

1-16-2023

A wind power dispatching strategy based on improved NashQ under multi-agent game

Hailin ZHENG

College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China

Zhenshan ZHU

College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China

Buying WEN

College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China

Zhimin WENG

College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China

Follow this and additional works at: <https://jepst.researchcommons.org/journal>

Recommended Citation

ZHENG, Hailin; ZHU, Zhenshan; WEN, Buying; and WENG, Zhimin (2023) "A wind power dispatching strategy based on improved NashQ under multi-agent game," *Journal of Electric Power Science and Technology*. Vol. 37: Iss. 6, Article 7.

DOI: 10.19781/j.issn.1673-9140.2022.06.007

Available at: <https://jepst.researchcommons.org/journal/vol37/iss6/7>

This Article is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology.

多主体博弈下基于改进 NashQ 算法的风电场调度策略

郑海林, 朱振山, 温步瀛, 翁智敏

(福州大学电气工程与自动化学院, 福建 福州 350108)

摘 要:针对可再生能源消纳与发电市场的博弈问题,研究不同场景中风电的调度策略,提出基于改进 NashQ 算法的风电调度策略模型。首先,在市场上博弈环境下建立风电优化调度模型,计及风电上网的预测偏差考核惩罚、风力发电经济效益与环境效益,考虑可再生能源的弃电限制,在这一基础上,对比风电独立运行、风—光、风—储联合运行下的风电调度策略;其次,采用 JS 散度优化各个智能体的学习率,提高多智能体强化学习的收敛效率;最后,在 Matlab 中搭建电网模型进行分析,仿真结果验证:改进 NashQ 方法相较于 NashQ、NETRL 算法的收敛速度有明显提升,风—车联合运行模式在多主体博弈下有较好吸引力。

关 键 词:改进 NashQ;多智能体强化学习;多主体博弈;风电调度;电动汽车充电站

DOI:10.19781/j.issn.1673-9140.2022.06.007 中图分类号:TM734 文章编号:1673-9140(2022)06-0062-11

A wind power dispatching strategy based on improved NashQ under multi-agent game

ZHENG Hailin, ZHU Zhenshan, WEN Buying, WENG Zhimin

(College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China)

Abstract: Aiming at the problem of renewable energy gaming in power generation market, this paper studies wind power dispatching strategies in different scenarios, and proposes an improved NashQ wind power dispatching strategy. Firstly, a wind power optimal dispatch model under a game environment is established. In the dispatch model, the punishment for wind power forecast error, the environmental and economic benefits of wind power, and the cost of the curtailment of renewable energy are all considered. On this basis, the dispatch strategies are compared for the independent wind power operation mode, wind-vehicle operation mode, and wind-storage joint operation mode. Secondly, The Jensen-Shannon divergence is introduced for the learning rate of the intelligent agents. The convergence efficiency of multi-agent reinforcement learning is then improved. Finally, a microgrid model is constructed in Matlab for simulation. It is shown that the improved NashQ method has a significantly higher convergence speed than the NashQ and NETRL algorithms, and the wind-vehicle joint operation model has a better performance in the market games.

Key words: improve NashQ; multi-agent reinforcement learning; multi-agent game; wind power dispatch; electric vehicle station

收稿日期:2021-08-02;修回日期:2021-11-19

基金项目:福建省教育厅中青年教师教育科研项目(JAT190043);福州大学科研启动项目(510901)

通信作者:朱振山(1989—),男,博士,讲师,主要从事电力系统运行与控制的研究;E-mail:zhenshan198@126.com

随着“碳达峰”与“碳中和”目标的提出,要实现中国能源结构的转变,可再生能源机组的装机必将逐步增长,未来风电也将大量接入电网。但风力发电的波动性、间歇性以及随机性为其大规模并网带来了极大的挑战^[1-2]。受政策补贴的影响,电网公司优先保证风电上网^[3-4],但随着 2016 年来中国陆上风电 4 类风区上网指导价连续 4 次下调,风电的政策补贴红利逐步减少,参与竞争性电力市场成为未来风电的发展趋势。因此,考虑多方投资主体博弈下的风电调度策略成为当今重点关注的研究方向之一。

近年来,为了平抑风电出力的波动性与不确定性,中国各省份陆续出台了一系列强制风电企业配置储能的政策,但储能高昂的配置成本、运行成本以及辅助服务市场尚未完全建成一定程度上影响了风电投资主体配置储能的积极性。此外,电动汽车(electric vehicle, EV)因其环保、经济等优势而被广泛应用^[5],极大地推动了电动汽车充电站和充电桩的建设^[6]。相比于短时间内为风电场配置大量成本高昂的储能,利用已有的电动汽车充电站资源参与风电场联合调度的思路更具有可行性^[7]。

目前关于风电优化调度的研究主要有 2 类:一类是从风电运行效益的角度设定目标函数,建立优化调度模型,从而优化风电运行,如:可再生能源弃电量最小化^[8-9]、风电波动与预测偏差的最小化^[10-11]、可再生能源的环境效益^[12]以及政策补贴^[13]等;另一类将风电视作电网的一部分,从电网运行效益的角度设定运行目标,从而优化风电调度策略,如:电网运行成本最小化^[14]与用户的用电成本最小化^[15]、电网的环境效益^[16]以及电网的电压稳定性^[17]等。但由于电网中的发电设施是由不同的投资主体建设的,将电网视为整体进行调度,未能充分考虑各投资主体的效益,不利于调动各方的积极性。从风电角度优化风电调度策略的文献仅考虑了风电场投资方效益的最大化,忽视了其他投资主体的效益,无法反映风电进入电力市场后的真实效益。从长远的角度来看,不利于风电场进入竞争性电力市场后的可持续发展。

考虑到风电机组出力的不确定性以及电网内各方投资主体决策的相互影响,无法通过简单的优化方法来获得最佳的运行策略。现有文献通过 Nash 均衡博弈与强化学习的结合,采用多智能体强化学

习来求解这一复杂问题。文献[18-19]引入多智能体强化学习方法来求解多智能体博弈问题,从而优化电网内多方主体的调度策略;文献[20]在这一基础上,通过强化学习 Q 矩阵的迁移从而提高算法的在线学习速度;文献[21]引入资格迹更新技术提高多智能体博弈和迁移学习的收敛速度。以上文献所采用的多智能体强化学习算法中,智能体的学习往往局限在自身的学习经验,对外界环境不敏感,使得收敛速度较慢,会产生较大的计算成本。

