5.1 结论.....	82
5.2 局限和展望.....	82
参考文献.....	83
致 谢.....	91
作者简历及在学研究成果.....	92
独创性说明.....	93
关于论文使用授权的说明.....	94
学位论文数据集.....	95

1 绪论

1.1 研究背景和意义

人均汽车拥有量增加，导致环境恶化和能源危机。2021 年中国石油消耗 6.8 亿吨，进口占 85%，交通消耗 2.2 亿吨。占总量的 32.4%^[1]。许多研究人员已经研究了减少能源消耗的方法，这几年来，我国的科技发展有了质的飞跃，车辆领域迎来了新的阶段。智能化、电动化正在重新定义汽车，无人驾驶和能源消耗也迎来了新的机遇。

智能矿业的概念在全球范围内引起了广泛关注，原因是传统矿业行业存在安全风险和人员老龄化问题。此外，矿业运输的燃料消耗在该行业的温室气体排放中占据了很大比例。为了应对这些挑战，提高运营效率，人们开始致力于开发人工智能技术，以建立绿色和可持续的矿业环境。

在智能矿业中，先进的传感器和分析技术的应用使得矿区能够创建高清地图。这些地图为矿业运输系统的速度和轨迹优化技术提供基础，从而实现能源消耗的降低。同时，人工智能和工业物联网（IIoT）的融合使得智能矿业能够高效自动化地进行运营。数据驱动的自动化已经开始改变核心的矿业运营方式。在矿业运营中，卡车调度活动的重点是如何高效地将开采的材料从采矿工作面通过铲车运送到指定的目的地。有效的调度决策不仅可以极大地提高运输资源的利用率和效率。

非公路矿用自卸车通常在封闭且环境恶劣的矿山园区工作。矿山道路硬件基础设施较差、坡度变动频繁，自身负载载荷变化较大，通常矿山园区内自卸车的工作时速要求并不高，一般在 30km/h。在相对固定的路线中，矿用自卸车在装货点和卸货点之间来回穿梭，公路坡度和矿用自卸车荷载的巨大变化导致了较高燃料消耗和较低运输效率。因此，在露天矿山运输系统中，矿石运输是整个矿山成本中最重要的部分。运输所产生的各种费用占了矿山总体生产成本的一半，甚至在一些大型露天矿中占到 60%以上^[2]。矿山的无人封闭环境和巨大的经济潜力促使对车辆无人驾驶和最优速度行驶策略的改进至关重要，以提高车辆的燃油经济性和运输效率。

在采矿作业中，装载排队等待占据了卡车潜在空闲时间的最大比例，主要原因是铲车-卡车系统组织不够高效^[3]。智能车队管理在采矿行业中起着关键作用，研究表明通过有效的车队调度可以显著提高生产力。自上世纪 80 年代以来，研究人员已经开发了许多车队调度系统，基于露天矿调度算法，以建立调度员和卡车操作员之间的联系。在新型矿山无人车辆中，既包括燃油车辆，还包括电动车辆和混合动力车辆。像电动和混合

动力汽车，再生动能的能力可以减少能耗。另一种提高车辆系统效率的方法是改进驱动技术，提高传动系统的效率。第三种方法是调整驾驶员的行为，车辆按着经济性速度行驶，这种不改变车辆配置实现消耗最小化的速度行驶策略称为经济性速度策略^[4]。车辆的运输过程是整个运输系统中最值得进行优化的环节，对于露天矿山调度系统的优化目标可以归纳为降低运输成本以及提高运输效率。但是在实际的应用场景中，降低运输成本和提高运输效率是两个互相矛盾的方面。因此，设计一种多目标的最优的速度策略和调度策略似乎是至关重要的。

与过去的研究相比，燃油经济性将不再是唯一的主要的矿车性能指标，运输时间以及其他消耗费用也被认为可以降低运输成本。并且随着新型矿山车辆的出现，能量的消耗不仅仅局限于燃油经济性本身，同样还有电量消耗。本文提出了一种基于固定运输路线的矿用自卸车在固定运输路线下的车辆速度优化方案，并且基于该速度策略训练出学习型的车队调度策略。基于 MATLAB 和 Pytorch 平台，建立了汽车的各个零部件（包括动力系统和传动系统）以及矿山等场景模型。针对矿山路网中的无人驾驶车辆，设置合理的优化目标，采用多种优化方法对不同的车型制定最优的行驶策略，实现单车能耗和效率相对最优。在尽可能保证运输效率的情况下，训练车队调度策略，研究不同因素对调度策略的影响，比较优化后运输成本的改善。

1.2 经济型速度策略研究现状

在无人驾驶条件下，露天矿山单个汽车运行的最优问题，应该以矿用汽车在该路径下矿石运输成本最低为优化目标。为了达成这一目标，需要综合考虑矿用汽车在该路径下的燃油消耗和所用时间（运输效率）。对于内燃机动力汽车运行燃油消耗最低的问题，学者们的经验性研究得到的结论主要有：避免频繁的加速和制动，避免激进加速和过于缓慢的加速过程，避免长时间怠速。

来自 Johns Hopkins 大学的 E Gilbert^[5] 首次对巡航工况进行研究，通过数学严格证明了周期控制优于“准松弛稳态(QRSS)”法。就机理而言，匀速行车策略也算是一种 QRSS，但通常被误认为“油耗最优”。这项研究首次推翻了这一观念。同时，有学者指出发动机的万有特性并非全局“凸”性^[6]。因此在有限时间内，最优解是一系列切换控制，而非固定油门开度策略。Li 等人^[7]在研究中通过最优控制发现了前车匀速行驶中的新方法，提出了“加速-滑行”式行车策略，并给出了发动机控制准则，揭示了定量化分析在节油规律探索中的潜力。J. Ma^[8]等人针对起伏的山岭地区，测试并进一步验证了松彭特里亚金极小值的方法在车辆经济性驾驶策略上可行性，经过对比，该优化后的策略可以节约约 20%的燃油消耗，为后续的研究奠定了理论基础。此外 D. Shen^[9]等人采用彭特里亚金

极小值的方法对变速箱换挡策略进行分析,考虑怠速、换挡等状态,最后得到的阶段最优控制经过验证后油耗节约 4.8%。徐少兵等人^[10]利用 Legendre 伪谱法研究了 CVT 型车辆的经济性加速策略,并通过数值求解得到了经济性加速度的量化表达。他们发现,“加速-滑行”操作能够提高负荷率,改善发动机效率,从而降低油耗^[11]。此外,他们还研究了挡位离散型车辆的经济性加速策略,并提出了节油潜力验证的方法^[12]。

通常情况下,经济性的驾驶策略不仅与车辆本身动力结构的特点有关,而且还取决于道路环境和交通状态。经济性驾驶策略的获取方式通常可分为经验型和理论型两类。

经验型策略是基于对实际驾驶经验和道路数据的分析总结而得出的节油操作规律。然而,这种策略存在一些缺点:首先,由于驾驶经验的个体差异和车辆特性的影响,这些规律往往缺乏普适性,难以达到最优节油效果;其次,缺乏清晰的节油机理,使得这些规律不利于驾驶辅助系统的设计和车辆控制;经验型操作策略还包括车辆的理保养和规范用车。车辆良好状态是节能的前提,因此合理保养和规范用车至关重要^[13]。典型的措施包括定期更换润滑油、空气滤清器,以及合理调整轮胎等,这些措施有助于提高发动机效率和降低行驶阻力。尽管本文重点是研究速度行驶策略和能耗关系,对车辆保养和维护的讨论将不再展开。此外,获取这些规律需要进行大量的实车实验,成本高且难以迁移应用到不同车型中。

相比之下,理论型策略基于车辆的物理模型和数学优化,通过计算得到最优的节能驾驶策略^[6, 14-17]。这种策略具有成本低、最优性高的优势,能够更好地优化节能效果。通过大量的数学计算,理论型策略能够充分利用车辆动力学和发动机模型,最大程度地提高节能能力。

目前,理论性驾驶策略主要围绕典型或固定的驾驶情景展开研究,如坡道、跟车、多信号灯等。研究流程包括:对车辆的内部动力结构、运行工况和路况进行数学化建模;使用数学优化工具算法进行复杂的策略辨识和规划求解;针对计算得到的驾驶策略进行量化定性分析,总结规律性的操控原则。然而,在实际研究中,提取操控准则和规律性的难度较大,特别是量化节油规律的挑战十分艰巨。尽管已有一些针对燃油车辆的提取方法,但对电动车辆和混合动力车辆的经济性速度策略研究尚不充分。

新能源汽车的发展对动力系统提出了更高的要求。单个动力电池组成的动力系统存在诸多问题:(1)比功率低和爬坡等大功率场景瞬时尖峰电流放电影响电池循环寿命。(2)提高动力电池的放电电流和功率上限,影响汽车布局^[18],同时增加成本,也不利于轻量化。(3)续驶里程短。

全局优化计算量大,不具备实时性,DP 等全局优化算法可以作为线上控制的规则库设计基础,可以作为验证其它控制算法有效性的基准。Bianchi 等人^[19]提取动态规划

方法的规则并应用于混合动力车辆，与动态规则方法进行了比较验证。长安大学的张盟阳^[20]针对长 12 米的串联式混合动力汽车，采用动态规划逆向优化过程，相对于恒温器控制策略减少能耗 2.38%~4.54%。

J. Hu 等人^[21]利用彭特里亚金极小值的方法对智能混合动力车辆在丘陵地形下进行经济性驾驶策略优化。Qunya Wen^[22]针对非道路混合动力车辆提出了自适应等效燃油消耗策略，节省燃油高达 6.36%。徐少兵^[23]针对并联混合动力车辆，利用伪谱法进行能量优化研究，发现最节油的操作仍是速度波动型巡航。梁志豪^[24]针对 PHEV 车辆，采用动态规划和伪谱法优化发动机油耗和动力电池电耗，发现伪谱法在求解精度、收敛速度和约束处理方面更有优势。甄玉珂^[25]针对混动车辆，使用伪谱法根据经验和等效燃油消耗模型进行模式划分，验证其在燃油经济性和适用性方面具有良好效果。

随着科技的进步，智能算法的迅速崛起，机器学习也很快被应用到车辆领域。Huang X^[26]等人将机器学习算法用于识别不同驾驶情况，通过多特征提取，创建多个子模型进而实现经济型控制。该系统能够根据选择的驾驶信息有明显影响的特征值对车辆进行实时控制。

1.3 矿山车队智能调度研究现状

自动驾驶系统通过深度学习架构在各种任务中取得了显著进展。以往的研究，例如 Kiran 等人^[27]的工作，已经探索了强化学习（Reinforcement Learning, RL）在自动驾驶中的应用，涵盖了 RL 框架内的关键概念、扩展和挑战。值得注意的是，RL 算法，包括用于离散动作空间的深度 Q 网络（Deep Q-Network）和用于连续动作的深度确定性演员-评论家（Deep Deterministic Policy Gradient, DDPG），Sallab^[28]已经在 TORCS 模拟器中进行了评估并应用于车道保持场景。这些 RL 方法结合了 Q 学习和模型预测控制，在建模误差较大的情况下表现出比传统方法更好的鲁棒性。因此，研究工作集中在多智能体强化学习（Multi-Agent Reinforcement Learning, MARL）方法上，以应对自动驾驶任务的多智能体特性。例如，在合成和真实交通场景中，Chu 等人^[29]比较了基于演员-批判（Advantage Actor Critic, A2C）方法的可扩展和分散 MARL 算法与独立 A2C 和独立 Q 学习算法，结果显示在现有的分散 MARL 算法上具有更好的优化性、鲁棒性和样本效率。MARL 方法还被应用于高层次决策和协调自主车辆在高速公路场景中的行为，展示了在自动驾驶中学习安全策略的潜力。Luo 等人^[30]开发了一个 MARL 框架，用于模拟共享电动出行系统中的均衡任务，并考虑到电动车的续航里程和充电特性。

总之，强化学习、多智能体强化学习和深度 Q 网络技术的应用在推动自动驾驶系统方面展示了巨大的潜力，解决了决策、协调和优化等方面的挑战，最终实现了更安全、

更高效的自动驾驶体验。这些技术还为在采矿场景中自主卡车的调度策略开发提供了新的机会。

早期的方法,如 Elbrond^[31]和 Soumis^[32]等人提出的方法,通过解决非线性规划(NLP)网络问题来解决车队调度问题。他们的目标是找到装载点和卸载点之间的最短路径,考虑可用的卡车,并在考虑重量限制的同时最大限度地提高生产。Temeng、Otuonye 等人^[33]利用 Dijkstra 算法通过找到源和汇点之间的最短路径来确定最佳路线。最近,混合整数线性规划(MILP)方法在车队调度中变得流行起来,旨在最小化延迟和成本。多项研究采用了 MILP 技术,包括 Topal 和 Ramazan^[34]、Ta、Ingolfsson 等人^[35]、Chang、Ren 等人^[36]以及 Upadhyay 和 Askari-Nasab^[37]。排队理论也被应用于矿山运输系统中,其中 Koenigsberg^[38]是该领域的先驱之一。Kappas 和 Yegulalp^[39]提出了一个排队理论模型,将卡车-铲车系统视为一个生产网络,将卡车视为顾客,其他服务区域视为服务器。在后续研究中,Ercelebi 和 Bascetin^[40]开发了一个排队理论模型,根据卡车数量、处理工厂吞吐量和等待时间等性能参数来分配卡车。遗传算法也被用于优化露天矿卡车调度问题。He、Wei 等人^[41]采用了遗传算法方法来解决这个问题。此外,多目标遗传算法专门用于解决在露天矿山作业中分配一队异质卡车的任务,由 Alexandre、Campelo 等人^[42]展示了这一点。在采矿工业中采用了各种方法,从非线性规划和排队理论模型到混合整数线性规划和遗传算法,以解决车队调度挑战。这些努力旨在优化卡车分配,减少延迟并提高露天矿山作业的生产力。

传统的露天矿山作业调度系统往往缺乏对卡车故障和铲车延迟等意外事件的实时响应能力。因此,这些传统方法在采矿环境发生变化时无法适应和优化调度模型。为了解决这个局限性,人工智能(AI)算法,特别是强化学习(RL)和实时调度系统已经成为智能车队调度的有效解决方案。de Carvalho 和 Dimitrakopoulos^[43]实施了一个深度 Q 学习强化学习模型来提出一种新颖的卡车调度策略。这种方法利用了调度经验所得的知识,并展示了生产效率和车队管理的改进。类似地,Zhang、Odonkor 等人^[44]和 Huo、Sari 等人^[45]利用 Q-Learning 作为无模型的强化学习算法来提高车队生产力和减少等待时间。强化学习技术也被应用于电动车队,以最小化客户等待时间、电力成本和运营费用^[46]。在车队管理的背景下,Holler、Vuorio 等人^[47]提出了一种深度强化学习方法,将个体驾驶员视为智能体,一个中央车队管理智能体负责代表所有驾驶员进行决策。为了解决大规模车队管理问题,提出了一个包含上下文深度 Q 学习和上下文多智能体演员-评论家的上下文多智能体强化学习框架^[48]。Ahumada、Riveros 等人^[49]描述了基于多智能体系统(MAS)的分布式方法,其中智能体代表铲车和卡车并相互交互。Wang、Dai 等人^[50]开发了一个用于异构车队的实时调度模型,目标是 minimize 卡车等待时间、与计划路径流量的偏差和运输成本。

然而，传统的调度系统在有效处理随机性和不确定性方面往往存在局限性。这些系统往往在短期规划能力上有缺陷，并且对影响车队管理的不确定性考虑不足。此外，目前对不依赖人工驾驶员的自主卡车车队调度系统的研究有限。Ma 等人^[51]在一个连接和自动化车辆平台上使用 Relaxed Pontryagin's Minimum Principle (RPMP) 对新开发的算法进行了测试和验证。实验数据表明，实施这些算法可以节省燃料消耗超过 20%。该算法基于矿山现场的详细道路信息计算和优化自主卡车的速度和轨迹，从而提高效率和节约能源。因此，有必要针对自主卡车专门设计智能车队调度系统，促进采矿行业向更智能、更高效的车队管理过渡。本研究旨在通过集成和评估基于人工智能的智能调度系统和自主卡车速度优化，解决这些问题，并应用于露天矿山作业中。

1.4 本文主要研究内容

智能矿业的概念在全球范围内引起了广泛关注，原因是传统矿业行业存在安全风险和人员老龄化问题。此外，矿业运输的燃料消耗在该行业的温室气体排放中占据了很大比例。在采矿作业中，装载排队等待占据了卡车潜在空闲时间的最大比例。为了应对这些挑战，提高运营效率，设计一种最优的速度策略似乎是至关重要的。与此同时智能的车队管理在采矿过程中也起着关键的作用。本文针对上述问题，利用动态规划优化算法求解出最适合矿山成本管理的行驶速度，针对特定的道路和车型进行计算优化。基于最优的行驶速度策略上，训练出基于人工智能强化学习的智能车队调度系统。

第一章：阐述本文所研究的矿山无人驾驶车队管理调度的意义，介绍了运输过程中的车辆行驶速度优化的现状，以及基于多智能体调度优化模型方法，最后结合本文研究内容对论文整体结构介绍。

第二章：本章提出了一种考虑行驶时间、能量消耗和电池寿命的多目标优化方法，采用动态规划对无人采矿车辆的速度轨迹优化。降低运输成本和提高运输效率是两个互相矛盾的方面，设计一种适合矿山企业运营的最优的速度策略似乎是至关重要的。

第三章：本章通过将调度问题形式化为多智能体强化学习问题，结合真实矿山参数创建了一个高度适应性的基于事件的采矿模拟器。提出了一种新颖的综合方法，将深度强化学习与速度优化相结合，用于在采矿作业中调度无人驾驶卡车，研究表明有效的车队调度可以显著提高生产力和降低能耗。

第四章：本章对矿山运输调度系统进行概述，总结归纳在运输效率、成本控制、单车利用率和运输灵活性的评价指标，结合工程经验制定矿山运输评价体系。并利用评价指标本文提出的速度优化的学习型调度系统全面评估测试，为运输工作评价提供了依据。

本文研究路线如图 1-1 所示。

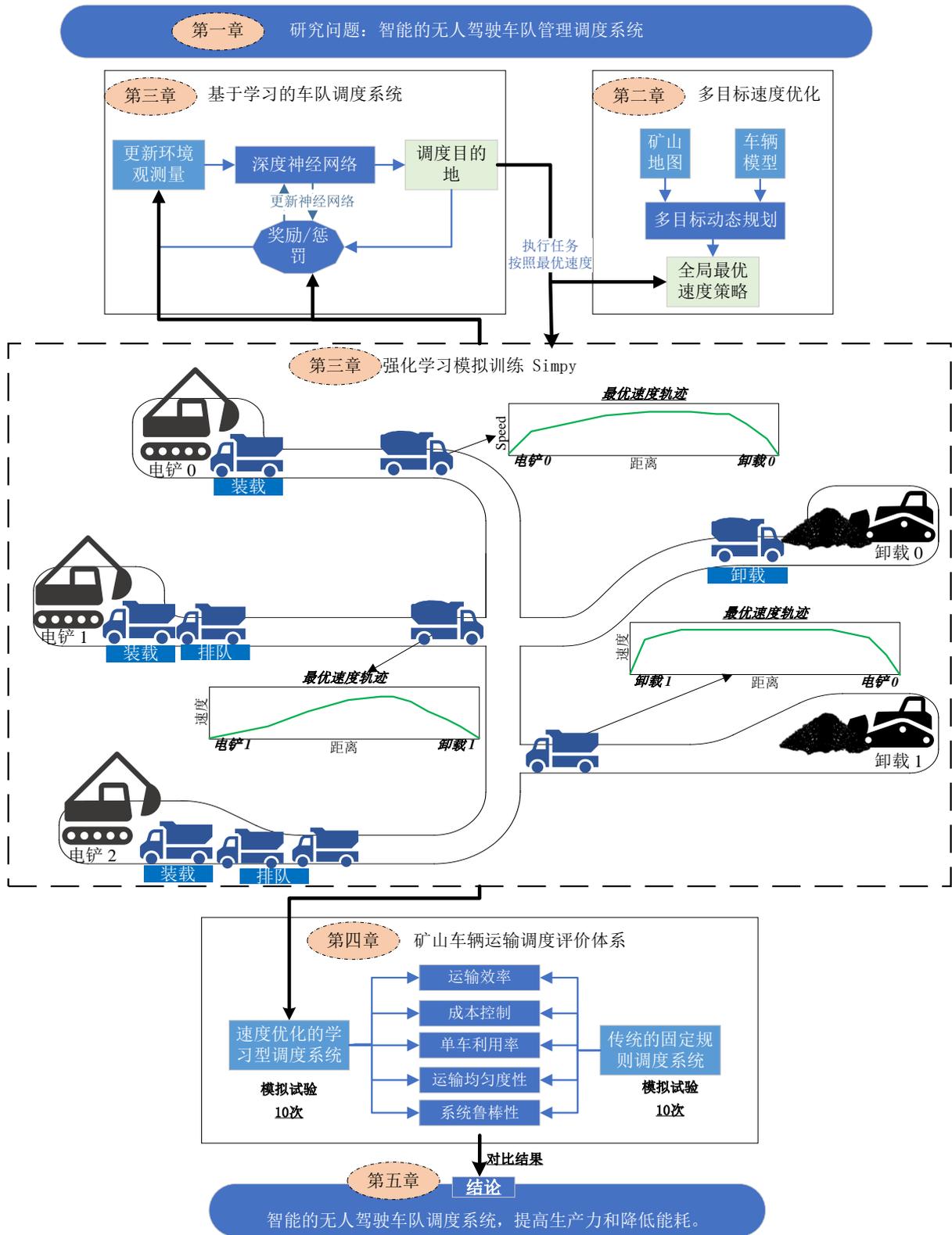


图 1-1 本文研究路线

Fig.1-1 Research line of this paper.

2 矿用无人驾驶车辆多目标行驶速度策略优化方法

2.1 引言

无人采矿运输是一种智能交通控制系统，相比传统的运输系统，可以提供更好的经济效益。有人驾驶车辆的速度轨迹取决于驾驶员的驾驶风格，但在自动驾驶条件下，可以利用数学方法进行优化。本文以载重 50 吨的燃油和电动矿用车辆为研究对象，提出了一种考虑时间、能量消耗和电池寿命的多目标优化方法。采用动态规划算法，在两种不同的矿区条件模拟下解决了具有不同优化目标的最佳速度轨迹。通常采用单一目标（能量消耗）优化的速度轨迹常常采用加速滑行（PnG）方法，导致电池容量损失增加和行驶时间增加。因此，提出了一种多目标优化方法。对于电动车辆，通过多目标方法优化的轨迹可以减少电池容量损失 22.01%，减少时间消耗 41.28%，导致能量消耗增加 42.12%。对于燃油车辆，可以减少时间消耗 32.54%，导致能量消耗增加 7.68%。这种速度轨迹更加平稳，波动较小，可以更好地满足采矿运输的要求，并对优化封闭区域的自主运输成本具有特殊的参考价值。

为了在有限的资源（如车辆数量和时间消耗）下最大化经济效益和总运输量，在传统的有人驾驶车辆中，整体运输成本主要与驾驶员的驾驶习惯和经验有关。车辆的速度优化是主要的生态驾驶技术之一，通过最小化能量消耗，可以在最大程度上扩展动力传动系统的能力和纵向控制。图 2-1 显示了在特定道路条件下的真实驾驶员速度曲线，时间消耗与距离之间存在明显的线性关系，并且在整个行程中有大量的成本优化空间。自主采矿运输系统与传统运输系统的一个重要区别是自主车辆的速度轨迹可以通过详细的路线信息和车辆特性进行优化。因此，在自主采矿运输系统中，一些研究人员已经证明在不同的优化方法基础上，能源消耗、排放、运输效率和其他性能方面有很大的改进空间。

无论研究者关注哪个因素，核心问题都是解决从起点到终点的速度轨迹优化和能量分配问题。然而，在实际驾驶场景中优化整个速度轨迹是困难的，因为这是一个具有多个因素的非常复杂的问题。以往的速度和能量优化研究很少考虑不同的成本组成部分。此外，整个采矿班次的时间消耗是提高效率和经济性的重要因素，电池健康状况也被视为维持电动汽车长寿命的优化目标，许多因素存在冲突。降低能源消耗会导致行驶时间增加，刹车能量回收系统也会对电池健康造成更多损害。因此，由于复杂的因素和成本函数，速度优化是一个多目标和非线性的最优控制问题。

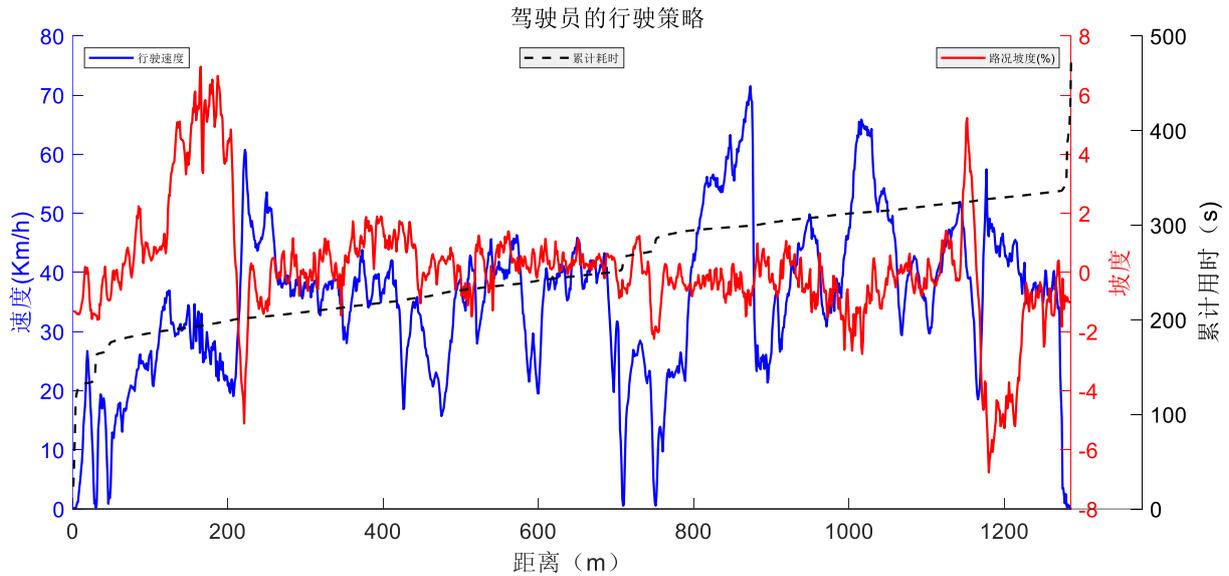


图 2-1 真实驾驶员速度曲线

Fig.2-1 Real drivers speed curve

本章提出了一种多目标优化策略。此方法的主要贡献从三个方面来看：

- (1) 考虑能量、效率和电池健康状况，基于物理约束条件构建了一种燃油车和电动车的多目标速度轨迹优化方法。
- (2) 采用动态规划方法处理多目标速度轨迹优化问题，以确保全局最优性。
- (3) 分析了通过单一能耗油耗和多目标优化得到的速度轨迹，揭示了能量、时间和电池损耗之间的权衡关系，有助于降低所需的采矿运输成本。

2.2 速度轨迹优化问题设计

由于矿山车辆的条件受到资源限制（例如车辆数量和工作时间），因此车辆应在已有的条件下进行高效运营。行驶时间越短，速度越高，通常能消耗越高。此外，过高的负荷会对燃油车和电动车的动力系统造成一定的损害。因此，必须优化行驶速度和能源消耗，以实现经济和合理的状态。

2.2.1 矿山工作场景简介

露天矿山运输是一个周期性的操作，包括装载、满载运输、卸载和空载运输，可以分为两种类型：满载上坡和满载下坡。如图 2-2 和 2-3 所示，道路坡度通常不超过 12%。由于独特的工作环境，本研究为长时间上坡和下坡工况建立了两种类型的固定驾驶路线。

这两条路径的总里程约为 4900 米，最大坡度为 12%，并且是往返操作。在第一种道路条件下，车辆以空载爬升到山顶，停下来装载，然后满载返回起点下坡。另一种情况是车辆以空载爬升到山顶，停下来卸载，然后满载返回起点下坡。矿区车辆的最大速度为 30km/h。详细的道路数据如表 2-1 所示。目标是找到一个理想的解决方案，使车辆在适当和经济的状态下运行。



图 2-2 露天矿真实场景

Fig.2-2 Real scene of open-pit mine

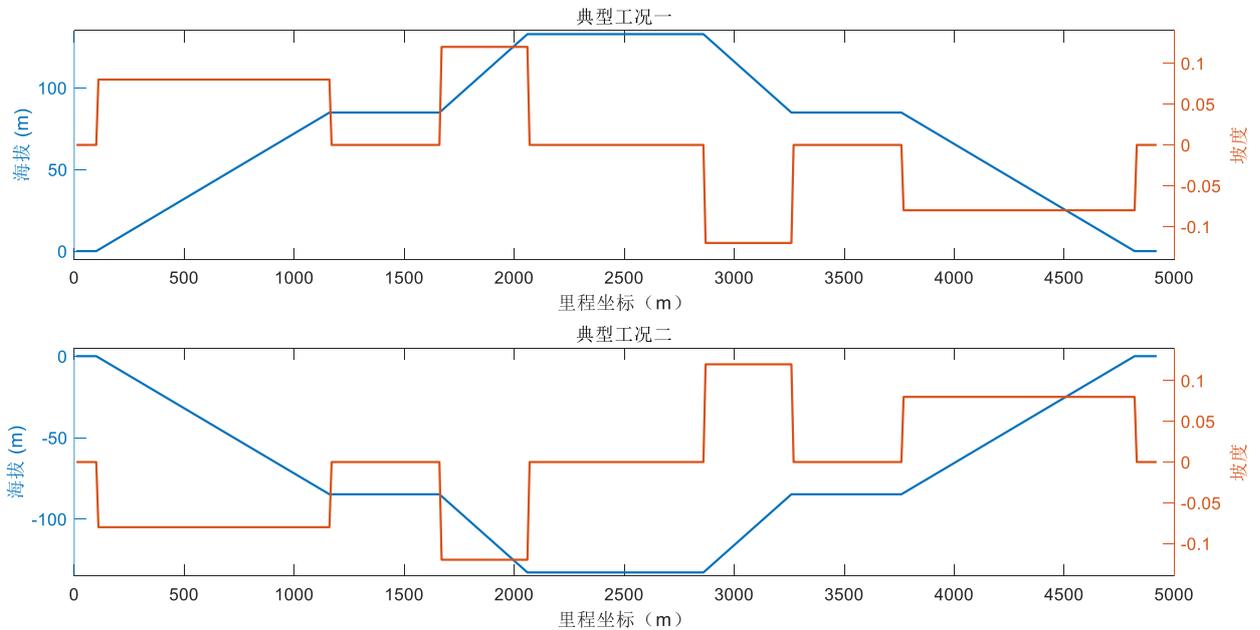


图 2-3 两种矿山道路

Fig.2-3 Two kinds of mine roads

表 2-1 两种矿山道路的具体参数

距离 [m]	线路 1: 坡度	线路 2: 坡度	装载状态
0-100	0	0	空载
100-1160	0.08	-0.08	空载
1160-1660	0	0	空载
1660-2060	0.12	-0.12	空载
2060-2460	0	0	空载
2460-2860	0	0	满载
2860-3260	-0.12	0.12	满载
3260-3760	0	0	满载
3760-4820	-0.08	0.08	满载
4820-4920	0	0	满载

2.2.2 车辆经济型行驶策略的优化方法

目前常用的经济性驾驶策略计算方法包括仿真法、最优控制和滚动时域优化。这些方法分别适用于模拟车辆行为、寻找性能极值和连续优化驾驶策略。仿真法是基于实验或训练的模型化方法，可用于理解问题和探索解决方案，但需要大量实验优化参数。最优控制旨在寻找控制使系统性能达到极值的方案^[52]，在车辆经济性驾驶策略中通常以油耗或电量消耗为性能指标。动态规划是由美国学者 R.贝尔曼在 1957 年提出的^[53]，而极大值原理则是前苏联学者 L.S.庞特里亚金在 1958 年提出的^[54]。这两个理论对最优控制理论的形成和发展产生了重大推动作用。在车辆经济性驾驶策略中，最优控制通常以发动机油耗、电量消耗等作为性能指标。有的学者会参考车辆的舒适性和车辆本身的寿命损耗作为性能指标 $J(x, u)$ 。以车辆动力学模型建立状态空间方程 $f(x, u, t)$ ，动力学模型的建立的方法会在后文中展开介绍，满足车辆性能、道路坡度和环境的约束 $C(x, u, t) = 0$ ，简记为式 2-1:

$$\begin{aligned} & \min J(x, u) \\ & s. t.: \dot{x} = f(x, u, t) \\ & C(x, u, t) = 0 \end{aligned} \quad (2-1)$$

这些方法致力于寻找最佳的驾驶控制策略，涉及油门开度、发动机扭矩转速、电机扭矩转速、挡位、制动力等参数的优化，以在完成驾驶任务的同时实现经济性。核心在于构建和解决问题模型，需要将车辆、道路和交通等因素数学化建模。然而，由于涉及的模型通常是强非线性的，求解过程计算量大，复杂的模型会导致难以求解^[55]。因此，这些方法通常只适用于固定且相对封闭的系统，无法满足对动态环境的实时要求，仅能针对确定的工况进行离线计算。

滚动时域优化方法，也称为模型预测控制，采用滚动式有限时域优化策略，实现随时间推移的在线优化^[56,57]。它在每个采样时刻根据当前测量信息求解一个有限时域开环优化问题，并将得到的控制序列应用于被控对象。随后，根据新的测量值刷新优化问题并重新求解，实现持续的在线优化。Richalet 等在 1978 年提出了预测控制算法的三个要素：预测模型、滚动优化和反馈控制^[58]。Kamel 等则设计了非线性 MPC 控制器，用于在坡道工况下的节油巡航，以 10 秒的预测时域、0.1 秒的步长在线优化车辆速度轨迹^[59]。刘焕峰则通过将 MPC 的性能指标设为多目标，包括安全跟踪性、燃油经济性和舒适性，探讨了这三个目标之间的权重关系，以最大程度提升巡航车辆的燃油经济性^[60]。这种方法的优势在于其灵活性和适应性，能够应对实时交通环境中的不确定性因素^[61]。然而，其在实际交通应用中受到信息量和计算能力的限制，尤其是在面对复杂的交通场景时，实时在线计算的挑战较大。

2.2.3 优化算法：动态规划

动态规划（DP）是一种数值方法，用于解决优化问题。由美国数学家贝尔曼^[53]于 1951 年提出。其核心思想是将子问题的解存储起来，以便后续使用，如图 2-4 所示，通过递归地将复杂问题分解为简单子问题。

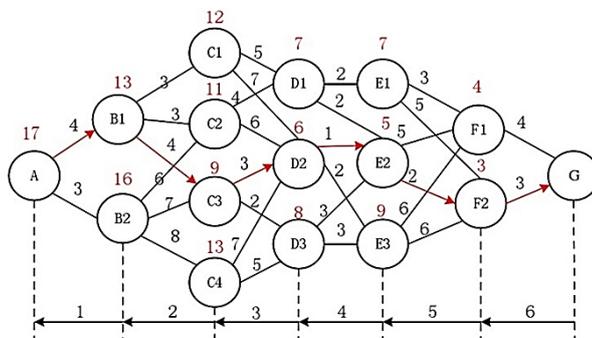


图 2-4 动态规划的基本概念

Fig.2-4 Basic concept of dynamic programming

首先，根据距离将速度优化问题划分为子问题。如果分段距离太小，将产生更多的子问题，并导致计算量过大。为了考虑计算时间和准确性之间的关系，通过最小速度变化将路径划分为多个小节。接下来，对速度进行离散化，如图 2-5 所示。每个路段的初始和最终速度可以自由组合。最后，可以计算出所有可能条件下该路段的加速度、时间和能量消耗，并同时得到目标函数的值。每个路段的最佳结果将被存储，并作为下一个路段的初始状态添加。在计算完所有路段后，选择最后一个路段结果的最小值作为线索，反向找出前面各路段的值。

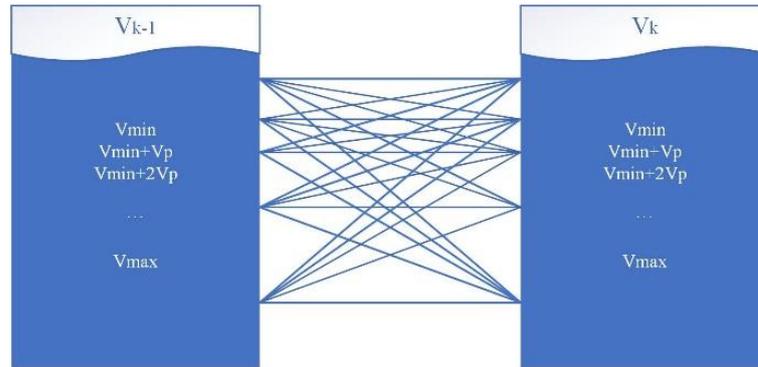


图 2-5 速度的离散化

Fig.2-5 Discretization of velocity

2.3 车辆模型概述

在本节中,我们建立了燃油车和电动车的动态模型,并在 MATLAB 中进行了模拟。下面讨论了车辆组件的关键参数。

2.3.1 车辆动力学系统建模方法

基于车辆动力学模型,采用相应的模型化描述,开发建立整车中的各零部件模型是探索经济性驾驶策略的基础之一。在各个零件之间的工作状况的基础上,可以进行简单的控制设计,减少不必要的研究成本。

