

# 基于深度强化学习的微能源系统优化调度\*

张波, 冯国礼, 郭景维, 王敏, 秦振威  
(国网宁夏电力有限公司 信息通信公司, 宁夏 银川 750001)

**摘要:** 微能源系统是城市配网终端的重要聚合部分,其应对源荷随机特性的能力为城市配网稳定运行提供了有效支撑。针对城市工业园微能源系统提出一种考虑源荷随机波动的动态调度方法。考虑工业园多种可调度资源对微能源系统的经济调度构建数学模型,然后将构建的微能源系统经济调度模型表示为具有连续动作调节的深度强化学习(DRL)模型,最后采用双延迟深度确定性策略梯度算法获取 DRL 模型下的动态连续调度策略。所提方法不仅避免对源荷随机波动的不确定性进行建模,同时也避免了离散  $Q$  学习的可调节设备出力不连续性。仿真结果表明所提出的动态调度方法具有更好的经济性和自适应性。

**关键词:** 微能源系统; 深度强化学习; 优化调度; 源荷随机波动

中图分类号: TM73 文献标志码: A 文章编号: 1673-6540(2022)11-0063-08

doi: 10.12177/emca.2022.121

## Optimal Scheduling of Micro-Energy System Based on Deep Reinforcement Learning\*

ZHANG Bo, FENG Guoli, GUO Jingwei, WANG Min, QIN Zhenwei  
(Information & Communication Company, State Grid Ningxia Electric Power Co., Ltd., Yinchuan 750001, China)

**Abstract:** Micro-energy system is an important aggregation part of urban distribution network terminals, and its ability to cope with the random characteristics of source-load provides an effective support for stable operation of urban distribution network. An intelligent dynamic scheduling method considering the random fluctuation of source-load is proposed for the micro-energy system in urban industrial park. A mathematical model is constructed for the economic dispatch of micro-energy system considering multiple dispatchable resources in the industrial park. Then, the constructed economic dispatch model of micro-energy system is represented as a deep reinforcement learning model with continuous action regulation. Finally, a dual delayed deep deterministic policy gradient algorithm is used to obtain the dynamic continuous dispatch policy under the deep reinforcement learning model. The proposed method not only avoids modeling the uncertainty of random fluctuation of source-load, but also avoids the discontinuity of adjustable equipment output with discrete  $Q$ -learning. Simulation results show that the proposed dynamic scheduling method has better economy and adaptivity.

**Key words:** micro-energy system; deep reinforcement learning (DRL); optimal scheduling; random fluctuation of source-load

收稿日期: 2022-08-02; 收到修改稿日期: 2022-09-21

\* 基金项目: 国网宁夏电力有限公司科技项目(5229XT20003T)

作者简介: 张波(1986—),男,硕士,高级工程师,研究方向为计算机应用技术。

冯国礼(1984—),男,本科,高级工程师,研究方向为计算机应用技术、网络安全。

郭景维(1984—),男,本科,工程师,研究方向为计算机应用技术。

## 0 引言

国际可再生能源署发布的《能源转型之电网灵活性》报告中指出,到 2050 年,全球风电、光伏等可再生能源在未来电力系统中占比将达到 85%<sup>[1]</sup>。国家发改委和国家能源局发布的《能源生产和消费革命战略(2016—2030)》中指出,到 2023 年实现非化石能源发电量占全部发电量的比重力争达到 50%<sup>[2]</sup>。为实现我国能源革命战略的顺利实施,中央财经委员会第九次会议提出构建以新能源为主体的新型电力系统<sup>[3]</sup>,其基于可再生能源和清洁能源、以主干电网和微网构成主要架构,是一种可持续的电力发展模式<sup>[4-6]</sup>。

在新型电力系统中,微网不仅是规模化接入和消纳可再生能源的重要环节,而且作为城市配网终端的主要聚合单元,通过自治运行能够实现对城市配网调度运行性能的优质支撑<sup>[7]</sup>。微网中存在多种能源供给和用能需求,通过经济安全的能量管理实现清洁能源的消纳和多种能源的协调互补是重要的研究内容之一。文献[8]针对工业微网提出了两阶段多时间尺度调度策略,通过协调供能、蓄热装置实现多设备互补运行和成本降低;文献[9-10]针对工业园微能源网构建了考虑生产约束的经济性最优调度模型,实现了工业生产调度时序性和用能需求多元性的结合;文献[11]针对炼油工业提出一种降低用能成本的需求侧能量调度方案;文献[12]针对工业园空调系统提出一种双层优化模型,减小了用户空调系统运行费用,并增加了用户代理商利润;文献[13]针对工业园综合能源系统提出考虑供能可靠性的能量调度方法,提升了综合能源系统的供能可靠性和经济性。上述研究虽然从微网不同的方面提升调度运行的经济性,但均未有效考虑可再生能源和负荷波动特性的影响。

