

DOI: 10.19783/j.cnki.pspc.241593

基于深度强化学习的电力系统暂态稳定快关汽门 紧急控制策略

孙正龙¹, 陈威翰¹, 耿鑫地², 王思璇¹, 杨浩¹, 潘超¹, 蔡国伟¹

(1. 现代电力系统仿真控制与绿色电能新技术教育部重点实验室(东北电力大学), 吉林 吉林 132012;

2. 国网河北省电力有限公司衡水供电分公司, 河北 衡水 053000)

摘要: 快关汽门是提升电力系统暂态稳定性的经典控制方式之一, 但其控制变量具有高维度、离散化的特点, 且参数整定不合理将引发功角后续摇摆失稳, 控制策略制定的复杂性致使快关汽门难以在线应用与实时决策。为此, 提出了基于深度强化学习的快关汽门控制决策方法。首先, 构建基于深度强化学习的紧急快关汽门决策制定框架。然后, 将快关汽门控制问题转化为马尔可夫决策过程(Markov decision process, MDP), 以综合考虑最优稳定控制效果及最小化稳控代价为目标设置奖励函数, 并采用近端策略优化(proximal policy optimization, PPO)算法求解, 得到快关策略的合理配置。最后, 通过改进的电科院 SG-77 系统验证所提方法的有效性。仿真结果表明所提方法在保证快关汽门策略有效性与时效性的同时, 可实现在预案式失配场景下作出正确决策, 提高了电力系统的暂态稳定性和动态响应能力。

关键词: 快关汽门决策; 深度强化学习; 暂态稳定; 近端策略优化算法

Fast valving emergency control strategy for power system transient stability based on deep reinforcement learning

SUN Zhenglong¹, CHEN Weihang¹, GENG Xindi², WANG Sixuan¹, YANG Hao¹, PAN Chao¹, CAI Guowei¹

(1. Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education (Northeast Electric Power University), Jilin 132012, China; 2. Hengshui Power Supply Branch, State Grid Hebei Electric Power Co., Ltd., Hengshui 053000, China)

Abstract: Fast valving is one of the classic control methods to improve transient stability in power systems. However, its control variables are high-dimensional and discrete, and improper parameter tuning may trigger subsequent power angle oscillations and instability. The complexity of the control strategy development makes it difficult to apply fast valving closure in online and real-time decision-making. To address this challenge, a fast valving control decision method based on deep reinforcement learning is proposed. First, a deep reinforcement learning-based emergency fast valving decision-making framework is constructed. Then, the fast valving control problem is transformed into a Markov decision process (MDP). A reward function is designed to balance optimal stability control performance and minimized control cost, and the proximal policy optimization (PPO) algorithm is used to solve it, yielding a rational configuration of the fast valving strategy. Finally, the effectiveness of the proposed method is verified using the improved SG-77 system developed by CEPRI. Simulation results show that the proposed method ensures both the effectiveness and timeliness of the fast valving strategy, enabling correct decision-making under mismatched contingency scenarios, and improves transient stability and dynamic response capability of power systems.

This work is supported by the National Natural Science Foundation of China (No. 52277084).

Key words: fast valving decision-making; deep reinforcement learning; transient stability; proximal strategy optimization algorithm

基金项目: 国家自然科学基金项目资助(52277084); 吉林省国际科技合作项目资助(20230402074GH)

0 引言

随着新能源并网装机量的快速增长、电力电子设备的广泛应用以及特高压交直流输电工程陆续建成投运,我国电力系统规模不断扩大,网架结构愈发复杂,运行工况更加多变^[1-3],现代电力系统暂态稳定面临巨大挑战^[4-6]。

随着电网技术的不断发展,电力系统的保护与控制策略也在逐步演进。作为传统的紧急控制,切机和切负荷在某些情况下能够维护系统的稳定性,但机组和负荷再并网的时间较长,可能会使电力系统的效率降低或增加运营成本。同时,尽管现代电网正积极地引入并大规模部署可再生能源,以满足可持续发展和环境保护的需求,但汽轮机在电力系统中仍然扮演着不可或缺的角色^[7]。作为传统的大型发电设备,汽轮机为电网提供了稳定的基荷电力和必要的系统惯性,这对于维护电网的暂态稳定和应对突发事件至关重要。汽轮机快关汽门作为一种有效的控制手段,在实际工程上已经得到了广泛应用,如沙角电厂快关试验^[8-9]、双鸭山发电厂原苏制210 MW 机组快关装置的试验^[10]等,利用汽轮发电机组的数字式电液控制系统,预先设定控制方式实现机组汽门快关,从而迅速调整汽轮机的输出,以响应电网的实时需求和变化^[11],通过增大汽轮机机械功率短暂下跌过程中的减速面积,以增强系统暂态稳定性^[12]。快关汽门防止了不必要的机组停机,减少了对机组冲击造成的不利影响,避免过多切机影响系统稳定,提高了电力系统的灵活性和韧性^[13]。低惯量可再生能源与储能系统的广泛应用,也使得使用汽轮机快关汽门维持电力系统暂态稳定变得更加重要。

快关汽门的效果依赖其参数整定是否合适,传统的汽轮机制造商会基于汽轮机的设计特性与预期的操作条件提供几个预设的快关汽门参数,快关汽门控制可从制造商预定义的3~5个初始设置中选择其一^[7],并基于机械功率与电磁功率的不平衡逻辑启动快关汽门控制^[14-15],文献[16]在制定快关汽门控制方案时引入李雅普诺夫理论。文献[17]在原控制启动逻辑上增加了对转子转速的判断,但其只优化了控制逻辑,极有可能造成超调或者调整不足。为了弥补此类不足,可以通过协调快关汽门和控制制动电阻改善发电机电压特性,有效抑制转子角度的摆动,提高电力系统的暂态稳定性^[18-19],然而,多种稳定控制方式的协调配合又增加了策略制定的复杂性。

除了优化控制逻辑和整合稳控资源,对快关汽门进行最佳参数整定,同样是解决该问题的关键方

向。文献[20]提出了一种计算汽轮机功率输出的方法,为准确确定快关汽门的控制参数提供了依据。文献[21-22]分别采用庞特里雅金最大值原理和多维插值全局优化的方法进行测试,结果验证了所提方法的有效性,但是对于大型电网,传统的最优控制策略求解方法的计算时间可能会大幅增加,难以保证控制的时效性^[23]。

近年来,数据驱动的方法为解决实际电力系统控制问题提供了广阔的应用前景,通过数据驱动方法可以有效应对电力系统中的不确定性因素。文献[24]提出了一种基于配电网量测数据的有源配电网无模型电压控制方法,实现了对运行电压的实时控制优化。文献[25]提出了数据驱动框架,采用Bagging 算法对模型进行实时更新,实现对电力系统暂态稳定的精准评估。深度强化学习(deep reinforcement learning, DRL)在电力系统稳定控制中应用广泛,主要包括电压稳定^[26-28]、功角稳定^[29-30]和频率稳定^[31-32]。针对电压失稳问题,文献[26-28]分别基于深度Q网络(deep Q-network, DQN)、深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法以及反事实多智能体策略梯度(counterfactual multi-agent policy gradient, COMA)算法,实现了低压减载紧急控制决策。针对暂态功角稳定问题,文献[29-30]分别提出了基于DRL的紧急切机与切负荷决策方法,并验证了方法的可行性和正确性,其中文献[30]将紧急切机控制相关知识经验融合到智能体的训练过程中,提升了智能体的学习效率与决策效果。文献[31]针对频率稳定问题,利用单智能体和多智能体方法制定切负荷策略。文献[32]中提出基于多智能体深度强化学习的风场频率控制方法,解决了风电并网后系统频率调节的问题。目前尚未涉及有关利用强化学习解决快关汽门控制的问题。

