

引用格式:张春雁,窦真兰,张继行,等.电-热-气耦合虚拟电厂参与主辅联合市场的博弈竞价策略[J].电力科学与技术学报,2026,41(2):156-166.

Citation: ZHANG Chunyan, DOU Zhenlan, ZHANG Jihang, et al. Game bidding strategy for electric thermal gas coupled virtual power plants participating in main and auxiliary joint markets[J]. Journal of Electric Power Science and Technology, 2026, 41(2): 156-166.

电-热-气耦合虚拟电厂参与主辅联合市场的 博弈竞价策略

张春雁¹, 窦真兰¹, 张继行², 王玲玲², 蒋传文²

(1. 国网上海市电力公司, 上海 200023; 2. 上海交通大学电气工程学院, 上海 200240)

摘要:随着综合能源系统的发展,多能耦合虚拟电厂成为管理综合能源系统的有效途径,但由于其内部能源种类众多,耦合关系复杂,参与市场时无法获取竞争对手信息,采用传统数学方法求解其竞价策略困难。提出一种基于约束感知多代理强化学习的电-热-气耦合虚拟电厂参与主辅联合市场的博弈竞价策略。首先,搭建多能耦合虚拟电厂的数学模型,描述其内部电-热-气多能耦合的能量流动关系。其次,构建多能耦合虚拟电厂以价格制定者参与不完全信息主辅联合市场的博弈模型,并允许虚拟电厂之间进行电能量和辅助服务交易。最后,对约束感知强化学习进行改进,提出约束感知多代理强化学习算法进行模型求解,通过在算法中设置看门狗模块有效保证了虚拟电厂策略符合电网安全约束,并和理论最优解进行对比,验证所提方法的有效性。

关键词:多能耦合虚拟电厂;约束感知;多代理强化学习;主辅联合市场;博弈竞价策略

DOI: 10.19781/j.issn.1673-9140.2026.02.014 中图分类号: TM732 文章编号: 1673-9140(2026)02-0156-11

Game bidding strategy for electric thermal gas coupled virtual power plants participating in main and auxiliary joint markets

ZHANG Chunyan¹, DOU Zhenlan¹, ZHANG Jihang², WANG Lingling², JIANG Chuanwen²

(1. State Grid Shanghai Municipal Electric Power Company, Shanghai 200023, China; 2. School of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: As integrated energy systems develop, multi-energy coupled virtual power plants have become an effective way to manage integrated energy systems. However, due to their numerous internal energy types, complex coupling relationships, and inability to obtain competitor information when participating in the market, it is difficult to solve their bidding strategies by adopting traditional mathematical methods. A game bidding strategy for electric thermal gas coupled virtual power plants participating in main and auxiliary joint markets based on constraint-aware multi-agent reinforcement learning is proposed. Firstly, a mathematical model of a multi-energy coupled virtual power plant is built and the energy flow relationship of its internal electric thermal gas multi-energy coupling is described. Secondly, a game model of multi-energy coupled virtual power plants is built, in which price setters participate in the incomplete information main and auxiliary joint market, and virtual power plants are allowed to trade electricity and auxiliary services. Finally, constraint-aware reinforcement learning is improved, and the constraint-aware multi-agent reinforcement learning algorithm is proposed for solving the model. By setting a watchdog module in the algorithm, the virtual power plant strategy is effectively ensured to comply with power grid safety constraints, and compared with the theoretical optimal solution to verify the effectiveness of the proposed method.

Key words: multi-energy coupled virtual power plant; constraint-aware; multi-agent reinforcement learning; main and auxiliary joint markets; game bidding strategy

收稿日期: 2025-02-27; 修回日期: 2025-06-04

基金项目: 国家电网有限公司科技项目(5400-202317213A-1-1-ZN)

通信作者: 王玲玲(1993—), 女, 博士, 助理研究员, 主要从事电力系统优化运行、综合能源系统等方面的研究; E-mail: wanglingling1993@sjtu.edu.cn

虚拟电厂作为一种资源聚合技术方案,可对多种能源进行协同互补调控,以较低成本挖掘资源潜力。而综合能源系统虚拟电厂除涵盖电力资源外,还可对气、热网系统进行统筹优化,实现多种能源间的耦合互补运行,提高综合能源系统效率并减少能源浪费。除此之外,新电改明确提出,须促进调频、备用等辅助服务市场与电力现货市场的有效衔接,并推动两者联合出清;同时,应促进新型储能、分布式资源以及综合能源系统、虚拟电厂等新型市场主体参与主辅联合市场。新型市场主体的加入提升了市场参与积极性,还可为新型电力系统提供灵活性支撑。在此背景下,开展电-热-气耦合虚拟电厂优化调度及市场交易策略研究具有重要的理论与工程意义。

目前,关于多能耦合虚拟电厂的研究已较为丰富,文献[1]考虑虚拟电厂内部包含电、气、热能源网络,通过统一控制平台实施多能互补优化调度,有效促进可再生能源消纳并降低碳排放量。文献[2]考虑多能耦合虚拟电厂中多种资源的不确定性,基于鲁棒优化方法开展优化调度运行研究,有效降低了不确定性带来的虚拟电厂运行风险,提升了其运行经济性与效率。文献[3]考虑虚拟电厂聚合碳捕集装置、垃圾焚烧等资源,并对其实施优化调度。文献[4]考虑虚拟电厂内部电、热、冷多种能源的互补耦合特性,结合储能设备的时域特性对其实施优化调度,所得调度结果充分考虑了多能协同运行特性,提高了虚拟电厂的运行效率与经济效益。文献[5]构建了包含电气耦合设备、负荷响应等资源的虚拟电厂多维度优化调度运行模型,其运行结果的效率更优。文献[6]考虑虚拟电厂聚合气-电耦合设备形成多能耦合虚拟电厂,重点研究了优化调度运行策略并开展风险评估。文献[7]考虑电力系统与热力系统耦合形成虚拟电厂,综合考虑各主体利益实施互补协同调度,其运行结果减少了能源浪费并提升了效益。文献[8]构建了包含多种能源类型及多主体的虚拟电厂运营框架与通信方案,并针对多能源间的互补耦合特性实施优化调度。

虚拟电厂参与电力市场的研究也较为丰富。文献[9]考虑市场环境下含分布式发电、电动汽车及储能设备的虚拟电厂多目标协同优化运行问题,并采用混合整数线性规划方法求解。文献[10]提出了虚拟电厂内部灵活性资源参与辅助服务市场的契合度指标,基于该指标构建虚拟电厂参与主辅

联合市场的博弈竞价模型,进而求解其最优竞价策略。文献[11]考虑虚拟电厂内部资源的不确定性,开展虚拟电厂参与主辅联合市场联合出清研究,有效提升了虚拟电厂的总体收益并降低了运行风险。然而,上述文献的研究对象均为单个虚拟电厂,未考虑多个市场主体间策略性竞价行为的相互影响。文献[12]针对含多个综合能源虚拟电厂的主动配电网,提出了双层博弈优化调度策略,其中上层模型描述主动配电网的出清行为,下层模型描述多个综合能源虚拟电厂的博弈竞价行为,并通过迭代方法求解博弈均衡解。文献[13]在主辅联合市场环境下聚合电、热、气资源,形成多区域综合能源虚拟电厂,结合多区域间的多能互补特性实施优化调度,其调度结果可有效降低系统总成本。文献[14]采用了对角化算法求解多个虚拟电厂参与主辅联合市场的策略性投标博弈均衡问题。文献[15]考虑日前、实时两阶段预测精度的影响,构建了多时间尺度虚拟电厂交互运行调度模型,并采用滚动优化方法求解,显著提升了虚拟电厂的效益。文献[16]提出了含楼宇冷负荷的综合能源虚拟电厂参与调频辅助服务的博弈竞价策略,在保证楼宇温度舒适度的同时,有效提高了虚拟电厂的收益。文献[17]将虚拟电厂视为价格制定者,构建了虚拟电厂参与电力市场的双层优化模型,所得市场均衡解提升了电力系统的整体效益。

