

# 基于深度强化学习的有源配电网电压分层控制策略

杜婉琳<sup>1</sup>, 王玲<sup>1</sup>, 罗威<sup>2</sup>, 朱远哲<sup>1</sup>, 吕鸿<sup>1</sup>, 马潇男<sup>3</sup>, 周霞<sup>3\*</sup>

(1. 广东电网有限责任公司电能质量重点实验室(广东电网有限责任公司电力科学研究院), 广东省 广州市 510080; 2. 广东电网有限责任公司梅州供电局, 广东省 梅州市 514021; 3. 南京邮电大学自动化学院、人工智能学院, 江苏省 南京市 210023)

## Voltage Hierarchical Control Strategy of Active Distribution Network Based on Deep Reinforcement Learning

DU Wanlin<sup>1</sup>, WANG Ling<sup>1</sup>, LUO Wei<sup>2</sup>, ZHU Yuanzhe<sup>1</sup>, LÜ Hong<sup>1</sup>, MA Xiaonan<sup>3</sup>, ZHOU Xia<sup>3\*</sup>

(1. Key Laboratory of Power Quality of Guangdong Power Grid Co., Ltd. (Electric Power Research Institute of Guangdong Power Grid Co., Ltd.), Guangzhou 510080, Guangdong Province, China; 2. Meizhou Power Supply Bureau of Guangdong Power Grid Co., Ltd., Meizhou 514021, Guangdong Province, China; 3. College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, Jiangsu Province, China)

**摘要:**【目的】分布式电源发电的随机性和波动性, 给有源配电网(active distribution network, ADN)的电压控制带来了严峻的挑战, 在此背景下, 亟需一种高效的电压控制策略来保证ADN的安全运行。【方法】基于深度强化学习方法, 提出了一种双层区域配电网电压控制策略。首先, 以调压设备的调节特性和可控元素复杂化的特点为前提, 针对ADN辐射网架结构, 设计了区域协调控制区域和本地自治控制区域, 分别构建每个区域的电压控制模型; 然后, 通过深度Q网络(deep Q-network, DQN)算法和深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法对该模型进行求解, 以实现实时跟踪电压变化的目的, 有效解决了ADN运行过程中电压控制问题; 最后, 通过IEEE 33节点仿真算例对该方法进行了验证。【结果】利用DQN算法和DDPG算法分别求解协调控制区域和本地自治区域的控制变量, 实现了ADN系统电压调节的实时决策, 解决了ADN潮流双向流动、电压复杂多变的问题。【结论】所提控制策略控制电压偏差效果明显, 具有很强的准效性和实用性。

**关键词:** 有源配电网 (ADN); 区域协调控制; 本地自治控制; 深度强化学习; 电压控制策略

**ABSTRACT:** [Objectives] The randomness and volatility

基金项目: 国家自然科学基金项目(52207009); 南方电网公司科技项目(GDKJXM20200331)。

Project Supported by National Natural Science Foundation of China (52207009); Science and Technology Projects of China Southern Power Grid Corporation(GDKJXM20200331).

of distributed power generation poses significant challenges for the voltage control in active distribution network (AND). In this context, there is an urgent need for an efficient voltage control strategy to ensure the safe operation of ADN. [Methods] Based on the deep reinforcement learning method, a voltage control strategy for double-layer regional distribution networks was proposed. First, based on the adjustment characteristics of voltage regulating equipment and the complexity of controllable elements, a regional coordinated control area and a local autonomous control area were designed for the radiating grid structure of the ADN, and the voltage control model of each area was constructed. Then, the model was solved by deep Q-Network (DQN) algorithm and deep deterministic policy gradient (DDPG) algorithm to achieve the purpose of tracking voltage changes in real time, and effectively solve the voltage control problem during the operation of the ADN. Finally, the method was verified by IEEE 33-bus simulation examples. [Results] The DQN algorithm and the DDPG algorithm were used to solve the control variables in the coordinated control region and the local autonomous region respectively, realizing real-time decision-making of voltage regulation in the ADN system, and solving the problems of bidirectional flow of ADN power flow and complex and changeable voltage. [Conclusions] The proposed control strategy has obvious effect on controlling voltage deviation, and has strong accuracy and practicality.

**KEY WORDS:** active distribution network (ADN); regional coordination control; local autonomous control; deep reinforcement learning; voltage control strategy

## 0 引言

与传统的被动配电网不同,有源配电网(active distribution network, ADN)可以实现电网的双向流动和信息的互联互通<sup>[1]</sup>,同时,ADN还能提高分布式电源的利用效率和经济性,促进可再生能源开发,符合国家能源转型的发展方向<sup>[2-7]</sup>。但是,该方法也对电网的电压控制提出了更高的要求<sup>[8]</sup>,因为不恰当的电压水平会导致设备发生过电流,进而引起用户设备的损坏<sup>[9-10]</sup>。所以,应当将有效控制ADN的电压波动作为后续研究的重点之一<sup>[11-12]</sup>。

ADN电压波动控制问题较为复杂,是拥有众多目标函数、变量和约束条件的高维非线性问题<sup>[8]</sup>。目前,高维非线性问题的解决思路主要采用启发式搜索算法<sup>[13-15]</sup>和二阶锥松弛技术<sup>[16-17]</sup>。文献[18-19]采用启发式搜索算法中的粒子群算法和遗传算法对配电网进行协调控制,以改善配电网运行过程中的电压偏移现象,但是,该方法计算量大,时间成本较高,需要耗费大量的计算资源;文献[20-21]利用二阶锥技术求解可控设备模型来实现电压偏差最小化,但是,此类技术依赖于源荷预测数据和精确的电力系统优化模型,难以实现配电网电压实时调控的要求。