1) 发动机模型

发动机是一个复杂的非线性系统,受多种因素影响,如转速、节气门开度、点火提前角、冷却水温度等^[62]。发动机油耗是评价速度行驶策略经济性的关键指标,其喷油率由多种因素决定,如油门、温度、进气等^[63]。不同领域对发动机模型精度要求不同,在交通运输研究中,简化模型可满足要求,而在发动机电控中对模型精度要求更高^[63]。对于经济性驾驶策略研究,需要准确的油耗模型,但不需关注每次喷油量。目前研究主要集中在功率需求模型和基于万有特性的油耗模型。

a) 功率需求型模型。

该模型广泛应用于车辆控制和系统优化中。它根据车速、加速度和道路坡度等信息计算出发动机的瞬时需求功率,并建立了功率与喷油量的函数关系,从而得到行驶中的油耗。典型的功率需求模型包括 SIDRA、VSP 和 Biggs 模型^[64, 65]。这些模型简单易用,适合车辆控制器设计和系统优化,但未考虑变速器挡位对发动机工作状态的影响,导致在某些场景下精度较低。

b) 基于万有特性图的油耗模型。

为获取发动机的稳态输出特性，通常会使用发动机试验台架进行测试。这些测试涉及不同节气门开度和转速下的稳态输出转矩和燃油消耗率，并绘制发动机负荷曲线。利用插值方法拟合出燃油消耗率与发动机转速以及转矩的关系曲面，通常称为万有特性图。这种方法简单易用，通过查询或插值估计瞬时喷油率，计算速度快，且精度较高^[4]，但忽略了动态影响。Pelkmans 等人进行的测试表明，发动机动态工况与稳态工况的油耗误差约为 4% 以内^[66]。虽然存在一定误差，但是在经济性驾驶、车辆控制器设计等领域已成为主要模型^[6, 67-69]。

2) 电机模型

一个完整的电机模型包括很多电机参数，比如磁通量、定子绕阻、机对数等，并且精确的电机模型要求计算步长很短，会严重影响仿真运算的速度；另外电机是一个十分复杂物理过程，电机内部的结构参数的获取较为困难。

建立电机动态数学模型通常需要多种电机参数，常见的建模方法包括数学建模法、解析建模法和有限元建模等^[70]。传统数学方法建立的模型过于理想化，忽略了许多非线性因素，电机内部参数常被视为常数。为提高模型精度，有学者考虑引入损耗模型^[71, 72]，但电机内部参数变化复杂，需要其他方法获取。解析法可描述电机参数和特性之间的关系，但只适用于特殊磁场边界情况，且通常进行了简化^[73, 74]。有限元法在考虑非线性因素时精度较高^[75]，能够获取电机内部参数的变化特性，因此在电机设计软件中得到广泛应用。

本文主要关注转矩、转速和效率映射关系，寻找经济的工作状态。所以电机模型可以适当的简化，电机内部的损耗以及电磁学相关的问题不是本研究的重点，无需考虑转矩产生的原理以及内部的复杂原理，仅需要得到扭矩、转速和效率的映射关系即可。

在电动车辆驾驶策略的优化领域内，电机的模型和发动机模型类似，一般根据大量的电机试验台架数据绘制发电机效率的 MAP 图、驱动效率 MAP 图。一种是直接通过扭矩、转速二维 MAP 图查表、插值计算得到对应的效率^[76]，该方法使用简单、运行速度较快，但不能满足绝大多数的工作场景，不能应对复杂多变的工况；另一种则根据已有的图表数据建立转速、转矩和电机效率的高阶函数映射关系。文献^[25]中对转速、转矩和效率之间的关系使用 2 维 5 次多项式拟合的方式表达电机内部的损耗，该方法较为简单能够满足基本的研究需求。

3) 电池模型

电池是新能源车辆中重要的能量存储系统，不仅对于电动车量承担着重要的责任，而且对于混合动力车辆也是常用的能量存储系统，具有储存能量、释放电能、制动回收

等功能，电池的等效模型需要较好的反应电池的动态特性，能够满足多变的工作状态。

电池模型在电动车辆和混合动力车辆研究中至关重要，决定了车辆的性能、续航里程和使用寿命。目前，研究人员主要采用三种不同的建模方式：电化学模型、数据驱动模型和等效电路模型^[77]。电化学模型基于电化学反应动力学和传质传热理论，描述锂电池的充放电行为^[78-80]。目前电化学模型包括准二维模型、单粒子模型和简化的准二维模型等^[81,82]，但计算量大，应用受限。数据驱动模型利用实验数据训练机器学习算法，模拟锂电池的动态特性，但精度受电池数据和机器学习方法影响^[77]。等效电路模型使用电器元件模拟电池充放电过程^[83]，结构简单且精度相对较高，被广泛采用。研究比较了不同模型，发现在 DST 测试下，DP 模型性能最优^[84]。

考虑电池模型在实时计算中的计算量问题，并且主要针对矿用自卸车的经济性驾驶策略的速度优化设计，车辆驾驶策略主要使用等效电路模型展开研究。

4) 动力学系统模型

车辆动力学模型用于描述车辆的运动和受力情况。在经济性驾驶策略中，安全行驶是首要考虑因素。驾驶策略主要关注车辆纵向动力学特性，以加速度和功率需求为基础估计能耗，并优化驾驶速度。与横向的动力学特性不同，驾驶策略的速度优化不考虑车辆轮胎的动力学特性、传动系统的高阶特性、转向性能等^[6, 85-87]。动力学模型的仿真可分为前向和后向两种^[88]。前向仿真适用于系统零部件的动态响应仿真，但较为复杂且速度较慢；后向仿真则更适合车辆能量管理策略的评估，并具有更快的反应速度。在研究经济性驾驶策略时，通常选择相对简单的后向仿真模型。

2.3.2 动力系统设置

我们采用了两种车辆研究，它们的载荷都为 45 吨。一种是燃油车，具有发动机、变速器和液压制动系统。变速器采用了六速 AMT。燃油车的关键参数如表 2-2 所示。另一种是电动车，其动力系统由电池、电动机、变速器和制动系统组成。变速器有 4 个档位。电动机在制动时作为发电机工作进行能量回收。电动车的详细信息如表 2-3 所示。

然而，在制动方面，燃油车和电动车存在差异。对于燃油车，它通过机械液压制动系统进行制动。而对于电动车，在制动程度轻或中等时，车辆处于再生制动状态，电动机将能量回收并存储在电池中。然而，当制动程度较高并超过能量回收能力时，机械制动系统被激活，与发电机一起工作，以确保车辆的安全。

需要说明的是，该两款车型采用的变速器均采用自动变速，不对挡位进行优化。其换挡模式均使得电机和发动机工作在效率最高的区间，自适应换挡。

表 2-2 燃油车的基本参数

Tab.2-2 Basic parameters of fuel vehicle

类别	数值	其他
自重	34000kg	
载重	45000kg	
发动机	389kw	最大功率
变速箱	6 挡	4 2 2.68 2.01 1.35 1 0.67
传动总减速比	17.83	
轮胎有效半径	0.873m	21.00-35

表 2-3 电动车的基本参数

Tab.2-3 Basic parameters of electric vehicle

类别	数值	其他
自重	37000kg	
载重	45000kg	
发动机	-	
变速器	4 挡	8.39 4.34 2.27 1
传动总减速比	17.83	
轮胎有效半径	0.873m	21.00-35
电动机	450Kw, 2400Nm	
电池	228Ah	3 并 192 串

2.3.3 车辆纵向动力学模型

本文研究的采矿车辆具有较大的载荷能力、低速度和简单的工况。忽略了车辆的行驶稳定性和舒适性。因此，车辆的纵向动力学模型如图 2-6 所示。

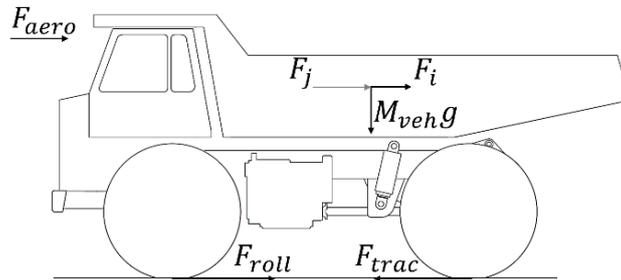


图 2-6 车辆纵向动力学模型

Fig.2-6 The vehicle longitudinal dynamics model

F_i 为坡道阻力， F_j 为加速阻力， F_{aero} 为空气阻力， F_{roll} 为滚动阻力，滚动阻力系数的数值取 0.02。在车辆的各项受力中，凡是和车辆速度行驶方向一致的，均取正值，驱动力 F_{trac} 的最终条件为：

$$F_{trac} \geq F_i + F_j + F_{aero} + F_{roll} \quad (2-2)$$

只有在满足此条件下才能正常行驶。

- 1) $F_{trac} > 0$ 时，车辆处于驱动模式，如果车辆的各个阻力和逐渐变大，使得车辆阻力和超出驱动力时，车辆逐渐减速。
- 2) $F_{trac} = 0$ 和 $F_{trac} < 0$ 时，车辆处于滑行、制动模式，车辆在此时会减速，此时整车需求功率为负值。对于电车来讲，此时电机开始发电，为电池补充电量。

2.3.4 发动机模型

发动机的机械限制包括扭矩和转速，并且燃油消耗可用这两个因素的函数来表示。如方程(2-3)所示， T_{ICE} 和 n_{ICE} 其中分别发动机扭矩和转速。发动机的外特性曲线和单位功率燃油消耗图如图 2-7 所示。

$$Fuel = f(T_{ICE}, n_{ICE}) \quad (2-3)$$

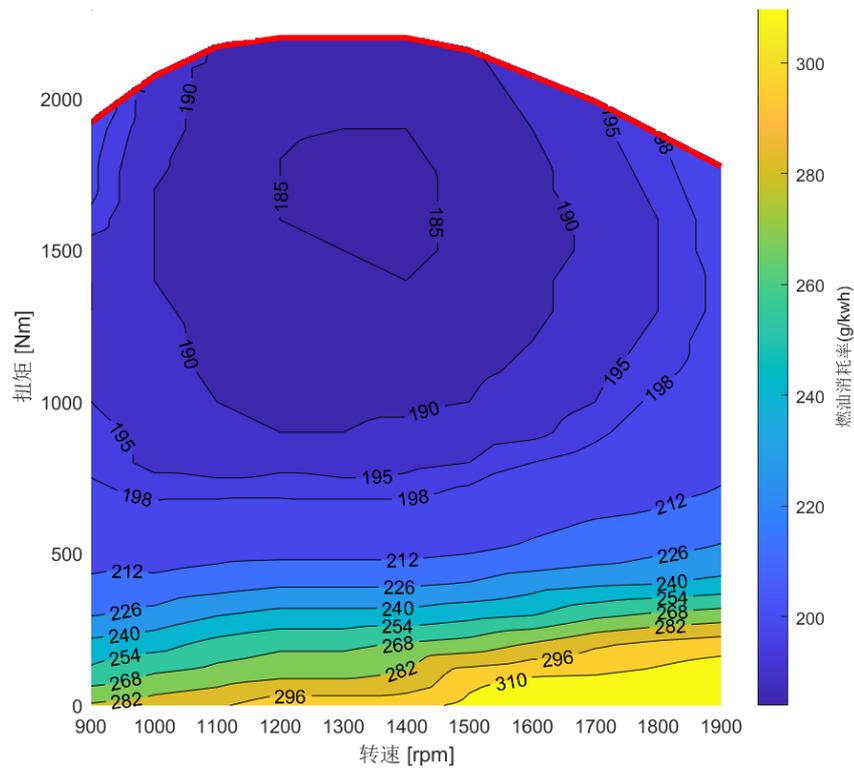


图 2-7 发动机外特性曲线和燃油消耗曲线

Fig.2-7 Engine external characteristic curve and fuel consumption curve

2.3.5 电机模型

电动机可以作为发电机，在车辆制动时回收能量。电动机从电池中提取的电功率可以建模为：

$$P_{Gen/Tra} = \eta_{Mot} T_{Gen/Tra} n_{Gen/Tra} \quad (2-4)$$

其中 P_{Tra} 是牵引电机的功率， P_{Gen} 是发电机工作时的功率。 η_{Mot} 是电动机的效率。 $T_{Gen/Tra}$ 和 $n_{Gen/Tra}$ 分别是电动机的扭矩和转速。电机的效率根据其转速和扭矩进行动态调整，如方程 (2-5) 所示。电机的外特性曲线和效率图如图 2-8 所示。

$$\eta_{Mot} = f(T_{Gen/Tra}, n_{Gen/Tra}) \quad (2-5)$$

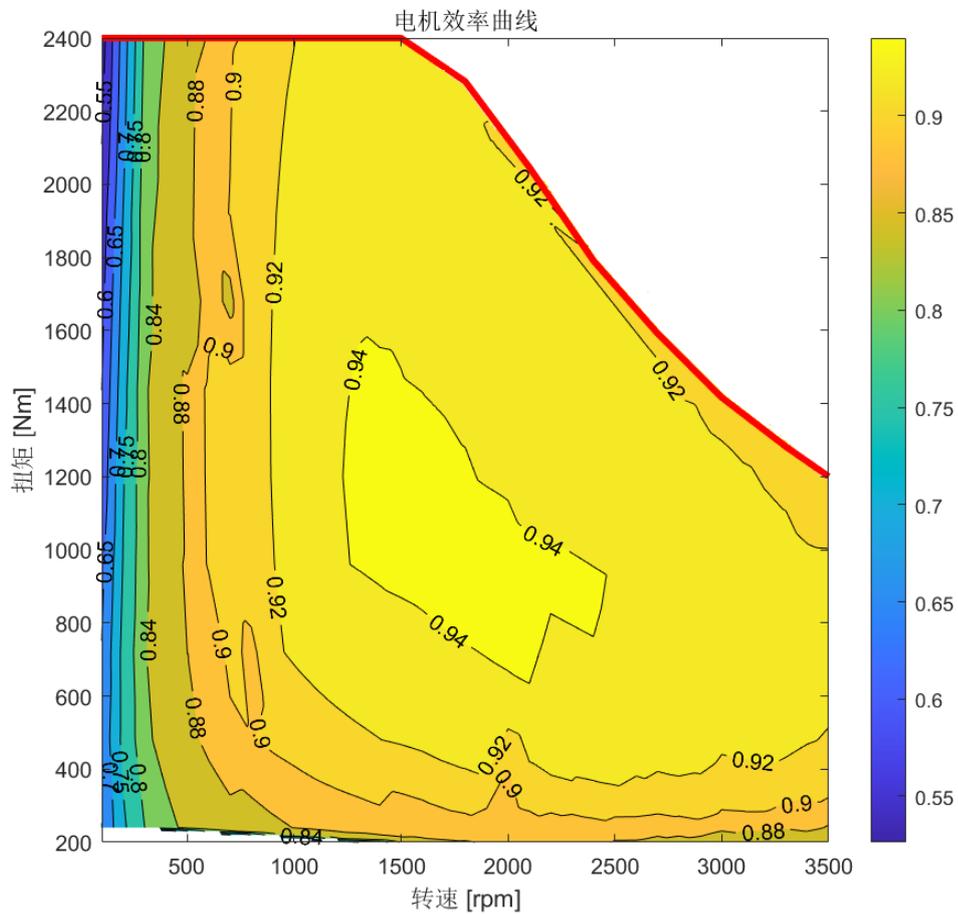


图 2-8 电机效率曲线

Fig.2-8 The motor efficiency curve

2.4 优化算法设计与分析

此外，速度轨迹优化涉及许多方面，如行驶时间、能量消耗和电池损耗。它们在交通系统中相互关联，尤其重要。Xu 等人^[89]提出了一种双层速度优化方法，利用 Dijkstra

算法和实时计算考虑了交通信号信息，以减少燃料消耗和行程时间。Zhu 等人^[90]提出了一种考虑交通信息的电动车节能路径规划方法。实验结果表明，与基于距离的路径相比，它具有更低的能量消耗和较短的行程时间；与基于时间的路径相比，它具有更低的能量消耗和较长的行程时间。Patterson 等人^[91]证明了等待时间和卡车选择之间的相互关系可以最小化矿山的能量消耗，并提出了一种原创的混合整数线性规划模型，调度运输活动以减少卡车和铲车的能量消耗。Dehkordi 等人^[92]开发了一种多目标优化方法，通过基于网络级别的安全指标，同时最小化燃料消耗、行程时间和安全风险，改善燃料效率并减少安全风险。然后，Liu 等人^[93]提出了一种基于拟谱方法的多目标再生制动控制策略，带有终端约束。该控制策略通过优化速度和制动扭矩分配，同时减少能量消耗和电池容量损失，有助于找到理想的平衡解。

在本节中，首先在采矿场景下进行考虑能量消耗的单目标仿真。然后，将建立多目标优化数学模型，以“时间消耗”、“能量消耗”和“电池寿命”作为优化目标。因此，采用多目标方法对速度进行优化。

2.4.1 单一目标优化

运输系统成本的构成非常复杂。主要包括运输时间、能量消耗、磨损和设备折旧。考虑能量消耗的单目标函数如下所示：

$$J = J_{en} = \int_{s_0}^{s_f} P_{out} \frac{ds}{u} \quad (2-6)$$

其中， P_{out} 是车辆的输出功率， s_f 是总行驶距离， u 代表车辆速度。除此之外，燃油车和电动车的动力传动组件是不同的。 T_{ICE} 、 T_{Mot} 分别是发动机和电动机的输出扭矩。 n_{ICE} 、 n_{Mot} 表示发动机和电动机的转速。 i_g 是传动比。燃油车和电动车的物理约束如方程(2-7)和(2-8)所示。

$$\begin{cases} T_{ICE_min} < T_{ICE} < T_{ICE_max} \\ n_{ICE_min} < n_{ICE} < n_{ICE_max} \\ i_g \in \{i_{g1}, i_{g2}, i_{g3}, i_{g4}, i_{g5}, i_{g6}\} \end{cases} \quad (2-7)$$

$$\begin{cases} T_{Mot_min} < T_{Mot} < T_{Mot_max} \\ n_{Mot_min} < n_{Mot} < n_{Mot_max} \\ SOC_{min} < SOC < SOC_{max} \\ i_g \in \{i_{g1}, i_{g2}, i_{g3}, i_{g4}\} \end{cases} \quad (2-8)$$

$$\begin{aligned}
 J &= J_{Fuel} = \int_{s_0}^{s_f} Fuel \cdot P_{out} \frac{ds}{u} \\
 \text{s. t.} \\
 \left\{ \begin{aligned}
 \frac{du}{ds} &= \frac{F_{trac} - F_i - F_j - F_{aero} - F_{roll}}{mu} \\
 a &= \frac{udu}{ds} \\
 T_{ICE} &= \frac{F_{trac} R_{tire}}{i_g i_f} \\
 T_{ICE_min} &< T_{ICE} < T_{ICE_max} \\
 n_{ICE} &= \frac{u i_g i_f}{2\pi R_{tire}} \\
 n_{ICE_min} &< n_{ICE} < n_{ICE_max} \\
 i_g &\in \{i_{g1}, i_{g2}, i_{g3}, i_{g4}, i_{g5}, i_{g6}\} \\
 a_{min} &< a < a_{max} \\
 0 &< u < u_{max}
 \end{aligned} \right. \quad (2-9)
 \end{aligned}$$

$$\begin{aligned}
 J &= J_{Elec} = \int_{s_0}^{s_f} \frac{U_{0c} - \sqrt{U_{0c}^2 - 4P_{bat} R_{eq}}}{2R_{eq} Q_{max}} \frac{ds}{u} \\
 \text{s. t.} \\
 \left\{ \begin{aligned}
 \frac{du}{ds} &= \frac{F_{trac} - F_i - F_j - F_{aero} - F_{roll}}{mu} \\
 a &= \frac{udu}{ds} \\
 T_{Mot} &= \frac{F_{trac} R_{tire}}{i_g i_f} \\
 T_{Mot_min} &< T_{Mot} < T_{Mot_max} \\
 n_{Mot} &= \frac{u i_g i_f}{2\pi R_{tire}} \\
 n_{Mot_min} &< n_{Mot} < n_{Mot_max} \\
 P_{bat} &= \frac{T_{Mot} n_{Mot}}{\eta_{Mot}} \\
 i_g &\in \{i_{g1}, i_{g2}, i_{g3}, i_{g4}\} \\
 a_{min} &< a < a_{max} \\
 0 &< u < u_{max}
 \end{aligned} \right. \quad (2-10)
 \end{aligned}$$

燃料消耗和 SOC 变化是运输成本的重要因素，可以根据发动机、电动机和电池的特性通过功率输出来计算。因此，燃油车和电动车的单目标最优控制问题可以表示为方程 (2-9) 和 (2-10)。其中 m 是车辆的质量， R_{tire} 是轮胎的半径， i_f 是最终传动比， a 代

表加速度。在这种采矿条件下，最小和最大加速度分别为 $-0.7m/s^2$ 和 $0.7m/s^2$ ，最大速度为 $30km/h$ 。

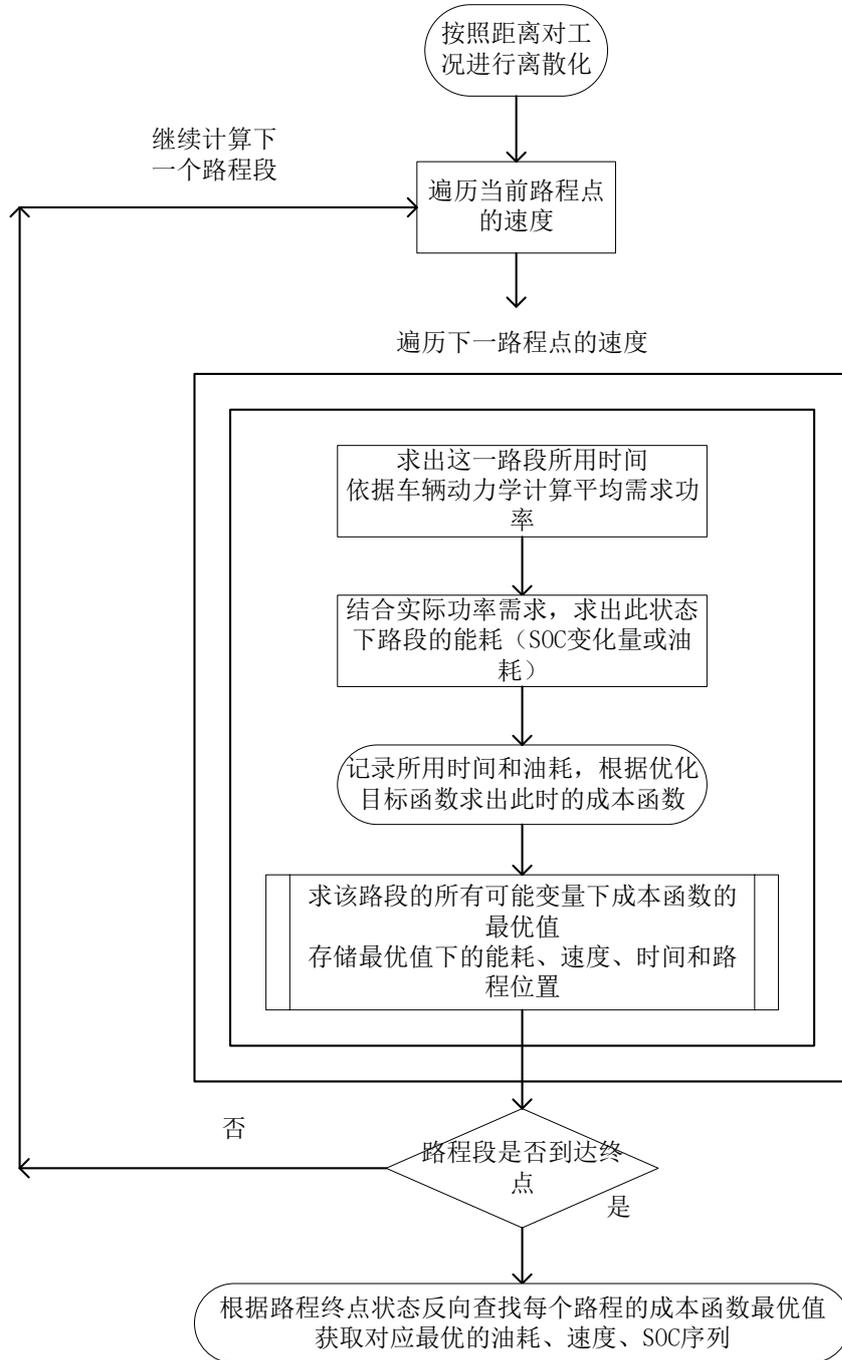


图 2-9 动态规划优化流程

Fig.2-9 The process of dynamic programming optimization

使用动态规划方法来解决这个问题。最小速度变化 v_p 被设定为 $5km/h$ 。如图 2-9 所示，每个最终速度的最佳结果以及详细的速度、时间和能量消耗将被存储并作为下一部分的初始状态添加进去。在计算完所有的部分之后，选择最后一个部分结果的最小值作为线索，以逆向找出前面部分的值。然后，从所有的优化的部分成功地提取出速度轨迹。

图 2-10 和图 2-11 显示了燃油车在燃料消耗目标下的速度轨迹。对于路线 1，时间和燃料消耗分别为 924.6s 和 3660.1g；对于路线 2，时间和燃料消耗分别为 1236.7s 和 7587.4g。这两条速度轨迹首先加速到达最大速度，然后经常采用 PnG 策略，使发动机在高效工作条件下运行。

电动车在电能消耗目标下优化的速度轨迹也采用了 PnG 策略，对于两条路线都使电动机尽可能地在高效范围内工作，如图 2-12 和图 2-13 所示。通过频繁的滑行能量回收，实现了电能的减少。这与通过能量消耗目标优化的燃油车的轨迹类似。对于路线 1，电池 SOC 增加了 0.0679，但对于路线 2，电池 SOC 减少了 0.2205。完成路线 1 需要 1156.4s，完成路线 2 需要 1879.0s。对于路线 1 来说，下坡滑行时回收的能量足以弥补上坡行驶时的能量消耗，因此电池 SOC 最终增加。

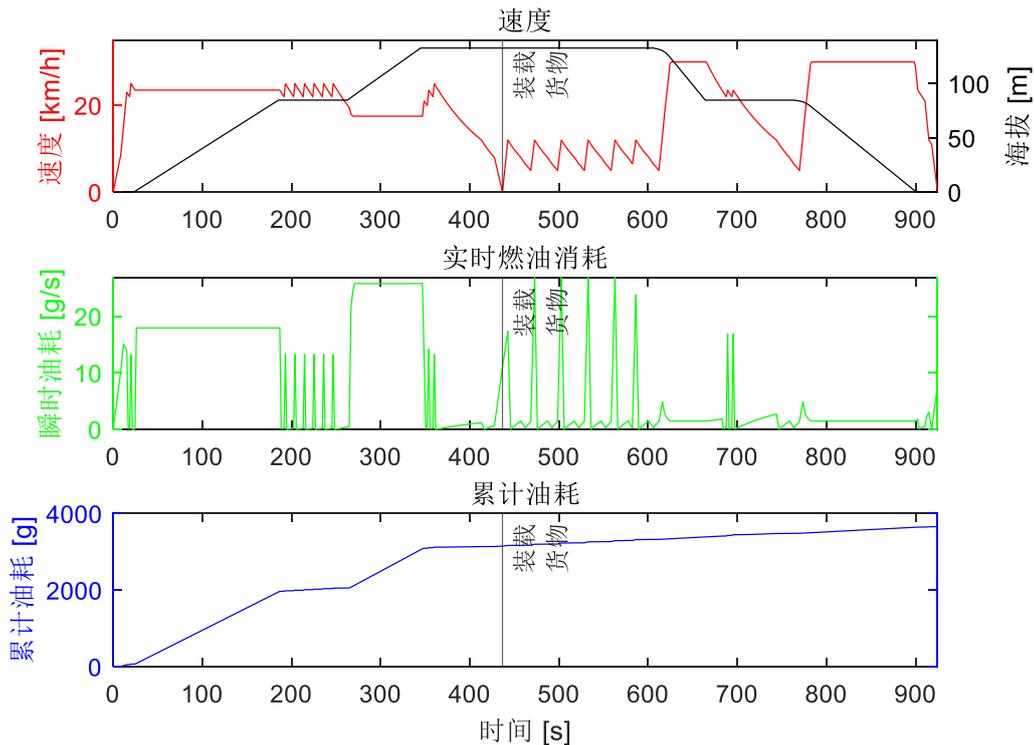


图 2-10 针对路况一，对燃油车进行能耗优化后的策略

Fig.2-10 Fuel vehicle driving strategy after energy consumption optimization for road condition 1.

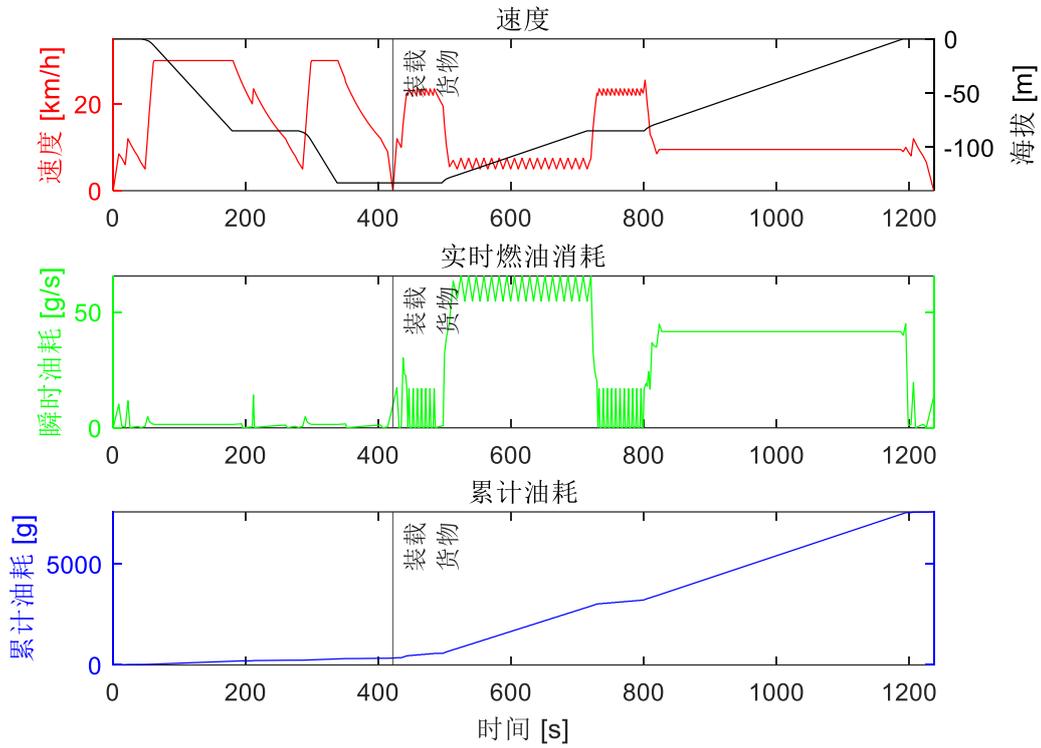


图 2-11 针对路况二，对燃油车进行能耗优化后的策略

Fig.2-11 Fuel vehicle driving strategy after energy consumption optimization for road condition 2.

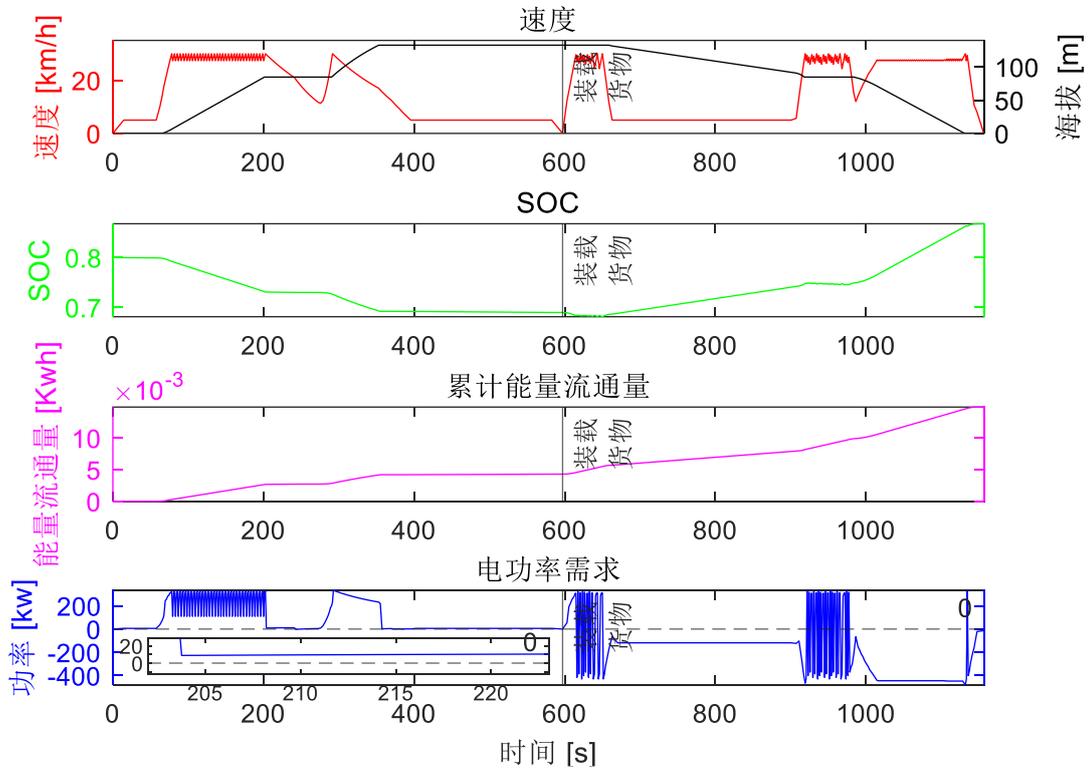


图 2-12 针对路况一，对电动车进行能耗优化后的策略

Fig.2-12 Electric vehicle driving strategy after energy consumption optimization for road condition 1.

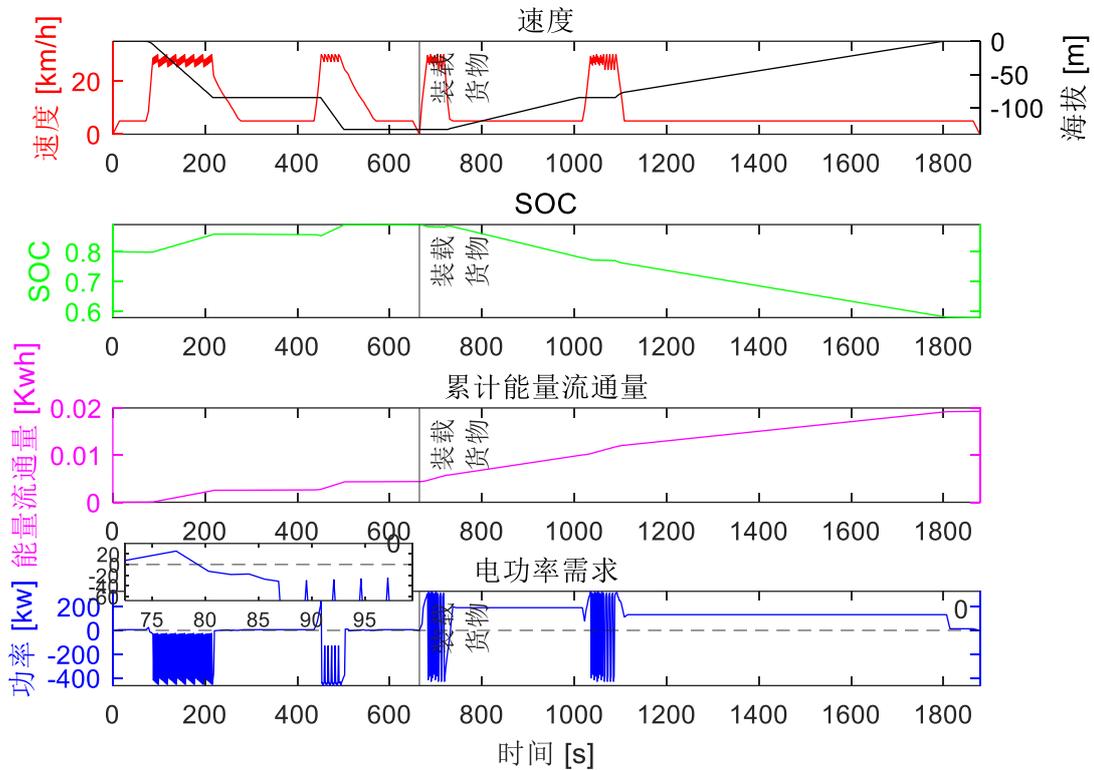


图 2-13 针对路况二，对电动车进行能耗优化后的策略

Fig.2-13 Electric vehicle driving strategy after energy consumption optimization for road condition 2.

