

计及碳排放的电动汽车充电站优化定价策略

尹力¹, 盛俊杰¹, 袁杰¹, 冯燕钧²

(1. 国网湖北省电力有限公司 武汉供电公司, 湖北 武汉 430000;

2. 东南大学 电气工程学院, 江苏 南京 210096)

摘要:在碳中和背景下,提出了一种计及碳排放的电动汽车充电站优化定价策略。首先,构建电动汽车用户价格响应特性模型;其次,建立以碳排放最低、负荷峰谷差最小以及充电站收益最高为目标的定价优化模型并将其转化为马尔科夫决策过程;然后,提出一种基于时间差分误差的改进近端策略优化算法,以提高算法的效率与稳定性;最后,通过算例分析表明,所提定价策略能够降低配电网负荷峰谷差,提高充电站经济效益与低碳水平。

关键词:充电站;碳排放;定价策略;深度强化学习;改进近端策略优化算法

中图分类号:TM73 **文献标识码:**A **DOI:**10.19457/j.1001-2095.dqed26207

Optimal Pricing Strategy of Electric Vehicle Charging Stations Considering Carbon Emissions

YIN Li¹, SHENG Junjie¹, YUAN Jie¹, FENG Yanjun²

(1. Wuhan Power Supply Company, State Grid Hubei Electric Power Company Limited, Wuhan 430000, Hubei, China; 2. School of Electrical Engineering, Southeast University, Nanjing 210096, Jiangsu, China)

Abstract: In the context of carbon neutrality, an optimal pricing strategy for electric vehicle charging stations considering carbon emissions was proposed. Firstly, a price response characteristic model of electric vehicle users was constructed. Secondly, an optimal pricing model with the lowest carbon emission, the minimum load peak-valley difference and the highest charging station revenue as the objectives was established and transformed into a Markov decision process. Then, an improved proximal policy optimization (PPO) algorithm based on the time difference error was proposed to improve the efficiency and stability of the algorithm. Finally, the example analysis shows that the proposed pricing strategy can reduce the peak-valley difference of distribution network load and improve the economic benefit and low carbon level of charging station.

Key words: charging station; carbon emission; pricing strategy; deep reinforcement learning; improved proximal policy optimization (PPO) algorithm

电动汽车保有量的持续增长给电网与充电站的经济、安全运行带来了挑战^[1-2],用户在晚间的聚集充电行为将引起负荷“峰上加峰”问题,如何引导电动汽车群体的充电行为成为了当前的热点问题^[3-5]。

目前,国内外学者开展了大量基于电价引导电动汽车充电行为的研究。文献[6]提出了一种双层充电定价策略,其中,上层基于K-means聚类算法划分电价的峰谷平时段,下层基于快速非支配排序遗传算法制定电价以提高充电站利润。然而,该模型仅能提供帕累托解集,在实际应用

中需要人工决策的介入。文献[7]从充电站-用户主从博弈的角度,建立了充电站实时定价模型,并基于遗传算法进行寻优求解,以提高充电站调峰效果与售电利润。文献[8]考虑可再生能源与储能系统的影响,提出了基于随机动态规划和贪婪算法的充电定价策略以减少对电网影响,并提高充电站利润与客户满意度。文献[9]针对电动出租车建立了充电引导及定价策略,在实时更新电动出租车的状态并预测充电需求的基础上,通过制定充电电价以提高充电站设备利用效率。

为了提高电动汽车充电引导的实时性与鲁

基金项目:国家电网公司科技项目(5400-202322566A-3-2-ZN)

作者简介:尹力(1989—),女,硕士,高级工程师,主要研究方向为智能配用电运行分析与管理,Email:Liyin_wh@outlook.com

棒性,部分学者探究了深度强化学习算法在充电定价引导领域的应用^[10-11]。文献[12]采用 Logistic 函数刻画用户在电价响应下的负荷转移,并基于深度确定性策略梯度函数(deep deterministic policy gradient, DDPG)进行充电引导以减小净负荷峰谷差。为了提高用户经济收益,文献[13]基于多智能体 DDPG 算法控制车辆在不同电价下的充放电功率。文献[14]则建立了一种基于深度强化学习的动态定价模型,将电动汽车充电行为建立为马尔科夫决策过程(Markov decision process, MDP),并依据充电站到达率与排队容量制定实时电价。然而,该模型仅考虑了充电站方面的信息,缺少对配电网信息的综合感知。

基于上述分析,本文构建了一种基于深度强化学习的充电站优化定价策略以提高充电站收益并降低碳排放与负荷峰谷差。在此基础上,提出一种基于时间差分误差(time difference error, TD-error)的改进近端策略优化(proximal policy optimization, PPO)算法进行模型的离线训练与在线求解,以提高算法的训练效率与收敛稳定性,最后基于算例仿真对所提策略的有效性进行验证。