人工智能技术有望解决先前诸多方法的局限性,所以,诸多学者探讨了基于新一代人工智能技术<sup>[22]</sup>的配电网电压波动调控方法<sup>[23-24]</sup>。其中,强化学习(reinforcement learning, RL)方法作为人工智能技术的重要分支被广泛应用。文献[25-26]研究了单层尺度RL方法在配电网中的应用问题,该方法可以实时提供灵活的控制决策,针对电压波动问题控制效果良好。但是,文献[27]的研究发现,单层尺度RL容易受到环境变化的影响,出现过拟合现象。针对此问题,文献[28-29]基于长时间尺度RL开展了研究,在配电网电压控制方面具有灵活性、自适应性。然而,上述方法均未考虑控制变量和AND拓扑结构的特点,所以存在模型不确定性,导致算法的控制效果不稳定。

因此,基于上述电压波动控制技术中所存在的问题,本文提出了基于深度强化学习的有源配

电网电压分层控制策略。分层建立有功-无功协调优化模型后,采用深度强化学习算法求解;此外,考虑控制变量和AND拓扑结构的类型,提出利用深度Q网络(deep Q-network, DQN)算法求解区域协调控制的离散型设备变量,利用深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法求解本地自治区域的连续型设备变量,以达到算法和物理模型有机整合。最后,基于IEEE 33节点划分多层次电压控制区域,验证了本文所提出的算法的有效性。

## 1 电压分层协调控制总框架

### 1.1 控制指标

将配电网的电压可以分为3个不同的状态区间:正常状态区间、预警状态区间和紧急状态区间,如图1所示。

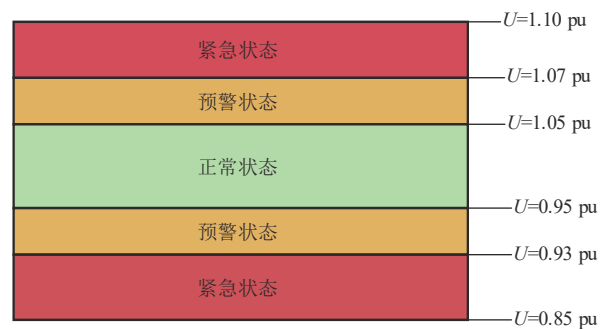


图1 配电网电压分布状态图

Fig. 1 Voltage distribution state diagram of distribution network

#### 1) 正常状态

安全阈值设置为[0.95 pu, 1.05 pu]。

#### 2) 紧急状态

紧急状态表达式为

$$[f_{VSmin,i}, f_{VSmax,i}] = [1.07 \text{ pu}, 1.10 \text{ pu}] \quad (1)$$

$$[f_{VXmin,i}, f_{VXmax,i}] = [0.85 \text{ pu}, 0.93 \text{ pu}] \quad (2)$$

式中:  $f_{VSmin,i}$  和  $f_{VSmax,i}$  分别为第  $i$  个区域上紧急状态下限值和上限值;  $f_{VXmin,i}$  和  $f_{VXmax,i}$  分别为第  $i$  个区域下紧急状态下限值和上限值。

#### 3) 预警状态

预警状态表达式为

$$[f_{YVSmin,i}, f_{YVSmax,i}] = [1.05 \text{ pu}, 1.07 \text{ pu}] \quad (3)$$

$$[f_{YVXmin,i}, f_{YVXmax,i}] = [0.93 \text{ pu}, 0.95 \text{ pu}] \quad (4)$$

式中： $f_{YVSmin,i}$ 和 $f_{YVSmax,i}$ 分别为第 $i$ 个区域上预警状态下限值和上限值； $f_{YVXmax,i}$ 和 $f_{YVXmin,i}$ 分别为第 $i$ 个区域下预警状态上限值和下限值。

## 1.2 控制框架

本文核心是分区对不同类型的控制设备进行控制，图2为配电网分层控制方法的结构图。

主动配电网电压分层协调控制系统由主动配电网管理系统(active distribution management system, ADMS)、区域协调控制器和本地自治控制器3部分组成。控制器被安装在每个区域中，并且各个负荷节点处的电压传感器与控制器通信网络相邻布置。ADMS通过采集到的节点电压数据进行计算，当判断电压偏差平均标么值处于紧急状态时，将信号传给区域协调控制器，采用DQN进行训练，得出调压器分接头的最佳档位位置；若区域协调控制器没有收到信号或者通过区域协调控制器控制以后，电压仍然处于非正常状态，则通过本地

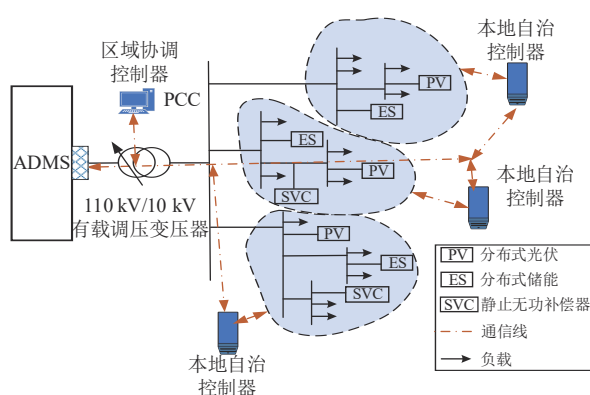


图2 配电网分层控制方法结构图

Fig. 2 Structure diagram of hierarchical control method of distribution network

自治控制器对光伏逆变器、静态无功补偿装置(static var compensator, SVC)和储能进行调节，采用DDPG算法进行训练，获得最优控制策略后将控制信号下达到各个调压设备，实现电压实时自治控制，整体控制流程如图3所示。

