

基于DDPG的新型配电网调频功率分配仿真与控制研究

李明¹, 魏承志¹, 郭小易¹, 赵瑞峰², 卢建刚², 陈益哲³, 高宜凡³, 甘锴³

(1. 南方电网科学研究院有限公司, 广东 广州 510630; 2. 广东电网有限责任公司 电力调度控制中心, 广东 广州 510062; 3. 广东电网有限责任公司 肇庆供电局, 广东 肇庆 526020)

摘要: 新型电力系统中, 新能源的发电方式与运行特性显著区别于传统同步机, 频率安全稳定运行面临新的挑战 and 机遇。当电网出现功率扰动引发的频率波动时, 为了调节系统频率, 需要及时地调整机组出力, 调频机组功率分配的合理性对能源利用效率和电网稳定性具有重要意义。因此, 基于深度强化学习中的深度确定性策略梯度算法 (DDPG), 以区域系统的频率偏差和区域控制误差设计奖励函数, 设置多个场景训练智能体, 按照扰动大小赋予场景权重, 输出通用策略集的加权功率分配因子, 最终实现机组调频功率的合理分配。建立基于 Python 调用 PSS/E 的联合仿真平台, 以北欧地区 44 节点 IEEE 标准电力系统模型为仿真算例, 利用联合仿真平台验证了所提优化策略的可行性与优越性。

关键词: 新型电力系统; 频率稳定; 调频功率分配; 深度确定性策略梯度算法

中图分类号: TM73 **文献标识码:** A **DOI:** 10.19457/j.1001-2095.dqed26523

Research on Simulation and Control of Frequency Regulation Power Distribution in New Distribution Network Based on DDPG

LI Ming¹, WEI Chengzhi¹, GUO Xiaoyi¹, ZHAO Ruifeng², LU Jiangang², CHEN Yizhe³, GAO Yifan³, GAN Kai³

(1. Electric Power Research Institute, China Southern Power Grid Co., Ltd., Guangzhou 510630, Guangdong, China; 2. Power Dispatching and Control Center, Guangdong Power Grid Co., Ltd., Guangzhou 510062, Guangdong, China; 3. Zhaoqing Power Supply Bureau, Guangdong Power Grid Co., Ltd., Zhaoqing 526020, Guangdong, China)

Abstract: In the new power system, the generation method and operating characteristics of new energy are significantly different from traditional synchronous machines, and the safe and stable operation of frequency faces new challenges and opportunities. When there is a large frequency fluctuation caused by high-power disturbances in the power grid, in order to adjust the system frequency, it is necessary to manually adjust the output of the units in a timely manner. The rationality of power allocation of frequency modulation units is of great significance to energy utilization efficiency and power grid stability. Based on the deep deterministic policy gradient (DDPG) algorithm in deep reinforcement learning, a reward function was designed based on the frequency deviation and regional control error of the regional system. Multiple scenarios were set to train the intelligent agent, and scene weights were assigned according to the disturbance size. The weighted power allocation factor of the general policy set was output, ultimately achieving a reasonable allocation of unit frequency modulation power. A joint simulation platform was established based on Python calling PSS/E, and the 44 node IEEE standard power system model in the Nordic region was used as a simulation example to verify the feasibility of using the joint simulation platform to generate optimization strategy results.

Key words: new power system; frequency stabilization; frequency modulation power distribution; depth deterministic policy gradient (DDPG) algorithm

基金项目: 国家自然科学基金(U22A20221); 南方电网重点科技项目 031200KK52222027(GDKJXM20222224)

作者简介: 李明(1977—), 男, 硕士, 教授级高级工程师, 主要研究方向为高压直流输电、交直流控制保护与自动化系统,

Email: liming@csg.cn

“十四五”时期,我国将加快推进以新能源为主体的新型电力系统建设,加快构建以新能源为主体的新型电力系统,是推动我国能源生产和消费革命、构建清洁低碳安全高效能源体系的重要手段^[1]。新型电力系统在规划、运行和控保方面与传统电力系统存在的差异,也使得电力系统已存在的问题更加凸显,频率安全问题就是其中的典型代表^[2]。而且,具有极强的间歇性和不确定性的规模新能源接入将给电网带来强烈的随机扰动,使得频率控制变得复杂且困难^[3]。自动发电控制(automatic generation control, AGC)技术可以解决这一问题,电网的频率控制可以分为一次、二次和三次,而二次调频则是通过发电机的调频器来调整,实现无差调节,即通过对机组的有功输出进行调节,实现对电力系统的频率和互联电网间交流功率的动态稳定控制,对保证新型电力系统的安全可靠运行具有重要作用^[4]。

近年来,AGC一直是国内外学术界和工业界的研究热点,伴随电力系统发展演变诞生了大量学术创新和成果^[5]。文献[6]剖析了传统的以PI算法为主的集中式AGC,但由于集中式AGC需要精准调参或详细的系统模型,在新型电力系统中实现难度显著提高;文献[7]对分布式AGC进行了探索,提出一种面向新能源并网的协同控制强化学习算法获取分布式多区域电网的协同控制;文献[8-10]研究光伏、风力发电机组和电池储能系统等快速调频机组参与AGC调频服务,助力调频厂完成上级调度的指令,提高电力系统在大规模新能源发电接入下的稳定性和安全性;文献[11]针对异质调频机组之间的协同配合问题,提出了一种分布式协同AGC方法,有效地提高了系统的调频性能,且能够在设计的时间内实现期望的有功功率分配;文献[12]针对实际电网中发生过多次AGC振荡问题,提出了一种AGC振荡模式辨识方法,可用于AGC振荡的风险评估;文献[13]针对传统AGC控制策略调频容量不足、无法充分利用互联网网各类调频资源抑制频率快速波动的问题,提出了兼顾电网调频性能、煤耗量和SOC恢复能力的火储联合AGC控制策略,使互联电网煤耗量和调频性能都得到了优化;文献[14]提出了一种级联PD-模糊PID控制器,并采用灰狼优化算法等技术对该控制器的缩放参数进行了优化,解决了传统混合电源统一电力系统自动发电控制问题,可为新型电力系统的AGC提供

