矿用混合动力装载机能量管理迁移学习策略研究*

张杰[1,2,3,4]，梁龙[1,2,4]，李龙[1,4,5]，张威[6]

（1.长沙矿山研究院有限责任公司， 湖南 长沙 410012；
2.金属矿山安全技术国家重点实验室， 湖南 长沙 410012；
3.北京科技大学， 北京 100083；
4.深部金属矿安全智能开采国家矿山安全监察局重点实验室， 湖南 长沙 410012；
5.战略性稀有金属矿产高效开发国家技术创新中心， 湖南 长沙 410012；
6.重庆工商大学 机械工程学院， 重庆 400060）

摘要：为了提升矿用混合动力装载机能量管理策略对多变循环工况的适应性，提出了一种深度Q网络（Deep Q-Network，DQN）与迁移学习（Transfer Learning，TL）融合的能量管理策略DQN-TL。该方法的核心思想在于利用装载机V型作业的内在循环特性，将这种可重复的模式规律视为一种可提炼、可复用的高价值知识。策略首先利用DQN在数据丰富的源域中学习并提炼出应对典型作业循环的通用控制规律；随后通过参数迁移，将相关规律特征引入目标域模型，并结合少量数据微调以适应特定工况特征，最后通过建立工况的结构化循环特征与迁移学习的内在联系，提升井下复杂工况适应性。仿真结果表明，与基于规则的控制策略相比，所提出的DQN-TL策略的燃油消费率为189.21 g/(kW·h)，燃油经济性提升了13.8%，其功率分配策略更接近于全局最优方案，表现出了较好的工程应用潜力。

关键词：混合动力装载机；深度Q网络；迁移学习；能量管理策略

中图分类号：TD56 **文献标识码**：A **文章编号**：1005-2763(2026)09-0257-09

Research on Transfer Learning Strategy for Energy Management of Mining Hybrid Loaders

Zhang Jie[1,2,3,4], *Liang Long*[1,2,4], *Li Long*[1,4,5], *Zhang Wei*[6]

(1.Changsha Institute of Mining Research Co., Ltd., Changsha, Hunan 410012, China;
2.State Key Laboratory of Safety Technology of Metal Mines, Changsha, Hunan 410012, China;
3.University of Science and Technology Beijing, Beijing 100083, China;
4.Key Laboratory of National Mine Safety Administration for Safe and Intelligent Mining of Deep Metal Mines, Changsha, Hunan 410012, China;
5.National Technological Innovation Center for Efficient Development of Strategic Rare Metal Mineral Resources, Changsha, Hunan 410012, China;
6.School of Mechanical Engineering, Chongqing Technology and Business University, Chongqing 400060, China)

Abstract: To improve the adaptability of the energy management strategy for mining hybrid loaders to variable cyclic operating conditions, an energy management strategy integrating Deep Q-Network (DQN) and Transfer Learning (TL), named DQN-TL, was proposed. The core idea of this method is that the inherent cyclic characteristic of the loader's V-type operation is utilized, and such repeatable patterns are regarded as extractable and reusable high-value knowledge. First, DQN is utilized by the strategy to learn and distill the general control laws for typical operating cycles in a data-rich source domain. Then, through parameter transfer, the relevant law features are introduced into the target domain model, which is then fine-tuned with a small amount of data to adapt to specific operating conditions. Finally, by establishing the intrinsic relationship between the structured cyclic characteristics of the operating conditions and transfer learning, the adaptability to complex underground conditions is improved. The simulation results show that, compared with a rule-based control strategy, the fuel consumption rate of the proposed DQN-TL strategy is 189.21 g/(kW·h), the fuel economy is improved by 13.8%, its power allocation strategy is closer to the global optimal solution, and good potential for engineering application is demonstrated.

Key words: Hybrid loader, Deep Q-Network, Transfer learning, Energy management strategy

* **收稿日期**：2026-04-10
基金项目：中国五矿集团有限公司全国重点实验室专项资金项目(2025GZKJ05)
作者简介：张杰(1987—)，男，江苏无锡人，硕士，高级工程师，主要从事矿山机电设备的研究与开发。E-mail:379148701@qq.com

0 引言

随着全球能源短缺与环境污染问题日益严峻,工程机械作为能源消耗和碳排放的重要领域,其节能减排技术受到广泛关注与研究[1]。混合动力技术通过集成多种能源,能够有效提高能源利用效率、降低排放,已成为工程机械绿色化发展的重要方向之一[2]。其中,混合动力系统的能量管理(Energy Management Strategy, EMS)是直接影响混合动力系统节能潜力的核心技术,直接影响整机的燃油经济性、动力性以及排放性能[1-2]。

现有混合动力工程机械 EMS 大体可分为基于规则、基于优化和基于学习的 3 种策略。基于规则的策略(Rule-Based Strategy, RBS)主要涉及逻辑门限、状态机及模糊控制等方法,其结构清晰、计算量小、易于工程实现[3],但控制参数通常依赖经验标定,面对物料密度、道路坡度、驾驶操作和作业节拍变化时容易出现适应性不足。针对矿用运输装备,Feng Yanbiao 等[4-5]从燃油消耗、动力电池寿命平衡以及全生命周期运行成本角度分别研究了大型混合动力矿用车的最优能量管理问题;Li Hongliang 等[6]针对液压混合动力矿用车构建了规则控制与参数优化相结合的策略。这些研究表明,将部件寿命、运行成本和工况约束纳入 EMS 有助于改善综合性能,但规则策略仍难以保证跨工况最优性。以动态规划(Dynamic Programming, DP)为代表的基于优化的策略可在已知完整工况的条件下给出全局最优基准[7],但是无法直接用于策略的在线优化。等效燃油消耗最小策略和模型预测控制则通过瞬时优化或滚动时域优化提高在线可实施性;然而,这类方法普遍依赖较准确的模型、工况先验或预测信息,其计算负担和参数鲁棒性制约了其在复杂矿山现场的直接应用。

