

DOI: 10.19783/j.cnki.pspc.230828

考虑运行状态信息的综合能源系统图强化学习优化调度

吕金玲, 王小君, 窦嘉铭, 孙庆凯, 刘 璿, 和敬涵

(北京交通大学电气工程学院, 北京 100044)

摘要: “双碳”背景下, 异质能源的耦合加剧迫使综合能源系统(integrated energy system, IES)拓扑朝着更复杂、更灵活的方向不断演变。然而, 现有优化调度方法对非欧网络拓扑知识及其异质潮流约束考虑不足。针对这一问题, 提出一种基于图强化学习的综合能源系统优化调度方法。首先, 基于图理论在保证节点多样状态的情况下, 将异质能源网络拓扑转换为网络图模型。其次, 通过建立基于真实图映射的状态-动作-奖励的框架, 利用图强化学习的方法学习图模型的非欧拓扑信息, 将异质潮流知识加入系统节点运行状态, 从而实现 IES 的安全优化调度。最后, 利用某工业园区的真实数据进行仿真验证, 所提方法相对于传统方法有效缓解了节点电压越限的问题。结果表明, 所提方法能够在考虑 IES 真实拓扑运行状态信息和异质潮流安全的情况下实现 IES 的优化调度。

关键词: 综合能源系统; 优化调度; 图强化学习; 运行状态; 安全约束

Optimal dispatch of an integrated energy system based on graph reinforcement learning considering operation state information

LÜ Jinling, WANG Xiaojun, DOU Jiaming, SUN Qingkai, LIU Zhao, HE Jinghan

(School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: With the dual-carbon goals, the intensified coupling of heterogeneous energy sources is driving integrated energy system (IES) topology to evolve towards greater complexity and flexibility. However, the existing optimal scheduling methods do not sufficiently consider the knowledge of non-European network topology and its heterogeneous power flow constraints. To address this issue, this paper proposes an optimal dispatch method based on graph reinforcement learning. First, by guaranteeing diverse node states, the heterogeneous energy network topology connections are converted into network graph models using graph theory. Second, a state-action-reward framework based on real graph mapping is established, and the graph reinforcement learning method is employed to learn the non-Euclidean knowledge and heterogeneous flow constraints brought by the graph model, thus achieving safe and optimal scheduling of IES. Finally, the real data of an industrial park is used for simulation verification. Compared with the traditional method, the proposed method effectively alleviates the node voltage over-limit problem. The results indicate that the proposed method can achieve optimal dispatch of IES while considering the actual topology operation state information and heterogeneous flow safety.

This work is supported by the National Natural Science Foundation of China (No. 51977005).

Key words: integrated energy system; optimal dispatch; graph reinforcement learning; operation state; security constraints

0 引言

在全球化石能源大量消耗和能源结构迫切转型的趋势下, 综合能源系统(integrated energy system, IES)作为一种经济高效、绿色集约的能源利用方案, 其不同能源之间的互联互动可实现能量的梯级利用,

逐渐受到社会各界的广泛关注^[1-3]。优化调度作为 IES 运营的基础, 其结果的合理性对于 IES 的规划设计与运行优化具有重要意义^[4-7]。

近年来, 随着对 IES 优化调度问题的深入研究, 大量学者提出了诸如数学解析法^[8-10]、启发式算法^[11-12]等不同求解方法, 且已在不同类型和规模的 IES 中取得了显著的成果。其中, 文献[13]采用混合整数线性规划方法解决了天然气组合能源系统

基金项目: 国家自然科学基金项目资助(51977005)

的安全约束联合扩展规划问题；文献[14]利用二阶锥规划方法求解考虑双向能量流的电-气 IES，以最大限度地降低 IES 的日前运营成本；文献[15]通过模糊数学方法将热电联供型微网系统的多目标优化模型转换为单目标模型，解决了电动汽车和分布式电源接入微电网场景下 IES 的优化调度问题。然而，这类解决方案在发展中已经略显疲态，主要存在以下两方面的限制：1) 数学解析法通常依赖于精确的数学模型和参数，导致其扩展性较低；2) 启发式算法求解效率低，容易陷入局部最优解，影响实时调度的效率。随着 IES 内部供能与转换设备类型不断增多、产业规模不断扩大，致使上述传统方法在求解具有复杂网络拓扑结构信息的实际 IES 系统时略显疲态，存在计算复杂度高、陷入局部最优解、模型适应性不足等问题。

目前，人工智能技术在信息化产业领域的逐渐兴起，为其在能源行业的优化运行提供了一种新的思路，部分学者尝试采用强化学习^[16]求解 IES 运行调度问题。其中，文献[17]提出一种利用部分模型信息的微电网优化调度安全策略学习方法，解决了源荷不确定性下的安全问题。文献[18]基于近端策略优化算法重新设计奖励函数，以引导智能体生成满足交流潮流及 $N-1$ 安全约束的调度策略。文献[19]采用深度强化学习求解考虑安全约束的 IES 优化调度模型，通过将电网与热网潮流约束融入模型，实现了对调度策略的安全校验。虽然上述文献尝试利用强化学习优化考虑拓扑信息的 IES 运行调度问题，但其只能处理小规模 IES，且只是将拓扑信息转化为约束条件放入惩罚函数作为软约束存在，无法考虑能源系统中的运行状态信息和网络拓扑关系的制约关系，迫切需要进一步深入探讨。

因此，本文将具有处理复杂拓扑网络信息能力的图学习与具有自主探索能力的强化学习相结合，提出一种图强化学习(graph reinforcement learning, GRL)求解方法。基于 GRL 的 IES 优化调度方法可将系统实际拓扑和状态信息纳入智能体动作策略^[20-21]，使智能体在面临复杂多变、灵活动态的调度环境时，能够考虑系统的整体效益，并提高系统的整体经济性与安全性。

1 图强化学习算法设计

本节主要提出一种考虑 IES 运行状态和拓扑结构的 GRL 算法，并介绍了其总体结构。

本文所提出的 GRL 算法由图卷积神经网络(graph convolutional network, GCN)与深度确定性策略梯度(deep deterministic policy gradient, DDPG)构

成。GRL 算法包括图策略网络和图价值网络，其总体结构框图如图 1 所示。

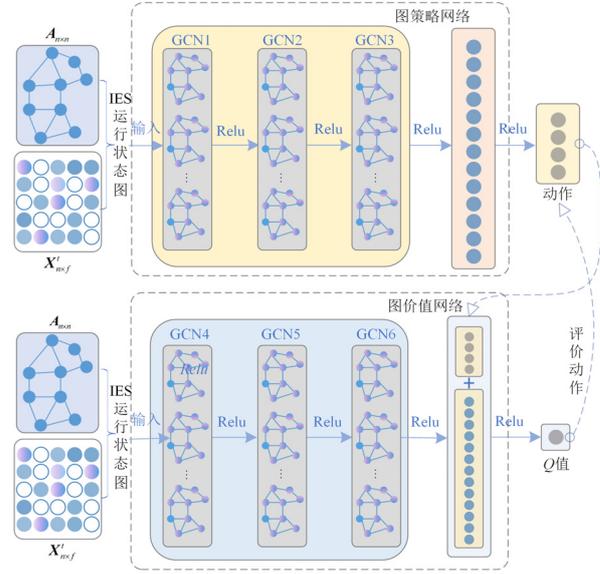


图 1 GRL 算法总体结构示意图

Fig. 1 Schematic diagram of overall structure of GRL algorithm

1) 图策略网络的输入是考虑强非线性潮流知识的 IES 运行状态图，包含表征网络拓扑结构的邻接矩阵 $A_{n \times n}$ 及节点电压、功率等信息的特征矩阵 $X_{n \times f}^I$ ，输出为 IES 拓扑图对应节点的设备出力。

2) 图价值网络初始输入与图策略网络相同，图卷积层后的第一个全连接层(fully connected layer, FC)的输入由图卷积层输出的特征和图策略网络生成的动作拼接而成，图价值网络的最终输出为一维数据，即为该环境状态下所施加动作的 Q 值，用于对动作进行评价。

在图策略网络 θ^π 和图价值网络 θ^Q 中，节点特征信息以隐藏特征的形式在 IES 真实的拓扑层中流动，第 $l+1$ 层的传播特征 $Z^{(l+1)}$ 可以表示为

$$\begin{cases} Z^{(l+1)} = f(Z^{(l)}, A_{n \times n}) = \sigma(\bar{A}Z^{(l)}W^{(l)}) \\ \bar{A} = \tilde{D}^{-\frac{1}{2}}(A_{n \times n} + E)\tilde{D}^{-\frac{1}{2}} \end{cases} \quad (1)$$

式中： $Z^{(l)}$ 为第 l 层的输入特征； $W^{(l)}$ 为第 l 层的权重矩阵； \tilde{D} 为邻接矩阵 $A_{n \times n}$ 的度矩阵； $\sigma(\cdot)$ 为激活函数； E 为单位矩阵。

因此，GRL 算法层间传递拓扑特征信息的规则如式(2)、式(3)所示。

$$\begin{cases} Z^{(l+1)} = \text{Relu}(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}Z^{(l)}W_{\text{GCN}}^{(l)} + b_{\text{GCN}}^{(l)}) \\ \tilde{A} = A + I \end{cases} \quad (2)$$

$$Z^{(l+1)} = \text{Relu}(Z^{(l)}W_{\text{FC}} + b_{\text{FC}}) \quad (3)$$

式中: $\mathbf{W}_{\text{GCN}}^{(l)}$ 、 $\mathbf{b}_{\text{GCN}}^{(l)}$ 分别为第 l 层 GCN 的权重矩阵和偏置向量; $\text{Relu}(\cdot)$ 为激活函数; \mathbf{W}_{FC} 、 \mathbf{b}_{FC} 分别为全连接层的权重矩阵、偏置向量。

