

DOI:10.14044/j.1674-1757.pcrpc.2024.04.013

基于强化学习的串联超级电容器组非能耗均衡方法

宋倩¹, 蓝俊欢²

(1. 河池学院大数据与计算机学院, 广西 河池 546300; 2. 广西电网有限责任公司河池供电局, 广西 河池 546300)

摘要:超级电容器作为新兴储能元件的代表,在创新驱动发展战略实施背景下使用日益广泛,为解决串联超级电容器组使用中各单体电压偏差大、利用效率不高的问题,充分利用人工智能技术,提出了一种基于强化学习的串联超级电容器组非能耗均衡方法,将近端策略优化算法(PPO)应用于串联超级电容器组电压均衡中,在 Matlab/Simulink 仿真平台中搭建了串联超级电容器组和强化学习环境并验证了该算法的有效性,相较于传统的能耗型电压均衡方法,均衡效率较好,超级电容器的损耗较小,实验结果表明了 PPO 算法可实现串联超级电容器组电压的均衡。

关键词:强化学习; PPO 算法; 人工智能; 超级电容器; 储能

Non-energy-consumption Equalization Method for Series Super Capacitor Banks Based on Reinforcement Learning

SONG Qian¹, LAN Junhuan²

(1. College of Big Data and Computer Science, Hechi University, Guangxi Hechi 546300, China;
2. Hechi Power Supply Bureau, Guangxi Power Grid Co., Ltd., Guangxi Hechi 546300, China)

Abstract: As a representative of emerging energy storage components, super capacitor are increasingly used in the context of the implementation of the innovation-driven development strategy. For solving such problems as large voltage deviation and low utilization efficiency of each single unit in the use of series-connected super capacitor banks and fully using artificial intelligent technology, a kind of reinforcement learning-based non-energy equalization method for series super capacitor banks is proposed, and the proximal policy optimization operator (PPO) is applied to the voltage equalization of series super capacitor banks. The super capacitor bank and reinforcement learning environment are set up in Matlab/Simulink simulation platform and the effectiveness of the algorithm is verified. The super capacitor, compared to the traditional energy consumption voltage equalization method, has good equalization efficiency and small loss. The experimental results show that the PPO algorithm can equalize the voltage of series super capacitor bank.

Keywords: reinforcement learning; PPO algorithm; artificial intelligence; super capacitor; energy storage

0 引言

在新时代人工智能正大踏步向我们走来,其中机器学习、深度学习、大数据技术吸引了学者和民众最多的目光。人工智能覆盖从科技、艺术创作,甚至法律、社会治理,人工智能的研究和应用边界正在急速扩展^[1],这是百年未有之大变局下中国的科技和社会里的一个新浪潮。我国十四五规划中,人工智能是科技发展和贯彻新发展理念的重点,文中提到“智能”

“智慧”不下五十处,深刻表明了党和国家对人工智能技术发展壮大的紧迫感,这事关我国建设创新型国家和推动经济高质量发展^[2-3]。强化学习作为人工智能的代表之一,其在电压均衡方面有一定的优势^[4]。

超级电容器对比常用的锂离子电池,其电能的存储并不需要化学反应,而是一种电荷的纯物理迁移^[5-8]。无论充电或者放电,理论上超级电容器都不会消耗或耗散能量,所以拥有极快的充放电速度。超级电容器充放电容量大、效率高、循环寿命长,能

收稿日期:2023-03-20

基金项目:2022年广西高校中青年教师科研基础能力提升项目(2022KY0606);2023年河池学院校级科研平台(2023XJPT001)。

耐超低温,能量回收效率高,在未来的储能系统中极具发展潜力^[9-11]。但超级电容器单体额定电压低,一般不超过3 V,因此常需将数个单体串联起来使用,为了使每个单体电压保持一致,不产生多余的能量损耗,需要对各单体的电压进行均衡,均衡技术是串联电容器组应用的关键一环^[7,12]。

串联超级电容器组目前在储能领域广泛应用^[13-15],人工智能技术也正如火如荼地发展。本文将强化学习中的近端策略优化算法应用于串联超级电容器组的电压均衡中,将传统电压均衡问题定义为强化学习问题。相较于传统的电压均衡方法,如并联电阻、电感法,飞度电容法、变压器均衡法^[16-18],基于近端策略优化算法的串联超级电容器组非能耗均衡方法所需的开关器件少,驱动电路和控制策略简单,均衡效率较好,超级电容器的损耗较小,通过实验验证了PPO算法应用于串联超级电容器组电压均衡中的有效性。

