

基于改进SAC算法的微电网优化调度

雷强,武鹏荣,李振文

(西安工程大学 电子信息学院,陕西 西安 710600)

摘要:针对风电出力的间歇性与不确定性导致微电网风电消纳率低、系统购电成本高问题,提出一种基于类残差软演员评论家(R-SAC)算法的微电网智能调度策略。首先,将调度问题建模为部分可观测马尔可夫决策过程(PO-MDP),并结合长短期记忆网络(LSTM)与注意力机制对风电出力及负荷变化进行短期预测。然后,在传统软演员评论家(SAC)算法的基础上引入类残差网络结构,构建改进型R-SAC算法以提升策略的收敛速度与探索效率。最后,基于西北地区某实际微电网数据进行仿真实验,结果验证了所提策略的有效性及其优越性。

关键词:微电网调度;风电消纳;深度强化学习;状态估计;策略网络优化

中图分类号:TM73 **文献标识码:**A **DOI:**10.19457/j.1001-2095.dqcd26592

Optimal Scheduling of Microgrid Based on Improved SAC Algorithm

LEI Qiang, WU Pengrong, LI Zhenwen

(School of Electronic Information, Xi'an Polytechnic University, Xi'an 710600, Shaanxi, China)

Abstract: In response to the intermittency and uncertainty of wind power generation, which result in low wind power utilization and high electricity purchase costs in microgrids, an intelligent dispatch strategy based on a residual-like soft actor-critic (R-SAC) algorithm was proposed. The scheduling problem was formulated as a partially observable Markov decision process (PO-MDP), and short-term predictions of wind power output and load variations were achieved by combining long short-term memory (LSTM) networks with the attention mechanism. A residual-like network structure was then incorporated into the traditional soft actor-critic (SAC) framework to construct an improved R-SAC algorithm, which enhanced the convergence speed and exploration efficiency of the policy. Finally, simulation experiments based on data from an actual microgrid in the northwest region validated the effectiveness and superiority of the proposed strategy.

Key words: microgrid scheduling; wind power utilization; deep reinforcement learning (DRL); state estimation; policy optimization

随着“双碳”战略的持续推进及分布式新能源的广泛应用,风力发电在微电网系统中的渗透率持续增长^[1]。然而,风电出力的高度间歇性与不可预测性使得传统调度机制难以实现其有效接纳,对调度策略的自适应性及鲁棒性提出了更高求^[2-3]。

为应对上述挑战,研究者已提出多种调度优化方法。传统的动态规划(dynamic programming, DP)、模型预测控制(model predictive control, MPC)和混合整数线性规划(mixed-integer linear programming, MILP)等数学优化方法,在模型准确且可预测扰动较小时具备良好解析性,但在面

对高维动态系统和外部扰动下的实时调度时,鲁棒性与灵活性较差^[4-5]。

近年来,深度强化学习(deep reinforcement learning, DRL)以其模型无关性、自适应策略搜索能力及在线优化特性,在微电网调度问题中得到广泛关注^[6-7]。代表性算法如深度确定性策略梯度(deep deterministic policy gradient, DDPG)、双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)和柔性演员-评论家(soft actor-critic, SAC),在处理高维状态空间与连续动作控制问题中性能突出,已被应用于储能控制、可调负荷管理以及电价响应等任务中^[8-9]。

但当前方法普遍存在对状态完全可观测性的依赖,忽略预测误差和输入扰动可能引发的策略不稳定问题;同时,策略网络多采用传统堆叠结构,缺乏结构优化机制,导致训练过程易出现梯度消失与策略震荡,影响调度性能^[10-12]。

基于此,本文提出一种融合状态估计与结构优化的深度强化学习微电网调度策略,在建模层面,基于PO-MDP构建统一调度框架,刻画系统中风电、负荷与电价等状态变量不可完全观测的现实特性;在感知层面,设计融合LSTM与注意力机制的状态估计模块,以提升系统对关键变量变化趋势的预测精度与时序依赖建模能力;在策略优化层面,构建引入类残差网络结构的改进型R-SAC算法,强化高维输入下的策略表达能力与训练收敛稳定性。

