

# 基于深度强化学习的大规模电动汽车充换电负荷优化调度

刘敦楠<sup>1</sup>, 王玲湘<sup>1</sup>, 汪伟业<sup>1</sup>, 李 华<sup>1</sup>, 王 文<sup>2</sup>, 刘明光<sup>1</sup>

(1. 华北电力大学经济与管理学院, 北京市 102206; 2. 国网电动汽车服务有限公司, 北京市 100053)

**摘要:** 针对大规模充换电站的聚合优化调度问题,提出一种基于SAC深度强化学习的充换电负荷实时优化调度策略。该策略充分考虑了负荷调控过程中的用户因素、系统因素和市场因素,能够实现大规模电动汽车与各类电力系统主体的友好互动。首先,考虑充换电站的发展规模和调度性能建立联合运行框架;其次,提出考虑多重用户特征的可调性识别模型对电动汽车的实际可调性进行判断;进而,考虑充换电站优化调度的多重时空特征,构建不同场景下可调充换电负荷的优化调度模型;然后,基于SAC算法求解并网充换电负荷的实时调度方案;最后,以电动汽车聚合优化虚拟电厂负荷为例,验证了SAC算法应用于大规模电动汽车充换电负荷实时优化调度的经济性和高效性。

**关键词:** 深度强化学习; 电动汽车; 充换电协调; 实时优化调度; 电动汽车聚合商

## 0 引言

2020年9月,国家发展改革委发布《关于扩大战略性新兴产业投资 培育壮大新增长点增长极的指导意见》<sup>[1]</sup>,鼓励智能汽车与智慧城市协同发展。电动汽车(EV)作为重要的灵活性可调资源,能够与多种能源协同调度<sup>[2-3]</sup>,为电力系统的经济可靠运行提供支撑<sup>[4-5]</sup>。然而,在电动汽车灵活性资源的实际应用过程中也面临着一些重要问题,其中电动汽车可调资源分析<sup>[6-7]</sup>和智能协调调度方法设计<sup>[8-9]</sup>是两项关键问题。

目前,电动汽车灵活性负荷的研究通常围绕电动汽车充放电的优化调度展开。例如,文献[10]采用分群优化的方法控制电动汽车充放电过程,从而实现电动汽车对电网负荷的调节;文献[11]通过控制规模化电动汽车的充放电,实现电动汽车与风电、火电的协同优化;文献[12]基于用户充电距离等因素对用户满意度的影响,对电动汽车入网的可调性进行判断。但上述研究都是直接对电动汽车的优化调度过程进行研究,未结合用户可调性识别,也并未考虑到决策主体是电动汽车聚合商(electric vehicle aggregator, EVA),而EVA的优化调度过程中还具

有另一种可调节资源——换电资源,随着充换电一体化建设进程的推进,换电负荷将在聚合调度中发挥重要的作用<sup>[13-15]</sup>。

在电动汽车智能协调调度方法的研究中,广泛使用的算法有交替方向乘子法(alternating direction method of multipliers, ADMM)<sup>[16-17]</sup>、拟牛顿法(BFGS)、启发式算法<sup>[18-19]</sup>、约束优化求解算法<sup>[20]</sup>、人工智能算法<sup>[21]</sup>等。其中,ADMM、BFGS、启发式算法和约束优化求解算法都是基于全局环境状态对现有决策变量进行求解,具有计算复杂度高、决策空间大、容易陷入局部最优、无法区分个体调度方案、环境数据需实时更新等缺陷,无法求解大规模电动汽车的实时调度问题。而强化学习等人工智能算法通过神经网络对大数据的拟合能力,避免了传统算法重复的高复杂度的计算<sup>[22]</sup>,可以在未知全局用户接入的情况下对单个用户做出合理的全局较优的调度安排,解决了传统算法对数据的实时依赖性高的缺陷,具有优化效果好、鲁棒性强、计算效率高的特征,适用于大规模电动汽车实时优化调度的求解。

基于上述研究内容,考虑目前充换电一体化建设进程,本文首先从EVA的角度设计了电动汽车充换电过程的联合运行框架;进而,考虑用户参与电动汽车负荷优化调度意愿和可调节潜力,设计考虑多重用户特征的可调性识别模型;然后,提出电动汽车充换电负荷的实时优化调度模型;最后,基于SAC(soft actor-critic)强化学习算法对充换电负荷进行

收稿日期: 2021-05-23; 修回日期: 2021-06-22。

上网日期: 2021-08-30。

国家社科基金重大项目(19ZDA081); 国家电网公司科技项目(5418-202018247A-0-0-00)。

实时优化调度,并采用真实数据对所提模型的有效性进行验证。

## 1 电动汽车充换电站联合运行框架

电动汽车充换电负荷联合优化调度的基本框架如图1所示,其中EVB表示换电站电池。

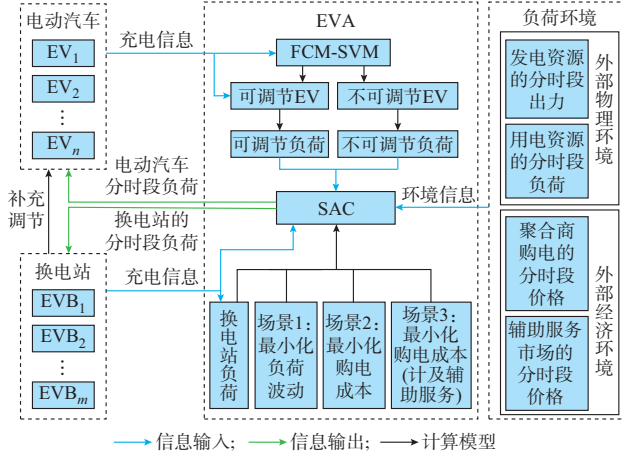


图1 充电桩与换电站联合优化调度框架

Fig. 1 Joint optimal scheduling framework of charging pile and battery swapping station

本文所提优化调度框架主要从EVA的角度出发,考虑充电桩与换电站两大资源主体进行资源优化调度,其特征具体如下:

1)在EVA内外部关系方面,调度框架包括EVA内部资源、EVA和外部电力系统,EVA通过聚合调度内部充换电负荷实时响应外部电力系统运行的变化情况。

2)EVA内部调度方法主要包括3个阶段:(1)充换电负荷信息录入;(2)可调节电动汽车负荷识别;(3)电动汽车充换电负荷协调调度。

3)EVA与外部环境的协调主要设置3种典型场景:场景1是EVA作为微网或者相关联合体中的一部分进行调度,以平抑负荷波动;场景2是EVA作为独立个体,仅接受分时价格价格的引导,调节负荷以实现最小充电成本;场景3是EVA作为参与辅助服务的市场主体,兼顾辅助服务市场效益与购电成本进行调度,以实现EVA经济性最优。

据此研究电动汽车充换电负荷的实时优化调度策略,并分析其对于电力系统经济可靠运行的贡献。

## 2 电动汽车充电站调度策略

### 2.1 可调节电动汽车用户识别

电动汽车接入充电设备后具有负荷调节的可能

性。因此,本文考虑一种基于模糊C均值(fuzzy C-means, FCM)可调隶属关系确认与支持向量机(support vector machine, SVM)智能识别的电动汽车用户可调性识别方法,以智能识别可调节电动汽车和不可调节电动汽车。

