

基于多智能体深度强化学习的大容量电池储能电站功率分配策略

彭寒梅¹, 赵长桥¹, 谭貌^{1,2}, 陈颢^{1,2}, 李辉^{1,2}

(1. 湘潭大学自动化与电子信息学院, 湖南湘潭411105; 2. 湖南省多能协同控制技术工程研究中心(湘潭大学), 湖南湘潭411105)

摘要: 大容量电池储能电站功率分配的决策变量多, 且策略需考虑多个优化目标及能自动适应场景的不确定性。为此, 提出了一种基于多智能体深度强化学习(multi-agent deep reinforcement learning, MADRL)的电池储能电站功率分配决策方法。首先, 基于大容量电池储能电站结构及其功率分配特性构建基于MADRL的功率分配决策框架, 每个储能单元设置一个功率分配智能体, 多个智能体构成合作关系; 然后, 设计考虑储能电站有功功率损耗、荷电状态(state of charge, SOC)一致性和健康状态损失最小优化目标的功率分配智能体模型, 采用深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法去中心化训练各智能体网络参数, 算法收敛后得到储能子系统充放电功率值。最后, 算例验证了所提方法的有效性, 能在有效提高储能子系统SOC均衡性的同时降低有功功率损耗、健康状态损失和充放电切换次数。

关键词: 电池储能电站; 功率分配; 多智能体; 深度强化学习; SOC一致性

Power Allocation Strategy for Large-Capacity Battery Energy Storage Power Station Based on Multi-Agent Deep Reinforcement Learning

PENG Hanmei¹, ZHAO Changqiao¹, TAN Mao^{1,2}, CHEN Jie^{1,2}, LI Hui^{1,2}

(1. College of Automation and Electronic Information, Xiangtan University, Xiangtan, Hunan 411105, China; 2. Hunan Engineering Research Center of Multi-Energy Cooperative Control Technology, Xiangtan University, Xiangtan, Hunan 411105, China)

Abstract: The decision variables for power allocation of large-capacity battery energy storage power station are numerous, and their strategies need to consider multiple optimization objectives and the uncertainty of automatically adapting to the scenario. Therefore, this paper proposes a power allocation decision-making method for battery energy storage power stations based on multi-agent deep reinforcement learning (MADRL). Firstly, based on the structure and power allocation characteristics of large-capacity battery energy storage power stations, a power allocation decision framework based on MADRL is constructed. Each energy storage unit is equipped with a power allocation intelligent agent, and multiple agents form a cooperative relationship. Then, a multi-agent DRL model for power allocation is designed, considering the optimization objectives of active power loss, state of charge (SOC) consistency, and state of health loss of battery energy storage power stations. The deep deterministic policy gradient (DDPG) algorithm is used to decentralize the training of network parameters for each agent. After the algorithm converges, the charging and discharging power values of the energy storage subsystem are obtained. Finally, the effectiveness of the proposed method is verified by the example, which can effectively improve the SOC balance of the energy storage subsystem while reducing active power loss, state of health loss, and charging and discharging switching times.

Key words: battery energy storage power station; power allocation; multi-agent; deep reinforcement learning; SOC consistency

基金项目: 国家自然科学基金资助项目(51777179); 湖南省自然科学基金资助项目(2023JJ50241)。

Foundation item: Supported by the National Natural Science Foundation of China (51777179); the Natural Science Foundation of Hunan Province (2023JJ50241).

0 引言

电化学储能电站是一种重要的电能存储形式,具有响应速度快、灵活可控、无地理环境约束等优势,可为电力系统提供强大的削峰填谷手段和惯量支撑手段^[1-3]。近年来,电化学储能电站及其技术得到了快速发展,建设数量不断攀升^[4-7]。大容量电化学储能电站一般采用电化学电池作为储能元件,通常包含多个电池系统。电池储能电站运行过程中,电池系统充放电功率会直接影响储能子系统的荷电状态(state of charge, SOC)一致性、容量损耗和储能电站的运行损耗。因此,储能子系统充放电功率的合理分配是保证电池储能电站安全稳定与高效运行的关键^[8-11]。

大容量电池储能电站运行中,储能子系统的SOC一致性差会导致电池的安全性和充放电效率降低^[12-13];同时,随着充放电次数不断增加,储能子系统可用容量衰减,健康状态(state of health, SOH)损失,易出现过充、过放和过电流等破坏电池内特性的现象,导致电池寿命缩短^[14-15]。大容量电池储能电站的功率分配具有以下特点:1)决策变量和运行约束多;2)需兼顾SOC均衡、系统运行损耗和SOH损失等多个优化目标;3)需应对参与功率分配的储能单元选择带来的不确定性。由此,大容量电池储能电站的功率分配决策方法需能兼顾多目标优化,并能自动适应系统的不确定性。

目前,电池储能系统的功率分配方法主要有:1)基于SOC占比的传统功率分配策略;2)基于优化目标方程及其求解的功率分配策略。传统功率分配策略下,调控后SOC均衡性差,且所有储能子系统均参与工作,难以考虑充放电对电池寿命和运行经济性的影响^[16-18]。基于优化目标方程及其求解的功率分配策略下,通常建立SOC均衡度最优、系统损耗最低等目标的功率分配优化方程,采用最优化算法求解,且先进行参与功率分配的储能子系统选择^[19-23]。文献[19]提出以总运行成本、容量损失和SOC一致性为优化目标的锂离子电池储能系统功率分配策略,采用金枪鱼群优化算法求解。文献[20]建立以SOC均衡度和系统功率损耗为目标的大容量钒电池储能系统功率优化分配模型,采用虚拟粒子-自适应差分进化算法求解。文献[21]采用多目标优化方法确定配电系统中储能单元的最佳充电、

放电时间,能有效降低运行成本并提升安全性。但是,参与功率分配的储能子系统选择及电网下达的需求功率指令均具有不确定性,而参考文献[19-22]的功率分配最优化方法,强调事先预置各种工况的应对方案,且难以较好地处理多个优化目标。由此,基于SOC占比及基于优化目标方程及其求解的功率分配策略不能同时满足大容量电池储能电站功率分配对多目标优化和不确定性自适应能力的要求。

深度强化学习(deep reinforcement learning, DRL)通过不断与环境交互,自动探索到最优策略,具有无模型学习、能灵活响应不确定性环境的特点^[24-25]。储能子系统SOC与上一时刻充放电功率有关,可将功率分配作为序贯决策,采用DRL进行求解。然而,大容量电池储能电站功率分配具有决策变量和运行约束多的特点,采用单一智能体DRL方法会导致因动作和状态空间维度高而收敛性较差。多智能体深度强化学习(multi-agent deep reinforcement learning, MADRL)通过各智能体间的有限信息交流,依据特定的奖惩机制能达到全局最优^[26-27]。目前已有研究将MADRL应用于能源管理及电力系统优化调度决策方面^[28-30],但较少用于储能电站及其调控领域。本文提出采用多智能体方法将多个储能单元的功率分配转换为多个智能体的决策,以降低智能体的动作维度和状态维度,提高在线训练速度与收敛性,并兼顾不确定性和多目标优化需求。

本文针对大容量电池储能电站功率分配的特性,提出一种基于MADRL的功率分配决策方法。首先,基于大容量电池储能的结构,分析其功率分配特性,构建基于MADRL的功率分配决策框架,每个储能单元设置一个功率分配智能体,多个智能体构成合作关系。然后,设计考虑储能电站运行损耗、SOC一致性和SOH损失最小优化目标的功率分配多智能体DRL模型,再采用深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法去中心化训练各智能体网络参数,算法收敛后得到充放电功率分配策略。最后,通过大容量电池储能电站算例进行验证。

