

DOI: 10.19783/j.cnki.pspc.231360

## 基于安全强化学习的电网稳控策略智能生成方法

邱建<sup>1</sup>, 朱煜昆<sup>2,3</sup>, 张建新<sup>1</sup>, 朱益华<sup>2,3</sup>, 徐光虎<sup>1</sup>, 涂亮<sup>2,3</sup>

(1. 中国南方电网有限责任公司, 广东 广州 510663; 2. 直流输电技术全国重点实验室(南方电网科学研究院有限责任公司), 广东 广州 510663; 3. 广东省新能源电力系统智能运行与控制企业重点实验室, 广东 广州 510663)

**摘要:** 新型电力系统的“双高”趋势改变了电力系统经典稳定特性, 导致稳定机理更复杂, 系统稳定模式更多样, 因此基于典型运行方式的在线稳定控制策略面临挑战。为解决新型电力系统的功角稳定问题, 提出了基于安全强化学习的稳控策略智能生成方法。首先, 建立了电力系统稳控问题的含约束马尔可夫模型, 归纳并提出了紧急控制切机动作涉及的安全约束。其次, 为了提高对于电网暂态响应的时空特征提取能力, 构建了基于图卷积层和长短期记忆单元的特征感知网络。然后, 为了提高稳控策略智能体的训练效率, 提出了基于内嵌领域知识约束的近端策略优化算法稳控策略训练框架。最后, 在 IEEE 39 节点系统和某实际电网中进行测试验证。结果表明, 所提方法能够根据系统运行状态和故障响应自适应生成切机稳控策略, 其决策效果和效率均优于现有的稳控策略。

**关键词:** 稳控策略; 安全强化学习; 时空特征; 领域知识

### Intelligent generation method of power system stability control strategy based on safe reinforcement learning

QIU Jian<sup>1</sup>, ZHU Yukun<sup>2,3</sup>, ZHANG Jianxin<sup>1</sup>, ZHU Yihua<sup>2,3</sup>, XU Guanghu<sup>1</sup>, TU Liang<sup>2,3</sup>

(1. China Southern Power Grid Company Limited, Guangzhou 510663, China; 2. State Key Laboratory of HVDC, Electric Power Research Institute, CSG, Guangzhou 510663, China; 3. Guangdong Provincial Key Laboratory of Intelligent Operation and Control for New Energy Power System, Guangzhou 510663, China)

**Abstract:** The trend of a “higher proportion of renewable energy and power electronics” in the new power system has changed the classical stability characteristics of the system. The stability mechanism is more complex, and the system stability modes are more diverse. Online stability control strategies based on typical operating modes face a challenge. Considering the rotor angle stability problem of the new power system, an intelligent generation stability control strategy based on safe reinforcement learning is proposed. First, a constrained Markov model for power system stability control problems is established, and the safety constraints involved in rotor angle stability control are summarized and proposed. Secondly, to improve the ability to extract spatial and temporal features of the power grid’s transient response, a feature perception network based on graph convolutional layers and long short-term memory units is constructed. Then, to improve the training efficiency of the stability control agent, a training framework of stability control strategies using proximal policy optimization algorithm based on embedded domain knowledge constraints is proposed. Finally, a case study is performed on the IEEE 39-bus system and a practical power grid. The results show that the proposed method can adaptively generate unit tripping strategies based on the system operating state and fault response, and its decision-making effectiveness and efficiency are superior to existing stability control strategies.

This work is supported by the Key Science and Technology Project of China Southern Power Grid Co., Ltd. (No. 000000KK52210139).

**Key words:** stability control strategy; safety reinforcement learning; temporal and spatial characteristics; domain knowledge

## 0 引言

在新型电力系统建设背景下, 电力系统“双高”

特征日益显著, 给电网的安全稳定带来了严峻的挑战<sup>[1-2]</sup>。现有电网稳控策略是根据系统运行典型方式进行故障集扫描并构建相对保守的稳控措施。然而, 随着新型电力系统的逐步推进, 电网潮流方式<sup>[3-4]</sup>愈加多样, 不断挤压相对保守的运行空间。如何在兼顾安全性和经济性的前提下, 构建基于典型

基金项目: 南方电网公司重点科技项目资助(000000KK52210139)

运行方式的稳控策略成为难题。因此,为了满足未来新型电力系统运行需求,亟需提出适用电网潮流多样性与不确定性、稳定模态多样性与机理复杂性的稳控策略生成方法<sup>[5-6]</sup>。

根据稳控策略生成的时刻和执行方式的差异,针对稳控策略的研究可分为3种类型:离线预决策、实时匹配;在线预决策、实时匹配;实时决策、实时控制<sup>[7]</sup>。

目前,电网稳控主要采用第一种模式,即离线进行海量运行方式的故障集仿真并建立稳控策略表。对于传统电网运行方式相对确定、电网发展变化较慢的阶段而言,该模式具有较好的实践效果。

电网量测与通信系统的发展使得获取电网实时状态数据成为可能。同时,基于高性能仿真和高精度电网仿真模型的实时仿真平台也快速发展。这两者共同使得“在线预决策、实时匹配”的稳控模式成为可能。通过对实时电网的滚动仿真分析,制定稳控策略预案,当实际发生故障后按照最新的稳控策略执行。相比于第一种模式,这种模式提高了策略的时效性和准确性。目前,后者在广东、山东、陕西等地已有实际应用。

