

基于深度强化学习的氢能综合能源系统优化调度方法

张磊¹, 吴红斌¹, 何叶¹, 徐斌², 张明星³, 丁明¹

(1. 新能源利用与节能安徽省重点实验室(合肥工业大学), 安徽省合肥市 230009;

2. 国网安徽省电力有限公司电力科学研究院, 安徽省合肥市 230601;

3. 国网安徽省电力有限公司六安供电公司, 安徽省六安市 237000)

摘要: 为实现碳减排目标, 氢能与综合能源系统的结合成为最具潜力的发展方向之一。针对当前氢能综合能源系统调度策略灵活性不足、复杂系统多目标优化求解困难等问题, 提出一种基于深度强化学习的氢能综合能源系统优化调度方法。首先, 采用耦合设备的变工况模型, 构建风-光-氢-冷-热-电综合能源系统, 拓展设备联合供能空间。其次, 考虑系统运行成本、碳排放量、系统自供给平衡度和新能源利用率, 基于最优解距离构建多目标优化模型, 激发智能体探索性。然后, 通过时序片段表征优化深度强化学习算法, 增强了智能体对系统状态变化的估计精度。最后, 在源荷实测数据的基础上设计仿真算例。结果表明, 所提方法可以有效提高氢能综合能源系统调度的灵活性, 充分挖掘氢能的碳减排潜力, 实现调度经济性和环保性的双重优化。

关键词: 综合能源系统; 氢能; 优化调度; 深度强化学习; 多目标优化; 可再生能源

0 引言

为了实现碳减排目标, 必须深入推动能源系统的绿色革命, 将传统电力系统与新型能源形式耦合, 构建综合能源系统(integrated energy system, IES)^[1]。IES需要协调不同能量形式之间的耦合运行, 具备源荷双侧不确定性管理能力, 并兼顾系统运行的灵活性、高效性和经济性^[2-3]。因此, IES能量调度管理成为实现碳减排目标的关键技术。

氢能综合利用具备绿色低碳、储存运输便利、多能源整合等优势, 被视为最具潜力的 IES 发展方向之一^[4]。文献[5]基于阶梯式碳交易优化了两阶段电转气过程, 利用氢能转化设备实现了低碳性与经济性。文献[6]建立了含电转气的能量耦合设备梯级利用模型, 提出随机优化调度方法, 实现了氢能高品位利用。文献[7]构建燃料电池(fuel cell, FC)模型, 提出一种滚动时域双层调度方法, 验证了氢能系统作为备用能源的优势。现有文献为了避免高维非线性模型求解, 往往采用固定效率的设备模型, 但会

限制设备输出灵活性, 不利于系统能量调度。

为了平衡碳减排和经济运行之间的矛盾, 多目标优化成为制定 IES 调度方案的关键要素^[8]。文献[9]将成本、负荷曲线方差和可再生能源弃电量作为优化目标, 采用粒子群优化算法求解 Pareto 最优解集。文献[10]考虑系统经济性和新能源利用率, 采用非支配排序遗传算法进行多目标优化。文献[11]以成本最低和能效最高为目标, 采用改进 ϵ -约束法得到 Pareto 前沿集, 以实现多目标优化。但 Pareto 形式的多目标函数更关注前沿解集而不是某一个最优解, 限制了求解寻优算法的选择。

变工况模型会增加系统的非线性, 而多目标优化也会提高求解的复杂性^[12]。此外, 源侧新能源的深度参与、负荷侧需求形式的相互转化都会加深系统中的双端不确定性, 增加求解难度^[13]。传统方法在处理上述问题时难以避免对原始问题结构信息的改变, 或需要依赖先验知识进行模型假设, 进而导致调度性能不足^[14-15]。而具备无模型特征的深度强化学习(deep reinforcement learning, DRL)可以避免上述问题, 成为 IES 调度优化的研究热点。文献[16]测试了多种 DRL 方法, 验证了其解决综合能源调度问题的优越性。文献[17]采用深度确定性策略梯度(deep deterministic policy gradient, DDPG)进行连续状态和动作空间下的动态调度决策。文献[18]基于 DDPG 提出一种调峰模型以协调天然气机组和发电

收稿日期: 2024-01-02; 修回日期: 2024-03-21。

上网日期: 2024-07-03。

国家自然科学基金区域创新发展联合基金资助项目(U19A20106); 安徽省自然科学基金资助项目(2108085UD05); 安徽省科技重大专项计划资助项目(202203f07020003)。

机,平滑系统的净负荷曲线。虽然现有研究肯定了DRL方法相对于传统方法的有效性^[16-18],但IES中的源荷参数作为随机变量具有部分可观测性,导致传统DRL方法直接应用于IES调度时决策性能的下降。

针对上述问题,本文首先采用定工况与变工况模型构建风-光-氢-冷-热-电耦合的IES,分析模型差异对调度的影响;然后,考虑经济性、碳排放量、新能源消纳率和系统自给平衡度,基于最优解距离提出多目标优化函数;最后,采用样本时序片段表征改进双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)算法应对调度模型的非线性与不确定性,并训练智能体获得最终调度策略,实现调度计划在经济性和环保性上的双重优化。

1 风-光-氢-冷-热-电耦合系统建模

1.1 系统结构

本节构建了以质子交换膜燃料电池(proton exchange membrane fuel cell, PEMFC)为核心的冷-热-电联产机制,并建立了基于母线的IES模型^[19]。所研究的系统属于园区IES,不涉及能量传输对系统的影响,具体结构如图1所示。

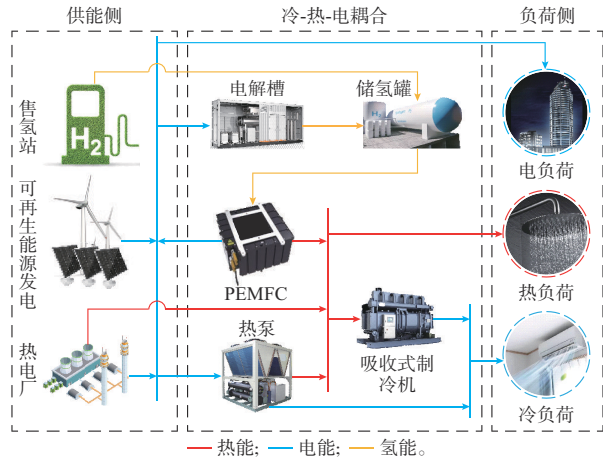


图1 IES结构
Fig.1 Structure of IES

该系统主要包含3个部分:能源供给部分、能源转换与储存部分和用能负荷。整个系统同时考虑用户的冷、热、电需求,以风力机(wind turbine, WT)和光伏(photovoltaic, PV)作为主要能量来源。热电厂和售氢站属于上级能源网,与IES连接作为保证负荷需求的能源后备。系统的中间部分主要存在两套完整的能量耦合机制:一是由电解槽(electric tank,

