

基于深度强化学习的电动汽车有序充电优化方法

喻磊¹, 谈竹奎², 王扬², 刘通¹, 肖宁², 肖小兵², 欧俊杰², 林心昊¹, 陈千懿¹

(1. 南方电网科学研究院, 广州 510663; 2. 贵州电网有限责任公司电力科学研究院, 贵阳 550002)

摘要: 针对大规模电动汽车无序接入电网引发的用户充电开销大和电网负荷波动加剧等问题, 提出了基于深度强化学习(deep reinforcement learning, DRL)的电动汽车充电行为优化方法。首先, 以最小化电网负荷波动和用户充电费用为目标, 建立了电动汽车有序充电优化调度模型。其次, 将电动汽车的充电行为构建为马尔科夫决策过程(Markov decision process, MDP), 根据电网负荷预测信息和分时电价对充电时段进行优先级评定, 并根据优先级控制电动汽车充电行为。通过双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)算法对电动汽车有序充电策略进行快速优化。最后, 通过算例验证了所提方法在减少用户的充电开销和配电网的负荷波动方面的有效性。

关键词: 电动汽车; 深度强化学习; 有序充电; 优先级评定

Deep Reinforcement Learning Based Optimization Method for Ordered Charging of Electric Vehicle

YU Lei¹, TAN Zhukui², WANG Yang², LIU Tong¹, XIAO Ning², XIAO Xiaobing², OU Junjie², LIN Xinhao¹, CHEN Qianyi¹

(1. Electric Power Research Institute, CSG, Guangzhou 510663, China; 2. Electric Power Research Institute of Guizhou Power Grid Co., Ltd., Guiyang 550002, China)

Abstract: The large-scale disordered integration of electric vehicles (EVs) into power grid poses many problems, such as enlarged power fluctuation for grid and increased charging cost for users. To solve these problems, this paper proposes an optimized EV charging method based on deep reinforcement learning (DRL). Firstly, an EV ordered charging scheduling model is established aiming at minimizing the power fluctuation and user charging cost. Secondly, the EV charging behaviour is formulated as a Markov decision process (MDP) which evaluates the priority for each charging period based on load prediction information and time-of-use electricity tariff. The charging behaviour of EVs is controlled by the priority. The twin delayed deep deterministic policy gradient (TD3) algorithm is adopted to solve the MDP problem, which quickly optimizes EV ordered charging strategy. Finally, the effectiveness of the proposed method in reducing charging cost and load fluctuation of distribution network are verified by case studies based on numerical examples.

Key words: electric vehicles; deep reinforcement learning; ordered charging; priority evaluation

0 引言

2020年国务院印发《新能源汽车产业发展规划(2021-2035)》^[1], 指出我国应推动新能源汽车产业

高质量发展, 加快建设汽车强国。与传统汽车相比, 电动汽车(electric vehicle, EV)具有高效、清洁的特点。在国家政策的推动下我国EV产业蓬勃发展, EV的数量急剧增长^[2]。与此同时, 大规模的EV于高峰时段接入电网会加剧配电网的负荷波动, 进一步加大峰谷差, 导致“峰上加峰”等问题, 并带来设备过载、损耗增加等不良影响^[3-5]。

2021年, 国家发展改革委要求各地结合当地情况推行分时电价(time-of-use tariff)机制。分时电价可在需求侧有效引导用户规范用电行为, 从而保障

基金项目: 国家重点研发计划资助项目(2022YFE0205300); 贵州电网有限责任公司科技项目(GZKJXM20210484)。

Foundation item: Supported by the National Key Research and Development Program of China (2022YFE0205300); the Science and Technology Project of Guizhou Power Grid Co., Ltd. (GZKJXM20210484).

电力供应与需求的均衡运行。同样地,充电站也可根据分时电价合理调控EV充电行为,以减少配电网的负荷波动和用户的充电费用^[6-7]。因此,实时优化EV充电行为具有重要的研究意义。

传统的EV充电行为优化调度策略是以多主体利益、降低负荷曲线方差、提高风电消纳能力等为目标构建优化调度模型^[8-9],然后采用运筹学或进化算法^[10-12]进行求解。但这些算法往往高度依赖实时阶段EV充电的预测信息,且求解高维问题时会耗费大量资源,效率较低。文献[13]计及用户响应度构建电动汽车充放电优化调度模型并设计基于粒子群优化算法的模型求解器。文献[14]研究了电动汽车充电站参与调压的建模方法及其控制策略。

近年来,强化学习(reinforcement learning, RL)凭借其在序列决策方面的优势发展迅速,同时也给电动汽车充放电决策实时优化领域的研究带来了新的机遇^[15-18]。文献[19]考虑了EV接入配电网、可再生能源和负荷所带来的不确定性,使用Q-learning算法进行日前训练,日内不需要迭代计算即可进行实时调度。文献[20]将15 min设定为一个时间步,使用马尔科夫决策过程(Markov decision process, MDP)对每个时间步中的EV的充电行为进行分析,通过构造奖励函数引导用户EV进行充电选择。为解决传统深度Q学习算法在EV充电控制优化中动作和状态维度上绝对数值相差过大的问题,文献[21]提出了基于竞争深度Q网络(dueling deep Q network, DDQN)的深度强化学习优化方法,兼顾了用户出行需求和配电网电压调控,将每辆EV单独作为优化变量进行控制实现简单,但该方法只适用于EV数量较少的情况,EV数量过多容易导致“维度灾难”问题。文献[22]引入了EV集群的思想,建立了EV集群的能量上下边界模型,使用软行动者-评论家(soft actor-critic, SAC)^[23]算法进行模型训练和实时调度。文献[24]对单辆EV充电过程进行MDP建模,使用双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)^[25]算法从负荷聚合商层面优化EV集群的充电行为。文献[26]按照配网节点和充电结束时间分配EV集群,提出考虑充电成本和充电不足惩罚的充电调度策略,并使用深度DRL算法求解。上述研究需构建复杂的电动汽车集群充电站环境,即充电优化过程需由智能体分布式部署实现,因此

