

DOI: 10.19783/j.cnki.pspc.220050

基于深度强化学习的含高比例可再生能源配电网 就地分散式电压管控方法

徐博涵, 向月, 潘力, 方梦秋, 彭光博, 刘友波, 刘俊勇

(四川大学电气工程学院, 四川 成都 610065)

摘要: 含有可再生能源并网的区域电网存在通信条件差、量测设备不足、不同节点的电压管控设备难以协同等问题, 因此提出一种基于深度强化学习的分散式就地电压管控方法。该方法首先将缺少量测数据的电压管控问题转化为部分可观的马尔科夫决策问题, 构建了以网络损耗最小为优化目标的多智能体分散式电压管控框架。然后采用多智能体深度确定性策略梯度算法对智能体进行离线训练, 并使用训练完成的智能体进行在线电压管控。最后, 基于改进的 IEEE33 节点系统进行了算例仿真和分析。结果表明, 各智能体可以根据各自节点的电气信息求解出近似的全局最优解。

关键词: 多智能体; 电压管控; 量测数据不足; 多智能体深度确定性策略梯度算法

Local decentralized voltage management of a distribution network with a high proportion of renewable energy based on deep reinforcement learning

XU Bohan, XIANG Yue, PAN Li, FANG Mengqiu, PENG Guangbo, LIU Youbo, LIU Junyong

(College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

Abstract: A multi-agent decentralized local voltage control method based on the deep reinforcement learning is proposed. This is needed because there are some problems in the regional grid with renewable energy, such as poor communication conditions, insufficient measurement equipment, and difficult coordination of voltage control equipment at different nodes. First, this method transforms the voltage control problem lacking measurement data into a partial observable Markov decision problem, and a multi-agent decentralized voltage control framework with the optimization goal of minimizing network loss is constructed. Then, a multi-agent deep deterministic policy gradient algorithm is used to train the agents offline, and the trained agents are used for online voltage control. Finally, an example is simulated and analyzed based on the improved IEEE33 bus system. The results show that each agent can solve the approximate global optimal solution according to the electrical information of its own node.

This work is supported by the Science and Technology Project of State Grid Corporation of China (No. SGTJDK00DWJS2100039).

Key words: multi-agent; voltage control; insufficient measurement data; multi-agent deep deterministic policy gradient algorithm

0 引言

近年来, 为了应对环境问题, 世界各国都在推动能源系统的低碳化、清洁化转型。在此背景下, 我国提出了“双碳”要求, 力争在 2030 年前实现碳

达峰, 在 2060 年前实现碳中和^[1]。在电力行业, 由于风电和光伏等新型可再生能源在发电过程中有着零碳排放和零边际成本等优点, 其发电占比正在快速上升^[2]。但是, 随着可再生能源装机容量的增大, 电网的电压波动变得更为剧烈, 无功电压补偿技术成为了一种重要的电压管控手段^[3]。由于大多数可再生能源发电都是通过逆变器进行并网的, 具有利用空闲容量提供无功补偿的能力, 探索可再生能源

基金项目: 国家电网科技项目资助“基于群体智能的能源互联网多源协同运行关键技术研究”(SGTJDK00DWJS2100039)

发电的无功管控方法对加强电网电压管控能力具有现实意义^[4]。

目前电压管控领域的研究方向主要分为集中式电压管控、分布式电压管控以及就地电压管控3种。集中式管控通过设置一个中心处理器,来接收整个电网的全部信息,作出全局最优决策。文献[5]基于动态规划法实现了 SVC 设备的无功电压管控,文献[6-7]基于混合整数二阶锥规划法实现了 SVC 和投切电容器的协同电压管控。文献[8]基于粒子群算法实现了无功电压管控,具有更强的全局寻优能力。文献[9]基于鲁棒算法实现了考虑不确定性辐射状配电网的电压管控。但是集中式电压管控方法都需要良好的通信条件,而且中心处理器面临着巨大的计算压力。

为了克服集中式电压管控的计算压力,分布式算法开始应用于电压管控领域。文献[10]基于分布模拟退火法实现了区域自治的无功电压管控。文献[11]基于分散决策实现了无功电压优化,降低了电网损耗。文献[12]基于交替方向乘子算法实现了邻域信息交互下的自主电压管控。但是,分布式算法依然依赖于完备的系统信息以及良好的通信条件。在通信条件较差、量测数据不完备的区域电网中,一般采用就地管控策略,如 $Q(V)$ 下垂控制等,但是就地管控策略存在可调节资源利用不充分、网络损耗难以优化等问题^[13]。

为了解决上述问题,国内外学者对基于数据驱动的电压管控技术进行了大量研究,并取得了初步成果。文献[14]基于深度确定性策略梯度算法进行电压管控,实现了基于数据的离线算法训练,求解速度较基于模型求解的算法有大幅的提升。为了增强算法的鲁棒性,文献[15]采用多智能体算法设置多个智能体分别对电网各个区域进行电压管控,实现了基于数据的分布式管控。但是,该方法依然需要完备的系统观测数据。为此,文献[16]提出了一种基于双层深度确定性策略梯度算法的电压管控方法,训练两种智能体来分别完成目标电压设置和目标电压逼近的任务,实现了在少量量测数据下的无功电压管控。文献[17]提出使用卷积神经网络拟合潮流无功优化函数,实现了低感知态势下含高比例清洁能源的电压管控,但是上述方法依然无法摆脱对通信条件的依赖。

基于上述研究现状,本文设计了一种基于深度强化学习的含高比例可再生能源配电网就地分散式电压管控。首先建立了可再生能源并网的电压管控机理模型,为智能体提供交互环境。然后,将电压管控问题转化为部分可观的马尔科夫决策过程(partially

observable markov decision process, POMDP),构建了智能体的观测空间、动作空间以及回报函数,引导各智能体协同进行电压管控,降低网络损耗,并防止电压越界。之后,使用多智能体深度确定性策略梯度(multi-agent deep deterministic policy gradient, MADDPG)算法对多智能体进行训练,实现了各智能体间决策的协同进化,训练完成的智能体在进行决策时可以计及其他智能体决策的耦合影响,具有更好的决策稳定性。最后,使用改进的 IEEE33 节点系统进行算例仿真和分析,验证了本文所提管控方法的有效性。

