

DOI: 10.19783/j.cnki.pspc.230160

# 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

高发骏<sup>1</sup>, 王怀远<sup>1</sup>, 党然<sup>2</sup>

(1. 新能源发电与电能变换重点实验室(福州大学), 福建 福州 350108; 2. 陕西飞机工业有限责任公司, 陕西 汉中 723000)

**摘要:** 深度学习算法在电力系统暂态稳定性评估问题上有着优秀的表现, 然而模型评估结果的不可知性与决策过程的不可控性阻碍了其在实际中进一步的应用。构建了基于 Transformer 编码器的暂态稳定评估模型, 尝试通过模型对于特征量的注意力权重, 解释和分析模型所关注和学习到的规则。在此基础上, 结合可解释性结果提出了一种利用物理信息指导模型优化的模型更新方法。从模型的损失函数出发, 通过微调的方式修正模型对特征量的注意力权重分布, 加强对于样本失稳模式的挖掘。在微调模型的过程中, 引入注意力引导函数提高对特定失稳模式关键机组的关注度, 以减少对于特定失稳模式样本的误判, 进一步提高整体的预测精度。在 IEEE39 节点系统和华东电网系统的仿真均验证了该方法的有效性。

**关键词:** Transformer; 暂态稳定性评估; 可解释性; 注意机制; 损失函数

## Interpretability analysis and model update research of a transient stability assessment model based on Transformer

GAO Fajun<sup>1</sup>, WANG Huaiyuan<sup>1</sup>, DANG Ran<sup>2</sup>

(1. Key Laboratory of New Energy Generation and Power Conversion (Fuzhou University), Fuzhou 350108, China;  
2. Shaanxi Aircraft Industry Limited Liability Company, Hanzhong 723000, China)

**Abstract:** Deep learning algorithms have excellent performance in power system transient stability assessment, but the incomprehensibility of the assessment results and the uncontrollability of the decision-making process hinder their practical adoption by industry. A transient stability assessment model based on the Transformer encoder is proposed. The rules that the model focuses on and learns can be interpreted and analyzed by the attention weights of the features. Thus a model updating method is proposed which employs physical information in combination with interpretable results to guide model optimization. From the perspective of the loss function, the attention weight distribution of the model to the features is adjusted in a fine-tuned way to enhance the mining for the instability patterns. In the process of fine-tuning, an attention-guiding function is introduced to increase the attention weights to the key generators of specific instability patterns, so as to reduce the misclassification of specific instability patterns. In this way the overall prediction accuracy can be improved. The performance of the proposed method is verified on the IEEE39-bus system and the East China power grid system.

This work is supported by the Natural Science Foundation of Fujian Province (No. 2022J01113).

**Key words:** Transformer; transient stability assessment; interpretability; attention mechanism; loss function

## 0 引言

近年来交直流混联电网的形成<sup>[1]</sup>、大规模不确定性能源的接入<sup>[2-3]</sup>, 给电力系统的安全带来了巨大

挑战, 这对电力系统暂态稳定评估提出了更高的要求。传统的暂态稳定评估方法包括时域仿真法<sup>[4-5]</sup>和直接法<sup>[6-7]</sup>, 但由于其快速性的限制, 无法应用于电力系统在线暂态稳定性评估。

机器学习算法具有很强的数据处理能力, 其计算速度快、精度高, 近年来在电力系统在线暂态稳

基金项目: 福建省自然科学基金项目资助(2022J01113)

定性评估中得到了快速发展<sup>[8-10]</sup>。常用的机器学习算法包括人工神经网络<sup>[11]</sup>、支持向量机(support vector machine, SVM)<sup>[12]</sup>、决策树(decision tree, DT)<sup>[13-14]</sup>等。深度学习作为机器学习的一个重要分支,与传统的机器学习方法相比,具有更加优秀的特征提取能力和更高的准确度。目前已有许多深度学习模型被应用于暂态稳定评估,其中包括深度置信网络<sup>[15-16]</sup>、堆叠自动编码器<sup>[17]</sup>、卷积神经网络(convolutional neural network, CNN)<sup>[18]</sup>、长短期记忆网络(long short-term memory, LSTM)<sup>[19-20]</sup>等。同时许多学者根据暂态稳定评估中面临的各种问题进行了相关的研究。文献[21]提出基于一维 CNN 的暂态稳定评估方法,以解决传统的机器学习方法难以捕捉时序信息的问题。文献[22]提出了自适应代价敏感赋值方法,以解决不稳定样本误分类的代价比稳定样本更高的问题,并且能够消除异常样本对评估规则的影响。文献[23]提出基于迁移学习的暂态稳定评估方法,其能自适应跟踪系统的运行方式和拓扑结构的变化,有效地更新评估模型。