硬件成本偏高。

大规模EV接入配电网后带来了诸多不确定性因素,为解决大规模EV的有序充电问题,本文提出了一个简单高效的基于DRL的EV有序充电优化方法。首先,搭建充电场景,并建立以用户充电费用最少、配电网总负荷标准差最小为目标函数的EV充电行为优化模型。其次,提出优先级评定的思想,通过评定各时段充电优先级简单地控制每辆EV的充电行为。然后,使用TD3算法进行模型训练和求解。最后,通过算例表明,本文所提策略在满足用户出行的充电需求条件下可大幅降低用户的充电费用,并有效减小配电网的负荷波动。

1 场景构建

1.1 场景描述

本文构建图1所示的EV充电站场景对EV充电优化问题进行建模。该场景类似于工业园区、居民小区附近的公共慢充站,并在此基础上进行电动汽车有序充电优化。EV用户驶入充电站,并设定预计离去时间,充电桩功率固定。作为该场景的控制主体,智能体响应分时电价,整合负荷预测信息与EV电量等信息,调控充电桩的充电时段。智能体需在满足用户出行需求的基础上最大程度地减少用户的充电开销和配电网的负荷波动,起到削峰填谷的作用。

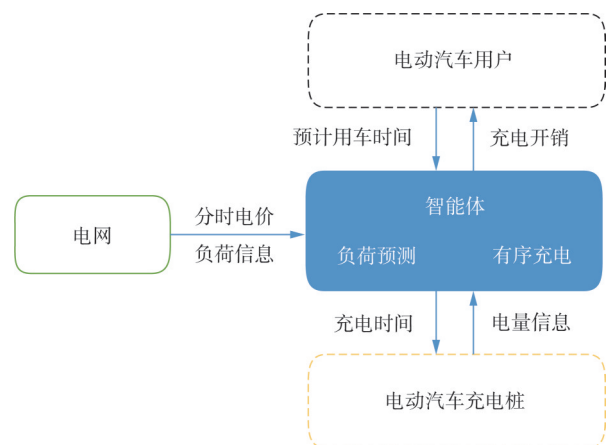


图1 场景构建

Fig. 1 Scenario construction

1.2 实时优化调度模型

为满足用户出行需求的同时还需保证配电网的安全、经济运行, EV的充电行为要受到多种约束

条件的限制。若要兼顾配电网和用户之间的利益,则必须从多个目标函数之间寻找平衡,以最小化目标函数 f 为例,其数学模型可表示为:

$$\begin{cases} \min f = \{f_1(x), f_2(x), \dots, f_M(x)\} \\ \text{s.t. } g(x) \leq 0, h(x) = 0 \end{cases} \quad (1)$$

式中: $f_M(x)$ 为第 M 个待优化的目标函数; x 为待优化的变量; $g(x) \leq 0$ 为 x 的不等式约束; $h(x) = 0$ 为 x 的等式约束。

1.2.1 目标函数

实时优化调度模型的目标是尽可能减少用户充电开销,并保证配电网的安全运行。结合式(1)设定EV的充电行为的优化目标为:

$$f_1 = \min \sum_{t=1}^T \sum_{i=1}^{N_t} \lambda_t P_{i,t} \Delta t \quad (2)$$

$$f_2 = \min \sqrt{\frac{1}{T} \sum_{t=1}^T (P_t - P_{av})^2} \quad (3)$$

式中:优化目标函数 f_1 为用户日开销最小; T 为时间步数; N_t 为用户数量(即EV数量); Δt 为时间步长; $P_{i,t}$ 为 t 时刻第 i 辆车的充电功率; λ_t 为 t 时刻的电价; f_2 为负荷标准差最小; P_t 为该区域在时刻 t 的负荷; P_{av} 为该区域当日平均负荷。

1.2.2 约束条件

充电站须满足用户的出行需求,即用户驶离充电站时EV电池的SOC应充满。本文对充电时刻、充电功率和EV电池的SOC进行了约束,以避免充电过程中出现欠充和过充的情况。

充电时刻约束为:

$$t_{i,arr} \leq t \leq t_{i,dep} \quad (4)$$

式中 $t_{i,arr}$ 和 $t_{i,dep}$ 分别为第 i 辆EV的到达时刻和预计离去时刻。

充电功率约束为:

$$0 \leq P_{i,t} \leq P^{\max} \quad (5)$$

式中: $P_{i,t}$ 为第 i 辆EV在时刻 t 的充电功率; $P^{\max}$ 为EV最大充电功率。

EV电池的SOC约束为:

$$p_{i,t+1}^{\text{SOC}} = p_{i,t}^{\text{SOC}} + \eta p_{i,t} \Delta t \quad (6)$$

$$p_{i,tarr}^{\text{SOC}} \leq p_{i,t}^{\text{SOC}} \leq P^{\max} \quad (7)$$

$$t = t_{i,dep}, p_{i,t}^{\text{SOC}} = P^{\max} \quad (8)$$

式中: $P_{i,t}^{\text{SOC}}$ 和 $P_{i,t+1}^{\text{SOC}}$ 分别为第 i 辆EV在时刻 t 和 $t+1$ 的电量大小; η 为EV的充电效率; $p_{i,tarr}^{\text{SOC}}$ 和 $P^{\max}$ 分别为第 i 辆EV来到充电站的起始电量和最大电量。

1.2.3 优先级控制充电时段

考虑EV在电价低谷时段充电时间越久充电开销越小,因此,本文将优化目标转化为在每辆EV的充电时段为到达充电站的EV安排好充电计划。本文提出了充电优先级评定的思想,通过分时电价和负荷预测信息对充电时段进行优先级评定,并根据优先级控制每辆EV的充电时段。

EV依据优先级充电控制步骤如下。

1)设定用户告知充电站EV离站时间,且离站时电量充满,根据已知的EV驶入充电站时间、离开充电站时间以及电量,判断EV是否可以在离开充电站时充满电,若是则执行步骤2),否则全程充电。

2)选取最高优先级的时段,先判断是否在该EV的充电区间内,若是则该时段充电,否则该时段不充电。接着选取次优先级的时段,以此类推,直到电量充满。如此便得到各时段的充电功率。

