



基于改进 TD3 算法的储能系统多目标优化调度策略研究

樊 荣

(驻马店职业技术学院科研与信息化处, 河南 驻马店 463000)

摘 要: 针对源荷波动下传统储能调度适应性、收敛性不足问题, 本文提出基于改进 TD3 的多目标优化调度策略。以 TD3 算法为基础, 引入优先级经验回放机制, 构建兼顾经济性、储能损耗与电网波动的多目标奖励函数, 搭建三层智能化储能调度框架。依托微电网仿真算例开展验证, 结果表明, 所提算法收敛速度显著提升, 可降低系统运行成本、削减电网交互功率峰值, 同时减少储能频繁启停损耗。

关键词: TD3 算法; 储能系统; 多目标优化

doi: 10.19799/j.cnki.2095-4239.2026.0775

中图分类号: TM 73

文献标志码: A

文章编号: 2095-4239 (2026) 09-3633-03

Research on multi-objective optimization scheduling strategies for energy storage systems based on the improved TD3 algorithm

FAN Rong

(Zhumadian Vocational and Technical College, Scientific Research and Informatization Division, Zhumadian 463000, Henan, China)

Abstract: To address the shortcomings of adaptability and convergence of traditional energy storage scheduling under source-load fluctuations, this paper proposes a multi-objective optimal scheduling strategy based on an improved TD3 algorithm. Based on the TD3 algorithm, a prioritized experience replay mechanism is introduced, a multi-objective reward function that takes into account economy, energy storage loss and power grid fluctuations is constructed, and a three-layer intelligent energy storage scheduling framework is established. Verification is carried out based on a microgrid simulation example, and the results show that the proposed algorithm significantly improves the convergence speed, can reduce system operation costs, cut the peak value of power grid interaction power, and meanwhile reduce the loss caused by frequent start-stop of energy storage.

Keywords: TD3 algorithm; energy storage systems; multi-objective optimization

能源是国民经济发展和社会运作的重要基础, 合理开发和利用能源一直是社会研究热点^[1]。随着人工智能技术的快速发展, 深度强化学习技术在机器人控制、资源调度等领域取得极大进展, 并逐步应用于电力系统, 为储能系统智能化调度提供了全新技术路径^[2]。TD3 算法适用于储能充放电功率等

收稿日期: 2026-09-07。

基金项目: 2025 年度河南省软科学研究计划项目 (252400410368); 河南省二〇二四年科技发展计划软科学研究项目 (242400410491)。

作者简介: 樊荣 (1994—), 男, 硕士, 研究方向为计算机应用技术, E-mail: fanrong2025@163.com。

引用本文: 樊荣. 基于改进 TD3 算法的储能系统多目标优化调度策略研究[J]. 储能科学与技术, 2026, 15(9): 3633-3635.

Citation: FAN Rong. Research on multi-objective optimization scheduling strategies for energy storage systems based on the improved TD3 algorithm[J]. Energy Storage Science and Technology, 2026, 15(9): 3633-3635.

连续动作决策,但仍存在样本利用率低、收敛较慢及多目标协调能力不足等问题。本文改进TD3算法,构建多目标优化调度模型,搭建智能化调度框架,通过仿真对比验证所提策略的优越性,为储能系统高效稳定运行提供新的技术方案。

1 储能系统多目标优化调度建模

1.1 储能系统运行数学模型

储能系统运行体现为荷电状态SOC的动态变化,并受充放电功率、转换效率与设备物理的边界约束。本文以电化学储能为研究对象,忽略储能内部复杂电化学过程,建立功率-SOC时序演化模型。储能充放电状态下荷电状态迭代计算公式表示为式(1)。

$$\text{Soc}(t+1) = \text{Soc}(t) + \frac{[\eta_{\text{ch}} P_{\text{ch}}(t) - P_{\text{dis}}(t)/\eta_{\text{dis}}] \Delta t}{E_N} \quad (1)$$

式中, $\text{Soc}(t)$ 为 t 时刻储能荷电状态; $P_{\text{ch}}(t)$ 、 $P_{\text{dis}}(t)$ 分别为 t 时刻充电、放电功率; η_{ch} 、 η_{dis} 为充电、放电效率; Δt 为仿真时间步长; E_N 为储能额定容量。储能运行需满足充放电功率与SOC边界约束,并保证同一时刻充、放电动作互斥。

1.2 系统运行优化目标及约束条件

本文将系统运行经济性、储能寿命保护和电网友好性作为多目标优化调度的三个子目标,综合考虑购电成本、峰谷套利收益以及储能损耗代价^[9],并通过加权方式将其转化为综合优化目标。系统运行首先需满足功率平衡约束,微网内新能源出力、储能充放电功率、负荷功率与电网交互功率应保持平衡。其次为储能本体运行约束,储能充放电功率需维持在安全运行区间,并满足充放电互斥约束。电网交互功率不得超过限值。上述目标与约束共同构成储能系统多目标优化调度问题的数学边界,并为改进TD3智能体的状态空间、动作空间和奖励函数设计提供依据。

2 基于改进TD3算法的多目标优化调度策略

2.1 针对储能调度场景的TD3算法改进

传统深度强化学习算法应用于储能调度时,存在样本利用效率低、奖励信号稀疏,训练过程易出现震荡收敛慢的问题。本文在TD3算法中引入带权重的优先级经验回放机制,提高高价值调度样本的