1 电动汽车用户价格响应特性

电动汽车用户前往充电站补充电能的本质是一种服务/商品购买行为,而影响用户消费行为的最主要因素即为价格。在面对较小的价格差距时,用户往往存在一个心理阈值,当价格差低于该阈值时,用户基本不会做出响应;当超过该阈值时,用户的响应意愿随着价格刺激而增大;当价格差值达到饱和阈值后,用户响应意愿不会随之增大,而是达到了响应极限期。基于此,本文建立考虑死区的电动汽车用户价格响应模型如下式所示:

$$\mu = \begin{cases} 0 & \Delta\lambda \leq \Delta\lambda_1 \\ \frac{\mu_{\max}(\Delta\lambda - \Delta\lambda_1)}{\Delta\lambda_2 - \Delta\lambda_1} & \Delta\lambda_1 < \Delta\lambda < \Delta\lambda_2 \\ \mu_{\max} & \Delta\lambda \geq \Delta\lambda_2 \end{cases} \quad (1)$$

式中: μ 为电动汽车用户响应意愿; μ_{\max} 为最大响应意愿; $\Delta\lambda$ 为充电站新的分时电价与原电价的价格差,元/(kW·h); $\Delta\lambda_1, \Delta\lambda_2$ 分别为死区阈值与饱和阈值。

基于式(1)可以得到电动汽车用户充电电价响应特性示意图如图1所示。可见,当价格差 $\Delta\lambda$ 小于死区阈值 $\Delta\lambda_1$ 时,用户将不会做出响应,仍按

照原充电计划前往目标充电站充电;当价格差 $\Delta\lambda$ 处于线性区时, $\Delta\lambda$ 越大,用户响应意愿越强烈,更愿意前往价格低的充电站进行充电;当价格差 $\Delta\lambda$ 大于饱和阈值 $\Delta\lambda_2$ 时,用户的响应意愿达到最大值 μ_{\max} 。

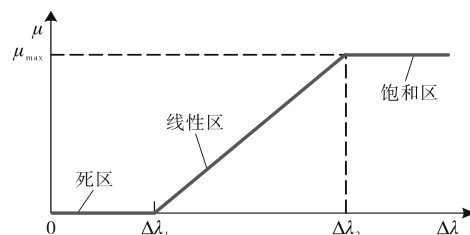


图1 电动汽车用户充电电价响应特性图

Fig.1 Response characteristics of charging electricity prices for electric vehicle users

2 计及碳减排的电动汽车充电站协调定价模型

为充分发挥电动汽车的碳减排效益并提高配电网的低碳化水平,本文将电动汽车碳排放最小、配电网负荷峰谷差最小以及充电站收益最大设为优化目标,建立计及碳减排的电动汽车充电站协调定价模型。同时,考虑到实际工况可能遇到的随机干扰与实时性问题,本文进一步将该优化模型转换为MDP模型,便于后续深度强化学习算法的训练求解。

2.1 计及碳减排的充电站定价模型

2.1.1 目标函数

1) 电动汽车碳排放最小。考虑电网在不同时段的碳排放因子以及电动汽车群体的移动储能特性,通过合理的定价策略能够调整区域充电负荷时空分布,从而降低电动汽车群体充电行为所带来的碳排放量,如下式所示:

$$\min f_1 = \sum_{t=1}^T \gamma_t^{\text{CO}_2} P_t^{\text{EV}} \Delta t \quad (2)$$

式中: T 为一天总时间段; $\gamma_t^{\text{CO}_2}$ 为电网在 t 时段的碳排放因子,kg/(kW·h); P_t^{EV} 为不同时刻电动汽车总充电负荷; Δt 为时间步长。

2) 配电网负荷峰谷差最小。所提协调定价策略的主要目的之一是平滑负荷曲线,提高电网削峰填谷能力,因此将下式所示的负荷峰谷差最小设为子目标函数:

$$\min f_2 = P_{\max}^{\text{DN}} - P_{\min}^{\text{DN}} \quad (3)$$

式中: $P_{\max}^{\text{DN}}, P_{\min}^{\text{DN}}$ 分别为配电网负荷峰值与谷值。

3) 充电站收益最高。充电站依靠从电网购电并向电动汽车用户售电盈利,因而优化策略应

保证充电站收益不低于实施分时电价前的收益,并尽可能提高充电站收益水平,如下式所示:

$$\max f_3 = \sum_{t=1}^T \sum_{i=1}^N (\lambda_{i,t}^{CS} - \lambda_t^{TOU}) P_{i,t}^{CS} \Delta t \quad (4)$$

式中: N 为充电站数量; λ_t^{TOU} 为工业分时电价; $\lambda_{i,t}^{CS}$ 为充电站 i 在 t 时段的充电电价; $P_{i,t}^{CS}$ 为充电站 i 在 t 时段的充电负荷大小。

2.1.2 约束条件

1) 配电网潮流约束如下:

$$-P_{i,t}^{CS} - P_{i,t}^L = V_{i,t} \sum_{j \in i} V_{j,t} (G_{ij} \cos \theta_{ij,t} + B_{ij} \sin \theta_{ij,t}) \quad (5)$$

$$-Q_{i,t}^{CS} - Q_{i,t}^L = V_{i,t} \sum_{j \in i} V_{j,t} (G_{ij} \sin \theta_{ij,t} - B_{ij} \cos \theta_{ij,t}) \quad (6)$$

式中: $P_{i,t}^L, Q_{i,t}^L$ 分别为配电网的有功与无功负荷; $P_{i,t}^{CS}, Q_{i,t}^{CS}$ 分别为充电站的有功与无功负荷; $V_{i,t}$ 为节点电压; G_{ij}, B_{ij} 分别为支路 $i-j$ 的电导与电纳; $\theta_{ij,t}$ 为相角差。