目前离线紧急控制决策制定方式存在效率不高、严重依赖专家经验等问题,为更加合理地匹配快关汽门的控制资源,提升其控制效果,本文提出了一种基于深度强化学习的智能快关汽门决策制定方法。首先,分析了快关汽门对系统暂态稳定性的影响,从能量角度解释了快关汽门控制引起的功角后续失稳问题。其次,针对快关汽门决策制定问题,构建了智能决策框架,设计了快关汽门控制的DRL模型,为更好地观测功角后续摆动状态,将能量函数加入奖励函数。然后,通过与电力系统仿真环境的交互,训练基于近端策略优化的智能体,智能体接收环境中的观测结果并提供控制策略。最后,通过改进的电科院SG-77系统进行仿真验证,结果表明所提方法能够快速生成有效的快关汽门控制策

T_A 一定的条件下, 汽门关闭持续时间 T_B 对功角后续摇摆稳定性的影响, 故设 t_2 时刻的系统暂态能量为 V_{t_2} 。在汽门持续快关期间, 机械功率保持恒定, 系统的暂态能量也保持不变, 因此汽门重新开启时的系统暂态能量与汽门完全关闭时保持一致。

将 t_3 到 t_4 的汽门开启过程划分为 n 等份, 即 $t_3 = t_3^1$ 、 $t_4 = t_3^n$, t_3^i 的机械功率由 P_{m3}^i 变为 P_{m3}^{i+1} , 同时考虑到功角与角速度不能突变的性质, t_3^i 时刻系统的能量表示为

$$V_{t_3^i}^{i+} = V_{t_3^i}^{i-} + P_{m3}^i(\delta_{t_3^i}^i - \delta_{s3}^i) + P_e(\cos \delta_{t_3^i}^i - \cos \delta_{s3}^i) - P_{m3}^{i+1}(\delta_{t_3^i}^i - \delta_{s3}^{i+1}) - P_e(\cos \delta_{t_3^i}^i - \cos \delta_{s3}^{i+1}) \quad (4)$$

式中: $V_{t_3^i}^{i-}$ 为 t_3^i 时刻前系统的能量; $\delta_{t_3^i}^i$ 为 t_3^i 时刻发电机功角; δ_{s3}^i 为 t_3^i 时刻发电机功角的稳定平衡点; δ_{s3}^{i+1} 为 t_3^{i+1} 时刻发电机功角的稳定平衡点; $\delta_{t_3^i}^{i+1}$ 为 t_3^{i+1} 时刻发电机功角。

在 t_3^i 到 $t_3^{(i+1)-}$ 期间, 机械功率与稳定平衡点保持不变, $V_{t_3^i}^{i+} = V_{t_3^{(i+1)-}}^{i+}$ 。考虑到 $V_{t_3^1}^{1-} = V_{t_2}$, 进一步推导 t_4 即 t_3^n 时刻系统暂态能量的表达式为

$$V_{t_4} = V_{t_2} + (P_{m2}\delta_{t_3} - P_{m4}\delta_{t_4}) + \sum_{i=1}^{n-1} P_{m3}^i(\delta_{t_3^i}^{i+1} - \delta_{t_3^i}^i) + (P_{m4}\delta_{s4} - P_{m2}\delta_{s2}) + P_e(\cos \delta_{s4} - \cos \delta_{s2}) \quad (5)$$

式中: δ_{t_3} 和 δ_{t_4} 分别为 t_3 和 t_4 时刻发电机功角; δ_{s2} 和 δ_{s4} 分别为 t_2 和 t_4 时刻发电机功角的稳定平衡点; P_{m2} 和 P_{m4} 分别为 t_2 和 t_4 时刻机械功率。

V_{t_2} 由汽门关闭过程决定; $(P_{t_4}\delta_{s4} - P_{t_2}\delta_{s2})$ 项与 $P_e(\cos \delta_{s4} - \cos \delta_{s2})$ 项由汽门重新开启过程前后的机械功率和系统的稳定平衡点决定, 与汽门关闭持续时间无关; 其余两项与汽门重新开启前后的功角相关。当不考虑阻尼等外部因素时, 功角的振荡周期是恒定的。当汽门开启时间 T_C 一定时, δ_{t_4} 与 δ_{t_3} 存在一定的映射关系, 故 V_{t_4} 的值取决于 δ_{t_3} , 即取决于汽门关闭持续时间的长短。由于功角呈周期性变化, V_{t_4} 图像也表现出相应的周期性特征。若汽门持续关闭时间设置合理, V_{t_4} 取极小值或者比较小的值且未超出稳定域时, 功角后续摇摆是稳定的。反之, 如果设置不合理, V_{t_4} 超出了稳定域, 则后续摇摆将会失去稳定。

2 基于深度强化学习的电力系统暂态稳定快关汽门控制

2.1 马尔可夫决策与紧急控制

暂态稳定紧急控制旨在以最小代价确保系统的

稳定运行。在紧急控制中, 需要依据系统实时状态, 生成离散化的稳定控制动作序列。然而, 随着控制动作组合数量增加, 控制变量维度也显著增加, 使得紧急控制问题变得复杂, 成为一个包含离散变量的多维序列决策问题。同时, 每一项决策不仅直接影响系统的即时状态, 还对系统的长期稳定性产生影响。

马尔可夫决策过程通常由一个五元组 $\{S, A, P, R, \gamma\}$ 构成。其中: S 为状态空间; A 为动作空间; P 为转移概率; R 为奖励函数; γ 为用于计算累积奖励的折扣因子。

智能体与电力系统环境的强化学习交互过程如图 3 所示。智能体是决策的主体, t 时刻将观测的环境状态 s_t 作为输入, 根据策略 π 选择输出的动作 a_t , 环境在接受动作后转移到状态 s_{t+1} , 同时对智能体的动作行为进行评价, 给出反馈奖励 r_{t+1} 。智能体通过与环境的持续交互, 逐步优化其动作策略, 实现累积奖励最大化, 最终得到特定目标下的最优控制策略 π 。

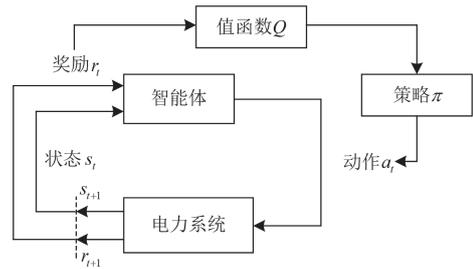


图 3 智能体与电力系统环境的强化学习交互过程

Fig. 3 Reinforcement learning interaction process between the agent and power system environment

强化学习的目标是最大化累积奖励, 智能体的回报反映一定时间内各时刻智能体累积奖励 R 的总和, 定义回报 G_t 为 t 时刻后马尔科夫链的总折扣奖励, 其表达式为

$$G_t = R_t + \gamma R_{t+1} + \gamma^2 R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (6)$$

式中: k 为 t 时刻后智能体与环境交互的次数。

动作价值函数(即值函数 Q)用于评估智能体在遵循某一策略时, 在状态 s 下执行动作 a 所能获得的期望累积回报。

$$Q_{\pi}(s, a) = E_{\pi} [G_t | S_t = s, A_t = a] \quad (7)$$

在马尔可夫决策过程中, 智能体获取最优控制策略 π 的关键在于准确识别最优动作价值函数 $Q_{\pi}(s, a)$, 即当采取某动作得到的函数值最大时, 令选择该动作的概率为 1, 其余动作的概率为 0。

$$\pi(a|s) = \begin{cases} 1, & a = \arg \max Q_{\pi}(s, a) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

为了应用强化学习方法, 可将电力系统稳定控制过程建模为马尔可夫决策过程。严格来说, 电力系统往往难以完全满足马尔可夫性质。但是, 如果在短时间内推导稳定控制策略, 电力系统未来状态将受当前状态影响, 从而形成马尔可夫过程^[37-38]。因此, 基于无模型强化学习深度分析运行环境, 探索环境特性, 获取控制策略, 可规避仿真过程中简化模型和不确定性因素导致的电网控制决策效果偏差, 提升决策的准确性和有效性。

