

DOI: 10.12158/j.2096-3203.2024.05.024

融合注意力机制与 SAC 算法的虚拟电厂多能流低碳调度

俞晓荣¹, 徐青山², 杜璞良³, 王冬²

(1. 国网江苏省电力有限公司泰州供电分公司, 江苏 泰州 225300; 2. 东南大学电气工程学院, 江苏 南京 210096; 3. 上海电力大学经济与管理学院, 上海 200090)

摘要:虚拟电厂(virtual power plant, VPP)作为多能流互联的综合能源网络,已成为中国加速实现双碳目标的重要角色。但 VPP 内部资源协同低碳调度面临多能流的耦合程度紧密、传统碳交易模型参数主观性强、含高维动态参数的优化目标在线求解困难等问题。针对这些问题,文中提出一种融合注意力机制(attention mechanism, AM)与柔性动作评价(soft actor-critic, SAC)算法的 VPP 多能流低碳调度方法。首先,根据 VPP 的随机碳流特性,面向动态参数建立基于贝叶斯优化的改进阶梯型碳交易机制。接着,以经济效益和碳排放量为目标函数构建含氢 VPP 多能流解耦模型。然后,考虑到该模型具有高维非线性与权重参数实时更新的特征,利用融合 AM 的改进 SAC 深度强化学习算法在连续动作空间对模型进行求解。最后,对多能流调度结果进行仿真分析和对比实验,验证了文中方法的可行性及其相较于原 SAC 算法较高的决策准确性。

关键词:虚拟电厂(VPP);多能流;改进碳交易机制;深度强化学习;注意力机制(AM);柔性动作评价(SAC)算法
中图分类号: TM732 **文献标志码:** A **文章编号:** 2096-3203(2024)05-0233-14

0 引言

能源行业的低碳化是实现双碳目标的关键^[1-3]。虚拟电厂(virtual power plant, VPP)作为多能流互联的综合能源网络,将分散在多种能源网络中的资源进行整合与协同优化^[4-6]。目前 VPP 内部资源协同优化低碳调度面临传统碳交易模型参数主观性强、多能流的耦合程度紧密、待优化变量维度高、含动态参数的优化目标在线求解困难等问题^[7-9],同时随着氢流供给体系的加入和双碳目标的提出,模型维度和动态参数数量增加,而现有的算法未能充分考虑高维模型求解的难点,因此亟须研究一种面向高维待优化变量的含氢多能流 VPP 在线低碳调度方法,实现提升经济效益的同时降低碳排放量。

氢能凭借其高热值、环境友好、来源及应用广泛的特点吸引了学术界的高度关注。利用富余新能源电能电解制氢并结合成熟的储氢技术,可以有效促进未来电网大规模风电消纳,同时满足氢能高需求,具有巨大的应用潜力。文献[10]建立海上风电-氢能系统配置优化模型,讨论了各种敏感性因素对容量配置和经济性的影响;文献[11]基于电解制氢技术将氢能经济引入电力系统以期促进弃风消纳,建立以典型日运行成本最低为目标的虚拟弃风-制氢

联合体经济调度模型;文献[12]为充分发挥氢能在综合能源系统中的效用,提出一种考虑电热柔性负荷及氢能精细化建模的氢能综合能源系统(integrated energy system with hydrogen energy, H₂-IES)低碳运行方法;文献[13]提出风-火耦合多能系统的多尺度评价指标及其数学模型,并基于模糊层次分析-熵权法确定各指标的权重。上述文献多侧重于电气互联系统中的协同运行框架和双能源系统净负荷峰谷差的减小,未考虑双碳背景下氢能的广泛应用和传统碳交易机制的参数固定性问题。

另一方面,在 VPP 实际调度中,能源物理禀赋与负荷需求都具有较明显的动态特性。文献[14]采用强区间线性规划法中的两阶段分解算法处理模型中 VPP 系统的动态特性;文献[15]设计 VPP 有功功率-频率调节模式,但是提出的综合响应决策法只能进行分段求解;文献[16]提出基于在线分布式优化的自趋优运行方法,减少每一个控制时段的运算与通信时间;文献[17]提出事件驱动的刺激-反馈控制机制,根据负荷跟踪误差的偏差触发事件,事件启动进程引导分布式能源及时独立决策。然而,使用传统优化方法对 VPP 调度问题进行求解存在实时性问题,无法对 VPP 完成实时在线资源调度。

近年来,以深度强化学习为代表的人工智能算法在求解在线调度问题中应用广泛。文献[18]利用交替方向乘子法(alternating direction method of multipliers, ADMM)将原问题的目标函数等价地分解为若干个可求解的子问题,然后并行求解各子问

收稿日期:2024-03-03;修回日期:2024-05-11

基金项目:江苏省重点研发计划资助项目“基于物联网融合的用能互联网运行与交易应用科技示范”(BE2020688)

题,最后根据在线协调子问题的解得到原问题的全局解;文献[19]提出考虑高维不确定性的热电联产VPP两阶段随机鲁棒优化调度策略,开发了加速列与约束生成算法来对此优化调度问题进行求解。这类方法在优化任务中能得出实时性较优解,但是由于在线求解变量的高维性,求解结果往往不够精确。

文中基于现有研究,提出一种融合注意力机制(attention mechanism, AM)与柔性动作评价(soft actor-critic, SAC)算法的VPP多能流低碳调度方法。首先,针对VPP的实际碳排放量,建立一种考虑动态优化参数的奖惩阶梯型碳交易机制,充分调动VPP运营商参与碳交易市场的积极性^[20-21]。然后,基于碳捕集技术和含氢多能流网结构,以经济效益和碳排放量为目标函数建立优化调度模型,并利用AM-SAC深度强化学习算法在连续动作空间对问题进行求解。考虑到该模型具有高维非线性与权重参数实时更新的特征,传统SAC算法中的策略网络很容易忽略模型中的关键信息从而导致整体算法训练过程中的奖励值偏低,因此将AM加入SAC算法的策略网络中,利用AM捕捉状态空间中的重要信息,帮助策略网络作出最优的动作预测,相较于传统SAC算法能够给出更准确的优化调度结果。最后,进行仿真分析并对比多种优化算法性能,研究结果验证了优化调度模型的可行性与高效性。

1 考虑碳交易市场的多能流VPP优化调度架构

文中建立的考虑碳交易市场的多能流VPP优

化调度架构如图1所示。该架构由考虑改进碳交易机制的VPP模型与基于AM-SAC深度强化学习算法的实时决策模型构成。图中, s_t 为状态空间; a_t 为动作空间; r_t 为奖励值; s_{t+1} 为下个时刻状态。

考虑改进碳交易机制的VPP模型分为供给侧、多能流网和负荷侧。供给侧综合考虑了天然气市场、电网、碳市场,共同向多能流网内部供能,但不支持能量返送,碳市场与多能流网内部的碳排放配额量进行交易。负荷侧含热、冷、电、氢负荷,其中热负荷由热电联产机组与燃气锅炉共同供热,冷负荷由冰蓄冷空调和吸收式制冷机共同制冷,电负荷由光伏、风电、蓄电池、热电联产机组、电转气设备共同供电,氢负荷由储氢罐经过氢燃料电池与电解槽供氢。多能流网包含了上述热电联产机组、燃气锅炉、冰蓄冷空调、吸收式制冷机等设备。对传统火电机组的改造和电转气的应用是该VPP系统的主要碳减排方式,火电机组经过碳捕集及封存技术改造成碳捕集电厂,捕捉到的CO₂配额可以向碳市场交易。电转气设备可以购买碳市场低价的碳原料进行电转气以消纳多余的光伏、风电。

基于深度强化学习算法的实时决策模型将VPP系统的电、冷、热、氢负荷需求、风电和光伏的预测功率、当前所处调度时间断面及该断面下购电、购碳、产氢价格作为状态变量输入基于AM-SAC的深度学习底层框架。为了将输入层中各状态变量的经济目标与碳排放目标解耦,引入最大熵原理并采用2个双层神经网络互相配合,从经济效益和碳排放量两方面综合设计奖励函数,基于参数更新高效引导智能体探索,最终输出多能流VPP调度的最优策略。

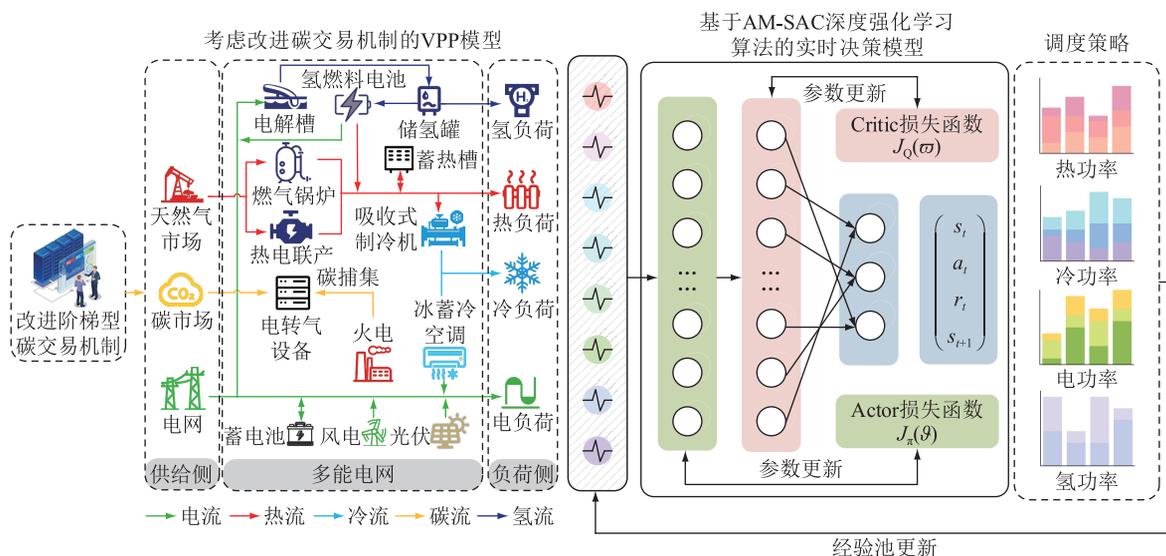


图1 考虑碳交易市场的多能流VPP优化调度架构

Fig.1 Optimal scheduling architecture of multi-energy flow VPP considering carbon market