基于学习的策略通过智能体与环境交互直接学习状态到控制动作的映射,为降低 EMS 对精确模型和人工规则的依赖提供了新途径[8]。Liu Teng 等[9]和 Zou Yuan 等[10]较早将 Q 学习及工况转移概率用于混合动力履带车辆,实现了面向不同任务循环的自适应功率分配;Han Xuefeng 等[11]进一步采用双深度 Q 学习缓解 Q 值高估问题,提高了履带车辆能量管理的学习性能。在工程机械领域,Zhang Wei 等[12]构建了融合直接学习与间接学习的混合动力工程机械智能能量管理架构,随后又利用环境模型复用和在线学习形成面向不同作业循环的广义能量管理框架。相关成果说明,矿用及工程装备的重复作业模式能够为数据驱动能量管理提供有效经验,但传统表格型强化学习在连续、高维状态空间中仍易受到维数灾难、收敛速度和样本效率的限制。

深度强化学习将深度神经网络的特征表征能力与强化学习的序贯决策能力相结合,其中深度 Q 网络(Deep Q-Network, DQN)通过经验回放和目标网络提高了价值函数学习的稳定性[13]。Wu Juda 等[14]将深度 Q 学习用于功率分流式混合动力客车的连续能量管理;Tan Huachun 等[15]采用 Actor-Critic 框架处理连续状态与连续动作空间;Wu Yuankai 等[16]和 Li Yuecheng 等[17]分别引入交通信息与道路坡度信息,提高了策略对外部工况的利用能力;Du Guodong 等[18]验证了深度强化学习在混合动力汽车实时能量管理中的优化潜力;Lian Renzong 等[19]通过嵌入专家规则改善了深度强化学习的训练效率和可解释性。文献[20]指出,深度强化学习 EMS 具备在线决策和非线性映射优势,但仍面临训练样本需求量大、收敛过程耗时、训练稳定性不足以及未见工况下泛化能力有限等问题。对于矿用装载机而言,真实工况采集成本高、车型及作业对象差异明显,若每次工况变化均从零训练网络,将显著增加策略开发和标定成本。

迁移学习(Transfer Learning, TL)旨在将源域中获得的知识迁移至相关目标域,从而减少目标域数据需求和重复训练成本[21]。Guo Xiaowei 等[22]将预训练的深度确定性策略梯度网络迁移至新的履带车辆行驶工况,验证了深度强化学习与迁移学习融合的可行性;Chen Hao 等[23]基于真实行驶循环数据构建了 PPO 与迁移学习融合的能量管理策略;Huang Ruchen 等[24]进一步提出跨车型深度迁移强化学习框架,提升了知识复用范围。综上所述,现有迁移强化学习 EMS 主要面向道路车辆、履带车辆或通用行驶循环,针对串联式矿用装载机“发动机-发电机-超级电容”系统及 V 型往复作业循环的研究仍较少,尤其是如何把铲装、运输和卸料阶段的周期重复规律转化为可迁移知识,并在目标工况数据有限时兼顾初始性能、收敛效率与超级电容 *SoC* 约束,尚缺乏有针对性的研究。

基于此,本文以超级电容为辅助能源的串联式矿用混合动力装载机为对象,提出一种深度 Q 网络与参数迁移相结合的能量管理策略(DQN-TL)。首

先,依据实测 V 型作业循环及动力系统特性建立装载机能量管理马尔可夫决策模型,并在数据较充分的源域中训练 DQN,提取典型循环下发动机发电功率与超级电容功率协同分配规律;其次,将源域 Q 网络参数用于目标域网络初始化,并利用少量目标域数据和较小学习率进行微调,使策略在保留通用循环知识的同时适应工况差异;最后,基于实测负载数据,将所提 DQN-TL 策略与基于规则的策略及 DP 全局最优基准进行对比,从发动机输出功率、超级电容 *SoC*、发动机工作点和燃油经济性等方面评价策略性能。与已有主要面向一般行驶循环或跨车型迁移的研究不同,本文重点建立矿用装载机 V 型循环结构特征与知识迁移机制之间的联系,为多变矿山作业条件下能量管理策略的快速构建与复用提供方法参考。

1 模型建立

为研究矿用混合动力装载机能量管理策略,首先需要建立其动力系统的数学模型。这里研究的串联式混合动力装载机主要由发动机、发电机、超级电容组、驱动电机、传动系统及主控制器等组成,其系统结构如图 1 所示。发动机带动发电机发电,产生的电能一部分供给驱动电机驱动车辆及液压系统作业,另一部分为超级电容充电或由超级电容放电补充。能量管理策略的核心在于合理分配发动机发电功率以及超级电容的充放电功率,以满足整车需求功率,同时优化燃油经济性并维持超级电容 *SoC* 在合理范围。