为应对微网中源荷不确定性,文献[14]采用随机规划方法对不确定性因素进行概率分布分析,而文献[15-16]采用鲁棒优化方法对不确定性因素建模,虽然这些方法能够有效刻画不确定性,但是受到不确定性场景数目和最坏场景约束的影响,仍旧无法应对源荷不确定性对微网经济调度的影响。随着人工智能技术的发展,基于数据驱动的机器学习方法逐渐在优化运行方面得到

应用和发展。深度强化学习(DRL)作为机器学习的一种典型代表,结合了深度学习强大的信息表征能力和强化学习的序列决策优化能力,为多种优化调度问题提供了一种新的解决途径<sup>[17-18]</sup>,同时 DRL 的自主学习能力和自适应能力在应对不确定因素的影响时,具有明显的优势。文献[19-20]提出一种基于 Q-learning 方法的综合能源微网优化调度方法以应对光伏出力和负荷需求的不确定性;文献[21]采用深度确定性策略梯度算法(DDPG)提出一种自适应不确定性经济调度方法,实现了任意场景下的电力系统动态经济调度。然而,基于 Q-learning 方法的强化学习算法只能处理离散动作的问题,而现实中众多研究问题均为连续动作问题,同时基于 Q-learning 方法和 DDPG 方法的强化学习算法存在过估计的问题。

基于上述分析,针对以工业园为例的微能源系统,为应对源荷不确定性产生的影响同时提升微能源系统的运行经济性,提出一种基于双延迟深度确定性策略梯度(TD3)<sup>[22]</sup>的动态优化调度方法。首先构建了工业园微能源系统的经济调度模型,然后基于行动器-评判器(Actor-Critic)框架将经济调度模型表示为具有连续动作调节的 DRL 模型,并使用双延迟深度确定性策略梯度算法获取 DRL 模型下的动态连续调度策略,最后,通过对比算例分析,验证所提方法的优越性。

## 1 工业园微能源系统优化调度模型

### 1.1 工业园微能源系统

工业园微能源系统由配电网和配气网提供外部能量输入,由屋顶分布式光伏提供内部能量输入。工业园微能源系统架构图如图 1 所示,能量转换设备有热电联产机组(CHP)、燃气锅炉(GB)、电锅炉(EB),能量存储设备有蓄电池(BES)、空气压缩储能系统(CAES),用能设备有工业电负荷、工业热负荷和气负荷。

#### 1.1.1 热电联产机组模型

热电联产机组输出的电功率和热功率为

$$P_{\text{CHP}}(t) = Q_{\text{CHP}}(t)H_{\text{NG}}\eta_{\text{CHP}} \quad (1)$$

$$H_{\text{CHP}}(t) = mP_{\text{CHP}}(t) \quad (2)$$

式中: $P_{\text{CHP}}(t)$ 和 $H_{\text{CHP}}(t)$ 分别为热电联产机组输出的电功率和热功率; $Q_{\text{CHP}}(t)$ 为热电联产机组消耗的天然气量; $H_{\text{NG}}$ 为天然气热值; $\eta_{\text{CHP}}$ 和 $m$ 分别

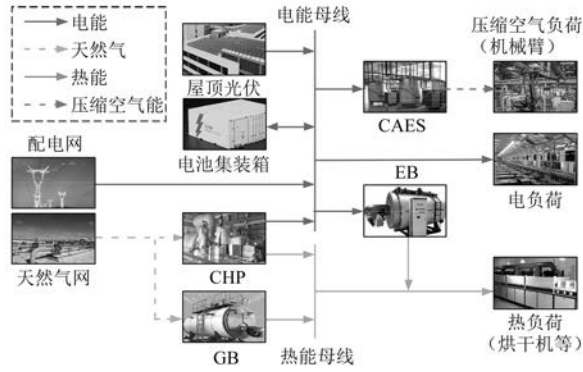


图1 工业园微能源系统架构图

为热电联产机组产电效率和热电比值。

热电联产机组需要满足运行上下限约束,如下所示:

$$P_{CHP}^{\min} \leq P_{CHP}(t) \leq P_{CHP}^{\max} \quad (3)$$

式中: $P_{CHP}^{\min}$  和  $P_{CHP}^{\max}$  分别为热电联产机组输出电功率的最小和最大值。

### 1.1.2 燃气锅炉模型

燃气锅炉通过消耗天然气产生热能,燃气锅炉输出的热功率如下所示:

$$H_{GB}(t) = Q_{GB}(t)H_{NG}\eta_{GB} \quad (4)$$

式中: $H_{GB}(t)$  为燃气锅炉输出的热功率; $Q_{GB}(t)$  为燃气锅炉消耗的天然气量; $\eta_{GB}$  为燃气锅炉产热效率。

燃气锅炉需要满足运行上下限约束,具体如下所示:

$$H_{GB}^{\min} \leq H_{GB}(t) \leq H_{GB}^{\max} \quad (5)$$

式中: $H_{GB}^{\min}$  和  $H_{GB}^{\max}$  分别为燃气锅炉输出热功率的最小和最大值。

### 1.1.3 电锅炉模型

电锅炉输出的热功率如下所示:

$$H_{EB}(t) = P_{EB}(t)\eta_{EB} \quad (6)$$

式中: $H_{EB}(t)$  为电锅炉输出的热功率; $P_{EB}(t)$  为电锅炉消耗的电能; $\eta_{EB}$  为电锅炉电热效率。

电锅炉需要满足运行上下限约束,如下所示:

$$P_{EB}^{\min} \leq P_{EB}(t) \leq P_{EB}^{\max} \quad (7)$$

式中: $P_{EB}^{\min}$  和  $P_{EB}^{\max}$  分别为电锅炉输入电热功率的最小和最大值。

### 1.1.4 蓄电池模型

蓄电池通过存储和释电能实现可再生能源的消纳、电能使用高峰时段转移等。蓄电池的荷

电状态(SOC)如下所示:

$$S_{BES}(t) = S_{BES}(t-1) - \frac{\eta_{BES}P_{BES}(t)\Delta t}{E_{BES}} \quad (8)$$

式中: $S_{BES}(t)$  和  $S_{BES}(t-1)$  分别为  $t$  时刻和上一时刻蓄电池的 SOC; $\eta_{BES}$  为蓄电池的充放电效率; $P_{BES}(t)$  为蓄电池的充放电功率; $E_{BES}$  为蓄电池容量; $\Delta t$  为调度时间间隔。

蓄电池需要满足运行上下限约束和 SOC 约束,如下所示:

$$\begin{cases} E_{BES\_start} = E_{BES\_end} \\ P_{BES}^{\min} \leq P_{BES}(t) \leq P_{BES}^{\max} \\ S_{BES}^{\min} \leq S_{BES}(t) \leq S_{BES}^{\max} \end{cases} \quad (9)$$

式中: $E_{BES\_start}$  和  $E_{BES\_end}$  分别为调度周期初始和结束时刻的蓄电池容量; $P_{BES}^{\min}$  和  $P_{BES}^{\max}$  分别为蓄电池充放电功率的最小和最大值; $S_{BES}^{\min}$  和  $S_{BES}^{\max}$  分别为蓄电池 SOC 的最小和最大值。

### 1.1.5 CAES 模型

CAES 通过消耗电能将空气压缩至储气罐中,并通过调节储气罐中的压力将空气压缩输送到工厂气动系统,如动力机械臂<sup>[9]</sup>。CAES 储气量如下所示:

$$E_{CA}(t) = E_{CA}(t-1) - D_{CA}(t)\Delta t + \eta_{CA}P_{CA}(t)\Delta t \quad (10)$$

式中: $E_{CA}(t)$ 、 $E_{CA}(t-1)$  和  $D_{CA}(t)$  分别为 CAES 在  $t$  时刻、上一时刻的储气量和 CAES 在  $t$  时刻的用气量; $P_{CA}(t)$  为 CAES 中空压机  $t$  时刻的输出功率; $\eta_{CA}$  为 CAES 压缩机效率。

CAES 需要满足气量连续性约束、气罐压力约束和 CAES 启停约束,如下所示:

$$\begin{cases} E_{CA\_start} = E_{CA\_end} \\ P_{CA}^{\min} \leq P_{CA}(t) \leq P_{CA}^{\max} \\ P_{CA}^{\min}V_{CA} \leq E_{CA} \leq P_{CA}^{\max}V_{CA} \\ E_{CA}(t) + u_{CA}(t)p_{CA}^{\max}V_{CA} \leq 1.1p_{CA}^{\max}V_{CA} \\ 1.1p_{CA}^{\min}V_{CA} \leq E_{CA}(t) + u_{CA}(t)p_{CA}^{\min}V_{CA} \end{cases} \quad (11)$$

式中: $E_{CA\_start}$  和  $E_{CA\_end}$  为调度周期初始和结束时刻的储气量; $P_{CA}^{\min}$  和  $P_{CA}^{\max}$  分别为压缩机功率的最小和最大值; $p_{CA}^{\min}$  和  $p_{CA}^{\max}$  分别为储气罐允许的最小和最大压强; $V_{CA}$  为储气罐体积; $u_{CA}(t)$  为  $t$  时刻 CAES 中空压机的运行状态, $u_{CA}(t) = 0$  或  $1$ , 分别表示空压机停止运行和启动运行。

