

引用格式:高琴,徐光虎,夏尚学,等.基于深度强化学习的电力系统紧急切机稳控策略生成方法[J].电力科学与技术学报,2025,40(1):39-46.

Citation:GAO Qin,XU Guanghu,XIA Shangxue, et al.Policy generation method for power system stability control during emergent tripping of unit based on deep reinforcement learning[J]. Journal of Electric Power Science and Technology,2025,40(1):39-46.

基于深度强化学习的电力系统紧急切机 稳控策略生成方法

高琴¹,徐光虎¹,夏尚学²,杨欢欢¹,赵青春²,黄河¹

(1.中国南方电网有限责任公司,广东广州 510663;2.南京南瑞继保电气有限公司,江苏南京 211102)

摘要:电力系统快速发展的同时也改变着电力系统的结构,使得系统稳定机理变得更加复杂。为解决新能源电力系统存在的功角稳定问题,提出基于深度强化学习的电力系统紧急切机稳控策略生成方法。首先,归纳并提出电力系统紧急控制切机动作策略以及涉及的安全约束,并将电力系统稳控模型转换为马尔科夫决策过程,再采用特征评估与斯皮尔曼(Spearman)等级相关系数方法筛选出最典型的特征数据;随后,为提高稳控策略智能体的训练效率,提出基于深度确定性策略梯度(deep deterministic policy gradient,DDPG)算法的稳控策略训练框架;最后,在IEEE 39节点系统和某实际电网中进行测试验证。研究结果显示,所提方法能够根据系统的运行状态和对故障的响应,自动调整生成切机稳控策略,在决策效果和效率方面都表现出更好的性能。

关键词:新能源电力系统;稳控策略;强化学习;深度确定性策略梯度算法;马尔科夫模型

DOI:10.19781/j.issn.1673-9140.2025.01.004 中图分类号:TM723 文章编号:1673-9140(2025)01-0039-08

Policy generation method for power system stability control during emergent tripping of unit based on deep reinforcement learning

GAO Qin¹, XU Guanghu¹, XIA Shangxue², YANG Huanhuan¹, ZHAO Qingchun², HUANG He¹

(1.China Southern Power Grid Co., Ltd., Guangzhou 510663, China; 2.NR Electric Co., Ltd., Nanjing 211102, China)

Abstract: The rapid development of the power system has been changing its structure, making the system stability mechanism more complex. To ensure power angle stability in the new energy power system, a policy generation method for power system stability control during emergent tripping of units based on deep reinforcement learning is proposed. Firstly, the policies for emergent tripping of units of the power system are summarized, as well as the security constraints involved. The power system stability control model is then transformed into a Markov decision process. Next, the most typical feature data are selected by feature evaluation and the Spearman rank correlation coefficient method. To improve the training efficiency of the intelligent agent of the stability control policy, a training framework for the stability control policy based on the deep deterministic policy gradient (DDPG) is put forward. Finally, tests are performed in the IEEE 39 node system and a real-life power grid for validation. The results show that the proposed method can automatically adjust and generate a stability control policy for tripping of units according to the system's running states and fault responses, confirming its enhanced decision-making effect and efficiency.

Key words: new energy power system; stability control policy; reinforcement learning; depth deterministic policy gradient algorithm; Markov model

收稿日期:2024-04-02;修回日期:2024-07-24

基金项目:中国南方电网有限责任公司科技项目(000005KK52220027)

通信作者:夏尚学(1978—),男,硕士,教授级工程师,主要从事电网安全稳定控制方面的研究;E-mail:p344717084_1006x@126.com

高比例新能源和电力电子设备接入新型电力系统给电网安全稳定带来了严峻的挑战^[1-3],导致传统稳控策略处理系统失稳问题变得更加严重^[4-5]。为了应对新型电力系统运行的需求,迫切需要建立一种适用于电网潮流多样性、稳定模态多样性和机理复杂性的稳控策略生成模型^[6-7]。

新能源的不确定性明显改变了电网的瞬态稳定特性,成为稳控策略制定中不可忽视的要素^[8-10]。针对电网暂态稳定分析与控制问题,目前已有相关研究取得了初步进展,主要包括基于物理模型、实时仿真、人工智能算法的三类方法^[9]。从物理机理角度出发,文献[11-13]通过建立水、光互补系统模型,提出了基于等面积法则的紧急切机方案;文献[14]进行了双馈型风电机组接入对电力系统暂态功角稳定性的影响分析。此类方法增加了系统的复杂性,需要进行额外改进,以确保系统的协同工作。从实时仿真技术角度出发,文献[15]采用现场可编程逻辑门阵列进行实时仿真任务分解;文献[16]通过利用电网节点拓扑特征,采用功率支援与协调切机的方法实施仿真;文献[17]提出了一种基于元件的启发式系统剖分方法。此类方法导致工程实践中对于新能源电力系统的切机量计算缺乏理论指导,且单纯依赖时域仿真切机方案的制定效率有待提升。从人工智能算法应用角度出发,现有研究主要分为基于启发类算法^[18]和强化学习^[19-20]两类。基于Q启发类算法,文献[21]建立随机森林暂态稳定裕度回归模型,并定义切机和切负荷灵敏度,基于遗传算法求取最优紧急控制策略,此方法依赖于大量具有切机标签的样本,且需要极大的仿真计算量;从强化学习的角度出发,文献[22]提出通过深度学习提取电网与环境信息特征,并采用Q学习算法训练最优的切机策略;文献[23]根据等面积定则建立基于深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法的切机决策模型,以应对“双高”电力系统特征。