ET)、储氢罐、PEMFC构成的热-电联产机制;二是由热泵(heat pump, HP)和吸收式制冷机(absorption chiller, AC)构成的冷-热联供机制。

变工况设备模型的联合供能一方面可以更贴近现实的设备出力特性,另一方面也可以拓展系统供能空间,增加调度灵活性。以HP和PEMFC的联合供热为例,附录A图A1对比了定工况模型和变工况模型的出力差异。可以看出,变工况模型联合供热可以覆盖更大范围的热能需求。下文给出了所有耦合设备的变工况数学模型,而各设备对应的定工况模型形式见附录B。

1.2 PEMFC模型

PEMFC具有高效、低污染、低温操作和快速启动等优势,已经成为一种被广泛研究和应用的燃料电池技术。但现有调度研究中多采用定工况模型表达PEMFC的出力特征,如附录B式(B1)所示。为了更准确反映PEMFC的运行状态,采用半经验方程描述出力变化过程,构建PEMFC变工况数学模型^[20-22]。

PEMFC的输出形式包括电能与热能两部分。忽略空气散热的影响,电、热功率输出关系表示如下:

$$\begin{cases} P_{e,t}^{\text{FC}} = V_t^{\text{FC}} I_t^{\text{FC}} \\ P_{h,t}^{\text{FC}} = \mu_t H_{\text{H}_2} - P_{e,t}^{\text{FC}} \end{cases} \quad (1)$$

式中:下标 t 表示 t 时隙; $P_{e,t}^{\text{FC}}$ 为PEMFC的电功率; $P_{h,t}^{\text{FC}}$ 为输出的热功率; V_t^{FC} 为PEMFC的输出电压; I_t^{FC} 为输出电流; μ_t 为氢气质量流速; H_{H_2} 为氢气高位热值。

其中,电流与氢气质量流速有如下关系:

$$I_t^{\text{FC}} = \frac{2\mu_t F}{M_{\text{H}_2}} \quad (2)$$

式中: F 为法拉第常数; M_{H_2} 为氢气摩尔质量。

输出电压主要由能斯特开路电压、活化极化过电压、欧姆过电压和浓差过电压组成,表达式如下:

$$V_t^{\text{FC}} = N_{\text{cell}} (V_t^{\text{N}} - V_t^{\text{ohm}} - V_t^{\text{act}} - V_t^{\text{con}}) \quad (3)$$

式中: N_{cell} 为燃料电池组包含的电池数目; V_t^{N} 为能斯特开路电压; V_t^{ohm} 为欧姆过电压损耗; V_t^{act} 为活化极化过电压损耗; V_t^{con} 为浓差过电压损耗。上述4种电压均可以由 t 时隙的电流计算得到,具体公式见附录B式(B2)至式(B10)。

根据式(1)至式(3)可以看出,PEMFC的电、热输出功率可以转化为和氢气质量流速直接相关的非线性变工况模型。该模型的电、热效率由氢气质量流速决定,能够充分反映PEMFC的柔性特征。

1.3 ET模型

系统采用质子交换膜ET,能量转化效率受设备额定容量和输入功率的影响,可由式(4)计算得到^[23]:

$$P_i^{H_2} = P_{e,t}^{ET} \left[\zeta_1^{ET} \left(\frac{P_{e,t}^{ET}}{P_{rated}^{ET}} \right)^2 + \zeta_2^{ET} \left(\frac{P_{e,t}^{ET}}{P_{rated}^{ET}} \right) + \zeta_3^{ET} \right] \quad (4)$$

式中: $P_i^{H_2}$ 为ET产出氢气的等效电能; $P_{e,t}^{ET}$ 为ET的输入功率; P_{rated}^{ET} 为ET的额定输入功率; ζ_1^{ET} 、 ζ_2^{ET} 、 ζ_3^{ET} 为拟合曲线参数。

1.4 HP模型

HP拥有制冷和制热两种工作模式,其本质是逆卡诺循环的不同应用,制冷和制热能力即为HP搬运热能的能力,其数学模型如下所示^[24]:

$$P_i^{HP} = P_{e,t}^{HP} \eta \left[\zeta_1^{HP} + \zeta_2^{HP} \frac{P_{e,t}^{HP}}{P_{rated}^{HP}} + \zeta_3^{HP} \left(\frac{P_{e,t}^{HP}}{P_{rated}^{HP}} \right)^2 + \zeta_4^{HP} \left(\frac{P_{e,t}^{HP}}{P_{rated}^{HP}} \right)^3 \right] \quad (5)$$

$$\begin{cases} P_{h,t}^{HP} = P_i^{HP} (1 - M) \\ P_{c,t}^{HP} = P_i^{HP} M \end{cases} \quad (6)$$

式中: P_i^{HP} 为HP搬运热量的功率; $P_{e,t}^{HP}$ 为HP电功率; η 为额定功率下的最大性能系数; P_{rated}^{HP} 为HP额定功率; ζ_1^{HP} 、 ζ_2^{HP} 、 ζ_3^{HP} 、 ζ_4^{HP} 为经验拟合系数; $P_{h,t}^{HP}$ 为HP提供的热能; $P_{c,t}^{HP}$ 为HP提供的冷能; M 表示HP工作模式的0-1变量,取值为1表示制冷,取值为0表示制热。

1.5 AC模型

AC可以实现低品位热能利用,将冗余热能用于满足用户的制冷需求,其模型如下所示^[25]:

$$P_{c,t}^{AC} = P_{h,t}^{AC} \left[\zeta_1^{AC} + \zeta_2^{AC} \frac{P_{h,t}^{AC}}{P_{rated}^{AC}} + \zeta_3^{AC} \left(\frac{P_{h,t}^{AC}}{P_{rated}^{AC}} \right)^2 + \zeta_4^{AC} \left(\frac{P_{h,t}^{AC}}{P_{rated}^{AC}} \right)^3 \right] \quad (7)$$

式中: $P_{c,t}^{AC}$ 为AC搬运热量的功率; $P_{h,t}^{AC}$ 为AC的有功功率; P_{rated}^{AC} 为AC的额定功率; ζ_1^{AC} 、 ζ_2^{AC} 、 ζ_3^{AC} 、 ζ_4^{AC} 为经验拟合系数。

2 多目标优化调度模型

优化目标的选择对调度问题的求解至关重要。由于氢能设备高昂的成本和新能源的不确定性,单目标优化会导致DRL制定的调度计划在碳减排和经济运行之间失衡。为了引导智能体在学习过程中