1 电压管控机理模型

本节首先对电压协同控制框架中的可再生能源设备进行建模,基于不同设备的无功补偿特性定义其运行区间,然后给出了潮流模型,用于下文的智能体训练过程。

1.1 风电出力模型

文献[18]表明风速服从韦伯分布,因此可以使用式(1)来表示风电机组的有功功率。同时,风电机组一般配备有连续的无功补偿装备^[19],所以风电机组的实际无功功率可以用式(2)表示。

$$P_{WT} = \begin{cases} 0, & v < v_{run}, v > v_{cut} \\ P_{WT} \frac{v - v_{run}}{v_{rate} - v_{run}}, & v_{run} \leq v \leq v_{rate} \\ P_{WT}, & v_{rate} \leq v \leq v_{cut} \end{cases} \quad (1)$$

$$\begin{cases} P_{WT} \tan \varphi + Q_C^{\min} \leq Q_{WT} \leq P_{WT} \tan \varphi + Q_C^{\max} \\ \cos \varphi_{\min} \leq \cos \varphi \leq \cos \varphi_{\max} \end{cases} \quad (2)$$

式中: v_{run} 、 v_{cut} 、 v_{rate} 分别为风电机组运行的投入风速、切出风速和额定风速; v 表示实际风速, $v \sim W(\lambda, b)$; P_{WT} 、 P_{WT} 表示风电机组的实际出力和额定出力; Q_{WT} 为风电机组的实际无功功率; Q_C^{\min} 、 Q_C^{\max} 分别为无功补偿装置的无功功率下限和上限; $\cos \varphi$ 为风电机组的功率因数; $\cos \varphi_{\max}$ 、 $\cos \varphi_{\min}$ 分别为功率因数的上限和下限,一般取 ± 0.95 。

1.2 光伏出力模型

文献[20]表明光照强度服从贝塔分布,因此可以使用式(3)来表示光伏电源的有功功率。

$$P_{PV} = \begin{cases} P_{Pv} \frac{I}{I_r}, & I \leq I_r \\ P_{Pv}, & I > I_r \end{cases} \quad (3)$$

式中: I 、 I_r 分别表示实际的光照强度和光伏电源的额定光照强度, $I \sim B(\alpha, \beta)$; P_{PV} 、 P_{Pv} 分别表示

光伏电源的有功功率和额定有功功率。

同时, 根据文献[21]的研究表明, 光伏电源可以通过逆变器优化(optimal inverter dispatch, OID)策略来控制光伏电源输出的无功功率, 使用 OID 策略的光伏电源运行区间可表示为

$$\begin{cases} 0 \leq P_{PV} \leq P_{PVav} \\ P_{PV}^2 + Q_{PV}^2 \leq S_{PV}^2 \\ -P_{PV} \tan \theta \leq Q_{PV} \leq P_{PV} \tan \theta \end{cases} \quad (4)$$

式中: P_{PV} 、 Q_{PV} 表示光伏电源的有功功率和无功功率; P_{PVav} 为 OID 策略下光伏电源的最大有功功率; S_{PV} 表示光伏电源并网逆变器的设备容量; θ 为 OID 策略下光伏电源的最大功率角。

1.3 负荷的不确定模型

文献[22]表明可以使用正态分布来表示负荷的不确定性, 如式(5)一式(8)所示。

$$f(P_L) = \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left[-\frac{(P_L - \mu_p)^2}{2\sigma_p^2}\right] \quad (5)$$

$$f(Q_L) = \frac{1}{\sqrt{2\pi}\sigma_q} \exp\left[-\frac{(Q_L - \mu_q)^2}{2\sigma_q^2}\right] \quad (6)$$

$$\mu_p \sim U(\mu_{p,\min}, \mu_{p,\max}) \quad (7)$$

$$\mu_q \sim U(\mu_{q,\min}, \mu_{q,\max}) \quad (8)$$

式中: P_L 、 Q_L 表示负荷的有功功率和无功功率; σ_p 、 σ_q 表示负荷有功功率和无功功率的标准差; μ_p 、 μ_q 表示负荷有功功率和无功功率的期望, 服从平均分布; $\mu_{p,\min}$ 、 $\mu_{p,\max}$ 和 $\mu_{q,\min}$ 、 $\mu_{q,\max}$ 表示 μ_p 和 μ_q 平均分布的区间下限和上限。

1.4 潮流模型

本文仅考虑稳态下的电压管控, 潮流模型如式(9)一式(11)所示。

$$P_{ij} = G_{ij}V_i^2 - G_{ij}V_iV_j \cos \theta_{ij} - B_{ij}V_iV_j \sin \theta_{ij}, \forall ij \in \mathcal{E} \quad (9)$$

$$Q_{ij} = -B_{ij}V_i^2 + B_{ij}V_iV_j \cos \theta_{ij} - G_{ij}V_iV_j \sin \theta_{ij}, \forall ij \in \mathcal{E} \quad (10)$$

式中: V_i 为节点 i 的电压幅值; P_{ij} 、 Q_{ij} 分别为从节点 i 流向节点 j 的有功功率和无功功率; θ_{ij} 为节点 i 和节点 j 之间的相角差; $G_{ij} + jB_{ij}$ 为节点 i 和节点 j 之间的支路导纳。

同时, 由功率守恒定理可得

$$G_iV_i^2 + \sum_{ij \in \mathcal{E}} P_{ij} = \begin{cases} -P_{Di}, i \in N/N_{IR} \\ P_{Gi} - P_{Di}, i \in N_{IR} \end{cases} \quad (11)$$

$$-B_iV_i^2 + \sum_{ij \in \mathcal{E}} Q_{ij} = \begin{cases} -Q_{Di}, i \in N/N_{IR} \\ Q_{Gi} - Q_{Di}, i \in N_{IR} \end{cases} \quad (12)$$

式中: $G_i + jB_i$ 为节点 i 的并联导纳; P_{Di} 、 Q_{Di} 、 P_{Gi} 、 Q_{Gi} 分别为节点 i 的有功负荷、无功负荷、可再生能源发电的有功功率、无功功率; N_{IR} 和 N 分别表示可再生能源节点的集合和电网全部节点的集合。

2 量测数据不足下的多智能体电压管控模型

本节首先对智能体模型中涉及的马尔科夫决策过程进行介绍, 然后将多智能体就地分散式电压管控问题转化为 POMDP 问题, 建立了多智能体电压管控模型。

2.1 马尔科夫决策过程

在机器学习领域中, 一般采用马尔科夫决策过程来描述智能体与环境的交互过程。在一个马尔科夫决策过程中, 智能体可以从外界环境中获取观测数据, 并使用自身策略来根据观测数据作出决策。受到决策动作的逆向影响, 环境状态发生更新, 智能体根据新的环境状态获取回报, 并开始新一轮的马尔科夫决策过程。马尔科夫决策过程的流程图如图 1 所示。