随着新能源并网容量逐步提升,我国西北西南地区已经出现了新能源并网容量超过传统电源的情况。新能源出力的不确定性显著改变了电网暂态稳定特性,新能源运行工况不确定性成为了稳控策略制定中不可避免的环节<sup>[8]</sup>。稳控策略制定需要对大量的运行场景进行分析计算,离线/在线预决策的模式难以满足紧急控制的时效性要求。因此,对于“实时决策、实时控制”模式下暂态稳定分析与控制亟待研究。

针对电网暂态稳定分析与控制问题<sup>[9-10]</sup>,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、人工智能三类方法。

从物理机理角度出发,现有研究将稳定分析理论方法应用于新型电力系统中,并考虑高比例新能源并网后的模型变化,修正了稳定分析模型。文献[11]以扩展等面积法则(extended equal area criterion, EEAC)为基础,考虑了新能源出力短时波动性对等值发电机功角轨迹的影响,从而计算紧急控制策略。文献[12]将新能源与直流接入电网后对频率稳定的影响进行了归一化映射,实现了高维多因素的降维处理。文献[13]构建了一种二次型 Lyapunov 函数,利用 Lyapunov 判据对孤岛微电网进行暂态稳定分析。

从实时仿真技术角度出发,文献[14]利用现场可编程逻辑门阵列(field programmable gate array,

FPGA)进行实时仿真任务的分解,实现了高效并行的新型电力系统仿真。文献[15]基于电网节点拓扑特征进行了网络方程的稀疏化方法,提高了仿真计算效率。文献[16]利用内存计算技术提高电网在线仿真计算速度,并在国调4万节点规模系统中进行了验证测试。目前,在实际电网中已有成熟应用,如中国电科院的电力系统全数字实时仿真系统已在多个省市部署应用<sup>[17]</sup>。

从人工智能算法应用角度出发,现有研究主要分为基于监督学习<sup>[18]</sup>和强化学习<sup>[19-22]</sup>两类。前者利用深度神经网络建立电网暂态故障响应特征与紧急控制的映射关系,经过大量样本训练后提高决策正确率。后者依赖电网镜像环境,利用智能体进行电网控制策略的自主寻优,通过与镜像环境的交互反馈,不断优化决策质量。文献[20]针对暂态稳定预防控制问题,构建了深度确定性策略梯度模型,利用发电机功率控制和极端电压控制,提升系统应对大扰动的暂态稳定性。文献[21]利用切负荷动作进行频率紧急控制,通过对动作值函数分布的拟合提升了智能体训练效率。文献[22]利用李雅普诺夫函数拟合惯性机组和电力电子设备的差异化动态特性,应对“双高”电力系统特征,利用循环神经网络预测频率变化趋势。

目前,上述两类研究均存在不足。基于监督学习的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。此外,随着电网的变化,样本集需要不断更新,并重新进行深度神经网络训练。基于强化学习的方法在高维状态空间和高维动作空间内进行自由探索时,可能会陷入局部最优或陷入无法收敛的情况,导致智能体训练效率较低。

为应对以上问题,本文提出一种基于安全强化学习的稳控策略生成方法。该方法利用图卷积神经网络(graph convolutional network, GCN)<sup>[23]</sup>的拓扑特征识别能力和长短期记忆单元(long short-term memory, LSTM)<sup>[24]</sup>的时间序列规律挖掘能力,实现了电网高维时空特征降维,解决了电网高维量测信息感知问题。同时,所提方法利用电网安全约束,指引智能体策略探索的方向,提升训练效率。

对比本领域现有研究,本文贡献主要体现在三个方面:(1)提出了电网时空特征感知网络架构,能够实现电网海量量测数据的降维,并提取与暂态稳定强相关的空间特征与时间序列趋势;(2)构建了具有高比例新能源的电网镜像环境,能够模拟新型电力系统强不确定性运行方式下的暂态稳定紧急控制;(3)基于数字化表达的电网安全运行的约束条件,建立了电网稳控安全强化学习智能体模型,提高了训

练过程的收敛速度与决策的安全性。

## 1 电网稳控模型构建

当电网遭遇严重故障时,可能引发暂态稳定问题。为了有效保障系统安全稳定运行,稳控系统会采取切机、切负荷等紧急控制措施。本文旨在研究稳控策略,利用强化学习技术优化稳控策略效果的同时降低稳控措施的附加影响,包括切机选择和切负荷容量等。本节分别从暂态稳定紧急控制原理和含约束马尔可夫模型两方面进行介绍。

### 1.1 暂态稳定紧急控制原理

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

$$\begin{cases} \frac{d\omega}{dt} = \frac{P_M - P_E}{M} \\ \frac{d\delta}{dt} = \omega - \omega_0 \end{cases} \quad (1)$$

式中: $\omega$ 为发电机转子转速(rad/s); $\omega_0$ 为发电机初始转速(rad/s); $P_M$ 和 $P_E$ 分别为原动机功率和发电机电磁功率; $\delta$ 为发电机功角; $M$ 为发电机转动惯量。当事故发生后系统 $P_M$ 短时间内不变,而 $P_E$ 下降,造成发电机转子加速,导致功角失稳问题。

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。切机与切负荷分别对应着降低 $P_M$ 和增加 $P_E$ ,这两种方式均可降低系统功率差额,从而提高功角稳定裕度。然而,两者在实际工程中动作优先级不同。根据《电力系统安全稳定导则》规定的不同稳定事件等级<sup>[25]</sup>,电力系统承受大扰动后,第一级标准可以切机但不能失负荷,第二、三级标准可以同时进行切机切负荷。为了不失一般性,本文仅采取切机措施进行紧急控制。