平衡碳减排和经济运行,本文采用多目标优化构建调度模型,提出以下4种调度结果评估指标。

2.1 评估指标

2.1.1 系统运行成本 C_t

运行成本 C_t 主要包含考虑全生命周期的设备运行成本、外部购能成本和储氢容量变化成本,表示如下:

$$C_t = \sum_{n_D=1}^{N_D} C_{n_D}^{OP} P_{n_D,t} + \sum_{n_E=1}^{N_E} C_{n_E}^E P_{n_E,t} + C_{H_2}^S P_{Tank}^{RC} (c_{t-1}^{H_2} - c_t^{H_2}) \quad (8)$$

式中: $C_{n_E}^E$ 为系统向外部购买第 n_E 种能源的单位能量成本,分时电价见附录C表C1,其余能源价格视为恒定; N_E 为系统内能源形式的总数; $P_{n_E,t}$ 为从外部购入的第 n_E 种能量量; $C_{n_D}^{OP}$ 为第 n_D 种设备的平准化单位能量运行成本,可由附录D式(D1)和式(D2)计算得到^[26]; N_D 为系统内设备数量总数; $P_{n_D,t}$ 为第 n_D 种设备的输入或输出能量; $C_{H_2}^S$ 为系统内部产生1 kW·h等效氢气消耗的成本; P_{Tank}^{RC} 为储氢系统的额定等效电能容量; $c_t^{H_2}$ 为储氢系统容量百分比。 $c_{t-1}^{H_2} - c_t^{H_2}$ 表示氢能的容量变化,其值为负表示容量增加,系统获得了收益。

2.1.2 系统碳排放量 $E_t^{CO_2}$

本文系统中碳排放主要来源于热电厂,对电网供电和热网供热产生的碳排放量计算如下:

$$E_t^{CO_2} = \frac{\beta_1 P_{e,t}^{grid}}{1 - \alpha_1} + \frac{\beta_2 P_{h,t}^{grid}}{1 - \alpha_2} \quad (9)$$

式中: $P_{e,t}^{grid}$ 和 $P_{h,t}^{grid}$ 分别为由热电厂提供的电能和热能; α_1 、 α_2 为网络传输的能量损耗系数; β_1 、 β_2 为单位能量碳排放系数。

2.1.3 供需自给不平衡度 R_t^{im}

对于包含新能源设备的IES而言,系统可以实现自给自足且不存在弃风/弃光量是最理想的运行状态。但由于系统中的不确定性,实际运行状况往往与调度计划存在偏差。因此, R_t^{im} 被用来反映调度结果和理想运行状态的偏离程度, R_t^{im} 越低,则调度方案性能越好,其计算公式如下:

$$R_t^{im} = \frac{|P_{e,t}^{wast} - P_{e,t}^{grid}| + |P_{h,t}^{wast} - P_{h,t}^{grid}| + |P_{c,t}^{im}| + |P_{H_2,t}^{wast} - P_{H_2,t}^{pur}|}{L_{e,t} + L_{h,t} + L_{c,t}} \quad (10)$$

式中: R_t^{im} 为系统内负荷不平衡程度; $P_{H_2,t}^{pur}$ 为来自上级能源网络的氢气; $P_{e,t}^{wast}$ 、 $P_{h,t}^{wast}$ 和 $P_{H_2,t}^{wast}$ 分别为调度方案对于实际需求产生的冗余电能、冗余热能、冗余氢能; $P_{c,t}^{im}$ 为调度计划偏差导致的冷负荷供需不平衡

量; $L_{e,t}$ 、 $L_{h,t}$ 、 $L_{c,t}$ 分别为用户的电、热、冷负荷需求。

2.1.4 新能源利用率 R_t^{RE}

R_t^{RE} 主要反映调度方案对新能源的利用程度,其值越大,则调度方案性能越好,表示如下:

$$R_t^{RE} = \frac{P_t^{PV} + P_t^{WT}}{P_t^{PV, \max} + P_t^{WT, \max}} \quad (11)$$

式中: P_t^{PV} 和 P_t^{WT} 分别为PV、WT在调度计划中的实际出力; $P_t^{PV, \max}$ 和 $P_t^{WT, \max}$ 分别为PV、WT的最大可能出力。

2.2 多目标优化的目标函数

评估指标是构建多目标优化目标函数的基础,而DRL方法需要获得一个单一的奖励值来训练智能体。因此,采用优劣解距离法(technique for order preference by similarity to ideal solution, TOPSIS)的思想整合4种指标以构建目标函数。

TOPSIS需要预先给出调度结果的最优评估值和最劣评估值,通过计算当前调度结果和最优、最劣解的距离进行方案性能排序^[27]。但TOPSIS直接应用于DRL方法会限制智能体在训练阶段的探索性,分析过程见附录D第D2章。因此,本文仅选取最优解距离按式(12)定义了目标函数 f ,函数值越大,则调度方案效果越好。

$$f(I_t) = -\|I_t - I_{\text{best}}\|_2 \quad (12)$$

式中: I_t 为调度方案在指标空间中的高维点; I_{best} 为调度方案的理论最优解。

运行成本和碳排放量难以确定其可能的边界,且不一致的指标量纲会影响空间距离的计算。因此,采用Sigmoid函数对这两个指标进行变换,将 C_t 与 $E_t^{\text{CO}_2}$ 的数值映射到介于0-1之间的 C_t^* 与 $E_t^{\text{CO}_2*}$,具体见附录D式(D4)。此时,在 t 时隙可以计算得到一个四维的指标 $I_t = [C_t^*, E_t^{\text{CO}_2*}, R_t^{\text{im}}, R_t^{\text{RE}}]$,并有最优解 $I_{\text{best}} = [0, 0, 0, 1]$ 。模型约束条件见附录D第D4章。

3 基于时序片段表征的DRL调度方法

3.1 马尔可夫决策过程

为了满足系统的能量需求,综合能源调度需要管理与分配多种能源资源,故引入DRL方法来应对大规模复杂问题的求解。该方法通过建立马尔可夫决策过程来描述智能体与环境的交互,基本理论与交互过程见附录E第E1章^[28]。本文奖励函数与目标函数式(12)一致。

3.2 系统状态的时序片段表征

在本文研究中, t 时隙的系统状态包括天气参数、负荷参数、时间参数和储氢系统容量百分比。 t

时隙的系统状态 s_t 可以表示为:

$$s_t = \{W_t^{\text{GHI}}, W_t^{\text{WS}}, W_t^{\text{TEM}}, L_{e,t}, L_{h,t}, L_{c,t}, N_t^{\text{day}}, N_t^{\text{hour}}, c_t^{\text{H}_2}\} \quad (13)$$

