

DOI: 10.19783/j.cnki.pspc.240501

融合电网拓扑信息的分支竞争 Q 网络智能体紧急切负荷决策

潘晓杰¹, 胡泽², 姚伟², 兰宇田², 徐友平¹, 王玉坤¹, 张慕婕¹, 文劲宇²

(1. 国家电网公司华中分部, 湖北 武汉 430070; 2. 强电磁技术全国重点实验室
(华中科技大学电气与电子工程学院), 湖北 武汉 430074)

摘要: 暂态电压失稳事件紧急控制措施制定是电力系统仿真分析中的一个重要内容, 离线预先制定紧急切负荷决策, 在线匹配执行决策方案。但该工作目前主要依赖专家分析海量仿真数据得到, 耗时耗力。因此提出了一种融合电网拓扑信息的分支竞争 Q 网络智能体的电力系统紧急切负荷决策方法, 以提高离线紧急切负荷决策的效率。首先, 建立了一种基于事件驱动的马尔科夫决策过程, 可以有效指导深度强化学习智能体的训练。其次, 设计了一种分支竞争 Q 网络智能体, 相比传统无分支网络, 基于分支竞争 Q 网络智能体具有更强的训练效率和决策能力。然后, 为进一步增强智能体的训练效率和决策性能, 通过图卷积增强将电力系统拓扑信息融入到智能体的训练过程。最后, 在中国电力科学研究院 8 机 36 节点系统进行了验证。相比于无分支网络和无拓扑信息融入的深度强化学习智能体, 所提方法具有更高的训练效率和决策性能。

关键词: 仿真分析; 暂态电压失稳; 紧急切负荷决策; 深度强化学习; 分支竞争 Q 网络; 电网拓扑信息; 图卷积增强

Emergency load shedding decision-making using a branching dueling Q-network integrating grid topology information

PAN Xiaojie¹, HU Ze², YAO Wei², LAN Yutian², XU Youping¹, WANG Yukun¹, ZHANG Mujie¹, WEN Jinyu²
(1. Central China Branch of State Grid Corporation of China, Wuhan 430070, China; 2. State Key Laboratory of
Advanced Electromagnetic Engineering and Technology (School of Electrical and Electronic Engineering,
Huazhong University of Science and Technology), Wuhan 430074, China)

Abstract: The formulation of emergency control measures for transient voltage instability events is a crucial aspect of power system simulation analysis. Traditionally, emergency load shedding decisions are pre-determined offline and matched for execution in real-time. However, this process heavily relies on expert analysis of massive amounts of simulation data, which is both time-consuming and labor-intensive. To improve the efficiency of offline emergency load shedding decision-making, this paper presents a method for power system emergency load shedding decisions that integrates power grid topology information into a branching dueling Q-network (BDN) agent. First, an event-driven Markov decision process (MDP) is established to effectively guide the training of the deep reinforcement learning agents. Second, a BDN agent is designed, which exhibits superior training efficiency and decision-making capability compared to traditional non-branching networks. Then, to further enhance the agent's training efficiency and decision-making performance, power grid topology information is integrated into the agent's training process through graph convolutional networks (GCN). Finally, the proposed method is validated on the 8-machine 36-node system of the China Electric Power Research Institute. Compared to non-branching networks and deep reinforcement learning agents without integrated topology information, the proposed method demonstrates higher training efficiency and better decision-making performance.

This work is supported by the National Natural Science Foundation of China (No. U22B20111).

Key words: simulation analysis; transient voltage instability; emergency load shedding decision-making; deep reinforcement learning; branching dueling Q-network; power grid topology information; graph convolution enhancement

基金项目: 国家自然科学基金项目资助 (U22B20111); 国家电网公司科技项目资助“基于数据-知识混合驱动的电网安全稳定分析关键技术研究” (52140023000S)

0 引言

仿真分析广泛应用于电力系统运行、规划、安全稳定防御等领域,是电力系统一项重要的基础性支撑技术^[1-2],具体包括仿真计算和仿真数据分析两部分。仿真数据分析中的一个重要内容是基于仿真数据进行电力系统暂态电压失稳时^[3]的紧急控制措施制定。目前,根据决策逻辑和实施方式的不同,电力系统暂态电压失稳紧急决策可以分为3类^[4-5],其中由人工专家根据预想故障离线预决策控制措施,写入操作手册,实际故障发生时快速匹配执行的方法最为常用^[6]。随着电力系统规模不断扩大,运行工况日益复杂,预想暂态电压失稳故障数量激增。而传统的基于专家经验反复试验以寻找最优决策措施的方法已难以满足离线预整定对实时性的需求^[7]。

为了提高人工预决策的效率,一些研究提出了启发式优化算法来制定电力系统暂态电压失稳时的紧急控制决策。文献[8]基于节点电压相对于控制变量的轨迹灵敏度指标选择负荷进行切除;文献[9]同样基于轨迹灵敏度指标选择发电机或负荷进行切除;文献[10]结合拓展等面积法计算各紧急措施的性能指标,采用贪心算法进行决策尝试。但这些方法都有一个共同问题,当控制对象增多时,这些方法均耗时且不易收敛。

近年来,人工智能技术快速发展,为紧急切负荷预决策带来新的机遇^[11]。其中,深度强化学习(deep reinforcement learning, DRL)技术已应用到电力市场交易^[12]、电网边缘计算资源分配^[13]和电力系统控制器参数优化领域^[14]。相比人工决策方法,其决策效率得到了极大的提升。因此,利用 DRL 技术进行暂态电压稳定紧急切负荷决策,将有望极大地减轻专家依赖,提升决策效率。目前,已有一些研究利用 DRL 技术进行切负荷决策,文献[15]提出了一种基于风险规避的 DRL 紧急切负荷算法,考虑控制代价的同时提高决策成功率。文献[16]提出了一种基于响应的 DRL 紧急切负荷算法,考虑多种控制场景以提高算法鲁棒性。然而,已有方法仍存在以下不足。

首先,现有的马尔科夫决策过程(Markov decision process, MDP)对应于基于响应的紧急切负荷决策过程,鲜有文献涉及到基于事件的 MDP,现有的 MDP 存在不适用的情况^[17]。其次,对于离散紧急切负荷决策,智能体的决策空间随着可控负荷的增加呈指数爆炸增长。当可控负荷较多时,智能体几乎无法训练。虽然有研究利用多智能体决策来减少决策空间,但是多智能体的训练和协调较为困难^[18]。最后,现有文献对于故障引起的电力系统拓扑变化考

虑不足,通常采用卷积神经网络等提取暂态电压信息^[19],DRL 智能体决策能力受限。

为了解决以上现有方法的不足,本文提出了一种融合电网拓扑信息的分支竞争 Q 网络智能体紧急切负荷决策方法,以提高离线紧急切负荷决策的效率。首先,基于暂态稳定仿真建立了深度 Q 网络(deep Q network, DQN)智能体与电力系统环境的基于事件的 MDP,以指导 DQN 智能体的训练。其次,提出了一种分支竞争网络(branching dueling network, BDN),将决策空间维度由指数降为线性,极大地提升了 DQN 智能体训练效率和决策能力。然后,设计图卷积神经网络(graph convolutional neural network, GCN)将故障前后的电力系统拓扑变化融入到智能体的训练过程,进一步提升智能体的训练效率和决策能力。最后,在中国电力科学研究院(China electric power research institute, CEPRI)8 机 36 节点系统验证了所提融合电网拓扑信息的分支竞争 Q 网络智能体紧急切负荷决策方法(GCN-BDN-DQN)的优势。

