

基于深度强化学习的孤立多微电网系统 频率和电压综合控制

谢黎龙^{1,2}, 李勇汇^{1,2}, 范培潇^{1,2}, 万黎³, 黄梦旗^{1,2}, 杨军^{1,2}

(1. 交直流智能配电网湖北省工程技术研究中心, 湖北 武汉 430072;

2. 武汉大学 电气与自动化学院, 湖北 武汉 430072; 3. 国网湖北省电力科学研究院, 湖北 武汉 430077)

摘要:分布式电源出力不确定性和负荷功率扰动给孤立多微电网系统稳定带来较大威胁。提出基于多智能体柔性动作评价(MA-SAC)算法的孤立多微电网负荷频率控制器(LFC),同时采用柔性动作评价(SAC)算法对自动电压调节器(AVR)的比例积分(PI)控制参数进行优化调整。建立了多微电网LFC和AVR组合模型。对于电压和频率控制器的设计,分别根据SAC算法和多智能体深度强化学习(MA-DRL)框架建立各自的状态、动作空间与奖励函数。选择合适的神经网络与训练参数经过预学习生成深度强化学习控制器。最后通过仿真分析,基于SAC算法优化的PI控制器能更快跟踪电压参考值;多微电网系统遭遇功率扰动时,MA-SAC控制器可以快速维持频率稳定。

关键词:多微电网系统;柔性负荷;负荷频率控制;自动电压调节;MA-SAC算法

中图分类号:TM712

文献标志码:A

DOI:10.16081/j.epae.202311013

0 引言

可再生能源发电出力的随机性和不确定性降低了其在电网中的利用效率。为解决这一难题,微电网被提出并得到了发展^[1]。频率和电压是衡量电能质量的重要指标,有功功率不平衡引起频率波动,无功功率变化影响电压。负荷频率控制器(load frequency controller, LFC)是解决电力系统频率偏移的重要方法之一^[2],电压的变化可以通过发电机的励磁来控制^[3],上述两方面研究对提高微电网系统电能质量具有重要作用。

将多个微电网相互耦合连接构成多微电网群,可以提高整个系统的抗干扰能力。孤岛微电网惯性较小^[4],互联而成的多微电网拓扑结构更加复杂,这为控制器的设计增加了难度^[5]。在实际场景中,自动电压调节器(automatic voltage regulator, AVR)回路与LFC回路弱交叉耦合对系统频率稳定也有一定的影响。

许多学者研究了LFC对互联系统的影响。文献[6]对互联系统展开负荷频率控制,然而固定参数的比例积分(proportional integral, PI)控制器难以适应多微电网系统非线性、时变不确定性、强干扰性等特

性。文献[7-9]采用灰狼优化算法、回溯搜索算法、花卉授粉算法对互联系统的PI和比例积分微分(proportional integral differential, PID)控制器参数进行优化整定。上述研究仅限于LFC,并未考虑AVR回路。文献[10]介绍了单区域LFC和AVR组合模型。文献[11]采用模拟退火算法对多区域LFC和AVR的PID控制器进行优化调整。文献[12]采用闪电搜索算法对带微分滤波器的积分双微分控制器参数进行优化。上述研究通过优化算法确定PID控制器增益。当运行场景变化时,优化求解控制参数会产生较大的实时计算负担。同时上述研究对不同区域系统间的相互支撑考虑较少,未充分发挥多区域微电网间的协同互助能力。

随着人工智能的发展,深度强化学习应用于电力领域成为当今研究热点^[13]。将神经网络引入强化学习,使强化学习具备了解决高维问题的能力^[14]。文献[15]基于深度Q学习提出了互联系统控制策略。文献[16]采用深度Q学习对线性有源抗扰控制器参数进行实时调整。然而上述研究方法中智能体动作为离散空间,在解决连续动作问题时,动作的离散程度对训练收敛和控制效果有较大影响。针对连续动作问题,基于actor-critic框架下深度强化学习得以发展^[17]。文献[18]将深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法应用于单区域微电网的电压和频率综合控制,提高了系统的动态性能,然而研究局限于单区域微电网。文献[19]采用DDPG算法提高了双区域电力系统的动态性能,然而研究采用1个控制器控制2片区域,没有考虑多个智能体的协调配合,降低了系统的灵活性。

收稿日期:2023-01-18;修回日期:2023-08-14

在线出版日期:2023-12-18

基金项目:国家自然科学基金资助项目(51977154);国网湖北省电力有限公司科技项目(B31532226407)

Project supported by the National Natural Science Foundation of China(51977154)and the Science and Technology Program of State Grid Hubei Electric Power Company Limited (B31532226407)

针对多区域电力系统,基于多智能体深度强化学习(multi-agent deep reinforcement learning, MA-DRL)方法被提出。文献[20]将多智能体DDPG算法应用于多区域电力系统LFC的设计。上述连续动作算法主要集中在单智能体,对于分布式电源种类较多、拓扑结构复杂的多微电网系统,单智能体难以满足控制要求。同时在控制器的设计算法上,当前研究集中在基于确定性策略的强化学习算法,对基于随机策略的强化学习算法缺少研究和应用。而以DDPG算法为代表的确定性策略算法参数较多,训练不够稳定,较难收敛。柔性动作评价(soft actor critic, SAC)算法作为一种先进的基于策略的强化学习算法,探索能力强,易于收敛^[21],基于多智能体柔性动作评价(multi-agent soft actor critic, MA-SAC)算法设计控制器是提高多微电网系统暂态性能的有效方式。

基于此,本文采用SAC算法对孤立多微电网中的AVR进行改进,在MA-DRL框架下基于MA-SAC算法对LFC进行设计。建立了多源LFC和AVR组合的多微电网模型,考虑两回路间弱交叉耦合。引入了柔性负荷这一调频资源,考虑了其容量随机变化导致的输出功率增量约束变化。采用SAC算法对电压PI控制器参数进行优化调整,加快了电压跟踪参考值的速度,减小了超调量。基于MA-SAC算法设计LFC,训练时加入 L_2 参数范数惩罚项与梯度裁剪操作,并设计终止函数来挖掘子微电网间协同互助的潜力。最后通过仿真算例证明所提基于数据驱动的频率和电压综合控制的有效性,进一步提升了子微电网间的相互支撑能力。

1 多微电网模型

本文研究的微电网均包含柴油发电机、飞轮储能、柔性负荷和风电,单个微电网结构如图1所示。图中:变量下标 $i, j(i, j=1, 2, \dots, n, i \neq j, n$ 为微电网总数)表示第 i, j 个微电网,后同; ΔP_{tieij} 为微电网间的联络线功率变化; ΔP_{Gi} 为柴油发电机的功率增量; ΔP_{Wi} 为风电的功率增量; ΔP_{Ei} 为飞轮储能的功率增量; ΔP_{Hi} 为柔性负荷的功率增量; ΔP_{Li} 为单个微电网受到的负荷扰动。

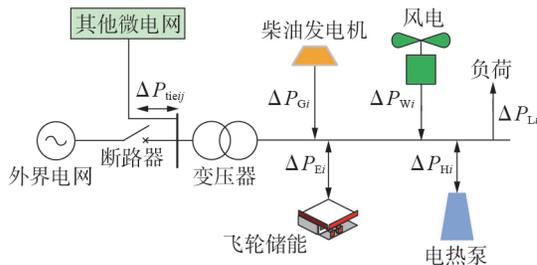


图1 单个微电网示意图

Fig.1 Schematic diagram of single microgrid

1.1 调频电源负荷频率控制模型

1.1.1 柴油发电机负荷频率控制模型

柴油发电机模型包括一个调速器和一个柴油发电机^[22],其动态模型表达式如式(1)所示。

$$\begin{cases} \Delta \dot{P}_{Gi} = \frac{1}{T_{ti}} \Delta X_{Gi} - \frac{1}{T_{ti}} \Delta P_{Gi} \\ \Delta \dot{X}_{Gi} = \frac{1}{T_{fi}} \Delta u_{Gi} - \frac{1}{T_{fi}} \Delta X_{Gi} - \frac{1}{R_{fi} T_{fi}} \Delta f_i \end{cases} \quad (1)$$

式中: ΔX_{Gi} 为调速器的输出信号; Δu_{Gi} 为柴油发电机的控制信号; T_{ti}, T_{fi} 分别为调速器和柴油发电机的时间常数; R_{fi} 为柴油发电机的调差系数; Δf_i 为频率偏移。考虑到柴油发电机的发电速率约束和发电功率增量变化约束,非线性功率特性见附录A式(A1)。

1.1.2 飞轮储能负荷频率控制模型

飞轮储能负荷频率控制模型为一阶传递函数,如图2所示。图中: Δu_{Ei} 为飞轮储能控制器的控制信号; T_{Ei} 为飞轮储能控制器的时间常数; σ 为调节死区系数; ΔP_{Eimaxi} 和 ΔP_{Eimin} 分别为飞轮储能功率增量的上限和下限。