式中: W_t^{GHI} 为全球水平辐照度(global horizontal irradiance, GHI); W_t^{WS} 为风速; W_t^{TEM} 为温度; N_t^{day} 和 N_t^{hour} 分别为日编号和日内时隙编号。

决策动作 a_t 为主要设备运行状态,可表示为:

$$a_t = \{A_t^{\text{ET}}, A_t^{\text{PV}}, A_t^{\text{WT}}, A_t^{\text{HP}}, A_t^{\text{HPM}}, A_t^{\text{AC}}, \mu_t\} \quad (14)$$

式中: A_t^{ET} 为ET的负载率; A_t^{PV} 为PV出力百分比; A_t^{WT} 为WT出力百分比; A_t^{HP} 为HP负载率; A_t^{HPM} 为HP工作模式; A_t^{AC} 为AC负载率。

DRL通常假设系统满足马尔可夫决策原则,即下一时隙的状态仅由当前时隙的状态决定,不依赖于以往任何状态。因此,从信息熵的角度分析,DRL任务的假设可写为:

$$H(S'|s) = -\sum_{i=1}^n P(s'_i|s) \log_2(P(s'_i|s)) = 0 \quad (15)$$

式中: S' 为下一时隙所有可能状态的集合; $P(s'_i|s)$ 为 s 状态下一时隙状态为集合中第 i 个状态 s'_i 的概率; n 为用于计算信息熵的样本总数; $H(S'|s)$ 为信息熵,其越接近0,则状态信息越完备,DRL性能越好。

然而,在IES调度问题中存在着较强的状态参数时序相关性,即 $H(S'|s)$ 值较高,直接应用DRL的调度性能并不理想。通过信息熵分析可以看出,提高系统状态中信息的完备程度是应对该问题的有效方案之一^[29]。

天气与负荷状态值的时序相关性是随着时间步长的增加而递减的。可以认为,对于 t 时隙的时序状态参数仅受到 $t-1$ 至 $t-n$ 历史时隙的显著影响, $t-n$ 之前历史时隙的影响可以忽略。因此,可以以时间序列片段代替单一数值来表征特定时隙的系统状态,提高状态信息的完备程度。

在本文考虑的调度问题中没有额外预测模型的支持,在做 t 时隙动作决策时无法提取、获得 t 时隙的系统状态,仅能依据历史信息进行决策。以GHI为例,时序片段表征可以写为:

$$W_t^{\text{GHI}*} = \{W_{t-n}^{\text{GHI}}, W_{t-n+1}^{\text{GHI}}, \dots, W_{t-1}^{\text{GHI}}\} \quad (16)$$

此时, t 时隙系统的时序片段状态可以表示为:

$$s_t^* = \{W_t^{\text{GHI}}, W_t^{\text{WS}}, W_t^{\text{TEM}}, L_{e,t}, L_{h,t}, L_{c,t}, N_t^{\text{day}}, N_t^{\text{hour}}, c_t^{\text{H}_2}\} \quad (17)$$

采用TD3算法进行训练,该方法属于连续动作决策方法,满足高精度动作控制的需求^[30-31]。方法细节见附录E第E2章。针对所研究的问题,所提时序片段表征的TD3调度方法的训练流程如图2所示。

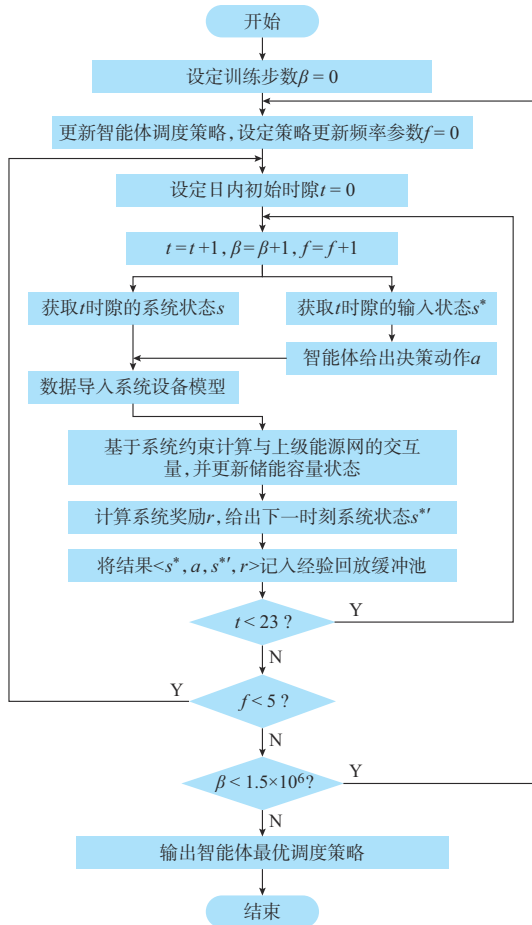


图2 DRL训练过程流程图
Fig. 2 Flow chart of DRL training process

4 算例分析

4.1 数据描述

所用负荷与天气数据均来自中国安徽省合肥市某园区,原始数据集时间范围为1年,分辨率为1h,包含GHI、风速、温度、电/热/冷负荷需求和采样时刻信息。测试集共计28d,分别取自四季中的某一星期,训练集为剩余全部数据。后续分析中的典型夏季日和冬季日分别为测试集中的13d和23d。系统中关键设备参数见附录F表F1。所有原始数据都进行了归一化处理,归一化信息见附录F表F2,统计规律如附录F图F1所示。

为验证所提方法的有效性,分别设置6组对比方案,其中,TS-TD3代表所提方案。各方案细节如表1所示。表2显示了各方案最优调度策略在测试集上的日均性能指标。为了更直观地反映新能源发电情况,将新能源利用率转化为新能源发电量进行比较。各方案计算耗时见附录F第F2章。

表1 各方案具体细节
Table 1 Specific details of each scheme

方案	设备模型	求解算法	时序片段	多目标方法
TS-TD3	变工况	TD3	有	最优解距离
方案1	定工况	TD3	有	最优解距离
方案2	变工况	TD3	有	TOPSIS
方案3	变工况	TD3	无	最优解距离
方案4	变工况	DDPG	有	最优解距离
方案5	变工况	DDPG	有	最优解距离

表2 各方案调度性能平均指标

Table 2 Average indicators of scheduling performance for each scheme

方案	日均奖励值	总成本/元	CO ₂ 排放量/kg	自供给不平衡度	新能源发电量/(kW·h)
TS-TD3	-17.0	1 933.7	78.5	0.339	1 504.800
方案1	-24.3	672.2	259.3	0.762	1 390.600
方案2	13.8	737.4	224.4	0.795	1 509.944
方案3	-23.5	2 579.5	108.1	0.505	296.900
方案4	-20.3	2 138.1	106.8	0.509	1 291.900
方案5	-19.6	2 053.9	118.1	0.473	1 355.800