2 考虑改进碳交易机制的 VPP 数学模型

碳交易机制主要是利用市场手段对生产过程中产生的碳排放进行控制,合理有效的碳交易机制能充分激发 VPP 运营商参与碳市场的积极性。文中首先根据 VPP 无偿得到的初始碳排放配额建立奖惩阶梯型碳交易成本模型,并引入贝叶斯参数优化器对重要参数进行优化;接着以最小化 VPP 运行成本和最小化碳排放量惩罚成本为目标函数构建优化调度模型;最后给出各单元的约束条件。

2.1 奖惩阶梯型碳交易机制

2.1.1 初始碳排放配额模型

VPP 中的碳排放源主要为热电联产机组、燃气锅炉、火电机组和向配电网购电。假设向配电网购买的电能均来自火电机组,在该机制中,VPP 可以在初始阶段无偿获得定量的碳排放配额,初始碳排放配额模型如式(1)一式(5)所示。

$$E_{CAF} = E_{ADN} + E_{CHP} + E_{GB} + E_{TC} \quad (1)$$

$$E_{ADN} = \sum_{t=1}^T (\lambda_{p1} + \lambda_{p2} P_{ADN,t} + \lambda_{p3} P_{ADN,t}^2) \quad (2)$$

$$E_{CHP} = \sum_{t=1}^T \lambda_{h2} (\zeta P_{CHP,t} + H_{CHP,t}) \quad (3)$$

$$E_{GB} = \sum_{t=1}^T (\lambda_{h1} + \lambda_{h2} H_{GB,t} + \lambda_{h3} H_{GB,t}^2) \quad (4)$$

$$E_{TC} = \sum_{t=1}^T (\lambda_{p1} + \lambda_{p2} P_{TC,t} + \lambda_{p3} P_{TC,t}^2) \quad (5)$$

式中: E_{CAF} 为 VPP 运营商无偿获得的所有碳排放配额; E_{ADN} 、 E_{CHP} 、 E_{GB} 、 E_{TC} 分别为向配电网购电、热电联产机组、燃气锅炉、常规火电机组的无偿碳排放总额; λ_{p1} 、 λ_{p2} 、 λ_{p3} 为单位电量的碳排放分配额计算系数; λ_{h1} 、 λ_{h2} 、 λ_{h3} 为单位热量的碳排放分配额计算系数^[22]; $P_{ADN,t}$ 、 $P_{CHP,t}$ 、 $P_{TC,t}$ 分别为 t 时刻向配电网购电功率、热电联产机组发电功率和常规火电机组发电功率; $H_{CHP,t}$ 、 $H_{GB,t}$ 分别为热电联产机组发热功率和燃气锅炉发热功率; ζ 为发电量折算成供热量的折算系数,文中取 $6 \text{ MJ}/(\text{kW} \cdot \text{h})$ ^[23]; $T=24 \text{ h}$ 。

2.1.2 基于贝叶斯参数优化的奖惩阶梯型碳交易模型

为提高 VPP 运营商参与碳交易市场的积极性,文中在传统阶梯型碳交易模型的基础上引入奖惩机制。同时,为了使阶梯型碳交易机制的参数选取更加合理,文中使用了一种基于贝叶斯参数的碳交易机制参数优化器。

需要明确的是,当 VPP 的碳排放量未达到碳排放配额时,可以向碳市场出售多余的碳配额并获得

一定额度的奖励收益;反之,当 VPP 的实际碳排放量大于碳排放配额时,须向碳市场购买额外的碳配额,交易价格与碳排放量正相关^[24]。奖惩阶梯型碳交易成本如式(6)一式(8)所示。

$$E_{CO_2} = E_{total} - E_{P2G} - E_{CS} \quad (6)$$

$$\Delta E = E_{CO_2} - E_{CAF} \quad (7)$$

$$F_{CO_2} =$$

$$\begin{cases} -\chi_{CO_2}(1+2\alpha)(\Delta E-d) & \Delta E > d \\ -\chi_{CO_2}(1+2\alpha)h - \chi_{CO_2}(1+\alpha)\Delta E & 0 < \Delta E \leq d \\ \chi_{CO_2}\Delta E & -d < \Delta E \leq 0 \\ \chi_{CO_2}d + \chi_{CO_2}(1+\beta)(\Delta E-h) & -2d < \Delta E \leq -d \\ \chi_{CO_2}(2+\beta)h + \chi_{CO_2}(1+2\beta)(-\Delta E-2h) & -3d < \Delta E \leq -2d \\ \chi_{CO_2}(3+3\beta)h + \chi_{CO_2}(1+3\beta)(-\Delta E-3d) & \Delta E \leq -3d \end{cases} \quad (8)$$

式中: E_{CO_2} 为 VPP 实际碳排放量; E_{total} 为热电联产机组、燃气锅炉和火电机组的碳排放总量; E_{P2G} 为电转气设备吸收的 CO_2 量; E_{CS} 为处于碳封存中的 CO_2 量; ΔE 为可出售或须购买的碳配额; F_{CO_2} 为 VPP 的碳交易成本; χ_{CO_2} 为碳市场上碳交易基价; α 、 β 分别为奖励机制参数和惩罚机制参数; d 为碳排放量的区间长度。

阶梯型碳交易机制中涉及的参数包括碳交易基价 χ_{CO_2} 、奖励机制参数 α 、惩罚机制参数 β 和区间长度 d ,合理设置参数对于奖惩机制作用的发挥至关重要。参数优化目标为碳排放成本计算误差最小化,优化参数设置为碳交易基价 χ_{CO_2} 、奖励机制参数 α 、惩罚机制参数 β 和区间长度 d ,并将其记为参数组合。优化过程中首先使用高斯过程作为概率代理模型拟合实际的目标函数;然后设优化过程中第 x 个元素抽样点为 P_x ,实际目标函数评价后的修正抽样点为 P'_x , P 为抽样点集合,基于概率代理模型的后验证信息构造采集函数 $F(x)$,即式(9),具体方式为求取目标值的高斯分布累计密度函数;最后使用采集函数 $F(x)$ 根据后验概率分布得到下一个较优的抽样点。算法流程如图 2 所示。

$$F(x) = p(f(x) \leq f(x^*) - \gamma) = \varphi\left(\frac{f(x^*) - \gamma - \mu(x)}{\sigma(x)}\right) \quad (9)$$

式中: $f(x)$ 为高斯过程的目标值; $f(x^*)$ 为高斯过程中当下最优目标值; γ 为平衡参数; $\varphi(\cdot)$ 为高斯分布累计密度函数; $\mu(x)$ 为目标函数的均值; $\sigma(x)$ 为目标函数的方差。

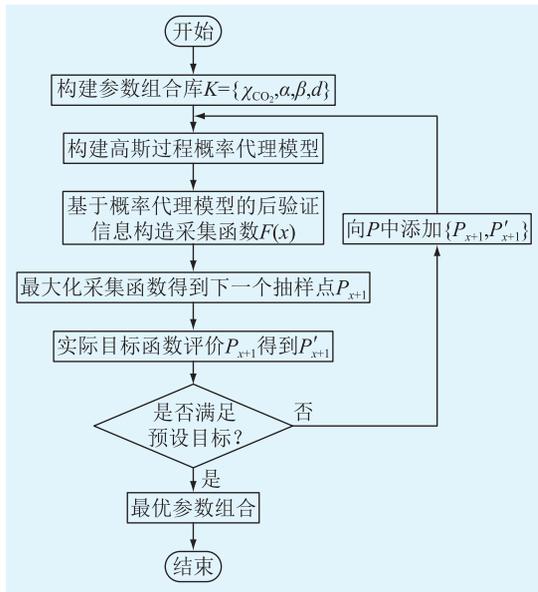


图2 参数优化器工作流程

Fig.2 Workflow of the parameter optimizer

2.2 设备建模

(1) 电解槽建模。电解槽通过电解水的方式制氢,所输出的氢能储存到储氢罐中以供应氢负荷及氢燃料电池。

$$G_{P2H,t} = P_{P2H,t} \eta_{P2H} \quad (10)$$

$$G_{P2H,t} + G_{HC,t} = G_{L,t} + G_{HFC,t} + G_{HS,t} \quad (11)$$

式中: $G_{P2H,t}$ 为 t 时刻电解槽的产氢功率; $P_{P2H,t}$ 为 t 时刻电解槽的用电功率; η_{P2H} 为产氢效率; $G_{HC,t}$ 为 t 时刻储氢罐供氢功率; $G_{L,t}$ 为 t 时刻氢负荷的耗氢功率; $G_{HFC,t}$ 为 t 时刻氢燃料电池因产热产电而产生的耗氢功率; $G_{HS,t}$ 为 t 时刻储氢罐储氢功率。

(2) 氢燃料电池建模。氢燃料电池化学特性,如热力学方程,见式(12)。

$$\begin{cases} P_{HFC,t} = \eta_{HFC} G_{H2P,t} \\ C_{HFC} \frac{dT_{HFC,t}}{dt} = G_{H2P,t} - P_{HFC,t} - \\ \frac{T_{HFC,t} - T_{ENV,t}}{R_{HFC}} - H_{HFC,t} \end{cases} \quad (12)$$

式中: $P_{HFC,t}$ 为 t 时刻氢燃料电池发电功率; $G_{H2P,t}$ 为 t 时刻氢燃料电池耗氢功率; η_{HFC} 为氢燃料电池发电转换效率; $T_{HFC,t}$ 为 t 时刻氢燃料电池的工作温度; C_{HFC} 、 R_{HFC} 分别为氢燃料电池的热容、热阻; $T_{ENV,t}$ 为 t 时刻环境温度; $H_{HFC,t}$ 为 t 时刻氢燃料电池的冷却循环水吸热功率。

