

DOI: 10.19783/j.cnki.pspc.230381

# 融合图神经网络模型与强化学习的综合能源系统优化调度

王新<sup>1</sup>, 张良<sup>2</sup>, 任晓龙<sup>1</sup>, 曾逸舟<sup>2</sup>, 司恒斌<sup>1</sup>, 陈曦<sup>1</sup>, 杨乐<sup>1</sup>, 张志宏<sup>2</sup>

(1. 国网陕西省电力公司信息通信公司, 陕西 西安 710065; 2. 厦门大学, 福建 厦门 361005)

**摘要:** 随着人工智能技术特别是强化学习在能源优化调度领域的深入研究, 将系统状态表示为向量用于学习的模式, 其训练效率与信息利用率较低。针对这一问题, 提出了一种融合图神经网络模型与强化学习的综合能源系统优化调度方法。首先, 将电-热-气综合能源系统建模为图结构数据, 充分利用系统的拓扑信息。其次, 提出了基于图神经网络架构的强化学习模型, 使其可以充分利用图结构信息实现更快的训练速度, 获得更大的探索空间。最后, 将表示系统状态的图结构信息送入该模型进行训练, 算例仿真验证了该方法的训练效率与探索能力。

**关键词:** 电-热-气综合能源系统; 优化调度; 深度强化学习; 图神经网络模型

## Optimal scheduling of integrated energy systems by fusing a graph neural network model and reinforcement learning

WANG Xin<sup>1</sup>, ZHANG Liang<sup>2</sup>, REN Xiaolong<sup>1</sup>, ZENG Yizhou<sup>2</sup>, SI Hengbin<sup>1</sup>, CHEN Xi<sup>1</sup>, YANG Le<sup>1</sup>, ZHANG Zhihong<sup>2</sup>

(1. State Grid Shaanxi Information and Telecommunication Company, Xi'an 710065, China;

2. Xiamen University, Xiamen 361005, China)

**Abstract:** In in-depth research on artificial intelligence technology, especially reinforcement learning in the field of energy optimization scheduling, the training efficiency and information utilization rate of the system state expressed as a vector for the learning mode, is low. In response to this problem, this paper proposes an integrated energy system optimization scheduling method that integrates the graph neural network model and reinforcement learning. First, the electricity-heat-gas integrated energy system is modeled as graph structure data, making full use of the topological information of the system. Second, a reinforcement learning model based on graph neural network architecture is proposed, so that it can make full use of graph structure information to achieve a faster training speed and obtain a larger exploration space. Finally, the graph structure information representing the system state is sent to the model for training, and the example simulation verifies the training efficiency and exploration ability of the method.

This work is supported by the National Natural Science Foundation of China (No. 62176227 and No. U2066213).

**Key words:** electricity-heat-gas integrated energy system; optimal scheduling; deep reinforcement learning; graph neural network model

## 0 引言

随着当前能源需求增加与低碳经济的矛盾日益突出, 区域综合能源系统的建设与研究开始受到广泛的关注。与传统单一能源网络相比, 区域综合能源系统通过在多种环节的交互, 可以实现能源间的协同互补, 从而达到提高能源利用率、降低运行成

本、减少碳排放的目标。

电-热-气综合能源系统的优化调度作为多能流分析的核心问题之一, 对于节能减排、实现能源的充分利用具有重要意义<sup>[1-6]</sup>。文献[7]研究了电-热-气-冷系统耦合的多场景多目标的区域综合能源系统的优化调度方法。文献[8]从直接负荷控制的角度, 提出了电-热-气系统协调优化模型。由于现有多能流工作多为稳态模型, 文献[9]考虑热网和天然气网的动态特性来提高调度灵活性, 提出了包含直流电网的电-热-气耦合系统的线性优化模型。

近些年新能源电站的广泛铺设, 其出力的随机

**基金项目:** 国家自然科学基金项目资助(62176227, U2066213); 中央高校基本科研业务费项目资助(20720210047); 国家电网陕西省电力公司科技项目资助(SGSNXT00GCJS2200106)

性为优化工作带来了新的困难<sup>[10-11]</sup>。文献[12]从需求响应的角度, 提出电-气-热能源系统优化调度模型, 以提高新能源的消纳率。文献[13]针对风电出力与汽车充电负荷的不确定性, 建立随机模型并提取典型场景, 提出协同规划模型, 在一定程度上缓解了随机规划方法面对大量场景的求解困难问题。文献[14]基于(p,w)-norm 不确定集下的鲁棒线性优化理论, 将随机模型转化为确定性模型, 提出了直接负荷控制的鲁棒优化调度模型。由于鲁棒优化需考虑极端情况下的最优解, 可能会使调度结果过于保守。文献[15]提出了基于 Hausdorff 距离的电-热-气系统两阶段分布鲁棒优化模型, 克服传统鲁棒优化过于保守的局限性。

随着近些年机器学习理论的发展与广泛应用, 强化学习(reinforcement learning, RL)作为机器学习的重要分支, 在组合优化问题的求解工作上取得了突出的进展<sup>[16]</sup>, 并在电力系统优化调度<sup>[17]</sup>领域得到了广泛应用。其核心思想为通过探索与反馈机制来学习并得到优化问题的解, 广泛的探索可避免陷入局部最优解。此外, 相比于传统基于松弛或启发式的求解算法, 强化学习与深度学习的结合, 具有更高的自适应能力, 在完成模型训练后, 无须每次针对新的系统状态进行求解, 在面向大规模系统时具有更快的响应速度。在综合能源系统的最优经济调度问题上, 已开展了基于深度Q网络(deep Q-learning network, DQN)、深度确定性策略梯度(deep deterministic policy gradient, DDPG)等强化学习方法的研究<sup>[18-20]</sup>。但由于算法本身的局限性, 在训练效率与优化结果上并不十分突出。SAC(soft actor critic)<sup>[21]</sup>算法作为离线随机策略梯度算法, 在训练效率与探索能力上优势明显, 因此一经提出便受到了广泛关注与应用, 在电-热联合能源系统的协调优化<sup>[22-23]</sup>、电-气联合能源系统的优化调度<sup>[24]</sup>问题上取得了显著成效。

但现有基于强化学习模型的优化调度工作少有面向电-热-气综合能源系统的研究, 且上述所有工作将系统状态表示为向量作为输入进行训练, 忽略了系统本身的拓扑结构, 这与需要利用拓扑连接求解最优潮流的传统思想相违背。图神经网络(graph neural network, GNN)<sup>[25]</sup>作为近些年的热门研究领域, 能够充分利用系统的拓扑信息, 有效刻画节点间复杂的非线性关系, 在电力系统的无功优化问题上取得了相较于传统的多层感知机(multi-layer perceptron, MLP)更好的效果<sup>[26]</sup>。因此认为使用强化学习方法学习综合能源系统的最优调度策略, 需要充分利用系统的拓扑结构信息。

