

基于高维协同软演员-评论家的多智能体自动发电控制

柳丹^{1,2}, 任建宇¹, 席磊^{1,3}, 刘治洪⁴, 全悦⁵, 施宇¹

(1. 三峡大学电气与新能源学院, 湖北 宜昌 443002; 2. 国网湖北省电力有限公司电力科学研究院, 武汉 430077; 3. 梯级水电站运行与控制湖北省重点实验室(三峡大学), 湖北 宜昌 443002; 4. 输变电装备技术全国重点实验室(重庆大学), 重庆 400044; 5. 北京交通大学电气工程学院, 北京 100044)

摘要: 随着新能源渗透率不断提高, 风光等新能源出力的强随机性导致电网频率不稳定及控制性能变差。为此, 从自动发电控制角度探索一种多智能体强化学习方法, 即高维协同软演员-评论家算法。所提算法通过在最大熵框架下鼓励智能体进行随机探索, 以解决 Q 学习及其衍生算法无法快速更新 Q 表以适应环境变化的问题。同时利用高斯分布策略生成连续动作值, 使算法可在高维连续状态空间中寻找协同最优解, 以解决传统强化学习高维“状态-动作”的维数灾问题, 从而有效应对新能源出力强随机性所导致的频率不稳定及控制性能变差的问题。通过在改进的 IEEE 标准两区域负荷频率控制模型及华中三区域负荷频率控制模型上的仿真实验, 验证了该算法的有效性, 且相较于其他算法, 具有更优的控制性能及频率稳定性。

关键词: 自动发电控制; 高维协同; 软演员-评论家; 多智能体; 高斯分布

Multi-Agent Automatic Power Generation Control Based on High Dimensional Collaborative Soft Actor-Critic

LIU Dan^{1,2}, REN Jianyu¹, XI Lei^{1,3}, LIU Zhihong⁴, QUAN Yue⁵, SHI Yu¹

(1. College of Electrical and New Energy, China Three Gorges University, Yichang, Hubei 443002, China; 2. Electric Power Research Institute, State Grid Hubei Electric Power Co., Ltd., Wuhan 430077, China; 3. Hubei Provincial Key Laboratory for Operation and Control of Cascaded Hydropower Station, Yichang, Hubei 443002, China; 4. State Key Laboratory of Power Transmission Equipment Technology, Chongqing University, Chongqing 400044, China; 5. School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: As renewable energy penetration continues to increase, the strong randomness of wind and solar renewable energy output leads to instability of grid frequency and deterioration of control performance. To address this, a multi-agent reinforcement learning approach is explored from the perspective of automatic generation control, that is the high dimensional cooperative soft actor-critic algorithm. The proposed algorithm encourages agents to engage in random exploration within a maximum entropy framework to address the inability of Q-learning and its derivative algorithm to rapidly update the Q-table in response to environmental changes. It also utilizes a Gaussian distribution strategy to generate continuous action values, enabling the algorithm to find collaborative optimal solutions in high dimensional continuous state spaces, thus solving the traditional reinforcement learning curse of dimensionality "state-action". Frequency instability and declining control performance caused by the strong randomness of renewable energy outputs are effectively addressed. Through simulation experiments on an improved IEEE standard two-area load frequency control model and the Central China three-area load frequency control model, the effectiveness of the algorithm is validated. And the proposed algorithm has superior control performance and frequency stability compared to other algorithms.

Key words: automatic power generation control; high dimensional collaboration; soft actor-critic; multi-agent; Gaussian distribution

基金项目: 国家自然科学基金资助项目(52277108, 52477104)。

Foundation item: Supported by the National Natural Science Foundation of China (52277108, 52477104).

0 引言

新能源为主体的新型电力系统模式下,规模化新能源随机出力的强时空差异导致电力系统难以维持净负荷量平衡,影响电力系统的频率稳定^[1]及控制性能。学者们试图从自动发电控制(automatic generation control, AGC)角度进行探索,以解决新型电力系统模式下规模化新能源并网所带来的电力系统频率^[2]不稳定及控制性能变差的问题。

基于马尔科夫随机决策过程的强化学习方法^[3-4]在解决随机问题方面具有优势。文献[5]将 $R(\lambda)$ 模仿学习策略引入AGC,解决了随机环境下AGC自适应问题。文献[6]将控制性能标准(control performance standard, CPS)下的互联电网看作一个随机系统,将其转化为马尔科夫随机决策问题,进而采用Q学习来解决,实现了AGC系统的动态优化。文献[7]提出了一种基于多时间尺度的风机调频控制策略,通过优化风电机组频率响应系数和储能出力设置,提升了整个电网的频率稳定性。然而上述文献所采用控制模式为集中式控制。在大规模新能源接入的分布式多区域新型电力系统背景下,由于集中式控制无法进行多区域协同,不能完全适用于分布式新型电力系统。

传统强化学习尤以Q学习及其衍生算法应用更为广泛。学者们试图从具有模型无关特性的Q学习及其衍生算法角度探索分布式AGC模式,以适应分布式新型电力系统^[8]的协同控制,进而获取更优的控制性能。文献[9]提出了一种融入资格迹的双Q学习算法,通过统一双Q学习和双Expected-Sarsa(λ),避免了Q值高估问题,消除了分布式电网中的频率不稳定问题。文献[10]提出了一种具有动作自寻优的双Q学习算法以应对大规模分布式能源接入电网时的强随机扰动问题。文献[11]将乐观原理融入双Q学习算法解决Q值高估问题,有效解决了大规模风机并网的分布式电网强随机扰动问题。文献[12]将差分进化变参数向量融入双Q学习算法,提高了多区域集成能源系统的收敛速度。文献[13]利用已训练的神经网络替换Q学习算法中的动作选择机制,解决了智能发电的频率控制问题。文献[14]将资格迹引入Q学习中解决了最优碳流问题。文献[15]提出

了分层强化学习方法,通过对AGC中的发电指令派遣动态优化任务进行分解和引入时间变化协调因子以降低AGC调节成本。

上述Q学习及其衍生算法需要较长时间来更新Q表以适应环境变化^[16],这使得调频机组无法快速作出反应,导致电力系统频率突变,最终威胁运行安全^[17-19]。本文引入软更新演员评论家(soft actor-critic, SAC)算法^[20],通过在每个决策步骤考虑动作的熵,有利于探索最优动作,从而在环境突变时快速找到适应新环境的策略。

然而,随着新能源^[21-23]的不断接入,系统的状态和控制变量随之增加,并且由于电力系统中各机组存在交互和耦合关系,使得系统的状态进一步增加。而SAC算法由于系统状态和控制变量的增加而导致状态空间的广度和深度增加,高维状态-动作的可能组合呈现指数级增长,使学习过程变得困难且计算量增大,即维数灾问题^[24](这将在本文的实验部分证实)。此问题导致智能体无法在有限时间内获取最优策略,以至于分布式电力系统在面对具有强随机性扰动时,不能及时做出调整,从而影响电力系统的稳定性。

文献[25]采用神经网络近似价值函数,通过参数化策略函数减轻了维数灾。文献[26]采用Beta分布策略生成连续的动作值以提高控制精度从而减轻维数灾问题。但上述方法在面对动作幅度变化大及控制精度要求高的分布式电力系统时,智能体无法寻找最优策略。因此,本文采用具有高维协同属性的高斯分布策略^[27],其参数化的均值和方差分别表示在当前策略下最可能的动作及其探索程度,这允许智能体在面对动作幅度变化大的情况时执行最可能的动作且能考虑其他近似的动作,从而探索到更精确的最优动作以解决维数灾问题。

因此,本文针对新能源规模化接入电力系统所带来的频率不稳定及控制性能变差的问题,提出了一种能将智能体执行策略建模为连续动作空间并探索高维状态-动作协同最优解的多智能体AGC算法,即基于高维协同的软演员-评论家(high dimensional synergy of soft actor-critic, HDS-SAC)的多智能体AGC算法。所提算法通过在最大熵框架下鼓励智能体进行随机探索以快速适应环境变化,并采用具有高维协同属性的高斯分布策略在连续状态空间中定义每个动作的概率使智能体能快速探索到最优动

作以解决维数灾问题。在改进的IEEE标准两区域负荷频率控制(load frequency control, LFC)模型和华中三区域电网模型进行仿真,验证了HDS-SAC的有效性。

1 HDS-SAC算法

面向新型电力系统提出了基于HDS-SAC的多智能体AGC算法。在SAC框架中利用高斯分布策略对连续动作空间进行建模,策略学习不依赖于动作空间的离散化,而是利用高斯分布的连续性来生成动作以解决高维“状态-动作”的维数灾问题。

1.1 SAC算法

强化学习^[28]是通过大量数据在特定场合下在线学习一个最优的执行策略,此策略能够最大化智能体在一定策略下的预期累计回报。以Q学习及其衍生算法为主流的经典强化学习算法依靠价值矩阵在离散动作空间学习到一个最优控制策略。但面对新能源接入的分布式电力系统时,因其无法快速更新Q表以迅速响应具有强随机性的新能源出力,即存在无法快速适应环境变化的问题。

SAC引入熵正则化使智能体执行探索行为,确保回报最大化的同时保持策略的随机性。通过在每个决策步骤考虑动作的熵,有利于探索最优动作,从而在环境突变时快速找到适应新环境的策略。加入熵正则化的价值函数如式(1)所示。

$$Q^\pi(s, a) = r(s, a) + \gamma E_{s', a'} [Q^\pi(s', a') - \alpha \log(\pi(a'|s'))] \quad (1)$$