影响电动汽车响应能力的主要因素包括充电功率、荷电状态(SOC)、充电时长(达到理想荷电状态的最短时长)和接入时长(入网到离网时间)。而用户行为偏好则通过车型、电池容量、日均电力需求、日均入网时长、用户积分/优惠券使用情况等特征参数来进行反映。基于此,设置电动汽车充电可调性特征元组如式(1)所示。

$$U = (u_1, u_2, u_3, u_4, u_5) \quad (1)$$

式中: $u_1$ 为车型特征参数; $u_2$ 为电池容量特征参数; $u_3$ 为电力需求特征参数; $u_4$ 为用户接入时长特征参数; $u_5$ 为用户激励特征参数,为0-1变量。对应的计算方式如式(2)所示。

$$u_i = \frac{\bar{g}_i}{G} \quad i = 1, 2, \dots, 5 \quad (2)$$

式中: $\bar{g}_i$ 为第*i*个特征参数下用户对特征类别的车-网友友好互动的平均响应次数; $G$ 为EVA组织的车-网友友好互动总次数。

采用FCM对典型电动汽车充电可调性特征元组的样本进行聚类分析,可以得到有意愿参与响应与无意愿参与响应的聚类中心。聚类过程中目标函数见式(3),约束条件见式(4)。

$$\min \sum_{w=1}^W \sum_{v=1}^2 \alpha_{w,v}^\gamma \|U_w - Z_v\|^2 \quad (3)$$

$$\sum_{v=1}^2 \alpha_{w,v}^\gamma = 1 \quad w = 1, 2, \dots, W \quad (4)$$

式中: $W$ 为样本总数; $w$ 为样本数标识; $v$ 为类别数; $U_w$ 为具有5维特征的样本; $Z_v$ 为具有5维特征的第*v*簇的中心; $\alpha_{w,v}$ 为样本 $U_w$ 属于第*v*类的隶属度,单个样本 $U_w$ 对每个簇的隶属度之和为1; $\gamma$ 为隶属度因子; $\|U_w - Z_v\|$ 表示负荷样本到聚类中心的欧氏距离。

利用式(3)和式(4)可推导出隶属度 $\alpha_{w,v}$ 的计算式(5)和簇中心 $Z_v$ 的计算式(6),推导过程见附录A。

$$\alpha_{w,v} = \frac{1}{\sum_{k=1}^2 \left( \frac{\|U_w - Z_k\|}{\|U_w - Z_v\|} \right)^{\frac{2}{\gamma-1}}} \quad (5)$$

$$Z_v = \frac{\sum_{w=1}^W (\alpha_{w,v}^\gamma U_w)}{\sum_{w=1}^W \alpha_{w,v}^\gamma} \quad (6)$$

式中: $k$ 为类别标识,实际含义与 $v$ 一致。

迭代至满足式(7)时终止。

$$\max_{w,v} (\alpha_{w,v}^{(n+1)} - \alpha_{w,v}^{(n)}) < \varepsilon \quad (7)$$

式中: $n$ 为迭代次数; $\varepsilon$ 为误差阈值。

得到簇中心分别如式(8)与式(9)所示。

$$Z_1 = (z_{1,1}, z_{1,2}, z_{1,3}, z_{1,4}, z_{1,5}) \quad (8)$$

$$Z_2 = (z_{2,1}, z_{2,2}, z_{2,3}, z_{2,4}, z_{2,5}) \quad (9)$$

据此,定义电动汽车参与响应隶属度与不参与响应隶属度分别如式(10)与式(11)所示。

$$d_1 = \sqrt{\sum_{i=1}^5 \frac{(u_i - z_{1,i})^2}{z_{1,i}}} \quad (10)$$

$$d_2 = \sqrt{\sum_{i=1}^5 \frac{(u_i - z_{2,i})^2}{z_{2,i}}} \quad (11)$$

式中: $z_{1,i}$ 和 $z_{2,i}$ 分别表示簇中心 $Z_1$ 和 $Z_2$ 的第 $i$ 个元素。

进而,以样本电动汽车用户充电可调性特征元组作为输入,以样本电动汽车的实际可调属性作为输出,即可构造映射关系。采用典型数据训练SVM后<sup>[23-24]</sup>,即可利用训练后的FCM-SVM模型对电动汽车接入情况进行实时可调性判断。

## 2.2 不可调节与可调节负荷

基于可调性判断结果可对电动汽车的可调节负荷和不可调节负荷进行测算。

### 2.2.1 不可调节负荷

不可调节负荷的计算方法如式(12)所示。

$$Q_{u,t} = \sum_{s \in M} E_{u,s,t} \quad (12)$$

式中: $Q_{u,t}$ 为 $t$ 时段的不可调节负荷; $E_{u,s,t}$ 为 $t$ 时段第 $s$ 辆不可调节电动汽车的充电电量; $M$ 为不可调节电动汽车集合。

对于不可调节电动汽车来说,其充电过程具有如式(13)和式(14)所示特征。

#### 1) 电量特征

$$E_{u,s,t} = p_{\text{fast}} u_{2,s,t} \quad (13)$$

式中: $p_{\text{fast}}$ 为最大充电功率; $u_{2,s,t}$ 为 $t$ 时段第 $s$ 辆不可调节电动汽车的充电时长。

#### 2) 功率特征

不可调节电动汽车的充电功率按照电动汽车车主设置或者最大充电功率充电。

### 3) 充电时间特征

不可调节电动汽车的充电时长如式(14)所示,接入时长由车主意愿决定。

$$T_{s,\text{in}} = \frac{E_{u,s}^{\text{need}}}{p_{\text{fast}}} \quad (14)$$

式中: $E_{u,s}^{\text{need}}$ 为第 $s$ 辆不可调节电动汽车的最小充电需求; $T_{s,\text{in}}$ 为第 $s$ 辆不可调节电动汽车的充电总时长。

### 2.2.2 可调节负荷

可调节负荷的计算方法如式(15)所示。

$$Q_{c,t} = \sum_{p \in N} E_{c,p,t} \quad (15)$$

式中: $Q_{c,t}$ 为 $t$ 时段的可调节负荷; $E_{c,p,t}$ 为 $t$ 时段第 $p$ 辆可调节电动汽车的充电电量; $N$ 为可调节电动汽车集合。

对于可调节电动汽车来说,其充电过程具有如式(16)至式(18)所示特征。

#### 1) 充电电量特征

$$0 \leq E_{c,p,t} \leq p_{\text{fast}} u_{2,p,t} \quad (16)$$

式中: $u_{2,p,t}$ 为 $t$ 时段第 $p$ 辆可调节电动汽车的充电时长。

#### 2) 功率特征

$$0 \leq p_{c,p,t} \leq p_{\text{fast}} \quad (17)$$

式中: $p_{c,p,t}$ 为 $t$ 时段第 $p$ 辆可调节电动汽车的充电功率。

#### 3) 电动汽车电量特征:

$$E_{c,p}^{\text{need}} \leq \sum_{T_{p,\text{in}} \leq t \leq T_{p,\text{out}}} E_{c,p,t} = \int_{T_{p,\text{in}} \leq t \leq T_{p,\text{out}}} p_{c,p,t} dt \leq (1 - u_{4,p}) E_{c,p}^{\text{max}} \quad (18)$$