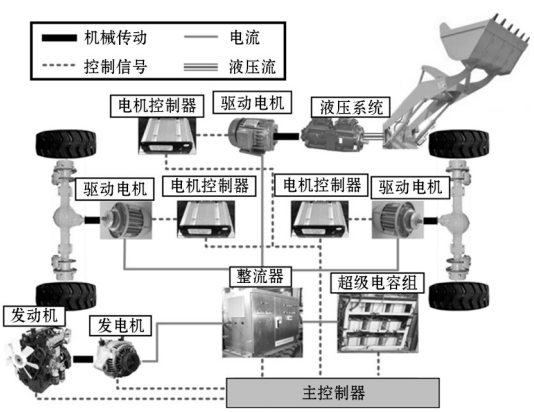


图 1 串联式混合动力装载机系统结构

Fig.1 Structure of series hybrid loader system

1.1 需求功率模型

外部环境及内部系统构型的复杂性使得矿用混合动力装载机的动态模型难以建立。因此,本文对该装载机模型进行了简化,忽略电子负载和风阻,将加载和卸载对应的质量变化简化为两个常数,即空载和满载值。基于动力学分析,工程车辆动力总成模型如式(1)所示[12]:

$$\begin{cases} P_{\mathrm{d}}=\dfrac{60F_{\mathrm{d}}v\eta_{\mathrm{d}}}{4\pi^{2}}+\displaystyle\sum_{i=1}^{i=N}\frac{p_i^h q_i^h}{60\eta_i^h} \\ F_{\mathrm{d}}=[F_{fx}+F_{rx}]-[R_{fx}+R_{rx}]- \\ \qquad m(t)g\sin(\theta)-R_{\mathrm{s}}+\kappa m(t)\dfrac{\mathrm{d}v}{\mathrm{d}t} \end{cases} \tag{1}$$

式中:P_{d} 为系统总(需求)功率;F_{d} 为总(需求)驱动力;v 为装载机当前行驶速度;η_{d} 为驱动电机传动效率;F_{fx} 和 F_{rx} 分别为前后轮的驱动力;R_{fx} 和 R_{rx} 分别为前后轮阻力;$m(t)$ 为负载质量;g 为重力加速度;t 为时间;θ 为路面坡度;R_{s} 为纵向铲阻力;κ 为旋转质量的换算系数;p_i^h、q_i^h 以及 η_i^h 分别表示液压泵的压力、流量和效率。

1.2 发动机-发电机-整流器模型

基于图 2 所示发动机-发电机-整流器等效模型对能量的产生与转化过程进行简化,如式(2)所示[12]:

$$\begin{cases} U_{\mathrm{g}}=C_{\mathrm{e}}\omega_{\mathrm{g}}-C_{\mathrm{x}}\omega_{\mathrm{g}}I_{\mathrm{g}} \\ T_{\mathrm{eg}}=C_{\mathrm{g}}I_{\mathrm{g}}-C_{\mathrm{x}}I_{\mathrm{g}}^{2} \\ \dfrac{T_{\mathrm{e}}}{i_{\mathrm{eg}}}-T_{\mathrm{m}}=\dfrac{2\pi}{60}\left(\dfrac{J_{\mathrm{e}}}{i_{\mathrm{eg}}^{2}}+J_{\mathrm{g}}\right)\dfrac{\mathrm{d}n_{\mathrm{g}}}{\mathrm{d}t} \\ n_{\mathrm{g}}=i_{\mathrm{eg}}n_{\mathrm{e}} \end{cases} \tag{2}$$

式中:U_{g} 为发电机输出电压;C_{e} 为发电机电动势系数;I_{g} 为发电机电流;ω_{g} 为发电机角速度;$C_{\mathrm{x}}\omega_{\mathrm{g}}$ 为等效电阻,电阻系数 $C_{\mathrm{x}}=3P_{\mathrm{N}}L/\pi$;$P_{\mathrm{N}}$ 为极对数;L 为电枢同步电感;T_{eg} 为发电机电磁转矩;C_{g} 为感应电动势系数;T_{e} 为发动机输出扭矩;i_{eg} 为发动机-发电机传动比;n_{e} 为发动机转速;n_{g} 为电机转速;J_{e} 为发动机转动惯量;J_{g} 为发电机转动惯量。

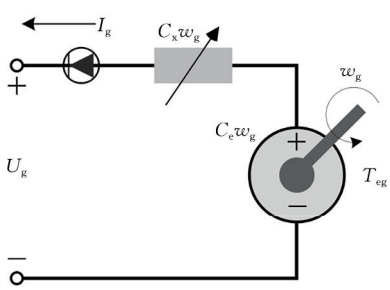


图 2 发动机-发电机-整流器等效模型

Fig.2 Equivalent model of engine-generator-rectifier

1.3 辅助能量源模型

基于RINT等效电路对辅助能量源的超级电容器组进行建模,如式(3)所示[10]:

$$SoC(t)=\frac{U_{id}-i_{uc}R_{uc}}{U_{max}} \tag{3}$$

式中:$SoC(t)$为超级电容在t时刻的荷电状态;U_{id}为理想电容电压;i_{uc}为超级电容电流;R_{uc}为超级电容内阻;U_{max}为最大端电压。

2 基于迁移学习的能量管理策略

针对矿用混合动力装载机在V型往复作业循环下的能量管理问题,传统强化学习方法面临训练数据需求量大、收敛速度慢以及"Sim-to-Real Gap"等挑战。针对上述问题,本文提出了一种基于迁移学习的能量管理策略,旨在利用源域的先验知识加速目标域策略的学习过程,提高策略的性能和适应性。

2.1 能量管理问题描述与强化学习建模

矿用混合动力装载机的能量管理目标是在满足整车功率需求、保证作业性能的前提下,最小化燃油消耗,并维持超级电容SoC在合理范围内。这可以被建模为一个马尔可夫决策过程(Markov Decision Process, MDP),通过强化学习求解,关键要素定义如下。