这些文献均假设多能耦合虚拟电厂处于完全信息市场环境,即每个多能耦合虚拟电厂的控制器除已知其控制范围内的资源信息外,还掌握其他市场成员的信息,该假设与实际市场情况不符。针对不完全信息市场环境下虚拟电厂参与市场优化调度的问题,多代理人工智能方法更为有效,其多代理特性可对其他市场成员的竞价模式进行估计,并模拟多个市场成员间的博弈均衡过程。文献[18]提出点对点能源交易模式,通过多代理强化学习方法求解,得到了多个市场主体的报价策略与调度计划。文献[19]在不完全信息市场环境下,模拟发电集团的博弈竞价行为,最终实现市场均衡并提升了市场整体收益。在模型求解方面,传统数学方法采用卡罗需-库恩-塔克(Karush-Kuhn-Tucker, KKT)条件进行模型转化,但该过程要求下层模型为非凸模型;在求解多能耦合虚拟电厂优化调度问题时,模型中存在大量时间耦合约束及非线性、离散决策变量,导致传统方法求解难度较大。另一方面,传统多代理强化学习算法在训练过程中,为满足约束

条件,通常在目标函数中设置超约束惩罚项;该措施并非强制约束,在实际执行过程中仍会出现超约束现象,导致调度结果不满足电网安全约束,引发严重后果。

针对这些问题,本文提出一种约束感知多代理深度确定性策略梯度(mixed-integer programming multi-agent deep deterministic policy gradient, MIP-MADDPG)方法,求解电-热-气耦合虚拟电厂参与主辅联合市场的双层博弈竞价策略,本文的主要贡献如下:

1) 搭建电-热-气耦合虚拟电厂数学模型,描述其内部多种能源之间的耦合关系。

2) 提出多个虚拟电厂参与主辅联合市场双层博弈模型,市场模型中每个市场成员均作为价格影响者以平等的地位参与市场,每个市场成员均按照自身的报价策略进行量-价信息申报,为更加贴近真实市场情况设置了不完全信息市场环境,即竞争对手的报价决策信息是未知的,将不完全信息市场这一具体情况抽象为MADDPG算法中每个智能体无法获悉其余智能体动作选择网络中的参数,通过近似的方式估计对手的动作选择策略,较好地模拟现实市场环境。

3) 对约束感知强化学习进行改进,提出MIP-MADDPG算法进行模型求解,将多虚拟电厂的博弈竞价行为映射到算法动作值,有效保证了虚拟电厂策略符合电网安全约束,通过和理论最优解进行对比证明了方法有效性。

1 多能耦合虚拟电厂参与主辅联合市场交易模型

本节介绍多能耦合虚拟电厂参与主辅联合市场交易模型,包括多能耦合虚拟电厂结构模型和市场双层博弈模型,考虑多个虚拟电厂之间、虚拟电厂和电力市场之间均可进行交易且虚拟电厂作为价格制定者进行报价决策。交易结构示意图如图1所示。

多能耦合虚拟电厂涉及电、热、气3种类型能源,包括风电、光伏、电储能、热储能、燃气锅炉、热电联产等机组。在不完全信息市场中,每个市场主体均无法获取其余市场主体的市场信息。各虚拟电厂根据自身资源特性,优化能源价格、交易电量与系统内机组出力,以满足自身负荷和备用需求。

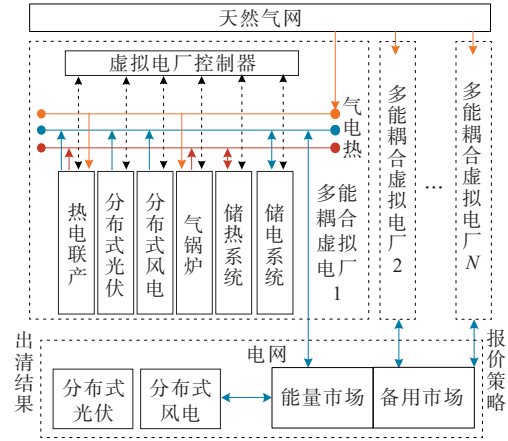


图1 多能耦合虚拟电厂参与主辅联合市场交易模型

Figure 1 Trading model of multi-energy coupled virtual power plants participating in main and auxiliary joint markets

1.1 多能耦合虚拟电厂决策模型

1.1.1 目标函数

多能耦合虚拟电厂根据上层电网制定的电能量与备用价格进行内部优化,并将电能与备用容量上报给电网。在日前优化过程中,以1h为调度颗粒度,目标函数为多能耦合虚拟电厂经济效益最大化,表达式如下:

$$\max C_i^c = \sum_{t=1}^{24} \pi_w \sum_{w=1}^{n_w} [C_{MEVPP,i,t,w} + C_{i,t,w}^{\text{trad}} - C_{i,t,w}^{\text{gas}} - C_{ES,i,t,w} - C_{TS,i,t,w} - \sum_k (C_{k,t,w}^{\text{up}} R_{k,t,w}^{\text{up}} + C_{k,t,w}^{\text{down}} R_{k,t,w}^{\text{down}})] \quad (1)$$

式中, π_w 为场景 w 发生的概率; $C_{MEVPP,i,t,w}$ 为在 t 时刻场景 w 下多能耦合虚拟电厂 i 与电网购售电、备用收益; $C_{i,t,w}^{\text{trad}}$ 为在 t 时刻场景 w 下多能耦合虚拟电厂 i 之间购售电、备用收益; $C_{i,t,w}^{\text{gas}}$ 为在 t 时刻场景 w 下多能耦合虚拟电厂 i 采购天然气的费用; $C_{ES,i,t,w}$ 、 $C_{TS,i,t,w}$ 分别为在 t 时刻场景 w 下多能耦合虚拟电厂 i 的电储能与热储能成本; $R_{k,t,w}^{\text{up}}$ 、 $R_{k,t,w}^{\text{down}}$ 分别为在 t 时刻场景 w 下第 k 类备用资源提供的上备用与下备用容量; $C_{k,t,w}^{\text{up}}$ 、 $C_{k,t,w}^{\text{down}}$ 分别为在 t 时刻场景 w 下第 k 类的上备用与下备用价格。

目标函数中各变量表达式如下:

$$C_{i,t,w}^{\text{trad}} = p_{ij,t,w}^e q_{ij,t,w}^e + p_{ij,t,w}^{r,\text{up}} q_{ij,t,w}^{r,\text{up}} + p_{ij,t,w}^{r,\text{down}} q_{ij,t,w}^{r,\text{down}} \quad (2)$$

$$\begin{cases} C_{ES,i,t,w} = C_E P_{ES,i,t,w} \\ C_{TS,i,t,w} = C_T P_{TS,i,t,w}^h \end{cases} \quad (3)$$

$$\sum_k (C_{k,t,w}^{\text{up}} R_{k,t,w}^{\text{up}} + C_{k,t,w}^{\text{down}} R_{k,t,w}^{\text{down}}) = C_{CHP,i,t,w}^{\text{up}} R_{CHP,i,t,w}^{\text{up}} + C_{CHP,i,t,w}^{\text{down}} R_{CHP,i,t,w}^{\text{down}} + C_{ES,i,t,w}^{\text{up}} R_{ES,i,t,w}^{\text{up}} + C_{ES,i,t,w}^{\text{down}} R_{ES,i,t,w}^{\text{down}} \quad (4)$$

