

DOI: 10.19783/j.cnki.pspc.241453

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

张宁¹, 杨凌霄², 李炫浓³, 胡存刚¹, 孙秋野⁴

(1. 安徽大学电气工程与自动化学院, 安徽 合肥 230601; 2. 安徽大学人工智能学院, 安徽 合肥 230601;
3. 国网肥西县供电公司, 安徽 合肥 231299; 4. 东北大学智能电气科学与技术研究院, 辽宁 沈阳 110819)

摘要: 虚拟电厂(virtual power plant, VPP)作为一种新型能源管理模式, 将分布式能源资源进行智能化集成和优化, 其对于促进可再生能源消纳、能源结构的优化和能源系统的绿色化具有重要意义。以多能虚拟电厂为研究对象, 以实现能源区域消纳为研究目的, 提出了一种基于分层深度强化学习的多能虚拟电厂区域消纳优化调度方法。首先, 提出了一种非直接多能虚拟电厂区域消纳运行框架, 确保用户参与自主性的同时避免用户信息公开化。其次, 基于多能耦合以及多时间尺度特性构建虚拟电厂内的联合交易机制, 避免了因忽略能源传输特性导致的交易失败, 实现跨能源类型的灵活匹配, 在完成区域自消纳的同时提高自身收益。最后, 提出基于分层深度强化学习的优化求解策略, 以解决所提模型由于大规模状态动作空间以及稀疏奖励特性带来的求解难题。通过仿真算例验证了所提方法的有效性, 表明所提虚拟电厂调度策略可以有效实现区域自消纳。

关键词: 虚拟电厂; 多能交易; 多时间尺度; 分层深度强化学习

Regional consumption optimization strategy for multi-energy virtual power plants based on hierarchical deep reinforcement learning

ZHANG Ning¹, YANG Lingxiao², LI Xuannong³, HU Cungang¹, SUN Qiuye⁴

(1. School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China; 2. School of Artificial Intelligence, Anhui University, Hefei 230601, China; 3. State Grid Feixi County Power Supply Company, Hefei 231299, China; 4. Intelligent Electrical Science and Technology Research Institute, Northeastern University, Shenyang 110819, China)

Abstract: As a novel energy management paradigm, the virtual power plant (VPP) enables the intelligent integration and optimization of distributed energy resources, playing a significant role in promoting renewable energy consumption, optimizing energy structures, and facilitating greener energy systems. Focusing on multi-energy VPPs with the objective of achieving regional energy consumption, this paper proposes a regional consumption optimization and scheduling method based on hierarchical deep reinforcement learning. First, a non-direct regional consumption operation framework for multi-energy VPPs is proposed, which ensures user autonomy in participation while avoiding the disclosure of private information. Second, a joint trading mechanism within the VPP is designed, considering multi-energy coupling and multi-timescale characteristics. This avoids trading failures caused by neglecting energy transmission constraints, enables flexible matching across different energy types, and enhances VPP revenues while realizing regional self-consumption. Finally, an optimization strategy based on hierarchical deep reinforcement learning is developed to overcome the challenges posed by the large-scale state-action space and sparse reward characteristics of the model. Simulation case studies validate the effectiveness of the proposed method, demonstrating that the scheduling strategy can effectively achieve regional self-consumption.

This work is supported by the National Natural Science Foundation of China (No. 62203004 and No. 62303006).

Key words: virtual power plant; multi-energy trading; multi-timescale; hierarchical deep reinforcement learning

0 引言

随着社会经济的快速发展和能源需求的增长,

传统的电力供应模式已无法满足需求, 人们对能源供应的环保性、高效性等方面提出了更高的要求。虚拟电厂(virtual power plant, VPP)作为一种新型能量管理模式, 能够聚合分布式电源、储能系统、电动汽车等设备以及不同能源类型进行协调运行^[1-5]。

基金项目: 国家自然科学基金项目资助(62203004, 62303006)

其可以以一个特殊的电厂参与能源市场和能源网络运行,从而实现资源的合理调配,进而促进低碳发展^[6-9]。因此,虚拟电厂在实现我国“碳达峰、碳中和”目标中起到至关重要的作用^[10]。

由于虚拟电厂能够有效整合区域内的异构能源设备和类型,因而其结构和运行特性与区域能源系统密切相关,呈现出多样性、灵活性的特点^[11]。此外,随着用户的行为变得更加灵活和自主,用户间端对端的能源交易也成为调节供需的重要手段^[12]。文献[13]基于价格匹配机制在电力系统里实现了用户间的交易。文献[14]提出一种支撑多微网端对端交易的配电系统协同优化调度模型,通过两层间的迭代实现配电系统和微网的协同优化运行。文献[15]在考虑可再生能源出力不确定性的情况下提出了用户端对端的电能交易策略。文献[16]构建了基于纳什议价理论的虚拟电厂电碳端对端交易机制模型。上述研究深入讨论了用户间的端对端能源市场构建,然而上述能源市场通常需要用户间进行信息交换以促进交易。在该机制中,其他用户通常可以获得诸如出价和需求等信息,存在用户信息泄露的隐患。因此,构建可以保证用户自主性同时避免用户信息泄露的端对端交易机制是一个重要课题。

除了以电能为主的虚拟电厂模型,能够有效整合多能源类型的虚拟电厂也开始被越来越多的学者研究。文献[16-22]深入讨论了多能虚拟电厂中不同能源类型对调度的影响,文献[17]提出了一种虚拟电厂参与日前电力现货市场投标的主从博弈多元模型,实现虚拟电厂收益及社会效益的最大化。文献[18]构建了基于碱性电解槽宽功率适应模型的风光氢热虚拟电厂,通过风光氢热能源优化配置,实现系统运行成本最低和碳排放最少。文献[19]提出一种考虑了需求响应的电-气互联的虚拟电厂模型,提高了系统的经济效益和清洁能源消纳率。文献[20]提出了一种多能虚拟电厂的预测优化方法,提高了虚拟电厂容纳高比例可再生能源的灵活性。文献[21]考虑电、热多种能源,对风速、光照、市场价格及负荷的不确定性进行建模,实现了虚拟电厂利润最大化和碳排放最小化。以多能虚拟电厂为主体的优化运行已经被广泛研究,但考虑多能虚拟电厂内用户间交互的研究相对较少。文献[22]同时考虑了自能源与不同能源网络以及自能源用户间的交易,构建了相应的博弈竞价模型,然而不同能源类型用户之间的交易依然是被禁止的。同时,鉴于不同能源类型的传输速度存在显著差异,能够应对区域多能系统的多时间尺度问题,并实现不同能源类型用户间自由交易的多能虚拟电厂模型尚需进一步深入研究。

随着新型电力系统的自主性与灵活性的提升,用户行为、可再生能源、价格波动带来的不确定性对电力系统运行产生的影响愈发显著,对应的精确环境模型难以构建^[23]。强化学习作为机器学习的一个重要分支,通过智能体与环境的互动来学习最优行为策略,展现出处理复杂动态环境、自主学习以及实时决策等显著特点。相较于其他优化控制算法,强化学习的独特优势在于其能够在没有精确环境模型的情况下完成决策,对于包含不确定性、环境模型难以精确构建的问题具有较强的适应性^[24-25]。该特性在虚拟电厂的研究与应用中显得尤为重要。文献[26]采用多步强化学习算法解决分布式能源系统的自动发电控制问题,由此减轻了可再生能源并网产生的随机扰动对控制性能的影响。因此,采用强化学习算法解决虚拟电厂中的序贯决策问题,成为了一种有效的思路并得到广泛应用^[27-28]。然而,近端策略优化等强化学习方法在处理具有大状态空间和稀疏奖励特征的多能调度模型时可能存在挑战^[29]。因此,对于具有处理相应问题能力的虚拟电厂求解策略仍值得进一步研究。针对上述研究现状,本文主要贡献如下:

