

基于状态-响应框架的有源配电网运行优化策略

胡柳君¹, 董红¹, 曾繁宏¹, 张军¹, 张勇军^{1,2}, 高毓群¹

(1. 广东电网有限责任公司广州供电局, 广州 510620; 2. 华南理工大学电力学院, 广州 510640)

摘要: 为提高配电网的运行效率和电压质量, 结合深度强化学习和分布式广义快速对偶上升(SAC-GFD)算法, 提出了一种基于状态-响应框架的优化策略。首先, 利用软演员-评论家(soft actor-critic, SAC)算法将配电网运行优化问题建模为马尔可夫决策过程, 智能体在含有可再生能源波动和负荷不确定性的环境中进行交互与探索, 获得对不确定性环境具有鲁棒性控制策略。将配电网运行优化问题转化为马尔可夫决策过程, 从而训练出能够快速输出配电网设备最优有功功率和无功功率的智能体。其次, 计算当前配电网的潮流分布、节点电压状态以及有功功率和无功-电压灵敏度矩阵。然后, 用户基于当前配电网状态, 采用分布式方法计算自身负荷的最优调整值, 确保配电网的安全运行。最后, 在 IEEE 33 节点系统上的仿真结果表明, 相较于传统的深度强化学习算法, 所提算法能更有效地降低网络损耗和节点电压偏差, 且具有更快的训练速度和更好的优化结果。

关键词: 状态-响应; 深度强化学习; 软演员-评论家; 广义快速对偶上升法; 有源配电网

Active Distribution Network Operation Optimization Strategies Based on State-Response Framework

HU Liujun¹, DONG Hong¹, ZENG Fanhong¹, ZHANG Jun¹, ZHANG Yongjun^{1,2}, GAO Yuqun¹

(1. Guangzhou Power Supply Bureau of Guangdong Power Grid Co., Ltd., Guangzhou 510620, China; 2. School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: In order to improve the operational efficiency and voltage quality of distribution networks, an optimization strategy is proposed based on a state-response framework, combining deep reinforcement learning with the distributed generalized fast dual ascent (SAC-GFD) algorithm. Firstly, the soft actor-critic (SAC) algorithm is employed to model the distribution network operation optimization problem as a Markov decision process (MDP). The agent interacts and explores in an environment with renewable energy fluctuations and load uncertainties, obtaining a control strategy that is robust to uncertain environments. The distribution network operation optimization problem is thus transformed into a Markov decision process, enabling the training of an agent capable of quickly outputting the optimal active and reactive power for distribution network equipment. Secondly, the current distribution network's power flow distribution, node voltage states, and active and reactive voltage sensitivity matrices are calculated. Then, based on the current state of the distribution network, users apply a distributed method to calculate their optimal load adjustments to ensure the safe operation of the network. Finally, the simulation results on the IEEE 33-node system demonstrate that compared to traditional deep reinforcement learning algorithms, the proposed method more effectively reduces network losses and node voltage deviations while achieving faster training speeds and better optimization results.

Key words: state-response; deep reinforcement learning; soft actor-critic; generalized fast dual ascent; active distribution network

基金项目: 国家自然科学基金面上项目(52177085); 广东电网有限责任公司广州供电局科技项目(GZHKJXM20210043(080041KK52210002))。

Foundation item: Supported by the General Program of National Natural Science Foundation of China (52177085); the Science and Technology Project of Guangzhou Power Supply Bureau of Guangdong Power Grid Co., Ltd. (GZHKJXM20210043(080041KK52210002))。

0 引言

随着分布式可再生能源的大规模接入,配电网在确保安全运行的基础上,还面临着日益严峻的网损增加、电压超限及可再生能源消纳困难等挑战^[1-2]。有源配电网中的各类有功功率、无功功率调节设备和用户端的灵活性资源^[3-4],为解决这些问题提供了丰富的手段,但亟需建立合理的运行优化策略,从而协调各类资源之间的控制,保障配电网经济高效与安全运行^[5]。

目前,配电网的运行控制策略主要分为两种:机理模型策略和数据驱动控制策略。机理模型策略主要从配电网的物理层面出发^[6],基于线路参数、网架结构建立数学模型^[7-8],并且利用启发式^[9]、二阶锥规划^[10]、鲁棒优化^[11]等方法进行求解,通常须依赖精确的数学参数。然而在实际工程中这些数据往往难以准确获取,直接制约了优化结果的实用价值^[12-13]。此外,基于机理模型的策略控制方式主要分为集中式和分布式两种^[14]。集中式的算法能求得全局最优解,但随着设备规模的增加,其通信压力和计算时间会呈指数级增长,且单点故障可能会导致系统的瘫痪运行^[15]。而分布式算法可拓展性好,通信负担小,鲁棒性强^[16],但各节点之间的协调挑战性更大。

为此,部分学者采用了包括深度强化学习算法在内的数据驱动策略^[17]优化配电网系统的运行过程。深度强化学习(deep reinforcement learning, DRL)无需依赖精准的数学模型^[13],而是通过智能体与环境进行交互,逐步适应环境参数的不确定性。文献[18]提出了一种图卷积网络的近端策略优化算法,优化配电网运行过程中的电压状态。文献[19]基于一种的双级强化学习方法对配电网中的慢速和快速设备进行双时间尺度调节。文献[20]利用多智能体强化学习在弱通信条件下实现区域自治-全局协同的电压控制。

然而,无论是多智能体还是单智能体的DRL架构,虽然可以较好地适应有源配电网的特性并优化其运行状态,但是其普遍存在无法有效处理电压约束的问题。现有的DRL策略对于电压约束大多依据电压超限的程度大小,转化为奖励值中的惩罚项,较为依赖于惩罚系数的整定^[21]。当惩罚系数过大时,容易导致智能体控制策略过于保守,使得网

损较大,影响优化效果;当惩罚系数过小时,则可能导致控制策略过于激进,使得配电网在运行过程中出现较多的电压超限状态^[22]。

针对上述问题,本文提出了一种基于状态-响应框架的双层配电网运行优化策略。在状态层,利用软演员-评论家(soft actor-critic, SAC)算法优化配电网中有功和无功设备的动作值,智能体在不确定性环境中进行交互与学习,获得对环境具有鲁棒性的控制策略。与传统方法不同,本文不再依赖惩罚系数来处理电压约束,而是在潮流计算后,将当前电压状态传递到响应层。在响应层,用户根据各节点的电压状态,采用广义快速对偶上升法^[23]调整自身负荷出力,从机理上直接对节点电压进行响应优化,有效避免了电压超限问题的发生,最终保证有源配电网的安全运行。

1 有源配电网的状态-响应框架

本文假设该有源配电网呈辐射状网络拓扑结构,并接入有分布式光伏发电设备(photovoltaic, PV)、储能系统(energy storage system, ESS)、用户负荷设备。配电网需要尽可能地消纳分布式光伏的发电功率,并且在满足用户用电需求的同时保证系统运行在安全稳定的状态。

在运行过程中,设备不同的有功、无功功率出力情况,会导致配电网运行状态的差异变化。不同状态下的配电网,对用户负荷参与需求响应的要求程度也不尽相同。本文将同时考虑配电网状态与用户响应控制,提出一种面向有源配电网的状态-响应框架的双层优化策略,如图1所示。