式中, $p_{ij,t,w}^e, p_{ij,t,w}^{r,up}, p_{ij,t,w}^{r,down}$ 分别为在 t 时刻场景 w 下多能耦合虚拟电厂 i 和 j 之间交易的电价、上备用价格、下备用价格; $q_{ij,t,w}^e, q_{ij,t,w}^{r,up}, q_{ij,t,w}^{r,down}$ 分别为在 t 时刻场景 w 下多能耦合虚拟电厂 i 和 j 之间交易的电能、上备用容量、下备用容量; C_E, C_T 分别为电储能、热储能充放电成本系数; $P_{ES,i,t,w}, P_{TS,i,t,w}^h$ 分别为在 t 时刻场景 w 下多能耦合虚拟电厂 i 的电储能、热储能充放电功率; $C_{CHP,i,t,w}^{up}, C_{CHP,i,t,w}^{down}, C_{ES,i,t,w}^{up}, C_{ES,i,t,w}^{down}$ 分别为在 t 时刻场景 w 下多能耦合虚拟电厂 i 中热电联产和电储能提供的上、下备用容量成本系数; $R_{CHP,i,t,w}^{up}, R_{CHP,i,t,w}^{down}, R_{ES,i,t,w}^{up}, R_{ES,i,t,w}^{down}$ 分别为在 t 时刻场景 w 下多能耦合虚拟电厂 i 中热电联产和电储能机组提供的上、下备用容量。

1.1.2 约束条件

1) 功率平衡约束。

$$\begin{cases} P_{i,t,w}^{buy} + P_{new,i,t,w} + P_{CHP,i,t,w} + P_{ES,i,t,w}^{out} = \\ P_{i,t,w}^{sell} + P_{ES,i,t,w}^{in} + g_{load,i,t} + \sum_{j \neq i}^M q_{ij,t,w} \end{cases} \quad (5)$$

$$\begin{cases} P_{CHP,i,t,w}^h + P_{GB,i,t,w}^h + P_{TS,i,t,w}^{h,out} = \\ P_{TS,i,t,w}^{h,in} + g_{load,i,t}^h \end{cases} \quad (6)$$

$$g_{ub,i,t,w}^g = P_{CHP,i,t,w}^g + P_{GB,i,t,w}^g + P_{MT,i,t,w}^g + g_{load,i,t}^g \quad (7)$$

式中, $P_{i,t,w}^{buy}, P_{i,t,w}^{sell}$ 分别为多能耦合虚拟电厂 i 与电网之间的购售电量; $g_{load,i,t}, g_{load,i,t}^h, g_{load,i,t}^g$ 分别为虚拟电厂内部电负荷、热负荷和气负荷; $P_{CHP,i,t,w}^h, P_{GB,i,t,w}^h$ 分别为虚拟电厂内部热电联产机组和电锅炉机组产热功率; $P_{TS,i,t,w}^{h,out}, P_{TS,i,t,w}^{h,in}$ 分别为虚拟电厂内部热储能的放热功率和储热功率。

2) 储能设备运行约束。

$$0 \leq P_{CH,t,w} + P_{CH,t,w}^{down} \leq P_{CH}^{max} I_{s,t,w} \quad (8)$$

$$0 \leq P_{DC,t,w} + P_{DC,t,w}^{up} \leq P_{DC}^{max} I_{r,t,w} \quad (9)$$

$$I_{s,t,w} + I_{r,t,w} \leq 1 \quad (10)$$

$$S_{e,t,w} = S_{e,t-1,w} + [(P_{CH,t,w} + P_{CH,t,w}^{down}) \cdot \eta_e^{CH} - (P_{DC,t,w} + P_{DC,t,w}^{up}) / \eta_e^{DC}] \quad (11)$$

$$S_e^{min} \leq S_{e,t,w} \leq S_e^{max} \quad (12)$$

$$S_{e,0,w} = S_{e,T,w} \quad (13)$$

式中, $P_{CH}^{max}, P_{DC}^{max}$ 分别为电储能最大充电和放电功率; $I_{s,t,w}, I_{r,t,w}$ 分别为在 t 时刻场景 w 下电储能充、放电 0-1 状态变量; $S_{e,t,w}$ 为在 t 时刻场景 w 下的电储能容量; η_e^{CH}, η_e^{DC} 分别为充、放电效率; S_e^{max}, S_e^{min} 分别为电储能运行时允许储存电量的最大值与最小值; $S_e(0), S_e(T)$ 分别为调度周期初始与结束时的荷电状态。由于热储能与电储能约束条件一致, 本文不再赘述。

3) CHP 机组运行约束。

$$P_{CHP,i}^{min} I_{CHP,i,t,w} \leq P_{CHP,i,t,w} \leq P_{CHP,i}^{max} I_{CHP,i,t,w} \quad (14)$$

$$P_{CHP,i,t,w} - P_{CHP,i,t-1,w} \leq \quad (15)$$

$$r_{CHP,i}^{up} (1 - u_{CHP,i,t,w}) + r_{CHP,i}^{down} u_{CHP,i,t,w}$$

$$P_{CHP,i,t-1,w} - P_{CHP,i,t,w} \leq \quad (16)$$

$$r_{CHP,i}^{down} (1 - v_{CHP,i,t,w}) + r_{CHP,i}^{down} v_{CHP,i,t,w}$$

式中, $P_{CHP,i}^{min}, P_{CHP,i}^{max}$ 分别为多能耦合虚拟电厂 i 中热电联产机组最小和最大输出功率; $r_{CHP,i}^{down}, r_{CHP,i}^{up}$ 分别为热电联产机组爬坡率上、下限值。

4) 燃气锅炉运行约束。

$$Q_{i,t,w}^{GF} = \eta_i^{GF} V_{i,t,w}^{GF} \quad (17)$$

$$H_{GF,i}^{min} \leq H_{GF,i,t,w} \leq H_{GF,i}^{max} \quad (18)$$

式中, η_i^{GF} 为燃气锅炉热功率转换系数; $V_{i,t,w}^{GF}$ 为 t 时刻场景 w 下多能耦合虚拟电厂 i 中燃气锅炉输入的天然气功率; $H_{GF,i}^{min}, H_{GF,i}^{max}$ 分别为多能耦合虚拟电厂 i 中燃气锅炉最小和最大热功率输出值。

5) 多能耦合虚拟电厂交易功率约束。

$$\sum_{i=1}^m P_{e,i,t,w}^{out} = \sum_{i=1}^m P_{e,i,t,w}^{in} \quad (19)$$

$$\sum_{i=1}^m P_{r,i,t,w}^{up,out} = \sum_{i=1}^m P_{r,i,t,w}^{up,in} \quad (20)$$

$$\sum_{i=1}^m P_{r,i,t,w}^{down,out} = \sum_{i=1}^m P_{r,i,t,w}^{down,in} \quad (21)$$

$$P_{e,i,t,w}^{out} \geq 0 \quad (22)$$

$$P_{r,i,t,w}^{up,out} \geq 0 \quad (23)$$

$$P_{r,i,t,w}^{down,out} \geq 0 \quad (24)$$

$$q_{ij,t,w}^e + q_{ij,t,w}^{r,up} \geq q_{ij}^{max} \quad (25)$$

$$q_{ij,t,w}^e - q_{ij,t,w}^{r,down} \geq 0 \quad (26)$$

式中, q_{ij}^{max} 为多能耦合虚拟电厂 i 与 j 之间最大传输量; $q_{ij,t,w}^e, q_{ij,t,w}^{r,up}, q_{ij,t,w}^{r,down}$ 分别为 t 时刻场景 w 下多能耦合虚拟电厂 i 与 j 之间交易的电能量和上备用、下备用容量。