因此,针对现有研究的不足,本文提出一种多主体博弈环境下基于改进 NashQ 算法的风电调度策略,综合考虑风电场、火电厂、储能电站以及电动汽车充电站各方的效益,兼顾风电场的预测偏差考核、绿证收益与售电效益,构建多方博弈下风电场调度、风—储联合调度与风—车联合调度的优化模型。采用多智能体强化学习算法进行求解,并引入 JS (Jensen-Shannon) 散度对比 Nash 均衡分布与历史经验分布来优化 NashQ 算法中各智能体学习率,提升算法的求解效率,用以求解多主体序贯决策问题。此外,通过对算法中 Q 矩阵的训练,提升算法在线求解的速度,使得算法可以灵活适应不同博弈环境。算例仿真结果验证本文算法的求解性能以及风电场与电动汽车充电站联合调度策略的优越性。

1 多主体博弈建模

1.1 电网系统结构

电网系统结构主要由火电厂、风电场、储能电站以及电动汽车充电站和其他用电负荷等组成,本文所讨论的电网系统为输电网层级,其结构如图 1 所示。

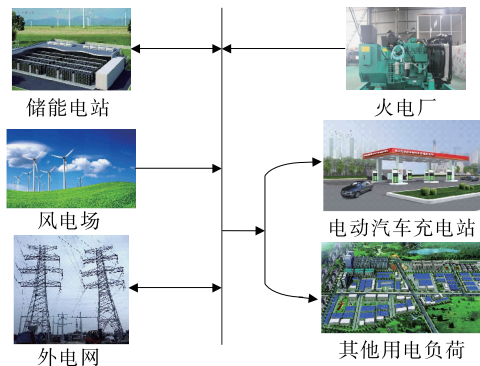


图 1 输电网系统结构

Figure 1 Schematic diagram of transmission network

1.2 博弈主体基本模型

1) 风电场主体。

为了保证电网内功率动态实时平衡,要求可再生能源电站的实际出力与预测出力曲线应基本一致,偏差的发电量或电量不足需要缴纳较高的罚金。风电场主体独立运行时目标函数为

$$\max C_w = \sum_{t=1}^T (S_{w,t} - C_{w,t}) - C_{ws} \quad (1)$$

$$S_{w,t} = (p_w + p_{gr}) \cdot (P_{w,t} - P_{aba,t}) \cdot \Delta t \quad (2)$$

$$C_{w,t} = ((P_{w,t} - P_{aba,t}) \cdot p_{wo} + |P_{w,t} - P_{aba,t} - P_{pre,t}| p_w \eta) \cdot \Delta t \quad (3)$$

式(1)~(3)中 $S_{w,t}$ 为风电场的售电效益;; $C_{w,t}$ 为风电场的成本; C_{ws} 为风电场分摊的为风电场分摊的辅助服务费用; $P_{w,t}$ 为风电场 t 时刻气象条件下最大出力; $P_{aba,t}$ 为 t 时刻风电场弃风功率; $(P_{w,t} - P_{aba,t}) \cdot p_{wo}$ 为风电场的运行成本; $|P_{w,t} - P_{aba,t} - P_{pre,t}| p_w \eta$ 为风电场实际出力与计划出力偏差的惩罚,风电场的实际出力等于在该时刻风速下的风电场最大出力减去风电场弃风功率; p_{wo} 为风电运维成本; $P_{pre,t}$ 为 t 时刻的预测风电出力; p_w 为风电上网的价格; p_{gr} 为风电的绿证价格; η 为偏差考核比例;风电的弃风比例应不超过 μ , 限制条件为 $P_{aba,t} \leq P_{w,t} \cdot \mu$ 。

2) 电动汽车充电站主体。

当电网内发电机组出力不足以满足负荷需求时,充电站通过补助 EV 用户,削减充电需求,满足电网供需平衡。充电站的目标函数为

$$\max S_{ev} = \sum_{t=1}^T P_{EV} \cdot N_{EV} \cdot (P_{ev,t} \cdot (p_{el} - \Delta p_t) - p_{evo}) \cdot \Delta t \quad (4)$$

$$p_{evo} = p_{lc} + p_{sc} \quad (5)$$

式(4)、(5)中 p_{evo} 为单个充电桩单时段下的运维成本,由人工成本 p_{lc} 与场地成本 p_{sc} 两部分组成; $P_{ev,t}$ 为 t 时刻 EV 用户调度意愿; N_{EV} 为充电站最大容纳量; P_{EV} 为每辆电动汽车的充电功率; p_{el} 为换电站参与调度的辅助服务补贴; Δp_t 为 t 时刻充电站对 EV 补助;EV 参与调控的意愿 $P_{ev,t}$ 基于文献[22]的调研数据,拟合调度意愿与补偿价格的关系为 $P_{ev,t} = -1.71\Delta p_t^5 + 6.93\Delta p_t^3 - 7.2\Delta p_t^2 + 2.97\Delta p_t$

3) 火电厂主体。

火电厂的收益一部分为直接售电的效益,另一部分则是火电机组响应电网内功率需求调整自身出力提供 AGC 服务的收益,目标函数为

$$\max C_{MT} = \sum_{t=1}^T (S_{MT,t} - C_{MT,t}) - C_{MTs} \quad (6)$$

式中 $S_{MT,t}$ 为火电机组的运行收益; $C_{MT,t}$ 为火电机组的运行成本; C_{MTs} 为火电机组分摊的辅助服务费用。

成本函数为

$$C_{MT,t} = \sum_{n=1}^N \left\{ (aP_{MTi,t}^2 + bP_{MTi,t} + c) \cdot \Delta t + (p_{up} \Delta P_{MTui,t} + p_{down} \Delta P_{MTdi,t}) + (p_{on}(h_{i,t} - h_{i,(t-1)}) + p_{off}(h_{i,(t-1)} - h_{i,t})) \right\} \quad (7)$$

式中 第 1 项为机组的发电成本;第 2 项为机组的爬坡成本^[21];第 3 项为机组的启停成; a 、 b 、 c 为微型火电系统发电成本系数; $P_{MTi,t}$ 为微型火电系统第 i 台机组 t 时刻发电量; $i = 1, 2, \dots, N$ 为微型火电机组数; $\Delta P_{MTui,t}$ 、 $\Delta P_{MTdi,t}$ 为微型火电机组上、下爬坡成本; p_{up} 、 p_{down} 为微型火电机组上、下爬坡量; $h_{i,t}$ 为第 i 台机组 t 时刻的启停状态,1 为启机,0 为停机; p_{on} 、 p_{off} 为启停机成本。

火电厂的收益为

$$S_{MT,t} = p_{MT} \cdot \sum_{i=1}^N P_{MTi,t} \cdot \Delta t + p_{AGC} \cdot \Delta P_{AGC,t} \quad (8)$$

式中 第 1 项为火电机组的售电效益;第 2 项为火电机组响应 AGC 服务补偿; p_{MT} 为火电机组上网电价; p_{AGC} 为 AGC 辅助服务单位电量补贴,当火电的出力调整与电网中功率偏差相反时, $\Delta P_{AGC,t}$ 取为正,即火电机组获取 AGC 响应补偿;反之则取为负,作为调节偏差的惩罚。