图价值网络通过最小化损失函数优化参数, 损失函数 $L(\theta^Q)$ 如式(4)、式(5)所示。

$$L(\theta^Q) = E(y_i - Q(s_i(\mathbf{A}_{n \times n}, \mathbf{X}_{n \times f}^t), \mathbf{a}_i | \theta^Q))^2 \quad (4)$$

$$y_i = r_i + \gamma Q' \left(\begin{array}{l} s_{i+1}(\mathbf{A}_{n \times n}, \mathbf{X}_{n \times f}^t), \\ \pi'(s_{i+1}(\mathbf{A}_{n \times n}, \mathbf{X}_{n \times f}^t) | \theta^\pi) | \theta^Q \end{array} \right) \quad (5)$$

式中: $s_i(\mathbf{A}_{n \times n}, \mathbf{X}_{n \times f}^t)$ 为与图邻接矩阵和图特征矩阵相关的状态向量; y_i 为目标 Q 值; $E(\cdot)$ 为期望函数; $Q(\cdot | \theta^Q)$ 为参数 θ^Q 下的价值网络; \mathbf{a}_i 为动作向量, $\mathbf{a}_i = \pi(s_{i+1}(\mathbf{A}_{n \times n}, \mathbf{X}_{n \times f}^t) | \theta^\pi)$; r_i 为即时奖励; γ 为折扣因子; $Q'(\cdot | \theta^Q)$ 为参数 θ^Q 下目标图价值网络的目标 Q 值; $\pi'(\cdot | \theta^\pi)$ 为参数 θ^π 下目标图策略网络的目标策略值。

图价值网络与图策略网络的参数 θ^Q 和 θ^π 根据式(6)、式(7)进行更新, 对于图策略网络, 其提供梯度信息 $\nabla_a Q(s_i(\mathbf{A}_{n \times n}, \mathbf{X}_{n \times f}^t), \mathbf{a}_i | \theta^Q)$ 作为动作改进的方向, 并采用采样策略梯度进行更新。

$$\theta^Q \leftarrow \theta^Q - \mu_Q \nabla_{\theta^Q} L(\theta^Q) \quad (6)$$

$$\theta^\pi \leftarrow \theta^\pi + \mu_\pi \nabla_{\theta^\pi} \pi \quad (7)$$

式中: ∇ 为表示梯度计算的函数; $\nabla_{\theta^\pi} \pi = \nabla_a Q(s_i(\mathbf{A}_{n \times n}, \mathbf{X}_{n \times f}^t), \mathbf{a}_i | \theta^Q) \nabla_{\theta^\pi} \pi(s_i | \theta^\pi)$; μ_Q 为图价值网络学习率; μ_π 为图策略网络学习率。

采用软更新技术对目标价值网络参数和目标策略网络参数进行更新, 如式(8)、式(9)所示。

$$\theta^Q \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'} \quad (8)$$

$$\theta^\pi \leftarrow \tau \theta^\pi + (1 - \tau) \theta^{\pi'} \quad (9)$$

式中, τ 为软更新系数。

本文将物理对象及所提 GRL 框架应用于 IES 优化调度问题, 得到考虑节点运行状态信息的 IES 优化调度模型, 物理对象及其模型将在第 2 节中详细阐述。

2 考虑运行状态信息的 IES 优化调度模型

本节首先构建了考虑目标函数与网络约束条件等调度知识的 IES 优化调度数学模型。其次, 将其进一步转化为马尔可夫决策过程, 为 GRL 算法在优化调度中的快速求解奠定基础。

2.1 电热耦合 IES 介绍

本文以某工业园区电热耦合 IES^[19]为研究对

象, 其中电力系统由电源、电网、储能(battery energy storage, BES)和有功无功负荷组成, 电源包括上级电网、热电联产机组(combined heat and power, CHP)以及 PV 光伏阵列。热力系统由热源、热网和热负荷组成, 热源包括 CHP、电锅炉(electric boiler, EB)和燃气锅炉(gas boiler, GB)。CHP 与 EB 是电力系统和供热系统的耦合元件。

1) 热电联产机组

热电联产机组消耗天然气产生电能和热能, 其运行特性为

$$P_{\text{CHP}}^{i,t} = M_{\text{CHP}}^{i,t} \eta_{\text{CHP}} \quad (10)$$

$$Q_{\text{CHP}}^{i,t} = \tan(\arccos \varphi_{\text{CHP}}) P_{\text{CHP}}^{i,t} \quad (11)$$

$$P_{\text{CHP}}^{\min} \leq P_{\text{CHP}}^{i,t} \leq P_{\text{CHP}}^{\max} \quad (12)$$

式中: $P_{\text{CHP}}^{i,t}$ 为 t 时刻 i 节点热电联产机组输出的有功功率; $M_{\text{CHP}}^{i,t}$ 为 t 时刻 i 节点热电联产机组天然气的消耗量; η_{CHP} 为热电联产机组的电转化效率; $Q_{\text{CHP}}^{i,t}$ 为 t 时刻 i 节点热电联产机组输出的无功功率; φ_{CHP} 为 CHP 机组的功率因数; P_{CHP}^{\min} 和 P_{CHP}^{\max} 分别为热电联产机组输出电功率的最小值和最大值。

热电联产机组的“电热特性”一般采用定热电比, 表示为

$$b_{\text{CHP}} = \frac{H_{\text{CHP}}^{i,t}}{P_{\text{CHP}}^{i,t}} \quad (13)$$

式中: $H_{\text{CHP}}^{i,t}$ 为 t 时刻 i 节点热电联产机组输出的热功率; b_{CHP} 为热电比。

2) 电锅炉

电锅炉消耗电能产生热能, 其运行特性为

$$H_{\text{EB}}^{i,t} = P_{\text{EB}}^{i,t} \eta_{\text{EB}} \quad (14)$$

$$H_{\text{EB}}^{\min} \leq H_{\text{EB}}^{i,t} \leq H_{\text{EB}}^{\max} \quad (15)$$

式中: $H_{\text{EB}}^{i,t}$ 为 t 时刻 i 节点电锅炉输出的热功率; $P_{\text{EB}}^{i,t}$ 为 t 时刻 i 节点电锅炉消耗的电功率; η_{EB} 为电锅炉的转化效率; H_{EB}^{\min} 和 H_{EB}^{\max} 分别为电锅炉输出热功率的最小值和最大值。

3) 燃气锅炉

燃气锅炉消耗天然气产生热能, 其运行特性为

$$H_{\text{GB}}^{i,t} = M_{\text{GB}}^{i,t} \eta_{\text{GB}} \quad (16)$$

$$H_{\text{GB}}^{\min} \leq H_{\text{GB}}^{i,t} \leq H_{\text{GB}}^{\max} \quad (17)$$

式中: $H_{\text{GB}}^{i,t}$ 为 t 时刻 i 节点燃气锅炉输出的热功率; $M_{\text{GB}}^{i,t}$ 为 t 时刻 i 节点燃气锅炉天然气的消耗量; η_{GB} 为燃气锅炉的转化效率; H_{GB}^{\min} 和 H_{GB}^{\max} 分别为燃气锅炉输出热功率的最小值和最大值。

4) 电储能

电储能的荷电状态(state of charge, SOC)及运行特性为

$$C_{\text{SOC}}^{i,t} = C_{\text{SOC}}^{i,t-1} - \frac{\eta_{\text{BES}} P_{\text{BES}}^{i,t} \Delta t}{Q_{\text{BES}}} \quad (18)$$

$$C_{\text{SOC}}(0) = C_{\text{SOC}}^{\text{ini},i} \quad (19)$$

$$C_{\text{SOC}}^{\text{min}} \leq C_{\text{SOC}}^{i,t} \leq C_{\text{SOC}}^{\text{max}} \quad (20)$$

$$P_{\text{BES}}^{\text{min}} \leq P_{\text{BES}}^{i,t} \leq P_{\text{BES}}^{\text{max}} \quad (21)$$

式中: $C_{\text{SOC}}^{i,t}$ 、 $C_{\text{SOC}}^{i,t-1}$ 分别为 t 、 $t-1$ 时刻 i 节点电储能的荷电状态; Q_{BES} 为电储能的容量; Δt 为时间间隔; $C_{\text{SOC}}^{\text{ini},i}$ 为 i 节点处电储能初始时刻的荷电状态; $C_{\text{SOC}}^{\text{min}}$ 和 $C_{\text{SOC}}^{\text{max}}$ 分别为电储能设备荷电状态的最小值和最大值; $P_{\text{BES}}^{\text{min}}$ 和 $P_{\text{BES}}^{\text{max}}$ 分别为电储能设备充/放电功率的最小值和最大值; $P_{\text{BES}}^{i,t}$ 为 t 时刻 i 节点电储能的充/放电功率, $P_{\text{BES}}^{i,t}$ 为正时表示放电, $P_{\text{BES}}^{i,t}$ 为负时表示充电; η_{BES} 为电储能充/放电系数, 如式(22)所示。

$$\eta_{\text{BES}} = \begin{cases} \eta_{\text{ch}} & P_{\text{BES}}^{i,t} < 0 \\ \frac{1}{\eta_{\text{dis}}} & P_{\text{BES}}^{i,t} \geq 0 \end{cases} \quad (22)$$

式中, η_{ch} 和 η_{dis} 分别为电储能的充、放电效率。

5) 光伏阵列

光伏阵列的无功出力和有功出力满足式(23)。

$$(P_{\text{PV}}^{i,t})^2 + (Q_{\text{PV}}^{i,t})^2 \leq (S_{\text{PV}})^2 \quad (23)$$

式中: $P_{\text{PV}}^{i,t}$ 、 $Q_{\text{PV}}^{i,t}$ 分别为光伏阵列输出的有功功率、无功功率; S_{PV} 为光伏阵列的视在功率。

