

基于进化深度强化学习的含氢虚拟电厂低碳经济调度

彭春华, 陈 俐, 曾心志, 孙惠娟

(华东交通大学 电气与自动化工程学院, 江西 南昌 330013)

摘要:为提升能源效率、促进可再生能源消纳,推动虚拟电厂实现低碳经济运行,提出一种以氢能为主的多能源耦合运行模式,引入碳捕集设备和阶梯碳交易机制,构建含氢虚拟电厂低碳经济调度模型,以充分挖掘系统的碳减排潜力,优化可再生能源消纳能力;同时,针对模型变量众多、约束条件复杂的特性,提出了一种基于进化软行动器-评判器深度强化学习方法,基于所提模型与算法构建含氢虚拟电厂的深度强化学习优化调度框架,设计了调度动作、状态空间和奖励函数,使智能体与环境形成交互体系,学习最优运行策略。通过设置不同情景和算法对比进行算例仿真,验证了所提方法能够减少含氢虚拟电厂运行成本与碳排放量。

关键词:虚拟电厂;氢能;低碳经济调度;进化策略;深度强化学习;多能源耦合

中图分类号:TM73;TP18

文献标志码:A

DOI:10.16081/j.epae.202601009

0 引言

近年来,虚拟电厂(virtual power plant, VPP)正迅速向聚合资源更广、智能化程度更高、市场化能力更强的新型VPP演进^[1]。氢能作为一种高效的能源耦合、储存和低碳手段,与新型VPP结合可以形成含氢虚拟电厂(hydrogen virtual power plant, HVPP),该系统能够有效应对可再生能源占比提升带来的挑战,为新型VPP下增强电网对波动性新能源的接纳以及消纳能力提供新思路^[2]。

目前已有一些HVPP的相关研究。文献[3]采用地源热泵-碳捕集-电转气联合运行方式,有效改善了系统碳排放,但氢能利用途径较单一;文献[4]考虑两段式电转气(power to gas, P2G)与燃气掺氢机组联合运行,拓宽了氢能产用途径,减少了能量梯级损耗和VPP碳排放;文献[5]以系统日运行成本最低为目标,提出了可再生能源碱性-质子交换膜联合制氢系统多时间尺度优化策略;文献[6]提出了计及电解槽功率波动的光制氢系统能量管理优化方法。上述研究大多集中于可再生能源制氢、储氢等技术,然而,可再生能源制氢的产量与系统实际氢能需求之间存在供需不平衡现象,使得系统难以确保稳定的氢能供应。

现有研究大多集中于单一的氢能利用途径,如电制氢或电转气,未能充分利用氢能可在多能源系统中的耦合优势以降低碳排放。文献[7]提出电转氢

和火电机组掺氢燃烧相结合的技术,提高了风光利用率,降低了碳排放并拓宽了氢能的利用途径;文献[8]构建含碳捕集电厂与氢能多元利用框架,提高了系统能源利用效率并有显著的低碳经济效益;文献[9]考虑绿证交易,有效提高了发电商利润及VPP收益。随着氢能产业的不断发展,针对传统绿证机制难以适配氢能产业的问题,文献[10]提出了绿证-碳交易联动模型,通过外购绿证抵扣碳排放来降低园区综合能源系统碳交易成本。此外,为响应“双碳”政策,一些研究利用碳交易机制降低系统碳排放。文献[11-12]引入阶梯碳交易机制,在有效提升HVPP整体用能效率的同时,大幅减少了系统碳排放。本文综合上述研究成果,同时考虑阶梯碳交易和氢能绿证交易,建立了HVPP的多元利用模式,在降低VPP运行成本的同时进一步挖掘系统降碳潜能。

近年来,深度强化学习凭借其在复杂场景决策下的高维数据处理能力和灵活的策略调整特性^[13-14],已逐步应用于VPP优化调度中。文献[15]针对含分布式风-光-储单元的VPP,提出一种基于深度确定性策略梯度强化学习的交互式优化调度模型;文献[16]使用异步优势行动-评判算法(asynchronous advantage actor-critic, A3C)解决了VPP经济调度问题,该算法使用贪心策略进行探索,而这种探索策略可能会限制算法对状态空间的探索能力,导致陷入局部最优解或者收敛到次优解的情况;文献[17]基于软行动器-评判器(soft actor-critic, SAC)和双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)算法实现了VPP中电动汽车的充放电调度。上述算法虽具备处理连续动作空间的能力,但在实际应用中,智能体在探索环境时策略存在过度保守以及在高维空间探索效率欠佳,这在很大程度上影响了智能体的训练效率。针对上述问

收稿日期:2025-04-16;修回日期:2025-09-09

在线出版日期:2026-02-02

基金项目:国家自然科学基金资助项目(52267007,52567008);
江西省自然科学基金资助项目(20242BAB26070)

Project supported by the National Natural Science Foundation of China(52267007,52567008) and Jiangxi Provincial Natural Science Foundation(20242BAB26070)

题,本文考虑进化策略在全局搜索和跳出局部最优解方面的优势,提高对策略探索的多样性^[18],同时引入自动变异机制,提出一种进化软行动器-评判器(evolutionary soft actor-critic, ESAC)算法,该算法使用进化策略对行动器网络的权重空间进行探索,扩大搜索范围,可在很大程度上避免局部最优解,从而提高算法的收敛速度和稳定性。

综上所述,本文基于以氢能为核心的多能源耦合运行模式,构建了HVPP低碳经济调度模型。该模型在运行层面采用电制氢、P2G、气转电、氢转氨等多种技术,拓宽氢能应用场景,实现氢能在VPP中的多元化应用,提高能源的利用效率;在低碳机制层面引入阶梯碳交易和氢能绿证交易,强化源侧减碳效果。针对模型高维变量多与约束复杂的特点,本文设计了深度强化学习环境、动作空间、状态空间及奖励函数,并采用ESAC算法进行训练,通过离线生成最优调度策略,最终实现HVPP低碳经济运行。

1 多能源耦合运行模式和碳排放模型

HVPP的能源生产侧由电制氢单元模块构成,整合了电能、热能、氢能、天然气和氨气5种异质能源。供给侧以火电机组、混氢燃气轮机(hydrogen-mixed gas turbine, HGT)、混氢燃气锅炉(hydrogen-mixed gas boiler, HGB)、电制氢设备、电转氨装置、甲烷反应槽、电锅炉(electric boiler, EB)以及储能装置等核心装置,实现多能流的高效转换与协同利用。电制氢单元基于新能源发电实现氢能的规模化生产,储能系统则通过氢、热和电储能增强系统灵活性, HVPP运行框架如附录A图A1所示。