1.2 主辅联合市场博弈出清模型

1.2.1 目标函数

在主辅联合市场博弈出清模型中, 电网以总成本最小为目标, 包括向上级电网购电、备用成本, 与多能耦合虚拟电厂交易成本, 目标函数为

$$\min C_{ADN} = \sum_{t=1}^{24} \pi_w \sum_{w=1}^{n_w} (C_{IESP,t,w} + C_{UG,t,w}) \quad (27)$$

$$C_{IESP,t,w} = u_{e,t,w}^b P_{e,t,w}^{b,IESP} + u_{r,t,w}^{b,up} P_{r,t,w}^{b,IESP,up} + u_{r,t,w}^{b,down} P_{r,t,w}^{b,IESP,down} - u_{e,t,w}^s P_{e,t,w}^s - \quad (28)$$

$$u_{r,t,w}^{s,up} P_{r,t,w}^{s,IESP,up} - u_{r,t,w}^{s,down} P_{r,t,w}^{s,IESP,down}$$

$$C_{UG,t,w} = \lambda_{e,t}^s P_{e,t,w}^{b,UG} + \quad (29)$$

$$\lambda_{r,t}^{s,up} P_{r,t,w}^{b,UG,up} + \lambda_{r,t}^{s,down} P_{r,t,w}^{b,UG,down}$$

式中, $u_{e,t,w}^b, u_{r,t,w}^{b,up}, u_{r,t,w}^{b,down}$ 分别为场景 w 下电网制定的

购电、上备用、下备用价格； $u_{e,t,w}^s, u_{r,t,w}^{s,up}, u_{r,t,w}^{s,down}$ 分别为场景 w 下电网制定的售电、上备用、下备用价格； $P_{e,t,w}^{b,MEVPP}, P_{r,t,w}^{b,MEVPP,up}, P_{r,t,w}^{b,MEVPP,down}$ 分别为场景 w 下电网购电、上备用、下备用容量； $P_{e,t,w}^{s,MEVPP}, P_{r,t,w}^{s,MEVPP,up}, P_{r,t,w}^{s,MEVPP,down}$ 分别为场景 w 下电网售电、上备用、下备用容量； $\lambda_{e,t}^s, \lambda_{r,t}^{s,up}, \lambda_{r,t}^{s,down}$ 分别为场景 w 下向上级电网购电、上备用、下备用价格； $P_{e,t,w}^{s,UG}, P_{r,t,w}^{s,UG,up}, P_{r,t,w}^{s,UG,down}$ 分别为场景 w 下向上级电网购电、上备用、下备用容量。

1.2.2 约束条件

1) 功率平衡约束。

$$P_t^L + P_{t,w}^{s,IESP} = P_{t,w}^{b,IESP} + P_{t,w}^{b,UG} + P_t^{new} \quad (30)$$

2) 网络潮流约束。

$$\begin{cases} P_{G,i,t,w} - P_{L,i,t,w} = \\ U_{i,t,w} \sum_{j \in i} U_{j,t,w} (G_{ij} \cos \theta_{ij,t,w} + B_{ij} \sin \theta_{ij,t,w}) \\ Q_{G,i,t,w} - Q_{L,i,t,w} = \\ U_{i,t,w} \sum_{j \in i} U_{j,t,w} (G_{ij} \sin \theta_{ij,t,w} - B_{ij} \cos \theta_{ij,t,w}) \\ i = 1, 2, \dots, n \end{cases} \quad (31)$$

式中, n 为节点数； $P_{G,i,t,w}, Q_{G,i,t,w}$ 分别为节点 i 在 t 时段输入的有功、无功功率； $P_{L,i,t,w}, Q_{L,i,t,w}$ 分别为节点 i 在 t 时段所带负荷的有功、无功功率； $U_{i,t,w}, U_{j,t,w}$ 分别为节点 i, j 在 t 时段的节点电压； G_{ij}, B_{ij} 分别为支路 ij 的电导、电纳； $\theta_{ij,t,w}$ 为 i, j 节点在 t 时段电压之间的相角差。

3) 网络节点电压约束。

$$U_i^{\min} \leq U_{i,t,w} \leq U_i^{\max} \quad (32)$$

式中, U_i^{\min}, U_i^{\max} 分别为节点 i 电压的上、下限。

4) 联络线传输容量约束。

$$(P_{e,t,w}^{s,IESP} + P_{r,t,w}^{s,IESP,up}, P_{e,t,w}^{b,IESP} + P_{r,t,w}^{b,IESP,up}) \leq P^{\max} \quad (33)$$

$$0 \leq P_{e,t,w}^{b,UG} + P_{r,t,w}^{b,UG,up} \leq P^{\max,UG} \quad (34)$$

$$P_{e,t,w}^{b,UG} - P_{r,t,w}^{b,UG,down} \geq 0 \quad (35)$$

式中, P^{\max} 为电网与多能耦合虚拟电厂的最大传输功率限值。

2 基于 MIP-MADDPG 的求解方法

本文所涉及的不完全信息市场环境下的多市场主体博弈均衡求解问题包含大量非线性、离散变量,且求解流程复杂,市场主体之间信息不互通,各过程之间存在大量的变量传递,导致采用传统方法求解较为困难。无模型的人工智能方法如多代理强化学习可以学习市场主体的决策模式,对竞争对手的决策信息进行拟合判断,有效地解决了不完全

信息市场环境博弈问题。然而,强化学习算法在求解过程中约束仅以目标函数惩罚的形式存在,部分结果无法满足电网安全运行约束,造成严重后果。因此,本文提出了一种 MIP-MADDPG 进行模型求解。

2.1 算法原理

MIP-MADDPG 基于 MADDPG 算法进行改进,集成看门狗模块,即对不符合约束的迭代轮次进行退出处理,保证每次迭代解均符合约束,actor 网络对应动作选择过程, critic 网络根据全局已知信息进行动作选择评价并指导网络参数更新。对于数量为 N 的代理,其策略集 $\pi = \{\pi_1, \pi_2, \dots, \pi_N\}$ 被 $\theta = \{\theta_1, \theta_2, \dots, \theta_N\}$ 参数化,此时的期望的回报梯度可以表示为

$$\begin{aligned} \nabla_{\theta} J(\theta_i) = \\ E_{s,a \sim D} [\nabla_{\theta} \mu_i(a_i | o_i) \nabla_{a_i} Q_i^{\mu}(s, a_1, a_2, \dots, a_N) |_{a_i = \mu_i(o_i)}] \end{aligned} \quad (36)$$

式中, $Q_i^{\mu}(s, a_1, a_2, \dots, a_N)$ 为进行训练的动作值函数,其输入所有代理的动作和状态信息,输出每个代理的动作值函数值;状态信息包含所有代理的观察值 $\{o_1, o_2, \dots, o_N\}$; $D = \{s, s', a_1, a_2, \dots, a_N, r_1, r_2, \dots, r_N\}$ 为经验回放池。

进行神经网络参数更新时,由于各个代理之间的信息不互通,须通过参数估计值 $\mu_{\phi_i}^j$ 对其他代理的策略进行近似,参数更新方程为

$$L(\phi_i^j) = -E_{o_j, a_j} [\log \mu_i^j(a_j | o_j) + \lambda H(\mu_i^j)] \quad (37)$$