借鉴;文献[15]针对互联新型电力系统经济运行问题,提出了一种新的AGC优化调配方法—混沌遗传算法,有效降低了AGC机组的能耗费用。综上可知,已有文献虽然已对AGC策略进行了很大的优化,但仍存在部分缺陷,如集中式AGC仍无法保证各个区域之间的协同一致性,储能资源参与AGC的经济性和补偿机制有待深入研究,新兴的以智能优化算法为主的分布式AGC还不具备强大的自动寻优能力等^[16-17]。

对于具有大量随机性、波动性、间歇性新能源的新型电力系统频率稳定控制,采用人工智能是一个有效手段,可以实现电力系统控制的自动化、智能化和高效化^[18-19]。其中,强化学习是人工智能的一个重要分支,其在参数设置方面具有简单性、灵活性及普适性,可利用其强大的感知和逻辑推演能力来解决电网调控问题,成为电网新兴的研究热点和应用方向,间接推动了强化学习在AGC方面的快速发展^[20-21]。经典强化学习中应用最普遍的是Q学习及其衍生算法,由于其具有更好的动态寻优能力,故与智能优化算法相比,在电网稳定控制领域的应用具有更高的效率与准确性^[22]。然而,Q学习体系算法存在收敛性较差问题,而且目前鲜有将其应用到新型电力系统稳定控制的相关研究报道^[23-24]。

鉴于此,本文从自动发电控制角度出发,针对新型电力系统频率稳定控制问题,提出将适应能力更好、收敛速度更快的深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法应用于AGC的机组功率分配与控制中,有助于更准确、高效、实时地提高新型电力系统频率的稳定性。本文基于网络学习率、梯度更新策略以及目标网络的软更新方式,设置多个场景为DDPG训练学习以及按照扰动大小权衡不同场景的权重,合理设计奖励函数引导智能体进行优化分配;最后,建立了Python调用PSS/E的联合仿真平台,验证了利用联合仿真平台生成优化策略结果的可行性和优越性。

1 深度确定性策略梯度算法

深度确定性策略梯度算法是深度强化学习的方法之一,算法引入双网络结构,降低了神经网络更新过程中的扰动,提高了算法的稳定性。算法通过策略梯度算法,能够根据学习经验选择最优或较优的策略,并采用经验回放机制,打乱

历史数据间的相关性,提高学习效率,加快收敛速度。

1.1 DDPG的网络结构—Actor-Critic网络

DDPG算法的实现框架如图1所示,其结构形式为Actor-Critic网络,分为Actor策略网络和Critic价值网络,二者之间通过交互对自身网络参数进行更新。这一双重网络之间的激励性策略可在某种程度上改善学习策略的质量与效率。

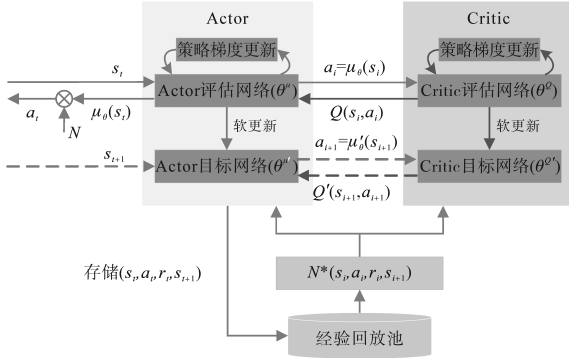


图1 DDPG算法框架

Fig.1 DDPG algorithm framework

两个网络的主要功能如下:Actor网络的作用是选择智能体与环境交互的动作,根据当前状态 s_t 获取确定性动作函数 $\mu_\theta(s)$,并基于概率分布结果,输出所有动作的执行概率,最终选出动作 a_t ;而Critic网络用于评估Actor部分所选动作的优劣,在当前状态 s_t 下,执行动作 a_t 时,用价值函数 $Q_\theta(s, a_t)$ 来评估动作的优劣。

借鉴深度强化学习网络(deep Q-network, DQN)的思想,Actor策略网络和Critic价值网络各自再分为评估网络和目标网络,即有Actor评估神经网络 $\mu_\theta(s)$,Critic评估神经网络 $Q(s_t, a_t | \theta^Q)$ 、以及Actor目标神经网络 $\mu'_{\theta}(s)$,Critic目标神经网络 $Q'(s_t, a_t | \theta^{Q'})$,共4个网络。

Critic评估网络通过神经网络拟合状态动作价值函数,输入当前状态和动作,输出对应的 Q 值;Critic目标网络的输入为Actor目标网络生成的未来动作,输出为下一状态动作的 Q 值评估。Critic评估网络 $Q(s_t, a_t | \theta^Q)$ 可以利用最小化损失进行更新,损失函数如下:

$$L(\theta^Q) = \text{MSE}[Q(s_t, a_t | \theta^Q), y_t] \\ = \frac{1}{N} \sum_{i=1}^N [y_i - Q(s_t, a_t | \theta^Q)]^2 \quad (1)$$

其中 $y_t = r_t(s_t, a_t) + \gamma Q'[s_{t+1}, \mu(s_{t+1}) | \theta^{Q'}]$ (2)
式中: y_t 为以目标网络拟合的下一状态的 Q 值和当前状态的奖励得到的 Q 值目标; $r_t(s_t, a_t)$ 为智能

体从环境转换中获得的奖励值; γ 为折扣系数。

Actor网络则将状态与动作之间的关系确定化,即根据系统目前的状况,给出一种确定化的动作策略,DDPG算法通常利用下式确定性动作函数表示动作选择:

$$a_t = \mu(s_t, \theta^\mu) \quad (3)$$

另外,为了给动作的选择带来随机性,DDPG还可在探索过程中引入随机噪声,从而提高探索环境的效率,如下式所示:

$$a_t = \mu(s_t, \theta^\mu) + K \quad (4)$$

式中: K 为随机噪声参数。

Actor网络用梯度上升法引导网络参数更新以得到更高的评价,如下式所示:

$$\nabla_{\theta^\mu} J(\theta^\mu) \approx \frac{1}{N} \sum_i [\nabla_a Q(s_i, a_i | \theta^Q) \cdot \nabla_{\theta^\mu} \mu(s_i | \theta^\mu)] \quad (5)$$