2) 线路安全约束如下:

$$V_{i,\min} \leq V_{i,t} \leq V_{i,\max} \quad (7)$$

$$I_{ij,\min} \leq I_{ij,t} \leq I_{ij,\max} \quad (8)$$

$$P_{ij,\min} \leq P_{ij,t} \leq P_{ij,\max} \quad (9)$$

$$Q_{ij,\min} \leq Q_{ij,t} \leq Q_{ij,\max} \quad (10)$$

式中: $V_{i,\max}, V_{i,\min}$ 分别为节点电压上、下限; $I_{ij,t}$ 为支路电流; $I_{ij,\max}, I_{ij,\min}$ 分别为支路电流上、下限; $P_{ij,t}$ 为支路有功功率; $P_{ij,\max}, P_{ij,\min}$ 分别为支路有功功率上、下限; $Q_{ij,t}$ 为支路无功功率; $Q_{ij,\max}, Q_{ij,\min}$ 分别为支路无功功率上、下限。

3) 充电站定价约束如下:

$$\lambda_t^{TOU} < \lambda_{i,t}^{CS} \leq \chi \lambda_t^{TOU} \quad (11)$$

式中: χ 为定价限幅系数。

4) 充电站收益约束如下:

$$\sum_{t=1}^T \sum_{i=1}^N \vartheta_{i,t}^{CS} \geq \sum_{t=1}^T \sum_{i=1}^N \tilde{\vartheta}_{i,t}^{CS} \quad (12)$$

式中: $\vartheta_{i,t}^{CS}, \tilde{\vartheta}_{i,t}^{CS}$ 分别为实施所提分时电价之后与之前的充电站收益。

2.2 基于MDP的充电站协调定价策略

深度强化学习智能体与环境交互架构如图2所示。作为一种基于Actor-Critic架构的深度强化学习算法, PPO智能体从配电网环境中观测实时状态 s_t 并基于Actor网络做出实时决策动作 a_t , 在执行动作 a_t 后智能体将得到即时奖励 r_t 。通过上述交互不断累积历史样本并更新网络参数, 最终帮助智能体学习到环境状态-最优动作的映射关系。

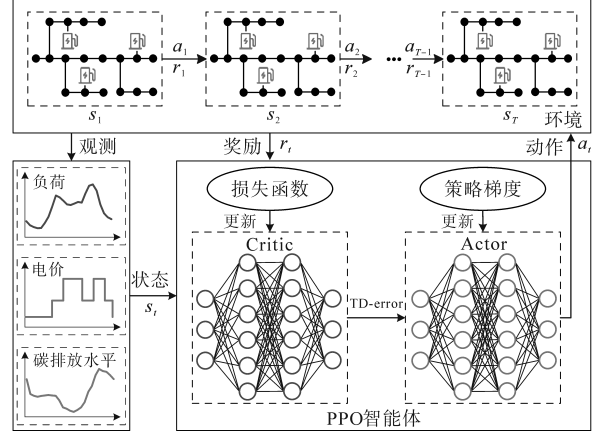


图2 深度强化学习智能体与环境交互架构

Fig.2 Architecture of deep reinforcement learning agent and environment interaction

1) 状态。状态代表智能体对环境实时信息的感知, 用以辅助智能体决策。在本文中, 环境状态由配电网与充电站两方面组成, 如下式所示:

$$s_t = \{ \underbrace{P_{i,t}^L, \gamma_t^{CO_2}}_{DN}, \underbrace{P_{j,t}^{CS}, \lambda_{j,t}^{CS}}_{CS} \} \quad (13)$$

2) 动作。动作代表智能体依据环境信息做出的决策方案, 即确定各充电站实时电价。

$$a_t = \{ \lambda_{j,t}^{CS} \} \quad (14)$$

3) 奖励。奖励是在执行动作 a_t 后环境给予的实时反馈, 用以引导智能体学习更优秀的决策方案。奖励 r_t 由碳排放成本 $r_t^{CO_2}$ 、负荷峰谷差惩罚 r_t^{DN} 以及充电站收益奖励 r_t^{CS} 3项组成, 如下式所示:

$$r_t = r_t^{CO_2} + r_t^{DN} + r_t^{CS} \quad (15)$$

其中

$$r_t^{CO_2} = -\lambda^{CO_2} \gamma_t^{CO_2} P_t^{EV} \Delta t \quad (16)$$

$$r_t^{DN} = \begin{cases} -\lambda^{PD} (P_{\max}^{DN} - P_{\min}^{DN}) & s_t \text{ 为终点} \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

$$r_t^{CS} = \sum_{i=1}^N (\lambda_{i,t}^{CS} - \lambda_t^{TOU}) P_{i,t}^{CS} \Delta t \quad (18)$$