式中, ϕ_i^j 为近似参数; H 为策略分布的熵。

图2为 MIP-MADDPG 算法的求解流程。

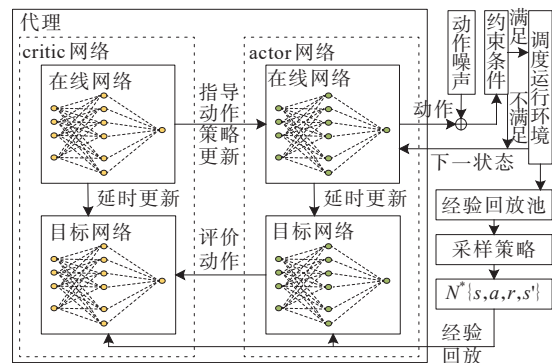


图2 MIP-MADDPG 算法求解流程

Figure 2 Process of MIP-MADDPG algorithm

2.2 模型求解

采用 MIP-MADDPG 求解本文模型的过程如下:主辅联合市场竞价博弈模型中,每个多能耦合

虚拟电厂均为 MIP-MADDPG 算法中的一个代理,不完全信息的主辅联合市场对应算法的环境,多能耦合虚拟电厂的投标决策对应代理的动作,电网出清结果和各个机组运行情况为环境的状态,整个市场的总利润为奖励,代理与环境不断交互,最终达到均衡状态。图 3 为本文所提模型和 MIP-MADDPG 求解算法之间的对应关系示意图。

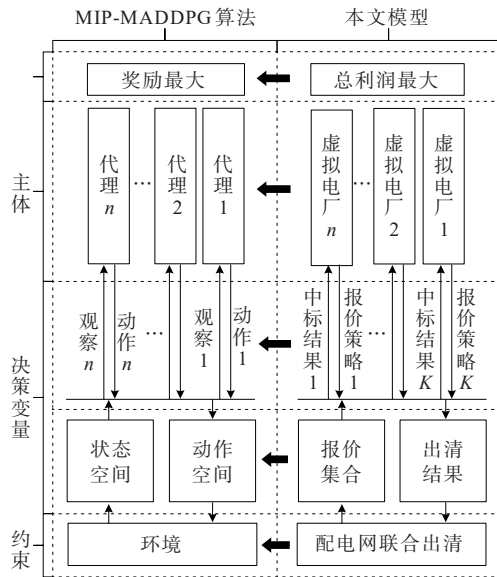


图 3 本文模型与 MIP-MADDPG 算法对应关系示意图
Figure 3 Correspondence between the proposed model and MIP-MADDPG algorithm

表 1 为 MIP-MADDPG 算法求解的伪代码,展现了本文的求解流程。

表 1 基于 MIP-MADDPG 算法的模型求解流程
Table 1 Model solving process based on MIP-MADDPG algorithm

| 步骤 | 求解流程 |
|----|--|
| 1 | 初始化代理 critic 网络 $Q_{\theta^c}(s, a)$ 和 actor 网络 $\mu_{\theta^a}(o)$ 参数 |
| 2 | 用参数 θ^Q 和 θ^μ 初始化目标网络 |
| 3 | for 迭代轮次 = 1 to 最大迭代轮次 do: |
| 4 | 初始化主辅联合市场出清结果 |
| 5 | for $t = 1$ to T do: |
| 6 | 观察当前市场状态 $s = \{o_1, o_2, \dots, o_{J+K}\}$ |
| 7 | 为每个虚拟电厂选择动作 $a = \mu(o_t \theta^a)$ |
| 8 | 完成一次市场出清,得到下一状态 s' 和回报 r |
| 9 | if 虚拟电厂决策量不满足系统安全运行约束 do: |
| 10 | 结束本次迭代,从步骤 5 开始新一轮循环 |
| 11 | Break |
| 12 | 存储状态转移矩阵 $\{s, a, r, s'\}$ 到经验回放池 |
| 13 | for $k = 1$ to K do: |
| 14 | 从经验回放池中随机抽取样本 $\{s, a, r, s'\}$ |

续表 1

| 步骤 | 求解流程 |
|----|---|
| 15 | 计算 TD 目标 $y = r_t + \gamma Q_{\theta^c}(s', a_t', \dots, a_N') _{a_j' = \mu_j'(o_t)}$ |
| 16 | 通过最小化损失函数更新 critic/actor 参数 |
| 17 | end |
| 18 | 更新每个代理的目标网络 |
| 19 | $\theta^Q = (1 - \tau)\theta^Q + \tau\theta^Q$; $\theta^\mu = (1 - \tau)\theta^\mu + \tau\theta^\mu$ |

3 算例分析

3.1 参数设置

算例部分采用 IEEE 14 节点标准测试系统,电网内部包含风电机组、光伏机组和 3 个多能耦合虚拟电厂,其中虚拟电厂 1 和虚拟电厂 2 内部可再生能源为光伏,虚拟电厂 3 内部可再生能源为风电。算例系统的拓扑结构如图 4 所示。电网除了与多能耦合虚拟电厂之间进行主辅联合交易外,也可以向上级电网申请购售电交易,上级电网电能价格和备用价格如表 2 所示。各个机组参数如表 3 所示。电网系统中天然气价格如表 4 所示。

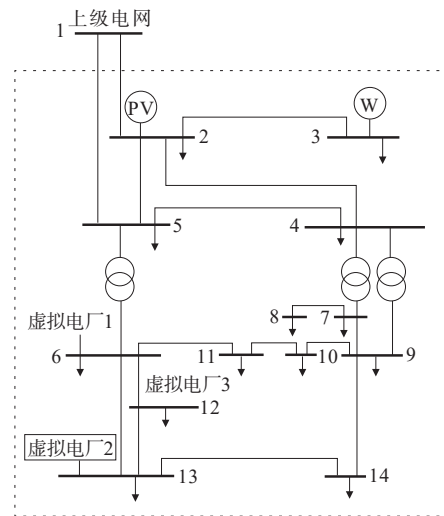


图 4 算例网络拓扑结构示意图
Figure 4 Network topology of calculation example

表 2 上级电网电能价格和备用价格

| 时段 | superior power grids 美元/千瓦时 | |
|---------------------------------------|-----------------------------|----------|
| | 上级电网电能价格 | 上级电网备用价格 |
| 01:00—06:00, 23:00—24:00 | 0.39 | 0.16 |
| 07:00—09:00, 15:00—17:00, 21:00—22:00 | 0.74 | 0.22 |
| 10:00—14:00, 18:00—20:00 | 1.21 | 0.42 |

表 3 多能耦合虚拟电厂内部机组参数

Table 3 Internal unit parameters of a multi-energy coupled virtual power plant

| 虚拟电厂 | CHP | | | | | 燃气锅炉 | | | 电/热储能 | | | | | | 成本系数 | |
|------|-------|------|------|-------|-------|-------|------|------|-------|---------|---------|---------|---------|----------|----------|-----------|
| | 容量/MW | 最小出力 | 爬坡速率 | 电效率/% | 热效率/% | 容量/MW | 最小出力 | 效率/% | 容量/MW | 最小容量/MW | 初始容量/MW | 充电速率 | 放电速率 | 充电效率 | | 放电效率 |
| 1 | 2 | 0 | 0.6 | 50 | 40 | 1 | 0 | 85 | 2/2 | 0.2/0.2 | 0.8/1.2 | 1.8/1.2 | 1.8/1.6 | 0.95/0.9 | 0.95/0.9 | 0.05/0.05 |
| 2 | 2 | 0 | 0.6 | 45 | 35 | 2 | 0 | 90 | 2/2 | 0.2/0.2 | 0.8/1.2 | 1.8/1.2 | 1.8/1.6 | 0.95/0.9 | 0.95/0.9 | 0.05/0.05 |
| 3 | 2 | 0 | 1.0 | 45 | 35 | 3 | 0 | 85 | 3/3 | 0.2/0.3 | 0.8/1.2 | 1.8/1.8 | 1.8/2.4 | 0.95/0.9 | 0.95/0.9 | 0.05/0.05 |