然而,目前阻碍深度学习方法在电力系统中进一步被应用的原因有:1)对评估结果的不可知性<sup>[24]</sup>;2)对于决策过程的不可控性。针对评估结果的不可知性,已有学者进行了相关的研究。文献[14, 25]使用基于 DT 的“白盒”模型解释和分析模型的决策过程。然而 DT 是一种浅层学习模型,结构简单,准确度难以满足需求。文献[26]通过与模型无关的局部可解释性方法(local interpretable model-agnostic explanations, LIME)来解释基于 CNN 模型的预测。文献[27]使用基于实例的机器学习模型解释工具(shapley additive explanation, SHAP)挖掘了影响评估结果的关键因素。但 LIME 只能解释单个预测,而 SHAP 的计算时间过长,并且这两种方法生成的解释结果都不够稳定<sup>[28]</sup>。文献[29]提出的 Transformer 模型为可解释性分析提供了新思路。该模型仅依靠注意力机制来计算输入与输出的关系,能动态地增强高信息价值特征的表达并抑制冗余特征。同时随着训练的过程生成较为稳定的特征注意力分布,直观地展示特征量对暂态稳定评估结果的重要程度<sup>[30-31]</sup>,以达到准确性和透明性的平衡。

目前的研究主要集中在分析评估结果的不可知性,而决策过程的不可控性还鲜有研究关注。在实际应用中,评估结果的可知性与决策过程的可控性对于提高评估模型的可靠性和可信度都至关重要。传统的深度学习模型的训练方法往往无法控制模型的决策过程,导致某些样本的评估规则不合理。当可解释结果表明模型决策过程不合理时,目前还

缺少一种合适方式修正模型的评估规则。因此,本文聚焦于此,将物理信息与可解释性结果相结合以修正模型的决策过程,引导模型学习正确的物理知识并作出正确评估。

本文提出了基于 Transformer 编码器的暂态稳定评估方法。通过可视化特征量的注意力权重分布,解释模型学习到的评估规则。在此基础上,本文主要针对模型决策过程的可控性问题,提出了一种利用物理信息指导模型优化的方法。通过注意力引导函数引导模型更加聚焦于关键失稳机组,以便充分学习到样本中的稳定性相关信息,正确捕捉到样本的主导失稳模式,提高模型的合理性和准确性。

## 1 模型原理

### 1.1 Transformer 原理

标准的 Transformer 模型<sup>[29]</sup>主要分为编码器和解码器。针对暂态稳定评估问题,本文采用 Transformer 编码器作为评估模型的核心,具体结构如图 1 所示,Transformer 编码器由位置编码和多个相同结构的子模块组成。子模块中包括多头注意力层、前馈网络、残差连接和层标准化。位置编码可以帮助模型识别特征向量中的位置关系。特征矩阵经过多头注意力层和前馈网络层后均会采用残差连接和层标准化进行处理。残差连接有效地解决了梯度消失和网络退化的问题。层标准化的作用是加快训练速度和模型收敛。

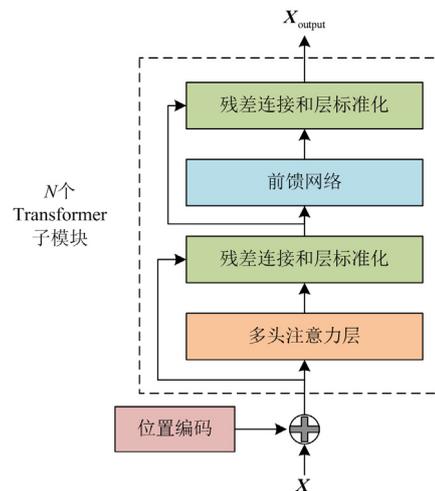


图 1 Transformer 编码器的结构

Fig. 1 Structure of Transformer encoder

多头注意力机制是 Transformer 编码器的核心,将模型分为多个头,形成多个子空间,可以让模型去关注不同方面的信息,最后将各个方面的信息综合起来,这样有助于网络捕捉到更丰富的特征信息,

计算公式如式(1)—式(4)所示。

$$\begin{cases} \mathbf{Q} = \mathbf{W}^Q \mathbf{X} \\ \mathbf{K} = \mathbf{W}^K \mathbf{X} \\ \mathbf{V} = \mathbf{W}^V \mathbf{X} \end{cases} \quad (1)$$

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V} \quad (2)$$

$$\mathbf{H}_i = \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \quad (3)$$

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = [\mathbf{H}_1, \dots, \mathbf{H}_h]\mathbf{W}^O \quad (4)$$

式中: 查询矩阵  $\mathbf{Q}$ 、键矩阵  $\mathbf{K}$ 、值矩阵  $\mathbf{V}$  由特征向量矩阵  $\mathbf{X}$  变换生成;  $\mathbf{W}^Q$ 、 $\mathbf{W}^K$  和  $\mathbf{W}^V$  均为线性变换矩阵;  $d$  为缩放因子;  $\mathbf{W}_i^Q$ 、 $\mathbf{W}_i^K$  和  $\mathbf{W}_i^V$  分别是将  $\mathbf{Q}$ 、 $\mathbf{K}$ 、 $\mathbf{V}$  投影到第  $i$  个子空间的转换矩阵,  $i=1,2,\dots,h$ ,  $h$  为子空间总数;  $\mathbf{H}_i$  为第  $i$  个子空间的单头注意力值;  $\mathbf{W}^O$  为转换矩阵, 用于连接所有子空间的注意力值。