(1) 状态空间。

状态变量应能充分描述系统当前的工作状态。对于装载机混合动力系统,状态空间$s\in S$定义为:

$$s=[v,SoC(t),P_r,P_h]\in S \tag{4}$$

式中:P_r为整车总需求功率;P_h为液压系统需求功率。

(2) 动作空间。

动作是能量管理控制器在每个决策时刻t可以选择的控制指令$a\in A$。对于串联式混合动力系统,控制变量通常是发动机的输出功率或发电机的功率。选择发动机目标发电功率P_{tar}作为动作a,如式(5)所示:

$$a=[P_{tar}]\in A \tag{5}$$

(3) 奖励函数。

奖励函数R用于评估在状态s下执行动作a后转移到下一个状态s'的好坏,引导智能体学习最优策略,其设计直接影响最终策略的性能。本研究中奖励函数主要考虑燃油消耗和超级电容SoC的维持,如式(6)所示:

$$R=-\beta m_f-\eta P_v \tag{6}$$

式中:β和η分别为燃油消耗和SoC的惩罚权重系数,用于平衡不同优化目标;m_f为发动机瞬时燃油消耗率;P_v为SoC越界惩罚函数;当SoC超出设定合理区间$[SoC_{min},SoC_{max}]$时给予较大负惩罚,其具体表达为:

$$\begin{cases}\text{当 } SoC(t)<SoC_{min}\text{时},P_v=p_{low}[SoC_{min}-SoC(t)]^2\\ \text{当 } SoC(t)>SoC_{max}\text{时},P_v=p_{high}[SoC(t)-SoC_{max}]^2\\ \text{当 } SoC_{min}\leqslant SoC(t)\leqslant SoC_{max}\text{时},P_v=0\end{cases} \tag{7}$$

式中:p_{low}和p_{high}分别为SoC过低或过高时的惩罚系数。

2.2 基于DQN迁移学习的能量管理策略

所提出的迁移学习能量管理策略框架如图3所示,主要包括源域(预训练)、目标域(知识迁移)和模型微调三部分。

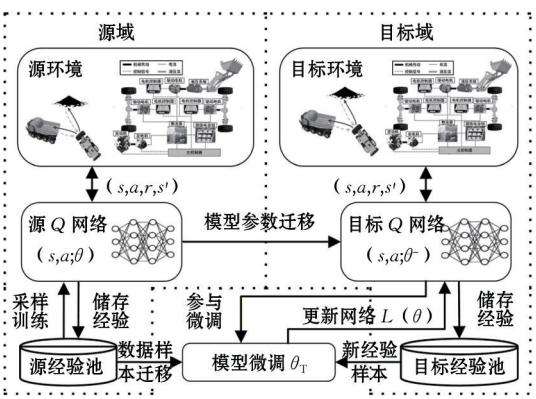


图3 基于迁移学习的能量管理策略框架

Fig.3 Framework of Transfer Learning based Energy Management Strategy

2.2.1 源域(预训练)

源域可以是一个高保真的混合动力装载机仿真模型,或者是一个数据充足的相似工况。此阶段的目标是训练一个具备通用能量管理知识的DQN智能体。

(1) 构建源域环境。基于第1章建立的混合动力装载机模型,构建一个详细的仿真环境,该环境能准确模拟V型往复作业循环的典型特征。

(2) DQN策略训练。DQN算法使用一个深度神经网络$Q(s,a;\theta)$来近似最优动作值函数$Q^*(s,a)$。其中,θ为网络参数,其训练过程主要包括:第一步,经验回放,即将智能体与环境交互产生的经验元组(s,a,r,s')存储在一个固定大小的源经验池$\mathfrak{I}$中,其中s'为执行动作a后下一时刻对应状态,训练

时,从中随机采样小批量数据,以打破数据相关性;第二步,使用一个独立的、参数为 θ^- 的目标网络 $Q(s,a;\theta^-)$ 来计算目标 Q 值,以提高训练的稳定性,目标网络的参数会定期从主网络复制而来;第三步,进行 Q 网络更新,通过最小化损失函数 $L(\theta)$ 来更新主网络参数 θ,如式(8)所示:

$$L(\theta)=E_{(s,a,r,s')\sim\Im}[(y-Q(s,a;\theta))^2] \tag{8}$$

其中,目标 Q 值 y 由目标网络计算得出:

$$y=r+\gamma \max_a Q(s',a;\theta^-) \tag{9}$$

式中,γ 为折扣因子。

(3) 知识提取。训练收敛后,源域 DQN 策略的主网络参数 θ_s 便蕴含了关于能量管理的有效知识。

2.2.2 目标域(知识迁移)

知识迁移的核心是将源域学到的知识应用于目标域。目标域可能是一个参数略有差异的新型号装载机模型,或一个与源域工况有所不同的实际作业场景。

采用参数迁移方法如下。

(1) 网络结构复用,目标域 DQN 模型的网络结构与源域保持一致。

(2) 参数初始化,将源域训练好的 Q 网络参数 θ_s 直接作为目标域 Q 网络的初始参数 $\theta_{T,init}$。如此,可确保目标域的智能体在开始学习前,可通过源域的训练具备价值判断能力,从而避免从零开始的盲目探索。