1) 本文提出了一种基于价格匹配的区域消纳型多能虚拟电厂模型,所提模型在避免用户信息公开化的同时有效确保用户参与的自主性,提高了区域内用户自消纳的能力。

2) 考虑不同能源类型的物理特性对虚拟电厂调度的影响,构建了“多能-多时间尺度”联合交易机制。该机制基于能源转化情况、不同能源传输速率以及节点距离等因素计算用户间实际交易的能源传输时间。基于此,重新设计了虚拟电厂内用户间交易约束,有效避免了用户在市场匹配后因忽略能量传输过程而无法完成交易的情况。

3) 在联合混合交易机制的基础上构建调度模型,所提模型可以利用能量转换和储能设备实现跨能源类型的灵活匹配,并结合内部补贴积极响应能源网络需求以提高自身收益,有效提高虚拟电厂运行的灵活性与经济性。

4) 考虑到所提模型大规模状态动作空间以及稀疏奖励特性带来的求解难题,提出基于分层深度强化学习的优化求解策略。所提分层深度强化学习算法结合离线策略,通过对动作空间的分解有效地降低了计算复杂度,并规避了在环境奖励稀疏的情况下求解对应虚拟电厂调度问题所带来的挑战。

1 多能虚拟电厂模型

随着能源和信息技术的不断发展,能源系统正

在更深层次地向用户开放, 用户参与度日渐增加。与此同时, 区域能源就地消纳可以有效降低能量损耗。因此, 构建一个合理的区域消纳型虚拟电厂并设计相应的运行机制至关重要。本文设计了基于价格匹配的虚拟电厂运行模型, 模型中涵盖了电-气能源的交易运行模式。图1展示了该运行模型。如图所示, 区域消纳型虚拟电厂模型由用户群体和负责组织区域交易的电厂运营商构成。同时, 所提虚拟电厂模型内部还包含相应的能量转换设备和储能设备, 电厂运营商可以对能量转换设备以及储能设备进行调度以便更好地进行区域内的用户交易。其中, 能源转换设备包括微型燃气轮机(micro steam turbine, MST)和电转气(power to gas, P2G)设备。因此, 电力和天然气可以实现相互转换, 以便实现跨能源类型交易。

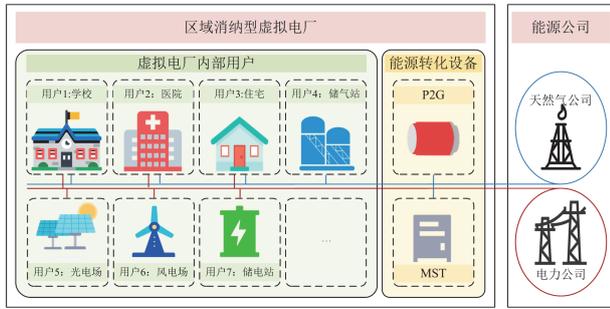


图1 多能虚拟电厂结构示意图

Fig. 1 Sketch map of proposed multi-energy VPP

所设计的虚拟电厂模型是基于用户间的日内交易。用户在所提虚拟电厂模型中以买家和卖家的身份参与, 而电厂运营商则基于用户需求和价格进行匹配。每个用户将根据自身的具体情况制定交易策略, 包括价格、能量和时间要求等。随后用户将策略提交给电厂运营商, 运营商会集合所有的信息, 并选择合适的配对。无法在虚拟电厂管理下完成区域消纳的用户将会与相应的能源供应公司进行交易, 以实现供需平衡。

在虚拟电厂调度过程中, 用户参与能量、时间以及能源投标价格在用户层不被公开。同时, 由于不同用户供给或需求的能源在能源市场上不具有特殊性, 部分匹配是由电厂运营商通过储能及能量转换设备实现。因此, 即使在虚拟电厂内完成匹配的用户也无需知道交易双方的信息, 避免用户信息公开化。此外, 用户可以完全根据自身意愿参与所提虚拟电厂调度, 虚拟电厂运行商无权改变用户基于自身偏好所做出的决策, 保障了用户的自主权。

2 虚拟电厂多维交易机制

2.1 用户竞价策略

考虑能源公司的存在, 在用户参与虚拟电厂调度过程中, 设定价格应加以限制, 以确保区域消纳型虚拟电厂中的能源购买方和出售方均获得利益, 对于虚拟电厂中的能源出售方, 出售价格的范围可以表示为

$$P_{acq}^{\theta} \leq p^{s^{\theta}} \leq P_{sell}^{\theta} - \zeta^{\theta} \quad (1)$$

式中: P_{acq}^{θ} 和 P_{sell}^{θ} 分别表示能源公司的能源收购价格和销售价格; p 表示用户的投标价格; s 表示虚拟电厂的能源出售方; θ 包含 e、g 两个符号, 分别表示电、气两种能源。能源差额 ζ 需要确保转换后的能源价格仍比直接与能源公司交易更具优势。对于虚拟电厂中的能源购买方, 收购价格的范围如式(2)所示。

$$P_{acq}^{\theta} + \zeta^{\theta} \leq p^{b^{\theta}} \leq P_{sell}^{\theta} \quad (2)$$

式中: b 表示在虚拟电厂中的能源购买方。

根据式(1)可知, 能源出售方进入虚拟电厂时出售价格不低于能源公司的收购价格, 保证了其进入虚拟电厂获得的利润不会低于其直接出售给能源公司的利润。同时, 由于能源购买方不会在虚拟电厂内以高于能源公司售价的金额购买能源, 因此能源出售方出售价格不能高于能源公司的出售价格。同理, 基于上述两方面的原因, 本文设立了式(2)所描述的能源购买方的价格约束。用户参与虚拟电厂调度的时间间隔设置为 $24/n^{\theta}$ 。其中 n 是对应能量在虚拟电厂一天中的交易次数。虚拟电厂允许用户进入后停留一定的时间间隔, 超出等待时间后用户会退出虚拟电厂。因此虚拟电厂运营商需要根据用户设置的等待时间对用户状态进行记录, 即参数 $J_{i,t}^{s^e}$ 、 $J_{i,t}^{s^g}$ 和 $J_{i,t}^{b^e}$ 、 $J_{i,t}^{b^g}$ 。 J 代表用户的剩余等待时间, i 和 \bar{i} 分别表示电能的购买方和出售方对应节点, l 和 \bar{l} 分别表示天然气的购买方和出售方对应节点, t 表示时间。当时间间隔从 $t \rightarrow t+1$ 时, 记录将从 $J_{i,t}^{s^{\theta}} \rightarrow J_{i,t+1}^{s^{\theta}}$ 减少, 其中 $J_{i,t+1}^{s^{\theta}} = J_{i,t}^{s^{\theta}} - 1$, $J_{i,t}^{s^{\theta}} = 0$ 时用户将从虚拟电厂退出。

2.2 计及用户交易的虚拟电厂调度模型

在虚拟电厂内的交易中, 所有用户都应根据自己的决定设计投标策略。鉴于用户的定价策略不受虚拟电厂运营商的影响, 并考虑到式(1)和式(2)中提出的价格机制, 只要用户之间的交易在虚拟电厂内完成, 无论是能源购买者还是能源出售者都能实现利润超过与能源公司交易的利润。此外, 虚拟电厂运营商必须采取适当动作, 确保在市场内最大限度