(3) 碳捕集电厂建模。碳捕集电厂的碳捕集能耗包括固定能耗和运行能耗,因此碳捕集电厂 t 时刻净输出功率 $P_{CCG,t}$ 为:

$$\begin{cases} P_{CCG,t} = P_{\Sigma CCG,t} - P_S - P_R \\ P_R = \frac{1}{n_{E4CO_2,t}} E_{E4CO_2,t} \end{cases} \quad (13)$$

式中: $P_{\Sigma CCG,t}$ 为碳捕集电厂总输出功率; P_S 、 P_R 分别为碳捕集固定能耗和运行能耗; $n_{E4CO_2,t}$ 为 t 时刻消耗单位电能所处理的 CO_2 量; $E_{E4CO_2,t}$ 为 t 时刻碳捕集电厂正在处理的 CO_2 量。

(4) 电转气设备建模。电转气设备 t 时刻消耗的 CO_2 总量 $E_{P2G,t}$ 及生成天然气的体积 $V_{P2G,t}$ 为:

$$E_{P2G,t} = P_{P2G,t} n_{P2G,t} \quad (14)$$

$$V_{P2G,t} = \frac{3.6 \eta_{P2G}}{H_{GAS}} P_{P2G,t} \quad (15)$$

式中: $P_{P2G,t}$ 为 t 时刻电转气设备消耗的电功率; $n_{P2G,t}$ 为 t 时刻电转气设备消耗单位电能所需要吸收的 CO_2 量; η_{P2G} 为电转气设备的转换效率; H_{GAS} 为天然气热值。

(5) 储能设备建模。蓄电池 t 时刻的储电容量 $S_{BESS,t}$ 为:

$$S_{BESS,t} = S_{BESS,t-1} (1 - \rho_{BESS}) + \eta_{BTC} P_{BTC,t} - \frac{P_{BTD,t}}{\eta_{BTD}} \quad (16)$$

式中: ρ_{BESS} 为蓄电池的能量自损系数; $P_{BTC,t}$ 、 $P_{BTD,t}$ 分别为蓄电池 t 时刻的充电、放电功率; η_{BTC} 、 η_{BTD} 分别为蓄电池充电、放电效率。

(6) 燃气锅炉设备建模。燃气锅炉 t 时刻发热功率 $H_{GB,t}$ 为:

$$H_{GB,t} = \eta_{GB} H_{GAS} V_{GB,t} \quad (17)$$

式中: η_{GB} 为燃气锅炉的效率; $V_{GB,t}$ 为 t 时刻燃气锅炉消耗天然气的体积。

(7) 热电联产机组设备建模。热电联产机组 t 时刻发电功率 $P_{CHP,t}$ 及发热功率 $H_{CHP,t}$ 为:

$$\begin{cases} P_{CHP,t} = \eta_{CHPE} H_{GAS} V_{CHP,t} \\ H_{CHP,t} = \eta_{CHPH} H_{GAS} V_{CHP,t} \end{cases} \quad (18)$$

式中: η_{CHPE} 、 η_{CHPH} 分别为热电联产机组的发电、发热效率; $V_{CHP,t}$ 为 t 时刻热电联产机组消耗天然气的体积。

(8) 冰蓄冷空调设备建模。文中考虑并联式冰蓄冷空调,在用户使用冷能高峰期融冰制冷,在用户使用冷能低谷期制冰蓄冷。 t 时刻蓄冷罐中储存的冷却能 $T_{IS,t}$ 为:

$$T_{IS,t} = (1 - \rho_{ISAC}) T_{IS,t-1} + \eta_{IS} Q_{IS,t} - \frac{Q_{IM,t}}{\eta_{IM}} \quad (19)$$

式中: ρ_{ISAC} 为冰蓄冷空调设备的自损系数; η_{IS} 、 η_{IM} 分别为蓄冰和融冰系数; $Q_{IS,t}$ 、 $Q_{IM,t}$ 分别为 t 时刻的蓄冰与融冰功率。

(9) 吸收式制冷机设备建模。吸收式制冷机通过吸收热能制冷, t 时刻输出冷功率 $Q_{AR,t}$ 为:

$$Q_{AR,t} = \eta_{AR} H_{AR,t} \quad (20)$$

式中: $H_{AR,t}$ 为吸收式制冷机输入热功率; η_{AR} 为制冷效率。

2.3 目标函数

文中从经济效益与碳排放量两个角度出发, 建立计及奖惩碳交易机制的 VPP 优化调度模型, 并基于简单线性加权 (simple additive weighting, SAW) 模型综合考虑最小化 VPP 运行成本和最小化 VPP 碳排放量惩罚成本为优化目标。

2.3.1 最小化 VPP 运行成本

$$\min f_1 = \min (F_{H_2} + F_{HS} + F_{P2H} + F_{HFC} + F_{GRID} + F_{GAS} + F_{CO_2} + F_{CCG} + F_{CS} + F_{P2G} + F_{\delta}) \quad (21)$$

式中: f_1 为 VPP 运行成本; F_{H_2} 、 F_{HS} 、 F_{P2H} 、 F_{HFC} 分别为制氢成本和储氢罐、电解槽、氢燃料电池的运维成本; F_{GRID} 、 F_{GAS} 分别为 VPP 购电和购气成本; F_{CO_2} 为碳交易成本; F_{CCG} 为碳捕集电厂的燃料成本; F_{CS} 为碳封存成本; F_{P2G} 为电转气设备成本; F_{δ} 为 VPP 其他设备运维成本。

(1) 制氢成本及运维成本。氢能相关成本包括制氢成本和储氢罐、电解槽、氢燃料电池的运维成本。

$$F_{H_2} = \sum_{t=1}^T \lambda_{H_2} m_{H_2,t} \quad (22)$$

$$F_{HS} = \sum_{t=1}^T \lambda_{HS} m_{HS,t} \quad (23)$$

$$F_{P2H} = \sum_{t=1}^T \lambda_{P2H} P_{P2H,t} \quad (24)$$

$$F_{HFC} = \sum_{t=1}^T \lambda_{HFC} P_{HFC,t} \quad (25)$$

式中: $m_{H_2,t}$ 为 t 时刻产氢质量; $m_{HS,t}$ 为 t 时刻储氢罐中氢气的质量; λ_{H_2} 、 λ_{HS} 、 λ_{P2H} 、 λ_{HFC} 分别为单位制氢成本和储氢罐、电解槽、氢燃料电池的单位运维成本。

(2) VPP 购电和购气成本。

$$F_{GRID} = \sum_{t=1}^T P_{ADN,t} \lambda_{GRID,t} \quad (26)$$

$$F_{GAS} = \sum_{t=1}^T (V_{CHP,t} + V_{GB,t}) \lambda_{GAS,t} \quad (27)$$

式中: $\lambda_{GRID,t}$ 、 $\lambda_{GAS,t}$ 分别为 t 时刻电价和天然气价格。

(3) 碳捕集电厂的燃料成本。

$$F_{CCG} = \sum_{t=1}^T (\lambda_{CCG1} P_{CCG1,t}^2 + \lambda_{CCG2} P_{CCG2,t} + \lambda_{CCG3}) \quad (28)$$

式中: λ_{CCG1} 、 λ_{CCG2} 、 λ_{CCG3} 为 t 时刻碳捕集电厂的燃料成本系数。

(4) 碳封存成本。

$$F_{CS} = \sum_{t=1}^T \lambda_{CS,t} E_{CS,t} = \sum_{t=1}^T \lambda_{CS,t} (E_{total,t} - E_{CO_2,t} - E_{P2G,t}) \quad (29)$$

式中: $\lambda_{CS,t}$ 为 t 时刻碳封存成本系数。

(5) 电转气设备成本。电转气设备的成本包括运行成本和原料成本^[25]。

$$F_{P2G} = \sum_{t=1}^T (\lambda_{CO_2,t} E_{P2G,t} + \delta_{P2G} E_{P2G,t}) \quad (30)$$

式中: $\lambda_{CO_2,t}$ 为 t 时刻单位 CO_2 原料的购买成本; δ_{P2G} 为电转气设备运行成本系数。

(6) VPP 其他设备运维成本。

$$F_{\delta} = \sum_{t=1}^T \sum_{j=1}^M \delta_{j,t} P_{j,t} \quad (31)$$

式中: M 为 VPP 中供能设备的数量, 包括光伏设备、风电设备、燃气锅炉、热电联产机组、吸收式制冷机设备等; $\delta_{j,t}$ 为 t 时刻设备 j 输出单位功率的运维成本; $P_{j,t}$ 为 t 时刻设备 j 输出功率。

2.3.2 最小化 VPP 碳排放量惩罚成本

衡量 VPP 碳排放量时, 文中综合考虑碳捕集电厂与购电产生的碳排放量。为统一多个目标函数的量纲以便于线性加权, 文中引入碳排放量惩罚系数, 将碳排放量与惩罚系数相乘得到惩罚成本, 从而将最小化 VPP 碳排放量等价转化为最小化 VPP 碳排放量惩罚成本。

$$\min f_2 = \min \mu_{CO_2} (E_{CO_2} + \kappa_{GRID-CO_2} P_{ADN,t}) \quad (32)$$

式中: f_2 为碳排放量惩罚成本; μ_{CO_2} 为碳排放量惩罚系数; $\kappa_{GRID-CO_2}$ 为购电的碳排放系数。

2.3.3 SAW 目标函数

文中所提调度模型的优化目标为在最低碳排放量水平下的最小化运行成本。直接求解双目标优化调度模型会存在伪有效解、漂移现象等不完美的理论问题。为此, 文中引入 SAW 模型转化多目标优化模型。SAW 模型缩放表达式为:

$$f_i^* = \frac{f_i^{\max} - f_i}{f_i^{\max} - f_i^{\min}} \quad (33)$$

式中: f_i^* 为缩放后目标函数, $i = 1, 2$; f_i 为原目标函数; f_i^{\max} 、 f_i^{\min} 分别为目标函数 f_i 的最大值和最小值。

仅用缩放后的值代表原目标使得求解相对容易, 同时也能保留解的精度。将缩放后的单目标函数进行线性加权, 用一个综合的效用函数来代表总体优化的目标, 根据实际应用地的经济需求和环保政策选择权重。文中选取权重相等的特殊情况, 得