4.2 变工况模型对调度结果的影响性分析

从表2中方案1和TS-TD3方案的对比可以看出采用定工况模型后,日均奖励值下降高达42.9%。在各项调度指标中,尽管总运行成本有着明显的下降,但对应的碳排放量和自供给不平衡度大幅增加,新能源发电量明显缩减。这意味着方案1的调度策略存在严重的能量分配不合理性,系统会产生大量冗余能量或过度依赖于上级能源网,进而导致碳排放量大幅增加。

选取测试集中夏季、冬季典型日,对两种方案的

调度结果进行分析,各方案调度结果的能量耦合关系如图3所示。图中:每个矩形表示IES中的能量交互环节;每个矩形左侧的数字表示流向该环节的能量,而右侧的数字则表示流向其他环节的能量,单位为kW·h。从图3(a)和(b)的对比可以看出,定工况模型中ET、FC、HP和AC出力显著减少,智能体倾向于减少对耦合设备的调用。作为电-热耦合核心的PEMFC,其出力接近于0,以此为例展开分析。

高额运行成本会减少智能体调用该设备的倾

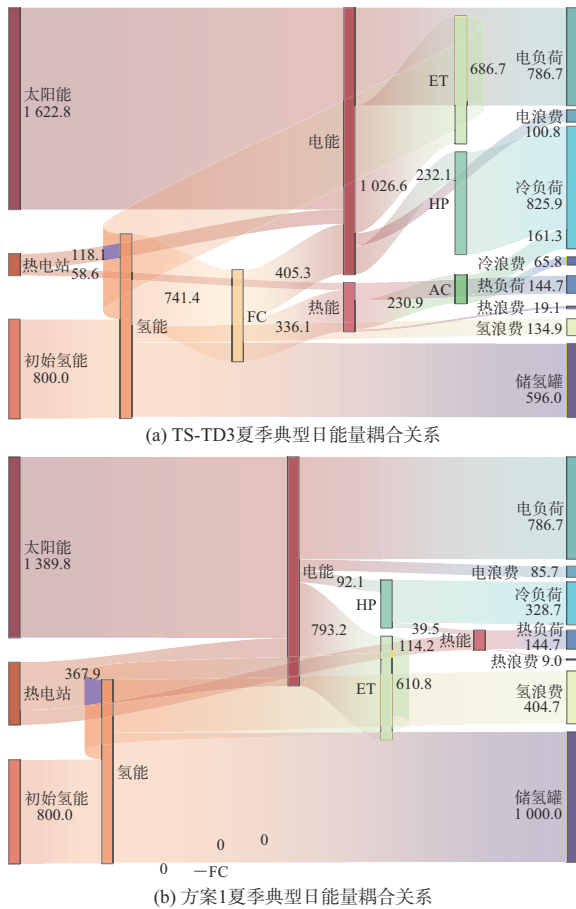


图3 夏季典型日调度结果的能量耦合关系

Fig. 3 Energy coupling relationship of scheduling results in a typical day of summer

向。对于PEMFC而言,其成本主要由两方面组成:一方面,设备本身的购买价格十分昂贵,导致平准化单位能量运行成本折算得到的单位能量运行成本较高;另一方面,绿氢在生产过程中使用新能源电力,且存在能量损耗,这导致氢能单价会显著高于热电厂供能单价。但是,PEMFC作为电-热耦合核心与氢储能利用关键环节,具有灵活的电-热输出机制和碳减排能力。为了优化碳排放指标,智能体最终策略应当在碳排放与成本之间取得平衡。

然而,方案1采用的定工况模型打破了这种平衡,僵化的运行模式限制了PEMFC调度灵活性。设备工作效率固定后,联合供能范围受到了限制,抑制了碳减排能力的发挥。因此,智能体出于成本考虑,选择近乎完全弃用PEMFC。作为PEMFC缺失供能部分的补充,方案1在夏季典型日中热电站的总体出力增加了172.8%。

全耦合设备的定工况模型会进一步加剧这种失衡。这样的现象在冬季典型日中再一次得到验证,方案1的智能体同样抑制了耦合设备的调用,如附

录F图F2所示。

在整个测试集上,TS-TD3方案相较于方案1,碳排放减少69.7%,自供给不平衡度减少42.3%,新能源发电量增加8.2%。可以证明,对于综合能源调度问题,更为精细化的变工况模型可以有效提高DRL调度策略性能,降低系统碳排放,提高系统自给平衡度和新能源发电量。

4.3 多目标方法对调度结果的影响性分析

为了与方案2进行比较,将TS-TD3的调度结果按TOPSIS重新计算后,日均奖励值为16.02,相比于方案2提高了16.1%。从表2中TS-TD3和方案2的4种性能指标对比可以看出,尽管原因不同,但方案2的调度结果陷入了和方案1相似的失衡状态。

选取测试集中的夏季、冬季典型日进行细致分析,调度结果如图4所示。以冬季为例,从图4(a)和(b)的供需曲线对比可以看出,两种方案的调度结果都可以大致分为0~5 h、6~16 h和17 h以后3个时段。

方案2在0~5 h制定了高额的风电出力计划,却没有调用足够的ET出力对冗余电能进行转化,导致了高额的供需不平衡。在6 h以后,由于该日的风能资源丰富,系统供能基本由风电满足,而储氢罐容量逐渐饱和,大量冗余能量被浪费。在17 h以后的时段,大量依赖热网供热,导致系统碳排放增加。智能体这样的调度计划会大幅度降低成本,同时提高新能源发电量,但相应的系统自供给不平衡度和碳排放量会大幅增加。

TS-TD3方案在0~5 h基本维持供需平衡,尽管该时段存在丰富的风能资源,但TS-TD3方案的智能体并没有盲目增加WT的出力。在6~16 h的时段中,过量的风资源和储氢容量的限制同样导致大量冗余能量。但相比于方案2,TS-TD3方案的能量浪费减少了37.7%。而在17 h以后,智能体通过PEMFC进行供热,在满足负荷需求的同时,实现了碳减排。

IES调度的不同目标之间往往存在冲突,寻求各目标之间的均衡性是多目标优化的关键。DRL需要经历一个试错学习过程,尝试尽可能多的方向,明确的方案性能排序会过早地确定智能体的求解方向。而TOPSIS的直接应用会促使智能体偏向于某几个指标,忽略各指标之间的均衡。正是因为这个原因,方案2的智能体在制定调度计划时会显著倾向于降低系统运行成本,提高新能源发电量,降低了对供需自平衡度和碳排放量的关注。同样的现象可以在夏季典型日调度结果中得到印证,如附录F图