2.2 IES 优化调度数学模型

1) 目标函数

本文中所要实现的调度目标是要在考虑运行状态信息不越限的情况下, 尽量使得其运行成本达到最优。因此, 本文的目标函数是使得运行成本 F 最小化, 即 $\min F$ 。

一般来说, IES 的运行总成本 $C_{\text{cost}}(s_t, a_t)$ 包括拓扑中 i 节点的机组与电网交互的购电成本、购气成本, i 节点处连接的电储能充放电折旧成本^[22]。故目标函数表示为

$$\min F = \sum_{t=1}^T C_{\text{cost}}(s_t, a_t) = \sum_{t=1}^T \left(\sum_{i \in \Omega} \varepsilon_{\text{E}}^t P_{\text{grid}}^{i,t} + \sum_{i \in \Omega} \varepsilon_{\text{NG}}^t M_{\text{NG}}^{i,t} \right) \Delta t + \sum_{t=1}^T \sum_{i \in \Omega_{\text{BES}}} \rho_{\text{BES}} |P_{\text{BES}}^{i,t}| \quad (24)$$

$$M_{\text{NG}}^{i,t} = M_{\text{CHP}}^{i,t} + M_{\text{GB}}^{i,t} \quad (25)$$

式中: T 为一个运行周期内的时段数; Ω 为购电与购气节点的集合; Ω_{BES} 为电储能节点的集合; $P_{\text{grid}}^{i,t}$ 为 t 时刻 i 节点与上级电网交互的电功率, $P_{\text{grid}}^{i,t}$ 为正

时表示从上级电网购电, $P_{\text{grid}}^{i,t}$ 为负时表示向上级电网售电; ε_{E}^t 为 t 时刻的电价, $P_{\text{grid}}^{i,t}$ 为正时电价取购电电价, 为负时电价取售电电价; $\varepsilon_{\text{NG}}^t$ 为 t 时刻的天然气管价; $M_{\text{NG}}^{i,t}$ 为 t 时刻系统购买天然气的总量; ρ_{BES} 为电储能折旧成本系数。

2) 运行状态约束条件

除 2.1 节中设备自行满足的约束(式(10)一式(23))外, IES 动态调度问题的运行约束包括电网潮流约束、热网潮流约束、节点电压越限约束和各设备运行约束。

(1) 电网潮流约束

在节点注入型的潮流模型(式(26)一式(28))中, 电网的任一节点应满足节点功率平衡约束(式(27)、式(28)), 表示为

$$\begin{cases} P_i^t = U_i^t \sum_{j=1}^N U_j^t (G_{ij} \cos \delta_{ij}^t + B_{ij} \sin \delta_{ij}^t) \\ Q_i^t = U_i^t \sum_{j=1}^N U_j^t (G_{ij} \sin \delta_{ij}^t - B_{ij} \cos \delta_{ij}^t) \end{cases} \quad (26)$$

$$P_i^t = P_{\text{grid}}^{i,t} + P_{\text{CHP}}^{i,t} + P_{\text{BES}}^{i,t} + P_{\text{PV}}^{i,t} - P_{\text{EB}}^{i,t} - P_{\text{load}}^{i,t} \quad (27)$$

$$Q_i^t = Q_{\text{grid}}^{i,t} + Q_{\text{CHP}}^{i,t} + Q_{\text{PV}}^{i,t} - Q_{\text{load}}^{i,t} \quad (28)$$

式中: P_i^t 、 Q_i^t 分别表示 t 时刻 i 节点的有功功率和无功功率; U_i^t 表示 t 时刻 i 节点的电压; $Q_{\text{grid}}^{i,t}$ 为 t 时刻 i 节点与电网交互的无功功率; $P_{\text{load}}^{i,t}$ 为 t 时刻 i 节点处的有功负荷; $Q_{\text{load}}^{i,t}$ 为 t 时刻 i 节点处的无功负荷; G_{ij} 为 i 节点与 j 节点之间的电导; B_{ij} 为 i 节点与 j 节点之间的电纳; δ_{ij}^t 为 t 时刻 i 节点与 j 节点之间的相角差。

(2) 节点热功率平衡约束

本文的热力系统采用质调节方式, 即水流量固定, 水温可调以此保障热力负荷的供应。系统的热网稳态模型约束如式(29)一式(33)所示^[19]。

$$C_p A_{\text{down}} \mathbf{M} \mathbf{A}_{\text{down}}^T \mathbf{T}_n^t - C_p A_{\text{up}} \mathbf{M} \mathbf{T}_c^t = \mathbf{H}_n^t \quad (29)$$

$$\mathbf{H}_n^t = \mathbf{H}_{\text{CHP}}^t + \mathbf{H}_{\text{EB}}^t + \mathbf{H}_{\text{GB}}^t - \mathbf{H}_{\text{load}}^t \quad (30)$$

$$\mathbf{E} (\mathbf{A}_{\text{down}}^T \mathbf{T}_n^t - \mathbf{T}_a^t) + \mathbf{T}_a^t = \mathbf{T}_c^t \quad (31)$$

$$\mathbf{M} = \text{diag}(m_1, m_2, m_3, \dots, m_i) \quad (32)$$

$$\mathbf{E} = \text{diag}(\mathbf{e}, \frac{\lambda_1 L_1}{C_p m_1}, \mathbf{e}, \frac{\lambda_2 L_2}{C_p m_2}, \dots, \mathbf{e}, \frac{\lambda_i L_i}{C_p m_i}) \quad (33)$$

式中: C_p 为水的比热容; A_{up} 、 A_{down} 分别为热网节点与管道支路的上、下关联矩阵; \mathbf{M} 为管道流量对角矩阵; \mathbf{E} 为温度衰减系数对角矩阵; \mathbf{H}_n^t 为 t 时刻流入节点的热功率矩阵; $\mathbf{H}_{\text{CHP}}^t$ 、 \mathbf{H}_{EB}^t 、 \mathbf{H}_{GB}^t 、 $\mathbf{H}_{\text{load}}^t$

分别为 t 时刻 CHP、EB、GB 及负荷的热功率矩阵; T_n^t 、 T_e^t 及 T_a^t 分别为 t 时刻节点温度矩阵、管道末端温度矩阵及管道外部环境温度矩阵; m_i 、 λ_i 及 L_i 分别为 i 管道的流量、单位长度的导热率及长度。

(3) 节点电压越限约束

为保证系统安全运行, 节点电压 U_i^t 满足约束条件式(34)。

$$U_i^{\min} \leq U_i^t \leq U_i^{\max} \quad (34)$$

式中, U_i^{\min} 和 U_i^{\max} 分别为系统安全电压的最小值和最大值。

2.3 考虑节点运行状态的马尔可夫决策过程

基于上述具有针对性的优化调度过程, 本节通过将数学模型转换为马尔可夫决策过程进一步得到面向 IES 优化调度的 GRL 模型, 表 1 给出了考虑运行状态信息的马尔可夫决策过程与数学模型之间的变量对应关系。

表 1 考虑运行状态信息的马尔可夫决策过程与数学模型之间的变量对应关系

Table 1 Correspondence between variables of Markov game model based on graph process and mathematical model considering operation state information

基于图的马尔可夫过程要素	IES 智能体 GRL 优化调度模型变量
环境	包含物理拓扑的 IES 系统实际运行环境
状态空间 S	反映当前拓扑状态下 IES 实际运行状态图, 包含 $A_{n \times n}$ 、 $X_{n \times f}^t$
动作空间 A	IES 调度问题相关决策变量的集合, 包含: $A = (P_{\text{CHP}}^{i,t}, H_{\text{EB}}^{i,t}, P_{\text{BES}}^{i,t}, Q_{\text{PV}}^{i,t})$
奖励函数 r_t	系统运行成本、机组约束惩罚及网络拓扑约束惩罚
状态图转移概率 P	执行动作后状态对之间的转移概率, 无模型图强化学习的 P 由 IES 环境确定

1) 基于图运行状态信息的状态空间 S

状态图 $s_t(A_{n \times n}, X_{n \times f}^t)$ 包括拓扑结构信息 $A_{n \times n}$ 和节点特征(电压、功率等)信息 $X_{n \times f}^t$ 。

(1) 图邻接矩阵 $A_{n \times n}$: $A_{n \times n}$ 用来表示图中节点间的连接关系。

(2) 节点特征矩阵 $X_{n \times f}^t$: $X_{n \times f}^t$ 中每一行代表一个节点特征向量, 即 $X_{n \times f}^t = [U_i^t, P_i^t, Q_i^t, C_{\text{SOC}}^{i,t}]$ 。

2) 基于图运行状态信息的动作空间 A

对于 IES 的优化调度问题, 智能体在每个 t 时刻可根据当前策略网络输入的当前状态图 $s_t(A_{n \times n}, X_{n \times f}^t)$ 做出动作 a_t (动作用 i 节点处的相应设备出力表示)。由于 CHP 热出力、EB 输入电功率可分别由

式(13)、式(14)得到, 电网交互功率可由式(26)—式(28)得到, GB 热出力可由式(30)并结合式(29)、式(31)—式(33)得到, 因此本文定义动作空间为 $A = (P_{\text{CHP}}^{i,t}, H_{\text{EB}}^{i,t}, P_{\text{BES}}^{i,t}, Q_{\text{PV}}^{i,t})$ 。

3) 状态图转移概率 P

状态图转移概率指智能体由当前运行状态图 $s_t(A_{n \times n}, X_{n \times f}^t)$ 执行所选动作 a_t 后, 转移至新状态图 $s_{t+1}(A_{n \times n}, X_{n \times f}^{t+1})$ 的可能性。针对 IES 动态调度问题, 智能体做出动作后, 下一时刻状态中节点的特征信息需要智能体与环境交互后才能获得, 并观测模拟环境在受到此动作影响后的下一时刻状态, 以实现包含环境随机性的状态图转移概率 P 。