采样概率,降低无效样本对网络参数更新的干扰。同时,考虑储能多目标优化调度需求,构造加权多目标联合奖励函数,兼顾运行经济性与储能寿命损耗。改进后的奖励表达式为式(2)。

$$R_t = \lambda_1 C_{\text{cost}}(t) - \lambda_2 C_{\text{loss}}(t) - \lambda_3 |P_{\text{grid}}(t)| \quad (2)$$

式中, $C_{\text{cost}}(t)$ 为 t 时刻系统购电成本; $C_{\text{loss}}(t)$ 为储能等效损耗代价; $P_{\text{grid}}(t)$ 为微网与电网交互功率; λ_1 、 λ_2 、 λ_3 分别为各项目标对应的权重系数。通过调整权重系数协调经济性、储能寿命与电网功率波动目标,提高算法的收敛速度和策略稳定性。

2.2 多目标优化调度框架构建

本文基于改进TD3算法,搭建环境感知、智能决策、设备执行三层储能调度框架。环境感知层以15 min为调度步长,采集微网负荷、新能源出力、分时电价及储能SOC的数据,并进行降噪和归一化处理。智能决策层依托多目标奖励约束机制和源荷波动动态输出最优充放电功率指令。设备执行层接收指令驱动储能变流器动作,响应延时控制在百毫秒级,实时平抑系统功率偏差。该框架可平衡电网运行稳定性、储能使用寿命与系统运行经济性,满足复杂微电网场景的储能多目标优化调度需求。

2.3 算法训练流程

本文改进TD3算法采用离线训练、在线部署的方式,共训练1200轮,单轮迭代包含96个调度时序步长。训练初期,读取环境感知层预处理后的源荷、电价、储能状态数据,初始化Actor网络、两个Critic网络及其对应的目标网络,并初始化优先级经验回放池。在每个时序步长中,智能体根据当前状态输出储能充放电动作,计算机多目标奖励,并将状态、动作、奖励及下一状态样本存入优先级经验回放池。训练过程中按优先级采集经验,更新目标网络参数。当连续100轮平均奖励值趋于稳定时,判定模型收敛,保存最优网络参数。训练后的Actor网络模型部署至储能调度系统,用于在线输出充放电功率指令。

3 算例仿真与结果分析

3.1 算例场景及参数设置

为验证基于改进TD3算法的储能系统多目标优化调度策略的有效性,搭建含光伏、负荷、储能及并网接口的微电网仿真场景,以典型日间源荷时序数据作为仿真输入,调度步长为15 min,单日共

96个调度时段。仿真选用额定容量500 kWh、额定功率250 kW的电化学储能系统，储能充放电效率均取0.92，SOC安全运行区间为0.2~0.8。电价采用工商业分时电价，峰段电价0.85元/kWh、谷段电价0.33元/kWh。改进TD3算法共训练1200轮，网络学习率设为 1×10^{-3} ，经验回放池容量为10000，权重系数 λ_1 、 λ_2 、 λ_3 分别取值0.6、0.2、0.2。同时设置传统固定阈值调度、原始TD3算法作为对比方案，保证仿真对比的客观性与全面性。

3.2 仿真结果分析

本文从算法收敛性能、运行经济性与电网适配性三个维度，对三种案进行仿真结果对比。收敛结果表明，相较于原始TD3算法，改进算法收敛轮次提前约220轮，训练震荡幅度显著降低，样本利用率与策略稳定性提升。在运行经济层面，所提策略充分利用峰谷电价差套利，日均购电成本较传统阈值调度降低12.6%，较原始TD3算法降低5.3%。在电网波动抑制方面，微网与外网交互功率峰值削减18.4%。仿真结果验证了改进算法可兼顾系统经济运行、设备寿命保护与电网稳定运行的多重需求。

4 结 语

针对源荷波动下传统储能调度适应性、收敛性

不足问题，本文提出基于改进TD3算法的储能系统多目标优化调度策略。通过搭建储能系统数学模型，设计兼顾经济、寿命与电网波动的多目标奖励函数，构建分层智能调度框架，改善传统算法收敛慢、稳定性差的问题。结果表明，该策略可有效加快模型收敛速度，降低微电网运行成本，平抑外网功率波动，并减少储能启停损耗，兼顾了调度经济性与系统稳定性。后续可引入多智能体算法，开展混合储能协同调度研究，进一步提升复杂场景适配能力。

参 考 文 献

- [1] 梁海峰, 闫峰, 尚隽, 等. 基于深度强化学习的综合能源系统优化调度[J]. 内蒙古电力技术, 2025, 43(4): 21-29.
LIANG H F, YAN F, SHANG J, et al. Optimized scheduling of integrated energy system based on deep reinforcement learning [J]. Inner Mongolia Electric Power, 2025, 43(4): 21-29.
- [2] 林伟. 低压微电网的储能系统配置与优化调度[J]. 科技与创新, 2025 (19): 126-129. DOI:10.15913/j.cnki.kjycx.2025.19.036.
LIN W. Configuration and optimal scheduling of energy storage system in low-voltage microgrid[J]. Science and Technology & Innovation, 2025(19): 126-129. DOI: 10.15913/j. cnki. kjycx.2025.19.036.
- [3] 张墨阳, 周伟香. 基于深度强化学习的微电网能量管理策略优化分析[J]. 集成电路应用, 2025, 42(12): 660-661.
ZHANG M Y, ZHOU W X. Analysis of microgrid energy management strategy optimization based on deep reinforcement learning[J]. Application of IC, 2025, 42(12): 660-661.