式中: λ^{CO_2} 为碳排放成本系数, 元/kg; λ^{PD} 为峰谷差惩罚系数, 元/MW。

4) 状态转移概率。设 t 时刻的环境状态为 s_t , 执行动作 a_t 后进入 $t+1$ 时刻, 环境状态会随之转移到 s_{t+1} , 这一过程可由状态转移概率函数 Ψ 表示, 如下式所示:

$$s_{t+1} = \Psi(s_t, a_t) \quad (19)$$

在实际应用中这一转移过程常伴随着实时随机干扰, 转移模型是无法精确量化的。为此, 本文引入深度强化学习算法, 其不依赖于特定的系统物理转移模型, 而是基于大量训练与探索过程来学习环境的概率转移模型。

3 改进深度强化学习 PPO 算法

作为信赖域策略优化(trust region policy optimization, TRPO)的改进算法, PPO 算法通过约束策略更新的幅度从而保证其训练稳定性^[15-16]。然而,该方法忽略了不同样本的重要性区别,在实际训练中存在未充分挖掘样本价值与训练效率低的问题,无法有效平衡算法的稳定性与训练速度。为此,本文提出了一种基于时间差分误差的改进 PPO 算法,基于样本偏差动态调整 PPO 算法裁减率,提高算法的效率与稳定性。

3.1 PPO 算法基本原理

在 PPO 算法中, t 时刻的优势函数 $\hat{A}_t(a, s)$ 如下式所示:

$$\hat{A}_t(a, s) = -V_\phi(s_t) + r_t + \gamma V_\phi(s_{t+1}) + \dots + \gamma^{K-t+1} r_{K-1} + \gamma^{K-t} V_\phi(s_K) \quad (20)$$

式中: $V_\phi(s_t)$ 为状态价值函数; γ 为折旧率; K 为视野步长。

在式(20)的基础上, PPO 算法的训练目标 L 可表述为获得能使期望优势函数最大的智能体网络参数,如下式所示:

$$\begin{aligned} L &= \max_{\theta} E_t \left[\frac{\pi_{\theta^{t+1}}(a|s)}{\pi_{\theta^t}(a|s)} \hat{A}_t(a, s) \right] \\ &= \max_{\theta} E_t [\tau_t \hat{A}_t(a, s)] \end{aligned} \quad (21)$$

式中: π_{θ^t} 为策略函数; τ_t 为新旧策略比率。

为了保证训练稳定性, PPO 算法对 τ_t 进行了限幅以保证其大小在 $[1 - \varepsilon, 1 + \varepsilon]$ 范围内,而目标函数可改为下式:

$$L_{\text{clip}} = \max_{\theta} E_t [\min(\tau_t \hat{A}_t, \text{clip}(\tau_t) \hat{A}_t)] \quad (22)$$

其中

$$\text{clip}(\tau_t) = \begin{cases} 1 - \varepsilon & \tau_t < 1 - \varepsilon \\ \tau_t & 1 - \varepsilon \leq \tau_t \leq 1 + \varepsilon \\ 1 + \varepsilon & \tau_t > 1 + \varepsilon \end{cases} \quad (23)$$

式中: ε 为裁减率。

3.2 PPO 算法改进策略

本文所提改进 PPO 算法的核心思想是基于样本的优劣动态调整裁减率 ε , 从而平衡算法面对不同样本下的训练速度与稳定性。所提方法如图 3 所示, 当样本 TD-error 较大时, 意味着当前策略效果较差, 该样本能够提供较高的经验贡献, 因此设置较大的裁减率以提高算法更新幅度与训练效率, 寻找更优秀的策略; 反之, 意味着当前策略已较为优秀, 因此应减小裁减率以保证收

敛稳定性。其中, 动态裁减率可由下式计算得到:

$$\varepsilon = \begin{cases} \varepsilon_{\max} & \delta_t \geq \bar{\delta}_t \\ \varepsilon_{\min} & \delta_t < \bar{\delta}_t \end{cases} \quad (24)$$

$$\delta_t = r_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t) \quad (25)$$

式中: $\varepsilon_{\max}, \varepsilon_{\min}$ 分别为最大、最小裁减率; δ_t 为时间差分误差; $\bar{\delta}_t$ 为时间差分误差的均值。