面向包含风力电站的电-热-气综合能源系统, 本文提出了一种基于 GNN 架构的 SAC 强化学习模型。首先, 介绍基于 GNN 的 SAC 模型理论与架构。然后, 搭建强化学习环境, 并介绍了将综合能源系统建模为图结构数据。最后进行了对比实验分析, 结果表明该方法相较于其他强化学习算法以及基于 MLP 的 SAC 算法, 在满足安全约束的前提下, 能够更快地收敛并得到更为经济的调度方案, 且在含储能单元的场景下表现良好。

## 1 基于图神经网络架构的强化学习模型

### 1.1 图神经网络模型

图结构数据可以定义为  $G=(V, E)$ , 其中  $V$  表示系统中的节点集,  $E$  表示系统中的边集。在神经网络模型训练中, 通过节点间的信息传递聚合进行节点的更新与表示<sup>[27]</sup>。

$$\mathbf{h}_i^k = \psi^k \left( \mathbf{h}_i^{k-1}, \sum_{j \in \mathcal{N}(i)} \phi^k(\mathbf{h}_i^{k-1}, \mathbf{h}_j^{k-1}, \mathbf{e}_{i,j}) \right) \quad (1)$$

式中:  $\mathbf{h}_i^k$  表示节点  $i$  在经过第  $k$  层神经网络后的向量表示;  $\mathcal{N}(i)$  表示节点  $i$  的邻域节点;  $\psi$ 、 $\phi$  表示不同的可微函数;  $\mathbf{e}_{i,j}$  为边的特征向量。

图神经网络模型在图表示学习中的一大局限是当堆叠过多的网络层数时, 会因为节点表示趋于一致而导致训练效果下降, 一般将其称为过平滑。为更好地利用图中信息并防止过平滑的出现, 本文采用注意力机制<sup>[28]</sup>来克服这一问题, 即在节点信息聚合时为邻域节点分配不同的权重, 并在模型训练中迭代更新权值, 具体为

$$\mathbf{h}_i^k = \alpha_{i,i} \mathbf{W} \mathbf{h}_i^{k-1} + \sum_{j \in \mathcal{N}(i)} \alpha_{i,j} \mathbf{W} \mathbf{h}_j^{k-1} \quad (2)$$

式中:  $\mathbf{W}$  表示神经网络参数矩阵, 来对节点特征进行线性变换;  $\alpha_{i,j}$  为注意力系数。

$$\alpha_{i,j} = \frac{\exp(\text{GELU}(\mathbf{a}^T [\mathbf{W} \mathbf{h}_i \parallel \mathbf{W} \mathbf{h}_j \parallel \mathbf{W}_e \mathbf{e}_{i,j}]))}{\sum_{k \in \mathcal{N}(i) \cup \{i\}} \exp(\text{GELU}(\mathbf{a}^T [\mathbf{W} \mathbf{h}_i \parallel \mathbf{W} \mathbf{h}_k \parallel \mathbf{W}_e \mathbf{e}_{i,k}]))} \quad (3)$$

式中: 向量  $\mathbf{a}$  为注意力网络的参数向量;  $\mathbf{W}_e$  为对边信息进行线性变换的参数矩阵; GELU 为激活函数; “ $\parallel$ ” 为向量连接符。

### 1.2 强化学习

强化学习有别于监督学习与无监督学习, 其核心思想是通过探索的方式进行学习, 通过反馈来调整策略, 并最终得到在当前环境下的最优解。具体来说, 智能体从环境中获取当前状态  $s_t$  并输出动作  $a_t$ , 该动作作用于环境获得对应的奖励  $r(s_t, a_t)$ 。智能体根据反馈的奖励值来调整策略及学习网络参数, 目标是获得最大的累积奖励, 如式(4)所示。

$$J = \sum_{t=0}^T \mathbb{E}[r(s_t, a_t)] \quad (4)$$

Actor-Critic(AC)算法作为被广泛使用的强化学习算法，分为两个相对独立但进行交互的网络：Actor 和 Critic。其基本架构如图 1 所示，Actor 通过策略网络  $\pi(a_t | s_t)$  给出当前状态  $s_t$  下的动作  $a_t$ ，Critic 基于状态  $s_t$  和奖励  $r$ ，通过价值网络  $Q$  给出对当前策略的评估，其中策略网络与价值网络的参数在模型训练中学习。

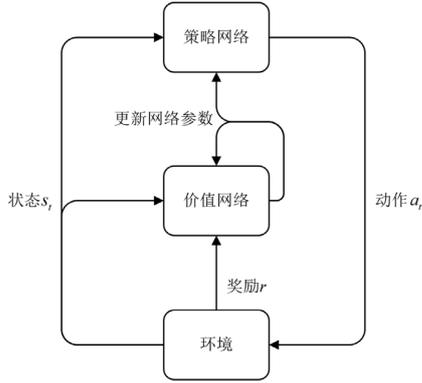


图 1 Actor-Critic 算法框架  
Fig. 1 Actor-Critic algorithm framework

为了提升算法的探索能力，SAC 算法在 AC 框架的基础上增加了熵项，目标是最大化奖励的同时最大化熵值，如式(5)所示。

$$J = \sum_{t=0}^T \mathbb{E}[r(s_t, a_t) + \eta \cdot \mathcal{H}(\pi(a_t | s_t))] \quad (5)$$

式中：熵项  $\mathcal{H}(\pi(a_t | s_t))$  表示在状态  $s_t$  下策略  $\pi(a_t | s_t)$  的熵值； $\eta$  为温度系数，调整熵项在累积奖励中的权重。增加熵项的目的是让策略随机化，使每一个动作的概率分布尽可能分散，对应的熵更大，从而保证了探索的随机性，扩大探索范围，避免陷入局部最优解。

在对策略进行评估时，使用 Soft Q 值函数及修正贝尔曼算子  $\mathcal{T}^\pi$ 。

$$\mathcal{T}^\pi Q(s_t, a_t) = r(s_t, a_t) + \gamma \mathbb{E}[V(s_{t+1})] \quad (6)$$

$$V(s_t) = \mathbb{E}_{a_t \sim \pi}[Q(s_t, a_t) - \log \pi(a_t | s_t)] \quad (7)$$