在状态层中,配电网控制中心拥有对配电网设备的控制权,在考虑节点电压、潮流平衡等约束条件下,以最小化成本为运行目标,并向参与需求响应的用户传递当前配电网状态。该过程通过数据驱动的方式实现,智能体根据目标函数利用SAC算法进行离线训练与在线部署。智能体能够根据配电网实时状态,以毫秒级速度生成光伏、储能的控制策略,并更新设备动作后的配电网状态。

在响应层中,用户在接收到来自配电网的状态信息后调节自身设备运行工况,参与需求响应,并计算对电网状态产生的影响。为了保护用户的隐私信息,用户无需提交自身具体的设备参数情况,而是采用分布式的方法在保证系统安全的前提下最大

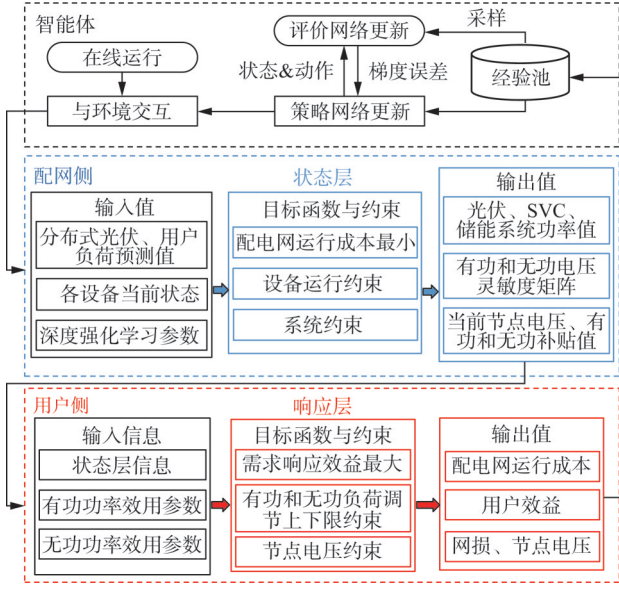


图 1 状态-响应框架图

Fig. 1 State-response framework diagram

化用电效益。

2 配电网双层优化模型

2.1 状态层优化

配电网状态层的优化环节考虑多个设备之间协同调度，以最小化运行成本 Φ_t 为目标函数，如式(1)所示。

$$\begin{aligned} \min \sum_{i \in T} \Phi_t \\ = C_{\text{GRID}} \sum_{i \in T} \sum_{l \in \mathcal{L}} P_{i,l}^{\text{LOSS}} + \\ C_{\text{ESS}} \sum_{i \in T} \sum_{m \in \mathcal{M}_{\text{ESS}}} \left[(P_{i,t,m}^{\text{PD}})^2 + (P_{i,t,m}^{\text{PC}})^2 \right] + \\ C_{\text{PV}} \sum_{i \in T} \sum_{m \in \mathcal{M}_{\text{PV}}} (P_{i,t,m}^{\text{PV}} - P_{i,t,m}^{\text{PRE}})^2 \end{aligned} \quad (1)$$

式中： C_{GRID} 为电网电价； C_{ESS} 为储能系统的损耗成本系数； C_{PV} 为分布式资源的弃风弃光惩罚成本系数； $P_{i,l}^{\text{LOSS}}$ 为系统网损功率； $P_{i,t,m}^{\text{PC}}$ 、 $P_{i,t,m}^{\text{PD}}$ 分别为位于节点 i 的第 m 台储能系统 t 时段的充电和放电有功功率； $P_{i,t,m}^{\text{PV}}$ 、 $P_{i,t,m}^{\text{PRE}}$ 分别为分布式光伏的实际有功功率及其预测出力功率； $\mathcal{L}$ 为配电网线路集合， $l \in \mathcal{L}$ ； $\mathcal{M}_{\text{ESS}}$ 、 $\mathcal{M}_{\text{PV}}$ 分别为储能系统和分布式光伏的设备集合； T 为一天 24 时段。

在状态层的优化过程中，需要满足的约束条件如下，其中式(2)–(3)为潮流计算等式约束；式(4)为节点电压约束；式(5)–(9)为 PV、静止无功补偿器(static var compensator, SVC)、ESS 的运行

和容量上下限约束。

$$\begin{aligned} V_{e,i,t} \sum_{j \in \mathcal{N}} (G_{ij} V_{e,j,t} - B_{ij} V_{f,j,t}) + \\ V_{f,i,t} \sum_{j \in \mathcal{N}} (G_{ij} V_{f,j,t} - B_{ij} V_{e,j,t}) + \end{aligned} \quad (2)$$

$$\begin{aligned} P_{i,t,m}^{\text{PD}} - P_{i,t,m}^{\text{PC}} + P_{i,t,m}^{\text{PV}} - P_{i,t,m}^{\text{LOAD}} = 0 \\ \sum_{j \in \mathcal{N}} (G_{ij} V_{f,j,t} + B_{ij} V_{e,j,t}) + \end{aligned} \quad (3)$$

$$\begin{aligned} Q_{i,t,m}^{\text{PV}} + Q_{i,t,m}^{\text{SVC}} - Q_{i,t}^{\text{LOAD}} = 0 \\ V_{\min} \leq V_{i,t} \leq V_{\max} \end{aligned} \quad (4)$$

$$-\sqrt{(S_{i,t,m}^{\text{PV}})^2 - (P_{i,t,m}^{\text{PV}})^2} \leq Q_{i,t,m}^{\text{PV}} \quad (5)$$

$$\leq \sqrt{(S_{i,t,m}^{\text{PV}})^2 - (P_{i,t,m}^{\text{PV}})^2} \quad (6)$$

$$Q_{\min}^{\text{PV}} \leq Q_{i,t,m}^{\text{PV}} \leq Q_{\max}^{\text{PV}} \quad (7)$$

$$Q_{\min}^{\text{SVC}} \leq Q_{i,t,m}^{\text{SVC}} \leq Q_{\max}^{\text{SVC}} \quad (8)$$

$$S_{i,t}^{\text{ESS}} = S_{i,t-1}^{\text{ESS}} + P_{i,t,m}^{\text{PC}} \eta - P_{i,t,m}^{\text{PD}} / \eta \quad (9)$$

$$\begin{aligned} [P_{i,t,m}^{\text{PC}}, P_{i,t,m}^{\text{PD}}, S_{i,t,m}^{\text{ESS}}] \geq [0, 0, S_{\min}^{\text{ESS}}] \\ [P_{i,t,m}^{\text{PC}}, P_{i,t,m}^{\text{PD}}, S_{i,t,m}^{\text{ESS}}] \leq [P_{\max}^{\text{PC}}, P_{\max}^{\text{PD}}, S_{\max}^{\text{ESS}}] \end{aligned} \quad (9)$$