火电机组运行需要满足机组的爬坡约束以及机组最大、最小功率的约束,即

$$\begin{cases} P_{MTi,min} \leq P_{MTi,t} \leq P_{MTi,max} \\ 0 \leq \max\{\Delta P_{MTui,t}, \Delta P_{MTdi,t}\} \leq \Delta P_{MTi,max} \end{cases} \quad (9)$$

式中 $P_{MTi,min}$ 、 $P_{MTi,max}$ 分别为微型火电机组功率最小、最大值; $\Delta P_{MTi,max}$ 为机组爬坡功率上限。

4) 储能电站主体。

储能电站的目标函数为

$$\max C_{\text{ess}} = \sum_{t=1}^T (S_{\text{ess},t} - C_{\text{ess},t}) \quad (10)$$

储能电站的收益为

$$S_{\text{ess},t} = \begin{cases} (p_{g,t} \cdot P_t^{\text{dis}} - p_{g,t} \cdot P_t^{\text{cha}}) \cdot \Delta t + \\ p_{\text{as}} \cdot \Delta P_{\text{es},t} \cdot \Delta t \end{cases} \quad (11)$$

式中 第 1 项为储能峰谷套利;第 2 项为储能向电网提供调频服务补偿; P_t^{cha} 、 P_t^{dis} 为电池储能系统 t 时刻的充放电电量; $p_{g,t}$ 为 t 时刻电网内的电能价格; p_{as} 为储能提供辅助服务的补贴;当储能电池充放电与电网中功率偏差相反时, $\Delta P_{\text{es},t}$ 取为正,即储能电池获取调频补偿;反之取负值作为惩罚。

储能电站的成本为

$$C_{\text{ess},t} = \begin{cases} (P_t^{\text{cha}} \eta_{\text{cha}} + P_t^{\text{dis}} \eta_{\text{dis}}) \cdot p_{g,t} \cdot \Delta t + \\ \frac{i(1+d)^n}{L_{\text{max}}((1+d)^n - 1)} \cdot C_{\text{es}} \cdot \Delta t \end{cases} \quad (12)$$

式中 第 1 项储能充放电损耗成本;第 2 项为储能全寿命周期成本; η_{cha} 、 η_{dis} 为电池储能的充放电效率; d 为贴现率; n 为电池使用年限; C_{es} 为电池本体成本; L_{max} 为电池年平均最大使用次数。

电池储能系统运行需要满足充放电最大功率、荷电状态约束,约束条件为

$$\begin{cases} 0 \leq P_t^{\text{cha}} \leq P_{\text{max}}^{\text{cha}} \\ 0 \leq P_t^{\text{dis}} \leq P_{\text{max}}^{\text{dis}} \\ S_{\text{min}} \leq S_t \leq S_{\text{max}} \end{cases} \quad (13)$$

式中 $P_{\text{max}}^{\text{cha}}$ 、 $P_{\text{max}}^{\text{dis}}$ 为电池储能功率充、放功率的上限; S_t 为 t 时刻电池储能的荷电状态; S_{max} 、 S_{min} 分别为电池储能荷电状态的上限、下限。

5) 辅助服务费用分摊模型。

电网内的辅助服务费用由发电企业按发电量进行分摊^[23],即

$$C_{\text{ser}} = \sum_{t=1}^T (p_{\text{AGC}} \cdot \Delta P_{\text{AGC},t} + p_{\text{as}} \cdot \Delta P_{\text{es},t} \cdot \Delta t + S_{\text{ev}} + p_{\text{et}} |P_{\text{et},t}|) \quad (14)$$

式中 第 1 项为火电机组 AGC 服务补偿,第 2 项为储能调频服务补偿,第 3 项为 EV 用户调节需求响应补偿,第 4 项为联络线功率传输费用; $P_{\text{et},t}$ 为 t 时刻电网内功率不平衡引起的联络线上功率响应,本文设定联络线的传输容量约束为 2 MW; p_{et} 为 t 单位功率传输成本。

常规火电系统承担费用 C_{MTs} 与风电场承担费用 C_{ws} 为

$$\begin{cases} W = \sum_{i=1}^N \sum_{t=1}^T P_{\text{MTi},t} + \sum_{t=1}^T (P_{w,t} - P_{\text{aba},t}) \\ C_{\text{MTs}} = \frac{\sum_{i=1}^N \sum_{t=1}^T P_{\text{MTi},t}}{W} C_{\text{ser}} \\ C_{\text{ws}} = \frac{\sum_{t=1}^T (P_{w,t} - P_{\text{aba},t})}{W} C_{\text{ser}} \end{cases} \quad (15)$$

式中 W 为发电企业的总发电量。

1.3 多主体联合博弈

1) 多主体博弈模式。各主体基于某一时刻下各方的状态,决策下一时刻动作,且各主体的策略受到其他主体影响,为混合策略博弈。

模式 1 风电场与电动汽车充电站联合参与多主体博弈,参与博弈的主体有火电站、储能电站以及风—车联合运行主体。风—车联合运行的优势在于当风电出力小于预测出力时,可以通过减少 EV 充电需求来减少风电的预测偏差惩罚。

模式 2 风电场配置对应的储能联合参与多主体博弈,参与博弈的主体有火电站、储能电站以及风—储联合运行主体。当风电出力小于预测出力时,电池储能对外放电;反之则对储能电池充电。

2) 风—车联合运行模型。风—车联合运行效益为

$$\begin{aligned} \max C_{\text{we}} = & \sum_{t=1}^T (S_{w,t} (P_{w',t}) - C_{w,t} \cdot \\ & (P_{w',t})) + \sum_{t=1}^T S_{\text{ev}} (P_{\text{ev}',t}) \end{aligned} \quad (16)$$

式中 第 1 项为 EV 减少充电需求后风电场等效出力下的运行效益,第 2 项为电动汽车充电站减少风电场预测偏差后的调节补偿; $P_{w',t}$ 为风—车联合运行下风电场实际出力, $P_{w',t} = P_{w,t} + P_{\text{evl},t}$, 其中 $P_{\text{evl},t}$ 为 EV 减少充电需求用于减少风电出力偏差部分; $P_{\text{ev}'}$ 为风—车联合运行下充电站的实际调节电量, $P_{\text{ev}',t} = P_{\text{EV}} \cdot N \cdot P_{\text{ev},t} - P_{\text{evl},t}$ 。参与博弈的主体效用函数包括式(6)、(10)、(16)。

3) 风—储联合运行模型。风—储联合运行效益为

$$\max C_{ws} = \sum_{t=1}^T \left\{ \begin{aligned} & (S_{w,t}(P_{w'',t}) - C_{w,t}(P_{w'',t})) + \\ & S_{ess,t}(P_{cha',t}, P_{dis',t}) - \\ & C_{ess,t}(P_{cha,t}, P_{dis,t}) \end{aligned} \right\} \quad (17)$$