3)若用户未告知车辆离站时间,除获悉用户用车习惯并结合车辆到站时间(如00:00点到站),可不考虑用户快速充电需求,尽量在电价低谷时期充电,以追求最低的充电开销,否则充电站优先以最快速度充满电,即EV驶入充电站后立即充电,直至电量充满。

2 EV充电控制的DRL方法

2.1 强化学习

2.1.1 核心概念与术语

在强化学习中,智能体(agent)在与环境(environment)不断交互的过程中学习如何采取动作以最大化累计奖励。强化学习问题可建模成一个用元组 (S, A, P, R, γ) 表示的马尔可夫决策过程。 S 为状态空间, A 为动作空间, P 为状态转移概率, R 为奖励, $\gamma \in (0, 1)$ 为折扣因子。强化学习算法的目的是找到一个策略最大化未来奖励之和 $R_t = \sum_{i=t}^T \gamma^{i-t} r(s_i, a_i)$ 。该策略实施期间,在 i 时刻开始,在 T 时刻结束, $r(s_i, a_i)$ 为在 t 时刻对应的状态 s_i 下采取动作 a_i 所获得的奖励。

强化学习中使用状态值函数 V^π 或动作值函数 Q^π 来评估策略,其中动作值函数 Q^π 应用更为广泛。 $Q^\pi(s, a)$ 为智能体按照策略 π 在某一状态下执行动作得到的累计奖励的估计值,如式(9)(也称作Bell-

man 公式)所示。

$$Q^\pi(s, a) = r(s, a) + \gamma \sum_{s' \in S} p(s'|s, a) \sum_{a' \in A} \pi(s', a') Q^\pi(s', a') \quad (9)$$

式中: $p(s'|s, a)$ 为状态 s 下采取动作 a 的状态转移概率; $\pi(s', a')$ 为策略 π 在状态 s' 下采取动作 a' 的概率; $Q^\pi(s', a')$ 为智能体按照策略 π 在状态 s' 下执行动作 a' 的累计奖励的估计值。

式(10)为最优 Bellman 公式, 即对应最优策略 π^* 的动作值函数 $Q^*(s, a)$ 。

$$Q^*(s, a) = r(s, a) + \gamma \sum_{s' \in S} p(s'|s, a) \{ \max_{a' \in A} Q^*(s', a') \} \quad (10)$$

2.1.2 要素定义

1) 状态空间。第 i 辆 EV 驶入充电站时的状态如下式(11)所示。

$$s_i = [l_{i,1}^p, l_{i,2}^p, \dots, l_{i,t_p}^p, \lambda_1, \lambda_2, \dots, \lambda_{t_p}, i, t_{i,arr}, t_{i,dep}] \quad (11)$$

式中: l_{i,t_p}^p 为第 i 辆 EV 在 t_p 时刻的预测负荷值; λ_{t_p} 为第 i 辆 EV 在 t_p 时刻的分时电价; $t_{i,arr}$ 和 $t_{i,dep}$ 分别为第 i 辆 EV 的到达时刻和预计离去时刻。考虑私家车使用习惯, EV 于当日傍晚驶入充电站并于次日上午驶离, 因此, 每辆 EV 驶入充电站时须得到当日及次日 48 h 的负荷预测信息及分时电价, 即 t_p 取值为 48 h。

2) 动作空间。RL 算法在第 i 辆 EV 的动作 a_i 定义为:

$$a_i = [v_1, v_2, \dots, v_{t_p}] \quad (12)$$

式中 v_{t_p} 为 t_p 时刻第 i 辆 EV 动作的优先级评定值。

值得注意的是, 每辆 EV 驶入充电站的时间和预定离开充电站的时间并不相同, 即充电时段各不相同。为方便控制, 本文对全部充电时段进行优先级评定, 之后根据 EV 各充电时段优先级控制 EV 的充电行为。

3) 奖励函数。模型的优化目标为最小化 EV 的充电开销以及所在区域的负荷波动, 故对任意选取的第 i 辆车, 设计奖励函数 r_i 为:

$$r_i = -w_1 m_i - w_2 \Delta \sigma_{i,p} \quad (13)$$

式中: m_i 为第 i 辆 EV 的充电开销; $\Delta \sigma_{i,p}$ 为第 i 辆 EV 接入后引起的负荷标准差变化量预测值; w_1 和 w_2 分别为权重值。

4) 状态转移。 s_{i+1} 为第 $i+1$ 辆 EV 的状态, 故需考虑第 i 辆 EV 接入所引起的负荷波动。本文采取

的措施是将第 i 辆 EV 的充电负荷与 l_{i,t_p}^p 直接相加作为下一状态的负荷预测值, 即 l_{i+1,t_p}^p 。

2.2 TD3 算法及求解过程

深度确定性策略梯度^[27] (deep deterministic policy gradient, DDPG) 算法以行动者-评论家 (actor-critic) 框架为基础, 可用于解决连续动作空间上的强化学习问题。DDPG 算法的策略梯度 $\nabla_\theta J(\mu_\theta)$ 如式(14)所示。

$$\nabla_\theta J(\mu_\theta) = \int \rho^\mu(s) \nabla_\theta \mu_\theta(s) \nabla_a Q^\mu(s, a) \Big|_{a=\mu_\theta(s)} ds \quad (14)$$

式中: μ 为待优化的确定性策略; μ_θ 为参数化表示的确定性策略, 其中 θ 为其参数; $\mu_\theta(s)$ 为在 s 状态下的确定性策略; $J(\mu_\theta)$ 为该行为策略的衡量函数; $\rho^\mu(s)$ 为行为策略的概率; $Q^\mu(s, a)$ 为策略 μ 在 s 状态下采取动作 a 的累计奖励的估计值。