地实现区域消纳,同时提高虚拟电厂的收益。虚拟电厂运营商的目标函数设计为

$$\text{Max}F = B - S - Y + A \quad (3)$$

式中: B 是能源购买方购买能源时支付给虚拟电厂运营商的总金额; S 是能源出售方收取的总费用,该费用将从电厂运营商处扣除; Y 表示电厂运营商使用转换设备的成本; A 是虚拟电厂运行时的政策补贴,政策补贴和区域内的消纳程度呈正相关,确保经营者可以最大程度地完成用户间的匹配,实现区域自消纳。 B 可表示为

$$B = \sum_{t=1}^T \sum_{i \in N_i^{b^c}} v_i^{b^c}(t) p_i^{b^c}(t) + \sum_{t=1}^T \sum_{i \in N_i^{b^g}} v_i^{b^g}(t) p_i^{b^g}(t) \quad (4)$$

式中: T 表示虚拟电厂的整个调度周期; N 表示相应的参与者集合; v 表示参与方的总交易量。

S 可以表示为

$$S = S^{\text{dir}} + S^{\text{con}} \quad (5)$$

$$S^{\text{dir}} = \sum_{t=1}^T \sum_{i \in N_i^{s^c}} v_i^{s^c, \text{dir}}(t) p_i^{s^c}(t) + \sum_{t=1}^T \sum_{i \in N_i^{s^g}} v_i^{s^g, \text{dir}}(t) p_i^{s^g}(t) \quad (6)$$

$$S^{\text{con}} = \sum_{t=1}^T \sum_{i \in N_i^{s^c}} v_i^{s^c, \text{con}}(t) \hat{p}_i^{s^c}(t) + \sum_{t=1}^T \sum_{i \in N_i^{s^g}} v_i^{s^g, \text{con}}(t) \hat{p}_i^{s^g}(t) \quad (7)$$

式中: \hat{p} 是能源经过转换设备后的价格; 上标 dir 和 con 分别表示直接交易的能量和转换装置转换的能量。 Y 可表示为

$$Y = \sum_{t=1}^T \sum_{i \in N_i^{s^c}} \alpha_i c^{\text{P2G}} v_i^{s^c, \text{con}}(t) + \sum_{t=1}^T \sum_{i \in N_i^{s^g}} \beta_i c^{\text{MST}} v_i^{s^g, \text{con}}(t) \quad (8)$$

式中: c 是相应设备的转换成本; α 和 β 是用来表示对应能源转换设备是否工作的二进制变量。 A 可以表示为

$$A = \lfloor \delta / U z \rfloor \mathcal{G} \quad (9)$$

式中: δ 表示进入虚拟电厂且完成匹配的参与者的数量; U 表示所有用户在虚拟电厂中的总参与量; z 为政府补贴系数; \mathcal{G} 表示虚拟电厂完成部分区域消纳时的补贴单价,当有用户完成匹配便会给予 \mathcal{G} 的补贴。政府按分档机制进行补贴,为保证匹配的成功率,采用 $\lfloor \delta / U z \rfloor$ 向下取整确定响应的补贴档位。进一步,考虑到区域消纳型虚拟电厂交易过程中能源购买方接收到的所有能量均来自能源出售方,因此相关约束表示为

$$\sum_{i \in N_i^{b^c}} v_i^{b^c}(t) - \sum_{i \in N_i^{s^c}} v_i^{s^c}(t) = 0 \quad (10)$$

对于天然气, $v_i^{b^g}(t)$ 和 $v_i^{s^g}(t)$ 之间的关系与式(10)相似。与包含单一能源类型的虚拟电厂相比,所提区域消纳型虚拟电厂包含多种能源类型,同时允许参与者在虚拟电厂内调用能量转换设备。因此,应

设计相应的约束条件以确保有效运行。部分相关约束条件表示为

$$v_i^{s^c, \text{dir}}(t) + v_i^{s^c, \text{con}}(t) = v_i^{s^c}(t) \quad (11)$$

$$v_i^{s^c, \text{dir}}(t) = v_i^{b^c, \text{dir}}(t) \quad (12)$$

$$\sum_{i \in N_i^{b^c, \text{dir}}} v_i^{b^c, \text{dir}}(t) + \sum_{i \in N_i^{b^c, \text{con}}} v_i^{b^c, \text{con}}(t) = \sum_{i \in N_i^{b^c}} v_i^{b^c}(t) \quad (13)$$

$$v_i^{s^g, \text{con}}(t) \eta^{\text{MST}} = v_i^{b^c, \text{con}}(t + \Delta \tau_1^{i\bar{i}}) \quad (14)$$

式中: η 表示设备的运行效率,其上标表示对应设备; $\Delta \tau_1$ 是气体传输造成的延迟时间,其上标表示传输的两个端点。

2.3 基于能源传输特性的虚拟电厂多时间尺度能量约束

考虑区域消纳型虚拟电厂中不同能源的固有特性,在虚拟电厂内部交易过程中,需要考虑能量传输所需时间。如式(14)所示,在所提虚拟电厂模型中, t 是能源出售方传输能量的时间,在这种能量传输时间中,天然气的交易约束可表示为

$$v_i^{s^g, \text{dir}}(t) + v_i^{s^g, \text{con}}(t) = v_i^{s^g}(t) \quad (15)$$

$$v_i^{s^g, \text{dir}}(t) = v_i^{b^g, \text{dir}}(t + \Delta \tau_1^{i\bar{i}}) \quad (16)$$

$$\sum_{i \in N_i^{b^g, \text{dir}}} v_i^{b^g, \text{dir}}(t) + \sum_{i \in N_i^{b^g, \text{con}}} v_i^{b^g, \text{con}}(t) = \sum_{i \in N_i^{b^g}} v_i^{b^g}(t) \quad (17)$$

$$v_i^{s^c, \text{con}}(t) \eta^{\text{P2G}} = v_i^{b^g, \text{con}}(t + \Delta \tau_1^{i\bar{i}}) \quad (18)$$

由于天然气的传输速度相较于电能慢,因此虚拟电厂内天然气的交易间隔会更长。 $\Delta \tau_1$ 值由天然气传输速度与用户间的距离决定。因此, $\Delta \tau_1$ 可表示为

$$\Delta \tau_1 = \left\lceil \frac{D_g}{V_g \Delta t} \right\rceil \quad (19)$$

式中: D_g 表示能量转换设备与节点 l 间的距离; Δt 表示单位时间; V_g 表示天然气的传输速度。为了确保天然气能够在规定时间内顺利交付给购买方,出售方的出售时间与购买方的使用时间之间的时间差应当不小于实际的运输时间。因此,在计算天然气运输所需时间时,式(19)对实际运输时间进行向上取整,以确保计算出的运输时间大于或等于实际传输时间。而若采用四舍五入等方法,则可能会忽略部分运输时间,进而导致天然气无法按时交付。由此,本文考虑了不同类型能源的传输速度所导致的用户之间不同的交易延迟,从而构建出区域消纳型虚拟电厂中的多时间尺度交易模式。

式(14)一式(18)所描述的虚拟电厂内部交易模型考虑了不同能源的多时间特性。不同能源的传输固有特性导致其交易周期也存在差异性。此外,所

提虚拟电厂模型中还包括能源转换装置, 可以利用不同能源之间的耦合关系有效促进区域内用户自消纳。同时也考虑了设备转换中不同能源的传输差异导致的传输延时。基于上述设计, 所提虚拟电厂模型更好地反映了能源交易动态的实际情况。

需要注意, 多能系统作为一个多时间常数系统, 不同文献的多时间尺度有不同解释, 部分文献讨论了在多个时间周期进行优化或规划^[30], 也有部分文献讨论了多种能源传输响应的的时间尺度差异对控制与优化的影响^[31-32]。本文的多时间尺度主要体现了多种能源传输上的时间尺度差异性, 基于这种多时间尺度差异性讨论传输延时对虚拟电厂调度结果的影响。