## 1.2 经济调度模型

### 1.2.1 目标函数

工业园微能源系统在保证工业生产任务要求的前提下,通过合理安排可控机组的调度计划,可实现综合运行成本最低的目标。微能源系统运行成本主要由购能成本、设备维护成本和碳排放成本构成。微能源系统的经济调度目标函数如下所示:

$$\begin{cases} \min F = \min(C_b + C_o + C_{ce}) \\ C_b = \sum_{t=1}^T \{c_e(t)P_E(t) + c_g[Q_{GB}(t) + Q_{CHP}(t)]\} \Delta t \\ C_o = \sum_{t=1}^T [c_b P_{BES}(t) + c_{ca} P_{CA}(t)] \Delta t \\ C_{ce} = \sum_{t=1}^T \{c_{ce}[Q_{GB}(t) + Q_{CHP}(t)]\} \Delta t \end{cases} \quad (12)$$

式中: $C_b$ 、 $C_o$  和  $C_{ce}$  分别为微能源系统的购能成本、设备维护成本和碳排放成本; $c_e$ 、 $c_g$ 、 $c_b$ 、 $c_{ca}$  和  $c_{ce}$  分别为市电电价、天然气单位价格、蓄电池折损费用系数、CAES 维护费用系数和碳排放费用系数,其中  $c_e$  由上网电价  $c_{sale}$  和购电电价  $c_{buy}$  构成; $T$  为调度周期数。

### 1.2.2 约束条件

工业园微能源系统优化调度约束包括电能、热能、气能平衡约束、购能约束和设备运行约束,如下所示:

$$\begin{cases} P_{PV} + P_E + P_{BES,d} + P_{CHP} = P_L + P_{EB} + P_{CA} + P_{BES,c} \\ H_{CHP} + H_{EB} + H_{GB} = H_L \\ D_{CA} = D_{LCA} \end{cases} \quad (13)$$

$$\begin{cases} P_E^{\min} \leq P_E(t) \leq P_E^{\max} \\ Q_G^{\min} \leq Q_G(t) \leq Q_G^{\max} \\ Q_G(t) = Q_{CHP}(t) + Q_{GB}(t) \end{cases} \quad (14)$$

式中: $P_{PV}$  为屋顶光伏发电功率; $P_E$  为微能源系统向配电网购售功率; $P_{BES,c}$  和  $P_{BES,d}$  分别为蓄电池充、放电功率; $P_L$ 、 $H_L$  和  $D_{LCA}$  分别为工业园微能源系统的电负荷、热负荷和生产产线用气量; $P_E^{\min}$  和  $P_E^{\max}$  分别为微能源系统向配电网购售功率的最小和最大值; $Q_G$  为微能源系统向配气网购买的天然气量; $Q_G^{\min}$  和  $Q_G^{\max}$  分别为购买天然气的最小和最大值。设备运行约束如式(3)、式(5)、式

(7)、式(9)和式(11)所示。

## 2 基于 DRL 的经济调度方法

### 2.1 经济调度的强化学习框架设计

强化学习是机器学习中一种通过智能体和环境交互学习来制定最佳状态-动作策略的方法。强化学习的核心是智能体和环境的交互。智能体通过观测环境的状态根据策略函数给出环境动作,并基于状态和动作计算每一步奖励,环境执行智能体给定的动作并将新的状态提供给智能体。智能体依据每一步的奖励来寻找使累计奖励最大化的状态-动作策略。图 2 所示为强化学习的核心框架。

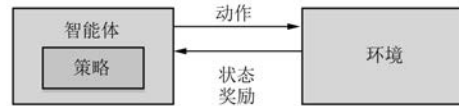


图 2 强化学习核心框架

强化学习本质可使用马尔科夫决策过程(MDP)表述,即下一时刻状态仅与当前时刻状态和动作有关。MDP 常表示为一个元组  $(S, A, T_p, R)$ , 其中: $S$  表示状态空间; $A$  表示动作空间; $T_p$  表示状态转移函数,  $T_p: S \times A \times S' \rightarrow [0, 1]$ , 即状态  $S$  执行动作后转移到下一个状态  $S'$  的概率; $R$  表示奖励函数,  $R: S \times A \rightarrow R$ , 即发生状态转移时环境给出的即时奖励。

(1) 状态空间  $S$ 。在工业园微能源系统的经济调度模型中,环境的观测状态包括负荷需求、屋顶光伏发电功率、蓄电池 SOC 和 CAES 储气状态。状态  $S$  可表示为

$$S = [P_L \ H_L \ E_L \ P_{PV} \ S_{BES} \ E_{CA}] \quad (15)$$

(2) 动作空间  $A$ 。智能体的动作包括热电联产机组输出的电热功率、燃气锅炉输出的热功率、电锅炉输入电功率和输出的热功率、蓄电池的充放电功率、CAES 输入电功率和储用气量以及微能源系统与配电网交互功率。依据式(2)、式(4)、式(6)和式(13),热电联产机组输出的电功率、燃气锅炉输出的热功率和蓄电池的充放电功率是智能体必不可少的动作,其他动作均可由相应的计算式得到。动作空间  $A$  可表示为

$$A = [P_{CHP} \ H_{GB} \ P_{BES}] \quad (16)$$

(3) 状态转移函数  $f$ 。智能体的状态转移函数如下所示:

$$s_{t+1} = f(s_t, a_t, \sigma_t) \quad (17)$$

式中:  $s_t$  和  $a_t$  分别为  $t$  时刻的状态和动作;  $\sigma_t$  为  $t$  时刻的随机项, 即体现屋顶光伏发电功率和微能源系统负荷需求导致的随机影响。