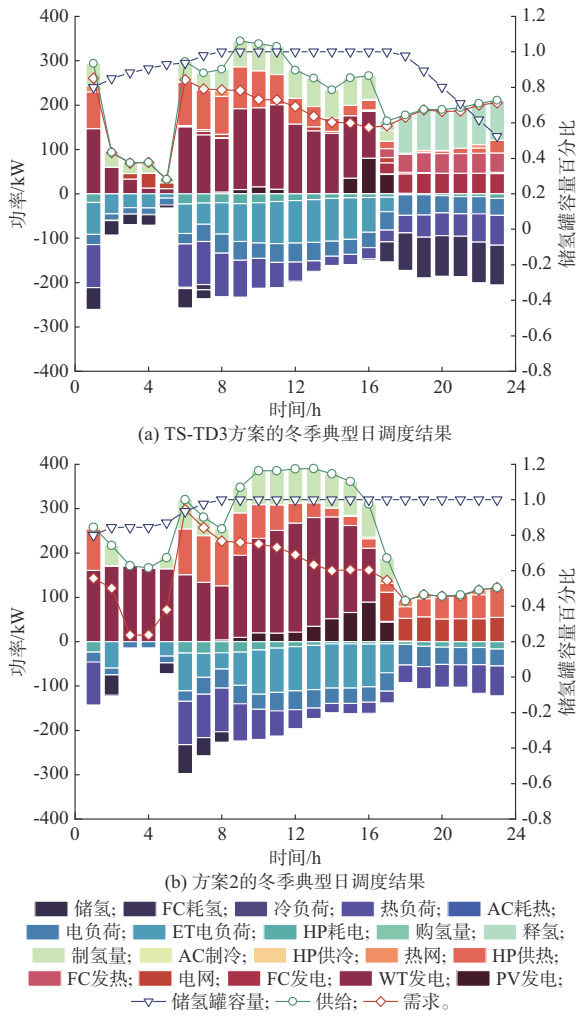


图4 冬季典型日调度结果
Fig. 4 Scheduling results of a typical day in winter

F3所示。可以看出,方案2的智能体在夏季典型日中依然大幅限制耦合设备的出力,甚至在17 h以后放弃了对冷负荷的供给。

由表2可以看出,TS-TD3方案更充分地利用了IES的能量转化与调度潜力。相比于方案2,其碳排放量减少了65%,系统自供给不平衡度降低了45.6%。由此可见,基于最优距离构建的多目标方法比TOPSIS更适合DRL优化调度,有效提高了智能体对求解空间的探索能力。

4.4 时序片段表征对调度结果的影响性分析

为了明确TD3算法本身的特征,以及时序片段表征结合TD3算法的优势,设计方案3至5与TS-TD3方案进行对比,分析不同算法的差异性。传统的DDPG方法被视为基准方法,并通过方案4探究了时序片段表征和DDPG结合的可能性。各方案在测试集上的日均平均奖励值变化曲线与详细分析见附录F图F4。由表2可见,TS-TD3方案相对于

方案3至5,奖励值分别提高了27.7%、16.3%和13.3%。在4项分指标上,TS-TD3方案全面优于方案3至5。

图5所示为各方案在测试集上的性能指标,以不同颜色对测试日所属季节进行了区分。从图上可以直观观察到,TS-TD3方案在整个测试集上具有普遍的性能优势。

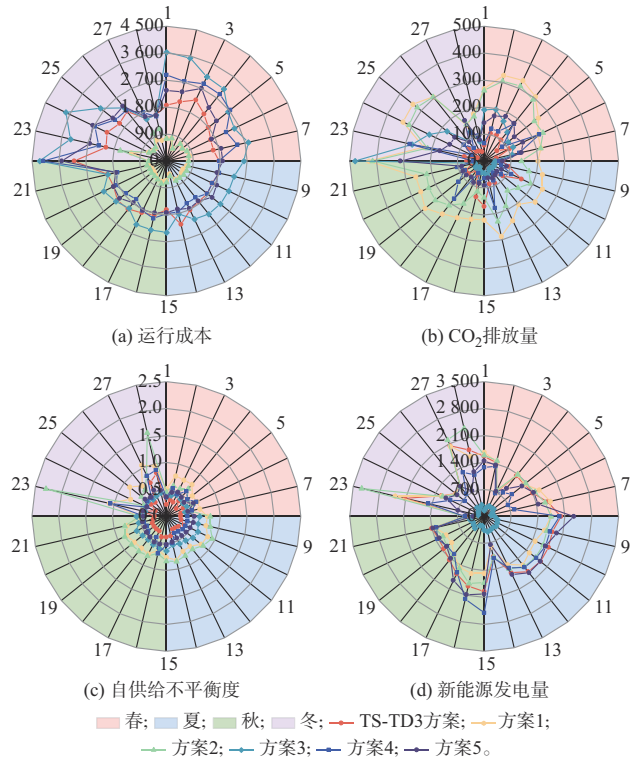


图5 各方案在28日测试集上的性能指标
Fig. 5 Performance indicators of each scheme on 28-day test set

从图5(d)可以看出,方案3对新能源的利用十分有限,其日均新能源发电量比TS-TD3方案降低了80.3%。表2显示,虽然方案3对新能源的利用极为有限,但相比于方案1和2,碳排放和自供给不平衡度指标得到了控制。充分发挥耦合设备的性能实现了这一点,而代价是方案3的运行成本最高,整体调度性能评价最低。

本文中各方案的智能体输入均为历史数据,这就意味着智能体需要在未来时隙新能源发电量和负荷需求未知的情况下制定调度方案。智能体不得不在反复多次的试错学习中,通过状态转移概率函数实现对未来源荷变化趋势的估计。而源荷变化具有很大的不确定性,由于传统的TD3算法中双Critic网络机制会倾向于更保守的调度决策,故在学习过程中,方案3的智能体选择通过尽可能减少对新能源的利用来确保动作值函数的稳定。而时序片段表

征正是针对这一点做出了改进。较高的信息完备程度减少了源荷不确定性对智能体的影响,使智能体敢于制定更为积极的调度计划。

方案4和5在各项指标上都较为接近,在运行成本、CO₂排放量和新能源发电量方面具备较为明显的季节性特征。在夏秋两季的指标性能普遍优于春冬两季。这是因为风能比太阳能的不确定性更强,而夏秋两季太阳能资源充沛,春冬两季风能资源充沛,故春冬季的系统不确定性更强。DDPG倾向于更激进的调度决策,更高的信息完备程度反而激化了这种倾向。因此,方案4的性能要整体劣于方案5。相比之下,TS-TD3方案的季节性差异更小,调度策略也更为稳健。