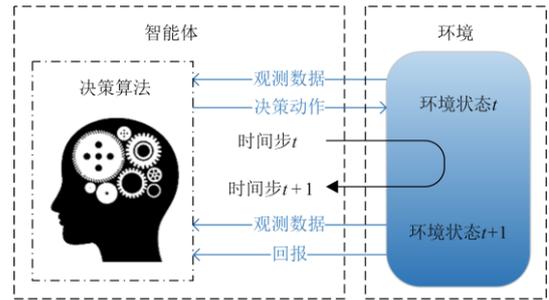


图 1 马尔科夫决策过程流程图

Fig. 1 Flow chart of Markov decision process

多智能体马尔科夫决策过程的全部信息可以整合为一个元组, 如式(13)所示。

$$(S, [O_i]_N, \gamma, [A_i]_N, \rho, [R_i]_N) \quad (13)$$

式中: S 为智能体的环境状态空间, 包含外界环境中可以用于马尔科夫决策的全部信息; O_i 为智能体 i 通过观测取得的观测数据, 当 $O_i = S$ 时称为完全可观的马尔科夫决策过程, 反之则为部分可观的马尔科夫决策过程; $[\cdot]_N$ 为 N 个智能体的数据集合; γ 为折扣因子, 表示当前动作对下一时间步智能体获得回报的影响效果; A_i 表示智能体 i 决策的动作值, 由智能体自身策略和观测数据决定, 可表示为

$$A_i = \pi_i(O_i) \quad (14)$$

式中, $\pi_i(\cdot)$ 为智能体 i 的策略函数。

ρ 为环境转移率, 表示在受到智能体动作的影响后当前环境状态更新为下一环境状态的概率, 可

表示为

$$P(S_{t+1}|S_t \times [A_t]_N) = \rho \quad (15)$$

式中: t 为时间步, 表示进行马尔科夫决策的次数; S_{t+1} 为新的环境状态。

r_t 表示全部智能体完成动作后, 分配给智能体 i 的回报, 由新的环境状态决定, 表示为式(16)。

$$r_t = f_i(S_{t+1}) \quad (16)$$

式中, $f_i(\cdot)$ 为智能体 i 的回报函数, 需要根据智能体的任务目标进行设置。

在多次的马尔科夫决策过程中, 智能体使用 Q 值来表示自己当前动作获取的总收益期望, 并以最大化 Q 值为目标来更新自身的策略函数。

$$\max_{\pi} Q_{\text{value},i} = E \left(\sum_{t=1}^T \gamma^t r_{i,t} \right) \quad (17)$$

式中: $Q_{\text{value},i}$ 为智能体 i 的 Q 值; T 为马尔科夫决策过程的总时间步数。

2.2 基于 POMDP 的多智能体电压管控模型

本节将量测数据不足下的电压协同管控问题转化为 POMDP 问题, 对其环境状态空间、观测空间、动作空间和回报函数进行定义, 引导多智能体根据少量观测数据控制分布式电源协同进行电压管控。

2.2.1 环境状态空间

电压管控问题的物理环境为网络潮流, 所以智能体的环境状态空间应包含潮流计算所需要的全部信息, 如式(18)所示。

$$S = \{P_i, Q_i, V_i, \theta_i\}, i \in N \quad (18)$$

式中, P_i 、 Q_i 、 θ_i 分别为节点 i 的有功功率、无功功率和电压相角。

2.2.2 观测空间

由于本文考虑的是分散就地电压管控, 各智能体仅根据各自管控的可再生能源设备接入节点的电气信息(节点电压及有功、无功功率)进行决策, 因此设计智能体的观测空间如式(19)所示。

$$O_i = \{P_i, Q_i, V_i, \theta_i\}, i \in N_{\text{IR}} \quad (19)$$

2.2.3 动作空间

动作空间为电压管控问题中智能体可以控制的设备变量, 本文为可再生能源设备发电的无功功率。同时, 为了满足式(2)和式(4)的无功功率上下限约束, 本文使用 \tanh 函数作为神经网络的激活函数, 将神经网络的输出限制在 $[-1, 1]$, 并线性放缩至设备的无功功率限制内, 如式(20)所示。

$$A_i = Q_{Gi} = \pi_i(O_{i,t} | \delta^{\pi_i}) \cdot Q_{Gi}^{\text{max}}, i \in N_{\text{IR}} \quad (20)$$

式中: Q_{Gi} 、 Q_{Gi}^{max} 为节点 i 处可再生能源发电的无功功率和最大无功功率; δ^{π_i} 为策略函数 $\pi_i(O_i | \delta^{\pi_i})$ 的网络参数。

2.2.4 回报函数

由于智能体会以最大化获得回报为向来更新策略函数, 因此设计如式(21)的回报函数, 引导智能体降低网络损耗并防止电压越限。

$$r_t = -\sum_{ij \in \mathcal{E}} G_{ij} (V_i^2 + V_j^2 - 2V_i V_j \cos \theta_{ij}) - c \sum_{j=1}^{N_{\text{vio},i}} (V_j - 1) \quad (21)$$

式中: $N_{\text{vio},i}$ 为受节点 i 无功管控的节点中出现电压越限问题的节点集合; c 为电压违规惩罚系数。

由于电压管控后的环境状态可以通过潮流计算确定, 且电压管控动作不具备长期效益, 所以设置 $\rho=1$, $\gamma=0$ 。至此完成了多智能体电压管控问题到 POMDP 的转化, POMDP 框架如图 2 所示。

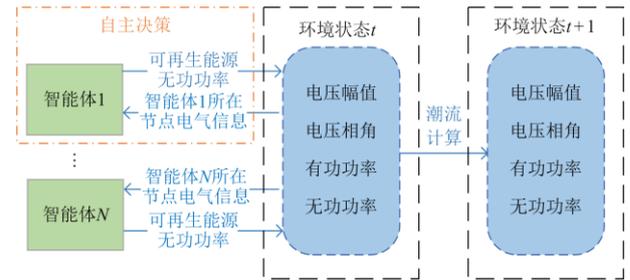


图 2 多智能体电压管控 POMDP 框架

Fig. 2 POMDP framework of multi-agent voltage control

3 基于 MADDPG 的多智能决策算法

3.1 多智能体的集中式离线训练

MADDPG 算法是一种新兴的基于多智能体 Actor-Critic 网络的深度强化学习算法。它引入了双层多智能体 Actor-Critic 网络的概念, 在智能体进行集中式训练前复制一份初始的多智能体 Actor-Critic 网络作为目标网络, 并使用软更新技术对双层网络的参数进行异步更新, 大幅度提升了多智能体训练的速度和稳定性^[23]。MADDPG 算法的流程图如图 3 所示。