1.1 多能源耦合运行模式

本文提出HVPP框架的以氢能为主的多能源耦合运行模式见附录A图A2,核心环节包括电解水制氢单元及氢能多元转化模块。

传统的P2G技术整体能源转化效率较低,为提高氢能利用效率的同时促进氢能的多途径利用,将P2G拆分为电制氢和甲烷化2个环节。电制氢过程由可再生能源供能,产生的氢气通过各途径进行梯级利用:优先直供混氢燃气管网供给燃气设备使用;其次将氢能输送至甲烷化装置合成天然气,实现碳循环利用;少量为电转氨机组提供原料以生产注入燃煤机组的掺氢燃料;剩余部分储存于储氢罐中,以满足系统灵活用氢需求。当系统氢能供应不足时,通过氢能绿证交易购氢维持系统氢能供给,若系统用氢需求低于产氢量,富余氢气则通过氢能绿证交易出售,显著提升能源利用经济性并降低系统运行成本。

1.2 碳排放模型

系统中产生CO₂的设备包括火电机组、HGT和

HGB。为推进低碳化改造,在产能侧引入了碳捕集技术。系统内HGT、HGB以及火电机组所产生的CO₂通过烟气旁路集中收集,其中一部分经由多级分流装置导入碳捕集单元进行处理,剩余部分排放至大气环境。进入碳捕集装置中的CO₂,采用胺液吸收-蒸汽再生工艺,部分CO₂经吸收塔处理后,被甲烷槽用于燃料替代,剩余部分则通过地质封存或矿物化封存实现长期隔离。系统各部分碳排放模型具体表达式如附录A式(A1)所示。

2 HVPP调度模型

2.1 系统运行模型

在本文中,电制氢装置利用系统富余风能制取氢能,以满足系统对于氢能的需求。同时,HGT、HGB利用混氢燃气输出电、热功率,EB消耗电能输出热功率,电转氨装置在制氨过程中通过哈勃合成氨反应并释放能量,生成的氨气供给火电机组。各设备建模如下所示。

1) HGT和碳捕集设备。

$$\begin{cases} P_t^{\text{HGT}} = (P_t^{\text{HGT},g} + P_t^{\text{HGT},h}) \eta_p^{\text{HGT}} \\ H_t^{\text{HGT}} = (P_t^{\text{HGT},g} + P_t^{\text{HGT},h}) \eta_h^{\text{HGT}} \\ P_t^{\text{TO}} = P_t^{\text{HGT}} - P_t^{\text{CE}} - P_t^{\text{F}} \\ P_t^{\text{CE}} = \omega^c M_t^{\text{C}} \\ M_t^{\text{C}} = \partial \delta_t^s (M_t^{\text{TP}} + M_t^{\text{HGT}} + M_t^{\text{HGB}}) \end{cases} \quad (1)$$

式中: P_t^{HGT} 和 H_t^{HGT} 分别为HGT在 t 时刻的供电功率和供热功率; $P_t^{\text{HGT},g}$ 和 $P_t^{\text{HGT},h}$ 分别为HGT在 t 时刻消耗的天然气和氢气功率; η_p^{HGT} 和 η_h^{HGT} 为HGT的电和热效率; P_t^{TO} 为 t 时刻HGT电功率净出力; P_t^{CE} 为 t 时刻碳捕集运行能耗; P_t^{F} 为 t 时刻碳捕集固定能耗; ω^c 为碳捕集单位CO₂碳捕能耗; M_t^{C} 为 t 时刻碳捕集处理CO₂的量; ∂ 为碳捕集效率; δ_t^s 为 t 时刻烟气分流比; M_t^{TP} 、 M_t^{HGT} 和 M_t^{HGB} 分别为 t 时刻火电机组、HGT和HGB的碳排放量。

2) 混氢燃气锅炉。

$$\begin{cases} P_t^{\text{HGB}} = P_t^{\text{HGB},g} + P_t^{\text{HGB},h} \\ H_t^{\text{HGB}} = \eta^{\text{HGB}} P_t^{\text{HGB}} \end{cases} \quad (2)$$

式中: P_t^{HGB} 为HGB在 t 时刻的总出力; $P_t^{\text{HGB},g}$ 和 $P_t^{\text{HGB},h}$ 分别为HGB在 t 时刻消耗的天然气和氢气功率; η^{HGB} 为HGB的热效率; H_t^{HGB} 为 t 时刻HGB的热功率。

3) 电制氢机组。

$$Q_t^{\text{EH}} = \eta^{\text{EH}} P_t^{\text{EH}} \quad (3)$$

式中: Q_t^{EH} 为 t 时刻电制氢所制取的氢气功率; η^{EH} 为电制氢的产氢效率; P_t^{EH} 为 t 时刻电制氢耗电量。

4) 电锅炉。

$$H_t^{\text{EB}} = \eta^{\text{EB}} P_t^{\text{EB}} \quad (4)$$

式中: H_t^{EB} 和 P_t^{EB} 分别为 t 时刻EB的产热功率和耗电

功率; η^{EB} 为EB电热转换效率。

5)电转氨设备。

在传统火电机组中采用掺氨燃烧技术,可通过氨气与煤粉混合燃烧,利用氨燃料的零碳特性降低单位发电碳排放。氨燃烧产物主要为氮气和水,不产生 CO_2 ,碳排放削减率与氨燃料热值替代率等量。

$$\begin{cases} H_t^{NH_3} = \eta_h^{NH_3} \nu M_t^{NH_3} \\ Q_t^{NH_3} = \eta^{NH_3} P_t^{NH_3} \\ M_t^{NH_3} = P_t^{NH_3} / \omega^{NH_3} \end{cases} \quad (5)$$

式中: $H_t^{NH_3}$ 为 t 时刻电转氨机组向系统提供的热功率; $\eta_h^{NH_3}$ 为电转氨机组用于供热的热释放比例; ν 为生成单位 NH_3 所释放的热功率; $M_t^{NH_3}$ 为 t 时刻电转氨机组生产的氨的质量; $Q_t^{NH_3}$ 为 t 时刻产氨功率; η^{NH_3} 为电转氨设备产氨效率; $P_t^{NH_3}$ 为 t 时刻电转氨设备消耗的电功率; ω^{NH_3} 为电转氨设备制氨的单位能耗。

6)甲烷反应槽。

碳捕集装置将捕集的部分 CO_2 作为高品质原料输送至甲烷反应槽,用于合成天然气。

$$M_t^{MH} = \eta^{MH} \delta^{CO_2} Q_t^{MH} \quad (6)$$

式中: M_t^{MH} 为甲烷槽在 t 时刻虚拟固碳量; Q_t^{MH} 为 t 时刻甲烷槽消耗的氢气; δ^{CO_2} 为生成单位功率天然气所需的 CO_2 量; η^{MH} 为甲烷槽产天然气效率。