为了解决这些问题,本文提出基于深度强化学习的电力系统紧急切机稳控策略生成方法。首先针对电力系统稳控问题定义约束条件,并将其转化为马尔科夫决策过程;随后通过电气量特征评估和斯皮尔曼(Spearman)相关系数方法精选出最具代表性的特征数据;最后采用DDPG算法对智能体进行训练,以提高稳控策略的智能化水平。

1 稳控模型的构建

本文针对电网在面临重大故障时可能出现的

暂态稳定问题采用紧急控制手段,如切机和切负荷。通过强化学习技术优化这些稳控策略,同时减小其附加影响,包括切机选择和切负荷容量。

1.1 暂态稳定控制原理

以单机无穷大系统为例解释电网暂态过程。该系统的发电机功角转速方程为

$$\begin{cases} \frac{d\omega}{dt} = \frac{P_K - P_D}{M} \\ M \cdot \frac{d^2\delta}{dt^2} = P_K - P_D \\ P_D = (E \cdot V \sin \delta) / X_D \end{cases} \quad (1)$$

式中, ω 为发电机转子转速,rad/s; P_K 、 P_D 分别为原动机发电功率和发电机电磁功率; δ 为发电机功角; M 为发电机转动惯量; E 为发电机内部电势; V 为系统电压; X_D 为发电机同步电抗。

当事故发生后,系统中 P_K 短时间内不变,而 P_D 下降,造成发电机转子加速,导致功角稳定问题的产生。为了分析紧急控制措施效果,需要在模型中加入切机和切负荷的稳控措施。切机可以有效降低 P_K ,切负荷时会增加系统电压 V ,发电机功角、发电机内部电势和发电机同步电抗相对稳定,因此可以有效增加 P_D 。采用这2种方式可以减小系统功率差,从而增强功角稳定裕度。为了确保适用于各种环境,本文选择紧急控制中采取切机措施,以避免过多特定条件的局限性。

1.2 含约束的马尔科夫模型

为了使用强化学习方法解决电网瞬态稳定性切换的紧急控制问题,需要将电网的数学模型转化为包含约束的马尔科夫模型,用于描述数据交互和决策过程。该模型由五元组 $(S^i, A^i, P, R^i, \gamma)$ 来描述,即模型状态空间、动作空间、约束条件、奖励函数和折扣系数。在电力系统稳控问题中,五元组各元素的定义如下。

1.2.1 状态空间

在稳控智能体中,状态向量的设定需要反应电网的暂态过程特性以及发电机转子的功角特性,状态向量可表示为

$$S^i = [\delta_i^i, \omega_i^i, a_i^i, I_i^i, V_i^i], i \in G, t \in T \quad (2)$$

式中, G 为发电机数量; T 为训练时间段; δ_i^i 、 ω_i^i 、 a_i^i 、 I_i^i 和 V_i^i 分别为第 i 个发电机 t 时刻的功角、转速、转子加速度、母线电流和母线电压。

1.2.2 动作空间

本文采取的紧急控制动作为切机,设置可控对象为电网内 m 台发电机,定义紧急控制动作为

$$A_t^i = [c_t^1, c_t^2, \dots, c_t^G] \quad (3)$$

式中, c_t^i 为第 i 台发电机 t 时刻切机指令, 取值为 1 或 0, 分别表示切机和不切机。

1.2.3 奖励函数

为了客观评估切机动作, 本文将奖励函数划分为 2 个方面: 功角稳定控制效果和切机机组数量的合理性, 即

$$R_t^i = \alpha \cdot r_t^s - \beta \cdot r_t^a \quad (4)$$

$$r_t^s = \begin{cases} 1, & \max |\delta_t^i - \delta_t^j| < \pi \\ -1, & \max |\delta_t^i - \delta_t^j| \geq \pi \end{cases} \quad \forall i, j \in \{1, 2, \dots, G\}$$

$$r_t^a = A_t^i \cdot S_t^i$$

式中, r_t^s 为功角稳定奖励, 当发电机最大功角差小于 180° 时取值为 1, 当功角失稳时取值为 -1; r_t^a 为切机量惩罚; α, β 为奖励函数系数。

1.2.4 安全约束

为了在紧急控制中确保动作的安全可靠性, 本文概括安全约束并将其作为调控智能体不可违反的规范。

1) 切机动作时序约束。切机操作必须在故障切除时刻之后进行, 且一旦机组完成切机, 将不再考虑后续并网操作。约束条件为

$$A_t^i = 0, t \leq T, i \in \{1, 2, \dots, G\} \quad (5)$$

$$A_t^i - A_{t-1}^i \geq 0, \forall i \in \{1, 2, \dots, G\}, \forall t \in [0, T] \quad (6)$$