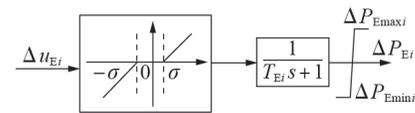


图2 飞轮储能负荷频率控制模型

Fig.2 Load frequency control model of flywheel energy storage

1.1.3 柔性负荷频率控制模型

在电力系统中,很多负荷比如冰箱、电热泵等都有参与系统频率调控的潜能,因此可以将集群形式的柔性负荷参与微电网的频率调控中。柔性负荷计算中心根据指令进行取整,决定电热泵的关断,结合此时的柔性负荷输出功率增量上下限约束确定最终的柔性负荷功率输出增量。由于不同类别的柔性负荷控制模型基本相同,本文采用电热泵建立柔性负荷频率控制模型,如图3所示。图中: Δu_{Hi} 为柔性负荷控制模型的控制信号; T_{Hi} 为柔性负荷控制模型的时间常数; ΔP_{Himaxi} 和 ΔP_{Himin} 分别为柔性负荷功率增量的上限和下限, ΔP_{Himaxi} 和 ΔP_{Himin} 的推导见附录A式(A2),柔性负荷功率增量 ΔP_{Hi} 的表达式及约束条件见附录A式(A3)。

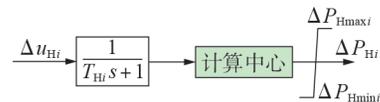


图3 柔性负荷频率控制模型

Fig.3 Flexible load frequency control model

1.2 LFC与AVR组合模型

AVR是维持发电机端电压稳定的局部闭环系

统, AVR模型见附录A图A1。负荷频率模型和联络线模型分别见附录A式(A4)、(A5)。多微电网LFC和AVR组合模型见附录A图A2。由模型结构可以看出,柴油发电机的电压调节过程对频率控制有一定影响。

2 控制算法原理

2.1 SAC算法

与传统深度强化学习不同, SAC算法引入了动作熵。智能体需要将奖励与动作熵值全部最大化, 如式(2)所示。

$$\pi^* = \operatorname{argmax}_{\pi} E_{(s_t, a_t) \sim p_{\pi}} \left[\sum_t R(s_t, a_t) + \alpha H(\pi(\cdot | s_t)) \right] \quad (2)$$

式中: π^* 为最优策略; $E[\cdot]$ 为数学期望; s_t, a_t 分别为智能体在 t 时刻的状态与采取的动作; $(s_t, a_t) \sim p_{\pi}$ 表示在策略 π 下轨迹 (s_t, a_t) 的分布; $R(s_t, a_t)$ 为智能体在 s_t 下采取 a_t 获得的回报值; α 为温度系数; $H(\cdot)$ 为动作熵值的计算函数。

SAC算法共包含3个神经网络, 分别为actor网络、critic网络、目标critic网络, 对应的参数分别为 ϕ, θ, θ' 。由式(3)~(7)更新网络参数。

$$J(\phi) = E_{s_t \sim D, a_t \sim \pi_{\phi}} \left[\ln \pi_{\phi}(a_t | s_t) - \frac{1}{\alpha} Q_{\theta}(a_t | s_t) \right] \quad (3)$$

$$J(\theta) = E_{(s_t, a_t, s_{t+1}) \sim D, a_{t+1} \sim \pi_{\theta}} \left[\frac{1}{2} (Q_{\theta}(a_t | s_t) - y)^2 \right] \quad (4)$$

$$y = r(s_t, a_t) +$$

$$\gamma E_{a_{t+1} \sim \pi_{\theta}} \left[Q_{\theta'}(a_{t+1} | s_{t+1}) - \alpha \ln \pi_{\phi}(a_{t+1} | s_{t+1}) \right] \quad (5)$$

$$\theta' = \tau \theta + (1 - \tau) \theta' \quad (6)$$

$$J(\alpha) = E_{a_t \sim \pi_{\phi}} \left[-\alpha \ln \pi_{\phi}(a_t | s_t) - \alpha H_0 \right] \quad (7)$$

式中: $J(\phi), J(\theta), J(\alpha)$ 分别为以参数 ϕ 、输出值 Q 下参数 θ 以及系数 α 为自变量的损失函数; D 为经验回放池; $Q_{\theta}(a_t | s_t)$ 为critic网络在 s_t 下采取 a_t 的输出值; $\pi_{\phi}(a_t | s_t)$ 为actor网络参数 ϕ 在 s_t 下采取 a_t 的概率; $r(s_t, a_t)$ 为智能体在 s_t 下采取 a_t 后获得的奖励函数; $Q_{\theta'}(a_{t+1} | s_{t+1})$ 为目标critic网络在 s_{t+1} 下采取 a_{t+1} 后的输出值; τ 为更新系数; H_0 为熵的限制; γ 为折扣系数。

2.2 基于MA-DRL框架的MA-SAC算法

在MA-DRL框架下的多微电网负荷频率控制问题中, 每个智能体都基于神经网络的建立。对于第 i 个微电网智能体, actor网络参数 ϕ_i 的优化目标为:

$$J(\phi_i) = E_{\{s_i^t\} \sim D, \{a_i^t\} \sim \{\pi_{\phi_i}\}} \left[\ln \pi_{\phi_i}(a_i^t | s_i^t) - \frac{1}{\alpha} Q_{\theta_i}(\{a_i^t\} | \{s_i^t\}) \right] \quad (8)$$

式中: $J(\phi_i)$ 为第 i 个智能体中策略 π_i 下 actor 网络参数 ϕ_i 的损失函数; 回放池 D 用于存储 t 时刻状态量集合 $\{s_i^t\}$ 、 $t+1$ 时刻观察量集合 $\{s_i^{t+1}\}$ 、 t 时刻动作集合 $\{a_i^t\}$ 和奖励集合 $\{r_i^t\}$; $Q_{\theta_i}(\{a_i^t\} | \{s_i^t\})$ 为第 i 个智能体 critic 网络 t 时刻在状态 s_i 下采取动作 a_i 时的输出值; $\pi_{\phi_i}(a_i^t | s_i^t)$ 为第 i 个智能体 critic 网络 t 时刻在状态 s_i 下采取动作 a_i 的概率。

第 i 个智能体 critic 网络参数 θ_i 通过最小化贝尔曼误差进行更新:

$$J(\theta_i) = E_{\{s_i^t\}, \{a_i^t\}, \{r_i^t\}, \{s_i^{t+1}\} \sim D} \left[\frac{1}{2} (Q_{\theta_i}(\{a_i^t\} | \{s_i^t\}) - z)^2 \right] \quad (9)$$

$$z = r_i + \gamma E_{\{a_i^{t+1}\} \sim \{\pi_{\phi_i}\}} \left[Q_{\theta_i}(\{a_i^{t+1}\} | \{s_i^{t+1}\}) - \alpha \ln \pi_{\phi_i}(a_i^{t+1} | s_i^{t+1}) \right] \quad (10)$$

式中: $J(\theta_i)$ 为第 i 个智能体中以参数 θ_i 为自变量的损失函数; s_i^{t+1}, a_i^{t+1} 分别为第 i 个智能体在 $t+1$ 时刻的状态量、采取的动作; $Q_{\theta_i}(\{a_i^{t+1}\} | \{s_i^{t+1}\})$ 为第 i 个智能体目标 critic 网络 $t+1$ 时刻在状态 s_i 下采取动作 a_i 时的输出值。

第 i 个智能体目标 critic 网络参数 θ'_i 同样通过软更新方式进行更新, 温度系数 α 通过式(7)随着训练不断更新。采用自适应随机梯度下降算法对神经网络进行更新, 可见附录A式(A6)。

3 多微电网控制器设计

对于图A2中多微电网频率控制, 传统控制器难以发挥多个微电网协同互助的能力, 抗干扰能力较弱。本文基于MA-SAC算法设计多微电网深度强化学习频率控制器对频率进行控制, 增强子微电网间的协同互助。对于图A1中的AVR设计, 大多数研究采用固定参数的PI控制器, 虽然结构简单, 但是恒定的增益导致控制器适应性较差。本文采用SAC算法对电压PI控制器进行实时调节。综合控制设计框图如图4所示。