7)火电机组。

$$\begin{cases} M_t^{TP} = \mu^{TP} m_t^{c,p} \\ m_t^{c,p} = a(P_t^{TP})^2 + bP_t^{TP} + c - M_t^{NH_3} Q_v^{NH_3} / Q_v^c \end{cases} \quad (7)$$

式中: M_t^{TP} 为掺氨火电机组的碳排放量; μ^{TP} 为单位燃煤燃烧的碳排放量; $m_t^{c,p}$ 为 t 时刻燃煤消耗量; a 、 b 、 c 为火电机组的耗量特性系数; P_t^{TP} 为火电机组的发电量; $Q_v^{NH_3}$ 和 Q_v^c 分别为氨和煤的低位热值。

为保证HGT和混氢燃气锅炉的安全运行,当HGT的混氢比为10%~20%^[19]时,燃烧器可安全且稳定燃烧。混氢燃气锅炉使用混氢燃气时,氢气的摩尔质量比例需保持在2%~20%。为精准刻画HGT与HGB在不同工况下的掺氢比变化规律,建立HGT和HGB的掺氢比与甲烷槽产天然气模型,如式(8)所示。

$$\begin{cases} \kappa_t^{HGT} = \frac{P_t^{HGT,h}}{H^h} / \left(\frac{P_t^{HGT,h}}{H^h} + \frac{P_t^{HGT,g}}{H^g} \right) \\ \kappa_t^{HGB} = \frac{P_t^{HGB,h}}{H^h} / \left(\frac{P_t^{HGB,h}}{H^h} + \frac{P_t^{HGB,g}}{H^g} \right) \\ V_t^{MH} = \frac{3.6\eta^{MH} Q_t^{MH}}{H^g} \end{cases} \quad (8)$$

式中: H^h 为氢气热值; H^g 为天然气热值; κ_t^{HGT} 和 κ_t^{HGB} 分别为 t 时刻HGT与HGB的掺氢比; V_t^{MH} 为 t 时刻甲烷槽产生的天然气体积; η^{MH} 为甲烷槽产天然气效率。

蓄电池、储热罐、储氢罐等储能设备模型类似,故对储能设备进行统一建模,参考文献[20]。

2.2 阶梯碳交易和氢能绿证交易模型

2.2.1 阶梯碳交易

碳配额分配方式主要分为有偿分配和无偿分配2种类型,本文中的碳配额采用多数文献采用的无偿分配,并利用基准线法作为碳排放配额的分配方法,其中碳配额基准值参考文献[21],碳配额分配对象有火电机组、HGT和HGB,具体表达式如附录A式(A2)所示。

为更有效控制碳排放,本文基于统一型碳交易框架,引入阶梯式碳交易机制,通过划分碳排量区间确定交易成本。同时,为进一步激励减排行为,引入补偿系数加大减排奖励力度,以此增强碳市场的约束效力。阶梯碳交易成本的计算模型见文献[22]。

2.2.2 氢能绿证交易

当前绿氢制备成本较高,但其环保价值显著。为此,本文将电力市场中的绿证交易机制引入氢能市场,并根据绿证供需关系制定交易价格,以促进绿氢的经济发展和环境效益的实现。氢能绿证交易的具体模型如下:

$$\begin{cases} C^{hgc} = C^{hgc,max} - o \sum_{i=1}^T Q_t^{EH} \\ 0 \leq C^{hgc} \leq C^{hgc,max} \end{cases} \quad (9)$$

式中: C^{hgc} 为氢能绿证的出清价格; $C^{hgc,max}$ 为氢能绿证的最高可接受价格; o 为氢能绿证的需求比; T 为总时段数。

2.3 目标函数和约束条件

2.3.1 目标函数

针对考虑碳交易和氢能绿证交易的HVPP,为实现经济型的同时兼顾环境效益,构建低碳经济目标函数,如式(10)所示。

$$\min F^{HVPP} = \min(C^{OP} + C^{CO_2} + C^F + C^{CD} + C^W + C^B + C^{TP} - C^H) \quad (10)$$

式中: F^{HVPP} 为HVPP总运行成本; C^{OP} 为机组运行成本; C^{CO_2} 为阶梯碳交易成本; C^F 为碳封存成本; C^{CD} 为碳捕集设备成本; C^W 为弃风惩罚成本; C^B 为购气成本; C^{TP} 为火电机组的启停和煤耗成本; C^H 为氢能绿证交易收益。每个部分具体表达式如下。

1)机组运行成本 C^{OP} 。

$$C^{OP} = \sum_{t=1}^T \sum_{j=1}^N c_{j,t}^{op} P_{j,t} \quad (11)$$

式中: $c_{j,t}^{op}$ 为 t 时刻第 j 种设备的运行维护成本; $P_{j,t}$ 为 t 时刻除火电机组外第 j 种设备的出力; N 为设备种类。

2)阶梯碳交易成本 C^{CO_2} 。

$$C^{CO_2} = \sum_{t=1}^T f_t^{CO_2} \quad (12)$$

式中: $f_t^{\text{CO}_2}$ 为 t 时刻HVPP的碳交易成本。

3)碳封存成本 C^{F} 。

$$C^{\text{F}} = \sum_{t=1}^T c^{\text{seq}} M_t^{\text{Cse}} \quad (13)$$

式中: c^{seq} 为碳封存单位成本; M_t^{Cse} 为 t 时刻碳封存量。

4)碳捕集设备成本 C^{CD} 。

$$C^{\text{CD}} = \sum_{t=1}^T C_t^{\text{cap}} = C^{\text{FL}} \frac{(1+r)^{N^{\text{Z}}} r}{365 \times 24 [(1+r)^{N^{\text{Z}}} - 1]} \quad (14)$$

式中: C_t^{cap} 为 t 时刻碳捕集设备折旧成本; C^{FL} 和 N^{Z} 分别为碳捕集设备总成本和折旧年限; r 为碳捕集电厂项目贴现率。

5)弃风惩罚成本 C^{W} 。

$$C^{\text{W}} = \sum_{t=1}^T \alpha^{\text{waste}} P_t^{\text{waste}} \quad (15)$$

式中: α^{waste} 为弃风惩罚系数; P_t^{waste} 为 t 时刻经消纳后的弃风量。

6)购气成本 C^{B} 。

$$C^{\text{B}} = \sum_{t=1}^T c_t^{\text{buy, g}} V_t^{\text{buy}} \quad (16)$$

式中: $c_t^{\text{buy, g}}$ 为 t 时刻天然气的购买单价; V_t^{buy} 为 t 时刻购气体积。

7)火电机组启停和煤耗成本 C^{TP} 。

$$\begin{cases} C^{\text{TP}} = C^{\text{TP1}} + C^{\text{TP2}} \\ C^{\text{TP1}} = G \sum_{t=2}^T [\beta_t^{\text{TP}} (1 - \beta_{t-1}^{\text{TP}}) + \beta_{t-1}^{\text{TP}} (1 - \beta_t^{\text{TP}})] \\ C^{\text{TP2}} = \sum_{i=1}^T m_i^{c, \text{tp}} \end{cases} \quad (17)$$