综上所述,时序片段表征和TD3算法的结合可以有效提高智能体对系统不确定性的估计能力和调度策略的精确程度,进而提升整个系统的多目标优化效果。

5 结语

为提高氢能IES的调度性能,优化不同能量供需之间的耦合机制,本文提出了一种基于DRL的IES优化调度方法,通过理论分析和仿真验证得出以下结论:

1)多能耦合设备联合供能的变工况模型可以提高系统运行的灵活性,结合DRL算法优势,有利于系统降碳和新能源消纳。

2)基于最优解距离构建的多目标优化函数更适合应用于DRL算法,提升智能体对氢综合能源系统调度潜力的挖掘。

3)时序片段表征可以提高TD3算法对系统不确定性的估计能力,进而从运行成本、碳排放量、自供给平衡性和新能源消纳方面全面提升智能体所制定调度策略的性能。

后续研究可对智能体网络模型进行有针对性的选择,分析不同网络类型应对调度问题的特性,进一步提高智能体学习能力。

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>),扫英文摘要后二维码可以阅读网络全文。

参考文献

- [1] 陈胜,张景淳,卫志农,等.面向能源转型的电-气-氢综合能源系统规划与运行[J].电力系统自动化,2023,47(19):16-30.
CHEN Sheng, ZHANG Jingchun, WEI Zhinong, et al. Energy transition oriented planning and operation of electricity-gas-hydrogen integrated energy system [J]. Automation of Electric Power Systems, 2023, 47(19): 16-30.
- [2] 曾鸣,刘英新,周鹏程,等.综合能源系统建模及效益评价体系综述与展望[J].电网技术,2018,42(6):1697-1708.
ZENG Ming, LIU Yingxin, ZHOU Pengcheng, et al. Review and prospects of integrated energy system modeling and benefit evaluation [J]. Power System Technology, 2018, 42(6): 1697-1708.
- [3] BERJAWI A E H, WALKER S L, PATSIOS C, et al. An evaluation framework for future integrated energy systems: a whole energy systems approach [J]. Renewable and Sustainable Energy Reviews, 2021, 145: 111163.
- [4] 郜捷,宋洁,王剑晓,等.支撑中国能源安全的电氢耦合系统形态与关键技术[J].电力系统自动化,2023,47(19):1-15.
GAO Jie, SONG Jie, WANG Jianxiao, et al. Form and key technologies of integrated electricity-hydrogen system supporting energy security in China [J]. Automation of Electric Power Systems, 2023, 47(19): 1-15.
- [5] 陈锦鹏,胡志坚,陈颖光,等.考虑阶梯式碳交易机制与电制氢的综合能源系统热电优化[J].电力自动化设备,2021,41(9):48-55.
CHEN Jinpeng, HU Zhijian, CHEN Yingguang, et al. Thermoelectric optimization of integrated energy system considering ladder-type carbon trading mechanism and electric hydrogen production [J]. Electric Power Automation Equipment, 2021, 41(9): 48-55.
- [6] 高玉,王琦,陈严,等.考虑需求响应和能量梯级利用的含氢综合能源系统优化调度[J].电力系统自动化,2023,47(4):51-59.
GAO Yu, WANG Qi, CHEN Yan, et al. Optimal dispatch of integrated energy system with hydrogen considering demand response and cascade energy utilization [J]. Automation of Electric Power Systems, 2023, 47(4): 51-59.
- [7] HAGGI H, SUN W, FENTON J M, et al. Proactive rolling-horizon-based scheduling of hydrogen systems for resilient power grids [J]. IEEE Transactions on Industry Applications, 2022, 58(2): 1737-1746.
- [8] 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.
- [9] 马燕峰,谢家荣,赵书强,等.考虑园区综合能源系统接入的主动配电网多目标优化调度[J].电力系统自动化,2022,46(13):53-61.
MA Yanfeng, XIE Jiarong, ZHAO Shuqiang, et al. Multi-objective optimal dispatching for active distribution network considering park-level integrated energy system [J]. Automation of Electric Power Systems, 2022, 46(13): 53-61.
- [10] 曾博,徐富强,刘裕,等.考虑可再生能源与需求响应协同增效的能量枢纽多目标区间优化规划方法[J].中国电机工程学报,2021,41(21):7212-7225.
ZENG Bo, XU Fuqiang, LIU Yu, et al. Multi-objective interval optimization approach for energy hub planning with consideration of renewable energy and demand response