1 强化学习理论

强化学习基本流程见图1,基本分类见图2。PPO算法,即近端策略优化算法,属于策略优化中一种典型的算法。

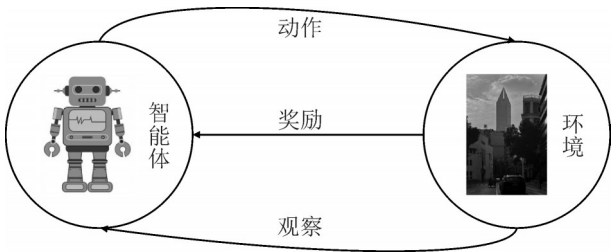


图1 强化学习基本流程
Fig. 1 Basic process of reinforcement learning

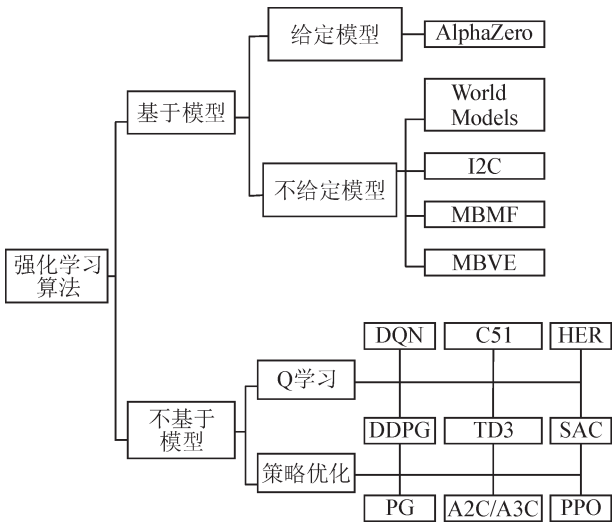


图2 强化学习基本分类
Fig. 2 Basic classification of reinforcement learning

PPO算法具有TRPO算法的部分优点,同时具备更好的通用性、更易实施的优势。PPO由actor网络和critic网络2个网络组成。critic网络为了获取即时动作的优势函数,需要对环境和奖励之间的关系进行学习,为达到获得高奖励值概率的目的,actor网络需要基于优势函数对策略的参数进行实时调整。

PPO算法缩短了新旧策略间距离,在于优势函数的剪切机理得到了有效应用,实现了新策略在原旧策略的近端输出。PPO算法将actor的损失函数更改为公式(1)

$$L^{CLIP}(\theta) = E_t \left[\min \left(r_t(\theta) \hat{A}_t, \text{clip} \left(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon \right) \hat{A}_t \right) \right] \tag{1}$$

式中: $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$; π_θ 、 $\pi_{\theta_{old}}$ 分别为新策略和旧策略; θ 为策略参数; $r_t(\theta)$ 为新旧两策略的比值。clip是裁剪函数,作用是提高算法的整体效率和性能,在降低算法实现难度的前提下,也保留了TRPO算法的稳定性优点,因其界定了新旧策略的变化范围在 $(1 - \varepsilon, 1 + \varepsilon)$ 之间。 $\hat{A}_t$ 是优势函数估计值。另一种计算方法如式(2),两个概率分布差异的程度用KL来表征,KL即散度也称相对熵。

$$L^{KL PEN}(\theta) = E_t \left[r_t(\theta) \hat{A}_t - \beta KL \left[\pi_{old}(\cdot|s_t), \pi_\theta(\cdot|s_t) \right] \right] \tag{2}$$

式中, β 为超参数,其更新策略可借鉴下式

$$d = E_t \left[KL \left[\pi_{\theta_{old}}(\cdot|s_t), \pi_\theta(\cdot|s_t) \right] \right], \tag{3}$$

$$\text{if } d < \frac{d_{targ}}{1.5}, \beta = \frac{\beta}{2} \text{ or if } d > 1.5d_{arg}, \beta = 2\beta$$

critic网络参数更新参考公式为

$$L_t^{VF} = \left(V_\theta(s_t) - V_t^{targ} \right)^2 \tag{4}$$

式中, $V_\theta(s_t)$ 、 V_t^{targ} 分别是状态估计价值和目标价值奖励。

综上所述得出如表1所示伪代码。

表1 PPO算法伪代码
Table 1 Pseudo-code of PPO algorithm

Algorithm 1 PPO, Actor-Critic style
for iteration = 1, 2, ... do
for actor = 1, 2, ..., N do
run policy $\pi_{\theta_{old}}$ in environment for T timesteps
Compute advantage estimate $\hat{A}_1, \dots, \hat{A}_t$
end for
Optimize surrogate L wrt θ , with K epochs and minibatch size $M \leq NT$
$\theta_{old} \leftarrow \theta$
end for