(4) 奖励函数  $r$ 。为实现工业园微能源系统的最小调度成本, 将微能源系统的经济调度模型中目标函数式(12)改写为奖励函数如下所示:

$$r_t = -k[C_b(s_t, a_t) + C_o(s_t, a_t) + C_{ce}(s_t, a_t)] - \xi \quad (18)$$

式中:  $r_t$  为  $t$  时刻的奖励函数值;  $k$  为比例因子;  $\xi$  为惩罚因子, 当约束条件不满足时, 惩罚因子为常数, 当约束条件满足时, 惩罚因子为 0。

(5) 状态-动作值函数  $Q_\pi(s, a)$ 。智能体的策略  $\pi$  为状态  $S$  到动作  $A$  的映射, 智能体采用状态-动作值函数  $Q_\pi(s, a)$  来衡量策略  $\pi$  的优劣程度, 如下所示:

$$Q_\pi(s, a) = E_\pi \left[ \sum_{t=0}^T \gamma^t r_t \mid s=s_t, a=a_t \right] \quad (19)$$

式中:  $\gamma$  为折扣因子,  $\gamma \in [0, 1]$ , 表示未来奖励在累积奖励中所占比重;  $r_t$  为  $t$  时刻的奖励;  $E_\pi[\cdot]$  为策略  $\pi$  的期望。

最优策略  $\pi^*$  可表示如下:

$$\pi^* = \operatorname{argmax} Q_\pi(s, a) \quad (20)$$

### 2.2 经济调度的 DRL 部署

经济调度的强化学习框架设计中, 智能体的动作  $A$  在现实场景中均为连续调节的变量, 且状态转移函数  $f$  中包含不确定性因素影响的随机项  $\sigma_t$ , 仅通过强化学习应对动作的连续调节和不确定性因素的影响将大大降低强化学习效率和适用性。深度学习是机器学习中一种基于对数据进行表征学习的方法, DRL 将深度学习的感知能力和强化学习的决策能力相结合, 有效提升强化学习的学习效率并极大地拓展了强化学习的使用范围。

TD3 算法基于 Actor-Critic 框架, 由  $Q$  值网络和策略网络两部分构成, 策略网络实现状态  $S$  到动作  $A$  的映射,  $Q$  值网络实现对策略网络建立映射的量化评估, 即状态-动作值函数  $Q_\pi(s, a)$ 。TD3 算法通过增加  $Q$  值网络和软更新的方式抑制  $Q_\pi(s, a)$  的过高估计。  $Q$  值网络将给出两个状

态-动作值函数  $Q_{\pi_1}(s, a)$  和  $Q_{\pi_2}(s, a)$ , 并取两者的最小值作为最终  $Q$  值网络的估计值, 同时通过构建与  $Q$  值网络和策略网络相对应的目标  $Q$  值网络和目标策略网络, 采用软更新的方式使  $Q$  值网络和策略网络向目标  $Q$  值网络和目标策略网络传递参数, 延缓  $Q_\pi(s, a)$  的更新速度。

基于 DRL 的工业园微能源系统经济调度智能体部署 TD3 算法, 通过 TD3 算法的训练可使智能体的策略网络具备任意环境状态  $s_t$  下产生最优动作  $a_t$ 。TD3 算法训练智能体的具体步骤如图 3 所示。