3 基于分层深度强化学习的虚拟电厂调度策略

上文从运营商的角度出发, 构建了多能虚拟电厂的市场模型, 涵盖了运营商和用户的收益、虚拟电厂内用户间的交易限制, 以及设备运行和能流传输等约束条件。然而, 为了保障用户的自主性, 与用户行为相关的变量(如用户进入市场的时间、需求和报价)对于运营商具有不确定性, 对应的环境模型难以精确构建。强化学习通过智能体与环境的交互来学习最优行为策略, 适合求解包含不确定性且环境模型难以精确构建的问题^[24]。因此本文选择强化学习算法作为求解对应虚拟电厂调度策略的基本方法。此外, 所构建的虚拟电厂模型在运行过程中也会产生与大型状态动作空间和稀疏奖励相关的问题。为了应对上述问题, 本文提出了一种结合离线策略的分层深度强化学习算法来求解虚拟电厂调度策略。

3.1 马尔可夫决策过程

所提区域消纳型虚拟电厂的马尔可夫决策过程(Markov decision process, MDP)模型包括 4 个基本组成部分, 即 \hat{s} 、 a 、 \tilde{p} 和 R 。其中: \hat{s} 表示有限的环境状态; a 表示有限的动作集; \tilde{p} 表示从状态 \hat{s} 到状态 \hat{s}' 的状态转换概率; R 表示从环境中获得的瞬时奖励函数。 $\hat{s} \times a \times \hat{s}' \rightarrow R$ 表明奖励空间 R 与状态和动作之间的关系。在所提虚拟电厂 MDP 模型中, 具体的动作集 a 由能源转换量、交易价格等模型的决策变量决定, 同时也和参与者数量相关, 如式(20)所示。

$$a = \left\{ \begin{array}{l} v_i^{b^c}, p_i^{b^c}, v_i^{b^s}, p_i^{b^s}, v_i^{s^c, dir} \\ p_i^{s^c}, v_i^{s^c, dir}, p_i^{s^s}, v_i^{s^c, con} \\ \hat{p}_i^{s^c}, v_i^{s^c, con}, \hat{p}_i^{s^s}, \alpha_i, \beta_k \end{array} \right\} \quad (20)$$

虚拟电厂 MDP 模型状态由 4 个参数向量组成, 其中包括用户的投标参数、市场不匹配量、交易周期和现有用户的剩余停留周期。

奖励函数 R 用于表示虚拟电厂动作获得的奖励, 在设定约束条件下, 系统运行的目标是获得更多的奖励反馈。

$$R = \begin{cases} 0, & \text{如果违反约束} \\ F, & \text{其他情况} \end{cases} \quad (21)$$

所提虚拟电厂模型在求解的过程中不仅表现出多动作和多状态特性, 而且状态和动作数量与用户量直接相关, 因此虚拟电厂模型具有大状态-动作空间特性。此外, 由式(21)可以看出, 在满足约束的情况下, 本文所提算法的奖励值和虚拟电厂运营商的收益直接相关。在虚拟电厂运营商收益函数中考虑了在市场内完成匹配的订单, 并根据其定义了不同层级的政府补贴, 即式(9)描述的内容, 使得所提虚拟电厂模型仅在满足限制要求的小部分情况下即可以获得对应的奖励值, 对奖励结果影响较大且呈现出稀疏性。上述两个特点均给使用强化学习技术求解虚拟电厂模型调度策略带来了挑战。

3.2 结合离线策略的分层深度强化学习算法

由于模型奖励稀疏导致在学习策略的过程中可能长期得不到正反馈, 传统强化学习算法在求解所提虚拟电厂调度模型时会遇到困难。为解决上述难题, 本文提出结合离线策略的分层深度强化学习算法。所提算法将策略分为不同级别的子策略, 每个子策略在学习过程中都会获得上一级策略的奖励, 通过分层的结构, 上层控制器给下层控制器设定目标的同时也会根据下层控制器的策略好坏反馈一个对应的内在奖励, 保证了即使在下层控制器无法获得正反馈的外部奖励, 下层控制器依然能够获得奖励, 由此提升样本的利用效率, 从而在一定程度上缓解了奖励稀疏的问题。如图 2 所示, 所提分层深度强化学习算法采用双层结构, 包括上层策略和下层策略。上层策略在抽象层上运行, 并为下层策略设定目标, 这些目标直接对应于下层策略的目标状态。在下层策略中使用深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法进行训练。由于虚拟电厂模型中没有涉及储能等设备的影响, 相应的动态空间是连续的。同时 DDPG 作为一种深度强化学习方法, 通过智能体与环境的交互来学习最优行为策略, 适合求解包含不确定性且环境模型难以精确构建的问题^[24], 可以有效应对本文所构建的虚拟电厂模型。经验转换设置为 $(s_0, g_t, a_t, r_t, s_{t+1}, g_{t+1})$, 其中 g_t 是由上层策略产生的上层目标。下层策略需要接收上层策略迭代输出的目标 g , 因此对应的目标函数设为式(22)。

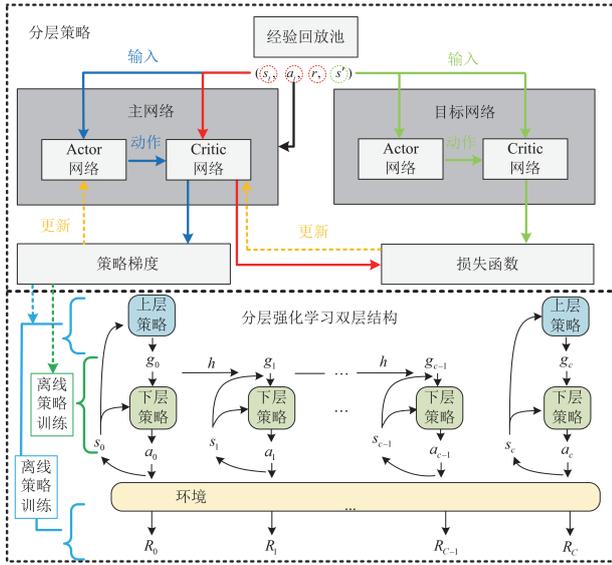


图 2 所提分层深度强化学习运行过程

Fig. 2 Operation process of the proposed HRL

$$\varepsilon^i(s_t, a_t, s_{t+1}) = \begin{pmatrix} Q_{\lambda}^i(s_t, g_t, a_t) - r(s_t, g_t, a_t, s_{t+1}) \\ -\gamma Q_{\lambda}^i(s_{t+1}, g_{t+1}, \mu_{\phi}^i(s_{t+1}, g_{t+1})) \end{pmatrix} \quad (22)$$

式中： ε 为贝尔曼误差函数； Q 是状态-动作函数，即 Q 函数，其表示从状态开始执行某一动作 s_t 的预期值； γ 为折扣因子； λ 和 ϕ 是梯度更新参数； μ 表示确定性神经网络策略；上标 i 表示下层策略。该策略在训练过程中通过添加高斯噪声来收集经验。上层策略在 s_0 产生一个目标 g_0 ，引导下层策略在步骤 c 之后到达 s_{c-1} 。上述关系可以表示为 $s_0 + g_0 = s_{c-1}$ 。状态变为 s_1 时， $s_0 \rightarrow s_1$ ， $s_0 + g_0 \rightarrow s_1 + g_0 \neq s_{c-1}$ 。为了保持目标的绝对位置不变，无论状态如何变化，设置目标转换模型 \tilde{h} 进行迭代，如式(23)所示。

$$\tilde{h}(s_t, g_t, s_{t+1}) = s_t + g_t - s_{t+1} \quad (23)$$

同时，内在奖励函数是根据当前观测值与目标观测值之间的距离定义的，如式(24)所示。通过最大化奖励函数， $s_t + g_t$ 和 s_{t+1} 之间的差值将变小。

$$r(s_t, g_t, a_t, s_{t+1}) = e_t - \|s_t + g_t - s_{t+1}\|^2 \quad (24)$$

式中： e_t 表示虚拟电厂的运行经济效益，由式(3)确定。因此，在该方法中，下层动作所获得的奖励既考虑了虚拟电厂的经济效益，又确保了下一个时刻的状态朝着上层训练预期的方向发展。由此，基于这种奖励函数，下层训练既考虑了经济方面，也考虑了系统中的用户交易，从而避免了过度追求交易获得的效益而导致用户在系统中无法完成交易，进而不能获得较大的市场补贴，从而影响总收益。