式中: a 和 a' 分别为当前动作和下一个时刻的动作; s 和 s' 分别为当前状态和下一时刻的状态; r 为奖励函数; γ 为折扣因子; $Q^\pi(s, a)$ 为在遵循策略 π 时,在当前状态 s 执行动作 a 得到的期望回报; $\log(\pi(a'|s'))$ 为策略 π 在状态 s' 下选择动作 a' 的对数概率; α 为熵正则化系数; $E_{s', a'}$ 为未来状态 s' 和未来动作 a' 的期望值。

因此,根据式(1)可知在状态和动作空间受限的条件下该策略评估可以收敛到策略 π 的Q函数。则改进策略如式(2)所示。

$$\pi_{\text{new}} = \arg \min_{\pi \in \Pi} D_{\text{KL}} \left(\pi(\cdot|s_t) \parallel \frac{e^{\frac{1}{\alpha} Q_{\text{soft}}^\pi(s_t, \cdot)}}{Z_{\text{soft}}^\pi(s_t)} \right) \quad (2)$$

式中: π_{new} 为新策略; s_t 为智能体在时间步 t 时的状态; $\pi(\cdot|s_t)$ 为当前策略在状态 s_t 下的动作概率分布;

$Q_{\text{soft}}^\pi(s_t, \cdot)$ 为策略 π_{old} 在状态 s_t 下的软状态-动作值函数; D_{KL} 为Kullback-Leibler(KL)散度,衡量两个分布之间的差异; $Z_{\text{soft}}^\pi(s_t)$ 为旧策略 π_{old} 在状态 s_t 下的分布值函数; Π 为所有可能的策略空间。

1.2 高斯分布策略

在最大熵强化学习框架下, SAC的策略更加随机。当环境复杂度增加时,状态空间维度增加导致SAC算法的状态-动作的组合增加进而出现维数灾。为此本文采用具有高维协同属性的高斯分布策略,其参数化的均值和方差分别表示在当前策略下最可能的动作和动作的探索程度,这允许智能体即便在面对动作幅度变化大的情况下,在执行最可能的动作的同时仍能考虑其他近似的动作,从而探索到更精确的最优动作以解决维数灾问题。

高斯分布策略用于连续动作空间,其中策略 π 是一个输出动作的概率分布,而非单一动作。假设给定状态下的动作服从多元高斯分布,其数学形式表示为 $\pi(als)$ 。假设在某一状态 s 下策略输出的动作服从高斯分布。

$$\pi(als) = \frac{1}{\sqrt{2\pi} \sigma(s)} e^{-\frac{[a - \mu(s)]^2}{2\sigma^2(s)}} \quad (3)$$

式中: $\pi(als)$ 为在状态 s 下执行动作 a 的概率; $\mu(s)$ 为在状态 s 下采取动作 a 的均值; $\sigma(s)$ 为在状态 s 下采取动作 a 的标准差。

若用神经网络近似 $\mu(s)$ 和 $\sigma(s)^2$ (此处近似方差对数 $\rho(s; \theta) = \ln \sigma(s)^2$), 则策略近似为:

$$\pi(als; \theta) = \frac{1}{\sqrt{2\pi} e^{\rho(s; \theta)}} e^{-\frac{1}{2} \left[\frac{a - \mu(s; \theta)}{e^{\rho(s; \theta)}} \right]^2} \quad (4)$$

式中: $\pi(als; \theta)$ 为给定状态 s 和策略网络参数 θ 的条件下采取动作 a 的概率密度函数; $\mu(s; \theta)$ 为均值函数, $\rho(s; \theta)$ 为方差函数。

1.3 HDS-SAC算法

基于SAC并融合高斯分布策略,本文提出了HDS-SAC算法。在HDS-SAC中,从状态-动作对 (s_t, a_t) 中定义策略 π 的软状态-动作回报如式(5)所示。

$$Z^\pi(s_t, a_t) = r_t + \gamma V^\pi(s_{t+1}) \quad (5)$$

$$V^\pi(s) = E_{a \sim \pi} [Q^\pi(s, a) - \log \pi(als)] \quad (6)$$

$$Q^\pi(s, a) = E_{s' \sim p} [r(s, a) + \gamma V^\pi(s')] \quad (7)$$

式中: $Z^\pi(s_t, a_t)$ 为智能体在状态 s_t 下按照策略 π 采取动作 a_t 获得的价值分布; $V^\pi(s)$ 为在状态 s 下按照

策略 π 选择动作的预期回报； r_t 为奖励函数； $\log \pi(a|s)$ 为在状态 s 下采取动作 a 的对数概率； $\sim$ 为服从某种分布； P 为状态转移概率； $E_{a \sim \pi}(\cdot)$ 为根据策略 π 下选择动作 a 的期望值； $E_{s' \sim P}(\cdot)$ 为状态转移概率 P 下对下一动作 s' 的期望值。

则在最大熵框架下的 Bellman 算子的软状态-动作分布值函数如式(8)所示。

$$T_R^\pi Z^\pi(s, a) = r + \gamma [E_{s' \sim \beta} (Z^\pi(s', a') - \alpha E_{a' \sim \pi} \log \pi(a'|s'))] \quad (8)$$

式中： T_R^π 为作用于价值分布 $Z^\pi(s, a)$ 的 Bellman 更新算子； R 为熵正则化的算子； $Z^\pi(s', a')$ 为状态 s' 、动作 a' 下的分布值函数； $E_{a' \sim \pi}(\cdot)$ 为根据策略 π 选择动作 a' 的期望值； $E_{s' \sim \beta}(\cdot)$ 为在状态空间 β 中选择一个状态 s' ，并在策略 π 选择动作 a 的期望值。

假设通过两个独立的函数 $Q_\theta(s, a)$ 和 $\sigma_\psi(s, a)$ 近似高斯分布的均值(即 Q 值)和协方差，即 $F_{\theta, \psi}(\cdot|s, a) = N(Q_\theta(s, a), \sigma_\psi^2(s, a))$ ，高斯分布为：

$$D_{KL}(F_{\text{target}}(\cdot|s, a), F_{\theta, \psi}(\cdot|s, a)) = \log \frac{\sigma_\psi(s, a)}{\sigma_{\text{target}}} + \frac{\sigma_{\text{target}}^2 + (Q^\pi(s, a) - Q_\theta(s, a))^2}{2\sigma_\psi^2(s, a)} - \frac{1}{2} \quad (9)$$

式中： $Q_\theta(s, a)$ 为预测的动作值函数； $\sigma_\psi(s, a)$ 为预测的标准差； $F_{\text{target}}(\cdot|s, a)$ 为智能体在状态 s 下采取动作 a 获得未来累计奖励的理想分布； Ψ 为标准差函数的参数； $F_{\theta, \psi}(\cdot|s, a)$ 为参数化价值分布； σ_{target} 为目标分布的标准差； σ_{target}^2 为目标分布的方差； $N(Q_\theta(s, a), \sigma_\psi^2(s, a))$ 为均值为 $Q_\theta(s, a)$ 、方差为 $\sigma_\psi^2(s, a)$ 的高斯分布。

由于策略网络和动作概率分布建模为高斯分布，且在连续状态空间下高斯分布对采样动作的过程不可导，因此需要将 $J_\pi(\varphi)$ 重参数化^[29]。HDS-SAC 通过重参数化进行策略梯度更新，即更新策略网络 π_φ 中的参数 φ 如式(10)所示。

$$J_\pi(\varphi) = E_{s \sim \beta, a \sim \pi} [-\alpha \log(\pi_\varphi(f(\varepsilon_t; s_t)|s_t)) + (Q_\theta(s_t, f(\varepsilon_t; s_t)) - \alpha \log(\pi_\varphi(f(\varepsilon_t; s_t)|s_t))) f_\phi(\varepsilon_t; s_t)] \quad (10)$$

式中： $J_\pi(\cdot)$ 为策略 π 的目标函数； ε_t 为随机噪声变量； $\log(\pi_\varphi(f(\varepsilon_t; s_t)|s_t))$ 为智能体在状态 s_t 下选择动

作 $f(\varepsilon_t; s_t)$ 的对数概率； $Q_\theta(s_t, f(\varepsilon_t; s_t))$ 为智能体在状态 s_t 下选择动作 $f(\varepsilon_t; s_t)$ 预期回报； $f(\cdot)$ 为重参数化技巧生成的随机动作； $\pi_\varphi(a|s)$ 为智能体在策略 π 时状态 s 下选择动作 a 的策略函数，策略 π 由参数 φ 决定； $a_t = f_\varphi(\varepsilon_t; s_t) = \mu_\varphi(s_t) + \varepsilon_t \sigma_\varphi(s_t)$ ， $\mu_\varphi(\cdot)$ 和 $\sigma_\varphi(\cdot)$ 分别为策略网络输出的均值函数和标准差函数，策略网络由参数 φ 决定。

2 基于 HDS-SAC 的 AGC 设计

HDS-SAC 由智能体、状态、动作、奖励和环境组成，其中 AGC 控制器作为智能体，而环境为分布式电力系统。智能体通过执行动作与环境进行交互。本文设置的动作集为 $A = [-50, 50]$ MW。

对于规模化新能源接入的分布式多区域互联电力系统，智能体在与环境互动过程中会生成新的状态信息，即区域控制误差(area control error, ACE)，并将其作为输入状态进行处理。在此架构下，智能体根据动作和环境反馈持续更新状态。

本文将新型电力系统划分为多个区域，采用分布式控制架构^[30]，每个区域由独立的智能体控制，智能体之间保持平等协作关系。这种多智能体系统(multi-agent system, MAS)通过智能体间的信息共享实现动作策略的更新。所有区域的智能体使用相同的奖励函数，意味着它们的目标一致，构成了一个典型的完全合作任务。