2.4.2 多目标优化

不同工况下，车辆具有不同的运输成本，这些成本主要包括运输时间、能量消耗、磨损和设备折旧。电池的功耗会导致电池寿命的减少。因此，电动车的能量消耗目标必须平衡电池寿命损失和功耗。这是一个帕累托最优问题。基于单目标优化的结果，提出的问题的目标是在终端制动距离、终端速度以及动力传动组件的物理约束条件下同时最小化能量耗散、行驶时间和电池容量损失。以下构建了以“时间消耗”、“能量消耗”和“电池寿命”作为多目标最优化数学模型。多目标优化框架如图 2-14 所示。

本文提出的多目标函数采用线性形式，并考虑了行驶时间、能量消耗和电池健康状况，归一化的多目标函数如下所示：

$$J = \omega_T \frac{J_T}{J_{Tmin}} + \omega_{E_{health}} \frac{J_{E_{health}}}{J_{E_{hmin}}} + \omega_{en} \frac{J_{en}}{J_{enmin}} \quad (2-11)$$

其中 J_T , $J_{E_{health}}$ 和 J_{en} 分别表示目标函数中的行驶时间、电池损失和能量消耗的代价，

ω_T , $\omega_{E_{health}}$ 和 ω_{en} 表示它们的权重。如果是燃油车, 可以将 $\omega_{E_{health}}$ 设置为 0。如第 3.4.1 节所讨论的, J_{Tmin} , $J_{E_{hmin}}$ 和 J_{enmin} 是参考值, 它们是通过单目标优化得到的最小成本。 J_{en} 的函数形式取决于车辆类型, 如方程 2-9 和 2-10 所示。对于燃油车, J_{en} 表示燃料消耗。然而, 对于电动车, 它表示 SOC 的变化。

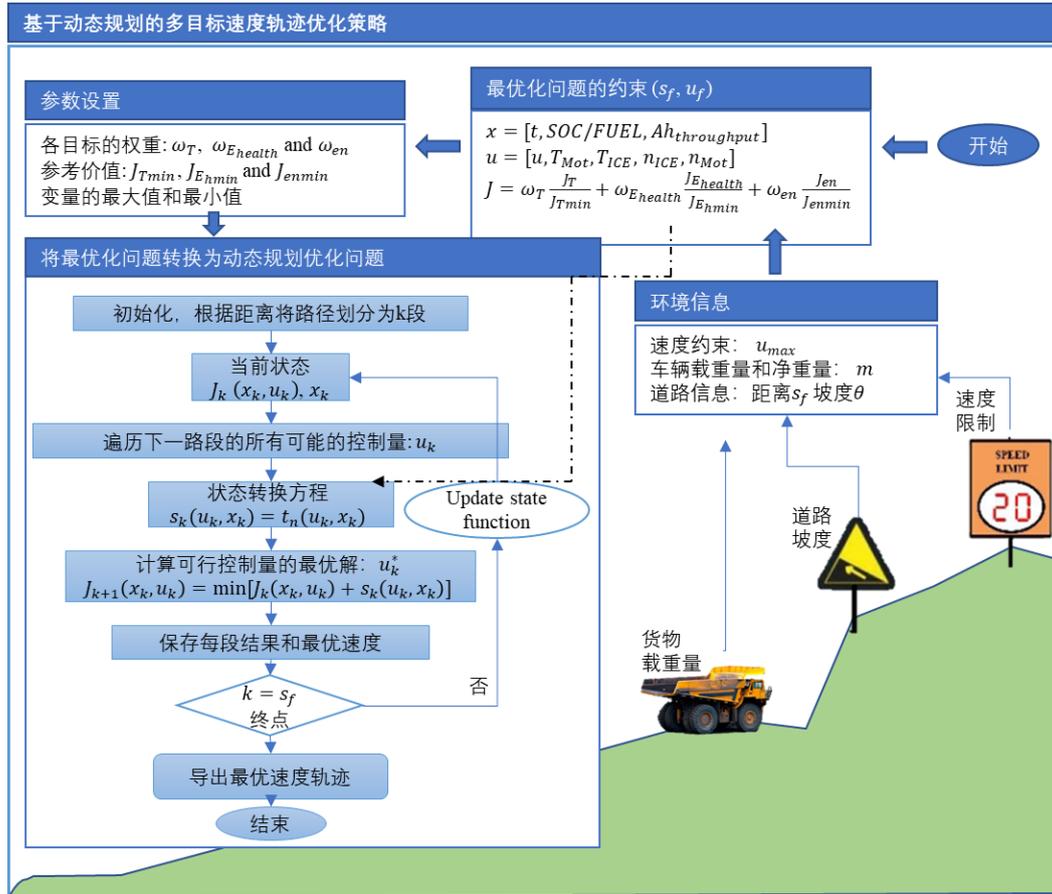


图 2-14 多目标优化速度策略的框架结构

Fig.2-14 Frame structure of multi-objective speed optimization strategy.

J_T 代表行驶时间的目标函数, 其计算公式如 2-12 所示。当车辆行驶轨迹所消耗的时间最少时, 即 $J_T = J_{Tmin}$ 。此时行驶时间的目标函数值为 1。而在其他情况下, 时间消耗 J_T 大于 J_{Tmin} 。行驶时间的目标函数大于 1。因此, $\frac{J_T}{J_{Tmin}}$ 的值越小, 越接近 1, 行驶时间就越短, 速度轨迹就越高效。类似地, 可以对车辆的能量消耗和电池损耗进行归一化, 得到相应的公式。 J_T , $J_{E_{health}}$ 和 J_{en} 的值越接近参考值, 速度轨迹策略的目标函数就越低。

$$J_T = \int_{s_0}^{s_f} \frac{ds}{u} \quad (2-12)$$

研究表明，电池损耗的衰减与累积安时流通量^[94, 95]有关。累积安时流通量可以衡量电池老化的程度。电池的累积安时流通量 $Ah_{throughput}$ 表达式如下所示。

$$Ah_{throughput} = n_b \int_{s_0}^{s_f} |i_b| \frac{ds}{u} \quad (2-13)$$

式中 n_b 代表电池组的串联线路的数量， i_b 代表电池的每个串联线路的电流。考虑到实际运算中电流的计算量比较大，在这里，电池组的累计能量流通量 $E_{throughput}$ 作为衡量电池损耗的依据，其单位为 Kwh。能量流通量和安时流通量的关系表示如下：

$$\begin{aligned} E_{throughput} &= \int_{s_0}^{s_f} |P_{bat}| \frac{ds}{u} = n_b \int_{s_0}^{s_f} |V_{oc} i_b| \frac{ds}{u} \\ &= n_b \overline{V_{oc}} \int_{s_0}^{s_f} |i_b| \frac{ds}{u} = \overline{V_{oc}} Ah_{throughput} \end{aligned} \quad (2-14)$$

电池寿命的目标函数如公式 2-15 所示。因此，燃油车和电动车在现有约束条件下的多目标最优控制问题可以表示为公式 2-16 和 2-17。

$$J_{Ehealth} = \int_{s_0}^{s_f} |P_{bat}| \frac{ds}{u} \quad (2-15)$$

$$\begin{aligned} J_{Fuel} &= \omega_T \frac{\int_{s_0}^{s_f} \frac{ds}{u}}{J_{Tmin}} + \omega_{en} \frac{\int_{s_0}^{s_f} Fuel \cdot P_{out} \frac{ds}{u}}{J_{enmin}} \\ &\quad s. t. \\ &\quad \left\{ \begin{array}{l} \frac{du}{ds} = \frac{F_{trac} - F_t - F_j - F_{aero} - F_{roll}}{mu} \\ a = \frac{udu}{ds} \\ T_{ICE} = \frac{F_{trac} R_{tire}}{i_g i_f} \\ T_{ICE_min} < T_{ICE} < T_{ICE_max} \\ n_{ICE} = \frac{u i_g i_f}{2\pi R_{tire}} \\ n_{ICE_min} < n_{ICE} < n_{ICE_max} \\ i_g \in \{i_{g1}, i_{g2}, i_{g3}, i_{g4}, i_{g5}, i_{g6}\} \\ a_{min} < a < a_{max} \\ 0 < u < u_{max} \end{array} \right. \end{aligned} \quad (2-16)$$

$$J_{Elec} = \omega_T \frac{\int_{s_0}^{s_f} \frac{ds}{u}}{J_{Tmin}} + \omega_{Ehealth} \frac{\int_{s_0}^{s_f} |P_{bat}| \frac{ds}{u}}{J_{Ehmin}} + \omega_{en} \frac{\int_{s_0}^{s_f} \frac{U_{OC} - \sqrt{U_{OC}^2 - 4P_{bat}Req} ds}{2ReqQ_{max}}}{J_{enmin}}$$

$$\left. \begin{aligned}
 & \text{s. t.} \\
 & \frac{du}{ds} = \frac{F_{trac} - F_t - F_j - F_{aero} - F_{roll}}{mu} \\
 & a = \frac{udu}{ds} \\
 & T_{Mot} = \frac{F_{trac} R_{tire}}{i_g i_f} \\
 & T_{Mot_min} < T_{Mot} < T_{Mot_max} \\
 & n_{Mot} = \frac{u i_g i_f}{2\pi R_{tire}} \\
 & n_{Mot_min} < n_{Mot} < n_{Mot_max} \\
 & P_{bat} = \frac{T_{Mot} n_{Mot}}{\eta_{Mot}} \\
 & i_g \in \{i_{g1}, i_{g2}, i_{g3}, i_{g4}\} \\
 & a_{min} < a < a_{max} \\
 & 0 < u < u_{max}
 \end{aligned} \right\} \quad (2-17)$$

在这种采矿场景中， a_{min} 和 a_{max} 分别为 $-0.7m/s^2$ 和 $0.7m/s^2$ ， u_{max} 为 30km/h。如第 2.4.1 节所讨论的，燃油车和电动车的 J_{Tmin} ， J_{Ehmin} 和 J_{enmin} 参考值是通过它们各自目标下的单目标优化结果得到的。Route 1 和 Route 2 的详细优化结果如表 2-4 和表 2-5 所示，其中下划线的数字可以作为参考值，结果的分析将在下一节中进行讨论。因此，对于燃油车的 Route 1， J_{Tmin} 和 J_{enmin} 分别设定为 628.6 和 3660.1。同时，对于 Route 2，分别设定为 825.6 和 7587.4。对于电动车，情况类似。Route 1 的 J_{Tmin} ， J_{Ehmin} ，和 J_{enmin} 分别为 635.5、28.7 和 0.0679。而 Route 2 的参考值分别为 1403.3、55.0 和 0.2205。

在 MATLAB 中进行了多次仿真，结果证明了多目标动态规划方法的有效性。 ω_T ， $\omega_{Ehealth}$ 和 ω_{en} 均设为 1，但燃油车的 $\omega_{Ehealth}$ 设为 0。对燃油车进行多目标动态规划优化结果如图 2-15 和图 2-16 所示。如图所示，速度轨迹在上坡情况下也显示出轻微的加速滑行 PnG 现象，以平衡时间和燃料消耗的优化目标。

对电动车进行多目标动态规划优化结果如图 2-17 和图 2-18 所示，在上坡和下坡时，优化后的速度轨迹仍然采用了加速滑行 PnG 轨迹，这在明显不同。它提供了速度轨迹与坡度之间的具体关系；随着坡度的变化，它将采用不同的加速滑行 PnG 策略频率。关于能量消耗、行驶时间和电池损耗的详细结果显示在表 2-4 和表 2-5 中。

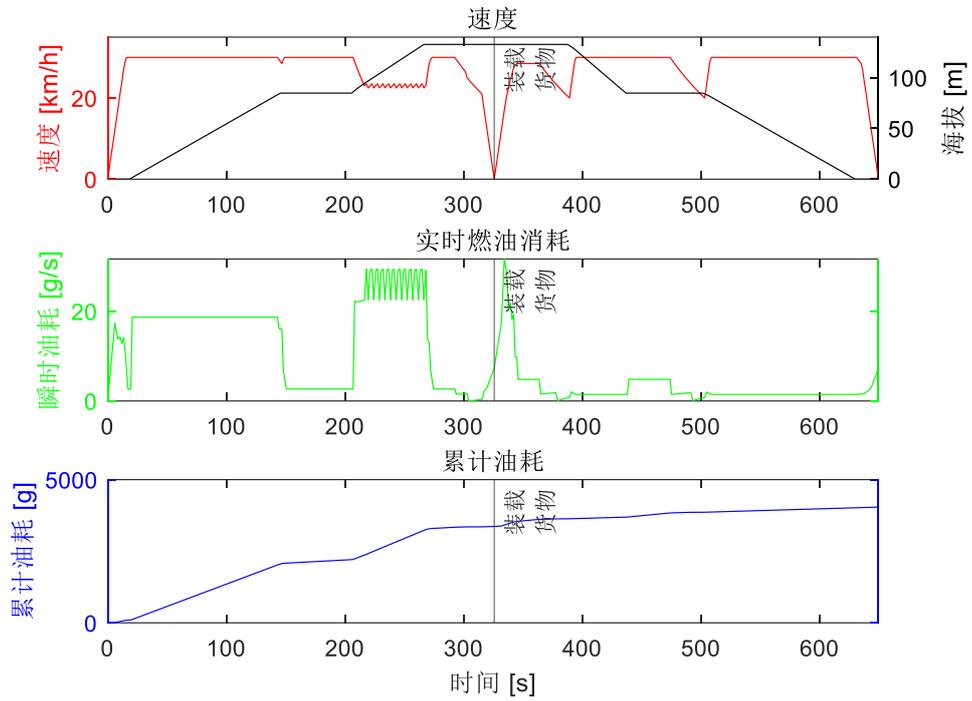


图 2-15 针对路况一，对燃油车进行多目标优化后的策略

Fig.2-15 Fuel vehicle driving strategy optimized by multi-objective for road condition 1.

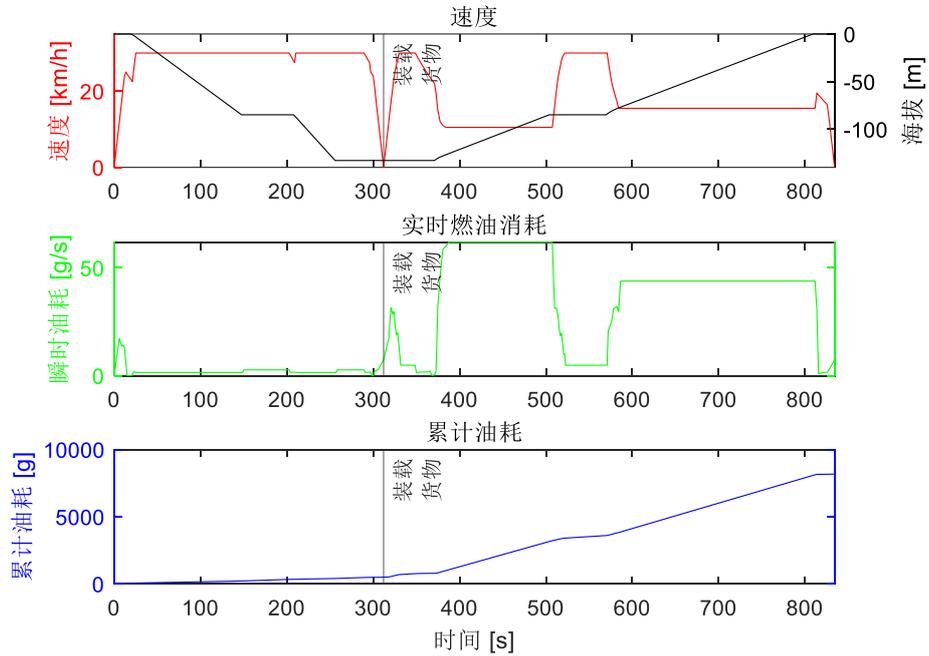


图 2-16 针对路况二，对燃油车进行多目标优化后的策略

Fig.2-16 Fuel vehicle driving strategy optimized by multi-objective for road condition 2.

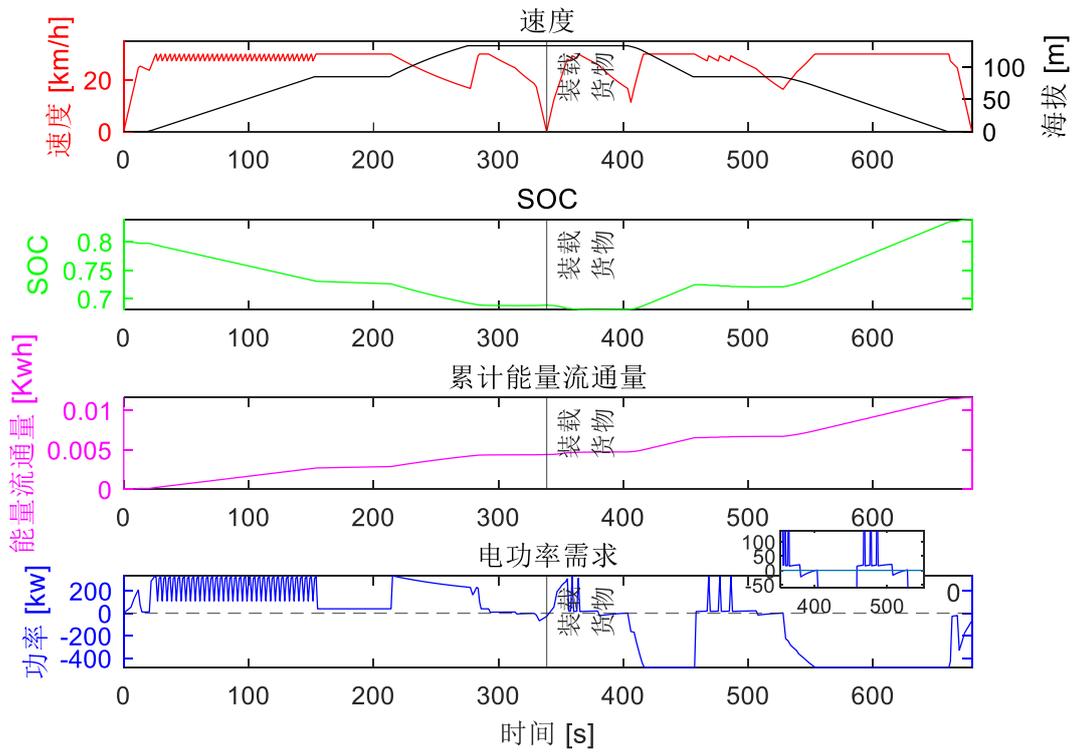


图 2-17 针对路况一，对电动车进行多目标优化后的策略

Fig.2-17 Electric vehicle driving strategy optimized by multi-objective for road condition 1.

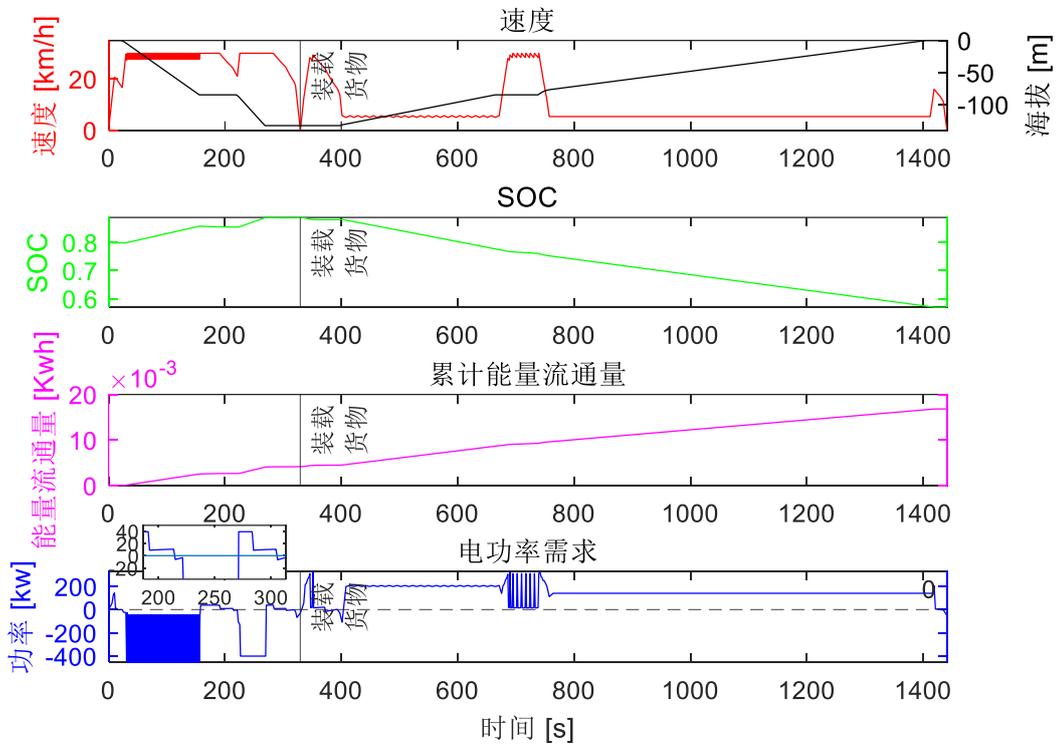


图 2-18 针对路况二，对电动车进行多目标优化后的策略

Fig.2-18 Electric vehicle driving strategy optimized by multi-objective for road condition 2.

2.5 仿真结果与分析

为了应用多目标公式，需要通过单目标优化获得一些参考值，即 J_{Tmin} ， $J_{E_{hmin}}$ 和 J_{enmin} 。因此，分别针对两种车辆模型，使用第 2.4.1 节中讨论的单目标方法进行了关于行驶时间和电池损耗目标的仿真，如“时间消耗”和“电池寿命”。本文考虑了一个采矿道路场景，其约束条件与第 2.4.1 节中描述的条件相同。通过“时间消耗”和“电池寿命”进行优化的车辆的速度轨迹如图 2-19 至图 2-24 所示。

图 2-19 到图 2-22 的优化目标是关于时间消耗。可以发现，当时间消耗被设置为单一优化目标时，两种车辆的速度轨迹趋向于达到最大速度，以尽快通过整个路线并在最小的波动下减少行驶时间。发动机或电动机始终在外特性曲线允许的最大扭矩下运行。然而，在上坡时，为了保护发动机和电动机免受需求功率超过功率限制的影响，会减小速度。图 2-23 和图 2-24 显示，与其他结果相比，以累计能量流通量为优化目标时，优化后的累积能量流明显减少，这可以延长电池的寿命。该优化的速度策略也采用了加速滑行 PnG 行为。然而，它与其他优化结果不同。该加速滑行轨迹旨在尽可能使电池保持非工作状态，既不充电也不放电。

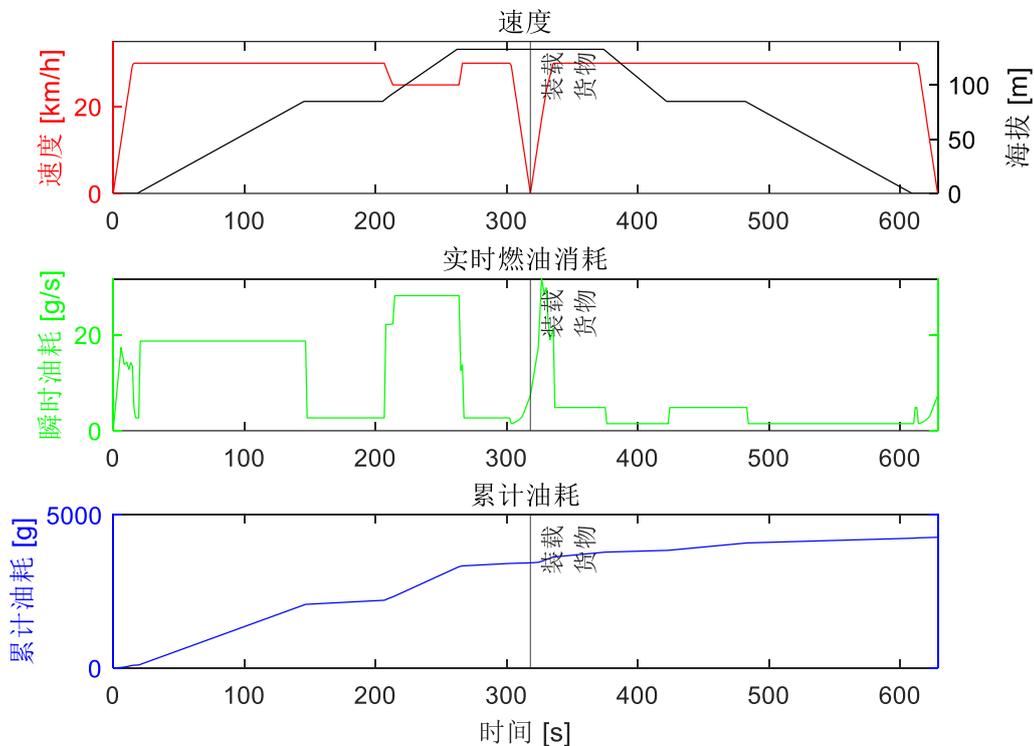


图 2-19 针对路况一，对燃油车进行行驶时间优化后的策略

Fig.2-19 Fuel vehicle driving strategy after travel time optimization for road condition 1.

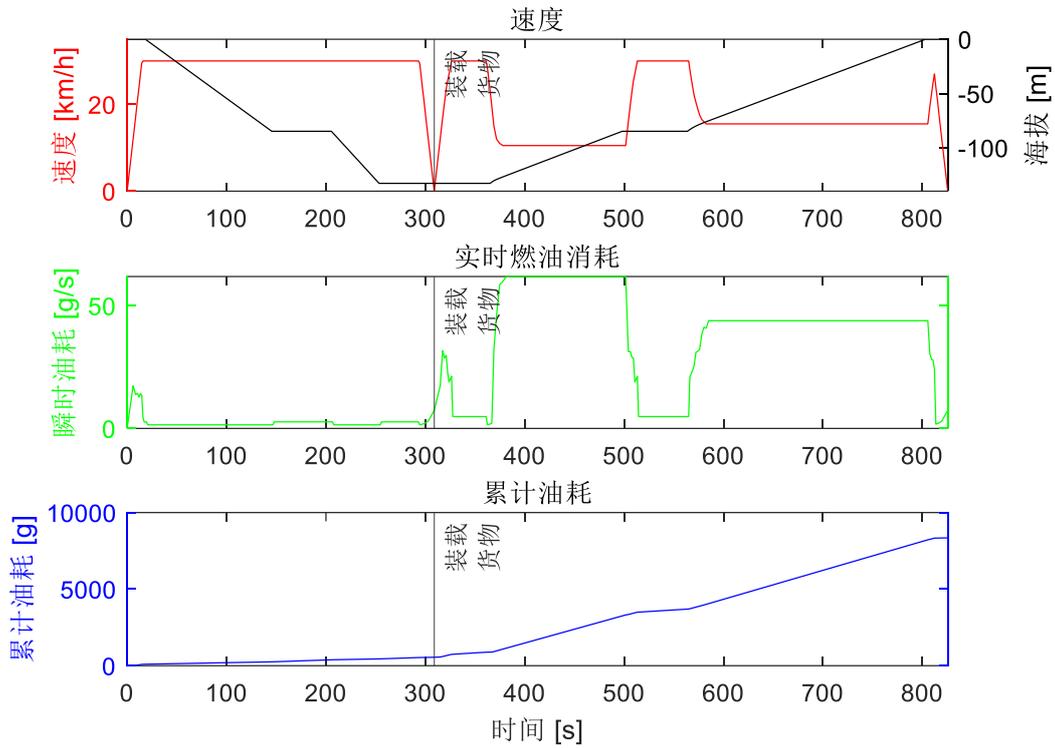


图 2-20 针对路况二，对燃油车进行行驶时间优化后的策略

Fig.2-20 Fuel vehicle driving strategy after travel time optimization for road condition 2.

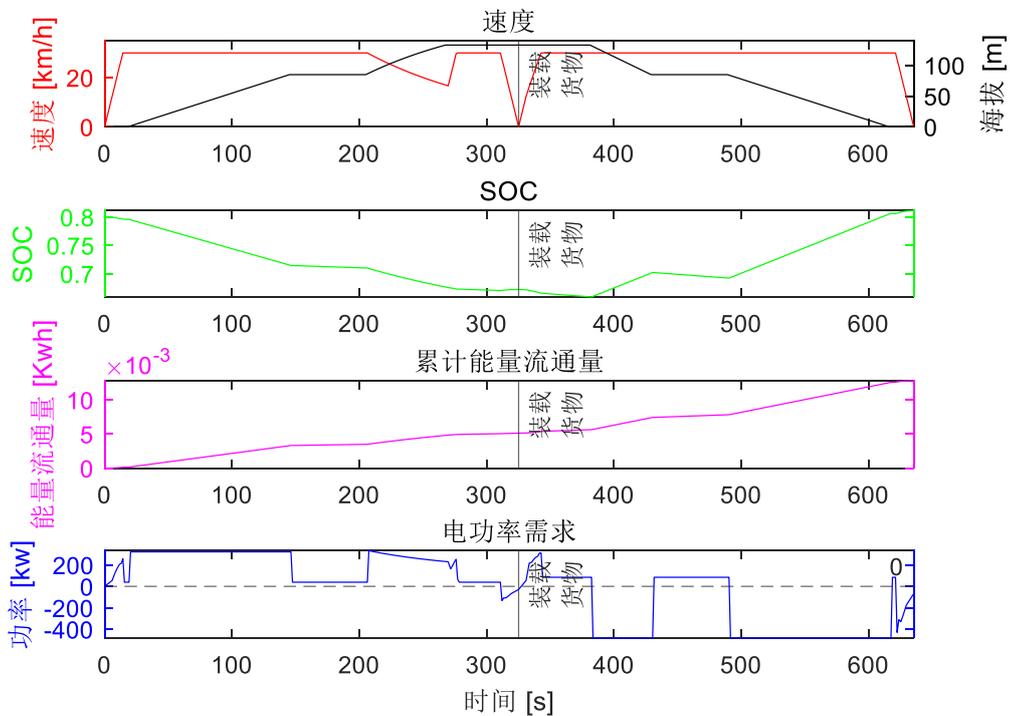


图 2-21 针对路况一，对电动车进行行驶时间优化后的策略

Fig.2-21 Electric vehicle driving strategy after travel time optimization for road condition 1.

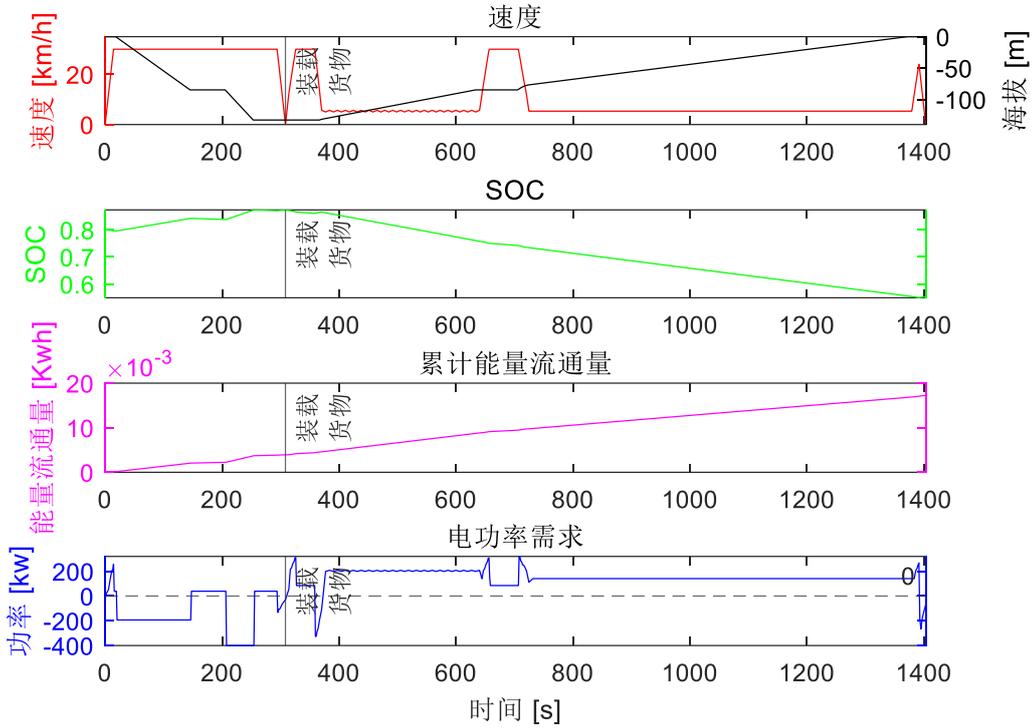


图 2-22 针对路况二，对电动车进行行驶时间优化后的策略

Fig.2-22 Electric vehicle driving strategy after travel time optimization for road condition 2.

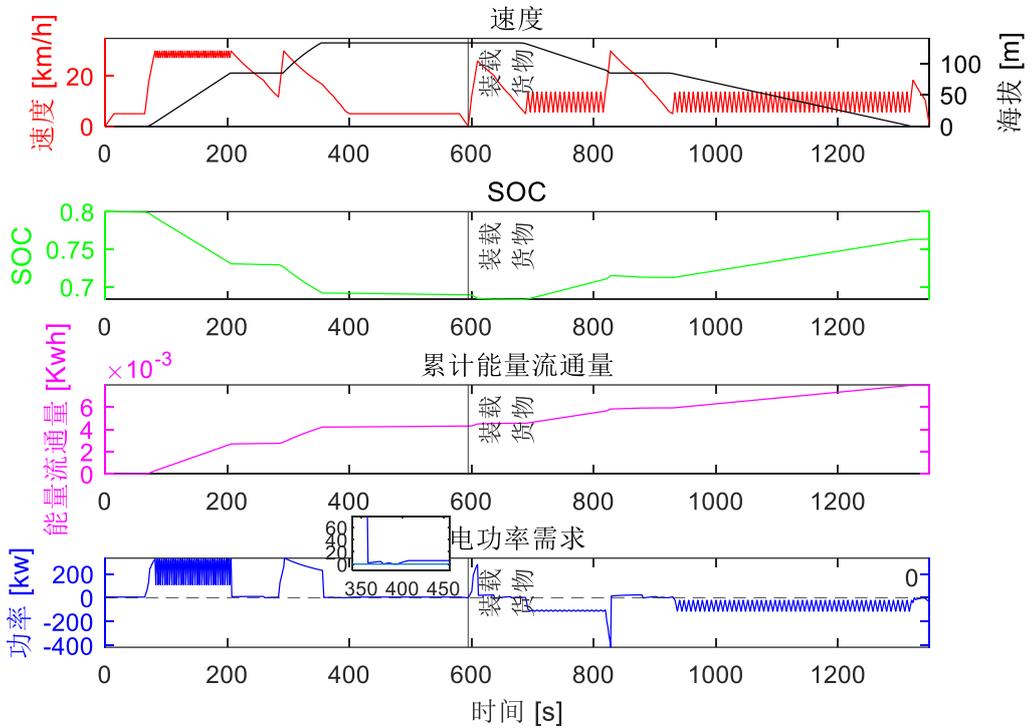


图 2-23 针对路况一，对电动车进行电池寿命优化后的策略

Fig.2-23 Electric vehicle driving strategy after battery lifetime optimization for road condition 1.

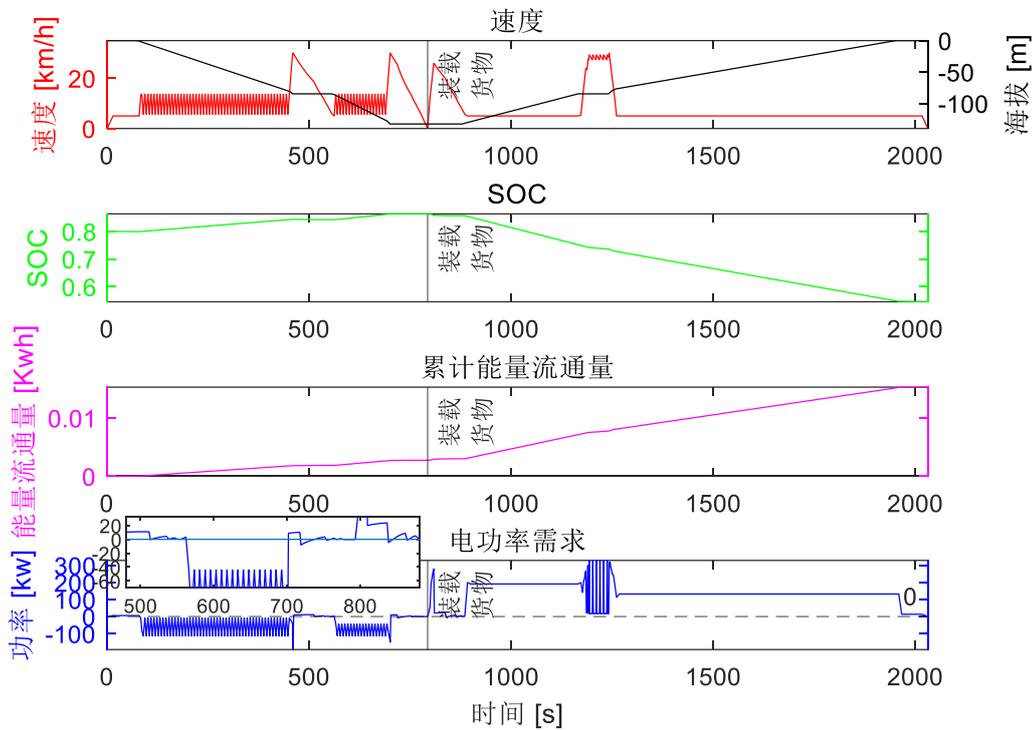


图 2-24 针对路况二，对电动车进行电池寿命优化后的策略

Fig.2-24 Electric vehicle driving strategy after battery lifetime optimization for road condition 2.

基于仿真结果，将具有不同目标的两种车辆的详细优化结果编制成表 2-4 和表 2-5，下划线标出的结果是在单目标条件下的最小值，可以作为第 2.4.2 节中多目标优化公式的参考值。此外，时间和能量消耗的变化在图 2-25、图 2-26、图 2-27 和图 2-28 中绘制在一起。

表 2-4 两种车型在路况一下优化结果

Tab.2-4 Optimization results of the two models under road condition 1.

燃油车辆优化结果			
优化目标	累计油耗值	耗时	
时间	4267. 0g	<u>628. 6s</u>	
油耗	<u>3660. 1g</u>	924. 6s	
多目标优化	4041. 5g	649. 6s	
电动车辆优化结果			
优化目标	始末 SOC 变化量	累计能量流通量	耗时
时间	+0. 0113	46. 1Kwh	<u>635. 5s</u>
电量	<u>+0. 0679</u>	53. 6Kwh	1156. 4s
电池能量流通	-0. 0361	<u>28. 7Kwh</u>	1349. 9s
多目标优化	+0. 0393	41. 8Kwh	679. 0s

表 2-5 两种车型在路况二下优化结果

Tab.2-5 Optimization results of the two models under road condition 2.