1 基于 GCN-BDN-DQN 的紧急切负荷决策框架

如图 1 所示,基于 GCN-BDN-DQN 的紧急切负荷决策框架包括 3 部分。第一部分是暂态电压失稳事件数据输入,第二部分是 GCN-BDN-DQN 的训练过程,第三部分是训练好的 GCN-BDN-DQN 智能体根据暂态电压失稳事件生成紧急切负荷决策。数据输入部分,首先根据预想的暂态电压失稳事件,例如电力系统三相短路故障造成的暂态电压失稳,并结合时域仿真获取母线电压和发电机功角的时序仿真数据,再利用 GCN 提取母线电压和电力系统拓扑特征。然后,智能体依据母线电压和电力系统拓扑特征对决策空间进行探索,和电力系统环境进行交互,并对智能体参数进行更新。最后,在线应用阶段,当检测到暂态电压失稳事件发生后,将电压图特征和发电机功角采样数据输入到已经离线预训练好的智能体中,立即获取紧急切负荷决策,匹配控制措施立即执行。总的来说,在线匹配执行较为容易,难点在于离线时智能体的训练决策。

离线训练时,GCN-BDN-DQN 智能体在与电力系统环境的 MDP 交互过程中逐渐学习到有效的紧急切负荷决策。每一轮迭代中,电力系统环境根据预想的暂态电压失稳事件,调用暂态稳定仿真计算程序,获取节点电压和发电机功角的仿真数据,并利用 GCN 提取电力系统拓扑特征,作为当前状态输入到 GCN-BDN-DQN 智能体。智能体根据当前状态

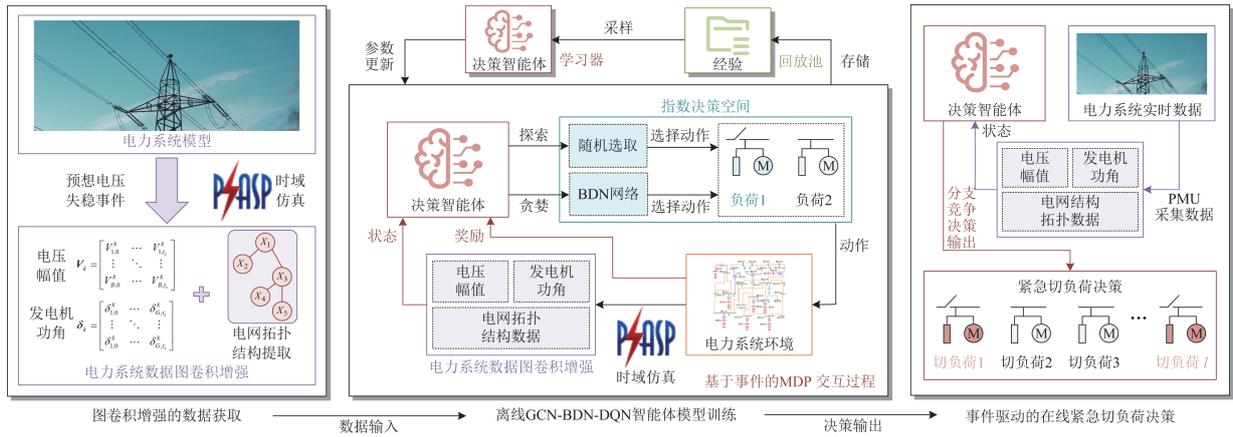


图 1 基于 GCN-BDN-DQN 的紧急切负荷决策框架

Fig. 1 Emergency load shedding decision framework based on GCN-BDN-DQN

利用贪婪或探索策略决策当前紧急切负荷动作。特别地, 智能体的贪婪策略决策网络为设计的 BDN, 相比传统无分支 DQN, 决策能力更强。最后, 电力系统环境执行智能体的切负荷决策, 调用暂态稳定仿真更新当前电压图特征和功角状态, 并根据决策措施的有效性反馈奖励。交互过程存储到经验回放池, 智能体定期从经验池抽取一部分经验更新其 BDN 网络参数。

2 事件驱动的 MDP 元素设计

MDP 用于指导 DRL 智能体的训练, 由图 1 可以看到, 整个 MDP 过程根据暂态电压失稳事件展开, GCN-BDN-DQN 智能体在预想暂态电压失稳事件发生后立刻一次性给出紧急切负荷决策, 而电力系统环境根据预想暂态电压失稳事件和 GCN-BDN-DQN 智能体的决策, 更新当前电压图特征和功角状态。

MDP 中的状态、动作、奖励、完成标志以及状态转移如下所述。

1) 状态: 状态 s_k 可以描述为

$$s_k = [V_k, \delta_k] \quad (1)$$

式中: V_k 表示步骤 k 处的母线电压幅值; δ_k 表示步骤 k 处的发电机功角。母线电压幅值和功角通过暂态稳定仿真得到, 它们分别具有以下形式。

$$V_k = \begin{bmatrix} V_{1,0}^k & \cdots & V_{1,t_s}^k \\ \vdots & & \vdots \\ V_{B,0}^k & \cdots & V_{B,t_s}^k \end{bmatrix} \quad (2)$$

$$\delta_k = \begin{bmatrix} \delta_{1,0}^k & \cdots & \delta_{G,t_s}^k \\ \vdots & & \vdots \\ \delta_{1,0}^k & \cdots & \delta_{G,t_s}^k \end{bmatrix} \quad (3)$$

式中: V_{B,t_s}^k 表示在 t_s 时刻步骤 k 处母线 B 的电压幅值; δ_{G,t_s}^k 表示在 t_s 时刻步骤 k 处发电机 G 的功角。

为方便训练, 需要对变量进行归一化处理, 可将 δ_k 中所有数据通过式(4)归一化到 $[-1,1]$, 式中, 向左箭头表示对原始数据进行处理并赋予新值。

$$\delta_{G,t_s}^k \leftarrow \tanh(\delta_{G,t_s}^k / 2\pi) \quad (4)$$

2) 动作: 动作 a_k 采用 ϵ -greedy 策略。

$$a_k = [a_k^1, \cdots, a_k^i, \cdots, a_k^m] \quad (5)$$

式中, a_k^i 表示在步骤 k 处负荷 i 的子动作。动作 a_k 由策略网络决定, 如式(6)所示。

$$\begin{cases} N_k = \arg \max_j Q(s_k; j), & \epsilon_k \geq \epsilon_0 \\ N_k = \text{random}, & \epsilon_k < \epsilon_0 \\ \epsilon_k = \epsilon_c + (\epsilon_0 - \epsilon_c) e^{-n_k / n_{\text{decay}}} \end{cases} \quad (6)$$