2) 切机机组选择约束。当电网发生重大扰动后, 通常由于继电保护动作导致电网负荷减少, 从而形成系统电磁功率过剩。为了解决这一问题, 通常在工程实践中选择从功角加速度最大的机组开始切除, 即优先切除功率过剩最严重的机组。基于这一领域知识, 本文制定动作约束: 当机组的功角加速度小于某一阈值 ε 时, 不切除。约束条件为

$$A_t^i = 0, \forall a_t^i \leq \varepsilon, i \in \{1, 2, \dots, G\} \quad (7)$$

3) 切机量约束。在故障发生后, 如果切机动作执行过度, 可能导致另一种极端情况, 即供电不足的可能性增加。因此, 在执行功角稳定控制时, 必须谨慎考虑切除容量的合理性。约束条件为

$$\sum_{i=1}^G (1 - A_t^i) \cdot P_t^i \geq \eta \cdot P_t^{\text{load}} \quad (8)$$

式中, η 为负荷功率供给系数, 本文取值为 0.98; P_t^i 为第 i 个发电机 t 时刻的输出功率; P_t^{load} 为 t 时刻的负荷所需功率。

2 面向暂态过程的电网特征提取

本文采用特征评估与 Spearman 等级相关系数

方法筛选出最典型的特征数据, 从而简化马尔科夫决策模型的状态表示, 提高模型的准确性。

2.1 特征量评估

特征量评估主要是评估在不同状态下各个电气量数据的可区分性, 用于筛选特征数据的特征量评估方法原理如下。

假设第 i 类状态有 N_i 个样本, 用 x_{ij} 表示第 i 类状态中第 j 类电气量数据样本, 距离度量为 $d(x_{ij}^{(k)}, x_{ij}^{(l)})$, 类内距离为

$$D_{ij} = \frac{1}{N_i(N_i - 1)} \sum_{k=1}^{N_i} \sum_{l=1, l \neq k}^{N_i} d(x_{ij}^{(k)}, x_{ij}^{(l)}) \quad (9)$$

式中, k, l 分别为第 j 类电气量数据的第 k, l 个样本。

平均类内距离为

$$D_j = \frac{1}{M} \sum_{i=1}^M D_{ij} \quad (10)$$

式中, M 为状态类别数。

计算第 i 类状态中第 j 类电气量数据均值, 即

$$X_{ij} = \frac{1}{N_i} \sum_{k=1}^{N_i} x_{ij}^k \quad (11)$$

则可计算第 j 类电气量数据在所有状态类别下的平均类间距离为

$$D_j' = \frac{1}{M(M-1)} \sum_{k=1}^M \sum_{l=1, l \neq k}^M d(x_{aj}, x_{bj}) \quad (12)$$

式中, x_{aj}, x_{bj} 分别为第 a, b 类状态中第 j 类电气数据的均值。

定义区分度因子 δ 为

$$\delta_{ab}^j = \frac{d(x_{aj}, x_{bj})}{|D_{aj} + D_{bj}|} \quad (13)$$

式中, D_{aj}, D_{bj} 分别为第 a, b 类状态中第 j 类电气数据的类型。 δ_{ab}^j 决定了第 j 类电气量数据对第 a, b 类状态的可区分性, 比值越大可区分性越强。

2.2 Spearman 相关系数法

针对具有相同变化特点的数据类型, 本文采用 Spearman 等级相关系数进行剔除。Spearman 等级相关系数是度量变量相关程度的一种方法, 具体步骤如下。

1) 数据转换为等级。对于每个变量, 将其值从小到大排列, 并为每个值分配一个等级(排名)。如果有重复值, 将其排名取平均值。

2) 计算等级差。对于每一对数据点, 计算其在 2 个变量中的等级差。等级差即为 2 个变量对应排名之差。

3) 计算 Spearman 等级相关系数。通过计算等

级差的秩相关系数来衡量2个变量之间的关系。秩相关系数的取值范围为 $-1\sim 1$,其中,1表示完全的正相关:一个变量的增加总是伴随着另一个变量的增加; -1 表示完全的负相关:一个变量的增加总是伴随着另一个变量的减少; 0 表示无相关性:2个变量之间没有线性关系。

Spearman等级相关系数的计算公式为

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (14)$$

式中, d_i^2 为每对数据点的等级差; n 为数据点的数据量。

3 基于强化学习的稳控策略智能体

电网稳控模型是作为强化学习的环境,与智能体进行信息交互,对于每一个时间段 t ,环境模型采集电网数据构成状态空间传递给智能体,智能体经过训练将动作空间信息传递给环境模型,环境模型经过反馈将奖励值传递给智能体,强化学习交互过程如图1所示。强化学习智能体的目标是在有限的行动步数内,通过最大化累积奖励值来寻找最优的决策策略。考虑电力系统稳控策略的安全性要求,常规强化学习算法难以在电网安全域内进行有效的探索。