燃油车辆优化结果			
优化目标	累计油耗值	耗时	
时间	8347.4g	<u>825.6s</u>	
油耗	<u>7587.4g</u>	1236.7s	
多目标优化	8170.3g	834.3s	
电动车辆优化结果			
优化目标	始末 SOC 变化量	累计能量流通量	耗时
时间	-0.2478	61.9Kwh	<u>1403.3s</u>
电量	<u>-0.2205</u>	69.4Kwh	1879.0s
电池能量流通	-0.2545	<u>55.0Kwh</u>	2030.8s
多目标优化	-0.2298	60.6Kwh	1440.9s

从燃油车和电动车的多目标优化结果中可以得出一些规律。与能量优先的速度轨迹相比，多目标优化的速度轨迹花费的时间更少，行驶更平稳，速度波动更少。此外，它可以减少电动车的电池容量损失，延长寿命。但它会导致能耗增加。对于燃油车而言，在 Route 1 上，时间消耗可以减少 29.74%，在 Route 2 上可以减少 32.54%。对于电动车而言，多目标优化的轨迹速度策略可以在 Route 1 上减少 22.01% 的电池损失，在 Route 2 上节省 41.28% 的时间。

如图 2-25 至图 2-28 所示，当将运输时间设置为单一目标时，两种车辆的行驶时间与距离之间存在明显的线性关系。曲线的斜率较小且波动不大，这意味着车辆以尽可能快的恒定速度通过路线，但这会导致高能量消耗。当能量消耗或电池寿命被设定为单一目标时，最优行驶速度轨迹与坡度有特定的关联。时间消耗和距离的曲线具有非线性特征，在不同的坡度下呈现不同的斜率。能量消耗处于相对较低的状态。在进行多目标优化时，两种车辆都采用了加速滑行 PnG 轨迹，这与任何单一目标优化结果都不同。时间消耗曲线是非线性的，并在时间优化和能量优化之间寻求权衡。

能量消耗和时间消耗在一定程度上具有对立的特性，并且与车辆速度直接相关。减少能量消耗需要牺牲行驶时间，反之亦然。这与电池损耗类似。通过频繁的能量回收来减少电量，会导致累积能量流的增加和电池寿命的损失。然而，通过减少充电和放电的频率来优化电池寿命。电池寿命的减少与电量消耗之间存在妥协。

采矿车辆运输的目标非常复杂，很难找到一个固定的公式来计算。它在它们之间实现了不同的帕累托最优权衡。对多目标解进行分析，展示了“时间消耗”、“能量消耗”和“电池寿命”之间的权衡，有助于找到期望的解决方案。

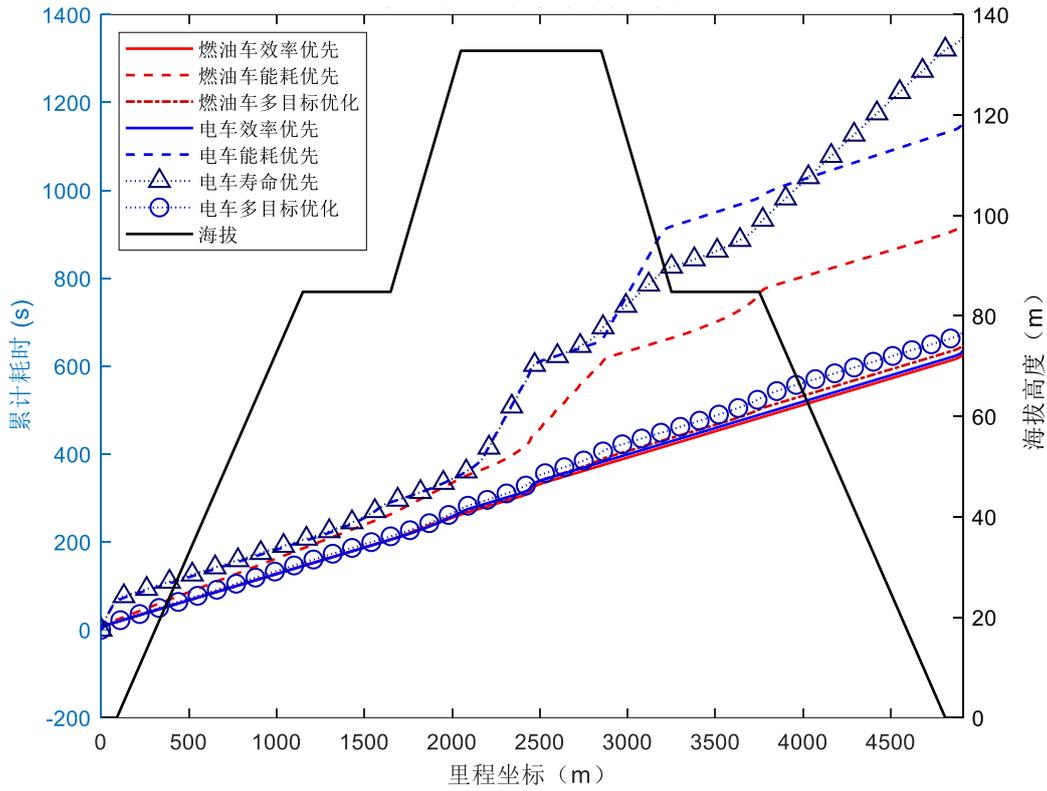


图 2-25 针对路况一，不同行驶策略的耗时对比

Fig.2-25 Time consumption comparison of different driving strategies for road condition 1.

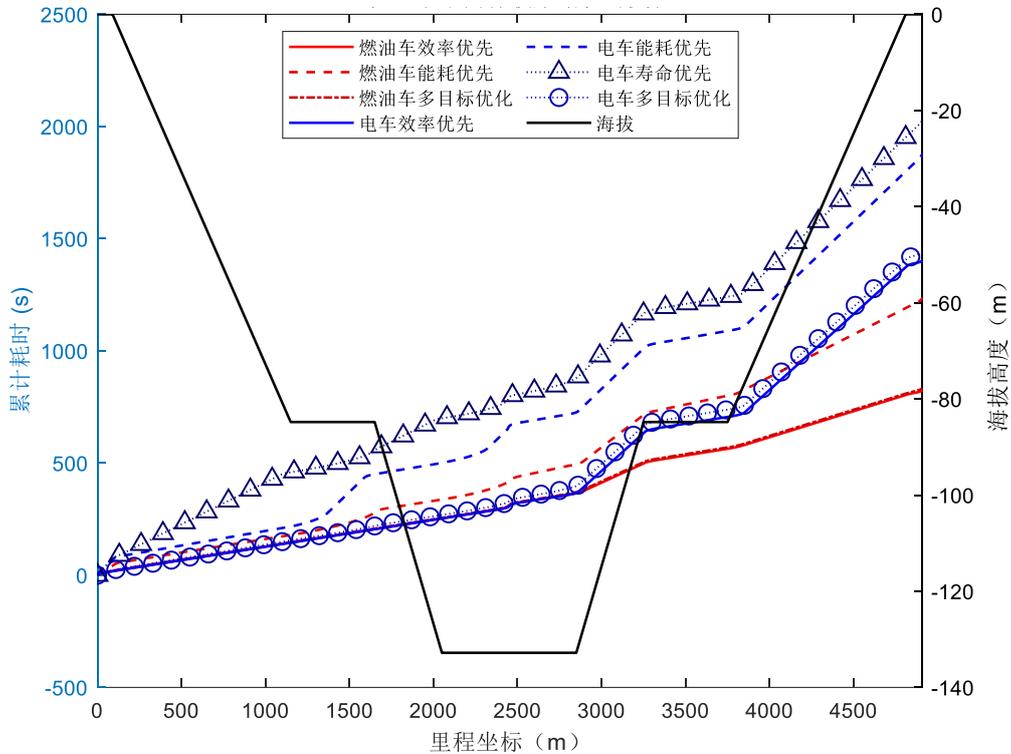


图 2-26 针对路况二，不同行驶策略的耗时对比

Fig.2-26 Time consumption comparison of different driving strategies for road condition 2.

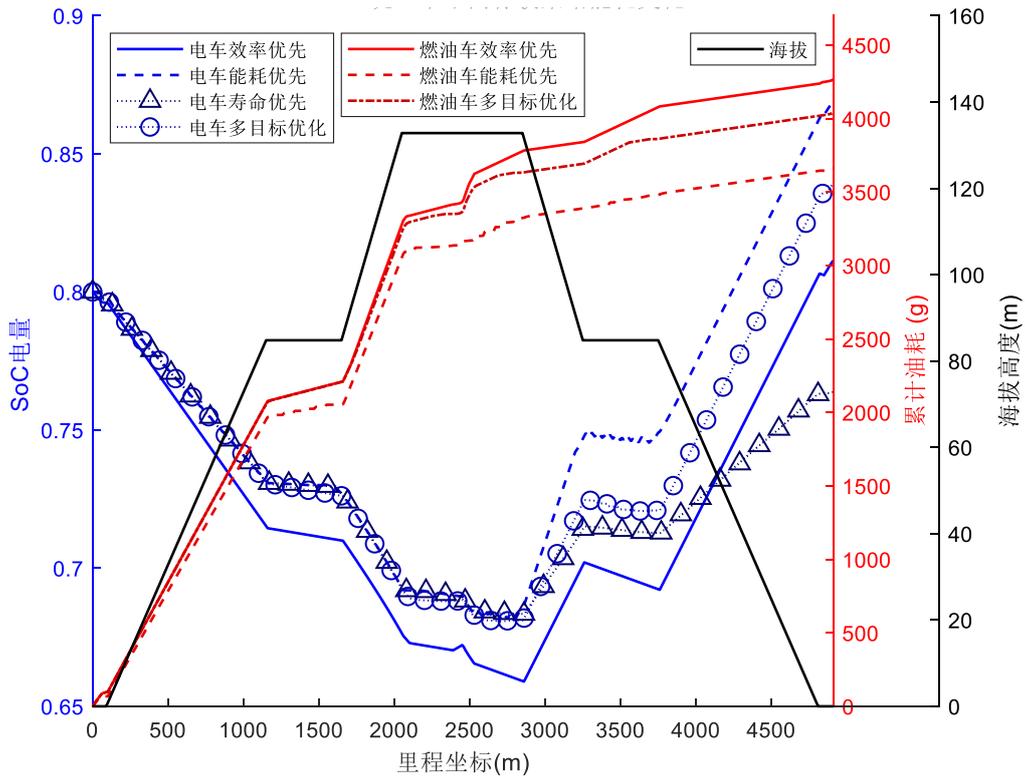


图 2-27 针对路况一，不同行驶策略的能耗对比

Fig.2-27 Energy consumption comparison of different driving strategies for road condition 1.

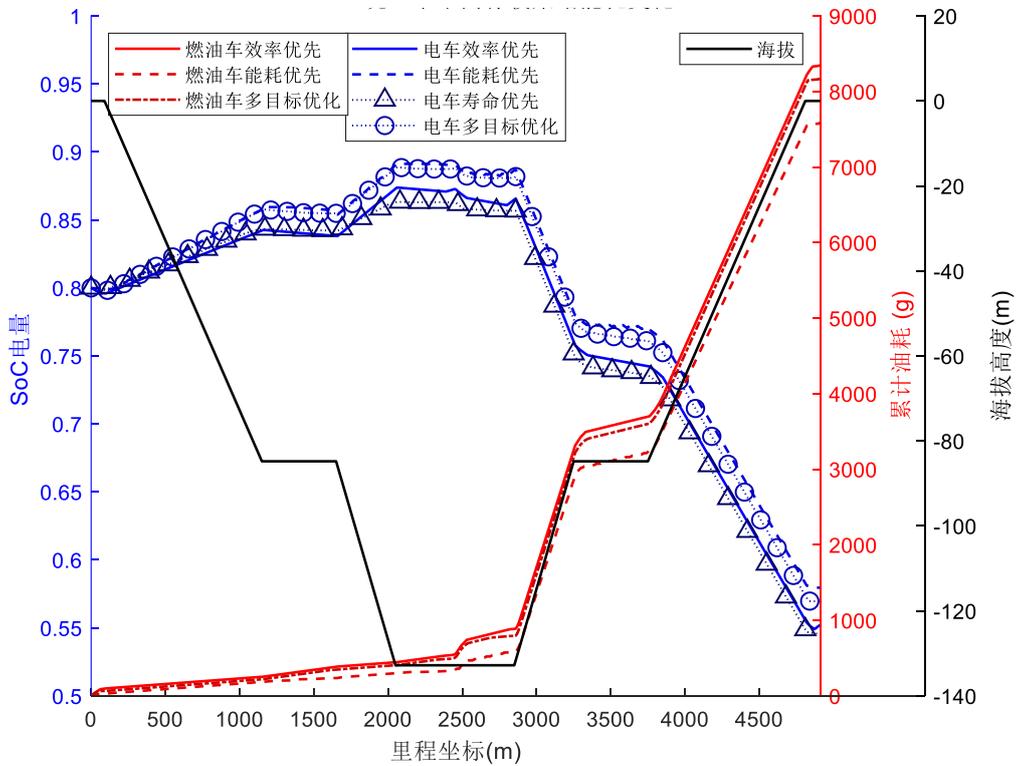


图 2-28 针对路况二，不同行驶策略的能耗对比

Fig.2-28 Energy consumption comparison of different driving strategies for road condition 2.

2.6 本章小结

在无人驾驶条件下，矿山运输车辆的行驶速度轨迹具有更多的优化空间。它尚未被纳入以往的矿山运输节能研究。本节讨论了自主采矿车辆的速度轨迹优化。从以上研究可以得出以下结论：

1) 提出了一种考虑行驶时间、能量消耗和电池寿命的多目标优化方法，采用动态规划进行优化。其在燃油车和电动车模型上验证了其有效性和适用性。因此，在实际封闭露天矿条件下，它更适合矿山运输的成本管理。

2) 利用动态规划在两种采矿场景下进行了多目标优化，包括上坡和下坡条件。发现速度轨迹与路线坡度之间存在特定的关联。随着路线坡度的变化，采用不同频率的加速滑行 PnG 方法。

3) 与能量优先这种单一目标优化的速度轨迹相比，多目标优化的行驶速度更快，花费的时间更少，行驶更平稳，速度波动更少。此外，它可以减少电动车的电池容量损失，延长寿命。但会导致能能耗增加。这是一种帕累托最优权衡。

在未来的工作中，应基于准确的矿山实证数据利用大数据和机器学习更多地研究目标的组成和权重。多目标优化的结构可以自适应地应用于不同的运输场景。

3 基于深度强化学习的矿山无人驾驶车队调度系统

3.1 引言

在矿山运输领域，传统的调度和人工控制方法往往会导致效率降低和次优结果，包括资源浪费、能源消耗增加和安全风险。本章建立了一个新的露天采矿运输模拟环境通过将 Deep Q-Network (DQN) 与动态规划轨迹优化方法相结合，可以提高采矿操作的效率，从而减少等待时间和能源消耗。所提出的方法旨在训练车队在任务管理、排队时间和等待车辆数量方面做出更明智的决策，该方法在模拟器中进行了多次验证。

考虑到传统采矿业普遍存在的安全风险和劳动力老龄化问题，智能采矿的概念已成为全球关注的热门话题。采矿运输的燃料消耗占该行业温室气体排放的一半以上。此外，解决安全问题、提高运营效率和增强盈利能力的需求促使了人工技术的发展，旨在实现绿色和可持续的采矿环境。

先进的传感器和分析技术的部署使得能够创建采矿场地的高清地图。这有利于在运输系统中实施速度和轨迹优化技术，从而减少能源消耗。人工智能和工业互联网 (IIoT) 的集成使得智能采矿能够实现高效和自动化操作。

矿山运输中的卡车调度活动包括铲车将材料从采矿工作面高效地运输到指定目的地，既要考虑数量也要考虑质量。有效的调度决策不仅可以大大提高运输资源的利用率和效率，还可以增加收入和提高满意度。真实矿山环境中的运输周期如图 3-1 所示。

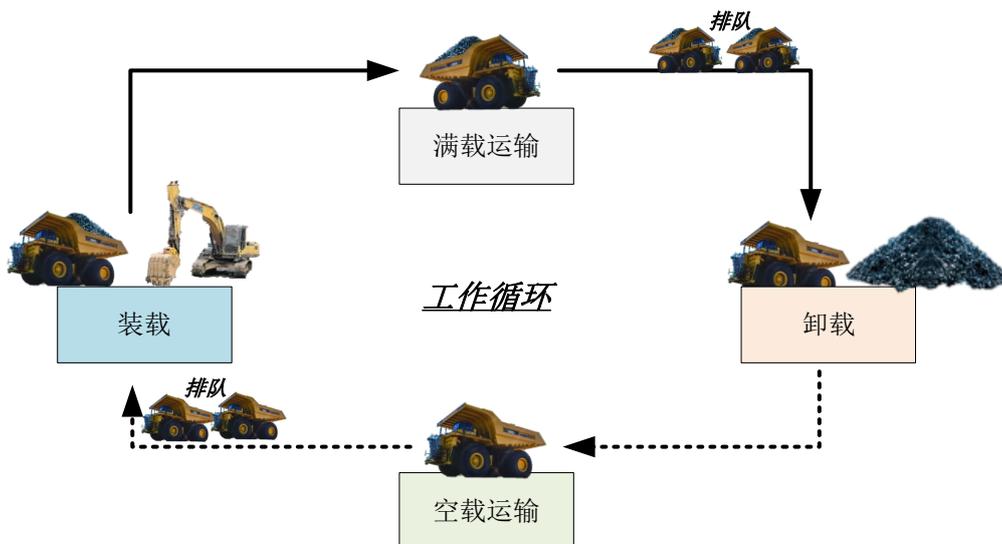


图 3-1 真实的采矿工作循环

Fig.3-1 Real mining work cycle

在矿山运输中，智能车队管理在提高生产率方面起着至关重要的作用。传统的调度解决方案由于其短期计划的性质，往往会出现次优性能。本节提出了基于强化学习的调度系统，并结合速度优化。主要工作包括以下几点：

(1) 建立了一个新的露天采矿运输模拟环境，可以测试和评估各种调度方法。

(2) 引入了 DQN 强化学习调度系统，该系统结合了动态规划速度优化方法，以提高操作效率。

(3) 比较了学习型的调度系统和传统系统之间的性能，展示了矿山运输中生产率和效率的改进。

3.2 车队调度的强化学习方法

强化学习是一种机器学习方法，通过智能体与环境的互动学习如何做出最优决策，以最大化长期奖励，而不需要预先标记的数据集。强化学习适用于那些涉及序列决策和不确定性的问题，例如自动驾驶、机器人控制、游戏策略和资源管理等领域。强化学习的优势在于它能够处理复杂的、动态的环境，并能够在面对不确定性和长期目标时做出决策。它已经在许多领域取得了显著的应用进展，并在自主控制、智能游戏、自适应系统和优化等领域展现出潜力。

如图 3-2，在强化学习中，智能体通过观察环境的状态，强化学习涉及智能体选择适当的动作与环境互动。每个时间步，智能体收到奖励信号评价其行为。其目标是找到一种最优策略，以最大化长期累积奖励。



图 3-2 强化学习基本框架

Fig.3-2 The basic framework of Reinforcement Learning

强化学习通过值函数或策略函数指导智能体的行为决策。值函数评估特定状态下的长期累积奖励期望，策略函数定义了智能体在状态下如何选择动作。算法通过更新值函数或策略函数来优化决策策略。主要概念包括马尔可夫决策过程 (MDP)、Q-learning、策略梯度方法和深度强化学习。MDP 描述序列决策问题，基于马尔可夫性质建模环境，定义了状态、动作、状态转移概率、奖励函数和折扣因子。

1) 状态 (State) S : 描述环境的特征或情境，可以是离散的或连续的。

- 2) 动作 (Action) A : 在特定状态下智能体可以执行的操作或决策, 根据问题的性质可以是离散的或连续的。
- 3) 状态转移概率 (State Transition Probability) P : 定义了给定状态和动作下智能体从一个状态转移到另一个状态的概率分布, 反映了环境的动态特性。
- 4) 奖励函数 (Reward Function) R : 将状态和动作映射到数值奖励, 表示智能体在特定状态下采取某个动作的立即回报, 可以鼓励或惩罚智能体的行为。
- 5) 折扣因子 (Discount Factor) γ : 介于 0 和 1 之间的值, 衡量未来奖励的重要性, 影响智能体对即时奖励和延迟奖励的权衡。

累计奖励的计算方式有很多种, 常用的有平均累计奖励, 如式 3-1。其中 G_t 表示累计奖励值, γ 表示折扣因子, 较小的折扣因子更加关注即时奖励, 而较大的折扣因子更加重视长期累积奖励。状态转移概率的定义公式为式 3-2。同时奖励值也是不固定的, 智能体在某状态通过动作到达下一状态的奖励值也是不固定的, 公式为式 3-3。

$$G_t = \sum_{i=0}^{T-t} \gamma^i r_{t+i+1}, \quad \gamma \in [0,1] \quad (3-1)$$

$$p(s'|s, a) = P(s_{t+1}|s_t, a_t) \quad (3-2)$$

$$p(s', r|s, a) = P(s_{t+1}, r_{t+1}|s_t, a_t) \quad (3-3)$$

MDP (马尔可夫决策过程) 中, 核心是优化策略, 常用 $\pi(a|s)$ 表示。它表示在状态 s 时, 选择某个动作的概率分布。状态价值函数 $V_\pi(s)$ 定义为在状态 s 下按照策略 π 获得的预期累积回报, 表示如下:

$$V_\pi(s) = \mathbb{E}_\pi[G_t|s_t] = \mathbb{E}_\pi[\sum_{i=0}^{\infty} \gamma^i r_{t+i+1}|s_t] \quad (3-4)$$

价值函数 $Q_\pi(s, a)$ 定义智能体在状态 s 下依据策略 π 执行动作 a 后, 获得的累计回报, 表达如下:

$$Q_\pi(s, a) = \mathbb{E}_\pi[G_t|s_t, a_t] = \mathbb{E}_\pi[\sum_{i=0}^{\infty} \gamma^i r_{t+i+1}|s_t, a_t] \quad (3-5)$$

根据式(3-4)和(3-5), 状态价值函数和状态动作价值函数可以通过以下计算:

$$\begin{cases} V_\pi(s) = \sum_{a \in A} \pi(a|s) Q_\pi(s, a) \\ Q_\pi(s, a) = \sum_{s', r} p(s', r|s, a) [r + \gamma V_\pi(s')] \end{cases} \quad (3-6)$$

进一步推导为递归形式, 得到贝尔曼公式:

$$\begin{cases} V_\pi(s) = \sum_{a \in A} \pi(a|s) \sum_{s', r} p(s', r|s, a) [r + \gamma V_\pi(s')] \\ Q_\pi(s, a) = \sum_{s', r} p(s', r|s, a) [r + \gamma \sum_{a' \in A} \pi(a'|s') Q_\pi(s', a')] \end{cases} \quad (3-7)$$

因此, 强化学习目的是找出能使得状态值函数 $V_*(s)$ 或者状态动作值函数 $Q_*(s)$ 最大化的策略 π_* :

$$V_*(s) = \max V_\pi(s) \quad (3-8)$$

$$Q_*(s) = \max Q_\pi(s, a) \quad (3-9)$$

当采用最优策略 π_* 时，马尔可夫决策过程（MDP）能够达到最优解和最优值函数。最优值函数表示在给定最优策略下，从每个状态出发所能获得的最大期望累积回报。为了获得最优值函数，可以采用时间差分学习方法或蒙特卡洛方法进行迭代更新。时间差分学习方法基于当前估计的值函数，通过与实际观测到的回报进行比较来更新值函数的估计值。这种方法可以在每个时间步进行增量更新，逐渐收敛到最优值函数。蒙特卡洛方法则通过采样多个完整的轨迹或序列，并计算累积回报来估计值函数的期望值。这种方法的更新是基于完整的回报序列，而不是单个时间步的差异。

蒙特卡洛方法是一种基于采样轨迹的强化学习方法，用于估计值函数的期望值或动作值函数的期望值。它通过采样多个完整的轨迹或序列，并计算累积回报来估计值函数的期望值。蒙特卡洛方法具有无模型的特点，可以直接从交互中学习，并且适用于部分可观测环境和连续动作空间。更新过程如式 3-10，其中 α 代表状态 s_t 出现的次数。

$$V(s_t) \leftarrow V(s_t) + \alpha(G_t - V(s_t)) \quad (3-10)$$

时间差分（TD）学习是一种基于差异的强化学习方法，用于估计值函数和学习最优策略。与蒙特卡洛方法相比，TD 学习方法可以在单个时间步骤上进行增量更新，无需等待完整的轨迹。相比于蒙特卡洛方法，TD 学习方法的收敛速度更快，且不需要等待完整的轨迹。TD 的更新过程如式 3-11，其中 α 代表学习率， $r_{t+1} + \gamma V(s_{t+1})$ 代表 TD 的目标值。

$$V(s_t) \leftarrow V(s_t) + \alpha(r_{t+1} + \gamma V(s_{t+1}) - V(s_t)) \quad (3-11)$$

强化学习求解本质是找到最优策略，智能体通过策略选择最优动作，从而得到最高长期奖励。

3.2.1 深度强化学习求解

当前的强化学习方法主要分为基于价值的、基于策略的，以及结合两者的方法。

1) 基于价值的方法

基于价值的方法通过估计状态值函数或动作值函数来学习最优策略，例如 Q-Learning 和 DQN。这些方法适用于处理离散动作空间的任務。如表 3-1 所示，通过构建状态和动作表格来存储状态动作 q 值，每次执行动作时，查阅并选取当前状态下最大值的动作。

表 3-1 Q-Table 简介

Q-Table	a1	a2
s1	q(s1,a1)	q(s1,a2)
s2	q(s2,a1)	q(s2,a2)
s3	q(s3,a1)	q(s3,a2)

强化学习的目的就是在优先的状态和动作空间内最大化式 3-6，采取贪心策略不断选择动作，结合时间差分的学习方法不断更新 Q-Table，Q-Learning 的更新公式为 3-12，其中 α 代表学习率， γ 为衰减系数， γ 越大表示越重视未来的影响。

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r_{t+1} + \gamma \max_{a'} Q(s', a') - Q(s, a)) \quad (3-12)$$

但在 Q-Learning 进行强化学习的过程中，很多场景会很复杂。一些任务有着巨大的状态空间和动作空间，Q-Table 难以存储足够的 q 值来用于更新和训练。因此以 Q-Table 为基础的学习方式并不能解决一些复杂的问题。例如式 3-13，其中 ϕ 代表神经网络的权重矩阵参数，通过神经网络来代理 Q-Table 进行数值拟合，解决了 Q 值不充分的问题，并能处理大规模的动作状态空间。

$$Q_\phi(s, a) = Q(s, a) \quad (3-13)$$

该网络结构由卷积层和全连接层组成，能够灵活地处理各种输入。如图 3-3，Double DQN 使用经验回放缓冲区存储智能体与环境交互的经验，训练时随机抽取样本以提高训练的稳定性。为了提高稳定性，DQN 采用了两个神经网络：主网络用于计算当前状态下的动作值函数，目标网络则用于计算目标动作值函数。目标网络的参数是主网络参数的副本，并定期更新，以减少训练过程中目标值的变动，增强稳定性。DQN 通过不断的迭代训练，使神经网络逼近最优的动作值函数，学习到最优策略。

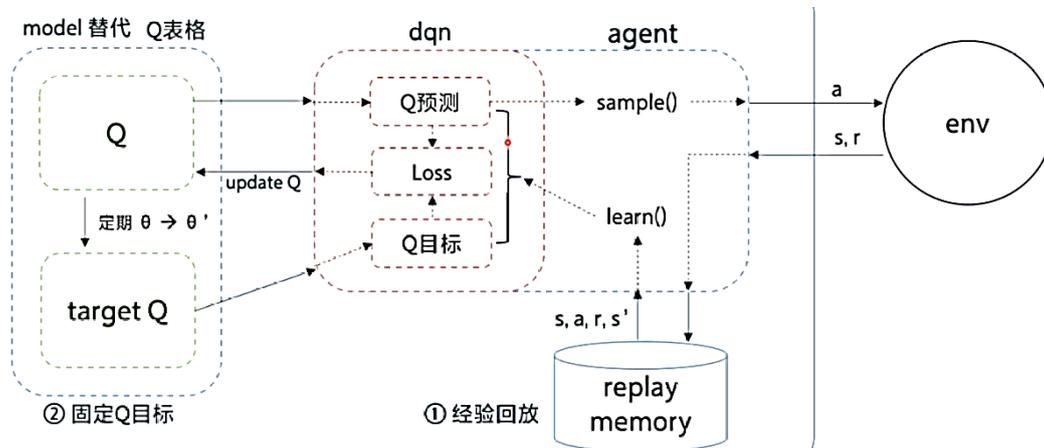


图 3-3 DDQN 算法框架

Fig.3-3 Framework of DDQN algorithm

$$Q_{\varphi}(s, a) \leftarrow Q_{\varphi}(s, a) + \alpha(r_{t+1} + \gamma \max_{a'} Q_{target}(s', a') - Q(s, a)) \quad (3-14)$$

2) 基于策略的方法

基于策略的强化学习方法专注于学习和优化策略，与基于值函数的方法不同。它们直接优化策略，以获取最佳行为策略。常见的基于策略的方法包括策略梯度方法（Policy Gradient）和演员评判方法（Actor-Critic）。这些方法在处理连续动作空间、高维状态空间和非线性策略等问题上具有优势。通常，策略被描述为包含参数 θ 的策略 π ，如式 3-15，这些方法通常使用梯度上升法来寻找最优策略。首先，将优化目标设定为初始状态下的预期回报，见式 3-16。其中 τ 代表利用策略 π_{θ} 采样的一段轨迹，其中包含 $\{s_1, a_1, \dots, s_T, s_T\}$ ； $R(\tau)$ 代表这段轨迹 τ 所带来的回报； $P_{\theta}(\tau)$ 代表轨迹 τ 出现的概率； N 是利用当前策略 π_{θ} 采样的一段轨迹的条数； T 为轨迹 τ 的步长。

$$\pi_{\theta}(s, a) = P(a|s, \theta) \quad (3-15)$$

$$\begin{aligned} J(\theta) &= \sum_{\tau} P_{\theta}(\tau) R(\tau) \\ &= \sum_{\tau} (p(s_{t=0}) \prod_{t=1}^T p_{\theta}(a_t|s_t) p(s_{t+1}|s_t, a_t) \sum_{t=1}^T r(s_t^n, a_t^n)) \\ &\approx \frac{1}{N} \sum_{n=1}^N R(\tau^n) \end{aligned} \quad (3-16)$$

对参数 θ 求导可得：

$$\nabla_{\theta} J(\theta) = \frac{1}{N} \sum_{n=1}^N R(\tau^n) \sum_{t=1}^{T_n} \nabla_{\theta} \log p_{\theta}(s_t^n, a_t^n) \quad (3-17)$$

最常见的基于策略的方法为蒙特卡洛策略梯度算法，通过使用价值函数 $V_{\pi}(s)$ 来替代公式中的 $Q_{\pi_{\theta}}(s, a)$ ，则更新公式为 3-18。其中 $V_t(s)$ 为采用蒙特卡洛方法计算的在 t 位置时刻的状态价值。使用梯度上升来最大化期望奖励，如式 3-18。

$$\theta = \theta + \alpha \nabla_{\theta} J(\theta) \quad (3-18)$$

然而，基于蒙特卡洛的方法通常具有较高的方差，需要更多的样本来进行训练，参数更新不一定是最优的方向。因此，在实践中，根据具体问题的特点和需求，选择合适的方法进行强化学习建模和求解是很重要的。

3) 演员评判类（Actor-Critic）方法

演员评判类（Actor-Critic）方法结合了值函数估计（Critic）和策略优化（Actor）。它同时学习策略网络（Actor）和值函数网络（Critic），以提高学习效率和稳定性。如图 3-4，演员评判类方法通过策略梯度方法更新策略。Actor 根据当前策略选择动作，智能体得到指令后与环境进行交互，观察下一个状态和奖励值。然后，Actor 网络根据奖励值和动作概率分布的梯度来更新参数，以便选出奖励值高的动作。Critic 网络用于估计策略的价值，并提供对策略的评估。Critic 网络的参数通过奖励值与估计的值函数之间

的差异来进行更新。这可以使用时序差分学习（TD 学习）或蒙特卡洛方法。通过交替进行 Actor 网络和 Critic 网络更新，演员评判类方法可以自适应地改进策略和优化值函数。这种相互作用和合作使得演员评判类方法能够在学习过程中逐步优化策略，并在不断的迭代中提高性能。

A2C, A3C 和 DDPG 是常见的演员评判类方法。其中，DDPG 是最常见的方法之一，其算法结构如图 3-5 所示。DDPG 使用四个网络进行训练和学习，包括演员网络（Actor Network）、评判网络（Critic Network）、目标演员网络（Target Actor Network）和目标评判网络（Target Critic Network）。通过交替更新演员网络和评判网络的参数，并使用目标演员网络和目标评判网络进行目标估计，DDPG 可以逐渐优化策略和值函数，并在连续动作空间的强化学习任务中实现良好的性能。

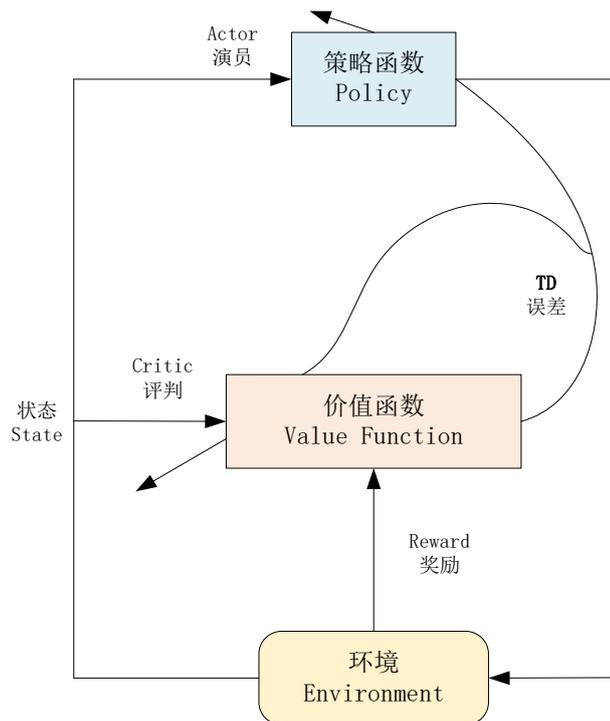


图 3-4 演员评判类算法框架

Fig.3-4 Framework of Critic-Actor algorithm

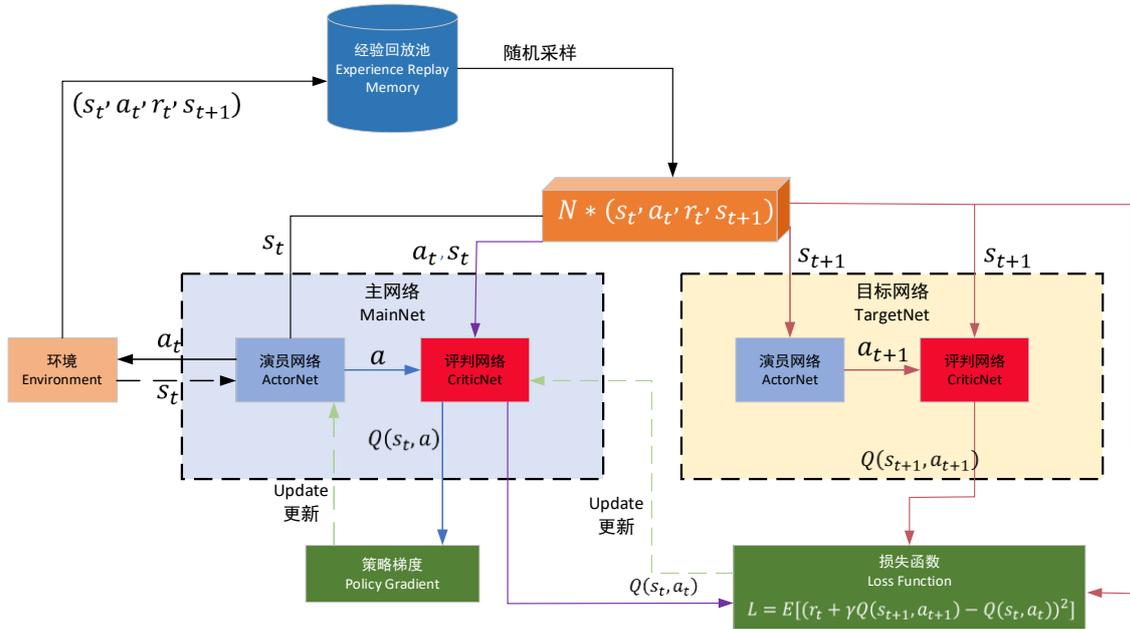


图 3-5 Deep Deterministic Policy Gradient (DDPG) 算法框架
Fig.3-5 Framework of Deep Deterministic Policy Gradient (DDPG) algorithm

3.2.2 多智能体深度强化学习

强化学习中的多智能体系统（Multi-Agent Reinforcement Learning, MARL）是由分布式人工智能演变而来，其智能体具有独立性和自主性，通过并行协调解决问题以降低系统复杂性。多智能体强化学习将问题建模为随机博弈，其中智能体在多个阶段中进行决策，并与其他智能体交互。

MARL 的目标是通过智能体之间的合作、竞争或协调来解决复杂的任务。如图 3-6，在于环境的交互过程中，环境中含有不断进行学习和更新的其他智能体，在任何一个智能体的视角下，环境是非稳态的，环境迁移的分布会发生变化。

一个随机博弈可以定义为元组 $(N, \mathcal{S}, \mathcal{A}, \mathcal{R}, p, \gamma)$ ，其中 N 是智能体的数目， \mathcal{S} 是所有智能体的状态集合， $\mathcal{A} = A_1 \times A_2 \times \dots \times A_N$ 代表所有智能体的动作集合； $\mathcal{R} = r_1 \times r_2 \times \dots \times r_N$ 是所有智能体的奖励函数集合； $p: \mathcal{S} \times \mathcal{A} \rightarrow \Omega(\mathcal{S})$ 是环境转移的概率， $\Omega(\mathcal{S})$ 代表分布的集合； γ 代表折扣因子， $\gamma \in (0, 1)$ 。