2.2.3 模型微调

在目标域中,利用少量真实数据或新的仿真数据对迁移过来的模型进行微调,使其适应新环境的特性,主要包括 3 个步骤。

(1) 目标域环境交互。智能体在目标域环境中进行交互,收集新的(s,a,r,s')样本,存入目标域的经验池 $\Im_T$。

(2) 继续训练。使用与源域相同的 DQN 算法对网络进行更新。关键在于,微调阶段的学习率 η_T 通常会设置得比源域训练时的学习率 η_S 更小,表达如式(10)所示:

$$\theta_T \leftarrow \theta_T-\eta_T\nabla_{\theta_T}L(\theta_T)\ (\eta_T<\eta_S) \tag{10}$$

式中,θ_T 为微调后的网络参数。使用较小的学习率 η_T 有助于在学习新知识的同时,最大程度地保留从源域学到的宝贵经验,防止“灾难性遗忘”。

(3) 完成微调。当策略在目标域的性能达到预期或收敛时,停止训练。最终得到由参数 θ_T 构成的优化后的能量管理决策网络。

3 仿真及讨论

为验证所提出的基于深度强化学习与迁移学习的能量管理策略(DQN-TL)的有效性和先进性,在 MATLAB/Simulink 与 Python 平台进行联合仿真试验。仿真模型中的整车参数、动力部件(发动机、发电机、超级电容)特性及作业负载循环数据均来自改装后的矿用混合动力串联式装载机试验平台实测数据,以确保仿真结果的可靠性与现实指导意义。

3.1 测试平台与试验设计

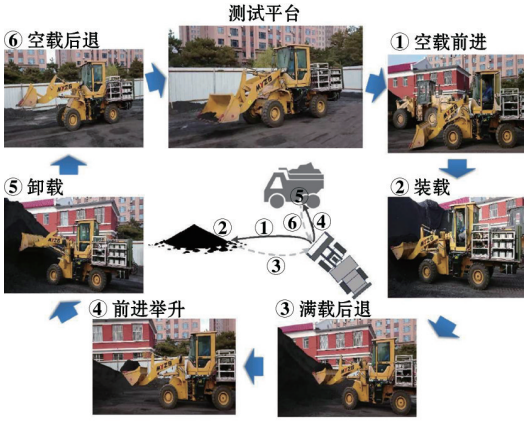
实测数据来自改装后的矿用串联式混合动力装载机测试平台,平台主要参数见表 1。

表 1 测试平台主要参数

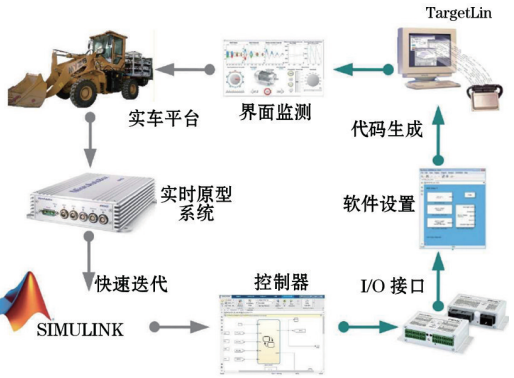
Table 1 Main parameters of the test platform

整车质量/kg	发动机额定功率/kW	发动机最大转速/(r/min)	发电机额定功率/kW	发电机最大转速/(r/min)	电机最大功率/kW	电机最大转速/(r/min)
2 000	60	2 500	50	1 500	40	9 000
超级电容容量/F	超级电容最大电压/V	最大 *SoC*	最小 *SoC*	电机工作电压/V	整流器输出电压/V	整流器输出电流/A
17.5	360	0.95	0.5	290～380	0～400	0～180

为模拟矿用装载机的典型作业过程,实车测试采用经典 V 型往复作业循环,如图 4(a)所示。一个完整循环依次包含 6 种工作模式:①空载前进;②装载;③满载后退;④前进举升;⑤卸载;⑥空载后退。该循环覆盖行驶、铲装、举升和卸料等主要负载阶段,可反映驱动系统与液压系统功率需求的周期性变化。数据监测与样本采集基于 dSPACE 控制原型系统进行,其结构如图 4(b)所示。实车平台运行信号经实时原型系统接入 Simulink 环境,用于模型快速迭代与控制策略验证;控制算法通过 TargetLink 完成代码生成,经 I/O 接口部署至控制器,并通过上位机界面进行在线监测,从而形成“实车平台—实时原型系统—控制器—实车平台”的闭环采集与验证流程。



(a) V型作业实测方案



(b) 数据采集系统框架

图 4 V型作业循环与数据采集系统

Fig.4 V-type operating cycle and data acquisition system

所采集样本采用交叉验证法进行划分,其中源域选取 20 个典型 V 型循环作为 DQN 预训练工况,数据总时长约 2 200 s,对应约 2.2×10^4 个采样点;目标域选取另一独立采集批次中的 5 个循环用于迁移后微调,总时长约为 550 s,对应约 5.5×10^3 个采样点,目标域数据规模约为源域的 25%。此外,另选择 20 个未参与源域训练及目标域微调的完整循环构成独立测试工况,单次对比仿真时长为 2 200 s。

3.2 仿真工况与对比策略设置

为评估策略在典型应用场景下的性能,选取了一段包含多个完整 V 型往复作业循环的实测工况数据作为仿真测试条件,部分数据片段如图 5 所示。设定初始 *SoC* 为 75%,可用工作区间为[0.5, 0.9]。