在上层策略中，DDPG 与离线策略被用于训练

策略 μ^h ，每 c 步更新一次。经验转换设置为 $(s_t, g_t, \sum R_{t+c}, s_{t+c})$ 。其中， R_{t+c} 表示序列 R_t, \dots, R_{t+c} 。上级策略的目标函数设为

$$\varepsilon^h(s_t, g_t, s_{t+c}, g_{t+c}) = (Q_{\omega}^h(s_t, g_t) - \sum R_{t+c} - \gamma Q_{\omega}^h(s_{t+c}, \mu_{\phi}^h(s_{t+c}, g_{t+c})))^2 \quad (25)$$

式中： ω 和 ϕ 是梯度更新的参数；上标 h 表示上层策略。在训练过程中，下层策略 μ^l 在随机环境中不断更新。在相同的状态 s_t 中， μ^l 对于相同的目标 g_t ，奖励 $\sum R_{t+c}$ 和下一个状态 s_{t+1} 是不断变化的。因此，为保持奖励不变，需要对离线策略进行修正。选择较高层次的目标 \tilde{g}_t 可以最大化在缓冲区中获得的较低层次动作的概率。对数概率函数 $\ln \mu^l(a_{t+c-1} | s_{t+c-1}, \tilde{g}_{t+c-1})$ 可按以下方式计算。

$$\ln \mu^l(a_{t+c-1} | s_{t+c-1}, \tilde{g}_{t+c-1}) \propto - \frac{1}{2} \sum_{i=t}^{t+c-1} \|a_i - \mu^l(s_i, g_i)\|_2^2 + C \quad (26)$$

$$\tilde{g}_t = \arg \max \mu^l(a_{t+c-1} | s_{t+c-1}, \tilde{g}_{t+c-1}) \quad (27)$$

式中： C 为函数固定量。在实际应用过程中，为接近最大值，通常会计算一组随机抽样目标 \tilde{g}_t 的对数概率，并选择最高值的目标对经验进行重新标记。

在分层深度强化学习算法中，奖励函数用于定义一组下层策略，其中每个策略都进行训练以将观察到的状态与预期目标相对应。上层策略选择这些目标，并采用一种针对离线策略的修正方法，将不同实际案例的下层策略经验纳入其中。所提基于分层深度强化学习的虚拟电厂求解策略运行过程如图 2 所示。所提强化学习算法可以提高采样效率，使其更普遍地适用于解决与大型状态-动作空间和稀疏奖励相关的问题。上层控制器为下层控制器提供目标，同时根据下层控制器的策略提供相应的内在奖励，避免出现在奖励稀疏环境下进行训练长期得不到正反馈的情况，最终得到虚拟电厂最优调度策略。

4 算例分析

4.1 系统描述

本文的 4 个算例均基于一台配置为 Intel Core i9-13900H 5.4 GHz CPU、16 GB RAM 以及 RTX4060 的计算机上完成。在这些算例中，交易时间是以响应速度最快的电能为基准来设定的，而价格周期则是基于 2022 年山东莱州某示范工程的电力市场数据，该市场的交易周期为 15 min。具体的价格以及设备参数等数据，则是参考了该示范工程 2021 年 12 月—2022 年 6 月的脱敏数据以及文献[22]来共同设置的。

算例对应系统中, MST 和 P2G 设备的功率上限均为 700 kW, 转化效率分别为 0.5 和 0.4。为保证所有交易能在一天内完成, 23:00—24:00 时段暂停交易, 15 min 为一个周期, 全天共计 92 个交易周期。算法核心参数中, 批量训练大小为 24, 由于所构建的多能虚拟电厂多维交易策略中, 允许用户设定等待时间, 同时天然气传输本身需要时间, 因此设定折扣因子为 0.99, 以在分层深度强化学习算法寻找最优策略过程中追求更为长远的奖励, critic 和 actor 网络的学习率分别为 0.001 和 0.0001, 噪声过程参数为 0.2, 经验池大小为 10 000, 隐藏层的大小为(256, 256, 18, 64)。由于用户行为的不确定性, 虚拟电厂 MDP 模型无法构建精确的环境模型, 因此采用经验回放机制来近似状态转移概率。同时考虑到模型的稀疏奖励特性, 采用优先经验回放获得状态转移概率。

为了证明所提模型的有效性, 在算例 1 中评估

了所提虚拟电厂模型在正常工作条件下的性能。算例 2 展示了是否考虑多时间尺度的虚拟电厂模型的性能差异。算例 3 则讨论了能源转换设备对虚拟电厂模型运行效率的影响。算例 4 对比了不同算法间的差异性。

4.2 算例结果

4.2.1 正常工作条件

在正常运行状态时, P2G 和 MST 这两个能量转换设备的运行结果如图 3 所示, 其中子图(a)和(d)分别表示两个转换设备在各个时刻的运行状态。(b)和(e)分别表示不同迭代次数下两种转换设备随迭代次数变化的运行状态。其中(a)、(c)和(d)、(f)分别对应(b)和(e)的截面。对比(c)和(f)可以看出, 迭代约 1600 次时结果趋于稳定, 表明所提基于分层深度强化学习算法的虚拟电厂调度方法在给定场景下通过迭代计算实现了收敛。