表 4 天然气价格情况

Table 4 Natural gas price 美元/千瓦时

| 时段 | 天然气价格 |
|---------------------------------------|-------|
| 24:00—次日 05:00 | 0.25 |
| 06:00—07:00, 13:00—16:00, 20:00—23:00 | 0.36 |
| 08:00—12:00, 17:00—19:00 | 0.43 |

各个多能耦合虚拟电厂的负荷情况如图 5 所示。

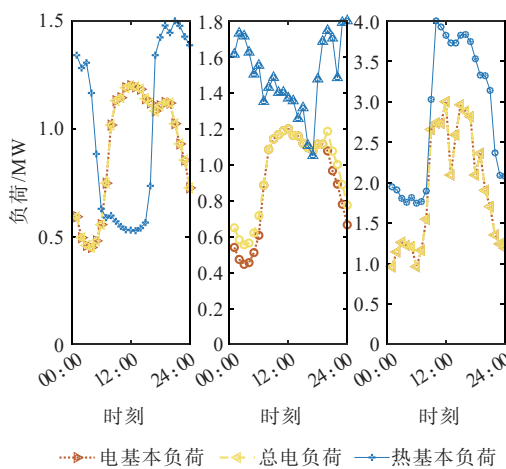


图 5 多能耦合虚拟电厂的负荷情况

Figure 5 Load of multi-energy coupled virtual power plants

MIP-MADDPG 算法训练数据集选取 2020 年夏季中国上海某海上风电场和 2022 年夏季上海连民村综合能源示范区每日实际运行数据(包含 6、7、8 月份共 92 d),包括负荷需求、分布式风电、光伏和楼宇空调功率;MIP-DDPG 算法在训练阶段依靠训练数据集学习负荷需求和各种灵活性资源出力模式,对训练集中的每个场景做 200 轮次迭代,更新其内部神经网络参数并将训练完成的参数保存,用于实际优化运行过程,MIP-MADDPG 算法训练过程如图 6 所示。

由图 6 可知,MIP-MADDPG 算法在迭代轮次到达 20 次之前奖励值波动较大,表明算法正在进行不同策略之间的探索,并不断在经验回放池保存经验。随着迭代轮次的增加,算法在 25 次迭代之后奖

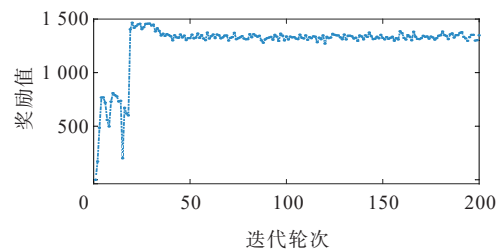


图 6 MIP-MADDPG 算法训练过程

Figure 6 Training process of MIP-MADDPG algorithm

励值趋于平稳,表明算法完成了收敛,证明了本文所提算法的稳定性。

3.2 结果分析

根据算法训练完成的网络参数进行在线执行,得到的多能耦合虚拟电厂 1 内部机组出力情况、备用情况和热平衡情况如图 7~9 所示。

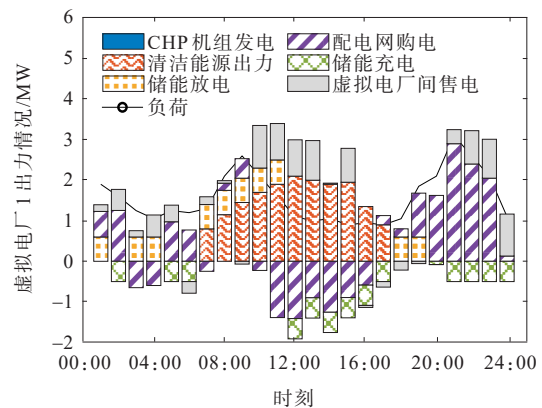


图 7 多能耦合虚拟电厂 1 内部机组出力情况

Figure 7 Internal unit output of multi-energy coupled virtual power plant 1

从多能耦合虚拟电厂 1 进行电能量交易时内部机组的处理情况可知,由于虚拟电厂 1 内部包含光伏机组,其出力集中在日间且中午 12:00 左右达到峰值,由于清洁能源出力消纳优先级较高,且此时虚拟电厂 1 产生了较多电能量富余,在时间段 08:00—16:00 中虚拟电厂 1 向电网进行售电以获取收益。而夜间光伏机组出力为 0,为了满足负荷需求,选择向电网购电。

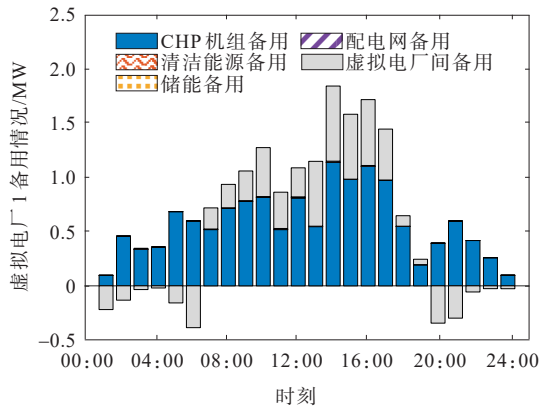


图 8 多能耦合虚拟电厂 1 备用情况

Figure 8 Standby situation of multi-energy coupled virtual power plant 1

观察多能耦合虚拟电厂 1 备用情况可知,其内部备用需求主要依靠热电联产机组和虚拟电厂之间交易备用。且由于热电联产机组需要提供热能,备用能力有限,在 06:00—18:00 备用需求高峰期对其余多能耦合虚拟电厂采购备用容量。

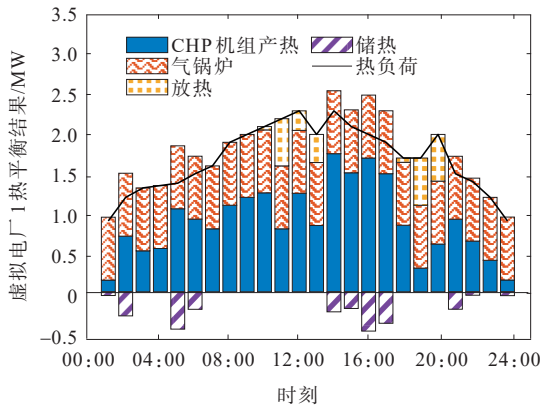


图 9 多能耦合虚拟电厂 1 热平衡情况

Figure 9 Thermal balance of multi-energy coupled virtual power plant 1

热电联产机组在电能量市场没有出力,从虚拟电厂 1 热平衡结果可知,热电联产机组主要进行产热以满足热负荷需求,配合内部的清洁能源机组实现电-热-气多能源的协同最优运行。

多能耦合虚拟电厂 2 在进行电能量交易时的内部机组出力情况如图 10 所示。多能耦合虚拟电厂 3 内部机组出力情况如图 11 所示,其能耦合虚拟电厂 3 的出力和备用情况。虚拟电厂 2 和虚拟电厂 1 内部包含资源情况相同,分析方法类似,在此不再赘述。