4) 奖励函数 r_t

奖励函数由两部分组成, 即系统运行成本和包含拓扑网络约束罚函数。

(1) 系统运行成本 $C_{\text{cost}}(s_t, a_t)$

强化学习的目标是通过智能体与环境进行交互来学习一种策略, 使得智能体获得最大的累计回报。基于此, 将 IES 动态经济调度总成本最小化问题转化为强化学习奖励函数最大化形式, 并对数值进行缩放, 防止奖励函数值过大导致模型不收敛^[23], 具体参考式(24)—式(25)。

(2) 带有强非线性潮流知识的约束惩罚函数 $C_{\text{pun}}(s_t, a_t)$

系统内各设备的运行约束、带有潮流平衡的节点电压越限所构成的网络约束, 采用罚函数的形式添加到奖励函数中, 具体如式(35)所示。

$$C_{\text{pun}}(s_t, a_t) = P_{\text{BES}}^{i,t}(s_t, a_t) + P_{\text{CHP}}^{i,t}(s_t, a_t) + P_{\text{SOC}}^{i,t}(s_t, a_t) + P_{\text{PV}}^{i,t}(s_t, a_t) + P_{\text{EB}}^{i,t}(s_t, a_t) + P_{\text{GB}}^{i,t}(s_t, a_t) \quad (35)$$

以 BES 为例, t 时刻 BES 违反设备运行约束的罚函数 $P_{\text{BES}}^{i,t}(s_t, a_t)$ 表示为

$$P_{\text{BES}}^{i,t}(s_t, a_t) = |P_{\text{BES}}^{i,t} - P_{\text{BES}}^{\max}| + |P_{\text{BES}}^{\min} - P_{\text{BES}}^{i,t}| \quad (36)$$

t 时刻 CHP、PV、GB 和 EB 违反运行约束的罚函数项与 BES 具有相同的形式, 不再赘述, 违反节点电压约束的罚函数 $P_{U_i}^t(s_t, a_t)$ 如式(37)所示。

$$P_{U_i}^t(s_t, a_t) = \sum_{i=1}^T \sum_{i=1}^n |U_i^t - U_i^*| \quad (37)$$

式中, U_i^* 为电压标幺值, 取值 1.0 p.u.。

故而, 智能体在 t 时刻获得的奖励函数是 $C_{\text{cost}}(s_t, a_t)$ 与 $C_{\text{pun}}(s_t, a_t)$ 的加权和, 表示为

$$r_t(s_t, a_t) = -\lambda(C_{\text{cost}}(s_t, a_t) + C_{\text{pun}}(s_t, a_t)) \quad (38)$$

式中, λ 为缩放系数, 取值 0.001。

3 基于 GRL 的 IES 动态优化调度框架

基于第 1 节的 GRL 算法与第 2 节所提的考虑运行状态信息的 IES 优化调度模型, 本节设计一个将 GRL 算法应用于该模型快速求解的优化调度框架, 以解决具有复杂物理结构的 IES 优化调度问题。

所提的框架包括三层两阶段, 如图 2 所示。“三层”即物理实体层、拓扑映射层及智能体交互层; “两阶段”即离线训练阶段、在线应用阶段。两阶段详细阐述了三层之间的交互关系。下面分别对框架中的内容进行详细介绍。

“物理实体层”即某工业园区型 IES, 其具有高能耗、多能互补、热电联供、安全可靠的特点, 是推进能源转型和可持续发展的重要手段之一^[24-26], 如图 2 中第一层所示; “物理实体层”将 IES 真实系统映射为有权无向拓扑图^[27-28], 如图 2 中第二层所示; 其包括代表物理知识的邻接矩阵和代表强非线性潮流知识等系统状态信息的节点特征矩阵; 由“拓扑映射层”得到的 IES 运行状态图输入图 2 中的“智能体交互层”, 该层中的图策略网络及图价值网络执行离线训练及在线应用过程, 使智能体最大限度地

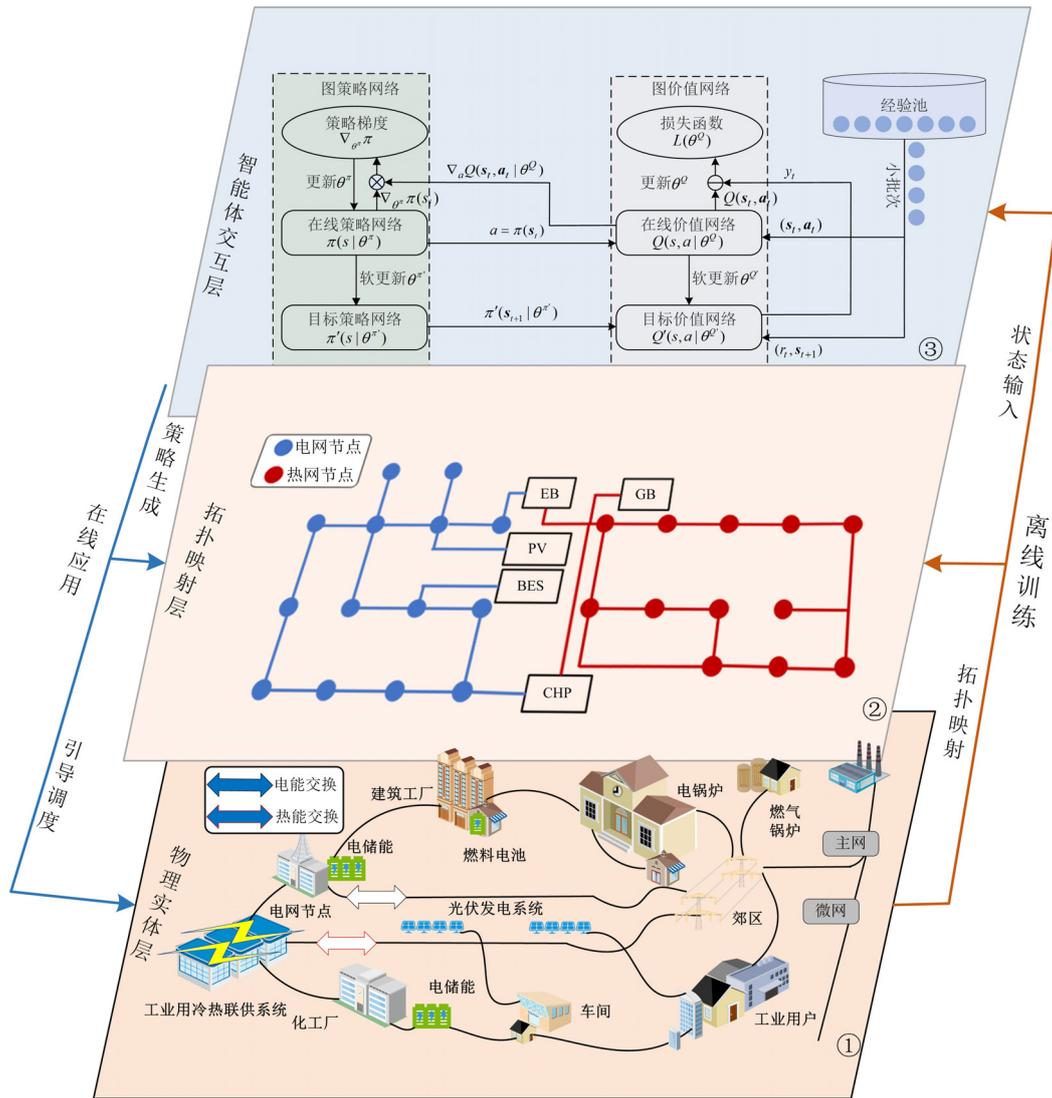


图 2 基于 GRL 的 IES 优化调度框架

Fig. 2 IES optimal schedule framework based on GRL

所提框架的主要特点包括:

1) 能够充分感知 IES 产生的海量非结构化、非欧式潮流数据;

2) 能够学习系统所包含的能源知识, 从而提高决策能力。

在构建上述 GRL 模型后, 采用“离线训练+在

线应用”的方式求解包含系统运行状态信息和系统拓扑结构的 IES 动态经济调度模型。

3.1 离线训练过程

智能体与 IES 交互的 MDP 包括 T 个时间步数, 在时刻 t , 首先由负荷等环境信息生成当前时刻的运行状态图, 并将其输入图策略网络以实现 IES 复杂图结构信息的提取与感知, 其次, 图策略网络输出 t 时刻智能体的设备出力动作, 智能体经过环境交互更新系统下一时刻的状态图, 并根据式(10)—式(38)得到 t 时刻的奖励函数 r_t 。此时, 四元组 $[s_t, a_t, r_t, s_{t+1}]$ 作为一个样本放入经验池中。智能体继续根据 s_{t+1} 得到 $t+1$ 时刻的设备出力动作 a_{t+1} , 循环上述过程直到 MDP 结束。经过若干次 MDP 过程得到大量样本数据, 根据式(1)—式(9)更新参数, 直到循环满足最大训练次数 T_{\max} , 保存获得最大奖励回报下的 GRL 网络模型, 即可结束训练。