式中: $E_{c,p}^{\text{need}}$ 为第 $p$ 辆可调节电动汽车的最小充电需求; $T_{p,\text{in}}$ 为第 $p$ 辆可调节电动汽车的接入时间; $T_{p,\text{out}}$ 为第 $p$ 辆可调节电动汽车的离开时间; $u_{4,p}$ 为第 $p$ 辆可调节电动汽车的入网时刻的荷电状态; $E_{c,p}^{\text{max}}$ 为第 $p$ 辆可调节电动汽车的电池容量。

## 2.3 可调节电动汽车的最优化调度模型

基于用户的充电需求与用户特征元组 $U$ ,可以对接入的电动汽车进行实时优化调度。据此,设置根据电动汽车在3种典型场景下的聚合优化调度目标函数。

### 1) 场景1

基于电动汽车聚合平抑负荷波动的优化调度目标函数如式(19)所示。该场景下可以基于现有负荷基线(外部物理环境总负荷与不可调节负荷的和),

每加入1辆电动汽车,为其设计最小化负荷波动的电动汽车充电方案。

$$\min \frac{\sum_{t=1}^{96} \left( Q_{u,t} + Q_{\text{load},t} + \sum_{p \in N} E'_{c,p,t} - Q_{\text{ave}} \right)^2}{Q_{\text{ave}}^2} \quad (19)$$

式中: $Q_{\text{load},t}$ 为 $t$ 时段不计电动汽车的系统负荷值; $E'_{c,p,t}$ 为调节后的电动汽车在 $t$ 时段的负荷值; $Q_{\text{ave}}$ 为系统所有负荷的均值,其计算式见式(20)。

$$Q_{\text{ave}} = \frac{\sum_{t=1}^{96} (Q_{u,t} + Q_{\text{load},t} + Q_{c,t})}{96} \quad (20)$$

### 2) 场景 2

基于EVA购电成本最低的优化调度目标函数见式(21)。基于原有负荷曲线的充电成本,每加入1辆电动汽车,为其设计最小化充电成本的电动汽车充电方案。

$$\min \left[ \left( Q_{u,t} + \sum_{p \in N} E'_{c,p,t} \right) p_{a,t} \right] - \sum_{t=1}^{96} (Q_{\text{tar},t} p_{a,t}) \quad (21)$$

式中: $p_{a,t}$ 为电力市场在 $t$ 时段的购电价格; $Q_{\text{tar},t}$ 为电动汽车在 $t$ 时段的原始负荷值。

### 3) 场景 3

考虑辅助服务市场的优化调度目标函数见式(22),每加入1辆电动汽车,对其设计最小化充电成本的充电方案。

$$\min \left[ \left( Q_{u,t} + \sum_{p \in N} E'_{c,p,t} \right) p_{a,t} \right] - \sum_{t=1}^{96} (Q_{\text{tar},t} p_{a,t}) - \sum_{t \in T_{\text{au}}} \left[ \left( Q_{u,t} + \sum_{p \in N} E'_{c,p,t} \right) p_{\text{au},t} \right] + \sum_{t \in T_{\text{au}}} (Q_{\text{au},t} p_{\text{au},t}) \quad (22)$$

式中: $T_{\text{au}}$ 为辅助服务市场的开市时段; $p_{\text{au},t}$ 为 $t$ 时段的辅助服务市场价格; $Q_{\text{au},t}$ 为 $t$ 时段的电动汽车参与辅助服务市场的负荷基线。

## 3 电动汽车换电站调度策略

### 3.1 换电站行为约束

换电站充换电过程中电池总量保持不变。因此,令换电站内所有电池具有编号 $o$ ,将换电站内换出电池编号赋予换入电池。

换电站总负荷计算如式(23)所示。

$$Q_{b,t} = \sum_{o=1}^O (E_{\text{ex},o,t} - E_{\text{ex,ou},o,t}) \quad (23)$$

式中: $Q_{b,t}$ 为 $t$ 时段换电站的总负荷; $E_{\text{ex},o,t}$ 为 $t$ 时段第 $o$ 组电池的充电电量; $O$ 为换电站电池总组数;

$E_{\text{ex,ou},o,t}$ 为 $t$ 时段第 $o$ 组电池的放电电量。

换电站运行过程中,具有如式(24)至式(32)所示的特征约束。

#### 1) 电池的充电情况约束

$$p_{\min} \omega_o \Delta t \leq E_{\text{ex},o,t} \leq p_{\max} \omega_o \Delta t \quad (24)$$

$$\omega_o = \begin{cases} 1 & \text{电池接入充电机} \\ 0 & \text{电池未接入充电机} \end{cases} \quad (25)$$

式中: $\omega_o$ 为第 $o$ 组电池是否接入充电机的标识; $p_{\min}$ 和 $p_{\max}$ 分别为充电机的最小、最大充电功率; $\Delta t$ 为每时段的时长。

#### 2) 换电站电池放电约束

$$p_{\min, \text{out}} \omega_o \Delta t \leq E_{\text{ex,ou},o,t} \leq p_{\max, \text{out}} \omega_o \Delta t \quad (26)$$

$$C_{\text{out},o,t} \leq R_{\text{out},o,t} \quad (27)$$

式中: $p_{\min, \text{out}}$ 和 $p_{\max, \text{out}}$ 分别为充电机的最小、最大放电功率; $C_{\text{out},o,t}$ 为 $t$ 时段第 $o$ 组电池的放电成本; $R_{\text{out},o,t}$ 为 $t$ 时段第 $o$ 组电池的放电收入。放电过程中的成本与收益计算式见附录B。

#### 3) 电池的接入情况约束

$$\sum_{o=1}^O \omega_o \leq I \quad (28)$$

式中: $I$ 为换电站充电机总数。

#### 4) 电池电量约束

$$S_{\text{SOC},o,t} = S_{\text{SOC},o,t-1} + \frac{E_{\text{ex},o,t} - E_{\text{ex,ou},o,t}}{Q_o} \quad (29)$$

$$0 \leq S_{\text{SOC},o,t} \leq 1 \quad (30)$$

$$E_{\text{ex},o,t} + E_{\text{ex,ou},o,t} \neq 0 \quad \forall t \in T_o \quad (31)$$

式中: $S_{\text{SOC},o,t}$ 为 $t$ 时段第 $o$ 组电池的SOC; $Q_o$ 为第 $o$ 组电池的电池容量; $T_o$ 为第 $o$ 组电池接入充电设备的总时间。式(30)保证电池充放电不超出电池容量约束;式(31)保证电池接入充电机过程中不造成资源浪费,未充满电的电池不离网。

#### 5) 换电站充放电电池供应约束

$$C_{\text{ex},t} \leq B_{\text{ex},t} \quad (32)$$

式中: $C_{\text{ex},t}$ 为 $t$ 时段换电站需完成的换电电池量; $B_{\text{ex},t}$ 为 $t$ 时段换电站已有的满电电池量。