2.1 奖励函数

控制器的性能通过北美电力可靠性委员会的 CPS^[31] 标准来评估，HDS-SAC 控制器应保持 CPS 指标、频率($\Delta f/\text{Hz}$)指标、区域控制偏差(area control error, ACE)稳定。CPS 指标分为 CPS1、CPS2、CPS3。其中 CPS1 用于定量表示控制区域对其所属的互联电网可靠性的贡献，CPS2 用于区域电网在较长时间尺度内对 ACE 的控制精度的评估，CPS3 为 CPS1 和 CPS2 的加权值，用于整体评估区域电网的控制性能。

为保证分布式多区域电力系统的平稳运行且 CPS 的长期收益最大，奖励函数表达式如式(11)所示。

$$r(t) = -[\delta(t)^2 + 100(\omega(t) - 200)^2] \quad (11)$$

式中： $r(t)$ 为第 t 次迭代的奖励值； $\delta(t)$ 为 CPS1 第 t 次迭代时刻的瞬时值； $\omega(t)$ 为 ACE 在第 t 次迭代的瞬时值。

2.2 参数设置

1) 熵正则化系数 α ($0 < \alpha < 1$): 决定算法探索程度。 α 越大越有助于探索新状态空间; α 越小则会导致收敛至次优策略, 本文 α 取 1。

2) 折扣因子 γ ($0 < \gamma < 1$): 调节算法对即时奖励与未来奖励的权衡。 γ 接近 1 时算法考虑长期收益; γ 接近 0 时算法偏向短期奖励, 易学到次优策略, 本文 γ 取 0.99。

3) 学习率 ν ($0 < \nu < 1$): 决定算法在学习过程中参数更新的速度。 ν 越高时算法会出现振荡, 难以收敛至最优解; ν 越低则会导致学习进程变慢。本文 ν 取 0.000 001。

4) 策略网络的学习率 ν_{actor} ($0 < \nu_{\text{actor}} < 1$): 决定策略网络在梯度下降过程中参数更新步长大小。 ν_{actor} 越大则学习速度越快; ν_{actor} 越小则学习速度越慢。本文 ν_{actor} 取 0.000 01。

5) 价值网络的学习率 ν_{critic} ($0 < \nu_{\text{critic}} < 1$): 决定每次迭代中参数更新幅度。 ν_{critic} 越高则价值函数更新速度越快, 易适应环境变化。但价值函数会在最优解附近震荡, 难以收敛。 ν_{critic} 越小将可避免过度震荡, 但价值函数的更新速度变慢, 学习时间增加。本文 ν_{critic} 取 0.001。

6) 经验回放池的大小 R : R 越大, 存储经验越多, 有助智能体学习复杂策略; R 越小, 智能体无法充分利用历史经验, 易导致学习效率降低。本文 R 取 1 000 000。

7) 软更新参数 ε ($0 < \varepsilon < 1$): 决定目标网络参数向主网络参数靠拢的速度。 ε 越大, 参数更新越慢, 但能增强学习的稳定性; ε 越小, 参数更新越迅速, 能加速学习过程, 但会导致学习过程中的震荡和不稳定。本文 ε 取 0.005。

8) 最小批样本数 τ : 影响训练效率和效果。 τ 越小, 可加快模型更新频率, 但会导致训练过程出现高方差; τ 越大, 会导致计算效率降低, 会导致模型陷入局部最优。本文 τ 取 64。

2.3 算法框架图

HSD-SAC 从策略网络输出的均值和标准差, 经过高斯分布采样输出动作, 即机组所需的功率值。电力系统则根据该动作返回新的状态和奖励。智能体将状态转换、动作、奖励以及新状态的组合存储在经验回放池中。整体框架见图 1。

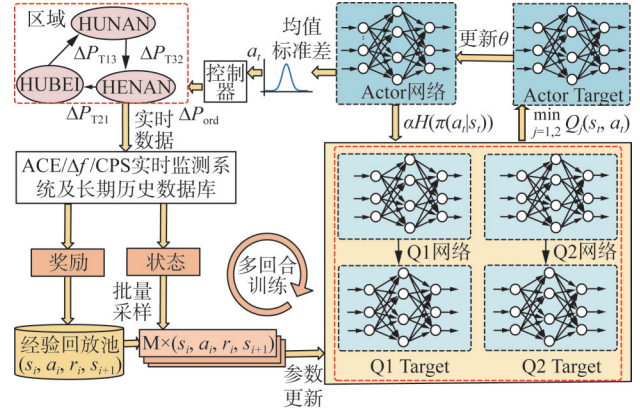


图1 HDS-SAC算法整体框架图

Fig. 1 Overall frame diagram of HDS-SAC algorithm

3 算例分析

3.1 改进的 IEEE 标准两区域 LFC 模型

当风电大规模接入电网时, 其有功功率扰动必会引起电网频率的波动。而储能技术^[32]因机组启停迅速、运行灵活且能降低系统峰谷差, 能有效应对电力系统频率波动。因此, 本文在 IEEE 标准两区域 LFC 模型^[33]的基础上, 搭建了包含火电^[34]、水电^[35]、风电^[36]、抽水蓄能^[37]、超导磁储能^[38]的 LFC 模型, 相关参数参考文献[39-40]。并对基于 HDS-SAC 的 AGC 控制器进行仿真分析, 控制器的采样时间为 4 s, 模型详见图 2。

仿真算例均在 MATLAB R2022b 环境下运行, HDS-SAC 算法采用 Python3.9 编写, 仿真采用的个人计算机的主要参数分别为: 1) CPU: AMD Ryzen 7 5800 H; 2) 主频: 3.2 GHz; 3) 内存: 32 GB。

火电机组模型如图 3 所示, T_{fg} 为火电机组调速器时延常数, 本文取 0.08 s; T_n 为火电机组时间常数, 取 0.03 s。

水电机组模型如图 4 所示, T_{gw} 为水电机组调速器时间常数, 本文取 0.041 s; T_{rs} 、 T_{rh} 分别为水轮机时间常数, 本文分别取 -1 s 和 0.5 s; T_{wls} 为水锤时间常数, 本文取 0.513 s。

简化后的风机模型如式(12)所示, v_w 为风机轮毂最高处风速, $v_{\min}$ 、 v_e 、 $v_{\max}$ 分别为风机的切入风速、额定风速、切除风速, P_e 为风机额定输出功率, $P_w(v_w)$ 为风机在风速 v_w 下输出总功率。

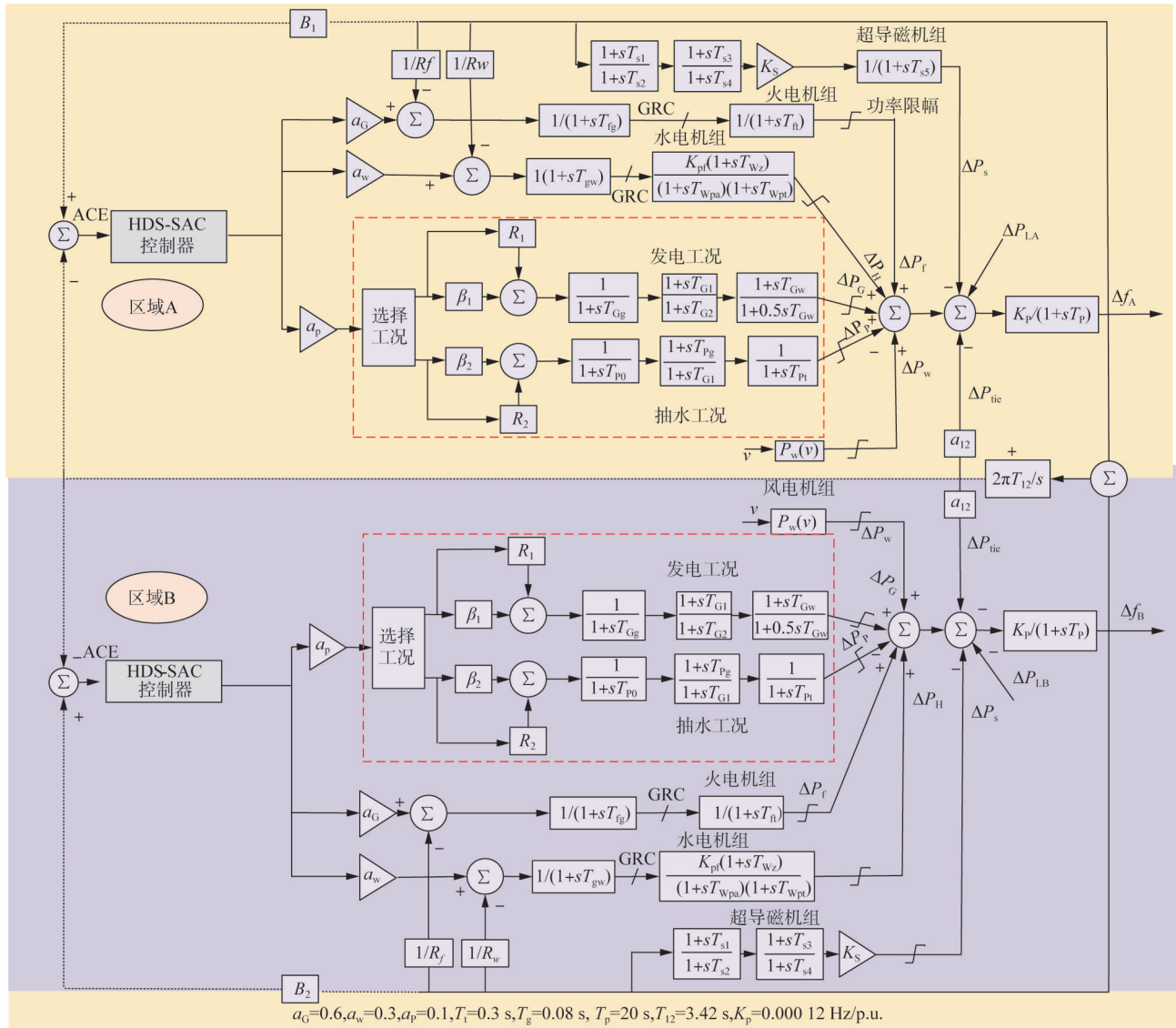


图2 IEEE标准两区域LFC模型

Fig. 2 LFC models in IEEE standard two-region

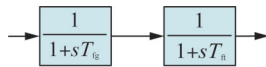


图3 火电机组模型

Fig. 3 Thermal power units model

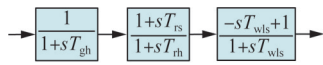


图4 水电机组模型

Fig. 4 Hydropower units model

$$P_w(v_w) = \begin{cases} 0, & 0 \leq v_w \leq v_{\min} \\ P_e \frac{v_w^2 - v_{\min}^2}{v_e^2 - v_{\min}^2}, & v_{\min} \leq v_w \leq v_e \\ P_e, & v_e \leq v_w \leq v_{\max} \\ 0, & v_{\max} < v_w \end{cases} \quad (12)$$