1 微电网优化建模

1.1 微电网结构与数学模型

1.1.1 微电网结构

本文所研究的微电网系统由风力发电系统、储能系统和负荷组成,负荷主要包括普通负荷与温控负荷(thermostatically controlled loads, TCLs)。微电网通过公共耦合点(point of common coupling, PCC)与主电网相连接,使得微电网能够与外部电网进行能量交换。图1展示了该微电网的系统结构。

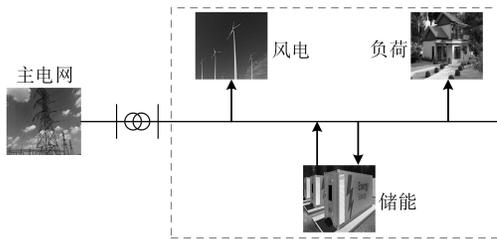


图1 微电网系统结构

Fig.1 Structure of the microgrid system

1.1.2 风电出力模型

风电系统的功率 P_w 输出受到风速 v_w 变化的影响。在微电网系统中,通常使用分段函数来描述风速与风电功率之间的关系。具体来说,风电功率 P_w 可表示为

$$P_w = \begin{cases} 0 & v_w \leq v_{ci} \\ 0 & v_w \geq v_{co} \\ P_r & v_r \leq v_w \leq v_{co} \\ P_r \left(\frac{v_w - v_{ci}}{v_r - v_{ci}} \right) & v_{ci} \leq v_w \leq v_r \end{cases} \quad (1)$$

式中: P_r 为风机的额定功率,kW; v_{ci} 为启动风速,m/s; v_{co} 为停机风速,m/s; v_r 为额定风速,m/s。

1.1.3 储能系统模型

储能系统的工作状态由其充放电状态(state of charge, SOC)来描述,其状态更新公式为

$$SOC_{t+1} = \begin{cases} SOC_t + \frac{\eta_c P_b}{E_{max}} & P_b \geq 0 \\ SOC_t - \frac{P_b}{\eta_d E_{max}} & P_b < 0 \end{cases} \quad (2)$$

式中: P_b 为储能的充放电功率; E_{max} 为储能系统的最大容量; η_c, η_d 分别为储能系统的充电效率和放电效率。

1.1.4 温控负荷模型

TCLs系统能够响应外部的电力需求和风电的波动。其热动力学方程如下:

$$T_{t+1} = T_t + \Delta t \left[-\frac{1}{\tau} (T_t - T_a) + \frac{P_{TCL}}{C} \right] \quad (3)$$

式中: T_t 为当前时刻的室内温度,°C; τ 为热时间常数,h; C 为热容量,kW·h/°C; P_{TCL} 为TCLs运行功率,kW; T_a 为环境温度,°C。

1.2 目标函数

为了同时优化微电网的经济效益和风能消纳率,提出目标函数描述如下:

1)目标1:最小化购电成本。购电成本是指微电网从主电网购买电力的费用,目标函数可以表示为

$$\min C_{grid}(a) = \sum_{t=0}^T [\lambda_{buy}(t) \cdot P_{buy}(t) \cdot \Delta t] \quad (4)$$

式中: a 为决策变量; $\lambda_{buy}(t)$ 为时间 t 的购电价格,元/(kW·h); $P_{buy}(t)$ 为时间 t 的购电功率,kW; Δt 为时间步长,h。

2)目标2:最大化风电消纳率。为提升风电资源的本地消纳能力,目标函数设计如下:

$$\max U_{wind}(a) = \frac{\sum_{t=0}^T [P_w(t) \cdot \Delta t]}{\sum_{t=0}^T [P_w^{max}(t) \cdot \Delta t]} \quad (5)$$