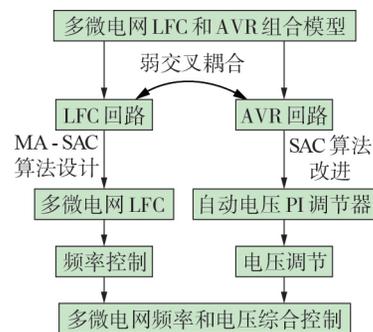


图4 多微电网综合控制设计框图

Fig.4 Block diagram of integrated control design of multi-microgrid

3.1 多微电网频率控制器设计

在MA-DRL框架中,每个智能体训练时需要获取其他智能体的状态信息与动作信息,故在多微电网中,每个微电网调频电源的输出和频率信息需要通过通信线传递到其他微电网,如图5所示。图中: X_i 、 Q_i 分别为第*i*个智能体的训练样本和动作价值输出值。

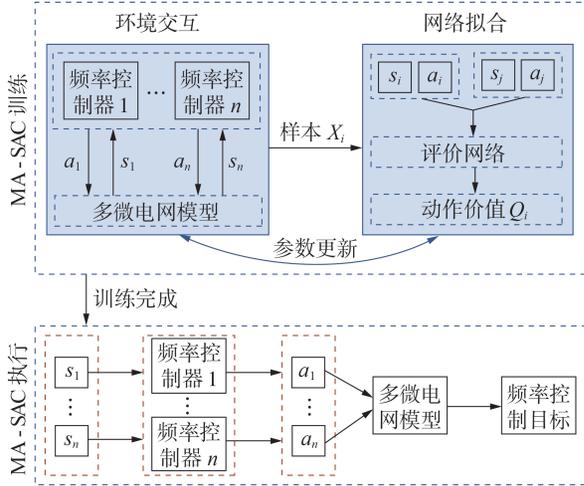


图5 基于MA-SAC算法多微电网频率控制框图

Fig.5 Block diagram of multi-microgrid frequency control based on MA-SAC algorithm

在训练时,每个微电网智能体获取其他微电网的频率信息与控制器动作信息,具体详细信息见3.1.1节。每个智能体在获取本地微电网和其他微电网的频率信息和动作信息后,通过均方损失函数,更新评价网络参数,降低评价网络输出 Q 值的误差,同时根据 Q 值更新动作网络参数和智能体的频率控制策略。当评价网络与动作网络训练结束后,每个微电网智能体在执行时通过本地的频率信息采取合理动作,维持频率的稳定。

考虑到多微电网系统子微电网间的信息传输成本会随着距离的增加而明显增加,因此本文设计的多微电网系统中智能体间的信息传输距离须保持在合理的范围,即足够小的通信时延,可以在系统建模过程中忽略时延的影响。

3.1.1 状态变量与动作变量

本文基于MA-SAC算法设计多微电网频率控制器,对于第*i*个微电网智能体,选择实际频率 f_i 与额定频率的偏移量 Δf_i 以及频率误差的积分值 $\int \Delta f_i dt$ 作为状态量 s_i ,如式(11)所示。

$$s_i = \left(f_i, \Delta f_i, \int \Delta f_i dt \right) \quad (11)$$

式中: $\Delta f_i = f_i - f_N$, f_N 为额定频率。式(11)反映了智能体得到的外界环境信息。

动作变量 a_i 选取为第*i*个微电网调频电源的控

制信号,即:

$$a_i = (\Delta u_{Gi}, \Delta u_{Ei}, \Delta u_{Hi}) \quad (12)$$

在控制器设计过程中,没有将其他微电网的动作变量直接输入本地智能体,而是选择其他子微电网*j*调频电源的功率输出增量作为动作信息 A_j ,如式(13)所示。

$$A_j = (\Delta P_{Gj}, \Delta P_{Ej}, \Delta P_{Hj}) \quad (13)$$

根据2.4节内容,第*i*个微电网智能体的训练样本单元 X_i 为:

$$X_i = \left\{ (s_i^t, s_j^t), (a_i^t, A_j^t), r_i^t, (s_i^{t+1}, s_j^{t+1}) \right\} \quad (14)$$

3.1.2 奖励函数

奖励值是十分重要的超参数,通过奖励函数,频率控制目标问题才能转化为奖励最大化问题。智能体通过获得的奖励值才能判断采取动作的好坏,进而采取合适的动作,考虑到 $(-0.03, 0.03)$ Hz的频率调节死区,本文设计第*i*个微电网智能体频率控制器建立的奖励函数 r_i 为:

$$r_i = \begin{cases} \eta & |\Delta f_i| < 0.03 \text{ Hz} \\ \mu_1 |\Delta f_i| & 0.03 \text{ Hz} \leq |\Delta f_i| < 0.05 \text{ Hz} \\ \mu_2 |\Delta f_i| & 0.05 \text{ Hz} \leq |\Delta f_i| < 0.08 \text{ Hz} \\ \mu_3 |\Delta f_i| & 0.08 \text{ Hz} \leq |\Delta f_i| < 0.12 \text{ Hz} \\ \mu_4 |\Delta f_i| & 0.12 \text{ Hz} \leq |\Delta f_i| \leq 0.2 \text{ Hz} \\ B & |\Delta f_i| > 0.2 \text{ Hz} \end{cases} \quad (15)$$

式中: $\mu_1 - \mu_4$ 为不同频率偏移区间对应的惩罚系数; η 为正奖励值,本文设计当且仅当频率偏移小于0.03 Hz时,智能体获得较大的正奖励值,频率偏移越大,智能体获得的奖励值越小; B 为巨大惩罚值,当频率偏移超过0.2 Hz时,智能体将受到巨大惩罚。

为了提高训练的速度,引入终止函数。通过终止函数,智能体可以判断是否进入终止状态。如果智能体进入终止状态,则结束本轮迭代,重新初始化状态开展新一轮的训练。如果智能体未进入终止状态,则继续本轮训练。终止状态一般是比较差的状态,本文对应为频率偏移很大的情况。结合奖励函数将终止状态的奖励设置很低以减少智能体进入终止状态的次数。本文频率控制器设计的终止函数 $f_{\text{end}}(\Delta f_i)$ 为:

$$f_{\text{end}}(\Delta f_i) = \begin{cases} T_{\text{termination}} & |\Delta f_i| > 0.2 \text{ Hz} \\ T_{\text{non-termination}} & |\Delta f_i| \leq 0.2 \text{ Hz} \end{cases} \quad (16)$$

式中: $T_{\text{termination}}$ 、 $T_{\text{non-termination}}$ 分别为本集训练结束、继续信号。

由于是多个智能体集中训练,当一个微电网的频率偏移满足进入终止状态的条件时,所有智能体都将结束本轮训练,开启新一轮的迭代,即其他智能体无法继续采取动作取得高额回报。因此,每个智

能体间会相互协调配合,使每个微电网频率偏移尽量远离终止状态,以获取高额回报,进而大幅度提高子微电网间协同互助的能力。

3.1.3 神经网络

神经网络强大的拟合能力使强化学习具备了解决高维问题的能力。神经网络的种类较多,包括卷积神经网络、全连接网络、循环神经网络等。神经网络的拟合能力与网络层数 b 和第 e ($e=1,2,\dots,g$)层神经元的数目 u_e 密切相关。由于频率等状态量为一维向量,且为了避免神经网络参数更新时出现梯度消失的问题,选用全连接网络和长短期记忆(long short-term memory, LSTM)网络构建智能体,如附录A图A3所示。

为了避免训练时出现梯度消失的问题,神经网络层与层之间选用线性整流函数(rectified linear units, ReLU)。同时,为了避免训练时出现梯度爆炸的问题,本文加入了梯度裁剪的操作:当网络参数梯度的 L_2 范数大于梯度阈值 ρ 时,基于该阈值重新缩放梯度,并且不改变梯度的方向。

为了防止神经网络训练时出现过拟合的问题,提高网络的泛化能力,本文加入正则化处理。智能体根据损失函数式(8)、(9),采用随机梯度下降方法更新网络参数时加入 L_2 参数范数惩罚项 L_2 ,如式(17)所示。

$$L_2 = \frac{\lambda}{2m} \omega^T \omega \quad (17)$$

式中: ω 为神经网络权重向量; λ 为正则化因子,包括动作网络正则化因子 λ_{actor} 与评价网络正则化因子 λ_{critic} ; m 为随机采样的样本数。

3.1.4 算法流程

多微电网系统中多智能体通过随机策略采取一系列动作与环境进行交互,并将经验存储在经验回放池 D 中。采用随机梯度下降更新网络参数时,智能体从经验回放池 D 中随机抽取 m 个训练样本 X ,根据损失函数进行训练。设多智能体最多训练的集数为 T ,每集训练的最大时间步长为 T_m ,每步时间为 t_s 。具体算法流程如附录A图A4所示。

在每集训练开始时,需要对第一个状态进行初始化。设 p_i 为第 i 个微电网的功率干扰量,在每集训练开始时对每个子微电网施加功率干扰。 p_i 变化范围在 $[-L, L]$ 之间, L 为边界值。由于加入微电网的扰动在该范围内完全随机,每个子微电网遭遇的功率干扰也不尽相同,通过训练多微电网MA-SAC LFC可以应对各种复杂的功率扰动情况。