3.1.1 Actor 神经网络迭代过程

MADDPG 算法中 Actor 神经网络负责拟合 POMDP 中的策略函数 $\pi_i(\cdot)$ 。同时, 为了防止智能体过早地收敛到局部最优解, MADDPG 算法在 Actor 神经网络中加入了一个随着学习过程逐渐衰减的噪声来增强智能体对最优策略的探索能力, 如式(22)所示。

$$A_{i,t} = \pi_i(O_{i,t} | \delta^{\pi_i}) + v_{i,t} \quad (22)$$

$$v_{i,t+1} = (1 - \tau_v) v_{i,t} \quad (23)$$

式中: $A_{i,t}$ 、 $O_{i,t}$ 、 $v_{i,t}$ 为智能体 i 在 t 时间步的动作值、观测值和噪声; τ_v 为噪声的衰减系数。

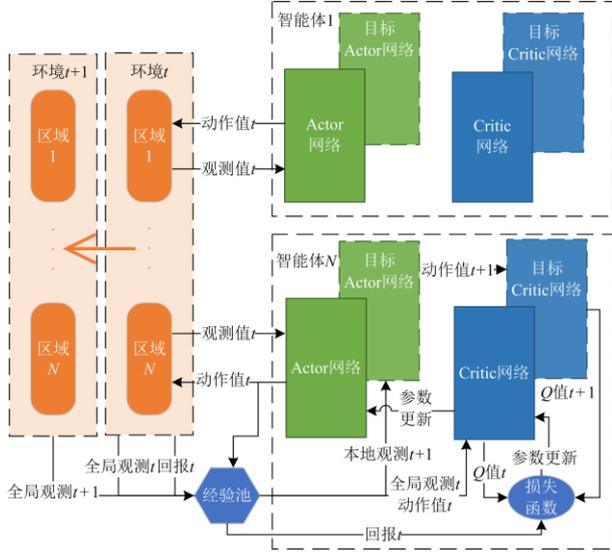


图3 MADDPG 算法流程图

Fig. 3 Flow chart of MADDPG algorithm

Actor 神经网络参数的更新梯度由 Critic 神经网络计算的 Q 值决定, 并根据预设的更新系数进行更新, 可表示为

$$\frac{\partial \pi_i}{\partial \delta^{\pi_i}} = \frac{\partial C_i(O_t, A_t | \delta^{C_i})}{\partial A_t} \cdot \frac{\partial \pi_i(O_{i,t} | \delta^{\pi_i})}{\partial \delta^{\pi_i}} \quad (24)$$

$$\delta_{t+1}^{\pi_i} = \delta_t^{\pi_i} + \varepsilon_{\pi} \frac{\partial \pi_i}{\partial \delta^{\pi_i}} \quad (25)$$

式中: $C_i(\cdot | \delta^{C_i})$ 为智能体 i 的 Critic 神经网络函数, 用于预测智能体 i 的 Q 值; δ^{C_i} 为 $C_i(\cdot | \delta^{C_i})$ 的网络参数; O_t 、 A_t 分别为全部智能体在时间步 t 的观测值集合和动作值集合; ∂ 为偏导符号; ε_{π} 为 Actor 网络参数的更新系数。

3.1.2 Critic 网络迭代过程

MADDPG 算法根据贝尔曼公式构建了 Critic 网络参数的误差函数, 为 Critic 网络参数提供了更新梯度, 计算公式为

$$L(\delta^{C_i}) = (y_{i,t} - C_i(O_t, A_t | \delta^{C_i}))^2 \quad (26)$$

$$y_{i,t} = r_{i,t} + \gamma(C_i(O_{t+1}, \pi_i(O_{i,t+1} | \delta^{\pi_i}) | \delta^{C_i})) \quad (27)$$

$$\frac{\partial L(\delta^{C_i})}{\partial \delta^{C_i}} = 2(y_{i,t} - C_i(O_t, A_t | \delta^{C_i})) \cdot \frac{\partial C_i(O_t, A_t)}{\partial \delta^{C_i}} \quad (28)$$

$$\delta_{t+1}^{C_i} = \delta_t^{C_i} + \varepsilon_C \frac{\partial L(\delta^{C_i})}{\partial \delta^{C_i}} \quad (29)$$

$$\delta_{t+1}^{\pi_i} = \tau_{\pi} \delta_t^{\pi_i} + (1 - \tau_{\pi}) \delta_t^{\pi_i} \quad (30)$$

$$\delta_{t+1}^{C_i} = \tau_C \delta_t^{C_i} + (1 - \tau_C) \delta_t^{C_i} \quad (31)$$

式中: $L(\delta^{C_i})$ 为 Critic 神经网络参数的误差函数, 用于为网络参数更新提供梯度; $y_{i,t}$ 为贝尔曼公式计算的 Q 值; $\pi_i(\cdot)$ 、 $C_i(\cdot)$ 分别为智能体 i 的 Actor 目标网络和 Critic 目标网络; ε_C 为 Critic 神经网络参数的更新系数; δ^{π_i} 、 δ^{C_i} 为智能体 i 目标网络的神经网络参数; τ_{π} 、 τ_C 为目标网络参数的软更新系数, 远小于其他网络的更新系数, 以实现参数的异步更新。

此外, MADDPG 算法设置了经验池来消除训练数据的关联性, 进一步提高算法性能。算法的详细流程如表 1 的伪代码所示, 表中 T 为训练总次数, N 为智能体总数, M 为采样数量。

表 1 MADDPG 算法伪代码

Table 1 Pseudo-code of MADDPG algorithm

MADDPG 算法
随机生成参数建立初始的 Actor-Critic 网络
复制网络参数建立初始的目标网络
设置噪声
For $t = 1, T$:
初始化设备随机状态, 式(1)一(8)
初始化潮流计算, 式(9)一(12)
For $i = 1, m$:
获取智能体 i 的初始观测值, 式(20)
求取动作值, 式(22)
End for
根据全部动作值 A_t 进行潮流计算获取 S_{t+1} 和 O_{t+1} , 式(9)一(12)
计算智能体的回报值, 式(21)
将 (O_t, A_t, R_t, O_{t+1}) 存入经验池
For $i = 1, n$:
随机从经验池中抽取 M 个样本 $[O_i, A_i, R_i, O_{i+1}]_M$
更新智能体 i 的 Critic 网络的参数, 式(26)一(29)
更新智能体 i 的 Actor 网络的参数, 式(24)、式(25)
软更新 Actor-Critic 网络的目标网络参数, 式(30)、式(31)
更新噪声, 式(23)
End for
End for
保存全部智能体的 Actor 神经网络参数 $[\delta^{\pi_i}]_v$ 用于在线决策