到 SAW 模型加权表达式。

$$f = \min \sum f_i^* \quad (34)$$

2.4 约束条件

(1) 功率平衡约束。

$$\begin{cases} G_{P2H,t} + G_{HC,t} = G_{L,t} + G_{HFC,t} + G_{HS,t} \\ P_{L,t} + P_{ISAC,t} + P_{BTC,t} = P_{HFC,t} + P_{PV,t} + P_{WT,t} + \\ \quad P_{CCG,t} + P_{CHP,t} + P_{BTD,t} + P_{ADN,t} \\ H_{L,t} + H_{AR,t} + H_{HSTC,t} = \\ \quad H_{CHP,t} + H_{HFC,t} + H_{GB,t} + H_{HSTD,t} \\ Q_{L,t} = Q_{AR,t} + Q_{ISAC,t} + Q_{IS,t} \end{cases} \quad (35)$$

式中： $P_{L,t}$ 、 $H_{L,t}$ 、 $Q_{L,t}$ 分别为 t 时刻用户电负荷、热负荷与冷负荷； $P_{ISAC,t}$ 为 t 时刻冰蓄冷空调耗电功率； $P_{PV,t}$ 、 $P_{WT,t}$ 分别为 t 时刻光伏和风电上网功率； $H_{HSTC,t}$ 、 $H_{HSTD,t}$ 分别为 t 时刻蓄热槽的储、放热功率； $Q_{ISAC,t}$ 为 t 时刻冰蓄冷空调的输出冷功率。

(2) 电解槽约束。

$$P_{P2H,min} \leq P_{P2H,t} \leq P_{P2H,max} \quad (36)$$

$$|P_{P2H,t+1} - P_{P2H,t}| \leq \Delta P_{P2H} \quad (37)$$

式中： $P_{P2H,max}$ 、 $P_{P2H,min}$ 分别为电解槽用电功率上、下限； ΔP_{P2H} 为电解槽的爬坡出力约束。

(3) 碳捕集电厂约束。

$$P_{CCG,min} \leq P_{CCG,t} \leq P_{CCG,max} \quad (38)$$

$$|P_{CCG,t+1} - P_{CCG,t}| \leq \Delta P_{CCG} \quad (39)$$

$$\kappa_{CCG} = \frac{E_{CS,t} + E_{P2G,t}}{0.96P_{CCG,t}} \quad (40)$$

式中： $P_{CCG,max}$ 、 $P_{CCG,min}$ 分别为碳捕集电厂净输出功率上、下限； ΔP_{CCG} 为碳捕集电厂的爬坡出力约束； κ_{CCG} 为碳捕集系统的烟气分流比， $0 \leq \kappa_{CCG} \leq 1$ 。

(4) 电转气设备约束。

$$0 \leq P_{P2G,t} \leq P_{P2G,max} \quad (41)$$

式中： $P_{P2G,max}$ 为电转气设备的最大运行功率。

(5) 储能设备约束。

$$S_{BESS,min} \leq S_{BESS,t} \leq S_{BESS,max} \quad (42)$$

$$0 \leq P_{BTC,t} \leq P_{BTC,max} \quad (43)$$

$$0 \leq P_{BTD,t} \leq P_{BTD,max} \quad (44)$$

$$\sum_{t=1}^T \eta_{BTC} P_{BTC,t} - \sum_{t=1}^T \frac{P_{BTD,t}}{\eta_{BTD}} = 0 \quad (45)$$

式中： $S_{BESS,max}$ 、 $S_{BESS,min}$ 分别为蓄电池储电容量的上、下限； $P_{BTC,max}$ 、 $P_{BTD,max}$ 分别为蓄电池充电、放电功率的最大值。热储能约束与电储能约束在形式上相同。

(6) 燃气锅炉设备约束。

$$H_{GB,min} \leq H_{GB,t} \leq H_{GB,max} \quad (46)$$

式中： $H_{GB,max}$ 、 $H_{GB,min}$ 分别为燃气锅炉输出热功率上、下限。

(7) 热电联产机组设备约束。

$$P_{CHP,min} \leq P_{CHP,t} \leq P_{CHP,max} \quad (47)$$

$$H_{CHP,min} \leq H_{CHP,t} \leq H_{CHP,max} \quad (48)$$

$$|P_{CHP,t+1} - P_{CHP,t}| \leq \Delta P_{CHP} \quad (49)$$

式中： $P_{CHP,max}$ 、 $P_{CHP,min}$ 分别为热电联产机组发电功率上、下限； $H_{CHP,max}$ 、 $H_{CHP,min}$ 分别为热电联产机组发热功率上、下限； ΔP_{CHP} 为热电联产机组爬坡出力约束。

(8) 冰蓄冷空调设备约束。

$$Q_{ISAC,min} \theta_{ISAC,t} \leq Q_{ISAC,t} \leq Q_{ISAC,max} \theta_{ISAC,t} \quad (50)$$

$$0 \leq Q_{IS,t} \leq Q_{ISAC,max} \theta_{IS,t} \quad (51)$$

$$Q_{ISAC,min} \leq Q_{ISAC,t} + Q_{IS,t} \leq Q_{ISAC,max} \quad (52)$$

$$0 \leq Q_{IM,t} \leq Q_{IM,max} \theta_{IM,t} \quad (53)$$

$$\theta_{IS,t} = 0 \quad t \notin V \quad (54)$$

$$\theta_{IM,t} = 0 \quad t \in V \quad (55)$$

式中： $\theta_{ISAC,t}$ 、 $\theta_{IS,t}$ 、 $\theta_{IM,t}$ 分别为 t 时刻输出冷功率、蓄冰、融冰对应的状态标记； $Q_{ISAC,max}$ 、 $Q_{ISAC,min}$ 分别为冰蓄冷空调输出冷功率上、下限； $Q_{IM,max}$ 为蓄冷罐融冰功率上限； V 为电价谷时时刻集合。

(9) 吸收式制冷机设备约束。

$$0 \leq H_{AR,t} \leq H_{AR,max} \quad (56)$$

式中： $H_{AR,max}$ 为吸收式制冷机输入热功率上限。

3 基于 AM-SAC 深度强化学习算法的实时决策模型

使用 SAW 模型综合后的 VPP 优化调度模型总体优化目标面临优化变量维度高且需要在线决策的挑战，传统的深度强化学习算法并不适用。文中首先介绍基于深度强化学习的 SAC 算法^[26]解决优化问题的通用模型，然后给出 AM-SAC 算法求解 VPP 优化调度问题的流程。

3.1 SAC 模型与 AM

3.1.1 SAC 通用模型

非确定性策略算法 SAC，不同于以最大化智能体奖励为目标的传统深度强化学习算法，其引入最大熵原理^[27-28]以最大化智能体奖励和动作的熵值，防止智能体的策略过早收敛于局部最优解。同时，还可以在奖励函数中鼓励智能体探索，使得算法更加稳定。因此，与其他深度强化学习算法相比，SAC 算法在优化变量维度高、动作空间连续的 VPP 优化调度场景中具有优越性。

SAC 算法包含 1 个策略网络和 2 组 Q 网络。其中，策略函数 π 根据当前状态空间 s_t 输出概率分

布的均值 μ 和方差 σ ,同时结合标准正态分布采样得到的噪声 τ ,输出动作空间 a_t 。2组Q网络用于评估当前状态下采取的动作价值,计算并选取2组Q网络输出的最小值,从而有效防止Q值的过高估计。

3.1.2 AM

文中设计的VPP优化调度问题中,目标函数为最小化VPP运行成本以及最小化碳排放量惩罚成本,AM-SAC算法中的状态空间包括电能、热能、冷能、氢能的多能流负荷需求序列、分时电价、分时气价等。在状态空间的众多状态量中必然存在对优化调度最终结果影响较大的因素,同样也会存在影响较小的因素,多个状态量对深度强化学习实时优化调度结果影响的权重分布处于实时动态变化之中。SAC算法中的策略网络很容易忽略其中的关键信息,导致整体算法训练过程中的奖励值偏低。因此,文中考虑将AM加入SAC算法的策略网络中,策略网络利用AM捕捉状态空间中的重要信息,自动关注如多能流负荷需求序列等重要的状态量以增强网络的感知能力,帮助策略网络作出最优的动作预测,包括除风力和光伏以外的各设备的功率、电池的转化功率和向配网购电功率等,相较于传统SAC算法能够给出更准确的优化调度结果。

AM以网络结构模型实现,包括2个隐含层:第一层为特征层,由关键特征向量、值特征向量和选择特征向量构成;第二层为注意力层,注意力层节点与值特征向量相连至终端节点层形成输出层。AM的网络结构如图3所示。其中, N_1 、 N_2 、 N_3 、 N_4 分别为输入向量、注意力向量、选择特征向量和输出特征向量个数。

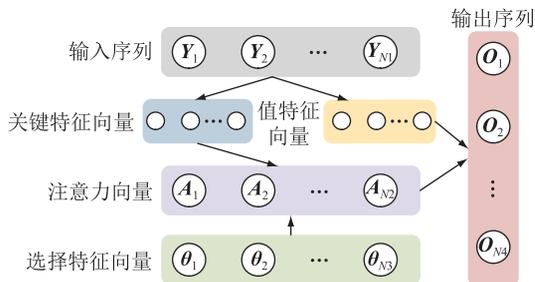


图3 AM网络结构

Fig.3 Network architecture of AM

在AM网络结构中,假设 Y 为输入层的状态向量,特征层中的关键特征向量 K 、值特征向量 V 分别为:

$$K = W_K \cdot Y \quad (61)$$

$$V = \text{Sigmoid}(W_V \cdot Y) \quad (62)$$

式中: W_K 、 W_V 为权值矩阵; $\text{Sigmoid}(\cdot)$ 为非线性激

活函数。

注意力层节点与值特征向量相连,并结合值特征向量作为输出层,第 n 个注意力层节点注意力表达与最终网络输出如式(63)、式(64)所示。

$$A_n = e^{\theta_n K} / \left[\left(\sum_{m=1}^N e^{\theta_m K} \right) - e^{\theta_n K} \right] \quad (63)$$

$$O = A \cdot V = \sum_{m=1}^N (A_m \cdot V_m) \quad (64)$$

式中: A_n 、 θ_n 分别为第 n 个注意力向量和选择特征向量; N 为输入层向量含有的元素数量; O 为AM网络最终输出; V_m 为第 m 个注意力向量的值特征向量。