IES 动态优化调度模型的离线训练流程如图 3 所示, 设定 $N=96$ 、 $T=24$ 、 $T_{\max}=5000$ 。

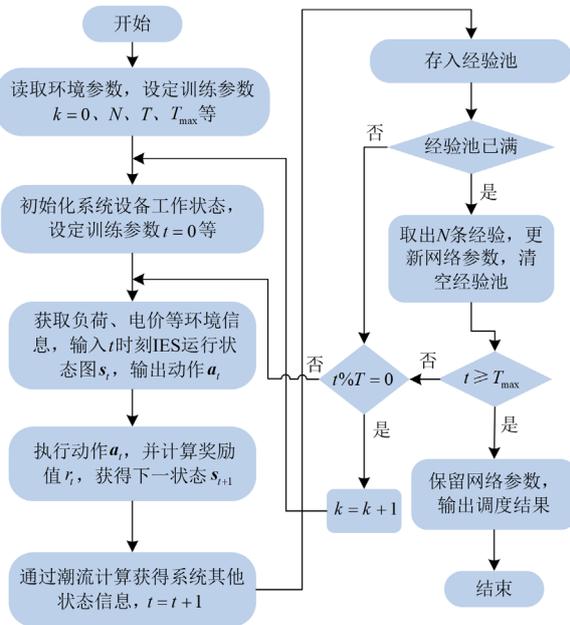


图 3 离线训练流程

Fig. 3 Process of offline training

3.2 在线应用过程

训练结束后, 将训练好的模型应用于实际调度任务中, 即为在线应用过程。

针对 IES 的动态调度问题, 每次调度任务来临时, 智能体获取 s_t , 然后利用训练好的图策略网络选择据 a_t , 该动作与环境交互, 随后智能体感知 s_{t+1} , 同时计算当前时刻的奖励 r_t 。继而智能体采集下一时刻的状态, 直至达成 T 个时间步数的最终目标。

4 算例验证与分析

4.1 参数设置与算例设置

4.1.1 参数设置

1) 物理参数设置

系统内各机组的参数见表 2、表 3 所示。

表 2 设备运行参数

Table 2 Device operation parameters

设备类型	最小电(热)功率/MW	最大电(热)功率/MW	效率
热电联供机组	0	30	0.35
电储能	-10	10	0.95
燃气锅炉	0	25	0.8
电锅炉	0	15	0.95

表 3 其他参数

Table 3 Other parameters

设备类型	参数	值
热电联供机组	热电比	1.2
	功率因数	0.9
光伏阵列	视在功率/MVA	20
电储能	荷电状态最小值	0.2
	荷电状态最大值	0.8
	折旧系数	0.05
	容量/MWh	10

算例基于文献[29]中某工业园区的电热耦合系统进行修改, 系统内各设备安装情况如图 4 所示, 节点 1 处接入上级电网。

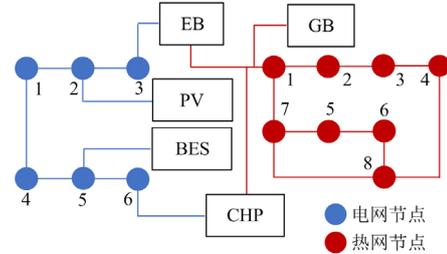


图 4 电热耦合系统 14 节点拓扑结构图

Fig. 4 Combined power and heat system topology nodes of 14 nodes

电力系统线路参数如表 4 所示, 基准容量 $S=100$ MVA, 基准电压 $U=23$ kV。在热力系统中, 热源节点为节点 1, 供水温度恒定为 100 $^{\circ}\text{C}$, 管道外部环境温度为 10 $^{\circ}\text{C}$, 热力网络管道参数如表 5 所示。天然气单价为 0.4 元/kWh。算例中电热负荷需求及光伏出力基于文献[19]进行修改, 如图 5 所示。采用分时电价如表 6 所示, 其中平时段为 07:00—12:00、19:00—23:00, 峰时段为 12:00—19:00, 谷时段为 23:00—07:00。

表 4 电力网络参数

Table 4 Power network parameters

支路编号	支路起点	支路终点	电阻/p.u.	电抗/p.u.
1	1	4	0.075	0.26
2	1	2	0.11	0.17
3	2	3	0.09	0.20
4	4	5	0.08	0.04
5	5	6	0.04	0.14

表 5 热力网络参数

Table 5 Thermal network parameters

管道编号	管道起点	管道终点	长度/m	流量/(kg/s)	导热率/(W/m·K)
1	1	2	350	265.8	0.12
2	2	3	175	241.4	0.12
3	3	4	175	143.58	0.12
4	2	5	75	24.4	0.12
5	3	6	175	97.82	0.12
6	7	1	350	265.8	0.12
7	5	7	75	24.4	0.12
8	8	7	175	241.4	0.12
9	6	8	175	97.82	0.12
10	4	8	175	143.58	0.12

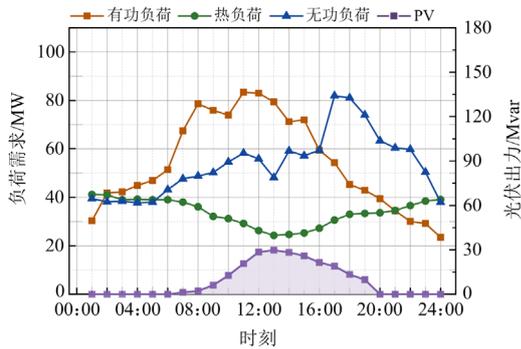


图 5 系统总电热负荷需求与 PV 出力曲线

Fig. 5 Daily curves of system total load and PV generation

表 6 分时电价

Table 6 Time-of-use electric price

时段	购电电价/(元/kWh)	售电电价/(元/kWh)
峰时段	0.98	0.5
平时段	0.49	0.2
谷时段	0.17	0

2) 图强化学习算法参数设置

本文基于文献[30]对 GRL 的超参数进行设计, Actor-Critic 网络采用 Adam 优化, 学习率分别为 0.0001 和 0.001。经过多次仿真实验本文设置 3 层图卷积层, 每层神经元的个数为 64, 激活函数采用 Relu, 其他参数如表 7 所示。

表 7 GRL 算法参数设置

Table 7 GRL algorithm parameter settings

参数	GRL
探索率	0.2
软更新系数	0.05
折现率	0.9
小批量大小	96
调度总周期/h	24
时间间隔/h	1
总训练次数	5000

4.1.2 仿真场景设置

为对本文所提的 GRL 算法进行有效对比和分析, 基于经典 DDPG 算法对 GRL 进行基准测试, 场景设置如下。

场景 1: DDPG 算法, 未考虑运行状态约束。

场景 2: GRL 算法, 未考虑运行状态约束。

场景 3(本文所提): GRL 算法, 考虑运行状态约束。

4.2 算例仿真结果分析

4.2.1 离线训练结果分析

基于本文所提的 GRL 算法对 IES 优化调度模型进行离线训练, 从奖励函数收敛性、损失函数及电压越限惩罚三个方面, 分析所提算法的优势。

1) 不同算法奖励函数收敛性分析

基于 GRL 算法的考虑拓扑网络前后的 IES 动态经济调度模型训练过程中奖励函数变化情况及 GRL 算法收敛性对比分析, 如图 6 所示。

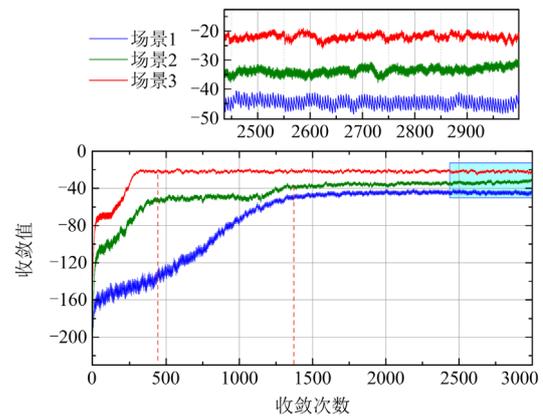


图 6 奖励函数收敛曲线

Fig. 6 Reward function convergence curve

从图 6 可知, GRL 算法(场景 2、场景 3)相对于 DDPG 算法(场景 1)在收敛速度方面缩短了约 41%, 最终收敛值增加了约 10%, 且增加运行状态约束后收敛速度更快, 收敛值更高, 这表明 GRL 算法能够更好地学习拓扑结构和能量传输规律, 并将其纳入

决策过程, 提升了收敛速度和收敛效果。

具体来说, 本文所提 GRL 算法能够捕捉系统中的结构化信息和网络拓扑关系, 加快收敛速度, 减少越限惩罚, 提高收敛值, 且本文所提 GRL 算法在经过约 400 次训练后已基本达到收敛, 并且随着训练的进行, 其对拓扑的学习能力不断提升, 总奖励回报的振荡保持在较小区间内。

2) 损失函数分析

为了进一步判断智能体训练效果, 对损失函数进行分析, 如图 7 所示。

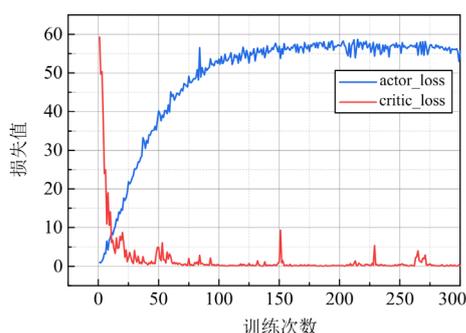


图 7 场景 3 损失函数曲线

Fig. 7 Loss function curve of scenario 3

由图 7 可知, 在训练初期, 图策略网络 actor_loss 和图价值网络 critic_loss 的损失值波动较大, 但随着训练次数的增加, 图策略网络逐渐学习到最佳策略, 并向着最佳策略方向增大, 图价值网络随之减小。在大约 100 次训练后, 这种波动逐渐减小, 并最终收敛到一个稳定状态, 这表明本文方法能够在较小的训练次数下即可有效地实现系统的策略选择。