本文将紧急快关汽门控制问题建模为 MDP 问题。对于汽门控制而言, 状态空间 S 代表电力系统的当前状态, 用于判断系统是否处于稳定状态以及是否需要进行汽门控制。动作空间 A 主要包括参与稳控的目标机组与可供选择的快关参数。在快关汽门控制决策中, 转移概率 P 作为强化学习中的关键参数, 不仅描述了智能体决策动作对系统暂态稳定的影响, 还为智能体的决策优化与控制策略的制定提供了基础。通过引入转移概率对紧急快关汽门控制进行建模, 智能体能够更精准地把握系统状态, 进而制定出更为有效的控制策略。奖励函数 R 则考虑了系统暂态稳定性的提升以及成本控制等多个方面, 当系统成功恢复到稳定状态时, 智能体获得较高的奖励。折扣因子 γ 决定了未来奖励在当前决策中的重要性, 快关汽门决策往往需要考虑长期的稳定性, 即不仅要关注角首摆稳定性, 还要考虑后续摆的稳定。因此, 选取较大的折扣因子有利于智能体考虑策略执行后系统可能面临的潜在风险。

2.2 基于近端策略优化的快关汽门决策制定框架

2.2.1 近端策略优化算法

近端策略优化 (proximal policy optimization, PPO) 算法是一种基于 Actor-Critic 框架的深度强化学习策略梯度方法, 图 4 为 Actor 网络与 Critic 网络结构图。

Actor 网络采用全连接层神经网络构建, 以环境观测状态为输入, 通过隐藏层的非线性映射拟合输出动作的概率分布, 再由输出层神经网络输出对应的动作概率向量。Actor 网络输出的策略如式(9)所示。

$$\pi_{\theta}(a|s) = \frac{e^{f_{\theta}(a,s;\theta)}}{\sum_{a'} e^{f_{\theta}(a',s;\theta)}} \quad (9)$$

式中: $\pi_{\theta}(a|s)$ 为给定状态 s 下采取动作 a 的概率; $f_{\theta}(a,s;\theta)$ 为网络对于状态 s 和动作 a 的输出得分, θ 为神经网络参数; a' 表示智能体所有可能执行的动作。

动作。Actor 网络最终选择概率最大动作作为输出。

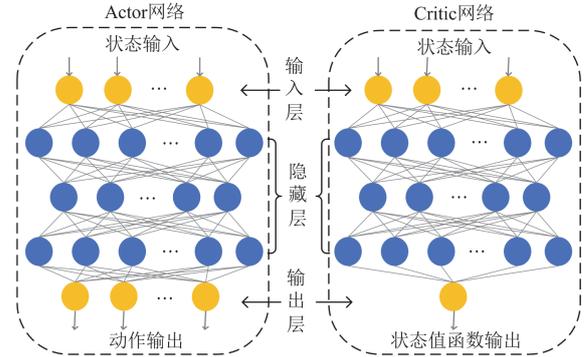


图 4 神经网络结构图

Fig. 4 Neural network structure diagram

同理, Critic 网络也采用全连接层神经网络构建, 以环境观测状态作为输入, 并输出对应的状态价值估计 $V(s_t)$, 结合实际获得的即时奖励 r 和下一个状态的价值估计 $V(s_{t+1})$ 得到优势函数 $A(s_t, a_t)$, 并反馈给 Actor 网络, 进而指导 Actor 网络选择更优的动作, 优势函数 $A(s_t, a_t)$ 公式为

$$A(s_t, a_t) \approx r + \gamma V(s_{t+1}) - V(s_t) \quad (10)$$

PPO 算法基于信赖域策略优化 (trust region policy optimization, TRPO) 算法发展而来, TRPO 算法采用 KL 散度 (Kullback-Leibler divergence) 来限制策略更新的幅度, 确保学习的稳定性, TRPO 算法优化目标及 KL 散度约束分别如式(11)、式(12)所示。

$$\underset{\theta}{\text{maximize}} \quad \hat{\mathbb{E}}_t \left[\frac{\pi_{\theta}(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} A(s_t, a_t) \right] \quad (11)$$

$$\text{s.t.} \quad \hat{\mathbb{E}}_t [\text{KL}[\pi_{\theta_{\text{old}}}(\cdot|s_t), \pi_{\theta}(\cdot|s_t)]] \leq \delta \quad (12)$$

式中: π_{θ} 表示采用新的网络参数 θ 时的新策略; $\pi_{\theta_{\text{old}}}$ 表示用旧的网络参数 θ_{old} 时的旧策略; $\hat{\mathbb{E}}_t[\cdot]$ 表示在有限批次样本上的经验平均值; KL 表示散度; δ 表示置信度, 用于限制策略更新的幅度。

TRPO 算法需要求解一个带有 KL 散度约束的优化问题, 不仅增加了计算负担, 还使算法的实现变得更加复杂。相比之下, PPO 算法引入了裁剪率, 通过控制新旧策略比率的变化幅度, 简化计算过程的同时, 保持了策略更新的稳定性。

$$L(\theta) = \hat{\mathbb{E}}_t [\min(r_t(\theta)A(s_t, a_t), \text{clip}(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon)A(s_t, a_t))] \quad (13)$$

式中: $L(\theta)$ 为 PPO 算法的目标函数; $r_t(\theta)$ 为新旧策略之间的概率比例, $r_t(\theta) = \pi_{\theta}(a_t|s_t)/\pi_{\theta_{\text{old}}}(a_t|s_t)$; $\text{clip}(\cdot)$ 为裁剪函数; ε 为裁剪参数, 且 $\varepsilon > 0$, 裁剪参数用于定义裁剪阈值, 将新旧策略变化控制在

$[1 - \varepsilon, 1 + \varepsilon]$ 内, 在保证策略更新安全性的同时, 有效降低了该算法的实现难度与计算成本, 从而显著提升了实用性和普适性。

PPO 算法训练流程图如图 5 所示, 其标准步骤如下。

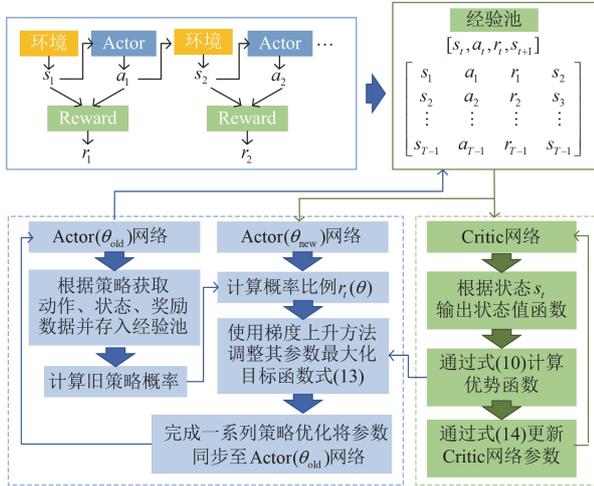


图 5 PPO 算法训练流程图
Fig. 5 PPO algorithm training flowchart

步骤 1: 对网络参数进行初始化设置。

步骤 2: 智能体通过观察环境状态 s_t 并基于 Actor(θ_{old}) 网络给出动作的概率分布, 执行概率最大的动作 a_t , 得到环境反馈的奖励 r_t 和下一个状态 s_{t+1} 。记录当前时刻的动作、状态和奖励 (s_t, a_t, r_t, s_{t+1}), 并存入经验池, 迭代进行直至经验池存满。

步骤 3: 从经验池中随机提取历史数据, 即状态集和动作集, 输入到 Critic 网络中, 同时 Critic 网络输出优势函数 $A(s, a)$, 再根据式(14)计算 Critic 网络的损失, 用以更新 Critic 网络参数。

$$L_{loss} = \frac{1}{N} \sum_{i=1}^N [A(s, a)]^2 \quad (14)$$

式中: N 表示从经验池中随机提取的样本数量。

步骤 4: Actor(θ_{new}) 网络和 Actor(θ_{old}) 网络接收经验池中状态 s 和动作 a 的输出集合, 得到相应的概率分布 π_{θ} 和 $\pi_{\theta_{old}}$ 。并根据式(11)更新 Actor(θ_{new}) 网络参数。