随着电网“双高”特征逐步增强,式(1)缺乏对新能源机组的弱惯性特征和直流功率突变特征的考虑。电网暂态响应的描述很难简单地利用转子运动方程进行描述,因此本文基于新型电力系统机电暂态仿真模型进行研究。

### 1.2 含约束马尔可夫模型

为了使用强化学习方法求解电网暂态稳定切机紧急控制问题,需要将该数学模型转化为含约束马尔可夫模型。它由五元组 $(S, A, P, R, \gamma)$ 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。

#### 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,本文将状态向量 $S$ 表示为

$$S = [\delta_{1:T}^{1:K} \ \omega_{1:T}^{1:K} \ a_{1:T}^{1:K} \ I_{1:T}^{1:N} \ V_{1:T}^{1:N}] \quad (2)$$

式中: $\delta_{1:T}^{1:K}$ 、 $\omega_{1:T}^{1:K}$ 、 $a_{1:T}^{1:K}$ 分别为编号1~K的发电机在1~T时刻的功角、角速度和转子加速度; $I_{1:T}^{1:N}$ 、 $V_{1:T}^{1:N}$ 为编号1~N的节点在1~T时刻的母线电流和电压。

为了便于神经网络训练,上述状态变量均归一化至 $[-1,1]$ 内,设原始数据为 $x$ ,则归一化公式为

$$\tilde{x} = \frac{2x_i - (x_{\max} + x_{\min})}{x_{\max} - x_{\min}} \quad (3)$$

式中: $\tilde{x}$ 为归一化后的数据; $x_i$ 为输入数据; $x_{\max}$ 、 $x_{\min}$ 分别为原始数据中的最大、最小值。

#### 1.2.2 动作空间

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

$$A_t = [c_t^1 \ c_t^2 \ \dots \ c_t^i \ \dots \ c_t^K] \quad (4)$$

式中: $A_t$ 表示紧急控制动作空间; $c_t^i$ 表示第i台发电机在t时刻的切机指令。

#### 1.2.3 奖励函数

为了客观评价切机动作,本文将奖励函数分为两部分——功角稳定控制效果和切机机组数量合理性。

$$\begin{cases} R_t = \alpha \cdot r_t^s - \beta \cdot r_t^a \\ r_t^s = \begin{cases} 1, \max |\delta_t^i - \delta_t^j| < \pi \\ -1, \max |\delta_t^i - \delta_t^j| \geq \pi \end{cases}, \forall i, j \in \{1, 2, \dots, K\} \\ r_t^a = A_t \cdot G_t^{1:K} \end{cases} \quad (5)$$

式中: $R_t$ 为t时刻的奖励函数值; $r_t^s$ 为功角稳定奖励,当发电机最大功角差小于 $180^\circ$ 时取1,当功角失稳时取-1; $r_t^a$ 为切机量惩罚,通过对切机容量的惩罚,尽可能地降低电网稳控动作成本和对电网的影响; $G_t^{1:K}$ 为K台发电机的功率列向量; $\alpha$ 和 $\beta$ 均为奖励函数系数,本文分别设置为10和0.1。

#### 1.2.4 安全约束

为了确保紧急控制动作安全可靠,本文将紧急控制中需要考虑的安全约束进行总结,并将其作为调控智能体需要遵循的硬约束,主要依据专家经验与电网实际控制规则,将先验知识融入智能体训练之中,以提高动作安全性和训练收敛效率。

1) 切机动作时序约束:切机时刻必须在故障切除时刻以后;机组切机以后不考虑后续并网操作。

$$A_t^i = 0, t \leq T^c, i \in \{1, 2, \dots, K\} \quad (6)$$

式中: $A_t^i$ 为第i台机组在t时刻的切机动作指令; $T^c$ 为故障切除时刻。

$$A_t^i - A_{t-1}^i \geq 0, \forall i \in \{1, 2, \dots, K\}, \forall t \in [0, T] \quad (7)$$

2) 切机机组选择约束: 在电网发生大扰动后, 继电保护动作通常导致电网负荷减少, 从而形成系统电磁功率过剩。为解决这一问题, 工程中一般选择从功角加速度最大的开始切除, 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  $\varepsilon$  时, 不切除, 如式(8)所示。

$$A_t^i = 0, \forall a_t^i \leq \varepsilon, i \in \{1, 2, \dots, K\}, \forall t \in [0, T] \quad (8)$$

式中:  $a_t^i$  为第  $i$  台发电机在  $t$  时刻的功角加速度;  $\varepsilon$  为阈值。

3) 切机量约束: 故障后, 切机动作若过度执行, 则可能造成功率供应不足的另一极端情况。因此, 功角稳定控制执行必须考虑切除容量的合理性。

$$\sum_{i=1}^K (1 - A_t^i) \cdot P_t^i \geq \eta \cdot P_t^{\text{load}} \quad (9)$$

式中:  $P_t^i$  为第  $i$  台发电机  $t$  时刻的有功功率;  $P_t^{\text{load}}$  为  $t$  时刻的总有功功率;  $\eta$  为负荷功率供给系数, 本文取 0.99。

### 1.2.5 镜像环境

针对上述电网暂态稳定紧急控制 CMDP 问题, 本文利用 Matlab 工具箱 Power system toolbox 进行电网机电暂态时域仿真, 将其作为智能体训练环境, 如图 1 所示。