式中 第 1 项为储能参与减少风电预测偏差后风电场等效出力下的运行效益,第 2 项为电池储能减少风电场预测偏差后的峰谷套利与调频补偿带来的收益,第 3 项为电池储能充放电产生的成本; $P_{w'',t}$ 为风—储联合运行下风电场实际出力, $P_{w'',t} = P_{w,t} + P_{cha,t}\eta_{cha} + P_{dis,t}/\eta_{dis}$, 其中 $P_{cha,t}$ 、 $P_{dis,t}$ 为电池储能用于满足风电出力偏差部分; $P_{es',t} = \{P_{cha',t}; P_{dis',t}\} = \{P_{cha,t} - P_{cha,t}; P_{dis,t} - P_{dis,t}\}$, $\{P_{cha',t}; P_{dis',t}\}$ 为风—储联合运行下储能电池的实际充放电电量。参与博弈的主体效用函数包括式(6)、(10)、(17)。

2 改进 NashQ 算法

2.1 强化学习

强化学习是学习状态与行为之间的映射关系,通过感知环境状态以及奖励学习和决策的过程。 Q 学习是一种普遍应用的强化学习算法,其递归方程为

$$Q(s, a)^{t+1} = (1 - \alpha)Q(s, a)^t + \alpha(r + \gamma \max_{a' \in A} Q(s', a')^t) \quad (18)$$

式中 α 为学习率; γ 为折扣率; r 为在执行动作 a 的回报; $Q(s, a)^t$ 为状态 s 下执行动作 a 的 Q 值; s' 为执行动作 a 后的状态; $\max_{a' \in A} Q(s', a')^t$ 为状态 s' 下的最优策略; A 为智能体的动作空间。

2.2 NashQ 算法

当博弈中其他智能体的策略均给定时,智能体不再改变自身策略,即为 Nash 均衡,有

$$\nu_i(S_i, \pi_1^*, \dots, \pi_i^*, \dots, \pi_n^*) \geq \nu_i(S_i, \pi_1^*, \dots, \pi_i^j, \dots, \pi_n^*) \quad (19)$$

式中 $\nu_i(S_i, \pi_1^*, \dots, \pi_i^*, \dots, \pi_n^*)$ 为第 i 个智能体采取策略 π_i^* 时的收益, $\pi_1^*, \dots, \pi_i^*, \dots, \pi_n^*$ 为智能体 $i=1, 2, \dots, n$ 的 Nash 均衡策略; n 为参与博弈的智能体个数。

引入多智能体博弈环境, Q 值的值函数及迭代

公式为

$$\begin{aligned} Q_i(s, a_1, a_2, \dots, a_n)^{t+1} &= (1 - \alpha) \cdot \\ Q_i(s, a_1, a_2, \dots, a_n)^t &+ \alpha \cdot \\ (r_i^t + \gamma Q_{Nash,i}(s')^t) \end{aligned} \quad (20)$$

式中 $Q_i(s, a_1, a_2, \dots, a_n)^t$ 为 t 时刻状态 s 下各智能体动作为 $a_1, a_2, \dots, a_n$ 的 Q 值; $Q_{Nash,i}(s')^t$ 为各个智能体 t 时刻动作后进入状态 s' 的 Nash 均衡解, $Q_{Nash,i}(s') = v_i(s', \pi_1^*, \dots, \pi_n^*)$ 。

2.3 改进 NashQ 算法

学习率会影响智能体的学习速度,当学习率取值较大时,智能体对于新尝试结果的占比越高,对于历史经验遗忘率也就越高;当学习率取值较小时,智能体接受新尝试结果比例较小,使得学习效率下降。而大多数研究强化学习的文献往往将学习率取为定值,把较优的学习经验与较差的学习经验混合,大大降低了学习速度,影响智能体的收敛速度。

KL 散度(kullback-leibler divergence)又称为相对熵,是 2 个概率分布 P 和 Q 差别的非对称性的度量,用于衡量 2 种分布的相似度。由于 KL 散度是不对称的且不满足三角不等式,故

$$D_{KL}(P, Q) \neq D_{KL}(Q, P) \quad (21)$$

$$D_{KL}(P, Q) > D_{KL}(Q, S) + D_{KL}(S, Q) \quad (22)$$

JS 散度是 KL 散度基础上的变体,可以有效地解决 KL 散度存在的问题。当 2 个概率分布较远时,KL 散度没有意义,而 JS 散度仍可以衡量 2 个概率分布的相似度。

引入 JS 散度优化各智能体的学习率后的 NashQ 算法流程如图 2 所示。引入 JS 散度优化各智能体的学习率 $\alpha_{i,t}$, 即

$$\begin{aligned} \alpha_{i,t} = D_{JS}(\hat{P}_i \parallel P_i) &= \frac{1}{2} [D_{KL}(\hat{P}_i \parallel \frac{P_i + \hat{P}_i}{2}) + \\ D_{KL}(P_i \parallel \frac{P_i + \hat{P}_i}{2})] \end{aligned} \quad (23)$$

$$P_i(a_j | s_j) = \frac{\exp(Q_i(s_j, a_j))}{\sum_{a_j \in A_i} \exp(Q_i(s_j, a_j))} \quad (24)$$

$$\hat{P}_i(a_j | s_j) = \frac{\pi_i^*(a_j | s_j)}{\sum_{a_j \in A_i} \pi_i^*(a_j | s_j)} \quad (25)$$

式中 P_i 为智能体的历史经验概率分布; $\hat{P}_i$ 为智能体 Nash 均衡概率分布; $D_{JS}(\hat{P}_i \parallel P_i)$ 为历史经

验概率分布与混合策略分布的 JS 散度; $Q_i(s_j, a_j)$ 为智能体 i 面对状态 s 时选择动作 a 的 Q 值; $\pi_i^*(a_j | s_j)$ 为智能体 i 的 Nash 均衡的策略。