1 电池储能电站的结构

大容量电池储能电站的结构及其充电能量流、

信息流示意图如图 1 所示。1) 包含 N 个储能单元, 储能单元经变压器并入 10 kV 总馈线, 再与电网并联。每个储能单元包含 J 个储能子系统, 储能子系统包括电池系统、储能变流器 (power conversion system, PCS) 和电池管理系统 (battery management system, BMS), 其中电池系统由大量单体电池串并联而成, 负责储存电能和释放电能; PCS 负责电池系统的直流电与交流电的双向转换。电池系统安装在电池舱, PCS 和变压器安装在逆变变压一体舱, 为一次系统的核心设备。2) 二次系统主要包括 BMS 和能量管理系统 (energy management system, EMS), EMS 分为总控 EMS 和单元 EMS。总控 EMS 负责对整个储能电站的能量流动进行宏观调控, 单元 EMS 对各个储能单元进行局部管理。各单元之间通过通信线路进行实时信息传递, 确保系统能够按照预定的控制策略高效、稳定地运行。

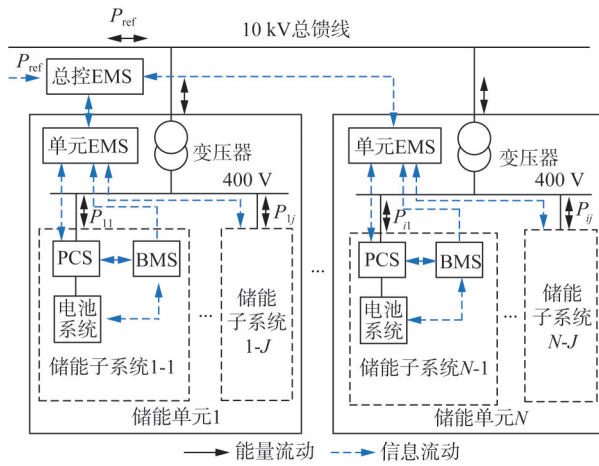


图 1 大容量电池储能电站结构示意图

Fig. 1 Structure schematic diagram of the large-capacity battery energy storage power station

2 多智能体 DRL 功率分配决策框架

2.1 功率分配分析

图 1 中大容量电池储能电站的运行调控流程为: 总控 EMS 接收电网调度中心下发的需求功率指令 P_{ref} , 进行全局调度决策, 并将决策结果下发给单元 EMS; 单元 EMS 根据总控 EMS 下发的信息、各个储能子系统 BMS 传递过来的 SOC、SOH 等信息, 进行功率分配决策, 决策出各个储能子系统的充放电功率 P_{ij} , 并下发给 PCS; 由 PCS 控制电池系统中的电池充放电。

为提高延长电池系统使用寿命, 总控 EMS 接收到需求功率指令后, 需进行参与充放电储能单元的选择^[31]。通常采用基于 SOC 值的储能单元参与的选择方法^[32]: 对储能单元 SOC 均值进行排序, SOC 均值较低者充电优先级更高, SOC 均值较高者则放电优先级更高; 再按需求功率指令大小判断并选定调度时段 t 下参与功率分配的储能单元组合, 应满足:

$$\sum_{i=1}^{M_t} \sum_{j=1}^J |P_{\text{Bmij}}| - |P_{\text{ref},t}| \geq 0 \quad (1)$$

式中: $P_{\text{ref},t}$ 为时段 t 下发的需求功率指令; P_{Bmij} 为参与本次功率分配的储能单元 i 第 j 个储能子系统的最大充放电功率; M_t 为时段 t 参与功率分配的储能单元数量; J 为每个储能单元包含的储能子系统的数量。

选定参与充放电的储能单元后, 需对选定的储能单元进行功率分配。根据大容量电池储能电站及电池系统的运行特性, 功率分配策略需兼顾储能电站运行损耗最低、SOC 一致性最好和 SOH 损失最小等多个优化目标。设置功率分配的目标函数为:

$$\min \sum_{t=1}^T (P_{\text{loss},t} + \Delta S_{\text{SOC},t} + \Delta S_{\text{SOH},t}) \quad (2)$$

式中: $P_{\text{loss},t}$ 为时段 t 储能电站有功功率损耗; $\Delta S_{\text{SOC},t}$ 为时段 t 储能子系统的 SOC 一致性; $\Delta S_{\text{SOH},t}$ 为时段 t 储能子系统的 SOH 损失; T 为调度时段数量。

储能电站有功功率损耗包括电池系统、PCS 和变压器的有功功率损耗, 为:

$$\begin{cases} P_{\text{loss},t} = \sum_{i=1}^{M_t} \sum_{j=1}^J (P_{\text{BLij},t} + P_{\text{PCSLij},t}) + \sum_{i=1}^{M_t} P_{\text{TLi},t} \\ P_{\text{BLij},t} = n(cR_{\text{os}} + R_{\text{ps}}\alpha_{\text{con}}\alpha_{\text{SOC}}\alpha_{\text{T}})(P_{ij,t}/U_{\text{B}})^2 + ncP_{\text{BL}} \\ P_{\text{PCSLij},t} = (1 - \eta_{\text{PCS}})|P_{ij,t}| \\ P_{\text{TLi},t} = P_{\text{T0}} + P_{\text{TK}} \left(\frac{\sqrt{P_{i,t}^2 + Q_{i,t}^2}}{C_{\text{TN}}} \right)^2 \\ P_{i,t} = \sum_{j=1}^J (|P_{ij,t}| + P_{\text{PCSLij},t}) \end{cases} \quad (3)$$

式中: $P_{ij,t}$ 为时段 t 储能单元 i 的第 j 个储能子系统的充放电功率, 即功率分配的决策变量, 设定 $P_{ij,t} > 0$ 表示放电, $P_{ij,t} < 0$ 表示充电; $P_{\text{BLij},t}$ 、 $P_{\text{PCSLij},t}$ 分别为时段 t 储能单元 i 的第 j 个储能子系统中电池系统、PCS 的有功功率损耗; $P_{\text{TLi},t}$ 、 $P_{i,t}$ 、 $Q_{i,t}$ 分别为时段 t

储能单元 i 的变压器有功功率损耗、低压侧有功功率和无功率; n 、 c 分别为电池系统的并联电池串数量及每个电池串中单体电池的数量; U_B 为电池串两端电压; P_{BL} 为单体电池电流为 0 时的损耗; R_{os} 、 R_{ps} 分别为单体电池的欧姆内阻和极化内阻; α_{con} 、 α_{SOC} 、 α_T 分别为电池串构成修正系数、电池 SOC 状态修正系数和电池工作状态修正系数; η_{PCS} 为 PCS 转换效率; P_{T0} 、 P_{Tk} 、 C_{TN} 分别为变压器运行铁损、额定铜损和额定容量。

储能子系统的 SOC 一致性用储能子系统的剩余荷电状态方差之和表示, 为:

$$\begin{cases} \Delta S_{SOC, t} = \sum_{i=1}^{M_i} \sum_{j=1}^J (S_{SOC, ij, t} - \frac{\sum_{i=1}^{M_i} \sum_{j=1}^J S_{SOC, ij, t}}{MJ})^2 / MJ \\ S_{SOC, ij, t} = S_{SOC, ij, t-1} - \frac{\eta_{ij, t} P_{ij, t} \Delta t}{S_{SOH, ij, t} C_{BN}} \\ \eta_{ij, t} = 1 - \frac{P_{BL, ij, t}}{P_{ij, t}}, P_{ij, t} < 0; \eta_{ij, t} = \frac{P_{ij, t}}{P_{ij, t} - P_{BL, ij, t}}, P_{ij, t} > 0 \end{cases} \quad (4)$$

式中: $S_{SOC, ij, t}$ 、 $S_{SOC, ij, t-1}$ 分别为时段 t 、 $t-1$ 下储能单元 i 的第 j 个储能子系统的 SOC 值; $S_{SOH, ij, t}$ 为时段 t 储能单元 i 的第 j 个储能子系统的 SOH 值; $\eta_{ij, t}$ 为储能子系统的转换效率; Δt 为调度时段的时间间隔; C_{BN} 为储能子系统的额定容量。

时段 t 储能子系统的 SOH 损失表示为:

$$\begin{cases} \Delta S_{SOH, t} = \sum_{i=1}^{M_i} \sum_{j=1}^J \frac{L_{ij, t}}{C_{BN}} \times 100 \% \\ L_{ij, t} = Ae^{\frac{-31700 + 370.3 P_{ij, t} / P_{BN}}{HT_e}} (S_{SOC, ij, t-1} - S_{SOC, ij, t})^{0.55} \times (\frac{P_{ij, t}}{P_{BN}})^2 \Delta t \end{cases} \quad (5)$$