式中： G_{ij} 、 B_{ij} 分别为导纳矩阵的实部和虚部； $V_{e,j,t}$ 和 $V_{f,j,t}$ 分别为 t 时刻内 j 节点电压的实部和虚部； $\mathcal{N}$ 为配电网节点集合， $i, j \in \mathcal{N}$ ； $S_{i,t,m}^{\text{PV}}$ 、 $Q_{i,t,m}^{\text{PV}}$ 分别为节点 i 的第 m 台分布式光伏的视在功率和无功功率； $V_{i,t}$ 为节点 i 的电压值； $V_{\max}$ 、 $V_{\min}$ 分别为节点 i 电压的上下限； $P_{i,t}^{\text{LOAD}}$ 、 $Q_{i,t}^{\text{LOAD}}$ 分别为节点 i 的用户在 t 时刻的负荷有功和无功功率； $S_{i,t,m}^{\text{ESS}}$ 为节点 i 的第 m 台 ESS 的视在功率； $S_{i,t}^{\text{ESS}}$ 为节点 i 的 ESS 视在功率值； η 为容量系数； $Q_{i,t,m}^{\text{SVC}}$ 为位于节点 i 的第 m 台 SVC 的无功功率； $Q_{\max}^{\text{PV}}$ 、 $Q_{\min}^{\text{PV}}$ 、 $Q_{\max}^{\text{SVC}}$ 、 $Q_{\min}^{\text{SVC}}$ 分别为 PV 和 SVC 的容量上下限； $P_{\max}^{\text{PC}}$ 、 $P_{\max}^{\text{PD}}$ 、 $S_{\max}^{\text{ESS}}$ 、 $S_{\min}^{\text{ESS}}$ 分别为储能系统的充放电功率上限和容量上下限。

2.2 响应层优化

用户主要通过调整有功和无功功率来参与需求响应过程，因此，建模时重点考虑用户的调节成本和由此产生的经济效益。因此，本文使用开口向下的拟二次函数对用户参与响应的行为特征进行建模^[16]，其二阶导为负，最优值出现在一阶导数为 0 的点，用于表示用户调整功率时的最优决策值，使得用户的用电效益最大化。第一项是有功功率的调节代价；第二项是有功功率调节带来的经济效益；第三项是无功功率的调节代价。对于每个用户而言，其目标函数如下。

$$\begin{aligned} & \max \sum_{i \in \mathcal{N}_{\text{user}}} \sum_{t \in T} U_{i,t} \\ & = \sum_{i \in \mathcal{N}} \sum_{t \in T} [\alpha_{i,t} (P_{i,t}^{\text{LOAD}})^2 + \beta_{i,t} P_{i,t}^{\text{LOAD}} + \gamma_{i,t} (Q_{i,t}^{\text{LOAD}} - Q_{i,t-1}^{\text{LOAD}})^2] \end{aligned} \quad (10)$$

式中: $U_{i,t}$ 为节点 i 的用户在 t 时刻的用电效益; $\alpha_{i,t}$ 、 $\beta_{i,t}$ 、 $\gamma_{i,t}$ 为用户有功和无功功率的效用参数; $\mathcal{N}_{\text{user}}$ 为用户集合。

在用户参与响应过程中应同时满足如下约束。

$$P_{\min}^{\text{LOAD}} \leq P_{i,t}^{\text{LOAD}} \leq P_{\max}^{\text{LOAD}} \quad (11)$$

$$Q_{\min}^{\text{LOAD}} \leq Q_{i,t}^{\text{LOAD}} \leq Q_{\max}^{\text{LOAD}} \quad (12)$$

$$V_{\min} \leq V_{i,t} + \Delta V_{i,t}^{\text{USER}} \leq V_{\max} \quad (13)$$

式中: $P_{\max}^{\text{LOAD}}$ 、 $P_{\min}^{\text{LOAD}}$ 、 $Q_{\max}^{\text{LOAD}}$ 、 $Q_{\min}^{\text{LOAD}}$ 分别为用户有功和无功功率调整的上下限; $\Delta V_{i,t}^{\text{USER}}$ 为由于用户功率调整而引起节点 i 的电压变化量。

3 基于强化学习-快速对偶上升算法原理

在配电网状态层和用户响应层的优化过程中,当配电网的状态不同时,需要用户参与响应的程度也不尽相同。因此,本文采用一种基于深度强化学习和分布式优化计算结合的算法,对状态-响应的协调优化过程进行求解。

3.1 马尔可夫决策过程

在基于状态-响应框架的配电网运行优化问题中,为了最小化运行成本,该问题被建模成为一个可观测的马尔可夫决策过程,并且用一个五元组来表示: $\langle S, A, O, R, \gamma \rangle$ 。其中, S 为状态空间集合, A 为动作空间集合, O 为状态转移概率, R 为奖励函数, γ 为折扣因子,用于当前和未来的奖励权重。

1) 状态空间 S

选择配电网中的可观测状态值,作为训练过程中智能体所处环节的输入变量,记为 $S = P \times Q \times V \times G$, S 包含了反映环境不确定性的关键变量,如分布式光伏的实时出力 P^{PV} 和负荷的实时需求 P^{LOAD} 。这些变量的动态变化使得智能体需要具备适应性的策略。

其中, $P = \{P^{\text{PV}}, P^{\text{PD}}, P^{\text{PC}}, P^{\text{LOAD}}\}$ 为配电网内部所有光伏系统、储能系统、负荷的有功功率集合; $Q = \{Q^{\text{PV}}, Q^{\text{SVC}}, Q^{\text{LOAD}}\}$ 为配电网内部所有光伏系统、SVC、负荷的无功功率集合; $V = \{V\}$ 为配电网所内部所有节点电压赋值的集合; $G =$

$\{g^{\text{P}}, g^{\text{Q}}\}$ 为配电网中的用户调整单位有功功率和无功功率的补贴价格集合。因此,配电网系统在 t 时刻下的状态空间集合如式(14)所示。

$$s = \left\{ P^{\text{PV}}, P^{\text{PD}}, P^{\text{PC}}, P^{\text{LOAD}}, V, Q^{\text{PV}}, Q^{\text{SVC}}, Q^{\text{LOAD}}, g^{\text{P}}, g^{\text{Q}} \right\}, s \in S \quad (14)$$

如果目标函数发生改变,如进一步考虑其他类型的经济成本、碳排放量、电压偏差程度等,也可以纳入在本研究中,作为相关变量输入到状态空间。

2) 动作空间 A

有源配电网通过控制设备的有功和无功功率,实现最小化运行成本。动作空间如下。

$$a = \{a_i^{\text{PV}}, a_i^{\text{SVC}}, a_i^{\text{ESS}}, a_i^{\text{P}}, a_i^{\text{Q}}\}, a \in A \quad (15)$$

式中 a_i^{PV} 、 a_i^{SVC} 、 a_i^{ESS} 、 a_i^{P} 、 a_i^{Q} 分别为光伏系统的有功功率、SCV的无功功率值、储能系统功率值、有功功率补偿价格、无功功率补偿价格。

3) 状态转移函数 O

配电网在采用不同的动作时,会以不同的概率转移到不同的新状态。因此利用状态转移函数 $O(s'|s, a)$ 描述配电网中智能体在状态 s 出发,采取动作 a 后,转移到下一个状态 s' 的概率分布。

4) 奖励 R

奖励函数 R 定义了智能体在 MDP 过程中每个状态-动作对下所获得的奖励,从而量化每个决策的价值。奖励函数需要被精心设计,以确保智能体的策略在满足所有约束的同时,能够最大化累积奖励。