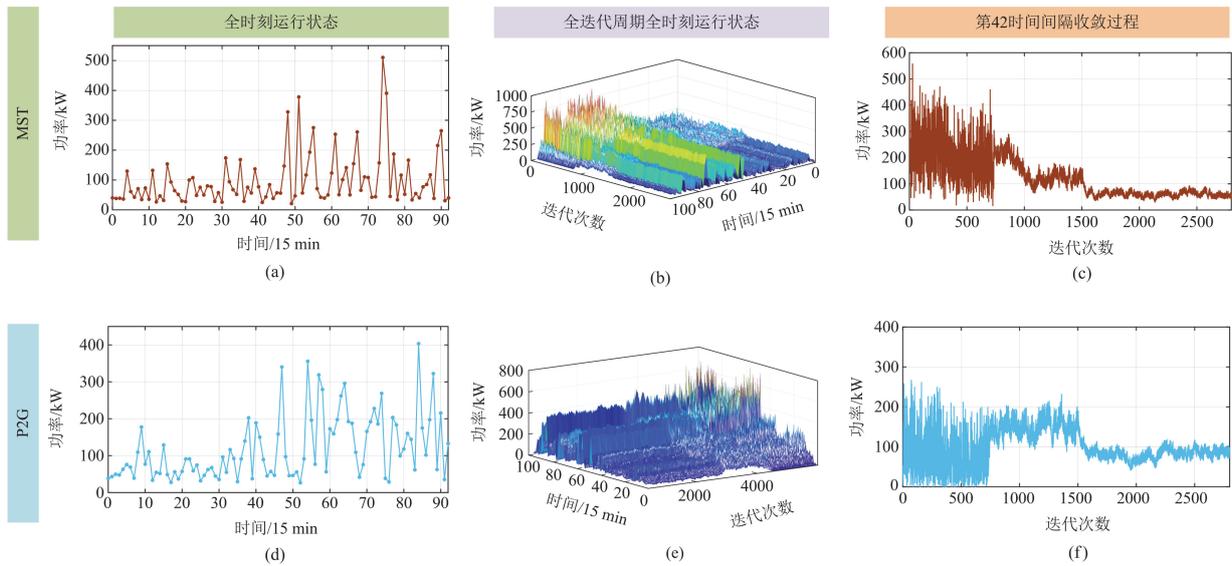


图 3 全天能源转换设备运行状态

Fig. 3 Operating status of energy conversion devices during a day

图 4 展示了在正常工作情况下的虚拟电厂内各时段的交易量和传输量。和图 3 对比可知, 由于虚拟电厂运营商利用 P2G 和 MST 这两个能量转换设备来增加用户的匹配度, 因此这两个转换设备工作情况取决于有多少用户需要进行转换, 因此能源转换的工作情况与用户匹配情况相关但不直接对应。除此之外, 用户的需求时间和用户进入虚拟电厂的时间可以不同, 同时涉及天然气交易也需要能量传输时间, 因此图 4 中交易量和传输量之间存在差异。

4.2.2 时间尺度影响分析

图 5 展示了不考虑能源时间尺度差异带来的影

响时, 虚拟电厂能量交易情况和损失的能量。对比图 4 和图 5 可以发现, 算例 1 所示情况的在一天的总交易量为 74 803 kWh, 算例 2 所示的情况在一天的总交易量为 77 368 kWh。由于算例 2 中没有考虑时间尺度差异造成的传输时延, 减少了虚拟电厂调度中的一些约束, 导致交易量略有增加。在这种情况下, 用户之间的需求匹配没有计算能源传输时间因为距离、能源类型等差异而造成的偏差, 天然气的传输延迟会导致无法在交易要求时间内完成能量传输, 进而导致能量损失。如图 5 所示, 在大部分时段都会有能量损失。因此, 考虑多时间尺度特性

对多能源交易的影响至关重要。

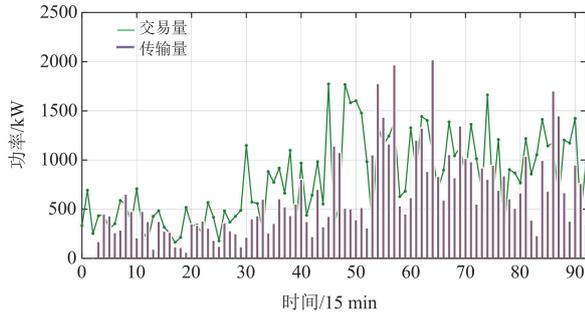


图4 算例1中能源交易量与能源传输量

Fig. 4 Volume of energy deals and transmissions of case 1

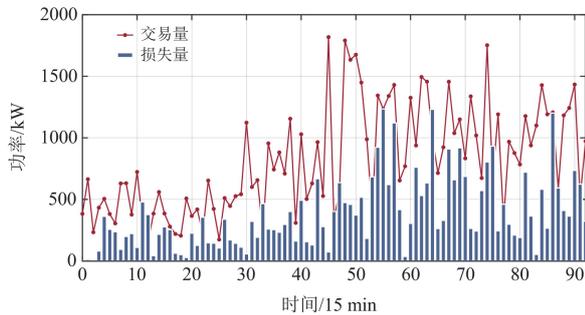


图5 算例2中能源交易量与能源损失量

Fig. 5 Volume of energy deals and transmissions of case 2

4.2.3 能源转换设备影响分析

图6展示了在有能量转换设备和没有能量转换设备两种情况下所提出的虚拟电厂之间的奖励值比较。从图中可以看出，有转换设备的虚拟电厂的回报始终高于无转换设备的虚拟电厂。事实上，由于能源转换设备存在能量损耗，从理论上讲，若用户能够完全不依赖此类设备而完全通过直接匹配实现能量供需平衡，虚拟电厂运营商将获得最高收益。然而，当前算例结果表明，市场运营者通过引入能源转换设备反而提高了用户参与市场匹配的成功率。

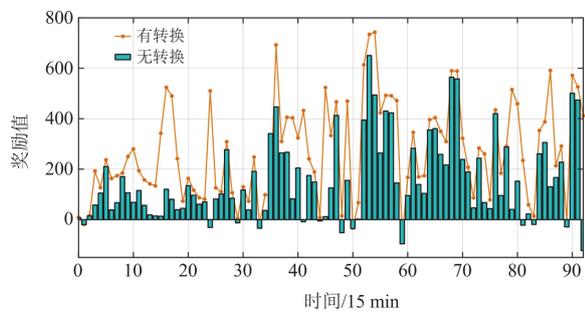


图6 有无能量转换设备两种情况下各时段奖励值对比

Fig. 6 Contrast of rewards obtained in the presence and absence of energy conversion devices

为了避免因设备损耗能量造成的收益降低，利用能源转换设备完成匹配是虚拟电厂运营商的后备方案，这种性质决定了除特殊情况外，使用能源转换设备只会提高匹配率而不会对用户匹配率造成负面影响。而从结果来看，有转换设备的虚拟电厂调度结果的奖励值为 25 116，而无转换设备的虚拟电厂调度奖励值为 14 563。因此，通过合理地使用能源转换设备可以有效提高虚拟电厂内的用户匹配率以及运营商的收益。

4.2.4 不同算法对比

算例4将本文所采用的分层深度强化学习方法、DDPG和基于模型的传统优化方法进行对比。传统优化算法采用MATLAB编写模型并利用CPLEX求解器求解。图7展示了DDPG和本文算法在相同条件下的运行结果。可以发现，相较于DDPG，本文所采用的分层深度强化学习算法表现更好。本文采用的分层深度强化学习算法是由两层DDPG算法构成，相较于单一DDPG，分层结构使得智能体可以在不同的层次上分别学习，每个层次关注的任务更为具体和简单。这有助于智能体更快地积累经验并优化策略，提高探索能力和训练稳定性，对稀疏奖励问题也有更好的表现。本文所构建的模型中状态与动作和参与市场的用户息息相关，在大量用户接入的情况下，系统状态维数大量增加，极大地提高了训练复杂度。同时本文针对市场内未完成匹配的订单，定义了政府补贴项，分档补贴机制对奖励结果影响较大且呈现出稀疏性，导致采用DDPG面临高方差的情况，从而影响了性能稳定性以及学习速度。对应到结果上即为DDPG在3000次迭代过程中均呈现出了结果波动大、收敛速度慢的情况。而本文所采用的分层深度强化学习算法可以更有效地解决模型对应的复杂状态空间和奖励稀疏性的情况。

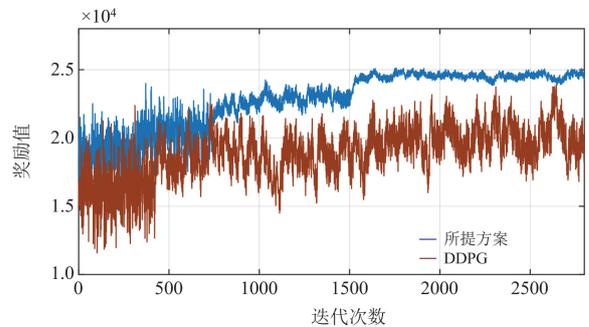


图7 不同强化学习算法迭代结果

Fig. 7 Iterative results of different reinforcement learning algorithms

图8展示了所提分层深度强化学习方法和基于模型的传统优化方法在各时段的奖励值表现。由于基于模型的优化方法需要给定信息求解最优策略, 而诸如用户进入虚拟电厂的时间、价格、需求等信息具备不确定性, 环境模型难以精确构建。在缺乏后续时间段内相关信息的情况下, 本算例中传统优化方法采用交易不等待的机制。和传统方法相比, 由于所提分层深度强化学习方法能更好地利用等待机制, 其奖励值在大多数时候都更高。尽管在诸如1、2、6、9等多个时段时传统优化方法的回报高于所提方法, 但采用传统优化方法时会有更多用户无法在市场内完成匹配, 进而造成更大的损失。采用传统优化方法的总奖励值为15 055, 比本文所采用算法总奖励值低了10 061。上述两个对比仿真结果表明, 与传统优化方法以及其他强化学习方法相比, 本文采用的分层深度强化学习算法在所提出的虚拟电厂模型中具有更好的表现。