式中: j 表示输出神经元的序号; N_k 表示在步骤 k 处选择的输出神经元的序号, 其序号值对应某一负荷的某一动作 a_k^i , 将在后续介绍; ϵ -greedy 采用指数衰减策略, ϵ_0 和 ϵ_c 分别为初始和最小探索概率, 而 ϵ_k 表示在步骤 k 处的探索概率; n_k 表示总迭代次数; n_{decay} 为衰减系数。

$Q(s_k, a_k^i)$ 根据 Q 网络计算。对于无分支的 DQN 算法, Q 网络参数的更新方法如下。

首先, 计算单个样本的时间差分误差。

$$e(\theta_w) = \gamma Q'(s_{k+1}, \arg \max_a Q(s_{k+1}, a; \theta_w); \theta_w') + r_{k+1} - Q(s_k, a_k; \theta_w) \quad (7)$$

式中: Q 和 Q' 分别表示 Q 网络和目标网络; θ_w 和 θ_w' 分别表示第 w 次更新时 Q 网络和目标网络的网络参数; a 表示下一步骤中所有可能的动作; $e(\theta_w)$ 表示第 w 次更新时 Q 网络评估的 Q 值与目标网络之间

的时序差分误差; γ 是折扣因子; r_{k+1} 表示下一步骤的总奖励, 将在后续介绍。

然后, 计算批量损失并根据梯度下降法更新 Q 网络的参数。

$$L(\theta_w) = \sum_{M=1}^{\Psi} (e_M(\theta_w))^2 \quad (8)$$

$$\theta_{w+1} = \theta_w - \alpha \nabla_{\theta_w} L(\theta_w) \quad (9)$$

式中: $L(\theta_w)$ 表示第 w 次更新时 Q 网络评估的 Q 值与目标网络之间的批量损失; $e_M(\theta_w)$ 表示第 M 个样本的时序差分误差; Ψ 表示从经验回放区中抽取的样本; α 表示学习率; $\nabla(\cdot)$ 表示梯度计算。

3) 结束标志: 结束标志用于评判智能体紧急切除负荷决策的有效性并指导奖励的设置。本文采用暂态电压严重性指数(transient voltage severity index, TVSI)来判断暂态电压稳定性^[20]。

$$T_{\text{VDI}(v,t)} = \begin{cases} \frac{V_{v,0} - V_{v,t}}{V_{v,0}}, & \text{if } \frac{|V_{v,0} - V_{v,t}|}{V_{v,0}} \geq \beta \\ 0, & \text{otherwise} \end{cases} \quad \forall t \in [t_2, t_f] \quad (10)$$

$$T_{\text{VSI}} = \frac{\sum_{v=1}^B \sum_{t=t_2}^{t_f} T_{\text{VDI}(v,t)}}{B \times (t_f - t_2)} \quad (11)$$

式中: $T_{\text{VDI}(v,t)}$ 表示第 v 条母线在 t 时刻的电压波动指标; $V_{v,t}$ 表示第 v 条母线在 t 时刻的电压幅值; 特别的, $V_{v,0}$ 表示第 v 条母线电压幅值初值; t_2 表示故障清除时刻; t_f 表示总仿真时间; B 表示母线总数目; β 表示可接受的电压偏差程度, 一般取 10%; T_{VSI} 越大, 说明电压波动越明显, 暂态电压稳定性越低, 一般取 6 为判断阈值。

完成标志 F_k^{done} 设置为

$$F_k^{\text{done}} = \begin{cases} 1, & T_{\text{VSI}} \leq T_{\text{thre}} \cap 0 \leq k \leq K_{\text{max}} \\ 0, & T_{\text{VSI}} > T_{\text{thre}} \cup k > K_{\text{max}} \end{cases} \quad (12)$$

式中: K_{max} 表示允许的最大迭代次数, 大于 K_{max} 表示无解; T_{thre} 是瞬态电压不稳定的阈值, 可取 6。

4) 奖励: 奖励考虑两部分, 一部分是成功决策奖励, 另一部分是紧急切除负荷决策量惩罚, 以减少不必要的过切负荷。

$$R_{\Delta} = -r_{\Delta} \times \left(\sum_{i=1}^m a_k^i - \sum_{i=1}^m a_{k-1}^i \right) \quad (13)$$

式中: R_{Δ} 表示紧急切除负荷决策量惩罚; $\sum_{i=1}^m a_k^i$ 表示

步骤 k 所有的子动作之和; $\sum_{i=1}^m a_{k-1}^i$ 表示步骤 $k-1$ 时

所有子动作的总和; r_{Δ} 是负奖励系数。

另一部分是成功切除负荷的奖励 R_T 来指导训练:

$$R_T = r_T \times F_k^{\text{done}} \quad (14)$$

式中, r_T 为正奖励系数。

总奖励 r_k 设计为

$$r_k = R_{\Delta} + R_T \quad (15)$$

5) 状态转移: 状态转移是通过电力系统分析综合程序(power system analysis synthesis program, PSASP)调用暂态稳定仿真来实现的。计算从 t_0 到 t_f 的下一个状态 V_k 和 δ_k 。然后截取 t_0 到 t_s 的数据作为智能体的状态输入。 t_0 表示仿真开始时间, t_s 表示智能体输入数据的时间上限, t_1 表示故障发生时间, t_2 表示故障清除时间, t_3 表示紧急切除负荷决策投入时间。

GCN-BDN-DQN 在智能体和环境交互过程中获得 MDP 的每个元素。这些元素作为交互样本放入到样本池中。智能体定期抽取样本以更新其 BDN 网络。交互样本 M 定义为

$$M = [s_k, \mathbf{a}_k, s_{k+1}, r_k, F_k^{\text{done}}] \quad (16)$$

3 基于 GCN 的电力系统拓扑特征提取

电力系统拓扑结构表示节点之间的连接关系。在发生导致暂态电压失稳的三相短路严重故障后, 故障线路将会被切除, 节点之间断开连接。因此, 故障前后电力系统拓扑结构会发生变化。拓扑结构的变化表征了直接的故障特征。若能将拓扑变化融入到智能体的训练过程之中, 预期能提升智能体的训练效率和决策性能^[21]。

如图 2 所示, GCN 和 CNN 均用于提取电压信息, 不同的是, GCN 不仅关注节点本身的电压特征(图中深色部分), 而且关注故障前后节点与邻近节点的连接特征(图中浅色部分)。CNN 仅关注节点本身的电压特征, 忽略了故障前后电力系统拓扑特征