式中: $P_w(t)$ 为 t 时的实际风电功率输出,kW; P_w^{max} 为 t 时的风电机组的最大理论功率输出,kW。

3)定义综合优化目标函数如下:

$$\min J(a) = \beta_1 \cdot C_{grid}(a) - \beta_2 \cdot U_{wind}(a) \quad (6)$$

其中, β_1 和 β_2 为目标权重因子,权值满足 $0 \leq \beta_i \leq 1$,且 $\beta_1 + \beta_2 = 1$ 。

2 深度强化学习框架

2.1 部分可观测马尔可夫决策过程

在传统的强化学习框架中,通常假设智能体能够完全观测到环境状态,并能够获取所有影响决策的关键信息。然而,在许多实际应用中,特别是在微电网调度中,这一假设并不成立。

为解决这个问题,本节建立了部分可观测马尔可夫决策过程(PO-MDP)作为优化框架。与传统的马尔可夫决策过程(MDP)模型不同,PO-MDP允许智能体在部分信息缺失的情况下,通过历史观测推测环境的当前状态。这种方法使智能体能够在不完全观测的情境下进行有效决策,显著提高了其在面对不确定性时的适应能力和决策效率。

2.1.1 状态和观测空间

系统的状态包含风电发电量、电池的充电状态、TCLs的温度状态以及主电网的电力需求等。定义系统状态 S_t 为

$$S_t = [D_t \ P_{W_t} \ E_{B_t} \ T_{TCLst}] \quad (7)$$

式中: D_t 为 t 时刻的电力需求,kW; P_{W_t} 为 t 时刻风电发电量,kW; E_{B_t} 为 t 时刻储能荷电状态,取值范围[0,1]; T_{TCLst} 为TCLs的温度状态,℃。

值得说明的是,主电网电价是由市场供需关系决定的外部因素,其变化独立于智能体的控制行为,且不可直接观测或预测。因此,未将其纳入系统状态的定义中,而是作为外部扰动变量进行建模,影响系统的状态转移过程。

状态转移函数为

$$S_{t+1} \sim P(S_{t+1} | S_t, a_t, \xi_t) \quad (8)$$

其中, ξ_t 表示主电网在 t 时刻的分时电价。虽然 ξ_t 未被纳入系统状态的定义,但作为外部扰动变量被引入状态转移函数中,并直接参与状态演化机制。具体而言, ξ_t 的取值将在每一时刻决定购电成本,从而影响动作 a_t 中与电网交互相关的行为,如购电功率的选择或储能设备的充放电决策。在执行动作 a_t 后,系统的多个状态变量将根据 ξ_t 的取值进行更新。如较高的 ξ_t 可能促使智能体减少从主电网的购电,增加本地储能的放电,从而导致下一状态中的储能状态升高、主电网购电量降低。

定义每个时刻的观测 O_t 为

$$O_t = [h_{D_{t-1}} \ h_{W_{t-1}} \ E_{B_{t-1}} \ T_{TCLst-1}] \quad (9)$$

式中: $h_{D_{t-1}}$ 为截至 $t-1$ 时刻的历史电力需求,kW;

$h_{W_{t-1}}$ 为截至 $t-1$ 时刻的风电数据,kW; $E_{B_{t-1}}$ 为截至 $t-1$ 时刻的储能状态取值范围[0,1]; $T_{TCLst-1}$ 为截至 $t-1$ 时刻的TCLs温度,℃。

2.1.2 动作空间

在每个时刻 t ,智能体需要选择一个适当的动作空间 A_t 来控制微电网的负荷调度和储能管理,确保最小化购电成本和最大化风电的消纳。定义动作空间 A_t 包含以下三个控制量:

$$A_t = [SOC_t \ P_{W_t}^c \ P_{TCLst}] \quad (10)$$

式中: SOC_t 为储能系统的充放电控制量,kW; $P_{W_t}^c$ 为风电消纳量,kW; P_{TCLst} 为TCLs负荷的调节量,kW。

2.1.3 奖励函数

智能体的目标是 minimized 运行成本并最大化风电消纳。设计奖励函数如下:

$$r(S_t, A_t) = -\beta_1 \cdot C_{grid}(t) + \beta_2 \cdot U_{wind}(t) \quad (11)$$