步骤 5: 多次重复步骤 4 后, 将 Actor(θ_{new}) 网络参数同步至 Actor(θ_{new}) 网络。

步骤 2—步骤 5 为一轮训练(Episode)。

2.2.2 整体框架设计

为了提供合适的快关汽门控制方案, 保证计算速度和系统功角稳定, 本文采用深度强化学习方法获取暂态稳定快关汽门控制策略。所提方法的总体

框架如图 6 所示。

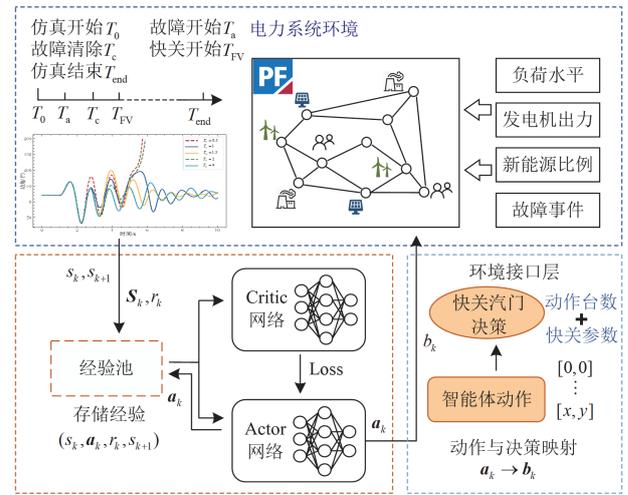


图 6 整体架构设计图
Fig. 6 Design diagram of overall architecture

框架上半部分是需要进行快关汽门决策制定的电力系统环境, 框架下半部分是具有 Actor-Critic 架构的 DRL 智能体, 整体为智能体与电力系统环境交互过程。智能体由 Actor 神经网络与 Critic 神经网络构成。Actor 网络生成第 k 步交互时系统的状态 s_k 下每个动作的概率分布, Critic 网络负责估计累积奖励或值函数并矫正 Actor 网络。当前结构下, 智能体在学习最佳动作策略的同时, 还能评估其动作的长期效果。智能体训练过程涉及大量的数据和复杂的计算, 需要在模拟的电力系统环境中不断尝试和学习, 最终获取最优的控制策略。电力系统环境部分是智能体学习和测试的场所, 模拟了实际系统的运行工况, 包括负荷水平、发电机输出、新能源发电以及故障事件等, 旨在尽可能真实地反映实际电力系统的复杂性与不确定性。智能体通过与电力系统环境的交互来了解不同动作对系统稳定性的影响。

在训练过程中, 智能体在每回合根据当前的状态信息作出决策, 并执行时域仿真。若系统功角保持稳定, 则该回合结束; 否则, 智能体将继续探索新的决策方案。通过与电力系统环境的持续交互训练, 智能体不断优化决策能力。最终, 智能体在失稳工况中寻求保障系统功角稳定的最优快关汽门控制策略。

在测试阶段, 完成训练后, 将智能体置于各种预设和未预设的故障场景中, 以测试其对电力系统暂态稳定的控制能力。测试结果将被用来评估智能体的决策效率和电力系统响应能力, 进一步验证快关汽门策略的暂态稳定控制效果。

最后, 该框架不仅限于离线训练和测试。涉及在线应用时, 智能体可通过在线训练更新策略, 将电网的真实运行情况反馈给智能体进一步训练, 以适应电网的实际状况; 离线预训练将提升智能体在线训练效率。

2.3 紧急快关汽门控制的 DRL 模型设计

在将快关汽门控制问题建模为 DRL 问题的过程中, 智能体状态空间、动作空间和奖励函数的设置直接影响 DRL 模型的性能。合理设计状态、奖励、动作可有效提升智能体探索最优策略的能力。

2.3.1 状态空间

智能体的输入观测状态 \mathbf{S}_k 包含反映功角稳定和快关汽门决策的关键信息。本文选择故障后和稳控后对应时间 $[T_a, T_{fv1}) \cup [T_{fv1}, T_{o1}) \cup \dots \cup [T_{fvk}, T_{ok}]$ 内的发电机相对功角矩阵 δ_k 以及发电机端口能量 W_k 作为智能体观测对象, 观测状态 \mathbf{S}_k 如式(15)所示。

$$\mathbf{S}_k = [\delta_k, W_k] \quad (15)$$

$$\delta_k = [\delta_1^k, \dots, \delta_m^k] = \begin{bmatrix} \delta_{1,0}^k & \dots & \delta_{m,0}^k \\ \vdots & & \vdots \\ \delta_{1,t}^k & \dots & \delta_{m,t}^k \end{bmatrix} \quad (16)$$

$$W_k = [W_1^k, \dots, W_m^k] = \begin{bmatrix} W_{1,0}^k & \dots & W_{m,0}^k \\ \vdots & & \vdots \\ W_{1,t}^k & \dots & W_{m,t}^k \end{bmatrix} \quad (17)$$

式中: $\delta_{m,t}^k$ 为第 k 步下同步机 m 在 t 时刻的相对功角; $W_{m,t}^k$ 为第 k 步下同步机 m 在 t 时刻的端口能量。

2.3.2 动作空间

在进行快关汽门决策时, 智能体的决策实际上是由参与稳控的机组数、快关控制的目标机组以及快关汽门控制参数组成的三元组, 智能体通过训练能够参与稳控发电机并调整其控制参数, 通过与电网环境的交互来确定最优解。考虑到实际电力系统中快关汽门控制的应用通常是针对单个机组的, 故设置参与快关控制的机组数量上限为 2 台。由于快关汽门控制具有暂态特性——快关汽门关闭时间 T_A 主要影响功角首摆, 且越快首摆越稳定, 焖缸时间 T_B 主要影响功角的后摆, 故动作空间中快关参数仅考虑焖缸时间。Actor 网络输出动作 \mathbf{a}_k 在环境接口层通过动作-决策映射对应决策 \mathbf{b}_k 。

$$\mathbf{a}_k = [a_1^k, a_2^k, \dots, a_n^k, a_{n+1}^k] \quad (18)$$

$$\mathbf{b}_k = [b_1^k, b_2^k, \dots, b_{m_g}^k, b_{para}^k] \quad (19)$$

式中: a_n^k 表示智能体在第 k 步输出的第 n 个动作; b_i^k 表示第 k 步下选择第 m 台参与快关控制的机组,

共设置 m_g 台机组可参与快关控制; b_{para}^k 表示第 k 步下参与快关控制的机组快关参数。

2.3.3 奖励函数

智能体与环境的交互过程中需要制定一个指标来衡量系统的暂态功角稳定程度, 能量函数不仅能够表征系统的暂态稳定程度, 还可清晰反映决策前后发电机端口能量的变化趋势, 进而揭示发电机功角的后续摇摆情况。故本文设计奖励函数包含电力系统暂态稳定指标(transient stability index, TSI)和能量奖励函数两部分, 如式(20)所示。

$$R = \begin{cases} -p, & T_{TSI} < \frac{1}{3} \\ \lambda_{TSI} \times T_{TSI} + \lambda_w \times \Delta W, & T_{TSI} \geq \frac{1}{3} \end{cases} \quad (20)$$

$$T_{TSI} = \frac{360^\circ - \Delta\delta_{\max}}{360^\circ + \Delta\delta_{\max}} \quad (21)$$

$$W(s_t) = \sum_{i=1}^m \int P_i d\theta_i + Q_i d(\ln U_i) \quad (22)$$

$$\Delta W = |W(s_t) - W_{ref}| \quad (23)$$

式中: p 为稳控后失稳的惩罚值; λ_{TSI} 为快关汽门决策效果奖励系数; λ_w 为能量函数惩罚系数; ΔW 为端口能量之和 $W(s_t)$ 与设定能量基准值 W_{ref} 的差值的绝对值。当 $T_{TSI} \geq 1/3$ 即 $\Delta\delta_{\max} \leq 180^\circ$ 时, 系统判定为稳定, 反之判定为失稳; $\Delta\delta_{\max}$ 为发电机的最大功角差; m 为机组的总台数; P_i 和 Q_i 分别为第 i 台发电机发出的有功功率和无功功率; U_i 和 θ_i 分别为第 i 台发电机母线的电压幅值和相角。