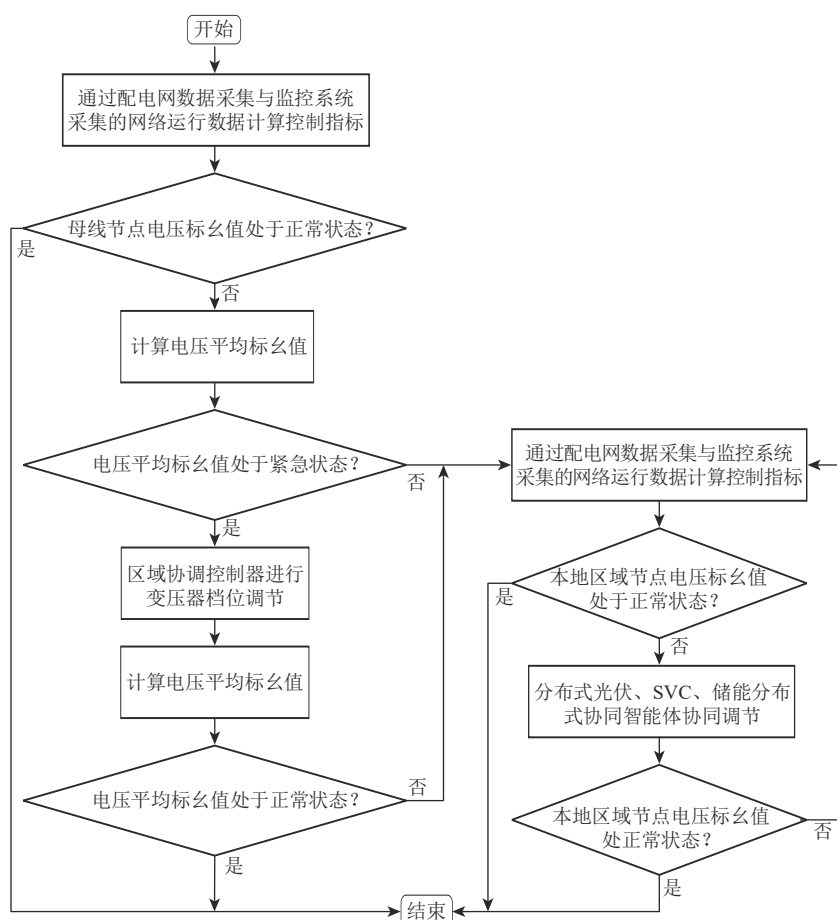


图3 整体控制流程图

Fig. 3 Overall control flow chart

## 2 深度强化学习算法

### 2.1 DQN算法

DQN算法通过使用神经网络来估计当前状态下每个行动的 $Q$ 值,输入是一个状态 $s$ ,输出是一个向量,其中每个元素表示一个可能的动作 $a$ 的 $Q$ 值:

$$Q(s, a) \approx Q(s, a; \theta) \quad (5)$$

式中: $\theta$ 为神经网络的权重参数; $Q(s, a)$ 为真实的 $Q$ 值; $Q(s, a; \theta)$ 为神经网络估计的 $Q$ 值。

目标是最小化 $Q$ 值的平方误差损失,即:

$$L(\theta) = E \{ [r(t) + \gamma \max_{a'} Q(s', a'; \theta_{\text{target}}) - Q(s, a; \theta)]^2 \} \quad (6)$$

式中: $r(t)$ 为当前状态和行动的奖励; $\gamma$ 为折扣因子; $s'$ 为下一个状态; $a'$ 为下一个状态下所有可能行动的最大 $Q$ 值对应的行动; $\theta_{\text{target}}$ 为目标网络的权重参数。

DQN算法使用随机梯度下降来更新神经网络的权重参数 $\theta$ 。可以计算损失函数 $L(\theta)$ 的梯度:

$$\begin{aligned} \nabla \theta L(\theta) &= r_{\text{target}} + \gamma \max_{a'} Q\left(\frac{s', a'}{\theta_{\text{target}}}\right) - \\ &Q\left(\frac{s, a}{\theta}\right) \times \nabla \theta Q\left(\frac{s, a}{\theta}\right) \end{aligned} \quad (7)$$

式中: $s'_i$ 为第 $i$ 组数据的下一个状态; $a'_i$ 为第 $i$ 组数据下一个状态下所有可能行动的最大 $Q$ 值对应的行动; $r_i$ 为第 $i$ 组数据下状态和行动的奖励。

根据梯度下降算法,更新神经网络的权重参数 $\theta$ ,即

$$\theta_{i+1} = \theta_i - \alpha \nabla \theta L(\theta) \quad (8)$$

式中 $\alpha$ 为学习率。

DQN算法使用 $\epsilon$ -greedy策略来选择行动,该策略以 $\epsilon$ 的概率选择随机行动。其中, $\epsilon$ 为探索率,通常在训练早期设置较高的值,然后逐渐降低到一个较小的值。

### 2.2 深度确定性策略梯度算法

DDPG算法将经验数据 $(s_t, a_t, r_t, s_{t+1})$ 存储到经验池中,通过同时建立Actor目标网络和Critic目标网络来输出目标 $Q$ 值,并通过最小化目标 $Q$ 值和Critic网络输出 $Q$ 值的差值完成模型的优化训练。

使用高斯分布表示输出的概率分布,即

$$a = \mu(s; \theta^\mu) + \epsilon(1) \quad (9)$$

式中: $a$ 为输出的动作; $\mu$ 为策略网络; $\epsilon(1)$ 为高

斯噪声; $\theta^\mu$ 为策略网络的权重参数。

使用策略梯度算法来训练策略网络,即最大化长期回报的期望值 $J(\theta^\mu)$ :

$$\nabla J(\theta^\mu) = E_{\pi} [\nabla_a Q(s, a | \theta^Q) |_{s=s_t, a=\mu(s_t)} \nabla \theta^\mu \mu(s | \theta^\mu)] \quad (10)$$

式中: $E_{\pi}$ 表示在策略 $\pi$ 下的期望; $Q(s, a | \theta^Q)$ 是值函数网络,即状态-动作值函数; $\theta^Q$ 为Critic目标网络的权重参数公式; $\nabla_a Q(s, a | \theta^Q)$ 表示状态-动作值函数相对于动作 $a$ 的梯度。