式中: $L_{ij, t}$ 为时段 t 储能单元 i 的第 j 个储能子系统的容量损耗; P_{BN} 为储能子系统的额定功率; A 为常数; H 为气体常数; T_e 为温度值。

式(2)目标函数的约束为:

$$\begin{cases} P_{ref, t} = \sum_{i=1}^{M_i} P_{i, t} \\ 0 < P_{i, t} \leq P_{TN} \\ (-P_{BN}) \leq P_{ij, t} \leq P_{BN}, S_{SOC, min} \leq S_{SOC, ij, t} \leq S_{SOC, max} \end{cases} \quad (6)$$

式中: P_{TN} 为变压器的额定功率; $S_{SOC, min}$ 、 $S_{SOC, max}$ 分别为储能子系统允许的 SOC 最小值与最大值。

由式(1)~(6)可知: 不同调度时段 t 参与功率

分配的储能单元个数 M_t 及具体的储能单元变化, 具有较强的不确定性, 使得式(2)~(6)求解困难; 大容量电池储能电站含多个储能子系统, 使得功率分配的决策变量(即时段 t 储能单元 i 的第 j 个储能子系统的充放电功率 $P_{ij, t}$)及约束多; 功率分配策略需均衡有功功率损耗、储能子系统 SOC 不一致性、储能子系统 SOH 损失最小 3 个优化目标。这对功率分配调度的决策精度和决策速度提出了挑战。为此, 本文提出采用 DRL 在线决策功率分配优化调度, 并采用多智能体方法, 将多个储能单元的储能子系统充放电功率决策转换为多个智能体的决策。

2.2 功率分配决策框架

本文提出基于 MADRL 的功率分配决策框架如图 2 所示, DRL 难以通过与大容量电池储能电站物理系统交互进行优化决策, 需与系统模型进行交互。

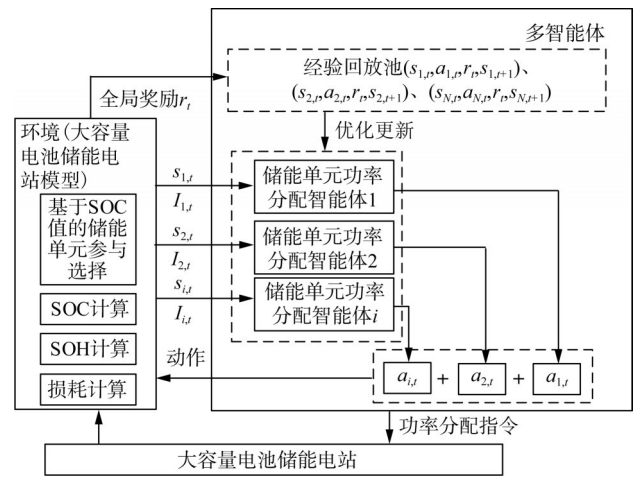


图2 基于 MADRL 的电池储能电站功率分配决策框架

Fig. 2 Power allocation decision framework for battery energy storage power stations based on MADRL

1) 每个储能单元 EMS 中设置一个功率分配智能体, 储能单元功率分配智能体 1、储能单元功率分配智能体 2、储能单元功率分配智能体 i , $i=1, 2, \dots, N$; 与智能体交互的环境为包括基于 SOC 值的储能单元参与选择、SOC 值计算、SOH 计算及损耗计算等的大容量电池储能电站模型。为兼容储能单元参与功率分配的不确定性, 用 $I_{i, t}$ 表示储能单元 i 在第 t 个调度时段的参与标志, 其值取为 1、0 时分别表示参与功率分配和不参与功率分配。 $I_{i, t}$ 值由基于 SOC 值的储能单元参与选择得到。

2) 各储能单元功率分配智能体为合作关系, 各

智能体共一个环境，环境每调度时段反馈给各智能体相同的全局奖励 r_t 。第 t 个时段，储能单元功率分配智能体 i 感知所在储能单元的参与标志 $I_{i,t}$ 及状态 $s_{i,t}$ ；当 $I_{i,t}=0$ 时，通过设计特殊值机制使该智能体“禁用”，当 $I_{i,t}=1$ 时，根据状态 $s_{i,t}$ 利用策略网络输出功率分配决策 $a_{i,t}$ ；环境执行所有智能体功率分配决策，输出下一个状态 $s_{i,t+1}$ 及全局奖励 r_t ，并进入下一个调度时段的决策。各储能单元功率分配智能体由各自的经验回放池($s_{i,t}$, $a_{i,t}$, r_t , $s_{i,t+1}$)进行优化更新。

3) MADRL 采用“在线学习”模式，MADRL 与大容量电池储能电站模型不断交互进行学习和更新策略，继而演进到最优功率分配策略。MADRL 算法收敛后，MADRL 输出动作即功率分配指令发给实际大容量电池储能电站，支配其充放电运行，否则，继续学习。

3 功率分配多智能体 DRL 模型设计

3.1 DRL 模型设计

本文提出的基于 MADRL 的大容量电池储能电站功率分配决策框架中含多个储能单元功率分配智能体，其 DRL 模型设计如下。

1) 动作空间设计

动作空间为功率分配的相关决策变量。储能单元功率分配智能体 i 的动作空间为其所在储能单元内的储能子系统的充放电功率，表示为：

$$a_{i,t} = \{P_{ij,t}, i = 1, 2, \dots, J\} \quad (7)$$

2) 状态空间设计

状态空间是智能体所感知的环境信息。时段 t 储能单元功率分配智能体 i 的状态空间包括：当前所处的调度时段 t ，该时段本储能单元的储能子系统的 SOC 值、SOH 值及上一时段的 SOC 值，表示为：

$$s_{i,t} = \{t, S_{\text{SOC } ij,t}, S_{\text{SOH } ij,t}, S_{\text{SOC } ij,t-1}, j = 1, 2, \dots, J\} \quad (8)$$

3) 奖励函数设计

全局奖励 r_t 包括目标奖励和约束条件奖励。由式(2)可知，目标奖励包括储能电站有功功率损耗奖励、储能子系统的 SOC 一致性奖励和储能子系统的 SOH 损失奖励，表示为：

$$r_{\text{C1},t} = -P_{\text{loss},t}, r_{\text{C2},t} = -\Delta S_{\text{SOC},t}, r_{\text{C3},t} = -\Delta S_{\text{SOH},t} \quad (9)$$

式中 $r_{\text{C1},t}$ 、 $r_{\text{C2},t}$ 、 $r_{\text{C3},t}$ 分别为储能电站有功功率损耗

奖励、储能子系统的 SOC 一致性奖励和储能子系统的 SOH 损失奖励。

由式(3)–(5)计算得到式(9)中的 3 个目标奖励。

将式(6)的 1 个平衡约束、3 个不等式约束形成的约束奖励分别表示为 $r_{1,t}$ 、 $r_{2,t}$ 、 $r_{3ij,t}$ 、 $r_{4ij,t}$ ，当不满足约束时对应的奖励取值设为-1，否则取值为 0。时段 t 的约束条件奖励 $r_{\text{Rall},t}$ 为：

$$r_{\text{Rall},t} = -(r_{1,t} + \sum_{i=1}^N r_{2i,t} + \sum_{i=1}^N \sum_{j=1}^J (r_{3ij,t} + r_{4ij,t})) \quad (10)$$

式中 N 为储能单元数量。

形成时段 t 的立即全局奖励 r_t 为：

$$r_t = c_1 r_{\text{C1},t} + c_2 r_{\text{C2},t} + c_3 r_{\text{C3},t} + c_4 r_{\text{Rall},t} \quad (11)$$

式中 c_1 、 c_2 、 c_3 和 c_4 为各项奖励权重系数。

3.2 DDPG 算法

由基于 MADRL 的大容量电池储能电站功率分配框架及设计的多智能体 DRL 模型可知，各储能单元功率分配智能体为合作关系，共享一个环境，所有智能体的奖励都是相同的。进一步地，储能单元功率分配多智能体 DRL 采用“去中心化训练+去中心化决策”方式，并基于 DDPG 算法在智能体本地训练各智能体网络参数。