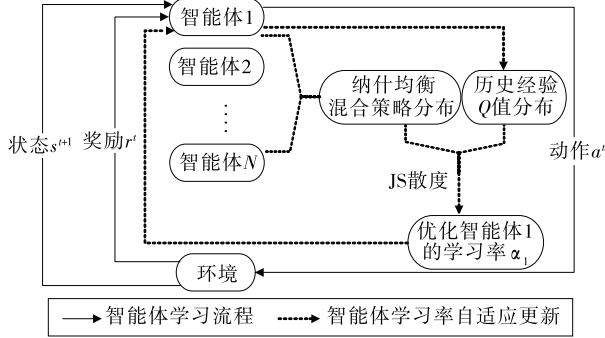


图 2 改进 NashQ 算法流程

Figure 2 The process of improved NashQ algorithm

3 基于改进 NashQ 算法的风电调度策略

3.1 状态空间

在多主体博弈下,电网中各个主体状态空间参数包括风电实际最大出力 $P_{w,t}$ 、负荷电量 $P_{load,t}$ 、风电预测出力 $P_{pre,t}$ 、储能荷电状态 S_t 、机组 i 的启停状态 $h_{i,t}$ 与前一时刻出力 $P_{MTi,(t-1)}$ 以及 t 时刻的电动汽车可调度数量 $N_{ev,t}$; t 时刻风电场的状态集为 $S_w^t = \{P_{w,t}, P_{pre,t}\}$ 、电动汽车充电站的状态集为 $S_{ev}^t = \{P_{g,t}, N_{ev,t}\}$ 、火电厂的状态集为 $S_m^t = \{h_{1,t}, h_{2,t}, P_{MT1,(t-1)}, P_{MT2,(t-1)}\}$ 、储能电站的状态集为 $S_{es}^t = \{P_{g,t}, S_t\}$ 。以上变量均为连续变量,为了构建改进 NashQ 算法中的 Q 矩阵,将所有状态变量离散化。风电场、火电厂、储能电站以及电动汽车充电站各主体的状态空间分别划分为 96、4、9、1 个离散空间。

3.2 动作空间

风电场的动作变量包括 t 时刻的风电弃风电量 $P_{aba,t}$;电动汽车充电站的动作变量为用户补偿价格 Δp_t ;火电系统的动作变量由机组 1、2 的出力 $P_{MT1,t}, P_{MT2,t}$ 组成;储能的充、放电量分别为 P_t^{cha}, P_t^{dis} 。风电场、火电厂、电动汽车充电站以及储能电池的动作集分别为 $A_w^t = \{P_{aba,t}\}$ 、 $A_m^t = \{P_t^{cha}, P_t^{dis}\}$ 、 $A_{ev}^t = \{c_{ev,t}\}$ 、 $A_{es}^t = \{P_t^{cha}, P_t^{dis}\}$ 。以上动作变量同样离散化,各主体的动作空间分别划分为 5、9、5、3 个

离散空间。

3.3 算法流程

改进 NashQ 算法中采用 ϵ -greedy 贪婪算法对动作空间进行探索,动作的探索为

$$a_i = \begin{cases} \arg \max_{a \in A} Q_i(s, a), \epsilon < 1 - \epsilon_0 \\ a_{i,rand}, & \text{其他} \end{cases} \quad (26)$$

式中 $\epsilon_0 \in [0, 1]$ 为贪婪搜索算法的探索率,取较小正数; $a_{i,rand}$ 为智能体 i 随机选择的动作。

判断多智能体强化学习过程中各个智能体的收益是否趋于收敛,收敛判据为

$$\Delta Q_{i,t} = \frac{\|Q_{i,t} - Q_{i,t-1}\|_2}{\|Q_{i,t}\|_2} \leq \sigma \quad (27)$$

其中, σ 为较小正数,本文取为 0.015; $Q_{i,t}$ 为第 i 个智能体第 t 次循环得到的 Q 值表。若各个智能体 Q 矩阵均收敛则结束学习,输出最优策略;反之则继续循环。

基于改进 NashQ 的风电调度策略的流程如图 3 所示。

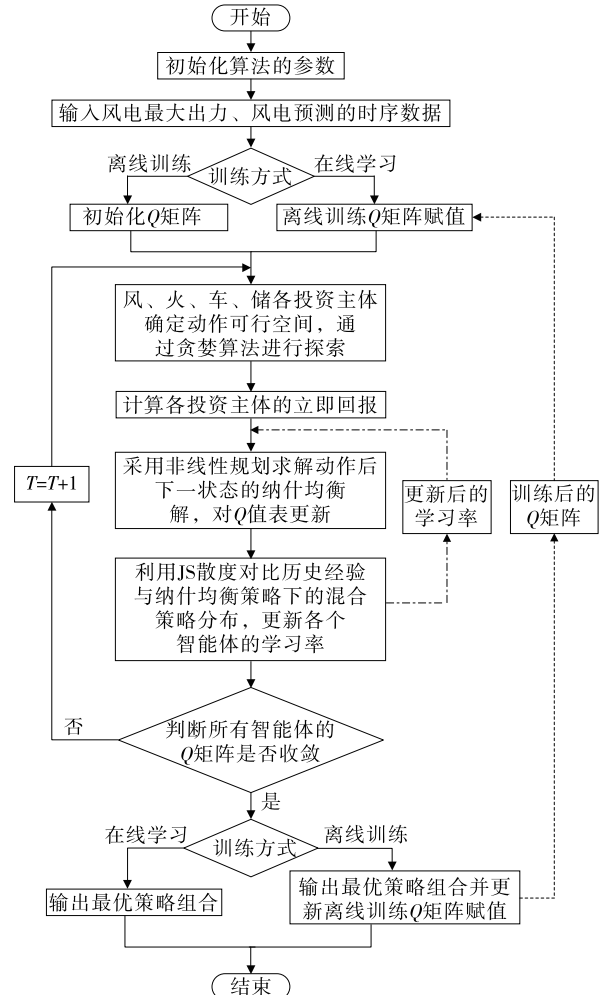


图 3 风电调度策略的流程

Figure 3 The process of wind power dispatching strategy

4 算例分析

4.1 算例概况

本文以某个风电场所在地区的电网为例,由 250 MW 风电场、8 MW·h 储能电站、2 台火电机组的火电厂以及 1 个电动汽车充电站组成,辅助服务补偿价格、电网分时电价以及各设备参数分别如表 1~3 所示;算法参数设置如表 4 所示;算法典型日风电、预测以及用户负荷曲线如图 4 所示。

表 1 辅助服务补偿价格

Table 1 Auxiliary service compensation price			
设备	补偿价格/ (元/MW)	设备	补偿价格/ (元/MW)
火电机组 AGC	80	储能电站	200
风电场	120	电动汽车充电站	180

表 2 电网分时电价

Table 2 Time-of-use tariff 元/(kW·h)	
分时电价时段	电价
谷时段(00:00—07:00、23:00—24:00)	0.17
平时段(08:00—10:00、16:00—18:00、22:00—23:00)	0.49
峰时段(11:00—15:00、19:00—21:00)	0.83

表 3 设备参数

Table 3 Equipment parameters										
火电机组 AGC										
设备 编号	最大出 力/MW	最小出 力/MW	爬坡速率/ (MW/15 min)	发电成本系数			启机 成本/元	停机 成本/元	爬坡成本/(元/MW)	
				<i>a</i>	<i>b</i>	<i>c</i>			上	下
1	50	20	5	0.003 29	27.62	1 340	2 880	0	50	10
2	36	12	4	0.002 41	29.13	1 139	2 340	0	50	10
储能电站										
电池成本/ (元/(kW·h))		其他成本/ (元/(kW·h))		充电 效率/%	放电效率/%	贴现率/%	电池 寿命/a	储能电站 寿命周期/a	储能电池平均 充放电次数	
800		1 200		93.5	94.4	6	5	20	0.003 29	
风电场						电动汽车充电站				
单台风电机组 装机容量/MW		风电机组 总量/台	运维成本/ (元/(MW·h))	风电上网价格/ (元/(MW·h))	风电最大弃风 比例/%	单辆电动汽车电 池容量/(kW·h)		充电桩/个	充电桩充电 功率/kW	
2		125	29.6	570	2	50		40	22.5	