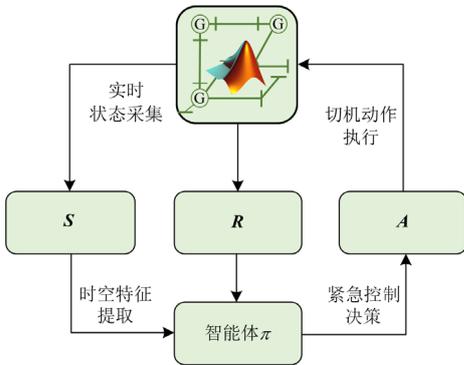


图 1 紧急控制智能体训练框架

Fig. 1 Emergency control agent training framework

根据用户设置的运行方式参数和故障参数, 对镜像环境进行时域仿真, 并实时输出状态和提取时空特征, 将其输入智能体。智能体基于当前策略网络  $\pi$  进行决策, 并在镜像环境中模拟执行。根据镜像环境执行的效果, 计算奖励函数并反馈至智能体, 以优化提升策略。

## 2 面向暂态过程的电网时空特征提取

在电网遭受大扰动后, 电压、电流信号从故障

位置沿输电线路向整个电网传播并逐渐衰减。同时, 在故障传播路径上, 如果电气量超出电网保护定值或系统运行边界时, 可能引发连锁故障, 造成严重后果<sup>[26]</sup>。因此, 对于暂态过程的空间传播和电气量的时序变化趋势进行特征分析, 可以有效掌握电网暂态稳定状态。本节分别从空间和时间角度提出了暂态过程的特征提取方法。

### 2.1 基于 GCN 的空间特征提取

电网的基本结构可以抽象为负荷、发电厂、变电站等若干节点和输电线路构成的网络。本文研究的电网暂态稳定问题考虑了节点特征, 包括转子功角  $\delta$ 、转子转速  $\omega$ 、转子加速度  $a$ 、母线电压  $V$  和母线电流  $I$  等。

针对这一非欧空间的复杂网络, 利用 GCN 进行特征提取。对于任意一层  $H^l$  的计算, 可表示为

$$\begin{cases} H^{l+1} = f(\mathbf{D}, H^l) \\ H^0 = S_t \\ f(\mathbf{D}, H^l) = \sigma(\mathbf{D}H^lW^l) \end{cases} \quad (10)$$

式中:  $H^l$  为第  $l$  层图神经网络层;  $\mathbf{D}$  为邻接矩阵;  $f$  为图卷积层计算函数;  $H^0$  为第 0 层图神经网络层;  $S_t$  为  $t$  时刻的状态变量;  $\sigma$  为非线性激活函数 ReLU;  $W^l$  为网络权重系数。

电网结构通过节点与边构成的邻接矩阵  $\mathbf{D}$  中元素 0 表示节点间无直接联系, 元素 1 表示节点间直接相连。基于 GCN 提取电网暂态过程空间特征如图 2 所示。

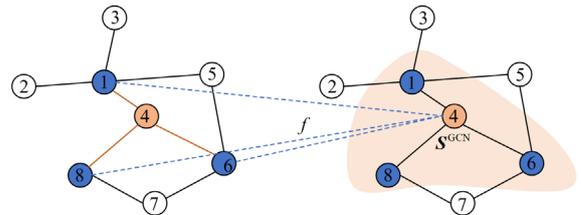


图 2 基于 GCN 提取电网暂态过程空间特征

Fig. 2 GCN extraction of grid transient process spatial features

经过上述计算后, 电网中不同节点之间的交互影响和不同特征之间的相互耦合关系被抽象提取为空间特征向量  $S^{\text{GCN}}$ 。

### 2.2 基于 LSTM 的时间特征提取

电力系统暂态过程通常在秒级时间区间内发生, 从故障发生到机组失稳非常迅速。因此, 需要快速进行电网功角稳定态势分析, 并采取紧急措施以避免失稳现象。根据电力系统摇摆方程理论, 摇摆曲线具有正弦耦合特征, 可以通过线性系统和正则摄动技术进行近似拟合<sup>[27]</sup>。因此, 从数学角度来

看, 电力系统的暂态稳定过程具有时序规律, 可以通过神经网络进行学习和预测。

通过对电网状态在毫秒时间尺度上的时序变化趋势提取和预测, 可以反应功角扰动轨迹。以  $S^{GCN}$  为 LSTM 输入向量, 其时序特征提取过程如图 3 所示。

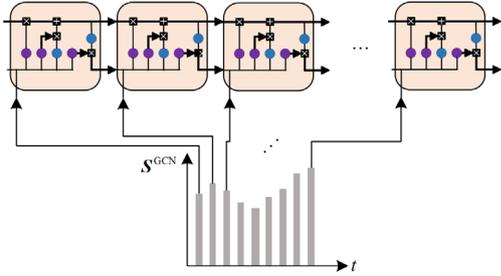


图 3 LSTM 提取电网暂态过程时序特征

Fig. 3 LSTM extraction of transient process of temporal features

图中展示了 LSTM 的遗忘门、输入门和输出门连接关系, 其计算过程为

$$\begin{cases} I_t = \sigma(S_t W_{xi} + H_{t-1} W_{hi} + b_i) \\ F_t = \sigma(S_t W_{xf} + H_{t-1} W_{hf} + b_f) \\ O_t = \sigma(S_t W_{xo} + H_{t-1} W_{ho} + b_o) \\ \tilde{C}_t = \tanh(S_t W_{xc} + H_{t-1} W_{hc} + b_c) \\ C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \\ H_t = O_t \odot \tanh(C_t) \end{cases} \quad (11)$$