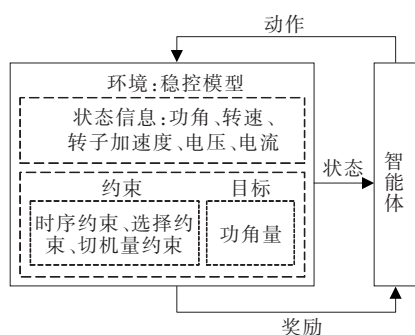


图1 强化学习交互过程

Figure 1 Reinforcement learning interaction process

3.1 目标函数

DDPG算法利用一阶优化近似表示其目标函数 $L^{\text{CLIP}}(\theta)$,并引用领域知识动作约束,从而提高算法的精准度,具体训练模型如下:

$$\begin{cases} \max_{\theta} L^{\text{CLIP}}(\theta) = E \left\{ \min \left[R_t(\theta) A_t, \text{clip} \left(R_t(\theta), 1, 1 - \epsilon, 1 + \epsilon \right) A_t \right] \right\} \\ \text{s.t. 式(5)~(8)} \end{cases} \quad (15)$$

式中, E 为数学期望, $R_t(\theta) = \pi_{\theta}(A_t|S_t)/\pi_{\theta}(A_t|S_t)$ 为更新后的策略与更新之前策略的概率比值;第2

项通过clip函数将 $R_t(\theta)$ 约束至 $[1 - \epsilon, 1 + \epsilon]$ 区间内,提高策略更新过程的收敛性。当优势函数 $A_t > 0$ 时,策略训练朝正向逼近;当 $A_t < 0$ 时,策略向更差的方向发展。

3.2 DDPG算法

由于电力系统的运行数据包含连续变量,如电压和电流,故决策所依赖的状态和动作空间都是连续的。这种情况下传统的离散动作空间的强化学习方法无法进行有效处理,而DDPG算法专门设计用于解决连续动作空间的问题,通过结合策略梯度方法和Actor-Critic架构,能够在高维连续空间中进行有效的策略优化。因此,选择DDPG算法进行模型训练不仅能够处理连续的状态和动作空间,还能提高决策的精度和稳定性,从而更好地满足电力系统稳控策略的需求。利用DDPG算法训练马尔科夫决策过程模型,并采用最终决策的动作价值函数来评估策略的优越性,即

$$Q_{\pi}(s_t, a_t) = E \left[r(s_t, a_t) + \gamma \cdot \max_{a'} Q_{\pi}(s', a') \right] \quad (16)$$

式中, $Q_{\pi}(s_t, a_t)$ 为某个策略 π 下状态 s_t 中动作 a_t 获得的奖励值; $r(s_t, a_t)$ 为状态 s_t 中动作 a_t 获得的奖励; γ 为折扣因子; $Q_{\pi}(s', a')$ 为整个训练过程中取某个状态 s' 中动作 a' 获得的奖励。

DDPG算法中加入了经验池回放机制,每一次通过环境交互后训练样本 $L = (r_{t+1}, s_{t+1})$,样本中 r_{t+1} 为 $t+1$ 次迭代的奖励, s_{t+1} 为 $t+1$ 次的状态量。为防止训练后期出现过拟合现象,本文引入正则化更新经验池中的数据,具体公式为

$$L_{t+1} = L + \lambda \sum_{i=1}^{t+1} |w_i| \quad (17)$$

式中, λ 为正则化系数; w_i 为奖励的权重系数。

DDPG算法框架如图2所示。通过最小化损失值(均方误差损失)来更新Critic网络的参数,Critic网络更新时的损失函数为

$$L_{\text{loss}} = \frac{1}{N} \sum_i \left[y_i - Q_{\pi}(s', a') \right]^2 \quad (18)$$

式中, y_i 为目标函数值。

Actor目标网络用于提供下一个状态的策略,Actor训练网络则是提供当前状态的策略,结合Critic训练网络的Q值函数可以得到Actor在参数更新时的策略梯度:

$$\nabla_{\theta} \alpha = \frac{1}{N} \sum_i \nabla_a Q(s, a|\theta_Q) \Big|_{s=s_i, a=\pi_{\theta}(s_i)} \nabla_{\theta} a_{t+1} \Big|_{s_i} \quad (19)$$

式中, $\nabla_{\theta} \alpha$ 为当前Actor网络 θ 对参数 π_{θ} 的梯度; $\nabla_a Q$

为当前 Critic 网络 Actor 的梯度;

DDPG 通过 soft update 保证 Actor 参数 θ_a' 和 Critic 参数 θ_Q' 可以缓慢更新,从而提高学习的稳定性,具体公式为

$$\begin{cases} \theta_a' \leftarrow \xi \theta_a + (1 - \xi) \theta_a' \\ \theta_Q' \leftarrow \xi \theta_Q + (1 - \xi) \theta_Q' \end{cases} \quad (20)$$