DDPG 算法是一种采用 actor-critic 框架的 DRL 算法，融合神经网络计算与确定性策略梯度的连续优化特性，在连续动作空间任务中表现出优良的决策性能^[32]。本文采用 DDPG 算法对储能单元功率分配智能体进行去中心化训练，每个智能体上部署 actor 网络和 critic 网络。DDPG 算法包含 4 个深度神经网络：分别使用 actor 网络和 critic 网络来近似策略及状态-动作价值函数，同时还有对应结构的目标 actor 和目标 critic 网络。目标网络的作用是更新 actor 和 critic 网络的参数，以改善算法的收敛性。critic 网络则通过梯度下降来更新其参数^[33-34]。训练优化目标为损失函数式(12)最小。

$$\begin{cases} L_Q = \frac{1}{K} \sum_k (y_k - Q(s_k, a_k | \theta_Q))^2 \\ y_k = r_k + \gamma Q'(s_{k+1}, \mu'(s_{k+1} | \theta_{\mu'}) | \theta_{Q'}) \end{cases} \quad (12)$$

式中： L_Q 为损失函数； $Q(s_k, a_k | \theta_Q)$ 、 θ_Q 分别为 critic 网络及其参数； K 、 k 分别为从经验回放池中抽取的样本数量和第 k 个样本； s_k 、 a_k 分别为第 k 个样本的状态和动作； $Q'(s_{k+1}, \mu'(s_{k+1} | \theta_{\mu'}) | \theta_{Q'})$ 、 $\theta_{Q'}$

分别为目标 critic 网络及其参数; $\mu'(s_{k+1}|\theta_{\mu'})$ 、 $\theta_{\mu'}$ 分别为目标 actor 网络及其参数; r_k 、 y_k 分别为对应的立即奖励和实际 Q 值; γ 为折扣因子。

actor 网络采用梯度更新网络参数, 最大化当前状态-动作价值如式(13)所示。

$$\max L_{\mu} = Q(s_k, a_k | \theta_Q) \Big|_{a_k = \mu(s_k | \theta_{\mu})} \quad (13)$$

式中: L_{μ} 为当前状态-动作价值函数; $\mu(s_k | \theta_{\mu})$ 、 θ_{μ} 分别为 actor 网络及其参数。

采用式(14)更新目标网络参数, 为:

$$\begin{cases} \tau\theta_Q + (1-\tau)\theta_{Q'} \rightarrow \theta_Q \\ \tau\theta_{\mu} + (1-\tau)\theta_{\mu'} \rightarrow \theta_{\mu'} \end{cases} \quad (14)$$

式中 τ 为目标网络学习率, $0 < \tau < 1$ 。

3.3 训练流程

采用 DDPG 算法训练储能单元功率分配智能体, 其流程如图3所示, 训练收敛后各智能体输出动作得到功率分配策略, 即储能子系统的充放电功率值, 主要步骤如下。

1) 初始化各储能单元功率分配智能体的网络参数, 初始化各自的经验回放池, 初始化回合数 $m=0$ 。

2) 各储能单元功率分配智能体从实际的大容量电池储能电站中获取式(8)所示的本储能单元初始状态。

3) 大容量电池储能电站模型根据电网调度中心下发的需求功率指令 $P_{ref,t}$ 及储能单元状态得到储能单元 i 在第 t 个调度时段的参与标志 $I_{i,t}$ 。当储能单元功率分配智能体感知 $I_{i,t}=0$ 时, 其禁用, 即不输出; 当感知 $I_{i,t}=1$ 时, 智能体根据当前状态 $s_{i,t}$ 输出动作 $a_{i,t}$ 。

4) 大容量电池储能电站模型执行动作 $a_{1,t}$ 、 $a_{2,t}$, $a_{i,t}$, 输出给各智能体共同的立即全局奖励 r_t 以及下一状态 $s_{1,t+1}$ 、 $s_{2,t+1}$ 、 $s_{i,t+1}$; 将新样本 $(s_{1,t}, a_{1,t}, r_t, s_{1,t+1})$ 、 $(s_{2,t}, a_{2,t}, r_t, s_{2,t+1})$ 、 $(s_{i,t}, a_{i,t}, r_t, s_{i,t+1})$ 分别存入各自的经验回放池; $t=t+1$ 。

5) 各智能体基于式(12)~(13)分别更新各自的 actor 网络和 critic 网络参数, 由式(14)更新目标网络参数。

6) 每与大容量电池储能电站模型交互 T 次, 所有智能体进行 1 次网络参数的学习; 每训练网络 1 次, 进行 1 次收敛条件判断: 同时达到以下两个条件时, 算法收敛。条件 1: 计算各智能体在 1— T 时

段内立即全局奖励的总和, 即累积全局奖励, 判断累积全局奖励在连续 M 次学习中的标准差低于设定的阈值 δ ; 条件 2: 连续 M 个回合所有智能体均不违背任何约束, 即约束条件奖励为 0。

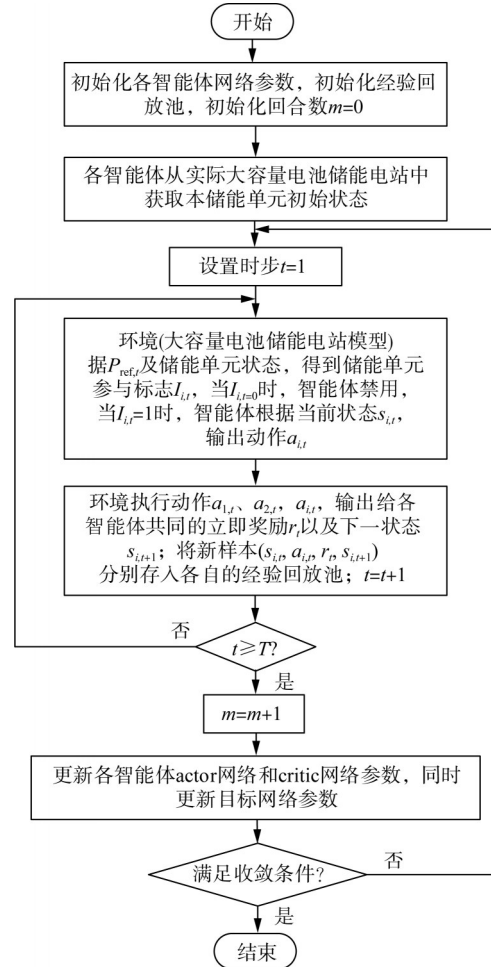


图3 储能单元功率分配智能体训练流程

Fig. 3 Flow chart of energy storage unit power allocation agent training

4 算例分析

4.1 算例系统

基于 i7-8550 U CPU、16GB 内存的笔记本电脑硬件环境, 及基于 Python3.7、Pytorch1.12.0 及 Pycharm 的软件环境下, 编程实现本文提出的基于 MADRL 的大容量电池储能电站功率分配策略。设置大容量电池储能电站算例: 额定容量为 10 MW、20 MWh, 含 4 个储能单元, 每个储能单元中含 4 个储能子系统, 其参数见附表 A1 所示; 设置 PCS 转换效率为 0.96, 储能子系统允许的 SOC 最小值、

最大值分别为 10 % 和 90 %，SOC 初值如表 1 所示，SOH 初值均为 100 %。

表 1 储能子系统的 SOC 初值

Tab. 1 Initial SOC values of energy storage subsystem

储能单元 编号	储能子系 统编号	SOC 初 值/%	储能单元 编号	储能子系 统编号	SOC 初 值/%
储能单元 1	1-1	45	储能单元 3	3-1	47
	1-2	38		3-2	57
	1-3	51		3-3	51
	1-4	54		3-4	62
储能单元 2	2-1	60	储能单元 4	4-1	51
	2-2	55		4-2	43
	2-3	69		4-3	61
	2-4	45		4-4	45