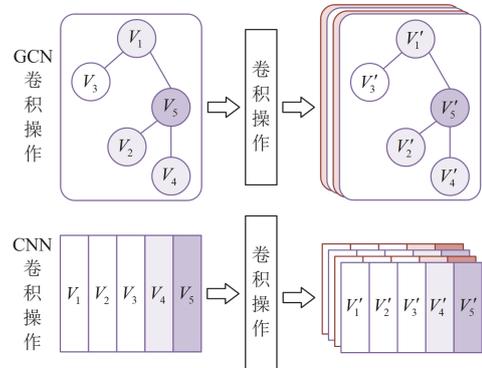


图 2 GCN 与 CNN 电压特征提取对比

Fig. 2 Comparison of voltage feature extraction between GCN and CNN

的变化, 恒定为数据输入时特征的排序。在严重故障导致的暂态电压失稳问题中, GCN 的潜在提取能力优于 CNN。

电网矩阵 \mathbf{G}_k 的维度为 $\mathbf{G}_k \in \mathbb{R}^{N \times N}$, 具体地, 可以用 g_k^{ij} 表示母线之间的连接关系。其中, 若母线 i 和母线 j 未连接, 则 $g_k^{ij} = 0$, 若母线 i 和母线 j 连接, 则 $g_k^{ij} = 1$ 。

多层图卷积运算中第 $l+1$ 层与第 l 层之间的传播关系为

$$\mathbf{V}_k^{(l+1)} = \sigma \left(\tilde{\mathbf{D}}_k^{-\frac{1}{2}} \tilde{\mathbf{G}}_k \tilde{\mathbf{D}}_k^{-\frac{1}{2}} \mathbf{V}_k^{(l)} \mathbf{W}_k^{(l)} \right) \quad (17)$$

式中: $\tilde{\mathbf{G}}_k$ 表示执行紧急切负荷决策 \mathbf{a}_k 后的电压图改进矩阵, $\tilde{\mathbf{G}}_k = \mathbf{G}_k + \mathbf{I}$, \mathbf{I} 为单位矩阵; $\tilde{\mathbf{D}}_k$ 表示 $\tilde{\mathbf{G}}_k$ 的度矩阵, 度矩阵为对角矩阵, 度矩阵 $\tilde{\mathbf{D}}_k$ 中的第 p 个对角元素值为 $\tilde{D}_k^{pp} = \sum_{q=1}^N q \tilde{G}_k^{pq}$, 其中, \tilde{G}_k^{pq} 表示电压图改进矩阵 $\tilde{\mathbf{G}}_k$ 中的第 p 行、第 q 列的值; $\mathbf{W}_k^{(l)}$ 表示计算电压图特征 \mathbf{T}_k 的多层图卷积运算的第 l 层网络的权重参数; $\sigma(\cdot)$ 表示非线性激活函数, 具体地, $\sigma(z) = \max(0, z)$, $\max(\cdot)$ 表示取最大值, z 表示待激活的函数值。

经过多层图卷积神经网络提取当前母线电压特征 \mathbf{V}_k 后, 得到当前电压图特征 \mathbf{T}_k 。其中, 多层图卷积运算的网络层数可以根据实际情况灵活调整。

经过 GCN 提取电力系统拓扑信息之后, 电压特征转化为电压图特征, 式(1)更新为

$$s_k = [\mathbf{T}_k, \delta_k] \quad (18)$$

4 分支竞争网络设计和智能体训练方法

4.1 分支竞争网络设计

对于 l 个负荷, 每个负荷 q 个动作的情况, 传

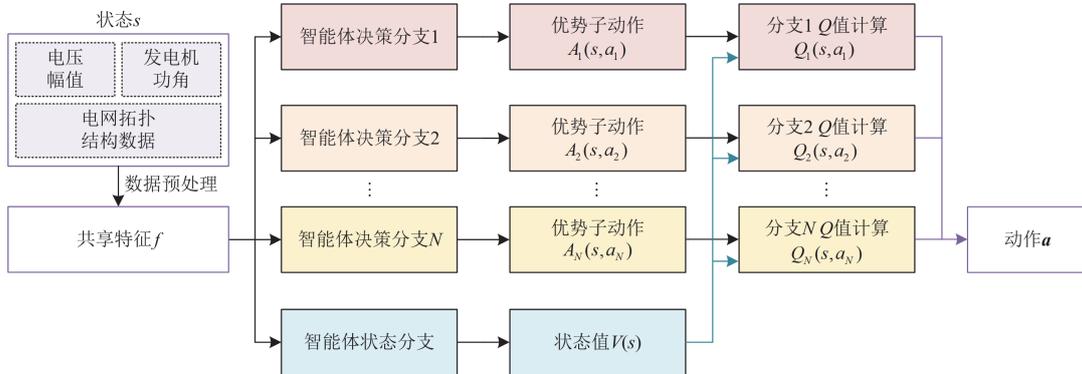


图3 BDN-DQN 策略网络设计

Fig. 3 Design of BDN-DQN strategy network

统无分支的 DQN 网络智能体输出神经元个数为 q^l 。随着系统增大, 可控负荷 l 增多时, DQN 智能体的输出神经元个数随着可控负荷呈指数爆炸增长, 智能体几乎无法训练。为了降低指数空间的维度, 可将单步决策问题转化为多步决策, 每轮只决策一个负荷的一个动作。输出神经元个数降为 $q \times l$ 。虽然决策空间维度降低, 但是 DQN 智能体迭代次数变多, 决策能力受限。因此, 本文设计了一种 BDN 网络, 输出神经元个数仅为 $q \times l$ 的同时, 能保留 DQN 智能体单步决策的能力。

假设最终的紧急切负荷决策是[1,2,3]。理想情况下, 单步决策的无分支 DQN 智能体需要 3 次迭代。第一次迭代选择负荷 1 的第一个子动作。第二次迭代选择负荷 2 的第二个子动作。在第三次迭代中选择负荷 3 的第三个子动作。理想情况下, 直接处理指数决策空间的 BDN-DQN 智能体仅需要 1 次迭代。分支 1 选择负荷 1 的第一个子动作。分支 2 选择负荷 2 的第二个子动作。分支 3 选择负荷 3 的第三个子动作。可以看出, 具有指数动作空间的分支 BDN-DQN 智能体的迭代次数较少, 并且决策能力更强。

如图 3 所示, 输入状态 s 通过特征预处理输出共享特征 f 。价值分支从共享特征 f 映射到状态值 $V(s)$, l 优势分支从共享特征 f 映射到优势子动作 $A_b(s, a_b)$ 。分支 b 的子动作 a_b 和其他可能的子动作 a'_b 属于 \mathcal{A}_b , $|\mathcal{A}_b|$ 表示 \mathcal{A}_b 的维度, 数值等于 q ; 每个优势分支 b 分别计算 $Q_b(s, a_b)$ 。

$$Q_b(s, a_b) = V(s) + \left[A_b(s, a_b) - \frac{1}{|\mathcal{A}_b|} \sum_{a'_b \in \mathcal{A}_b} A_b(s, a'_b) \right] \quad (19)$$

每个分支根据 Q 值和 ϵ -greedy 策略选择一个子动作。

$$\begin{cases} a_k^b = \pi^\varepsilon(s_k, \varepsilon_k) \\ a_k^b = \arg \max_j Q(s_k; j), \varepsilon_k \geq \varepsilon_0 \\ a_k^b = \text{random}, \varepsilon_k < \varepsilon_0 \\ \varepsilon_k = \varepsilon_c + (\varepsilon_0 - \varepsilon_c) e^{-n_k/n_{\text{decay}}} \\ u_k^b = a_k^b \times l_{\text{inte}}, 1 \leq b \leq l \end{cases} \quad (20)$$