如上文所述,本文基于有源配电网的运行成本,作为奖励值输入到智能体训练的过程中。

$$R = \sum_{t \in T} [\Phi_t - \sum_{i \in \mathcal{N}_{\text{user}}} U_{i,t}] \quad (16)$$

5) 折扣因子 γ

折扣因子 γ 的取值范围在 0~1 之间的常数,用于对未来奖励进行折扣,从而平衡实时奖励和未来奖励之间的权重。通过调整 γ 的值,可以有效影响配电网智能体的动作策略和学习效率。

3.2 强化学习-快速对偶上升算法

3.2.1 软演员-评论家算法

SAC^[22] 是一种基于策略梯度的强化学习算法,属于演员-评论家(actor-critic)方法的范畴。SAC 算法结合了策略梯度和熵正则化的优势,鼓励智能体探索更多可能的动作选择,用于解决连续动作空间

中的控制问题。这种高熵策略使得智能体在面对环境不确定性时,能够保持策略的多样性,从而提高对未知情况的适应能力,并获得更高的样本效率和更好的策略稳定性。具体表达如下。

$$V^\pi(s) = \mathbb{E}_{a \sim \pi} [Q^\pi(s, a) - \alpha \log \pi] \quad (17)$$

$$Q^\pi(s, a) = R + \gamma \mathbb{E}_{s' \sim P} [V^\pi(s')] \quad (18)$$

式中; $V^\pi(s)$ 为在策略 π 下的 Soft-Value 函数; $Q^\pi(s, a)$ 为 Soft-Q 函数; $\mathbb{E}_{a \sim \pi}$ 为策略 π 下对动作 a 的期望; $\mathbb{E}_{s' \sim P}$ 为在状态转移分布 P 下对下一状态的期望; $V^\pi(s')$ 为策略 π 下状态 s' 的价值函数。 α 为熵权重参数,数值越大,智能体探索能力越好。

SAC 中含有一个策略网络,用于参数化策略 $\pi_\theta(a|s)$,下标 θ 说明策略由神经网络生成,输出给定状态下的动作分布;两个 Q 网络,用于估计给定状态-动作对的动作值函数 $Q_{\phi_1}(s, a)$ 和 $Q_{\phi_2}(s, a)$,从而降低估计偏差;两个目标 Q 网络,用于计算目标值,提高训练的稳定性。其具体的网络更新如下。

1) Q 网络更新

使用经验回放缓冲区中的样本,通过最小化以下损失函数 J_Q 来更新 Q 网络参数 ϕ_1 和 ϕ_2 。

$$J_Q(\phi_i) = \mathbb{E}_{(s, a, r, s') \sim \mathcal{B}} \left[(Q_{\phi_i}(s, a) - y)^2 \right], i = 1, 2 \quad (19)$$

式中 $\mathbb{E}_{(s, a, r, s') \sim \mathcal{B}}$ 为从数据分布中采样的状态-动作-奖励-下一状态四元组上的期望; $\mathcal{B}$ 为经验池; y 为目标值。

其中,目标值 y 计算公式为:

$$y = r + \gamma \mathbb{E}_{a' \sim \pi_\theta} \left[\min_{i=1,2} Q_{\phi'_i}(s', a') - \alpha \log \pi_\theta(a'|s') \right] \quad (20)$$

式中 $\mathbb{E}_{a' \sim \pi_\theta}$ 为在策略 π_θ 下对采用动作 a' 的期望。

2) 策略网络更新

通过最大化以下目标函数更新策略网络参数 θ :

$$y = r + \gamma \mathbb{E}_{a' \sim \pi_\theta} \left[\frac{Q_{\phi'}(s', a') - \alpha \log \pi_\theta(a'|s')}{\alpha \log \pi_\theta(a'|s')} \right] \quad (21)$$

3) 目标 Q 网络更新

使用软更新方法来更新目标 Q 网络的参数 ϕ'_i :

$$\phi'_i \leftarrow \tau \phi_i + (1 - \tau) \phi'_i \quad (22)$$

式中 τ 为一个小于 1 的系数,用于控制目标网络的更新速度。

3.2.2 广义快速对偶上升法

广义快速对偶上升法^[23]是一种基于灵敏度矩阵

进行一阶近似化处理的分布式优化算法。对于节点 i 的用户而言,在调整有功和无功功率的过程中,除了会带来自身收益或成本上的变化,同时也会改变配电网的节点电压。如果使用神经网络对该过程进行拟合,会导致严重的维数灾难问题,因此,为了便于工程上进行计算,本文对配电网电压状态变化过程进行线性化处理,得到:

$$\Delta \mathbf{V} = [\mathbf{O}_P^V \ \mathbf{O}_Q^V] \cdot [\Delta \mathbf{P}^{\text{LOAD}} \ \Delta \mathbf{Q}^{\text{LOAD}}]^T \quad (23)$$

式中: $\Delta \mathbf{V}$ 为节点电压变化量矩阵; $\mathbf{O}_P^V$ 、 $\mathbf{O}_Q^V$ 分别为电压-有功功率、电压-无功灵敏度矩阵; $\Delta \mathbf{P}^{\text{LOAD}}$ 、 $\Delta \mathbf{Q}^{\text{LOAD}}$ 分别为负荷有功和无功功率的变化值。灵敏度矩阵可以在离线潮流计算过程中对节点有功和无功功率的微小修改,并分析电压变化量获得^[21],其具体表达式如下。

$$\mathbf{O}_P^V = \begin{bmatrix} \frac{\partial V_{1,t}}{\partial P_{1,t}^{\text{LOAD}}} & \cdots & \frac{\partial V_{1,t}}{\partial P_{N_{\text{user}},t}^{\text{LOAD}}} \\ \vdots & & \vdots \\ \frac{\partial V_{N_{\text{user}},t}}{\partial P_{1,t}^{\text{LOAD}}} & \cdots & \frac{\partial V_{N_{\text{user}},t}}{\partial P_{N_{\text{user}},t}^{\text{LOAD}}} \end{bmatrix}_{N_{\text{user}} \times N} \quad (24)$$

$$\mathbf{O}_Q^V = \begin{bmatrix} \frac{\partial V_{1,t}}{\partial Q_{1,t}^{\text{LOAD}}} & \cdots & \frac{\partial V_{1,t}}{\partial Q_{N_{\text{user}},t}^{\text{LOAD}}} \\ \vdots & & \vdots \\ \frac{\partial V_{N_{\text{user}},t}}{\partial Q_{1,t}^{\text{LOAD}}} & \cdots & \frac{\partial V_{N_{\text{user}},t}}{\partial Q_{N_{\text{user}},t}^{\text{LOAD}}} \end{bmatrix}_{N_{\text{user}} \times N} \quad (25)$$