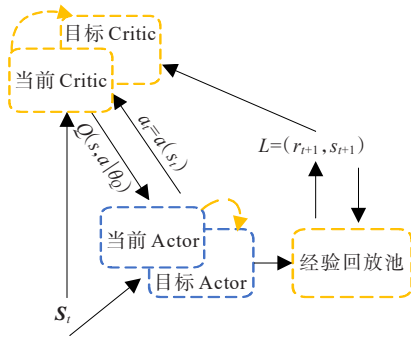


图 2 DDPG 算法框架

Figure 2 DDPG algorithm framework

3.3 策略更新

基于固定的暂态响应轨迹数据进行 Q 学习算法,每次从故障集中随机抽取部分数据,学习算法策略更新流程如图 3 所示。

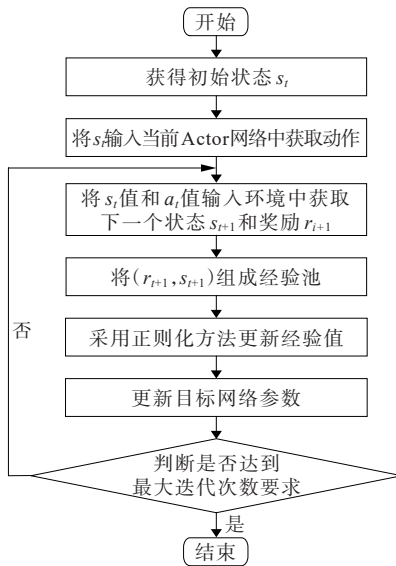


图 3 Q 学习算法策略更新流程

Figure 3 Update flowchart of Q learning algorithm policy

4 算例

4.1 算例设置

本文仿真测试在 IEEE 39 节点系统中进行,其中发电机 3 号节点位置添加 10 MW 的风电机组 G_{11} , 32 号和 35 号节点上的发电机 G_3 和 G_6 替换为等

容量的风电机组模型,以模拟高比例新能源渗透率电网运行场景;算例场景中新能源发电占比为 10%~20%。修改后的 IEEE 39 节点拓扑如图 4 所示;电网运行负荷水平、发电功率比例、故障位置和故障类型等参数如表 1 所示。

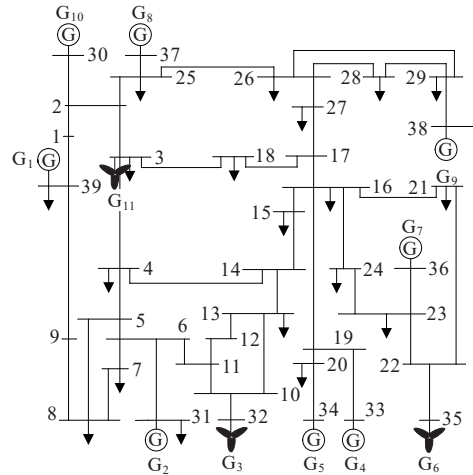


图 4 修改后的 IEEE 39 节点拓扑

Figure 4 Modified IEEE 39 node topology

表 1 仿真场景概率设置

Table 1 Simulation scenario probability settings

属性	设置区间	概率模型
负荷水平	[0.85, 1.15]	等比例分布
发电功率比例	[0.90, 1.15]	均匀分布
故障位置	所有线路	等概率模型
故障类型	相间短路,单相、两相、三相短路接地	等概率模型

随机抽取 4 000 组故障场景用于智能体训练,其中失稳场景为 1 785 组,稳定场景为 2 215 组,仿真时长为 10 s,仿真步长为 0.01 s。

4.2 对比方案

为了验证本文所提方法的紧急控制决策效果,分别从特征感知和领域知识融合两方面与现有研究进行对比,如表 2 所示,对比方法 1、2、3 分别为特征感知、领域知识、传统算法对照,上述方法的基础强化算法均为 DDPG 算法上展开进行。

表 2 算例对比方法

Table 2 Comparison of different methods

算法	说明
本文方法	特征量评估+Spearman 相关系数法+DDPG +SL
对比方法 1	DDPG+SL
对比方法 2	特征量评估+Spearman 相关系数法+DDPG
对比方法 3	DDPG 学习

4.3 性能测试

对比不同算法在训练阶段的奖励函数和损失下降曲线变化趋势,如图5所示,可以看出,在训练效率和收敛情况方面,对比方法3作为传统的DDPG学习算法相比其他方法均无优势;相比之下,本文方法在数据的提取和训练方面都做了改进,奖励函数值上升,训练效率明显优于其他方法,750轮时奖励函数值达到28.0,经过1000轮次左右迭代,奖励函数值逐渐收敛于29.2;对比方法1相比于本文方法的收敛效率较慢,经过1400轮迭代,奖励函数值收敛于24.1,且收敛值仅为本文方法的82%;对比方法2最后收敛情况与对比方法1类似,从收敛速度和收敛情况分析,与本文方法相比明显处于劣势状态,在经过2500轮次迭代才趋于收敛。