式中: C^{TP1} 为火电机组启停成本; C^{TP2} 为火电机组煤耗成本; G 为火电机组的启停成本系数; β_t^{TP} 为 t 时刻火电机组的状态,为0-1变量; β_{t-1}^{TP} 为上一时刻火电机组的状态。

8)氢能绿证收益 C^{H} 。

$$C^{\text{H}} = C^{\text{hgs}} \sum_{i=1}^T Q_i^{\text{SH}} + 0.5o \left(\sum_{i=1}^T Q_i^{\text{SH}} \right)^2 \quad (18)$$

式中: Q_i^{SH} 为 t 时刻通过氢能绿证交易出售的氢气功率。

2.3.2 约束条件

本文的HVPP调度模型约束条件包含机组出力约束、参数约束、爬坡约束、平衡约束、安全运行约束参考文献[23],具体表达式如附录A式(A3)~(A6)所示。

3 进化深度强化学习模型

3.1 奖励函数设计

本文构建的模型将降低VPP运行成本作为优化目标,智能体依据环境返回的奖励函数值来调整其行动策略,以追求奖励值的最大化。目标函数如式

(19)所示。

$$r(s_t, a_t) = -\vartheta F^{\text{HVPP}} + \varepsilon \quad (19)$$

式中: ϑ 为HVPP运行成本得比重系数; ε 为调节参数; s_t 为 t 时刻的状态; a_t 为 t 时刻的动作; $r(s_t, a_t)$ 为 t 时刻的奖励函数。

3.2 动作空间设计

本文通过功率平衡约束方程式(A6)来确定气站购气量、HGB供热功率及EB功率。由风电功率 P_t^{W} 、火电功率 P_t^{TP} 、HGT电功率 P_t^{HGT} 、甲烷反应槽产氢功率 Q_t^{MH} 、电制氢装置耗电功率 P_t^{EH} 、碳捕集设备烟气分离比率 δ_t^{S} 、燃气混氢比例 κ_t 、火电掺氨比例 γ_t 等8个动作变量构成动作空间 A_t ,如式(20)所示。

$$A_t = [P_t^{\text{W}}, P_t^{\text{TP}}, P_t^{\text{HGT}}, Q_t^{\text{MH}}, P_t^{\text{EH}}, \delta_t^{\text{S}}, \kappa_t, \gamma_t] \quad (20)$$

为保证所设计环境的动作能够适应HVPP模型的实际运行要求,动作必须满足模型中各类运行约束条件。

3.3 状态空间设计

在本文探讨的HVPP中,环境传递给深度强化学习智能体的信息一般涵盖电负荷、热负荷、风电出力以及天然气实时价格。此外,上一时刻 $t-1$ 的动作值也为状态量,具体如式(21)所示。

$$S_t = [H_t^{\text{e}}, H_t^{\text{h}}, f_t^{\text{w}}, \kappa_t^{\text{g}}, A_t] \quad (21)$$

式中: S_t 为状态空间; H_t^{e} 为 t 时刻电负荷需求; H_t^{h} 为 t 时刻热负荷需求; f_t^{w} 为风电预测值; κ_t^{g} 为 t 时刻天然气购买单价价格。

3.4 ESAC算法

进化策略(evolutionary strategies, ES)以高效并行搜索和连续参数优化著称,而基于最大熵原理的SAC算法在高维连续动作空间中平衡探索能力表现出色。然而,SAC因双 Q 网络估计 Q 值的偏差,可能导致策略过于集中,智能体易陷入局部最优,探索较为保守,难以拓展到更广泛的动作空间。为解决上述问题,本文采用ESAC算法,结合ES的种群优化思想,增强SAC的全局搜索能力。本文基于ESAC深度强化学习的HVPP优化调度过程如附录B图B1所示。为提高探索效率,引入了自动突变调整机制,通过动态调整变异率,使算法能够根据种群当前表现自适应调整扰动幅度,从而在复杂环境中实现鲁棒性更高和效率更高的学习。

3.4.1 算法目标

ESAC算法目标是寻找一个最优目标策略 π ,使得累计回报之和与策略的熵正则项的总期望值 $J(\pi)$ 最大化,在此期望值为智能体与环境交互的所有可能轨迹的统计平均值,如式(22)所示。

$$\begin{cases} J(\pi) = E_{\tau \sim \pi} \left[\sum_{t=0}^T \gamma^t \left(r(s_t, a_t) + \alpha \bar{H}(\pi(\cdot | s_t)) \right) \right] \\ \bar{H}(\pi(\cdot | s_t)) = -E_{a_t \sim \pi} [\log \pi(a_t | s_t)] \end{cases} \quad (22)$$

式中： $\tau=(s_0, a_0, s_1, a_1, \dots)$ 为由策略 π 生成的环境交互轨迹； $\gamma \in (0, 1)$ 为折扣因子； E 为期望函数； $\bar{H}(\pi(\cdot|s_t))$ 为策略 π 在状态 s_t 下的熵； α 为温度参数($\alpha > 0$)。

3.4.2 核心参数更新方法

ESAC算法网络结构主要包括1个行动器和1个评判器,其中评判器内含2个用于估计Q值的Q函数神经网络以及1个目标Q函数神经网络,用于增强算法的训练稳定性^[24]。详细参数更新方法见附录C。

3.4.3 进化策略

进化策略是一种启发式搜索过程,主要过程为:后代群体使用随机扰动进行变异,变异后评估种群中每个成员对应的适应度指标,并将得分较高的后代重新组合,形成下一代种群。本文根据其特性结合深度强化学习进行改进。

设进化策略的基准策略参数为 θ_{es} , n 为种群中个体数量,则每个进化策略参数个体可表示为 $\theta_{es,(n)}$,随机扰动表示为 $\epsilon_{(n)}$ 。随机扰动从高斯分布 $\mathcal{N}(0, 1)$ 中采样,以便变异 $\theta_{es,(n)}$ 并评估适应度目标,单次噪声扰动后的适应度值具有高度随机性,故取奖励值期望值作为最终的适应度 $F_{(n)}$,如式(23)所示。

$$F_{(n)} = E_{\epsilon_{(n)} \sim \mathcal{N}(0, 1)} [O(\theta_{es,(n)} + \sigma \epsilon_{(n)})] = \frac{1}{\sigma} E_{\epsilon_{(n)} \sim \mathcal{N}(0, 1)} [O(\theta_{es,(n)} + \sigma \epsilon_{(n)}) \epsilon_{(n)}] \quad (23)$$