式中: ∇ 为梯度计算; J 为策略 μ 的预期长期回报目标函数; N 为样本大小。

为了减少更新目标扰动带来的神经网络更新困难,两个网络各自的评估网络与目标网络之间通过软更新方式传递参数,以便平缓地更新目标网络^[4],其更新公式如下:

$$\theta^{\mu'} \leftarrow \alpha \theta^\mu + (1 - \alpha) \theta^{\mu'} \quad (6)$$

$$\theta^{Q'} \leftarrow \alpha \theta^Q + (1 - \alpha) \theta^{Q'} \quad (7)$$

式中: α 为滑动平均系数,一般取0.01保证更新平缓; θ^μ, θ^Q 分别为Actor网络和Critic网络的参数; $\theta^{\mu'}, \theta^{Q'}$ 分别为Actor目标网络和Critic目标网络的参数。

1.2 经验回放机制

DDPG算法采用了和DQN算法一样的经验回放机制。如图1所示,经验回放池是用来存储智能体在探索环境中获得的经验数据,并从这些数据中提取出一些有用的信息,用来训练神经网络的参数。由于采样的随机性,样本间原有的相关性将被打破。在此基础上,DDPG通过经验池能够同时学习当前和过去的经验数据,并且由于在学习过程中加入以往的经验数据,使得神经网络的学习效率有了很大的提高。因此,经验池的创建成功解决了数据相关性的问题,每一次训练过程中,智能体和环境互动得到的状态转移数据(当前状态、动作、奖励、下一个状态)都会被存储到经验回放池中,并且后续会按照随机抽样的方式提取这些数据用于训练,打乱数据之间的相关性,最终提高智能体学习的稳定性^[25]。

2 基于DDPG算法的调频机组功率分配策略设计

2.1 DDPG算法的调频机组功率分配原理

二次调频作为系统频率控制的核心环节,需要在多台调频机组间实现最优功率分配,以达到频率恢复与系统稳定的双重目标。这一功率分配问题本质上是一个高维、非线性、时变的连续决策过程,其过程具有连续状态和动作空间的特性。DDPG算法基于Actor-Critic框架,能够直接输出确定性连续动作,避免离散化误差,并利用深度神经网络通过数据驱动的方式处理高维、非线性的系统状态,学习高维非线性控制策略,其特性与二次调频功率分配问题高度匹配,因此,本文提出一种基于DDPG算法的功率分配策略。

基于DDPG算法的调频机组功率分配原理简述如下:首先对互联电网的各个区域进行实时监控,将每个区域的区域控制误差 ACE 和频率偏差 $|\Delta f|$ 的实时数据作为DDPG算法的环境状态,并依据设计好的奖励函数评估动作的优劣程度,将当前环境信息和对应的奖励输入到DDPG智能体中,智能体提前经过多重场景的训练迭代,使用多次训练收敛的策略网络对新的 ACE 和 $|\Delta f|$ 状态进行预测,最终向环境输出通用策略集的功率分配因子。DDPG向环境输出的通用策略集的策略权重为一组连续的动作空间,其意义为该区域电网调度端在面对新的负荷扰动场景时,以此计算各调频机组的功率分配因子,每台机组的调整功率由总的机组调频功率与对应分配因子之积决定。

DDPG为连续动作空间算法,在学习控制过程中输出的是一组连续的动作,结合调频控制系统对功率分配的精度要求,将动作设置为 n 台调频机组对应的 n 个功率分配系数。调频机组总的调控功率同每台调频机组对应功率分配系数的积即是每台调频机组的所需调整的功率,如下式所示:

$$\Delta P_j = a_j \sum_{j=1}^n \Delta P_j \quad (8)$$

式中: ΔP_j 为第 j 台机组的调频功率; a_j 为第 j 台机组的分配因子。

同时,为达到有功功率平衡的要求,各机组的功率分配系数应符合如下约束:

$$\sum_{j=1}^n a_j = 1 \quad 0 \leq a_j \leq 1 \quad (9)$$

由于本文采用联络线功率频率偏差控制方式,故将 ACE 和 Δf 纳入环境状态,以反映电网当前的状态,如下式所示:

$$s = [|\Delta f_i|, ACE_i] \quad (10)$$

$$ACE_i = (NI_{Ai} - NI_{Si}) + 10 \cdot B \cdot \Delta f_i \quad (11)$$

式中: Δf_i 为区域 i 的频率偏差; ACE_i 为区域 i 的区域控制误差; NI_{Ai} 为区域 i 的所有联络线实际功率之和; NI_{Si} 为区域 i 的所有联络线计划功率之和。

2.2 奖励函数的选择

基于以环境状态反映动作好坏的算法思维,为了使互联系统实现互补响应调整,本文采用联络线功率频率偏差控制方式,将系统频率偏差的瞬时绝对值 $|\Delta f(i)|$ 和区域控制误差 ACE 的瞬时值 $ACE(i)$ 的平方进行线性加权和作为奖励的目标函数,并对 Δf 和 ACE 的量纲进行标准化处理,则第 m 次训练时的奖励函数如下:

$$R(m) = -x|\Delta f(m)|^2 - (1-x)[ACE(m)]^2/1000 \quad (12)$$

式中: $x, 1-x$ 分别为 $|\Delta f(m)|$ 和 $ACE(m)$ 的权重系数,此处设置 $x=0.5$ 。

2.3 DDPG算法的调频机组功率分配策略模型

综合以上理论,构建基于DDPG算法的调频机组功率分配策略模型如图2所示。在PSS/E仿真平台中,建立具体电力系统的数学模型,其包含系统频率偏差与机组出力的调节比例关系;设置负荷扰动进行在线仿真,同时实时监测系统区域 i 的频率偏差 $|\Delta f(i)|$ 和联络线的区域控制误差信号 $ACE(i)$,使用模块函数将实时获取数据作为环境状态量传输至Python中。

在Python仿真平台中,读取Python传来的环境数据集,进行下一步DDPG算法的训练过程。具体地,DDPG算法调频机组功率分配的训练流程如图3所示。

1)DDPG训练过程开始前,按照负荷扰动点相同依次增大扰动功率的方式设置 S 个用于算法训练的负荷扰动场景集,分别为 $Scenaio_1-\Delta P_{load,1}$, $Scenaio_2-\Delta P_{load,2}$, $\dots$, $Scenaio_S-\Delta P_{load,S}$,其中, $Scenaio_1-\Delta P_{load,1}$ 表示设置负荷扰动功率为 $\Delta P_{load,1}$ 的场景1。本文44节点系统模型的仿真算例设置了4个负荷扰动场景。