3) 电压越限惩罚分析

对调度过程中的电压越限进行惩罚是保障系统安全稳定运行的关键策略。针对场景 3 考虑拓扑信息和运行状态约束的 GRL 算法, 其电压惩罚越限情况如图 8 所示。

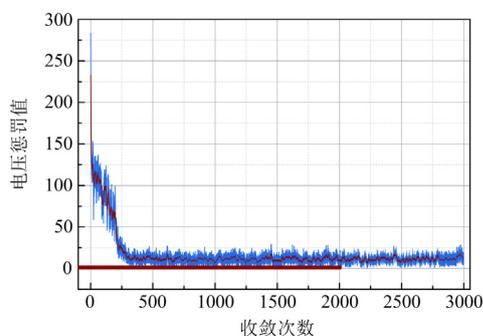


图 8 场景 3 电压越限惩罚

Fig. 8 Voltage over-limit punishment of scenario 3

由图 8 可知, 在训练初期, 智能体未与拓扑环境深入交互, 机组动作的电压安全域边界尚未建立, 系统电压越限明显, 在调度决策过程中机组提供功率与负荷需求不匹配程度较大, 容易引发系统运行安全性问题。但随着智能体与环境不断交互与试错, 智能体自适应输入的不规则非欧式结构数据, 能够提取节点间的潜在特征, 在 500 次迭代后所做出的优化调度决策基本不出现电压越限问题。本文所提方法在纳入拓扑结构和运行状态信息后能够将电压越限情况控制在合理范围内, 有效地避免机组功率与负荷需求不匹配的情况。

4.2.2 在线运行结果分析

经过上述针对 GRL 算法智能体的离线训练, 智能体已经学习到经验池中图数据的规律和特征, 将上述训练完成的模型在线应用, 可得到实际运行下 IES 动态经济调度结果, 下面针对场景 1、2、3 在机组出力结果、电压越限情况、园区经济效益以及拓扑变化适应能力方面进行对比分析, 并与其他方法进行对比分析。

1) 机组出力结果对比分析

3 种场景下 IES 的有功功率及热功率调度结果如图 9 所示。

由图 9 对比分析可知, 本文提出的图强化学习方法在保障拓扑安全的同时优化了综合能源系统的运行, 使得能量的产生和消耗得到了有效的平衡。

具体来说, 场景 1 和场景 2 在电价谷时段和平时段主要通过向主电网购电的方式来满足电负荷及电锅炉的需求, 电储能进行充电以备后续高峰时段使用, 热负荷需求主要由 EB 和 GB 共同满足; 在电价峰时段, CHP 增加了输出的有功功率, 而 EB 降低了有功功率的消耗量, 同时 BES 也在该时段进行放电; 热负荷需求主要由 CHP 机组和 GB 机组共同满足, 从而减少系统的运行成本。由于场景 2 考虑了能源系统中的结构化信息和网络拓扑关系, 因此机组出力变化趋势更加平缓。

场景 3 由于受运行状态约束条件的影响, 在电价谷时段和平时段, 系统虽仍以向电网购电方式满足电负荷需求, 但购电量明显下降; 在电价峰时段, CHP 机组小幅度增加了输出的有功功率, BES 进行放电, 同时 EB 降低消耗, 由此降低系统的运行成本; 热负荷需求由 CHP、EB 和 GB 共同满足, CHP 和 EB 受运行状态约束在电价平时段和峰时段输出热功率的变化幅度较小, GB 输出热功率的变化幅度较大。因而证明场景 3 在保证系统安全运行的同时具有一定经济性能。

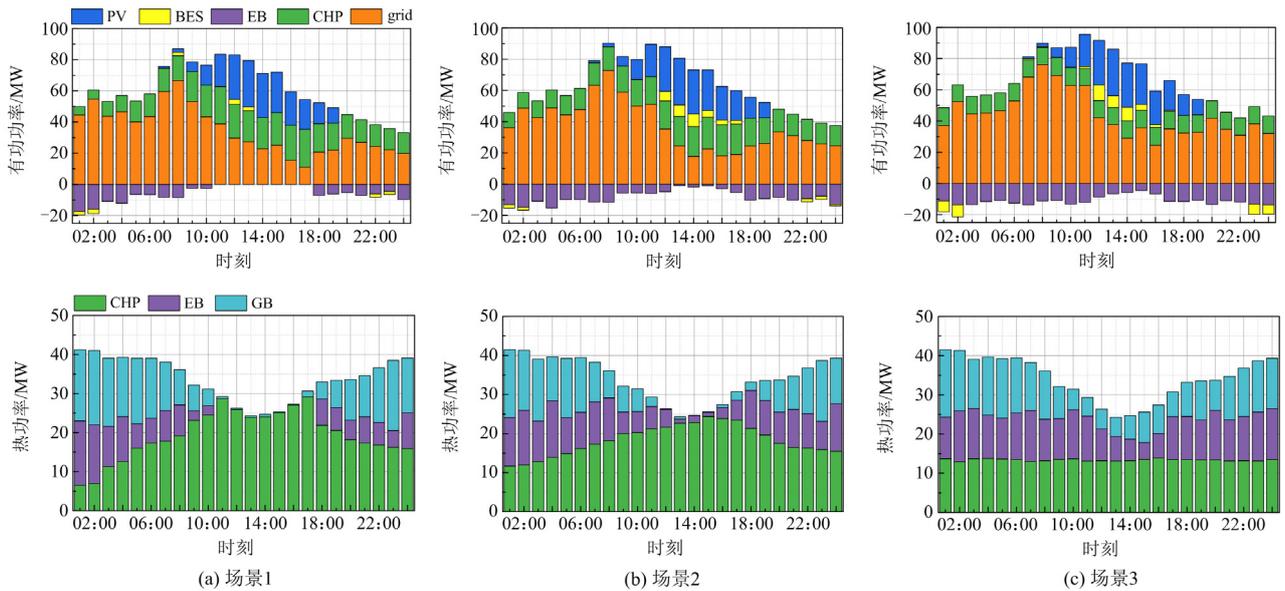


图9 不同场景下调度结果

Fig. 9 Dispatch results under different scenarios

由此可以看出，本文所提方法在保障系统安全稳定运行的同时，能够有效协调各机组的出力以满足系统的需求。

2) 电压越限对比分析

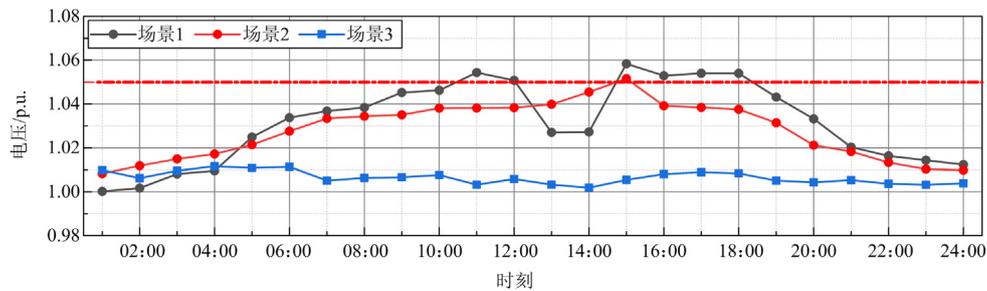
为进一步验证所提模型的有效性，对比分析 3 种场景的局部节点电压水平及相应的机组调度结果。考虑到 CHP 机组接入节点 6，因此对 CHP 的出力调度结果及节点 6 的电压水平进行分析，结果

如表 8 及图 10 所示。

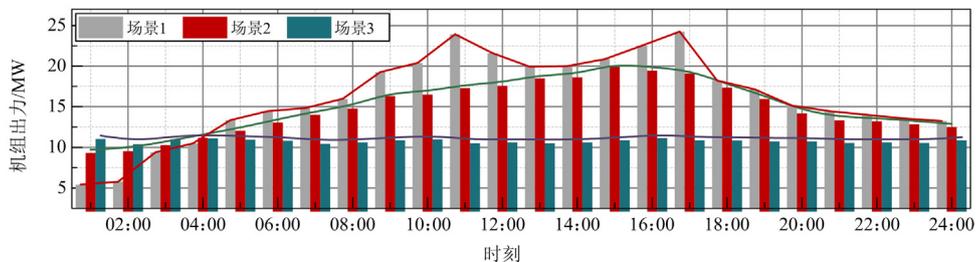
表 8 电压越限对比

Table 8 Comparison of voltage over the limit

场景	越限时段数/个	平均电压偏差/p.u.	最大电压偏差/p.u.
场景 1	6	0.039	0.084
场景 2	1	0.015	0.077
场景 3	0	0.006	0.025



(a) 节点6电压越限情况对比



(b) 节点6所接CHP机组出力情况对比

图 10 节点 6 电压水平及 CHP 调度结果

Fig. 10 Dispatch results of voltage level and CHP of node 6

场景 1 中 CHP 在 11:00—12:00 及 15:00—18:00 时段向电网注入的功率较高, 导致该时段出现电压越限; 同一时段场景 2 与场景 3 降低了 CHP 注入电网的功率, 但场景 2 仍有部分时段越限, 场景 3 系统运行在安全范围内, 电压更接近基准电压, 且电压变化幅度平缓, 从而证明本文所提方法的有效性。

3) 经济效益对比分析

从表 9 可以看出, 场景 2、场景 3 相比场景 1 的购电成本分别上升了约 4%、15%, CHP 机组运行成本分别下降了约 2%、15%, GB 运行成本变化不大。由此可见, 运行状态信息和安全约束都会限制 CHP 机组的出力, CHP 机组在电价峰时段只能小幅度增加出力来降低成本, 其中场景 3 同时受到拓扑信息和安全约束的影响, 因此下降幅度相比场景 2 有所增加。场景 2 相比场景 1 运行成本减少了 0.49%, 而场景 3 相比场景 2 运行成本增加了 2.7%, 说明场景 3 为满足拓扑约束将牺牲一部分经济效益。这与文献[19]所构建的基于 PPO 的 IES 电压越限情况以及文献[31]中环境-安全-经济“矛盾三角”形态具有一致性, 即虽然运行成本可能较高, 但在实际应用中, 该算法可能会产生更低的总体成本, 它可以更好地保证电力系统的稳定性和安全性, 从而减少系统故障和停机的时间成本。