式中： σ 为变异率； $O(\theta_{es,(n)} + \sigma \epsilon_{(n)})$ 为奖励值。

计算完所有个体的适应度后,进行软胜利者选择,即将所有个体适应度进行归一化、排序。选择前 w 个胜利者个体,构成胜利者个体集合, w 为:

$$w = ne \quad (24)$$

式中： e 为优势个体比例。

ES基准策略参数 θ_{es} 在执行完软胜利者选择后进行更新,更新公式如下:

$$\theta_{es}^{new} \leftarrow \theta_{es}^{old} + \frac{\alpha_{es}}{n\sigma} \sum_{i=1}^n F_{(i)} \epsilon_{(i)} \quad (25)$$

式中： θ_{es}^{old} 和 θ_{es}^{new} 分别为当前代和下一代种群的基准策略参数； α_{es} 为ES策略的学习率； i 为个体序号。

3.4.4 ESAC算法更新流程

HVPP深度强化学习调度训练流程图如附录D图D1所示。

训练步骤如下。

1)初始化参数与种群。

参数初始化:状态值函数 w_i 、目标值函数 \tilde{w}_i 、SAC策略参数 θ 、ES基准策略参数 θ_{es} 、Q函数 ϕ 、经验池 D 。

超参数设置:SAC智能体学习率 α_s 、ES种群学习率 α_{es} 、变异率 σ 、SAC更新概率 p_{sac} 、调整因子 ζ 、软

更新系数 τ 、优势个体比例 e 和固定间隔 g 。

种群初始化: $n_{pop} = \{\theta_{es,(1)}, \theta_{es,(2)}, \dots, \theta_{es,(n)}\}$,每个个体通过 $\theta_{es,(n)} = \theta_{es} + \sigma \epsilon_{(n)}$ 生成。

2)评估个体适应度。

由每个个体在HVPP环境中执行策略,基于当前状态 s_t 输出动作 a_t 。环境根据动作和随机源荷场景转移到新状态 s_{t+1} ,并计算奖励值 $r(s_t, a_t)$ 存入经验池 D 中,同时根据式(23)计算适应度 $F_{(n)}$,形成集合 $F = \{F_{(1)}, F_{(2)}, \dots, F_{(n)}\}$ 。

3)软胜利者选择、交叉和变异率更新。

进行软胜利者选择,选择适应度前 w 个胜利者个体,胜利者按照 F 排序选取,组成集合: $W = \{\theta_{es,(1)}, \theta_{es,(2)}, \dots, \theta_{es,(w)}\}$,在胜利者和子代之间进行交叉操作保留优势特征。通过式(C5)更新变异率,通过式(25)更新ES基准策略参数。

4)SAC梯度更新与经验回放。

在固定间隔 g ,从经验池 D 中采样数据,通过式(C1) — (C3)对评判器、行动器和温度参数进行更新。

5)种群更新与迭代。

新种群由交叉后胜者个体 $\theta_{es,(w)}$ 、SAC的策略参数 θ 和 θ_{es} 共同构成。

6)若算法的训练已达到最大回合数,则ESAC算法的训练过程结束,若未达到,则回到步骤2),开启新一轮的训练。

4 算例分析

4.1 算例描述

为验证本文如图A1所建立的模型,文中采用的HVPP系统包含1座风电厂、1座碳捕集电厂、1台电转氨机组、1台电制氢机组、1台燃气轮机、1台燃气锅炉、1台火电机组、1台电锅炉以及储氢、热、电设备,各设备参数如附录D表D1所示,电热负荷以及风电预测曲线如附录D图D2所示。

4.2 训练收敛性分析

本文算例仿真的硬件环境为:Inter i7-11800处理器。搭配16GB内存以满足计算需求,基于PyTorch平台进行编程,编程语言为Python 3.8。在仿真环境中,包括9种动作变量和13种状态变量,算法中各类超参数见附录D表D2。

ESAC算法在训练过程中奖励值变化曲线如图1所示。在训练的初始阶段,智能体聚焦于收集环境的交互信息,并将数据积累至经验池中,经验池饱和后,开始探索未知环境,智能体逐渐能从经验池中找出最大化奖励值的动作。此后,奖励值波动幅度显著减少,表明智能体的决策方向趋于稳定。在训练次数达到1700后,奖励值曲线基本趋于稳定,仅呈现较小波动,表明训练过程已接近收敛。

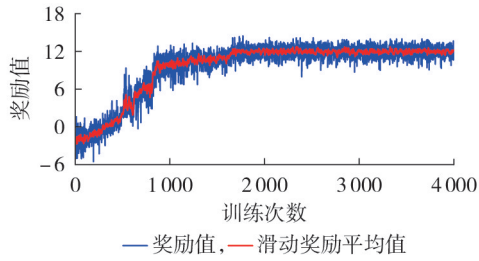


图1 奖励变化曲线

Fig.1 Curves of reward variation

4.3 调度结果分析

训练 HVPP 模型结束后,对图 A2 所示的电热负荷、风电出力场景进行测试,得到电负荷和热负荷优化调度结果。其中,电负荷优化调度结果如图 2 所示。

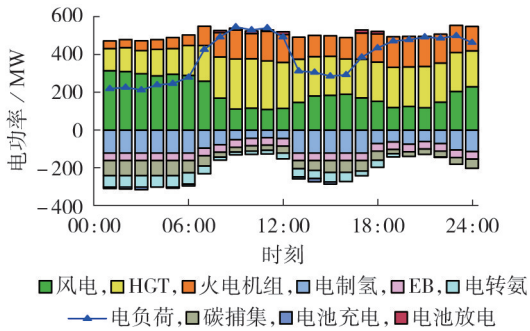


图2 电负荷优化调度结果

Fig.2 Optimized scheduling results of electrical load

在 01:00—06:00 时段,低谷电负荷下风电出力高,HGT 和火电机组以低出力平抑波动,充分利用风电驱动电制氢、碳捕集和电转氨,火电机组掺氨燃烧以降碳,储能充电储备电力。07:00—12:00 时段,电负荷高峰而风电下降,HGT 和火电机组大幅增加出力,降低电制氢等设备功率,维持碳捕集系统供电,储能放电缓解压力。13:00—18:00 时段,电负荷回落、风电恢复出力,HGT 出力减少,电制氢和碳捕集功率回升,储能先充电后放电平抑波动。19:00—24:00 时段,负荷较平稳、风电回升,HGT 和火电机组微调出力,电制氢随风电增加,电转氨后期停运。碳捕集系统烟气分流比率如附录 D 图 D3 所示。