式中： $\gamma$  为折扣因子，用以调整智能体对短期奖励与长期奖励的重视程度； $V(s_t)$  为状态值函数，来对当前状态进行评估。

价值网络  $Q$  与策略网络  $\pi$  的参数分别通过最小化贝尔曼残差与最小化 KL 散度来更新。

$$J_Q = \mathbb{E} \left[ \frac{1}{2} (Q(s_t, a_t) - r(s_t, a_t) + \gamma \mathbb{E}[V(s_{t+1})])^2 \right] \quad (8)$$

$$J_\pi = \mathbb{E} \left[ \text{KL} \left( \pi(a_t | s_t) \left\| \frac{\exp(Q(s_t, a_t))}{Z(s_t)} \right. \right) \right] \quad (9)$$

式中， $Z(\cdot)$  是对分布进行归一化的函数。

### 1.3 基于图神经网络结构的 SAC 强化学习模型

本文提出的基于图神经网络结构的强化学习模型，其中 Actor 网络结构对比如图 2 所示。

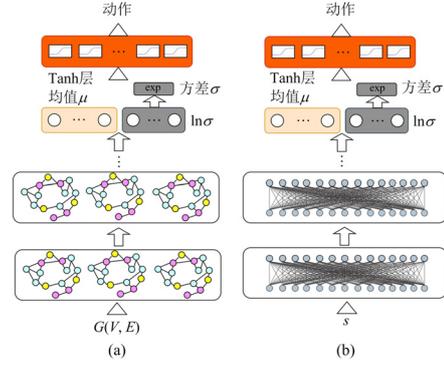


图 2 图神经网络结构和多层感知机结构的 Actor 网络  
Fig. 2 Actor network with GNN structure and MLP structure

如图 2(a)所示，基于图神经网络结构的 Actor 网络，输入为当前时刻  $t$  的状态图  $G(V, E)$ ，经过  $k$  层图神经网络，每一层的激活函数为 GELU，输出各动作的均值  $\mu$  与方差的对数  $\ln \sigma$ ，在对  $\ln \sigma$  进行指数变换后得到正态分布  $N(\mu, \sigma^2)$ 。采样并添加噪声后，通过 Tanh 层得到  $(-1, 1)$  内的值，并将其线性映射到动作范围内得到实际出力值。

## 2 面向综合能源系统的强化学习模型

### 2.1 环境模型

环境模型是与智能体进行交互的电-热-气综合能源系统模型，包含电力系统、热力系统、天然气系统与耦合系统。

#### 2.1.1 电力系统模型

交流电力系统潮流方程为

$$\begin{cases} P_i = U_i \sum_{j \in i} U_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \\ Q_i = U_i \sum_{j \in i} U_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \end{cases} \quad (10)$$

式中： $P_i$ 、 $Q_i$  分别表示节点  $i$  注入的有功功率与无功功率； $U_i$  表示节点  $i$  的电压幅值； $G_{ij}$  表示节点  $i$  与节点  $j$  之间的电导； $B_{ij}$  表示节点  $i$  与节点  $j$  之间的电纳； $\theta_{ij}$  表示节点  $i$  与节点  $j$  之间的相角差， $\theta_{ij} = \theta_i - \theta_j$ 。

对于系统中包含的风力电站，风电功率实时误差采用 Beta( $\alpha, \beta$ ) 分布<sup>[29]</sup>，其概率密度函数为

$$f(\Delta P, \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \Delta P^{\alpha-1} (1 - \Delta P)^{\beta-1} \quad (11)$$

式中:  $\Delta P$  为风电误差;  $\alpha$ 、 $\beta$  为 Beta 分布的形状参数,  $\alpha, \beta > 0$ 。 $\alpha$ 、 $\beta$  与真实风电历史数据的统计均值  $\mu_w$  和方差  $\sigma_w$  的关系为

$$\mu_w = \frac{\alpha}{\alpha + \beta} \quad (12)$$

$$\sigma_w^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (13)$$

### 2.1.2 热力系统模型

#### 1) 水力模型

水力模型由流量连续方程和回路压力方程构成。

$$\begin{cases} \mathbf{A}\mathbf{m} = \mathbf{m}_q \\ \mathbf{B}\mathbf{h}_t = 0 \end{cases} \quad (14)$$

式中:  $\mathbf{A}$  为节点-支路关联矩阵;  $\mathbf{m}$  为管道质量流量速率向量;  $\mathbf{m}_q$  为节点注入流量向量;  $\mathbf{B}$  为回路-支路关联矩阵;  $\mathbf{h}_t$  表示压头损失向量, 与管道的阻力系数和管道质量流量速率相关。

#### 2) 热力模型

热力模型包括节点功率模型、管道温度下降方程和介质混合方程。

$$\begin{cases} H_i = C_p m_{q,i} (T_i^{\text{in}} - T_i^{\text{out}}) \\ T_{j,i} = (T_{i,j} - T_e) e^{-\frac{\lambda L_{ij}}{C_p m_{ij}}} + T_e \\ \left( \sum_k^{|n_i|} m_{k,i} \right) T_i = \sum_k^{|n_i|} m_{k,i} T_{i,k} \end{cases} \quad (15)$$

式中:  $H_i$  表示节点  $i$  的热功率;  $C_p$  为水的比热容;  $m_{q,i}$  表示节点  $i$  的注入流量;  $T_i^{\text{in}}$ 、 $T_i^{\text{out}}$  分别为节点  $i$  的供水温度与回水温度;  $T_{i,j}$ 、 $T_{j,i}$  分别表示管道支路  $i-j$  的  $i$  端与  $j$  端的水温;  $T_e$  为外部环境温度;  $\lambda$  为导热系数;  $L_{ij}$  表示管道支路  $ij$  的长度;  $m_{ij}$  表示管道支路  $i-j$  的质量流量速率;  $m_{k,i}$  表示自节点  $k$  流至节点  $i$  的质量流量速率;  $T_{i,k}$  表示节点  $k$  流至节点  $i$  时的水温;  $T_i$  表示回水在节点  $i$  处混合后的水温;  $|n_i|$  表示所有流至节点  $i$  的节点总数。

### 2.1.3 天然气系统模型

天然气系统管道流量  $f_{ij}$  与节点压力  $p_i$ 、 $p_j$  间的关系为

$$\begin{cases} f_{ij} = \kappa_{ij} d_{ij} \sqrt{d_{ij} (p_i^2 - p_j^2)} \\ d_{ij} = \begin{cases} +1 & p_i > p_j \\ -1 & p_i < p_j \end{cases} \end{cases} \quad (16)$$