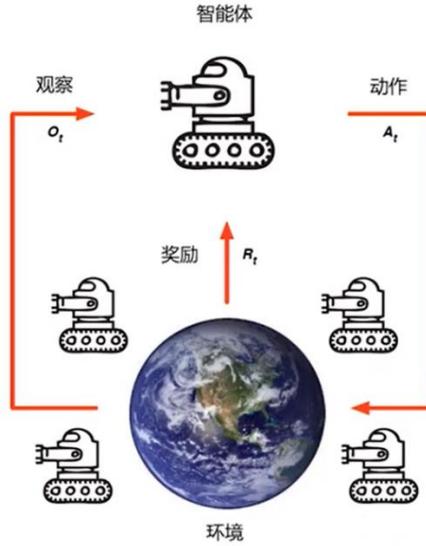


图 3-6 多智能体强化学习

Fig.3-6 Multi-agent reinforcement learning

对于一个随机博弈中的每个智能体 i ，其策略表示为：

$$\pi_i: \mathcal{S} \rightarrow \Omega(A_i) \quad (3-19)$$

所有的智能体联合策略可以表示为 $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_N]$ ；所有智能体的联合动作可表示为 $\boldsymbol{a} = [a_1, a_2, \dots, a_N]$ 。基于联合体策略 $\boldsymbol{\pi}$ ，智能体 i 在状态 s 下的价值为式 3-20；那么基于联合体策略 $\boldsymbol{\pi}$ ，智能体 i 采取动作 (s, \boldsymbol{a}) 下的价值为式 3-21。

$$V_i^\pi(s) = \sum_{t=0}^{\infty} \gamma^t \mathbb{E}_{\pi, p} [r_i^t | s_0 = s] \quad (3-20)$$

$$Q_i^\pi(s, \boldsymbol{a}) = \gamma_i(s, \boldsymbol{a}) + \gamma \mathbb{E}_{s' \sim p} [V_i^\pi(s')] \quad (3-21)$$

多智能体强化学习的方法主要分为三种，完全中心化的方法、完全去中心化的方法以及中心化训练去中心化执行方法。

完全中心化方法指的是所有智能体的决策和学习过程由一个中央控制器或决策者进行集中管理的方法。其优点是可以全局优化智能体的行动，协调智能体之间的合作或竞争关系，并且可以利用全局的状态信息来进行决策。此外，中央实体可以收集所有智能体的经验数据，进行集中的学习和更新。然而，完全中心化方法也存在一些局限性。由于所有智能体的决策和学习都由中央实体管理，这会增加计算和通信的复杂性，并且对中央实体的负载要求较高。此外，完全中心化方法可能面临可扩展性和实时性的挑战，特别是在涉及大规模智能体或需要实时决策的复杂环境中。

在强化学习中，完全去中心化的方法是指在多智能体环境中，每个智能体都具有独立的决策和学习能力，没有中央控制器或决策者进行集中管理的方法。每个智能体根据

自己的观测信息和局部的状态信息来进行决策，并根据自己的策略和价值函数进行学习。完全去中心化的方法的共同特点是每个智能体都具有独立的决策和学习能力，没有中央控制器来集中管理。每个智能体根据自己的观测信息和局部的状态信息进行决策，并独立地学习和更新自己的策略和值函数。这种去中心化的方法可以提高系统的鲁棒性和可扩展性，但也可能导致合作或竞争关系的协调和学习的困难。

“中心化训练去中心化执行”方法在训练时由中央控制器管理全局信息，协调智能体，但在执行阶段，智能体独立决策，不再需要中央控制。每个智能体可以使用在训练阶段学到的策略来选择行动，而无需与中央控制器进行通信。这样可以减少通信和计算负担，并使智能体能够在实际环境中独立地执行决策。其优点是可以在训练阶段充分利用全局信息进行学习和协调，同时在执行阶段实现去中心化的决策，提高系统的实时性和可扩展性。这种方法常用于涉及大规模智能体或需要实时决策的复杂环境中，以平衡训练和执行的需求。需要注意的是，这种方法需要确保在执行阶段智能体能够有效地利用训练阶段学到的策略，而无需依赖中央控制器。这可能需要对策略进行适当的修整以适应去中心化执行的环境。此外，还需要注意智能体之间的合作或竞争关系，以确保系统的整体性能。

从原理上讲，多智能体强化学习更加困难。该方法既要与环境进行交互，又要相互之间进行交互。简单的将其他智能体考虑成环境的一部分进行学习，会使得算法不收敛，学习不稳定。

3.3 矿山无人驾驶车队调度系统设计

露天矿的车队调度有两个主要目标：（1）有效地将卡车分配给铲车，以最大限度地提高生产率^[96]，（2）在满足需求的同时，最大限度地减少车队延误和成本^[48]。因此，车队的首要目标是将尽可能多的货物从铲车运输到卸载场，同时通过选择最佳路线来减少等待时间和成本。本研究针对的卡车车队均为无人驾驶车辆，仅考虑车辆的调度问题，假设车队在行驶运输过程中是无冲突的，不会产生任何碰撞，并且不考虑车辆在行驶过程中的道路横向偏差。为了实现最优的车队调度，提出了一种速度优化的学习型调度系统，包括基于动态规划的多目标速度轨迹优化方法和基于强化学习（RL）的车队调度系统。动态规划方法优化运输过程的速度，特别是在减少露天矿内的运输能耗和成本方面。基于 RL 的方法的主要优点在于其能够根据环境信息和经验进行学习和实时决策。在实际训练中，学习型调度系统的车辆采用动态规划中获得的优化速度进行训练。

3.3.1 基于强化学习的车队管理系统

在这项研究中，卡车被称为智能体。如图 3-7 所示，智能体通过累积奖励最大化来优化其决策，随着时间的推移逐渐学习到更优的行为^[97]。

本研究采用 DQN 强化学习算法来实现调度方案。DQN 算法通过训练智能体学习策略，每辆卡车被视为一个独立的智能体，并为所有智能体定义了一组状态变量，状态变量的定义对于智能体至关重要。

1) 智能体

在本章中，每辆卡车作为一个智能体。卡车车队由具有相同容量和装卸时间的相同卡车组成。然而，卡车的行驶速度是经过动态规划方法优化后的。此外，假设铲车和卸载点具有相同的参数和特征。

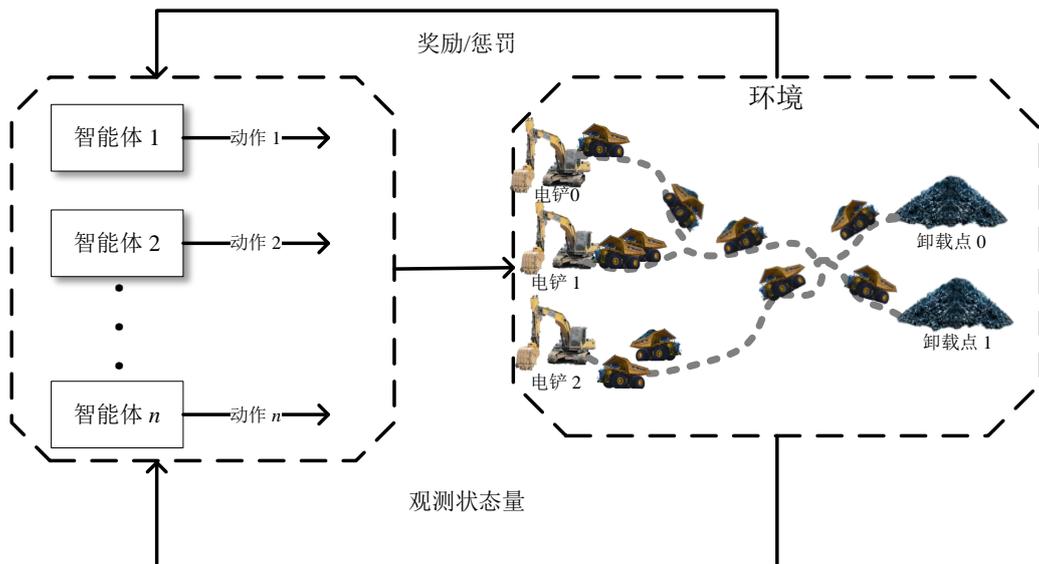


图 3-7 多智能体强化学习框架

Fig.3-7 The framework of multi-agent reinforcement learning

2) 状态空间

智能体观察当前环境以识别与以前状态的变化。在矿区内的每个卸载场和铲车处捕获记录相关卡车队列条件的局部状态 s_t 。在本研究中，状态向量 s_t 考虑以下因素。

a) 卡车载重量

为了适应异构卡车车队系统，卡车容量 C_T 被包括在状态空间中。这使得学习智能体能够意识到卡车容量的变化。通过在状态空间中加入卡车容量，可以训练决策网络有效地调度不同容量的卡车。目标是通过考虑卡车容量的因素做出最优调度决策，从而最大

限度地提高总产量。

b) 装载状态

为了获得矿山现场每辆卡车的装载状态，引入了一个二进制变量，表示为 L_t^T ，其中 t 代表时间。该变量区分卡车是否在装载站点已经装载完毕，或完成卸载点等待调度决策。当卡车已装载并等待时， L_t^T 被设置为1。另一方面，当卡车为空并在前往装载点的途中时， L_t^T 被设置为0。通过引入这个二进制变量，模型有效地表示了每辆卡车的装载状态，使学习智能体能够根据卡车的当前状态做出适当的调度决策。

c) 装载/卸载点的完成率

为了在装载点和卸载点之间实现平衡的生产，仅向智能体提供卡车容量和队列信息是不够的，因为这样只会将卡车派遣到最短的路径上。在本研究中，我们旨在确保所有装载位置保持相似的运输量，所有卸载点保持相似的卸载量。为了实现这一目标，我们引入了每个装载点和卸载站点的完成比率 CR_t^k 的概念。完成比率表示特定时间和特定站点的实际运输或卸载量与目标量之间的比率。它被表示为 CR_t^k ，其中 t 表示当前时间， k 表示装载点或卸载站点。完成比率使我们能够评估和控制装载点和卸载站点之间的生产平衡。通过考虑完成比率，我们旨在确保每个装载点和卸载站点保持类似的运输或卸载活动量，促进矿山现场的平衡生产。完成率的计算公式如下：

$$CR_t^k = \begin{cases} LW_t^k / TW_t^k & \text{前往装载点 } k \\ DW_t^k & \text{前往卸载点 } k \end{cases} \quad (3-22)$$

其中 LW_t^k 和 TW_t^k 分别表示卡车在时间 t 在装载点 k 装载并转移的产量，以及卡车在装载点 k 运输的目标产量。 DW_t^k 指的是在时间 t 在卸载点 k 卸载的产量。

d) 预计等待时间

对于每个装载点和卸载点，如果卡车被分配到该目的地，我们将计算可能的等待时间。为了计算预期等待时间，我们根据卡车的排队条件将其分为两类：“实际队列（Actual Queue，简称AQ）”和“途中队列（En-route Queue，简称EQ）”。“实际队列”包括那些在装载点或卸载点排队等待的卡车。“途中队列”包括那些已被派遣到该装载点或卸载点，但尚未到达的卡车。这两个队列的信息可以准确预测实际的等待时间，并促使调度系统将卡车派遣到队列较短的站点。因此，装载点/卸载点 k 在时刻 t 的预期等待时间 WT_t^k 可以通过方程（4-2）表示为：

$$WT_t^k = \sum_{i \in AQ^k} LD_i + \sum_{j \in EQ^{k*}} (LD_j + HL_j) + LD_T + HL_T \quad (3-23)$$

其中 LD_i 表示卡车 $i \in AQ^k$ （其中 AQ^k 是已到达装载点 k 的实际队列中的所有卡车）

的平均装载时间。方程的第二部分考虑了途中队列。特别地，它仅关注卡车 $j \in EQ^{k*}$ （其中 EQ^{k*} 指的是在装载点 k 的途中队列中预计比卡车 T 更早到达的所有卡车）的装载时间和剩余运输时间。以下不等式对于每个装载/卸载点始终成立： $EQ^{k*} \leq EQ^k$ 。方程中最后两项是卡车 T 的装载时间和运输时间。如果卡车 T 被派往卸载点， LD 将被卸载时间替代，并且卸载时间将被适当设置。

e) 排队等待卡车的总容量

仅考虑等待时间是不足以评估每个卸载/装载站点的队列情况，特别是在异构车辆调度系统中。一个队列可能在只有少数卡车实际排队时等待时间很长。因此，有必要计算实际队列 AQ^k 和在卡车 T 之前到达的途中队列 EQ^{k*} 中等待卡车的总容量，表示为 TC_t^k 。通过考虑等待卡车的总容量，智能体可以更好地了解每个卸载/装载站点的整体需求和拥堵情况，从而改善调度策略。

f) 被延误卡车的工作时间

“被延误的卡车”指的是那些已经在途中前往卸载/装载站点，但预计比卡车 T 到达时间晚的卡车。装载站点 k 上延迟卡车的数量可以通过以下方程获得：

$$DT^k = EQ^k - EQ^{k*} \quad (3-24)$$

卡车的工作时间指的是装载时间或卸载时间。如果卡车 T 被派往卸载站点， LD 将被卸载时间替代。因此，可以按以下方式计算活动时间：

$$AT_t^k = \sum_{i \in DT^k} LD_i \quad (3-25)$$

g) 被延误卡车的容量

仅观察延迟卡车的活动可能是不足够的。因此，计算延迟卡车的总容量 DC_t^k 可以训练智能体考虑其决策对其他卡车的影响。智能体将被鼓励在考虑整体生产的同时优先考虑自身利益。

智能体在时间 t 的状态空间可以表示如下。

$$s_t = [C_T, L_t^T, \langle CR_t^k, WT_t^k, TC_t^k, DT^k, AT_t^k \rangle_{k=1, \dots, N+M}] \quad (3-26)$$

如果矿山有 N 个装载站点和 M 个卸载站点，状态向量的长度为 $5 \times (N + M) + 2$ 。状态向量包含有关装载站点和卸载站点当前状态的信息。当卡车被分配到一个装载站点时，调度系统最关键的信息与装载站点的状态有关。因此，在智能体训练过程中，为了减轻计算负担，当卡车被指定到装载站点时，与卸载站点相关的状态向量组成部分可以被赋予零值。同样地，当卡车被分配到卸载站点时，与装载站点相关的组成部分可以被设为零。这种部分将状态向量置零的方法增强了训练过程的效率，使智能体能够专注于与其当前任务相关的信息，这将在下一部分讨论。此外，假设装载站点和卸载站点的特征和条件在矿山环境中保持不变。这个假设为智能体提供了一个稳定的环境来学习和做

出决策，因为它不需要考虑调度过程中装载站点和卸载站点状态的动态变化。

3) 动作空间和动作屏蔽

这个调度问题的行动空间涵盖了矿山中每辆卡车可用的所有可能目的地，包括所有的卸载站点和装载站点。在这个实验中，行动空间被设置为有限的离散空间。假设一个具有 N 个装载站点和 M 个卸载站点的矿山，行动空间可以用以下方程表示：

$$A = \{a_1, a_2, \dots, a_k\} (k = 1, \dots, N + M) \quad (3-27)$$

根据行动空间的定义，行动 a_k 表示卡车将被派往装载站点 k 或卸载站点 k 。当前在卸载站点的卡车被分配到另一个卸载站点是无效的。同样适用于装载站点的情况。因此，对于卸载站点上的卡车，唯一合适的行动是前往装载站点，反之亦然。为了增强学习过程的效率并降低调度系统的复杂性，在这个实验中实施了行动屏蔽。行动屏蔽过滤掉无效的行动，使智能体只能选择行动空间的一个子集作为可行的行动。剩下的无效行动被赋予无穷小的值，以防止它们被选择为有效行动。

4) 奖励函数

在训练系统时，智能体的性能是由奖励函数确定的。在获取行动后，卡车将在矿山环境中分散执行。因此，奖励与状态空间不同，并且将为每个智能体单独计算。它将鼓励正确的行动，并惩罚智能体进行低效的行动。当卡车被派往卸载站点时，在时间 t 时的奖励函数定义如下：

$$\mathcal{R}(s_t, a_t) = -\frac{t_{queue}}{\Delta t} - \frac{\sigma([CR_t^1, \dots, CR_t^M])}{\mu([CR_t^1, \dots, CR_t^M])} \quad (3-28)$$

其中 Δt 是完成行动 a_t 所需的时间，也指的是 a_t 和 a_{t+1} 之间的时间间隔。 t_{queue} 是执行 a_t 时的排队持续时间。 σ 和 μ 分别表示卸载站点个体的完成比例的标准差和均值。但是，如果卡车被派往装载站点，卸载站点的完整比例将被装载站点替代。在这种情况下，奖励方程包括两个组成部分。第一个组成部分衡量了卡车排队时间在总完成时间中所占的比例。这个组成部分允许调度系统优先考虑最小化排队时间，因为它对总完成时间有重要影响。奖励方程的第二个组成部分表示所有卸载/装载站点产量的变异系数 (coefficient of variance, CV)。这个组成部分旨在训练调度系统保持每个卸载/装载站点的平衡运输量。通过最小化变异系数，系统可以在卸载/装载站点之间均匀分配运输工作负荷，确保资源的更加公平利用。这种方法有助于调度系统内更高效和更公平的运营。

3.3.2 动态规划速度策略优化方法

尽管运输卡车的燃料消耗在露天矿中占总能源使用量的 30% 以上^[2]，但调度系统通

常不考虑运输的能耗。为了解决这个问题，在本节中引入了针对每条固定路线的轨迹优化方法。通常，在矿山中，包括距离和坡度在内的路线信息保持不变。因此，采用动态规划（DP）根据车辆特性（电机、电池和传动装置）和地理信息计算最小化能耗的最优速度和轨迹。动态规划如第 2 节所描述，以递归的方式将复杂问题分解为简单子问题。

然而，选择非常小的分段距离会导致子问题数量增加和计算要求过高。如算法 1 所示，路径根据最小速度变化被分割成不同的部分，以在计算时间和准确性之间取得平衡。接下来，对速度进行离散化，允许每个路段的初始速度和最终速度自由组合。随后，计算每个部分内所有可能条件下的加速度、时间和能耗。同时，获取目标函数值。每个部分的最佳结果被存储，并用作后续部分的初始状态。一旦计算完成所有部分，从最后一个部分的结果中选择最小值作为确定前面部分值的线索，以逆序方式进行。这个过程确保每个部分的最优值逐步确定，从而得到整体的最优解。

算法 1: 基于动态规划的速度轨迹优化

输入: 道路的数量 $n = 1, \dots, N$; 每个道路的具体分段信息（坡度、距离） $R_s, s = 1, \dots, S_n$; 车辆的具体参数。

输出: 每条道路上的一系列最优速度策略 T_n

初始化道路分段信息，设置速度初始变量 v_0, v_{end} 以及代价函数初始值 \mathcal{R} ，设置各个状态量的约束条件以及代价函数。

for $n = 1$ **to** N **do**

$V_{now} = v_0$

$\mathcal{R} = 0$

for $s = 1$ **to** S_n **do**

for v_i **in** V_{now} **do**

$V_{next} = \langle v_{min}, \dots, v_{max} \rangle$

对下一个路段中 R_s ，计算每一种可行速度 V_{next} 所要消耗的时间、能耗以及电池损耗

对每一种速度 V_{next} ，计算此阶段 s 代价方程 r ，并更新叠加到总代价方程

$\mathcal{R}(s_last, v_i)$ 。

记录更新的总代价方程的每种可行速度 v_i 情况，及其累计代价 $\mathcal{R}(s, v_i)$ 。

end

$V_{now} = \langle v_{min}, \dots, v_{max} \rangle$

在 V_{now} 的状态下，回溯查找累计代价 $\mathcal{R}(s, v_i)$ 中代价最低的情况和上一时刻的速度

$V'_s(v_{next})$ 。

end

遍历完毕，在最终的代价函数 $\mathcal{R}(S_n, v_{end})$ 中选取最优的值以及速度状态。

for $s = S_n$ **to** 1 **do**

反向回溯每段状态中最优代价函数和对应的速度值 $V'_s(v_i)$ 并填入最优速度轨迹序列 $\{T_n\}$ 。

end

end

当卡车得到强化系统的调度指令时，卡车根据已优化好的速度策略前往相应的目的地。根据上述讨论，现在我们在算法 2 中提出了将 DQN 与 DP 轨迹优化方法相结合的算法。

算法 2: 结合 DP 速度轨迹优化的强化学习调度方法

输入: 状态 s_t

输出: 动作 a

初始化记忆池 M 的容量 M_{max} 、动作价值函数的随机权重 θ ；设置环境仿真参数 $maxitr$ 、车辆参数 F 和仿真时间 TS 。

for $itr = 1$ to $maxitr$ **do**

 初始化模拟环境，获取初始状态值 s_0 。

for $t = 0$ to TS **do**

for $i = 1$ to F **do**

 获取当前状态观测值 s_t 。

if 车辆 T_i 需要分配目的地 **then**

 基于环境状态 s_t 按照 greedy 策略获取目的地动作 a_t

 在仿真环境中，按照算法 1 中的最优速度前往 a_t 目的地

 执行结束后计算奖励值 r_t 和下一时刻的状态 s_{tnext}

 存储数据 $\langle s_t, a_t, r_t, s_{tnext} \rangle$ 于记忆池 M 中

end

end

end

for $e = 1$ to E **do**

 从记忆池 M 中取出一批数据 $\langle s_t, a_t, r_t, s_{tnext} \rangle$

 计算目标的动作价值 $y_t = r_t + \gamma * \max_{atnext} Q(s_{tnext}, a_{tnext}; \theta')$

 计算误差 $err = y_t - Q(s_t, a_t; \theta)$

 更新 Q 网络的参数 $\theta' \leftarrow \theta + \nabla_{\theta} e^2$

end

end

3.4 仿真器设计

本文使用 Simpy 开发了一个用于矿山调度操作的仿真器。Simpy 是一个基于进程的离散事件仿真框架。装载点和卸载点被设计为具有固定容量和排队过程的资源。当卡车到达卸载点/装载点并需要调度时，状态空间将被获取并传递到神经网络中。Simpy 仿真器为开发调度策略提供了一个平台，使我们能够高效地测试不同的神经网络。图 3-8 说明了 Simpy 仿真器的结构图以及与调度学习器的交互过程。

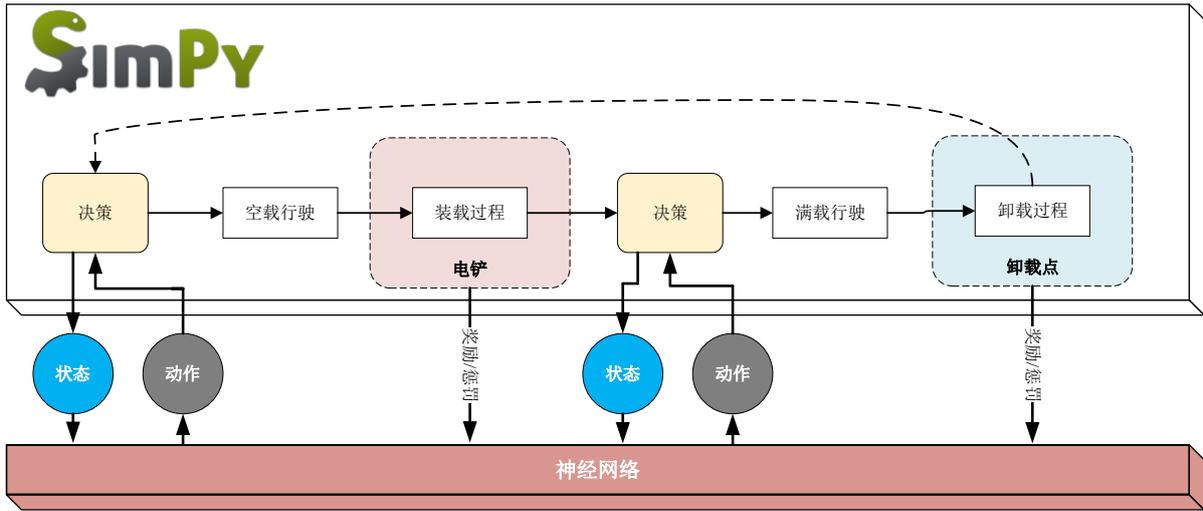


图 3-8 Simpy 仿真框架

Fig.3-8 The framework of Simpy simulation

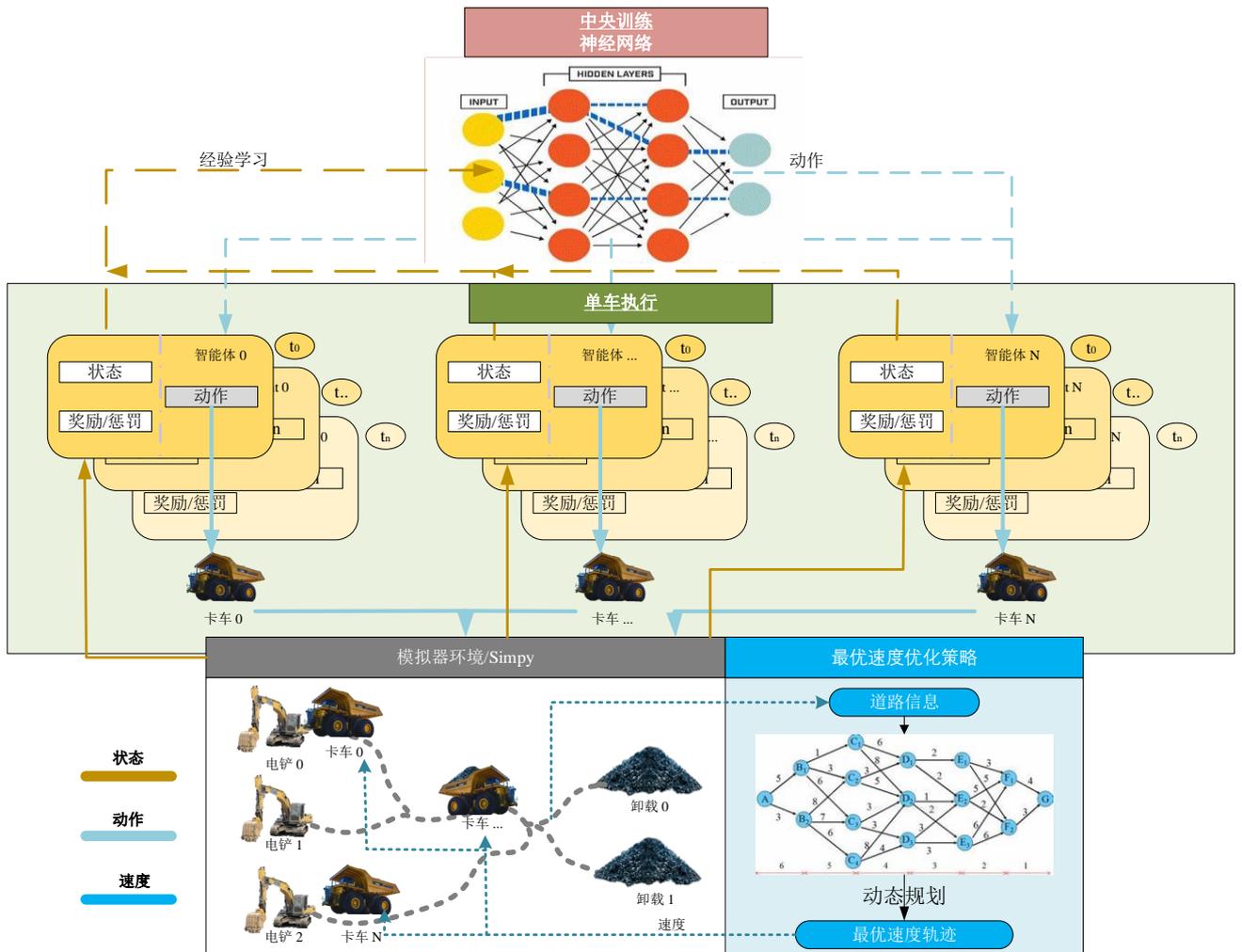


图 3-9 基于 DQN 的调度系统训练框架

Fig.3-9 Training framework of scheduling system based on DQN

由于本文使用的车队是同质的，因此装载和卸载等活动时间具有类似的特征，这些特征和能力是从实际数据中采样得到的。不同铲车和卸载点之间的运输时间消耗是通过动态规划算法优化的速度和轨迹得到的。图 3-9 说明了集成了 DP 轨迹优化的 DQN 调度系统的训练框架。为了研究所提出方法的能力，进行了大量实验，比较了在矿业中广泛使用的关键性能指标。

3.4.1 网络参数设置

该网络由三层组成，除了最后一层外，所有层都使用了 ReLU 激活函数。所有权重和偏置根据 PyTorch 的默认初始化进行初始化。优化算法采用 ADAM，学习率从 $1e-03$ 衰减到 $1e-05$ 。样本的批量大小 E 为 1024，记忆样本容量 M 为 $1e+07$ 。折扣因子 γ 为 0.9。特别地，采用变化的 epsilon-greedy 策略，从 80% 的随机行动几率衰减到 1%。

3.4.2 训练路网环境

正如图 3-8 和图 3-9 所示，该调度系统的训练平台部署在 Simpy 平台上，该智能调度系统集成了第三章所描述的多目标行驶速度策略优化方法。车辆前往调度的目的地时采用优化后的速度，默认车辆按照指定的速度轨迹前往目的地。在本调度模拟环境中假设所有道路和路口处的路面宽度足够宽，8h 作业时不考虑任何充换电、加油和维修等环节，在任务执行调度过程中不会产生任何冲突和拥堵，并且不考虑车辆在行驶过程中的道路横向偏差。

根据实验训练要求，我们采取国内某个实际的矿山生产场景作为训练环境，结合实际的道路信息，求解多目标下最优的速度轨迹，将最优的行驶时间作为仿真环境运行时间。实际的道路路网信息如下图 3-10 所示，实际道路里程如表 3-2 所示。

该道路数据是根据实际的路网提取得到，这套路网中包含三个转载点，两个卸载点，假设装载点的开采量最够充足并且开采速度能够满足不间断的装载任务。首先将道路的信息，包括高度、里程、坡度等信息，导入到多目标行驶速度策略优化算法中。算法原理见第二章中的图 2-14。其次，将优化后的速度和通过时间信息作为行驶时间的参数，进而将行驶时间作为训练环境的一部分。

3.4.3 车队参数设置

本章利用基于 Simpy 的矿山仿真器来训练智能体。如图 3-10 所示，训练环境设置

为 3 台铲车、2 个卸载点和 25 辆具有相同参数的电动卡车。电动卡车的具体动力学参数参照第二章的表 2-3，每辆卡车的容量为 45 吨，装载时间为 5 分钟，卸载时间为 1 分钟。初始化时，这 25 辆卡车被随机分配到 3 个装载点处，并具有不同的起始时间，假设在车辆行驶的过程中，不会产生任何冲突和横向偏差。为了模拟真实的作业时间，我们将每轮的仿真时间设置为 8 小时。

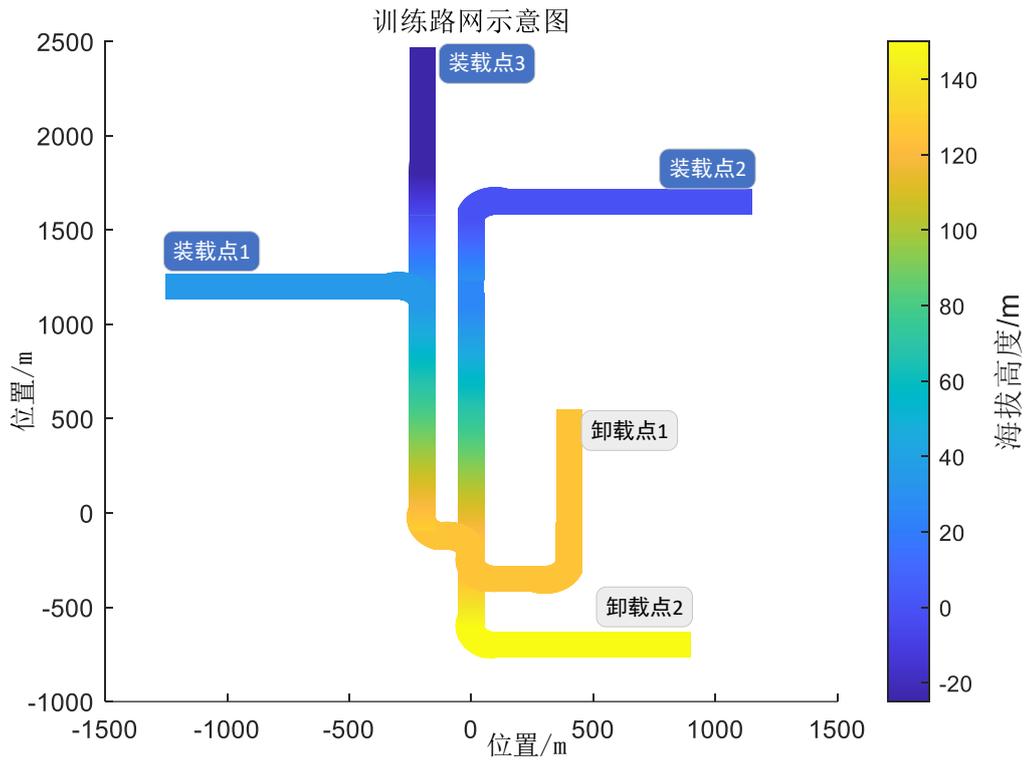


图 3-10 实际道路路网数据

Fig.3-10 The actual mining road network data

表 3-2 实际矿山道路运输里程表

Tab.3-2 The actual mining road network data

矿山道路里程表 (km)		
卸载点	卸载点 1	卸载点 2
装载点		
装载点 1	3.84	3.83
装载点 2	4.27	4.27
装载点 3	4.08	4.08

根据仿真实验的道路环境信息和选用的车辆动力学信息，我们求解在此种环境下的最优速度策略。正如第二章所述的多目标行驶速度策略的优化方法，我们将优化算法中

的目标方程式 2-17 的约束均简单设置为 1，多目标函数的归一化参考值 J_{Tmin} 、 $J_{E_{hmin}}$ 和 J_{enmin} 的值见本文 2.4.2。根据矿山最高速度 30km/h 限制，便可以求解到车辆在调度行驶中满载和空载过程中最优的行驶速度，具体的速度轨迹可以见图 3-11 和图 3-12。根据速度轨迹可以计算出车辆行驶时间，如表 3-3 和 3-4 所示，车辆的行驶时间便作为训练环境中车辆的运行任务的时间参数。于此同时，如表 3-5 和 3-6 所示，我们也可以通过计算得到整个车辆在执行最优速度策略速度的能量消耗，其中负数代表车辆在运输过程中采用了能量回收，最终出现了能量增加。

满载运输最优速度策略

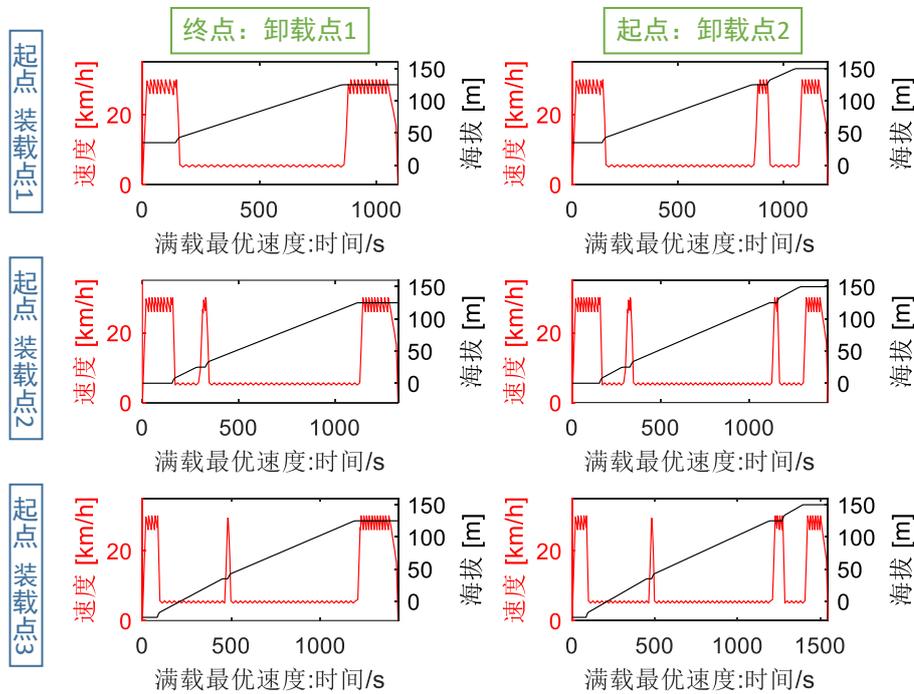


图 3-11 多目标优化后的满载速度轨迹

Fig.3-11 Full load velocity trajectory after multi-objective optimization

表 3-3 基于矿山道路优化后的满载运输行驶时间

Tab.3-3 Full load travel time after multi-objective optimization

优化后的矿山道路满载行驶时间 (s)		
卸载点	卸载点 1	卸载点 2
装载点		
装载点 1	1092.3	1211.4
装载点 2	1328.3	1446.4
装载点 3	1437.9	1546.2