分析虚拟电厂 3 运行结果可知,其电能量出力主要依靠热电联产机组,且除了 15:00—20:00 的所

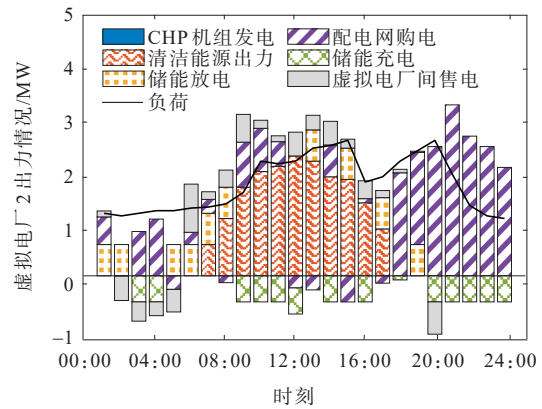


图 10 多能耦合虚拟电厂 2 内部机组出力情况

Figure 10 Internal unit output of multi-energy coupled virtual power plant 2

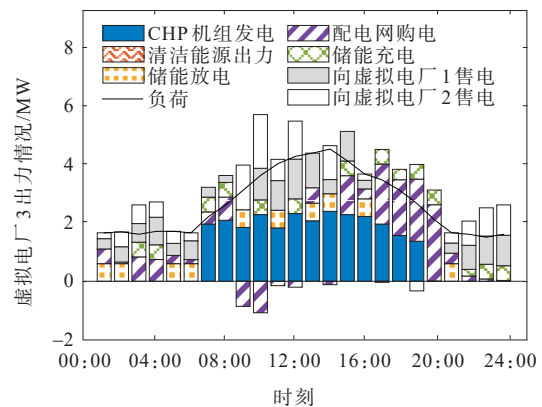


图 11 多能耦合虚拟电厂 3 内部机组出力情况

Figure 11 Internal unit output of multi-energy coupled virtual power plant 3

有时刻均向多能耦合虚拟电厂 1 和 2 售电,这是由于其内部风电资源丰富,在满足负荷需求前提下为了避免弃风惩罚选择向外售电。

为了研究多个虚拟电厂之间的交易情况,绘制 3 个虚拟电厂净出售的电量曲线及其交易价格,图 12、13 分别为多能耦合虚拟电厂之间进行交易的净电量和电价情况。

从多能耦合虚拟电厂之间进行的交易电量结果可知,由于虚拟电厂 1 和虚拟电厂 2 内部均包含光伏资源,在 08:00—16:00 时间段内光伏出力充足,在此阶段虚拟电厂 1 和 2 交易电量均为正值,虚拟电厂 3 交易电量为负值,说明虚拟电厂 1、2 在内部功率平衡之后,将富裕电量进行出售以获取利润,而虚拟电厂 3 选择在这一阶段从其余虚拟电厂购入相对便宜的清洁能源。观察价格变化趋势可知,08:00—16:00 时间段光伏充足,电量价格走低,而 16:00—20:00 时间段,虚拟电厂之间进行交易的电能量价格明显偏高,虚拟电厂 1 由于光伏等波动

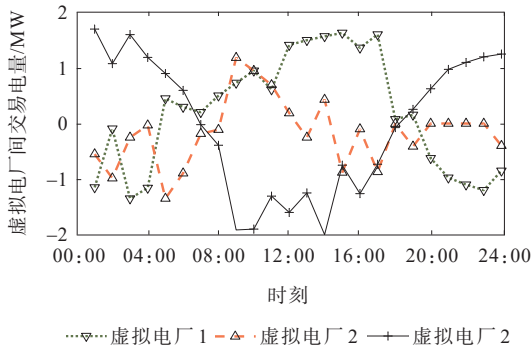


图 12 多能耦合虚拟电厂之间进行交易的净电量

Figure 12 Electricity trading between multi-energy coupled virtual power plants

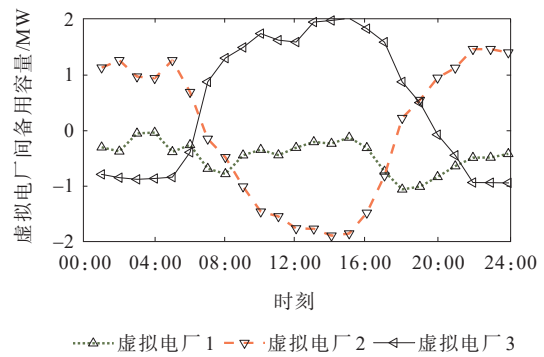


图 14 多能耦合虚拟电厂之间进行交易的净备用容量

Figure 14 Reserve capacity for trading between multi-energy coupled virtual power plants

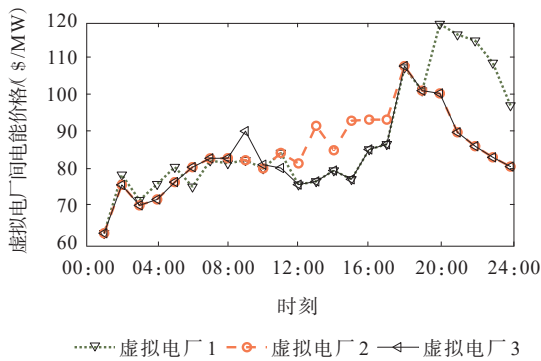


图 13 多能耦合虚拟电厂之间进行交易的电价

Figure 13 Electricity prices for trading between multi-energy coupled virtual power plants

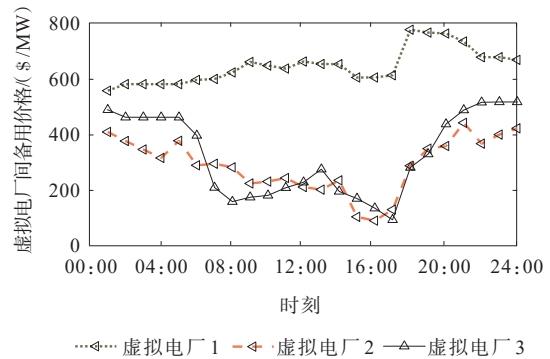


图 15 多能耦合虚拟电厂之间进行交易的备用价格

Figure 15 Reserve prices for trading between multi-energy coupled virtual power plants

能源出力骤减,需要向其余虚拟电厂购入电能,在 20:00 达到了价格最高点 120 \$/MW。

图 14、15 分别为多能耦合虚拟电厂之间进行交易的净备用容量和价格结果。

通过分析图 12、13 可知,在 08:00—16:00 时间段内虚拟电厂 1 和 2 将电量进行出售以获取利润,但是其内部出力不确定性较强。为了满足其备用需求,在该时段的备用容量交易为负,表示向其余虚拟电厂购入备用容量,而虚拟电厂 3 的情况相反。由以上分析可知,本文所提方法可以有效地针对各

个虚拟电厂内部的资源禀赋提出主辅联合市场的交易策略,通过电量、备用交易配合,从而获取最大利益。

为了验证本文所提方法的有效性,本节算例场景为场景 I。设置对比场景 II:进行完全信息市场环境下的博弈均衡分析,采用 GUROBI 求解器求解全局信息条件下的博弈优化问题作为理论最优解。设置对比场景 III:采用不考虑约束的 MADDPG 算法进行求解。设置对比场景 IV:采用传统 DQN 算法进行求解,与本文方法进行对比结果如表 5 所示。

表 5 不同算例场景对比情况

Table 5 Comparison of scenarios in different calculation examples

| 场景 | 总收益/ \$ | 主能量收益/ \$ | 辅助服务收益/ \$ | 总运行成本/ \$ | 弃风光成本/ \$ | 求解时间/ s | 违约约束量/ MW |
|-----|------------|--------------|---------------|--------------|--------------|------------|--------------|
| I | 16 843.24 | 10 254.69 | 6 588.55 | 3 682.45 | 349.21 | 10.3 | 0 |
| II | 16 984.21 | 10 363.78 | 6 620.43 | 3 676.38 | 340.24 | 15.2 | 0 |
| III | 16 736.14 | 10 156.87 | 5 963.75 | 4 258.31 | 553.64 | 9.6 | 10.26 |
| IV | 16 796.64 | 10 246.84 | 6 549.80 | 3 702.42 | 386.14 | 11.2 | 19.31 |