2 仿真曲面重构

为基于强化学习的串联超级电容器组电压均衡技术提供超级电容器放电数据,本文将利用基于Maxwell公司电池研发中心采集的超级电容器放电数据。由于对数据的采样方式是对间隔进行采样,因此无法得到精确的函数表达式,需要进行拟合。对所需的三维曲面, x 轴, y 轴, z 轴分别选取超级电容器放电时间、放电倍率、单体电压。

在一般情况下,表达式大多是未知的,曲面重构只能利用少量有限的的数据,已知表达式的情形是很少出现的。对此方面的相关研究也不少,像根据局部特征的大小使用非均匀采样的方式。它的特点之一是重建剩余点数量满足要求时,既提高了曲面重建速度,又减轻了其所需数据量。由于本文所用数据样本点之间存在关联,可用数据点数量也多,多项式拟合算法、局部加权回归算法是常用的曲面重构方法,下文简要分析。

多项式拟合的目的,对于离散的数据点,也能够用确定的函数表达式来描述,而且针对样本数据集中没有出现的点,相应的值也能够利用拟合的函数推算出结果。二次多项式曲面的公式为

$$f(x,y) = ax^2 + by^2 + cxy + dx + ey + f \quad (5)$$

式中, a 、 b 、 c 、 d 、 e 、 f 为各项系数,为了满足求解所达到的要求,需要不断连续地进行迭代更新,最小二乘法是常用的方法,原始数据和拟合数据的差值用误差平方和来衡量,数值越小越好,文中取 -0.05 、 -0.12 、 -2.35 、 -2.21 、 -1.52 。

另外局部加权回归散点平滑法也是一种常用的拟合方法,它的主要步骤是先选取一部分散点,以这部分散点为代表拟合出部分曲面,再利用部分到整体的思想拟合出全体的曲面。这一方法相较于多项式拟合法效果更佳,但也存在一定弊端,即选取散点数量多寡会显著影响拟合的效果,若选取数量不足则会明显削弱拟合效果,以致出现凸包和凹坑。

整体的多项式拟合算法在曲面重构上稳定性和效果更优于局部加权回归散点平滑法,前者凸包和凹坑较少,曲面更显滑顺。在强化学习中使用多项式拟合算法优势更突出,因其有可更新迭代的特点。

3 基于PPO算法的串联超级电容器组电压均衡控制方法

3.1 PPO算法基本框架

电压均衡技术是超级电容器组应用中一项重要的技术,对于提升超级电容器组的使用寿命和效

率都具有十分重要的意义^[19-22]。下文介绍对超级电容器组应用PPO算法进行电压均衡的方法。

常用的超级电容器组均衡方法有基于荷电状态和基于电压的均衡技术^[23],本文研究电压均衡技术,上节重构曲面是强化学习的环境,超级电容器的电压状态经过初始化后,通过PPO网络生成正态分布,输出每个动作的输出概率,同时将每一批的数据进行保存,并更新参数,经过critic网络和最小化优势函数。深度强化学习中的PPO算法,在解决连续动作空间和巨大离散动作空间的学习上具有比基于值函数更好的效果。