随后,将用户参与需求响应的目标函数转化为二次规划问题的形式,如下所示。

$$\max \mathbf{U} = -\frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} - \mathbf{W}^T \mathbf{x} \quad (26)$$

$$\text{s.t. } \mathbf{A} \mathbf{x} \leq \mathbf{B}$$

$$\mathbf{x} = \begin{bmatrix} \Delta \mathbf{P}^{\text{LOAD}} \\ \Delta \mathbf{Q}^{\text{LOAD}} \end{bmatrix}, \mathbf{Q} = \begin{bmatrix} \mathbf{q}_P \\ \mathbf{q}_Q \end{bmatrix}, \mathbf{W} = \begin{bmatrix} \mathbf{w}_P \\ \mathbf{w}_Q \end{bmatrix} \quad (27)$$

$$\mathbf{A} = [\mathbf{O}_P^V \ \mathbf{O}_Q^V]^T, \mathbf{B} = [\mathbf{V}_{\max} - \mathbf{V}_{\min}]^T$$

式中 $\mathbf{q}_P$ 、 $\mathbf{q}_Q$ 、 $\mathbf{w}_P$ 、 $\mathbf{w}_Q$ 为从用户的拟二次函数提取的参数。

对偶函数为:

$$\inf_{\mathbf{x} \in \mathbb{R}^n} \left\{ \frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{W} \mathbf{x} + \boldsymbol{\lambda}^T (\mathbf{A} \mathbf{x} - \mathbf{B}) \right\} \quad (28)$$

式中 $\boldsymbol{\lambda}$ 为拉格朗日乘子。通过对拉格朗日函数求导并设其为 0,可以得到解析解 $\mathbf{x} = -\mathbf{Q}^{-1}(\mathbf{W} + \mathbf{A}^T \boldsymbol{\lambda})$,将其代入对偶问题,得到对偶问题 d 的形式:

$$\begin{cases} \min d = \frac{1}{2} \lambda^T A Q^{-1} A^T \lambda + (B + A Q^{-1} W)^T \lambda \\ \text{s.t. } \lambda \geq 0, \lambda = [\lambda_{\max} \lambda_{\min}]^T \end{cases} \quad (29)$$

梯度 ∇d 表示对偶函数相对于拉格朗日乘子 λ 的梯度, 其中 $\lambda_{\max}$ 、 $\lambda_{\min}$ 分别为上下限约束的乘子。

$$\nabla d = A Q^{-1} A \lambda + B + A Q^{-1} W \quad (30)$$

$$d(\lambda') \leq d(\lambda) + \nabla d(x)^T (\lambda' - \lambda) + \frac{L_2}{2} \|\lambda' - \lambda\|^2 \quad (31)$$

式中: L_2 为 Lipschitz 常数, 表示梯度的变化速率, 它由矩阵 $A Q^{-1} A$ 的最大特征值决定; λ' 为拉格朗日乘子的新迭代点。在二次规划中, 矩阵的最大特征值决定了最陡的曲率, 因此 Lipschitz 常数是对偶函数梯度变化的一个上界。为了优化计算, 文献[23]建议通过置换矩阵 $A Q^{-1} A$ 为对角矩阵, 并求解半正定规划问题 $L \geq A Q^{-1} A$, 以最小化 L 的逆, 从而得到一个有意义的 Lipschitz 常数矩阵, 作为对偶上升法中的步长, 有助于改进优化算法的收敛速度。

3.2.3 求解流程

在整个优化的过程中, 由于灵敏度矩阵会随着潮流变化而发生变化, 因此在 SAC 算法生成有源配电网中设备的出力值后, 需要以式(1)为目标函数、式(2)~(9)为约束, 计算当前系统的潮流分布, 并得到电压-有功功率、电压-无功灵敏度矩阵, 并向用户广播。随后, 基于广义快速对偶上升法, 以式(10)为目标函数、式(11)~(13)为约束, 求解用户的最优响应负荷值, 并再次计算配电网的潮流, 完成求解过程。综上, 本文所提的有源配电网状态-响应双层优化过程如表1所示。

4 仿真实验

4.1 基本环境

本文所使用的案例是修改后的 IEEE 33 节点系统, 如图2所示。分别在 23、26 节点接入 SVC, 18、25 节点接入分布式光伏, 在 10、16、19、23、32 节点接入参与需求响应的用户。本文从中国南方某区域的 1 000 d 的光伏和负荷数据中随机采样, 生成训练数据集, 充分体现环境的不确定性。节点电压幅值上下限设置为 1.04 p. u. 和 0.96 p. u.。实验过程基于 python3.10 和 Tensorflow2 作为基本环境, 并基于 Mosek 完成分布式优化的计算过程。所

表1 有源配电网状态-响应双层优化过程

Tab. 1 Active distribution network state-response two-layer optimization process

算法: 有源配电网状态-响应双层优化算法
输入: 配电网设备、用户负荷当前状态
S1: 初始化 Q 网络和策略网络参数 ϕ_1, ϕ_2, θ
S2: 初始化目标 Q 网络参数 ϕ'_1 , 令乘子 $\lambda = 0$
S3: 初始化经验池 $\mathcal{B}$ 、交互环境
S4: For $n \leftarrow 1$ to N_{train} do:
S5: For $t \leftarrow 1$ to T do:
S6: 策略网络生成动作 $a \sim \pi(a s)$
S7: 计算潮流与灵敏度矩阵 O_p^v, O_Q^v
S8: 向用户广播 O_p^v, O_Q^v, g^p, g^Q
S9: 用户根据式(28)参与响应, 调整负荷
S10: 配电网控制中心更新 λ
S11: 计算潮流并获得新的状态 s'
将数据储存到经验池 $[s, a, R, s'] \rightarrow \mathcal{B}$
S12: End For
S13: For each gradient step do
S14: 根据式(19)~(20)更新 Q 网络
S15: 根据式(21)更新策略网络
S16: 根据式(22)更新目标 Q 网络
S17: End For
S18: End For

有程序在配置 CPU 为 Intel(R) Core(TM) i7, 内存为 32 GB 的个人电脑执行。DRL 算法的相关参数具体如表2所示。

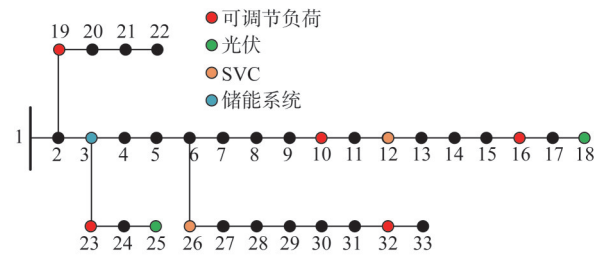


图2 IEEE 33节点系统

Fig. 2 IEEE 33-node system

同时, 为了说明本文模型在降低配电网运行成本和改善电压质量方面的有效性, 设置了3种运行模式进行实验与对比分析。

模式1: 采用传统的深度强化学习(DRL)方法, 仅对配电网状态层进行优化控制, 不考虑用户的需求响应。

模式2: 配电网和用户各自独立优化, 配网侧基于模式1进行优化控制, 而用户侧采用文献[23]

表 2 DRL 算法参数

Tab. 2 Parameters of DRL algorithm

参数类型	参数值	参数类型	参数值
Actor 和 Critic 神经元结构	[256, 256]	激活函数	ReLU
优化器	Adam	Actor 学习率	1×10^{-3}
折扣因子	0.99	Critic 学习率	5×10^{-4}
回合最大步数	24 h	批数据大小	64
步长	1 h	目标网络更新率	0.005
评价频率	3	熵系数	自动调整
经验池规模	5×10^5	噪声	Ornstein-Uhlenbeck