3.2 电压控制器设计

3.2.1 状态变量与动作变量

采用SAC算法对微电网 i 的电压控制器进行改进,状态变量 s 如式(18)所示。

$$s = \left(V_i, \Delta V_i, \int \Delta V_i dt \right) \quad (18)$$

式中: ΔV_i 为实际电压标幺值 V_i 与参考电压(1 p.u.)的偏差值 $\Delta V_i, \Delta V_i = V_i - 1$;

对PI控制参数进行实时调节,故动作变量 a 由比例系数 K_p 、积分系数 K_i 构成,如式(19)所示。

$$a = (K_p, K_i) \quad (19)$$

式中: $K_p \in [K_{pmin}, K_{pmax}]$, K_{pmax} 、 K_{pmin} 分别为 K_p 的上、下限; $K_i \in [K_{imin}, K_{imax}]$, K_{imax} 、 K_{imin} 分别为 K_i 的上、下限。

3.2.2 奖励函数

电压控制器的奖励函数建立原则是,电压偏移越大,获得的惩罚值越大,如式(20)所示。

$$r = \begin{cases} \chi_1 |\Delta V_i| + \varphi & |\Delta V_i| < 0.03 \text{ p.u.} \\ \chi_2 |\Delta V_i| & 0.03 \text{ p.u.} \leq |\Delta V_i| < 0.1 \text{ p.u.} \\ \chi_3 |\Delta V_i| & (-1 \text{ p.u.} \leq \Delta V_i \leq -0.1 \text{ p.u.}) \cup (0.1 \text{ p.u.} \leq \Delta V_i \leq 0.3 \text{ p.u.}) \\ C & \Delta V_i > 0.3 \text{ p.u.} \end{cases} \quad (20)$$

式中: χ_1 — χ_3 为不同电压偏移区间对应的惩罚系数; φ 为正奖励值; C 为智能体进入终止状态时获得的最大惩罚值。同样设置终止函数 $f_{end}(\Delta V_i)$ 为:

$$f_{end}(\Delta V_i) = \begin{cases} T_{termination} & \Delta V_i > 0.3 \text{ p.u.} \\ T_{non-termination} & \Delta V_i \leq 0.3 \text{ p.u.} \end{cases} \quad (21)$$

3.2.3 神经网络

电压控制器神经网络的结构和频率控制器智能体网络结构相同,同样引入梯度裁剪与 L_2 参数范数惩罚项训练技巧。

4 仿真结果与分析

4.1 仿真设置

在MATLAB/Simulink中搭建2个微电网互联的多微电网LFC和AVR组合模型。利用MATLAB中的Reinforcement Learning Toolbox对模型控制器进行设计。模型的仿真参数如附录B表B1所示,基准功率为200 kW,基准电压为10 kV,仿真中的电压和功率均为标幺值。神经网络选择见附录B表B2、B3。频率控制器智能体训练参数如附录B表B4所示,电压控制器智能体训练参数如附录B表B5所示。

4.2 预学习阶段

深度强化学习需要经过训练才能达到较好的控制效果,这个过程称之为预学习。本文先通过SAC算法对电压PI控制器进行优化,在训练得到基于SAC算法优化的电压控制器后,对多微电网频率深度强化学习控制器展开训练,最终得到多微电网系统的LFC。

由于可用于调控的电热泵数量具有随机性,柔性负荷的容量也会发生改变。在2个微电网智能体

共同训练时,根据某柔性负荷的输出容量变化数据设置了一组随时间发生概率随机变化的输出功率增量约束,见附录B图B1。2个微电网AVR回路参数与优化算法参数完全相同,且互不影响,微电网2与微电网1电压控制器的训练结果相同,本文电压仿真图形只展示微电网1的部分。微电网SAC算法对电压PI控制器参数优化时的训练过程如附录B图B2所示,2个子微电网频率控制器基于MA-SAC算法的训练过程如附录B图B3所示。

4.3 仿真结果

4.3.1 电压控制效果

为体现SAC算法中PI控制器(后文简称为SAC-PI控制器)的性能,引入传统PI控制器与模糊PI(fuzzy proportional integral, Fuzzy-PI)控制器,将3种控制器进行对比。微电网发电机机端电压波形见图6,对比分析结果见附录B表B6。

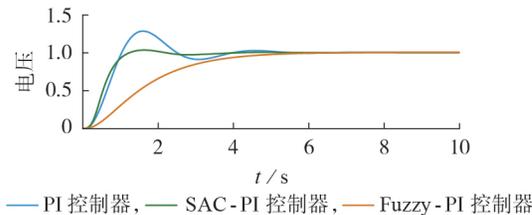


图6 3种控制器下微电网发电机机端电压波形

Fig.6 Waveforms of microgrid generator terminal voltage under 3 kinds of controller

由图6和表B6可以看出:传统PI控制具有较大的超调量,达到28.3%;Fuzzy-PI控制器的电压响应速度较慢;而SAC-PI控制器下的电压响应速度较快,调节时间很短,只有1.04s,且超调量很小,为3.3%,明显优于PI控制器。SAC算法对PI控制器参数的实时优化结果如附录B图B4所示。

4.3.2 AVR回路对LFC的影响

由于AVR回路和LFC回路存在一定程度上的耦合,电压的调节过程会给负荷频率控制带来一定的功率干扰。图7为AVR回路的扰动功率。

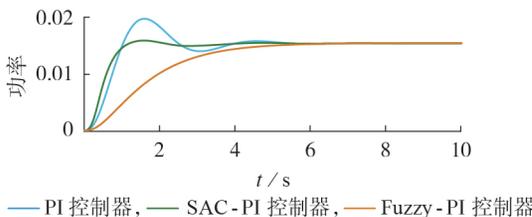


图7 AVR回路的扰动功率

Fig.7 Disturbance power of AVR circuit

对于MA-SAC控制器控制的多微电网系统,电压调节器采用SAC-PI控制器时,在初始时刻对微电网1施加0.03 p.u.的有功功率干扰,分析AVR回路对LFC的影响,如图8所示。

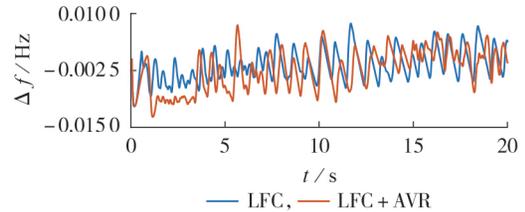


图8 是否考虑AVR耦合时的频率偏移

Fig.8 Frequency deviations with and without AVR coupling consideration

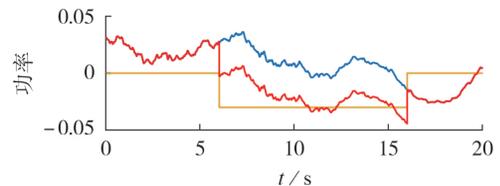
由图7、8可以看出,由于AVR回路增加了有功功率干扰,在微电网遭遇功率扰动时,负荷频率发生了更大的偏移,这对LFC提出了更高的要求。

4.3.3 频率控制效果

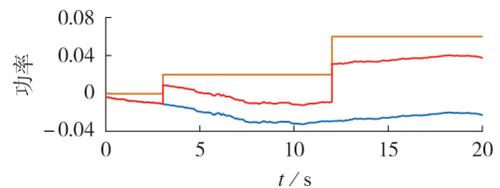
对于LFC和AVR组合模型,在考虑AVR回路耦合功率的影响下,对多微电网系统施加有功功率扰动,在如下场景中分析MA-SAC控制器的频率控制效果。

4.3.3.1 场景1:系统受到随机扰动与负载扰动

本文构建的多微电网系统中,设风电有功为输入微电网的随机扰动源,负荷有功扰动为阶跃扰动,本场景下探究随机扰动、阶跃扰动以及二者叠加而成的总扰动对频率的影响。功率扰动波形见图9。



(a) 微电网1受到的功率扰动



(b) 微电网2受到的功率扰动

— 风电扰动, — 负荷阶跃扰动, — 总扰动

图9 场景1多微电网受到的功率扰动

Fig.9 Power disturbances of multi-mircogrid under Scenario 1

为了体现MA-SAC控制器的性能,引入PI、Fuzzy-PI、多智能体深度确定性策略梯度(multi-agent deep deterministic policy gradient, MA-DDPG)控制器进行对比。其中MA-DDPG控制器的训练过程如附录B图B5所示,AVR回路中电压调节器均采用本文提出的SAC-PI控制器。各子微电网的频率偏移如图10所示,数据统计信息如附录B表B7所示。微电网1中各调频电源的功率输出增量如附录B图B6所示。