2)对于每个扰动场景,需对4个网络的参数进行初始化,具体为先对两个评估网络Actor和Critic进行初始化,再采用软更新方法分别对其目标网络进行参数初始化,并初始化用于经验回放

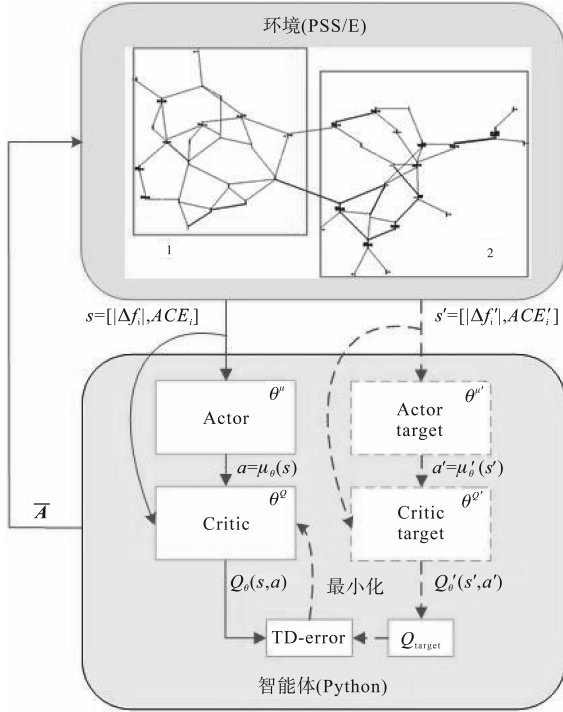


图2 基于DDPG算法的调频机组功率分配策略

Fig.2 Power allocation strategy if frequency modulation unit based on DDPG algorithm

的样本池 D 以及噪声过程参数 N 。进行算法的 M 轮训练迭代后,输出当下场景学习收敛的调频机组功率分配因子,最终得到 S 组不同负荷扰动场景下的功率分配因子,分别为

$$\begin{cases} A_1 = [a_{1,1} & a_{1,2} & \dots & a_{1,j}] \\ A_2 = [a_{2,1} & a_{2,2} & \dots & a_{2,j}] \\ \vdots \\ A_S = [a_{S,1} & a_{S,2} & \dots & a_{S,j}] \end{cases} \quad (13)$$

式中: A_i 为 Scenario- $\Delta P_{load,1}$ 场景下经算法训练后输出的一组功率分配因子向量; $a_{1,j}$ 为场景 1 下调频机组 j 的分配功率。

3) 对于每个训练轮次,首先需要初始化环境状态,然后对每个时间步都执行优化策略网络和价值函数网络的步骤。此处需要说明,本文调频控制策略的环境状态由系统频率偏差和区域控制误差信号组成。

4) 对于每个时间步内,优化策略网络和价值函数网络的步骤具体如下:

a) 根据当前 Actor 网络选择机组出力并执行机组出力调节动作,系统模拟运行后,返回系统频率 $|\Delta f'(i)|$ 和联络线的区域控制误差信号 $ACE'(i)$, 并将 $|\Delta f'(i)|$ 和 $ACE'(i)$ 作为新的环境状态量,计算奖励并将每个时间步的经验存储在

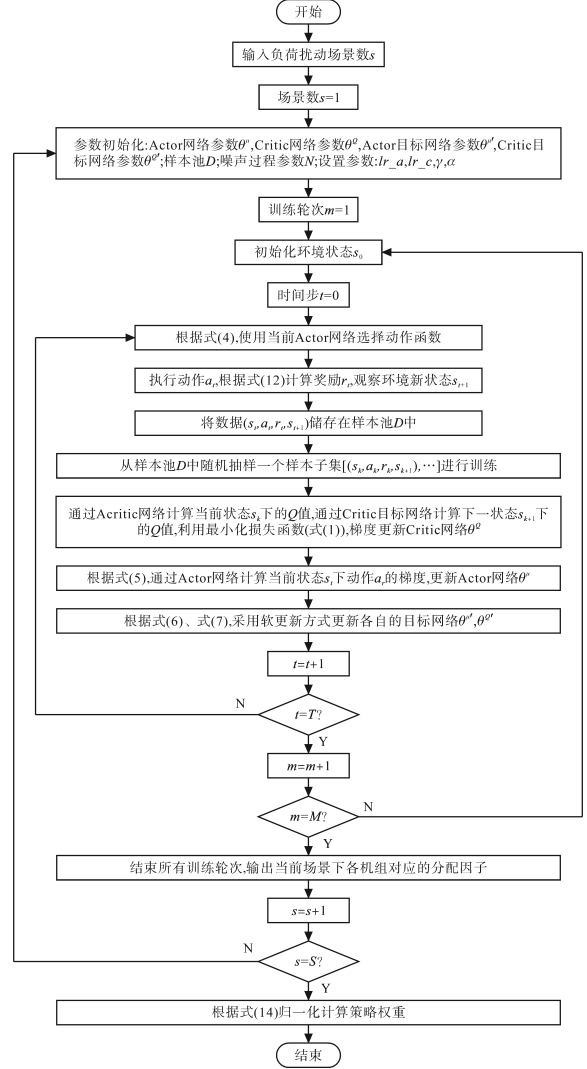


图3 调频机组功率分配策略的DDPG算法流程

Fig.3 DDPG algorithm flow of power distribution strategy of frequency modulation unit

样本池;

b) 从样本池中随机抽样一个样本子集进行训练;

c) 通过最小化损失函数进行梯度下降更新 Critic 网络;

d) 用 Actor 网络的确定性动作计算 Q 值,梯度上升更新 Actor 网络;

e) 采用软更新方式在线更新 Actor 和 Critic 目标网络。

根据 S 组不同场景的功率分配因子按下式归一化计算调频机组功率分配的策略权重,给对系统影响较大(负荷扰动较大)的场景赋予更高的权重,此策略权重综合了多个场景的学习情况,可用于求取通用策略集。

$$\bar{A} = \frac{\Delta P_{load,1}}{\sum_{s=1}^S \Delta P_{load,s}} A_1 + \frac{\Delta P_{load,2}}{\sum_{s=1}^S \Delta P_{load,s}} A_2 + \cdots + \frac{\Delta P_{load,S}}{\sum_{s=1}^S \Delta P_{load,s}} A_S \quad (14)$$