3.1.3 AM-SAC模型

AM-SAC算法由1个AM网络、1个策略网络和2组Q网络构成,在训练过程中根据当前时刻的输入信息生成新的权重参数分布,其结构如图4所示。首先策略网络根据当前状态空间 s_t 随机生成动作空间 a_t ,计算奖励值 r_t 和下一个时刻状态 s'_t ,并通过注意力网络获取下一时刻权重分布 O ;然后将数据组 $\{s_t, a_t, r_t, s'_t\}$ 储存至经验池中,当经验池中达到一定阈值时,抽取一批数据计算2组Q网络最小平均价值 $\min(Q_1, Q_2)$ 、策略梯度、网络损失值等,并更新各网络参数;最后,更新当前状态空间 s'_t 为 s_t ,重复上述过程直至收敛。

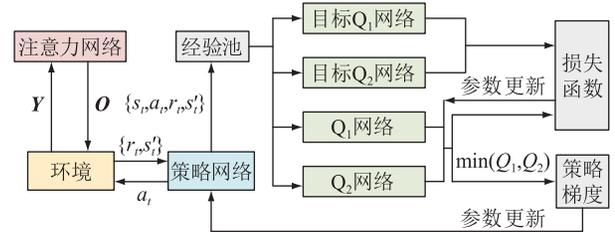


图4 AM-SAC算法结构

Fig.4 Structure of AM-SAC algorithm

在AM-SAC模型中,假设当前策略为 $\pi(\cdot | s_t)$,则AM-SAC算法的累计奖励函数可以表示为:

$$J(\pi) = \sum_{t=0}^K E_{(s_t, a_t) \sim \rho_\pi} (\omega(s_t) r(s_t, a_t) + \alpha H(\pi(\cdot | s_t))) \quad (65)$$

式中: E 为数学期望; K 为智能体每轮与环境交互的总时间步数; ρ_π 为策略 π 下轨迹 (s_t, a_t) 的分布; $\omega(s_t)$ 为注意力网络根据当前状态信息输出奖励函数的权重分布; $r(s_t, a_t)$ 为智能体执行动作 a_t 获得的奖励; α 为控制优化目标倾向的权重参数; $H(\pi(\cdot | s_t))$ 为当前策略的熵值, $H(\pi(\cdot | s_t)) =$

$E(-\lg \pi(\cdot | s_t))$ 。

最终,AM-SAC 算法求解得到的最优策略 π^* 可以表示为:

$$\pi^* = \operatorname{argmax} J(\pi) \quad (66)$$

3.2 算法设计

3.2.1 状态空间

在深度强化学习决策过程中,状态信息代表智能体所感知到的环境信息和动态变化,是深度强化学习算法生成决策和评估累计奖励的依据。

文中状态信息包括负荷需求(氢、电、冷、热负荷)、风力和光伏的输出功率以及当前所处调度时段。状态空间的具体定义表示为:

$$s_t = \{G_{L,t}, P_{L,t}, H_{L,t}, Q_{L,t}, P_{PV,t}, P_{WT,t}, \lambda_{GRID,t}, \lambda_{GAS,t}, \lambda_{CS,t}, t\} \quad (67)$$

3.2.2 动作空间

根据 VPP 设备建模与约束条件,设置动作空间,包括除风力和光伏以外的各设备的功率、电池的转化功率和向配网购电功率。在设计动作空间时设置动作的上、下界,使得生成的动作满足 2.4 节中的约束条件。动作空间的具体定义表示为:

$$a_t = \{G_{P2H,t}, G_{HFC,t}, P_{P2G,t}, P_{BTD,t}, P_{CHP,t}, H_{GB,t}, H_{CHP,t}, H_{AR,t}, Q_{IS,t}, Q_{AR,t}, P_{ADN,t}\} \quad (68)$$

3.2.3 奖励函数

奖励函数在一定程度上反映了训练的目的。根据目标函数优化要求,奖励函数从经济效益和碳排放量两方面进行设计,在最大化奖励的同时考虑实际系统约束条件,将约束条件作为惩罚项合并到奖励函数中。

$$r(a_t, s_t) = -\xi_0(f_t + \xi_1\psi_{1,t} + \xi_2\psi_{2,t}) \quad (69)$$

式中: ξ_0 为奖励函数缩放系数; f_t 为 t 时刻考虑经济效益和碳排放量的综合效用函数; ξ_1 、 $\psi_{1,t}$ 分别为违反功率平衡约束的惩罚系数及 t 时刻惩罚项; ξ_2 、 $\psi_{2,t}$ 分别为违反设备约束的惩罚系数及 t 时刻惩罚项。为体现功率平衡约束相较于设备上下限约束的严格性,设定 $\xi_1 > \xi_2$ 。

$$\psi_{1,t} = |G_{VAC,t}| + |P_{VAC,t}| + |H_{VAC,t}| + |Q_{VAC,t}| \quad (70)$$

$$\psi_{2,t} = \psi_{P2H,t} + \psi_{CCG,t} + \psi_{P2G,t} + \psi_{BESS,t} + \psi_{GB,t} + \psi_{CHP,t} + \psi_{ISAC,t} + \psi_{AR,t} \quad (71)$$

式中: $G_{VAC,t}$ 、 $P_{VAC,t}$ 、 $H_{VAC,t}$ 、 $Q_{VAC,t}$ 分别为氢能、电能、热能、冷能在 t 时刻的不平衡量; $\psi_{P2H,t}$ 、 $\psi_{CCG,t}$ 、 $\psi_{P2G,t}$ 、 $\psi_{BESS,t}$ 、 $\psi_{GB,t}$ 、 $\psi_{CHP,t}$ 、 $\psi_{ISAC,t}$ 、 $\psi_{AR,t}$ 分别为 t 时刻违反电解槽、碳捕集电厂、电转气设备、储能设备、燃气锅炉、热电联产机组、冰蓄冷空调、吸收式制冷设备约束的惩罚项。

以燃气锅炉为例, t 时刻燃气锅炉违反设备约束的惩罚项表达式为式(72),其余设备同理,因篇幅限制文中不再赘述。

$$\psi_{GB,t} = |H_{GB,t} - H_{GB,max}| + |H_{GB,min} - H_{GB,t}| \quad (72)$$

4 算例分析

4.1 算例设置

为验证文中所提计及碳捕集和多能流网的 VPP 多目标优化调度方法的效果,选取夏季某典型日进行分析,设置 VPP 优化调度的仿真时间为 24 h,仿真步长为 1 h。各设备参数如表 1 所示,采用分散管理方式,默认存在聚合路径,其示意如图 5 所示。

表 1 设备参数

Table 1 Equipment parameters

参数	数值	参数	数值
$\zeta / [MJ \cdot (kW \cdot h)^{-1}]$	6	$n_{E4CO_2,t} / [t \cdot (MW \cdot h)^{-1}]$	3.6
η_{IM}	0.71	η_{IS}	0.68
η_{P2G}	0.65	ρ_{BESS}	0.025
η_{BTC}	0.95	η_{BTD}	0.95
η_{CHPE}	0.33	η_{CHPH}	0.42
η_{P2H}	0.85	η_{HFC}	0.45
$P_{CCG,min}/MW$	120	$\Delta P_{P2H}/MW$	80
$P_{CCG,max}/MW$	430	$\Delta P_{CCG}/MW$	70
$P_{CHP,min}/MW$	0	$H_{CHP,min}/MW$	0
$P_{CHP,max}/MW$	150	$H_{CHP,max}/MW$	160
$S_{BESS,min}/MW$	10	$P_{P2H,min}/MW$	0
$S_{BESS,max}/MW$	130	$P_{P2H,max}/MW$	160

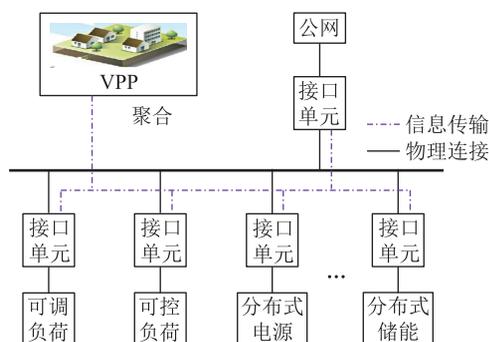


图 5 集中控制结构下的 VPP

Fig.5 VPP under centralized control structure

我国现行天然气定价采用成本加成法,由国家进行统一定价,不能随供需关系的波动及时调节^[29],因此文中算例考虑天然气峰谷分时定价,配网的分时电价以及分时气价如表 2 所示。AM-SAC 算法的学习率为 0.000 3,折扣因子为 0.95,训练批

次大小为 256 个,每次迭代的步长设置为 1 000 步,探索函数设置为随机采样。仿真实验的硬件配置为一台处理器为 R75800X、显卡为 3090Ti、内存为 32 GB 的服务器。

表 2 分时电价与分时气价
Table 2 Time-of-use electricity price and time-of-use gas price

时段	分时电价/ [元·(kW·h) ⁻¹]	分时气价/ (元·m ⁻³)
01:00—07:00	0.36	1.8
07:00—10:00	0.67	3.2
10:00—13:00	0.84	3.2
13:00—14:00	0.66	3.2
14:00—18:00	0.66	2.7
18:00—20:00	0.86	2.7
20:00—22:00	0.86	3.2
22:00—23:00	0.66	3.2
23:00—24:00	0.36	3.2