式中,  $\kappa_{ij}$  为管道常数。

由于天然气在传输过程中需要安装一定数量的压缩机以保证稳定运输, 对于包含压缩机的管道模型如图 3 所示。

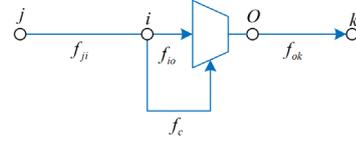


图 3 含燃气轮机驱动的压缩机的管道结构

Fig. 3 Piping structure with gas turbine driven compressor

对于含燃气压缩机的管道, 其模型<sup>[30]</sup>为

$$\begin{cases} f_c = \frac{k_c f_{io} T_{\text{gas}}}{q_{\text{gas}}} (k_c^{\frac{\tau-1}{\tau}} - 1) \\ f_{ji} = f_{io} + f_c \\ f_{ok} = f_{io} = \kappa_{ok} \sqrt{p_o^2 - p_k^2} \\ f_{ji} = \kappa_{ji} \sqrt{p_j^2 - p_i^2} \end{cases} \quad (17)$$

式中:  $f_c$  为燃气轮机的耗气量;  $p_j$ 、 $p_i$ 、 $p_o$ 、 $p_k$  分别为图 3 中 4 个节点处的压力;  $k_c$  为压缩比,  $k_c = p_o / p_i$ ;  $f_{ji}$ 、 $f_{ok}$ 、 $f_{io}$  分别为管道  $j-i$  的气流量、管道  $o-k$  的气流量与通过压缩机的气流量;  $T_{\text{gas}}$  为天然气温度;  $q_{\text{gas}}$  为天然气热值;  $\kappa_{ok}$ 、 $\kappa_{ji}$  分别为管道  $j-i$ 、 $o-k$  的管道常数;  $\tau$  为多变指数。

### 2.1.4 电-热-气系统耦合环节

针对电-热-气综合能源系统的耦合环节, 本文考虑可以同时发电和供热的燃气 CHP 机组来满足电力系统与热力系统的负荷需求。CHP 机组对应的多边形出力模型为

$$\min \{ P_{\min}^{\text{CHP}} - c_3 H^{\text{CHP}}, c_1 + c_2 H^{\text{CHP}} \leq P^{\text{CHP}} \leq P_{\max}^{\text{CHP}} - c_3 H^{\text{CHP}} \} \quad (18)$$

式中:  $P^{\text{CHP}}$ 、 $H^{\text{CHP}}$  分别表示 CHP 机组的电出力和热出力;  $P_{\max}^{\text{CHP}}$ 、 $P_{\min}^{\text{CHP}}$  分别表示 CHP 机组的电出力的上下限;  $c_1$ 、 $c_2$ 、 $c_3$  为多边形区域系数。

## 2.2 动作空间

智能体输出的动作共包括: 火力电站有功出力和 CHP 机组的电、热出力以及供热站热出力, 风电消纳系数以及天然气供气站的供气量。

$$A = \{ P_i^G, P_i^{\text{CHP}}, H_i^{\text{CHP}}, H_i^G, \alpha_i^W, f_i \} \quad (19)$$

各动作范围分别为

$$\begin{cases} P_{i,\min}^G \leq P_i^G \leq P_{i,\max}^G \\ P_{i,\min}^{\text{CHP}} \leq P_i^{\text{CHP}} \leq P_{i,\max}^{\text{CHP}} \\ H_{i,\min}^{\text{CHP}} \leq H_i^{\text{CHP}} \leq H_{i,\max}^{\text{CHP}} \\ H_{i,\min}^G \leq H_i^G \leq H_{i,\max}^G \\ 0 \leq \alpha_i^W \leq 1 \\ f_{i,\min} \leq f_i \leq f_{i,\max} \end{cases} \quad (20)$$

式中:  $P_i^G$  为火力电站  $i$  的有功出力;  $P_i^{\text{CHP}}$ 、 $H_i^{\text{CHP}}$  分别为 CHP 机组  $i$  的电、热出力;  $H_i^G$  为供热站热出力;  $\alpha_i^W$  为风电消纳系数;  $f_i$  为天然气供气站的供气量;  $P_{i,\max}^G$ 、 $P_{i,\min}^G$  分别为火力电站  $i$  有功出力的上下限;  $P_{i,\max}^{\text{CHP}}$ 、 $P_{i,\min}^{\text{CHP}}$  分别为 CHP 机组  $i$  电出力的上下限;  $H_{i,\max}^{\text{CHP}}$ 、 $H_{i,\min}^{\text{CHP}}$  分别表示 CHP 机组  $i$  热出力的上下限;  $H_{i,\max}^G$ 、 $H_{i,\min}^G$  分别为供热站  $i$  热出力的上、下限;  $f_{i,\max}$ 、 $f_{i,\min}$  分别为供气站  $i$  在一定时段内供气量的上下限。

### 2.3 状态空间

把电-热-气综合能源系统建模为图  $G(V, E)$ , 系统状态通过图中节点特征与边特征来反映。电力系统的节点特征与边特征分别为

$$\begin{cases} v_i^{\text{ele}} = (P_i^L) \\ e_{ij}^{\text{ele}} = (G_{ij}, B_{ij}) \end{cases} \quad (21)$$

式中,  $P_i^L$  为节点  $i$  的电负荷。

热力系统的节点特征与边特征分别为

$$\begin{cases} v_i^{\text{heat}} = (H_i^L) \\ e_{ij}^{\text{heat}} = (L_{ij}, m_{ij}) \end{cases} \quad (22)$$

式中,  $H_i^L$  为节点  $i$  的热负荷。

天然气系统的节点特征与边特征分别为

$$\begin{cases} v_i^{\text{gas}} = (f_i^L) \\ e_{ij}^{\text{gas}} = (l_{ij}, \kappa_{ij}) \end{cases} \quad (23)$$

式中:  $f_i^L$  为节点  $i$  的气负荷;  $l_{ij}$  为节点  $i$ 、 $j$  之间的管道长度。

### 2.4 奖励函数

对于电-热-气综合能源系统优化调度任务, 以最小化运行成本、不违背约束条件为目标。

#### 2.4.1 运行成本

##### 1) 火力电站运行成本

$$F_1 = \sum_{t=1}^T \sum_{i=1}^{|N_p|} (C_2^p P_{i,t}^2 + C_1^p P_{i,t} + C_0^p) \quad (24)$$

式中:  $T$  为总运行时间;  $|N_p|$  为火力发电站数量;  $P_{i,t}$  表示火力电站  $i$  在  $t$  时刻的有功出力;  $C_0^p$ 、 $C_1^p$ 、 $C_2^p$  为火力供电机组的耗量特性曲线参数。

##### 2) 供热站运行成本

$$F_2 = \sum_{t=1}^T \sum_{i=1}^{|N_H|} (C_2^H H_{i,t}^2 + C_1^H H_{i,t} + C_0^H) \quad (25)$$

式中:  $|N_H|$  为供热站数量;  $H_{i,t}$  表示火力电站  $i$  在  $t$  时刻的有功出力;  $C_2^H$ 、 $C_1^H$ 、 $C_0^H$  为火力供热机组的耗量特性曲线参数。