空载返回最优速度策略

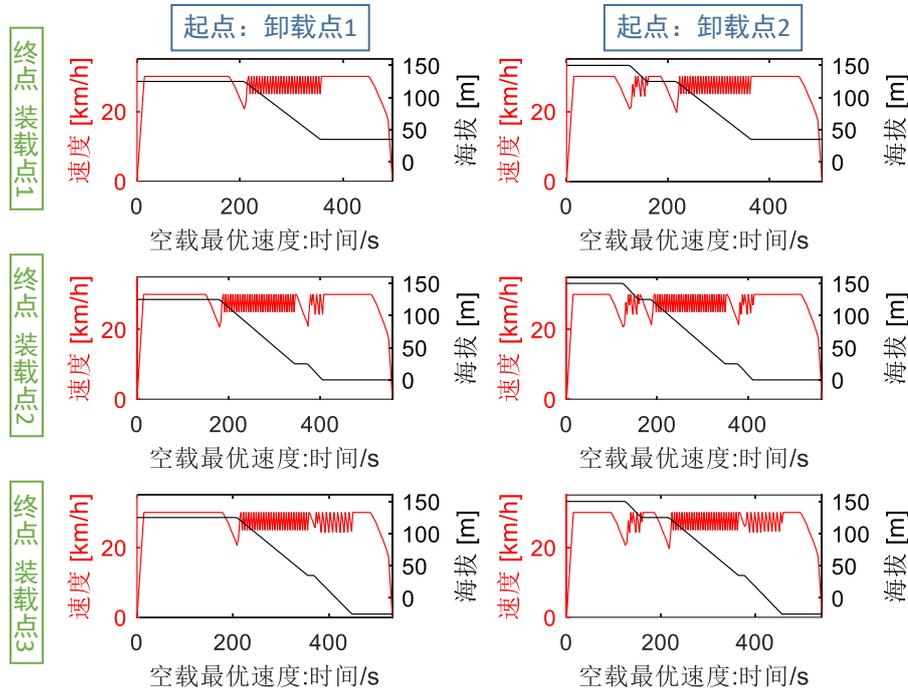


图 3-12 多目标优化后的空载速度轨迹

Fig.3-12 Empty load velocity trajectory after multi-objective optimization

表 3-4 基于矿山道路优化后的空载运输行驶时间

Tab.3-4 Empty load travel time after multi-objective optimization

优化后的矿山道路空载返回行驶时间 (s)		
卸载点	卸载点 1	卸载点 2
装载点 1	496.0	504.9
装载点 2	557.1	563.8
装载点 3	532.5	541.5

表 3-5 基于矿山道路优化后的满载运输的能量消耗

Tab.3-5 Full load energy consumption after multi-objective optimization

优化后的矿山道路满载行驶能量消耗 (kwh)		
卸载点	卸载点 1	卸载点 2
装载点 1	35.23	42.53
装载点 2	46.44	53.70
装载点 3	53.73	60.89

表 3-6 基于矿山道路优化后的空载返回运输能量消耗

Tab.3-6 Empty load energy consumption after multi-objective optimization

优化后的矿山道路空载返回行驶能量消耗 (kwh)		
卸载点	卸载点 1	卸载点 2
装载点 1	-6.20	-9.28
装载点 2	-9.99	-12.99
装载点 3	-12.25	-16.31

3.5 结果验证

如前文所述，该方法的过程可以在图 3-9 中展示。我们对 DQN 进行了 10 万次步骤的训练，并观察到总平均奖励的收敛。训练期间的总平均奖励如图 3-13 所示。可以观察到，在大约 10,000 个回合后，DQN 的总平均奖励达到了一个相对稳定的状态。

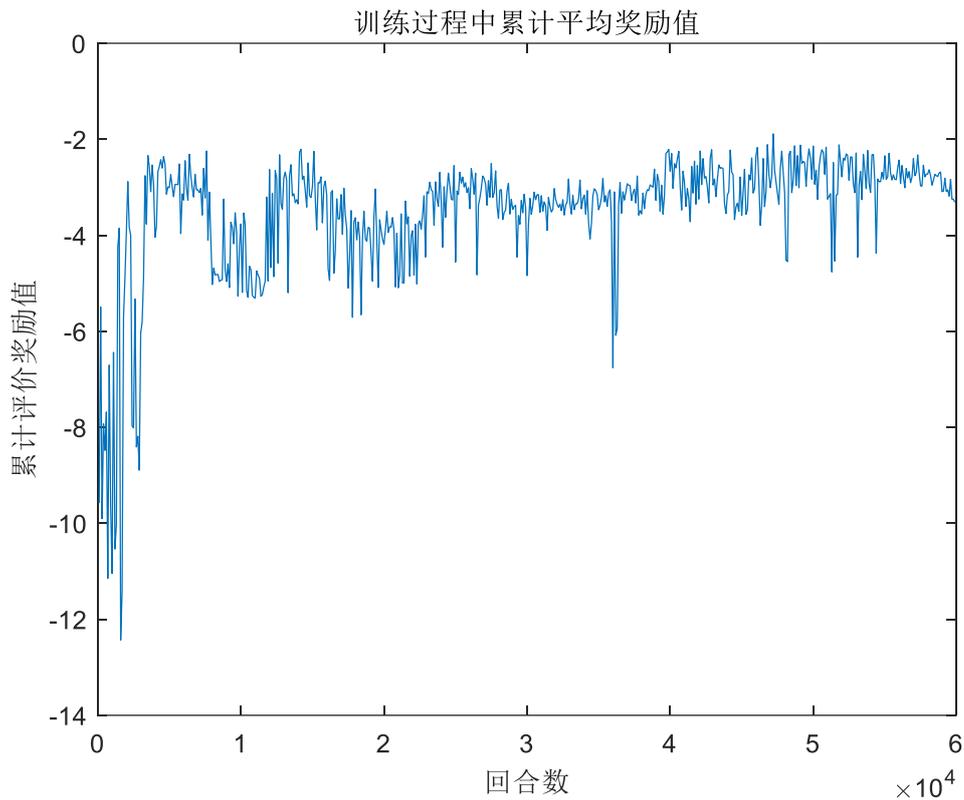


图 3-13 强化训练过程中的平均奖励值

Fig.3-13 The average reward value during reinforcement training

经过训练的 DQN 模型随后在模拟器中进行了 10 次实验，每次实验仿真持续 8 小时，涉及到一支由 25 辆卡车组成的车队。模拟验证环境中的路线和装、卸载点之间的情况和训练场景一致，详细数据可以见表 3-7 所示。各个站点运输完成量的详细数据见

表 3-8 所示，通过各个站点的运输量对比可以看出该运输调度系统下发任务的均匀合理程度。各个装卸载点运输量的标准差表示了任务分配的离散程度，不同装载点和卸载点的运输量可以不均匀的任务分配会导致拥堵和排队的情况，生产效率更低。

总体而言，结合速度轨迹优化的 DQN 调度系统经过训练后，算法能够生成高效的动态调度策略。如表 3-7 和 3-8 所示，车队能够完成正常的班次运输作业，对矿山的生产具有一定的指导意义，具体的运输性能和对比分析将会在第五章中进一步展开讨论。

表 3-7 基于速度优化的 DQN 调度系统的测试结果 1

Tab.3-7 Test results of DQN dispatching system based on speed optimization (1)

基于速度优化的 DQN 调度系统测试结果		
实验序号	8h 班次内 车队总运输量 (t)	8h 班次内 车队总能耗 (kwh)
1	10350	9259.62
2	10305	9437.18
3	10575	9534.87
4	11655	10280.21
5	10935	9828.50
6	10980	9761.00
7	10980	9799.30
8	10710	9769.99
9	10305	9323.21
10	10935	9869.16

表 3-8 基于速度优化的 DQN 调度系统的测试结果 2

Tab.3-8 Test results of DQN dispatching system based on speed optimization (2)

8h 班次内各个运输站点的运输完成量以及均匀程度							
实验序号	装载点 1 (t)	装载点 1 (t)	装载点 1 (t)	装载量 标准差	卸载点 1 (t)	卸载点 1 (t)	卸载量 标准差
1	3465	3915	2970	472.7	5625	5310	222.7
2	3960	3960	2385	909.3	5580	5310	190.9
3	3915	3960	2700	714.8	5670	5400	190.9
4	3960	3915	3780	93.7	6345	5895	318.2
5	3960	3915	3060	507.1	5940	5580	254.6
6	3960	3915	3105	481.2	5940	5625	222.7
7	3960	3915	3105	481.2	5895	5670	159.1
8	3915	3915	2880	597.6	5805	5490	222.7
9	3960	3960	2385	909.3	5535	5310	159.1
10	3960	3915	3060	507.1	5940	5580	254.6

3.6 小结

本文通过将调度问题形式化为多智能体强化学习问题。我们基于 `Simpy` 创建了一个高度适应性的基于事件的采矿模拟器，并结合真实矿山的参数。随后，我们提出了一种新颖的综合方法，将深度强化学习与动态规划轨迹优化相结合，用于在矿山运输中调度无人驾驶卡车。经过将近 10 万次的训练之后，在相同的训练环境中进行模拟试验。十次试验数据表明该系统能够生成高效的动态调度策略，车队能够完成正常的班次运输作业，完成无人运输调度。第五章将针对运输性能设置全面的评价体系，对比分析将会在第五章中进一步展开讨论和传统调度模式的优势。

然而，学习型调度系统依然面临限制，复杂的状态空间导致计算强度和学习时间增加。此外，真实的矿山现场通常包括许多交叉口，卡车可能会相互交叉或排队等待。

4 矿山无人驾驶车队运输评价体系

4.1 引言

在矿山运输系统中，矿业企业的运输成本非常高，车辆运输是整个矿山成本中最重要的部分。一般情况下运输所产生的各种费用占了矿山总体生产成本的一半，甚至在一些大型露天矿中占到60%以上^[2]。因此，随着时代发展和社会的进步，矿山场景下的车辆运输越来越受到各方的重视，我们急需建立起一套符合现有矿山运输效益的评价体系。矿山车辆运输评价体系直接受到评价指标和经济成本核算的影响，包括的指标数目和指标参数较多。本章主要根据矿山行业特点选择恰当的指标，根据实际工程经验建立起矿山车辆运输的评价体系，一是对本文提出的速度优化的学习型调度系统进行全方面的评估，二是为矿山企业的决策和管理提供更全面的支持。

4.2 矿山无人驾驶车队运输调度评价体系概述

随着技术的不断发展，矿山车辆运输正逐渐受益于新一代技术的应用。例如，自动驾驶技术、智能运输系统和远程监控技术等正在被引入，以提高运输效率、安全性和可持续性。矿山车辆运输评价体系是用于评估和衡量矿山车辆运输管理绩效的一套指标和方法体系。它旨在提供一个系统性的框架，帮助矿山管理者和相关利益相关者评估和监控车辆运输活动的效率、安全性、可持续性以及其他关键方面。随着数据收集和分析技术的进步，矿山车辆运输评价体系正朝着更科学、细致和全面的方向发展。

一个科学全面的矿山车辆运输评价体系不仅可以很好的评价目前的运输方案水平和成果，也可以有效的指导矿山运输业未来的发展。在建立的过程中，也可以不断总结分析现有的不足，进一步推动矿山智能运输的发展。建立矿山车辆运输评价体系对于提高绩效管理、环境保护、安全管理、决策支持和沟通等方面都具有重要意义。

构建矿山车辆运输评价体系时，遵循以下原则：

目标导向原则：评价体系应该有明确的目标，与矿山车辆运输的战略目标和业务需求相一致，以确保评价的准确性和有效性。

综合性原则：评价体系应该综合考虑多个方面的因素，如运输效率、成本控制、安全性、环境影响等，并针对不同因素采取相应的改进措施。

可量化原则：评价体系中的指标应该可量化和可测量，以便收集和分析相关数据，具有可行性和可操作性。

可比性原则：评价体系应该具有可比性，即不同时间段和不同矿山之间的运输绩效可以进行比较和对比。为了实现可比性，评价体系应该采用一致的指标和标准，并确保数据收集和分析的一致性和准确性。

参与性原则：评价体系的构建应该充分考虑各方面的参与和意见。包括管理层、运输人员、环境专家等在内的相关利益相关方应该参与进来，提供他们的专业知识和经验，以确保评价体系的合理性和可行性。

构建矿山车辆运输评价体系首先要明确矿山车辆运输评价体系的和目标目的。确定评价体系的主要目标是为了提高哪些方面的绩效，例如运输效率、成本控制、安全性、环境影响等。根据评价目标，确定适合的评价指标。评价指标应该能够量化矿山车辆运输的关键表现。例如，可以考虑运输燃料消耗、运输成本、运输量等指标。确保指标具有可量化、可测量和可比较性。确定评价指标后，需要收集相关的数据和信息来支持评价体系的实施。使用适当的分析方法和工具，对收集到的数据和信息进行分析和评估。陈奇^[98]的研究关注铁路货车车辆运用方面，构建了涵盖货物运输量、运用车保有量、工作量、货车周转时间和载重利用等指标的评价体系。而徐翔^[99]的工作则针对管道运输的低碳绿色发展，采用云模型构建了相应的评价指标体系。目前，还没有学者详细的描述矿山领域运输发展的指标体系。

本章根据矿山生产运输的理论基础，联系国内外管道运输的发展历程化及现阶段管道运输现状，我们重点对矿山运输中的运输效率、成本控制、车辆利用率、运输的均匀度和安全稳定性这五个方面进行梳理总结。

4.2.1 运输效率指标

矿山运输效率是指在矿山生产过程中，通过合理的运输管理和控制措施，实现高效、快速、安全地运输到目的地的能力和水平。它可以提高生产效率和产量，减少运输时间、提高车辆利用率和载重率。

1) 货物运输总量

货物运输总量指在一个班次内，矿山运输的货物总质量，与装车数、和载重量息息相关，直接反映了矿山运输某一时段内的货运组织情况。

2) 平均运输周转时间

矿山车辆运输的平均运输时间是衡量从矿山采矿现场到目的地，再次返回到装载现场的货物运输所需的平均时间。它表示了从开始运输到完成运输的平均时间，包括装载、运输和卸载等环节。较短的平均运输时间意味着更高的效率，因为货物可以更快速地从

矿山到达目的地。

4.2.2 成本控制指标

矿山运输的过程中，有效的控制和管理成本是矿山企业提高运营效益和盈利能力的重要任务。矿山的运输成本跟矿山的规模、距离、运输量、运输方式以及当地劳动力和燃料相关。

1) 运输总成本

衡量在一个班次运输过程中所产生的直接成本，包括燃料、人力、车辆维护、保险等费用。

2) 单位货物运输成本

衡量每单位货物在该矿山内运输所需的成本，等于运输总成本除以运输总重量。

3) 燃料成本

衡量单位货物运输所需的燃料消耗量，这涉及燃料的消耗，如柴油、汽油或电力等。燃料成本是矿山运输的主要费用之一。

4) 维修成本

衡量车辆维修和保养所需的成本。

4.2.3 单车利用率指标

车辆利用率用于评估车辆的闲置水平，充分利用车辆资源，过低的车辆利用率证明车辆冗余，带来效率低下，成本升高。

1) 单车闲置时间

闲置时间包括排队时间和空闲时间，评估车辆在运输任务中的实际工作时间与闲置时间的比率。提高车辆利用率可以减少单车闲置时间，优化资源利用，从而降低成本。

2) 单车运输量

每个车辆在规定班次内的运输量，用于比较单个车辆的运输完成量。

3) 单车运输均匀度

车辆个体运输均匀度可以通过离散程度来表示，单辆卡车运输量的变异系数（Coefficient of Variation，CV）是表示与车队平均产量相关的离散程度的指标。我们可以评估每辆车的运输量是否均匀分配。单辆卡车生产的变异系数越高，意味着有些卡车将处于闲置状态，而其他卡车则会更忙碌，这与我们在实际矿山现场的预期不符。其离

散度计算公式为式 4-1，其中 σ 和 μ 分别表示卸载站点个体的完成比例的标准差和均值， TP_t^M 代表在 t 时刻时，车队中的 M 车累计运输量。

$$CV = \frac{\sigma([TP_t^1, \dots, TP_t^M])}{\mu([TP_t^1, \dots, TP_t^M])} \quad (4-1)$$

4.2.4 运输均匀度指标

评估车辆运输系统对需求变化和调整的灵活性和响应能力。衡量车辆运输系统能否自主的应对和缓解交通拥堵的能力。

1) 道路使用均匀度

道路使用频次的变异系数（CV）显示了道路的使用或交通情况的不同水平。变异系数越低，所有道路的利用率和效率就越高。通过追踪和分析累积的行程次数，可以评估道路的受欢迎程度和效率。可以参考式 4-1，道路使用均匀度可以通过每条道路使用次数的标准差 σ 和均值 μ 的比值获得。

2) 装载点/卸载点的运输量均匀度

卸载点/装载点产量的变异系数（CV）代表了不同位置的完成情况的均匀程度。理想状态下，不同位置的生产速度应该保持一致，特别是对于装载点而言。因此，我们采用卸载点/装载点位置的生产量变异系数来评价生产节奏的匹配度。同样，各个位置的运输量的均匀度可以参考其他均匀的计算方式，通过各位置运输量的标准差 σ 和均值 μ 的比值获得。

4.2.5 安全稳定性指标

学习型的矿山调度方法的安全稳定性指的是该方法在面对各种不确定性和变化时的稳定性和可靠性。在矿山调度中，存在许多随机性和动态性的因素，尤其是车辆故障因素会对调度决策产生影响。

1) 事故率

事故率评估无人驾驶车队发生事故的频率和严重程度，包括车辆、人员以及周围环境事故。比如车辆的追尾、紧急停车、意外停止、车队通信故障等行为。定期检测并及时比较事故率，以便发现潜在的安全问题。

2) 自动驾驶稳定性

在车辆进行运输作业的过程中，超速、超车、转换目的地等违规行为属于自动驾驶车辆的异常行为。定期使用车载监控系统进行统计分析，及时发现并纠正错误。

3) 调度系统鲁棒性

无人驾驶车队调度系统的鲁棒性包括应急响应能力和环境适应能力。当面临突发事件或设备故障时，学习型调度系统能否对车队数目变化时依然保持稳定的性能。通过在仿真环境中进行实验和测试，可以评估该方法在各种情况下的性能表现。通过与基准方法对比鲁棒测试下的性能表现，可以帮助确定该方法在实际矿山环境中的可行性和适用性，从而为实际应用提供可靠的决策支持。

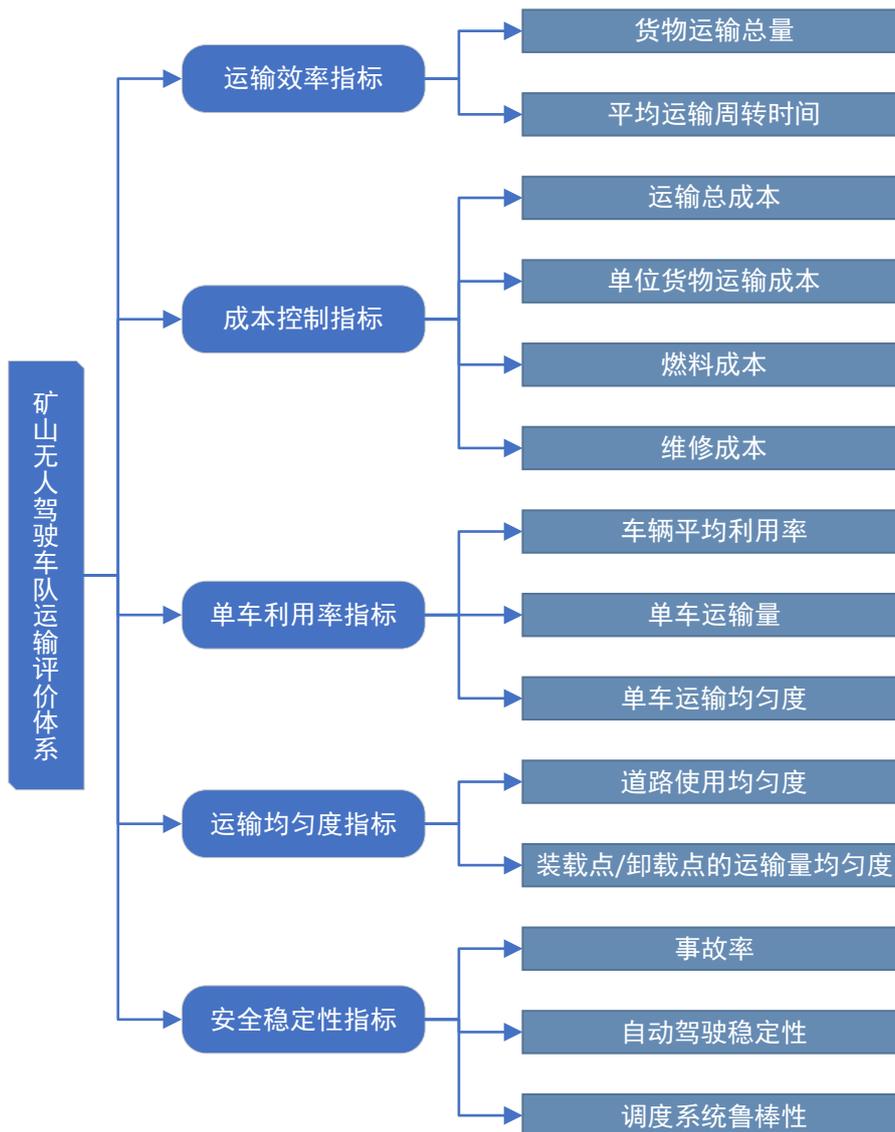


图 4-1 矿山无人驾驶车队运输评价体系

Fig.4-1 Evaluation system for mine autonomous fleet transportation

综上所述，根据科学的原则和方法，从货物运输效率、成本控制、车辆利用率以及运输的灵活适应性四个方面初步构建矿山无人驾驶车队调度考核体系如图 4-1 所示。

4.3 学习型矿山无人驾驶车队调度系统评价

本节基于以上评价指标体系对第四章提出的基于强化学习的矿山无人驾驶车队调度系统进行分析。需要说明的是本研究提出的矿山车队调度系统，没有实际生产运输的结果，只能在仿真环境中进行测试，因此无法全面评估学习型矿山无人驾驶车队调度系统的安全性。我们不考虑自动驾驶车辆在行驶过程中的冲突、故障和横向偏差问题，仅评估调度系统的鲁棒性。

首先，选定合适的路网环境，搭建相应的矿山道路环境，并进行可视化。利用第三章中的多目标行驶速度策略的优化方法求解该路网中最优速度，作为车辆在智能调度中的行驶速度。其次，选用合适的对比基准方法，将基于强化学习的矿山无人驾驶车队调度系统和基准方法部署在相应的矿山模拟环境中，运用本章提出的矿山车辆运输调度评价考核体系定量对比分析各指标的表现；最后，根据考核指标对各个环节进行分析，绘制图表，得出结论和建议。

4.3.1 参考基准模型

将传统的基准方法与本文章提出的结合速度优化的 DQN 强化学习调度方法进行比较，以评估所提出系统的性能。实验中使用了固定卡车分配策略（Fixed truck assignment strategy, FS）作为比较参考。固定的卡车分配策略在一次作业中将卡车分配到固定的铲车和卸载点之间的路线上，被广泛接受并在实践中持续使用^[45,96]。因为大多数研究人员普遍认可固定卡车分配方法仍在实践中使用，并且大多数自动驾驶卡车在矿山运输中以恒定速度行驶，并不会考虑实际的能耗优化，也不会针对实际的地形做出速度策略的改变。

4.3.2 仿真环境设置

1) 路网设置

本章利用基于 Simpy 的矿山仿真器来训练智能体，我们依然采取国内某个实际的矿山生产场景作为训练环境，对调度系统评价的路网环境和训练的环境一致。该路网有三个装载点和两个卸载点，详情请见第三节的 3.4.2。

2) 车队参数设置

本次仿真的车队包括 25 辆相同的电动车辆，电动卡车的具体动力学参数参照第三章的表 3-3，每辆卡车的容量为 45 吨。并且车队中的无人驾驶车辆在运输中均假设不会

产生任何冲突，不考虑行驶过程中的横向偏差。初始化时，这 25 辆卡车被随机分配到 3 个装载点处，并具有不同的起始时间。为了模拟真实的作业时间，我们将每轮的仿真时间设置为 8 小时。

对于结合速度优化的强化学习智能调度系统来说，本系统调度行驶速度是经过第二章的优化方法计算而得到的，他所采用的速度是随地形和成本需求优化而来的。此处验证采用的速度策略是和训练场景下的速度一样，我们将优化算法中的目标方程式 2-17 的不同因素均简单设置为 1，求解到车辆在调度行驶中满载和空载过程中最优的行驶速度，具体的行驶速度轨迹见图 3-11 和图 3-12，详细的行驶时间和能耗数据见表 3-3 至 3-6。

对于固定卡车的分配调度策略来说，该策略中的无人卡车在前往卸载点和装载点位置是固定的，且以恒定速度运行，速度不会做任何的更改或优化。考虑到实际矿山中车辆特性、地形限制以及工程经验，无人驾驶的车辆在满载和空载的行驶过程中分别保持 5 km/h 和 20 km/h 的速度，这些速度差异的实施是根据每个位置的具体要求和条件来确保安全运营。根据表 2-3 动力学特性和第 3.4.2 中的路网环境，车辆在道路恒定速度行驶过程中的满载和空载的油耗和行驶时间信息便可以计算得到，详细数据见表 4-1 至 4-4，其中能量消耗为负值表示在运输过程中采取了能量回收策略，能量有所增加。

表 4-1 恒定速度 5km/h 下的满载运输行驶时间

Tab.4-1 Full load travel time at a constant speed of 5km/h

车辆以恒定速度 5km/h 满载行驶时间 (s)		
卸载点	卸载点 1	卸载点 2
装载点 1	2779.2	2779.2
装载点 2	3096.0	3096.0
装载点 3	2952.0	2952.0

表 4-2 恒定速度 20km/h 的空载运输行驶时间

Tab.4-2 Empty load travel time at a constant speed of 20km/h

车辆以恒定速度 20km/h 空载返回行驶时间 (s)		
卸载点	卸载点 1	卸载点 2
装载点 1	694.0	694.0
装载点 2	774.0	774.0
装载点 3	738.0	738.0

表 4-3 恒定速度 5km/h 的满载运输的能量消耗

Tab.4-3 Full load energy consumption at a constant speed of 5km/h

车辆以恒定速度 5km/h 满载行驶能量消耗 (kwh)		
卸载点	卸载点 1	卸载点 2
装载点		
装载点 1	37.08	44.52
装载点 2	48.68	46.15
装载点 3	55.61	63.05

表 4-4 恒定速度 20km/h 的空载返回运输能量消耗

Tab.4-4 Empty load energy consumption at a constant speed of 20km/h

车辆以恒定速度 20km/h 空载返回行驶能量消耗 (kwh)		
卸载点	卸载点 1	卸载点 2
装载点		
装载点 1	-4.03	-6.43
装载点 2	-6.87	-9.27
装载点 3	-9.50	-11.90

4.3.3 指标分析对比

测试环境的参数设置完成后,将训练好的调度系统以及基准调度模型在矿山模拟器中分别进行 10 次试验,在试验过程中记录更全面综合的数据。每次试验的仿真时长均设置为 8 小时,与实际矿山现场操作保持一致,并且试验调度的车队为 25 辆相同的电车,详细对比数据如下。

1) 运输效率指标分析

运输效率指标包括一个班次内货物运输总量和运输周转时间。我们对这两种调度方法十次试验统计数据如表 4-5 所示,并绘制折线图 4-2。根据折线图可以直观的看出,速度优化后的学习型调度策略不论在运输周转时间方面,还是货物运输总量方面,均表现出更优秀的特点。采用速度优化的学习型调度策略拥有更高的产量以及更短的周转时间,表明运输效率更高。优化后的行驶速度会比定速更快,在一些特定的道路下会选择适当的降速,正常情况下会选择高速运行;但恒定速度的运输模式,管理方会考虑到全局安全行驶的因素人为的将速度稳定在满载 5km/h,空载 20km/h,避免在上下坡加减速的过程中产生危险。

2) 成本控制指标分析

矿山的运输成本包括很多,是一个复杂的经济问题。他的运输成本跟矿山路径的大

小、车辆行驶策略等都有关系。本试验中的调度过程是在虚拟仿真环境中进行，不能反映出真实的成本数据。在此次成本控制指标分析的过程中，我们采用的车辆均为电动车辆。因此，我们只考虑车辆能耗成本，忽略其他产生的成本。具体的能耗成本和单位运输能耗成本数据见表 4-6，并绘制成折线图 4-3。

表 4-5 运输效率指标统计数据
Tab.4-5 Test results of transport efficiency.

8h 班次内运输效率指标统计数据				
实验序号	学习型调度系统货物总运输量 (t)	固定调度策略货物总运输量 (t)	学习型调度系统平均运输周转时间 (s)	固定调度策略平均运输周转时间 (s)
1	10350	7065	3130.43	4585.99
2	10305	6885	3144.10	4705.88
3	10575	6750	3063.83	4800.00
4	11655	7065	2779.92	4585.99
5	10935	7065	2962.96	4585.99
6	10980	7065	2950.82	4585.99
7	10980	7065	2950.82	4585.99
8	10710	6975	3025.21	4645.16
9	10305	7065	3144.10	4585.99
10	10935	7065	2962.96	4585.99

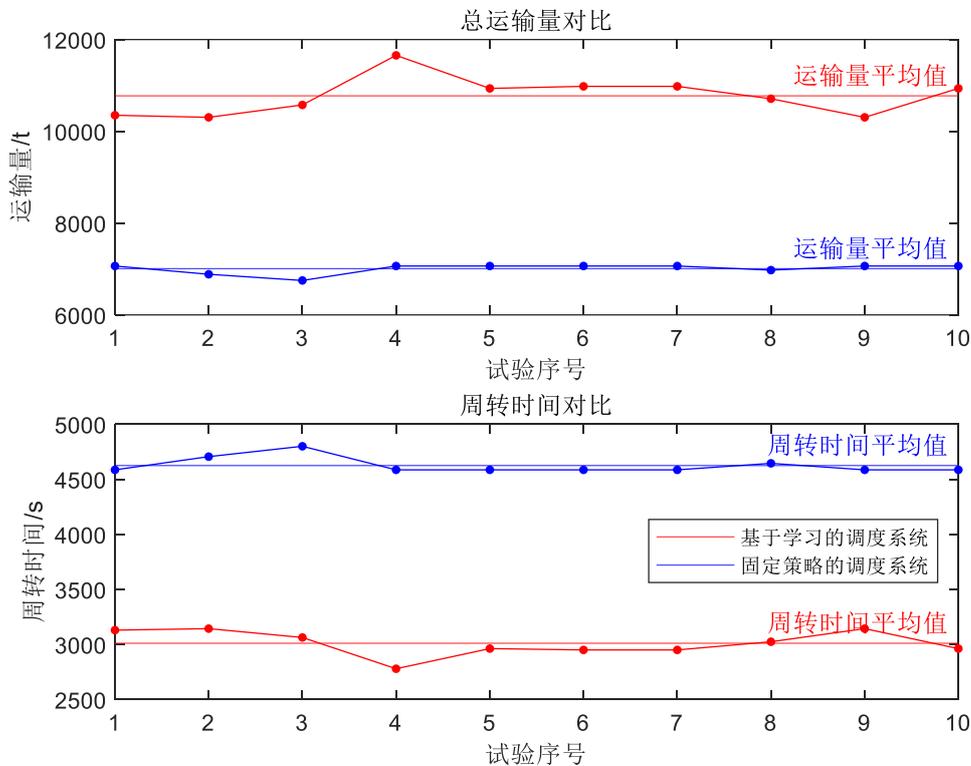


图 4-2 8h 班次内运输效率指标对比

Fig.4-2 Comparison of transport efficiency in 8h work shift

表 4-6 成本控制指标统计数据

Tab.4-6 Test results of cost management.

8h 班次内成本控制指标统计数据				
实验序号	基于学习的调度系统 运输总能耗(kWh)	固定调度策略 运输总能耗(kWh)	基于学习的调度系统 单位货物运输能耗 (kWh/t)	固定调度策略 单位货物运输能耗 (kWh/t)
1	9259.62	6804.93	0.89	0.96
2	9437.18	6856.32	0.92	1.00
3	9534.87	7159.22	0.90	1.06
4	10280.21	7035.39	0.88	1.00
5	9828.5	7129.95	0.90	1.01
6	9761	6881.07	0.89	0.97
7	9799.3	6829.61	0.89	0.97
8	9769.99	7097.92	0.91	1.02
9	9323.21	6985.35	0.90	0.99
10	9869.16	6990.93	0.90	0.99

从图 4-3、表 4-6 数据可以直观的看出，基于学习型的调度系统在一个班次内会产生更多的总能耗。但是系统采用了速度优化策略，使得车队的总运输次数增高，从而累计的导致了高能耗，但在单位货物运输能耗上表现很优异，均低于固定策略的单位运输能耗。因此，采用速度优化后的学习型调度系统虽然会提升速度导致运输量增大，累计能耗更高，但平均后的每吨运输能耗却降低约 10%，这更有利于成本的控制。

3) 单车利用率指标分析

单车利用率用来评估车辆的闲置水平，它包括单车闲置时间、单车的运输量以及单车的均匀度。通过检测单车的利用率水平，管理者可以有针对性的改善车队，实现利益最大化。

本研究一共做了 10 次试验，首先分别统计每辆车在工作循环 8h 中的排队时长和运输量。由于数据量过多，我们仅抽取其中一次试验的样本绘制为柱状图，如图 4-4 所示。然后，根据每辆车的实际运输量求解平均值和标准差，根据公式 4-1 求得变异系数，从而得到均匀度。由于这两种调度环境车队数量均为 25 辆车，单车运输量的平均值等同于运输效率中的货物运输总量。因此我们不对单车运输量过多对比，主要分析单车运输量的均匀度。按照此方法将 10 次试验的单车平均闲置时间、单车运输量均匀度绘制成图表，详细对比数据可以见表 4-7 和图 4-5。

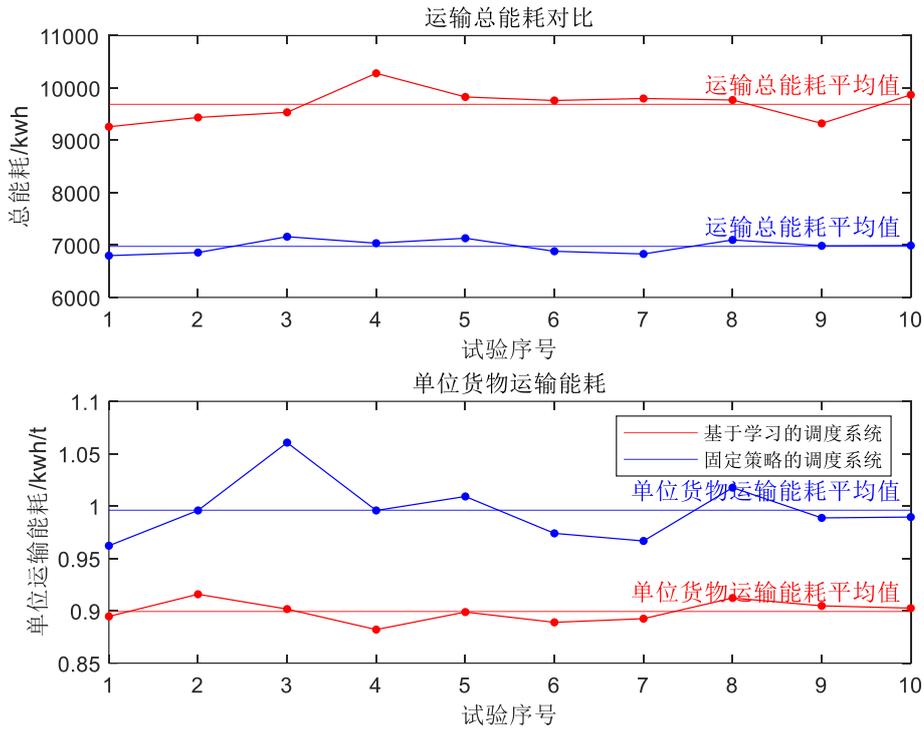


图 4-3 8h 班次内成本控制指标对比

Fig.4-3 Comparison of cost management in 8h work shift

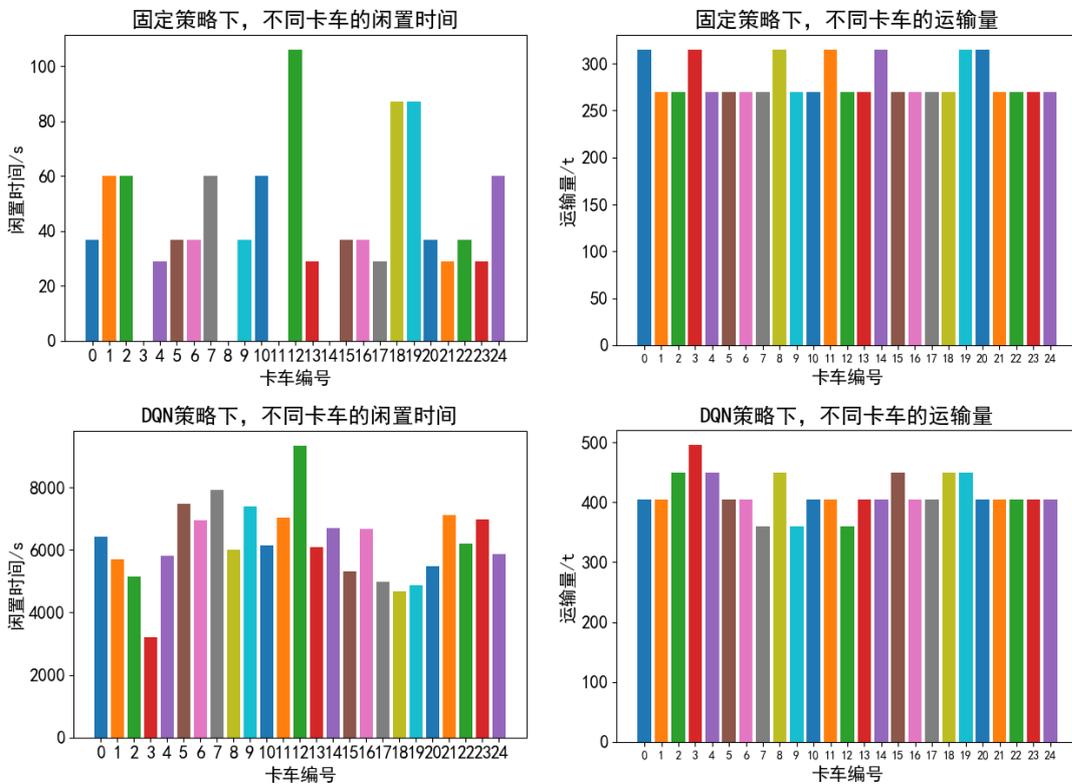


图 4-4 某一班次内单车闲置时间和运输量统计

Fig.4-4 Idle time and traffic volume of single vehicle in a shift

表 4-7 单车利用率指标统计数据

Tab.4-7 Test results of vehicle utilization.