3 算例分析

3.1 环境与智能体参数设置

3.1.1 仿真算例及场景设置

本文基于改进的电科院 SG-77 系统展开算例分析, 如图 7 所示。

该系统共包含 197 节点(500 kV 主网架为 85 个节点)、118 条交流输电线路、8 条直流输电线路、14 台发电机、25 个负荷。发电机采用 5 阶模型, 配有调速器、自动电压调节器以及电力系统稳定器。系统各类型负荷占比为: 恒阻抗负荷占比 53%, 恒电流负荷占比 34%, 恒功率负荷占比 13%。

本文算例采用的系统为含新能源的交直流混联系统, 新能源装机集中于系统区域 B, 根据以往仿真经验, 线路发生三相永久性 $N-1$ 故障易引发区域 B 同步机组功角失稳。具体设置预想故障工况包括暂态功角失稳工况 108 个, 并将其全部用于智能体

的训练, 详见表 1。

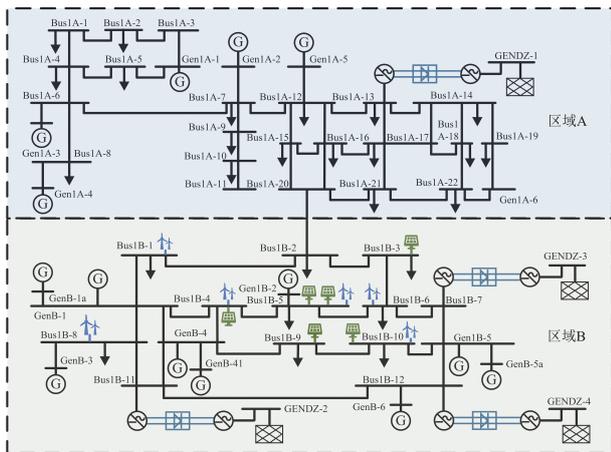


图 7 改进的电科院 SG-77 电力系统

Fig. 7 Improved CEPRI SG-77 power system

表 1 预想故障工况设置

Table 1 Setting of anticipated fault conditions

具体设置	
故障类型	三相短路接地故障
故障线路 (交流线路)	Bus1B-9 至 Bus1B-10(双回)、GenB-1 至 GenB-4(双回)、Bus1B-2 至 Bus1B-3(双回)、 GenB-5 至 Bus1B-5(双回)、Bus1B-4 至 Bus1B-5(双回)、Bus1B-9 至 GenB-4(双回)
故障持续时间/s	0.18、0.19
故障位置	25%、50%、75%
运行方式	高峰负荷、平峰负荷、低谷负荷

3.1.2 智能体训练参数设置

本文基于 DIgSILENT 软件搭建电力系统仿真环境。基于 PyTorch 搭建 DRL 模型并开展训练, 采用 Python 语言开展相关编程。本文智能体 Actor 网络和 Critic 网络的神经元结构均为[256, 256, 256], 训练总轮次设置为 20 000, 学习率为 0.000 636, 折扣因子 γ 为 0.9, 裁剪参数 ϵ 为 0.3。

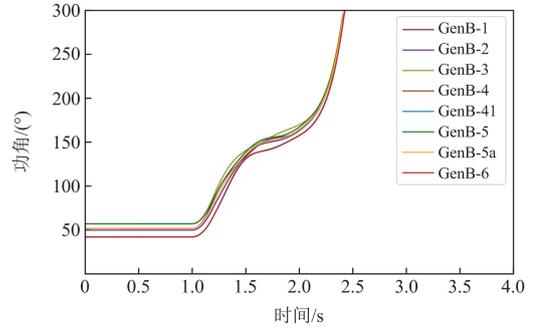
设定合适的惩罚系数能够引导智能体实现预期奖励的最大化, 故设置快关汽门决策效果奖励系数 $\lambda_{TSI} = 600$, 能量函数惩罚系数 $\lambda_w = -50$, 稳控后失稳的惩罚值 $p = 50$ 。

3.2 智能快关汽门的训练及决策结果

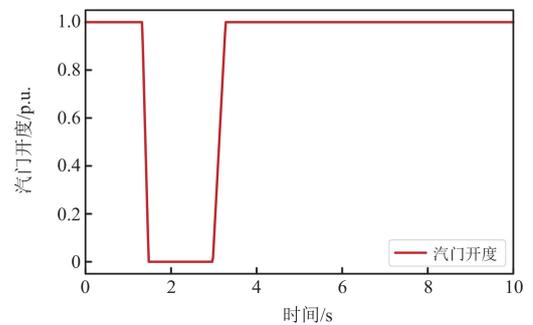
3.2.1 智能体快关汽门决策实例

以一个实际测试工况(高峰负荷的运行方式, 线路 GenB-1 至 GenB-4 一回线(线路 AC40)25%处发生三相短路接地故障, 故障持续 0.18 s 后切除)验证智能体输出的快关汽门决策效果。在没有施加快关决策之前, B 区域机组功角失稳, 如图 8(a)所示。智能体依据功角和端口能量的输入信息, 输出相应的

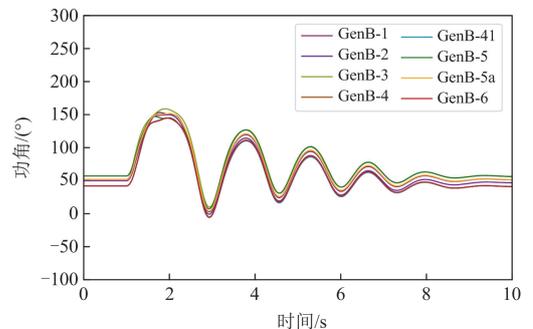
控制决策。选择对同步机 GenB-5, 执行快关汽门控制, 汽门关闭持续时间为 0.9 s, 汽门开度曲线如图 8(b)所示。快关汽门决策后最终系统恢复功角稳定, 如图 8(c)所示。



(a) 快关汽门前功角失稳



(b) 汽门开度



(c) 快关汽门后功角稳定

图 8 智能体的快关汽门决策制定

Fig. 8 Fast valving decision-making by the agent

图 9 提供了不同快关策略情况下, 系统最大功角差的摇摆情况。由此可见经训练的智能体在保证功角前摆稳定的同时, 能够确定合适的汽门持续关闭时间保证功角后续摆动的稳定。

本文统计了在 3 种运行方式下系统发生故障(线路 AC40 的 25%处发生三相短路接地故障, 故障持续 0.18 s 后切除)后, 采取与不采取快关汽门控制时的极限切除时间如表 2 所示。

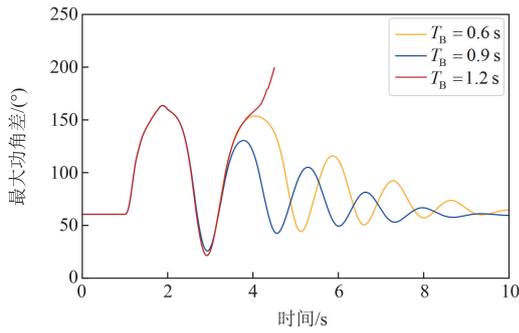


图 9 不同快关策略下的最大功角差

Fig. 9 Maximum rotor angle difference with different fast valving strategies

表 2 有无快关汽门控制时故障的极限切除时间
Table 2 Maximum clearing time for faults with/without fast valving control

运行方式	故障极限切除时间 (无快关汽门控制)	故障极限切除时间 (有快关汽门控制)
高峰负荷	0.17	0.194
平峰负荷	0.17	0.197
低谷负荷	0.174	0.206

在高峰负荷运行方式下, 若系统发生该故障后执行快关汽门控制, 可将极限切除时间提高至 0.194 s, 延长了故障切除时间 0.024 s; 平峰负荷运行方式下稳控执行后故障极限切除时间延长了 0.027 s; 低谷负荷运行方式下稳控执行后极限切除时间延长了 0.022 s。