式中:  $I_t$  为记忆门;  $F_t$  为遗忘门;  $O_t$  为输出门;  $\tilde{C}_t$  为候选记忆单元;  $C_t$  为记忆单元;  $C_{t-1}$  为上一个时刻的记忆单元;  $H_t$  为隐藏状态;  $H_{t-1}$  为上一个时刻的隐藏状态;  $W_{xi}$  为记忆门网络权重系数;  $W_{hi}$  为隐藏记忆门网络权重系数;  $b_i$  为记忆门网络偏置系数;  $W_{xf}$  为遗忘门网络权重系数;  $W_{hf}$  为隐藏遗忘门网络权重系数;  $b_f$  为遗忘门网络偏置系数;  $W_{xo}$  为输出门网络权重系数;  $W_{ho}$  为隐藏记忆门网络权重系数;  $b_o$  为输出门网络偏置系数;  $W_{xc}$  为记忆门网络权重系数;  $W_{hc}$  为隐藏记忆门网络权重系数;  $b_c$  为记忆门网络偏置系数; “ $\odot$ ” 为矩阵按元素位置相乘的计算符号。

通过多种门控记忆元的协同处理, 将时间序列中的长期趋势和短期波动综合分析, 得到暂态功角稳定态势感知的时空特征输出向量  $Y_t$ 。基于此, 可以进行紧急控制切机策略的智能分析决策。

### 3 稳控策略智能体

强化学习将感知与决策进行融合, 形成了具有

自主迭代优化能力的智能体。考虑到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探索。因此, 本文利用安全层对近端策略优化(proximal policy optimization, PPO)算法的结果进行安全校正, 并将其作为智能体训练算法。

#### 3.1 目标函数

PPO 算法是对信任域策略优化(trust region policy optimization, TRPO)算法的改进<sup>[28]</sup>, 该算法利用一阶优化近似表示其目标函数, 从而降低计算复杂度。

$$\max_{\theta} L^{\text{clip}}(\theta) = \mathbb{E}_t \left[ \min[R_t(\theta) A_t, \text{clip}(R_t(\theta), 1 - \epsilon, 1 + \epsilon) A_t] \right] \quad (12)$$

式中:  $L^{\text{clip}}(\theta)$  表示以  $\theta$  为网络参数集的损失函数;  $\mathbb{E}_t$  为  $t$  时刻的数学期望;  $R_t(\theta)$  表示更新后的策略与更新之前策略面临相同状态  $S_t$  时采取动作  $A_t$  的概率比值;  $A_t$  为优势函数, 表示 PPO 网络策略相比于随机策略的奖励函数优势。

式(12)的第二项通过 clip 函数将  $R_t(\theta)$  约束至  $[1 - \epsilon, 1 + \epsilon]$  区间内, 提高了策略更新过程的收敛性。当  $A_t > 0$  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 策略更新

PPO 算法训练基于固定的暂态响应轨迹数据进行, 每次从故障集中随机抽取部分数据。

表 1 PPO 算法策略更新流程

Table 1 Update process of PPO algorithm policy	
输入	训练样本, 训练回合数 $EP$
输出	智能体参数集 $\theta$
1	For $i = 1, 2, \dots$
2	For actor = 1, 2, ..., $EP$
3	在环境中执行 $T$ 步策略 $\pi_{\theta}$
4	计算优势函数 $A_1, A_2, \dots, A_t, \dots, A_T$
5	End for
6	利用训练样本优化网络参数集 $\theta'$
7	$\theta' \rightarrow \theta$
8	End for

#### 3.3 动作安全层

为了确保紧急控制动作符合 2.2.4 节所列举的安全约束, 策略网络输出层后级联一层安全层(safety layer, SL)<sup>[29]</sup>, 将输出动作  $A_t$  校核后输出。

图 4 中, 安全层通过计算符合安全约束的近似动作, 确保切机操作安全性与合理性, 其过程为

$$\begin{cases} \arg \min_{A^*} \frac{1}{2} \|A^* - \pi_{\theta}(Y_t)\|^2 \\ \text{s.t. 式(5)一式(8)} \end{cases} \quad (13)$$

式中： $A^*$  为神经网络输出后的安全动作； $\pi_\theta$  为神经网络参数集。

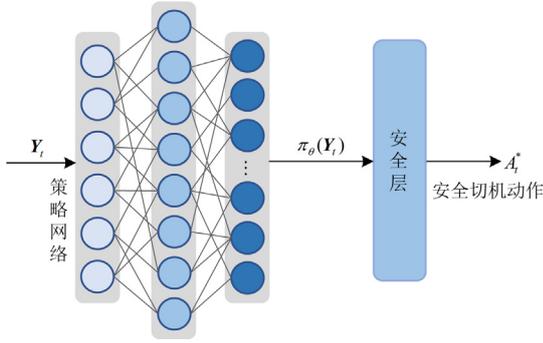


图 4 动作安全约束实施

Fig. 4 Implementation of action safety constraints

通过优化算法求解上述问题，输出符合安全约束且与原决策动作相似度最高的紧急控制动作。

## 4 算例分析

### 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 G1、G3 和 G9 替换为风电机组模型，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%。电网运行负荷水平、机组出力、故障位置和故障类型等参数如表 2 所示，共计 3200 组故障场景用于智能体训练，其中失稳场景为 1328 组，稳定场景为 1872 组，仿真时长为 10 s，仿真步长为 0.01 s。Power system toolbox 仿真平台计算硬件配置为：Intel Core (TM) i5-4210U CPU、NVIDIA RTX-3060 GPU 和 16 GB 内存。