换电站充放电过程中具有满电的电池数量计算式如式(33)、式(34)所示。

$$S_{o,t} = \begin{cases} 1 & S_{\text{SOC},o,t} = 1 \\ 0 & S_{\text{SOC},o,t} < 1 \end{cases} \quad (33)$$

$$B_{\text{ex},t} = \sum_{o=1}^O S_{o,t} \quad (34)$$

式中: $S_{o,t}$ 为 $t$ 时段第 $o$ 组电池是否满电的判断结果。

### 3.2 换电站优化调度模型

换电站通过更换电池来满足电动汽车电量需



求,具有充放电灵活性较高、电池更换速度快的特点。换电站协调充电站进行优化调度时,考虑其容量规模相对较小<sup>[25-26]</sup>,将其作为一种补充资源协调

$$\min \frac{\sum_{t=1}^{96} \left[ Q_{u,t} + Q_{load,t} + Q_{c,t} + \sum_{o \in O} (E_{ex,o,t} - E_{ex,out,o,t}) - Q'_{ave} \right]^2}{(Q'_{ave})^2} \quad (35)$$

式中: $Q'_{ave}$ 为 $t$ 时段加入换电站负荷后的负荷均值,其计算式如式(36)所示。

$$Q'_{ave} = \frac{\sum_{t=1}^{96} (Q_{u,t} + Q_{load,t} + Q_{c,t} + Q_{b,t})}{96} \quad (36)$$

对于场景2,其优化调度函数如式(37)所示。

$$\min \left[ \left( Q_{u,t} + Q_{c,t} + \sum_{o \in O} (E_{ex,o,t} - E_{ex,out,o,t}) \right) p_{a,t} \right] - \sum_{t=1}^{96} (Q'_{tar,t} p_{a,t}) \quad (37)$$

式中: $Q'_{tar,t}$ 为加入换电站负荷后 $t$ 时段的原始负荷值。

对于场景3,其优化调度函数如式(38)所示。

$$\min \left[ \left( Q_{u,t} + Q_{c,t} + \sum_{o \in O} (E_{ex,o,t} - E_{ex,out,o,t}) \right) p_{a,t} \right] - \sum_{t=1}^{96} (Q'_{tar,t} p_{a,t}) - \left[ \left( Q_{u,t} + Q_{c,t} + \sum_{o \in O} (E_{ex,o,t} - E_{ex,out,o,t}) \right) p_{a,t} \right] + \sum_{t \in T_{au}} (Q'_{au,t} p_{au,t}) \quad (38)$$

式中: $Q'_{au,t}$ 为加入换电站负荷后的辅助服务市场的负荷基线。

## 4 基于SAC强化学习的优化调度算法

SAC算法是由Haarnoja等人<sup>[27-28]</sup>提出的强化学习算法,通过引入最大熵鼓励策略,在提高算法的鲁棒性的同时加快了训练速度,能够在复杂的电力供需环境中为大规模电动汽车作出精准有效的调度决策。

### 4.1 SAC目标函数

SAC的目标函数要求策略在实现最大化累计回报的同时最大化策略熵,从而避免学习过程中的贪婪采样而陷入局部最优。据此,构建其目标函数 $\pi_{\max}^*$ 如式(39)所示。

$$\pi_{\max}^* = \arg \max_{\pi} \sum_{t=1}^T E_{(s_q, a_q) \sim p_{\pi}} (r(s_q, a_q) + \alpha \mathcal{H}(\pi(\cdot | s_q))) \quad (39)$$

式中: $E(\cdot)$ 为期望函数; $\pi$ 为策略; $s_q$ 为第 $q$ 辆电动汽车的状态空间(即系统基础负荷情况); $a_q$ 为第 $q$ 辆

充电站的充电过程,进一步优化负荷曲线。

对于场景1,其目标函数如式(35)所示。

电动汽车的动作空间(即充电情况); $r(s_q, a_q)$ 为第 $q$ 辆电动汽车的奖励函数,基于不同场景体现为不同的目标函数,如式(19)至式(22),是目标函数的相反数; $(s_q, a_q) \sim p_{\pi}$ 为策略 $\pi$ 所形成的状态-动作轨迹; $\alpha$ 为温度项,决定了熵对于奖励的影响程度; $\mathcal{H}(\pi(\cdot | s_q))$ 为状态 $s_q$ 时的策略的熵项,其计算方法如式(40)所示。

$$\mathcal{H}(\pi(a_q | s_q)) = - \int_{a_q} \pi(a_q | s_q) \log(\pi(a_q | s_q)) da_q = E_{a_q \sim p_{\pi}} (-\log(\pi(a_q | s_q))) \quad (40)$$

### 4.2 SAC迭代策略

强化学习过程中的价值函数 $Q(s_q, a_q)$ 如式(41)所示,用于策略价值评估;贝尔曼backup算子如式(42)所示,用于策略更新。

$$Q(s_q, a_q) = r(s_q, a_q) + \gamma E_{s_{q+1} \sim p}(Q(s_{q+1}, a_{q+1})) \quad (41)$$

$$\mathcal{T}^{\pi} Q(s_q, a_q) \triangleq r(s_q, a_q) + \gamma E_{s_{q+1} \sim p}(V(s_{q+1})) \quad (42)$$

式中: $E_{s_{q+1} \sim p}$ 为状态空间为 $s_{q+1}$ 的期望函数; $\mathcal{T}^{\pi}$ 为策略 $\pi$ 下的贝尔曼backup算子; $\gamma$ 为奖励的折扣因子; $V(s_{q+1})$ 为状态 $s_{q+1}$ 的新值函数,其计算方法如式(43)所示。

$$V(s_q) = E_{a_q \sim \pi}(Q(s_q, a_q) - \log \pi(a_q | s_{q+1})) \quad (43)$$

同时,有

$$Q^{k+1} = \mathcal{T}^{\pi} Q^k \quad (44)$$

式中: $Q^k$ 为第 $k$ 次计算时的价值函数。

将式(42)和式(44)不断进行迭代,即可实现式(45)。

$$\lim_{k \rightarrow \infty} Q^k = \hat{Q} \quad (45)$$

式中: $\hat{Q}$ 为软 $Q$ 值(soft  $Q$ -value)。

### 4.3 SAC策略更新

计算过程中的策略更新办法如式(46)所示。

$$\pi_{\text{new}} = \arg \min_{\pi \in \Pi} D_{\text{KL}} \left( \pi(\cdot | s_q) \parallel \frac{\exp\left(\frac{1}{\alpha} Q^{\pi_{\text{old}}}(s_q, \cdot)\right)}{Z^{\pi_{\text{old}}}(s_q)} \right) \quad (46)$$

式中: $D_{\text{KL}}$ 为KL散度(KL divergence); $\Pi$ 为策略集合; $Q^{\pi_{\text{old}}}(\cdot)$ 为旧策略 $\pi_{\text{old}}$ 下的价值函数; $Z^{\pi_{\text{old}}}(s_q)$ 为旧