2.1.4 状态转移概率

在PO-MDP中,状态转移通常具有不确定性,并且不完全由智能体的动作所决定。状态转移概率 $Pr(S_{t+1} | S_t, A_t)$ 反映了从一个时刻的状态到下一个时刻状态的演变规律。因此状态转移概率可以表示为

$$Pr(S_{t+1} | S_t, A_t) = Pr(D_{t+1} | D_t) \cdot Pr(PW_{dt+1} | PW_{dt}) \cdot Pr(E_{B_{t+1}} | E_{B_t}, D_t, PW_{dt}, AED_t) \cdot Pr(T_{TCLst+1} | T_{TCLst}, A_t) \quad (12)$$

负荷的转移概率 $Pr(D_{t+1} | D_t)$ 和风电的转移概率 $Pr(PW_{dt+1} | PW_{dt})$ 基于历史数据进行预测,而电池充电状态转移概率 $Pr(E_{B_{t+1}} | E_{B_t}, D_t, PW_{dt}, AED_t)$ 和TCLs的转移概率 $Pr(T_{TCLst+1} | T_{TCLst}, A_t)$ 则可以通过物理模型计算得到。

2.1.5 策略

策略 π 是从每个状态 S_t 到动作 A_t 的映射,决定了智能体如何根据观测选择最优的调度动作。策略可以是确定性的,也可以是随机的。智能体通过深度强化学习算法不断优化策略 π ,目标是最大化长期奖励,即最小化运行成本和最大化风电消纳。

2.2 基于LSTM和注意力机制的状态估计

在第2.1节的部分可观测马尔可夫决策过程模型中,已经描述了如何在部分可观测环境下利用历史信息来推断系统的当前状态。为了进一步提高调度决策的准确性,本节将结合长短时记忆网络(LSTM)和注意力机制,针对风电出力 and 负荷需求建立状态预测模型,用于弥补环境中这两类核心变量的不确定性。

LSTM网络作为适用于时间序列建模的递归神经网络,能够有效学习输入数据的时序依赖关系。为全面表达微电网系统的动态变化,模型将风电出力 P_{w_t} 、负荷 L_t 与储能状态 E_{Bt} 在每个时间步 t 拼接为联合特征向量:

$$\mathbf{x}_t = [P_{w_t}, L_t, E_{Bt}] \quad (13)$$

从而构成多维时间序列输入:

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \quad (14)$$

通过这种方式,LSTM能够学习多源信息在时间维度上的联合演化规律,从而更准确地捕捉系统状态的变化趋势。其状态更新过程如下:

$$f_t = \sigma(W_f \cdot [h_{t-1}, \mathbf{x}_t] + b_f) \quad (15)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, \mathbf{x}_t] + b_i) \quad (16)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_c \cdot [h_{t-1}, \mathbf{x}_t] + b_c) \quad (17)$$

$$h_t = O_t \cdot \tanh(c_t) \quad (18)$$

式中: \mathbf{x}_t 为当前时刻输入数据; h_{t-1} 为上一时刻的输出状态; h_t 为当前时刻输出状态; $\sigma(\cdot)$ 为激活函数; W_f, W_i, W_c 分别为遗忘门权重、输入门权重、候选状态权重; f_t, i_t, c_t 分别为当前时刻系统的遗忘门、输入门和单元状态; b_f, b_i, b_c 分别为遗忘门偏置、输入门偏置、候选状态偏置。

为了进一步增强对关键时刻历史状态的建模能力,在LSTM输出基础上引入注意力机制,按时间维度对隐状态序列进行加权融合。该机制根据不同时间步对当前状态估计的重要性分配权重,其计算方式如下:

$$\alpha_t = \frac{\exp(w^T h_t)}{\sum_{t'=1}^T \exp(w^T h_{t'})} \quad (19)$$

$$\hat{S}_t = \sum_{t'=1}^T \alpha_{t'} \cdot h_{t'} \quad (20)$$

式中: α_t 为第 t 时刻的注意力权重; w^T 为注意力机制参数; T 为时间序列的总长度; \hat{S}_t 为加权后的状态估计; $\alpha_{t'}$ 为第 t' 时刻的注意力权重; $h_{t'}$ 为第 t' 时刻的输出状态。