DDPG 算法在行动者部分引入确定策略梯度, 利用式(14)更新策略网络参数, 而在评论家部分利用误差函数 $E(\cdot)$ 按照式(15)对网络参数进行更新。

$$L(\theta^Q) = E[(Q(s_t, a_t | \theta^Q) - y_t)^2] \quad (15)$$

式中: $L(\theta^Q)$ 为损失函数; $Q(s_t, a_t | \theta^Q)$ 为评论家网络; θ^Q 为 Q 网络的参数; y_t 为优化目标值。

$$y_t = r(s_t, a_t) + \gamma Q(s_{t+1}, a_{t+1} | \theta^Q) \quad (16)$$

DDPG 算法创建新的评估网络 (行动者网络 $\mu'(s | \theta^{\mu'})$ 和评论家网络 $Q'(s, a | \theta^{Q'})$), 即目标网络。DDPG 算法采用“软更新”方式更新目标网络中的参数以提高算法的稳定性。其更新规则为:

$$\theta^{\mu'} \leftarrow \alpha \theta^{\mu'} + (1 - \alpha) \theta^{\mu'} \quad (17)$$

$$\theta^{Q'} \leftarrow \alpha \theta^{Q'} + (1 - \alpha) \theta^{Q'} \quad (18)$$

式中: $\theta^{\mu'}$ 和 $\theta^{Q'}$ 为实时网络的参数; $\theta^{\mu'}$ 和 $\theta^{Q'}$ 为目标网络的参数; α 为软更新速率因子 (远小于 1)。

此外, DDPG 算法采用经验回放机制, 把智能体与环境交互的数据存储到经验池中, 随后每次训练从中随机采样一个 minibatch 以解决经验数据相关性。在输出动作时 DDPG 算法通过添加随机噪声 ε 以更好地探索潜在的最优策略, 如式(19)所示。

$$\mu^N(s_t) = \mu(s_t | \theta_t^{\mu}) + \varepsilon \quad (19)$$

式中: $\mu(s_t | \theta_t^{\mu})$ 为在 θ_t^{μ} 网络参数及策略 μ 下采取的动作; $\mu^N(s_t)$ 为添加随机噪声后的动作。

TD3 算法是 DDPG 的优化版本, 融入了 Double Q-Learning 算法的思想。与 DDPG 算法相比, TD3 采用两套 Critic 网络, 因此共包括 6 个网络, 针对

DDPG 的不足做出了以下改进。

1) 裁剪双 Q 学习。TD3 算法采用双 Critic 网络去估算 Q 值, 并取二者中的较小值计算 TD 误差, 防止 Q 值的过估计, 优化目标值 y_t 如式 (20) 所示。

$$y_t = r(s_t, a_t) + \gamma \min_{n=1,2} Q_n(s_{t+1}, a_{t+1} | \theta^Q) \quad (20)$$

2) 策略延迟更新。每更新 d 次 Critic 网络就更新一次 Actor 网络, 以增加 TD3 算法训练的稳定性。

3) 添加随机噪声。计算目标值时, TD3 算法在目标动作中加入噪声, 以增加算法的鲁棒性, 优化目标值 y_t 如式 (21) 所示。

$$\begin{cases} y_t = r(s_t, a_t) + \gamma \min_{n=1,2} Q_n(s_{t+1}, a_{t+1} | \theta^Q + \varepsilon) \\ \varepsilon \sim \text{clip}(N(0, \sigma), -c, c) \end{cases} \quad (21)$$

式中: ε 为随机噪声; c 为随机扰动范围; $\text{clip}(\cdot)$ 为限定函数, 将按正态分布 σ 的噪声限定在 $[-c, c]$ 范围内, σ 为标准差。

3 算例分析

3.1 参数设置

为验证本文所提深度强化学习方法的优化效果, 设定 EV 充电桩的充电桩数量为 200。分时电价参考广东省(珠三角五市)一月份代理购电(一般工商业, 不满 1 kV)分时电价, 如图 2 所示。此外, 本文采用第九届“中国电机工程学会杯”的负荷数据模拟 EV 充电所在区域的用电负荷。

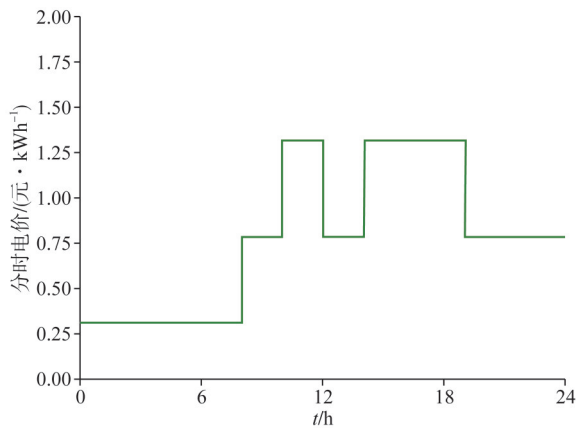


图2 分时电价

Fig. 2 Time-of-use electricity tariff

研究车辆选用市场销售量大的 BYD 系列的 BYDe6, EV 各项参数如表 1 所示。

假设 EV 一天出行结束时刻电量满足正态分布 $N(17.47, 3.41^2)$, 出行开始时刻电量满足正态分

布 $N(7.92, 3.24^2)$, 出行结束时刻电量满足正态分布 $N(0.4, 0.1^2)$ 。因此, 本文使用蒙特卡洛法生成每辆 EV 的充电需求。

表 1 BYDe6 参数

Tab. 1 BYDe6 parameters

参数	取值
电池容量/kWh	82
电池 SOC 值上限/%	90
电池 SOC 值下限/%	10
充放电效率/%	90
额定充放电功率/kW	7

3.2 训练过程

本文使用门控循环单元 (gate recurrent unit, GRU) 神经网络模型进行负荷预测, 其隐藏单元数设为 256。负荷数据集为 2 201 d 的 24 h 负荷数据, 以 5 : 5 的比例划分训练集和测试集, 使用前 72 h 负荷预测后 48 h 的负荷, 训练后的 GRU 模型在测试集上平均绝对百分比误差 (mean absolute percentage error, MAPE) 为 0.067。再将测试集以 7 : 3 的比例进行划分, 用于 TD3 模型的训练与验证。