##### 3) CHP 机组运行成本

$$F_3 = \sum_{t=1}^T \sum_{i=1}^{|N_{\text{CHP}}|} (C_0^{\text{CHP}} + C_1^{\text{CHP}} H_{i,t}^{\text{CHP}} + C_2^{\text{CHP}} P_{i,t}^{\text{CHP}} + C_3^{\text{CHP}} (H_{i,t}^{\text{CHP}})^2 + C_4^{\text{CHP}} (P_{i,t}^{\text{CHP}})^2 + C_5^{\text{CHP}} H_{i,t}^{\text{CHP}} P_{i,t}^{\text{CHP}}) \quad (26)$$

式中:  $|N_{\text{CHP}}|$  为 CHP 机组数量;  $P_{i,t}^{\text{CHP}}$ 、 $H_{i,t}^{\text{CHP}}$  分别表示 CHP 机组  $i$  在  $t$  时刻的电出力和热出力;  $C_0^{\text{CHP}} \sim C_5^{\text{CHP}}$  为 CHP 机组的耗量特性曲线参数。

##### 4) 天然气成本

$$F_4 = \sum_{t=1}^T \sum_{i=1}^{|N_{\text{gas}}|} C^{\text{Gas}} f_{i,t} \quad (27)$$

式中:  $|N_{\text{gas}}|$  为供气站数量;  $C^{\text{Gas}}$  为单位天然气价格;  $f_{i,t}$  为供气站在时刻  $t$  的供气量。

#### 2.4.2 约束条件

##### 1) 安全约束

电-热-气综合能源系统的稳定运行需要满足的安全约束条件有: 电力系统的电压约束、相角差约束和线路传输约束; 热力系统的节点温度与管道流量约束; 天然气系统的节点压力约束。

$$\begin{cases} U_{i,\min} \leq U_i \leq U_{i,\max} \\ -\theta_{ij}^A \leq \theta_{ij} \leq \theta_{ij}^A \\ |U_i^2 G_{ij} - U_i U_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij})| \leq P_l \\ T_{i,\min} \leq T_i \leq T_{i,\max} \\ m_{ij,\min} \leq m_{ij} \leq m_{ij,\max} \\ p_{i,\min} \leq p_i \leq p_{i,\max} \end{cases} \quad (28)$$

式中:  $U_{i,\max}$ 、 $U_{i,\min}$  分别为节点  $i$  的电压幅值的上下限;  $\theta_{ij}^A$  为相角差上限;  $P_l$  为线路传输功率上限;  $T_{i,\max}$ 、 $T_{i,\min}$  分别为节点  $i$  的供水温度的上下限;  $m_{ij,\max}$ 、 $m_{ij,\min}$  分别为管路  $ij$  的供水流量的上下限;  $p_{i,\max}$ 、 $p_{i,\min}$  分别为节点  $i$  的压力上下限。

##### 2) 爬坡约束

爬坡约束是指前后时刻动作的差值不能超过一定范围, 包括电力爬坡上下限和热力爬坡上下限。

$$\begin{cases} |P_{i,t}^G - P_{i,t-1}^G| \leq \Delta P_{i,\max}^G \\ |P_{i,t}^{\text{CHP}} - P_{i,t-1}^{\text{CHP}}| \leq \Delta P_{i,\max}^{\text{CHP}} \\ |H_{i,t}^{\text{CHP}} - H_{i,t-1}^{\text{CHP}}| \leq \Delta H_{i,\max}^{\text{CHP}} \\ |H_{i,t}^G - H_{i,t-1}^G| \leq \Delta H_{i,\max}^G \end{cases} \quad (29)$$

式中:  $\Delta P_{i,\max}^G$  为火力电站单位时刻电出力的爬坡上限;  $\Delta P_{i,\max}^{\text{CHP}}$ 、 $\Delta H_{i,\max}^{\text{CHP}}$  分别为 CHP 机组单位时刻电出力与热出力的爬坡上限;  $\Delta H_{i,\max}^G$  为供热站单位时刻热出力的爬坡上限。

综上, 奖励值包含系统运行成本与违反约束惩罚, 由于系统调度的目标是追求成本最小化, 而强化学习追求奖励最大化, 因此需要对奖励函数取负值。

$$r = -(F_1 + F_2 + F_3 + F_4 + \sum_{i=1}^{10} \delta_i |\cdot|) \quad (30)$$

式中:  $\delta_i$  为惩罚因子;  $|\cdot|$  为式(28)和式(29)中所列的 10 项安全约束与爬坡约束, 为保证训练结果满足约束条件, 一般设置较大的惩罚因子, 当满足约束时, 惩罚项为 0。

### 3 算法验证与分析

#### 3.1 算例分析

本文以修改融合的 6-6-6 节点的电-热-气综合能源系统算例<sup>[31]</sup>验证所提算法的训练效率与优化调度能力。系统结构如图 4 所示。

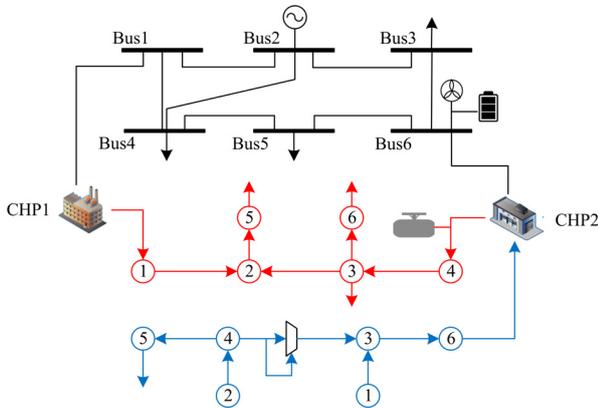


图 4 修改的电-热-气综合能源系统结构

Fig. 4 Structure of modified electricity-heat-gas integrated energy system

如图 4 所示, 黑色部分为电力系统, 节点 2 连接火力发电站; 节点 6 连接风力发电站; 红色部分为热力系统; 蓝色部分为天然气系统。CHP1 为燃煤热电联产机组, CHP2 为燃气热电联产机组。

为分析算法对于含储能场景的优化调度能力, 设置 3 种场景进行分析。

场景一: 无储能;

场景二: 含电储能, 蓄电池与电力系统节点 6 上的风电场相连;

场景三: 含热储能, 储热罐与热力系统节点 4 上的 CHP 机组相连。

#### 3.2 算法训练效率分析

在该算例场景一下, 使用网络层数与神经元个数相同的基于 MLP 架构的经典算法 SAC、DDPG 和双延迟深度确定性策略梯度算法(twin delayed DDPG, TD3)进行对比试验, 训练过程中的奖励值如