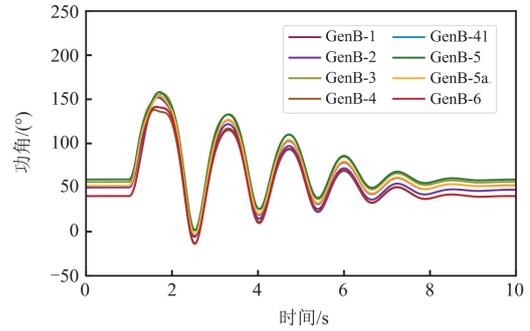
3.2.2 预案式失配案例

为验证本文所提控制策略的鲁棒性与适应性, 本节以故障场景 Bus1B-9 至 GenB-4 一回线(线路 AC121)的 50%处发生三相短路接地故障, 故障持续 0.18 s 后切除为例, 验证智能体制定的快关汽门策略的控制效果。

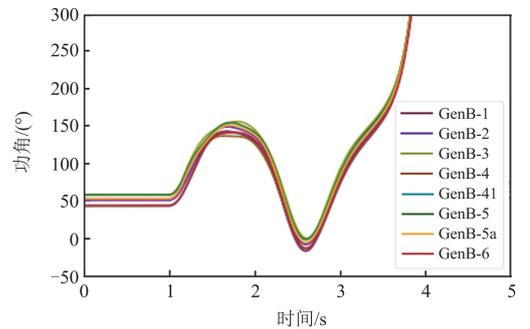
在平峰负荷运行方式下, 针对该故障的预案式稳控方案为: 同步机 GenB-4 执行快关汽门控制, 汽门关闭持续时间为 0.5 s。传统预案式方法下各机组功角曲线如图 10 所示。

在高峰负荷运行方式下, 伴随负荷水平的提升, 系统内发电机组出力也随之增加。当同步机 GenB-4 发生同样故障时, 同步机组加速面积远大于平峰负荷时机组的加速面积。执行相同的稳控方案后, 得到机组的功角曲线如图 10(b)所示, 同步机组功角失稳。由仿真结果可知, 传统的预案式控制方法无法完全覆盖所有的故障场景和电网运行状态, 系统负荷水平的变化会引起机组出力改变, 进而导致稳定

控制策略失配, 机组功角失去稳定。



(a) 平峰运行方式



(b) 高峰运行方式

图 10 传统预案式方法下的各机组功角曲线

Fig. 10 Rotor angle curves of each generator with traditional pre-planned method

针对高峰负荷运行方式下发生上述故障, 智能体给出的快关策略为: 同步机 GenB-4 执行快关汽门控制, 汽门关闭持续时间为 1 s。执行该策略后, 得到机组的功角曲线如图 11 所示, 系统功角可以保持稳定。由此可见, 本文所提方法通过学习历史数据, 能够自动识别和适应新的故障, 降低预案式失配的可能性。

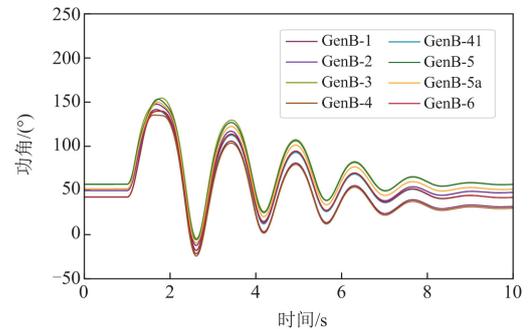


图 11 DRL 方法下的各机组功角曲线

Fig. 11 Rotor angle curves of each generator with DRL method

3.2.3 多机组协调决策实例

在电力系统中, 面对较严重的故障, 单机控制

的能力往往不够。且单机控制范围有限，难以独立应对大规模电网中的复杂故障情况。为使系统迅速恢复稳定，须考虑多机组之间的协调联合控制。本节以一个实际工况(平峰负荷的运行方式，线路 Bus1B-4 至 Bus1B-5 一回线(线路 AC112)的 75%处发生三相短路接地故障，故障持续 0.19 s 后切除)为例，验证在多机配合协调决策中本文方法的有效性。

针对该故障，智能体作出的决策为：同步机 GenB-1 执行快关汽门控制，汽门关闭持续时间为 0.8 s；GenB-2 执行快关汽门控制，汽门关闭持续时间为 1 s。单台机组控制与多台机配合控制效果对比如图 12 所示。

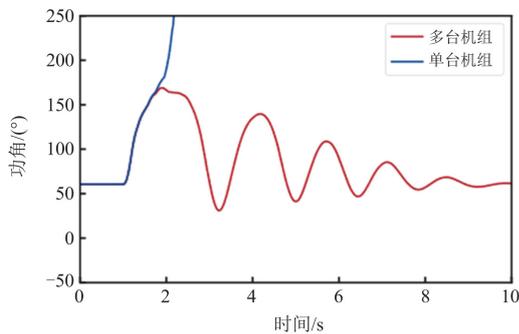


图 12 决策效果对比

Fig. 12 Comparison of decision-making effects

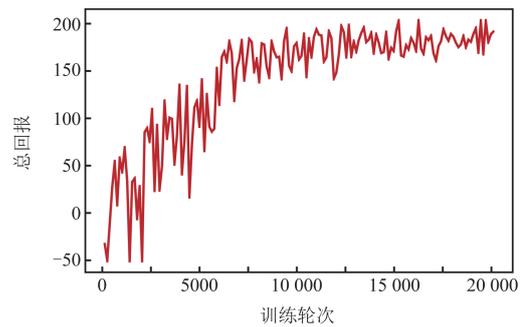
3.2.4 智能体的整体决策效果

为评估本文所提基于 PPO 算法的暂态稳定快关汽门紧急控制方法的整体决策效果，选择目前已广泛应用的基于 DQN 算法的紧急控制决策与本文所提方法进行对比，从算法收敛性与决策效果两方面验证所提算法的控制效果。

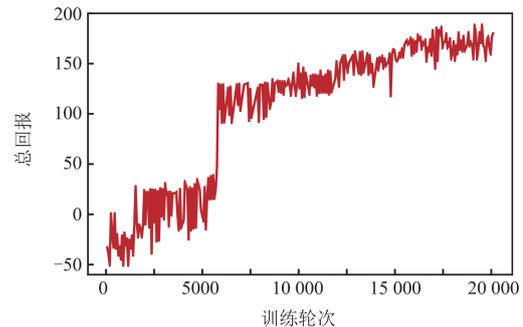
基于 PPO 算法与 DQN 算法的智能体在训练阶段获得的累计奖励总回报如图 13 所示。由图 13 可知，智能体在训练初始阶段的平均奖励值较低，随着训练过程的不断推进，智能体的性能逐渐提升，最终达到了一个稳定的高奖励状态，此时智能体在学习过程中不断试错并逐渐适应环境，后期实现更优的控制效果。通过对比两种方法的奖励曲线，基于 PPO 算法的决策方法所引入的裁剪机制能够有效保证训练过程的稳定性，并能够生成累计奖励值更高的策略。

决策效果对比如图 14 所示。本文通过对比两种方法在 108 种故障场景下所生成决策对应的系统最大功角差分布，验证了其决策有效性。由图 14 可知，基于 PPO 算法的决策方法给出的稳控策略可以将系统的最大功角差控制在更低水平，决策效果优于基于 DQN 算法的决策方法给出的控制策略，基于 PPO 算法的决策方法能够有效学习并找到大型电力

系统复杂决策问题的最优策略。



(a) 基于PPO算法的决策方法的奖励函数曲线



(b) 基于DQN算法的决策方法的奖励函数曲线

图 13 两种方法下的 Reward 收敛曲线

Fig. 13 Reward convergence curves with two methods

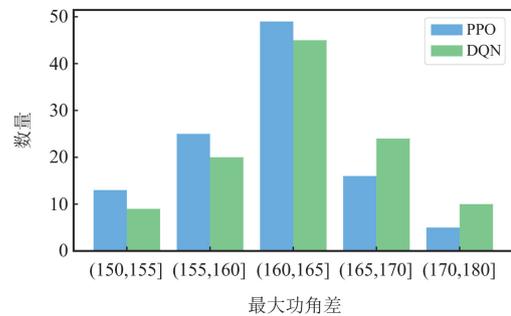


图 14 决策效果对比

Fig. 14 Comparison of decision-making effects

本文基于 108 种故障场景以及 1401 种稳控方式，开展强化学习与电力系统联合仿真训练。训练后的智能体可在几毫秒内生成控制措施，能够有效应对本文提出的所有故障场景，满足工程中暂态稳定的在线控制需求。为验证智能体在暂态稳定控制方面的优越性能，提取了 3 种运行方式下共 108 个故障场景的最大功角差和累计奖励。如图 15 所示，在基于 DRL 的快关汽门控制下，最大功角差保持在 150°~175°之间，在既定的安全范围内。智能体能够在极短时间内响应系统状态变化，生成准确有效的控制策略。