给定的电网调度中心下发的需求功率指令 $P_{\text{ref},t}$ 如图 4 所示，调度时段的时间间隔 $\Delta t=15 \text{ min}$ ，共 48 个调度时段。算例的功率分配 MADRL 模型中，含 4 个储能单元功率分配智能体，4 个智能体的动作空间维度数均为 4，状态空间维度数均为 12，累积全局奖励中的 T 取 4。DDPG 算法采用“在线学习”模式，参数为：3 层全连接的神经网络， γ 取 0.9，actor、critic 网络的学习率取 0.000 5，目标网络学习率 $\tau=0.005$ ，激活函数为 tanh，优化器为 Adam，经验回放池采用先进先出方式，大小取 4 800， K 取 60，噪声初始值为 0.35，噪声标准差以线性方式逐渐减少，最小值为 0.01。 M 取 50， δ 取 0.05。

4.2 训练过程

算例在下发的需求功率指令 $P_{\text{ref},t}$ 下采用本文提出的基于 MADRL 的功率分配方法，进行储能子系

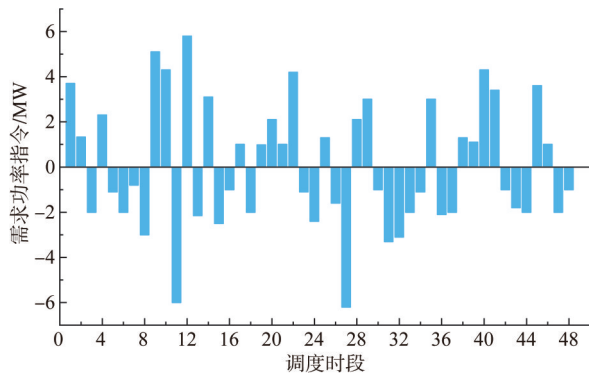


图 4 下发的需求功率指令

Fig. 4 Power demand instruction issued

统功率分配决策。其中，得到的调度时段 1~4 下多智能体训练收敛曲线如图 5 所示。

由图 5 可知：1) 训练初期，各智能体随机产生动作探索与环境交互获得经验，累积全局奖励波动较大，随着训练回合的增加，逐渐学习到最优功率分配策略，累积全局奖励在 400 回合之后收敛至较高奖励值，奖励值为 -202.320 7，在线学习时间约为 6.3 min。2) 经过多回合的学习训练，各智能体获得的累积全局奖励收敛，标准差不超过设定的阈值 δ ，且约束条件惩罚均为 0，形成了稳定的功率分配策略。

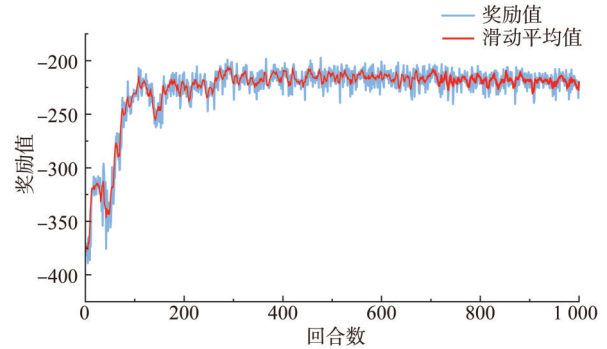
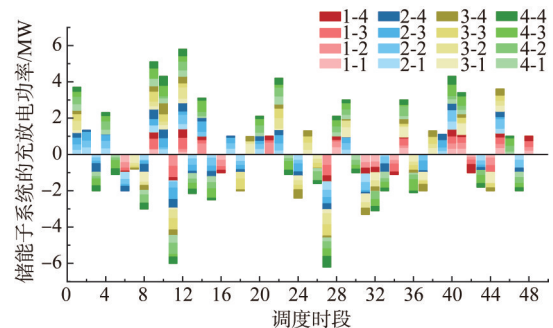


图 5 多智能体训练收敛曲线

Fig. 5 Convergence curves of multi-agent training

4.3 功率分配结果

多智能体通过 DDPG 算法训练并达到收敛后，学习到最优的功率分配策略。得到算例在设置的 48 个调度时段及储能子系统的 SOC 初值下的功率分配决策结果，即储能子系统的充放电功率值如图 6 所示，及功率分配优化目标结果如图 7 所示，得到 $P_{\text{ref},t}=5.1 \text{ MW}$ 、 4.3 MW 、 -6 MW 、 5.8 MW 对应调度时段 9~12 的功率分配结果如表 2 所示。



注：图例编号与表 1 中储能子系统编号对应。

图 6 功率分配决策结果

Fig. 6 Power allocation decision results

结果分析如下。

1)由图6可知,当电网调度中心下发的需求功率为正值时,储能电站放电,为负值时,储能电站充电;需求功率大小不同时,由大容量电池储能电站模型中的基于SOC值的储能单元参与选择得到参与功率分配的储能单元个数不同,使得某些调度时段下储能单元功率分配智能体通过感知 $I_{i,t}=0$,智能体禁用,即无动作输出,进而在满足电网调度中心下发的需求功率及储能电站运行条件下,MADRL功率分配策略能自动适应参与功率分配的储能单元的变化。

2)由表2可知,调度时段10下 $P_{\text{ref},t}$ 为4.3 MW时,储能单元1未参加功率分配。这是由于:调度时段9决策后储能单元1的SOC均值最小,根据基于SOC值的储能单元参与选择方法,调度时段10下储能单元2、3、4放电。调度时段9下 $P_{\text{ref},t}$ 为5.1 MW时,4个储能单元均参加功率分配,但储能单元4的储能子系统4-2的分配功率决策值为0。这是由于:储能子系统4-2的SOC值为储能单元4内的最低值,经过DRL优化决策后其分配功率为0,进而保障SOC均衡性的优化效果。

3)由图7可知,48个调度时段的储能电站有功功率损耗之和为3.182 MW,储能子系统的SOC一

致性由初始的68.250降低至5.943,储能子系统的SOH损失之和为 $109.328 \times 10^{-2}\%$ 。

上述功率分配结果及分析,验证了本文所提基于MADRL的大容量电池储能电站功率分配策略的有效性。

4.4 不同策略下的功率分配结果对比

设置3种不同的储能系统功率分配策略。策略1,基于SOC占比的传统功率分配策略^[16-18]:电池储能电站的储能单元均参与功率分配,储能子系统功率分配值按其当前SOC大小进行分配,即式(15)。策略2,基于优化目标方程及其求解的功率分配策略:首先进行储能单元参与选择,再建立并求解功率分配优化目标函数方程^[19-22];设置的功率分配优化目标函数及约束与本文相同,采用传统线性规划方法求解。策略3:本文所提功率分配策略。

$$\begin{cases} P_{ij,t} = \frac{1 - S_{\text{SOC } ij,t}}{\sum_{i=1}^{M_t} \sum_{j=1}^J S_{\text{SOC } ij,t}} P_{\text{ref},t}, P_{\text{ref},t} < 0 \\ P_{ij,t} = \frac{S_{\text{SOC } ij,t}}{\sum_{i=1}^{M_t} \sum_{j=1}^J S_{\text{SOC } ij,t}} P_{\text{ref},t}, P_{\text{ref},t} > 0 \end{cases} \quad (15)$$

采用3种不同策略对算例进行功率分配对比分析,得到48个调度时段的储能电站运行结果对比

表2 调度时段9—12的功率分配结果

Tab. 2 Power allocation results during scheduling period 9—12

储能子系统编号	功率分配结果/MW、决策后的SOC值/%			
	调度时段9	调度时段10	调度时段11	调度时段12
1-1	0.313, 48.05	0.000, 48.05	-0.065, 49.41	0.181, 45.70
1-2	0.019, 43.59	0.000, 43.59	-0.599, 55.88	0.157, 52.67
1-3	0.572, 42.62	0.000, 42.62	-0.622, 55.36	0.619, 42.67
1-4	0.343, 48.51	0.000, 48.51	-0.183, 52.25	0.465, 42.75
2-1	0.389, 48.62	0.319, 42.11	-0.065, 43.48	0.292, 37.52
2-2	0.002, 43.31	0.109, 41.06	-0.332, 47.83	0.508, 37.45
2-3	0.345, 46.42	0.506, 36.08	-0.625, 48.89	0.181, 45.20
2-4	0.524, 50.90	0.480, 41.08	-0.467, 50.63	0.457, 41.29
3-1	0.189, 61.68	0.241, 56.75	-0.110, 59.02	0.244, 54.04
3-2	0.250, 47.63	0.196, 43.63	-0.479, 53.42	0.625, 40.60
3-3	0.552, 34.83	0.397, 26.72	-0.625, 39.53	0.222, 35.00
3-4	0.296, 58.37	0.598, 46.12	-0.297, 52.18	0.367, 44.69
4-1	0.409, 43.90	0.276, 38.27	-0.432, 47.09	0.349, 39.97
4-2	0.000, 40.57	0.120, 38.10	-0.619, 50.79	0.486, 40.86
4-3	0.574, 54.47	0.570, 42.79	-0.140, 45.67	0.294, 39.66
4-4	0.323, 48.82	0.489, 38.82	-0.340, 45.77	0.353, 38.57