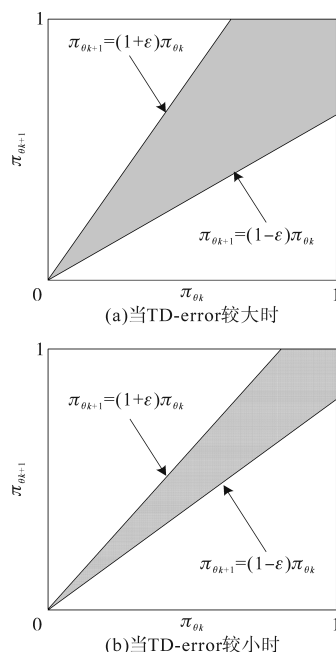


图3 动态调整裁减率示意图

Fig.3 Schematic of dynamic adjustment of clip ratio

3.3 改进 PPO 算法的训练流程

所提改进 PPO 算法训练流程如图 4 所示。在初始化训练环境后, 首先对配电网与充电站实时状态 s_t 进行观察, 并做出动作决策 a_t ; 其次, 计算

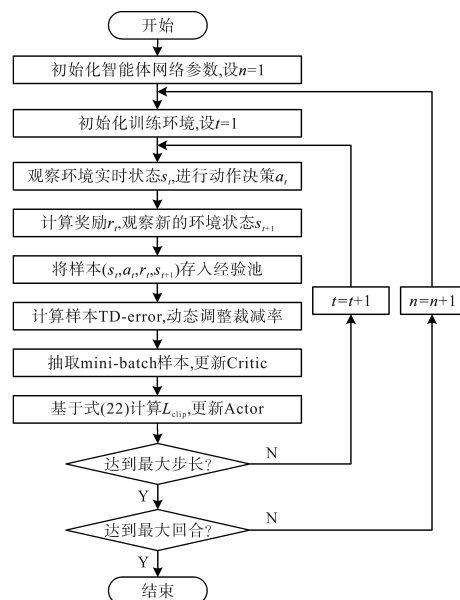


图4 改进 PPO 算法的训练流程图

Fig.4 Training flowchart of the improved PPO algorithm

所得奖励 r_t 并将样本 (s_t, a_t, r_t, s_{t+1}) 存放至经验池中;然后,计算样本的 TD-error,基于所提方法动态调整裁减率 ε ,并更新智能体网络参数;最后,重复上述步骤直至达到最大训练回合,最终输出训练完成的智能体以供在线部署。

4 算例分析

4.1 算例仿真参数

本文在 IEEE-33 节点配电网进行算例仿真,设该区域共 6 座充电站,分别接入节点 8、节点 12、节点 16、节点 20、节点 24、节点 30,且均配置了 10 根 60 kW 快充桩。设仿真时间步长为 1 h,用户最大响应意愿为 30%,死区阈值与饱和阈值分别取 0.1 元/(kW·h)与 1.5 元/(kW·h),充电站定价限幅系数为 2.6,碳排放成本系数为 0.05 元/kg,配电网峰谷差惩罚系数为 4 000 元/MW。工业分时电价如表 1 所示,配电网负荷曲线参考文献[17],电网不同时段碳排放因子参考文献[18]。

表 1 工业分时电价

Tab.1 Industrial time of use electricity price

	时段	电价/[元/(kW·h ⁻¹)]
峰	10:00—15:00, 18:00—21:00	1.025
平	07:00—10:00, 15:00—18:00, 21:00—23:00	0.615
谷	00:00—07:00, 23:00—24:00	0.305

4.2 离线训练过程

设算法的折扣率为 0.98, mini-batch 容量为 100,经验池容量为 2 000,最大与最小裁减率分别为 0.25 与 0.1,可得智能体离线训练奖励曲线如图 5 所示。由图可见,尽管前期智能体所获得的奖励较低,但得益于所提改进机制以较大的裁减率提高算法更新幅度与训练效率,智能体前期奖励波动式上升。在 400 回合之后,智能体奖励