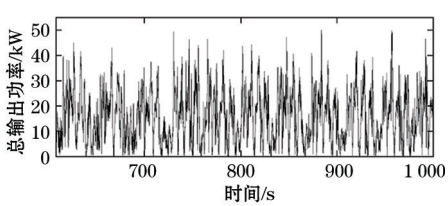


图 5 实测总功率曲线片段

Fig.5 Measured total power curve segment

为进一步评估所提出策略的性能,选取文献[12]中所采用的两类常见的 DP 和 RBS 策略与本文提出的 DQN-TL 策略进行对比分析。联合仿真采用 0.01 s 固定步长,DQN-TL、DP 和 RBS 策略的决策周期及工况数据更新周期均为 0.1 s,3 种策略使用相同的需求功率序列、部件模型参数和初始 *SoC*,以保证对比结果的一致性。

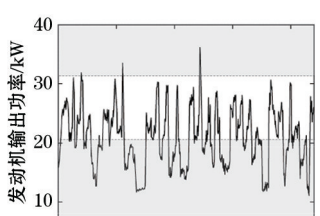
DP 算法能够在已知全局工况信息的前提下,计算出理论上燃油经济性最优的控制序列,其结果被视为评价其他次优策略性能的比较基准(Benchmark)。RBS 策略依据预设的 *SoC* 阈值和需求功率来控制发动机的启停和功率输出,是工程中应用广泛的基准策略,但其性能高度依赖于工况和经验标定。

3.3 仿真结果对比分析

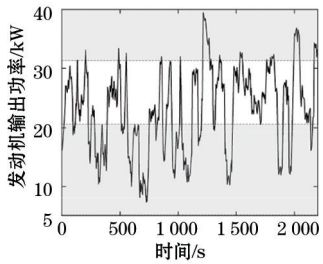
3.3.1 发动机输出功率对比

图 6 展示了 3 种不同能量管理策略下发动机输出功率随时间的变化曲线。图 6(a)中,DP 策略作为全局最优解,其对发动机的控制表现出极强的预见性。发动机一旦启动,便在较长时间内稳定运行于某一高效功率点,避免了频繁的模式切换,其功率曲线平滑且波动小,旨在全局范围内实现最优的能量分配。从图 6(b)可见,RBS 策略表现出典型的“阈值触发”特征。发动机的启停完全由 *SoC* 是否触及预设的上下限决定,导致其在整个工况中频繁启停,且功率输出呈现阶跃式变化,这不利于发动机的稳定运行和燃油经济性。

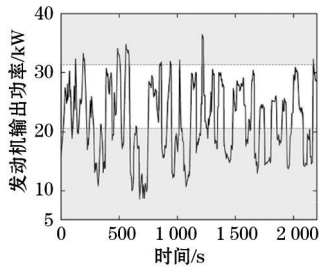
由图 6(c)可知,本文所提出的 DQN-TL 策略,其发动机功率分配趋势与 DP 策略高度相似。智能体通过学习,掌握了对未来短时域功率需求的预测能力,从而避免了规则策略的短视行为。发动机状态观测统计可知,在上述测试过程中 RBS 策略的发动机启停总次数达 12 次,而 DQN-TL 策略将其大幅削减至 4 次,且功率输出振荡幅值方差降低了 18.3%,有效抑制了功率的阶跃式剧烈波动。



(a) DP



(b) RBS



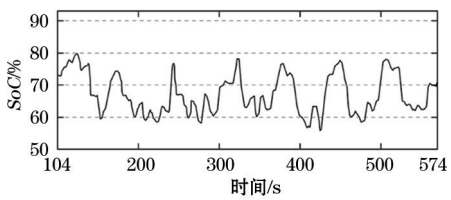
(c) DQN-TL

图 6 3 种策略下发动机输出功率对比

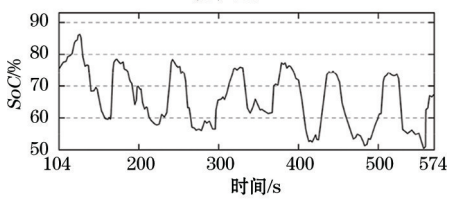
Fig.6 Comparison of engine output power under three strategies

3.3.2 超级电容 *SoC* 变化对比

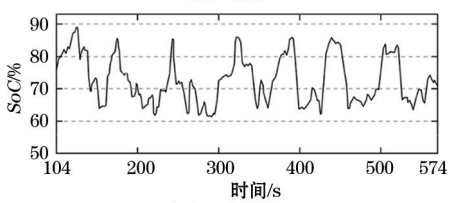
超级电容 *SoC* 的变化轨迹直接反映了策略对电能的利用方式。图 7 为 3 种策略下的 *SoC* 变化曲线。



(a) DP



(b) RBS



(c) DQN-TL

图 7 3 种策略下超级电容 *SoC* 变化曲线

Fig.7 *SoC* variation curves under three strategies

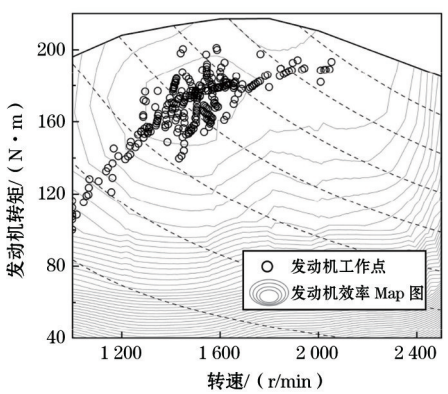
由图 7 可见,RBS 的 *SoC* 曲线在工作区间的上下限之间剧烈振荡。当 *SoC* 较低时,发动机会强制启动充电,当 *SoC* 较高时则停止工作,完全依赖电能,导致电能使用不均衡。

相比之下,DP 策略的 *SoC* 曲线呈现出一条近乎线性的、平缓的下降轨迹。它在整个行程中对电能进行了全局规划,确保在行程结束时恰好消耗掉所有可用电能,最大化地利用了电能的“削峰填谷”作用,这是其实现全局最优燃油经济性的关键。