值函数网络是一个连续函数,使用均方误差(mean squared error, MSE)作为损失函数,即

$$L(\theta^Q) = E \{ [y - Q(s, a | \theta^Q)]^2 \} \quad (11)$$

式中 $y = r + \gamma Q[s', \mu(s' | \theta^\mu) | \theta^Q]$ 表示目标值。

使用梯度下降算法来更新权重参数 $\theta^Q$ ,即

$$\theta^Q(t+1) = \theta^Q(t) - \alpha \nabla \theta L[\theta^Q(t)] \quad (12)$$

式中 $\nabla \theta L[\theta^Q(t)]$ 是损失函数相对于 $\theta^Q(t)$ 权重参数的梯度。

## 3 基于深度强化学习算法的分层调压控制模型

### 3.1 电压控制数学模型

在调节区域协调控制节点电压时,将控制的目标设置为:最小全域配电网母线节点电压越限值平均值,表达式为

$$\min F(x) = \begin{cases} \bar{U} - 1.05 \bar{U}_N, & \bar{U} \geq 1.05 \bar{U}_N \\ 0.95 \bar{U}_N - \bar{U}, & \bar{U} \leq 0.95 \bar{U}_N \end{cases} \quad (13)$$

式中: $\bar{U}$ 为全域配电网母线节点电压平均值; $\bar{U}_N$ 为配电网平均电压额定值。

在调节本地自治区域部分节点电压时,将控制目标设置为最小自治区域母线节点电压越限值,表达式为

$$\min F(x) = \begin{cases} \sum_{i=1}^M (U_i - 1.05 U_N), & U_i \geq 1.05 U_N \\ \sum_{i=1}^M (0.95 U_N - U_i), & U_i \leq 0.95 U_N \end{cases} \quad (14)$$

式中: $U_i$ 为第 $i$ 个节点的节点电压; $U_N$ 为该区域配电网额定电压值; $M$ 为该区域配电网母线节点数。

约束条件如下。

1) 潮流平衡约束

潮流平衡约束为



$$\begin{cases} \sum_{i=1}^M P_{i,L}(t) + P_{\text{Loss}}(t) = P_M(t) + P_{i,\text{PV}}(t) + P_{i,\text{ES}}(t) \\ \sum_{i=1}^M Q_{i,L}(t) + Q_{\text{Loss}}(t) = Q_M(t) + Q_{i,\text{PV}}(t) + Q_{i,\text{SVC}}(t) \end{cases} \quad (15)$$

式中： $P_{i,L}(t)$ 和 $Q_{i,L}(t)$ 分别为在 $t$ 时刻节点 $i$ 上负荷消耗的有功和无功功率； $P_{\text{Loss}}(t)$ 和 $Q_{\text{Loss}}(t)$ 分别为 $t$ 时刻线路中的有功损耗和无功损耗； $P_M(t)$ 和 $Q_M(t)$ 分别为 $t$ 时刻从主配网上传输的有功和无功功率； $P_{i,\text{PV}}(t)$ 和 $Q_{i,\text{PV}}(t)$ 分别为 $t$ 时刻分布式光伏的输出有功和无功功率； $P_{i,\text{ES}}(t)$ 为 $t$ 时刻节点 $i$ 上储能输出有功功率； $Q_{i,\text{SVC}}(t)$ 为 $t$ 时刻节点 $i$ 上SVC输出无功功率。

## 2) 光伏逆变器约束

光伏逆变器约束为

$$P_{\text{PV}}^{\min} \leq P_{i,\text{PV}}(t) \leq P_{\text{PV}}^{\max} \quad (16)$$

$$Q_{i,\text{PV}}^2(t) \leq S_{\text{inv}}^2 - P_{i,\text{PV}}^2(t) \quad (17)$$

式中： $P_{\text{PV}}^{\max}$ 、 $P_{\text{PV}}^{\min}$ 分别为光伏逆变器输出有功的上限和下限值； $S_{\text{inv}}^2$ 表示逆变器的容量。

## 3) 储能出力约束

储能出力约束为

$$P_{i,\min}^{\text{ES}}(t) \leq P_{i,\text{ES}}(t) \leq P_{i,\max}^{\text{ES}}(t) \quad (18)$$

$$\Delta P_{i,\min}^{\text{ES}}(t) \leq \Delta P_{i,\text{ES}}^{\text{ES}}(t) \leq \Delta P_{i,\max}^{\text{ES}}(t) \quad (19)$$

式中： $\Delta P_{i,\text{ES}}^{\text{ES}}(t)$ 为 $t$ 时刻节点 $i$ 上储能输出有功变化量； $P_{i,\min}^{\text{ES}}(t)$ 和 $P_{i,\max}^{\text{ES}}(t)$ 分别为 $t$ 时刻节点 $i$ 上储能输出有功的最小值和最大值； $\Delta P_{i,\min}^{\text{ES}}(t)$ 和 $\Delta P_{i,\max}^{\text{ES}}(t)$ 分别为 $t$ 时刻节点 $i$ 上储能输出有功最小和最大变化量。

## 4) SVC出力约束

SVC出力约束为

$$Q_{i,\min}^{\text{SVC}}(t) \leq Q_{i,\text{SVC}}(t) \leq Q_{i,\max}^{\text{SVC}}(t) \quad (20)$$

$$\Delta Q_{i,\min}^{\text{SVC}}(t) \leq \Delta Q_{i,\text{SVC}}^{\text{SVC}}(t) \leq \Delta Q_{i,\max}^{\text{SVC}}(t) \quad (21)$$

式中： $\Delta Q_{i,\text{SVC}}^{\text{SVC}}(t)$ 为节点 $i$ 上SVC输出无功变化量； $Q_{i,\max}^{\text{SVC}}(t)$ 和 $Q_{i,\min}^{\text{SVC}}(t)$ 分别为节点 $i$ 上SVC输出无功上下限； $\Delta Q_{i,\max}^{\text{SVC}}(t)$ 和 $\Delta Q_{i,\min}^{\text{SVC}}(t)$ 分别为 $t$ 时刻节点 $i$ 上SVC输出无功爬坡上下限。