图 5 所示。

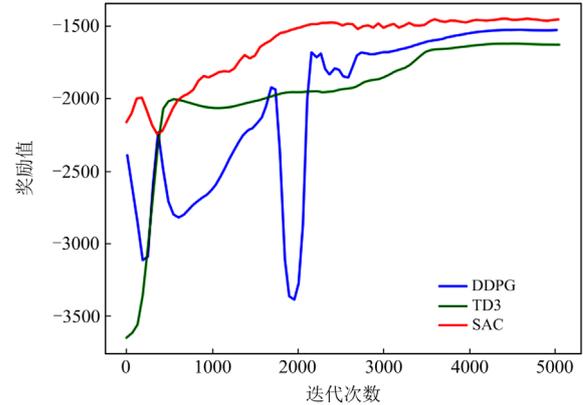


图 5 基于 MLP 架构的强化学习算法对比

Fig. 5 Comparison of RL algorithms based on MLP architecture

从图 5 可以看出, SAC 算法在经过 3500 轮左右的训练后收敛, 相较于其他强化学习算法, 在训练速度与优化结果上更具优势。为验证本文提出的基于 GNN 架构的强化学习算法, 在相同的网络参数设定下, 使用基于 GNN 与 MLP 架构的 SAC 算法进行了对比试验, 其中 Actor 网络的层数为 3, 各层神经元数量均为 96。Critic 网络的层数为 3, 各层神经元数量为 96, 各层激活函数为 GELU, 经验池数量为 24 000。通过 Adam 优化器来自动调整学习率。实验结果如图 6 所示。

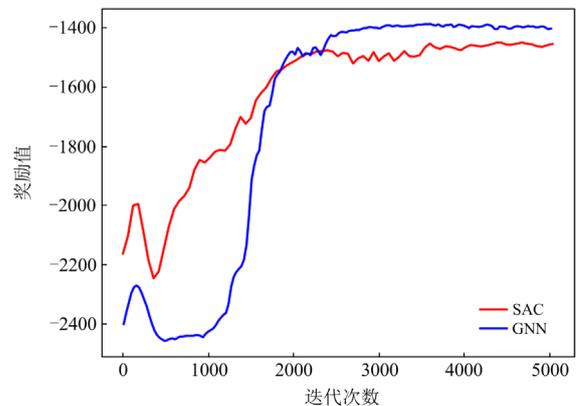


图 6 基于 GNN 与 MLP 架构的 SAC 算法对比

Fig. 6 Comparison of SAC algorithms based on GNN and MLP architectures

从图 6 中可以看出, 基于 GNN 架构的奖励值曲线经过 2500 轮左右的训练后收敛, 和在上一轮对比试验中的基于 MLP 架构 SAC 算法的奖励值曲线相比, GNN 架构的 SAC 算法在收敛速度与优化结果上优势明显。上述各方法在收敛后, 对应的系统运行成本如表 1 所示。可以看出, 基于 GNN 架构的 SAC 算法训练得到了最低成本的调度策略。

表 1 不同算法运行成本对比

算法	平均每小时成本/元
SAC_GNN	12 640
SAC_MLP	13 226
DDPG	13 920
TD3	15 582

### 3.3 电力系统调度结果分析

为分析算法在含储能场景下的调度能力，对场景一和场景二下的电力系统调度结果进行了对比，结果如图 7 所示。其中蓄电池的充电量显示在横坐标下方，放电时同其他出力源一样堆叠在横坐标上方。

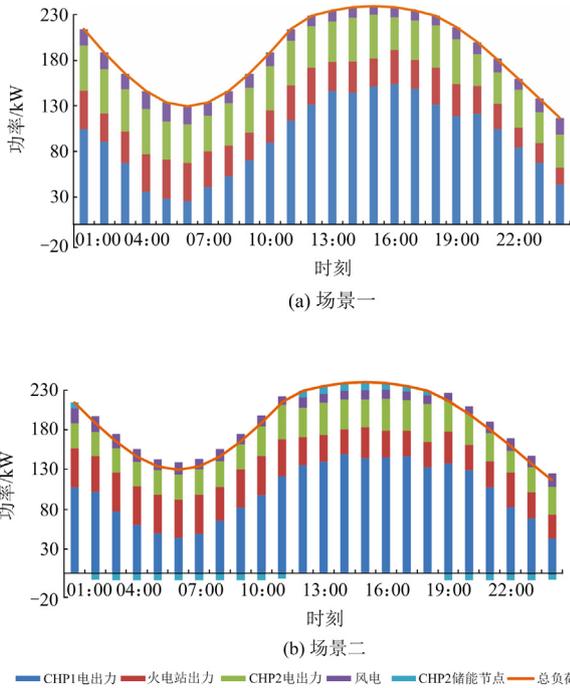


图 7 场景一与场景二的电力系统调度结果

Fig. 7 Power system scheduling results in scenarios 1&2

从图 7 中可以看出，各场景下总出力与总负荷一致，满足功率平衡条件。在场景一下，各出力节点满足出力范围约束，且由装机容量更大且调控更为灵活的 CHP 机组满足负荷缺口，CHP1 机组各时刻出力均满足爬坡约束。在场景二中，蓄电池在负荷低谷时段进行充电并在负荷高峰时段进行放电，从上下图对比来看，含蓄电池场景的调度方案实现了一定程度上的“削峰填谷”，各节点出力分布更为平稳。验证了算法在复杂系统中的调度能力以及对含储能单元的优化效果。