表 2 仿真场景概率设置

Table 2 Simulation scenario probability settings

属性	设置区间	概率模型
负荷水平	[0.75,1.25]	均匀分布
发电功率	[0.75,1.25]	均匀分布
故障位置	所有线路	等概率模型
故障类型	单相短路接地	等概率模型
	两相短路接地	
	相间短路	
	三相短路接地	

### 4.2 对比方法

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 3 所示。

上述方法的基础强化学习算法均为 PPO 算法，本文所提方法在此基础上融入了时空特征提取网络 (GCN + LSTM) 和领域知识动作约束 (SL)，并分别

构造了两组对照方法。具体测试对比结果见 4.3 节。

表 3 算例对比方法

Table 3 Comparison of different methods

方法	实验组与对照组
GCN+LSTM+PPO+SL	本文方法实验组
PPO+SL	特征感知对照组
GCN+LSTM+PPO	领域知识对照组
PPO	基础算法对照组

### 4.3 性能对比

首先，对比了不同算法在训练阶段的奖励函数和动作约束违反次数。从图 5 中可以看出，作为基准对比的 PPO 算法训练效率和收敛值相比其他方法均无优势。相比之下，本文方法的奖励函数值和收敛速度均优于其他算法，1000 轮时奖励函数达到 27.4，而在 2150 轮次左右逐渐收敛于 38.2。

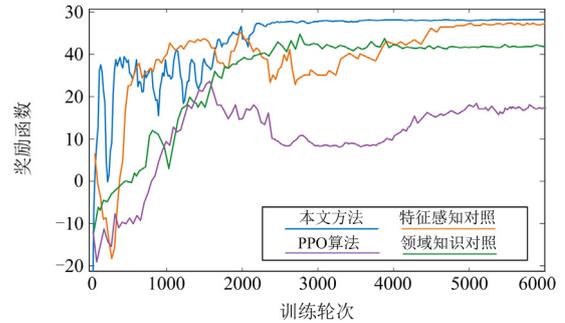


图 5 决策奖励函数值对比(滑动平均)

Fig. 5 Decision reward function value comparison (moving average)

特征感知对照算法的优化速度较慢，在 4250 轮左右奖励函数收敛于 36.3。领域知识对照算法不仅优化速度较慢，且收敛值小于前两种方法。最终第 2500 轮训练后收敛，收敛值仅为本文方法的 83.8%。

不同算法对于安全约束的训练过程如图 6 所示，采用了融入领域知识安全层的紧急控制智能体能够完全避免违反安全约束，极大提高了训练效率。而未考虑领域知识的纯探索型智能体，违反约束次数下降较慢且难以达到理想效果，最终收敛于 10 左右。

### 4.4 泛化分析

为了验证本文所提方法的泛化性能，本节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 0.18 s。

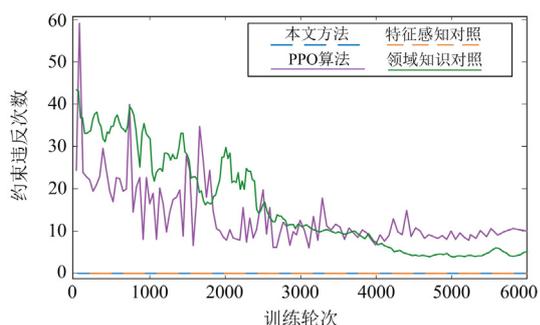


图6 安全约束违反次数对比(滑动平均)

Fig. 6 Comparison of security constraint violations (moving average)

如图7所示, 发电机1、2、3在故障后功角和转速波动, 若不采取动作则沿虚线路径演化造成功角失稳。智能体在故障切除后快速执行切机动作, 其余机组经过一段时间的振荡后恢复正常状态, 确保了系统安全。

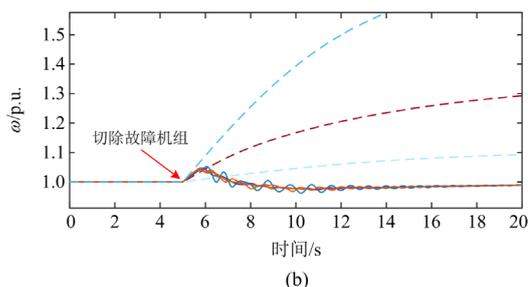
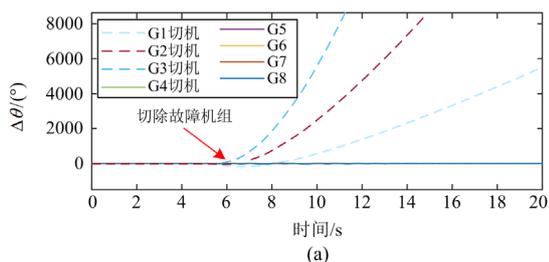


图7 智能体紧急切机控制效果

Fig. 7 Control effect of agent emergency control

#### 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, 本节将基于 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 237 个母线节点, 32 个发电机母线节点, 其中有 25 个火电机组并网节点, 5 个风电场并网节点, 2 个光伏电站并网节点, 算例场景中新能源功率渗透率为 10%~20%。

结果展示了与多个对比方法的训练与执行计算耗时, 如表 4 所示。

由于特征感知网络结构复杂, 本文方法单轮训练耗时较长, 但受益于高维特征感知能力, 本文的收敛轮数相比其他方法具有明显优势, 因而总体训练时间依然最快。稳定控制的应用场景对离线训练阶段的时间要求较低, 故本文方法的计算效率符合实际需要。