3 仿真验证

为了验证本文所提出基于 DDPG 算法的调频机组功率分配策略,基于前文描述的原理,在 PSS/E+Python 环境联合仿真平台上建立北欧地区 44 节点 IEEE 标准电力系统模型作为研究对象,其模型结构如图 4 所示。其中,左侧为区域 1,右侧为区域 2,两个区域之间有两条联络线相联系。调频机组的设置上,区域 1 内节点“5600”上的 4 台发电机为调频机组,可调容量为 3 256 MW;区域 2 内节点“7000”上的 9 台机组为调频机组,可调容量为 4 140 MW。同时,模拟负荷突然投入,为系统引入一个阶跃扰动。

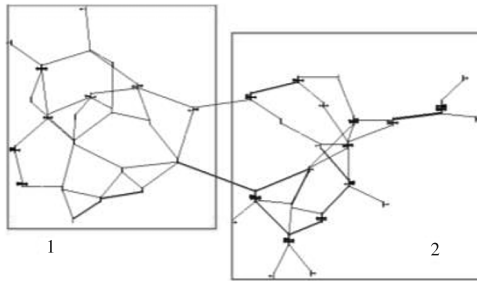


图 4 44 节点 IEEE 标准电力系统模型

Fig.4 44-node IEEE standard power system model

3.1 仿真设置

基于 DDPG 的调频控制系统,需要对某些参数进行初始化设置,如表 1 所示。

Actor 网络学习率 lr_a 是在学习过程中用于更新 Actor 参数的超参数。其值越大,系统的权重更新越快;反之可以减少权重的剧烈波动以提高稳定性。为了在保证稳定性的前提下提高网络的性能,经过多轮仿真调整,设置 lr_a 为 0.001。

Critic 网络学习率 lr_c 是在学习过程中用于更新 Critic 参数的超参数。其值越大,智能体寻优的速度越快;反之智能体寻优的速度越慢。在确保算法能够收敛的前提下用较快的运算速度,经过多轮仿真调整,设置 lr_c 为 0.002。

折扣因子 γ ($0 < \gamma < 1$) 是未来奖励的衰减值。 γ 过大时,由于距离现在较远,网络对未来奖励的评估精确度下降,预测结果不可信; γ 过小时,网络无法及时预见未来,只能看见当前的奖励,失去了前瞻性。经过多轮仿真调整,设置 γ 为 0.9。

滑动平均系数是反映目标网络更新其网络参数程度的参数。当 α 过大时,由于目标网络参数更新剧烈,容易给损失函数带来较大的扰动,造成神经网络更新的精确性下降; α 过小时,由于目标网络参数更新过于平缓,导致智能体的学习效率下降,算法收敛所需的时间变长,不利于系统控制,经过仿真调整,选取 $\alpha=0.01$ 。

表 1 参数设置

Tab.1 Parameter setting

参数	数值
lr_a	Actor 网络学习率 0.001
lr_c	Critic 网络学习率 0.002
γ	折扣因子 0.9
α	滑动平均系数 0.01

本文采用的 DDPG 算法采用经验回放机制,经验回放池容量为 1 000,算法通过与环境交互,将样本添加进经验回放池,在训练过程中,算法会从回放池中随机抽取 32 个样本进行批量学习。本文为确保实验结果一致,将环境随机种子固定为 1。

为验证所提方法在不同扰动场景下的适用性,本文在节点“5600”处设置不同大小的负荷阶跃扰动,如表 2 所示,仿真的起始时刻就是负荷施加扰动的时刻。

表 2 扰动场景设置

Tab.2 Disturbance scene setting

负荷扰动场景 Scenarios	扰动大小 $\Delta P_{load,s}/MW$	扰动节点
Scenario ₁ - $\Delta P_{load,1}$	500	节点“5600”
Scenario ₂ - $\Delta P_{load,2}$	1 000	节点“5600”
Scenario ₃ - $\Delta P_{load,3}$	1 500	节点“5600”
Scenario ₄ - $\Delta P_{load,4}$	2 000	节点“5600”

3.2 仿真结果分析

3.2.1 算例 1: 负荷扰动为 500 MW

负荷扰动为 500 MW 时,由图 5 可知,强化学习奖励值在剧烈波动后显著收敛。在初始阶段(迭代次数约 0~1 000)奖励值波动剧烈,这表明算法正处于探索阶段,通过尝试各种策略以寻找最优解,在探索过程中导致了奖励值的剧烈波动。大约在 1 000 次迭代后,波动幅度显著减小,奖励值稳定上升并收敛至-10 至-30 之间,这表明算法已找到较为稳定的优化策略。

图 6 显示了各机组的最终优化结果,可以看出,各机组的功率分配呈现出明显的差异化。其中,机组 1~3 的分配因子最高(约 0.16~0.17),机

组4较低一些(约0.14),机组10和机组13处于中等水平(分别约为0.12和0.08),而机组5、机组6的分配因子较小(约0.04~0.05);而机组7、机组8、机组9、机组11和机组12的分配因子极小(低于0.03)。这表明算法成功识别出了各机组的功率分配对二次调频的影响,在生成功率分配策略时优先调用调频效果更佳的机组,对调频效果较弱的机组则减少功率分配,以实现整体系统的优化运行。

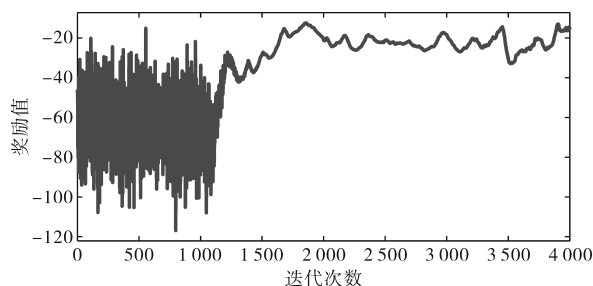


图5 500 MW 负荷扰动时 DDPG 训练过程奖励值

Fig.5 Reward values during DDPG training process under 500 MW load disturbance

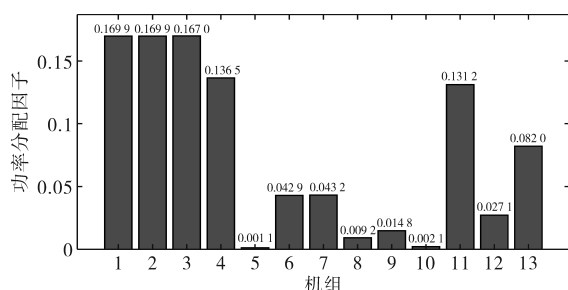


图6 500 MW 负荷扰动时各机组的功率分配因子

Fig.6 Power allocation factors of generating units under 500 MW load disturbance

3.2.2 算例2:负荷扰动为1 000 MW

负荷扰动为1 000 MW时,由图7可知,强化学习值同样在剧烈波动后显著收敛。在初始阶段(迭代次数约0~1 000)时奖励值波动剧烈,范围在-100至-450之间。大约在1 200次迭代后,奖励值显著提升并趋于稳定,最终收敛至-60至-80之间,波动幅度大幅降低,表明算法已成功找到接近最优的策略,训练效果显著。