- synergies[J]. Proceedings of the CSEE, 2021, 41(21): 7212-7225.
- [11] 丁煜蓉, 陈红坤, 吴军, 等. 计及综合能效的电-气-热综合能源系统多目标优化调度[J]. 电力系统自动化, 2021, 45(2): 64-73. DING Yurong, CHEN Hongkun, WU Jun, et al. Multi-objective optimal dispatch of electricity-gas-heat integrated energy system considering comprehensive energy efficiency[J]. Automation of Electric Power Systems, 2021, 45(2): 64-73.
- [12] 陈晚晴, 穆云飞, 贾宏杰, 等. 考虑设备变工况特性的区域综合能源系统优化调度方法[J]. 电网技术, 2021, 45(3): 951-958. CHEN Wanqing, MU Yunfei, JIA Hongjie, et al. Operation optimization method for regional integrated energy system considering part-load performances of devices [J]. Power System Technology, 2021, 45(3): 951-958.
- [13] 彭春华, 郑聪, 陈婧, 等. 基于置信间隙决策的综合能源系统鲁棒优化调度[J]. 中国电机工程学报, 2021, 41(16): 5593-5604. PENG Chunhua, ZHENG Cong, CHEN Jing, et al. Robust optimal dispatching of integrated energy system based on confidence gap decision[J]. Proceedings of the CSEE, 2021, 41(16): 5593-5604.
- [14] 聂欢欢, 张家琦, 陈颖, 等. 基于双层强化学习方法的多元园区实时经济调度[J]. 电网技术, 2021, 45(4): 1330-1336. NIE Huanhuan, ZHANG Jiaqi, CHEN Ying, et al. Real-time economic dispatch of community integrated energy system based on a double-layer reinforcement learning method [J]. Power System Technology, 2021, 45(4): 1330-1336.
- [15] 蔺伟山, 王小君, 孙庆凯, 等. 不确定性环境下基于深度强化学习的综合能源系统动态调度[J]. 电力系统保护与控制, 2022, 50(18): 50-60. LIN Weishan, WANG Xiaojun, SUN Qingkai, et al. Dynamic dispatch of an integrated energy system based on deep reinforcement learning in an uncertain environment [J]. Power System Protection and Control, 2022, 50(18): 50-60.
- [16] YI Z G, LUO Y S, WESTOVER T, et al. Deep reinforcement learning based optimization for a tightly coupled nuclear renewable integrated energy system [J]. Applied Energy, 2022, 328: 120113.
- [17] 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.
- [18] ZHANG B, HU W H, LI J H, et al. Dynamic energy conversion and management strategy for an integrated electricity and natural gas system with renewable energy: deep reinforcement learning approach [J]. Energy Conversion and Management, 2020, 220: 113063.
- [19] SONG D R, MENG W Q, DONG M, et al. A critical survey of integrated energy system: summaries, methodologies and analysis[J]. Energy Conversion and Management, 2022, 266: 115863.
- [20] KANDIDAYENI M, MACIAS A, KHALATBARISOLTANI A, et al. Benchmark of proton exchange membrane fuel cell parameters extraction with metaheuristic optimization algorithms [J]. Energy, 2019, 183: 912-925.
- [21] 随权, 马啸, 魏繁荣, 等. 计及燃料电池热-电综合利用的能源网日前调度优化策略[J]. 中国电机工程学报, 2019, 39(6): 1603-1613. SUI Quan, MA Xiao, WEI Fanrong, et al. Day-ahead dispatching optimization strategy for energy network considering fuel cell thermal-electric comprehensive utilization [J]. Proceedings of the CSEE, 2019, 39(6): 1603-1613.
- [22] BEHZADI A, ARABKOOHSAR A, GHOLAMIAN E. Multi-criteria optimization of a biomass-fired proton exchange membrane fuel cell integrated with organic Rankine cycle/thermoelectric generator using different gasification agents [J]. Energy, 2020, 201: 117640.
- [23] 朱兰, 王吉, 唐陇军, 等. 计及电转气精细化模型的综合能源系统鲁棒随机优化调度[J]. 电网技术, 2019, 43(1): 116-126. ZHU Lan, WANG Ji, TANG Longjun, et al. Robust stochastic optimal dispatching of integrated energy systems considering refined power-to-gas model [J]. Power System Technology, 2019, 43(1): 116-126.
- [24] ZHUANG C Q, CHOUDHARY R, MAVROGIANNI A. Uncertainty-based optimal energy retrofit methodology for building heat electrification with enhanced energy flexibility and climate adaptability[J]. Applied Energy, 2023, 341: 121111.
- [25] LUBIS A, JEONG J, GIANNETTI N, et al. Operation performance enhancement of single-double-effect absorption chiller[J]. Applied Energy, 2018, 219: 299-311.
- [26] SHEN W, CHEN X, QIU J, et al. A comprehensive review of variable renewable energy leveled cost of electricity [J]. Renewable and Sustainable Energy Reviews, 2020, 133: 110301.
- [27] 姜雅男, 于永进, 李长云. 基于改进 TOPSIS 模型的绝缘纸机-热老化状态评估方法[J]. 电工技术学报, 2022, 37(6): 1572-1582. JIANG Yanan, YU Yongjin, LI Changyun. Evaluation method of insulation paper deterioration status with mechanical-thermal synergy based on improved TOPSIS model[J]. Transactions of China Electrotechnical Society, 2022, 37(6): 1572-1582.
- [28] 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.
- [29] SUN Q K, WANG X J, LIU Z, et al. Multi-agent energy management optimization for integrated energy systems under the energy and carbon co-trading market [J]. Applied Energy, 2022, 324: 119646.
- [30] 李嘉文, 余涛, 张孝顺, 等. 基于改进深度确定性梯度算法的 AGC 发电功率指令分配方法[J]. 中国电机工程学报, 2021, 41(21): 7198-7212. LI Jiawen, YU Tao, ZHANG Xiaoshun, et al. AGC power generation command allocation method based on improved deep deterministic policy gradient algorithm [J]. Proceedings of the

CSEE, 2021, 41(21): 7198-7212.

[31] LI J W, YU T, ZHANG X S, et al. Efficient experience replay based deep deterministic policy gradient for AGC dispatch in integrated energy system [J]. Applied Energy, 2021, 285: 116386.

源系统预测、优化调度技术。E-mail: zhanglei950227@163.com

吴红斌(1972—),男,博士,教授,主要研究方向:智能配用电、分布式发电技术。E-mail: hfwuhongbin@163.com

何叶(1993—),女,通信作者,博士,主要研究方向:配电网规划。E-mail: heye1151358102@163.com

张磊(1995—),男,博士研究生,主要研究方向:综合能

(编辑 王梦岩)

Optimal Scheduling Method for Integrated Energy Systems with Hydrogen Based on Deep Reinforcement Learning

ZHANG Lei¹, WU Hongbin¹, HE Ye¹, XU Bin², ZHANG Mingxing³, DING Ming¹

(1. Anhui Province Key Laboratory of Renewable Energy Utilization and Energy Saving

(Hefei University of Technology), Hefei 230009, China;

2. Electric Power Research Institute of State Grid Anhui Electric Power Co., Ltd., Hefei 230601, China;

3. Lu'an Power Supply Company of State Grid Anhui Electric Power Co., Ltd., Lu'an 237000, China)

Abstract: In order to achieve carbon reduction targets, the combination of hydrogen energy and integrated energy systems has become one of the most potential development directions. Aiming at the problems such as the insufficient flexibility of scheduling strategy of hydrogen integrated energy system and difficulty in solving multi-objective optimization of complex systems, an optimal scheduling method for hydrogen integrated energy systems based on deep reinforcement learning is proposed. First, the variable operation condition model of coupled equipment is used to construct a wind-solar-hydrogen-cooling-heat-electricity integrated energy system, and expand the joint energy supply space of equipment. Secondly, considering the system operation cost, carbon emissions, system self-supply balance and renewable energy utilization rate, a multi-objective optimization model is built based on the optimal solution distance to stimulate the exploration of the agent. Then, the deep reinforcement learning algorithm is optimized by time segment characterization to enhance the estimation accuracy of the system state change. Finally, a simulation case is designed based on the measured data of the source and load. The results show that the proposed method can effectively improve the scheduling flexibility of the hydrogen integrated energy system, fully tap the carbon emission reduction potential of hydrogen energy, and realize the dual optimization of scheduling economy and environmental protection.

This work is supported by Regional Innovation and Development Joint Fund of National Natural Science Foundation of China (No. U19A20106), Anhui Provincial Natural Science Foundation of China (No. 2108085UD05) and Anhui Provincial Science and Technology Major Special Program of China (No. 202203f07020003).

Key words: integrated energy system; hydrogen energy; optimal scheduling; deep reinforcement learning; multi-objective optimization; renewable energy