当抽水蓄能机组工作时，其运行可分为发电运行和抽水运行两种方式。抽水蓄能等值模型运行过程见图5所示。

图5中， β 为偏差系数，其中 $\beta_1=\beta_2=\frac{20}{22}$ ； R 为下

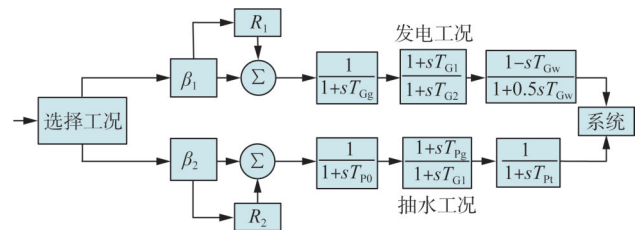


图5 抽水蓄能模型的运行过程

Fig. 5 Operation process of pumped storage model

垂系数, $R_1=R_2=0.05$; 发电工况下: T_{Gw} 为水启动时间常数, 本文取 0.2 s ; T_{G1} 为复位参数, 本文取 0.6 s ; T_{G2} 为补偿器参数, 本文取 10 s 。抽水工况: T_{P0} 为调速系数, 本文取 0.3 ; T_{Pg} 为电机增益系数, 本文取 0.8 ; T_{pi} 为惯性系数, 本文取 0.8 ; P_G 为发电工况; P_p 为抽水工况。

超导磁储能模型如图6所示, T_{s1} 、 T_{s2} 、 T_{s3} 、 T_{s4} 和 T_{s5} 为机组时间常数, 分别取 0.1 s 、 0.0665 s 、 0.0661 s 、 0.099 s 、 0.001 s ; K_s 为机组增益系数, 本文取 0.0665 。

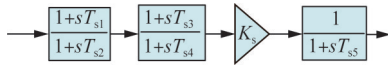


图6 超导磁储能模型

Fig. 6 Superconducting magnetic energy storage model

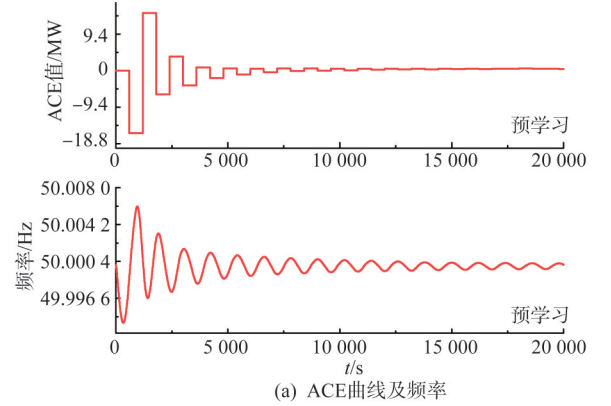
3.1.1 预学习

强化学习的核心机制为试错, HDS-SAC控制器在实际应用前须通过广泛试错完成预学习以获得最优的控制策略。本文将正弦负荷扰动(周期为 1200 s , 幅值为 50 MW , 时长为 20000 s)作为 HDS-SAC 控制器的学习样本。图7为 HDS-SAC 控制器预学习效果。

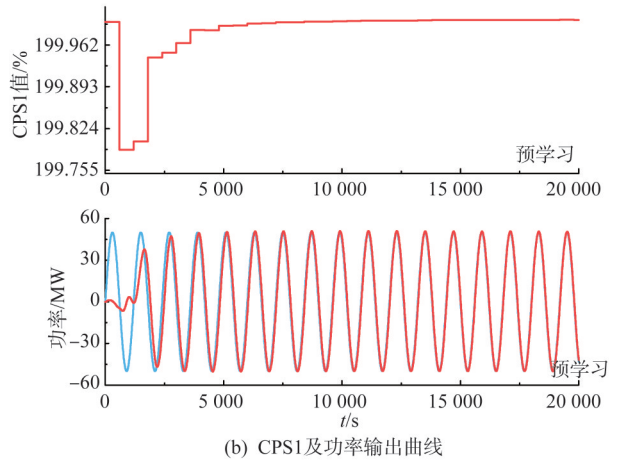
图7展示的 HDS-SAC 控制器的预训练过程表明, 经过大约 2700 s 的学习后控制器已能稳定响应正弦负荷的变化。HDS-SAC 控制器能快速地将系统频率偏差控制在 0.0055 Hz 范围内, 在 7200 s 后频率偏差降至 0.0007 Hz 以下, 其远小于电网的标准偏差 $\pm 0.2\text{ Hz}$ 。图7(a)所示 $|ACE|$ (ACE 绝对值的平均值) 最大误差为 15 MW , 但在经过 6000 s 之后控制器将 ACE 指标稳定在 $\pm 1\text{ MW}$ 之间; CPS1 (每 10 min CPS1 的平均值) 性能指标最大误差为 0.2% , 在 6600 s 后 CPS1 性能指标维持在 200% 以上。经分析, 本文所提的 HDS-SAC 控制器在预学习阶段指标已达要求, 可将其投入在线运行。

3.1.2 收敛性分析

为验证所提算法是否具备快速收敛特性, 将 HDS-SAC、SAC、(deep Q network, DQN)^[40]、(repeated update Q-Learning using lagrangian relaxation, RUQL-LR)^[41] 算法分别进行预学习(周期为 30000 s), 利用奖励函数进行对比分析。根据式(11)可知, 当奖励函数为 0 时可认为算法已达到最优。由图8可知, 经过约 2500 s 后 HDS-SAC 收



(a) ACE曲线及频率



(b) CPS1及功率输出曲线

图7 HDS-SAC 预学习效果

Fig. 7 Pre-learning effects of HDS-SAC

敛到最优, 相较于其他算法具有明显优势。

3.2 华中三区域互联 LFC 模型

本文所搭建的综合能源系统包括传统火电、水

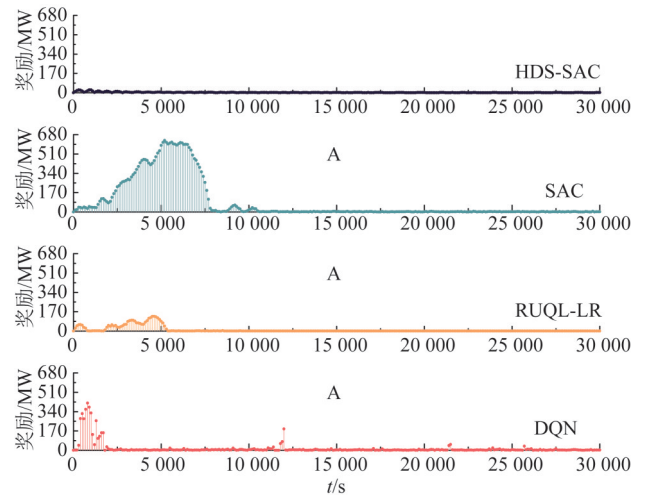


图8 预学习阶段的4种算法收敛特性分析

Fig. 8 Analysis of convergence characteristics of four algorithms in pre-learning stage

电机组、分布式能源^[42](风电机组、抽水蓄能机组)的华中三区域LFC模型见图9。

3.2.1 方波扰动

采用幅值为1 000 MW、周期为4 000 s、总时长为30 000 s的方波负荷扰动实验进行仿真,模拟

电网中单位时间内负荷连续突增突减情况,并与SAC、DQN、RUQL-LR3种不同算法进行对比。图10为方波扰动下4种算法指标对比。从图10(a)~(c)可知,HDS-SAC能够使 $|ACE|$ 的平均值下降67.29%~94.41%, $|\Delta f|$ (频率偏差绝对值的平均