分布式优化方法进行调整。用户根据自身的用电效用函数和电压约束，独立地优化自身的负荷调整。

模式 3：采用本文提出的状态-响应双层优化方法(SAC-GFD)，在配电网状态层和用户响应层进行协同优化。状态层利用 SAC 算法优化配电网设备的控制策略，生成初始的电压状态；响应层中，用户根据配电网的实时电压状态，采用分布式算法进行负荷调整，从而直接响应和改善节点电压。该模式旨在解决传统 DRL 方法在电压约束处理和用户需求响应方面的不足，实现配电网与用户侧的协同优化，确保配电网的安全稳定运行。

4.2 有效性分析

4.2.1 网络损耗分析

图 3 展示了 3 种运行模式在 24 h 内的网络损耗情况。通过对比可以看出，模式 3 在大多数时间段内的网络损耗均低于模式 1 和 2，特别是在高负荷时段(如上午 7 点—11 点和下午 17 点—21 点)，表现出明显的优势。模式 1 的总网络损耗为 1.18 MW，模式 2 降低至 1.04 MW，而模式 3 进一步降低至 1.02 MW。相较于模式 1 和模式 2，模式 3 分别减少了 13.6 % 和 1.9 % 的总网络损耗，说明其在整体网络损耗控制上最为有效。

4.2.2 电压波动分析

选取了 10、15、18、20 点 4 个负荷较高的时刻，对 3 种模式下的系统节点电压幅值进行对比，如图 4 所示。

通过分析各个节点的电压幅值可以看出，模式 3 在各个节点的电压控制方面均表现出显著的优势。模式 1 在所有时刻中均在系统末尾的节点出现了电压越下限的情况。模式 2 在引入需求响应之后，电压越下限的情况有所改善。模式 3 在大部分时间段

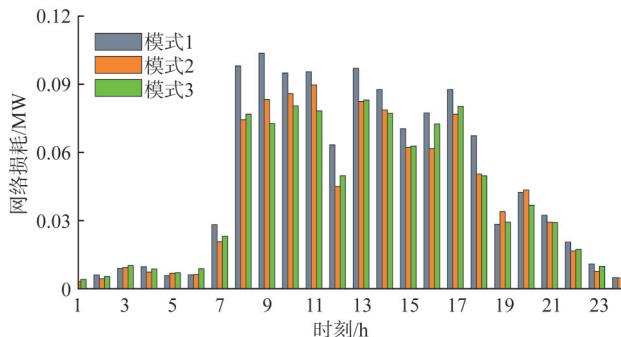


图 3 3 种模式下网络损耗对比

Fig. 3 Comparison of network losses in three modes

内的电压幅值高于模式 1 和模式 2，特别是在第 10 h 到第 24 h，模式 3 均没有出现电压越下限的情况。此外，整体而言，模式 3 下系统的节点电压幅值差较小，表现出更优异的电压控制能力。

4.2.3 效益结果分析

3 种模式下具体的日运行和效益结果如表 3—4 所示。储能调度情况如图 5 所示。

通过分析发现，3 种模式的储能损耗成本变化不大，模式 3 的弃风弃光成本比起模式 2 略微下降，但配电网运行成本的变化还是主要依赖于网损的变化。模式 3 的配电网运行成本为 1 929 元，相较于模式 1 和 2，分别降低了 4.45 %、5.96 %，折合年化成本为 3.13 万元、4.19 万元。

其次，在用户用电效益方面，模式 1 中用户不调整负荷时，其用电效益最低，而模式 2 和模式 3 分别为 6 302 元和 6 262 元，差异较小。

从网损角度看，模式 3 的网络损耗仅为 1.02 MW，低于模式 1 的 1.18 MW 和模式 2 的 1.04 MW。从电压安全角度看，模式 3 的最大电压差为 0.074 1 p. u.，显著优于模式 1 的 0.091 5 p. u. 和模式 2 的 0.081 7 p. u.，这意味着模式 3 在电压控制方面更为稳定。

最重要的是，模式 3 实现了电压越限次数为 0 次，而模式 1 和模式 2 分别为 79 次和 25 次，这表明模式 3 在电压安全性和稳定性方面的显著优势。虽然模式 2 的用户用电效益最好，但是在系统运行成本和安全方面的效果不如模式 3。

综上所述，模式 3 采用了本文所提的状态-响应双层优化框架，并结合了深度强化学习与广义快速对偶的 SAC-GFD 方法，在综合优化能力、动态适应性及电压控制等方面展现了较大优势，证明了其

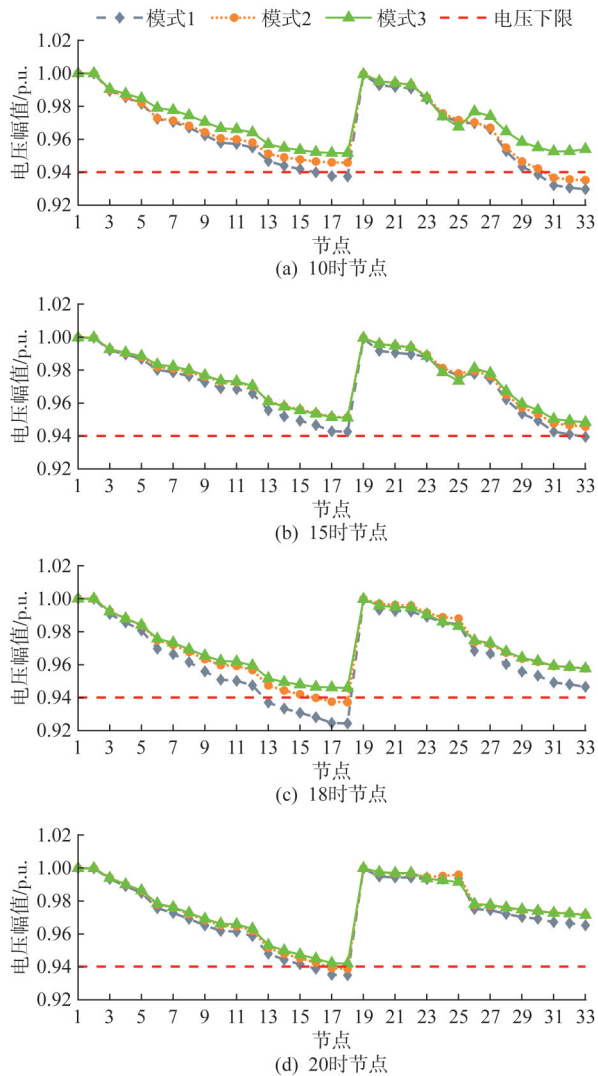


图 4 3 种模式下不同时刻的节点电压幅值对比

Fig. 4 Comparison of node voltage amplitude at different moments in three modes

表 3 3 种模式的日运行结果对比

Tab. 3 Comparison of daily operating results for the 3 modes

方法	总网络损耗/MW	最大电压差/p. u.	电压越限次数
模式 1	1.18	0.091 5	79
模式 2	1.04	0.081 7	25
模式 3	1.02	0.074 1	0