## 5) 有载调压变压器约束

有载调压变压器分接头调节范围的计算公式如下：

$$n_{\text{next}} \in [n_{\text{tpmin}}, n_{\text{tpmax}}] = [n_{\text{tpnow}} x_{\text{promin}}, n_{\text{tpnow}} x_{\text{promax}}] \quad (22)$$

$$[x_{\text{promin}}, x_{\text{promax}}] = \bigcap_{i=1,2,\dots,n} \left[ \frac{f_{\text{VSmin},i} / f_{\text{VXmin},i}}{U_{\text{eq},i}}, \frac{f_{\text{VSmax},i} / f_{\text{VXmax},i}}{U_{\text{eq},i}} \right] \quad (23)$$

式中： $n_{\text{tpmin}}$ 为有载变压器分接头最小调节值； $n_{\text{tpmax}}$ 为有载变压器分接头最大调节值； $x_{\text{promin}}$ 为有载变压器最小可调比例； $x_{\text{promax}}$ 为有载变压器最大可调比例； $n_{\text{tpnow}}$ 为有载变压器分接头当前所处位置； $n$ 为区域数量； $U_{\text{eq},i}$ 为等效电压。

## 3.2 深度强化学习算法设计

将深度强化学习应用于配电网的电压控制中，将电压控制问题转化为马尔科夫决策过程，并通过对智能体在环境中进行探索和学习来获得最优的控制行为。该算法将配电网潮流系统状态映射到控制动作，同时利用奖励函数和控制变量映射来实现控制目标，在满足约束条件的前提下确保配电网系统的安全运行。

### 3.2.1 区域协调控制的DQN算法

#### 1) 状态空间

将母线节点电压作为被控制对象，则状态空间为节点电压的集合 $S_{\text{DQN}}$ ：

$$S_{\text{DQN}} = \{v_1, v_2, \dots, v_i, \dots, v_N\} \quad (24)$$

式中： $v_i$ 表示节点 $i$ 电压的标么值； $N$ 为全域配电网母线节点数。

#### 2) 动作空间

将有载变压器分接头的位置设置为DQN算法的动作空间 $A_{\text{DQN}}$ ，假设每个设备有 $n$ 个档位，设置相同的调节范围： $i \times 1\% \text{pu}$ ， $-n \leq i \leq n$ ；档位调节的总范围 $\pm n \times 1\% \text{pu}$ ，即

$$A_{\text{DQN}} = \{-n \times 1\% \text{pu}, -(n-1) \times 1\% \text{pu}, \dots, 0\% \text{pu}, \dots, (n-1) \times 1\% \text{pu}, n \times 1\% \text{pu}\} \quad (25)$$

#### 3) 奖励函数

将点电压越限量作为控制目标，奖励函数设计为 $r_{\text{DQN},i}$ ：

$$r_{\text{DQN},i} = -\beta \Delta \bar{v}^2 \quad (26)$$

式中： $\beta$ 为权重系数； $\Delta \bar{v}$ 为全局配电网母线节点电压的平均标么值超过正常状态的数值。

### 3.2.2 本地自治控制的DDPG算法

#### 1) DDPG算法的状态空间

将各节点电压、有功功率和无功功率波动作为状态变量，构建配电网电压控制的状态空间的：

$$S_{DDPG,i}=\{v_1,\cdots,v_i,\cdots,v_M,dv_1,\cdots,dv_i,\cdots,dv_M,p_1,\cdots,\\p_i,\cdots,p_M,dp_1,\cdots,dp_i,\cdots,dp_M,q_1,\cdots,q_i,\cdots,\\q_M,dq_1,\cdots,dq_i,\cdots,dq_M\}$$
 (27)

式中： $dv_i$ 为节点*i*电压的波动； $p_i$ 为节点*i*的有功功率； $dp_i$ 为节点*i*有功功率波动； $q_i$ 为节点*i*的无功功率； $dq_i$ 为节点*i*无功功率波动； $1\leq i\leq M$ 。

2) DDPG 算法的动作空间

将并入的设备动作出力集合 $\{A_{DDPG,i}\}$ 可以定义为动作空间：

$$\{A_{DDPG,i}\}=[P_i^{PV}(t), Q_i^{PV}(t), P_i^{ES}(t), Q_i^{SVC}(t)]$$
 (28)

3) DDPG 算法的奖励函数

将节点电压越限量和设备出力调节量作为奖励函数 $R_{DDPG,i}$ 来更新 DDPG 算法的策略网络和值函数网络，即

$$R_{DDPG,i}=-\Delta v_i\cdot B_1\cdot\Delta v_i^T-\Delta P_i^{PV}\cdot C_1\cdot\Delta P_i^{PV^T}-\Delta Q_i^{PV}\cdot C_2\cdot\\ \Delta Q_i^{PV^T}-\Delta P_i^{ES}\cdot D_1\cdot\Delta P_i^{ES^T}-\Delta Q_i^{SVC}\cdot E_1\cdot\Delta Q_i^{SVC^T}$$
 (29)

式中： $\Delta v_i$ 为母线节点*i*的电压标么值的越限值； $\Delta Q_i^{PV}$ 表示母线节点*i*的无功变化量； $B_1$ 表示电压偏差项的惩罚系数； $C_1$ 和 $C_2$ 表示分布式光伏输出有功、无功调节量项的惩罚系数； $D_1$ 表示储能输出有功调节量项的惩罚系数； $E_1$ 表示 SVC 输出无功调节量项的惩罚系数。

4 仿真实验

4.1 仿真分析

采用标准 IEEE 33 节点配电网系统进行仿真，拓扑图如图 4 所示。系统中调压器位于节点 1，额定容量为 100 MV·A，调节范围根据具体情况确定，该算例具体包括分布式光伏、储能电池、SVC，设备参数如表 1 所示。