8h 班次内单车利用率指标统计数据				
实验序号	基于学习的调度系统 平均闲置时间/s	固定调度策略 平均闲置时间/s	基于学习的调度系统 单车运输均匀度	固定调度策略 单车运输均匀度
1	6224.6	40.8	0.075	0.071
2	6155.7	85.8	0.080	0.107
3	5802.1	45.0	0.067	0.082
4	3849.1	64.8	0.046	0.071
5	5037.9	72.6	0.068	0.071
6	5058.8	68.1	0.060	0.071
7	5081.0	56.0	0.060	0.071
8	5293.7	76.7	0.079	0.091
9	6212.3	56.6	0.067	0.071
10	4925.2	59.8	0.046	0.071

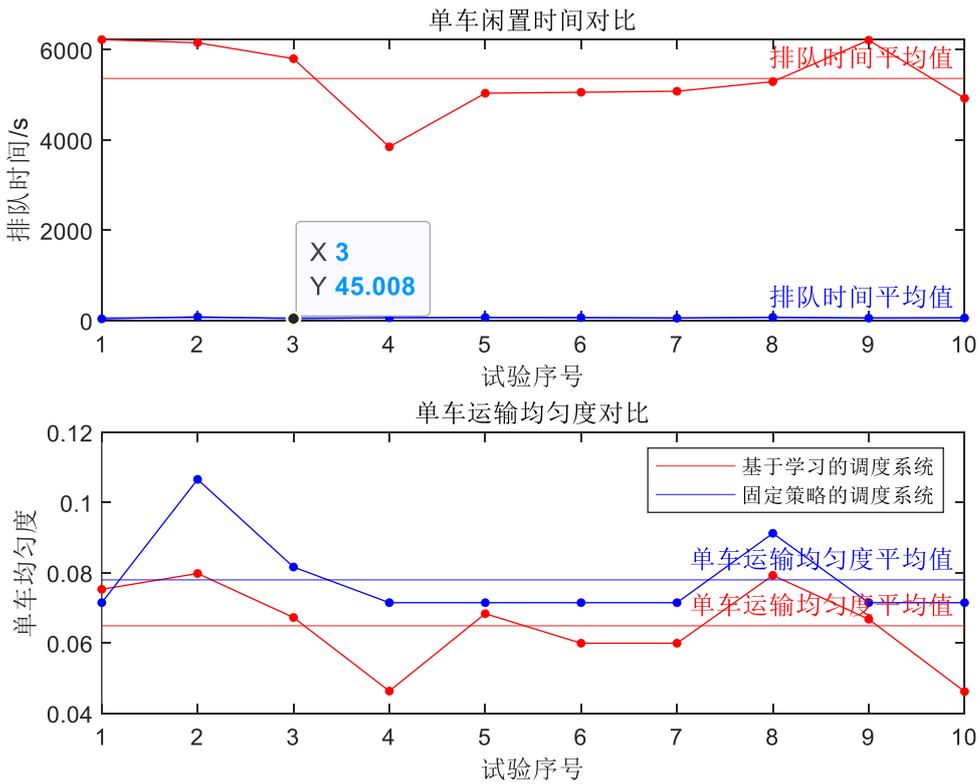


图 4-5 8h 班次内单车利用率指标对比

Fig.4-5 Comparison of vehicle utilization in 8h work shift

正如图 4-5 所示，在单车利用率方面，集成速度优化的学习型调度系统有更多的闲置时间，25 辆车连续 8h 作业时平均每辆车闲置 5364s，而规则的调度系统的闲置时间则仅有 63s。闲置时间过长的原因主要是优化后的速度过高，行驶时间对比策略采用

的恒定速度（满载和空载的行驶过程中分别保持 5 km/h 和 20 km/h 的速度）快，导致在到达装卸载的时间会产生冲突，产生排队闲置的情况更多。而对于单车运输均匀度方面，学习型的调度系统的单车运输量变异系数更低，这说明分配的任务更加平均，更适合生产调度。针对此种现象，应该结合实际情况减少速度优化的学习型调度系统中车队数量，或者降低优化后的速度，使其行驶时间更能匹配车辆装卸载的时间，减少排队的可能性。

4) 运输均匀度指标分析

运输均匀度指运输系统根据道路的需求来实时调整运输的任务，可以自主的缓解交通拥堵，他主要包括道路使用均匀度和各装载点/卸载点运输量的均匀度。首先，在试验过程中记录每条道路的使用次数和各个点的运输量，根据每辆车的实际数据求解平均值和标准差，然后根据公式 4-1 求得变异系数，从而得到均匀度。我们抽取其中一次试验的样本数据，包括各条道路的使用次数、各点位装卸载点的运输量，绘制成图 4-6。按照此方法将 10 次试验的各个点运输量均匀度、道路使用次数均匀度绘制成图表，详细对比数据可以见表 4-8 和图 4-7。

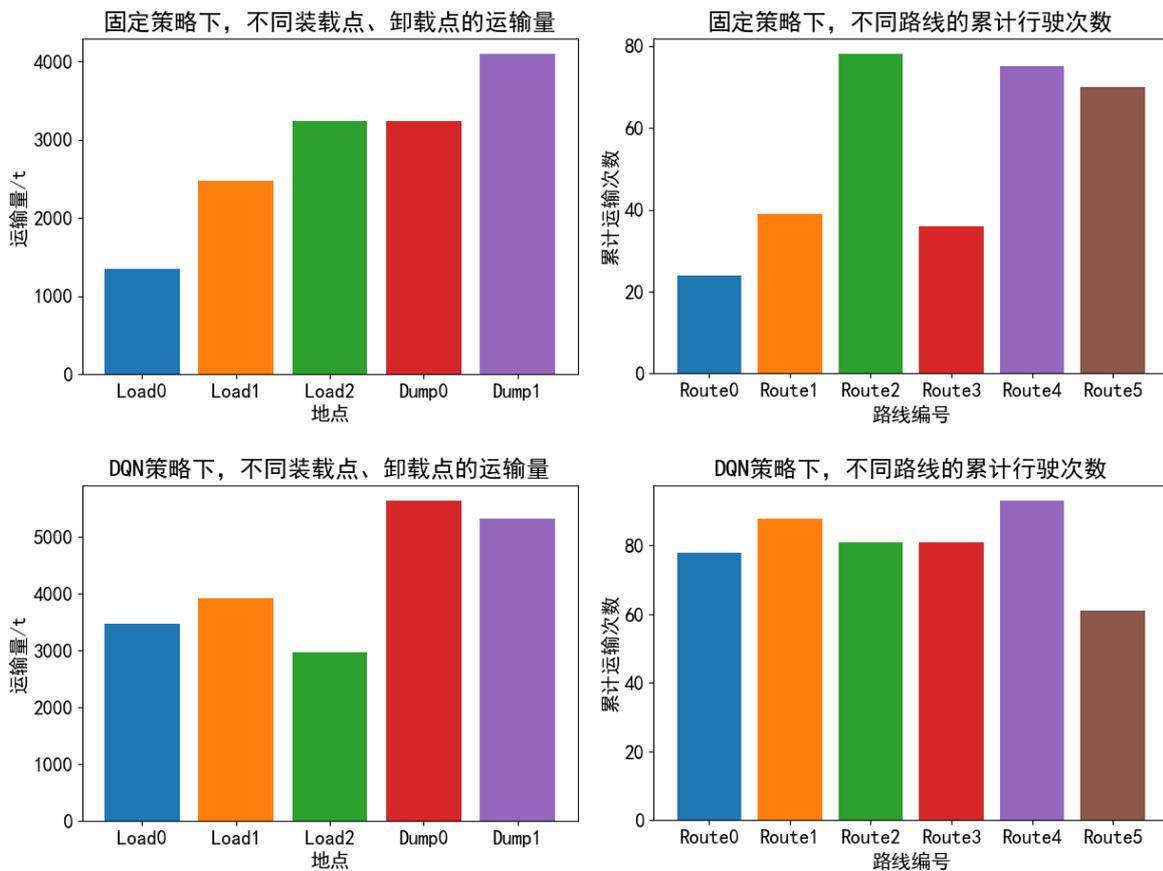


图 4-6 某一班次内不同地点运输量和道路使用量统计

Fig.4-6 Traffic volume and road usage at different locations in a shift

表 4-8 运输均匀度指标统计数据

Tab.4-8 Test results of transportation uniformity.

8h 班次内运输均匀度指标统计数据										
序号	1	2	3	4	5	6	7	8	9	10
学习型调度系统 道路使用均匀度	0.12	0.22	0.20	0.06	0.18	0.12	0.12	0.17	0.21	0.14
固定调度策略 道路使用均匀度	0.40	0.51	0.61	0.26	0.59	0.52	0.41	0.56	0.55	0.64
学习型调度系统 装载点运输量均匀度	0.11	0.22	0.17	0.02	0.11	0.11	0.11	0.14	0.22	0.11
固定调度策略 装载点运输量均匀度	0.33	0.44	0.55	0.24	0.14	0.30	0.38	0.27	0.47	0.19
学习型调度系统 卸载点运输量均匀度	0.03	0.02	0.02	0.04	0.03	0.03	0.02	0.03	0.02	0.03
固定调度策略 卸载点运输量均匀度	0.12	0.04	0.19	0.04	0.45	0.10	0.10	0.30	0.18	0.30

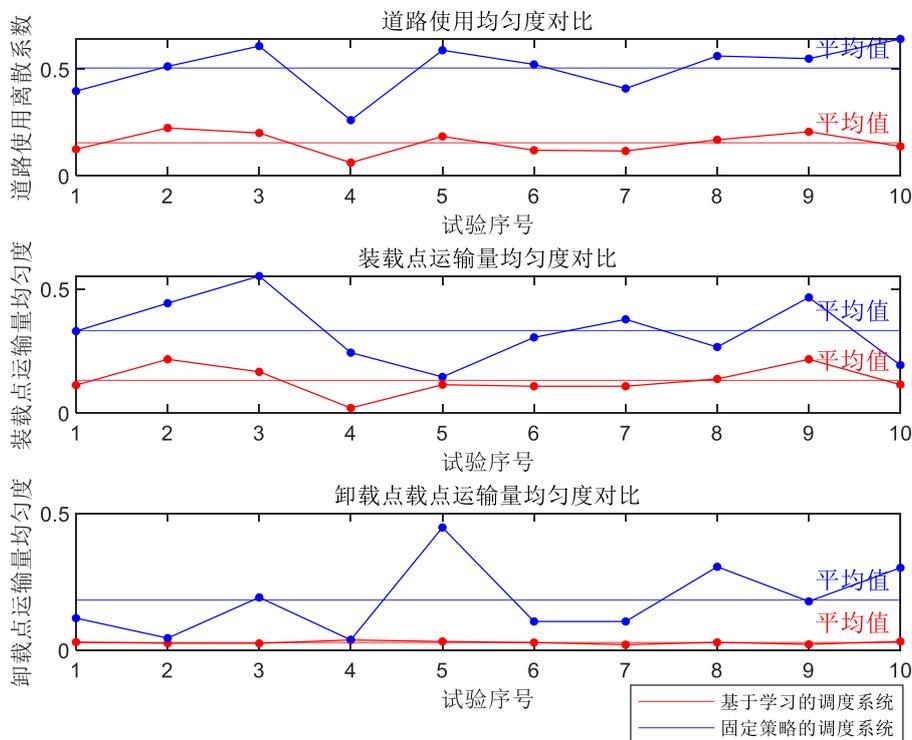


图 4-7 8h 班次内运输均匀度指标对比

Fig.4-7 Comparison of transportation uniformity in 8h work shift

如图 4-7 所示，单次样本的试验过程中，可以明显看出速度优化的学习型调度系统中每条道路的使用次数更平均，各个装载点和卸载点的任务量进度也更平均，这体现出更好的运输灵活性。因此我们可以得到以下结论，无论是从道路使用均匀度来看，还是装载点/卸载点均匀度来看，学习型的速度优化调度系统均表现出更好的性能，变异系数更低，运输的均匀度更强。

5) 安全稳定性分析

本研究提出的矿山车队调度系统，没有进行实际生产实践，只能在仿真环境中进行测试，因此很难全面的测试该车队调度系统的安全稳定性。我们不考虑自动驾驶车辆在行驶过程中的冲突、故障和横向偏差问题；仅评估调度系统的鲁棒性，特别是在需要将卡车临时故障关闭、进行维护或充电的一些事故情况下。在这种情况下，车队的其余部分需要临时合作，直到卡车恢复到车队中。为了评估该方法的鲁棒性并模拟意外情况，我们使用学习到的模型在相同的矿山模拟器中进行了测试，卡车数量不同（从 20 到 25 辆不等），模拟时间仍为 8 小时。

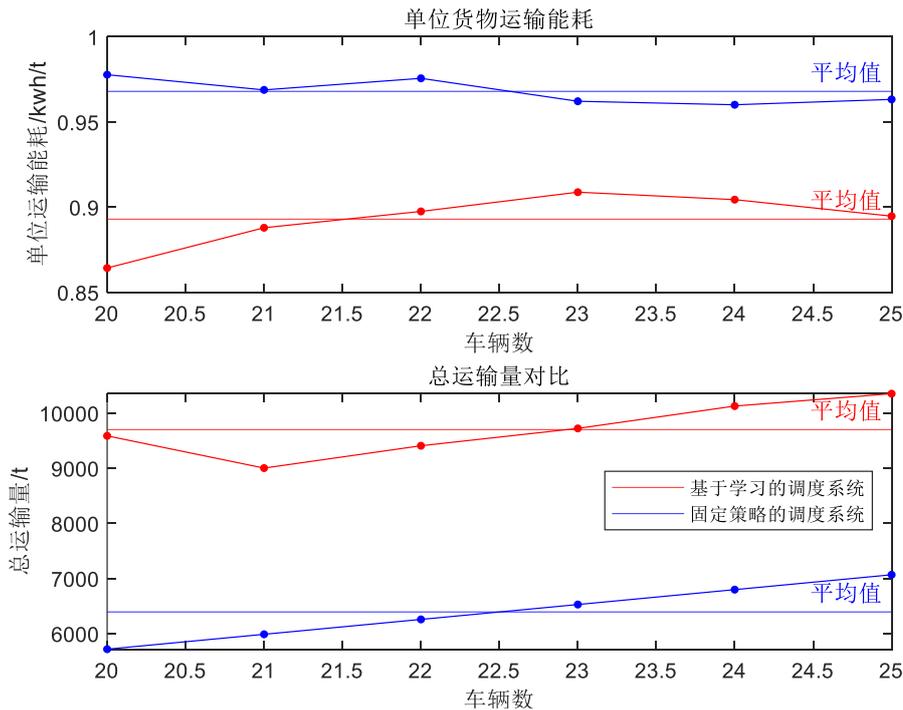


图 4-8 鲁棒性测试能耗和运输量对比图

Fig.4-8 Comparison of energy consumption and transport volume in robustness tests

图 4-8 至图 4-10 分别展示了鲁棒性测试不同性能的结果对比图，详细数据在表 4-9。结果表明该模型虽然在训练时使用的 25 辆卡车环境，但在具有不同车辆数目的环境中依然保持高且稳定的产量水平，并且优化后的行驶速度依然拥有更低能耗。但在单车利用率方面并没有表现出较好的性能。由于本调度方法集成了速度优化策略，优化后速度较快，导致平均等待闲置时间较高；在运输均匀度方面上，学习型调度系统无论是道

路运输均匀度上,还是在装/卸载点运输量均匀度上,在执行车队出现异常时,均表现出优异的性能,这能够保持矿山厂区内各个道路、站点均衡的生产节奏,充分利用道路。

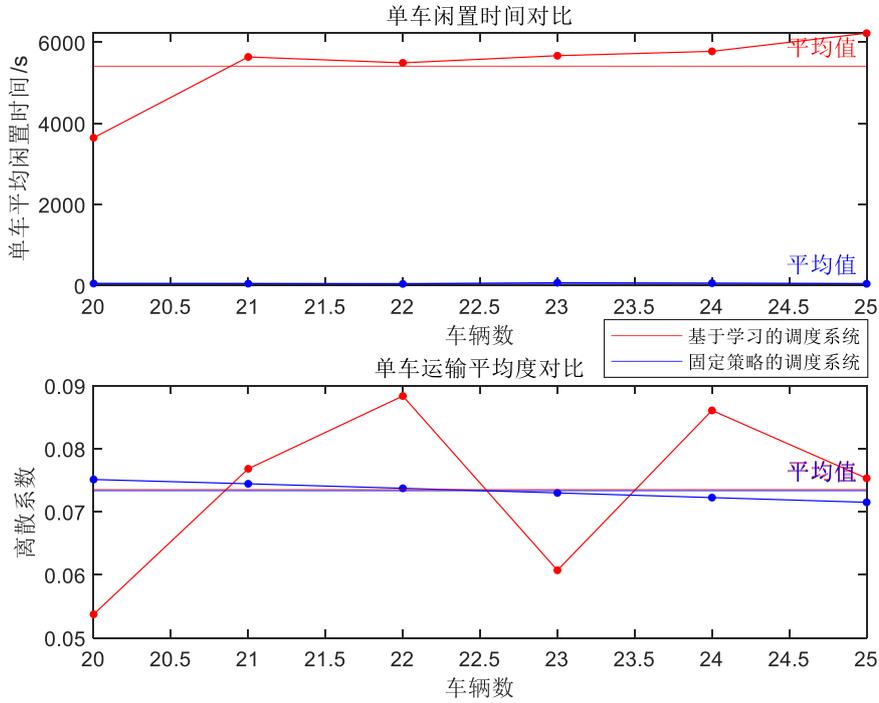


图 4-9 鲁棒性测试单车利用率对比图

Fig.4-9 Comparison of vehicle utilization in robustness tests

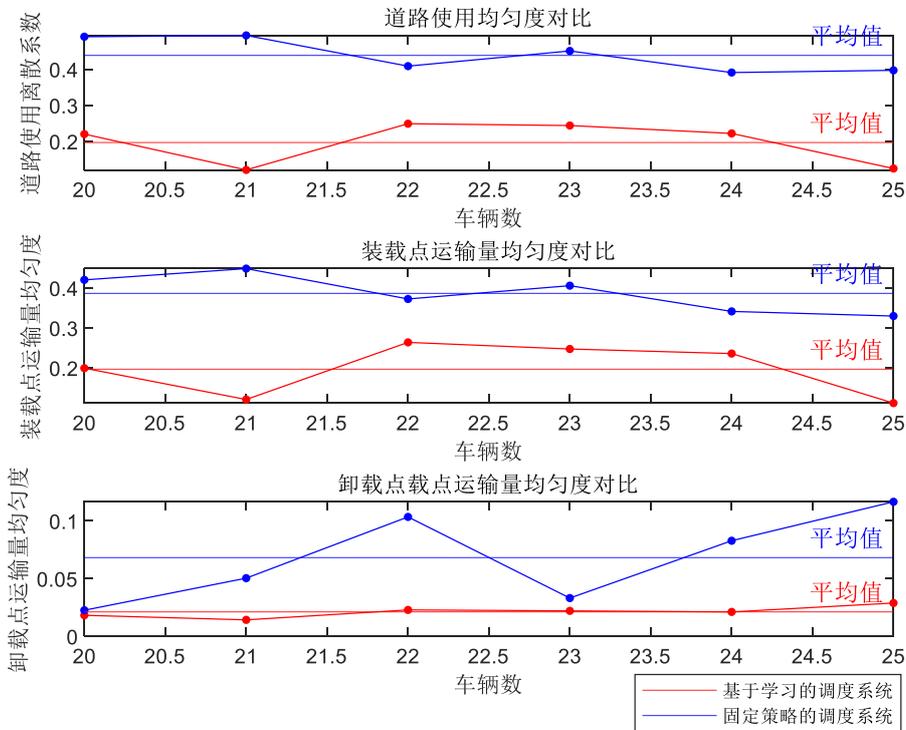


图 4-10 鲁棒性测试运输均匀度对比图

Fig.4-10 Comparison of transportation uniformity in robustness tests

重要的是,这种鲁棒性是在无需重新训练模型的情况下实现的,这与之前的方法不同,之前的方法在卡车数量发生变化时需要重新训练。此外,在6个指标中,结合速度优化的学习型的调度方式和固定策略之间的比较显示出,虽然在单车利用率性能上并没有多大的改善;由于速度优化的原因,但在运输单位货物的能耗方面所提出的方法优于固定策略。同样总运输量和运输均匀度方面比固定策略好得多。这表示该学习型算法可以满足在短时间内车辆临时意外退出的状况。这一观察结果突显了速度优化后的DQN学习型调度策略能够生成具有出色鲁棒性的高效动态调度策略。

表 4-9 调度系统的鲁棒性测试结果

Tab.4-9 Robustness test results of DQN dispatching system.

	基于学习的调度系统	固定调度策略	基于学习的调度系统	固定调度策略
车辆数	单位货物运输能耗 (kwh/t)	单位货物运输能耗 (kwh/t)	货物总运输量(t)	货物总运输量 (t)
25	0.89	0.96	10350	7065
24	0.90	0.96	10125	6795
23	0.91	0.96	9720	6525
22	0.90	0.98	9405	6255
21	0.89	0.97	9000	5985
20	0.86	0.98	9585	5715
车辆数	单车平均闲置时间(s)	单车平均闲置时间(s)	单车运输均匀度	单车运输均匀度
25	6224.62	40.77	0.08	0.07
24	5776.78	56.48	0.09	0.07
23	5669.52	66.42	0.06	0.07
22	5493.70	38.18	0.09	0.07
21	5638.40	45.04	0.08	0.07
20	3644.90	47.15	0.05	0.08
车辆数	装载点运输均匀度	装载点运输均匀度	卸载点运输均匀度	卸载点均匀度
25	0.11	0.33	0.03	0.12
24	0.24	0.34	0.02	0.08
23	0.25	0.41	0.02	0.03
22	0.26	0.37	0.02	0.10
21	0.12	0.45	0.01	0.05
20	0.20	0.42	0.02	0.02
车辆数	道路使用均匀度	道路使用均匀度		
25	0.12	0.40		
24	0.22	0.39		
23	0.24	0.45		
22	0.25	0.41		
21	0.12	0.49		

针对以上不同角度评价结果,我们最终对这十组测试结果做了综合性的对比。我们

将十次测试结果的各个指标求平均值，对比分析本文提出的结合速度优化的学习型调度系统的优势和劣势，并展开讨论和提出建议，详细数据请见表 4-10。

根据对基准模型固定策略调度方法性能的对比，我们对速度优化的学习型调度系统得到了全面的评估。如图 4-11，本研究提出的学习型算法在运输效率和运输灵活性这两个方面上表现出色，并且各个指标相互制约的关系如下。

其中运输效率得到改善的原因主要是采用了速度优化的行驶策略，车辆在行驶过程中根据路况采取不同的车速。车辆在爬坡时，降低车速；反而在平稳的路段时，采取更高的速度行驶；并且在下坡的路段时，车辆会开启能量回收模式，尽可能的储存更多的能量。因此车辆的行驶速度总体上升，运输周转时间下降，进而带来运输货物总量提升。

另一个表现突出的指标是运输灵活性。速度优化的学习型调度系统在训练过程中的奖惩函数包括了各个生产点完成量的变异系数，经过大量训练，调度系统会保证车辆平均分派在每个装载点和卸载点上，使得路网上的每条道路的使用量近乎均匀。这一点保证每个开采点生产节拍保持一致，生产进度协调；并且更均匀的道路使用量也会避免产生不必要的排队。

表 4-10 运输调度结果综合性能对比

Tab.4-10 Comprehensive performance comparison of DQN dispatching system.

8h 班次内综合对比评价				
一级指标	二级指标	学习型调度系统 平均值	固定策略 平均值	对比
运输效率	货物运输总量 (t)	10773	7006.5	提升
	运输周转时间 (s)	3011.5	4625.3	下降
成本控制	运输总能耗 (kwh)	9686.3	6977.07	提升
	单位货物运输能耗 (kwh/t)	0.9	1.0	下降
单车利用率	单车平均闲置时间 (s)	5364	62.6	提升
	单车运输均匀度	0.065	0.078	下降
运输灵活性	道路使用均匀度	0.15	0.50	下降
	装载点运输量均匀度	0.13	0.33	下降
	卸载点运输量均匀度	0.03	0.18	下降

正如前文所述，成本控制指标还与速度优化有关，车辆优化后的平均行驶速度会增高，但会在特定路况下尽可能的能量回收，采取能耗最佳的模式工作。这种策略下，虽然 8h 班次内车队运输量会增加，并且总能耗也会增加，但是单位货物的运输能耗会明显的降低，这一点更适合矿山企业的生产成本管理。

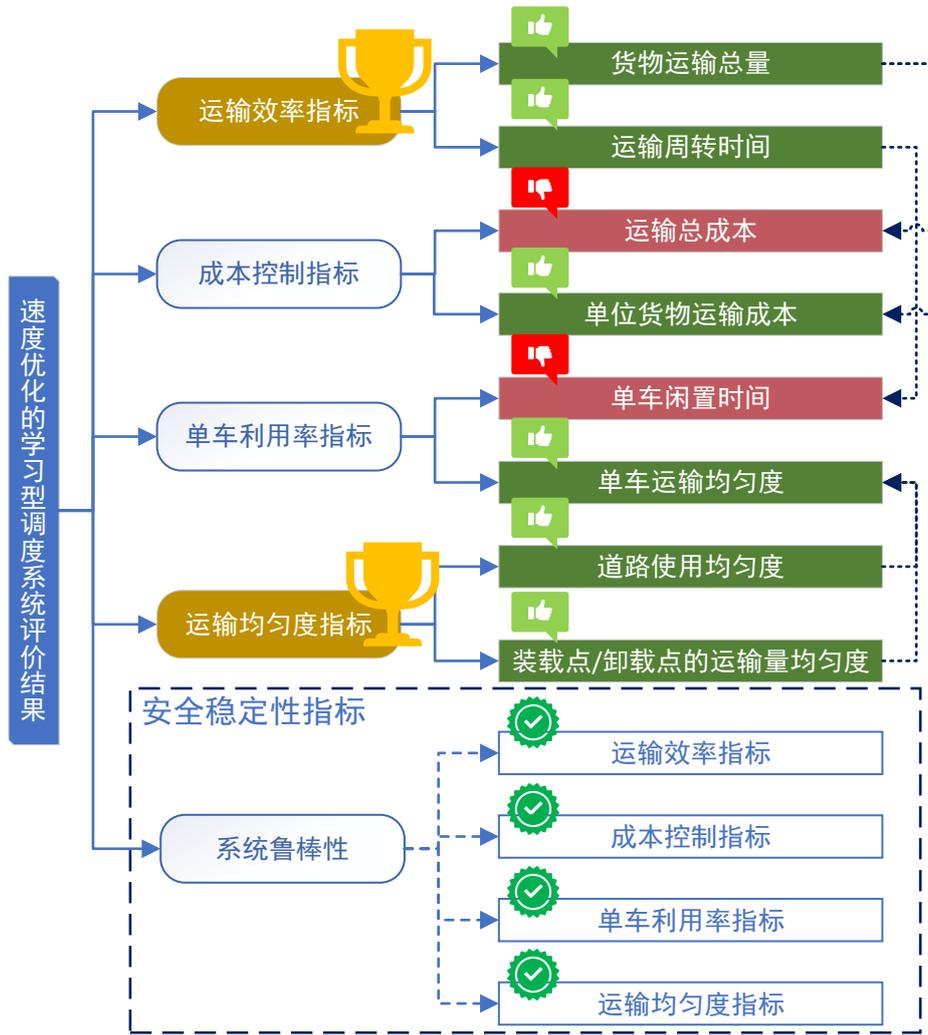


图 4-11 速度优化的学习型调度系统对比评价结果

Fig.4-11 Evaluation results of learning based dispatching systems with speed optimization

在单车利用率指标上，单车运输量均匀度出色的原因是因为学习型调度系统保证了任务在每个装载点和卸载点上近乎均衡，这间接了保证每辆车任务分配的均匀随机性，最终每辆车的运输量也达到了相对均衡的状态。另一方面，单车闲置时间过于长。速度优化后的策略导致运输周转时间变短，但是车辆的装载速度和卸载速度却没有发生变化。因此，这种状态在装载点和卸载点引起了供不应求的现象。车辆运输周转速度显著快于装载速度，导致多数情况下车辆不得不在装载点出等待。针对这种现象，建议改善装载速度，或适当对车辆行驶采取降速。

4.4 本章小结

首先，对矿山车辆运输调度进行理论概述，总结归纳出在运输效率、成本控制、单车利用率、运输灵活性四个方面的评价指标，根据相关工程经验和科学理论筛选合适的考核量并构建考核指标体系。其次，运用矿山运输评价考核体系对速度优化的学习型调度系统分析考核，采用规则的固定调度系统作为参考基准模型，对不同性能绘制图表、描绘总体对比情况和优劣势。

我们进行了 10 次试验用于评估调度系统的性能，根据学习型调度系统评价结果分析其运输生产意义，全面评价运输系统的优劣性。试验结果表明，速度优化后的学习型调度系统比传统模式拥有更好的运输效率和运输灵活性。在 8 小时班次内，单位货物运输成本减少，总产量增加；并且在各个卸载点和铲车装载位置之间展示出更均衡的任务调度，这意味着路线使用更加公平，开采速度保持一致。

此外，本文所提出的调度方法在处理意外的卡车故障方面一定的鲁棒性，从 20 到 25 辆卡车范围内，无需重新训练。这种能力在应对真实采矿作业中可能出现的意外事件时特别有价值。最后，根据指标性能分析调度系统的劣势及其原因，探究各个指标的内在联系。

5 结论与展望

5.1 结论

智能矿业的概念在全球范围内引起了广泛关注，无人驾驶智能车辆为传统矿业带来了新的机遇。有效的行驶策略和调度策略可以极大的提高运输资源利用率和效率。本文主要针对经济型速度行驶速度优化和学习型的自适应调度系统深入研究，主要工作及创新点如下：

- 1) 总结了国内外智能矿山车辆运输的发展现状，并对我国矿山无人驾驶车辆运输的关键环节进行论述，包括经济型行驶策略、强化学习调度系统。在此基础上，提出了通过速度优化和学习型调度来提升矿山运输的效率和降低成本。
- 2) 设计了一种考虑行驶时间、能量消耗和电池寿命的多目标速度优化方法，并在燃油车和电动车模型上验证了其有效性和适用性，它更适合矿山运输的成本管理。随着路线坡度的变化，车辆会采用不同频率的加速滑行 PnG 速度行驶策略。
- 3) 提出了基于深度强化学习的自适应车辆调度系统。本系统将速度优化的行驶策略与自主调度系统相结合，可以生成高效的动态调度策略。
- 4) 构建了矿山车辆运输调度评价体系，该评价体系对为矿山运输的发展提供了参考。同时我们利用该评价体系对本文提出的速度优化的学习型调度系统进行深度全面评价。

5.2 局限和展望

针对本文结论，有如下展望：

矿山车辆多目标行驶速度优化方法的各个目标函数之间的权重关系应该自适应调整，可以进一步基于矿山运输生产数据优化各个目标之间的权重。

本文采用 DQN 强化学习方法在处理复杂的状态空间，计算强度和学习时间也会增加。本文的调度系统车型单一，忽略了道路交叉排队和路面横向行驶位移偏差问题。在未来将考虑异构车队，并将路口冲突排队问题纳入考虑范围。此外可以探索行驶速度策略也纳入强化学习动作输出量。

无人驾驶车队的调度评价系统需要纳入对安全性能的评估，加强检测相关数据，通过机器学习不断修正评价指标，使之更加客观、合理。

参考文献

- [1] 国家统计局. 中国统计年鉴 2023 [J]. 统计, 2023.
- [2] SHAO S, LIU J, GENG Y, et al. Uncovering driving factors of carbon emissions from China's mining sector [J]. *Applied Energy*, 2016, 166(mar.15): 220-38.
- [3] KORYAGIN M, VORONOV A. Improving the Organization of the Shovel-Truck Systems in Open-Pit Coal Mines [J]. *Transport Problems*, 2017, 12(2): 113-22.
- [4] MENSING F, BIDEAUX E, TRIGUI R, et al. Trajectory optimization for eco-driving taking into account traffic constraints [J]. *Transportation Research Part D*, 2013, 18(JAN.): 55-61.
- [5] GILBERT E G. Vehicle cruise: Improved fuel economy by periodic control [J]. *Automatica*, 1976, 12(2): 159-66.
- [6] IVARSSON M, ASLUND J, NIELSEN L. Look-ahead control – consequences of a non-linear fuel map on truck fuel consumption [J]. *Proceedings of the Institution of Mechanical Engineers Part D Journal of Automobile Engineering*, 2016, 223(10): 1223-38.
- [7] LI S E, PENG H. Strategies to minimize fuel consumption of passenger cars during car-following scenarios [J]. *Proceedings of the Institution of Mechanical Engineers Part D Journal of Automobile Engineering*, 2011, 226(3): 2107 - 12.
- [8] MA J, HU J, LESLIE E, et al. An Eco-Drive Experiment on Rolling Terrains for Fuel Consumption Optimization with Connected Automated Vehicles; proceedings of the Transportation Research Board 97th Annual Meeting, F, 2018 [C].
- [9] SHEN D, KARBOWSKI D, ROUSSEAU A. Fuel-Optimal Periodic Control of Passenger Cars in Cruise Based on Pontryagin's Minimum Principle [J]. *Elsevier*, 2018, (31).
- [10] 徐少兵, 李升波, 赵芸辉, 等. CVT 型车辆经济性加速策略优化与分析 [J]. *机械工程学报*, 2015, 51(12): 10.
- [11] 徐少兵, 李升波, 成波. 无级变速器车辆经济性巡航策略的伪谱法优化 [J]. *控制理论与应用*, 2014, (6): 8.
- [12] 李升波, 徐少兵, 王文军, 等. 挡位离散型车辆经济性加速策略的伪谱法优化 [J]. *自动化学报*, 2015, 41(3): 11.
- [13] BARKENBUS J N. Eco-driving: An overlooked climate change initiative [J]. *Energy*

- policy, 2010, 38(2): 762-9.
- [14] KAMAL M, MUKAI M, MURATA J, et al. Ecological driver assistance system using model-based anticipation of vehicle-road-traffic information [J]. IET intelligent transport systems, 2010, 4(4): 244-51.
- [15] XU S, DENG K, LI S E, et al. Legendre pseudospectral computation of optimal speed profiles for vehicle eco-driving system; proceedings of the 2014 IEEE intelligent vehicles symposium proceedings, F, 2014 [C]. IEEE.
- [16] LI S, LI K, RAJAMANI R, et al. Model predictive multi-objective vehicular adaptive cruise control [J]. IEEE Transactions on control systems technology, 2010, 19(3): 556-66.
- [17] THOMAS G, VOULGARIS P G. Fuel minimization of a moving vehicle in suburban traffic; proceedings of the 2013 American Control Conference, F, 2013 [C]. IEEE.
- [18] CAMARA M B, GUALOUS H, GUSTIN F, et al. Design and new control of DC/DC converters to share energy between supercapacitors and batteries in hybrid vehicles [J]. IEEE Transactions on Vehicular Technology, 2008, 57(5): 2721-35.
- [19] BIANCHI D, ROLANDO L, SERRAO L, et al. A rule-based strategy for a series/parallel hybrid electric vehicle: an approach based on dynamic programming; proceedings of the Dynamic Systems and Control Conference, F, 2010 [C].
- [20] 张盟阳. 基于动态规划的 PHEV 能耗分析 [J]. 汽车实用技术, 2021.
- [21] HU J, SHAO Y, SUN Z, et al. Integrated optimal eco-driving on rolling terrain for hybrid electric vehicle with vehicle-infrastructure communication [J]. Transportation Research Part C: Emerging Technologies, 2016, 68: 228-44.
- [22] WEN Q, WANG F, CHENG M, et al. Adaptive equivalent consumption minimization strategy for off-road hydraulic hybrid vehicles: A cycle-to-cycle optimization approach [J]. IEEE Transactions on Vehicular Technology, 2021, 71(3): 2346-57.
- [23] 徐少兵. 车辆经济性驾驶的策略辨识与操控规律 [D]; 北京: 清华大学, 2016.
- [24] 梁志豪. 基于伪谱法的串并联式 PHEV 全局优化算法及参数匹配研究 [J]. 2018.
- [25] 甄玉珂. 基于伪谱法的 PHEV 全局优化能量管理策略在 WLTC 工况下的应用研究 [D]; 华南理工大学, 2019.
- [26] HUANG X, TAN Y, HE X. An intelligent multifeature statistical approach for the discrimination of driving conditions of a hybrid electric vehicle [J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 12(2): 453-65.
- [27] KIRAN B R, SOBH I, TALPAERT V, et al. Deep reinforcement learning for