由图 3 热负荷优化调度结果可知,01:00—06:00 时段,热负荷较高但逐渐下降,HGT 维持高出力,蓄热罐初期放热后转为储热,电转氨放热显著,利用副

产物热量辅助供热。07:00—12:00 时段,热负荷降至低点,HGT 略增出力,HGB 减至零以节约成本,蓄热罐持续储热,电转氨放热减弱。13:00—18:00 时段,HGT 调整出力,HGB 后期启用以应对热负荷增加,蓄热罐先微量储热后放热,电转氨放热随氨生产恢复增加。19:00—24:00 时段,热负荷先升后降,HGT 平稳,HGB 恢复出力,蓄热罐先放热后储热,电转氨放热后期降至 0。

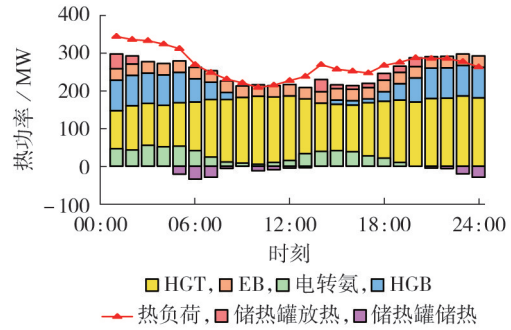


图3 热负荷优化调度结果

Fig.3 Optimized scheduling results of heat load

4.4 模型与算法性能验证

4.4.1 模型情景对比

为验证本文 HVPP 模型的优越性,设置以下情景开展验证。

情景 1:不考虑甲烷化,制取氢气仅用于燃气掺氢设备。

情景 2:在情景 1 的基础上,考虑甲烷化,甲烷化所需的 CO₂ 由碳捕集设备供给。

情景 3:在情景 2 的基础上,考虑电转氨设备以及储氢设备。

情景 4:在情景 3 的基础上,进一步引入氢能绿证交易机制。

4 种情景的优化调度结果如表 1 所示。由于不同方案的机组运行成本和碳捕集设备折旧成本波动较小,故不单独列出,但已计入总成本核算。由表 1 可知,情景 2 相较于情景 1 在多项指标上显著优化,总成本下降约 22.89%,碳排放量下降约 27.97%,购气成本下降约 2.83%,这主要由于情景 2 引入甲烷化技术,通过捕集系统产生的二氧化碳,与氢能反应生成甲烷,供给混氢燃气设备,从而减少购气和碳排放。

表 1 优化调度结果

Table 1 Optimal scheduling results

情景	总成本 / 万元	煤耗成本 / 万元	购气成本 / 万元	弃风惩罚成本 / 万元	碳交易成本 / 万元	碳排放 / t	氢能绿证收益 / 万元	弃风率 / %
1	689.12	81.74	301.47	9.24	135.32	3934.94	0	12.87
2	531.36	81.74	292.94	2.39	118.15	2834.3	0	10.06
3	487.86	72.11	288.05	0	116.06	2793.86	0	0
3	462.15	36.42	262.18	0	65.75	1582.49	9.58	0

情景3相较于情景2进一步优化,总成本下降约8.28%,碳排放量下降约1.43%,煤炭消耗成本下降约11.78%。这主要由于电解水制氢单元依托新能源发电规模化生产,通过氢储能优化调度完全消纳风电。煤炭消耗减少反映火电机组依赖降低,混氢燃气设备通过更高氢气掺混比例提升调峰能力,部分替代火电负荷,推动碳排放小幅下降。

情景4在情景3的基础上引入氢能绿证交易,整体经济性和低碳效果显著提升,总成本下降约5.17%,碳排放量下降约43.35%,天然气采购成本下降约8.98%,煤耗成本下降约49.48%。氢能绿证交易通过市场化机制将富余氢能转化为经济收益,同时提高电制氢单元规模化生产,减少对天然气和煤炭的依赖。氢能的多元化利用大幅减少化石燃料消耗,降低碳排放。

因此,借助以氢能为核心的多能源耦合模式的HVPP,可以有效提高风能消纳水平,减少系统运行成本并降低碳排放量。

4.4.2 算法优化性能对比

为评估本文所提ESAC算法在HVPP低碳经济调度中的优化性能,选择SAC和TD3作为对比算法。选择依据如下:ESAC基于SAC框架,通过引入进化策略和自动变异机制增强全局搜索和探索多样性,解决SAC在高维复杂环境中易陷入局部最优的问题,故以SAC为基准验证改进效果。TD3作为主流深度强化学习算法,适用于连续动作空间,通过双Q网络和延迟更新提高稳定性,但探索策略保守,收敛较慢,ESAC通过进化策略弥补此不足,适合对比以评估收敛速度和奖励稳定性。各算法奖励曲线变化趋势如图4所示。

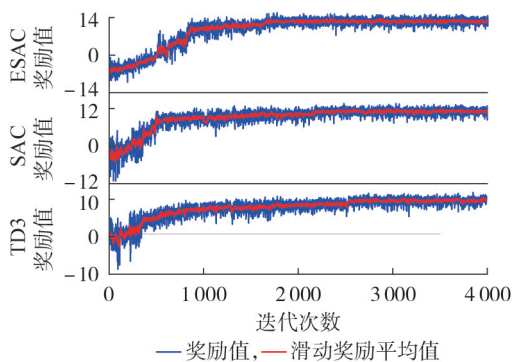


图4 深度强化学习算法奖励变化曲线

Fig.4 Reward variation curves for deep reinforcement learning algorithms

对比ESAC、SAC和TD3这3种算法的性能,结果表明ESAC算法在奖励函数值上表现最佳,且收敛速度最快。其奖励曲线在大约1800次训练周期后开始平稳,奖励值稳定在12左右。相比之下,SAC算法需要约2200次训练周期才能达到稳定状态,

奖励值约为10。而TD3算法表现较为逊色,需要约2500次训练周期后趋于平稳,奖励值约为8。

鉴于粒子群优化算法(particle swarm optimization, PSO)作为传统智能优化算法代表也常用于VPP调度模型的优化求解,故在此也将PSO优化结果进行了对比。各算法参数均根据10次以上的运行效果进行了最优配置,具体见附录E表E1。不同优化方法的对比结果如表2所示。

表2 不同优化方法对比

Table 2 Comparison of different optimization methods

优化方法	成本 / 元	优化时间 / s
PSO	4974652	297
TD3	4893381	1
SAC	4770642	1
ESAC	4621538	1

从表2中数据分析可知,PSO算法对高维问题适应能力较弱,其最终优化成本较TD3高出1.63%,且优化耗时较长。TD3算法由于采用延迟更新机制,训练过程稳定性欠佳,导致其成本较SAC增加了2.51%。与之相比,ESAC算法通过结合进化策略,解决了SAC算法效率较低的问题,最终调度成本较SAC下降了3.12%。上述结果充分证明了本文提出的ESAC算法在HVPP低碳经济优化调度中的卓越性能。