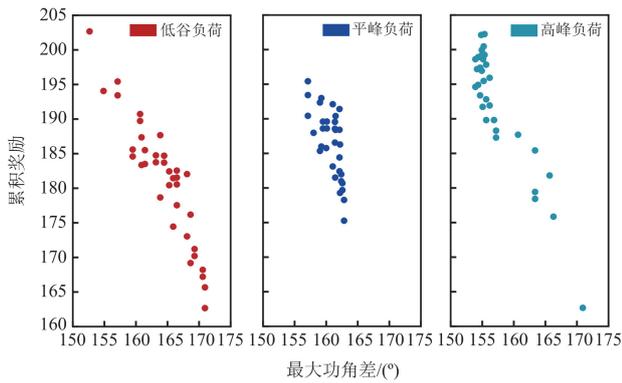


图 15 DRL 方法下的决策效果图

Fig. 15 Decision effect diagram based on DRL

3.2.5 模型泛化性分析

为验证所提模型具备泛化能力, 本文选择未参与训练的新故障场景来验证紧急快关汽门控制策略的有效性。线路 AC113 故障下的快关汽门控制策略如表 3 所示。

表 3 线路 AC113 故障下的快关汽门控制策略
Table 3 Fast-valve closure control strategy under AC113 line fault

AC113 line fault		
运行方式	故障场景	策略
高峰负荷	AC113 的 50% 处发生三相短路接地故障, 故障持续 0.18 s 后切除	GenB-6 进行快关汽门, $T_B = 0.7$ s
低谷负荷		GenB-6 进行快关汽门, $T_B = 0.7$ s
平峰负荷		GenB-5 进行快关汽门, $T_B = 0.7$ s

3 种运行方式下线路 Bus1B-3 至 Bus1B-6 一回线(线路 AC113)故障后进行快关汽门控制后的系统最大功角差曲线如图 16 所示。由图 16 可知, 经过训练的智能体能够对未经历过的故障场景给出合适的稳定控制措施, 保障系统暂态功角稳定性, 从而证明本文所提 DRL 模型具有泛化能力。

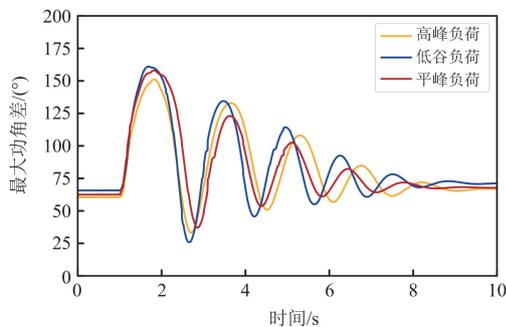


图 16 线路 AC113 故障下进行稳控后的最大功角差曲线

Fig. 16 Maximum rotor angle difference curve after stability control under AC113 line fault

4 结语

本文通过引入深度强化学习法, 提出了一种新的电力系统快关汽门控制策略。通过构建并训练基于马尔可夫决策过程的电力系统暂态稳定紧急快关汽门决策框架, 获取快关汽门控制策略, 并基于改进的电科院 SG-77 电力系统开展仿真分析。结果表明, 所提方法能在特定条件下快速提供有效的快关汽门策略, 有助于提高电力系统的暂态稳定性和动态响应能力。

然而, 需要说明的是, 本文方法仍在初步阶段, 其在不同电力系统环境中的普适性和长期效果尚需进一步验证, 未来的工作将集中在以下几个方面。

1) 模型和算法的优化: 继续改进深度强化学习模型, 优化算法参数, 以提高其在更多电力系统环境中的适应性和效能。

2) 广泛的实证测试: 在不同类型的电力系统中进行更多实证测试, 以评估所提方法的实际效果和适用范围。

综上, 本文为电力系统暂态稳定控制领域提供了一种可能的新方法, 但其应用前景和长期效果仍需在后续的研究中进一步探索和验证。

参考文献

- [1] 索之闻, 李晖, 张锋, 等. 高比例新能源直流送端系统分布式调相机优化配置[J]. 电力系统保护与控制, 2022, 50(23): 133-141.
- [2] 韩璐, 尹纯亚, 戴晨, 等. 高比例新能源送端系统暂态电压运行风险分析[J]. 电力系统保护与控制, 2024, 52(1): 23-34.
- [3] 乔鑫悦, 王丽馨. 数据驱动的新能源电力系统静态电压安全评估方法研究[J]. 东北电力大学学报, 2024, 44(5): 73-79.
- [4] 李楠, 张家恒. 基于 XGboost-DF 的电力系统暂态稳

定评估方法[J]. 电测与仪表, 2024, 61(10): 119-127.

LI Nan, ZHANG Jiaheng. A transient stability assessment method of power system based on XGboost-DF[J]. *Electrical Measurement & Instrumentation*, 2024, 61(10): 119-127.

[5] 刘颂凯, 党喜, 崔梓琪, 等. 针对样本类不平衡的深度残差网络电力系统暂态稳定评估方法[J]. *智慧电力*, 2024, 52(1): 116-123.

LIU Songkai, DANG Xi, CUI Ziqi, et al. Transient stability evaluation method for power systems with deep residual network considering class imbalance of samples[J]. *Smart Power*, 2024, 52(1): 116-123.

[6] 陆旭, 张理寅, 李更丰, 等. 基于内嵌物理知识卷积神经网络的电力系统暂态稳定评估[J]. *电力系统自动化*, 2024, 48(9): 107-119.

LU Xu, ZHANG Liyin, LI Gengfeng, et al. Transient stability assessment of power system based on physics informed convolution neural network[J]. *Automation of Electric Power Systems*, 2024, 48(9): 107-119.

[7] RUBAN N, KIEVETS A, ANDREEV M, et al. Turbine fast valving setting method based on the hybrid simulation approach[J]. *Energies*, 2023, 16(4).

[8] 谭伟. 660MW 单元机组“快关汽门”功能试验分析[J]. *广东电力*, 1998, 11(2): 38-41, 68.

TAN Wei. Analysis of “fast valving” function test for 660MW unit[J]. *Guangdong Electric Power*, 1998, 11(2): 38-41, 68.

[9] 蒙定中. 沙角 C 厂 2 号汽轮发电机组的快关试验[J]. *电网技术*, 1999, 23(12): 11-15, 23.

MENG Dingzhong. Fast valving test of steam turbine generator set No. 2 in Sha Jiao‘C’ power plant[J]. *Power System Technology*, 1999, 23(12): 11-15, 23.

[10] 都兴有. 关于快控汽门(快关)技术研讨介绍[J]. *中国电力*, 1999, 43(6): 25-29.

DU Xingyou. Introduction to the study of early valve actuation (fast valving) technology[J]. *Electric Power*, 1999, 43(6): 25-29.

[11] PATEL R, BHATTI T, KOTHARI D. Improvement of power system transient stability by coordinated operation of fast valving and braking resistor[J]. *IEE Proceedings on Generation, Transmission and Distribution*, 2003, 150(3): 311-316.

[12] KARADY G, MOHAMED M. Improving transient stability using fast valving based on tracking rotor-angle and active power[C]// *IEEE Power Engineering Society Summer Meeting*, July 21-25, 2002, Chicago, USA: 1576-1581.

[13] DONG Chao, WANG Bin, SUN Yongjian, et al. On effective rapid startup of 300 MW steam turbine unit[C]// *2020 Chinese Control and Decision Conference (CCDC)*, May 23-25, 2020, Hefei, China: 803-808.