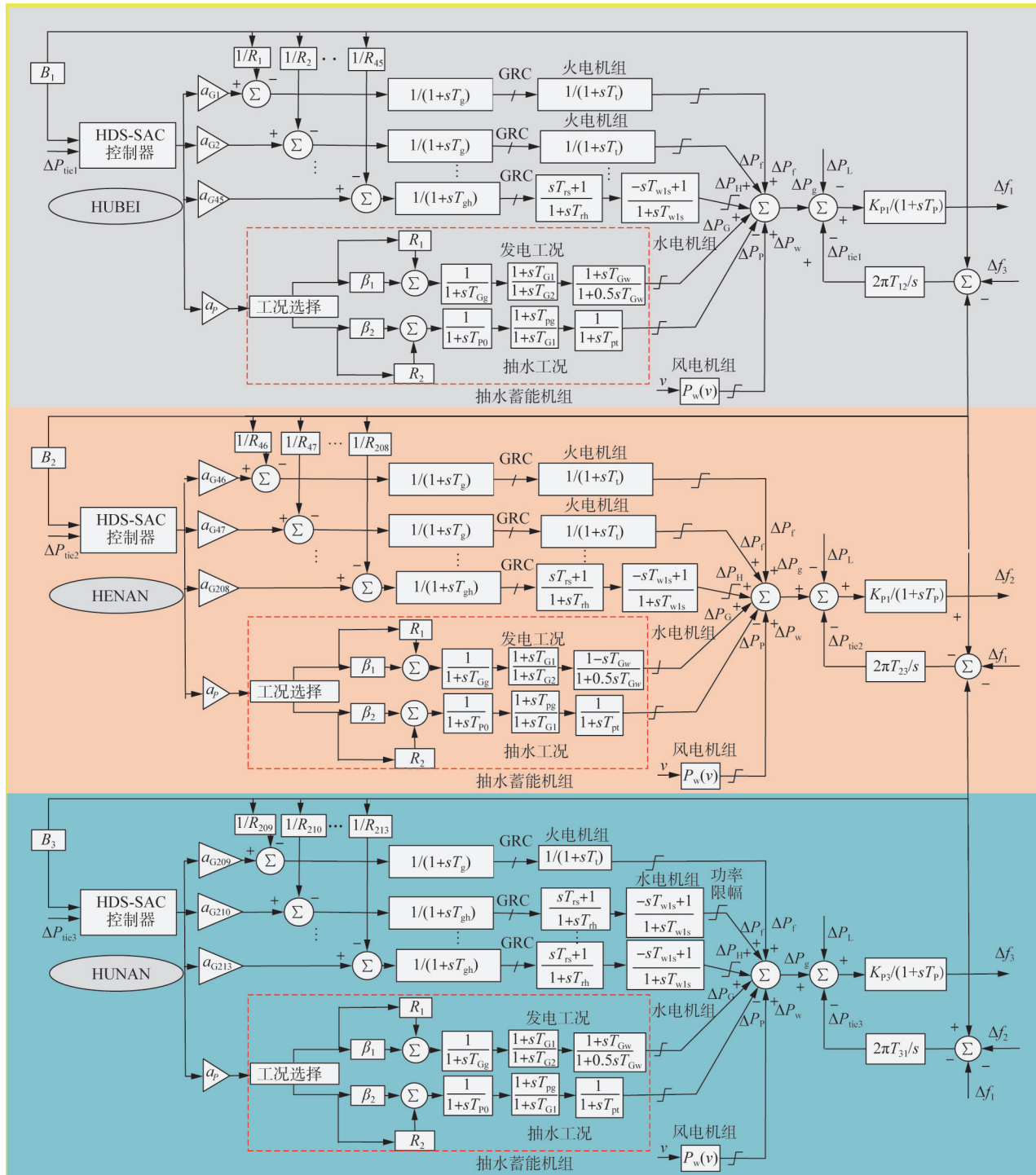


图9 华中电网三区域LFC模型

Fig. 9 Three-region LFC model in Central China power grid

值)的平均值降低47%~72%。在数据量化方面,HDS-SAC的标准偏差最小,比SAC降低77.41%,比RUQL-LR降低77.41%,比DQN降低85.77%。从图10(d)可知,HDS-SAC在电网发生负荷突增突减情况时能迅速调整发电机组出力跟随负荷扰动,HDS-SAC恢复稳态后不会产生振荡现象。相较于其他算法,HDS-SAC控制性能更优,可靠性更高。

3.2.2 阶跃扰动

采用阶跃负荷扰动(幅值1 000 MW、阶跃时间1 000 s、时长30 000 s)来模拟单位时间内负荷突增情况,以此来评估HDS-SAC、SAC、RUQL-LR、DQN这4种算法在面对负荷突增的快速反应能力。图11为4种算法的控制性能指标对比。由图11(a)所知,HDS-SAC可在电网发生突增负荷的情况下可迅速增加发电机组出力并在阶跃后稳定运行,不产生震荡现象。其超调量为0.023 7%,调整时间为98 s,峰值时间为83 s。由图11(b)可知,在负荷激增时HDS-SAC过渡过程时间最短,且频率阶跃后保持50 Hz不变。

表1为阶跃扰动下4种算法的指标对比。由表1可知,HDS-SAC能够使IACEI的平均值减少49.73%~89.23%;而4种算法的CPS1的平均值分别为199.950 2%、199.836 7%、199.883 8%、199.834 5%,HDS-SAC能够达到考核指标且超调量更低。

3.2.3 随机方波扰动

由于分布式能源所固有的间歇性和不可预测性,本文采用随机方波扰动(幅值1 000 MW、时长30 000 s)的情景来评估所提及的4种算法在复杂环境下负荷扰动无规律突增突减的适应效果。各算法指标对比如图12所示。各算法的 $|\Delta f|$ 平均值分别为0.001 7 Hz、0.005 9 Hz、0.003 8 Hz、0.013 2 Hz。与之相比,HDS-SAC的 $|\Delta f|$ 降低55.26%~87.12%,各算法的IACEI平均值分别为6.966 1 MW、24.656 6 MW、12.692 1 MW、50.186 3 MW,HDS-SAC的IACEI降低45.11%~86.12%。各算法的CPS 1平均值分别为199.345 9%、197.015 8%、198.350 9%、192.635 9%。HDS-SAC的标准误差比SAC降低90.32%、比RUQL-LR降低81.46%,比DQN降低93.77%。相较于其他算法,HDS-SAC的可靠性更高。

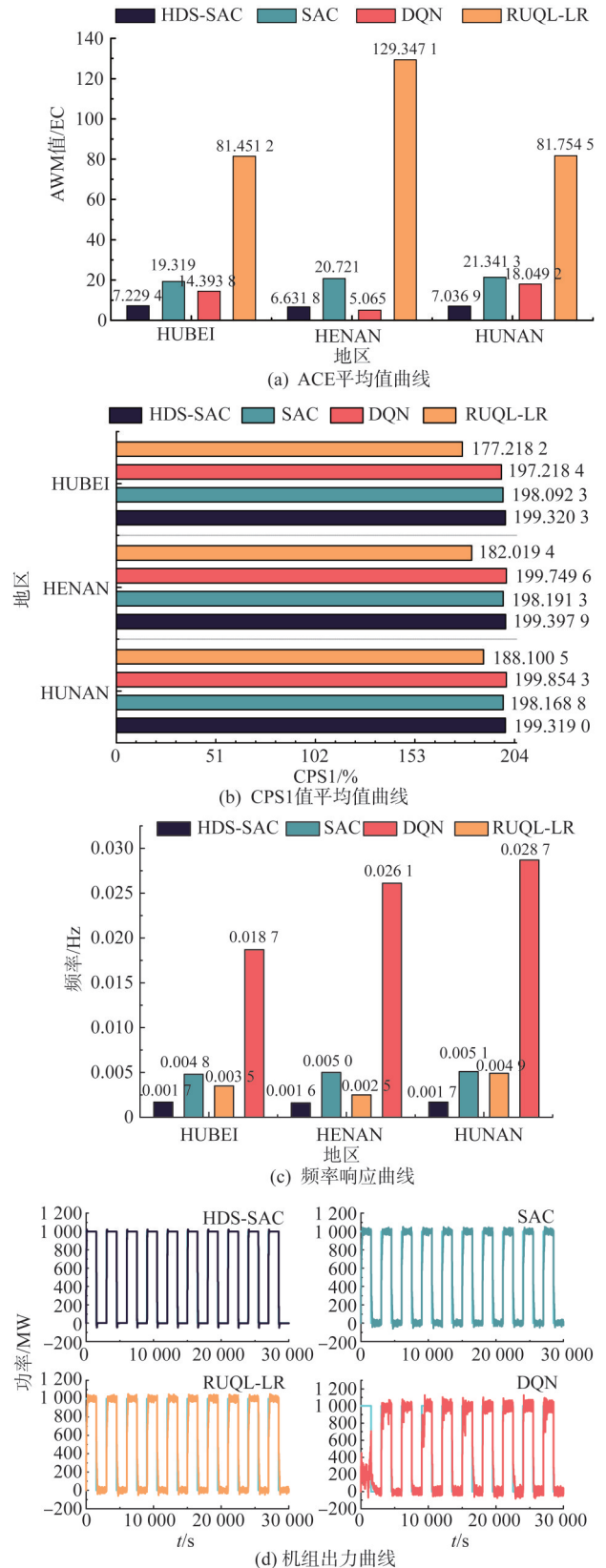


图10 方波扰动下不同算法的各项指标对比

Fig. 10 Comparison of various indexes of different algorithms under square waves perturbations