本文设定 TD3 算法中 Actor 网络和 Critic 网络均为四层全连接网络, 隐含层神经元个数分别为 512 和 256。训练过程中 Actor 网络和 Critic 网络学习率分别设为 0.001 和 0.002, 折扣率设为 0.95, 软更新速率因子设为 0.001, 奖励函数的两个权重值分别设为 0.1 和 0.05。训练回合数设置为 3 000, 在前 200 回合智能体随机选取动作, 200 回合后智能体开始学习。

3.3 结果分析

3.3.1 有序充电效果分析

表 2 展示了本文提出的有序充电方法 (DRL 有序充电) 在测试集上的表现。从表 2 可知, 充电时段优化后每辆车平均日开销为 17.08 元, 相对于无序充电的 37.97 元节约了 55.02%。同时, 负荷标准差由无序充电时的 1 658.28 降为 1 147.62, 下降了 30.79%。

图 3 为无序充电下与本文提出的有序充电方法下的电力负荷曲线对比图。显然, 大规模 EV 的无序充电行为会加剧负荷波动, 造成“峰上加峰”现象。本文所提出的有序充电方法可通过 TD3 算法对参与调度的 EV 充电时段进行优化, 可达到削峰填

谷的效果。

表 2 EV 充电优化结果

Tab. 2 Optimization results of EV charging

场景	充电日开销/元	负荷标准差
EV 接入前		1 518. 07
EV 无序充电	37. 97	1 658. 28
DRL 有序充电	17. 08	1 147. 62

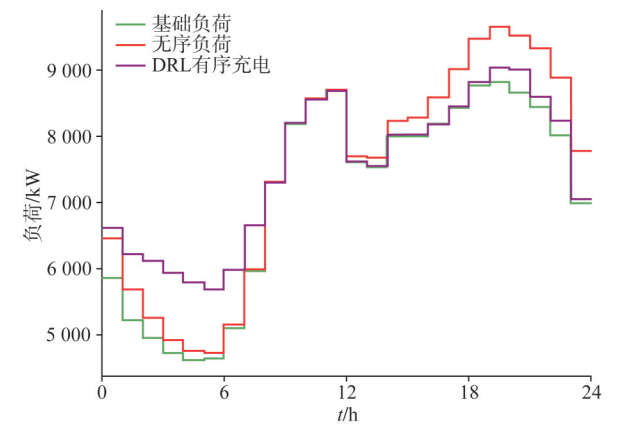


图3 充电行为优化前后负荷曲线对比

Fig. 3 Comparison of load curves before and after charging optimization

本文进一步面向EV用户参与有序调度的意愿开展深入研究,结果如表3所示。由表3可知,用户愿意参与充电优化调度的比例越高平均日开销就越少,也更能减少该区域的负荷波动。

表 3 考虑用户意愿的电动汽车调度

Tab. 3 EVs scheduling considering user's willingness

用户意愿比例/%	平均日开销/元	负荷标准差
0	37. 97	1 658. 28
25	32. 72	1 512. 75
50	27. 47	1 386. 01
75	22. 25	1 245. 15
100	17. 08	1 147. 62

3.3.2 有序充电方法对比分析

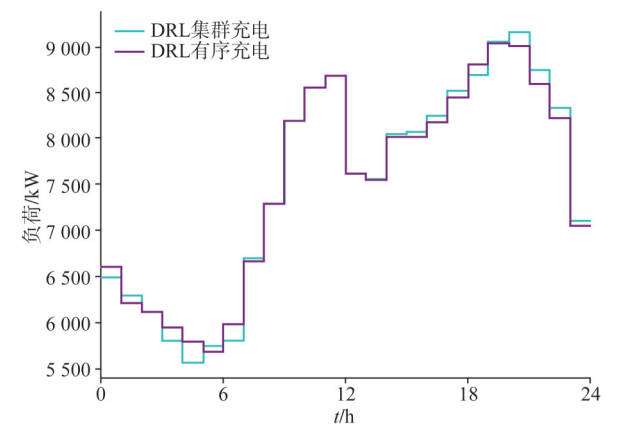
为进一步突出本文提出的有序充电方法的有效性与高效性,将本文提出的优化方法与基于DRL的集群充电优化方法^[20]进行对比,结果如表4和图4所示。

对比可知,本文提出的基于优先级思想的充电优化方法和基于集群的EV充电优化方法都能起到很好地减少用户开销和“削峰填谷”的效果。但本文提出的方法效果更好,与集群优化方法相比,可减

表 4 有序充电方法对比

Tab. 4 Comparison of ordered charging methods

优化方法	充电日开销/元	负荷标准差
DRL 集群充电	18. 75	1 194. 63
DRL 有序充电	17. 08	1 147. 62



注: DRL 有序充电代表本文所提优化方法。

图4 不同优化算法负荷曲线对比

Fig. 4 Comparison of load curves with different optimization algorithms

少8.91%的日开销和3.94%的负荷标准差。

4 结论

本文建立了考虑用户充电开销、配电网负荷标准差的EV有序充电优化模型,提出了优先级评定的思想,通过评定各时段充电优先级控制每辆EV的充电行为。本文使用TD3算法进行充电调度优化,并通过算例验证了该方法的有效性与高效性,主要结论如下。

1)本文提出优先级评定的思想,智能体整合负荷、电价、EV电量等信息,对各充电时段进行充电优先级评定,并根据此优先级控制充电桩的充电时段。该方法控制效果好,能够很好地满足用户的充电需求,不会出现充电不足或过充的情况。

2)算例结果表明,随着用户参与意愿比例的增大,EV有序充电的优化效果越好。通过控制EV在配电网负荷低谷时段有序充电,既满足了用户的出行需求,降低了用户开销,也保证了配电网的安全运行。