由图10可以看出,当发生强扰动时,在PI控制

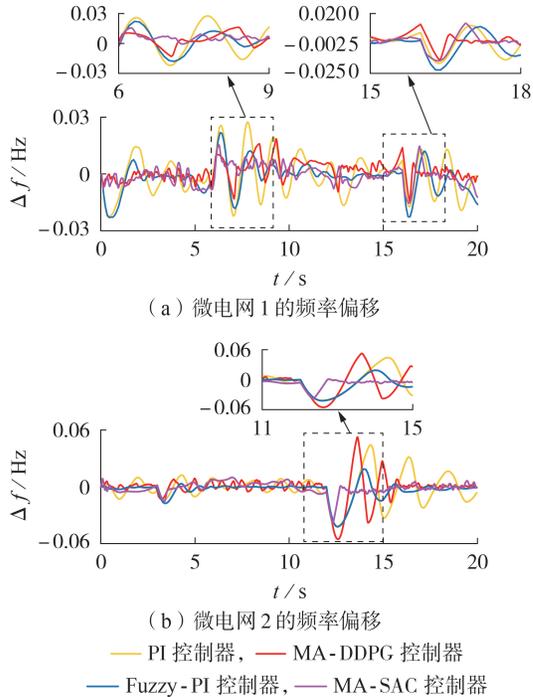


图10 多微电网的频率偏移

Fig.10 Frequency deviations of multi-microgrid

器和 Fuzzy-PI 控制器下微电网 1 中的频率偏移出现了较大波动。微电网 1 中 MA-DDPG 和 MA-SAC 这 2 种深度强化学习控制器对频率的控制效果更优,特别是 MA-SAC 控制器,频率最大偏移量只有 0.015 75 Hz,标准差也最小。

在 12 s 时微电网 2 中的有功负荷突然增加,MA-SAC 控制器的频率调节速率最快,其他 3 种控制器均出现了较长时间的频率波动,PI 控制器调节时间最长,4.48 s 后频率偏移量小于 0.03 Hz,且频率仍在波动。采用 MA-DDPG 控制器时微电网 2 的频率偏移量最大,为 0.055 4 Hz。出现这种情况的原因可能是 MA-DDPG 算法在训练时,只能确定一个固定动作,由于探索能力不足,没有找到最优动作,这说明 MA-DDPG 算法在训练收敛时存在困难。而基于 MA-SAC 算法的频率控制器具有优越的鲁棒性能,能够快速调节频率,抑制频率偏移。

4.3.3.2 场景 2: 单一微电网调频机组出现故障

多微电网系统中,单个子微电网中的调频电源可能出现故障,退出调频。同时调频能力下降的子微电网可能遭受极端功率扰动,进而威胁微电网安全运行。本节对这一异常情况进行仿真,用于验证所提方法的鲁棒性。假设微电网 1 中柔性负荷出现故障并退出调频,同时给微电网 1 加入极端功率扰动,如图 11 所示。微电网 1 遭遇的功率扰动最大已经达到 0.066 8 p.u.,同时考虑到 AVR 耦合功率的干扰,总的功率扰动已经大幅超过了此时微电网 1 的最大功率增量 0.055 p.u.。此场景下各子微电网的

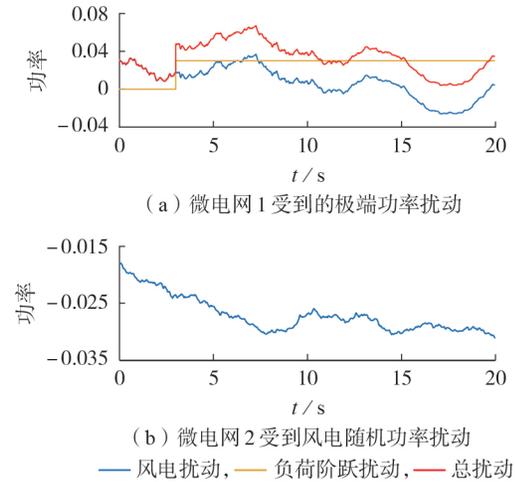


图 11 场景 2 多微电网受到的功率扰动

Fig.11 Multi-microgrid power disturbances under Scenario

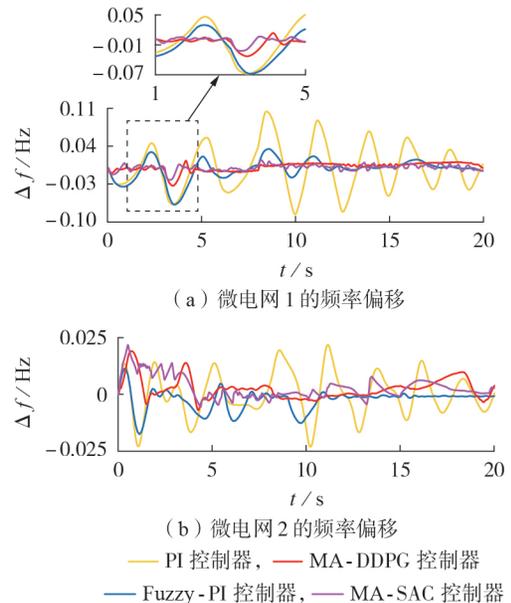


图 12 多微电网频率偏移

Fig.12 Frequency deviation of multi-microgrid

频率偏移如图 12 所示,数据统计信息如附录 B 表 B8 所示。微电网 1 中各调频电源功率输出增量如附录 B 图 B7 所示。

从图 12 可以看出,在微电网 1 柔性负荷退出调频,并遭遇极端功率干扰时,PI 控制器、Fuzzy 控制器下的频率出现了较大波动。MA-DDPG 控制器和 MA-SAC 控制器下频率偏移较小,通过仿真波形,结合平均值和标准差,可以得出 MA-SAC 控制器下微电网 1 的频率最为稳定,频率偏移最小,恢复时间最短。微电网 2 中,在随机风电扰动下,PI 控制器频率波动依旧最大,频率最大偏移量为 0.022 9 Hz,在整个仿真过程中频率一直存在波动。其他 3 种控制器在仿真过程中都只出现了小幅度的频率波动,而 MA-SAC 控制器在 2.91 s 将频率偏移量持续稳定在

0.01 Hz以内,恢复时间最短。强化学习控制器下频率偏移量出现小幅度波动的主要原因是奖励函数考虑了频率死区和微电网1遭遇了极端负荷扰动。可以看出,深度强化学习控制器下的多微电网系统,在微电网1投入重负荷,微电网2通过降低本地部分频率质量来帮助微电网1实现频率恢复,减小了频率偏移量。从图B7可以看出,当微电网1柔性负荷退出调频时,由剩下的调频源承担调频任务。传输线出现较大的功率变化,说明2个子微电网通过协同互助,提高了整个多微电网系统的稳定性。仿真结果说明在子微电网调频机组因故障退出调频且遭遇重负荷的异常情况下,MA-SAC控制器依旧可以维持整个多微电网系统的稳定,具有较强的鲁棒性与适应性,可以很好地应对系统结构的变化。

结合图B5所示MA-DDPG控制器的训练过程,可以看出MA-DDPG控制器的训练次数要远远高于MA-SAC控制器,且训练稳定性较差。MA-SAC控制器对训练参数的敏感度远低于MA-DDPG,使训练的收敛速度更快,并能获得更优的控制效果。