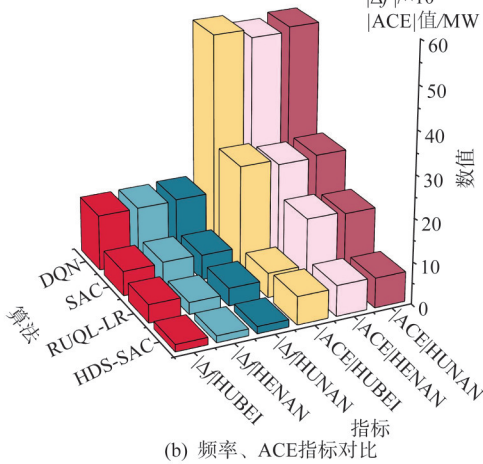
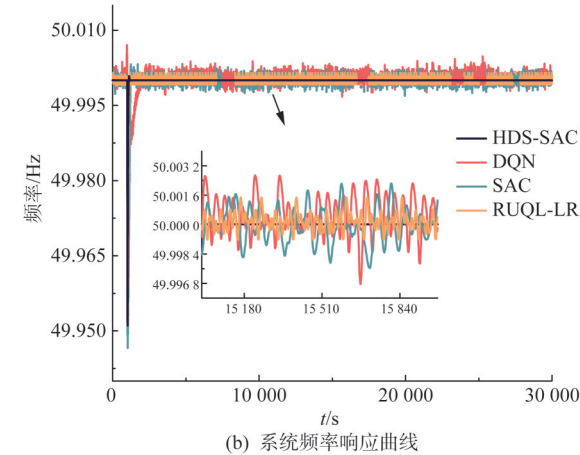
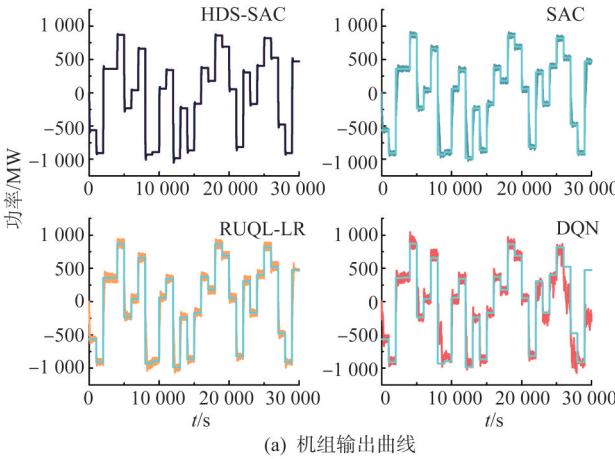
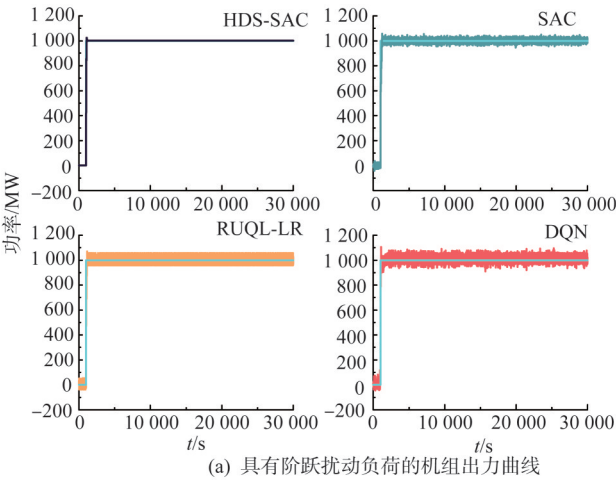


图 11 阶跃扰动下不同算法的各项指标对比
Fig. 11 Comparison of various indexes of different algorithms under step perturbations

表 1 阶跃扰动下不同算法的各项指标对比
Tab. 1 Comparison of various indexes of different algorithms under step perturbations

区域	算法	$ \Delta f /\text{Hz}$	$ \text{ACE} /\text{MW}$	CPS1 值/%	CPS2 值/%	CPS3 值/%
湖北	HDS-SAC	0.000 14	0.566 9	199.950 2	99.65	100
	SAC	0.001 00	4.325 1	199.836 7	99.02	100
	DQN	0.009 80	4.490 8	199.943 9	99.42	100
	RUQL-LR	0.000 61	3.193 7	199.918 0	99.40	100
河南	HDS-SAC	0.000 13	0.5960	199.955 5	99.64	100
	SAC	0.001 20	5.034 4	199.870 3	99.15	100
	DQN	0.030 10	157.72	199.952 3	67.76	100
	RUQL-LR	0.000 41	1.127 7	199.966 0	99.34	100
湖南	HDS-SAC	0.000 14	0.610 8	199.950 3	99.65	100
	SAC	0.001 30	5.265 8	199.853 9	99.10	100
	DQN	0.009 80	3.550 3	199.956 9	99.31	100
	RUQL-LR	0.000 48	1.764 8	199.834 5	99.46	100

图 12 随机方波扰动下 4 种算法指标对比
Fig. 12 Comparison of four algorithm metrics under random square wave disturbances

3.2.4 计算速度分析

为验证所提算法是否相较于 Q 衍生算法能快速适应环境及评估算法在面对计算量激增的表现，引入以下阶跃、方波、随机方波、变正弦、白噪声等扰动实验，以仿真时间 30 000 s 为考核标准记录算

法的真实仿真时间,实验时间如表2所示。从表2可以看出,在面对状态或动作空间维度增加而导致计算量增大时,所提算法在保证控制精度的前提下能大幅度提升计算速度。

表2 算法计算速度对比

Tab. 2 Comparison of algorithm computation speed

算法	HDS-SAC	SAC	RUQL-LR	DQN
阶跃扰动	155	216	187	150
方波扰动	161	235	184	172
随机方波扰动	161	219	178	153
变正弦扰动	151	237	192	164
白噪声扰动	150	224	190	152

3.2.5 随机白噪声扰动

针对大规模新能源并网下电力系统负荷每一时刻均在随机变化的情况,引入随机白噪声(幅值1 000 MW、时长30 000 s)负荷扰动,以仿真实验HDS-SAC控制器的性能。负荷跟踪曲线如图13所示。其他指标曲线如图14所示。图14中,HDS-SAC的频率平均值最小,相较于其他3种算法,HDS-SAC可减少 Δf 平均值62.92%~74.62%, $|\text{ACE}|$ 平均值减少62.50%~80.04%,显示出其在应对随机负荷扰动时的高鲁棒性。HDS-SAC的标准误差比SAC降低64.94%、比RUQL-LR降低70.17%,比DQN降低80.48%。相较于其他算法,HDS-SAC的可靠性高,跟踪过程更稳定。

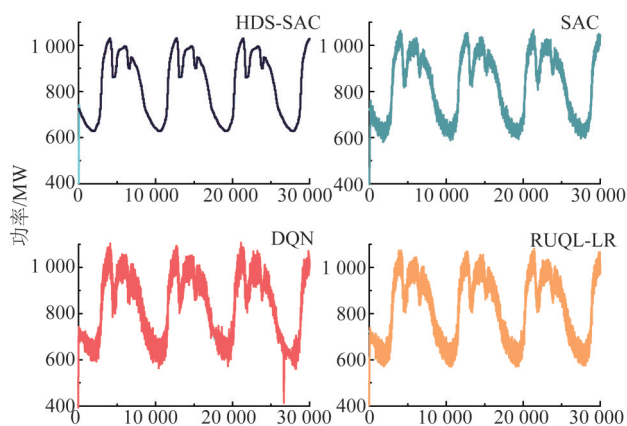


图13 随机白噪声扰动下4种算法的输出曲线

Fig. 13 Output curves of four algorithms under random white noise perturbations

3.2.6 变幅值正弦扰动

为了验证所提出算法在实时应对变化条件下的学习能力,并且评价其收敛性,引入变幅值正弦负

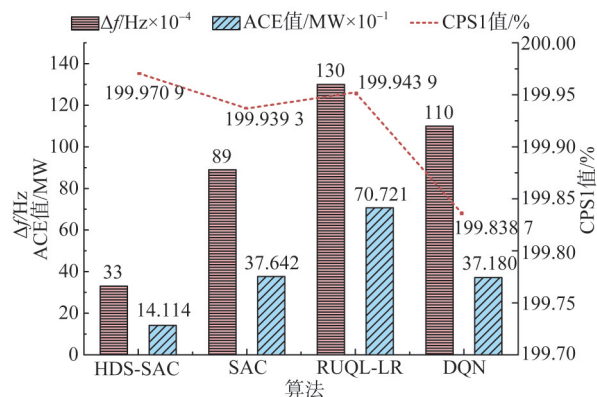


图14 随机白噪声扰动下算法性能指标对比

Fig. 14 Comparison of algorithm performance indexes under random white noise disturbances

荷扰动。图15为各算法的指标对比。图16为各算法的输出曲线。

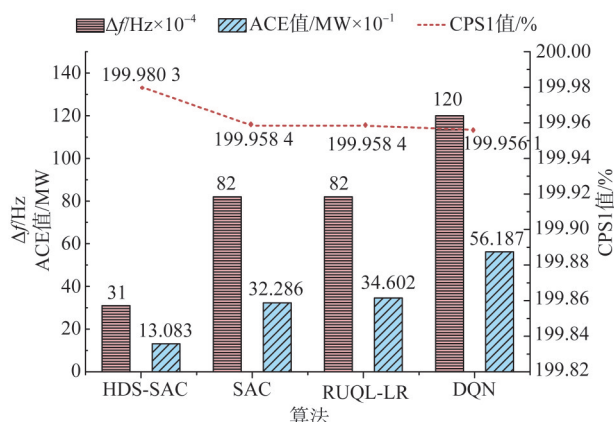


图15 变幅值正弦波扰动下算法性能指标对比

Fig. 15 Comparison of algorithm performance indexes under varying amplitudes sine wave disturbances

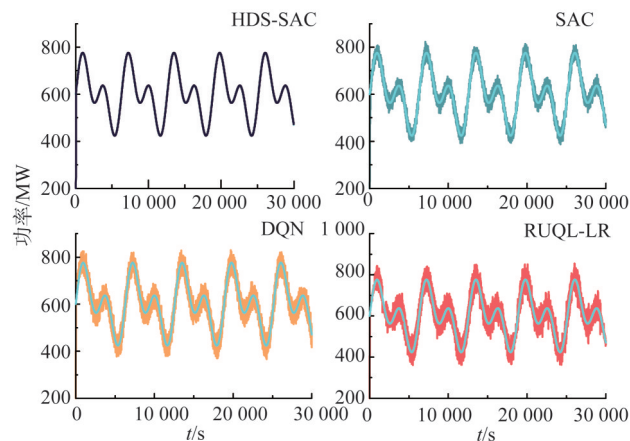


图16 变幅值正弦波扰动下4种算法的输出曲线

Fig. 16 Output curves of four algorithms under varying amplitudes sine wave perturbations