3.2 多智能体的在线分布式决策

多智能体的集中式训练过程已经离线完成了 MADDPG 算法中 Actor 网络参数的迭代, 在线运行时智能体可以直接使用迭代完成的网络参数生成策略函数进行独立决策, 不再需要同其他智能体进行信息交流, 消除了对通信条件的依赖, 决策过程如式(32)所示。

$$Q_{G_i} = \pi_i(O_i | \delta^{\pi_i, \text{off}}), \quad i \in N_{\text{IR}} \quad (32)$$

式中, δ_i^{off} 为离线训练后智能体 i 保存的 Actor 神经网络参数。

4 算例仿真与分析

本节基于改进的 IEEE33 节点系统^[24]进行算例仿真与分析, 系统结构如图 4 所示。系统加入 5 个分布式光伏电源以及 2 个风电机组, 并且仅在可再生能源接入节点设置了量测设备, 智能体仅根据各自接入节点的电气信息对可再生能源设备的无功功率进行控制, 可再生能源设备参数如表 2 所示。

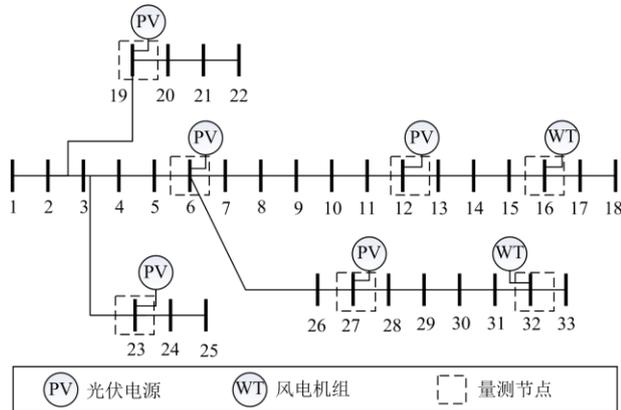


图 4 改进 IEEE33 节点系统结构

Fig. 4 Structure of improved IEEE 33-bus system

表 2 设备参数

Table 2 Equipment parameters

风电参数	值	光伏电源参数	值
最大功率因数	± 0.95	OID 策略最大角	$\pi/6$
额定功率/MW	2	额定功率/MW	1.5
配套无功设备容量/Mvar	0.5	设备容量/MVA	1.8

4.1 智能体训练结果分析

基于第 1 节的电压管控机理模型随机生成 60 天的运行场景, 并采用 MADDPG 算法为可再生能源接入节点训练智能体, 算法所设定的超参数如表 3 所示。

表 3 MADDPG 算法超参数

Table 3 Super parameters of MADDPG algorithm

参数	值	参数	值
τ_π	0.01	M	100
$\tau_{\pi'}$	0.0001	T	5760
τ_c	0.01	ϵ_{svc}	0.01
$\tau_{c'}$	0.0001	ϵ_{ib}	0.05
τ_v	0.0015		

图 5 为训练过程中各智能体的回报值, 由于 6、19、23、27 号节点位于线路首端, 不存在电压越限

现象, 因此具有相同的回报曲线。从回报曲线可以看出, 在前 10 个训练周期内智能体还在进行随机探索, 获取的回报很少。在第 10—第 30 周期, 智能体开始根据探索得到的经验更新自己的策略, 回报曲线快速上升, 只有 12、16、32 号节点偶尔会有轻微的电压越限, 导致回报函数略微降低。而第 30 个周期以后, 各智能体已经具有一致的回报函数, 且回报函数逐渐收敛, 说明各节点已经学习到降低网损的最优策略, 并且不再出现电压越限现象。

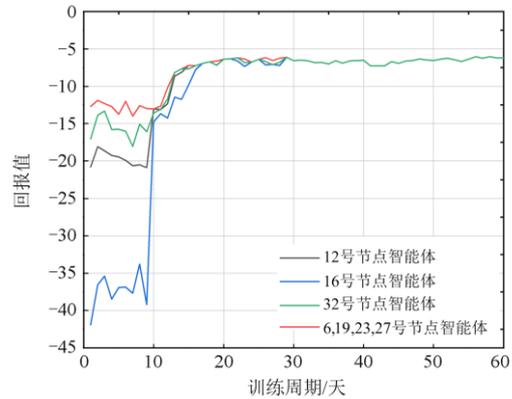


图 5 智能体训练平均回报曲线

Fig. 5 Average reward curve of agent training

随机生成 2000 次电网运行场景对训练完成的智能体的有效性进行验证。在 2000 次场景中, 各节点的波动区间如图 6 所示, 当无电压管控时, 电网电压的波动剧烈, 多个节点存在电压越限现象, 以可再生能源接入节点最为严重。而在智能体进行电压管控后, 电网整体的电压质量得到明显改善, 已经不存在电压越界现象, 各节点电压的波动区间普遍收紧, 最大电压偏差由 0.1 p.u. 降低至本文预设的 0.07 p.u. 以内, 证明了本文所提管控方法在电压管控方面的有效性。值得注意的是, 部分节点出现了最

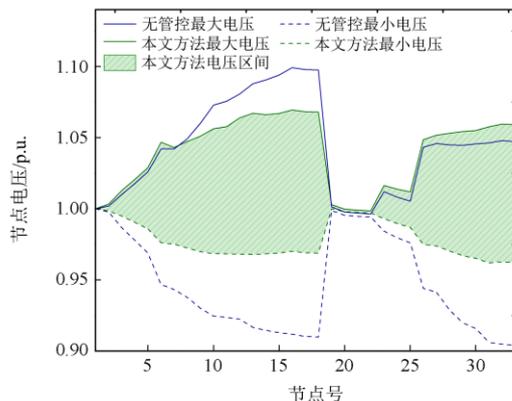


图 6 管控电压波动区间对比

Fig. 6 Comparison of voltage fluctuation range under control

大电压上升的现象,这是因为智能体在控制电压的同时要降低电网损耗,因此在允许的电压波动范围内进行了无功补偿。

4.2 电压管控性能分析

由于本文所提算法考虑的是量测设备不足、缺乏可靠通信条件的区域配电网的无功电压管控问题,在传统的电压控制策略中,只有分散式就地控制策略可以有效运行^[17]。因此,选择 $Q(V)$ 下垂控制策略进行算例对比,验证本文所提方法在控制电压和降低网损方面的有效性。