表4 不同算法计算效率对比

Table 4 Comparison of algorithm calculation efficiency

算法	IEEE 39 节点系统		某实际电网(237 节点)	
	单轮时间/s×收敛轮数=总时长/s	决策时间/ms	单轮时间/s×收敛轮数=总时长/s	决策时间/ms
本文	15.2×2150 = 32 680	13.4	16.6×5300 = 87 980	17.5
特征感知对照	13.5×2500 = 33 750	12.8	14.6×6700 = 97 820	15.3
领域知识对照	7.8×4250 = 33 150	8.3	10.5×10050 = 105 525	11.6
基础算法对照	6.6×5500 = 36 300	6.4	8.3×13500 = 112 050	8.9

此外, 由于安全层需要进行优化计算, 因而决策时间略长于无领域知识约束的算法。对比两个不同规模系统的测试结果, 本文方法的训练时长增长最少, 未出现“维数灾”问题, 具备迁移至更大规模系统的可行性。

## 5 结论

在新型电力系统建设背景下, 电网稳定机理和稳定模式愈加复杂, 传统紧急控制策略对于不同场景的适应性下降。本文首先将暂态功角稳定紧急控制建模为序贯决策问题, 构建了含约束马尔可夫模

型, 并利用强化学习智能体进行迭代交互, 不断训练优化切机决策。

在策略生成方面, 本文提出了 GCN 与 LSTM 级联而成的电网暂态稳定时空特征感知网络层, 提高了智能体对于电网暂态稳定趋势的理解能力。同时, 利用安全强化学习技术将电网运行安全约束加入智能体动作选择中, 提高了智能体训练效率, 并确保了切机动作的合理性和安全性。

通过 IEEE 39 节点系统和某实际电网的测试, 验证了所提方法的准确性和快速性。此外, 加入动作安全性的领域知识后, 智能体训练效率相比单一

的强化学习分别提高了 10%和 21%。

下一步的研究计划将考虑在紧急控制措施中加入切负荷控制, 以应对更加复杂多样的场景。

### 参考文献

- [1] 齐金山, 姚良忠, 廖思阳, 等. 高比例新能源电力系统静态电压稳定裕度在线概率评估[J]. 电力系统保护与控制, 2023, 51(5): 47-57.  
QI Jinshan, YAO Liangzhong, LIAO Siyang, et al. Online probabilistic assessment of static voltage stability margin for power systems with a high proportion of renewable energy[J]. Power System Protection and Control, 2023, 51(5): 47-57.
- [2] 骆钊, 聂灵峰, 田肖, 等. 基于模型预测控制的电解铝负荷参与电网频率稳定控制策略[J]. 电力建设, 2023, 44(10): 1-11.  
LUO Zhao, NIE Lingfeng, TIAN Xiao, et al. Control strategy of electrolytic aluminum load participating in power network frequency stability based on model predictive control[J]. Electric Power Construction, 2023, 44(10): 1-11.
- [3] 李群, 张宇宇, 王新宝, 等. 电力潮流灵活控制技术应用综述[J]. 电力工程技术, 2023, 42(1): 50-60.  
LI Qun, ZHANG Ningyu, WANG Xinbao, et al. Review on application of flexible power flow control technology[J]. Electric Power Engineering Technology, 2023, 42(1): 50-60.
- [4] 蒋玮, 撒鹏程, 贾俊, 等. 基于潮流追踪模块度的虚拟微电网分区规划[J]. 电力系统自动化, 2023, 47(24): 132-142.  
JIANG Wei, SA Pengcheng, JIA Jun, et al. Virtual microgrid partitioning planning based on modularity of power flow tracking[J]. Automation of Electric Power Systems, 2023, 47(24): 132-142.
- [5] 郭庆来, 兰健, 周艳真, 等. 基于混合智能的新型电力系统运行方式分析决策架构及其关键技术[J]. 中国电力, 2023, 56(9): 1-13.  
GUO Qinglai, LAN Jian, ZHOU Yanzhen, et al. Architecture and key technologies of hybrid-intelligence-based decision-making of operation modes for new type power systems[J]. Electric Power, 2023, 56(9): 1-13.
- [6] 刘映尚, 张建新, 徐光虎, 等. 南方区域复杂交直流互联电网系统运行特性与安全稳定控制[J]. 南方电网技术, 2020, 14(5): 44-50.  
LIU Yingshang, ZHANG Jianxin, XU Guanghu, et al. Operation characteristics and safety and stability control of complex AC/DC interconnected power grids in the southern region[J]. Southern Power System Technology, 2020, 14(5): 44-50.
- [7] 黄天罡. 统计分析与因果分析在暂态稳定量化评估中的融合[D]. 南京: 东南大学, 2018.  
HUANG Tiangan. Integration of statistical analysis and causal analysis in transient stability quantitative assessment[D]. Nanjing: Southeast University, 2018.
- [8] 张钢, 宗启航, 柯贤波, 等. 基于临界惯量和预想故障的含风电电力系统暂态功角稳定在线预警[J]. 电力系统保护与控制, 2023, 51(16): 72-83.  
ZHANG Gang, ZONG Qihang, KE Xianbo, et al. Online warning of transient power angle stability of power systems with wind power based on critical inertia and anticipated faults[J]. Power System Protection and Control, 2023, 51(16): 72-83.
- [9] 鲁广明, 张璐路, 马晶, 等. 基于时序特征选择与改进MSPCA算法的电网暂态稳定态势智能评估[J]. 电测与仪表, 2023, 60(6): 125-133.  
LU Guangming, ZHANG Lulu, MA Jing, et al. Intelligent transient stability situation assessment of powergrid based on time-series feature selection and improved MSPCA algorithm[J]. Electrical Measurement & Instrumentation, 2023, 60(6): 125-133.
- [10] 马海滨, 刘道伟, 赵高尚, 等. 基于交替方向乘子法的电网电压相量轨迹拟合及暂态稳定评估指标构建[J]. 电力建设, 2023, 44(12): 136-147.  
MA Haibin, LIU Daowei, ZHAO Gaoshang, et al. Power grid voltage phasor trajectory fitting and transient stability evaluation index construction based on alternate direction multiplier method[J]. Electric Power Construction, 2023, 44(12): 136-147.
- [11] 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.
- [12] 丁浩寅, 周磊, 黄志光, 等. 新能源与多馈入直流的受端电网频率紧急控制策略[J]. 电力系统及其自动化学报, 2023, 35(11): 74-83.  
DING Haoyin, ZHOU Lei, HUANG Zhiguang, et al. Frequency emergency control strategy for grids with new energy sources and multi-feed-in DC at the receiving end[J]. Proceedings of the CSU-EPSA, 2023, 35(11): 74-83.
- [13] 李锡林, 查晓明, 田震, 等. 频率突变影响下基于Lyapunov法的孤岛微电网暂态稳定性分析[J]. 电工技术学报, 2023, 38(增刊 1): 18-31, 55.  
LI Xilin, ZHA Xiaoming, TIAN Zhen, et al. Lyapunov based transient stability analysis of islanded microgrid under