电压均衡方法实现思路见图3。

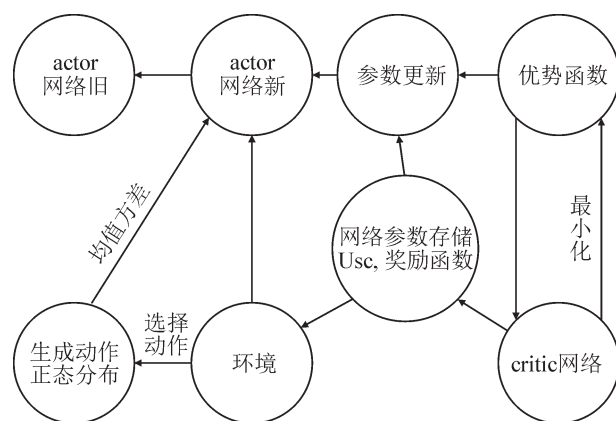


图3 PPO算法基本框架
Fig. 3 Basic framework of PPO algorithm

3.2 输入状态及输出动作的设定

输入状态 state 选取为3个超级电容器单体的电压 V_1 、 V_2 、 V_3 ,3个单体电压不同,期望3个单体的电压达到一致。

有了输入,还需要确定模型的输出动作,将动作设置为连续的动作空间,以更准确地模拟超级电容器实际工况下的放电情形,输出状态空间选取放电倍率,放电倍率需不同且连续,以确保actor网络的输出是不同放电倍率的选择概率,从而达到每个电压不同的超级电容器单体选择不同的放电速率,最后3个单体电压达到一致。

3.3 奖励函数定义

由于需要对超级电容器组中每节单体的电压进行均衡,则设计的奖励函数不仅需考虑到每节的单体电压,还需要考虑各单体电压之间的离散程度,离散度越小则电容器组间的电压差也越小,这时就需要最大化这一奖励,同时良好的奖励函数设计需要综合考虑效率和成本。设计的奖励函数为

$$r = -\beta \sqrt{\frac{1}{N} \sum_{i=1}^N (V_i - \mu)^2} + \delta \quad (6)$$

期望是超级电容器组单体的电压差逐渐减小,三个单体的电压趋于一致,则 $\frac{1}{N}\sum_{i=1}^N(V_i-\mu)^2$ 逐渐减小,同时希望奖励值 reward 是增大的,故式前的系数取负数,常数 δ 起调节奖励值大小的作用。

3.4 actor和 critic 的设置

ac 网络机制大体为:actor 神经网络的输入为超级电容器单体的电压,输出放电倍率并生成相应动作与环境发生交互, critic 神经网络的输入也同样是单体电压,通过使用价值函数对 actor 的行为进行评价打分,之后更新动作,同时对下个时刻的动作进行指导,TD 误差是 critic 网络对指标进行评估的依据,其公式为

$$\delta(t)=R_{t+1}+\gamma V(S_{t+1})-V(S_t) \tag{7}$$

式中: R_{t+1} 为实时奖励; $V(S_t)$ 为状态价值函数; γ 为折扣因子。

PPO 算法一般有一个 actor 旧网络和一个 actor 新网络和一个 critic 网络,经过一轮处理样本数据学习后,接续进行新 actor 网络和新 critic 网络的学习,同时把相应参数复制给旧的 actor 网络,最终完成整个网络参数的更新。actor 和 critic 网络的层数为 3 层,输入为 3 个超级电容器单体的电压,actor 网络的激活函数为 Relu 函数和 Sigmoid 函数,输出为均值与方差; critic 网络的激活函数为 Relu 函数,输出为损失函数。actor 和 critic 网络的结构和相应超参数的设置见表 2。

表2 PPO 算法主要超参数
Table 2 Main hyperparameters of PPO algorithm

超参数名称	参数描述	取值
Batch_size	每批处理样本数	24
EP_max	幕的数量	1 200
EP_len	每一幕的最大长度	100
A_LR	actor 网络学习率	10^{-4}
C_LR	critic 网络学习率	10^{-4}

4 实验结果与分析

为解决串联超级电容器组中存在的单体电压不均衡现象,本文采用强化学习中的 PPO 算法,选取 3 个超级电容器单体组成串联超级电容器组,各单体的初始电压分别为 2.7 V、2.4 V、2.2 V。强化学习算法的最终目的是使得 3 个单体的电压达到一致,即使得模型的奖励越来越大,下述的训练共包括一千两百个回合,以 24 个数据为一个 batchsize 进行训练,结果显示未发生参数更新弥散和爆炸,参数较适宜。

3 节超级电容器单体的初始电压为 2.7 V、

2.4 V、2.2 V,作为初始状态,在训练的一千两百个回合中,奖励值的数值变化见图 4,明显可见在前八百回合中,奖励值的波动很大,模型在不停调整输出动作以获得更高的奖励值。大概在一千回合左右,模型的奖励值达到稳定,得到一个较为合理和理想的值。