- autonomous driving: A survey [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(6): 4909-26.
- [28] SALLAB A E, ABDYOU M, PEROT E, et al. End-to-end deep reinforcement learning for lane keeping assist [J]. *arXiv preprint arXiv:161204340*, 2016.
- [29] CHU T, WANG J, CODECÀ L, et al. Multi-agent deep reinforcement learning for large-scale traffic signal control [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 21(3): 1086-95.
- [30] LUO M, DU B, ZHANG W, et al. Fleet rebalancing for expanding shared e-mobility systems: A multi-agent deep reinforcement learning approach [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(4): 3868-81.
- [31] ELBROND J, SOUMIS F. Towards integrated production planning and truck dispatching in open pit mines [J]. *International Journal of Surface Mining, Reclamation and Environment*, 1987, 1(1): 1-6.
- [32] SOUMIS F, ETHIER J, ELBROND J. Evaluation of the new truck dispatching in the mount wright mine [J]. *Application of Computers and Operations Research in the Mineral Industry*, 1989: 674-82.
- [33] TEMENG V A, OTUONYE F O, FRENDEWEY J O. A nonpreemptive goal programming approach to truck dispatching in open pit mines [J]. *Mineral Resources Engineering*, 1998, 7(02): 59-67.
- [34] TOPAL E, RAMAZAN S. A new MIP model for mine equipment scheduling by minimizing maintenance cost [J]. *European Journal of Operational Research*, 2010, 207(2): 1065-71.
- [35] TA C H, INGOLFSSON A, DOUCETTE J. A linear model for surface mining haul truck allocation incorporating shovel idle probabilities [J]. *European journal of Operational research*, 2013, 231(3): 770-8.
- [36] CHANG Y, REN H, WANG S. Modelling and optimizing an open-pit truck scheduling problem [J]. *Discrete Dynamics in nature and Society*, 2015.
- [37] UPADHYAY S, ASKARI-NASAB H. Truck-shovel allocation optimisation: a goal programming approach [J]. *Mining Technology*, 2016, 125(2): 82-92.
- [38] KOENIGSBERG E. Cyclic queues [J]. *Journal of the Operational Research Society*, 1958, 9(1): 22-35.
- [39] KAPPAS G, YEGULALP T M. An application of closed queueing networks theory in truck-shovel systems [J]. *International Journal of Surface Mining, Reclamation and*

- Environment, 1991, 5(1): 45-51.
- [40] ERCELEBI S G, BASCETIN A. Optimization of shovel-truck system for surface mining [J]. Journal of the Southern African Institute of Mining and Metallurgy, 2009, 109(7): 433-9.
- [41] HE M-X, WEI J-C, LU X-M, et al. The genetic algorithm for truck dispatching problems in surface mine [J]. Information technology journal, 2010, 9(4): 710-4.
- [42] ALEXANDRE R F, CAMPELO F, VASCONCELOS J A. Multi-objective evolutionary algorithms for the truck dispatch problem in open-pit mining operations [J]. Learn Nonlinear Models, 2019, 17(2): 53-66.
- [43] DE CARVALHO J P, DIMITRAKOPOULOS R. Integrating Production Planning with Truck-Dispatching Decisions through Reinforcement Learning While Managing Uncertainty [J]. Minerals, 2021, 11(6).
- [44] ZHANG C, ODONKOR P, ZHENG S, et al. Dynamic dispatching for large-scale heterogeneous fleet via multi-agent deep reinforcement learning; proceedings of the 2020 IEEE International Conference on Big Data (Big Data), F, 2020 [C]. IEEE.
- [45] HUO D, SARI Y A, KEALEY R, et al. Reinforcement Learning-Based Fleet Dispatching for Greenhouse Gas Emission Reduction in Open-Pit Mining Operations [J]. Resources, Conservation and Recycling, 2023, 188.
- [46] SHI J, GAO Y, WANG W, et al. Operating Electric Vehicle Fleet for Ride-Hailing Services With Reinforcement Learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(11): 4822-34.
- [47] HOLLER J, VUORIO R, QIN Z, et al. Deep reinforcement learning for multi-driver vehicle dispatching and repositioning problem; proceedings of the 2019 IEEE International Conference on Data Mining (ICDM), F, 2019 [C]. IEEE.
- [48] LIN K, ZHAO R, XU Z, et al. Efficient Large-Scale Fleet Management via Multi-Agent Deep Reinforcement Learning [Z]//LIN K, ZHAO R, XU Z, 等. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 1774-83.10.1145/3219819.3219993
- [49] AHUMADA G I, RIVEROS E, HERZOG O. An Agent-based System for Truck Dispatching in Open-pit Mines; proceedings of the ICAART (1), F, 2020 [C].
- [50] WANG X, DAI Q, BIAN Y, et al. Real-time truck dispatching in open-pit mines [J]. International Journal of Mining, Reclamation and Environment, 2023, 37(7): 504-23.
- [51] MA J, HU J, LESLIE E, et al. An eco-drive experiment on rolling terrains for fuel

- consumption optimization with connected automated vehicles [J]. *Transportation Research Part C: Emerging Technologies*, 2019, 100: 125-41.
- [52] 孙文瑜, 徐成贤, 朱德通. 最优化方法 [M]. 高等教育出版社, 2004.
- [53] BELLMAN R. Dynamic programming [J]. *Science*, 1966, 153(3731): 34-7.
- [54] PONTRYAGIN L S. The mathematical theory of optimal processes and differential games [J]. *Trudy Mat Inst Steklov*, 1985, 169: 119-58.
- [55] KURIYAMA M, YAMAMOTO S, MIYATAKE M. Theoretical study on eco-driving technique for an electric vehicle with dynamic programming; proceedings of the 2010 International Conference on Electrical Machines and Systems, F, 2010 [C]. IEEE.
- [56] OHTSUKA T. A continuation/GMRES method for fast computation of nonlinear receding horizon control [J]. *Automatica*, 2004, 40(4): 563-74.
- [57] WANG Y, BOYD S. Fast model predictive control using online optimization [J]. *IEEE Transactions on control systems technology*, 2009, 18(2): 267-78.
- [58] RAULT J, RICHALET A, TESTUD J, et al. Model predictive heuristic control: application to industrial processes [J]. *Automatica*, 1978, 14(5): 413-28.
- [59] YAMAGUCHI D, KAMAL M, MUKAI M, et al. Model predictive control for automobile ecological driving using traffic signal information [J]. *Journal of System Design and Dynamics*, 2012, 6(3): 297-309.
- [60] 刘焕峰. 基于 TPFM 和 MPC 的汽车经济型巡航控制方法研究 [D]; 吉林大学, 2018.
- [61] EBEN LI S, LI K, WANG J. Economy-oriented vehicle adaptive cruise control with coordinating multiple objectives function [J]. *Vehicle System Dynamics*, 2013, 51(1): 1-17.
- [62] 廖永康. 基于 P2 构型混合动力驱动模式切换扭矩协调控制研究 [D]; 重庆大学, 2017.
- [63] VAN BASSHUYSEN R, SCHÄFER F. Internal combustion engine handbook [M]. SAE International, 2016.
- [64] AKÇELIK R, SMIT R, BESLEY M. Calibrating fuel consumption and emission models for modern vehicles; proceedings of the IPENZ transportation group conference, Rotorua, New Zealand, F, 2012 [C].
- [65] BIGGS D, AKÇELIK R. Energy-related model of instantaneous fuel consumption [J]. *Traffic Engineering and Control*, 1986, 27(6): 320-5.
- [66] PELKMANS L, DEBAL P, HOOD T, et al. Development of a simulation tool to

- calculate fuel consumption and emissions of vehicles operating in dynamic conditions [R]: SAE Technical Paper, 2004.
- [67] BORHAN H, VAHIDI A, PHILLIPS A M, et al. MPC-based energy management of a power-split hybrid electric vehicle [J]. IEEE Transactions on Control Systems Technology, 2011, 20(3): 593-603.
- [68] BRANICKY M S, BORKAR V S, MITTER S K. A unified framework for hybrid control: Model and optimal control theory [J]. IEEE transactions on automatic control, 1998, 43(1): 31-45.
- [69] ALSABAAN M, NAIK K, KHALIFA T. Optimization of fuel cost and emissions using V2V communications [J]. IEEE Transactions on intelligent transportation systems, 2013, 14(3): 1449-61.
- [70] 张蓉, 何信华, 刘朝阳, 等. 永磁同步电机非线性实时模型建模设计与验证 [J]. 航天控制, 2018, 36(5): 5.
- [71] SHUTTA Y, TAKAHASHI Y, FUJIWARA K. Behavioral modeling of permanent magnet synchronous motor fed by PWM inverters considering iron losses due to carrier harmonics [J]. IEEE Transactions on Industry Applications, 2019, 139(11): 939-46.
- [72] 张兴华, 童歆渝, 刘伟. 考虑铁芯损耗的内置式永磁同步电机模型参数测量 [J]. 电力自动化设备, 2018, 38(3): 194-8.
- [73] ZHANG X, ZHANG C, YU J, et al. Analytical model of magnetic field of a permanent magnet synchronous motor with a trapezoidal Halbach permanent magnet array [J]. IEEE Transactions on Magnetics, 2019, 55(7): 1-5.
- [74] 陈滨斐. 表贴式永磁同步电机建模, 分析与设计 [D]; 天津: 天津大学, 2014.
- [75] SIZOV G Y, IONEL D M, DEMERDASH N A. Modeling and parametric design of permanent-magnet ac machines using computationally efficient finite-element analysis [J]. IEEE Transactions on Industrial Electronics, 2011, 59(6): 2403-13.
- [76] 冯彦彪. 串联式混合动力矿用自卸车性能及燃油成本分析 [D]; 北京: 北京科技大学, 2016.
- [77] 王少华. 电动汽车动力锂电池模型参数辨识和状态估计方法研究 [D]; 吉林大学, 2021.
- [78] ZHENG L, ZHANG L, ZHU J, et al. Co-estimation of state-of-charge, capacity and resistance for lithium-ion batteries based on a high-fidelity electrochemical model [J]. Applied energy, 2016, 180: 424-34.
- [79] GOPALUNI R B, BRAATZ R D. State of charge estimation in Li-ion batteries using an

- isothermal pseudo two-dimensional model [J]. IFAC Proceedings Volumes, 2013, 46(32): 135-40.
- [80] JOKAR A, RAJABLOO B, DÉSILETS M, et al. Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries [J]. Journal of Power Sources, 2016, 327: 44-55.
- [81] FULLER T F, DOYLE M, NEWMAN J. Relaxation phenomena in lithium - ion - insertion cells [J]. Journal of the Electrochemical Society, 1994, 141(4): 982.
- [82] FULLER T F, DOYLE M, NEWMAN J. Simulation and optimization of the dual lithium ion insertion cell [J]. Journal of the electrochemical society, 1994, 141(1): 1.
- [83] HAGEMAN S C. Simple PSpice models let you simulate common battery types [J]. EDN, 1993, 38(22): 117-&.
- [84] HE H, XIONG R, GUO H, et al. Comparison study on the battery models used for the energy management of batteries in electric vehicles [J]. Energy Conversion and Management, 2012, 64: 113-21.
- [85] KAMAL M A S, MUKAI M, MURATA J, et al. On board eco-driving system for varying road-traffic environments using model predictive control; proceedings of the 2010 IEEE International Conference on Control Applications, F, 2010 [C]. IEEE.
- [86] FRÖBERG A, NIELSEN L. Optimal fuel and gear ratio control for heavy trucks with piece wise affine engine characteristics [J]. IFAC Proceedings Volumes, 2007, 40(10): 335-42.
- [87] VAJEDI M, AZAD N L. Ecological adaptive cruise controller for plug-in hybrid electric vehicles using nonlinear model predictive control [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 17(1): 113-22.
- [88] 林歆悠, 孙冬野, 尹燕莉, 等. 基于随机动态规划的混联式混合动力客车能量管理策略 [J]. 汽车工程, 2012, (9): 830-6.
- [89] XU B, CHEN X, LI K, et al. Double-layer speed optimization for reducing fuel consumption with vehicle-to-infrastructure communication [J]. Journal of Intelligent Transportation Systems, 2019, 23(5): 513-24.
- [90] ZHU G, LIN J, LIU Q, et al. Research on the Energy-Saving Strategy of Path Planning for Electric Vehicles Considering Traffic Information [J]. Energies, 2019, 12(19): 3601.
- [91] PATTERSON S R, KOZAN E, HYLAND P. Energy efficient scheduling of open-pit coal mine trucks [J]. European Journal of Operational Research, 2017, 262(2): 759-70.
- [92] DEHKORDI S G, LARUE G S, CHOLETTE M E, et al. Including network level safety

- measures in eco-routing [J]. *Journal of Intelligent Transportation Systems*, 2022: 1-14.
- [93] LIU R, LIU H, HAN L, et al. A multi-objective regenerative braking control strategy combining with velocity optimization for connected vehicles [J]. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 2023, 237(6): 1465-74.
- [94] WANG J, LIU P, HICKS-GARNER J, et al. Cycle-life model for graphite-LiFePO₄ cells [J]. *Journal of power sources*, 2011, 196(8): 3942-8.
- [95] CORDOBA-ARENAS A, ONORI S, GUEZENNEC Y, et al. Capacity and power fade cycle-life model for plug-in hybrid electric vehicle lithium-ion battery cells containing blended spinel and layered-oxide positive electrodes [J]. *Journal of Power Sources*, 2015, 278: 473-83.
- [96] ZHANG L, XIA X. An Integer Programming Approach for Truck-Shovel Dispatching Problem in Open-Pit Mines [J]. *Energy Procedia*, 2015, 75: 1779-84.
- [97] SUTTON R S, BARTO A G. *Reinforcement learning: An introduction* [M]. MIT press, 2018.
- [98] 陈奇. 铁路运输企业车流分析及关键指标动态考核研究 [D]; 北京交通大学, 2022.
- [99] 徐翔. 基于云模型的管道运输低碳绿色发展研究 [D]; 大连海事大学, 2016.

致 谢

时光荏苒，岁月如梭，转眼间三年硕士研究生生涯即将画上句号。回顾这段难忘的旅程，我的内心感慨万千，既有对过去求学科研之路的深深眷恋，也有对未来的美好憧憬。当初，怀揣梦想，踏入了北京科技大学的校门。如今，我从一个懵懂少年成长为一名踌躇满志的青年。在这里，我收获了知识，更收获了宝贵的成长经历。在此，谨借此机会向所有给予我帮助和支持的人们致以最诚挚的谢意！

首先，我要衷心感谢我的家人。无论是从高中到北京科技大学，再到现在的研究生，你们始终如一地支持我，给予我最大的鼓励和无私的爱。你们是我坚强的后盾，是我前进的动力。没有你们的支持，就没有我今天的成绩。正是那片生我养我的家园，给予了我最初的梦想和力量。

其次，我要特别致敬我的导师杨珏教授。您的严谨治学、深厚学识，让我受益匪浅。在您的引领下，我不仅深入钻研专业知识，更学会了独立思考、解决难题。尤为珍贵的是，我曾跟随您出差，前往科研第一现场，那里的所见所闻，让我亲身体会了“纸上得来终觉浅，绝知此事要躬行”的真谛。您不仅是我学术上的引路人，更是我人生道路上的导师。我亲眼目睹了科研工作的艰辛与不易，感受到了团队合作的力量，也深刻体会到了理论与实践相结合的重要性。我也学会了如何面对挑战和困难，如何保持冷静和乐观，如何与人沟通和协作。这些宝贵的经验和教诲将伴随我走过未来的每一个阶段。

此外，我要感谢我的同窗好友们。我们一同度过了一段段难忘的时光，彼此鼓励、互相帮助。在科研的海洋中，我们共同探索、共同进步；在闲暇之余，我们畅谈理想、分享快乐。这些日子如诗如画，铭刻在我心间。

同时，北京科技大学为我提供了优质的学习环境和丰富的学术资源，让我得以在知识的海洋中畅游。在这里，我遇到了许多优秀的同学和老师，他们的陪伴和帮助让我在学术上不断取得进步。

回首北科大的七年，我深感自己的成长与蜕变。从莘莘学子到即将步入社会的青年，我的视野得到了极大的拓展，我的能力得到了全面的提升。这一切都离不开家人、导师和同学们的支持与帮助。展望未来，我将带着这份感激与收获，继续前行。在这个充满机遇与挑战的时代，我期待着在未来的道路上取得更多的成果与突破。同时，我也将所学知识应用于实际工作中，为社会的发展与进步贡献自己的绵薄之力。

最后，再次感谢所有在我求学之路上给予我帮助、关心和支持的人。你们的付出与奉献将永远铭记在我心中。在未来的日子里，我将以更加饱满的热情和更加坚定的信念迎接每一个挑战与机遇。

作者简历及在学研究成果

一、主要教育经历/工作经历（从大学起，到硕士入学止）

起止年月	学习或工作单位	备注
2017年09月 至2021年06月	在北京科技大学 车辆工程专业攻读学士学位	
2021年09月 至2024年06月	在北京科技大学 机械工程专业攻读硕士学位	

二、在学期间从事的科研工作

- 1) 取消驾驶室的无人纯电动矿用卡车及配套支持系统定制化开发研究与示范推广(基础理论研究)（2023-0101），主要参与人员，2023.01至今。
- 2) 无人驾驶车辆集群管理调度算法（2022-0716），主要参与人员，2022.09至今。
- 3) 申请专利：
杨珏,裴树伟,贾金龙,等. 一种无人驾驶矿用车辆调度方法及外置矿用车载雷达 [P]. 中国专利：CN116872852A,2023-10-13.

三、在学期间所获的科研奖励

罗伯托·罗卡（Roberto Rocca）奖学金二等奖

四、在学期间发表的论文

- [1] PEI S, YANG J. Multi-Objective Velocity Trajectory Optimization Method for Autonomous Mining Vehicles [J]. Int J Automot Technol, 2023, 24(6): 1627-41.
- [2] PEI S, YANG J. Learning-based Dispatching System with Trajectory Optimization for Autonomous Mining Transportation [J]. Journal of Intelligent Transportation Systems, Under Peer Review.

独创性说明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

作者签名：_____日期：_____年____月____日

关于论文使用授权的说明

本人完全了解北京科技大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

（保密的论文在解密后应遵循此规定）

签名：_____ 导师签名：_____ 日期：_____

学位论文数据集

关键词*	密级*	中图分类号*	UDC	论文资助
车队管理；能耗优化；	公开	U492.6+4	621	
学位授予单位名称*		学位授予单位代码*	学位类别*	学位级别*
北京科技大学		10008	工学	硕士
论文题名*		并列题名		论文语种*
速度优化的学习型矿用无人驾驶车队调度研究				
作者姓名*	裴树伟		学号*	M202110447
培养单位名称*		培养单位代码*	培养单位地址	邮编
北京科技大学		10008	北京市海淀区学院路 30 号	100083
学科专业*		研究方向*	学制*	学位授予年*
机械工程		智能车队管理	3	2024
论文提交日期*				
导师姓名*	杨珏		职称*	教授
评阅人	答辩委员会主席*		答辩委员会成员	
	刘立		徐晓明	孟宇
电子版论文提交格式 文本 (<input checked="" type="checkbox"/>) 图像 (<input type="checkbox"/>) 视频 (<input type="checkbox"/>) 音频 (<input type="checkbox"/>) 多媒体 (<input type="checkbox"/>) 其他 (<input type="checkbox"/>)				
推荐格式: application/msword; application/pdf				
电子版论文出版 (发布) 者		电子版论文出版 (发布) 地		权限声明
论文总页数*	110			
共 33 项, 其中带*为必填数据, 为 22 项。				

- [1] 国家统计局. 中国统计年鉴 2012 [J]. 统计, 2012.
- [2] SHAO S, LIU J, GENG Y, et al. Uncovering driving factors of carbon emissions from China's mining sector [J]. Applied Energy, 2016, 166(mar.15): 220-38.
- [3] KORYAGIN M, VORONOV A. Improving the Organization of the Shovel-Truck Systems in Open-Pit Coal Mines [J]. Transport Problems, 2017, 12(2): 113-22.
- [4] MENSING F, BIDEAUX E, TRIGUI R, et al. Trajectory optimization for eco-driving taking into account traffic constraints [J]. Transportation Research Part D, 2013, 18(JAN.): 55-61.
- [5] GILBERT E G. Vehicle cruise: Improved fuel economy by periodic control [J]. Automatica, 1976, 12(2): 159-66.
- [6] IVARSSON M, ASLUND J, NIELSEN L. Look-ahead control – consequences of a non-linear fuel map on truck fuel consumption [J]. Proceedings of the Institution of Mechanical Engineers Part D Journal of Automobile Engineering, 2016, 223(10): 1223-38.
- [7] LI S E, PENG H. Strategies to minimize fuel consumption of passenger cars during car-following scenarios [J]. Proceedings of the Institution of Mechanical Engineers Part D Journal of Automobile Engineering, 2011, 226(3): 2107 - 12.
- [8] MA J, HU J, LESLIE E, et al. An Eco-Drive Experiment on Rolling Terrains for Fuel Consumption Optimization with Connected Automated Vehicles; proceedings of the Transportation Research Board 97th Annual Meeting, F, 2018 [C].
- [9] SHEN D, KARBOWSKI D, ROUSSEAU A. Fuel-Optimal Periodic Control of Passenger Cars in Cruise Based on Pontryagin's Minimum Principle [J]. Elsevier, 2018, (31).
- [10] 徐少兵, 李升波, 赵芸辉, 等. CVT 型车辆经济性加速策略优化与分析 [J]. 机械工程学报, 2015, 51(12): 10.
- [11] 徐少兵, 李升波, 成波. 无级变速器车辆经济性巡航策略的伪谱法优化 [J]. 控制理论与应用, 2014, (6): 8.
- [12] 李升波, 徐少兵, 王文军, 等. 挡位离散型车辆经济性加速策略的伪谱法优化 [J]. 自动化学报, 2015, 41(3): 11.
- [13] BARKENBUS J N. Eco-driving: An overlooked climate change initiative [J]. Energy policy, 2010, 38(2): 762-9.
- [14] KAMAL M, MUKAI M, MURATA J, et al. Ecological driver assistance system using model-based anticipation of vehicle-road-traffic information [J]. IET intelligent transport systems, 2010, 4(4): 244-51.
- [15] XU S, DENG K, LI S E, et al. Legendre pseudospectral computation of optimal speed profiles for vehicle eco-driving system; proceedings of the 2014 IEEE intelligent vehicles symposium proceedings, F, 2014 [C]. IEEE.
- [16] LI S, LI K, RAJAMANI R, et al. Model predictive multi-objective vehicular adaptive cruise control [J]. IEEE Transactions on control systems technology, 2010, 19(3): 556-66.
- [17] THOMAS G, VOULGARIS P G. Fuel minimization of a moving vehicle in suburban traffic; proceedings of the 2013 American Control Conference, F, 2013 [C]. IEEE.
- [18] CAMARA M B, GUALOUS H, GUSTIN F, et al. Design and new control of DC/DC converters to share energy between supercapacitors and batteries in hybrid vehicles [J]. IEEE Transactions on Vehicular Technology, 2008, 57(5): 2721-35.
- [19] BIANCHI D, ROLANDO L, SERRAO L, et al. A rule-based strategy for a series/parallel hybrid electric vehicle:

- an approach based on dynamic programming; proceedings of the Dynamic Systems and Control Conference, F, 2010 [C].
- [20] 张盟阳. 基于动态规划的 PHEV 能耗分析 [J]. 汽车实用技术, 2021.
- [21] HU J, SHAO Y, SUN Z, et al. Integrated optimal eco-driving on rolling terrain for hybrid electric vehicle with vehicle-infrastructure communication [J]. *Transportation Research Part C: Emerging Technologies*, 2016, 68: 228-44.
- [22] WEN Q, WANG F, CHENG M, et al. Adaptive equivalent consumption minimization strategy for off-road hydraulic hybrid vehicles: A cycle-to-cycle optimization approach [J]. *IEEE Transactions on Vehicular Technology*, 2021, 71(3): 2346-57.
- [23] 徐少兵. 车辆经济性驾驶的策略辨识与操控规律 [D]; 北京: 清华大学, 2016.
- [24] 梁志豪. 基于伪谱法的串并联式 PHEV 全局优化算法及参数匹配研究 [J]. 2018.
- [25] 甄玉珂. 基于伪谱法的 PHEV 全局优化能量管理策略在 WLTC 工况下的应用研究 [D]; 华南理工大学, 2019.
- [26] HUANG X, TAN Y, HE X. An intelligent multifeature statistical approach for the discrimination of driving conditions of a hybrid electric vehicle [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2010, 12(2): 453-65.
- [27] KIRAN B R, SOBH I, TALPAERT V, et al. Deep reinforcement learning for autonomous driving: A survey [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(6): 4909-26.
- [28] SALLAB A E, ABDON M, PEROT E, et al. End-to-end deep reinforcement learning for lane keeping assist [J]. arXiv preprint arXiv:161204340, 2016.
- [29] CHU T, WANG J, CODECÀ L, et al. Multi-agent deep reinforcement learning for large-scale traffic signal control [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 21(3): 1086-95.
- [30] LUO M, DU B, ZHANG W, et al. Fleet rebalancing for expanding shared e-mobility systems: A multi-agent deep reinforcement learning approach [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(4): 3868-81.
- [31] ELBROND J, SOUMIS F. Towards integrated production planning and truck dispatching in open pit mines [J]. *International Journal of Surface Mining, Reclamation and Environment*, 1987, 1(1): 1-6.
- [32] SOUMIS F, ETHIER J, ELBROND J. Evaluation of the new truck dispatching in the mount wright mine [J]. *Application of Computers and Operations Research in the Mineral Industry*, 1989: 674-82.
- [33] TEMENG V A, OTUONYE F O, FRENDEWEY J O. A nonpreemptive goal programming approach to truck dispatching in open pit mines [J]. *Mineral Resources Engineering*, 1998, 7(02): 59-67.
- [34] TOPAL E, RAMAZAN S. A new MIP model for mine equipment scheduling by minimizing maintenance cost [J]. *European Journal of Operational Research*, 2010, 207(2): 1065-71.
- [35] TA C H, INGOLFSSON A, DOUCETTE J. A linear model for surface mining haul truck allocation incorporating shovel idle probabilities [J]. *European journal of Operational research*, 2013, 231(3): 770-8.
- [36] CHANG Y, REN H, WANG S. Modelling and optimizing an open-pit truck scheduling problem [J]. *Discrete Dynamics in nature and Society*, 2015, 2015.
- [37] UPADHYAY S, ASKARI-NASAB H. Truck-shovel allocation optimisation: a goal programming approach [J]. *Mining Technology*, 2016, 125(2): 82-92.
- [38] KOENIGSBERG E. Cyclic queues [J]. *Journal of the Operational Research Society*, 1958, 9(1): 22-35.
- [39] KAPPAS G, YEGULALP T M. An application of closed queueing networks theory in truck-shovel systems [J]. *International Journal of Surface Mining, Reclamation and Environment*, 1991, 5(1): 45-51.

- [40] ERCELEBI S G, BASCETIN A. Optimization of shovel-truck system for surface mining [J]. Journal of the Southern African Institute of Mining and Metallurgy, 2009, 109(7): 433-9.
- [41] HE M-X, WEI J-C, LU X-M, et al. The genetic algorithm for truck dispatching problems in surface mine [J]. Information technology journal, 2010, 9(4): 710-4.
- [42] ALEXANDRE R F, CAMPELO F, VASCONCELOS J A. Multi-objective evolutionary algorithms for the truck dispatch problem in open-pit mining operations [J]. Learn Nonlinear Models, 2019, 17(2): 53-66.
- [43] DE CARVALHO J P, DIMITRAKOPOULOS R. Integrating Production Planning with Truck-Dispatching Decisions through Reinforcement Learning While Managing Uncertainty [J]. Minerals, 2021, 11(6).
- [44] ZHANG C, ODONKOR P, ZHENG S, et al. Dynamic dispatching for large-scale heterogeneous fleet via multi-agent deep reinforcement learning; proceedings of the 2020 IEEE International Conference on Big Data (Big Data), F, 2020 [C]. IEEE.
- [45] HUO D, SARI Y A, KEALEY R, et al. Reinforcement Learning-Based Fleet Dispatching for Greenhouse Gas Emission Reduction in Open-Pit Mining Operations [J]. Resources, Conservation and Recycling, 2023, 188.
- [46] SHI J, GAO Y, WANG W, et al. Operating Electric Vehicle Fleet for Ride-Hailing Services With Reinforcement Learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(11): 4822-34.
- [47] HOLLER J, VUORIO R, QIN Z, et al. Deep reinforcement learning for multi-driver vehicle dispatching and repositioning problem; proceedings of the 2019 IEEE International Conference on Data Mining (ICDM), F, 2019 [C]. IEEE.
- [48] LIN K, ZHAO R, XU Z, et al. Efficient Large-Scale Fleet Management via Multi-Agent Deep Reinforcement Learning [Z]//LIN K, ZHAO R, XU Z, 等. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 1774-83.10.1145/3219819.3219993
- [49] AHUMADA G I, RIVEROS E, HERZOG O. An Agent-based System for Truck Dispatching in Open-pit Mines; proceedings of the ICAART (1), F, 2020 [C].
- [50] WANG X, DAI Q, BIAN Y, et al. Real-time truck dispatching in open-pit mines [J]. International Journal of Mining, Reclamation and Environment, 2023, 37(7): 504-23.
- [51] MA J, HU J, LESLIE E, et al. An eco-drive experiment on rolling terrains for fuel consumption optimization with connected automated vehicles [J]. Transportation Research Part C: Emerging Technologies, 2019, 100: 125-41.
- [52] 孙文瑜, 徐成贤, 朱德通. 最优化方法 [M]. 高等教育出版社, 2004.
- [53] BELLMAN R. Dynamic programming [J]. Science, 1966, 153(3731): 34-7.
- [54] PONTRYAGIN L S. The mathematical theory of optimal processes and differential games [J]. Trudy Mat Inst Steklov, 1985, 169: 119-58.
- [55] KURIYAMA M, YAMAMOTO S, MIYATAKE M. Theoretical study on eco-driving technique for an electric vehicle with dynamic programming; proceedings of the 2010 International Conference on Electrical Machines and Systems, F, 2010 [C]. IEEE.
- [56] OHTSUKA T. A continuation/GMRES method for fast computation of nonlinear receding horizon control [J]. Automatica, 2004, 40(4): 563-74.
- [57] WANG Y, BOYD S. Fast model predictive control using online optimization [J]. IEEE Transactions on control systems technology, 2009, 18(2): 267-78.
- [58] RAULT J, RICHALET A, TESTUD J, et al. Model predictive heuristic control: application to industrial processes [J]. Automatica, 1978, 14(5): 413-28.
- [59] YAMAGUCHI D, KAMAL M, MUKAI M, et al. Model predictive control for automobile ecological driving

- using traffic signal information [J]. *Journal of System Design and Dynamics*, 2012, 6(3): 297-309.
- [60] 刘焕峰. 基于 TPFM 和 MPC 的汽车经济型巡航控制方法研究 [D]; 吉林大学, 2018.
- [61] EBEN LI S, LI K, WANG J. Economy-oriented vehicle adaptive cruise control with coordinating multiple objectives function [J]. *Vehicle System Dynamics*, 2013, 51(1): 1-17.
- [62] 廖永康. 基于 P2 构型混合动力驱动模式切换扭矩协调控制研究 [D]; 重庆大学, 2017.
- [63] VAN BASSHUYSEN R, SCHÄFER F. *Internal combustion engine handbook* [M]. SAE International, 2016.
- [64] AKÇELIK R, SMIT R, BESLEY M. Calibrating fuel consumption and emission models for modern vehicles; proceedings of the IPENZ transportation group conference, Rotorua, New Zealand, F, 2012 [C].
- [65] BIGGS D, AKÇELIK R. Energy-related model of instantaneous fuel consumption [J]. *Traffic Engineering and Control*, 1986, 27(6): 320-5.
- [66] PELKMANS L, DEBAL P, HOOD T, et al. Development of a simulation tool to calculate fuel consumption and emissions of vehicles operating in dynamic conditions [R]: SAE Technical Paper, 2004.
- [67] BORHAN H, VAHIDI A, PHILLIPS A M, et al. MPC-based energy management of a power-split hybrid electric vehicle [J]. *IEEE Transactions on Control Systems Technology*, 2011, 20(3): 593-603.
- [68] BRANICKY M S, BORKAR V S, MITTER S K. A unified framework for hybrid control: Model and optimal control theory [J]. *IEEE transactions on automatic control*, 1998, 43(1): 31-45.
- [69] ALSABAAN M, NAIK K, KHALIFA T. Optimization of fuel cost and emissions using V2V communications [J]. *IEEE Transactions on intelligent transportation systems*, 2013, 14(3): 1449-61.
- [70] 张蓉, 何信华, 刘朝阳, 等. 永磁同步电机非线性实时模型建模设计与验证 [J]. *航天控制*, 2018, 36(5): 5.
- [71] SHUTTA Y, TAKAHASHI Y, FUJIWARA K. Behavioral modeling of permanent magnet synchronous motor fed by PWM inverters considering iron losses due to carrier harmonics [J]. *IEEJ Transactions on Industry Applications*, 2019, 139(11): 939-46.
- [72] 张兴华, 童歆渝, 刘伟. 考虑铁芯损耗的内置式永磁同步电机模型参数测量 [J]. *电力自动化设备*, 2018, 38(3): 194-8.
- [73] ZHANG X, ZHANG C, YU J, et al. Analytical model of magnetic field of a permanent magnet synchronous motor with a trapezoidal Halbach permanent magnet array [J]. *IEEE Transactions on Magnetics*, 2019, 55(7): 1-5.
- [74] 陈焯斐. 表贴式永磁同步电机建模, 分析与设计 [D]; 天津: 天津大学, 2014.
- [75] SIZOV G Y, IONEL D M, DEMERDASH N A. Modeling and parametric design of permanent-magnet ac machines using computationally efficient finite-element analysis [J]. *IEEE Transactions on Industrial Electronics*, 2011, 59(6): 2403-13.
- [76] 冯彦彪. 串联式混合动力矿用自卸车性能及燃油成本分析 [D]; 北京: 北京科技大学, 2016.
- [77] 王少华. 电动汽车动力锂电池模型参数辨识和状态估计方法研究 [D]; 吉林大学, 2021.
- [78] ZHENG L, ZHANG L, ZHU J, et al. Co-estimation of state-of-charge, capacity and resistance for lithium-ion batteries based on a high-fidelity electrochemical model [J]. *Applied energy*, 2016, 180: 424-34.
- [79] GOPALUNI R B, BRAATZ R D. State of charge estimation in Li-ion batteries using an isothermal pseudo two-dimensional model [J]. *IFAC Proceedings Volumes*, 2013, 46(32): 135-40.
- [80] JOKAR A, RAJABLOO B, DÉSILETS M, et al. Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries [J]. *Journal of Power Sources*, 2016, 327: 44-55.
- [81] FULLER T F, DOYLE M, NEWMAN J. Relaxation phenomena in lithium - ion - insertion cells [J]. *Journal of the Electrochemical Society*, 1994, 141(4): 982.
- [82] FULLER T F, DOYLE M, NEWMAN J. Simulation and optimization of the dual lithium ion insertion cell [J].

- Journal of the electrochemical society, 1994, 141(1): 1.
- [83] HAGEMAN S C. Simple PSpice models let you simulate common battery types [J]. EDN, 1993, 38(22): 117-8.
- [84] HE H, XIONG R, GUO H, et al. Comparison study on the battery models used for the energy management of batteries in electric vehicles [J]. Energy Conversion and Management, 2012, 64: 113-21.
- [85] KAMAL M A S, MUKAI M, MURATA J, et al. On board eco-driving system for varying road-traffic environments using model predictive control; proceedings of the 2010 IEEE International Conference on Control Applications, F, 2010 [C]. IEEE.
- [86] FRÖBERG A, NIELSEN L. Optimal fuel and gear ratio control for heavy trucks with piece wise affine engine characteristics [J]. IFAC Proceedings Volumes, 2007, 40(10): 335-42.
- [87] VAJEDI M, AZAD N L. Ecological adaptive cruise controller for plug-in hybrid electric vehicles using nonlinear model predictive control [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 17(1): 113-22.
- [88] 林歆悠, 孙冬野, 尹燕莉, 等. 基于随机动态规划的混联式混合动力客车能量管理策略 [J]. 汽车工程, 2012, (9): 830-6.
- [89] XU B, CHEN X, LI K, et al. Double-layer speed optimization for reducing fuel consumption with vehicle-to-infrastructure communication [J]. Journal of Intelligent Transportation Systems, 2019, 23(5): 513-24.
- [90] ZHU G, LIN J, LIU Q, et al. Research on the Energy-Saving Strategy of Path Planning for Electric Vehicles Considering Traffic Information [J]. Energies, 2019, 12(19): 3601.
- [91] PATTERSON S R, KOZAN E, HYLAND P. Energy efficient scheduling of open-pit coal mine trucks [J]. European Journal of Operational Research, 2017, 262(2): 759-70.
- [92] DEHKORDI S G, LARUE G S, CHOLETTE M E, et al. Including network level safety measures in eco-routing [J]. Journal of Intelligent Transportation Systems, 2022: 1-14.
- [93] LIU R, LIU H, HAN L, et al. A multi-objective regenerative braking control strategy combining with velocity optimization for connected vehicles [J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2023, 237(6): 1465-74.
- [94] WANG J, LIU P, HICKS-GARNER J, et al. Cycle-life model for graphite-LiFePO4 cells [J]. Journal of power sources, 2011, 196(8): 3942-8.
- [95] CORDOBA-ARENAS A, ONORI S, GUEZENNEC Y, et al. Capacity and power fade cycle-life model for plug-in hybrid electric vehicle lithium-ion battery cells containing blended spinel and layered-oxide positive electrodes [J]. Journal of Power Sources, 2015, 278: 473-83.
- [96] ZHANG L, XIA X. An Integer Programming Approach for Truck-Shovel Dispatching Problem in Open-Pit Mines [J]. Energy Procedia, 2015, 75: 1779-84.
- [97] SUTTON R S, BARTO A G. Reinforcement learning: An introduction [M]. MIT press, 2018.
- [98] 陈奇. 铁路运输企业车流分析及关键指标动态考核研究 [D]; 北京交通大学, 2022.
- [99] 徐翔. 基于云模型的管道运输低碳绿色发展研究 [D]; 大连海事大学, 2016.