表 4 算法参数

Table 4 Algorithm parameters

智能体初始学习率 α	折扣因子 γ	资格迹衰减因子 λ
0.01	0.8	0.6

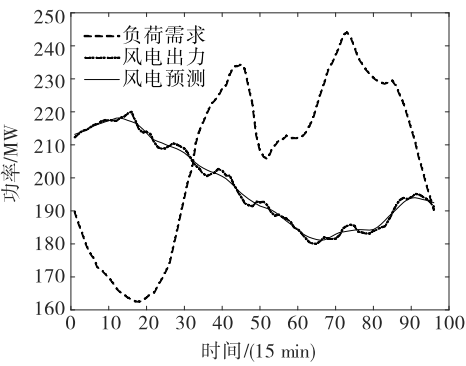


图 4 典型日风电、预测以及负荷曲线

Figure 4 Typical daily curve of wind power, forecast and load

将可再生能源功能状态按照可再生能源出力/电网内负荷需求之比、风电最大出力上限以及预测出力偏差量对其定义,划分为 1 天 96 个状态,火电机组包含启机与停机 2 种状态、储能包含充电/闲置/放电 3 种状态。火电厂、风电场、储能电站以及充电站的动作离散为 9、6、6、11 个离散空间。

4.2 Nash 均衡证明及改进 Nash 结果分析

NashQ 算法收敛的具体证明过程可参考文献[24]。为了保证 NashQ 算法能够找到 Nash 均衡解,则对于任一时间与状态下,各个智能体均能寻找到一个全局最优点或者鞍点,可用于更新 Q 矩阵。由于在电网中风电出力具有不确定性,无法保证每个时刻均可以找到全局最优点,在部分时刻仅存在鞍点,但鞍点的数目往往不止一个。

1)传统的 NashQ 算法在存在鞍点的算例中容

易出现运算结果在几个鞍点中徘徊选择,容易导致最终结果出现数个结果的情况。

2)文献[21]中引入资格迹,提出 NETRL 算法,通过智能体自身对于某一状态、动作的访问次数来优化 Q 值表的更新,虽然提高了智能体自身经验的感知,但忽视了智能体对于外部环境的感知,学习的效果有所提升但提升的有限。

3)本文提出的改进 NashQ 算法则可以较好地避免这一结果,在 NETRL 算法的基础上,通过 JS 散度对比智能体的自身经验与外部环境,提升智能体对外部环境的感知。

本文采用 Matlab2020a 软件进行编程求解,在 Intel i5-6300HQ(主频为 2.30 GHz)、内存 16 GB 的计算机上运行。利用风电发电系统智能体的收敛判据对比 3 种算法的收敛结果,收敛所需时间如表 5 所示,可以看出,本文所提出的改进 NashQ 算法求解所需时间为 NashQ 算法的 37.3%,也是文献[21]中 NETRL 算法的 46.9%,从而验证了本文所提出的改进 NashQ 算法的快速收敛性。收敛曲线如图 5 所示。

表 5 3 个算法收敛所需时间

Table 5 Converge time of the three algorithms			
算法	平均收敛次数	平均收敛时间/s	平均单次时间/s
NashQ	6 000	1 360.80	0.226 8
NETRL	4 800	1 080.96	0.225 2
本文算法	2 800	507.25	0.181 1

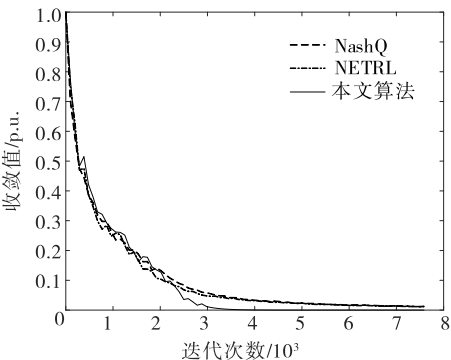


图 5 不同算法的收敛曲线

Figure 5 Convergence curves of different algorithms

4.3 仿真结果讨论分析

为对比风电不同运行方式下的经济效益,设定 3 种风电运行场景,并针对高比例可再生能源电网

系统运行模式给出优化风电的调度策略:①电网中风电场、电动汽车充电站、火电厂以及储能电站系统共同博弈;②风力发电系统与电动汽车充电站系统联合参与电网中的博弈;③电动汽车充电站转化为同等容量储能电池,风储联合参与电网中博弈。

3 种场景下风电场的弃风电量与减少的偏差考核如表 6 所示,风—车联合运行、风—储联合运行均能减少风电场的偏差考核,其中风—储联合运行下有着显著的优势。对比风电独立运行下弃风电量,风—车联合运行下减少了 43.4%,风—储联合运行减少了 44.7%。可以看出,风—储联合运行可以提高风力资源的利用率。

表 6 3 种场景下优化运行结果

Table 6 Optimization results under three scenarios		
场景	弃风电量/(MW·h)	减少偏差电量/MW
1	4.795 1	8.796 1
2	2.687 5	10.047 1
3	2.815 8	11.263 2

各个投资主体的经济性如表 7 所示,对比可以看出,风—车联合运行的经济效益优于风电场独立参与电网系统中多主体博弈的经济效益,提高收益 20 775.1 元。由于充电站系统的补贴成本比风力发电系统的偏差考核惩罚成本低,同时还能为其带来额外的售电效益。因此,风力发电系统与电动汽车充电站系统的联合运行具有实际意义。

表 7 3 种场景下各投资主体经济性

Table 7 Economics of investment subject under three scenarios				
场景	投资主体经济效益/元			
	火电厂	储能电站	风电场	充电站
1	185 700	581.92	2 197 000	1 224.9
2	172 200	711.4	2 219 000(风—车联合运行)	
3	197 600	630.4	2 180 000(风—储联合运行)	

3 种场景下风电场调度策略分别如图 6~8 所示。风—储联合运行由于储能电池系统较高昂的配置成本与运行成本,其经济效益甚至低于场景 1 中风电场独立运行的经济效益,在不考虑电厂侧储能参与辅助服务市场时,该风电场配置储能的运行模式难以吸引风力发电投资者为风电场配置储能。

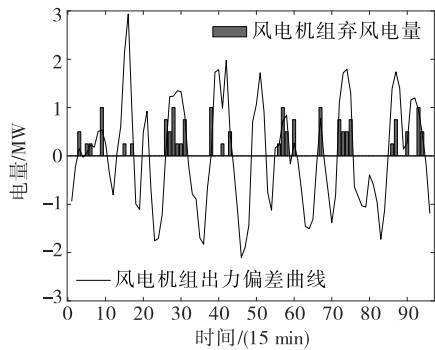


图 6 风电运行调度策略(场景 1)