式中: a_k^b 是步骤 k 中分支 b 的紧急切负荷子动作; j 是输出网络的神经元编号; u_k^b 表示第 b 个分支在第 k 步的紧急切负荷量; l_{inte} 表示紧急切负荷的间隔。

更新 Q 网络时, 可以采用双 Q 网络的方式计算分支 b 的目标值 Y_b 。

$$Y_b = r + \gamma Q_b \left(s', \arg \max_{a'_b \in \mathcal{A}_b} Q_b(s', a'_b; \theta); \theta' \right) \quad (21)$$

式中, θ 和 θ' 分别表示策略和目标网络的参数。

BDQ 的时间差分误差定义为所有分支的 Q_b 值与目标 Y_b 值之间的均方误差 $e_{\text{TD}}(\theta)$ 。

$$e_{\text{TD}}(\theta) = \mathbb{E} \left[\frac{1}{l} \sum_b (Y_b - Q_b(s, a_b; \theta))^2 \right] \quad (22)$$

4.2 智能体训练方法

对于 BDN-DQN 智能体, 其训练流程图如图 4 所示。设置总训练轮次 N_T , 批样本大小 Ψ , 目标网络同步周期 K , 学习率 α 和折扣因子 γ 。其次, 初始化 Q 网络 θ 和目标网络 θ' 。选择当前工况, 进行

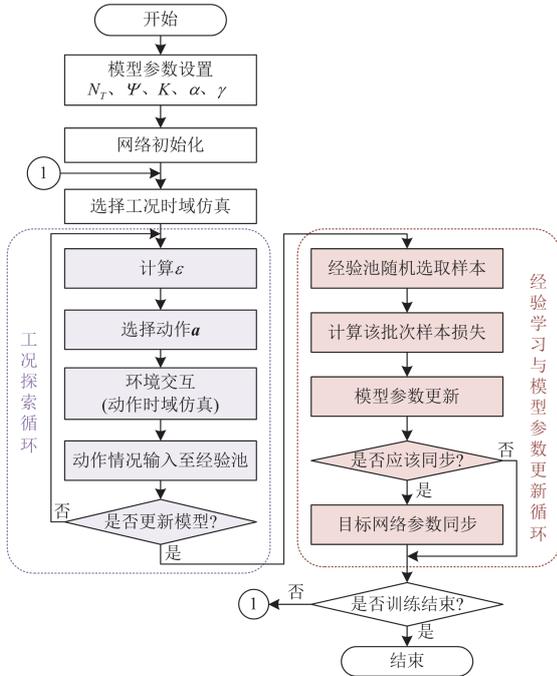


图 4 BDN-DQN 智能体训练流程

Fig. 4 Training process of BDN-DQN intelligent agents

暂态稳定仿真, 获取当前状态 s_0 , 如果完成标志 F_k^{done} 保持为 0, 则进行工况探索循环, 不断选择切负荷动作, 并和电力系统环境进行交互, 将返回结果、探索动作和完成标志 $(s_k, a_k, r_{k+1}, s_{k+1}, F_{k+1}^{\text{done}})$ 放入经验池中, 再进行下一轮循环直至完成标志置 1。之后从经验池中随机选择样本, 计算该批次样本的损失, 基于梯度反向传播对模型参数进行更新, 当更新轮数达到 K 的倍数次时, 将此时模型参数赋给目标网络。之后选择其他工况进行探索, 直到模型将所有工况都学习完毕。

5 算例验证

5.1 仿真环境与样本生成

本文提出了一种基于 GCN-BDN-DQN 的电力系统紧急切负荷决策方法。GCN 用于提取电力系统故障前后拓扑特征, BDN 用于降低巨额指数决策空间, 提升 DQN 智能体训练效率和决策性能。在 CEPRI-36 节点系统验证所提方法在训练效率和决策性能方面的优势。系统拓扑如图 5 所示。

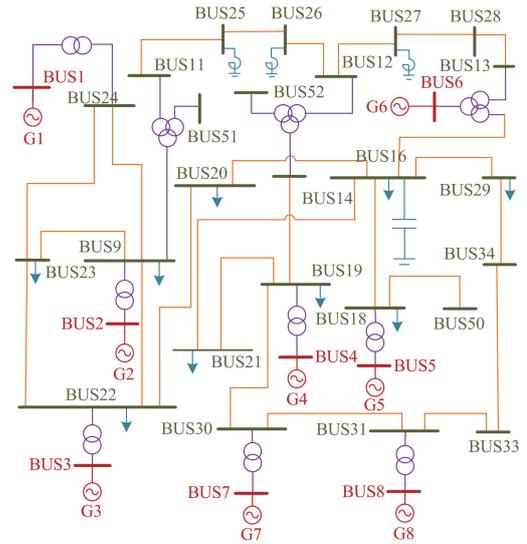


图 5 CEPRI-36 节点系统拓扑图

Fig. 5 Topology diagram of CEPRI-36 node system

算法研究和验证平台为 CPU Intel® Core(TM) i7-12700K @3.60 GHz、RAM 128 GB、GPU Nvidia Geforce RTX 3090。PSASP 用于暂态稳定仿真。Pytorch 框架用于智能体训练。训练参数为: 设置 t_0 为 0 s, t_1 为 0.1 s, t_s 为 6 s, t_f 为 10 s。最大切负荷比例设置为 0.5, 切负荷间隔设置为 $l_{\text{inte}} = 0.05$, $\varepsilon_0 = 0.9$, $\varepsilon_c = 0.05$, $n_{\text{decay}} = 10\ 000$, $r_{\Delta} = 1.5$, $r_T = 10$, $N_T = 20\ 000$, $\Psi = 128$, $K = 20$, $\alpha = 0.0001$, $\gamma = 0.2$, $T_{\text{thre}} = 6$ 。

如表 1 所示, 通过设置不同的潮流水平、负荷电动机比例、故障线路、故障位置、故障持续时间和故障类型, 依据文献[22]的 6 条专家经验, 生成了 197 个暂态电压失稳样本。其中 157 个样本用于智能体训练。40 个样本用于测试。

表 1 CEPRI-36 节点系统样本生成方案

条件	设置	种类
潮流水平	90%, 100%, 110%	3
负荷电动机比例	60%, 70%, 80%	3
故障线路	所有非变压器线路	26
故障位置	2%, 98%	2
故障持续时间	0.05 s, 0.15 s	2

5.2 基于事件的 MDP 的有效性验证

在基于事件的 MDP 的指导下, GCN-BDN-DQN 基于预想暂态电压失稳事件直接一次决策紧急切负荷控制措施, 电力系统环境直接根据紧急切负荷决策和预想暂态电压失稳事件, 调用暂态稳定仿真更新状态。如图 6 所示, 浅色为原始训练数据, 实线为经过高斯滤波平滑处理后的数据, 随着训练进行, GCN-BDN-DQN 智能体和环境交互的损失函数逐渐下降, 智能体趋于收敛。