4.4.3 算法泛化能力验证

为验证ESAC算法在不同风电场景及负荷波动强度下的泛化能力,本文设计了2种扰动实验场景,分别针对风电出力波动和负荷变化强度进行测试,并在每种场景下对比ESAC、SAC和TD3算法的优化性能。实验基于图D2的基准场景,具体扰动设置如下。

场景A,风电出力波动:在基准场景上,分别引入 $\pm 10\%$ 、 $\pm 20\%$ 和 $\pm 30\%$ 的随机扰动,模拟风电出力的不确定性和间歇性。扰动通过高斯分布 $\mathcal{N}(0, \chi^2)$ 生成,其中标准差 χ 分别对应预测值的10%、20%和30%。

场景B,负荷波动强度:在基准曲线上,分别施加 $\pm 15\%$ 和 $\pm 30\%$ 的随机扰动,模拟负荷高峰和低谷的波动。扰动同样采用高斯分布生成,标准差分别为负荷预测值的15%和30%。

在场景A和B下,分别采用ESAC、SAC和TD3算法进行仿真实验,记录总运行成本、碳排放量及弃风率,算法的鲁棒性和泛化能力。实验结果如附录E表E2所示。从结果中可以看出,在所有扰动场景下,ESAC算法的运行成本、碳排放量和弃风率均低于SAC和TD3算法,表明其在面对风电出力波动和负荷变化时具有更强的鲁棒性和泛化能力,特别是在高扰动强度下,ESAC算法的成本和碳排放增幅较

小,弃风率保持在较低水平,反映了其通过进化策略和自动变异机制有效增强了全局搜索能力和对复杂环境的适应性。

5 结论

针对新型VPP中大量出力不确定性可再生能源接入导致的弃能问题日益加剧以及传统火电机组碳排放较高的现状,本文提出一种HVPP低碳经济调度方法,建立以氢能为核心的多能源耦合机制,同时采用ESAC深度强化学习算法实现其低碳经济优化调度,并得到如下结论。

1)提出的HVPP经济调度模型以电解制氢技术为核心,通过将VPP中富余的风能转化为氢能并实现多元化应用,从而优化系统性能,显著提升了系统的经济效益和低碳特性。该模型在提升可再生能源消纳能力、优化运行成本及实现低碳目标方面展现出突出成效。

2)对传统深度强化学习算法进行优化改进,通过引入进化策略和自动变异机制,提出了ESAC算法,该算法在处理HVPP复杂调度问题时展现出更优异的性能。

未来的研究工作可进一步在现有模型中引入潮流计算约束,结合电力系统的网络拓扑结构和线路损耗特性,进一步优化调度策略。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 周保荣,高洪超,程韧俐,等. 第一性原理视角下虚拟电厂的技术体系构建[J]. 电力系统自动化,2025,49(20):3-15.
ZHOU Baorong,GAO Hongchao,CHENG Renli,et al. Technical system construction for virtual power plants from first principles perspective [J]. Automation of Electric Power Systems, 2025,49(20):3-15.
- [2] 向昱瑾,赵年年,李晓靖,等. 考虑源荷协调的电氢耦合微电网容量优化配置[J]. 智慧电力,2025,53(3):9-18.
XIANG Yujin,ZHAO Niannian,LI Xiaojing,et al. Capacity optimization configuration of electricity-hydrogen coupled micro-grid considering source-load coordination[J]. Smart Power, 2025,53(3):9-18.
- [3] 王秋杰,亓浩,谭洪,等. 考虑碳市场风险的热电联产虚拟电厂低碳调度[J]. 电力自动化设备,2024,44(10):8-15.
WANG Qiujie,QI Hao,TAN Hong,et al. Low-carbon scheduling of CHP virtual power plant considering carbon market risk[J]. Electric Power Automation Equipment,2024,44(10):8-15.
- [4] 武群丽,白佳怡. 考虑阶梯碳交易及两阶段电转气的综合能源系统优化调度[J/OL]. 华北电力大学学报(自然科学版).(2024-04-19)[2025-04-14]. <http://kns.cnki.net/kcms/detail/13.1212.TM.20240419.1446.002.html>.
- [5] 杨胜,樊艳芳,侯俊杰,等. 可再生能源 ALK-PEM 联合制氢系统多时间尺度优化运行策略[J]. 电力系统保护与控制,2025,53(3):68-80.
YANG Sheng,FAN Yanfang,HOU Junjie,et al. Multi-time scale optimization strategy of a renewable energy ALK-PEM combined hydrogen production system[J]. Power System Protection and Control,2025,53(3):68-80.
- [6] 陈磊磊,年珩,赵建勇,等. 计及电解槽功率波动的风光制氢系统能量管理优化[J]. 电力自动化设备,2025,45(9):20-27.
CHEN Leilei,NIAN Heng,ZHAO Jianyong,et al. Energy management optimization of wind-solar hydrogen production system considering power fluctuation of electrolyzers[J]. Electric Power Automation Equipment,2025,45(9):20-27.
- [7] 袁文腾,陈亮,王春波,等. 基于氨储能技术的电转氨耦合风光-火综合能源系统双层优化调度[J]. 中国电机工程学报,2023,43(18):6992-7003.
YUAN Wenteng,WANG Chunbo,et al. Bi-level optimal scheduling of power-to-ammonia coupling wind-photovoltaic-thermal integrated energy system based on ammonia energy storage technology[J]. Proceedings of the CSEE,2023,43(18):6992-7003.
- [8] 刘妍,胡志坚,陈锦鹏,等. 含碳捕集电厂与氢能多元利用的综合能源系统低碳经济调度[J]. 电力系统自动化,2024,48(1):31-40.
LIU Yan,HU Zhijian,CHEN Jinpeng,et al. Low-carbon economic dispatch of integrated energy system considering carbon capture power plant and multi-utilization of hydrogen energy [J]. Automation of Electric Power Systems,2024,48(1):31-40.
- [9] 柴超,刘松阳,孔维康,等. 考虑绿证交易的虚拟电厂运行策略优化研究[J]. 电力需求侧管理,2024,26(5):100-105.
CHAI Chao,LIU Songyang,KONG Weikang,et al. Operation strategy optimization of virtual power plant considering green card transaction [J]. Power Demand Side Management,2024,26(5):100-105.
- [10] 巴金宇,何川,南璐,等. 基于绿证-阶梯碳交易联动与分布鲁棒机会约束的含氢园区综合能源系统低碳经济调度[J]. 电力自动化设备,2025,45(10):169-176.
BA Jinyu,HE Chuan,NAN Lu,et al. Low-carbon economic scheduling of park-integrated energy system containing hydrogen based on green certificate-tiered carbon trading and distributionally robust chance constraint[J]. Electric Power Automation Equipment,2025,45(10):169-176.
- [11] 徐韵,徐耀杰,杨嘉禹,等. 基于阶梯碳交易的电转气虚拟电厂低碳经济调度[J]. 电力系统及其自动化学报,2023,35(7):118-128.
XU Yun,XU Yaojie,YANG Jiayu,et al. Low carbon economic dispatching for power-to-gas virtual power plant based on stepped carbon trading[J]. Proceedings of the CSU-EPSA,2023,35(7):118-128.
- [12] 赵振宇,李焯薪. 基于阶梯碳交易的碳捕集电厂-电转气虚拟电厂低碳经济调度[J]. 发电技术,2023,44(6):769-780.
ZHAO Zhenyu,LI Xinxin. Low-carbon economic dispatch based on ladder carbon trading virtual power plant considering carbon capture power plant and power-to-gas[J]. Power Generation Technology,2023,44(6):769-780.
- [13] 王仁浚,高红均,罗龙波,等. 基于深度强化学习的新型配电系统优化运行研究综述[J]. 电力自动化设备,2025,45(9):152-164.
WANG Renjun,GAO Hongjun,LUO Longbo,et al. Review of research on new distribution system optimization operation based on deep reinforcement learning[J]. Electric Power Automation Equipment,2025,45(9):152-164.
- [14] 张宁,杨凌霄,李炫浓,等. 基于分层深度强化学习的多能虚拟电厂区域消纳优化策略[J]. 电力系统保护与控制,2025,53(20):153-163.
ZHANG Ning,YANG Lingxiao,LI Xuannong,et al. Regional consumption optimization strategy for multi-energy virtual power