通过配电网控制区域划分方法，将配电网按图 4 中的划分结果进行划分，结果如表 2 所示。

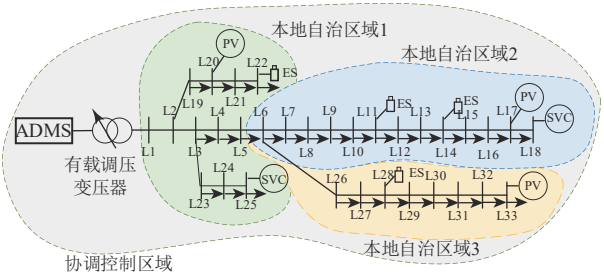


图4 IEEE 33节点配电网拓扑图

Fig. 4 IEEE 33-node distribution network topology

表1 设备参数

Tab. 1 Equipment parameters

序号	布置节点	类型	额定容量
1	L17、L19、L20、L33	分布式光伏	2.2 MW
2	L11、L14、L22、L28	储能电池	800 kW·h
3	L18、L25	SVC	2 MV·A

表2 控制区域划分结果

Tab. 2 Control zone division results

协调控制区域 节点编号	自治控制区域 编号	所含节点编号	控制节点编号
L1—L33	自治控制区域1	L1—L5、L19— L2、L3—L25	L20、L22、L25
	自治控制区域2	L6—L18	L11、L14、L17、 L18
	自治控制区域3	L26—L33	L28、L33

4.2 结果分析

IEEE 33 节点配电网在某一时刻引入功率扰动，部分节点越过正常状态区域限值，首先根据有载调压器的调节范围对 DQN 算法进行设计。

使用 DQN 智能体进行训练，每次训练 290 个样本，共进行 500 次训练。初始阶段，智能体奖励值较低，经过 230 次训练 episode 后，DQN 智能体逐渐收敛于最优策略，其奖励值稳定且震荡幅度减小。最终，DQN 智能体学习到了调节调压器分接头对电压控制的最优策略，具有良好的控制性能。DQN 智能体训练结果如图 5 所示。

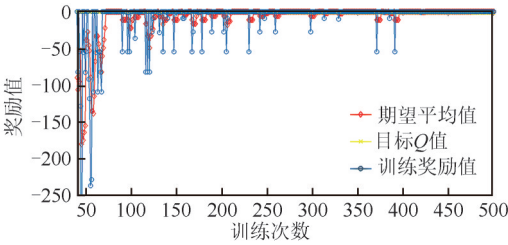


图5 DQN智能体训练过程

Fig. 5 DQN agent training process

图 6 为区域协调控制前后节点电压图。由此可见，各母线电压幅值有所下降，但仍然有部分母线节点电压标么值处于非正常状态安全阈值之间，因此，需要通过 DDPG 算法对电压越限节点所在区域做进一步的处理。DDPG 智能体训练结果如图 7 所示。

DDPG 智能体训练 1 000 个 episodes，每个 episode 在训练 300 个样本后结束，经过 70 个

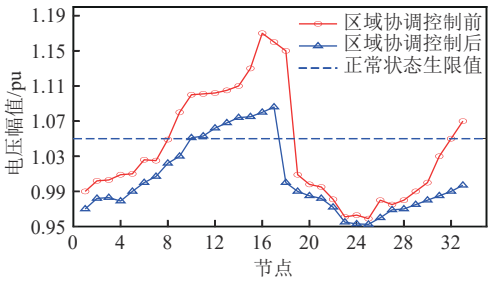


图6 区域协调控制前后节点电压图

Fig. 6 Node voltage diagram before and after regional coordination control

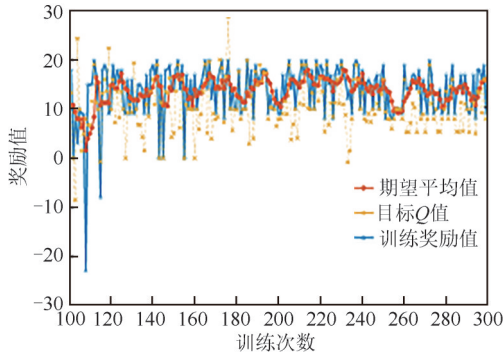


图7 DDPG智能体训练结果

Fig. 7 DDPG agent training results

episode后DDPG智能体的奖励值稳定，表明算法已收敛。图8为未削减光伏有功情况下节点10—17的控制效果。可见，控制区域内的节点电压并未控制到正常状态内，因此需要进一步削减有功功率。图9为通过光伏有功削减后节点10—17节点的电压，可快速有效地将母线节点电压控制在安全阈值内。

为了验证本文方法控制的优越性，将本文方法与传统的麻雀搜索(sparrow search algorithm,

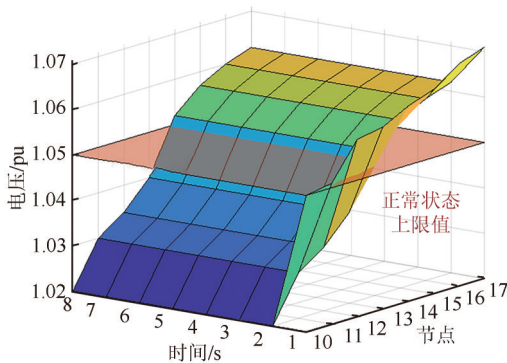


图8 无有功削减情况下的电压控制效果

Fig. 8 Voltage control effect without active power reduction

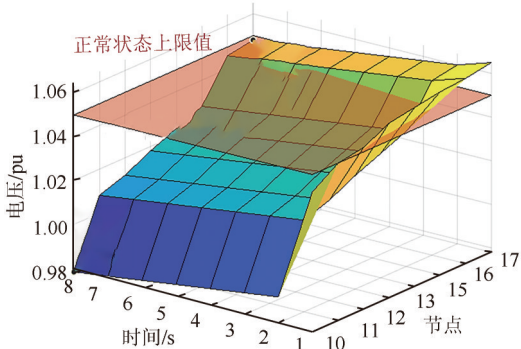


图9 加入有功削减情况下电压控制效果

Fig. 9 Voltage control effect with active power reduction

SSA) 算法和粒子群优化 (particle swarm optimization, POS)算法进行对比，不同策略下的控制结果如表3所示。SSA在控制调节上响应时间为5.52 s，平均电压偏差0.002 9 pu，POS在控制调节上响应时间为4.84 s，平均电压偏差0.003 1 pu。而采用强化学习控制方法进行调节时响应时间为0.11 s，平均电压偏差0.002 3 pu，可以看出，本文控制策略虽然与其他算法在平均电压偏差上相差不大，但是大幅度缩短了算法响应的的时间，充分显示了该强化学习算法的高效性和准确性。