5 结论

本文针对多微电网LFC和AVR组合模型,采用SAC算法对PI电压调节器控制参数进行改进,基于MA-SAC算法提出多微电网LFC。研究表明,SAC算法对比例系数和积分系数的实时优化提高了PI电压调节器的性能,发电机端电压可以更快达到参考值。本文提出的MA-SAC控制器有着极强的鲁棒性,当微电网遭遇随机功率波动和阶跃功率扰动时,系统可以快速恢复频率稳定。同时控制器具有很好的适应性,当子微电网中机组出现故障退出调频时,MA-SAC控制器可以实现子微电网间协同配合,维持整个系统的安全运行,提高了多微电网的抗干扰能力。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 祖国强,宋晨辉,张春晖,等. 蜂巢状综合能源配电系统的组网结构与运行方式[J]. 电力系统自动化,2023,47(12):47-55.
ZU Guoqiang, SONG Chenhui, ZHANG Chunhui, et al. Networking structures and operation modes of honeycomb integrated energy distribution system[J]. Automation of Electric Power Systems, 2023, 47(12): 47-55.
- [2] 项雷军,陈昊,郭新华,等. 基于模糊分数阶PID的含电动汽车的多能源微电网二次频率控制[J]. 电力自动化设备,2021,41(11):74-80.
XIANG Leijun, CHEN Hao, GUO Xinhua, et al. Secondary frequency control of multi-energy microgrid with electric vehicles based on fuzzy fractional-order PID[J]. Electric Power Automation Equipment, 2021, 41(11): 74-80.
- [3] 李国庆,张浩,李江,等. 基于分岔理论研究励磁饱和环节对系统电压稳定的影响[J]. 电力自动化设备,2015,35(3):1-5,46.
LI Guoqing, ZHANG Hao, LI Jiang, et al. Influence of excitation saturation element on power system voltage stability based on bifurcation theory[J]. Electric Power Automation Equipment, 2015, 35(3): 1-5, 46.
- [4] KERDPHOL T, RAHMAN F S, WATANABE M, et al. Robust virtual inertia control of a low inertia microgrid considering frequency measurement effects[J]. IEEE Access, 2019, 7: 57550-57560.
- [5] 边晓燕,孙明琦,许家玉,等. 计及灵活性储备的含风电多微电网系统分布式协调调控策略[J]. 电力自动化设备,2021,41(8):47-54,104.
BIAN Xiaoyan, SUN Mingqi, XU Jiayu, et al. Distributed coordinated dispatch and control strategy of multi-microgrid system with wind power considering flexibility reserve[J]. Electric Power Automation Equipment, 2021, 41(8): 47-54, 104.
- [6] SAIKIA L C, NANDA J, MISHRA S. Performance comparison of several classical controllers in AGC for multi-area interconnected thermal system[J]. International Journal of Electrical Power & Energy Systems, 2011, 33(3): 394-401.
- [7] GUHA D, ROY P K, BANERJEE S. Load frequency control of interconnected power system using grey wolf optimization [J]. Swarm and Evolutionary Computation, 2016, 27: 97-115.
- [8] GUHA D, ROY P K, BANERJEE S. Application of backtracking search algorithm in load frequency control of multi-area interconnected power system[J]. Ain Shams Engineering Journal, 2018, 9(2): 257-276.
- [9] MADASU S D, KUMAR M, SINGH A K. A flower pollination algorithm based automatic generation control of interconnected power system[J]. Ain Shams Engineering Journal, 2018, 9(4): 1215-1224.
- [10] 崔明勇,曹朋,朱大伟,等. 基于FOPI+FOPD控制器的单区域电力系统频率控制及电压调节[J]. 燕山大学学报,2022,46(2):157-165,176.
CUI Mingyong, CAO Peng, ZHU Dawei, et al. Frequency control and voltage regulation of single area power system based on FOPI+FOPD controller[J]. Journal of Yanshan University, 2022, 46(2): 157-165, 176.
- [11] VIJAYA CHANDRAKALA K R M, BALAMURUGAN S. Simulated annealing based optimal frequency and terminal voltage control of multi source multi area system [J]. International Journal of Electrical Power & Energy Systems, 2016, 78: 823-829.
- [12] RAJBONGSHI R, SAIKIA L C. Combined control of voltage and frequency of multi-area multisource system incorporating solar thermal power plant using LSA optimised classical controllers [J]. IET Generation, Transmission & Distribution, 2017, 11(10): 2489-2498.
- [13] 董运昌,王启明,曹杰,等. 基于过采样和级联机器学习的电网虚假数据注入攻击识别[J]. 电力系统自动化,2023,47(8): 179-188.
DONG Yunchang, WANG Qiming, CAO Jie, et al. Identification of false data injection attacks in power grid based on oversampling and cascade machine learning[J]. Automation of Electric Power Systems, 2023, 47(8): 179-188.
- [14] 赵鹏杰,吴俊勇,王焱,等. 基于深度强化学习的微电网优化运行策略[J]. 电力自动化设备,2022,42(11):9-16.
ZHAO Pengjie, WU Junyong, WANG Yi, et al. Optimal operation strategy of microgrid based on deep reinforcement learning[J]. Electric Power Automation Equipment, 2022, 42(11): 9-16.
- [15] YIN L F, YU T, ZHOU L. Design of a novel smart generation controller based on deep Q learning for large-scale interconnected power system [J]. Journal of Energy Enginee-

- ring, 2018, 144(3):04018033.
- [16] ZHENG Yuemin, SUN Qinglin, CHEN Zengqiang, et al. Deep Q -network based real-time active disturbance rejection controller parameter tuning for multi-area interconnected power systems[J]. *Neurocomputing*, 2021, 460:360-373.
- [17] KIUMARSI B, LEWIS F L. Actor-critic-based optimal tracking for partially unknown nonlinear discrete-time systems[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2015, 26(1):140-151.
- [18] FAN P X, KE S, KAMEL S, et al. A frequency and voltage coordinated control strategy of island microgrid including electric vehicles[J]. *Electronics*, 2021, 11(1):17.
- [19] 梁煜东, 陈峦, 张国洲, 等. 基于深度强化学习的多能互补发电系统负荷频率控制策略[J]. *电工技术学报*, 2022, 37(7):1768-1779.
- LIANG Yudong, CHEN Luan, ZHANG Guozhou, et al. Load frequency control strategy of hybrid power generation system: a deep reinforcement learning-based approach[J]. *Transactions of China Electrotechnical Society*, 2022, 37(7):1768-1779.
- [20] YAN Z M, XU Y. A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system[J]. *IEEE Transactions on Power Systems*, 2020, 35(6):4599-4608.
- [21] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[C]//35th International Conference on Machine Learning. Stockholm, Sweden: Jmlr-Journal Machine Learning Research, 2018:80.
- [22] RAO Yingqing, YANG Jun, XIAO Jinxing, et al. A frequency control strategy for multimicrogrids with V2G based on the improved robust model predictive control [J]. *Energy*, 2021, 222:119963.

作者简介:

谢黎龙(1999—),男,硕士研究生,主要研究方向为微电网控制(**E-mail**:lilongxie@whu.edu.cn);

李勇汇(1973—),男,博士,副教授,通信作者,主要研究方向为电网规划(**E-mail**:1679172758@qq.com)。

(编辑 王欣竹)

Deep reinforcement learning-based integrated frequency and voltage control for isolated multi-microgrid system

XIE Lilong^{1,2}, LI Yonghui^{1,2}, FAN Peixiao^{1,2}, WAN Li³, HUANG Mengqi^{1,2}, YANG Jun^{1,2}

(1. Hubei Engineering and Technology Research Center for AC/DC Intelligent Distribution Network, Wuhan 430072, China;

2. School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China;

3. State Grid Hubei Electric Power Research Institute, Wuhan 430077, China)

Abstract: The output uncertainty of distributed power source and the power disturbance caused by load pose a greater threat to the stability of isolated multi-microgrid systems. An isolated multi-microgrid load frequency controller(LFC) based on multi-agent soft actor critic(MA-SAC) algorithm is proposed, and the soft actor critic(SAC) algorithm is used to optimize and adjust the proportional integral(PI) control parameters of the automatic voltage regulator(AVR). A combined model of LFC and AVR for multi-microgrid is developed. For the design of voltage and frequency controllers, the corresponding states, action spaces and reward functions are established according to SAC algorithm and multi-agent deep reinforcement learning(MA-DRL) framework, respectively. The appropriate neural network and training parameters are selected to generate the deep reinforcement learning controller through pre-learning. Finally, through simulation analysis, the PI controller optimized based on SAC algorithm can track the voltage the reference value faster, and the MA-SAC controller can maintain the frequency stability quickly when the multi-microgrid system encounters power disturbance.

Key words: multi-microgrid system; flexible load; load frequency control; automatic voltage regulation; MA-SAC algorithm

$$\frac{d\Delta f_i}{dt} = (\Delta P_{Gi} + \Delta P_{Ei} + \Delta P_{Hi} - \Delta P_{tieij} - \Delta P_{Li} - \Delta P_{ei} - \Delta P_{Wi}) \frac{1}{H_i} - \Delta f_i \frac{D_i}{H_i} \quad (A4)$$