给定系统一天内的总负荷和可再生能源的总输出功率,如图 7 所示^[25]。选取电压波动最为明显的末端可再生能源接入节点(12、16、32 号节点)进行电压分析,不同控制策略下各节点的电压如图 8 所示。可以看到,两种控制方式下,各节点的电压波动均得到了有效抑制,不再出现电压越限问题,

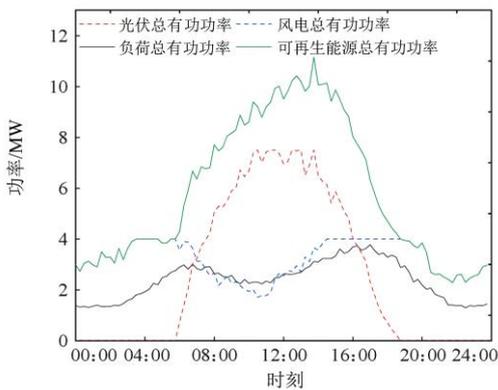


图 7 IEEE33 节点系统的日运行数据

Fig. 7 Daily operation data of IEEE 33-bus system

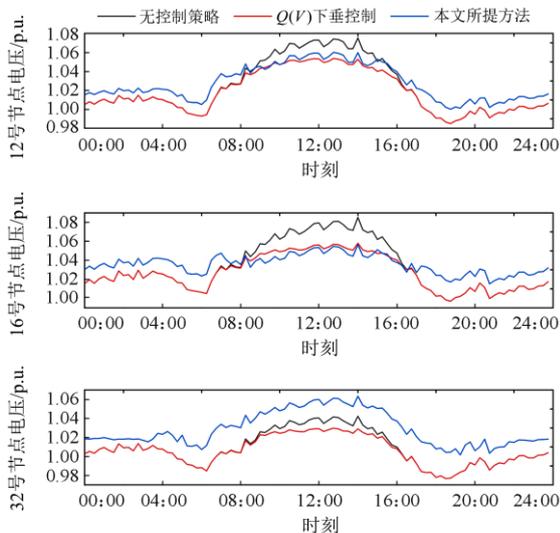


图 8 日运行状态下的电压

Fig. 8 Voltage deviation under daily operation

但是相较于下垂控制,本文所提方法在不同节点会适应性地调整节点所维持的电压等级,这主要是因为本文设定的回报函数下,智能体只有在控制电压水平的同时尽量降低网络损耗,才可以获得最大回报。

图 9 给出了 4 种控制方法一天内的网络损耗,其中,集中式控制使用完备的全局信息,可以认为是最优决策,而 $Q(V)$ 下垂控制和本文所提方法均只使用本地观测信息。从图 9 中可以看出,由于在 08:00 以前和 18:00 以后,光伏电源出力较少,电网电压偏移量不大,在此情况下, $Q(V)$ 下垂控制几乎不动作,电网损耗持平。而本文方法则会在电压允许范围内进行无功补偿以降低电网损耗,同时由于电网的安全压力较小,智能体专注于降低网损,几乎可以达到与集中式管控方法相近的精度。而在 10:00—16:00,光伏出力大幅增加,电网电压偏移量增大, $Q(V)$ 下垂控制为了控制节点电压开始吸收无功功率,但是各节点之间难以有效配合,因此出现了吸收无功功率过多、电网损耗增大的现象。相比之下,本文所提方法中各智能体的策略在训练过程中是协同进化的,各智能体会在保证不出现电压越界的前提下,合作降低网络损耗。但是此时电网的安全压力较大,由于各智能体仅可以使用本地信息进行决策,因此决策动作会较为保守,相较于集中式控制方法网络损耗略微增大。最终,本文管控方式下各可再生能源设备发出的无功功率如图 10 所示。

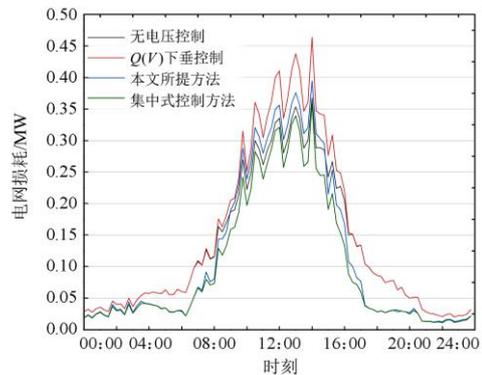


图 9 日运行状态下的网络损耗

Fig. 9 Power loss under daily operation

为了进一步对本文所提方法的设备协同管控性能进行分析,给出了 14:00 时 3 种控制方法下各可再生能源接入节点的无功功率,如表 4 所示。可以看到, $Q(V)$ 下垂控制仅可以根据节点电压偏移量线性地进行无功补偿,而本文所提方法中各节点的无功功率与集中式控制方法具有一致的方向,证明