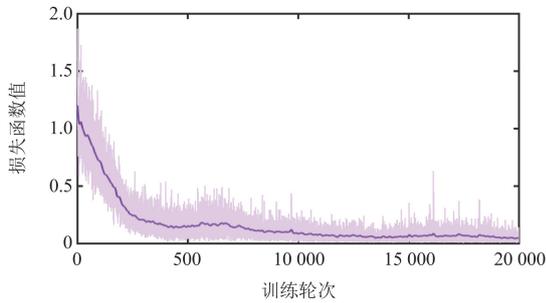


图 6 训练时 GCN-BDN-DQN 智能体损失函数变化

Fig. 6 Changes in the loss function of the GCN-BDN-DQN agent during training

进一步地, 每隔 200 次对智能体在测试集上面所有样本的奖励分布进行评估, 一共评估 10 次。奖励低于 0 表示 GCN-BDN-DQN 智能体未给出当前测试样本的有效紧急切负荷决策。单个样本的奖励上限为 10, 表示理想情况下暂态电压稳定而智能体未切除任何负荷。图 7 展示了随着训练进行 GCN-BDN-DQN 智能体奖励分布变化。可以看到, 随着训练进行, GCN-BDN-DQN 性能逐渐优化, 奖励的分布逐渐上升。在第 3 次评估时, GCN-BDN-DQN 智能体已能在所有测试样本上给出有效的紧急切负荷决策。

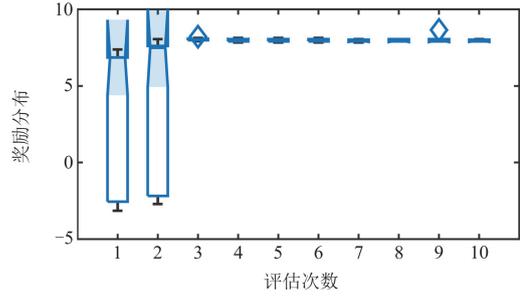


图 7 训练时 GCN-BDN-DQN 智能体奖励分布变化

Fig. 7 Changes in the reward distribution of the GCN-BDN-DQN agent during training

5.3 BDN 与图卷积增强的优势

BDN 网络中, DQN 智能体输出神经元个数随可控负荷仅呈线性增长, 且具有单步决策的能力。相比将多步决策转为单步决策的 DQN 智能体, 训练效率和决策能力均更强。基于 GCN 提取电力系统故障前后的拓扑变化特征, 能进一步增强 BDN-DQN 智能体的训练效率和决策能力。

5.3.1 不同 DRL 智能体计算复杂度分析

不同 DRL 智能体的决策空间不尽相同, 如果决策空间越小, 模型训练则更加容易。本文涉及两种不同决策空间的智能体: 指数级决策空间的 DQN 智能体^[23-24]和线性级决策空间的 BDN-DQN 智能体。

假设有 l 个负荷, 每个负荷 q 个动作, 每个输出通道对应的模型隐藏层需要 v 次计算。对于指数级决策空间的 DQN 智能体, 其输出通道具有 q^l 个, 因为其将不同负荷和不同动作的组合都视为一种输出结果, 此时一次决策需要模型计算 $v \times q^l$ 次, 模型时间复杂度为 $O(2^l)$ 。

对于本文采用的 BDN-DQN 智能体, 其输出通道仅有 $q \times l$ 个, 因为通过分支网络将单步决策问题转化为多步决策, 每轮只决策一个负荷的一个动作, 此时一次决策需要模型计算 $v \times q \times l$ 次, 模型复杂度仅为 $O(l)$ 。

5.3.2 不同 DRL 智能体在训练集上的性能变化

DRL 智能体在与环境的交互过程中, 会定期抽取一批样本池中的样本, 更新其策略网络参数, 这些被抽取到样本的总奖励称之为训练集批奖励。同时, 每轮训练时, 会对单个训练样本决策, 单个样本的奖励与迭代次数称之为训练集总奖励和训练集总迭代次数。训练集上面的批奖励上升, 总奖励上升, 总迭代次数下降说明智能体的决策能力变强, 而奖励上升越快说明智能体的训练效率越高。

如图 8 所示, 随着训练进行, 3 种智能体的批

奖励、每轮总奖励均逐步上升，每轮总迭代次数均逐步下降。相比无分支多步决策的 DQN 智能体，基于 BDN 的 DQN 智能体奖励上升更快，奖励值更大，说明 BDN 网络能极大地提升 DRL 智能体的训练效率和决策性能。相比仅提取电压时序特征的 DQN 智能体，融入了电压图特征的 DQN 智能体奖励上升更快，奖励值更大。说明 GCN 能进一步提升 BDN-DQN 智能体的训练效率和决策性能。

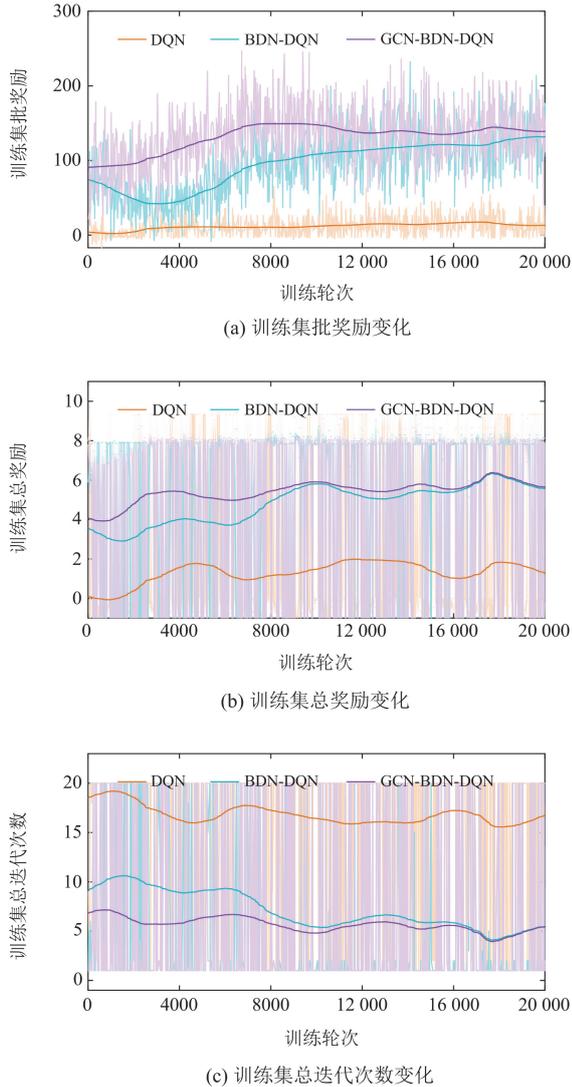


图 8 不同 DRL 智能体在训练集上的性能变化
Fig. 8 Changes in the performance of different DRL agents on the training set