策略  $\pi_{old}$  下的配分函数,用于归一化分布。

#### 4.4 SAC算法构建

SAC算法需要构建神经网络,包括Q值网络和策略网络。Q值网络通过几层神经网络输出单值,策略网络输出一个高斯分布。该过程中神经网络会进行更新,Q值网络参数具有如式(47)所示更新策略,策略网络参数具有如式(48)所示更新策略。

$$J_Q(\theta) = E_{(s_q, a_q, s_{q+1}) \sim D} \left( \frac{1}{2} \left( Q(s_q, a_q) - (r(s_q, a_q) + \gamma V_{\bar{\theta}}(s_{q+1})) \right)^2 \right) \quad (47)$$

$$J_{\pi}(\phi) = D_{KL} \left( \pi(\cdot | s_q) \parallel \exp \left( \frac{1}{\alpha} Q_{\theta}(s_q, \cdot) - \log Z(s_q) \right) \right) \quad (48)$$

式中: $\theta$ 为Q值网络参数; $\phi$ 为策略网络参数; $V_{\bar{\theta}}$ 和 $Q_{\theta}$ 分别为代入Q值网络参数后的新值函数和价值函数; $Z(s_q)$ 为状态 $s_q$ 时的配分函数。

同时,前文提到的温度参数 $\alpha$ 是辅助最大化熵的重要参数,能够最大化对动作空间的探索。合理的温度参数设置有助于实现对所有可行动作的迭代测试。因此,具有温度参数的更新如式(49)所示。

$$J(\alpha) = E_{a_q \sim \pi_q} (-\alpha \log \pi_q(a_q | \pi_q) - \alpha H_0) \quad (49)$$

式中: $\pi_q$ 为第 $q$ 辆电动汽车的充电策略; $H_0$ 为熵项。

式(47)至式(49)都是通过计算求最优化参数值。整个过程中,通过深度神经网络学习不断更新Q值网络参数、策略网络参数以及温度参数,可使模型不断收敛,求解最优策略。

#### 4.5 电动汽车充换电负荷优化调度整体流程

通过SAC可以为大规模电动汽车用户制定考虑用户特征的电动汽车充电方案。同时,考虑换电站具有更强的可调度性能,采用SAC算法、基于充电负荷的可调节情况对电动汽车换电站负荷进行进一步的优化调度,最终实现系统运行上的安全可靠性和市场运营上的经济性优化。整体的优化调度流程如图2所示。

### 5 算例分析

#### 5.1 算例设计

以典型日场景下A市充换电站聚合参与调度作为算例,收集A市2254个充电桩的12770条充电数据,包括接入时间、结束时间、充电电量、充电结束原因等数据内容。其中,电动汽车的出行规律分布情况见附录C图C1、图C2,原始充电电量分布情况见附录C图C3,容量分布情况见附录C图C4,电动汽车单日平均负荷规模约为12.5MW,样本电动

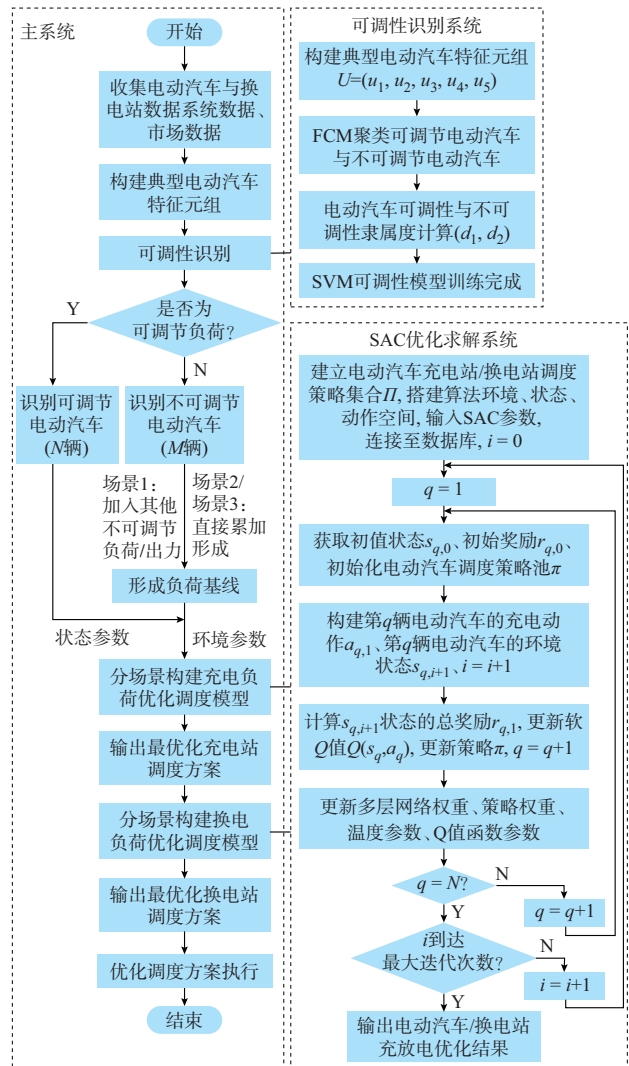


图2 基于SAC的电动汽车充换电站负荷优化调度流程图  
Fig. 2 Flow chart of optimal scheduling for SAC-based electric vehicle load for battery charging and swapping stations

汽车规模为8925辆。以A市典型换电站的换电负荷情况作为样本数据,换电站原始负荷分布情况见附录C图C5,设置换电站共有充电桩40组,每组最大充电功率为75kW,且具有备用电池80组。在目前用户行为偏好标签数据库还不完善的情况下,设定无可调潜力的电动汽车的用户可调节隶属度为0,有可调潜力的电动汽车用户的可调节隶属度为1,车主均希望在离开之前将充电桩电量充满电。

场景1的虚拟电厂负荷曲线见附录C图C6。场景中的主要经济参数如表1所示,其中,峰谷平时段划分的依据为基础负荷曲线的变化情况,分时段电价参照现行市场价格水平。辅助服务(调峰)开市时段为00:30—07:00以及12:30—16:00,辅助服务市场价格为0.2元/(kW·h)。

表1 电力市场参数  
Table 1 Electricity market parameters

时段	时段划分	分时段电价/(元·(kW·h) <sup>-1</sup> )
峰时段	早:09:00—12:00	1.082 4
	晚:17:00—22:00	
谷时段	早:08:00—09:00	0.416 4
	下午:12:00—17:00	
	晚:22:00—23:00	
平时段	晚:23:00—次日08:00	0.664 4

计算过程中SVM采用径向基(radial basis function, RBF)核函数,设置惩罚参数为1,核函数宽度参数为1/6。SAC模型以32条数据为一组训练批次,神经网络构建过程中设置3层隐藏层,64个神经元,学习率均为 $10^{-4}$ 。其他优化调度过程中的实施要点见附录C表C1。

## 5.2 计算结果

### 5.2.1 可调节与不可调节识别结果

利用FCM-SVM可对样本数据进行可调性识别,为验证模型效果,将FCM-SVM、*k*-means-SVM以及FCM-最邻近结点算法(KNN)模型的效果进行比较,对比结果如表2所示。其中,准确率为模型分类准确的样本占总样本个数,精准率为模型分类结果为可调的样本中真实可调的样本比例,召回率为真实可调的样本中归类正确的样本所占的比例。

表2 模型识别效果对比  
Table 2 Comparison of model identification effect

模型	计算时间/s	准确率/%	精准率/%	召回率/%
FCM-SVM	10.552 86	99.00	98.05	99.34
<i>k</i> -means-SVM	10.301 56	96.50	98.00	96.71
FCM-KNN	9.012 56	95.00	94.16	95.39