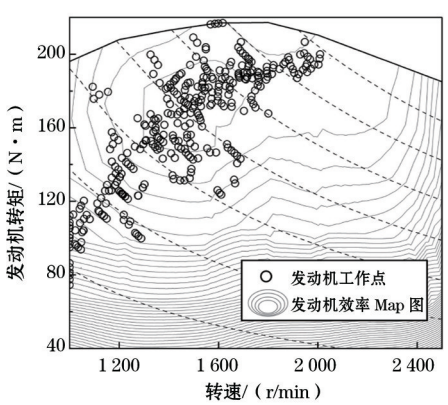
本文提出的 DQN-TL 策略的 *SoC* 曲线与 DP 策略的轨迹更加接近,同样展现了平滑下降的趋势。RBS 的 *SoC* 周期内循环变化浮动达 38%,而 DQN-TL 策略通过全局均衡分配,将其单次波动幅度平滑控制在 25%以内,显著改善了超级电容的充放电负荷。这表明基于 DQN-TL 的智能体通过学习和迁移,掌握了长期规划电能使用的能力,能够根据整体工况动态调整能量分配,而非仅仅响应瞬时状态,从而实现了对能量的优化利用。

3.3.3 发动机工作点分布对比

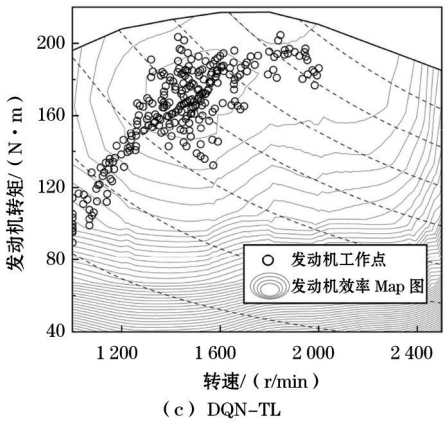
将发动机在整个工况中的工作点(转速-扭矩)叠加到其万有特性图(BSFC Map)上,可以直观地评价策略是否能使发动机工作在高效区,如图 8 所示。



(a) DP



(b) RBS



(c) DQN-TL

图 8 3 种策略下发动机工作点分布

Fig.8 Engine operating point distribution under three strategies

图 8(a)中,DP 策略对应结果作为最优基准,将绝大多数工作点都精确地集中在了万有特性图中的最低燃油消耗率区域(图中等高线表达的"盆地"核心区),最大化了发动机的运行效率。

由图 8(b)可知,RBS 策略作用下的发动机工作点分布较为分散,且有相当一部分工作点落在了燃油消耗率较高的区域(远离最低燃油消耗率区域),这将直接影响该策略的节能效果。

图 8(c)中,DQN-TL 策略作用下发动机工作点相比 RBS 策略呈现出较强向最低燃油消耗率区域聚集的趋势。这说明所提出的 DQN-TL 策略学习到了使发动机工作在经济区的有效方法。分布对比中,其效果显著优于 RBS 策略,并能够更好地逼近 DP 的全局最优解。

3.4 燃油经济性量化分析

为更精确地评估各策略的节油效果,将 3 种策略在完成相同作业任务后的关键性能指标进行了量化统计,结果见表 2。

表 2 3 种策略仿真结果对比

Table 2 Comparison of simulation results of the three strategies

策略	*SoC* 终值	平均燃油消耗率/(g/(kW·h))	节油率/%
RBS	0.41	219.50	—
DP	0.64	181.97	17.1
DQN-TL	0.58	189.21	13.8

由表 2 可知,在相同功率需求及工况特征下,RBS 策略平均燃油消耗率最高,为 219.50 g/(kW·h)。DP 策略作为理论最优解,平均燃油消耗率为 181.97 g/(kW·h),节油率达到 17.1%。

所提出的 DQN-TL 策略平均燃油消耗率为 189.21 g/(kW·h),节油率达到了 13.8%,其优化效果达到了 DP 全局最优潜力的 80%以上。这一结果充分证明,所提出的 DQN-TL 策略无须预知未来全部工况,仅通过对历史经验的学习和迁移,便能制定出性能高度接近全局最优解的控制策略,展现了其在实际应用中的巨大潜力和价值。

4 结论

针对矿用混合动力装载机能量管理问题,提出并验证了一种深度强化学习与迁移学习融合的创新策略。得到以下主要结论。

(1) 所构建的能量管理框架的有效性基于将装载机作业的循环模式向可迁移结构化知识的有效转化。通过"源域-知识迁移-模型微调"的范式,智能体得以继承应对 V 型循环的核心控制逻辑。

(2) 与传统策略相比,基于 DQN-TL 的策略在发动机功率分配、超级电容 *SoC* 规划上表现出显著的预见性和长期规划性,其控制行为对应了 V 型循环中铲装、运输、卸料等不同阶段的能量需求特征,最终使控制效果相比传统的基于规则的策略能够更好地逼近动态规划的全局最优基准。DQN-TL 策略的节油率提升至 13.8%,与 DP 全局最优节油率(17.1%)相差仅 3.3 个百分点,表现出了较大的经济性潜力。

(3) 综合对比表明,所提 DQN-TL 策略的优点在于无需预知全局完整工况及承受 DP 庞大的计算负担,仅依赖有限先验知识即可实现高度逼近全局最优的在线能量分配;其缺点在于若面临与 V 型循环特征差异巨大的随机无规律工况时,策略的泛化能力会受到限制。因此,该策略的适用范围重点面向具备高重复性作业节拍的矿山、港口等封闭或半封闭工程机械作业场景。

参考文献(References):

[1] Trinh Hoai-An, Truong Hoai Vu Anh, Do Tri Cuong, et al. Optimization-based energy management strategies for hybrid construction machinery: a review [J]. Energy Reports,2022,8:6035-6057.