Figure 6 The operation scheduling strategy of wind power(Scenario 1)

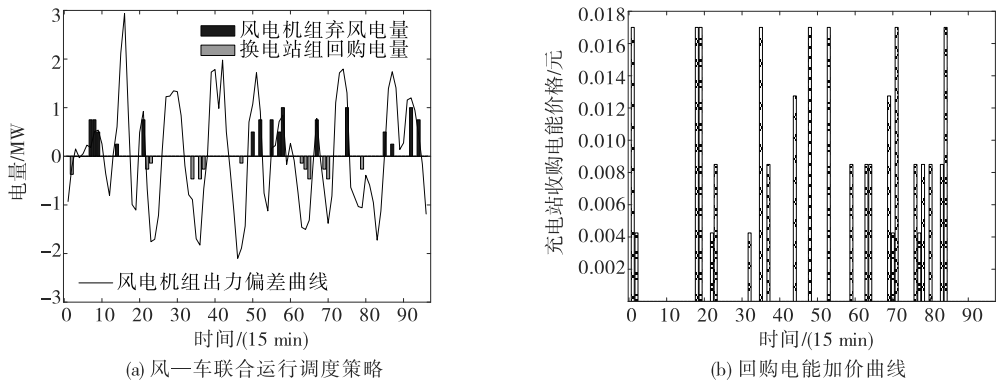


图 7 风—车联合运行调度策略及回购电能加价曲线(场景 2)

Figure 7 The operation scheduling strategy for the wind-mill(Scenario 2)

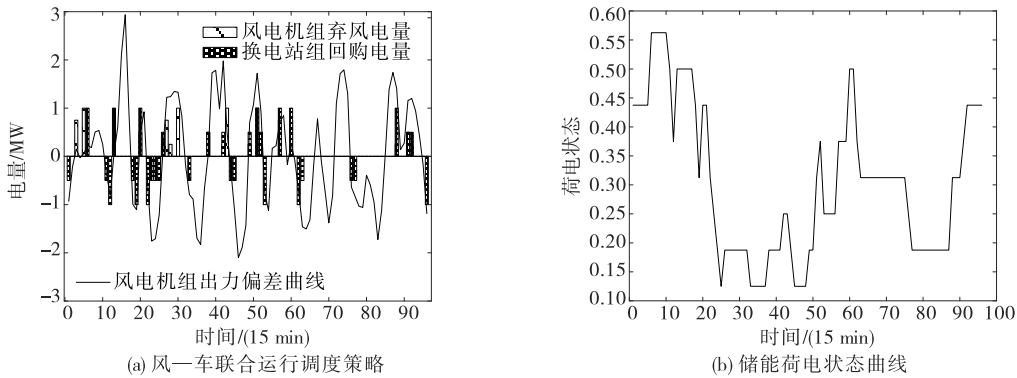


图 8 风—储联合运行调度策略及储能荷电状态曲线(场景 3)

Figure 8 The operation scheduling strategy for the wind-storage(Scenario 3)

4.4 在线仿真结果分析

本文基于 2020 年每个月抽取 3 d 的数据作为离线训练数据,训练集内选取 6 组数据($X_1 \sim X_6$),训练集外选取 6 组数据($C_1 \sim C_6$),检测本文算法的在线决策能力。在线学习与离线训练的对比如图 9 所示,可以看出,训练、非训练集数据在线学习所需平均时间分别为 319.21、427.07 s,相较于离线训练的所需平均时间(507.25 s)分别减少了 37.07%、15.81%。因此,通过 Q 值表的离线训练可以显著提升在线学习的求解效率。

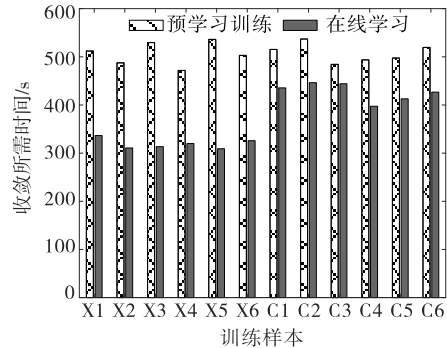


图 9 离线训练与在线学习的平均收敛时间对比

Figure 9 Comparison of average convergence time between off-line training and online learning

5 结语

本文以风电场作为研究对象,提出一种多主体博弈下基于改进 NashQ 算法的风电调度策略,分析对比了不同场景下的风力发电系统的效益。

1)综合考虑电网系统内不同投资主体的利益诉求,兼顾风电场的偏差考核、绿证效益以及售电效益,充分挖掘风电场、火电厂、储能电站与电动汽车充电站的调节能力,最大化风电场经济效益;

2)提出的改进 NashQ 算法与 NashQ、NETRL 算法对比,收敛所需时间分别缩短了 62.7%、53.1%,验证了改进 NashQ 算法的有效性;同时,采用离线训练方式训练 Q 值表可以显著提高在线学习的收敛效率,缩短收敛所需时间;

3)仿真结果表明,风—储、风—车的联合运行相较于风电池独立运行可以分别减少 44.7%、43.4%的弃风电量,但由于储能的高昂配置、运行成本,风—储联合运行经济效益较差;而风—车联合运行在减少弃风电量的同时,还提升了二者的经济效益,对于风电投资方与充电站投资方具有一定的吸引力。

后续研究会进一步考虑辅助服务市场中有偿提供辅助服务对于风电调度经济型的影响,为参与竞争电力市场的风电运营模式提供参考。

参考文献:

[1] 李灵聪,施嘉昊,武庆,等.含风电接入的电力系统脆弱性分析与量化评估[J].电网与清洁能源,2020,36(1):46-54.

LI Jiongcong,SHI Jiahao,WU Qing,et al.Vulnerability analysis and quantitative evaluation of the power system containing wind power[J].Power System and Clean Energy,2020,36(1):46-54.

[2] 任冲,柯贤波,樊国伟,等.大规模风电直流送出系统过电压抑制措施及控制方案优化研究[J].高压电器,2020,56(5):163-174.

REN Chong,KE Xianbo,FAN Guowei,et al.Transient voltage stabilization and control optimization for large-scale wind power UHV DC transmission system[J].High Voltage Apparatus,2020,56(5):163-174.

[3] 崔杨,姜涛,仲悟之,等.电动汽车与热泵促进风电消纳的区域综合能源系统经济调度方法[J].电力自动化设

备,2021,41(2):1-7.

CUI Yang,JIANG Tao,ZHONG Wuzhi,et al.Economic dispatch approach of RIES for electric vehicle and heat pump to promote wind power accommodation[J].Electric Power Automation Equipment,2021,41(2):1-7.