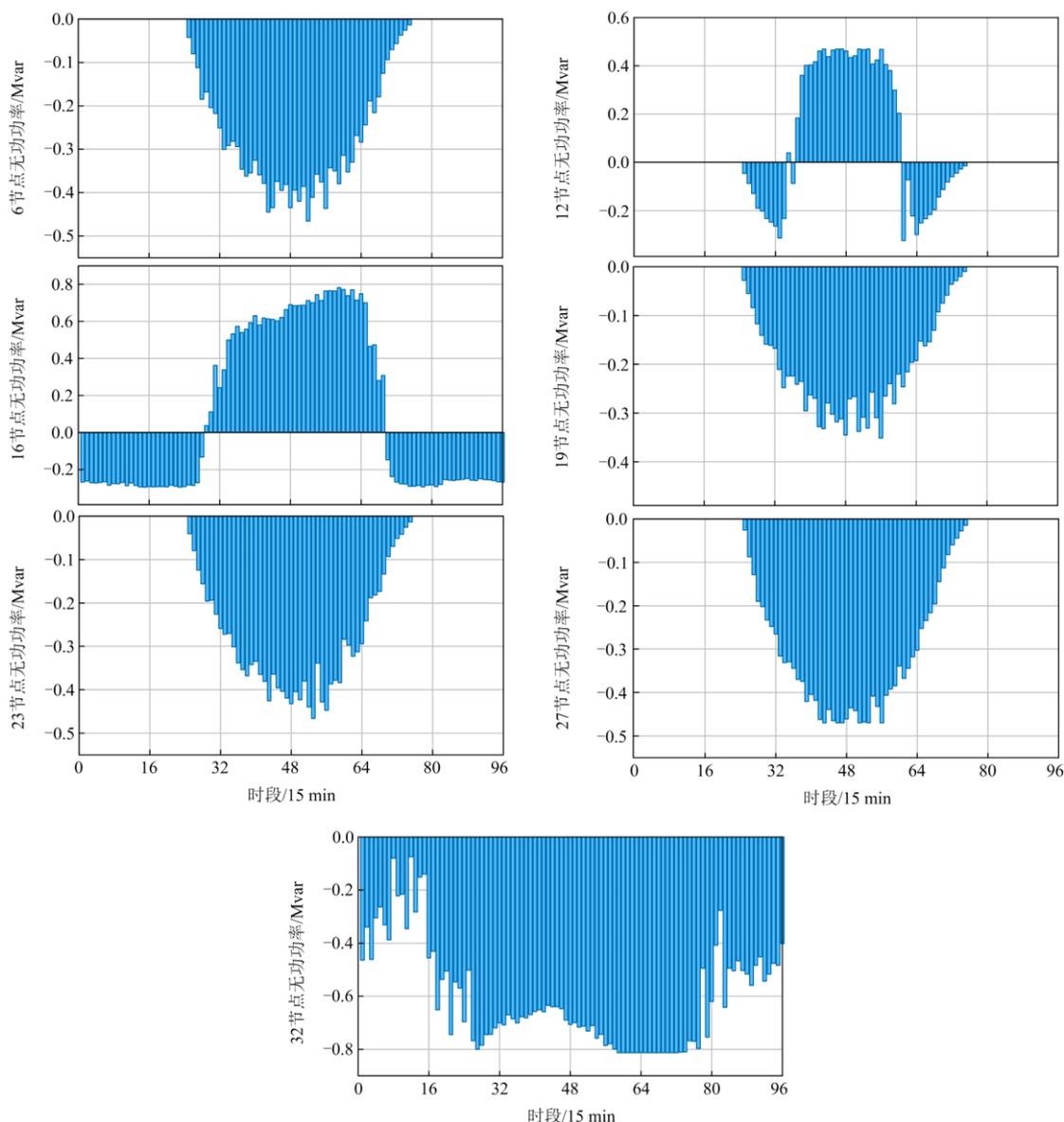


图 10 可再生能源发电的无功功率

Fig. 10 Reactive power of renewable energy power generation

表 4 可再生能源发电无功功率对比

Table 4 Comparison of reactive power of renewable energy power generation

节点 编号	可再生能源发电的无功功率/Mvar		
	$Q(V)$ 下垂控制	本文方法	集中式控制
6	0.1170	-0.4420	-0.4372
12	0.4420	0.3928	0.3412
19	0.000	-0.3203	-0.2463
23	0.000	-0.4420	-0.4417
27	0.1501	-0.4023	-0.3765
16	0.2642	0.6547	0.6167
32	0.0589	-0.7442	-0.7733

各个智能体间可以协同完成降低网络损耗和平抑电压波动的任务, 而不是单纯地降低本地节点的电压偏差量。如 32 号节点的智能体会在安全范围内发出无功功率来降低网络损耗, 而不是根据电压偏移量来吸收无功功率。从数值上来看, 相较于集中式管控, 本文方法的无功补偿功率偏大, 主要是因为智能体在无法获得其他智能体信息的情况下, 会更加倾向于适量地进行过量补偿, 以增加少量的网络损耗为代价, 来保证节点电压不发生越限。

最终, 本文所提方法、无管控方法、 $Q(V)$ 下垂控制和集中式控制一天内的总网络损耗分别为 2.76 MWh、3.09 MWh、3.48 MWh 和 2.54 MWh。

本文方法相较于 $Q(V)$ 下垂控制, 网络损耗下降了 20.7%, 可以证明本文所提方法具有与 $Q(V)$ 下垂控制相同的电压管控性能的同时, 可以大幅降低网络损耗, 具备有效性。同时, 虽然本文方法较集中式控制方法网络损耗略有增加, 但是本文方法本质上是一种就地控制策略, 不需要对通信设备和计算中心进行额外投资, 因此具有一定的现实意义。

5 结论

本文提出了一种基于多智能体深度确定性梯度算法的多智能体协同电压管控方法, 实现了在量测数据不足的情况下不依赖通信的多智能体就地分散式无功电压管控。通过算例分析, 证明了本文所提方法中智能体可以根据可再生能源接入节点的电气信息独立进行无功补偿功率决策, 有效地抑制电压波动, 防止节点电压越限。并与就地控制和集中式控制方法进行了性能对比, 本文方法相较于 $Q(V)$ 下垂控制, 可以大幅降低电网的网络损耗, 具有优越性。相较于集中式控制方法, 虽然本文方法的网损略有上升, 但是集中式控制方法需要完备的全局信息, 在量测数据不足的情况下无法求解, 因此本文方法在缺少在线监测设备和可靠通信条件的区域配电网中具有实用性。如何进一步提升智能体的管控精度, 将是本文后续研究的重点。