由表2可见,在本文场景中,FCM-SVM在大数据识别计算时具有明显的分类精度提升,这将有利于电动汽车优化调度方案的落实,所牺牲的计算速度也在可接受范围内,是适用于本文场景的分类策略。

图3为基于FCM-SVM的电动汽车可调度情况识别结果,展示了96个时段的可调节和不可调节电动汽车的负荷规模。

由图3可见,可调节负荷规模与不可调节负荷规模相当,可调节负荷主要分布于时段0至28、54至70、92至96,即每日的13:30—17:30和23:00—次日07:00。同时,从总体的负荷规模来看,样本所选充电桩站点的主要充电高峰有3个,即05:00—

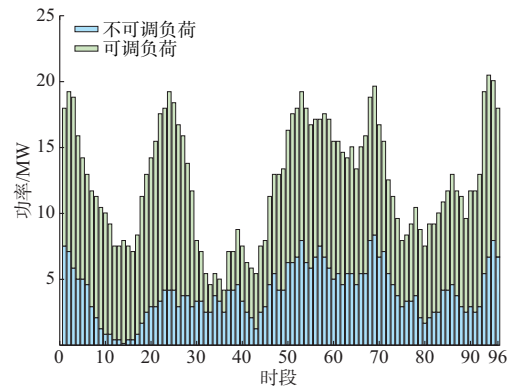


图3 电动汽车可调节与不可调节负荷规模  
Fig. 3 Adjustable and nonadjustable load scale of electric vehicles

07:00、12:30—18:00,20:00—次日01:00。

同时,随机抽样500个典型样本,样本中不同可调性用户在充电时长、接入时长、充电功率、荷电状态和调节时长的分布差异分别如附录D图D1至图D5所示。相对于不可调节负荷,可调节负荷具有充电时长和接入时长明显高、充电平均功率稍高、SOC偏高、可调节时长更长的特征。具体来说,可调节电动汽车负荷充电时长偏向于1 h上下,接入时长主要分布于2~5 h,充电功率大部分采用快充方式,SOC状态主要分布于60%~80%,可调节时长主要位于1~6 h;不可调节电动汽车充电时长大部分位于1 h以内,接入时长通常较短,荷电状态主要分布于70%~100%,可调节时长主要位于1~3 h。

### 5.2.2 电动汽车充电站调度结果

基于SAC深度强化学习算法求解可调节电动汽车负荷的优化调度方案。经过16 h、训练1 500次后,模型训练具有如附录E图E1所示收敛效果。为验证模型的实际调度效果,SAC强化学习、AC(actor-critic)强化学习以及PG(policy-gradient)强化学习算法的计算效果对比如表3所示,相关指标的说明见附录E表E1。

由表3可见,SAC算法对于各类型的场景都具有更低的波动性和相对优秀的收敛性,尽管每回合运行时间相对较长,但是其测试回合的运行时间也不超过1 s。从附录E图E1还可以发现,SAC具有更好的寻优效果,是适用于本文场景的优化求解算法。优化调度结果如图4所示。

由图4可见,3种场景下的优化调度方案都实现了不同程度的峰谷差控制和负荷平抑效果,结合市场环境还创造了一定的收益。具体来说,场景1的

表3 不同模型优化调度效果对比  
Table 3 Comparison of optimal scheduling effects of different models

场景	算法	波动性/%	训练时间/(轮·s <sup>-1</sup> )	测试时间/(轮·s <sup>-1</sup> )	至收敛最低训练回合
场景1	SAC	1.39	40	0.35	953
	AC	4.98	26	0.32	673
	PG	6.34	20	0.45	934
场景2	SAC	0.68	56	0.49	757
	AC	5.97	30	0.36	649
	PG	5.46	26	0.55	853
场景3	SAC	0.66	46	0.44	649
	AC	6.54	27	0.32	707
	PG	5.24	23	0.52	973

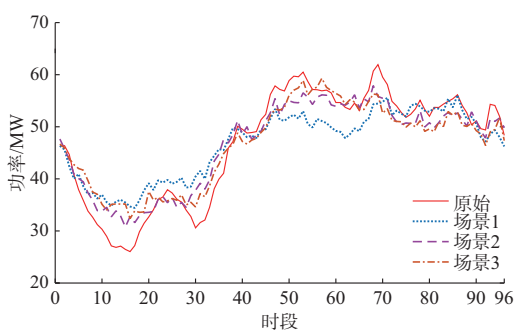


图4 基于SAC的优化调度结果  
Fig. 4 SAC-based optimal scheduling results

削峰填谷效果最为明显,降低了63%的负荷波动效果,EVA收益增加了1 966.93元/d,负荷峰谷差降低了510 kW,负荷率提升了7.84%。场景2下的电动汽车优化调度方案降低了41%的负荷波动效果,EVA收益增加了2 701.06元/d,负荷峰谷差降低了315 kW,负荷率提升了5.23%。场景3下的电动汽车优化调度方案受辅助服务市场的影响,具有最优的经济效益,降低了42%的负荷波动效果,EVA收益增加了3 223.51元/d,负荷峰谷差降低了321 kW,负荷率提升了3.31%。场景1的负荷实时优化调度是迎合电力系统调度需求的结果,场景2与场景3的负荷实时优化调度是迎合分时段电力市场价值变化的结果。整体协调调度的效果如表4所示。

表4 整体协调调度效果  
Table 4 Effect of overall coordinated scheduling

场景	负荷波动率/%	EVA收益/(元·d <sup>-1</sup> )	最大峰谷差/kW	负荷率/%
初始场景	5.16	0	1 287.590	75.45
场景1	1.87	1 966.93	777.690	83.29
场景2	3.05	2 701.06	972.950	80.68
场景3	2.83	3 223.51	966.703	78.76

附录E图E2展示了不同场景下电动汽车优化调度后的可调节电动汽车负荷变化情况,附录E图E3展示了电动汽车总负荷变化情况。综合图E2、图E3可见,场景1中电动汽车负荷跟踪了虚拟电厂聚合负荷的变化情况,使得系统的总体负荷曲线波动性变小,但是电动汽车负荷曲线的波动性增强。场景2和场景3中电动汽车的负荷变化充分响应了市场分时段电价的激励,使得EVA获得了最高市场效益,同时也使得电动汽车负荷曲线波动变大。

### 5.2.3 换电站进一步优化调度效果

在电动汽车优化调度完成之后,通过换电站的进一步优化调度可实现总负荷的进一步调整,优化调度完成后的系统总负荷情况如附录F图F1所示,换电站负荷经调节后的负荷变化情况如图5所示。