结果表明, HDS-SAC 展现出更优越的在线学习性能, 能够迅速适应负荷的波动。HDS-SAC 较其他算法收敛更快。综合分析, HDS-SAC 的 $|\Delta f|$ 平均值降低 62.20 %~74.17 %、 $|\Delta CEI|$ 的平均值降低 59.48 %~76.72 %。HDS-SAC 的标准误差比 SAC 降低 64.94 %、比 RUQL-LR 降低 70.17 %, 比 DQN 降低 80.48 %。综上, 相比于其他 3 种算法, HDS-SAC 的精度和可靠性更高。

4 结语

为解决具有强随机性的高比例大容量新能源接入电网带来的强随机扰动导致控制性能变差及频率不稳定问题, 本文提出了基于 HDS-SAC 的 AGC 算法。通过在最大熵框架下鼓励智能体进行随机探索, 解决了 Q 学习及其衍生算法无法快速更新 Q 表以适应环境变化的问题; 同时利用高斯分布策略生成连续动作值, 使算法可在高维连续状态空间中寻找协同最优解, 解决了传统强化学习高维“状态-动作”的维数灾问题。

通过在改进的 IEEE 两区域及华中三区域 LFC 模型上引入了多样化的负荷扰动进行仿真, 验证了所提算法所具有的连续控制特性保证了其能够快速精确地跟踪负荷扰动, 并有效减少了系统的超调现象。面对强随机扰动, 所提算法表现出了其他算法无法比拟的在线适应和学习能力, 特别是在应对变幅值正弦扰动时其机组出力曲线相比较其他算法更为平滑, 超调量更小。

然而通过引入目标熵所提算法需要在不同环境下细致地调整参数。为了优化算法在各种复杂环境中的表现, 未来的研究将集中于开发一种自适应的超参数调整机制以保证该算法能适应更复杂的环境。

参考文献

- [1] 张孝顺, 郑理民, 余涛. 基于多步回溯 $Q(\lambda)$ 学习的电网多目标最优碳流算法[J]. 电力系统自动化, 2014, 38(17): 118-123.
ZHANG Xiaoshun, ZHENG Limin, YU Tao. Multi-objective optimal carbon emission flow calculation of power grid based on multi step $Q(\lambda)$ learning algorithm [J]. Automation of Electric Power Systems, 2014, 38(17): 118-123.
- [2] 张孝顺, 余涛. 互联网网自动发电控制功率分配的改进逼近于理想解的排序-Q 多目标优化算法[J]. 控制理论与应用, 2015, 32(4): 497-503.
ZHANG Xiaoshun, YU Tao. Stochastic optimal generation command dispatch of interconnected power grids based on improved multi-objective technique for order preference similar to an ideal solution-Q algorithm[J]. Control Theory & Applications, 2015, 32(4): 497-503.
- [3] 李捷, 余涛, 潘振宁. 基于强化学习的增量配电网实时随机调度方法[J]. 电网技术, 2020, 44(9): 3321-3332.
LI Jie, YU Tao, PAN Zhenning. Real-time stochastic dispatch method for incremental distribution network based on reinforcement learning [J]. Power System Technology, 2020, 44(9): 3321-3332.
- [4] 陈俊斌, 余涛, 殷林飞, 等. 基于极限动态规划算法的微电网一体化调度与控制[J]. 控制理论与应用, 2019, 36(10): 1698-1706.
CHEN Junbin, YU Tao, YIN Linfei, et al. Integrated dispatch and control of microgrid based on extreme dynamic programming algorithm [J]. Control Theory & Applications, 2019, 36(10): 1698-1706.
- [5] 余涛, 梁海华, 周斌. 基于 $R(\lambda)$ 学习的孤岛微电网智能发电控制[J]. 电力系统保护与控制, 2012, 40(13): 7-13.
YU Tao, LIANG Haihua, ZHOU Bin. Smart power generation control for microgrids islanded operation based on $R(\lambda)$ learning [J]. Power System Protection and Control, 2012, 40(13): 7-13.
- [6] 余涛, 周斌, 陈家荣. 基于 Q 学习的互联网网动态最优 CPS 控制[J]. 中国电机工程学报, 2009, 29(19): 13-19.
YU Tao, ZHOU Bin, CHEN Jiarong. Q-learning based dynamic optimal cps control methodology for interconnected power systems [J]. Proceedings of the CSEE, 2009, 29(19): 13-19.
- [7] SHI Jiachuan, ZHAO Yujin, PENG Bo, et al. A coordinated fuzzy-based frequency control strategy of wind-storage system [C]// 2022 5th International Conference on Energy, Electrical and Power Engineering (CEEPE), April 22-24, 2022, Chongqing, China. New York: IEEE, 2022: 1017-1022.
- [8] 殷林飞, 余涛, 张泽宇, 等. 基于深度自适应动态规划的孤岛主动配电网发电控制与优化一体化算法[J]. 控制理论与应用, 2018, 35(2): 169-183.
YIN Linfei, YU Tao, ZHANG Zeyu, et al. Deep adaptive dynamic programming based integration algorithm for generation control and optimization of islanded active distribution network [J]. Control Theory & Applications, 2018, 35(2): 169-183.
- [9] 张孝顺, 余涛, 唐捷. 基于 CEQ(λ) 多智能体协同学习的互联网网性能标准控制指令动态分配优化算法[J]. 电工技术学报, 2016, 31(8): 125-133.
ZHANG Xiaoshun, YU Tao, TANG Jie. Dynamic optimal allocation algorithm for control performance standard order of interconnected power grids using synergetic learning of multi-agent CEQ(λ) [J]. Transactions of China Electrotechnical Society, 2016, 31(8): 125-133.
- [10] 唐捷, 张泽宇, 程乐峰, 等. 基于 CEQ(λ) 强化学习算法的微电网智能发电控制[J]. 电测与仪表, 2017, 54(1): 39-45.
TANG Jie, ZHANG Zeyu, CHENG Lefeng, et al. Smart generation control for micro-grids based on correlated equilibrium $Q(\lambda)$ learning algorithm [J]. Electrical Measurement & Instrumentation, 2017, 54(1): 39-45.
- [11] XI Lei, LI Haokai, ZHU Jizhong, et al. A novel automatic generation control method based on the large-scale electric vehicles and wind power integration into the grid. [J]. IEEE