4.2 贝叶斯参数选择

为验证贝叶斯超参数优化结果的有效性,以算法评价效用值为目标函数,阶梯型碳交易机制中待优化参数集包括碳交易基价 χ_{CO_2} 、奖励机制参数 α 、惩罚机制参数 β 和区间长度 d ,寻优区间为 $\chi_{\text{CO}_2} \in [0.15, 0.2]$, $\alpha \in [0.2, 0.3]$, $\beta \in [0.02, 0.08]$, $d \in [20\ 000, 25\ 000]$,设置总迭代次数为 20 次。迭代过程如图 6 所示。

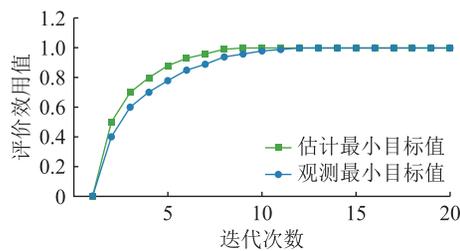


图 6 贝叶斯超参数优化迭代过程

Fig.6 Iterative process of Bayesian hyperparameter optimization

由图 6 可知,从第 12 次迭代开始观测最小目标值与估计最小目标值数值接近且趋于稳定,由此可以得到贝叶斯优化后最优超参数组合为 $\chi_{\text{CO}_2} = 0.168$, $\alpha = 0.21$, $\beta = 0.21$, $d = 21\ 560$ 。相较于未考虑贝叶斯参数优化所得的多能流 VPP 优化调度结果,优化后碳排放量惩罚成本减少了 256 元,运行成本减少了 2.24 万元,验证了在阶梯碳交易模型中引入贝叶斯参数优化算法能有效提高 VPP 的经济性和低碳性。

4.3 训练过程分析

文中算例选取 2021 年 7 月 20 日至 2021 年 7

月 25 日的 VPP 数据对调度模型进行离线训练。智能体接收多能流负荷需求、光伏出力、风电出力等训练数据后对深度强化学习模型进行训练,直至算法收敛至最大奖励值。测试集选取 2021 年 7 月 26 日的 VPP 数据,训练结束的调度模型直接应用于测试集,以充分验证文中所提基于 AM-SAC 算法的 VPP 多能流低碳调度方法的有效性。

为验证改进阶梯型碳交易机制和 AM 对改善 VPP 优化调度结果的有效性,文中在 SAC 算法中依次引入改进阶梯碳交易机制以及 AM,并对比不同场景的奖励值训练过程,数据结果如图 7 所示。考虑阶梯型碳交易机制的 AM-SAC 算法在迭代步数 1 500 左右已经达到收敛状态,并拥有最高的奖励值,而其他 2 种算法在迭代步数超过 2 500 后才趋于稳定,且奖励值偏低。由此可以得出结论,AM 和改进阶梯碳交易机制的引入有效地提升了 SAC 算法收敛效果以及平均奖励值。另外,由于 AM 能够对状态空间权重进行动态调整,不断适应不同数据带来的特征权重分布差异并使模型作出更优的选择,因此 AM-SAC 算法在应对与训练集数据不同的测试集时具有较好的泛化性能和适应性。

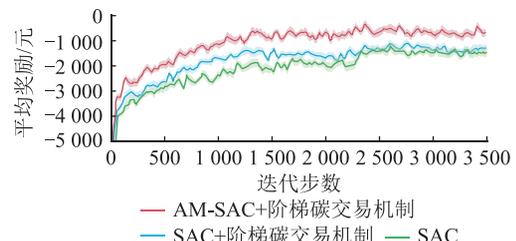


图 7 不同算法奖励值训练过程对比

Fig.7 Comparison of reward training process of different algorithms

4.4 VPP 优化调度结果分析

根据优化调度结果,VPP 内部的电、热、冷、氢能源调度平衡结果分别如图 8—图 11 所示。

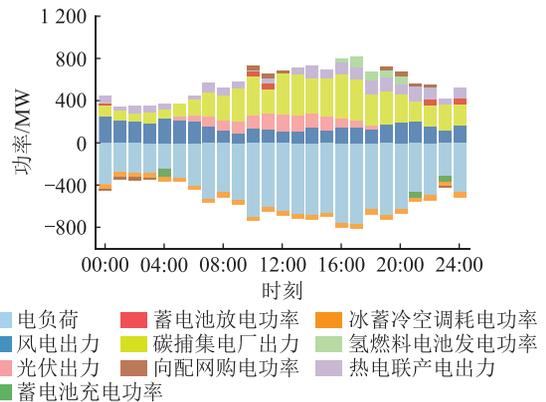


图 8 电能平衡

Fig.8 Electricity dispatching balance

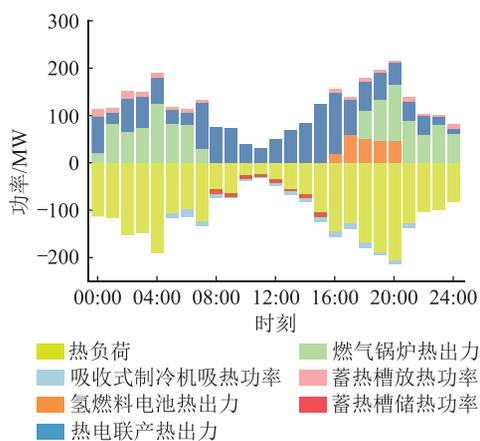


图9 热能平衡

Fig.9 Heat dispatching balance

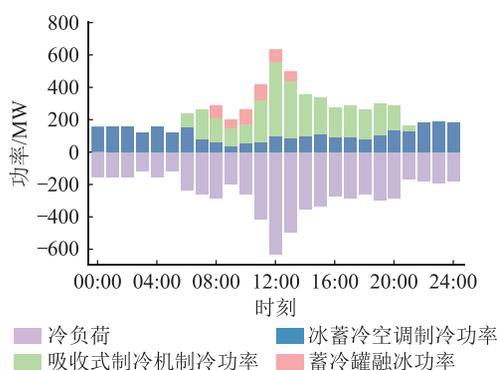


图10 冷能平衡

Fig.10 Cold dispatching balance

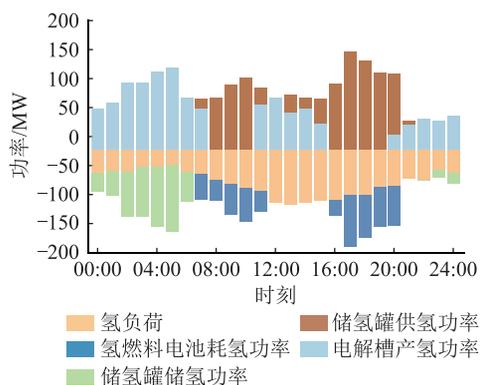


图11 氢能平衡

Fig.11 Hydrogen dispatching balance

4.4.1 电能调度结果

由图8可知,在00:00—06:00及23:00—24:00,电负荷需求量在400 MW左右,处于较低水平,此时平均70%的电负荷由风电提供,剩余的电负荷由碳捕集电厂提供,富裕的电能通过冰蓄冷空调转换为冷能,同时蓄電池处于充电状态。在07:00—09:00及13:00—17:00,电负荷主要由光伏、风电、碳捕集电厂以及热电联产机组提供。而在10:00—12:00及18:00—20:00,电负荷需求量较高,均超过

600 MW,除了光伏、风电、碳捕集电厂等提供能源之外,氢燃料电池进行发电,蓄電池进行放电,在电负荷供应不足的情况下VPP需要向配网购电。由电负荷调度平衡结果可推测,在夜间,风电出力大、发电成本更低且更环保,因此风电作为承担电负荷的主要能源,同时多余的风能可转换为冷能从而提高风电的消纳能力;而在白天,风电出力减少,因此主要由光伏、碳捕集电厂等提供电能。VPP向配网购电时段则集中在设备供应能力不足的电负荷高峰期。蓄電池在电价谷时段充电,在电价峰时段放电,氢燃料电池在电价峰时段出力,能够利用峰谷电价差有效降低向配网购电成本、减少碳排放。

4.4.2 热流调度结果

由图9可知,在00:00—04:00及18:00—24:00,热负荷需求量较大,每小时平均需求量为140 MW,主要由燃气锅炉、热电联产机组提供热出力,蓄热槽放热,同时氢燃料电池供应部分热负荷。在05:00—17:00,平均热负荷需求量较其他时段下降约57%,主要由热电联产机组提供热出力,部分热能通过吸收式制冷机转化为冷能,蓄热槽储热。通过分析热负荷调度平衡结果可以得知,在白天,热负荷需求量小而电负荷需求量大,为缓解电负荷的压力,热负荷主要由热电联产机组提供。蓄热槽在热负荷高峰时期放热、在热负荷低谷时期储热,同时氢燃料电池在热负荷高峰时期产热,能够降低购买天然气的成本、减少碳排放。

4.4.3 冷流调度结果

由图10可知,在00:00—05:00及21:00—24:00,冷负荷需求量约为160 MW,由冰蓄冷空调提供制冷功率。在06:00—20:00,平均冷负荷需求量大幅增加,约为308 MW,主要由吸收式制冷机、冰蓄冷空调提供冷能,同时蓄冰罐融冰制冷。冷负荷调度平衡结果显示,在夜间,电负荷需求量小,富裕的电能通过冰蓄冷空调转换为冷能,因此夜间主要由冰蓄冷空调提供制冷功率。同理,在白天,热负荷需求量减少、电负荷需求量增加,吸收式制冷机出力增加、冰蓄冷空调出力减少。

4.4.4 氢流调度结果

由图11可知,在00:00—06:00及21:00—24:00,氢负荷需求量较少,每小时需求量为42 MW,主要由电解槽提供产氢功率,同时储氢罐储存多余的氢能。在07:00—20:00,氢负荷平均需求增加至82 MW,主要由电解槽、储氢罐提供氢能,同时氢燃料电池消耗氢能产热产电。根据氢能调度平衡结果,为提高系统的整体效益、降低购能成本,