表 9 运行成本对比

Table 9 Comparison of operation cost

场景	元				
	购电成本	CHP 成本	GB 成本	BES 成本	总运行成本
场景 1	381 395.40	443 729.55	108357.91	0.83	933 483.69
场景 2	416 040.11	418 834.69	93996.74	2.02	928 873.56
场景 3	525 871.43	305 962.23	122565.62	3.51	954 402.79

4) 拓扑变化适应能力分析

为了应对紧急情况、定期检修等引起的拓扑结构改变问题, 将对本文所提算法进行扩展性校验, 以验证其对拓扑变化的适应能力。

当智能体离线训练到达 1000 次时, 对系统拓扑结构进行更改, 其具体变化情况如表 10 所示, 更改后系统的奖励函数收敛曲线如图 11 所示。从图 11 可以看出, GRL 表现出优异的泛化性能和可扩展性。当拓扑结构发生变化时, 经过训练的 GRL 策略仍然可行, 奖励值收敛曲线在迭代到 1000 次时出现波动, 但最终稳定在比原来较低的收敛值。这表明 GRL 拥有强大的状态感知和特征提取能力, 可以快速适应系统的变化, 并表现出良好的扩展性能。

5) 与其他方法对比

表 10 拓扑结构变化设置

Table 10 Topology change settings

支路运行方式	支路(节点-节点)	支路(节点-节点)
断开	1-2	2-3
连接	3-6	—

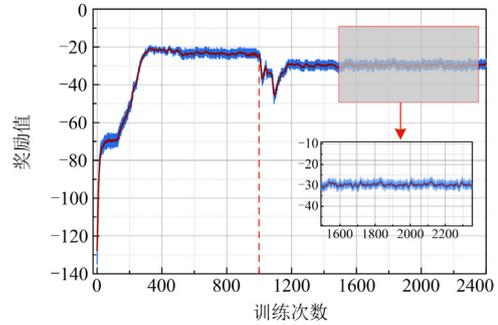


图 11 拓扑变化后奖励值收敛曲线

Fig. 11 Reward value convergence curve after topology change

为进一步验证本文所提方法的合理性和有效性, 与文献[19]所提考虑安全约束的优化方法进行对比分析。

方法 1: PPO 算法, 加入网络安全约束。

方法 2: 利用本文所提方法进行求解。

对比结果如表 11 所示。本文所提方法训练时间减少约 30 min, 平均电压偏差下降了 0.033, 电压越限时段减小为零, 其训练时间短、求解速度快, 运行总成本相比方法 1 增加不多(0.746%), 但更满足系统安全运行需求, 具有良好的实用性。

表 11 不同方法下的优化结果对比

Table 11 Comparison of optimization results under different methods

方法	训练时间/min	求解时间/s	越限时段数/个	平均电压偏差/p.u.	运行总成本/元
方法 1	63	0.80	3	0.039	947 286.31
方法 2	32	0.65	0	0.006	954 402.78

4.3 扩展算例分析

为验证本文所提方法的通用性和扩展性, 选取不同的算例并扩大系统规模进行分析, 采用图 12 所示“33 节点电网-26 节点热网”综合能源系统进行仿真验证, 网络模型及参数基于文献[32]进行修改, 系统总电热需求负荷及 PV 日出力有功数据如图 13 所示。基于本文方法验证上述扩展算例得到的调度结果, 如图 14 所示。表 12 为不同算例对模型性能影响的对比结果, 从求解速度的角度分析, 扩展算例的训练时间增加了约 30 min, 求解时间增加了 0.12 s, 虽然扩展算例的训练时间随着算例规模的扩大而增加, 但决策时间增加不大, 能够满足在

线应用的需要；从调度安全性的角度分析，该模型均不存在电压越限，平均电压偏差均在可接受范围内，调度结果能够满足系统的安全性要求。通过扩展算例分析表明，本文所提方法针对较大规模系统仍具有良好的适用性和扩展性。

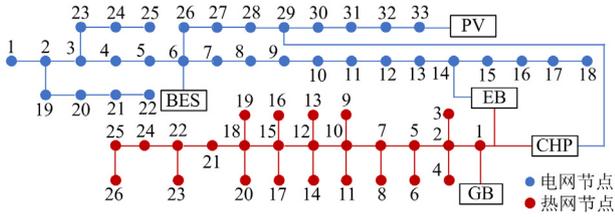


图 12 扩展算例拓扑结构

Fig. 12 Extended example topology

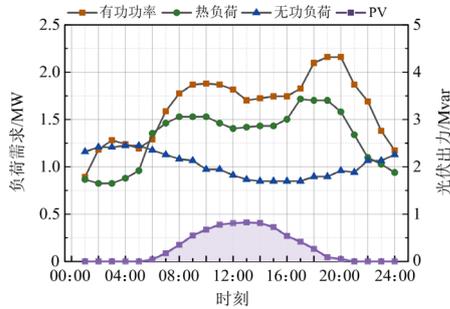
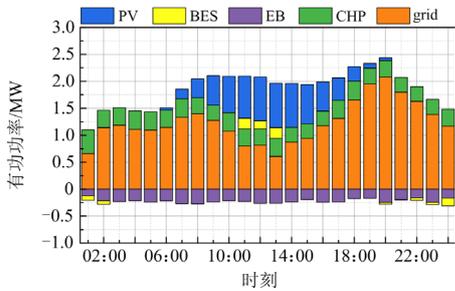
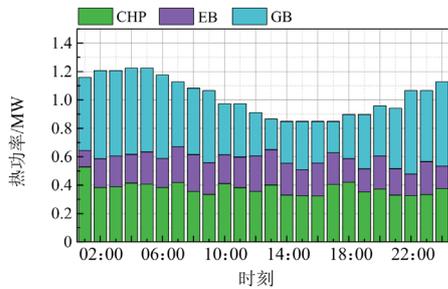


图 13 扩展算例总电热负荷需求与 PV 出力变化曲线

Fig. 13 Daily curves of total load and PV generation of extended example



(a) 扩展算例有功功率调度结果



(b) 扩展算例热功率调度结果

图 14 扩展算例调度结果

Fig. 14 Dispatch results of extended example

表 12 不同算例的模型性能对比

Table 12 Comparison of model performance between different examples

方法	训练时间/min	求解时间/s	越限时段数/个	平均电压偏差/p.u.
本文方法	32	0.65	0	0.006
扩展算例	61	0.77	0	0.019

5 结论

针对综合能源系统调度未能充分考虑结构化信息和真实化网络拓扑的问题，本文提出了一种基于图强化学习算法的 IES 动态经济调度方法。主要结论如下：

1) 从物理对象角度，由于考虑了网络拓扑约束和潮流限制，所提方法相比普通的深度强化学习方法，平均电压偏差更小，电压越限情况大幅减少，系统安全性能更高。

2) 从优化算法角度，所提方法能够在考虑异质网络拓扑与电热网络潮流的基础上实现综合能源系统的动态经济调度，加快了求解速度，且通过不同算例的优化调度结果证明本文方法具有较好的扩展性。

未来随着综合能源系统的发展，异质网络拓扑结构会变得更加复杂，基于人工智能技术指导园区进行更合理、更安全的动态经济调度具有重要意义，本文提出的图强化学习方法为实现这一目标提供了有益的方向，有助于推动 IES 在安全调度中发挥更大的价值。

参考文献

- [1] 王永真, 康利改, 张靖, 等. 综合能源系统的发展历程、典型形态及未来趋势[J]. 太阳能学报, 2021, 42(8): 84-95.
WANG Yongzhen, KANG Ligai, ZHANG Jing, et al. Development history, typical form and future of integrated energy system[J]. Acta Energiæ Solaris Sinica, 2021, 42(8): 84-95.
- [2] 汤东升, 钟伟, 慕小斌, 等. 基于鲁棒模型的多站融合场景下综合能源系统全局低碳策略[J]. 电测与仪表, 2022, 59(6): 113-121.
TANG Dongsheng, ZHONG Wei, MU Xiaobin, et al. Global low carbon strategy of integrated energy system in multi-station integration based on robust model[J]. Electrical Measurement & Instrumentation, 2022, 59(5): 113-121.
- [3] 焦亭, 白宇. 含电转热氢装置的综合能源系统经济运行[J]. 电测与仪表, 2022, 59(9): 1-7.