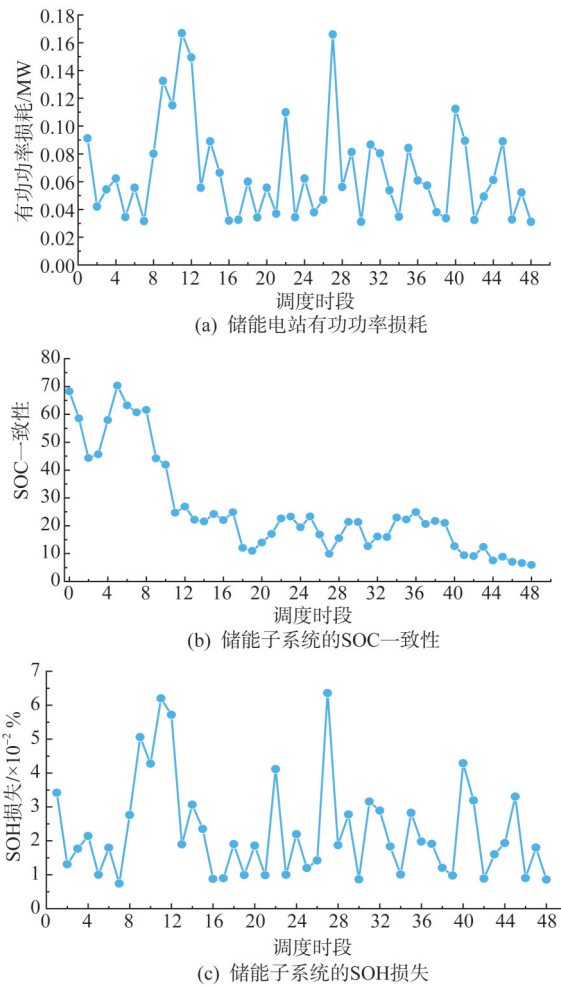


图7 功率分配优化目标结果

Fig. 7 Optimization target results of power allocation

(有功功率损耗之和、储能子系统的SOC一致性始末差, 储能子系统的SOH损失之和)如表3所示。SOC初值大小第1、2、3和第14、15、16的储能子系统, 在调度时段1—48下的SOC值变化如图8所示; 48个调度时段的储能子系统充放电切换次数分别为345次、263次、239次。

对比分析如下。

1)由表3可知, 相较于策略1、2, 本文策略下

表3 不同功率分配策略下的储能电站运行结果对比

Tab. 3 Comparison of operating results of the energy storage power station under different power allocation strategies

方法	有功功率损耗之和/MW	SOC一致性始末差	SOH的损失之和/%
策略1	3.716	48.075	1.312 37
策略2	3.417	51.276	1.177 51
策略3	3.182	62.307	1.093 28

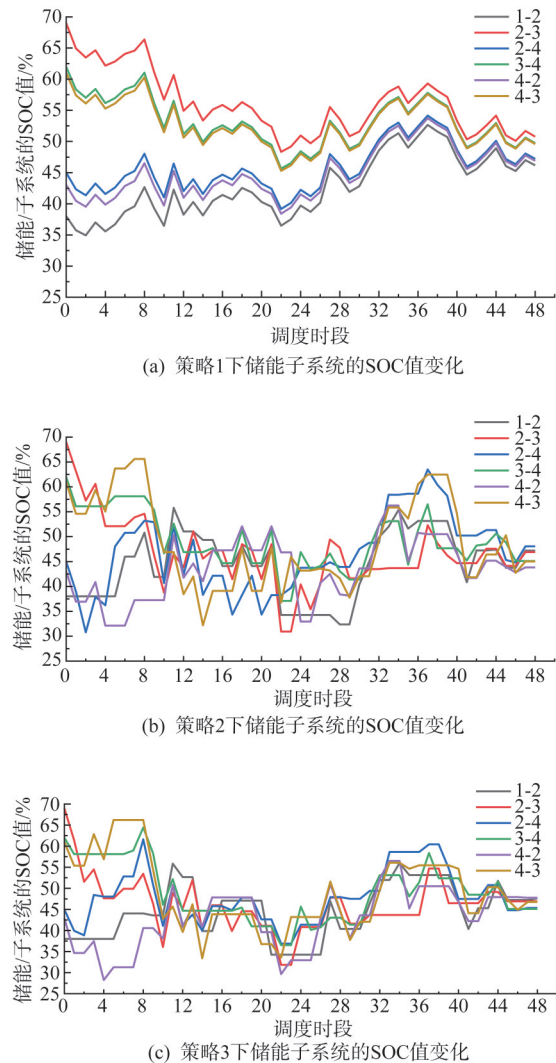


图8 不同策略下储能子系统的SOC值变化

Fig. 8 SOC variation of energy storage subsystems under different strategies

得到48个调度时段的储能电站有功功率损耗之和分别降低了14.370%和6.877%、储能子系统的SOC一致性始末差分别增加了29.604%和21.513%, 储能子系统的SOH损失之和分别减少了16.694%和7.153%; 此外, 储能子系统充放电切换次数分别减少了44.352%和9.125%。由图8可知, 本文策略下储能子系统SOC均衡效果最好。这是由于较传统策略、基于优化目标方程及其求解的功率分配策略, 本文策略采用MADRL方法中各储能单元功率分配智能体为合作关系, 协同优化包括储能电站有功功率损耗、储能子系统的SOC一致性和储能子系统的SOH损失最小的目标奖励, 进而能有效地利用功率分配策略优化, 实现大容量电

池储能电站储能子系统间的SOC均衡,并同时降低有功功率损耗、SOH损失和储能子系统充放电切换次数。

2)对比图8(a)与图8(c),本文策略下,当电网调度中心下发的需求功率变化时,有储能子系统的SOC值保持不变,这是因为DRL能自动适应需求功率变化带来的参与功率分配的储能单元变化,不参与功率分配的储能单元的储能子系统SOC值保持不变;而传统策略下,每个调度时段所有储能单元均参与功率分配,使得储能单元的储能子系统SOC值一直变化。

4.5 应对不确定性的决策结果

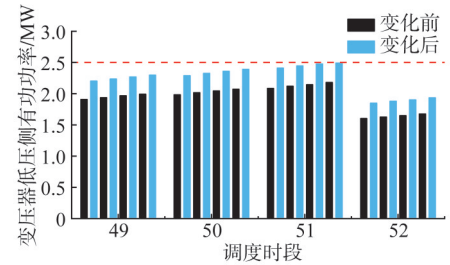
设置算例中PCS转换效率由于使用时间过长等因素降低为0.8,其他参数不变。给定电网调度中心下发的需求功率指令 $P_{ref,t}$ 为7.5 MW、7.8 MW、8.2 MW、6.3 MW,对应调度时段49—52,采用本文策略与传统策略对PCS转换效率变化前后进行功率分配决策,得到应对不确定性的变压器低压侧有功功率变化对比如图9所示。

由图9可知,当PCS转换效率降低,传统策略得到的储能单元2、3的变压器低压侧有功功率存在越限,本文策略下均未越限。这是由于:PCS转换效率发生变化,但是预置的传统的储能系统功率分配策略模型未随着转换效率的改变进行修正,从而使得当实际电池储能电站的PCS转换效率降低时其损耗增加,变压器低压侧有功功率增加,出现高于其额定功率的时段;本文“在线学习”模式的DDPG功率分配策略与映照实际情况的电池储能电站模型继续交互,能自动完成优化功率分配策略的调整,使变压器低压侧有功功率一直保持在合理运行区,验证了DDPG在线决策方法在应对电池储能电站设备不确定性上的有效性,较传统策略能自动应对电池储能电站设备的不确定性。