[2] Liu Jichao, Liang Yanyan, Chen Zheng, et al. Energy management strategies for hybrid loaders: classification, comparison and prospect[J]. Energies,2023,16(7):3018.

[3] 丁磊,吴一滔,刘永刚,等.串并联式混合动力汽车能量管理策略优化研究[J].重庆理工大学学报(自然科学),2023(37):37-43.

Ding Lei, Wu Yitao, Liu Yonggang, et al. Optimization of energy management strategy for series-parallel hybrid electric vehicles[J]. Journal of Chongqing University of Technology (Natural Science),2023(37):37-43.

[4] Feng Yanbiao, Dong Zuomin. Optimal energy management with balanced fuel economy and battery life for large hybrid electric mining truck[J]. Journal of Power Sources,2020,454:227948.

[5] Feng Yanbiao, Dong Zuomin. Optimal energy management strategy of fuel-cell battery hybrid electric mining truck to achieve minimum lifecycle operation costs[J]. International Journal of Energy Research,2020,44(13):10797-10808.

[6] Li Hongliang, Zhu Denglin, Shang Lihua, et al. Research on an energy management strategy and energy optimisation of hydraulic hybrid power mining trucks[J]. International Journal of Vehicle Design,2021,85(2-4):246-273.

[7] 王恒,李东风,许刘晴.新能源汽车能量管理的动态规划法应用及优化[J].汽车实用技术,2025(50):31-35.
Wang Heng, Li Dongfeng, Xu Liuqing. Application and optimization of dynamic programming in energy management for new energy vehicles[J]. Automobile Applied Technology, 2025(50):31-35.

[8] 唐香蕉,赵奕凡.动力电池-超级电容混动车辆能量管理研究[J].机械设计与制造,2024:5.
Tang Xiangjiao, Zhao Yifan. Energy management in hybrid vehicles with battery-supercapacitor hybrid systems[J]. Mechanical Design and Manufacturing,2024:5.

[9] Liu Teng, Zou Yuan, Liu Dexing, et al. Reinforcement learning of adaptive energy management with transition probability for a hybrid electric tracked vehicle[J]. IEEE Transactions on Industrial Electronics,2015,62(12):7837-7846.

[10] Zou Yuan, Liu Teng, Liu Dexing, et al. Reinforcement learning-based real-time energy management for a hybrid tracked vehicle[J]. Applied Energy,2016,171:372-382.

[11] Han Xuefeng, He Hongwen, Wu Jingda, et al. Energy management based on reinforcement learning with double deep Q-learning for a hybrid electric tracked vehicle[J]. Applied Energy,2019,254:113708.

[12] Zhang Wei, Wang Jinxin, Xu Zhenyu, et al. A generalized energy management framework for hybrid construction vehicles via model-based reinforcement learning[J]. Energy, 2022(260):124849.

[13] Mnih Volodymyr, Kavukcuoglu Koray, Silver David, et al. Human-level control through deep reinforcement learning [J]. Nature,2015,518:529-533.

[14] Wu Juda, He Hongwen, Peng Jiankun, et al. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus[J]. Applied Energy,2018,222:799-811.

[15] Tan Huachun, Zhang Hailong, Peng Jiankun, et al. Energy management of hybrid electric bus based on deep reinforcement learning in continuous state and action space [J]. Energy Conversion and Management, 2019, 195: 548-560.

[16] Wu Yuankai, Tan Huachun, Peng Jiankun, et al. Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus [J]. Applied Energy,2019,247:454-466.

[17] Li Yuecheng, He Hongwen, Khajepour Amir, et al. Energy management for a power-split hybrid electric bus via deep reinforcement learning with terrain information[J]. Applied Energy,2019,255:113762.

[18] Du Guodong, Zou Yuan, Zhang Xudong, et al. Deep reinforcement learning based energy management for a hybrid electric vehicle[J]. Energy,2020,201:117591.

[19] Lian Renzong, Peng Jiankun, Wu Yuankai, et al. Rule-interposing deep reinforcement learning based energy management strategy for power-split hybrid electric vehicle [J]. Energy,2020,197:117297.

[20] Hu Xiaosong, Liu Teng, Qi Xuewei, et al. Reinforcement learning for hybrid and plug-in hybrid electric vehicle energy management: recent advances and prospects [J]. IEEE Industrial Electronics Magazine,2019,13(3):16-25.

[21] Pan Sinno Jialin, Yang Qiang. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering,2010,22(10):1345-1359.

[22] Guo Xiaowei, Liu Teng, Tang Bangbei, et al. Transfer deep reinforcement learning-enabled energy management strategy for hybrid tracked vehicle [J]. IEEE Access, 2020, 8: 165837-165848.

[23] Chen Hao, Guo Gang, Tang Bangbei, et al. Data-driven transferred energy management strategy for hybrid electric vehicles via deep reinforcement learning [J]. Energy Reports,2023,10:2680-2692.

[24] Huang Ruchen, He Hongwen, Su Qichong. Towards a fossil-free urban transport system: an intelligent cross-type transferable energy management framework based on deep transfer reinforcement learning[J]. Applied Energy,2024, 363:123080.