观察表 5 所示对比结果可知,本文所提方法得到的虚拟电厂总收益可以达到理论最优解的 99.17%,求解时间方面,本文求解时间为 10.3 s,较场景 II 优化 47.6%,在保证不完全信息市场环境的情况下逼近最优解,且主能量市场收益、辅助服务收益和总运行成本等指标结果均近似等于理论最优解,最大误差不超过 2.64%,且相较于场景 III,有效避免了超约的情况,证明了本文方法的准确性和安全性。

4 结语

本文提出了一种基于约束感知多代理强化学习的电-热-气耦合虚拟电厂参与主辅联合市场的博弈竞价策略,并通过算例仿真得到了以下结论:

1) 搭建了多能耦合虚拟电厂的数学模型,描述了其内部电-热-气多能耦合的能量流动关系,通过算例仿真验证了其模型准确性,该模型可以准确地描述多能之间的耦合关系,且能较好地满足多能用户需求;

2) 将多能耦合虚拟电厂作为价格制定者参与主辅联合市场博弈,并允许虚拟电厂之间进行电能量和辅助服务交易,在不完全信息市场环境下求解博弈均衡状态。算例分析表明,本文所提方法在 50 轮次迭代之后完成了收敛,大幅提高了不完全信息市场博弈均衡问题的求解速度;

3) 对约束感知强化学习方法进行了改进,提出 MIP-MADDPG 算法进行模型求解,有效保证了虚拟电厂策略满足电网安全约束,通过和理论最优解进行对比,本文方法可逼近理论最优值的 99.17%,且有效避免了超约束现象,证明了本文方法有效性。

参考文献:

[1] 张宁,朱昊,杨凌霄,等.考虑可再生能源消纳的多能互补虚拟电厂优化调度策略[J].发电技术,2023,44(5):625-633.
ZHANG Ning, ZHU Hao, YANG Lingxiao, et al. Optimal scheduling strategy of multi-energy complementary virtual power plant considering renewable energy consumption[J]. Power Generation Technology, 2023, 44(5): 625-633.

[2] KONG X Y, XIAO J, LIU D H, et al. Robust stochastic optimal dispatching method of multi-energy virtual power plant considering multiple uncertainties[J]. Applied Energy, 2020, 279: 115707.

[3] 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[J].电网技术,2021,45(9):3534-3545.
SUN Huijuan, LIU Yun, PENG Chunhua, et al. Optimization scheduling of virtual power plant with carbon capture and waste incineration considering power-to-gas coordination[J]. Power System Technology, 2021, 45(9): 3534-3545.

[4] 刘蔚,张野,吴应双,等.基于改进价值分解网络的集成虚拟电厂的互联电网动态最优协作控制[J].供用电,2024,41(12):62-71.
LIU Wei, ZHANG Ye, WU Yingshuang, et al. Dynamic optimal collaborative control of interconnected power grids for integrated virtual power plants based on improved value decomposition networks[J]. Distribution & Utilization, 2024, 41(12): 62-71.

[5] JU L W, YIN Z, LU X L, et al. A Tri-dimensional Equilibrium-based stochastic optimal dispatching model for a novel virtual power plant incorporating carbon Capture, Power-to-Gas and electric vehicle aggregator [J]. Applied Energy, 2022, 324: 119776.

[6] 康靖,李雨桐,陈泉,等.基于建筑负荷聚合的虚拟电厂模式构建与运行探索[J].供用电,2025,42(4):59-68.
KANG Jing, LI Yutong, CHEN Quan, et al. Construction and operation exploration of virtual power plant mode based on building load aggregation[J]. Distribution & Utilization, 2025, 42(4): 59-68.

[7] 于松源.热电联产虚拟电厂协同优化调度研究[D].北京:华北电力大学,2021.
YU Songyuan. Coordinated optimal scheduling of combined heat and power-virtual power plants[D]. Beijing: North China Electric Power University, 2021.

[8] 肖杰.含多能源类型和多运营商的虚拟电厂优化调度研究[D].天津:天津大学,2019.
XIAO Jie. Research on optimal scheduling of virtual power plants with multiple energy types and operators [D]. Tianjin: Tianjin University, 2019.

[9] SOARES J, ALI FOTOUHI GHAZVINI M, VALE Z, et al. A multi-objective model for the day-ahead energy resource scheduling of a smart grid with high penetration of sensitive loads[J]. Applied Energy, 2016, 162: 1074-1088.

[10] 张艺,刘蕊.考虑不确定性风险的虚拟电厂优化调度模型研究[J].智慧电力,2024,52(8):9-18.
ZHANG Yi, LIU Rui. Virtual power plant optimal scheduling model considering uncertain risks[J]. Smart Power, 2024, 52(8): 9-18.

[11] 高峰,黄丽丽,陈洁,等.基于分层分区调控的虚拟电厂市场多时间尺度竞价策略[J].中国电力,2026,59(2):138-147.
GAO Feng, HUANG Lili, CHEN Jie, et al. Multi-time

- scale bidding strategy for virtual power plant markets based on hierarchical partition control[J]. *Electric Power*, 2026, 59(2): 138-147.
- [12] 李咸善, 马凯琳, 程杉. 含多区域综合能源系统的主动配电网双层博弈优化调度策略[J]. *电力系统保护与控制*, 2022, 50(1): 8-22.
- LI Xianshan, MA Kailin, CHENG Shan. Dispatching strategy of an active distribution network with multiple regional integrated energy systems based on two-level game optimization[J]. *Power System Protection and Control*, 2022, 50(1): 8-22.
- [13] 刘超波. 计及多区域综合能源系统的虚拟电厂协调优化调度研究[D]. 北京: 华北电力大学, 2023.
- LIU Chaobo. Research on coordinated optimal scheduling of virtual power plant considering multi-regional integrated energy system[D]. Beijing: North China Electric Power University, 2023.
- [14] SHAFIEKHANI M, BADRI A, SHAFIE-KHAH M, et al. Strategic bidding of virtual power plant in energy markets: a bi-level multi-objective approach[J]. *International Journal of Electrical Power & Energy Systems*, 2019, 113: 208-219.
- [15] WANG Y, AI X, TAN Z F, et al. Interactive dispatch modes and bidding strategy of multiple virtual power plants based on demand response and game theory[J]. *IEEE Transactions on Smart Grid*, 2016, 7(1): 510-519.
- [16] 湛归, 殷爽睿, 艾芊, 等. 智能楼宇型虚拟电厂参与电力系统调频辅助服务策略[J]. *电力工程技术*, 2022, 41(6): 13-20+57.
- ZHAN Gui, YIN Shuangrui, AI Qian, et al. A strategy for smart building-based virtual power plants participating in frequency regulation auxiliary service[J]. *Electric Power Engineering Technology*, 2022, 41(6): 13-20+57.
- [17] KARDAKOS E G, SIMOGLU C K, BAKIRTZIS A G. Optimal offering strategy of a virtual power plant: a stochastic bi-level approach[J]. *IEEE Transactions on Smart Grid*, 2016, 7(2): 794-806.
- [18] QIU D W, WANG J H, WANG J K, et al. Multi-agent reinforcement learning for automated peer-to-peer energy trading in double-side auction market[C]// *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*. Montreal, Canada. International Joint Conferences on Artificial Intelligence Organization, 2021: 2913-2920.
- [19] LIU D N, GAO Y, WANG W Y, et al. Research on bidding strategy of thermal power companies in electricity market based on multi-agent deep deterministic policy gradient[J]. *IEEE Access*, 2021, 9: 81750-81764.