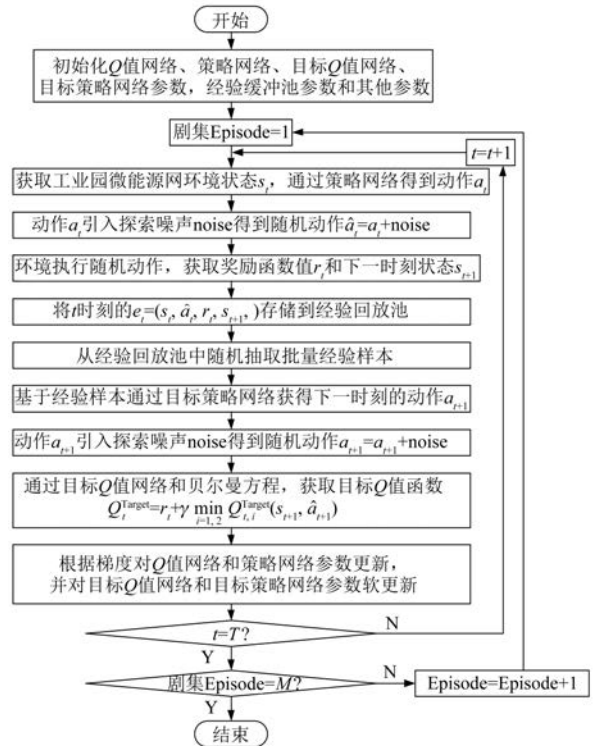


图 3 TD3 算法训练流程图

## 3 算例分析

### 3.1 参数设置

本文以西北地区某工业园 1#~3#生产厂房的微能源系统为例, 1#~3#厂房屋顶光伏装机容量为 2 MW, 配置热电联产机组 350 kW, 燃气锅炉 300 kW, 电锅炉 200 kW, 蓄电池 850 kW/850 kWh, CAES 300 kW, 储气罐容量 40 m<sup>3</sup>, 最大、最小压强 5、3.5 Pa。1#~3#生产厂房通过一台变压器与配电网相连, 微能源系统和配电网交互功率不大于 3 MW。配电网分时电价: 峰时段 (12:



00~19:00)0.59元、平时段(07:00~12:00,19:00~23:00)0.38元、谷时段(23:00~07:00)0.16元。配电网上网电价为0.2元,天然气价格为1.9元/m<sup>3</sup>。微能源系统设备参数如表1所示。

表1 微能源系统设备参数

参数名称	参数值	参数名称	参数值
$\eta_{CHP}$	0.32	$\eta_{BES\_d}$	0.92
$m$	1.25	$\eta_{CA}$	1
$\eta_{CB}$	0.82	$c_b$	0.05
$\eta_{EB}$	0.94	$c_{ca}$	0.01
$\eta_{BES\_c}$	0.92	$c_e$	0.1

工业园微能源系统的调度周期为24h,相邻两个调度时段的间隔为15min。以西北地区某工业园1#~3#生产厂房3月~5月日负荷数据和光伏发电数据作为历史数据对微能源系统经济调度智能体训练,使用6月的平均日负荷数据和平均光伏发电数据对智能体进行评估,如图4所示。基于TD3算法的DRL参数如表2所示。

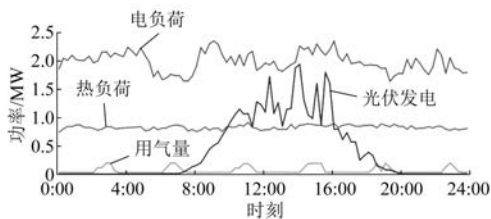


图4 平均日负荷曲线和光伏发电曲线

表2 基于TD3算法的DRL参数

参数名称	参数值
策略网络学习率	0.001
Q值网络学习率	0.001
折扣因子 $\gamma$	0.99
噪声标准差	0.01
批处理大小	64
比例因子 $k$	0.0004
惩罚因子 $\xi$	0.00012

### 3.2 基于DRL微能源系统经济调度分析

#### 3.2.1 基于TD3算法的微能源系统调度分析

使用历史数据训练智能体:智能体接收电、热、气负荷和光伏发电数据,并生成热电联产机组电功率、燃气锅炉热功率和蓄电池充放电功

率指令;根据图3对智能体的网络参数进行调整,并在平均奖励达到稳定值后,保存智能体形成的策略,即微能源系统的经济调度策略,智能体训练过程中奖励函数的变化如图5所示。然后采用图4所示的平均日负荷数据和平均光伏发电数据对智能体进行测试,基于TD3算法的微能源系统经济调度结果如图6~图8所示,其中图6为电能调度结果,图7为热能调度结果,图8为气能调度结果。