有效性。

4.3 收敛性分析

已有研究表明, 深度强化学习算法在电力系统领域有广泛应用。本文将重点比较 SAC 算法和深度确定性策略梯度 (deep deterministic policy gradient, DDPG) 算法, 这两种算法均用于连续动作优化, 并采用离策略学习和演员-评论家架构。DDPG 通过

表 4 3 种模式的日效益结果对比

Tab. 4 Comparison of daily benefit results for the 3 modes

方法	用户用电效益	配电网运行成本	网损成本	储能损耗成本	弃风弃光成本
模式 1	4 515	2 042.8	1 811	49.8	182
模式 2	6 302	2 014.8	1 752	50.8	212
模式 3	6 262	1 927.6	1 671	51.6	205

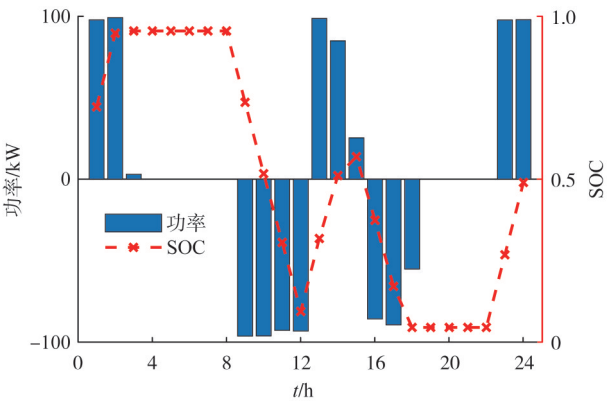


图 5 24 h 内储能调度情况

Fig. 5 Dispatch of energy storage in 24 hours

确定性策略网络直接预测动作, 并通过价值网络评估该动作的预期回报。SAC 算法通过提高策略熵来激励更广泛的探索, 而 DDPG 则通过在确定性策略输出中添加噪声 (如 Ornstein-Uhlenbeck 噪声) 来促进对环境的有效探索。

图 6 展示了 SAC-GFD 算法与 DDPG-GFD、传统的 SAC 算法在训练过程中的对比。

由图 6 可见, 虽然传统的 SAC 算法在训练初期可以得到较好的用户需求响应收益, 但随着训练的深入, SAC-GFD 算法的效果更好, 并且能更有效地降低了配电网的运行成本。这是因为传统的 DRL 算法只能得到趋于最优的控制结果, 而分布式算法可以从数学模型上得到最优值。此外, 从图 7(d) 可以看出, SAC-GFD 算法的弃光成本较低, 说明其在促进分布式光伏出力的本地消纳方面表现更优。同时, 图 7(e) 显示出 SAC-GFD 算法在降低总网络损耗成本上具有更好效果。总体而言, SAC-GFD 算法在训练速度和最终结果上均优于 DDPG-GFD、传统的 SAC 算法, 突显了其在电力系统优化中的潜在优势。

5 结语

本文基于 DRL 和分布式优化算法设计了一个

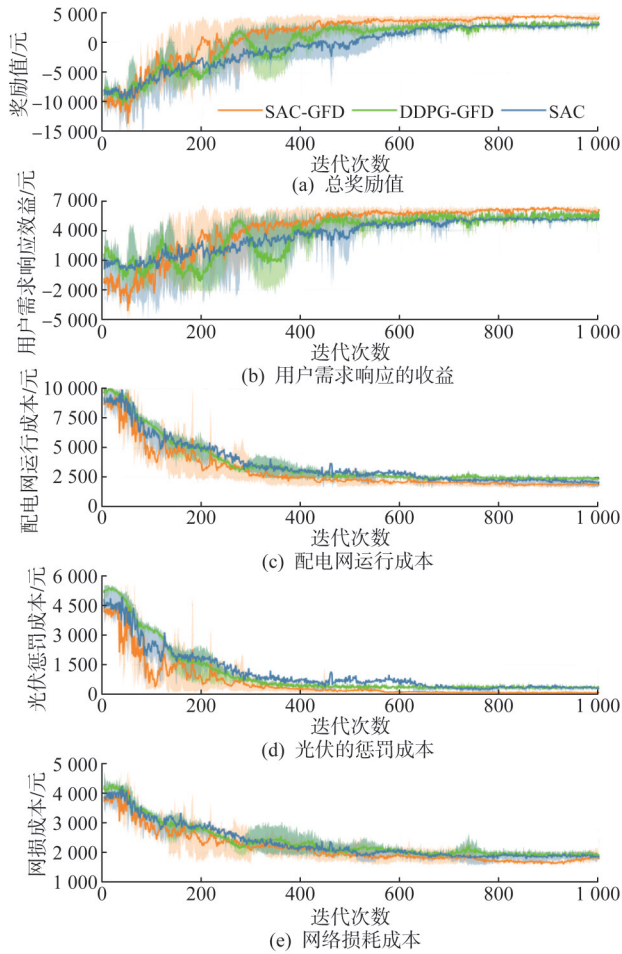


图6 3种算法训练结果对比图

Fig. 6 Comparison of the training results for three algorithms

状态-响应双层配电网运行优化策略。在状态层利用SAC算法优化光伏、储能、SVC设备的连续动作策略，在响应层基于分布式的广义加速对偶上升法优化各个用户的负荷出力。最后在改进的IEEE 33节点系统中进行仿真对比实验。结果表明：

1) 本文所提的模型将有模型分布式优化方法与无模型数据驱动法结合，能够在快速输出设备和用户最优动作值的前提下，保证结果不违反约束。

2) 本文所提算法能够保证较好的用户效益前提下，有效降低配电网运行成本、总网络损耗和节点电压偏差。

3) 与DDPG-GFD算法相比，SAC-GFD算法具有较快的训练速度和良好的优化结果。

参考文献

[1] 张勇军, 秦绍基, 孙健, 等. 低压台区透明化技术研究综述[J/OL]. 南方电网技术: 1-12. [2024-06-24]. <https://nfdwjs.csg.cn/gateway-web/zh/debutDetail.html?serialNum=20231204006>.

ZHANG Yongjun, QIN Shaoji, SUN Jian, et al. Research review on transparent technology of low-voltage substation area [J/OL]. Southern Power System Technology: 1-12. [2024-06-24]. <https://nfdwjs.csg.cn/gateway-web/zh/debutDetail.html?serialNum=20231204006>.

[2] CHENG R, LI Q, ZHANG Y, et al. Three-phase power cooperative management of distribution network with high distributed energy resources penetration [J]. Energy Reports, 2022(8): 239-250.

[3] 葛磊蛟, 刘航旭, 孙永辉, 等. 智能配电网多元电力用户群体特性精准感知技术综述[J]. 电力系统自动化, 2023, 47(20): 174-191.

GE Leijiao, LIU Hangxu, SUN Yonghui, et al. Review on accurate awareness technology for characteristics of diversified power user groups in smart distribution network [J]. Automation of Electric Power Systems, 2023, 47(20): 174-191.

[4] 朱志芳, 胡柳君, 陈丽萍, 等. 基于压力-状态-响应框架的有源配电网运行能效评价[J]. 广东电力, 2023, 36(12): 17-29.

ZHU Zhifang, HU Liujun, CHEN Liping, et al. Energy efficiency evaluation of active distribution network operation based on pressure-state-response framework [J]. Guangdong Electric Power, 2023, 36(12): 17-29.