与基于集群的强化学习优化方法相比,本文所提方法训练周期短,收敛速度快,且节省充电费用与降低负荷波动的效果更明显,能够更好地解决

EV 充电优化问题。

参考文献

- [1] 中华人民共和国国务院. 国务院办公厅关于印发新能源汽车产业发展规划(2021—2035年)的通知[EB/OL].[2012-11-02]. http://www.gov.cn/zhengce/content/2020-11/02/content_5556716.htm.
- [2] 朱继忠, 何晨可, 刘云, 等. 面向电动汽车高渗透率综合能源系统的换电模式综述与展望[J]. 南方电网技术, 2023, 17(10): 133-151.
ZHU Jizhong, HE Chenke, LIU Yun, et al. Review and prospect of swapping pattern for integrated energy system with high penetration of electric vehicles [J]. Southern Power System Technology, 2023, 17(10): 133-151.
- [3] 王斐, 李正烁, 叶萌, 等. 电动汽车充电对电网的影响及其优化调度研究述评[J]. 南方电网技术, 2016, 10(6): 70-80.
WANG Fei, LI Zhengshuo, YE Meng, et al. Review on research of impact of electric vehicles charging on power grids and its optimal dispatch [J]. Southern Power System Technology, 2016, 10(6): 70-80.
- [4] 马玲玲, 杨军, 付聪, 等. 电动汽车充放电对电网影响研究综述[J]. 电力系统保护与控制, 2013, 41(3): 140-148.
MA Lingling, YANG Jun, FU Cong, et al. Review on impact of electric car charging and discharging on power grid [J]. Power System Protection and Control, 2013, 41(3): 140-148.
- [5] 蔡黎, 商冰洁, 高乐, 等. 电动汽车有序充电研究与应用综述[J]. 南方电网技术, 2023, 17(11): 97-107.
CAI Li, SHANG Bingjie, GAO Le, et al. Review of research and application on orderly charging of electric vehicles [J]. Southern Power System Technology, 2023, 17(11): 97-107.
- [6] 杨健维, 苟方杰, 黄宇, 等. 基于不确定性测度的居民小区电动汽车充电分时电价制定策略[J]. 电网技术, 2018, 42(1): 96-102.
YANG Jianwei, GOU Fangjie, HUANG Yu, et al. Residential area electric vehicle charging pricing strategy based on uncertainty measure [J]. Power System Technology, 2018, 42(1): 96-102.
- [7] 李旭东, 杨烨, 李帆琪, 等. 计及电价不确定性和容量衰减的电动汽车充放电商业模式[J]. 中国电力, 2023, 56(1): 38-48.
LI Xudong, YANG Ye, LI Fanqi, et al. Business models of electric vehicle aggregators considering electricity price uncertainty and capacity decay [J]. Electric Power, 2023, 56(1): 38-48.
- [8] 张延宇, 饶新朋, 周书奎, 等. 基于深度强化学习的电动汽车充电调度算法研究进展[J]. 电力系统保护与控制, 2022, 50(16): 179-187.
ZHANG Yanyu, RAO Xinpeng, ZHOU Shukui, et al. Research progress of electric vehicle charging scheduling algorithms based on deep reinforcement learning [J]. Protection and Control of Power Systems, 2023, 17(11): 97-107.
- [9] 闫丽梅, 都彬彬, 徐建军. 基于聚类分层的规模化电动汽车协调充电策略[J]. 广东电力, 2022, 35(10): 65-73.
YAN Limei, DU Binbin, XU Jianjun. Coordinated charging strategy of large-scale electric vehicles based on clustering and stratification [J]. Guangdong Electric Power, 2022, 35(10): 65-73.
- [10] 段小宇, 胡泽春, 崔岩, 等. 长时间尺度下的电动汽车有序充电调度[J]. 电网技术, 2018, 42(12): 4037-4044.
DUAN Xiaoyu, HU Zechun, CUI Yan, et al. Optimal charging and discharging strategy for electric vehicles in large timescales [J]. Power System Technology, 2018, 42(12): 4037-4044.
- [11] 胡澄, 刘瑜俊, 徐青山, 等. 面向含风电楼宇的电动汽车优化调度策略[J]. 电网技术, 2020, 44(2): 564-572.
HU Cheng, LIU Yujun, XU Qingshan, et al. Optimal scheduling strategy for electric vehicles in buildings with wind power [J]. Power System Technology, 2020, 44(2): 564-572.
- [12] 齐先军, 李冬伟, 纪姝彦. 采用功率限制的住宅区电动汽车有序充电控制策略[J]. 电网技术, 2016, 40(12): 3715-3721.
QI Xianjun, LI Dongwei, JI Shuyan. A coordinated charging control strategy for electric vehicles in residential area with power limitation [J]. Power System Technology, 2016, 40(12): 3715-3721.
- [13] 李军, 梁嘉诚, 刘克天, 等. 计及用户响应度的电动汽车充放电优化调度策略[J]. 南方电网技术, 2023, 17(8): 123-132.
LI Jun, LIANG Jiacheng, LIU Ketian, et al. Optimal scheduling strategy for electric vehicles charging and discharging considering user responsiveness [J]. Southern Power System Technology, 2023, 17(8): 123-132.
- [14] 田增垚, 刘少午, 张强, 等. 电动汽车充电站参与电网调压的建模与控制方法[J]. 南方电网技术, 2022, 16(10): 113-119.
TIAN Zengyao, LIU Shaowu, ZHANG Qiang, et al. Modeling and control methods of charging stations of electric vehicles for voltage regulations [J]. Southern Power System Technology, 2022, 16(10): 113-119.
- [15] WANG S, DUAN J, SHI D, et al. A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning [J]. IEEE Transactions on Power Systems, 2020, 35(6): 4644-4654.
- [16] PARK K, MOON I. Multi-agent deep reinforcement learning approach for EV charging scheduling in a smart grid [J]. Applied Energy, 2022(328), 1-10.
- [17] 罗欣儿, 杜进桥, 田杰, 等. 基于深度强化学习的主动配电网高恢复力决策方法[J]. 南方电网技术, 2022, 16(1): 67-74.
LUO Xiner, DU Jinqiao, TIAN Jie, et al. High resilience decision-making method of active distribution network based on deep reinforcement learning [J]. Southern Power System Technology, 2022, 16(1): 67-74.
- [18] 陈卫东, 吴宁, 黄彦璐, 等. 基于深度学习的微电网优化调度辅助决策方法[J]. 南方电网技术, 2022, 16(1): 117-126.
CHEN Weidong, WU Ning, HUANG Yanlu, et al. Auxiliary decision-making method of optimal dispatching for microgrid based on deep learning [J]. Southern Power System Technology, 2022, 16(1): 117-126.
- [19] 李航, 李国杰, 汪可友. 基于深度强化学习的电动汽车实时调度策略[J]. 电力系统自动化, 2020, 44(22): 161-167.
LI Hang, LI Guojie, WANG Keyou. Real-time dispatch strategy for electric vehicles based on deep reinforcement learning [J]. Automation of Electric Power Systems, 2020, 44(22): 161-167.
- [20] 廖鑫, 李婧, 徐佳, 等. 基于MDP及激励需求响应的电动汽车有序充电控制[J]. 电力科学与技术学报, 2021, 36(5): 79-86.
LIAO Xin, LI Jing, XU Jia, et al. Research on coordinated charging control for electric vehicles based on MDP and incen-