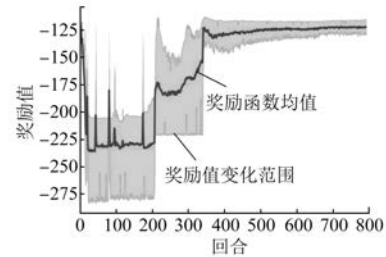


图5 基于TD3算法的微能源系统经济调度智能体训练图

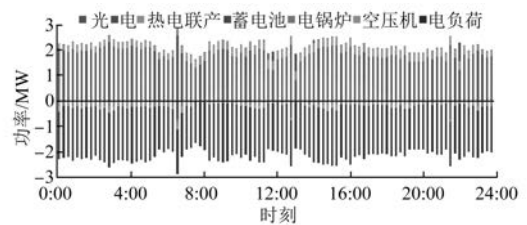


图6 基于TD3算法的电能调度结果

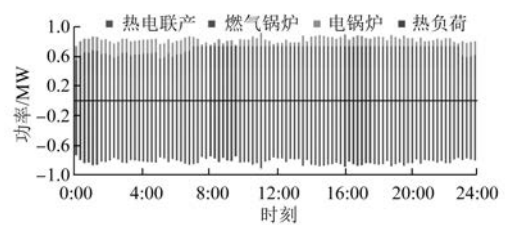


图7 基于TD3算法的热能调度结果

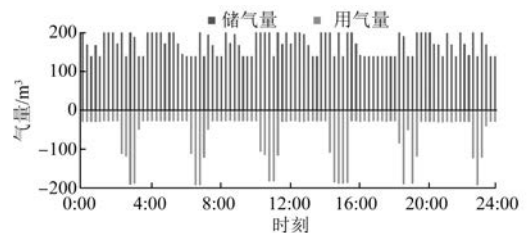


图8 基于TD3算法的气能调度结果

图 5 中阴影区和黑色曲线分别表示智能体训练过程中的奖励值变化范围和奖励函数的均值。通过图 5 可知基于 TD3 算法的智能体在 400 回合后逐渐收敛至稳定奖励值。图 6 展示了光伏发电功率、向电网购售功率、蓄电池充放电功率、热电联产机组和电锅炉耗电功率,空压机工作耗电功率和工业园 1#~3#生产厂房电负荷功率。如图 6 所示,微能源系统电能调度中蓄电池在谷价时充电并在其他时段放电,同时蓄电池会吸收光伏和热电联产机组产生的多余电能(12:00~12:15,13:30~13:45),此外电能调度中还会优先消纳光伏产生的电能以减低对市电的需求。图 7 展示了热电联产、燃气锅炉和电锅炉的热功率,以及工业园 1#~3#生产厂房热负荷功率。工业园由于生产工序特性,电、热、气负荷功率呈现较为规律的波动变化,且西北地区天然气存量丰富,成本经济。如图 7 所示,微能源系统热能调度中热电联产机组和燃气锅炉输出热功率占热负荷需求的比重超过 70%,而电锅炉输出的热功率受分时电价影响,在谷价时段比平价和峰价时段输出热功率更多。在峰价时段由于热电联产机组和燃气锅炉输出热功率已达最大值,热负荷需求不足的功率由电锅炉补足。生产产线用气量和储气罐储气量如图 8 所示,微能源系统的气能调度中储气罐在谷价时段通过空压机的工作尽可能维持最大压强,而在平价和峰价时段储气罐在满足生产产线用气量的基础上,减少维持最大压强的时刻以降低用电需求。通过上述调度结果分析,可知基于 TD3 算法的微能源系统调度智能体不仅能够获取全时段的最优奖励,同时每个调度时段也探索最优奖励并影响全时段的最优奖励。

### 3.2.2 不同调度方法对比分析

为验证提出的基于 TD3 算法的微能源系统经济调度方法的有效性,分别采用基于深度 Q 网络(DQN)的微能源系统经济调度方法和基于预测信息的调度方法进行比较。其中,基于 DQN 的微能源系统经济调度方法采用负荷需求、屋顶光伏发电功率、蓄电池 SOC 和 CAES 储气状态作为状态观测量,采用热电联产机组输出的电功率  $P_{\text{CHP}}$ 、燃气锅炉输出的热功率  $H_{\text{GB}}$  和蓄电池的充放电功率  $P_{\text{BES}}$  为动作量,并将  $P_{\text{CHP}}$ 、 $H_{\text{GB}}$ 、 $P_{\text{BES}}$  分别离散为  $\{0, 100, 200, 250, 300, 350\}$  kW、 $\{0, 100,$

$150, 200\}$  kW、 $\{-850, -450, -280, -50, 0, 50, 280, 450, 850\}$  kW。DQN 的折扣因子为 0.99,学习率为 0.001,批处理为 64。基于预测信息的调度方法采用神经网络对光伏发电功率和电、热、气负荷进行预测,然后采用优化求解器进行求解。三种方法均采用西北地区某工业园 1#~3#生产厂房 6 月平均日负荷数据和平均光伏发电数据进行调度结果比较,如表 3 所示。

表 3 不同调度方法比较 元

调度方法	调度成本	购能成本	维护成本	碳排成本
基于 TD3	49 785	48 306	60.263	1 418.2
基于 DQN	50 859	49 505	60.263	1 293.6
基于预测信息	50 177	48 743	60.263	1 373.6

由表 3 可知,基于 TD3 的微能源系统调度方法调度总成本为 49 785 元,其中购能成本占比最大,为 48 306 元,维护成本占比最小,为 60.263 元。基于 TD3 的调度方法比基于 DQN 的调度方法调度总成本节约 1 074 元,主要节约在购能成本;基于 TD3 的调度方法比基于预测信息的调度方法调度总成本节约 392 元,同样主要节约在购能成本。表 3 表明,基于 DQN 方法的动作为离散动作,无法对更细化的动作进行探索,基于预测信息的方法同实际信息存在误差,而基于 TD3 方法的动作连续,能够实现对动作边界范围内的所有值的选取和探索,同时其训练过程采用探索噪声扰动确保了不同误差情况下的有效探索,因而基于 TD3 的微能源系统经济调度方法减小了经济成本。