表3 不同策略下的控制结果

Tab. 3 Control results under different strategies

策略	优化前	强化学习算法	SSA	POS
响应时间/s	0	0.11	5.52	4.84
平均电压偏差/pu	0.024 0	0.002 3	0.002 9	0.003 1

5 结论

针对有源配电网内不同时间响应的调压设备特性，提出一种基于深度强化学习的有源配电网电压分层控制策略，具体结论如下：

1) 针对有源配电网电压特性，将有源配电网电压等级划分为3个不同的区域，有助于有源配电网系统的实时监测和管理，便于后续开展有源配电网电压控制。

2) 针对不同控制对象的响应特点，利用DQN算法和DDPG算法分别求解协调控制区域和本地自治区域的控制变量，实现了ADN系统电压调节的实时决策。

3) 经IEEE 33节点实验结果表明, 所提方法可以有效保证电压稳定在正常状态之内; 与SSA算法和POS算法相比, 所提方法在响应时间和响应效果方面均有更大的优势。

## 参考文献

- [1] 江道灼, 徐宁, 江崇熙, 等. 蜂巢状有源配电网构想、关键技术与展望[J]. 电力系统自动化, 2019, 43(17): 1-11.  
JIANG D Z, XU N, JIANG C X, et al. Conception, key technology and prospect of honeycomb-shape active distribution network[J]. Automation of Electric Power Systems, 2019, 43(17): 1-11.
- [2] 刘建伟, 李学斌, 刘晓鸥. 有源配电网中分布式电源接入与储能配置[J]. 发电技术, 2022, 43(3): 476-484.  
LIU J W, LI X B, LIU X O. Distributed power access and energy storage configuration in active distribution network[J]. Power Generation Technology, 2022, 43(3): 476-484.
- [3] 于淼, 闫旻睿, 万克斤, 等. 数据驱动的有源配电网运行态势智能感知方法[J]. 电力建设, 2024, 45(7): 34-53.  
YU M, YAN M R, WAN K T, et al. Data-driven intelligent situational awareness of active distribution networks[J]. Electric Power Construction, 2024, 45(7): 34-53.
- [4] 蔡欢, 袁旭峰, 熊炜, 等. 柔性互联配电网运行调度研究综述[J]. 智慧电力, 2022, 50(6): 92-99.  
CAI H, YUAN X F, XIONG W, et al. Review on operation scheduling of flexible interconnected distribution network[J]. Smart Power, 2022, 50(6): 92-99.
- [5] 闫群民, 董新洲, 穆佳豪, 等. 基于改进多目标粒子群算法的有源配电网储能优化配置[J]. 电力系统保护与控制, 2022, 50(10): 11-19.  
YAN Q M, DONG X Z, MU J H, et al. Optimal configuration of energy storage in an active distribution network based on improved multi-objective particle swarm optimization[J]. Power System Protection and Control, 2022, 50(10): 11-19.
- [6] 仲泽天, 李梦月, 王加澍, 等. 一种有源配电网分布式光伏消纳能力评估方法[J]. 电网与清洁能源, 2023, 39(2): 60-68.  
ZHONG Z T, LI M Y, WANG J S, et al. An assessment method for distributed photovoltaic absorption capacity of active distribution networks[J]. Power System and Clean Energy, 2023, 39(2): 60-68.
- [7] LOPION P, MARKEWITZ P, ROBINIUS M, et al. A review of current challenges and trends in energy systems modeling[J]. Renewable and Sustainable Energy Reviews, 2018, 96: 156-166.
- [8] 彭生江, 陆军, 张中丹, 等. 光伏接入对电网的影响研究[J]. 电力系统保护与控制, 2021, 49(5): 157-164.  
PENG S J, LU J, ZHANG Z D, et al. Research on the influence of photovoltaic access on a power grid[J]. Power System Protection and Control, 2021, 49(5): 157-164.
- [9] 段穰达. 有源配电网后评价指标体系及其综合评价方法[J]. 发电技术, 2021, 42(1): 86-93.  
DUAN R D. A post-evaluation index system of active distribution network project and its comprehensive evaluation method[J]. Power Generation Technology, 2021, 42(1): 86-93.
- [10] 张华赢, 艾精文, 汪伟. 基于约束型深度强化学习的主动配电网电压控制策略[J/OL]. 电测与仪表, 2022, <https://kns.cnki.net/kcms/detail/23.1202.TH.20220708.1459.010.html>.  
ZHANG H Y, AI J W, WANG W. Active distribution network voltage control strategy based on constrained deep reinforcement learning[J/OL]. Journal of Electrical Measurement and Instrumentation, 2022, <https://kns.cnki.net/kcms/detail/23.1202.TH.20220708.1459.010.html>
- [11] WECKX S, DRIESEN J. Optimal local reactive power control by PV inverters[J]. IEEE Transactions on Sustainable Energy, 2016, 7(4): 1624-1633.
- [12] 符杨, 廖剑波, 李振坤, 等. 考虑越限风险的主动配电网日前优化调度及运行[J]. 中国电机工程学报, 2017, 37(21): 6328-6338.  
FU Y, LIAO J B, LI Z K, et al. Day-ahead optimal scheduling and operating of active distribution network considering violation risk[J]. Proceedings of the CSEE, 2017, 37(21): 6328-6338.
- [13] 李鹏, 姜磊, 王加浩, 等. 基于深度强化学习的新能源配电网双时间尺度无功电压优化[J/OL]. 中国电机工程报: 1-15. <http://kns.cnki.net/kcms/detail/11.2107.TM.20220826.1757.024.html>.  
LI P, JIANG L, WANG J H, et al. Optimization of dual-time scale reactive voltage for distribution network with renewable energy based on deep reinforcement learning[J/OL]. Proceedings of the CSEE, 1-15. <http://kns.cnki.net/kcms/detail/11.2107.TM.20220826.1757.024.html>.
- [14] 蒋平, 梁乐. 基于内点法和遗传算法相结合的交直流系统无功优化[J]. 高电压技术, 2015, 41(3): 724-729.  
JIANG P, LIANG L. Reactive power optimization of hybrid AC/HVDC power system combining interior