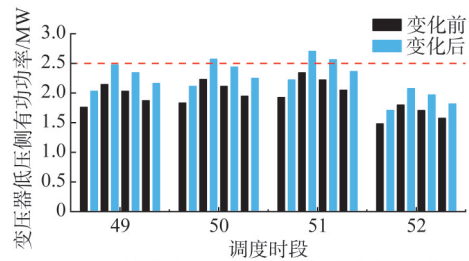
5 结语

本文针对大容量电池储能电站功率分配的决策变量多、需考虑多个优化目标,且其策略需能自动适应场景的不确定性,提出了一种基于MADRL的电池储能电站功率分配决策方法。算例结果及分析结果如下。

1)提出了基于MADRL的功率分配决策方法,每个储能单元设置一个功率分配智能体,且多个智



(a) 本文策略下变压器低压侧有功功率变化



(b) 传统策略下变压器低压侧有功功率变化

图9 应对不确定性的变压器低压侧有功功率变化对比

Fig. 9 Comparison of active power changes on the transformer low-voltage side in response to uncertainty

能体为合作关系。该方法可解决大容量电池储能电站功率分配场景包含众多变量及参与分配的储能单元不确定性带来的问题,兼顾了多个优化目标,降低了DRL动作和状态空间维度,且在有效提高储能子系统SOC均衡性的同时,降低有功功率损耗、SOH损失和充放电切换次数。

2)提出多智能体DRL采用“去中心化训练+去中心化决策”实现方式,基于“在线学习模式”DDPG算法在智能体本地训练各智能体网络参数,能自动适应包含不确定性因素的大容量电池储能电站功率分配场景,且可减小总能量管理系统的运算压力,减轻系统通信压力,提高模型与算法的实用性。

进一步提高应用于大容量电池储能电站功率分配决策的MADRL在线效率是下一步的研究重点。

参考文献

- [1] YANG Zhen, XIA Li, GUAN Xiaohong. Fluctuation reduction of wind power and sizing of battery energy storage systems in microgrids [J]. IEEE Transactions on Automation Science and Engineering, 2020, 17(3): 1195–1207.
- [2] LIU Yixue, HE Qing, SHI Xingping, et al. Energy storage in China: development progress and business model [J]. Journal of Energy Storage, 2023(72): 108240.1–108240.10.
- [3] ELALFY D A, GOUDA E, KOTB M F, et al. Comprehensive review of energy storage systems technologies, objectives, challenges, and future trends [J]. Energy Strategy Reviews,

- 2024(54): 101482.
- [4] 王凯亮, 孔慧超, 李俊辉, 等. 考虑配电网可靠性的储能系统选址定容优化[J]. 南方电网技术, 2022, 16(4): 21–29.
WANG Kailiang, KONG Huichao, LI Junhui, et al. Optimization of energy storage system location and capacity considering the reliability of distribution network [J]. Southern Power System Technology, 2022, 16(4): 21–29.
- [5] 齐彩娟, 李牧远, 吴奕辰, 等. 基于最大评价差异的储能多场景区间模糊层次选型方法[J]. 电力系统保护与控制, 2024, 52(22): 104–115.
QI Caijuan, LI Muyuan, WU Yichen, et al. A fuzzy hierarchical selection method for an energy storage multi scenario interval based on maximum evaluation difference [J]. Power System Protection and Control, 2024, 52(22): 104–115.
- [6] 刘禹彤, 赵琳, 王鑫太, 等. 基于SE-improved SGMD的混合储能容量优化配置方法研究[J]. 高压电器, 2024, 60(10): 78–85.
LIU Yutong, ZHAO Lin, WANG Xintai, et al. Research on optimal configuration method of hybrid energy storage capacity based on SE-improved SGMD [J]. High Voltage Apparatus, 2024, 60(10): 78–85.
- [7] 何耿生, 陈晖, 张世旭, 等. 储能提升高比例新能源基地低碳-经济-可靠性能的效益分析[J]. 南方电网技术, 2024, 18(4): 120–129.
HE Gengsheng, CHEN Hui, ZHANG Shixu, et al. Benefit analysis of energy storage to improve low carbon, economy and reliability performance of high proportion renewable energy base [J]. Southern Power System Technology, 2024, 18(4): 120–129.
- [8] 余东旭. 大型电化学储能电站设计创新及应用[J]. 电力建设, 2024, 45(6): 91–99.
YU Dongxu. Innovative design and application of a large-scale electrochemical energy storage power station [J]. Electric Power Construction, 2024, 45(6): 91–99.
- [9] 李建林, 王哲, 曾伟, 等. 百兆瓦级电化学储能电站能量管理研究综述[J]. 高电压技术, 2023, 49(1): 72–84.
LI Jianlin, WANG Zhe, ZENG Wei, et al. Review of energy management research on 100-Megawatt electrochemical energy storage power stations [J]. High Voltage Engineering, 2023, 49(1): 72–84.
- [10] 李建林, 李雅欣, 刘海涛, 等. 计及储能电站安全性的功率分配策略研究[J]. 电工技术学报, 2022, 37(23): 5976–5986.
LI Jianlin, LI Yaxin, LIU Haitao, et al. Research on power distribution strategy considering the safety of energy storage power station [J]. Transactions of China Electrotechnical Society, 2022, 37(23): 5976–5986.
- [11] 程苒, 段瑶, 吴亚雄, 等. 考虑等效循环放电深度约束的配电网储能规划[J]. 南方电网技术, 2024, 18(10): 54–63.
CHENG Ran, DUAN Yao, WU Yaxiong, et al. Energy storage planning of distribution network considering equivalent cyclic discharge depth constraints [J]. Southern Power System Technology, 2024, 18(10): 54–63.
- [12] 栗峰, 郝雨辰, 周昶, 等. 电网侧电化学储能调度运行及其关键技术[J]. 供用电, 2020, 37(6): 82–90.
LI Feng, HAO Yuchen, ZHOU Chang, et al. Dispatching operation and key technologies analysis of electrochemical energy storage on grid side [J]. Distribution & Utilization, 2020, 37(6): 82–90.
- [13] 郭向伟, 王晨, 钱伟, 等. 电池储能系统均衡方法研究综述[J]. 电工技术学报, 2024, 39(13): 4204–4225.
GUO Xiangwei, WANG Chen, QIAN Wei, et al. A review of equalization methods for battery energy storage system [J]. Transactions of China Electrotechnical Society, 2024, 39(13): 4204–4225.
- [14] 李相俊, 马锐. 考虑电池组健康状态的储能系统能量管理方法[J]. 电网技术, 2020, 44(11): 4210–4217.
LI Xiangjun, MA Rui. Energy management method of energy storage system considering the SOH of battery pack [J]. Power System Technology, 2020, 44(11): 4210–4217.
- [15] 徐珂, 余小平, 李凯, 等. 基于电池健康状态预测的储能系统功率分配[J]. 电力系统及其自动化学报, 2025, 37(5): 113–120.
XU Ke, SHE Xiaoping, LI Kai, et al. Power allocation method for energy storage systems based on battery health state prediction [J]. Proceedings of the CSU-EPSA, 2025, 37(5): 113–120.
- [16] 付华, 陆鹏, 张俊男. 基于A-SA-WOA算法的直流微电网全钒液流电池储能系统功率分配策略[J]. 电工技术学报, 2023, 38(7): 1826–1837.
FU Hua, LU Peng, ZHANG Junnan. Power allocation strategy of DC microgrid all vanadium redox flow battery energy storage system based on A-SA-WOA algorithm [J]. Transactions of China Electrotechnical Society, 2023, 38(7): 1826–1837.
- [17] 赵焱, 史学伟, 王莉斌, 等. 一种规模化电池储能电站功率分配策略[J]. 电网技术, 2022, 46(12): 5004–5012.
ZHAO Yi, SHI Xuewei, WANG Libin, et al. Power allocation strategy of large-scaled battery energy storage power station [J]. Power System Technology, 2022, 46(12): 5004–5012.
- [18] LI Xiangjun, WANG Shangxing. Energy management and operational control methods for grid battery energy storage systems [J]. CSEE Journal of Power and Energy Systems, 2021, 7(5): 1026–1040.
- [19] 王雨虹, 梁策, 付华. 基于PSTSO多目标优化储能系统功率分配[J]. 电力系统及其自动化学报, 2024, 36(8): 13–22.
WANG Yuhong, LIANG Ce, FU Hua. Optimize the power allocation strategy of energy storage system based on PSTSO multi-objective [J]. Proceedings of the CSU-EPSA, 2024, 36(8): 13–22.
- [20] 陈薇, 侯杨成, 张里, 等. 计及损耗的钒电池储能系统功率优化分配策略[J]. 电工技术学报, 2020, 35(19): 4038–4047.
CHEN Wei, HOU Yangcheng, ZHANG Li, et al. Power optimization allocation strategy for vanadium battery energy storage system considering loss [J]. Transactions of China Electrotechnical Society, 2020, 35(19): 4038–4047.
- [21] AZIZIVAHED A, BARANI M, RAZAVI S E, et al. Energy storage management strategy in distribution networks utilised by photovoltaic resources [J]. IET Generation, Transmission & Distribution, 2018, 12(21): 5627–5638.
- [22] LI Xining, LÜ Lyu, GENG Guangchao, et al. Power allocation strategy for battery energy storage system based on cluster switching [J]. IEEE Transactions on Industrial Electronics, 2022, 69(4): 3700–3710.
- [23] 黄婧杰, 李金成, 刘科明, 等. 含CVaR及增广 ε -约束法的多目标光储充电站容量优化配置[J]. 南方电网技术, 2023, 17(10): 94–103.
HUANG Jingjie, LI Jincheng, LIU Keming, et al. Optimal allocation of multi-objective photovoltaic energy storage charging station capacity with CVaR and augmented ε -constraint