## 4 结 语

本文针对城市工业园微能源系统提出了一种基于 TD3 的动态调度方法,首先依据马尔科夫决策过程,将微能源系统调度模型构建为强化学习框架,设计包含购能成本、维护成本和碳排成本的微能源系统经济调度奖励函数,然后采用 Actor-Critic 框架的 TD3 算法对智能体进行训练,最后以西北地区工业园微能源系统为例进行了有效性验证。所提方法不仅能够确保动作的连续性,避免奖励值的过估计,同时所提方法不依赖于预测信息和不确定性建模,能够保证源荷随机波动时的调度策略有效性。通过与基于 DQN 的调度方

法和基于预测信息的调度方法对比可知,所提基于TD3的微能源系统调度方法具有更好的经济性。

【参考文献】

[ 1 ] International Renewable Energy Agency. Power system flexibility for the energy transition, part1: overview for policy makers[R]. Bonn:IRENA,2018.

[ 2 ] 国家发展改革委,国家能源局. 能源生产和消费革命战略(2016—2030)[R]. 北京:国家发展改革委,国家能源局,2016.

[ 3 ] 舒印彪,陈国平,贺静波,等. 构建以新能源为主体的新型电力系统框架研究[J]. 中国工程科学, 2021,23(6): 61.

[ 4 ] 周孝信,曾嵘,高峰,等. 能源互联网的发展现状与展望[J]. 中国科学:信息科学,2017,47(2): 149.

[ 5 ] 周孝信,鲁宗相,刘应梅,等. 中国未来电网的发展模式和关键技术[J]. 中国电机工程学报,2014,34(29): 4999.

[ 6 ] 周孝信,陈树勇,鲁宗相,等. 能源转型中我国新一代电力系统的技术特征[J]. 中国电机工程学报, 2018,38(7): 1893.

[ 7 ] THALE S S, WANDHARE R G, AGARWAL V. A novel reconfigurable microgrid architecture with renewable energy sources and storage [J]. IEEE Transactions on Industrial Applications, 2015, 51(2): 1805.

[ 8 ] 王成山,吕超贤,李鹏,等. 园区型综合能源系统多时间尺度模型预测优化调度[J]. 中国电机工程学报,2019,39(23): 6791.

[ 9 ] 马昕,裴玮,肖浩,等. 考虑复杂生产约束的电池生产工业园区能源网络与生产管理综合优化运行[J]. 电网技术,2018,42(11): 3566.

[ 10 ] PEI W, MA X, DENG W, et al. Industrial multi-energy and production management scheme in cyber-physical environments: a case study in a battery manufacturing plant [J]. IET Cyber-Physical Systems:Theory & Applications,2019,4(1): 13.

[ 11 ] REKA S S, RAMESH V. Industrial demand side response modelling in smart grid using stochastic optimisation considering refinery process [J]. Energy & Buildings,2016,127: 84.

[ 12 ] 陈厚合,吴桐,李本新,等. 考虑建筑热惯性的园区代理商电价策略及用能优化[J]. 电力系统自动化,2021,45(3): 148.

[ 13 ] 黄宗宏,曹雨晨,胡志冰,等. 基于改进遗传算法的工业园区综合能源系统能量管理方法[J]. 电力建设,2019,40(8): 51.

[ 14 ] ZHENG J, KOU Y, LI M, et al. Stochastic optimization of cost-risk for integrated energy system considering wind and solar power correlated [J]. Journal of Modern Power Systems and Clean Energy, 2019,7(6): 1472.

[ 15 ] 郭帅,刘鲁嘉,杨萌,等. 计及可控负荷的微网自治调度模型[J]. 电网技术,2017,41(1): 48.

[ 16 ] LUO Z, GU W, WU Z, et al. A robust optimization method for energy management of CCHP microgrid [J]. Journal of Modern Power Systems and Clean Energy,2018,6(1): 132.

[ 17 ] YANG T, ZHAO L, LI W, et al. Reinforcement learning in sustainable energy and electric systems: a survey [J]. Annual Reviews in Control, 2020, 49: 145.

[ 18 ] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[M]. Cambridge: MIT Press,1998.

[ 19 ] 李怡瑾,唐昊,吕凯,等. 源荷不确定冷热电联供微网能量调度的建模与学习优化[J]. 控制理论与应用,2018,35(1): 56.

[ 20 ] ROCCHETTA R, BELLANI L, COMPARE M, et al. A reinforcement learning framework for optimal operation and maintenance of power grids [J]. Applied Energy,2019,241: 291.

[ 21 ] 彭刘阳,孙元章,徐箭,等. 基于深度强化学习的自适应不确定性经济调度[J]. 电力系统自动化, 2020,44(9): 33.

[ 22 ] FUJIMOTO S, VAN HOOF H, MEGER D. Addressing function approximation error in actor-critic methods [C]//International Conference on Machine Learning,2018.