- point algorithm and genetic algorithm[J]. High Voltage Engineering, 2015, 41(3): 724-729.
- [15] 李晓利, 高金峰. 用于配电网多目标无功优化的改进粒子群优化算法[J]. 电力自动化设备, 2019, 39(1): 106-111.
- LI X L, GAO J F. Improved particle swarm optimization algorithm for multi-objective reactive power optimization of distribution network[J]. Electric Power Automation Equipment, 2019, 39(1): 106-111.
- [16] 林少华, 吴杰康, 莫超, 等. 基于二阶锥规划的含分布式电源配电网动态无功分区与优化方法[J]. 电网技术, 2018, 42(1): 238-246.
- LIN S H, WU J K, MO C, et al. Dynamic partition and optimization method for reactive power of distribution networks with distributed generation based on second-order cone programming[J]. Power System Technology, 2018, 42(1): 238-246.
- [17] 肖遥, 别朝红, 黄格超, 等. 基于二阶锥最优潮流的机会约束分布鲁棒优化研究[J]. 电网技术, 2021, 45(4): 1505-1517.
- XIAO Y, BIE Z H, HUANG G C, et al. Chance-constrained distributional robust optimization based on second-order cone optimal power flow[J]. Power System Technology, 2021, 45(4): 1505-1517.
- [18] 李国庆, 陈蓉, 沈杰, 等. 大规模光伏接入的配电网动态无功优化研究[J]. 太阳能学报, 2018, 39(8): 2373-2380.
- LI G Q, CHEN R, SHEN J, et al. Research of dynamic reactive power optimization of large-scale photovoltaic access distribution network[J]. Acta Energetica Sinica, 2018, 39(8): 2373-2380.
- [19] 刘斌, 程杰, 肖荣, 等. 面向中低压一体化的配网无功优化方法[J]. 电气自动化, 2019, 41(6): 45-47.
- LIU B, CHENG J, XIAO R, et al. Reactive power optimization method for integrated medium and low voltage distribution networks[J]. Electrical Automation, 2019, 41(6): 45-47.
- [20] 褚国伟, 张友旺, 葛乐, 等. 自储能柔性互联配电网多时间尺度电压优化[J]. 电力系统自动化, 2021, 45(9): 71-79.
- CHU G W, ZHANG Y W, GE L, et al. Multi-time-scale voltage optimization of flexible interconnected distribution network with self-energy storage[J]. Automation of Electric Power Systems, 2021, 45(9): 71-79.
- [21] 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 Q Y, WU J K, MO C, 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.
- [22] 于成龙, 侯俊杰, 蒲洪波, 等. 新一代人工智能在国防科技领域发展探讨[J]. 国防科技, 2020, 41(4): 13-18.
- YU C L, HOU J J, PU H B, et al. Development of the new generation artificial intelligence in national defense science and technology[J]. National Defense Technology, 2020, 41(4): 13-18.
- [23] 蒲天骄, 乔骥, 韩笑, 等. 人工智能技术在电力设备运维检修中的研究及应用[J]. 高电压技术, 2020, 46(2): 369-383.
- PU T J, QIAO J, HAN X, et al. Research and application of artificial intelligence in operation and maintenance for power equipment[J]. High Voltage Engineering, 2020, 46(2): 369-383.
- [24] YAP K Y, SARIMUTHU C R, LIM J M Y. Artificial intelligence based MPPT techniques for solar power system: a review[J]. Journal of Modern Power Systems and Clean Energy, 2020, 8(6): 1043-1059.
- [25] 李琦, 乔颖, 张宇精. 配电网持续无功优化的深度强化学习方法[J]. 电网技术, 2020, 44(4): 1473-1480.
- LI Q, QIAO Y, ZHANG Y J. Continuous reactive power optimization of distribution network using deep reinforcement learning[J]. Power System Technology, 2020, 44(4): 1473-1480.
- [26] WANG Y. Grid voltage control method based on generator reactive power regulation using reinforcement learning[C]// 2020 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia). IEEE, 2020: 1060-1065.
- [27] 毕聪博, 唐聿劼, 罗永红, 等. 电力系统优化控制中强化学习方法应用及挑战[J]. 中国电机工程学报, 2024, 44(1): 1-21.
- BI C B, TANG Y J, LUO Y H, et al. Review on critical problems in reinforcement learning methods applied in power system optimization and control scenarios[J]. Proceedings of the CSEE, 2024, 44(1): 1-21.
- [28] 倪爽, 崔承刚, 杨宁, 等. 基于深度强化学习的配电网多时间尺度在线无功优化[J]. 电力系统自动化, 2021, 45(10): 77-85.
- NI S, CUI C G, YANG N, et al. Multi-time-scale online optimization for reactive power of distribution network based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2021, 45(10): 77-85.
- [29] 高一航, 徐晓春, 李鹏, 等. 面向高比例新能源并网的双时间尺度电网电压控制[J]. 电网与清洁能源, 2021, 37(5): 115-122.

GAO Y H, XU X C, LI P, et al. Dual-time-scale grid voltage control for high-proportion new energy grid-connected[J]. Power System and Clean Energy, 2021, 37(5): 115-122.

收稿日期: 2023-08-26。

修回日期: 2023-10-18。

作者简介:



杜婉琳

杜婉琳(1985), 女, 硕士, 高级工程师, 电力系统分析技术, 2788829207@qq.com;



马潇男

马潇男(1999), 男, 硕士研究生, 主要研究方向为配电网运行技术、电压管理, 344717084@qq.com;



周霞

周霞(1978), 女, 博士, 副教授, 从事电力通信、电力系统分析与控制研究, 本文通信作者, zhouxia@njupt.edu.cn。

(责任编辑 辛培裕)