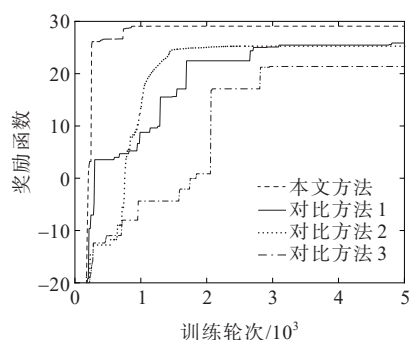


图5 决策奖励函数值对比(滑动平均)

Figure 5 Comparison of decision reward function values (moving average)

为了验证框架的有效性,针对不同方法的损失值函数如图6所示,对比方法3作为传统的算法学习,其损失值函数呈现较大的波动;由于本文方法考虑了领域知识,其损失值相较于未考虑领域知识约束的对比方法2表现出更好的性能;本文所提明显优于其他对比方案的紧急控制智能体策略,几乎消除了损失量,从而极大地提高了训练效率。

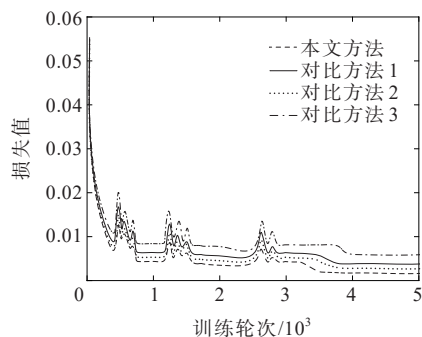


图6 损失值曲线

Figure 6 Loss value curve

4.4 案例分析

为了验证本文所提方法在故障后的有效性,选取训练集以外的代表性场景进行测试验证,分析其决策合理性。初始故障为母线2处的三相短路接地故障,持续时间为0.2 s。智能体未切机功角、转速波动如图7所示。

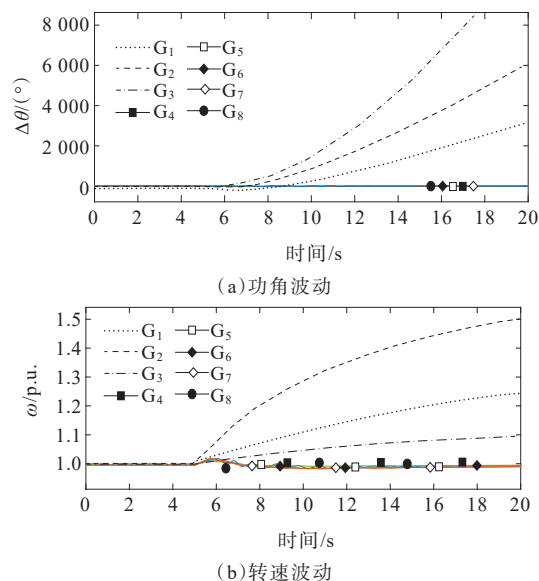


图7 智能体未切机功角、转速波动

Figure 7 Power angle and speed fluctuations of intelligent agent without tripping of units

根据图7,发电机1、2、3在故障后功角和转速波动,若不采取行动则沿虚线路径演化造成功角失稳。按照本文所提智能体故障切除后快速执行切机动作, G_1 、 G_2 、 G_5 执行切机策略,机组切除分别为75、70、105 MW,经过智能体切机振荡后恢复正常状态,如图8所示。

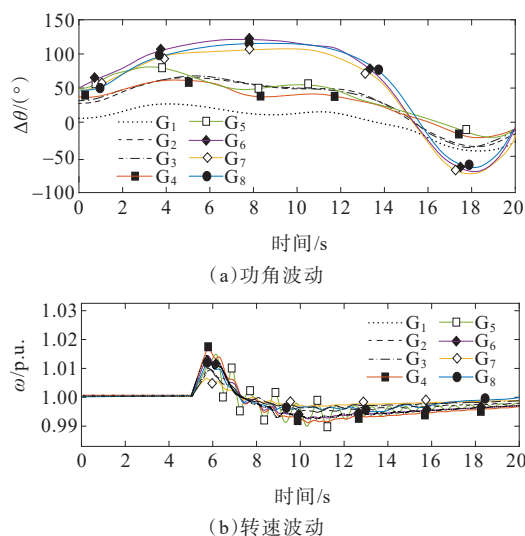


图8 智能体切机振荡后恢复情况

Figure 8 Recovery of intelligent agent after oscillation following tripping of units

由图8可知,系统经过一段时间的振荡后成功恢复正常状态,功角和转速稳定波动,不再出现波动失稳的情况,从而确保了系统的安全性。

4.5 计算效率

为了进一步验证本文所提方法的计算效率优势,本文将基于IEEE 39节点系统和某实际电网进行计算效率测试。该实际电网共245个母线节点;发电机母线节点32个,其中火电机组、风电场、光伏电站并网节点个数分别为24、5、3;算例场景中新能源功率渗透率为10%~20%。多个对比方法的训练与执行计算耗时如表3所示。