电解槽主要在电价谷时段以及光伏出力大的时段产氢,储氢罐在谷电时段储氢、在峰电时段供氢,氢燃料电池选择在峰电时段运行。

4.4.5 碳流调度结果

图12为VPP的碳量平衡结果。在00:00—06:00及23:00—24:00,碳捕集电厂产生的CO₂量均值为64 t,与均值为50 t的碳配额相近,且碳捕集电厂产生的CO₂中77.8%的部分被捕集进入电转气设备,富余的碳配额向碳市场出售。在07:00—09:00及15:00—22:00,碳捕集电厂产生的CO₂均值急剧增加至239 t,碳配额均值为177 t,CO₂捕集量均值也增加至122 t,此时碳配额仍有剩余的部分出售给碳市场。而在10:00—13:00,碳配额与碳捕集量不足以平衡碳捕集电厂产生的CO₂,因此VPP向碳市场购买共计142 t的碳配额。由碳量平衡数据可知,一天内绝大部分时段碳捕集电厂产生的大部分CO₂被碳捕集设备捕捉,因此剩余的碳配额可以向碳市场出售获取利润,只有在正午等少量时段VPP需要向碳市场购买碳配额。

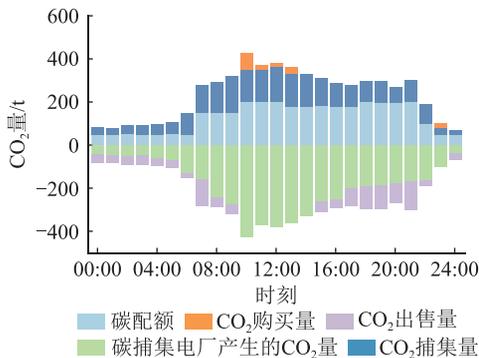


图12 碳量平衡

Fig.12 Carbon dispatching balance

为验证含氢VPP的优越性,即分析引入氢燃料电池对多能流VPP优化调度结果的影响,文中设置2个场景对比经济性和低碳性,如图13所示。场景1中不考虑氢燃料电池,电解槽产生的氢气经过储氢罐仅提供给氢负荷;场景2在场景1的基础上引入氢燃料电池,在峰电时刻消耗氢能以降低同时段火机组的出力,电解槽产生的氢气经过储氢罐一方面提供给氢负荷,另一方面提供给氢燃料电池进行产热产电。

由图13可知,在07:00—11:00以及16:00—20:00,场景2与场景1相比碳排放量惩罚成本平均降低了23.5%、运行成本降低了18.8%,而在其他时段2个场景在经济性和低碳性方面相差不大。结合图11中的氢能平衡结果,在07:00—11:00及16:00—20:00,氢能平衡氢燃料电池消耗氢能产热

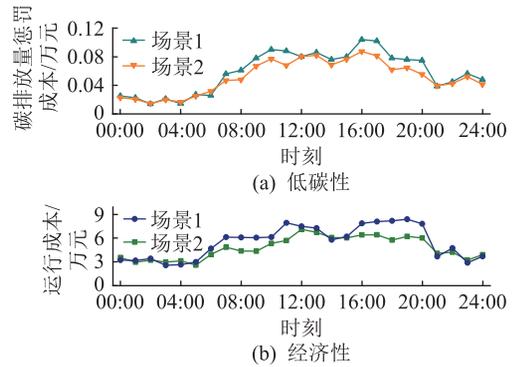


图13 不同场景经济性和低碳性对比

Fig.13 Comparison of economic viability and low carbon performance in different scenarios

产电,减轻了碳捕集电厂中火机组的压力,同时氢燃料电池较碳捕集电厂发电成本较低,因此引入氢燃料电池能显著提高多能流VPP优化调度结果的经济性和低碳性。

4.5 算法性能分析

为综合评价文中所提基于AM-SAC的VPP目标优化调度模型求解效果,将SAC算法、双重深度Q网络(double deep Q-network, Double DQN)算法、遗传算法(genetic algorithm, GA)与文中算法进行对比,对比结果如表3所示。GA通过模拟生物进化过程中的自然选择、基因重组和变异等机制,对解空间中的个体进行群体搜索和优化,然而容易陷入局部最优的困境,同时每次求解需要重复迭代,因此决策时间较长。Double DQN算法训练过程收敛性能较差,不适合求解状态和动作空间维度较高的问题,因此求解得到的运行成本与碳排放量惩罚成本处于较高水平。SAC算法能够收敛至最优策略,但与文中算法相比决策时间较长,这是因为SAC算法在150步左右陷入局部最优,直至1500步左右才跳出局部最优。而文中算法融合了AM,以训练时间为代价能够快速收敛至最优策略,具有显著的优势。

表3 AM-SAC算法与其他算法对比
Table 3 Comparison of AM-SAC algorithm with other algorithms

算法	运行成本/ 万元	碳排放量惩罚 成本/万元	训练时间/ min	决策 时间/s
GA	129.5	1.37		63.4
Double-DQN	139.2	1.53	133.5	21.9
SAC	123.2	1.24	145.7	23.7
AM-SAC	119.8	1.27	159.3	10.6

5 结语

文中提出一种双碳目标下基于深度强化学习

的多能流 VPP 优化调度方法。首先,根据 VPP 的随机碳流特性,建立一种面向动态优化参数的改进阶梯型碳交易机制。接着,以经济效益和碳排放量作为目标函数建立含氢 VPP 多能流解耦模型。然后,考虑到该模型具有高维非线性与权重参数实时更新的特征,利用 AM-SAC 算法在连续动作空间进行求解。最后,对多能流调度结果进行仿真分析和对比实验,结果验证了文中方法的可行性。主要结论如下:

(1) 面向动态参数建立基于贝叶斯优化的改进阶梯型碳交易机制,通过对碳交易基价、区间长度等重要模型参数进行动态实时优化,能有效提升 SAC 算法的收敛效果以及平均奖励值。同时,引入氢燃料电池可减少火电机组的使用,氢燃料电池发电成本较低,因此能显著提高多能流 VPP 优化调度结果的经济性和低碳性。

(2) 考虑到文中所提模型具有高维非线性与权重参数实时更新的特征,在 SAC 算法的基础上融入 AM 捕捉状态空间中的重要信息,帮助策略网络作出最优的动作预测,相较于传统 SAC 算法能够给出更准确的优化调度结果。

在后续研究中,将进一步考虑多个多能流 VPP 的协同优化调度问题,并对深度强化学习算法进行在线训练,提高深度强化学习算法的在线收敛能力,实现经济且低碳的 VPP 在线调度。

参考文献:

- [1] 王永真,潘崇超,韩恺,等. “双碳”目标下中国综合能源服务的态势辨析与理论架构[J]. 全球能源互联网,2023,6(4): 379-388.
WANG Yongzhen, PAN Chongchao, HAN Kai, et al. Situation identification and theory architecture of integrated energy services under dual carbon goal in china[J]. Journal of Global Energy Interconnection, 2023, 6(4): 379-388.
- [2] 文福拴,鲁刚,黄杰. 面向碳达峰、碳中和的综合能源系统[J]. 全球能源互联网,2022,5(2): 116-117.
WEN Fushuan, LU Gang, HUANG Jie. Integrated energy system towards carbon peak and neutrality targets[J]. Journal of Global Energy Interconnection, 2022, 5(2): 116-117.
- [3] 李贵君,赵兴勇,刘昊炆,等. 考虑梯级碳交易机制和电转气两阶段运行的综合能源系统优化[J]. 全球能源互联网, 2023,6(3): 331-340.
LI Guijun, ZHAO Xingyong, LIU Haoyang, et al. Integrated energy system optimization considering cascading carbon trading mechanism and two-stage operation of power-to-gas[J]. Journal of Global Energy Interconnection, 2023, 6(3): 331-340.
- [4] 王健,郑峻峰,黄际元,等. 虚拟电厂关键技术综述与发展展望[J]. 供用电,2023,40(12): 43-54, 86.
WANG Jian, ZHENG Junfeng, HUANG Jiyuan, et al. Key technology and development prospects of virtual power plants[J]. Distribution & Utilization, 2023, 40(12): 43-54, 86.
- [5] 高赐威. 虚拟电厂的概念思辨[J]. 供用电,2023,40(12): 2-6.
GAO Ciwei. Conceptual reflection on virtual power plants[J]. Distribution & Utilization, 2023, 40(12): 2-6.
- [6] 李彬,顾程镓,杜佳伦,等. 考虑电热转换的虚拟电厂资源调度优化策略[J]. 供用电,2023,40(12): 16-21.
LI Bin, GU Chengjia, DU Jialun, et al. Virtual power plant resource scheduling optimization strategy considering electrothermal conversion[J]. Distribution & Utilization, 2023, 40(12): 16-21.
- [7] 刘源,檀勤良,张兴平. 基于交互算法的多代理虚拟电厂调度优化及风险分析[J]. 电力工程技术,2022,41(6): 2-12.
LIU Yuan, TAN Qinliang, ZHANG Xingping. Multi-agent VPP coordinated control optimization and risk analysis based on the interactive algorithm[J]. Electric Power Engineering Technology, 2022, 41(6): 2-12.
- [8] 梁露,李佳承,韩长江,等. 运行风险与资源随机波动影响下的虚拟电厂优化调度方法[J]. 供用电,2023,40(12): 22-31.
LIANG Lu, LI Jiacheng, HAN Changjiang, et al. Optimal scheduling method of virtual power plant under the influence of operational risk and random fluctuation of resources[J]. Distribution & Utilization, 2023, 40(12): 22-31.
- [9] 郝建红,黄婷,许秋铭,等. 含多供能商的虚拟电厂多目标优化调度策略[J]. 供用电,2023,40(12): 32-42.
HAO Jianhong, HUANG Ting, XU Qiuming, et al. Virtual power plant multi-objective optimal scheduling strategy with multiple energy suppliers[J]. Distribution & Utilization, 2023, 40(12): 32-42.
- [10] 李健强,余光正,汤波,等. 考虑风光利用率和含氢能流的多能流综合能源系统规划[J]. 电力系统保护与控制, 2021,49(14): 11-20.
LI Jianqiang, YU Guangzheng, TANG Bo, et al. Multi-energy flow integrated energy system planning considering wind and solar utilization and containing hydrogen energy flow [J]. Power System Protection and Control, 2021, 49(14): 11-20.
- [11] 崔丽瑶,刘怀东,刘豪,等. 基于氢能经济的电网大规模风电消纳模式[J]. 电力系统及其自动化学报,2022,34(2): 108-115.
CUI Liyao, LIU Huaidong, LIU Hao, et al. Large-scale wind power accommodation mode of power grid based on hydrogen energy economy[J]. Proceedings of the CSU-EPSC, 2022, 34(2): 108-115.
- [12] 邓杰,姜飞,王文焯,等. 考虑电热柔性负荷与氢能精细化建模的综合能源系统低碳运行[J]. 电网技术,2022,46(5): 1692-1704.
DENG Jie, JIANG Fei, WANG Wenyue, et al. Low-carbon optimized operation of integrated energy system considering electric-heat flexible load and hydrogen energy refined modeling [J]. Power System Technology, 2022, 46(5): 1692-1704.