图8显示了各机组功率分配的最终优化结果。机组1~4功率分配因子最高(约0.17~0.18);机组10略低一些(约0.08);机组5~9和机组13功率分配因子中等(约0.03~0.04);而机组11和机组12功率分配因子最低(约0.01)。这表明了算法成功识别出了不同机组发出功率对二次调频的影响,通过分层配置差异化功率分配方案,以有效提升系统的整体调控性能。

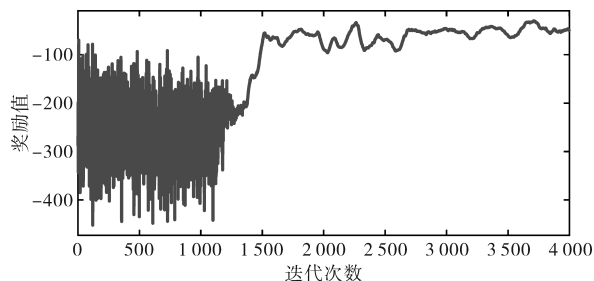


图7 1 000 MW 负荷扰动时 DDPG 训练过程奖励值

Fig.7 Reward values during DDPG training process under 1 000 MW load disturbance

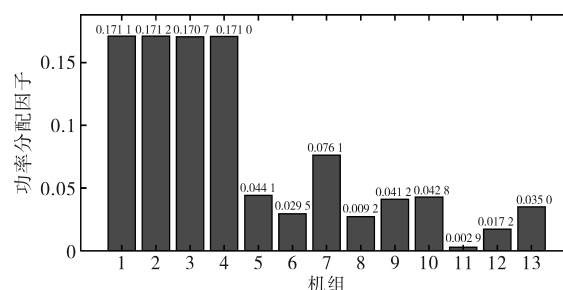


图8 1 000 MW 负荷扰动时各机组的功率分配因子

Fig.8 Power allocation factors of generating units under 1 000 MW load disturbance

3.2.3 算例3:负荷扰动为1 500 MW

负荷扰动为1 500 MW时的强化学习训练过程如图9所示。初始阶段时(迭代次数约0~1 000)奖励值波动剧烈,奖励值波动范围在-200至-1 000之间,表明算法处于初期探索阶段。大约在1 200~1 500次迭代后,奖励值显著提升并趋于稳定,并最终收敛到-10至-20之间,奖励值波动幅度大幅降低,这表明算法已成功找到最优的功率分配策略,训练效果显著。

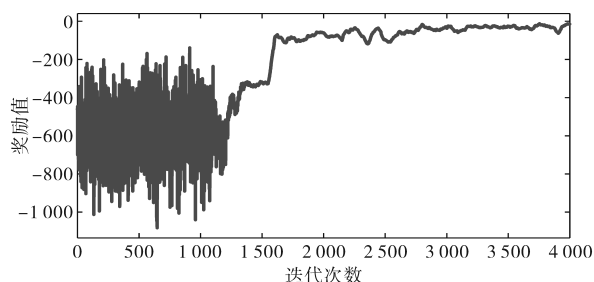


图9 1 500 MW 负荷扰动时 DDPG 训练过程奖励值

Fig.9 Reward values during DDPG training process under 1 500 MW load disturbance

图10展示了负荷扰动为1 500 MW时各机组功率分配的最终优化结果。机组1、机组3、机组4、机组5功率分配因子最高(约0.25左右);机组2功率分配因子中等(约0.13);而机组6~机组13功率分配因子最低(接近于0)。这表明算法成功识别出了不同的机组功率分配策略对二次调频

的影响,优先对调频性能更优的机组进行功率调控,以提升系统的整体调控性能。

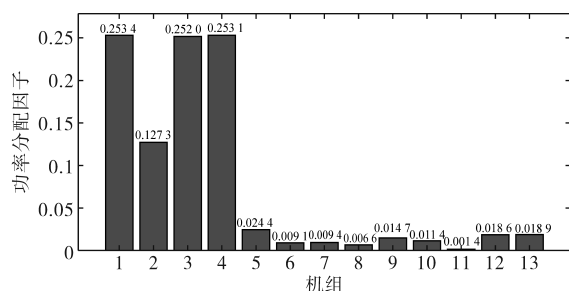


图10 1500 MW负荷扰动时各机组的功率分配因子

Fig.10 Power allocation factors of generating units under 1500 MW load disturbance

3.3.4 算例4:负荷扰动为2000 MW

负荷扰动为2000 MW时的强化学习训练过程如图11所示。初始阶段时(迭代次数约0~1000)奖励值波动剧烈,范围在-2000至-300之间。大约在1200~1500次迭代后,奖励值显著提升并最终收敛到-70至-60之间,表明算法已经成功找到了接近最优的功率分配策略,训练效果良好。