表3 多个对比方法的训练与执行计算耗时

Table 3 Training and execution calculation time of multiple methods

算法	IEEE 39节点系统		某实际电网(245节点)	
	总时长/ (10 ⁴ s)	决策时间/ ms	总时长/ (10 ⁴ s)	决策时间/ ms
本文方法	2.032 5	11.2	6.852 2	17.2
对比方法1	2.174 0	10.2	8.552 6	14.6
对比方法2	3.210 4	8.3	9.594 5	11.4
对比方法3	3.115 0	7.4	10.565 9	8.9

由表3可知,本文方法的训练决策时间较长,这是由于本文所提算法结构较为复杂。然而,在故障发生到系统恢复稳定状态所需的总时间上,本文方法却表现出最快的恢复速度。这充分说明,尽管训练决策时间相对较长,但本文方法能够选择出更优的切机策略。因此,本文方法在计算效率上符合实际需求。

5 结语

在新型电力系统建设的背景下,传统的紧急切机控制策略在应对不同场景时面临适应性降低、稳定精度不足等问题。因此,本文提出了基于深度强化学习的电力系统紧急切机稳控策略生成方法,该方法具有很高的实用性。

1) 本文将紧急控制下的暂态功角稳定视为序贯决策问题,建立了一个含有约束的马尔科夫模型,以改进切机决策的性能,并面对故障时能迅速、准确做出切机决策,从而确保了系统的安全性;

2) 本文提出了电气量特征评估和Spearman相关系数方法级联而成的暂态过程空间与时序特征感知网络层,能够迅速提取准确特征值,在总时长方面相对于其他对比方法减少了50%左右;

3) 通过对IEEE 39和某实际电网的测试,验证了本文所提方法在准确性和实用性的优势,相较于单一强化学习,智能体的训练效率在2种场景下分别提高了20%、31%。

参考文献:

- [1] 韩保军,高强,代飞,等.基于协同奖励函数多目标强化学习的智能频率控制策略研究[J].电力科学与技术学报,2023,38(2):18-29.
HAN Baojun, GAO Qiang, DAI Fei, et al. Intelligent frequency control strategy based on multi-objective reinforcement learning of cooperative reward function[J]. Journal of Electric Power Science and Technology, 2023, 38(2):18-29.
- [2] 李响,张丹,李秋燕,等.计及随机性的多分布式电源接入中压配电网承载能力评估[J].电力系统保护与控制,2024,52(17):150-160.
LI Xiang, ZHANG Dan, LI Qiuyan, et al. Hosting capacity evaluation of distributed generators accessing a medium voltage distribution network considering randomness[J]. Power System Protection and Control, 2024, 52(17):150-160.
- [3] 张磊光,陈海涛,杨军.基于SAC算法的含柔性负荷电-气互联系统的频率与气压协调控制策略[J].智慧电力,2024,52(4):8-14.
ZHANG Leiguang, CHEN Haitao, YANG Jun. Frequency-pressure coordinated control strategy of electrical-gas interconnection system based on SAC algorithm[J]. Smart Power, 2024, 52(4):8-14.
- [4] OJETOLA S, WOLD J, TRUDNOWSKI D. Multi-loop transient stability control via power modulation from energy storage devices[J]. IEEE Transactions on Power Systems, 2021, 36(6):5153-5163.
- [5] 李建林,胡筋扬,辛迪熙,等.基于参数自调节的电氢耦合系统调频控制策略研究[J].高压电器,2024,60(7):1-11.
LI Jianlin, HU Jiayang, XIN Dixi, et al. Research on frequency modulation control strategy of electric-hydrogen coupling system based on parameter self-regulation[J]. High Voltage Apparatus, 2024, 60(7):1-11.
- [6] 张柏林,李希德,魏博,等.基于改进的场景分类和去粗粒化MCMC的风电出力模拟方法[J].电测与仪表,2024,61(7):41-49.
ZHANG Bolin, LI Xide, WEI Bo, et al. Wind power output simulation method based on improved scene classification algorithm and time series correlation[J]. Electrical Measurement & Instrumentation, 2024, 61(7):41-49.
- [7] 张建新,常东旭,邱建,等.适应新型电力系统的安全稳定控制系统及装置设计方法[J].电网与清洁能源,2023,39(12):10-19.