- [4] KUSAKANA K.Optimal scheduling for distributed hybrid system with pumped hydro storage[J].Energy Conversion and Management,2016,111:253-260.
- [5] 陈岩,靳伟,王文宾,等.基于电动汽车分群的“风—网—车”联合消纳调度策略[J].中国电力,2021,54(4):107-118.
- CHEN Yan,JIN Wei,WANG Wenbin,et al.Scheduling strategy for “wind-network-vehicle” joint accommodation based on electric vehicle clustering[J].Electric Power,2021,54(4):107-118.
- [6] 欧名勇,陈仲伟,谭玉东,等.基于峰谷分时电价引导下的电动汽车充电负荷优化[J].电力科学与技术学报,2020,35(5):54-59.
- OU Mingyong,CHEN Zhongwei,TAN Yudong,et al.Optimization of electric vehicle charging load based on peak-to-valley time-of-use electricity price[J].Journal of Electric Power Science and Technology,2020,35(5):54-59.
- [7] 杨军伟,杜露露,刘夏,等.高风电渗透率下考虑需求侧管理策略的智能微电网调度方法[J].智慧电力,2021,49(3):32-39.
- YANG Junwei,DU Lulu,LIU Xia,et al.Dispatching method of smart microgrid considering demand-side management strategy under high wind power penetration rate[J].Smart Power,2021,49(3):32-39.
- [8] 陈玉敏,赵冬梅.含风电的多种形式储能协调调度多目标优化模型[J].电测与仪表,2020,57(23):71-78.
- CHEN Yumin,ZHAO Dongmei.Multi-objective optimization model for multi-form energy storage coordinated scheduling with wind power[J].Electrical Measurement & Instrumentation,2020,57(23):71-78.
- [9] 于德鳌,李慧,刘思嘉,等.基于改进萤火虫算法的含风电系统环境经济调度[J].电力科学与技术学报,2020,35(2):84-92.
- YU Deao,LI Hui,LIU Sijia,et al.Environmental economic dispatch of power system integrated with wind power based on an improved glowworm swarm optimization[J].Journal of Electric Power Science and Technology,2020,35(2):84-92.

- [10] 周任军,尹权,康信文,等.风电并入冷热电联供系统的优化运行控制[J].电力科学与技术学报,2014,29(3):45-51.
ZHOU Renjun, YIN Quan, KANG Xinwen, et al. Optimization operation control for wind power access to CCHP system[J]. Journal of Electric Power Science and Technology, 2014, 29(3): 45-51.
- [11] 任冲,程林,樊国伟,等.考虑风电预测及电压分布的风电场无功设备协调控制优化研究[J].电力电容器与无功补偿,2020,41(2):175-180+186.
REN Chong, CHENG Lin, FAN Guowei, et al. Study on coordination control optimization of reactive power equipment of wind farm considering wind power prediction and voltage distribution[J]. Power Capacitor & Reactive Power Compensation, 2020, 41(2): 175-180+186.
- [12] 谭鑫,徐秋磊,王洁雨,等.基于改进共原点灰色聚类的清洁能源消纳综合效益评估[J].广东电力,2021,34(2):28-35.
TAN Xin, XU Qiulei, WANG Jieyu, et al. Comprehensive benefit evaluation of clean energy consumption based on improved common origin gray clustering evaluation method[J]. Guangdong Electric Power, 2021, 34(2): 28-35.
- [13] 赖春羊,马光文,谢航,等.基于综合效益指标体系的风水互补系统经济调度[J].电网技术,2021,45(11):4319-4328.
LAI Chunyang, MA Guangwen, XIE Hang, et al. Economic dispatching of wind and hydropower complementary system based on comprehensive benefit index system[J]. Power System Technology, 2021, 45(11): 4319-4328.
- [14] HAKIMI S M, HASANKHANI A. Stochastic energy management of smart microgrid with intermittent renewable energy resources in electricity market[J]. Energy, 2020, 219(15): 119668.
- [15] ZHENG S L, SUN Y, LI B, et al. Incentive-based integrated demand response for multiple energy carriers under complex uncertainties and double coupling effects[J]. Applied Energy, 2020, 283(1): 116254.
- [16] ABU-ELZAIT S, PARKIN R. Economic and environmental advantages of renewable-based microgrids over conventional microgrids[C]// IEEE Green Technologies Conference, Lafayette, LA, United States, 2019.
- [17] XU J Y, WU G N, BAI Q Y. Research on the voltage stability of the microgrid[J]. Applied Mechanics & Materials, 2014, 521: 256-259.
- [18] 刘洪,李吉峰,葛少云,等.基于多主体博弈与强化学习的并网型综合能源微网协调调度[J].电力系统自动化,2019,43(1):40-50.
LIU Hong, LI Jifeng, GE Shaoyun, et al. Coordinated scheduling of grid-connected integrated energy microgrid based on multi-agent game and reinforcement learning[J]. Automation of Electric Power Systems, 2019, 43(1): 40-50.
- [19] SINGH V P, KISHOR N, SAMUEL P. Distributed multi-agent system based load frequency control for multi-area power system in smart grid[J]. IEEE Transactions on Industrial Electronics, 2017, 64(6): 5151-5160.
- [20] 陈艺璇,张孝顺,郭乐欣,等.基于多智能体迁移强化学习算法的电力系统最优碳—能复合流求解[J].高电压技术,2019,45(3):863-872.
CHEN Yixuan, ZHANG Xiaoshun, GUO Lexin, et al. Optimal carbon-energy combined flow in power system based on multi-agent transfer reinforcement learning[J]. High Voltage Engineering, 2019, 45(3): 863-872.
- [21] 李宏仲,王磊,林冬,等.多主体参与可再生能源消纳的Nash博弈模型及其迁移强化学习求解[J].中国电机工程学报,2019,39(14):4135-4150.
LI Hongzhong, WANG Lei, LIN Dong, et al. A Nash game model of multi-agent participation in renewable energy consumption and the solving method via transfer reinforcement learning[J]. Proceedings of the CSEE, 2019, 39(14): 4135-4150.
- [22] 罗敏,赵伟,林国营,等.基于电网峰谷分时电价联动的电动汽车有序充电电价研究[J].电器与能效管理技术,2015(24):78-82+92.
LUO Min, ZHAO Wei, LIN Guoying, et al. Research on coordinated charging price of electric vehicle based on users' price sensitivity[J]. Electrical & Energy Management Technology, 2015(24): 78-82+92.
- [23] 国家能源局南方监管局.南方区域发电厂并网运行管理实施细则(2017版)[EB/OL]. https://wenku.so.com/d/8286a30014d90ad6159e8fbba3f6ad73?src=www_rec, 2019-08-22.
- [24] HU J L, WELLMAN M P. Nash Q-Learning for general-sum stochastic games[J]. Journal of Machine Learning Research, 2004, 4(6): 1039-1069.