通过这一加权机制,模型能够突出历史时间序列中对当前状态预测贡献更大的时刻,从而有效提升状态估计的准确性。

3 改进型SAC算法调度策略

3.1 SAC算法

SAC算法是一种基于最大熵强化学习框架的离线策略深度强化学习算法,通过最大化策略的预期回报与熵的加权和,引导策略在获得高回

报的同时保持足够的随机性,从而提升训练效率与泛化能力。该算法采用双 Q 网络与目标网络缓解策略过估计问题,并引入随机策略进行动作采样,结合重参数技巧优化策略梯度,具有良好的收敛性和样本效率。

3.2 基于类残差网络的SAC算法改进

在微电网调度任务中,系统状态受风速、负荷、电价等多源扰动影响,具有高度的不确定性与动态性。传统SAC算法网络结构在处理此类高维、部分可观测输入时,易出现梯度消失、策略收敛缓慢甚至不稳定等问题。

针对上述挑战,提出一种改进型SAC算法—R-SAC算法,通过在策略网络和价值网络中引入类残差网络结构,优化原始网络设计。R-SAC算法在网络的每一层中增加跳跃连接,即将状态输入直接添加到每一层输出中,作为下一层网络的联合输入。

本文所提出的类残差网络结构是基于ResNet的残差块结构,与标准残差块不同之处在于跳跃连接的频率有所调整。标准残差块通常每两层进行一次跳跃连接,而类残差网络结构则在每一层都执行该操作,具体网络结构如图2所示。

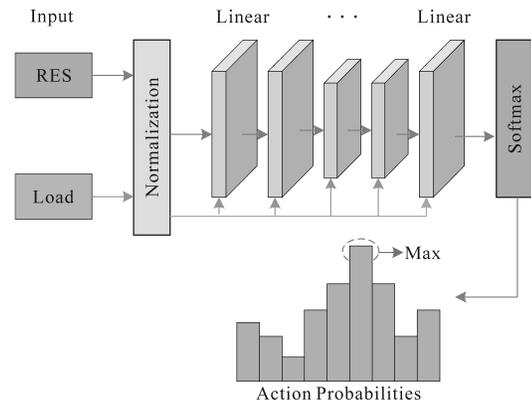


图2 类残差网络结构

Fig.2 Residual-like network architecture

类残差网络结构能够缓解深度网络训练问题的核心原因在于其优越的反向传播特性。假设第 i 层残差块的映射为

$$F(x_i, \omega_i) = x_i + f(x_i, \omega_i) \quad (21)$$

式中: $F(x_i, \omega_i)$ 为第 i 层的输出; $f(x_i, \omega_i)$ 为输入 x_i 和网络参数 ω_i 的非线性激活函数。

那么第 $i+1$ 层的输入为

$$x_{i+1} = F(x_i, \omega_i) = x_i + f(x_i, \omega_i) \quad (22)$$

循环这个计算,得到第 n 层的输入为

$$x_N = x_i + \sum_{n=i}^{N-1} F(x_n, \omega_n) \quad (23)$$

式(21)~式(23)显示了良好的反向传播性。假设损失为*l*,根据链式推导法则,可以得到:

$$\frac{\partial l}{\partial x_i} = \frac{\partial l}{\partial x_N} \frac{\partial x_N}{\partial x_i} = \frac{\partial l}{\partial x_N} \left[1 + \frac{\partial}{\partial x_i} \sum_{n=i}^{N-1} F(x_n, \omega_n) \right] \quad (24)$$

在式(24)中,损失函数*l*对输入*x_i*的梯度可分解为两部分:一部分为通过网络加权后的值,另一部分为恒等映射的值。这种线性组合确保了即使网络权重很小,梯度也不会消失,从而信息能够有效传递至浅层网络。

4 实验结果分析

4.1 实验设置与状态估计

为验证所提出的微电网优化调度方法的有效性,本文基于本地高性能计算平台构建了微电网仿真系统。实验平台配置包括 Intel Core i7-12700F 处理器、32 GB 内存和 NVIDIA GeForce RTX 3060 显卡,软件环境基于 Python 3.8 与 PyTorch 1.13。实验数据采集自中国西北地区某实际微电网,涵盖风速、负荷功率和分时电价等关键变量。