- ZHANG Jianxin, CHANG Dongxu, QIU Jian, et al. A design method of stability control device and system adapted to new power systems[J]. *Power System and Clean Energy*, 2023, 39(12): 10-19.
- [8] 张红丽,刘福锁,李威.考虑新能源短时波动的大电网在线暂态安全稳定控制方法[J].*电力系统自动化*,2023,47(20):166-173.
- ZHANG Hongli, LIU Fusuo, LI Wei. Online transient security and stability control method for large power grids considering short-term fluctuations of new energy sources[J]. *Automation of Electric Power Systems*, 2023, 47(20):166-173.
- [9] 李晨,张昊,周济,等.考虑频率电压空间分布特性的电网实时减载及负荷恢复策略[J].*电力科学与技术学报*, 2021,36(2):162-171+216.
- LI Chen, ZHANG Hao, ZHOU Ji, et al. Load shedding and load recovery strategy considering spatial distribution characteristics of frequency and voltage[J]. *Journal of Electric Power Science and Technology*, 2021, 36(2): 162-171+216.
- [10] 阚骏,董希建,王敏,等.基于MCMC的电网安全稳定控制系统动态可靠性评估方法[J].*电力工程技术*,2024,43(3):23-31.
- KAN Jun, DONG Xijian, WANG Min, et al. Dynamic reliability evaluation method for power grid safety and stability control system based on MCMC[J]. *Electric Power Engineering Technology*, 2024, 43(3): 23-31.
- [11] 耿华,何长军,刘浴霖,等.新能源电力系统的暂态同步稳定研究综述[J].*高电压技术*,2022,48(9):3367-3383.
- GENG Hua, HE Changjun, LIU Yushuang, et al. A review of transient synchronous stability research in new energy power systems[J]. *High Voltage Engineering*, 2022, 48(9): 3367-3383.
- [12] 陈韦韬,吴翔宇,许寅,等.面向暂态功角稳定提升的水光互补发电系统紧急切机控制[J].*电网技术*,2023,47(2):658-671.
- CHEN Weitao, WU Xiangyu, XU Yin, et al. Emergency cut-off control for hydro-photovoltaic complementary power generation systems to enhance transient angular stability[J]. *Power System Technology*, 2023, 47(2): 658-671.
- [13] 李雅晗,夏世威,马琳琳,等.交直流混联系统暂态功角稳定评估及特征量可解释性分析[J].*电力建设*,2024,45(2): 1-9.
- LI Yahan, XIA Shiwei, MA Linlin, et al. Transient power angle stability evaluation and interpretability analysis of AC/DC hybrid power system[J]. *Electric Power Construction*, 2024, 45(2): 1-9.
- [14] 孙仲卿,刘福锁,李威,等.基于特征匹配的暂态稳定紧急控制策略快速生成[J].*电力系统自动化*,2024,48(2):167-175.
- SUN Zhongqing, LIU Fusuo, LI Wei, et al. Rapid generation of transient stability emergency control strategy based on feature matching[J]. *Automation of Electric Power Systems*, 2024, 48(2): 167-175.
- [15] 付浩,李鹏,富晓鹏,等.面向多FPGA实时仿真器的资源优化配置方法[J].*电力系统自动化*,2023,47(11):88-100.
- FU Hao, LI Peng, FU Xiaopeng, et al. Resource optimization configuration method for real-time simulators based on multiple FPGAs[J]. *Automation of Electric Power Systems*, 2023, 47(11): 88-100.
- [16] 肖友强,邓步青,文云峰.计及送端电网频率稳定的直流联络线极限承载能力评估[J].*电力科学与技术学报*, 2021,36(2):3-12.
- XIAO Youqiang, DENG Buqing, WEN Yunfeng. Total transfer capability assessment of HVDC tie-lines based on frequency stability of the sending-end grid[J]. *Journal of Electric Power Science and Technology*, 2021, 36(2): 3-12.
- [17] 郝晓亮,付立军,马凡,等.中压直流综合电力系统建模与实时仿真实现方法[J].*电网技术*,2021,45(3):1100-1109.
- HAO Xiaoliang, FU Lijun, MA Fan, et al. Modeling and real time simulation of MVDC integrated power system [J]. *Power System Technology*, 2021, 45(3): 1100-1109.
- [18] 潘晓杰,徐友平,解治军,等.堆栈式集成学习驱动的电力系统暂态稳定预防控制优化方法[J].*发电技术*,2023, 44(6):865-874.
- PAN Xiaojie, XU Youping, XIE Zhijun, et al. Power system transient stability preventive control optimization method driven by stacking ensemble learning[J]. *Power Generation Technology*, 2023, 44(6): 865-874.
- [19] 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 emergency control of transient stability based on deep reinforcement learning[J]. *Automation of Electric Power Systems*, 2023, 47(5): 144-152.
- [20] XIE J, SUN W. Distributional deep reinforcement learning-based emergency frequency control[J]. *IEEE Transactions on Power Systems*, 2021, 37(4): 2720-2730.
- [21] 王彤,刘九良,朱劭璇,等.基于随机森林的电力系统暂态稳定评估与紧急控制策略[J].*电网技术*, 2020, 44(12): 4694-4701.
- WANG Tong, LIU Jiuliang, ZHU Shaoyuan, et al. Transient stability assessment and emergency control strategy based on random forest in power system[J]. *Power System Technology*, 2020, 44(12): 4694-4701.
- [22] CUI W, JIANG Y, ZHANG B. Reinforcement learning for optimal primary frequency control: a Lyapunov approach [J]. *IEEE Transactions on Power Systems*, 2022, 38(2): 1676-1688.
- [23] 卢恒光,林碧琳,温步瀛.基于深度强化学习的切机控制策略研究[J].*电器与能效管理技术*,2023(3):11-15+68.
- LU Hengguang, LIN Bilin, WEN Buying. Research on generator tripping control strategy based on deep reinforcement learning[J]. *Electrical & Energy Management Technology*, 2023(3): 11-15+68.