- Transactions on Neural Networks and Learning Systems, 2022, 35(5): 5824 – 5834.
- [12] YU Tao, XI Lei, YANG Bo, et al. Multiagent stochastic dynamic game for smart generation control [J]. Journal of Energy Engineering, 2016, 142(1): 4 – 15.
- [13] 殷林飞, 余涛. 基于深度Q学习的强鲁棒性智能发电控制器设计[J]. 电力自动化设备, 2018, 38(5): 12 – 19.
YIN Linfei, YU Tao. Design of strong robust smart generation controller based on deep Q learning [J]. Electric Power Automation Equipment, 2018, 38(5): 12 – 19.
- [14] 陈艺璇, 张孝顺, 郭乐欣, 等. 基于多智能体迁移强化学习算法的电力系统最优碳-能复合流求解[J]. 高电压技术, 2019, 45(3): 863 – 872.
CHEN Yixuan, ZHANG Xiaoshun, GUO Lexin, et al. Optimal carbon-energy combined flow in power system based on multi-agent transfer reinforcement learning [J]. High Voltage Engineering, 2019, 45(3): 863 – 872.
- [15] 张孝顺, 李清, 余涛, 等. 基于协同一致性迁移Q学习算法的虚拟发电部落AGC功率动态分配[J]. 中国电机工程学报, 2017, 37(5): 1455 – 1467.
ZHANG Xiaoshun, LI Qing, YU Tao, et al. Collaborative consensus transfer Q-learning based dynamic generation dispatch of automatic generation control with virtual generation tribe [J]. Proceedings of the CSEE, 2017, 37(5): 1455 – 1467.
- [16] 张孝顺, 余涛. 互联网AGC功率动态分配的虚拟发电部落协同一致性算法[J]. 中国电机工程学报, 2015, 35(15): 3750 – 3759.
ZHANG Xiaoshun, YU Tao. Virtual generation tribe based collaborative consensus algorithm for dynamic generation dispatch of AGC in interconnected power grid [J]. Proceedings of the CSEE, 2015, 35(15): 3750 – 3759.
- [17] 潘振宁, 张孝顺, 余涛, 等. 大规模电动汽车集群分层实时优化调度[J]. 电力系统自动化, 2017, 41(16): 96 – 104.
PAN Zhenning, ZHANG Xiaoshun, YU Tao, et al. Hierarchical real-time optimized dispatching for large-scale clusters of electric vehicles [J]. Automation of Electric Power Systems, 2017, 41(16): 96 – 104.
- [18] 陈艺璇, 张孝顺, 余涛. 基于纳什均衡迁移学习的碳-能复合流自律优化[J]. 控制理论与应用, 2018, 35(5): 668 – 681.
CHEN Yixuan, ZHANG Xiaoshun, YU Tao. Nash equilibrium inspired transfer learning for self-organizing optimal carbon-energy combined-flow [J]. Control Theory & Applications, 2018, 35(5): 668 – 681.
- [19] 王德志, 张孝顺, 刘前进, 等. 基于集成学习的孤岛微电网源-荷协同频率控制[J]. 电力系统自动化, 2018, 42(10): 46 – 52.
WANG Dezhi, ZHANG Xiaoshun, LIU Qianjin, et al. Ensemble learning for generation-consumption coordinated frequency control in an islanded microgrid [J]. Automation of Electric Power Systems, 2018, 42(10): 46 – 52.
- [20] DUAN Jingliang, YANG Guan, LI Shengbo, et al. Distributional soft actor-critic: off-policy reinforcement learning for addressing value estimation errors. [J] IEEE Transactions on Neural Networks and Learning Systems, 2020, 33(11): 6584 – 6598.
- [21] 王长春, 王果, 赵倩宇, 等. 考虑分时电价和充电利用率特征的大型电动汽车充电桩负荷短期预测方法[J]. 南方电网技术, 2024, 18(5): 75 – 84.
- WANG Changchun, WANG Guo, ZHAO Qianyu. A short-term load forecasting method for large scale electric vehicle charging stations considering characteristics of charging utilization rate and time-of-use electricity price [J]. Southern Power System Technology, 2024, 18(5): 75 – 84.
- [22] 张孝顺, 谭恬, 蒙蝶, 等. 基于光伏系统的动态代理模型最大功率点跟踪算法研究[J]. 综合智慧能源, 2021, 43(8): 1 – 10.
ZHANG Xiaoshun, TAN Tian, MENG Die, et al. Study on dynamic surrogate model for MPPT of PV systems [J]. Integrated Intelligent Energy, 2021, 43(8): 1 – 10.
- [23] 黄琳妮, 郑宝敏, 殷林飞. 基于ADMM的电-气综合能源系统多目标最优能量流分布式计算[J]. 广东电力, 2020, 33(10): 56 – 63.
HUANG Linni, ZHENG Baomin, YIN Linfei. Multi-objective optimal energy flow distributed calculation of electric-gas integrated energy system based on ADMM [J]. Guangdong Electric Power, 2020, 33(10): 56 – 63.
- [24] YU Tao, WANG Yangming, YE Wenjie, et al. Stochastic optimal generation command dispatch based on improved hierarchical reinforcement learning approach [J]. IET generation, transmission & distribution, 2011, 5(8): 789 – 797.
- [25] CHERIDITO P, JENTZEN A, ROSSMANNEK F. Efficient approximation of high-dimensional functions with neural networks [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(7): 3079 – 3093.
- [26] XIAO Qianhao, JIANG Li, WANG Manman, et al. An improved distributed sampling PPO algorithm based on beta policy for continuous global path planning scheme [J]. Sensors, 2023, 23(13): 6101 – 6125.
- [27] AI-DIGS A, CHEN Y. Generation and load balance using linear quadratic Gaussian control [C]//2016 North American Power Symposium (NAPS). September 18 – 20, 2016, Denver, CO, USA. New York: IEEE, 2016: 1 – 5.
- [28] 解立辉, 席磊. 强化学习在自动发电控制中的研究进展与展望[J]. 三峡大学学报(自然科学版), 2023, 45(5): 133 – 141.
XIE Lihui, XI Lei. Research progress and prospects of reinforcement learning in automatic generation control [J]. Journal of China Three Gorges University (Natural Sciences), 2023, 45(5): 133 – 141.
- [29] 张孝顺, 李清, 余涛, 等. 基于协同一致性迁移Q学习算法的虚拟发电部落AGC功率动态分配[J]. 中国电机工程学报, 2017, 37(5): 1455 – 1467.
ZHANG Xiaoshun, LI Qing, YU Tao, et al. Collaborative consensus transfer Q-learning based dynamic generation dispatch of automatic generation control with virtual generation tribe [J]. Proceedings of the CSEE, 2017, 37(5): 1455 – 1467.
- [30] SUN Changjian, GAO Fang. A tent marine predators algorithm with estimation distribution algorithm and Gaussian random walk for continuous optimization problems [J]. Computational Intelligence and Neuroscience, 2021(1): 1155 – 1172.
- [31] 余涛, 程乐峰, 张孝顺. 基于信息-物理-社会系统融合和群体机器学习的弱中心化微元网: 理论研究与关键科学问题分析[J]. 中国科学: 技术科学, 2019, 49(12): 1541 – 1569.
YU Tao, CHENG Lefeng, ZHANG Xiaoshun. The weakly-centralized Web-of-Cels based on cyber-physical-social

- systems integration and group machine learning: theoretical investigations and key scientific issues analysis [J]. *Scientia Sinica(Technologica)*, 2019, 49(12): 1541 – 1569.
- [32] 张东辉, 康重庆, 卢洵, 等. 高比例新能源系统中储能配置规模论证[J]. *南方电网技术*, 2022, 16(4): 3 – 11.
- ZHANG Donghui, KANG Chongqing, LU Xun, et al. Demonstration on the scale of energy storage deployment in high-proportion new energy power system [J]. *Southern Power System Technology*, 2022, 16(4): 3 – 11.
- [33] 唐建林, 余涛, 张孝顺, 等. 基于多搜索器优化算法的含可再生能源协同优化调度[J]. *控制理论与应用*, 2020, 37(3): 492 – 504.
- TANG Jianlin, YU Tao, ZHANG Xiaoshun, et al. Collaborative optimization scheduling with renewable energy based on multi-searcher optimization algorithm [J]. *Control Theory & Applications*, 2020, 37(3): 492 – 504.
- [34] 毛森茂, 瞿凯平, 陈艺璇, 等. 基于灰狼多目标算法的电网碳-能复合流优化调度[J]. *新型工业化*, 2016, 6(9): 11 – 17.
- MAO Senmao, QU Kaiping, CHEN Yixuan, et al. Grey wolf multi-objective optimizer for optimal carbon-energy combined-flow[J]. *The Journal of New Industrialization*, 2016, 6(9): 11 – 17.
- [35] 丁理杰, 刘姗姗, 史华勃. 水电高占比电网中水轮机模型对频率振荡特性影响及其适应性分析[J]. *电力系统保护与控制*, 2021, 49(12): 174 – 180.
- DING Lijie, LIU Shanmei, SHI Huabo. Influence of a turbine model on frequency oscillation characteristics and its adaptability in a high hydro-ratio power grid [J]. *Power System Protection and Control*, 2021, 49(12): 174 – 180.
- [36] 杨蕾, 李胜男, 黄伟, 等. 考虑风光新能源参与二次调频的多源最优协同控制[J]. *电力系统保护与控制*, 2020, 48(19): 43 – 49.
- YANG Lei, LI Shengnan, HUANG Wei, et al. Optimal coordinated control of multi-source for AGC with participation of wind and solar energy [J]. *Power System Protection and Control*, 2020, 48(19): 43 – 49.
- [37] 赵熙临, 周红玉, 付波, 等. 一种用于微网调频的风电与抽水蓄能综合控制方法[J]. *河南理工大学学报(自然科学版)*, 2023, 42(4): 121 – 129.
- ZHAO Xilin, ZHOU Hongyu, FU Bo, et al. A comprehensive control method for wind power and pumped storage in the frequency regulation of microgrid [J]. *Journal of Henan Polytechnic University(Natural Science)*, 2023, 42(4): 121 – 129.
- [38] 金建勋, 陈孝元. 面向智能电网能量调控应用的超导磁储能技术: 理论模型、装置特性、研究现状和应用展望[J]. *南方电网技术*, 2015, 9(12): 44 – 57.
- JIN Jianxun, CHEN Xiaoyuan. Theoretical modeling, device characteristic, research status and application prospect of SMES technology for smart grid energy management [J]. *Southern Power System Technology*, 2015, 9(12): 44 – 57.
- [39] AHN S J, NAM S, CHOI J, et al. Power scheduling of distributed generators for economic and stable operation of a microgrid[J]. *IEEE Transactions on Smart Grid*, 2013, 4(1): 398 – 405.
- [40] 沈国辉, 赵荣生, 董晓, 等. 基于多信息交互与深度强化学习的电动汽车充电导航策略[J]. *南方电网技术*, 2022, 16(1): 108 – 116.
- SHEN Guohui, ZHAO Rongsheng, DONG Xiao, et al. Electric vehicle charging navigation strategy based on multi-information interaction and deep reinforcement learning [J]. *Southern Power System Technology*, 2022, 16(1): 108 – 116.
- [41] 席磊, 刘治洪, 李彦营. 基于拉格朗日松弛强化学习算法的自动发电控制[J]. *中国电机工程学报*, 2023, 43(4): 1359 – 1369.
- XI Lei, LIU Zhihong, LI Yanying. Automatic generation control based on lagrangian relaxation reinforcement learning algorithm[J]. *Proceedings of the CSEE*, 2023, 43(4): 1359 – 1369.
- [42] 郭佩乾, 郝峰杰, 袁志昌, 等. 混合多类型储能的分布式能源系统运行优化方法[J]. *南方电网技术*, 2024, 18(8): 29 – 38.
- GUO Peiqian, HAO Fengjie, YUAN Zhichang, et al. Operational optimization method for distributed energy systems with hybrid multi-type energy storage [J]. *Southern Power System Technology*, 2024, 18(8): 29 – 38.

收稿日期: 2024-06-18; 网络首发日期: 2024-10-28

作者简介:

柳丹(1988), 女, 高级工程师, 硕士, 研究方向为新能源及储能运行控制技术、交直流微电网运行控制技术, dannyliu6@hotmail.com;

任建宇(1999), 男, 硕士研究生, 研究方向为自动发电控制, renjianyu1999@163.com;

席磊(1982), 男, 通信作者, 教授, 博士, 研究方向为信息物理系统网络攻击与防御、自动发电控制, xilei2014@163.com。