5.3.3 不同 DRL 智能体在测试集上的性能变化

每隔 200 次对智能体在测试集上所有样本的总奖励和总迭代次数进行评估。如图 9 所示，随着训练进行，3 种智能体的每轮总奖励均逐步上升，每轮总迭代次数均逐步下降。相比无分支多步决策的

DQN 智能体，基于 BDN 的 DQN 智能体奖励上升更快，奖励值更大，说明 BDN 网络能极大地提升 DRL 智能体的训练效率和决策性能。相比仅提取电压时序特征的 DQN 智能体，融入了电压图特征的 DQN 智能体奖励上升更快，奖励值更大。说明 GCN 能进一步提升 BDN-DQN 智能体的训练效率和决策性能。

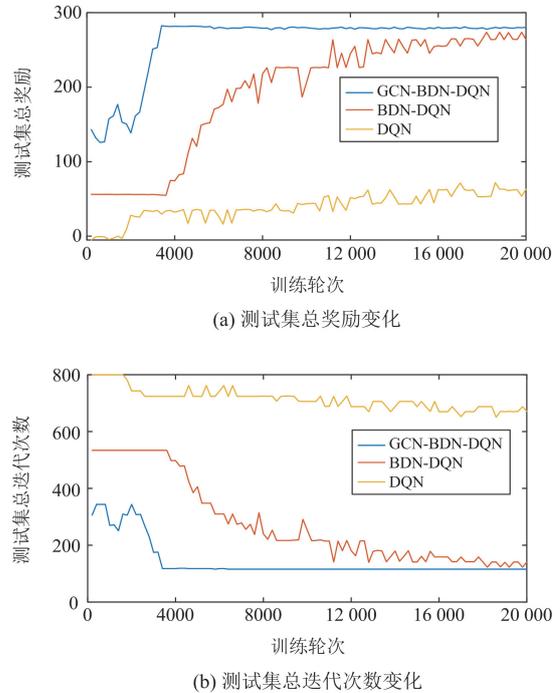


图 9 不同 DRL 智能体在测试集上的性能变化
Fig. 9 Changes in the performance of different DRL agents on the test set

5.3.4 不同 DRL 智能体实际紧急切负荷决策

以一个实际的暂态电压失稳事件中对不同 DRL 智能体的决策进行对比。暂态电压失稳事件为：第 0.1 s 时，线路 BUS12 和 BUS26 之间靠近 BUS12 端 2% 处发生三相短路故障，第 0.25 s 切除故障线路。负荷电动机比例为 80%，潮流运行于额定工况。

如图 10(a)所示，故障发生后若无紧急切负荷控制措施，母线电压持续下降，基于式(10)、式(11)计算出该工况的 T_{VSI} 指标为 33.19，表明该工况暂态电压失稳。如图 10(b)所示，单分支的 DQN 智能体决策后，母线暂态电压相比无决策有所恢复，但是 T_{VSI} 指标为 20.46，仍然不能达到暂态稳定。多分支 BDN-DQN 智能体决策后， T_{VSI} 指标降至 2.67，暂态电压恢复稳定，说明 BDN 网络能极大地提升 DRL 智能体的决策性能。多分支且融入拓扑信息的 GCN-BDN-DQN 智能体决策后， T_{VSI} 指标降至 2.65，暂态电压更加稳定，说明 GCN 能进一步提升 BDN-DQN

智能体的训练效率和决策性能。图 10(c)为 GCN-BDN-DQN 智能体决策后的暂态电压曲线。

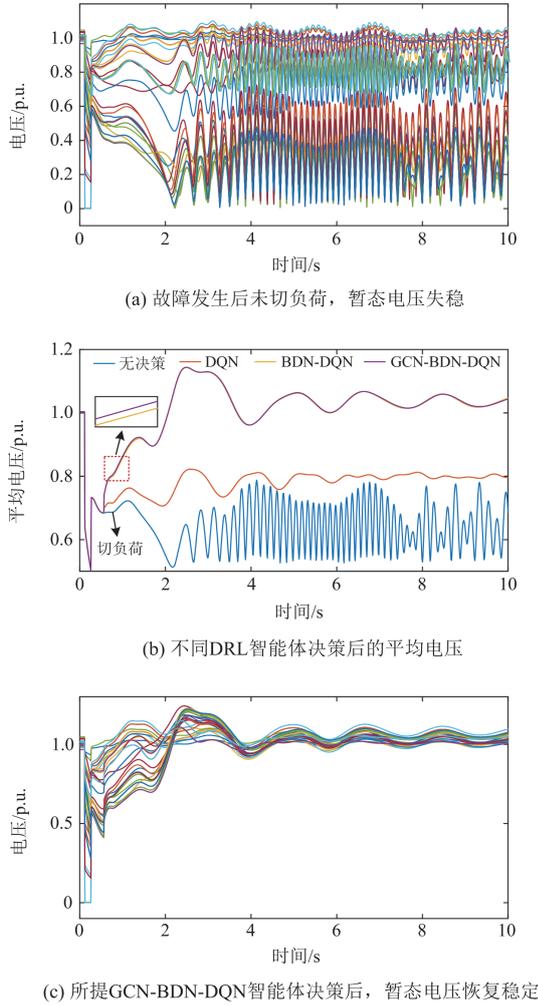


图 10 不同 DRL 智能体决策前后, 暂态电压变化情况

Fig. 10 Changes in transient voltage before and after decision-making by different DRL agents

6 结论

针对目前暂态电压失稳紧急切负荷决策措施制定严重依赖专家经验且耗时耗力, 而现有基于 DRL 的智能算法决策效率较低的问题, 本文提出了一种基于 GCN-BDN-DQN 智能体的紧急切负荷框架。首先, 基于暂态稳定仿真建立了智能体和电力系统环境基于事件的 MDP, 指导 GCN-BDN-DQN 智能体的训练。其次, 设计了一种 BDN 网络, 输出神经元由指数数量级降为线性数量级, 且智能体能够进行单步决策。最后, 利用 GCN 将故障前后拓扑特征变化融入智能体训练, 有效提升状态变量信息量。