- method [J]. Southern Power System Technology, 2023, 17 (10): 94 – 103.
- [24] 李鹏, 钟瀚明, 马红伟, 等. 基于深度强化学习的有源配电网多时间尺度源荷储协同优化调控[J]. 电工技术学报, 2025, 40(5): 1487 – 1502.
- LI Peng, ZHONG Hanming, MA Hongwei, et al. Multi-time scale optimal dispatch of source-load-storage coordination inactive distribution network based on deep reinforcement learning [J]. Transactions of China Electrotechnical Society, 2025, 40 (5): 1487 – 1502.
- [25] 梁涛, 柴露露, 谭建鑫, 等. 基于深度强化学习算法的氢耦合电-热综合能源系统优化调度[J]. 电力自动化设备, 2025, 45 (1): 59 – 66.
- LIANG Tao, CHAI Lulu, TAN Jianxin, et al. Optimal scheduling of hydrogen-coupled electrothermal integrated energy system based on deep reinforcement learning algorithm [J]. Electric Power Automation Equipment, 2025, 45 (1): 59 – 66.
- [26] 巨云涛, 陈希. 基于双层多智能体强化学习的微网群分布式有功无功协调优化调度[J]. 中国电机工程学报, 2022, 42 (23): 8534 – 8547.
- JU Yuntao, CHEN Xi. Distributed active and reactive power coordinated optimal scheduling of networked microgrids based on two-layer multi-agent reinforcement learning [J]. Proceedings of the CSEE, 2022, 42 (23): 8534 – 8547.
- [27] 董雷, 刘雨, 乔骥, 等. 基于多智能体深度强化学习的电热联合系统优化运行[J]. 电网技术, 2021, 45(12): 4729 – 4738.
- DONG Lei, LIU Yu, QIAO Ji, et al. Optimal dispatch of combined heat and power system based on multi-agent deep reinforcement learning [J]. Power System Technology, 2021, 45(12): 4729 – 4738.
- [28] 陈池瑶, 苗世洪, 姚福星, 等. 基于多智能体算法的多微电网-配电网分层协同调度策略[J]. 电力系统自动化, 2023, 47 (10): 57 – 65.
- CHEN Chiyao, MIAO Shihong, YAO Fuxing, et al. Hierarchical cooperative dispatching strategy of multi-microgrid and distribution networks based on multi-agent algorithm [J]. Automation of Electric Power Systems, 2023, 47 (10): 57 – 65.
- [29] 徐业琰, 姚良忠, 廖思阳, 等. 基于多智能体 Actor-double-critic 深度强化学习的源-网-荷-储实时优化调度方法[J]. 中国电机工程学报, 2025, 45(2): 513 – 527.
- XU Yeyan, YAO Liangzhong, LIAO Siyang, et al. Research on real-time optimal scheduling method of source-network-load-storage based on multi-agent Actor-Double-Critic deep reinforcement learning [J]. Proceedings of the CSEE, 2025, 45 (2): 513 – 527.
- [30] 孙国强, 殷岩岩, 卫志农, 等. 基于深度确定性策略梯度的主动配电网有功-无功协调优化调度[J]. 电力建设, 2023, 44 (11): 33 – 42.
- SUN Guoqiang, YIN Yanyan, WEI Zhinong, et al. Coordinated optimal dispatch of active and reactive power in active distribution networks using deep deterministic strategy gradient [J]. Electric Power Construction, 2023, 44 (11): 33 – 42.
- [31] 李建林, 武亦文, 王楠, 等. 吉瓦级电化学储能电站信息架构与安防体系综述[J]. 电力系统自动化, 2021, 45(23): 179 – 191.
- LI Jianlin, WU Yiwen, WANG Nan, et al. Review of information architecture and security system of gigawatt electrochemical energy storage power station [J]. Automation of Electric Power Systems, 2021, 45(23): 179 – 191.
- [32] 李翠萍, 王艺茗, 李军徽, 等. 基于多储能动态调控的光储黑启动控制策略[J]. 太阳能学报, 2024, 45(9): 611 – 622.
- LI Cuiping, WANG Yiming, LI Junhui, et al. Solar storage black start control strategy based on multi-storage dynamic regulation [J]. Acta Energetica Sinica, 2024, 45(9): 611 – 622.
- [33] 彭寒梅, 胡磊, 谭貌, 等. 基于深度强化学习的电-气区域综合能源系统安全校正控制决策方法[J]. 电网技术, 2024, 48 (3): 1019 – 1030.
- PENG Hanmei, HU Lei, TAN Mao, et al. Decision-making method for electricity-gas region integrated energy system security correction based on deep reinforcement learning [J]. Power System Technology, 2024, 48(3): 1019 – 1030.
- [34] 王琦, 杨毅远, 江季. Easy RL: 强化学习教程[M]. 北京: 人民邮电出版社, 2022.

收稿日期: 2024-12-23; 网络首发日期: 2025-04-22

作者简介:

彭寒梅(1979), 女, 通信作者, 副教授, 博士, 研究方向为综合能源系统控制与优化调度, penghanmei8@163.com;

赵长桥(2001), 男, 硕士研究生, 研究方向为大容量电池储能电站功率分配, zhaochangqiao2021@163.com;

谭貌(1981), 男, 教授, 博士, 研究方向为综合能源系统及其优化, mr. tanmao@gmail.com。

附录

表 A1 大容量电池储能电站算例参数

Tab. A1 Calculation parameters of large-capacity battery energy storage power station

设备	参数	数值	设备	参数	数值
储能电站	额定功率/MW	10	储能子系统	额定功率/MW	0.625
	额定容量/MWh	20		额定容量/MWh	1.25
	额定功率/MW	2.5		电池欧姆内阻/ Ω	1.9×10^{-4}
储能单元	额定容量/MWh	5	储能子系统	电池极化内阻/ Ω	3.81×10^{-4}
	变压器容量/MVA	2.5		P_{BL}/W	10
储能单元	储能PCS额定功率/MW	0.63	储能子系统	A_{con} ; $\alpha_{soc}; \alpha_T$	1; 1.75; 与上一时刻充放电状态相同取 1.05, 相反取 0.95