参考文献

- [1] 马国真, 林毓军, 张泽亚, 等. 计及源荷多重不确定性的综合能源系统鲁棒经济调度方法[J]. 电力系统保护与控制, 2021, 49(20): 43-52.
MA Guozhen, LIN Yujun, ZHANG Zeya, et al. A robust economic dispatch method for an integrated energy system considering multiple uncertainties of source and load[J]. Power System Protection and Control, 2021, 49(20): 43-52.
- [2] 李政洁, 撒奥洋, 周生奇, 等. 计及综合需求响应的综合能源系统优化调度[J]. 电力系统保护与控制, 2021, 49(21): 36-42.
LI Zhengjie, HAN Aoyang, ZHOU Shengqi, et al. Optimization of an integrated energy system considering integrated demand response[J]. Power System Protection and Control, 2021, 49(21): 36-42.
- [3] 闫梦阳, 李华强, 王俊翔, 等. 计及综合需求响应不确定性的园区综合能源系统优化运行模型[J]. 电力系统保护与控制, 2022, 50(2): 163-175.
YAN Mengyang, LI Huaqiang, WANG Junxiang, et al. Optimal operation model of a park integrated energy system considering uncertainty of integrated demand response[J]. Power System Protection and Control, 2022, 50(2): 163-175.
- [4] 郑能, 丁晓群, 管志成, 等. 基于场景法的配电网有功-无功协调优化[J]. 电网技术, 2019, 43(5): 1640-1651.
ZHENG Neng, DING Xiaqun, GUAN Zhicheng, et al. Active reactive power coordination optimization of distribution network based on scenario method[J]. Power System Technology, 2019, 43(5): 1640-1651.
- [5] 周晓华, 张银, 刘胜永, 等. 静止无功补偿器新型自适应动态规划电压控制[J]. 电力系统保护与控制, 2018, 46(12): 77-84.
ZHOU Xiaohua, ZHANG Yin, LIU Shengyong, et al. Voltage control of static var compensator based on novel adaptive dynamic programming[J]. Power System Protection and Control, 2018, 46(12): 77-84.
- [6] 郭清元, 吴杰康, 莫超, 等. 基于混合整数二阶锥规划的新能源配电网电压无功协同优化模型[J]. 中国电机工程学报, 2018, 38(5): 1385-1396.
GUO Qingyuan, WU Jiekang, MO Chao, et al. A model for multi-objective coordination optimization of voltage and reactive power in distribution networks based on mixed integer second-order cone programming[J]. Proceedings of the CSEE, 2018, 38(5): 1385-1396.
- [7] 卢姬, 常俊晓, 张云阁, 等. 考虑 DG 不确定性的主动配电网两阶段无功机会约束优化方法[J]. 电力系统保护与控制, 2021, 49(21): 28-35.
LU Ji, CHANG Junxiao, ZHANG Yunge, et al. Two stage reactive power chance constrained optimization method for active distribution network considering DG uncertainty[J]. Power System Protection and Control, 2021, 49(21): 28-35.
- [8] YOSHIDA H, KAWATA K. A particle swarm optimization for reactive power and voltage control considering voltage security assessment[J]. IEEE Transactions on Power Systems, 2000, 15(4): 1232-1239.
- [9] DARATHA N, DAS B, SHARMA J. Robust voltage regulation in unbalanced radial distribution system under uncertainty of distributed generation and loads[J]. International Journal of Electrical Power & Energy Systems, 2015, 73: 516-527.
- [10] VACCARO A, ZOBAA A F. Voltage regulation in active networks by distributed and cooperative meta-heuristic optimizers[J]. Electric Power Systems Research, 2013, 99(1): 9-17.
- [11] 王庆斌, 刘涌, 杨响, 等. 含分布式电源的 10 kV 配电网无功电压控制方法[J]. 电力电容器与无功补偿, 2018, 39(2): 125-130.
WANG Qingbin, LIU Yong, YANG Yun, et al. Reactive

- power voltage control method for 10 kV distribution network with distributed generation[J]. *Power Capacitor & Reactive Power Compensation*, 2018, 39(2): 125-130.
- [12] 梁俊文, 林舜江, 刘明波. 主动配电网分布式无功优化控制方法[J]. *电网技术*, 2018, 42(1): 230-237.
LIANG Junwen, LIN Shunjiang, LIU Mingbo. A method for distributed optimal reactive power control of active distribution network[J]. *Power System Technology*, 2018, 42(1): 230-237.
- [13] 蔡永翔, 唐巍, 徐鸥洋, 等. 含高比例户用光伏的低压配电网电压控制研究综述[J]. *电网技术*, 2018, 42(1): 220-229.
CAI Yongxiang, TANG Wei, XU Ouyang, et al. Review of voltage control research in LV distribution network with high proportion of residential PVs[J]. *Power System Technology*, 2018, 42(1): 220-229.
- [14] DUAN J, SHI D, DIAO R, et al. Deep reinforcement learning based autonomous voltage control for power grid operations[J]. *IEEE Transactions on Power Systems*, 2019, 35(1): 814-817.
- [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] ZHANG X, LIU Y, DUAN J, et al. DDPG-based multi-agent framework for SVC tuning in urban power grid with renewable energy resources[J]. *IEEE Transactions on Power Systems*, 2021, 36(6): 5465-5475.
- [17] 张宇精, 乔颖, 鲁宗相, 等. 含高比例分布式电源接入的低感知度配电网电压控制方法[J]. *电网技术*, 2019, 43(5): 1528-1535.
ZHANG Yujing, QIAO Ying, LU Zongxiang, et al. Voltage control for partially visible distribution networks with high DG penetration[J]. *Power System Technology*, 2019, 43(5): 1528-1535.
- [18] HETZER J, YU D C, BHATTARAI K. An economic dispatch model incorporating wind power[J]. *IEEE Transactions on Energy Conversion*, 2008, 23(2): 603-611.
- [19] 周瑜. 计及风电场无功支撑能力的电网无功优化调度研究[D]. 济南: 山东大学, 2021.
- ZHOU Yu. Research on optimal reactive power dispatching of power grid considering the reactive power support capacity of wind farm[D]. Jinan: Shandong University, 2021.
- [20] ZHANG S, CHENG H, LI K, et al. Multi-objective distributed generation planning in distribution network considering correlations among uncertainties[J]. *Applied Energy*, 2018, 226: 743-755.
- [21] DALL' ANESE E, DHOPLÉ S V, JOHNSON B B, et al. Optimal dispatch of residential photovoltaic inverters under forecasting uncertainties[J]. *IEEE Journal of Photovoltaics*, 2017, 5(1): 350-359.
- [22] ZHANG S, CHENG H, ZHANG L, et al. Probabilistic evaluation of available load supply capability for distribution system[J]. *IEEE Transactions on Power Systems*, 2013, 28(3): 3215-3225.
- [23] DESAI J P, MAKWANA V H. A novel out of step relaying algorithm based on wavelet transform and a deep learning machine model[J]. *Protection and Control of Modern Power Systems*, 2021, 6(4): 500-511.
- [24] BARAN M E, WU F F. Network reconfiguration in distribution systems for loss reduction and load balancing[J]. *IEEE Transactions on Power Delivery*, 1989, 4(2): 1401-1407.
- [25] 柴园园, 郭力, 王成山, 等. 含高渗透率光伏的配电网分布式电压控制[J]. *电网技术*, 2018, 42(3): 738-746.
CHAI Yuanyuan, GUO Li, WANG Chengshan, et al. Distributed voltage control in distribution networks with high penetration of PV[J]. *Power System Technology*, 2018, 42(3): 738-746.

收稿日期: 2022-02-12; 修回日期: 2022-03-31

作者简介:

徐博涵(1996—), 男, 硕士研究生, 主要研究方向为电力系统智能调度; E-mail: bohaxu_scu@163.com

向月(1987—), 男, 通信作者, 副教授, 博士, 主要研究方向为电力系统智能调度等。E-mail: xiang@scu.edu.cn

(编辑 许威)