式中： ΔP_{ei} 为 AVR 回路的耦合功率； H_i 为微电网惯性时间常数； D_i 为系统负荷阻尼系数。

5) 联络线功率模型。

在多微电网系统中，互联的子微电网间传输线功率模型为：

$$\frac{d\Delta P_{tieij}}{dt} = T_{ij}(\Delta f_i - \Delta f_j) \quad (A5)$$

式中： T_{ij} 为微电网 i 、 j 之间的同步系数。

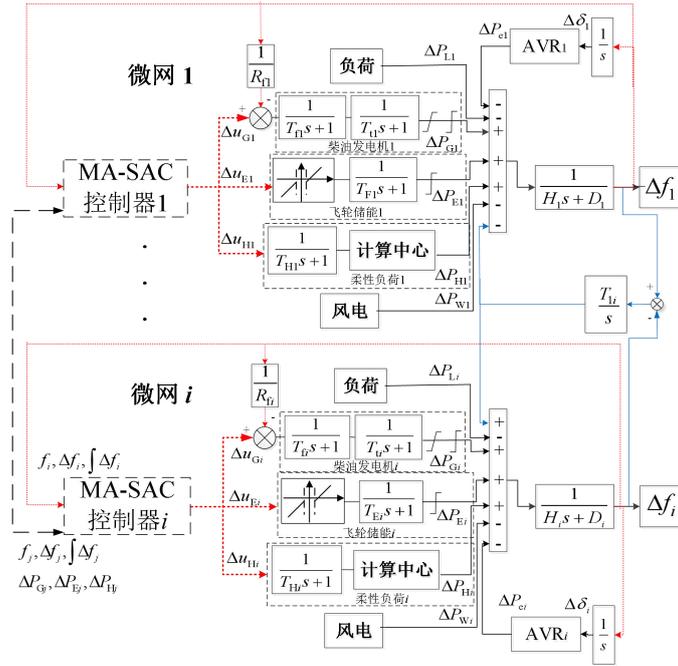


图 A2 多微电网 LFC 和 AVR 组合模型

Fig.A2 Multi-microgrid LFC and AVR combination model

6) 梯度下降优化算法。

损失函数的最小值通过梯度下降的方法得到，采用自适应随机梯度下降算法对神经网络参数进行更新，以动作网络参数 ϕ 为例，梯度更新方程为：

$$\begin{cases} m_l = \beta_1 m_{l-1} + (1 - \beta_1) \nabla_{\phi} J_{\pi_l}(\phi_l) \\ v_l = \beta_2 v_{l-1} + (1 - \beta_2) [\nabla_{\phi} J_{\pi_l}(\phi_l)]^2 \\ \phi_{l+1} = \phi_l - \frac{\xi m_l}{\sqrt{v_l} + d} \end{cases} \quad (A6)$$

式中： l 为迭代次数； β_1 为梯度移动平均衰减率； β_2 为平方梯度移动平均衰减率； d 为分母偏移； ξ 为学习率。

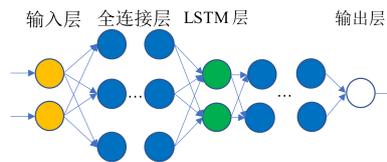


图 A3 神经网络结构

Fig.A3 Neural network structure

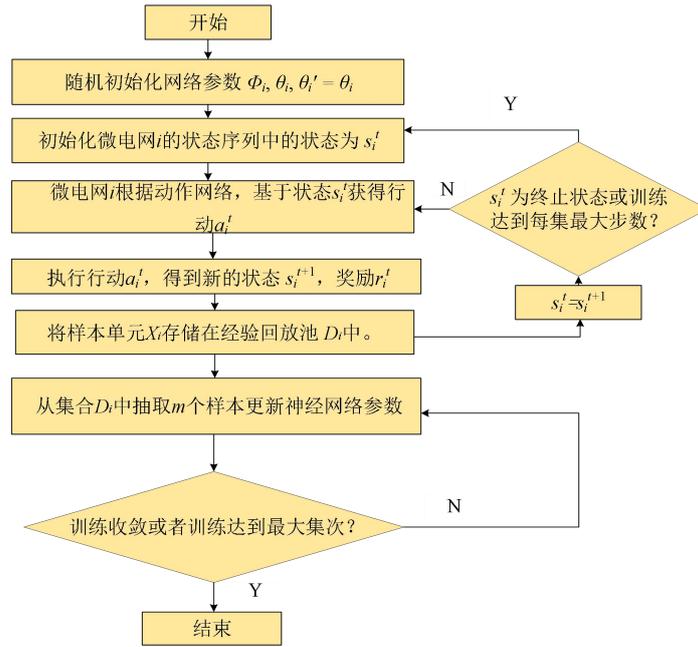


图 A4 基于 MA-SAC 算法多微电网训练流程
Fig.A4 Multi-microgrid training process based on MA-SAC algorithm

附录 B

表 B1 多微电网系统模型参数

Table B1 Multi-microgrid system model parameters

| 设备 | 参数 | 参数值 | 设备 | 参数 | 参数值 | |
|------------------------------|------------------|-----------------------|--------------------------|-----------------------|-----------------------|--------|
| 2 个微电网的柴油发电机 | K_{Ai} | 10 | 2 个微电网的飞轮储能 | T_{Fi}/s | 1 | |
| | K_{Ei}, K_{Si} | 1 | | $\Delta P_{E\max i}$ | 0.025 | |
| | K_{Gi} | 0.8 | | $\Delta P_{E\min i}$ | -0.025 | |
| | T_{Ai} | 0.1 | | σ | 0.01 | |
| | T_{Ei} | 0.4 | 微电网 1 柔性负荷 | T_{H1}/s | 1 | |
| | T_{Gi} | 1.4 | | $\Delta P_{H1\max,p}$ | 0.015 | |
| | T_{Si} | 0.05 | | $\Delta P_{H1\min,d}$ | -0.015 | |
| | $K_{1,i}$ | 0.02 | | M_1 | 500 | |
| | 2 个微电网的柴油发电机 | $K_{2,i}$ | -0.1 | 微电网 2 柔性负荷 | T_{H2}/s | 1 |
| | | $K_{3,i}$ | 1.3 | | $\Delta P_{H2\max,p}$ | 0.012 |
| | | $K_{4,i}$ | 1.4 | | $\Delta P_{H2\max,d}$ | -0.012 |
| | | P_{Si} | 1.5 | | M_2 | 400 |
| | 其他 | $R_G(\text{Hz/p.u.})$ | 0.05 | 其他 | $H_i(\text{p.u./Hz})$ | 14.22 |
| | | T_{Ei}/s | 10 | | $D_i(\text{p.u./Hz})$ | 0 |
| T_{Ei}/s | | 0.1 | $T_{12}(\text{p.u./Hz})$ | | 0.545 | |
| $\Delta P_{G\max i}$ | | 0.03 | f_n/Hz | | 50 | |
| $\Delta P_{G\min i}$ | | -0.03 | | | | |
| $\delta_{pf}(\text{p.u./s})$ | | 0.1 | | | | |
| $\delta_{df}(\text{p.u./s})$ | | -1 | | | | |

表 B2 频率控制器智能体网络结构参数

Table B2 Frequency controller agent network structure parameters

| 网络 | 路径 | 网络层数 b | 第 e 层神经元数目 u_e |
|------|---------|----------|----------------------------|
| 评价网络 | 状态路径 | 4 | $u_1=400, u_2=u_3=u_4=300$ |
| | 动作路径 | 3 | $u_1=u_2=u_3=300$ |
| | 公共路径 | 2 | $u_1=u_2=300$ |
| 动作网络 | 状态路径 | 3 | $u_1=400, u_2=u_3=300$ |
| | 平均值输出路径 | 3 | $u_1=u_2=u_3=300$ |
| | 标准差输出路径 | 3 | $u_1=u_2=u_3=300$ |

表 B3 电压控制器智能体网络结构参数

Table B3 Voltage controller agent network structure parameters

| 网络 | 路径 | 网络层数 b | 第 e 层神经元数目 u_e |
|------|---------|----------|----------------------|
| 评价网络 | 状态路径 | 3 | $u_1=64, u_2=u_3=48$ |
| | 动作路径 | 3 | $u_1=u_2=u_3=48$ |
| | 公共路径 | 1 | $u_1=48$ |
| 动作网络 | 状态路径 | 2 | $u_1=u_2=48$ |
| | 平均值输出路径 | 2 | $u_1=u_2=48$ |
| | 标准差输出路径 | 2 | $u_1=u_2=48$ |

表 B4 频率控制器智能体训练参数

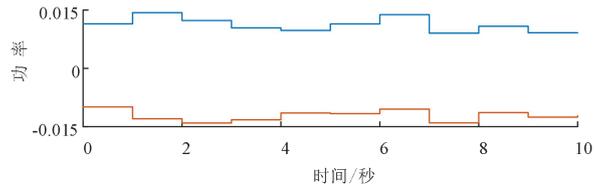
Table B4 Frequency controller agent training parameters

| 训练参数 | 参数值 | 训练参数 | 参数值 | 训练参数 | 参数值 |
|---------|-------|-------------------|-----------|--------------------|---------|
| η | 5 | β_1 | 0.9 | τ | 0.001 |
| μ_1 | -5 | d | 10^{-8} | D | 10^6 |
| μ_2 | -10 | λ_{actor} | 10^{-5} | m | 32 |
| μ_3 | -15 | γ | 0.999 | t_s/s | 0.05 |
| μ_4 | -20 | L | 0.065 | β_2 | 0.999 |
| B | -100 | T | 450 | ρ | 1 |
| ζ | 0.001 | T_m | 200 | λ_{critic} | 0.000 2 |