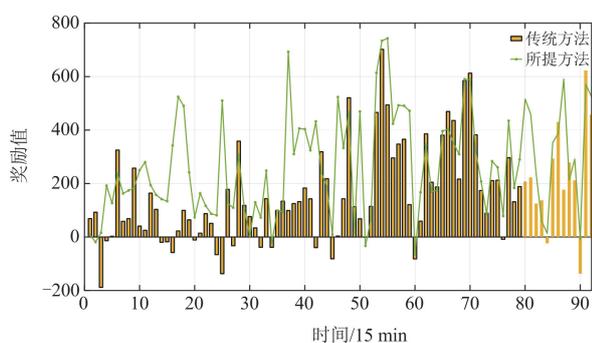


图8 不同方法各时段奖励值

Fig. 8 Reward values of different methods at various time periods

5 结论

本文以多能虚拟电厂为基础, 提出了一种基于分层深度强化学习的多能虚拟电厂区域消纳优化调度方法。所提方法在避免用户信息公开化的同时确保了用户参与的自主性, 在考虑运营商利益的基础上完成了区域自消纳。具体结论如下:

1) 考虑不同能源类型的物理特性对虚拟电厂调度的影响, 设计了基于多能耦合以及多时间尺度特性的联合交易机制。通过考虑不同能源类型的特点, 有效完成了匹配过程, 从而避免用户由于忽略能量传输过程而无法在虚拟电厂匹配后完成交易的情况。

2) 在所设计交易机制的基础上, 所提出的虚拟电厂模型采用能量转换设备实现跨能源类型的灵活匹配并提高匹配效率, 并通过内部补贴机制, 在完成区域自消纳的同时提高了虚拟电厂的自身收益,

从而有效提升了虚拟电厂运行的灵活性与经济性。

3) 本文提出了基于分层深度强化学习算法的虚拟电厂优化调度方法, 有效地解决了大状态动作空间和稀疏奖励所导致的求解难题。通过4个算例对不同情况下奖励值、交易量和传输量进行分析, 结果证明了所提调度方法的有效性。

本文所提方法有效提升了区域内用户的自我消纳能力, 但仍存在一些问题值得进一步研究。例如, 当用户出现违约情况时, 虚拟电厂应如何运行; 是否可以允许用户在进入虚拟电厂后的一个周期内调整其报价; 具备能源补充手段的运营商对于用户交易的灵活性会产生哪些影响等。基于上述问题, 对市场侧模型的调整以及市场机制的完善都是未来可以进一步研究的方向。

参考文献

- [1] 蔡光宗, 王伊晓, 袁智强, 等. 基于一致性算法的虚拟电厂调度指令动态跟踪策略[J]. 电力建设, 2024, 45(5): 71-79.
CAI Guangzong, WANG Yixiao, YUAN Zhiqiang, et al. Consensus-based dynamic dispatching instruction tracking strategy for virtual power plant[J]. Electric Power Construction, 2024, 45(5): 71-79.
- [2] 暴悦爽, 王金浩, 程雪婷, 等. 聚合电动汽车与碳捕集的虚拟电厂优化运营策略[J]. 全球能源互联网, 2025, 8(3): 299-309.
BAO Yueshuang, WANG Jinhao, CHENG Xueting, et al. Optimize operational strategies for virtual power plants integrating electric vehicles aggregation and carbon capture[J]. Journal of Global Energy Interconnection, 2025, 8(3): 299-309.
- [3] 仪忠凯, 侯朗博, 徐英, 等. 市场环境下灵活性资源虚拟电厂聚合调控关键技术综述[J]. 中国电力, 2024, 57(12): 82-96.
YI Zhongkai, HOU Langbo, XU Ying, et al. Aggregation and operation key technology of virtual power plant with flexible resources in electricity market environment: review[J]. Electric Power, 2024, 57(12): 82-96.
- [4] 梁博淼, 杨甲甲, 文福拴, 等. 兼顾公平与效率的聚合多元灵活资源虚拟电厂协同规划[J]. 电力系统自动化, 2025, 49(10): 67-76.
LIANG Bomiao, YANG Jiajia, WEN Fushuan, et al. Collaborative planning of virtual power plants aggregating multivariate and flexible resources and balancing fairness and efficiency[J]. Automation of Electric Power Systems, 2025, 49(10): 67-76.
- [5] 宋天琦, 吕志鹏, 宋振浩, 等. 虚拟电厂规模化灵活资源聚合调控框架研究与思考[J]. 中国电力, 2024, 57(1): 2-8.