- JIAO Ting, BAI Yu. Economic operation of integrated energy system containing power to heat and hydrogen device[J]. *Electrical Measurement & Instrumentation*, 2022, 59(9): 1-7.
- [4] 吕佳炜, 张沈习, 程浩忠, 等. 考虑互联互通的区域综合能源系统规划研究综述[J]. *中国电机工程学报*, 2021, 41(12): 4001-4021.
- LÜ Jiawei, ZHANG Shenxi, CHENG Haozhong, et al. Review on district-level integrated energy system planning considering interconnection and interaction[J]. *Proceedings of the CSEE*, 2021, 41(12): 4001-4021.
- [5] 许剑, 刘松阳, 卢建生, 等. 基于递阶模型的综合能源系统多时间尺度运行优化研究[J]. *供用电*, 2023, 40(1): 73-79.
- XU Jian, LIU Songyang, LU Jiansheng, et al. Research on multi-time scale operation optimization of integrated energy system based on hierarchical model[J]. *Distribution & Utilization*, 2023, 40(1): 73-79.
- [6] 郭杰, 吴杰康, 洪明志, 等. 综合能源系统碳交易优化调度模型[J]. *供用电*, 2023, 40(11): 94-104.
- GUO Jie, WU Jiekang, HONG Mingzhi, et al. Optimal scheduling model of carbon trading in integrated energy system[J]. *Distribution & Utilization*, 2023, 40(11): 94-104.
- [7] 蔡颖凯, 张冶, 曹世龙, 等. 面向综合需求响应的综合能源系统优化调度[J]. *电网与清洁能源*, 2022, 38(9): 65-72.
- CAI Yingkai, ZHANG Ye, CAO Shilong, et al. Optimal scheduling of the integrated electricity and natural gas system considering the integrated demand response[J]. *Power System and Clean Energy*, 2022, 38(9): 65-72.
- [8] 贲保记, 张恩硕, 张国, 等. 考虑综合需求响应与“双碳”机制的综合能源系统优化运行[J]. *电力系统保护与控制*, 2022, 50(22): 11-19.
- YUN Baoji, ZHANG Enshuo, ZHANG Guo, et al. Optimal operation of an integrated energy system considering integrated demand response and a “dual carbon” mechanism[J]. *Power System Protection and Control*, 2022, 50(22): 11-19.
- [9] 汪飞, 龚丹丹, 郭慧, 等. 计及动态氢价和不确定性的区域综合能源系统规划—运行两阶段优化[J]. *电力系统保护与控制*, 2022, 50(13): 53-62.
- WANG Fei, GONG Dandan, GUO Hui, et al. Two-stage optimization of regional integrated energy system planning-operation with dynamic hydrogen pricing and uncertainties[J]. *Power System Protection and Control*, 2022, 50(13): 53-62.
- [10] 王忠义, 聂凡杰, 徐扬, 等. 计及综合需求响应的电热集成系统优化运行[J]. *智慧电力*, 2023, 51(1): 86-93.
- WANG Zhongyi, NIE Fanjie, XU Yang, et al. Optimal operation of integrated electro-thermal system considering comprehensive demand response[J]. *Smart Power*, 2023, 51(1): 86-93.
- [11] 夏洪伟, 李坤, 韩丽. 考虑风电预测误差的电-热系统混合时间尺度调度[J]. *电力系统保护与控制*, 2022, 50(17): 86-96.
- XIA Hongwei, LI Kun, HAN Li. Hybrid time-scale dispatch of an electric-heating system considering wind power forecast error[J]. *Power System Protection and Control*, 2022, 50(17): 86-96.
- [12] 赵鹏翔, 李振, 王楠, 等. 基于源荷双侧主从博弈的园区综合能源系统运行优化策略[J]. *电力系统及其自动化学报*, 2021, 33(9): 109-116, 122.
- ZHAO Pengxiang, LI Zhen, WANG Nan, et al. Operation optimization strategy for district integrated energy system based on Stackelberg game between supply and demand sides[J]. *Proceedings of the CSU-EPSA*, 2021, 33(9): 109-116, 122.
- [13] ZHANG Yao, HU Yuan, MA Jin, et al. A mixed-integer linear programming approach to security-constrained co-optimization expansion planning of natural gas and electricity transmission systems[J]. *IEEE Transactions on Power Systems*, 2018, 33(6): 6368-6378.
- [14] SUN Yonghui, ZHANG Bowen, GE Leijiao, et al. Day-ahead optimization schedule for gas-electric integrated energy system based on second-order cone programming[J]. *CSEE Journal of Power and Energy Systems*, 2020, 6(1): 142-151.
- [15] 孙志攀, 刘会兰, 栗然, 等. 考虑电动汽车用户意愿的热电联供型微电网优化调度[J]. *电力建设*, 2018, 39(9): 95-104.
- SUN Zhipan, LIU Huilan, LI Ran, et al. Optimal dispatch for microgrids with CHP considering EV users' intention[J]. *Electric Power Construction*, 2018, 39(9): 95-104.
- [16] 罗建勋, 张玮, 王辉, 等. 基于深度强化学习的微电网优化调度研究[J]. *电力学报*, 2023, 38(1): 54-63.
- LUO Jianxun, ZHANG Wei, WANG Hui, et al. Research on optimal scheduling of micro-grid based on deep reinforcement learning[J]. *Journal of Electric Power*, 2023, 38(1): 54-63.
- [17] 刘林鹏, 朱建全, 陈嘉俊, 等. 基于柔性策略-评价网络的微电网源储协同优化调度策略[J]. *电力自动化设备*, 2022, 42(1): 79-85.
- LIU Linpeng, ZHU Jianquan, CHEN Jiajun, et al. Cooperative optimal scheduling strategy of source and storage in microgrid based on soft actor-critic[J]. *Electric Power Automation Equipment*, 2022, 42(1): 79-85.
- [18] 杨志学, 任洲洋, 孙志媛, 等. 基于近端策略优化算法的新能源电力系统安全约束经济调度方法[J]. *电网技术*, 2023, 47(3): 988-998.
- YANG Zhixue, REN Zhouyang, SUN Zhiyuan, et al. Security-constrained economic dispatch of renewable

- energy integrated power systems based on proximal policy optimization algorithm[J]. Power System Technology, 2023, 47(3): 988-998.
- [19] 蔺伟山, 王小君, 孙庆凯, 等. 计及安全约束的综合能源系统深度强化学习优化调度策略研究[J]. 电网技术, 2023, 47(5): 1970-1983.
LIN Weishan, WANG Xiaojun, SUN Qingkai, et al. Optimal dispatch strategy of integrated energy system based on deep reinforcement learning considering security constraints[J]. Power System Technology, 2023, 47(5): 1970-1983.
- [20] 范邦稷, 刘兴华, 丁涛, 等. 基于图注意力网络与深度确定性策略梯度的三相主动配电网供电恢复方法[J/OL]. 中国电机工程学报: 1-16[2023-05-15]. DOI:10.13334/j.0258-8013.pcsee.221311.
FAN Bangji, LIU Xinghua, DING Tao, et al. A GAT-DDPG based approach for three-phase active distribution system restoration[J/OL]. Proceedings of the CSEE: 1-16[2023-05-15]. DOI:10.13334/j.0258-8013.pcsee.221311.
- [21] HOSSAIN R R, HUANG Qiuhua, HUANG Renke, et al. Graph convolutional network-based topology embedded deep reinforcement learning for voltage stability control[J]. IEEE Transactions on Power Systems, 2021, 36(5): 4848-4851.
- [22] SUN Qingkai, WANG Xiaojun, LIU Zhao, et al. Dynamic energy dispatch strategy for integrated energy system based on constrained reinforcement learning[J]. CSEE Journal of Power and Energy Systems, 2021, 36(5): 4848-4851.
- [23] 杨挺, 赵黎媛, 刘亚闯, 等. 基于深度强化学习的综合能源系统动态经济调度[J]. 电力系统自动化, 2021, 45(5): 39-47.
YANG Ting, ZHAO Liyuan, LIU Yachuang, et al. Dynamic economic dispatch for integrated energy system based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2021, 45(5): 39-47.
- [24] 黎静华, 朱梦姝, 陆悦江, 等. 综合能源系统优化调度综述[J]. 电网技术, 2021, 45(6): 2256-2272.
LI Jinghua, ZHU Mengshu, LU Yuejiang, et al. Review on optimal scheduling of integrated energy systems[J]. Power System Technology, 2021, 45(6): 2256-2272.
- [25] 史喆, 梁毅, 李华, 等. 计及灵活性多目标电-热-交通综合能源系统区间优化运行[J]. 电力系统保护与控制, 2022, 50(21): 33-42.
SHI Zhe, LIANG Yi, LI Hua, et al. Interval optimal operation of a multi-objective electric-thermal-transportation integrated energy system considering flexibility[J]. Power System Protection and Control, 2022, 50(21): 33-42.
- [26] ZHU Mengting, XU Chengsi, DONG Shufeng, et al. An integrated multi-energy flow calculation method for electricity-gas-thermal integrated energy systems[J]. Protection and Control of Modern Power Systems, 2021, 6(1): 65-76.
- [27] CHEN Junbin, YU Tao, PAN Zhenning, et al. A scalable graph reinforcement learning algorithm based stochastic dynamic dispatch of power system under high penetration of renewable energy[J]. International Journal of Electrical Power & Energy Systems, 2023, 152.
- [28] CHEN Xinpei, YU Tao, PAN Zhenning, et al. Graph representation learning-based residential electricity behavior identification and energy management[J]. Protection and Control of Modern Power Systems, 2023, 8(2): 464-476.
- [29] LI Zhigang, WU Wenchuan, WANG Jianhui, et al. Transmission constrained unit commitment considering combined electricity and district heating networks[J]. IEEE Transactions on Sustainable Energy, 2016, 7(2): 480-492.
- [30] XING Qiang, CHEN Zhong, ZHANG Tian, et al. Real-time optimal scheduling for active distribution networks: a graph reinforcement learning method[J]. International Journal of Electrical Power & Energy Systems, 2023, 152.
- [31] 郭剑波. 构建新型电力系统是实现能源转型、达成“双碳”目标的有效途径[N]. 国家电网报, 2021-09-07(005).
GUO Jianbo. Building a new type of power system is an effective way to achieve energy transformation and achieve the “dual carbon” goal[N]. State Grid News, 2021-09-07(005).
- [32] 张义志, 王小君, 和敬涵, 等. 考虑供热系统建模的综合能源系统最优能流计算方法[J]. 电工技术学报, 2019, 34(3): 562-570.
ZHANG Yizhi, WANG Xiaojun, HE Jinghan, et al. Optimal energy flow calculation method of integrated energy system considering thermal system modeling[J]. Transactions of China Electrotechnical Society, 2019, 34(3): 562-570.

收稿日期: 2023-06-30; 修回日期: 2023-11-12

作者简介:

吕金玲(1998—), 女, 硕士研究生, 研究方向为人工智能在综合能源系统中的应用; E-mail: 21121459@bjtu.edu.cn

王小君(1978—), 男, 通信作者, 博士, 教授, 研究方向为电力系统分析与控制、综合能源系统优化运行; E-mail: xjwang1@bjtu.edu.cn

窦嘉铭(1996—), 男, 博士研究生, 研究方向为综合能源与人工智能、群智涌现与群智演化。E-mail: djmmjddjmmjd@163.com

(编辑 许威)