微电网各单元的出力范围与物理参数如表1所示,为调度系统的建模与仿真提供基础支撑。

表1 微电网单元参数

Tab.1 Parameters of microgrid components				
单元	功率上限/kW	功率下限/kW	容量/(kW·h)	初始温度/°C
WT	24	0	—	—
BEES	50	-50	300	—
TLCs	20	0	—	24

主电网采用分时电价机制,价格水平随不同时段进行调整,用以反映市场供需状态,具体购电与售电电价如表2所示。

表2 主电网实时电价

Tab.2 Real-time electricity prices of the main grid		
时段	购电电价/ [元·(kW·h) ⁻¹]	售电电价/ [元·(kW·h) ⁻¹]
00:00—08:00	0.2	0.1
08:00—11:00	0.95	0.5
11:00—18:00	0.5	0.2
18:00—23:00	0.95	0.5

图3展示了本文所提预测模型在风电出力预测任务中的表现。从整体趋势来看,风电功率在

03:00—08:00及20:00—22:00呈现明显波动上升趋势。预测曲线在两个主波动区段均与实际值保持高度一致,拐点位置和变化速率基本吻合,表明所建模型具有良好的动态响应能力与时序建模精度。

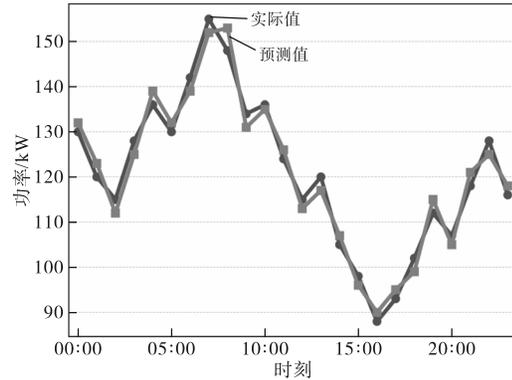


图3 风电出力预测与实际值对比

Fig.3 Comparison between predicted and actual wind power output

图4展示了负荷预测结果与实际值的对比。在典型高峰期(08:00—11:00,18:00—20:00)及低谷期(00:00—06:00),预测值与实际需求保持高度吻合,最大误差低于5%。在负荷突升突降的过渡时段,模型依然保持良好拟合效果,说明其对非平稳负荷信号具有较强的泛化能力。

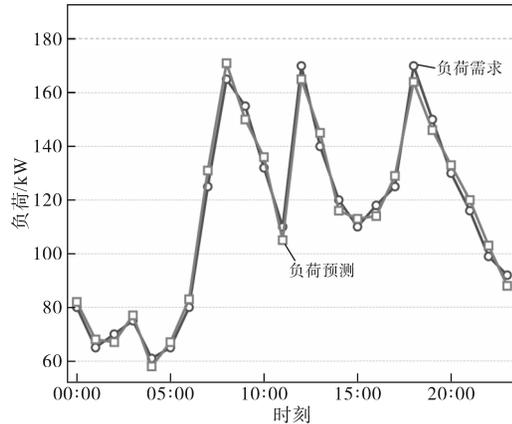


图4 负荷需求预测与实际值对比

Fig.4 Comparison between predicted and actual load demand

4.2 调度性能评估

本节围绕R-SAC策略在典型运行场景下的调度性能展开分析,重点从风电消纳率提升与购电成本控制两个优化目标出发,评估系统在不同时段的运行响应情况。

图5展示了风电出力及对应的售电与购电行为。从图中可见,风电功率在06:00—10:00与15:00—17:00出现两次上升波动,在07:00达到当日峰值。在风电富余阶段,系统优先本地使用风能,余电部分实现并网售电;售电量集中在风

电高峰期间,表明策略有效提升了风电利用率。在风电不足时段,系统仅在负荷高峰时有限购电,峰值购电功率控制在38 kW以内,表明在电价较高情况下系统成功抑制了对主网依赖,从而降低了购电成本。