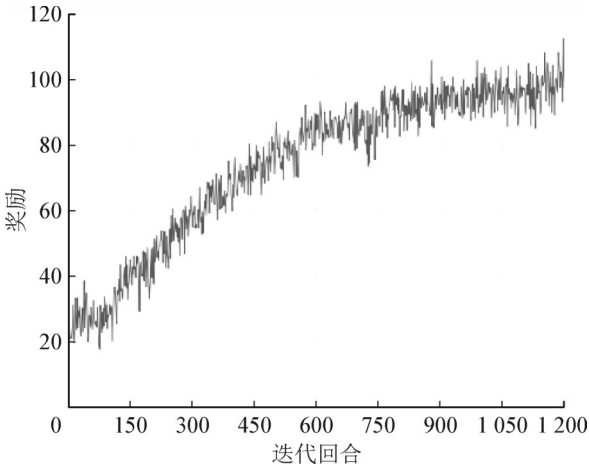


图4 奖励值变化图
Fig. 4 Variation of reward value

除了考虑奖励值之外,损失函数也是一个必不可少的指标,图 5 是 actor 网络的损失值变化情况,由图可知,actor 网络的损失值在经过波动后逐渐趋于零,actor 网络约在一千回合左右趋于零。显示了该算法的有效性,训练模型达到了预期的效果。

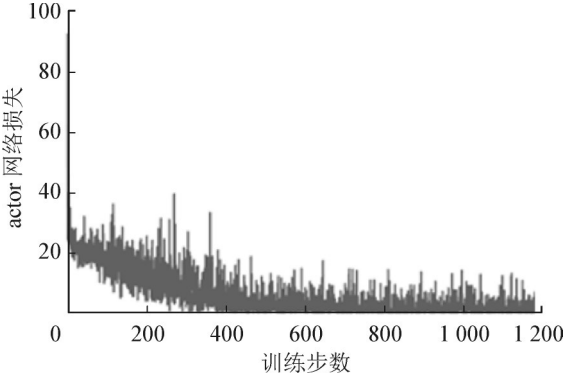


图5 actor 网络损失函数变化图
Fig. 5 Variation of loss function of actor network

串联超级电容器组 3 个单体电压的变化曲线图见图 6。

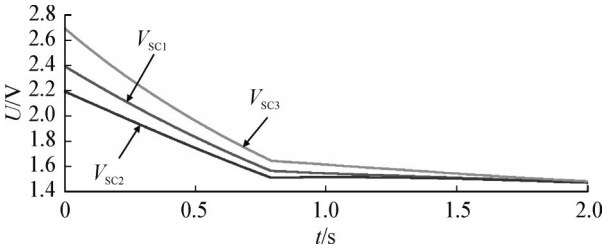


图6 串联超级电容器组单体电压变化图
Fig. 6 Voltage variation of single unit of series supercapacitor bank

3节单体的初始电压分别为2.7 V、2.4 V、2.2 V, SC1的电压下降较快, SC3的电压下降较慢, 在大约2 s时3节单体的电压趋于一致, 约为1.5 V, 表明了电压最高的超级电容器单体选取了最高的放电倍率, 电压低的单体相应地选择了较低的放电倍率, 因为不同单体都需要根据实时电压选择放电倍率以获取最大的奖励, PPO算法实现了这一目的, 通过不断实时地调整神经网络参数。相比串联超级电容器组能耗均衡方法, 此种均衡方法对超级电容器组的能量损耗较小, 能更好维持电容器组的寿命, 提高使用效率。最大压差值从0.5 V缩小到0.1 V以内, 方差也得到明显降低, 表明了PPO算法应用于串联超级电容器组电压均衡中的作用。

5 结语

串联超级电容器组目前在储能领域广泛应用, 人工智能技术也正如火如荼地发展, 将强化学习算法中的近端策略优化算法应用于串联超级电容器组的电压均衡中, 相较于传统的电压均衡方法, 如并联电阻、电感法, 飞度电容法, 变压器均衡法, 此方法均衡效率较好, 超级电容器的损耗较小, 通过实验验证了PPO算法应用于串联超级电容器组电压均衡中的有效性。本文的算法是基于数据集的电池数据, 最后的结果虽证明了该实验的有效性, 但应用到实际的电池管理系统中需进一步考虑更加复杂的串联超级电容器组运行工况。