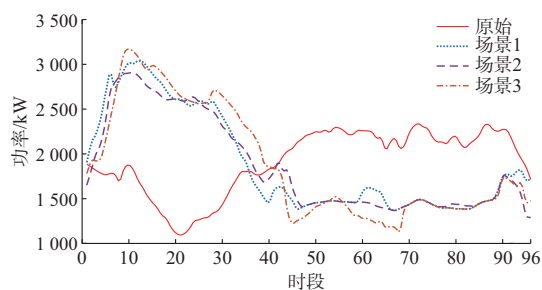


图5 换电站优化调度结果  
Fig. 5 Optimal scheduling results of battery swapping station

通过换电站的实时优化调度,进一步平滑了负荷曲线,场景1中,通过换电站进一步优化了15%的负荷波动,降低了509 kW的峰谷差与1.2%的负荷率;场景2中,通过换电站进一步优化了1.1%的负荷波动,降低了315 kW的峰谷差与1.1%的负荷率;场景3中,通过换电站进一步优化了0.9%的负荷波动,降低了320 kW的峰谷差与0.7%的负荷率。可见,换电站可对电动汽车负荷的优化调节结



果做出进一步的优化。

## 6 结语

为解决EVA促进电力系统经济可靠运行过程中的聚合资源实时优化调度问题,本文提出了基于SAC深度强化学习的大规模电动汽车充换电负荷实时优化调度策略,以实现最优化聚合运行目标,结合仿真算例可以得出以下结论:

1)本文以EVA为研究对象,从系统运行最优和经济效益最优的角度分别构建优化调度模型,可以实现平抑系统负荷波动、减少EVA的充电成本并降低峰谷差的效果。

2)FCM-SVM模型可实现电动汽车可调性识别,从中提取出可调节电动汽车负荷进行实时灵活调度;基于SAC模型可为电动汽车充电过程和换电站电池的充放电过程制定个性化方案,实现充换电站的智能协调调度。

3)通过分析SAC算法的收敛曲线并对比不同算法之间的结果稳定性、收敛速率、运行时间等内容,可以发现本文模型优化效果好、稳定性强,且运算速率较好,具有实际工程应用价值,适合解决大规模电动汽车接入的实时调度问题。

本文的研究成果可以应用于EVA对大规模电动汽车资源的实时优化调度,但其互动机制仅限于EVA与电动汽车资源之间,EVA之间的竞争关系、EVA内部的充电站和换电站之间的协调和平衡关系是后续研究的方向。

本文研究受到国网电动汽车服务有限公司的帮助,特此感谢!

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>),扫英文摘要后二维码可以阅读网络全文。

## 参考文献

- [1] 国家发展改革委.关于扩大战略性新兴产业投资 培育壮大新增长点增长极的指导意见[EB/OL].[2020-10-05].[http://www.gov.cn/zhengce/2020-09/24/content\\_5546618.htm](http://www.gov.cn/zhengce/2020-09/24/content_5546618.htm). National Development and Reform Commission. Guidelines on expanding investment in strategic emerging industries to foster and strengthen new growth points and poles[EB/OL]. [2020-10-05]. [http://www.gov.cn/zhengce/2020-09/24/content\\_5546618.htm](http://www.gov.cn/zhengce/2020-09/24/content_5546618.htm).
- [2] 黄森,陈鑫,文旭在,等.计及市场风险的含电动汽车风光互补微电网随机规划调度模型[J].电网技术,2020,44(8):2883-2890. HUANG Miao, CHEN Xin, WEN Xuzai, et al. Stochastic planning and scheduling model of electric vehicles with wind-solar complementary microgrid considering market risks [J]. Power System Technology, 2020, 44(8): 2883-2890.
- [3] ABBASI M H, TAKI M, RAJABI A, et al. Coordinated operation of electric vehicle charging and wind power generation as a virtual power plant: a multi-stage risk constrained approach [J]. Applied Energy, 2019, 239: 1294-1307.
- [4] MAHMUD K, HOSSAIN M J, RAVISHANKAR J. Peak-load management in commercial systems with electric vehicles [J]. IEEE Systems Journal, 2019, 13(2): 1872-1882.
- [5] 张新城,刘志珍,侯延进,等.计及冬季预热需求的居民区电动汽车负荷调度策略[J].电力自动化设备,2020,40(11):115-121. ZHANG Xincheng, LIU Zhizhen, HOU Yanjin, et al. Electric vehicle load scheduling strategy considering winter preheating demand in residential area [J]. Electric Power Automation Equipment, 2020, 40(11): 115-121.
- [6] SOVACOO B K, KESTER J, NOEL L, et al. Actors, business models, and innovation activity systems for vehicle-to-grid (V2G) technology: a comprehensive review[J]. Renewable and Sustainable Energy Reviews, 2020, 131: 109963.
- [7] REHMAN U. A decentralized dynamic marketing-based demand response using electric vehicles in smart grid[J]. Arabian Journal for Science and Engineering, 2020, 45(8): 6475-6488.
- [8] HUANG X, ZHANG Y, LI D, et al. An optimal scheduling algorithm for hybrid EV charging scenario using consortium blockchains [J]. Future Generation Computer Systems, 2019, 91: 555-562.
- [9] 王毅,陈进,麻秀,等.采用分群优化的电动汽车与电网互动调度策略[J].电力自动化设备,2020,40(5):77-84. WANG Yi, CHEN Jin, MA Xiu, et al. An interactive dispatching strategy for electric vehicles and power grid based on clustering optimization [J]. Electric Power Automation Equipment, 2020, 40(5): 77-84.
- [10] 刘东奇,王耀南,袁小芳.电动汽车充放电与风力/火力发电系统的协同优化运行[J].电工技术学报,2017,32(3):18-26. LIU Dongqi, WANG Yaonan, YUAN Xiaofang. Coordinated optimization operation of electric vehicle charge-discharge and wind/thermal power system [J]. Transactions of China Electrotechnical Society, 2017, 32(3): 18-26.
- [11] 张高,王旭,蒋传文.基于主从博弈的含电动汽车虚拟电厂协调调度[J].电力系统自动化,2018,42(11):48-55. ZHANG Gao, WANG Xu, JIANG Chuanwen. Coordinated dispatch of virtual power plants with electric vehicles based on master-slave game [J]. Automation of Electric Power Systems, 2018, 42(11): 48-55.
- [12] LIU J, ZHANG T, ZHU J, et al. Allocation optimization of electric vehicle charging station (EVCS) considering with charging satisfaction and distributed renewables integration [J]. Energy, 2019, 164: 560-574.
- [13] 实地探访新能源汽车换电站:换电模式风口来临政策持续“力挺” [EB/OL]. [2020-07-28]. <https://xw.qq.com/cmsid/20200728A0127100>. On-site visit to new energy vehicle electrical changing station: the arrival of electrical changing mode tuyere policy continued