- the influence of frequency abrupt change[J]. Transactions of China Electrotechnical Society, 2023, 38(S1): 18-31, 55.
- [14] 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.
- [15] XIAO T, WANG J, GAO Y, et al. Improved sparsity techniques for solving network equations in transient stability simulations[J]. IEEE Transactions on Power Systems, 2018, 33(5): 4878-4888.
- [16] 周二专, 冯东豪, 武志刚. 内存计算技术及在电网分析中的应用[J]. 电力系统自动化, 2017, 41(11): 1-7, 25.  
ZHOU Ezhuan, FENG Donghao, WU Zhigang. Memory computing technology and its application in power grid analysis[J]. Automation of Electric Power Systems, 2017, 41(11): 1-7, 25.
- [17] 朱旭凯, 周孝信, 田芳, 等. 基于电力系统全数字实时仿真装置的大电网机电暂态 - 电磁暂态混合仿真[J]. 电网技术, 2011, 35(3): 26-31.  
ZHU Xukai, ZHOU Xiaoxin, TIAN Fang, et al. Hybrid simulation of electromechanical transients and electromagnetic transients in large power grids based on a full digital real-time simulation device[J]. Power System Technology, 2011, 35(3): 26-31.
- [18] 高发骏, 王怀远, 党然. 基于 Transformer 的暂态稳定评估模型的可解释性分析与模型更新研究[J]. 电力系统保护与控制, 2023, 51(17): 15-25.  
GAO Fajun, WANG Huaiyuan, DANG Ran. Interpretability analysis and model update research of a transient stability assessment model based on Transformer[J]. Power System Protection and Control, 2023, 51(17): 15-25.
- [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] ZENG H, ZHOU Y, GUO Q, et al. A 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.
- [21] XIE J, SUN W. Distributional deep reinforcement learning-based emergency frequency control[J]. IEEE Transactions on Power Systems, 2021, 37(4): 2720-2730.
- [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] ZHANG S, TONG H, XU J, et al. Graph convolutional networks: a comprehensive review[J]. Computational Social Networks, 2019, 6(1): 1-23.
- [24] YU Y, SI X, HU C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.
- [25] 国家能源局. 电力系统安全稳定导则: GB38755—2019[S]. 北京: 中国标准出版社, 2019.  
National Energy Administration. Code on security and stability for power system: GB 38755—2019[S]. Beijing: Standards Press of China, 2019.
- [26] 胡福年, 陈灵娟, 陈军. 基于交流潮流的连锁故障建模与鲁棒性评估[J]. 电力系统保护与控制, 2021, 49(18): 35-43.  
HU Funian, CHEN Lingjuan, CHEN Jun. Cascading failure modeling and robustness evaluation based on AC power flow[J]. Power System Protection and Control, 2021, 49(18): 35-43.
- [27] 郝晓光, 李剑锋, 王智, 等. 基于 PLL 同步的弱连接 VSC 在严重电网故障下的解析注入电流稳定域构建[J]. 电力系统及其自动化学报, 2023, 35(5): 145-151.  
HAO Xiaoguang, LI Jianfeng, WANG Zhi, et al. Construction of analytical injected current transient stability region in weak grid-connected VSC based on PLL synchronization under severe grid faults[J]. Proceedings of the CSU-EPSA, 2023, 35(5): 145-151.
- [28] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[J]. arXiv preprint arXiv:1707.06347, 2017.
- [29] CEUSTERS G, PUTRATAMA M A, FRANKE R, et al. An adaptive safety layer with hard constraints for safe reinforcement learning in multi-energy management systems[J]. Sustainable Energy, Grids and Networks, 2023, 36: 101202.

收稿日期: 2023-10-20; 修回日期: 2024-02-15

作者简介:

邱建(1982—), 男, 硕士, 正高级工程师, 研究方向为电力系统安全稳定分析与控制; E-mail: qiu Jian@csg.cn

朱煜昆(1988—), 男, 通信作者, 硕士, 高级工程师, 研究方向为电力系统仿真与控制技术; E-mail: zhuyk@csg.cn

张建新(1983—), 男, 硕士, 正高级工程师, 研究方向为电力系统安全稳定分析与控制。E-mail: zhangjx@csg.cn

(编辑 姜新丽)