在 CEPRI-36 节点系统进行了仿真验证, 实验结果表明, 所提 MDP 能有效指导智能体的训练。

相比无分支网络, BDN 网络能提升训练效率和决策能力。相比仅提取电压时序信息, GCN 提取拓扑信息能进一步提升训练效率和决策能力。

参考文献

- [1] 程静, 苏乐, 岳雷. 双馈风电并网系统的宽频振荡机理分析与抑制[J]. 电力系统保护与控制, 2023, 51(12): 1-13.
CHENG Jing, SU Le, YUE Lei. Mechanism analysis and suppression of broadband oscillation in a DFIG wind power grid-connected system[J]. Power System Protection and Control, 2023, 51(12): 1-13.
- [2] 中国南方电网公司. 交直流电力系统仿真技术[M]. 北京: 中国电力出版社, 2007.
- [3] 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.
HAN Lu, YIN Chunya, DAI Chen, et al. Transient voltage operational risk of a high-proportion new energy sending system[J]. Power System Protection and Control, 2024, 52(1): 23-34.
- [4] 林湘宁, 卓毅鑫, 李正天, 等. 基于风电场动态电压安全决策树体系的 SVC 紧急控制策略[J]. 中国电机工程学报, 2018, 38(1): 41-50, 341.
LIN Xiangning, ZHUO Yixin, LI Zhengtian, et al. Wind farm SVC emergency control strategy based on dynamic voltage security decision trees[J]. Proceedings of the CSEE, 2018, 38(1): 41-50, 341.
- [5] XU Y, YANG Z, ZHANG J, et al. Real-time compressive sensing based control strategy for a multi-area power system[J]. IEEE Transactions on Smart Grid, 2017, 9(5): 4293-4302.
- [6] KUNDU P, PRADHAN A K. Enhanced protection security using the system integrity protection scheme (SIPS)[J]. IEEE Transactions on Power Delivery, 2015, 31(1): 228-235.
- [7] 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.
- [8] 刘水平, 刘明波, 谢敏. 基于轨迹灵敏度的电力系统模型预测切负荷控制方法[J]. 中国电机工程学报, 2010, 30(34): 28-35.
LIU Shuiping, LIU Mingbo, XIE Min. Model predictive control method for load shedding of power systems based on trajectory sensitivities[J]. Proceedings of the CSEE, 2010, 30(34): 28-35.
- [9] 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 Shaoxuan, 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.
- [10] 陈学通, 凌超, 薛峰, 等. 一种基于贪心算法的紧急控制策略优化搜索方法[J]. 电力系统保护与控制, 2017, 45(23): 74-81.
CHEN Xuetong, LING Chao, XUE Feng, et al. A fast method of searching optimal emergency control strategy based on greedy algorithm[J]. Power System Protection and Control, 2017, 45(23): 74-81.
- [11] SHI Z, YAO W, LI Z, et al. Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions[J]. Applied Energy, 2020, 278.
- [12] 程雄, 冯佳, 吕欣, 等. 兼顾激励相容的水电富集电网中长期电力市场交易模型[J]. 电力系统保护与控制, 2024, 52(3): 12-24.
CHENG Xiong, FENG Jia, LÜ Xin, et al. Medium and long-term electricity market trading model considering incentive compatibility in a hydropower-dominated grid[J]. Power System Protection and Control, 2024, 52(3): 12-24.
- [13] 尚立, 蔡硕, 崔俊彬, 等. 基于软件定义网络的电网边缘计算资源分配[J]. 电力系统保护与控制, 2021, 49(20): 136-143.
SHANG Li, CAI Shuo, CUI Junbin, et al. SDN-based MEC resource allocation of a power grid[J]. Power System Protection and Control, 2021, 49(20): 136-143.
- [14] 余涛, 张水平. 在策略 SARSA 算法在互联电网 CPS 最优控制中的应用[J]. 电力系统保护与控制, 2013, 41(1): 211-216.
YU Tao, ZHANG Shuiping. Optimal CPS control for interconnected power systems based on SARSA on-policy learning algorithm[J]. Power System Protection and Control, 2013, 41(1): 211-216.
- [15] LIU J, ZHANG Y, MENG K, et al. Real-time emergency load shedding for power system transient stability control: a risk-averse deep learning method[J]. Applied Energy, 2022, 307.
- [16] HUANG Q, HUANG R, HAO W, et al. Adaptive power system emergency control using deep reinforcement learning[J]. IEEE Transactions on Smart Grid, 2019, 11(2): 1171-1182.
- [17] 胡泽, 曾令康, 姚伟, 等. 电力系统两阶段紧急切负荷控制智能预决策[J]. 中国电机工程学报, 2024, 44(4): 1260-1272.
HU Ze, ZENG Lingkan, YAO Wei, et al. Intelligent pre-decision of two-stage emergency load shedding control in power system[J]. Proceedings of the CSEE, 2024, 44(4): 1260-1272.
- [18] WANG S, DUAN J, SHI D, et al. A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning[J]. IEEE Transactions on Power Systems, 2020, 35(6): 4644-4654.
- [19] 朱林, 张健, 陈达, 等. 面向暂态电压稳定评估的卷积神经网络输入特征构建方法[J]. 电力系统自动化, 2022, 46(1): 85-93.
ZHU Lin, ZHANG Jian, CHEN Da, et al. Construction method for input features of convolutional neural network for transient voltage stability assessment[J]. Automation of Electric Power Systems, 2022, 46(1): 85-93.
- [20] LI Q, XU Y, REN C. A hierarchical data-driven method for event-based load shedding against fault-induced delayed voltage recovery in power systems[J]. IEEE Transactions on Industrial Informatics, 2020, 17(1): 699-709.
- [21] 刘雨晴, 刘翌, 王小君, 等. 融合同步知识和时空信息的电力系统暂态稳定评估框架[J/OL]. 电网技术: 1-13[2024-01-30]. <https://doi.org/10.13335/j.1000-3673.pst.2023.1401>.
LIU Yuqing, LIU Zhao, WANG Xiaojun, et al. Power system transient stability assessment framework based on fusion of synchronization knowledge and spatial-temporal information[J/OL]. Power System Technology: 1-13[2024-01-30]. <https://doi.org/10.13335/j.1000-3673.pst.2023.1401>.
- [22] 石重托, 姚伟, 黄彦浩, 等. 基于 SE-CNN 和仿真数据的电力系统主导失稳模式智能识别[J]. 中国电机工程学报, 2022, 42(21): 7719-7731.
SHI Zhongtuo, YAO Wei, HUANG Yanhao, et al. Power system dominant instability mode identification based on convolutional neural networks with squeeze and excitation block and simulation data[J]. Proceedings of the CSEE, 2022, 42(21): 7719-7731.
- [23] HU Z, SHI Z, ZENG L, et al. Knowledge-enhanced deep reinforcement learning for intelligent event-based load shedding[J]. International Journal of Electrical Power & Energy Systems, 2023, 148.
- [24] 孙立钧, 顾雪平, 刘彤, 等. 一种基于深度强化学习算法的电网有功安全校正方法[J]. 电力系统保护与控制, 2022, 50(10): 114-122.
SUN Lijun, GU Xueping, LIU Tong, et al. A deep reinforcement learning algorithm-based active safety correction method for power grids[J]. Power System Protection and Control, 2022, 50(10): 114-122.

收稿日期: 2024-04-25; 修回日期: 2024-05-28

作者简介:

潘晓杰(1976—), 男, 博士, 教授级高工, 研究方向为电力系统稳定与控制、电力系统仿真技术; E-mail: pxj76@163.com

胡泽(2000—), 男, 硕士研究生, 研究方向为电力系统暂态电压失稳紧急切负荷控制; E-mail: huze@hust.edu.cn

姚伟(1983—), 男, 博士, 教授, 博士生导师, 研究方向为高比例新能源电网稳定分析与控制、新一代电力能源人工智能技术及应用、交直流混联电网稳定分析与控制; E-mail: w.yao@hust.edu.cn

兰宇田(2001—), 男, 通信作者, 硕士研究生, 研究方向为人工智能在电力系统中的应用。E-mail: lan_yutian2001@163.com

(编辑 魏小丽)