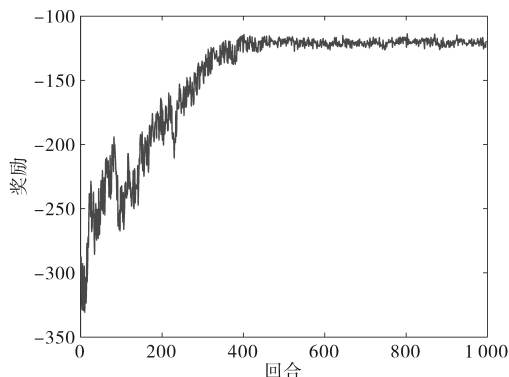


图 5 智能体离线训练奖励曲线

Fig.5 Reward curve for offline training of the agent

曲线逐渐收敛,此阶段裁减率较小,以提高算法的后期稳定性。最终,智能体奖励收敛于-120.76 附近,证明其学习到了最优的决策网络参数。

4.3 策略有效性分析

将离线训练后的智能体进行在线测试,可得优化定价前、后各充电站负荷空间分布与配电网负荷时间分布分别如图 6 与图 7 所示。

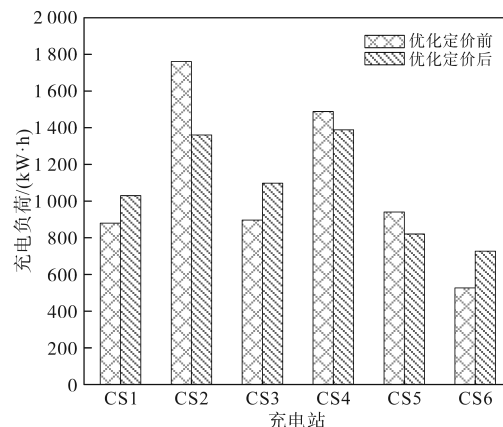


图 6 优化定价前后各充电站负荷

Fig.6 Charging load of each charging station before and after optimizing pricing

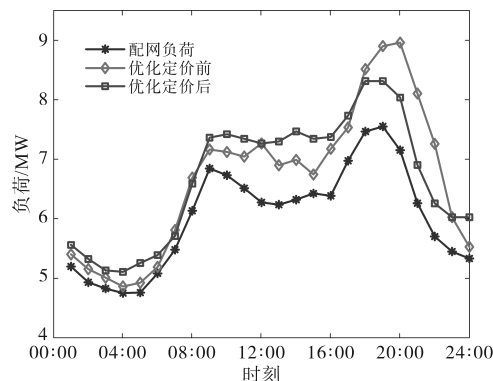


图 7 优化定价前、后配电网负荷

Fig.7 Distribution network load before and after optimizing pricing

由图 6 可见智能体的定价策略有效平衡了充电负荷的空间分布,各充电站间负荷标准差仅为 271.97 kW·h,相较于优化定价前降低了 39.89%,有效缓解了部分充电站充电拥堵的情况,提高了充电设施网络整体利用率。由图 7 可知,在时间分布上由于电动汽车在晚间负荷峰值期间的聚集充电行为,优化定价前配电网峰谷差增大,达 4.10 MW,而所提方法通过电价引导用户充电行为参与削峰填谷,负荷峰谷差仅为 3.21 MW,相较于前者降低了 21.71%,提高了配电网的稳定性和经济性。可见,所提方法能够综合感知环境信息并输出定价策略,通过空间上的充电站电价差与时间上的不同时段电价差能够有效引导充电负

荷实时响应。

为进一步分析所提策略的碳减排效益,表2给出了优化定价前、后的算例结果,图8则给出了全天不同时段碳排放成本曲线。由表2可见,优化定价后充电站收益为10 851.92元,较优化定价前提高了16.78%。同时,结合图8可见所提方法有效缓解了由电动汽车充电带来的碳排放问题,尤其在晚间负荷高峰时段。全天碳排放成本仅为754.74元,较优化定价前降低了16.44%。综上所述,所提协调定价策略能够有效引导电动汽车充电负荷的时间-空间转移,在保证充电站收益的基础上降低配电网峰谷差与碳排放成本。

表2 优化定价前、后的算例结果

Tab.2 Case study results before and after optimizing pricing

	优化定价前	优化定价后
碳排放成本/元	903.24	754.74
配电网峰谷差/MW	4.10	3.21
充电站收益/元	9 292.71	10 851.92

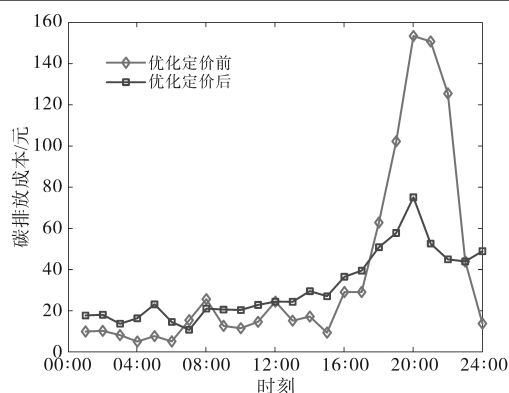


图8 优化定价前、后碳排放成本

Fig.8 Carbon emission cost before and after optimizing pricing

4.4 不同算法对比

为验证所提PPO算法改进策略的有效性,将DDPG算法、PPO算法与所提改进PPO算法进行训练对比,可得不同算法的训练奖励曲线与配电网电压幅值对比分别如图9与图10所示。由图9可见,PPO算法在约580回合实现收敛,但训练后期稳定性较差,平均奖励约为-131.52。DDPG算法收敛速度较快,但后期平均奖励较低,仅为-146.87,意味着算法的综合寻优能力较差。而所提改进PPO算法基于样本的优劣动态调整裁减率,使得智能体能够在训练前期提高更新幅度与训练效率,同时在训练后期保证收敛稳定性,平均奖励相较于PPO算法提升了8.18%。此外,由晚间用电高峰期的电压曲线可见,优化前各节点

电压分布显著不均且存在电压越限,最低电压为0.903(标么值),电压合格率仅为70.97%。所提算法通过充电定价策略引导用户充电行为的时空转移,显著抑制了电压越限情况。

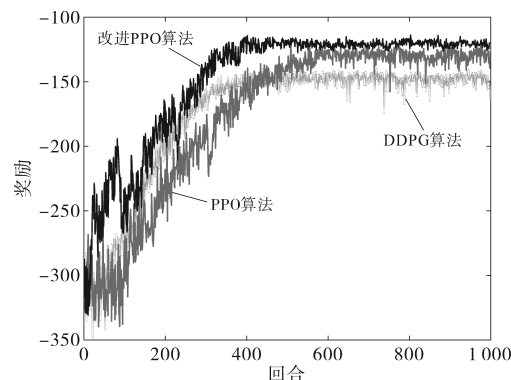


图9 不同深度强化学习算法训练曲线对比

Fig.9 Comparison of training curves for different deep reinforcement learning algorithms