- [13] 袁铁江,高玲玉,谢永胜,等. 基于氢能的风-火耦合多能系统设计及综合评估[J]. 电力自动化设备,2021,41(10):227-233,255.
YUAN Tiejiang, GAO Lingyu, XIE Yongsheng, et al. Design and comprehensive evaluation of wind-thermal power coupling multi-energy system based on hydrogen energy[J]. Electric Power Automation Equipment,2021,41(10):227-233,255.
- [14] 李东东,王啸林,沈运帷,等. 考虑多重不确定性的含需求响应及电碳交易的虚拟电厂优化调度策略[J]. 电力自动化设备,2023,43(5):210-217,251.
LI Dongdong, WANG Xiaolin, SHEN Yunwei, et al. Optimal scheduling strategy of virtual power plant with demand response and electricity-carbon trading considering multiple uncertainties[J]. Electric Power Automation Equipment,2023,43(5):210-217,251.
- [15] 潘明杰,解大,王西田. 计及分布式资源调节特性差异的虚拟电厂响应策略[J]. 电力系统自动化,2022,46(18):108-117.
PAN Mingjie, XIE Da, WANG Xitian. Response strategy for virtual power plant considering difference in regulation characteristics of distributed energy resources[J]. Automation of Electric Power Systems,2022,46(18):108-117.
- [16] 范帅,何光宇,郑湘明,等. 基于在线分布式优化的虚拟电厂自趋优运行方法研究[J]. 中国电机工程学报,2023,43(13):4935-4950.
FAN Shuai, HE Guangyu, ZHENG Xiangming, et al. Research on online distributed optimization-based self-approaching optimization operation method of virtual power plant[J]. Proceedings of the CSEE,2023,43(13):4935-4950.
- [17] 周欢,王芬,李志勇,等. 虚拟电厂自趋优负荷跟踪控制策略[J]. 中国电机工程学报,2021,41(24):8334-8349.
ZHOU Huan, WANG Fen, LI Zhiyong, et al. Load tracking control strategy for virtual power plant via self-approaching optimization[J]. Proceedings of the CSEE,2021,41(24):8334-8349.
- [18] 谢宏伟,严强,李扬,等. 市场模式下兼顾区域负荷特性的多虚拟电厂分布式协调优化[J]. 电力自动化设备,2023,43(5):199-209.
XIE Hongwei, YAN Qiang, LI Yang, et al. Distributed coordination optimization of multiple virtual power plants considering regional load characteristics in market[J]. Electric Power Automation Equipment,2023,43(5):199-209.
- [19] 元志伟,于松源,房方,等. 考虑高维不确定性的热电联产虚拟电厂优化调度[J]. 动力工程学报,2023,43(2):194-204.
YUAN Zhiwei, YU Songyuan, FANG Fang, et al. Optimal scheduling of combined heat and power-virtual power plant considering high-dimensional uncertainty[J]. Journal of Chinese Society of Power Engineering,2023,43(2):194-204.
- [20] 邱纯,应展烽,冯奕,等. 计及碳配额的混合储能综合微电网优化运行研究[J]. 电力工程技术,2022,41(2):119-127.
QIU Chun, YING Zhanfeng, FENG Yi, et al. Optimal operation of hybrid energy storage integrated micro-energy network considering carbon quote[J]. Electric Power Engineering Technology,2022,41(2):119-127.
- [21] 赵心怡,谢俊,周翠玉,等. 风-光-抽蓄零碳电力系统多时间尺度协调调度模型[J]. 电力工程技术,2023,42(3):121-129.
ZHAO Xinyi, XIE Jun, ZHOU Cuiyu, et al. A multi-time scale coordinated dispatching model of wind-photovoltaic-pumped storage zero-carbon power system[J]. Electric Power Engineering Technology,2023,42(3):121-129.
- [22] 唐巍,张瑜,宣东海,等. 计及阶梯式碳交易机制的低碳型园区经济运行调度策略[J]. 供用电,2021,38(9):10-18.
TANG Wei, ZHANG Yu, XUAN Donghai, et al. Low-carbon park economic operation scheduling strategy taking into account the tiered carbon trading mechanism[J]. Distribution & Utilization,2021,38(9):10-18.
- [23] 瞿凯平,黄琳妮,余涛,等. 碳交易机制下多区域综合能源系统的分散调度[J]. 中国电机工程学报,2018,38(3):697-707.
QU Kaiping, HUANG Linni, YU Tao, et al. Decentralized dispatch of multi-area integrated energy systems with carbon trading[J]. Proceedings of the CSEE,2018,38(3):697-707.
- [24] 骆钊,刘泓志,赵伟杰,等. 考虑碳-绿证联合交易的跨境综合能源系统协同运行优化[J]. 电力自动化设备,2023,43(11):1-8,25.
LUO Zhao, LIU Hongzhi, ZHAO Weijie, et al. Collaborative operation optimization of cross-border integrated energy system considering joint trading of carbon and green certificates[J]. Electric Power Automation Equipment,2023,43(11):1-8,25.
- [25] 孙惠娟,刘昀,彭春华,等. 计及电转气协同的含碳捕集与垃圾焚烧虚拟电厂优化调度[C]//2021年江西省电机工程学会年会论文集. 南昌,2022:88-95.
SUN Huijuan, LIU Yun, PENG Chunhua, et al. The optimal operation of the virtual power plant including carbon capture and waste incineration with co-operation of electric and rotary gas[C]//Proceedings of the Annual Meeting of Jiangxi Electrical Engineering Society in 2021. Nanchang,2022:88-95.
- [26] 彭姿余,王高才,农望. 基于SAC的多服务移动边缘计算中任务卸载和资源配置算法[J]. 计算机应用研究,2023,40(5):1497-1503.
PENG Ziyu, WANG Gaocai, NONG Wang. SAC based algorithm for task offloading and resource provisioning in multiple-services mobile edge computing[J]. Application Research of Computers,2023,40(5):1497-1503.
- [27] 罗文健,张靖,何宇,等. 基于优势柔性策略-评价算法和迁移学习的区域综合能源系统优化调度[J]. 电网技术,2023,47(4):1601-1615.
LUO Wenjian, ZHANG Jing, HE Yu, et al. Optimal scheduling of regional integrated energy system based on advantage

learning soft actor-critic algorithm and transfer learning [J]. Power System Technology, 2023, 47(4):1601-1615.

[28] 李曾琳,李波,白双霞,等. 基于 AM-SAC 的无人机自主空战决策[J]. 兵工学报, 2023, 44(9):2849-2858.

LI Zenglin, LI Bo, BAI Shuangxia, et al. UAV autonomous air combat decision-making based on AM-SAC [J]. Acta Armamentarii, 2023, 44(9):2849-2858.

[29] 史建勋. 中国天然气市场价格风险分析与防范建议[J]. 国际石油经济, 2022, 30(5):19-24.

SHI Jianxun. Price risk analysis and prevention suggestions of natural gas market in China [J]. International Petroleum Eco-

nomics, 2022, 30(5):19-24.

作者简介:



俞晓荣

俞晓荣(1979),女,硕士,高级工程师,从事配电网优化调度相关工作(E-mail:yuxiaorong_JS@163.com);

徐青山(1979),男,博士,教授,研究方向为智能配电系统与微网;

杜璞良(1992),男,博士,讲师,研究方向为虚拟电厂运行优化,城市电网低碳发展关键技术。

Optimizing multi-energy flow scheduling of hydrogen-inclusive virtual power plants based on deep reinforcement learning under dual-carbon targets

YU Xiaorong¹, XU Qingshan², DU Puliang³, WANG Dong²

(1. State Grid Taizhou Power Supply Company of Jiangsu Electric Power Co., Ltd., Taizhou 225300, China;

2. School of Electrical Engineering, Southeast University, Nanjing 210096, China;

3. School of Economics and Management, Shanghai University of Electric Power, Shanghai 200090, China)

Abstract: Virtual power plants, as a comprehensive energy network with multi-energy flow interconnection, have become an important player in China's accelerated pursuit of its dual carbon goals. However, it is difficult to coordinate internal resources with low-carbon emission when facing challenges such as tight coupling of multi-energy flows, subjectivity of traditional carbon trading model parameters and difficulty of online optimization with high-dimensional dynamic parameters. To address these issues, this paper proposes a virtual power plant multi-energy low-carbon dispatching method that integrates the attention mechanism (AM) and soft actor-critic (SAC) algorithm. Firstly, based on the random carbon flow characteristics of virtual power plants, an improved stepped carbon trading mechanism based on Bayesian optimization is established for dynamic parameters. Next, an economic benefit and carbon emission-based objective function is constructed for the decoupling model of multi-energy flows in virtual power plants. Considering the high-dimensional nonlinearity and real-time updating of weight parameters in this model, the improved SAC deep reinforcement learning algorithm with integrated attention mechanism is used to solve it in a continuous action space. Finally, simulation analysis and comparative experiments are conducted to verify the feasibility of the proposed method and its efficiency compared with SAC algorithm.

Keywords: virtual power plant (VPP); multi-energy flow; improved carbon trading mechanism; deep reinforcement learning; attention mechanism (AM); soft actor-critic (SAC) algorithm

(编辑 钱悦 方晶)