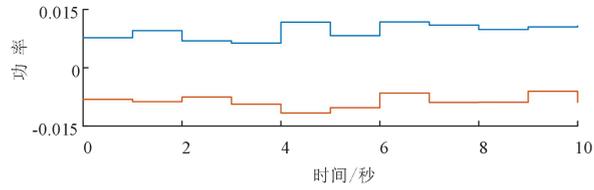
表 B5 电压控制器智能体训练参数

Table B5 Voltage controller agent training parameters

| 训练参数 | 参数值 | 训练参数 | 参数值 | 训练参数 | 参数值 |
|--------------------|-----------|----------------------|------------|----------------------|----------|
| χ_1 | -5 | ζ | 0.001 | τ | 0.001 |
| φ | 10 | β_1 | 0.9 | D | 10^6 |
| χ_2 | -5 | d | 10^{-8} | m | 32 |
| χ_3 | -10 | K_{pmin}, K_{imin} | 0.04, 0.04 | t_s/s | 0.05 |
| C | -200 | γ | 0.999 99 | β_2 | 0.999 |
| λ_{actor} | 10^{-5} | T | 1800 | ρ | 1 |
| λ_{critic} | 0.000 2 | T_m | 100 | K_{pmax}, K_{imax} | 0.6, 0.6 |



(a)微电网 1 柔性负荷输出功率增量约束



(b)微电网 2 柔性负荷输出功率增量约束

— 功率增量上限 — 功率增量下限

图 B1 柔性负荷输出功率增量约束

Fig.B1 Flexible load output power increment constraint

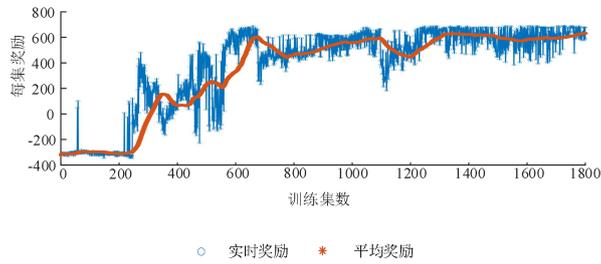
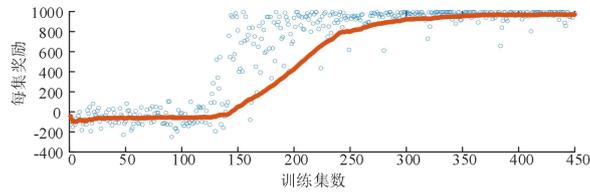
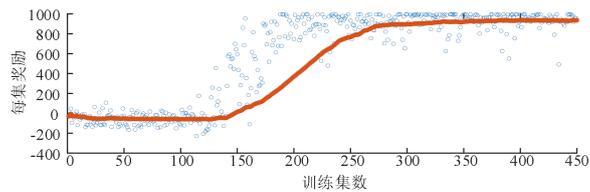


图 B2 基于 SAC 算法优化电压 PI 控制器参数训练流程

Fig.B2 Optimized voltage PI controller parameter training process based on SAC algorithm



(a)微电网 1 训练过程图



(b)微电网 2 训练过程图

○ 实时奖励 * 平均奖励

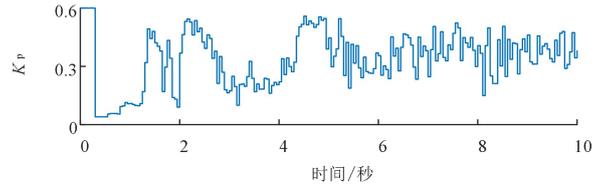
图 B3 多微电网频率控制器训练过程图

Fig.B3 Multi-microgrid frequency controller training process diagram

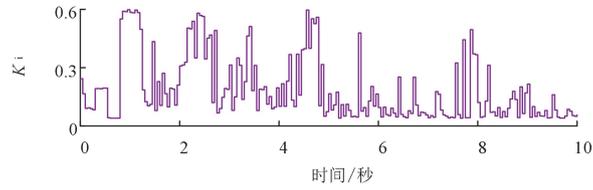
表 B6 电压仿真结果

Table B6 Simulative results of voltage

| 控制器类别 | 上升时间/s | 调节时间/s | 超调量 |
|----------|--------|--------|-------|
| PI | 0.98 | 3.58 | 28.3% |
| Fuzzy-PI | 4.25 | 4.25 | 0 |
| SAC-PI | 1.04 | 1.04 | 3.3% |

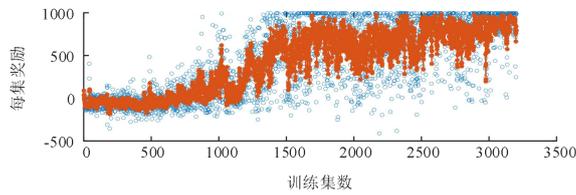


(a)比例系数

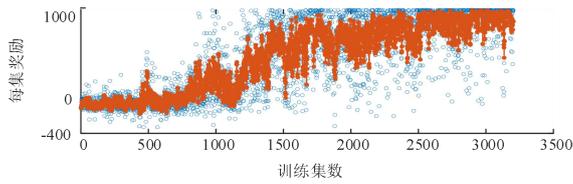


(b)积分系数

图 B4 PI 控制器优化参数
Fig.B4 PI controller optimization parameters



(a)微电网 1 训练过程图



(b)微电网 2 训练过程图

○ 实时奖励 * 平均奖励

图 B5 MA-DDPG 训练过程
Fig.B5 MA-DDPG training process

表 B7 场景 1 扰动下频率仿真结果

Table B7 Frequency simulation results under disturbance in Scenario1

| 控制器类别 | 绝对值最大值/Hz | 平均值/Hz | 标准差 |
|----------------|-----------|--------------|-----------|
| 微电网 1-PI | 0.027 76 | -0.000 509 2 | 0.010 160 |
| 微电网 1-Fuzzy-PI | 0.022 76 | -0.001 868 0 | 0.007 842 |
| 微电网 1-MA-DDPG | 0.018 85 | 0.001 524 0 | 0.005 306 |
| 微电网 1-MA-SAC | 0.015 75 | -0.000 589 6 | 0.005 041 |
| 微电网 2-PI | 0.044 28 | -0.000 259 5 | 0.014 410 |
| 微电网 2-Fuzzy-PI | 0.042 37 | -0.003 550 0 | 0.009 216 |
| 微电网 2-MA-DDPG | 0.055 41 | -0.000 950 6 | 0.012 920 |
| 微电网 2- MA-SAC | 0.038 05 | -0.001 199 0 | 0.006 556 |

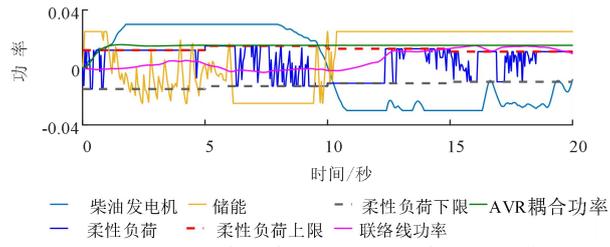


图 B6 场景 1 中微电网 1 各机组输出功率增量

Fig.B6 Output power increase of each unit of microgrid 1 under Scenario 1

表 B8 场景 2 扰动下仿真结果

Table B8 Frequency simulative results under disturbance in Scenario 2

| 控制器类别 | 绝对值最大值/Hz | 平均值/Hz | 标准差 |
|---------------|-----------|--------------|-----------|
| 微电网 1-PI | 0.104 2 | -0.000 918 4 | 0.045 00 |
| 微电网 1-Fuzzy | 0.067 0 | -0.002 716 0 | 0.019 29 |
| 微电网 1-MA-DDPG | 0.033 11 | 0.001 373 0 | 0.007 464 |
| 微电网 1-MA-SAC | 0.022 45 | -0.000 128 5 | 0.005 334 |
| 微电网 2-PI | 0.023 07 | 0.000 684 6 | 0.009 534 |
| 微电网 2-Fuzzy | 0.017 50 | -0.002 371 0 | 0.004 546 |
| 微电网 2-MA-DDPG | 0.019 16 | 0.002 497 0 | 0.004 791 |
| 微电网 2-MA-SAC | 0.021 89 | 0.003 423 0 | 0.005 147 |

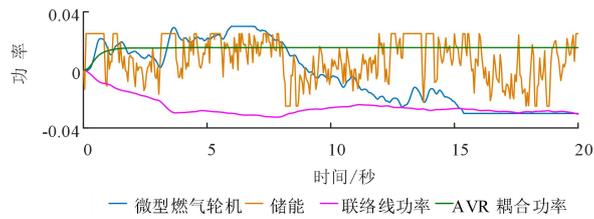


图 B7 场景 2 中微电网 1 各机组输出功率增量

Fig.B7 Output power increase of each unit of microgrid 1 under Scenario 2.