### 3.4 热力系统调度结果分析

对于热力系统，对场景一与场景三下的调度结果进行了对比，如图 8 所示。储热罐在蓄热时，蓄热量表示在横坐标下方，放热时同其他出力源一样堆叠在横坐标上方。

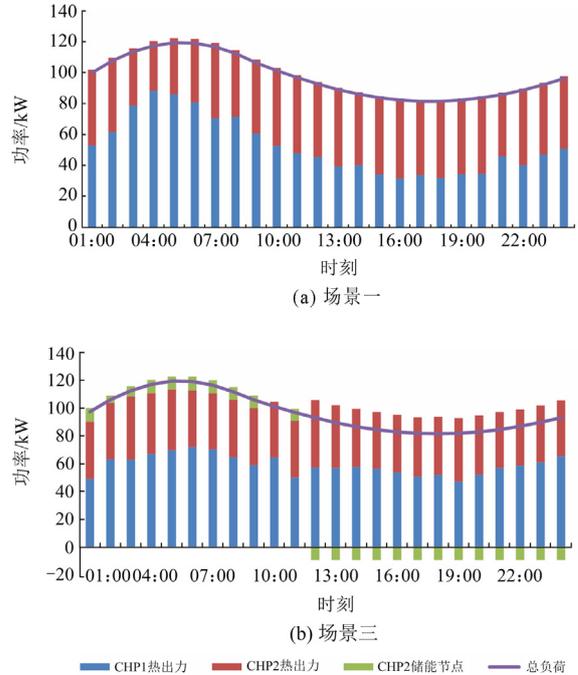


图 8 场景一与场景三的热力系统调度结果

Fig. 8 Heat system scheduling results in scenarios 1&3

从图 8 各时段的热力系统出力结果可以看到，总出力与负荷曲线基本吻合。各出力源均满足出力范围约束与爬坡约束。在场景三中，储热罐的设置使得系统可以在白天负荷低谷期进行蓄热，并在凌晨进行放热，有效实现了“削峰填谷”的功能。从 CHP1 和 CHP2 的出力来看，相较于场景一，各时刻的出力分布平稳，两者之和基本稳定在一条直线上，特别是 CHP1 的出力，场景一下的峰谷差异在 45 kW 左右，场景三下的峰谷差异只有不到 20 kW。验证了算法对热力系统的调度能力，说明可以通过调控储热罐的储/放来实现系统的平稳运行。

对于同时含蓄电池与储热罐的场景，由于其优化结果与只含一种储能单元的优化结果近似，故不再分析。

## 4 结论

本文将综合能源系统建模为图用于训练，并提出了基于 GNN 架构的 SAC 强化学习优化调度算法，从而充分利用图结构信息实现更为有效的探索学习。算例仿真实验展示了该算法的训练效率与有效

性。与基于 MLP 架构的强化学习算法相比, 所提方法具有更快的收敛速度, 得到了更好的优化结果。

### 参考文献

- [1] 孙宏斌, 潘昭光, 郭庆来. 多能流能量管理研究: 挑战与展望[J]. 电力系统自动化, 2016, 40(15): 1-8, 16. SUN Hongbin, PAN Zhaoguang, GUO Qinglai. Energy management for multi-energy flow: challenges and prospects[J]. Automation of Electric Power Systems, 2016, 40(15): 1-8, 16.
- [2] 张瑶, 王傲寒, 张宏. 中国智能电网发展综述[J]. 电力系统保护与控制, 2021, 49(5): 180-187. ZHANG Yao, WANG Aohan, ZHANG Hong. Overview of smart grid development in China[J]. Power System Protection and Control, 2021, 49(5): 180-187.
- [3] 朱浩昊, 朱继忠, 李盛林, 等. 电-热综合能源系统优化调度综述[J]. 全球能源互联网, 2022, 5(4): 383-397. ZHU Haohao, ZHU Jizhong, LI Shenglin, et al. Review of optimal scheduling of integrated electricity and heat systems[J]. Journal of Global Energy Interconnection, 2022, 5(4): 383-397.
- [4] 杨秀, 柴梓轩, 刘方, 等. 考虑热动态和碳交易的电-气-热综合能源系统协调调度[J]. 电测与仪表, 2021, 58(11): 49-58. YANG Xiu, CHAI Zixuan, LIU Fang, et al. Coordinated scheduling of electricity-gas-heat integrated energy system considering thermal dynamics and carbon trading[J]. Electrical Measurement & Instrumentation, 2021, 58(11): 49-58.
- [5] 田丰, 贾燕冰, 任海泉, 等. 计及用户行为及满意度的电-气综合能源系统优化调度[J]. 电测与仪表, 2021, 58(5): 31-38. TIAN Feng, JIA Yanbing, REN Haiquan, et al. Optimal dispatch of electricity-gas integrated energy system considering user behavior and satisfaction[J]. Electrical Measurement & Instrumentation, 2021, 58(5): 31-38.
- [6] 马喜平, 沈渭程, 甄文喜, 等. 基于低碳目标的电气综合能源系统优化调度策略研究[J]. 电网与清洁能源, 2021, 37(12): 116-122. MA Xiping, SHEN Weicheng, ZHEN Wenxi, et al. A study on the optimal scheduling strategy of electric-gas-thermal integrated energy system based on low carbon target[J]. Power System and Clean Energy, 2021, 37(12): 116-122.
- [7] 钟永洁, 孙永辉, 谢东亮, 等. 含电-热-气-冷子系统的区域综合能源系统多场景优化调度[J]. 电力系统自动化, 2019, 43(12): 76-84. ZHONG Yongjie, SUN Yonghui, XIE Dongliang, et al. Multi-scenario optimal dispatch of regional integrated energy system with power-heating-gas-cooling subsystems[J]. Automation of Electric Power Systems, 2019, 43(12): 76-84.
- [8] 艾欣, 陈政琦, 孙英云, 等. 基于需求响应的电-热-气耦合系统综合直接负荷控制协调优化研究[J]. 电网技术, 2019, 43(4): 1160-1171. AI Xin, CHEN Zhengqi, SUN Yingyun, et al. Study on integrated DLC coordination optimization of electric-thermal-gas coupling system considering demand response[J]. Power System Technology, 2019, 43(4): 1160-1171.
- [9] 陈瑜玮, 孙宏斌, 郭庆来. 综合能源系统分析的统一能路理论(五): 电-热-气耦合系统优化调度[J]. 中国电机工程学报, 2020, 40(24): 7928-7937. CHEN Yuwei, SUN Hongbin, GUO Qinglai. Energy circuit theory of integrated energy system analysis (V): integrated electricity-heat-gas dispatch[J]. Proceedings of the CSEE, 2020, 40(24): 7928-7937.
- [10] 张沈习, 王丹阳, 程浩忠, 等. 双碳目标下低碳综合能源系统规划关键技术及挑战[J]. 电力系统自动化, 2022, 46(8): 189-207. ZHANG Shenxi, WANG Danyang, CHENG Haozhong, et al. Key technologies and challenges of low-carbon integrated energy system planning for carbon emission peak and carbon neutrality[J]. Automation of Electric Power Systems, 2022, 46(8): 189-207.
- [11] 陈忠华, 高振宇, 陈嘉敏, 等. 考虑不确定性因素的综合能源系统协同规划研究[J]. 电力系统保护与控制, 2021, 49(8): 32-40. CHEN Zhonghua, GAO Zhenyu, CHEN Jiamin, et al. Research on cooperative planning of an integrated energy system considering uncertainty[J]. Power System Protection and Control, 2021, 49(8): 32-40.
- [12] 张涛, 郭玥彤, 李逸鸿, 等. 计及电气热综合需求响应的区域综合能源系统优化调度[J]. 电力系统保护与控制, 2021, 49(1): 52-61. ZHANG Tao, GUO Yuetong, LI Yihong, et al. Optimization scheduling of regional integrated energy systems based on electric-thermal-gas integrated demand response[J]. Power System Protection and Control, 2021, 49(1): 52-61.
- [13] 陈忠华, 高振宇, 陈嘉敏, 等. 考虑不确定性因素的综合能源系统协同规划研究[J]. 电力系统保护与控制, 2021, 49(8): 32-40. CHEN Zhonghua, GAO Zhenyu, CHEN Jiamin, et al. Research on cooperative planning of an integrated energy system considering uncertainty[J]. Power System Protection and Control, 2021, 49(8): 32-40.
- [14] 朱兰, 牛培源, 唐陇军, 等. 考虑直接负荷控制不确定性的微能源网鲁棒优化运行[J]. 电网技术, 2020, 44(4): 1400-1413. ZHU Lan, NIU Peiyuan, TANG Longjun, et al. Robust optimal operation for micro-energy grid considering uncertainties of direct load control[J]. Power System Technology, 2020, 44(4): 1400-1413.
- [15] 陈光宇, 张子祥, 李庆, 等. 基于 Hausdorff 距离的电-