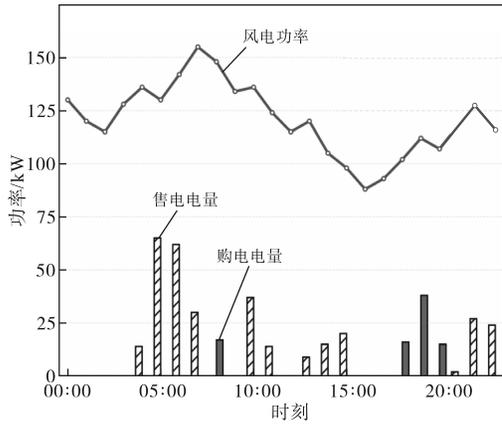


图5 风电出力与购售电功率时序图

Fig.5 Time series of wind power output and power trading with the main grid

图6综合展示了风电出力、负荷需求、储能SOC和TCLs调节行为。图中可见,用户负荷在08:00—11:00和18:00—21:00出现双峰结构,对应储能SOC明显下降,反映系统优先使用储能削减购电需求。在风电出力相对稳定的10:00—16:00,TCLs运行稳定,调节功率与其消耗值基本重合,表明其运行受控、舒适性保障良好。在系统供需失衡时段(如20:00后),TCLs协同调节参与负荷削峰,增强了微电网响应能力。

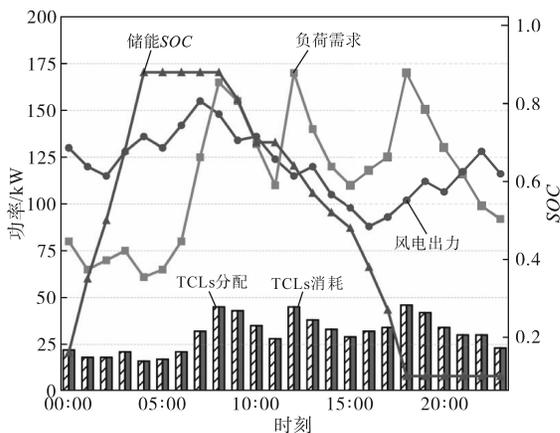


图6 微电网源-荷-储多维响应过程图

Fig.6 Multi-dimensional response process of source-load-storage in the microgrid

4.3 算法收敛对比分析

为评估各算法在训练过程中的收敛性与稳定性,本文将改进型算法R-SAC与基准算法SAC

和DDPG进行对比分析,如图7所示。所有算法均在相同的仿真环境与平台下运行,且训练过程采用一致的超参数配置:学习率设置为 $\eta = 10^{-3}$,批量大小 $B = 64$,经验回放池容量为10 000,折扣因子 $\gamma = 0.99$,软更新系数 $\tau = 0.005$,总训练回合数为3 000,优化器统一为Adam。

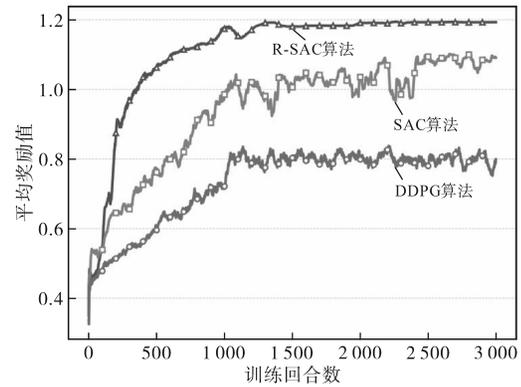


图7 不同算法训练收敛性对比

Fig.7 Comparison of training convergence among different algorithms

在网络结构方面,SAC与DDPG均使用包含两层、每层64单元的全连接神经网络。相比之下,改进算法R-SAC在此基础上引入类残差连接机制,通过在策略网络与价值网络中构建跳跃连接结构,增强了特征表达能力,并提升了训练过程的收敛速度与稳定性。

5 结论

本文基于深度强化学习框架,提出了一种融合状态估计与结构优化的风储TCLs系统联合调度策略。通过构建风电-储能-温控负荷协同模型,并引入改进型R-SAC算法,实现了对多源耦合系统的高效优化控制。基于实测数据的仿真实验表明,所提方法在调度性能与算法稳定性方面具有显著优势。具体结论如下:

1)基于PO-MDP建模框架,引入LSTM与注意力机制进行状态估计,增强了系统对环境不确定性的感知能力与策略鲁棒性。

2)在SAC算法结构中嵌入类残差网络结构以优化策略网络,提升了特征提取能力与训练稳定性。

3)在风电波动与电价峰谷显著场景下,所提策略可自适应协调储能与TCLs响应,实现风电优先利用与主网购电最小化,提升了经济性与可再生能源利用率。

参考文献

- [1] 冯斌,胡轶婕,黄刚,等. 基于深度强化学习的新型电力系统调度优化方法综述[J]. 电力系统自动化,2023,47(17):187-199.
FENG Bin, HU Yijie, HUANG Gang, et al. Review on optimization methods for new power system dispatch based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2023, 47(17): 187-199.
- [2] 周步祥,徐艺宾. 基于机器学习的源荷互动微电网优化调度[J]. 电力系统及其自动化学报,2022,34(2):144-150.
ZHOU Buxiang, XU Yibin. Optimal scheduling of source-load interactive micro-grid based on machine learning[J]. Proceedings of the CSU-EPSA, 2022, 34(2): 144-150.
- [3] MU C, SHI Y, XU N, et al. Multi-objective interval optimization dispatch of microgrid via deep reinforcement learning[J]. IEEE Transactions on Smart Grid, 2023, 12(2): 126-138.
- [4] QI J, LEI L, ZHENG K, et al. Optimal scheduling in IoT-driven smart isolated microgrids based on deep reinforcement learning [J]. IEEE Internet of Things Journal, 2023, 10(18): 16284-16299.
- [5] ZHANG Z, SHI J, YANG W, et al. Deep reinforcement learning based bi-layer optimal scheduling for microgrids considering flexible load control[J]. CSEE Journal of Power and Energy Systems, 2022, 9(3): 949-962.
- [6] ZHANG T, YUE D, YU L, et al. Joint energy and workload scheduling for fog-assisted multimicrogrid systems: a deep reinforcement learning approach[J]. IEEE Systems Journal, 2022, 17(1): 164-175.
- [7] ZHANG H, YUE D, DOU C, et al. PBI based multi-objective optimization via deep reinforcement elitelearning strategy for micro-grid dispatch with frequency dynamics[J]. IEEE Transactions on Power Systems, 2022, 38(1): 488-498.
- [8] LEI L, TAN Y, DAHLENBURG G, et al. Dynamic energy dispatch based on deep reinforcement learning in IoT-driven smart-isolated microgrids[J]. IEEE Internet of Things Journal, 2020, 8(10): 7938-7953.
- [9] YANG J, LIU J, XIANG Y, et al. Data-driven optimal dynamic-dispatch for hydro-PV-PHS integrated power systems using-deep reinforcement learning approach[J]. CSEE Journal of Power and Energy Systems, 2022, 9(3): 846-858.
- [10] LI X, YANG N, LI Z, et al. Confidence estimation transformer for long-term renewable energy forecasting in reinforcement learning-based power grid dispatching[J]. CSEE Journal of Power and Energy Systems, 2024, 10(4): 1502-1513.
- [11] 马临超,齐山成,刘毅,等. 考虑分布式可再生能源的配电网电力调度优化模型研究[J]. 可再生能源,2022,40(5):703-710.
MA Linchao, QI Shancheng, LIU Yi, et al. Research on distribution network power dispatching optimization model considering distributed renewable energy[J]. Renewable Energy Resources, 2022, 40(5): 703-710.
- [12] 张磊,吴红斌,何叶,等. 基于深度强化学习的氢能综合能源系统优化调度方法[J]. 电力系统自动化,2024,48(16):132-141.
ZHANG Lei, WU Hongbin, HE Ye, et al. Optimal scheduling method for integrated energy systems with hydrogen based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2024, 48(16): 132-141.

收稿日期:2025-04-24

修改稿日期:2025-06-08