- SONG Tianqi, LÜ Zhipeng, SONG Zhenhao, et al. Research and thinking on the aggregation and dispatching control framework of virtual power plant's large scale flexible resources[J]. *Electric Power*, 2024, 57(1): 2-8.
- [6] 孙秋野, 姚葭, 王一帆. 从虚拟电厂到真实电量: 虚拟电厂研究综述与展望[J]. *发电技术*, 2023, 44(5): 583-601.
- SUN Qiuye, YAO Jia, WANG Yifan. From virtual power plant to real electricity: summary and prospect of virtual power plant research[J]. *Power Generation Technology*, 2023, 44(5): 583-601.
- [7] GOUGH M, SANTOS S F, ALMEIDA A, et al. Blockchain-based transactive energy framework for connected virtual power plants[J]. *IEEE Transactions on Industry Applications*, 2021, 58(1): 986-995.
- [8] WANG S, WU W. Aggregate flexibility of virtual power plants with temporal coupling constraints[J]. *IEEE Transactions on Smart Grid*, 2021, 12(6): 5043-5051.
- [9] WANG T, HUANG Z, YING R, VALENCIA-CABRERA L. A low-carbon operation optimization method of ETG-RIES based on adaptive optimization spiking neural P systems[J]. *Protection and Control of Modern Power Systems*, 2024, 9(5): 162-177.
- [10] 张钧钊, 姜欣, 段世杰, 等. 虚拟电厂参与电-碳联合市场运行的竞价策略研究[J]. *电力系统保护与控制*, 2023, 51(11): 108-118.
- ZHANG Junzhao, JIANG Xin, DUAN Shijie, et al. Bidding strategy for a virtual power plant to participate in the power-carbon joint market[J]. *Power System Protection and Control*, 2023, 51(11): 108-118.
- [11] 彭超逸, 徐苏越, 顾慧杰, 等. 基于主从博弈的虚拟电厂参与多元竞争市场投标策略研究[J]. *电力系统保护与控制*, 2024, 52(7): 125-137.
- PENG Chaoyi, XU Suyue, GU Huijie, et al. Bidding strategy for a virtual power plant participating in a multiple competitive market based on the Stackelberg game[J]. *Power System Protection and Control*, 2024, 52(7): 125-137.
- [12] THE GRID WISE ARCHITECTURE COUNCIL. Grid wise transactive energy framework version 1.0[EB/OL]. [2021-12-18]. http://www.gridwiseac.org/pdfs/te_framework_report_pnnl-22946.pdf
- [13] CHEN T, SU W. Indirect customer-to-customer energy trading with reinforcement learning[J]. *IEEE Transactions on Smart Grid*, 2019, 10(4): 4338-4348.
- [14] 杨剑峰, 章剑光, 陈晓宇, 等. 支撑多微网端对端交易的配电系统协同优化调度模型[J]. *电力系统及其自动化学报*, 2024, 36(6): 92-100.
- YANG Janfeng, ZHANG Janguang, CHEN Xiaoyu, et al. Coordinated optimal dispatching model of power distribution system supporting multi-microgrid p2p transactions[J]. *Proceedings of the CSU-EPSA*, 2024, 36(6): 92-100.
- [15] YAN X, SONG M, et al. Peer-to-peer transactive energy trading of multiple microgrids considering renewable energy uncertainty[J]. *International Journal of Electrical Power and Energy Systems*, 2023, 152.
- [16] 钟荣豪, 张亚超, 朱蜀, 等. 基于电-碳点对点交易的虚拟电厂联盟与配电网协同低碳运行策略研究[J]. *电网技术*, 2024, 48(9): 3554-3563.
- ZHANG Ronghao, ZHANG Yachao, ZHU Shu, et al. Research on collaborative low-carbon operation strategy of electricity carbon peer-to-peer trading based virtual power plant union and power distribution network[J]. *Power System Technology*, 2024, 48(9): 3554-3563.
- [17] 彭超逸, 徐苏越, 顾慧杰, 等. 基于主从博弈的虚拟电厂参与多元竞争市场投标策略研究[J]. *电力系统保护与控制*, 2024, 52(7): 125-137.
- PENG Chaoyi, XU Suyue, GU Huijie, et al. Bidding strategy for a virtual power plant participating in a multiple competitive market based on the Stackelberg game[J]. *Power System Protection and Control*, 2024, 52(7): 125-137.
- [18] 刘雨佳, 樊艳芳, 郝俊伟, 等. 基于碱性电解槽宽功率率适应模型的风光氢热虚拟电厂容量配置与调度优化[J]. *电力系统保护与控制*, 2022, 50(10): 48-60.
- LIU Yujia, FAN Yanfang, HAO Junwei. Capacity configuration and optimal scheduling of a wind-photovoltaic-hydrogen-thermal virtual power plant based on a wide range power adaptation strategy for an alkaline electrolyzer[J]. *Power System Protection and Control*, 2022, 50(10): 48-60.
- [19] XUE G, WU C, NIU W, et al. Optimal scheduling of integrated power and gas energy systems considering demand response[C]// 2022 IEEE 6th Conference on Energy Internet and Energy System Integration (EI2), October 28-30, 2022, Chengdu, China: 2569-2573.
- [20] LIN X, LIN X, ZHONG W, et al. Predictive operation optimization of multi-energy virtual power plant considering behavior uncertainty of diverse stakeholders[J]. *Energy*, 2023, 280.
- [21] HADAYEGHPARAST S, FARSANGI A S, SHAYANFAR H. Day-ahead stochastic multi-objective economic/ emission operational scheduling of a large scale virtual power plant[J]. *Energy*, 2019, 172: 630-646.
- [22] YANG L, SUN Q, MA D, et al. Nash Q-learning based equilibrium transfer for integrated energy management

- game with WE-energy[J]. Neurocomputing, 2019.
- [23] 刘建行, 刘方. 基于深度强化学习的梯级水蓄风光互补系统优化调度策略研究[J]. 广东电力, 2024, 37(5): 10-22.
LIU Jianxing, LIU Fang. Research on optimized dispatching strategy of cascade hydropower-pumping-storage-wind-photovoltaic multi-energy complementary system based on deep reinforcement learning[J]. Guangdong Electric Power, 2024, 37(5): 10-22.
- [24] 孙秋野, 杨凌霄, 张化光. 智慧能源——人工智能技术在电力系统中的应用与展望[J]. 控制与决策, 2018, 33(5): 938-949.
SUN Qiuye, YANG Lingxiao, ZHANG Huaguang. Smart energy—applications and prospects of artificial intelligence technology in power system[J]. Control and Decision, 2018, 33(5): 938-949.
- [25] YIN X, LEI M. Jointly improving energy efficiency and smoothing power oscillations of integrated offshore wind and photovoltaic power: a deep reinforcement learning approach[J]. Protection and Control of Modern Power Systems, 2023, 8(2): 420-430.
- [26] ZHAO W, ZENG T, LIU Z, et al. Automatic generation control in a distributed power grid based on multi-step reinforcement learning[J]. Protection and Control of Modern Power Systems, 2024, 9(4): 39-50.
- [27] 陈嘉琛, 陈中, 李冰融, 等. 基于二阶随机动力学的多虚拟电厂自趋优能量管理策略[J]. 中国电机工程学报, 2024, 44(16): 6294-6306.
CHEN Jiachen, CHEN Zhong, LI Bingrong, et al. Energy management strategy for multi-virtual power plants with self-optimization based on second-order stochastic dynamics[J]. Proceedings of the CSEE, 2024, 44(16): 6294-6306.
- [28] 张继行, 张一, 王旭, 等. 基于多代理强化学习的多新型市场主体虚拟电厂博弈竞价及效益分配策略[J]. 电网技术, 2024, 48(5): 1980-1991.
ZHANG Jixing, ZHANG Yi, WANG Xu, et al. Game bidding and benefit allocation strategies for virtual power plants with multiple new market entities based on multi-agent reinforcement learning[J]. Power System Technology, 2024, 48(5): 1980-1991.
- [29] ARULKUMARAN A, DEISENROTH M P, et al. A brief survey of deep reinforcement learning[J]. IEEE Signal Processing Magazine, 2017(6): 26-38.
- [30] 李子晨, 夏杨红, 孙勇, 等. 考虑氢能长短周期储能特性的电氢综合能源系统容量配置方法[J]. 电网技术, 2025, 49(1): 12-21.
LI Zichen, XIA Yanghong, SUN Yong, et al. Optimal sizing of electricity-hydrogen integrated energy system considering multi-timescale operation of hydrogen storage system[J]. Power System Technology, 2025, 49(1): 12-21.
- [31] ZHANG N, SUN Q, YANG L, et al. Event-triggered distributed hybrid control scheme for the integrated energy system[J]. IEEE Transactions on Industrial Informatics, 2022, 18(2): 835-846.
- [32] 张苏涵, 顾伟, 陆帅, 等. 综合能源系统分析——从数值到解析(三): 混合时间尺度经济调度[J/OL]. 中国电机工程学报: 1-13[2024-12-29]. <https://link.cnki.net/urlid/11.2107.tm.20241202.1409.006>.
ZHANG Suhan, GU Wei, LU Shuai, et al. Integrated energy system analysis—from numerical to analytical (III): hybrid timescale economic dispatch[J/OL]. Proceedings of the CSEE: 1-13[2024-12-29]. <https://link.cnki.net/urlid/11.2107.tm.20241202.1409.006>

收稿日期: 2024-10-29; 修回日期: 2025-06-05

作者简介:

张宁(1993—), 男, 博士, 特聘副教授, 硕士生导师, 研究方向为综合能源系统建模与运行控制; E-mail: zhangning@ahu.edu.cn

杨凌霄(1992—), 女, 通信作者, 博士, 特聘副教授, 硕士生导师, 研究方向为人工智能及其在能源领域的应用研究; E-mail: yanglingxiao@ahu.edu.cn

李炫浓(2000—), 男, 助理工程师, 研究方向为电气系统智能控制等。E-mail: 1968918472@qq.com

(编辑 魏小丽)