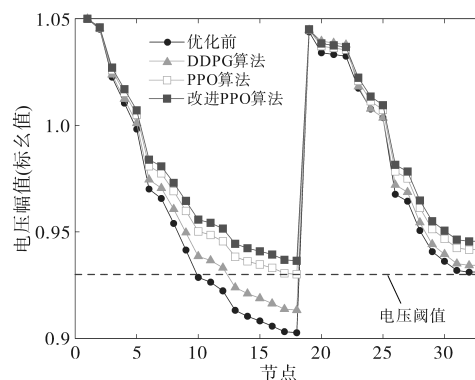


图10 配电网在20:00时刻的节点电压幅值对比

Fig.10 Comparison of node voltage amplitudes of distribution networks at 20:00

4.5 模型敏感性分析

最后,为分析超参数裁减率的设置对模型性能的影响,图11给出了不同最大与最小裁减率下智能体所得收敛奖励对比。

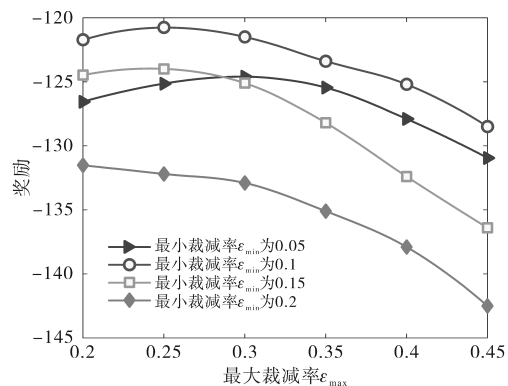


图11 不同最大与最小裁减率下算法奖励

Fig.11 Algorithm reward under different maximum and minimum clip ratio

由图 11 可见,随着 ε_{\max} 的增大,算法奖励整体呈下降趋势,其取值 0.45 时相较于取值 0.2 时奖励平均降低 6.76%。而随着 ε_{\min} 的增大,模型更新策略逐渐由保守转为激进。当 ε_{\min} 增大至 0.2 时,由图中最下方曲线可见算法奖励偏低,这是由于此时模型更新策略更加激进,可能破坏已验证的优秀策略。综上,当 ε_{\max} 与 ε_{\min} 分别取值 0.25 与 0.1 时算法奖励值最高,能够更好地权衡模型训练效率与收敛稳定性。

5 结论

本文基于电动汽车用户的价格响应特性,建立了计及碳排放的充电站优化定价模型,并将其转化为了 MDP 过程以便所提改进 PPO 算法进行训练求解。通过算法的离线训练与在线测试,得出如下结论:

1) 构建的基于 MDP 的充电站协调定价策略能够感知配电网-充电站环境状态,并基于深度强化学习制定实时充电电价引导电动汽车用户,有效提高了策略的实时性与鲁棒性。

2) 提出的基于 TD-error 的改进策略有效平衡了 PPO 算法的稳定性与训练速度,提高算法在应对不同样本情况下的训练效率与稳定性。

3) 算例结果表明,所提出的充电站优化定价模型能够在降低碳排放的基础上,提高充电站收益并降低配电网负荷峰谷差。

下一步研究将考虑区域级的规模化可再生能源以及常规火发电机组的运行特性,与电动汽车协调互补进一步提升碳减排效益。

参考文献

- [1] 钱涛,任孟极,邵成成,等. 基于深度学习考虑出行模式的电动汽车充电负荷场景生成[J]. 电力系统自动化, 2022, 46(12): 67-75.
QIAN Tao, REN Mengji, SHAO Chengcheng, et al. Deep-learning-based electric vehicle charging load scenario generation considering travel mode[J]. Automation of Electric Power Systems, 2022, 46(12): 67-75.
- [2] 周宇凡,高辉,龙羿. 计及动态电价的电动汽车定制化充电策略[J]. 电气传动, 2024, 54(7): 32-39.
ZHOU Yufan, GAO Hui, LONG Yi. Customized charging strategy for EV considering dynamic electricity prices[J]. Electric Drive, 2024, 54(7): 32-39.
- [3] 金志刚,胡怡,李根,等. 面向“代客加电”服务的电动汽车充电引导策略[J]. 电力系统保护与控制, 2022, 50(14): 76-84.
JIN Zhigang, HU Yi, LI Gen, et al. Charging guiding strategy for electric vehicles oriented to a valet charging service[J]. Power System Protection and Control, 2022, 50(14): 76-84.
- [4] 李东东,段维伊,林顺富,等. 实时电价条件下基于用户引导的电动汽车-充电桩匹配策略[J]. 电力系统自动化, 2020, 44(8): 74-82.
LI Dongdong, DUAN Weiwei, LIN Shunfu, et al. User guidance based matching strategy for electric vehicle-charging pile in condition of real-time electricity price[J]. Automation of Electric Power Systems, 2020, 44(8): 74-82.
- [5] ZHANG Yongmin, YOU Pengcheng, CAI Lin. Optimal charging scheduling by pricing for EV charging station with dual charging modes[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(9): 3386-3396.
- [6] 罗建竹,苏春. 基于用户行为和分时电价的充电定价策略优化[J]. 东南大学学报(自然科学版), 2021, 51(6): 1109-1116.
LUO Jianzhu, SU Chun. Optimization of charging pricing strategy based on user behavior and time-of-use tariffs[J]. Journal of Southeast University (Natural Science Edition), 2021, 51(6): 1109-1116.
- [7] 徐文法,王育飞,薛花,等. 基于博弈论的电动汽车充电实时定价策略研究[J]. 电测与仪表, 2021, 58(7): 36-43.
XU Wenfa, WANG Yufei, XUE Hua, et al. Research on real-time pricing strategy of electric vehicle charging based on game theory[J]. Electrical Measurement & Instrumentation, 2021, 58(7): 36-43.
- [8] LUO Chao, HUANG Yih-Fang, GUPTA Vijay. Stochastic dynamic pricing for EV charging stations with renewable integration and energy storage[J]. IEEE Transactions on Smart Grid, 2018, 9(2): 1494-1505.
- [9] 傅质馨,朱韦翰,朱俊澎,等. 动态路-电耦合网络下电动出租车快速充电引导及其定价策略[J]. 电力自动化设备, 2022, 42(4): 9-17.
FU Zhixin, ZHU Weiham, ZHU Junpeng, et al. Fast charging guidance and pricing strategy for electric taxis based on dynamic traffic-grid coupling network[J]. Electric Power Automation Equipment, 2022, 42(4): 9-17.
- [10] 范士雄,李立新,王松岩,等. 人工智能技术在电网调控中的应用研究[J]. 电网技术, 2020, 44(2): 401-411.
FAN Shixiong, LI Lixin, WANG Songyan, et al. Application analysis and exploration of artificial intelligence technology in power grid dispatch and control[J]. Power System Technology, 2020, 44(2): 401-411.
- [11] 赵晋泉,夏雪,徐春雷,等. 新一代人工智能技术在电力系统调度运行中的应用评述[J]. 电力系统自动化, 2020, 44(24): 1-10.
- [12] ZHAO Jinquan, XIA Xue, XU Chunlei, et al. Review on application of new generation artificial intelligence technology in power system dispatching and operation[J]. Automation of Electric Power Systems, 2020, 44(24): 1-10.
- [12] 王庆园,崔莉,王明深,等. 考虑快慢充负荷特性的电动汽车调峰定价策略[J]. 电力工程技术, 2023, 42(4): 31-40.