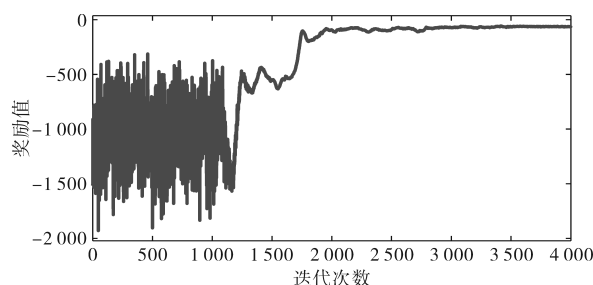


图11 2000 MW负荷扰动时DDPG训练过程奖励值

Fig.11 Reward values during DDPG training process

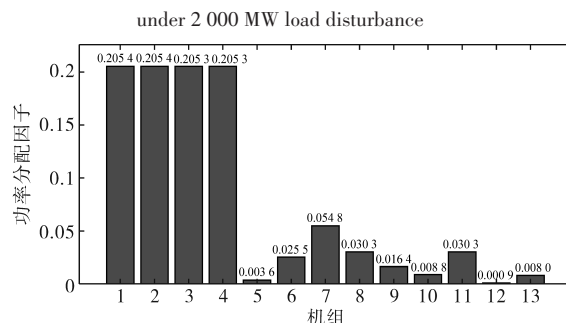


图12 2000 MW负荷扰动时各机组的功率分配因子

Fig.12 Power allocation factors of generating units under 2000 MW load disturbance

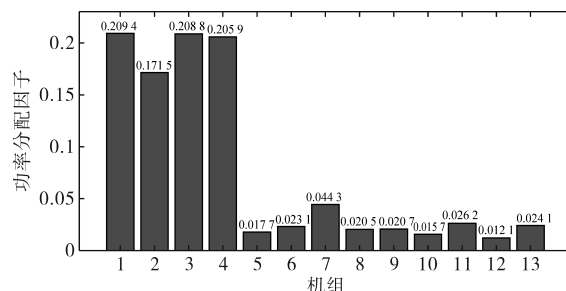


图13 通用策略集的加权机组功率分配因子

Fig.13 Unit weighted power distribution factor of general strategy set

表3 DDPG与其他强化学习算法性能对比

Tab.3 Performance comparison between DDPG and other reinforcement learning algorithms

算法	Q学习	DQN	REINFORCE	DDPG
动作空间	离散	离散	连续/离散	连续
学习机制	值函数	值函数	策略梯度	值函数+策略梯度
样本效率	低	中	低	高
算法稳定性	中	高	低	高
收敛速度	慢	慢	慢	快
抗干扰能力	弱	弱	强	强
功率分配精度	低	中	高	高

通过上述4个算例可以看出,本文所提方法能够在不同负荷扰动场景下都能训练得到一个功率分配策略,进一步地,可将多组负荷扰动场景的功率分配因子进行加权,最终得到通用策略集的加权功率分配因子,如图13所示。通用策略集综合了多个场景的学习情况,考虑了不同扰动场景下功率分配的需求和条件。可见,本文所提出的方法可以更好地应对各种负荷变化,具有较好的灵活性与普适性,适应于新型电力系统多种复杂应用场景,具有在新型电力系统中的进一步探索和应用价值。

为了进一步展示DDPG在本文所述功率分配任务中的优势,本文将DDPG和其他几种典型的强化学习算法进行了对比,如表3所示。

DDPG拥有连续的动作空间,能够实现功率分配因子的高精度分配,而Q学习、DQN等离散算法在动作量化的过程中都会产生一定程度上的精度损失,因此,DDPG可以更高精度地执行功率分配任务,更适用于本文所述的应用场景。从

算法结构来看,DDPG综合了价值函数和策略梯度两种学习机制,同时具有策略优化高效、学习过程稳定的优点。DDPG还具有较高的样本利用率,能够加快收敛速度并在一个较短的时间范围内生成质量较好的调频控制策略。综上所述可以看出,与其他强化学习算法相比,本文采用的DDPG算法更适用于功率分配任务,具有可进一步探索的应用价值。

4 结论

在新型配电系统中,具有极强的间歇性和不确定性的规模新能源接入将给电网带来强烈的随机扰动,使得频率控制变得复杂且困难。针对新型电力系统的调频需求,本文引入DDPG算法生成电网在频率扰动时期的机组分配策略。该策略考虑了不同的扰动位置、扰动方式,以扰动加权的形式设计奖励函数,促使DDPG引导智能体给出优化解,实现机组调频功率的合理分配与控制,保证电力系统频率稳定和安全运行。建立了基于Python调用PSS/E的联合仿真平台,验证了利用联合仿真平台生成优化控制策略的可行性与优越性,为指导工程实践奠定基础。