参考文献

- [1] 倪祥淦, 何志刚, 胡帅, 等. 基于BI-LSTM神经网络的宽采样频率电池SOH估算[J]. 车用发动机, 2022(5): 44-50.
NI Xianggan, HE Zhigang, HU Shuai, et al. Battery SOH estimation with wide sampling frequency based on BI-LSTM neural network[J]. Vehicle Engine, 2022(5): 44-50.
- [2] 孙彧, 曹雷, 陈希亮, 等. 多智能体深度强化学习研究综述[J]. 计算机工程与应用, 2020, 56(5): 13-24.
SUN Yu, CAO Lei, CHEN Xiliang, et al. Overview of multi-agent deep reinforcement learning[J]. Computer Engineering and Applications, 2020, 56(5): 13-24.
- [3] 赵冬梅, 陶然, 马泰屹, 等. 基于多智能体深度确定策略梯度算法的有功-无功协调调度模型[J]. 电工技术学报, 2021, 36(9): 1914-1925.
ZHAO Dongmei, TAO Ran, MA Taiyi, et al. A coordinated active-reactive power scheduling model based on multi-intelligent depth-determined policy gradient algorithm[J]. Journal of Electrical Engineering Technology, 2021, 36(9): 1914-1925.
- [4] 杨挺, 赵黎媛, 刘亚闯, 等. 基于深度强化学习的综合能源系统动态经济调度[J]. 电力系统自动化, 2021, 45(5): 39-47.
YANG Ting, ZHAO Liyuan, LIU Yachuang, et al. Dynamic economic dispatch of integrated energy systems based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2021, 45(5): 39-47.
- [5] 费陈, 宋树祥, 夏海英, 等. 基于XGBoost预测和ARIMA残差校正的锂离子电池SOC估算[J]. 武汉理工大学学报, 2019, 41(10): 94-101.
FEI Chen, SONG Shuxiang, XIA Haiying, et al. SOC estimation of Lithium-ion batteries based on XGBoost prediction and ARIMA residual correction[J]. Journal of Wuhan University of Technology, 2019, 41(10): 94-101.
- [6] 罗勇, 祁朋伟, 黄欢, 等. 基于容量修正的安时积分SOC估算方法研究[J]. 汽车工程, 2020, 42(5): 681-687.
LUO Yong, QI Pengwei, HUANG Huan, et al. Research on capacity correction-based SOC estimation method for ampere-time integration[J]. Automotive Engineering, 2020, 42(5): 681-687.
- [7] 刘晨扬, 陈思哲, 范元亮, 等. 一种基于重构电路的超级电容器快速电压均衡器[J]. 广东电力, 2021, 34(1): 30-40.
LIU Chenyang, CHEN Sizhe, FAN Yuanliang, et al. A fast voltage equalizer for supercapacitor based on reconfigured circuit[J]. Guangdong Electric Power, 2021, 34(1): 30-40.
- [8] 刘建刚, 杨胜杰, 王仲梅, 等. 基于协同控制的串联超级电容器电压均衡策略[J]. 控制与决策, 2021, 36(8): 1997-2001.
LIU Jiangang, YANG Shengjie, WANG Zhongmei, et al. Cooperative control-based voltage equalization strategy for series supercarpacitors [J]. Control and Decision Making, 2021, 36(8): 1997-2001.
- [9] 闫湖, 黄碧斌, 李琼慧, 等. 退役动力电池政策智能模拟方法研究[J]. 太阳能学报, 2022, 43(4): 30-35.
YAN Hu, HUANG Bibin, LI Qionghui, et al. Research on intelligent pre-evaluation of retirement power battery policy[J]. Acta Energaie Solaris Sinica, 2022, 43(4): 30-35.
- [10] 尚彦赓, 宋红为, 杨照光, 等. 基于二阶RC模型的锂电池充放电特性分析[J]. 高压电器, 2023, 59(7): 87-94.
SHANG Yanyun, SONG Hongwei, YANG Zhaoguang, et al. Charge and discharge characteristics analysis of lithium battery based on second-order RC model[J]. High Voltage Apparatus, 2023, 59(7): 87-94.
- [11] 谢国梁. 基于机器学习的锂电池组SOC预测及均衡控制研究[D]. 大连: 大连理工大学, 2021.
- [12] 李宋, 苗志林, 叶满园, 等. 七电平逆变器电容电压自均衡控制策略研究[J]. 高电压技术, 2022, 48(7): 2785-2793.
LI Song, MIAO Zhilin, YE Manyuan, et al. Research on self-balancing control strategy for capacitor voltage of seven-level inverter[J]. High Voltage Technology, 2022, 48(7): 2785-2793.