- WANG Qingyuan, CUI Li, WANG Mingshen, et al. Peak load regulation pricing strategy of electric vehicle considering fast and slow charging characteristics[J]. Electric Power Engineering Technology, 2023, 42(4): 31-40.
- [13] 王琼, 邹晴, 李乐, 等. 基于多智能体强化学习的电动汽车充放电调控算法[J]. 供用电, 2023, 40(9): 83-90.
- WANG Qiong, ZOU Qing, LI Le, et al. Multi-agent reinforcement learning algorithm for charging/discharging control of electric vehicles[J]. Distribution & Utilization, 2023, 40(9): 83-90.
- [14] ZHAO Zhonghao, LEE Carman K M. Dynamic pricing for EV charging stations: a deep reinforcement learning approach[J]. IEEE Transactions on Transportation Electrification, 2022, 8(2): 2456-2468.
- [15] 王子昊, 王旭, 蒋传文, 等. 基于近端策略优化算法的灾后配电网韧性提升方法[J]. 电力系统自动化, 2022, 46(21): 62-70.
- WANG Zihao, WANG Xu, JIANG Chuanwen, et al. Resilience improvement method for post-disaster distribution network based on proximal policy optimization algorithm[J]. Automation of Electric Power Systems, 2022, 46(21): 62-70.
- [16] 杨志学, 任洲洋, 孙志媛, 等. 基于近端策略优化算法的新能源电力系统安全约束经济调度方法[J]. 电网技术, 2023, 47(3): 988-998.
- ZHU Zhixue, REN Zhouyang, SUN Zhiyuan, et al. Security-constrained economic dispatch of renewable energy integrated power systems based on proximal policy optimization algorithm[J]. Power System Technology, 2023, 47(3): 988-998.
- [17] 马涛, 李津, 曹晓波, 等. 计及用户行为差异性和配电网潮流的电采暖负荷群优化调度策略研究[J]. 电力科学与技术学报, 2023, 38(1): 77-87.
- MA Tao, LI Jin, CAO Xiaobo, et al. Research on optimal dispatch strategy of electric heating load groups considering user behavior difference and distribution network power flow[J]. Journal of Electric Power Science and Technology, 2023, 38(1): 77-87.
- [18] 薛贵挺, 汪柳君, 刘哲, 等. 考虑碳排放的光储充一体站日前运行策略[J]. 电力系统保护与控制, 2022, 50(7): 103-110.
- XUE Guiting, WANG LiuJun, LIU Zhe, et al. Day-ahead operation strategy of an integrated photovoltaic storage and charging station considering carbon emissions[J]. Power System Protection and Control, 2022, 50(7): 103-110.
- 收稿日期: 2024-10-07
修改稿日期: 2024-11-17
- *****
- (上接第49页)
1306-1310.
- [16] 唐亚东, 刘寅杨, 杨维永. 基于等级保护网络安全体系的新型电力系统风险分析与防范[J]. 网络安全技术与应用, 2023(12): 130-133.
- TANG Yadong, LIU Yinyang, YANG Weiyong. Risk analysis and prevention of new power system based on hierarchical protection network security system[J]. Network Security Technology and Application, 2023(12): 130-133.
- [17] 潘小辉, 罗兴, 穆煜, 等. 挖掘分布式资源潜能的主-配-微一体化协同控制[J]. 电力大数据, 2024, 27(5): 1-8.
- PAN Xiaohui, LUO Xing, MU Yu, et al. Coordinated optimization control for main-distribution-microgrid interconnection utilizing the potential of massive distributed resources[J]. Power Systems and Big Data, 2024, 27(5): 1-8.
- 收稿日期: 2024-11-13
修改稿日期: 2024-12-30