参考文献

- [1] 李晖,刘栋,姚丹阳. 面向碳达峰碳中和目标的我国电力系统发展研判[J]. 中国电机工程学报, 2021, 41(18): 6245-6258.
LI Hui, LIU Dong, YAO Danyang. Analysis and reflection on the development of power system towards the goal of carbon emission peak and carbon neutrality[J]. Proceedings of the CSEE, 2021, 41(18): 6245-6258.
- [2] 马骏超,王晨旭,孙丹,等. 频率连续跌落故障下风电系统调频控制策略[J]. 电气传动, 2024, 54(12): 61-70, 78.
MA Junchao, WANG Chenxu, SUN Dan, et al. Frequency regulation control strategy for direct-drive wind power systems considering continuous frequency drop[J]. Electric Drive, 2024, 54(12): 61-70, 78.
- [3] 席磊,杜雄,李彦营,等. 基于具有强化学习思想的集成学习自动发电控制算法[J]. 南方电网技术, 2023, 17(7): 74-82.
XI Lei, DU Xiong, LI Yanying, et al. Automatic generation control algorithm based on ensemble learning with the idea of reinforcement learning[J]. Southern Power System Technology, 2023, 17(7): 74-82.
- [4] 胡泽春,罗浩成. 大规模可再生能源接入背景下自动发电控制研究现状与展望[J]. 电力系统自动化, 2018, 42(8): 2-15.
HU Zechun, LUO Haocheng. Research status and prospect of automatic generation control with integration of large-scale renewable energy[J]. Automation of Electric Power Systems, 2018, 42(8): 2-15.
- [5] 滕贤亮,谈超,昌力,等. 高比例新能源电力系统有功功率与频率控制研究综述及展望[J]. 电力系统自动化, 2023, 47(15): 12-35.
TENG Xianliang, TAN Chao, CHANG Li, et al. Review and prospect of research on active power and frequency control of high proportion new energy power system[J]. Automation of Electric Power Systems, 2023, 47(15): 12-35.
- [6] 王念,张靖,李博文,等. 自动发电控制研究综述[J]. 电测与仪表, 2020, 57(21): 1-8.
WANG Nian, ZHANG Jing, LI Bowen, et al. Research review of automatic generation control[J]. Electrical Measurement & Instrumentation, 2020, 57(21): 1-8.
- [7] 陈宋宋,张路涛,周颖,等. 面向新能源并网的分布式AGC协同算法[J]. 南方电网技术, 2023, 17(4): 58-68.
CHEN Songsong, ZHANG Lutao, ZHOU Ying, et al. Distributed AGC cooperative algorithm for new energy grid connection[J]. Southern Power System Technology, 2023, 17(4): 58-68.
- [8] 胡泽春,谢旭,张放,等. 含储能资源参与的自动发电控制策略研究[J]. 中国电机工程学报, 2014(29): 5080-5087.
HU Zechun, XIE Xu, ZHANG Fang, et al. Research on automatic generation control strategy incorporating energy storage resources[J]. Proceedings of the CSEE, 2014(29): 5080-5087.
- [9] 马斌,王亚军,栗维勋,等. 考虑调频能力的含储能风电场自动发电控制策略研究[J]. 电力系统及其自动化学报, 2022, 34(12): 66-71.
MA Bin, WANG Yajun, LI Weixun, et al. Research on automatic generation control strategy for wind farm with energy storage considering frequency modulation capability[J]. Proceedings of the CSU-EPSA, 2022, 34(12): 66-71.
- [10] 席磊,王昱昊,陈宋宋,等. 面向综合能源系统的多智能体协同AGC策略[J]. 电机与控制学报, 2022, 26(4): 77-88.
XI Lei, WANG Yuhao, CHEN Songsong, et al. Multi-agent collaborative AGC strategy for integrated energy system[J]. Electric Machines and Control, 2022, 26(4): 77-88.
- [11] 李忠文,柏宁宁,程志平,等. 新型电力系统中异质调频机组分布式协同AGC方法研究[J]. 电网技术, 2024, 48(6): 2327-2335.
LI Zhongwen, BAI Ningning, CHENG Zhiping, et al. Distributed collaborative AGC method for heterogeneous frequency regulation units in new power systems[J]. Power System Technology, 2024, 48(6): 2327-2335.
- [12] 李凯斌,陈磊,闵勇,等. 基于类噪声量测数据的自动发电控制振荡模式辨识方法[J]. 电力系统自动化, 2022, 46(23): 76-82.
LI Kaibin, CHEN Lei, MIN Yong, et al. Identification method for automatic generation control oscillation mode based on ambient measurements[J]. Automation of Electric Power Systems, 2022, 46(23): 76-82.
- [13] 黄伦俊,汪震,单煜,等. 计及SOC恢复的互联网火储联合AGC控制策略研究[J]. 能源工程, 2022, 42(2): 78-86.

- HUANG Lunjun, WANG Zhen, SHAN Yu, et al. An AGC control strategies based on thermal units and BESS of interconnected grid considering energy storage SOC recovery[J]. *Energy Engineering*, 2022, 42(2): 78–86.
- [14] 范璟. 级联PD-模糊PID控制器的自动发电控制方法[J]. *电子测量技术*, 2021, 44(3): 87–92.
- FAN Jing. Automatic generation control method based on cascade PD fuzzy PID controller[J]. *Electronic Measurement Technology*, 2021, 44(3): 87–92.
- [15] 李媛媛, 宋永华, 魏少岩. AGC机组调频经济性的混沌遗传算法研究[J]. *电力系统保护与控制*, 2008, 36(14): 17–21.
- LI Yuanyuan, SONG Yonghua, WEI Shaoyan. Research of chaos genetic algorithm on AGC units dispatchment[J]. *Power System Protection and Control*, 2008, 36(14): 17–21.
- [16] 范宏, 王兰坤, 邢梦晴, 等. 考虑频率稳定约束的电-氢互补多楼宇协调优化调度[J]. *上海交通大学学报*, 2023, 57(12): 1559–1570.
- FAN Hong, WANG Lankun, XING Mengqing, et al. Coordinated scheduling of multiple buildings with electric-hydrogen complementary considering frequency stability constraints [J]. *Journal of Shanghai Jiaotong University*, 2023, 57(12): 1559–1570.
- [17] PATEL R, LI C, MEEGAHAPOLA L, et al. Enhancing optimal automatic generation control in a multi-area power system with diverse energy resources[J]. *IEEE Transation on Power Systems*, 2019, 34(5): 3465–3475.
- [18] TANG Yong, HUANG Yanhao, WANG Hongzhi, et al. Framework for artificial intelligence analysis in large-scale power grids based on digital simulation[J]. *CSEE Journal of Power and Energy Systems*, 2018, 4(4): 459–468.
- [19] 许懿, 欧睿, 庞涛. 自动发电控制AGC关键问题分析与智能化改善[J]. *信息技术*, 2022, 46(9): 123–128.
- XU Yi, OU Rui, PANG Tao. Analysis and intelligent improvement of key issues of automatic power generation control AGC [J]. *Information Technology*, 2022, 46(9): 123–128.
- [20] 周毅, 周良才, 史迪, 等. 基于安全深度强化学习的电网有功频率协同优化控制[J]. *上海交通大学学报*, 2024, 58(5): 682–692.
- ZHOU Yi, ZHOU Liangcai, SHI Di, et al. Cooperative optimal control of active power frequency of power grid based on deep reinforcement learning [J]. *Journal of Shanghai Jiaotong University*, 2024, 58(5): 682–692.
- [21] JANG Hung-Chin, LEE Chun-Yee. Study on Q-Learning and deep Q-Learning in urban roadside parking space search[C]// 2022 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA: IEEE, 2022: 1248–1253.
- [22] LOW Ee Soong, ONG Pauline, CHEAH Kah Chun. Solving the optimal path planning of a mobile robot using improved Q-Learning[J]. *Robotics and Autonomous Systems*, 2019, 115: 143–161.
- [23] DONG Yuansheng, ZOU Xingjie. Mobile robot path planning based on improved DDPG reinforcement learning algorithm[C]// 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China: IEEE, 2020: 52–56.
- [24] 余璐. 基于深度强化学习的电力系统自动发电控制[D]. 宜昌: 三峡大学, 2020.
- YU Lu. Deep reinforcement learning for automatic generation control of power systems[D]. Yichang: China Three Gorges University, 2020.
- [25] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning[C]// 2016 International Conference on Learning Representations, San Juan, Puerto Rico: LCLR, 2016.

收稿日期: 2025-03-20

修改稿日期: 2025-07-31