- 热-气综合能源系统分布鲁棒优化调度[J]. 电网技术, 2022, 46(12): 4906-4918.
- CHEN Guangyu, ZHANG Zixiang, LI Qing, et al. Distributed robust optimal scheduling of electric thermal gas integrated energy system based on Hausdorff distance[J]. Power System Technology, 2022, 46(12): 4906-4918.
- [16] MAZYAVKINA N, SVIRIDOV S, IVANOV S, et al. Reinforcement learning for combinatorial optimization: a survey[J]. Computers & Operations Research, 2021, 134(1): 105400.
- [17] ZHANG Zidong, ZHANG Dongxia, QIU R C. Deep reinforcement learning for power system applications: an overview[J]. CSEE Journal of Power and Energy Systems, 2020, 6(1): 213-225.
- [18] 陈明昊, 孙毅, 胡亚杰, 等. 基于纵向联邦强化学习的居民社区综合能源系统协同训练与优化管理方法[J]. 中国电机工程学报, 2022, 42(15): 5535-5550.
- CHEN Minghao, SUN Yi, HU Yajie, et al. The collaborative training and management-optimized method for residential integrated energy system based on vertical federated reinforcement learning[J]. Proceedings of the CSEE, 2022, 42(15): 5535-5550.
- [19] 陈明昊, 孙毅, 谢志远. 基于双层深度强化学习的园区综合能源系统多时间尺度优化管理[J]. 电工技术学报, 2023, 38(7): 1864-1881.
- CHEN Minghao, SUN Yi, XIE Zhiyuan. The multi-time-scale management optimization method for park integrated energy system based on the bi-layer deep reinforcement learning[J]. Transactions of China Electrotechnical Society, 2023, 38(7): 1864-1881.
- [20] 王新迎, 赵琦, 赵黎媛, 等. 基于深度 Q 学习的电热综合能源系统能量管理[J]. 电力建设, 2021, 42(3): 10-18.
- WANG Xinying, ZHAO Qi, ZHAO Liyuan, et al. Energy management approach for integrated electricity-heat energy system based on deep Q-Learning network[J]. Electric Power Construction, 2021, 42(3): 10-18.
- [21] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[C]// International Conference on Machine Learning, July 10-15, 2018, Stockholm, Sweden: 1861-1870.
- [22] 刘雨, 董雷, 王春斐, 等. 基于柔性行动器-评判器的电-热综合能源系统协调优化[J]. 智能电网, 2021, 11(2): 107-117.
- LIU Yu, DONG Lei, WANG Chunfei, et al. Coordinated optimization of integrated electricity-heat energy system based on soft actor-critic[J]. Smart Grid, 2021, 11(2): 107-117.
- [23] 杨照, 黄少伟, 陈颖. 基于多智能体强化学习的多园区综合能源系统协同优化运行研究[J]. 电工电能新技术, 2021, 40(8): 1-10.
- YANG Zhao, HUANG Shaowei, CHEN Ying. Research on cooperative optimal operation of multi-park integrated energy system based on multi agent reinforcement learning[J]. Advanced Technology of Electrical Engineering and Energy, 2021, 40(8): 1-10.
- [24] 乔骥, 王新迎, 张擎, 等. 基于柔性行动器-评判器深度强化学习的电-气综合能源系统优化调度[J]. 中国电机工程学报, 2021, 41(3): 819-833.
- QIAO Ji, WANG Xinying, ZHANG Qing, et al. Optimal dispatch of integrated electricity-gas system with soft actor-critic deep reinforcement learning[J]. Proceedings of the CSEE, 2021, 41(3): 819-833.
- [25] WU Z, PAN S, CHEN F, et al. A comprehensive survey on graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1): 4-24.
- [26] 廖文龙, 于贇, 王煜森, 等. 基于图卷积网络的配电网无功优化[J]. 电网技术, 2021, 45(6): 2150-2160.
- LIAO Wenlong, YU Yun, WANG Yusen, et al. Reactive power optimization of distribution network based on graph convolutional network[J]. Power System Technology, 2021, 45(6): 2150-2160.
- [27] GILMER J, SCHOENHOLZ S S, RILEY P F, et al. Neural message passing for quantum chemistry[C]// International Conference on Machine Learning, August 6-11, 2017, Sydney, Australia: 1263-1272.
- [28] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[J]. arXiv preprint arXiv: 1710.10903, 2017.
- [29] 赵文猛, 刘明波, 周保荣, 等. 含风电接入多区域电力系统的分散式随机动态经济调度方法[J]. 中国电机工程学报, 2017, 37(24): 7087-7098.
- ZHAO Wenmeng, LIU Mingbo, ZHOU Baorong, et al. Decentralized stochastic dynamic economic dispatch for multi-area power systems with wind power integrated[J]. Proceedings of the CSEE, 2017, 37(24): 7087-7098.
- [30] 王英瑞, 曾博, 郭经, 等. 电-热-气综合能源系统多能流计算方法[J]. 电网技术, 2016, 40(10): 2942-2951.
- WANG Yingrui, ZENG Bo, GUO Jing, et al. Multi-energy flow calculation method for integrated energy system containing electricity, heat and gas[J]. Power System Technology, 2016, 40(10): 2942-2951.
- [31] IIT. Index of Data Illinois Institute of Technology [EB/OL]. <http://motor.ece.iit.edu/data/>.

收稿日期: 2023-04-08; 修回日期: 2023-06-13

作者简介:

王新(1979—), 男, 硕士, 高级工程师, 研究方向为能源互联网优化调度;

张志宏(1986—), 男, 通信作者, 博士, 副教授, 博士生导师, 研究方向为能源互联网与大数据分析技术. E-mail: zhihong@xmu.edu.cn

(编辑 魏小丽)