[5] 刘志勇, 曾庆彬, 吴晗. 高比例新能源接入的农村配电网网架多目标协同规划方法[J]. 广东电力, 2023, 36(9): 34-42.

LIU Zhiyong, ZENG Qingbin, WU Han. Multi-objective collaborative planning method for rural distribution network grid with high proportion of new energy [J]. Guangdong Electric Power, 2023, 36(9): 34-42.

[6] LIU Z, LIU S, LI Q, et al. Optimal day-ahead scheduling of islanded microgrid considering risk-based reserve decision [J]. Journal of Modern Power Systems and Clean Energy, 2021, 9(5): 1149-1160.

[7] DENG Z, LIU M, CHEN H, et al. Optimal scheduling of active distribution networks with limited switching operations using mixed-integer dynamic optimization [J]. IEEE Transactions on Smart Grid, 2019, 10(4): 4221-4234.

[8] 郭一帆, 欧阳森, 张晋铭, 等. 含小水电配电网重要负荷评估及其水光储优化配置方法[J]. 广东电力, 2024, 37(5): 32-42.

GUO Yifan, OUYANG Sen, ZHANG Jinming, et al. Assessment of critical load and optimization configuration method for small hydropower distribution network with integrated hydro-PV-energy storage [J]. Guangdong Electric Power, 2024, 37(5): 32-42.

[9] 黄伟, 刘斯亮, 王武, 等. 长时间尺度下计及光伏不确定性的配电网无功优化调度[J]. 电力系统自动化, 2018, 42(5): 154-162.

HUANG Wei, LIU Siliang, WANG Wu, et al. Optimal reactive power dispatch with long-time scale in distribution network considering uncertainty of photovoltaic [J]. Automation of Electric Power Systems, 2018, 42(5): 154-162.

[10] 陈伟, 陈龙康, 魏占宏, 等. 基于净能力及二阶锥规划的分布式光储多场景协同优化策略[J]. 电力自动化设备, 2024, 44(6): 26-34.

CHEN Wei, CHEN Longkang, WEI Zhanhong, et al. Multi-scenario collaborative optimization strategy of distributed

- photovoltaic and energy storage based on net ability and second-order cone programming [J]. Electric Power Automation Equipment, 2024, 44(6): 26–34.
- [11] 符杨, 武洁, 苏向敏, 等. 不平衡主动配电网电压无功两阶段多目标鲁棒优化[J]. 电力自动化设备, 2024, 44(5): 79–87. FU Yang, WU Jie, SU Xiangjing, et al. Two-stage multi-objective robust optimization of voltage/var for unbalanced active distribution network [J]. Electric Power Automation Equipment, 2024, 44(5): 79–87.
- [12] 林舜江, 冯祥勇, 梁炜焜, 等. 新能源电力系统不确定优化调度方法研究现状及展望[J]. 电力系统自动化, 2024, 48(10): 20–41. LIN Shunjiang, FENG Xiangyong, LIANG Weikun, et al. Research status and prospect of uncertain optimal dispatch methods for renewable energy power systems [J]. Automation of Electric Power Systems, 2024, 48(10): 20–41.
- [13] LIU H, WU W. Online multi-agent reinforcement learning for decentralized inverter-based volt-var control[J]. IEEE Transactions on Smart Grid, 2021, 12(4): 2980–2990.
- [14] 杨龙月, 任烜辰, 蔡智鹏, 等. 高光伏渗透率配电网电压控制策略研究综述[J]. 电网技术, 2024, 48(12): 5056–5070. YANG Longyue, REN Xuancheng, CAI Zhipeng, et al. Review of voltage control strategies for high PV penetration distribution networks[J]. Power System Technology, 2024, 48(12): 5056–5070.
- [15] 张军, 钟康骅, 张勇军, 等. 基于混合博弈的多区域电-储共享运营模式与经济效益分析[J]. 电力系统自动化, 2024, 48(3): 31–41. ZHANG Jun, ZHONG Kanghua, ZHANG Yongjun, et al. Hybrid game based sharing operation mode for multi-regional electric power and energy storage and its economic benefit analysis[J]. Automation of Electric Power Systems, 2024, 48(3): 31–41.
- [16] 姚星安, 刘文昊, 龚学良, 等. 基于改进对偶动态规划的产销者点对点交易全分散式出清方法[J]. 电网技术, 2023, 47(11): 4564–4575. YAO Xing'an, LIU Wenhao, GONG Xueliang, et al. Fully decentralized clearing for peer-to-peer trading of prosumers based on improved dual dynamic programming [J]. Power System Technology, 2023, 47(11): 4564–4575.
- [17] 李彬贤, 李佳勇, 海征, 等. 基于多智能体深度强化学习的配电网三相不平衡在线治理方法[J]. 中国电机工程学报, 2025, 45(5): 1729–1741. LI Binxian, LI Jiayong, HAI Zheng, et al. Online three-phase imbalance management method for distribution networks based on multi-agent deep reinforcement learning [J]. Proceedings of the CSEE, 2025, 45(5): 1729–1741.
- [18] 朱涛, 海迪, 李文云, 等. 基于图神经网络与强化学习的配电网电压与无功功率优化方法[J]. 南方电网技术, 2024, 18(11): 67–78. ZHU Tao, HAI Di, LI Wenyun, et al. Voltage and reactive power optimization method for distribution networks based on graph neural network and reinforcement learning [J]. Southern Power System Technology, 2024, 18(11): 67–78.
- [19] LIU H, WU W, WANG Y. Bi-level off-policy reinforcement learning for two-timescale volt/var control in active distribution networks [J]. IEEE Transactions on Power Systems, 2023, 38(1): 385–395.
- [20] 张玲玉. 基于多智能体强化学习的主动配电网分区协同电压控制研究[D]. 哈尔滨: 哈尔滨工程大学, 2024.
- [21] 梅铭洋, 寇鹏, 张智豪, 等. 面向主动配电网的安全多智能体深度强化学习电压优化控制[J]. 西安交通大学学报, 2023, 57(12): 157–167. MEI Mingyang, KOU Peng, ZHANG Zhihao, et al. Safe multi-agent deep reinforcement learning for optimal voltage control in active distribution networks [J]. Journal of Xi'an Jiaotong University, 2023, 57(12): 157–167.
- [22] SAYED A R, WANG C, ANIS H I, et al. Feasibility constrained online calculation for real-time optimal power flow: a convex constrained deep reinforcement learning approach [J]. IEEE Transactions on Power Systems, 2023, 38(6): 5215–5227.
- [23] FENG C, LIANG B, LI Z, et al. Peer-to-peer energy trading under network constraints based on generalized fast dual ascent [J]. IEEE Transactions on Smart Grid, 2023, 14(2): 1441–1453.

收稿日期: 2024-08-02; 网络首发日期: 2024-12-13

作者简介:

胡柳君(1996), 女, 工程师, 硕士, 研究方向电力系统规划研究及系统分析, 15116330584@163.com;

董红(1982), 女, 高级工程师, 硕士, 研究方向为电网规划、系统分析, donghong792@163.com;

曾繁宏(1994), 男, 工程师, 硕士, 研究方向为电力系统规划, 1206868572@qq.com。