- to “strongly support”[EB/OL]. [2020-07-28]. <https://xw.qq.com/cmsid/20200728A0127100>.
- [14] 田文奇,和敬涵,姜久春,等.电动汽车换电站有序充电调度策略研究[J].电力系统保护与控制,2012,40(21):114-119.  
TIAN Wenqi, HE Jinghan, JIANG Jiuchun, et al. Research on orderly charging scheduling strategy for electric vehicle electrical changing station [J]. Power System Protection and Control, 2012, 40(21): 114-119.
- [15] 刘佳佳.电动汽车换电调度及换电站充电策略研究[D].武汉:华中科技大学,2019.  
LIU Jiajia. Research on electric vehicle changing scheduling and charging strategy of electric changing station [D]. Wuhan: Huazhong University of Science and Technology, 2019.
- [16] 陈吕鹏,潘振宁,余涛,等.基于动态非合作博弈的大规模电动汽车实时优化调度[J].电力系统自动化,2019,43(24):32-40.  
CHEN Lüpeng, PAN Zhenning, YU Tao, et al. Real-time optimal scheduling of large-scale electric vehicles based on dynamic non-cooperative game [J]. Automation of Electric Power Systems, 2019, 43(24): 32-40.
- [17] 许刚,张丙旭,张广超.电动汽车集群并网的分布式鲁棒优化调度模型[J].电工技术学报,2021,36(3):565-578.  
XU Gang, ZHANG Bingxu, ZHANG Guangchao. Distributed and robust optimal scheduling model for large-scale electric vehicles connected to grid [J]. Transactions of China Electrotechnical Society, 2021, 36(3): 565-578.
- [18] 王惠中,黄金凤,杨家颜.基于MOPSO的电动汽车有序充电调度控制策略[J].兰州理工大学学报,2019,45(5):85-90.  
WANG Huizhong, HUANG Jinfeng, YANG Jiayan. Control policy of ordered charging scheduling of electric automobiles based on MOPSO [J]. Journal of Lanzhou University of Technology, 2019, 45(5): 85-90.
- [19] 李咸善,陈敏睿,程杉,等.基于双重激励协同博弈的含电动汽车微电网优化调度策略[J].高电压技术,2020,46(7):2286-2296.  
LI Xianshan, CHEN Minrui, CHENG Shan, et al. Optimal scheduling strategy of micro-grid with electric vehicle based on dual excitation cooperative game [J]. High Voltage Engineering, 2020, 46(7): 2286-2296.
- [20] 金国彬,潘狄,陈庆,等.考虑源荷不确定性的直流配电网模糊随机日前优化调度[J/OL].电工技术学报:1-12[2021-05-20]. <https://doi.org/10.19595/j.cnki.1000-6753.tces.201295>.  
JIN Guobin, PAN Di, CHEN Qing, et al. Fuzzy stochastic day-ahead optimal scheduling of DC distribution network considering source-charge uncertainty [J/OL]. Transactions of China Electrotechnical Society: 1-12 [2021-05-20]. <https://doi.org/10.19595/j.cnki.1000-6753.tces.201295>.
- [21] 周华嫣然,周羿宏,胡俊杰,等.人工智能技术支撑的集群电动汽车实时优化调度策略[J].电网技术,2021,45(4):1446-1459.  
ZHOU Huayanran, ZHOU Yihong, HU Junjie, et al. Real-time optimal scheduling strategy for clustered electric vehicles supported by artificial intelligence technology [J]. Power System Technology, 2021, 45(4): 1446-1459.
- [22] 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. Electric vehicle real-time scheduling strategy based on deep reinforcement learning [J]. Automation of Electric Power Systems, 2020, 44 (22) : 161-167.
- [23] 李向杰,张向文.基于人工蜂群和支持向量机的电动汽车制动意图识别方法[J/OL].中国机械工程:1-12[2021-06-18]. <http://kns.cnki.net/kcms/detail/42.1294.TH.20210329.1711.010.html>.  
LI Xiangjie, ZHANG Xiangwen. Braking intention identification method for electric vehicles based on artificial bee colony and support vector machine [J/OL]. The Chinese Mechanical Engineering: 1-12 [2021-06-18]. <http://kns.cnki.net/kcms/detail/42.1294.TH.20210329.1711.010.html>.
- [24] 姚春晓.广域可调节负荷的辨识方法研究[D].北京:华北电力大学,2019.  
YAO Chunxiao. Research on identification method of wide-area adjustable load [D]. Beijing: North China Electric Power University, 2019.
- [25] ZHAO T, LI Y, PAN X, et al. Real-time optimal energy and reserve management of electric vehicle fast charging station: hierarchical game approach [J]. IEEE Transactions on Smart Grid, 2018, 9(5): 5357-5370.
- [26] PEVEC D, BABIC J, KAYSER M A, et al. A data-driven statistical approach for extending electric vehicle charging infrastructure [J]. International Journal of Energy Research, 2018, 42(9): 3102-3120.
- [27] 黄煜梵,彭诺衡,林艳,等.基于SAC强化学习的多智能体车联网频谱资源动态分配[J/OL].计算机工程:1-12[2020-12-27]. <http://kns.cnki.net/kcms/detail/31.1289.TP.20201210.1648.003.html>.  
HUANG Yufan, PENG Nuoheng, LIN Yan, et al. Dynamic spectrum resource allocation of multi-agent network of vehicles based on SAC reinforcement learning [J/OL]. Computer Engineering: 1-12 [2020-12-27]. <http://kns.cnki.net/kcms/detail/31.1289.TP.20201210.1648.003.html>.
- [28] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[C]// 35th International Conference on Machine Learning, July 10-15, 2018, Stockholm, Sweden: 2976-2989.

刘敦楠(1979—),男,博士,教授,博士生导师,主要研究方向:电力系统自动化、电力市场。E-mail: 50601660@ncepu.edu.cn

王玲湘(1997—),女,通信作者,硕士研究生,主要研究方向:电动汽车聚合调度、电力市场。E-mail: wanglingxiang618@163.com

汪伟业(1997—),男,硕士研究生,主要研究方向:电力市场。E-mail: examplemy163email@163.com

(编辑 王梦岩)

## Optimal Scheduling of Electric Vehicle Load for Large-scale Battery Charging and Swapping Based on Deep Reinforcement Learning

LIU Dunman<sup>1</sup>, WANG Lingxiang<sup>1</sup>, WANG Weiye<sup>1</sup>, LI Hua<sup>1</sup>, WANG Wen<sup>2</sup>, LIU Mingguang<sup>1</sup>

(1. School of Economic and Management, North China Electric Power University, Beijing 102206, China;

2. State Grid Electric Vehicle Service Co., Ltd., Beijing 100053, China)

**Abstract:** For the optimal scheduling problem of the aggregation in large-scale battery charging and swapping stations, a real-time optimal scheduling strategy is proposed for battery charging and swapping loads based on soft actor-critic (SAC) deep reinforcement learning. The strategy fully considers the user, system and market factors in the load control process, and can realize the friendly interaction between large-scale electric vehicles and various types of power system subjects. Firstly, considering the development scale and scheduling performance of battery charging and swapping stations, a joint operation framework is established. Secondly, an adjustable identification model considering multi-user characteristics is proposed to judge the actual adjustable performance of electric vehicles. Furthermore, considering the multiple spatio-temporal characteristics of the optimal scheduling of battery charging and swapping stations, the optimal scheduling models of adjustable battery charging and swapping loads in different scenarios are constructed. Then, the real-time scheduling scheme of battery charging and swapping loads is solved based on SAC algorithm. Finally, the case of virtual power plant load for electric vehicle load aggregation optimization verifies the economy and high efficiency of the SAC algorithm applied in real-time optimal scheduling of electric vehicle loads for large-scale battery charging and swapping.

This work is supported by National Social Science Foundation of China (No. 19ZDA081) and State Grid Corporation of China (No. 5418-202018247A-0-0-00).

**Key words:** deep reinforcement learning; electric vehicle; battery charging and swapping coordination; real-time optimal scheduling; electric vehicle aggregator