- 2793.
- [13]廖力,徐雅雯,姜久春.串联锂离子电池组均衡策略研究[J].电源技术,2023,47(1):51-56.
LIAO Li, XU Yawen, JIANG Jiuchun. Research on equalization strategy of series Lithium ion battery pack[J]. Chinese Journal of Power Sources, 2023, 47(1): 51-56.
- [14]李建林,王哲,李雅欣,等.退役动力电池梯级利用分选技术研究[J].太阳能学报,2023,44(1):418-425.
LI Jianlin, WANG Zhe, LI Yaxin, et al. Research on sorting technology of cascade utilization of retires power batteries[J]. Acta Energaie Solaris Sinica, 2023, 44(1): 418-425.
- [15]李郑斐,阚加荣,冯敏星,等.一种基于多变压器的电压均衡电路[J].电子设计工程,2022,30(22):121-126.
LI Zhengfei, KAN Jiarong, FENG Minxing, et al. A voltage equalization circuit based on multi-transformers[J]. Electronic Design Engineering, 2022, 30(22): 121-126.
- [16]吴铁洲,徐瑜鸿,黄一恒,等.电池组电压-SOC分段均衡控制方法研究[J].电力电子技术,2022,56(7):91-94.
WU Tiezhou, XU Yuhong, HUANG Yiheng, et al. Study of battery pack voltage - SOC segmental equalisation control method [J]. Power Elect-ronics, 2022, 56(7): 91-94.
- [17]唐鑫,欧阳权,黄俊卉,等.基于深度强化学习的锂电池快速充电控制策略[J].机械工程学报,2022,58(22):69-78.
TANG Xin, OUYANG Quan, HUANG Lianghai, et al. Fast charging control for lithium-ion batteries based on deep reinforcement learning[J]. Journal of Mechanical Engineering, 2022, 58(22): 69-78.
- [18]杜海忠,骆滔,宋浩谊,等.基于电压平衡的锂电池主动均衡电路及策略[J].电气工程学报,2021,16(3):145-151.
DU Haizhong, LUO Tao, SONG Haoyi, et al. Active balancing circuit and strategy of Li-ion battery based on voltage balance[J]. Journal of Electrical Engineering, 2021, 16(3): 145-151.
- [19]熊蓉玲,段春怡,冉华明,等.基于深度强化学习的智能决策方法[J].电讯技术,2023,63(1):1-6.
XIONG Rongling, DUAN Chunyi, RAN Huaming, et al. An intelligent decision making method based on deep reinforcement learning[J]. Telecommunication Engineering, 2023, 63(1): 1-6.
- [20]王雨欣,王思怡,杨黎晖,等.超级电容储能和卸荷电路协调控制的永磁同步风电机组低电压穿越策略[J].高压电器,2023,59(4):177-185.
WANG Yuxin, WANG Siyi, YANG Lihui, et al. Low-voltage ride-through control strategy of permanent magnetic synchronous wind turbine with coordination of super capacitor energy storage and chopper circuit[J]. High Voltage Apparatus, 2023, 59(4): 177-185.
- [21]王京阳,韩子娇,董雁楠,等.含超级电容的光伏制氢系统建模与研究[J].东北电力技术,2022,43(10):44-48.
WANG Jingyang, HAN Zijiao, DONG Yannan, et al. Modeling and research on photovoltaic hydrogen production system with supercapacitor[J]. Northeast Electric Power Technology, 2022, 43(10): 44-48.
- [22]张智光,倪秋龙,廖培,等.计及新能源消纳的地区电网无功电压优化控制策略[J].浙江电力,2023,42(1):46-53.
ZHANG Zhiguang, NI Qiulong, LIAO Pei, et al. An optimal reactive voltage control strategy for regional power grids considering new energy consumption[J]. Zhejiang Electric Power, 2023, 42(1): 46-53.
- [23]张沛,朱驻军,谢桦.基于深度强化学习近端策略优化的电网无功优化方法[J].电网技术,2023,47(2):562-570.
ZHANG Pei, ZHU Zhujun, XIE Hua. Reactive power optimization based on proximal policy optimization of deep reinforcement learning[J]. Power System Technology, 2023, 47(2): 562-570.

作者简介:

宋倩(1995—),女,硕士,讲师,研究方向为强化学习和储能技术。E-mail:1711563632@qq.com。

蓝俊欢(1995—),男,硕士,助理工程师,研究方向为储能技术和电网运行。E-mail:Linklist_sq@163.com。