- plants based on hierarchical deep reinforcement learning[J]. Power System Protection and Control, 2025, 53(20):153-163.
- [15] 李明扬, 张智. 基于强化学习的含分布式风-光-储虚拟电厂优化调度[J]. 智慧电力, 2024, 52(8):50-56.
LI Mingyang, ZHANG Zhi. Optimal dispatch of distributed wind-solar-storage virtual power plants based on reinforcement learning[J]. Smart Power, 2024, 52(8):50-56.
- [16] LIN L, GUAN X, PENG Y, et al. Deep reinforcement learning for economic dispatch of virtual power plant in Internet of energy[J]. IEEE Internet of Things Journal, 2020, 7(7):6288-6301.
- [17] WANG J N, GUO C L, YU C S, et al. Virtual power plant containing electric vehicles scheduling strategies based on deep reinforcement learning[J]. Electric Power Systems Research, 2022, 205:107714.
- [18] KHADKA S, MAJUMDAR S, NASSAR T, et al. Collaborative evolutionary reinforcement learning[EB/OL]. (2019-05-06) [2025-12-12]. <https://doi.org/10.48550/arXiv.1905.00976>.
- [19] 崔耀欣, 刘晓佩, 陈明敏. F级重型燃气轮机燃烧器天然气掺氢全压试验研究[J]. 燃气轮机技术, 2021, 34(2):38-42.
CUI Yaoxin, LIU Xiaopei, CHEN Mingmin. Experimental study of natural gas mixed with hydrogen under full pressure of F-class heavy duty gas turbine burner[J]. Gas Turbine Technology, 2021, 34(2):38-42.
- [20] 任宇路, 程昱舒, 张雪瑞, 等. 碳交易下含氢储能的综合能源系统运行优化[J]. 电测与仪表, 2025, 62(12):11-19.
REN Yulu, CHENG Yushu, ZHANG Xuerui, et al. Operation optimization of integrated energy system with hydrogen storage under carbon trading[J]. Electrical Measurement & Instrumentation, 2025, 62(12):11-19.
- [21] 中华人民共和国生态环境部. 2023、2024年度全国碳排放权交易发电行业配额总量和分配方案[EB/OL]. (2024-10-25)[2025-09-06]. <https://www.gov.cn/zhengce/zhengceku/202410/P020241022275017526919.pdf>.
- [22] 陈登勇, 刘方, 刘帅. 基于阶梯碳交易的含P2G-CCS耦合和燃气掺氢的虚拟电厂优化调度[J]. 电网技术, 2022, 46(6):2042-2054.
CHEN Dengyong, LIU Fang, LIU Shuai. Optimization of virtual power plant scheduling coupling with P2G-CCS and doped with gas hydrogen based on stepped carbon trading[J]. Power System Technology, 2022, 46(6):2042-2054.
- [23] 张晓辉, 刘小琰, 钟嘉庆. 考虑奖惩阶梯型碳交易和电-热转移负荷不确定性的综合能源系统规划[J]. 中国电机工程学报, 2020, 40(19):6132-6142.
ZHANG Xiaohui, LIU Xiaoyan, ZHONG Jiaqing. Integrated energy system planning considering a reward and punishment ladder-type carbon trading and electric-thermal transfer load uncertainty[J]. Proceedings of the CSEE, 2020, 40(19):6132-6142.
- [24] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[EB/OL]. (2018-08-08)[2025-07-06]. <https://doi.org/10.48550/arXiv.1801.01290>.

作者简介:

彭春华(1973—),男,教授,博士研究生导师,博士,主要研究方向为电力系统优化运行、微电网优化运行与控制(E-mail:chinapch@163.com);

陈俐(1992—),女,博士研究生,主要研究方向为新能源电力系统建模与电力市场预测。

(编辑 丁欣欣)

Low-carbon economic scheduling of hydrogen virtual power plant based on evolutionary deep reinforcement learning

PENG Chunhua, CHEN Li, ZENG Xinzhi, SUN Huijuan

(School of Electrical & Automation Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: In order to improve energy efficiency, promote renewable energy consumption, and push the virtual power plant to realize low-carbon economic operation, a multi-energy coupling operation mode dominated by hydrogen energy is proposed, the carbon capture equipment and a tiered carbon trading mechanism is introduced, and a low-carbon economic scheduling model for hydrogen virtual power plants is constructed to fully exploit the system's carbon emission reduction potential and optimize the capacity for renewable energy consumption. Meanwhile, aiming at the characteristics of the model as numerous variables and complex constraints, a method based on evolutionary soft actor-critic deep reinforcement learning is proposed. Based on the proposed model and algorithm, a deep reinforcement learning optimization scheduling framework for hydrogen virtual power plants is constructed. The dispatch actions, state space, and reward function are designed to enable the agent to form an interactive system with the environment and learn the optimal operation strategy. Case simulations are conducted by setting different scenarios and conducting algorithm comparisons to verify that the proposed method can reduce the operating costs and carbon emissions of hydrogen virtual power plants.

Key words: virtual power plant; hydrogen energy low-carbon economic scheduling; evolutionary strategy; deep reinforcement learning; multi-energy coupling