- tive demand response [J]. Journal of Electric Power Science and Technology, 2021, 36(5): 79 – 86.
- [21] 杜明秋,李妍,王标,等.电动汽车充电控制的深度增强学习优化方法[J].中国电机工程学报,2019,39(14):4042 – 4049.
- DU Mingqiu, LI Yan, WANG Biao, et al. Deep reinforcement learning optimization method for charging control of electric vehicles [J]. Proceedings of the CSEE, 2019, 39(14): 4042 – 4048.
- [22] 赵小瑾,张开宇,冯冬涵,等.基于强化学习的电动汽车集群实时优化调度策略[J].智慧电力,2022,50(1):53 – 59,81.
- ZHAO Xiaojin, ZHANG Kaiyu, FENG Donghan, et al. Real-time optimal scheduling strategy for electric vehicle clusters based on reinforcement learning [J]. Smart Power, 2022, 50(1): 53 – 59, 81.
- [23] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor [C]// International Conference on Machine Learning. July 10 – 15, 2018, Stockholm, Sweden. Stockholm: PMLR, 2018: 1861 – 1870.
- [24] 赵星宇,胡俊杰.集群电动汽车充电行为的深度强化学习优化方法[J].电网技术,2021,45(6):2319 – 2327.
- ZHAO Xingyu, HU Junjie. Deep reinforcement learning based optimization for charging of aggregated electric vehicles [J]. Power System Technology, 2021, 45(6): 2319 – 2327.
- [25] FUJIMOTO S, VAN H H, MEGER D. Addressing function approximation error in actor-critic methods [C]// International Conference on Machine Learning. July 10 – 15, 2018, Stockholm, Sweden. Stockholm: PMLR, 2018: 1587 – 1596.
- [26] JIN J, XU Y. Optimal policy characterization enhanced actor-critic approach for electric vehicle charging scheduling in a power distribution network [J]. IEEE Transactions on Smart Grid, 2020, 12(2): 1416 – 1428.
- [27] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning [C]// 2016 the International Conference on Learning Representations, May 2 – 4, 2016, San Juan, Puerto Rico. San Juan: ICLR, 2016: 1 – 10.

收稿日期:2023-05-31;网络首发日期:2024-01-05

作者简介:

喻磊(1985),男,通信作者,高级工程师,博士,研究方向为分布式能源并网及微电网,智能配电技术, yulei@csg.cn;
谈竹奎(1975),男,高级工程师(教授级),硕士,研究方向为能源互联网、需求响应、柔性直流配网, tanzk@163.com;
王扬(1993),男,工程师,硕士,研究方向为低压交直流配电技术、分布式能源并网运行分析, yangwstu@163.com。

(上接第 106 页 Continued from Page 106)

- [24] YANG L, YANG J. Robust current control method for LCL-type shunt active power filters with inverter-side current feedback active damping [C]// 2018 IEEE Energy Conversion Congress and Exposition (ECCE), September 12 – 13, 2018, Hangzhou, China. New York: IEEE, 2018: 5706 – 5712.
- [25] 郭强,何黎鹏,肖蕙蕙,等.一种实现电流型PWM整流器直流侧电压波动抑制的控制方法[J].仪器仪表学报,2023,44(6):313 – 324.
- GUO Qiang, HE Lipeng, XIAO Huihui, et al. Control method for suppressing DC side voltage fluctuation of current-source PWM rectifier [J]. Chinese Journal of Scientific Instrument, 2023, 44(6): 313 – 324.
- [26] 阮新波,王学华,潘冬华,等. LCL 型并网逆变器的控制技术[M].北京:科学出版社,2015.
- [27] 王延伟,刘宇明,杨荣照,等.柔性直流输电系统高频谐振自适应抑制策略[J].南方电网技术,2024,18(2):1 – 11.
- WANG Yanwei, LIU Yuming, YANG Rongzhao, et al. Adaptive suppression strategy of high frequency resonance in MMC-HVDC transmission system [J]. Southern Power System Technology, 2024, 18(2): 1 – 11.
- [28] 中国电力科学研究院.分布式电源接入电网技术规定:Q/GDW 480—2010[S].北京:国家电网公司,2010.
- [29] PANG B, ZHU X, YANG J, et al. Voltage harmonics optimization for weak grid-tied doubly-fed induction generator with the capability of suppressing current harmonics [J]. IEEE Transactions on Energy Conversion, 2023, 38(2): 1452 – 1460.
- [30] 黄伟,刘蔚,黄润,等.利用PSS4B和GPSS附加阻尼控制抑制云南电网超低频振荡的对比分析[J].广东电力,2022,35(10):74 – 82.
- HUANG Wei, LIU Wei, HUANG Run, et al. Comparative analysis of suppressing ultra-low frequency oscillation of Yunnan Power Grid by using additional damping control of PSS4B and GPSS [J]. Guangdong Electric Power, 2022, 35(10): 74 – 82.