[14] 韩炜, 李蔚, 吴文健, 等. 汽轮机汽门快控功能仿真与参数辨识[J]. *动力工程学报*, 2019, 39(10): 804-810.

HAN Wei, LI Wei, WU Wenjian, et al. simulation and parameter identification for fast valving control of steam turbines[J]. *Journal of Chinese Society of Power Engineering*, 2019, 39(10): 804-810.

[15] 李文云. 汽轮发电机组快关汽门控制方案及实施[J]. *云南电业*, 2005, 11(3): 38-39.

LI Wenyun. Fast valving control scheme and implementation for steam turbine generator units[J]. *Yunnan Electric Power*, 2005, 11(3): 38-39.

[16] LIU G, LIN X, YANG Q, et al. Investigation of turbine valving control with Lyapunov theory[C]// *1997 Fourth International Conference on Advances in Power System Control, Operation and Management*, November 11-14, 1997, Hongkong, China: 505-508.

[17] CAI Yu, LI Wei, ZHANG Baorui, et al. Reliable analysis on fast valving of ultra-supercritical unit under transient fault conditions[C]// *ASME Power Conference*, June 26-30, 2017, Charlotte, USA.

[18] PATEL R, BHATTI T, KOTHARI D. A modified approach to transient stability enhancement with fast valving and braking resistor applications[J]. *International Journal of Electrical Power & Energy Systems*, 2006, 28(10): 729-738.

[19] YOON D, JANG G, MOON Y. A study of the application of fast valving and braking resistor for an intelligent SPS[J]. *International Journal of Electrical Power & Energy Systems*, 2013, 53: 818-823.

[20] MOISEICHENKOV A, SENYUK M, KOVALENKO P, et al. The technique of calculating a steam turbine power for characterization of the turbine fast valving[C]// *2021 International Conference on Electrotechnical Complexes and Systems (ICOECS)*, September 23-26, 2021, Ufa, Russian Federation: 79-82.

[21] 曹全喜, 于达仁, 于松海, 等. 采用最大值原理的汽轮发电机快速汽门控制[J]. *电力系统自动化*, 1993, 17(8): 17-22.

CAO Quanxi, YU Daren, YU Songhai, et al. Fast valve control of steam turbine generator using maximum principle[J]. *Automation of Electric Power Systems*, 1993, 17(8): 17-22.

[22] XIE Xueshu, LU Yan. Optimization control of fast turbine valve[C]// *Proceedings of 32nd IEEE Conference on Decision and Control*, December 15-17, 1993, San Antonio, USA: 3742-3743.

- [23] 邱建, 朱煜昆, 张建新, 等. 基于安全强化学习的电网稳控策略智能生成方法[J]. 电力系统保护与控制, 2024, 52(10): 147-155.
QIU Jian, ZHU Yukun, ZHANG Jianxin, et al. Intelligent generation method of power system stability control strategy based on safe reinforcement learning[J]. Power System Protection and Control, 2024, 52(10): 147-155.
- [24] 祝泳琪, 刘友波, 唐志远, 等. 基于数据驱动预测控制的有源配电网电压控制策略[J]. 电力系统自动化, 2024, 48(20): 100-108.
ZHU Yongqi, LIU Youbo, TANG Zhiyuan, et al. Voltage control strategy for active distribution network based on data-enabled predictive control[J]. Automation of Electric Power Systems, 2024, 48(20): 100-108.
- [25] 赵冬梅, 谢家康, 王闯, 等. 基于 Bagging 集成学习的电力系统暂态稳定在线评估[J]. 电力系统保护与控制, 2022, 50(8): 1-10.
ZHAO Dongmei, XIE Jiakang, WANG Chuang, et al. On-line transient stability assessment of a power system based on Bagging ensemble learning[J]. Power System Protection and Control, 2022, 50(8): 1-10.
- [26] HUANG Qihua, HUANG Renke, HAO Weituo, et al. Adaptive power system emergency control using deep reinforcement learning[J]. IEEE Transactions on Smart Grid, 2019, 11(2): 1171-1182.
- [27] LI Jian, CHEN Sheng, WANG Xinying, et al. Load shedding control strategy in power grid emergency state based on deep reinforcement learning[J]. CSEE Journal of Power and Energy Systems, 2021, 8(4): 1175-1182.
- [28] 张梓泉, 崔明建, 张程彬, 等. 基于反事实多智能体强化学习和有功无功协同控制的配电网电压优化[J]. 电力系统保护与控制, 2024, 52(18): 76-86.
ZHANG Zixiao, CUI Mingjian, ZHANG Chengbin, et al. Active and reactive power coordinated optimal voltage control of a distribution network based on counterfactual multi-agent reinforcement learning[J]. Power System Protection and Control, 2024, 52(18): 76-86.
- [29] XIE Jian, SUN Wei. Distributional deep reinforcement learning-based emergency frequency control[J]. IEEE Transactions on Power Systems, 2022, 37(4): 2720-2730.
- [30] 李舟平, 曾令康, 姚伟, 等. 基于知识融合和深度强化学习的智能紧急切机决策[J]. 中国电机工程学报, 2024, 44(5): 1675-1688.
LI Zhouping, ZENG Linggang, YAO Wei, et al. Intelligent emergency generator rejection schemes based on knowledge fusion and deep reinforcement learning[J]. Proceedings of the CSEE, 2024, 44(5): 1675-1688.
- [31] LIANG Yanchang, ZHAO Xiaowei, SUN Li, et al. A multiagent reinforcement learning approach for wind farm frequency control[J]. IEEE Transactions on Industrial Informatics, 2023, 19(2): 1725-1734.
- [32] CHEN Chunyu, CUI Mingjian, LI Fangxing, et al. Model-free emergency frequency control based on reinforcement learning[J]. IEEE Transactions on Industrial Informatics, 2021, 17(4): 2336-2346.
- [33] 黄其励. 电力工程师手册(电气卷)[M]. 北京: 中国电力出版社, 2000.
- [34] EDWARDS L, GREGORY J, OSBORN D, et al. Turbine fast valving to aid system stability: benefits and other considerations[J]. IEEE Transactions on Power Systems, 1986, 1(1): 143-153.
- [35] 陈珏, 戴克健, 杨昆. 汽轮发电机汽门快关时序对系统暂态稳定及轴系扭振的影响[J]. 华北电力学院学报, 1994, 20(1): 27-33.
CHEN Yu, DAI Kejian, YANG Kun. The impact of fast valving timing of steam turbine generators on system transient stability and shaft torsional vibration[J]. Journal of North China Institute of Electric Power, 1994, 20(1): 27-33.
- [36] 袁宇春, 张保会. 汽门快关的开启时刻对系统稳定性影响的研究[J]. 继电器, 1998, 25(2): 6-9.
YUAN Yuchun, ZHANG Baohui. Research on the impact of the opening moment of quick-closing throttle valves on system stability[J]. Relay, 1998, 25(2): 6-9.
- [37] 李宏浩, 张沛, 刘墨. 基于深度强化学习的暂态稳定紧急控制决策方法[J]. 电力系统自动化, 2023, 47(5): 144-152.
LI Honghao, ZHANG Pei, LIU Zhao. Decision-making method for transient stability emergency control based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2023, 47(5): 144-152.
- [38] ZENG Hongtai, ZHOU Yanzhen, GUO Qinglai, et al. Distributed deep reinforcement learning-based approach for fast preventive control considering transient stability constraints[J]. CSEE Journal of Power and Energy Systems, 2023, 9(1): 197-208.

收稿日期: 2024-11-29; 修回日期: 2025-06-05

作者简介:

孙正龙(1988—), 男, 通信作者, 博士, 教授, 博士生导师, 研究方向为新能源电力系统稳定与控制; E-mail: nedusunzl@necpu.edu.cn

陈威翰(1998—), 男, 博士研究生, 研究方向为新能源电力系统稳定与控制; E-mail: 2202100051@necpu.edu.cn

耿鑫地(1999—), 男, 硕士, 研究方向为新能源电力系统稳定与控制. E-mail: 1669765109@qq.com

(编辑 许威)