收稿日期:2024-01-11;网络首发日期:2024-06-01

作者简介:

刘可(1988),男,高级工程师,硕士,研究方向为电压管理与电能质量监测, 397553383@qq.com;
赵正奎(1988),男,博士生导师,博士,从事新能源并网系统建模及稳定性分析研究, zhaozhengkui@qhu.edu.cn;
张文倩(1994),女,工程师,硕士,从事电能质量监测研究, 505390676@qq.com。

征 稿 启 事

《南方电网技术》是由中国南方电网公司主管、南方电网科学研究院有限责任公司主办的电力技术类科技期刊,向国内外公开发行人。刊登电力系统的科研、规划、设计、工程、生产运行、设备和系统维护等方面的成果、经验和动态,尤其是科研创新方面的论文。设有特约专稿、高压直流输电、高电压与绝缘、系统分析与运行、控制与保护、调度与通信、输变电技术、电力技术经济、智能电网、分布式电源与微电网、低碳电力等栏目。热烈欢迎国内外的电力研究人员、工程技术人员、大专院校师生踊跃投稿。

1 文稿要求

1) 文稿应未在其他刊物上发表过,译文稿件须附上原作者的书面许可证明。内容要有科学性、创新性和实用性;论点明确、数据可靠、说理严谨、数学推导简明;语言流畅、文字简练、层次分明、重点突出。

2) 学术论文请按 GB 7713—1987《科学技术报告、学位论文和学术论文的编写格式》书写,篇幅以版面不超过 9 000 字(包括图表所占篇幅)为宜。同时请写明如下 a、b、c、d、e 五项:

a) 作者署名(不宜超过 7 位)、作者单位、地址、电话、email 信箱和邮政编码;

b) 论文题目和作者单位英译文,作者姓名的汉语拼音;

c) 作者的简历,包括姓名、出生年、性别、籍贯、职称、学位、当前从事的研究或工作等项;

d) 350 字左右的中文摘要,3~8 个关键词,以及摘要和关键词相应的英文文本。摘要采用第三人称写法,内容包括论文的目的、方法、结果和结论;

e) 若为国家基金资助项目或部、省级重大科研攻关项目,请提供项目名称及编号。

2 书写格式

1) 要求文稿正文字号为 5 号,汉字字体为宋体,英数字体为 Times New Roman,希腊字体为 Symbol。文稿须采用可编辑的电子文档形式。

2) 计量单位采用国家法定计量单位和符号,不能用“大气压”、“kgf/cm²”、“卡”、“ppm”等已废除的计量单位。

3) 文中和公式中容易混淆的字符请注明文种(希文、英文、罗马字等)、大小写、上下标及上下标字母含义,表示矢量及矩阵的字符用黑斜体表示或特别注明。

4) 文稿标题中不宜用缩略词(化学符号和公知公用者除外);摘要和正文中的缩略词在第一次出现时必须写出全称,后加括号附缩略词。

3 表格、插图及参考文献

1) 表格尽量采用“三线表”。表格的上方写表序和表名。表名应有自明性且中文、英文表名并列。表注放在表底,缩入 1 个汉字以“注:”起头排版。

2) 插图的下放应有图序和图名。图名应有自明性且中英文图名并列。工程图、电气图和函数图采用 AutoCAD、

Adobe Illustrator 或 Corel DRAW 软件绘制,工程图和电气图要符合国家标准规定,函数图要标明曲线序号及其注释,坐标轴上要有标值、标目,注明物理量和单位;数码照片图具备 200 万总像素以上。

3) 图片和表格中的文字、变量、单位和数字要齐全、清晰。

4) 参考文献应选用公开发表的资料,按引用的先后次序列表于文后,以[1]、[2]…标识序号,且与正文中的指示序号一致。按《信息与文献 参考文献著录规则,GB/T 7714—2015》和《中国学术期刊(光盘版)检索与评价数据规范,CAJ-CD B/T 1—2006 修订版试行稿》的要求著录文后参考文献。中文参考文献后要并列其英文译文(如示例[6]所示)。

文献类型及其标识为:普通图书[M];会议论文[C];报纸文章[N];期刊文章[J];学位论文[D];报告[R];标准[S];专利[P];汇编[G];档案[B];古籍[O];参考工具[K];其他未说明的文献类型,例如可公开的政府行政部门编号文件、行业或大公司的技术规范或工作手册[Z]。网上期刊[J/OL];网上电子公告[EB/OL]。电子文献尚需在载体标记后加上发表或更新日期(加圆括号)、引用日期(加方括号)和电子文献网址。

4 投稿

1) 请登录《南方电网技术》编辑部网站 <https://nfdwjs.csg.cn>,选择“作者登录”,登录或注册登录后选择投稿并按提示操作,登录方面出现问题请致电 020-36625646 咨询。

2) 投稿时务必附上联系电话(手机为佳)。

3) 编辑部不保证留存稿件,请作者自留底稿。稿件评审进程和结论可在《南方电网技术》编辑部网站上即时查询。作者也可致电 020-36625642、36625643 或发送电子邮件到 nfdwjs@csg.cn 向编辑部了解审稿情况。

4) 文章一经发表,编辑部便取得作者的自动授权,可将文章收录入本刊网站,及与本刊有合作关系的数据库。

参考文献著录格式示例:

- [1] 刘国钧,陈绍业,王凤翥. 图书馆目录[M]. 北京:高等教育出版社,1957:15-18.
- [2] 钟文发. 非线性规划在可燃毒物配置中的应用[C]//赵玮. 运筹学的理论与应用——中国运筹学会第五届大会论文集. 西安:西安电子科技大学出版社,1996:468-471.
- [3] 冯西桥. 核反应堆压力管道与压力容器的 LBB 分析[R]. 北京:清华大学核能技术设计研究院,1997.
- [4] 中国国家标准化管理委员会. 信息与文献 参考文献著录规则:GB/T 7714—2015[S]. 北京:中国标准出版社,2015.
- [5] 邓一刚. 全智能节电器:200610171314.3[P]. 2006-12-13.
- [6] 陈祥. 无线技术在现代电网管理中的应用与展望[J]. 南方电网技术,2009,3(6):85-89.
CHEN Xiang. Development of wireless technology in modern power grid management and its perspective [J]. Southern Power System Technology, 2009,3(6):85-89.
- [7] 王明亮. 关于中国学术期刊标准化数据库系统工程的进展[EB/OL]. (1998-08-16) [1998-10-04]. <http://www.cajcd.edu.cn/pub/wml.txt/980810-2.html>.

《南方电网技术》编辑部