## 1.2 注意力权重的计算

故障导致系统暂态失稳最常见的方式是两群失稳, 而后极少情况再发展为多群失稳。评估模型准确地评估出样本的稳定性, 就需要模型能够捕捉到样本的失稳模式。注意力权重可以在一定程度上反映模型对样本失稳模式的挖掘及其决策规则。因此, 通过对模型注意力矩阵分析, 运行人员可以结合运行经验判断决策是否合理可靠。

模型的输入数据选用发电机的特征。原始数据集包含不同机组在各个时间节点下的物理特征信息, 每个特征样本可表示为式(5)。

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{G1} \\ \mathbf{x}_{G2} \\ \vdots \\ \mathbf{x}_{GU} \end{bmatrix} = \begin{bmatrix} x_{G1,1} & x_{G1,2} & \cdots & x_{G1,T} \\ x_{G2,1} & x_{G2,2} & \cdots & x_{G2,T} \\ \vdots & \vdots & & \vdots \\ x_{GU,1} & x_{GU,2} & \cdots & x_{GU,T} \end{bmatrix} \quad (5)$$

式中:  $U$  为机组总数;  $T$  为总时间节点; 每一行表示一台机组的所有信息, 即  $\mathbf{x}_{Gu}$  代表第  $u$  台机组的全部物理特征信息,  $u=1,2,\dots,U$ ;  $x_{Gu,t}$  表示在第  $t$  个时间节点下第  $u$  台机组的特征数据,  $t=1,2,\dots,T$ 。

单头的注意力权重计算如式(6)所示。

$$A(\mathbf{Q}, \mathbf{K}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right) \quad (6)$$

为了考虑特征量对于整体的重要程度, 特征量所对应的注意力权重为所有子空间计算出的注意力权重的平均值<sup>[32]</sup>, 如式(7)所示。

$$\tilde{A}(\mathbf{Q}, \mathbf{K}) = \frac{1}{h} \sum_{i=1}^h A(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K) \quad (7)$$

根据式(7)可计算出特征样本对应的注意力权

重矩阵  $\mathbf{A}_{aw} \in \mathbb{R}^{N \times U}$ , 如式(8)所示。

$$\mathbf{A}_{aw} = \begin{bmatrix} a_{1,G1} & a_{1,G2} & \cdots & a_{1,GU} \\ a_{2,G1} & a_{2,G2} & \cdots & a_{2,GU} \\ \vdots & \vdots & & \vdots \\ a_{N,G1} & a_{N,G2} & \cdots & a_{N,GU} \end{bmatrix} \quad (8)$$

式中:  $N$  为样本总数;  $a_{k,Gu}$  代表第  $k$  个样本中第  $u$  台机组所对应的注意力权重,  $k=1,2,\dots,N$ 。每行表示一个样本中所有机组的注意力权重集合, 并且每行所有元素之和为 1, 可表示为

$$\sum_{u=1}^U a_{k,Gu} = 1 \quad (9)$$

注意力权重分布能直接反映特征对于暂态稳定评估结果的重要性。注意力权重的值越高, 则代表模型对这个特征量的关注度越大, 即这个特征量对模型的决策也越重要<sup>[30]</sup>。因此可以通过分析特征量的注意力权重得知模型的决策依据。

## 2 注意力权重的修正

对于失稳样本来说, 失稳机组的数据有更高的价值, 如果模型能够捕捉到失稳机组的数据变化, 正确识别其失稳模式, 那么就能更准确地评估。通过大量的仿真实验发现, 评估模型错误评估失稳样本的主要原因在于评估模型无法正确识别样本的主导模式, 对失稳机组关注度较小, 无法充分挖掘样本中的稳定性相关信息。如果能够根据这些误判样本的失稳模式对应调整其特征量的注意力权重分布, 那么就能进一步提升模型的预测精度。

因此, 本文尝试从损失函数出发, 在整体的损失函数中额外添加一项注意力引导函数作为附加项。注意力引导损失能将特定失稳模式样本中关键失稳机组的权重映射到损失函数中。在更新的过程中, 随着注意力引导函数项的降低, 关键失稳机组被分配到的权重就会增加。因此模型能够提高对于关键机组的关注度, 产生更合理的注意力权重分布, 从而更加准确地评估具有这些失稳模式的样本。

定义矩阵  $\mathbf{S} \in \mathbb{R}^{1 \times N}$ , 如式(10)所示, 其取值方法可表示为式(11)。

$$\mathbf{S} = [s_1, s_2, \dots, s_k, \dots, s_N] \quad (10)$$

$$s_k = \begin{cases} 1, & \text{特定样本} \\ 0, & \text{其他样本} \end{cases} \quad (11)$$

如果第  $k$  个样本为选定的失稳模式, 对应的  $s_k$  取值为 1, 否则取值为 0。 $\mathbf{S}$  的作用在于选择训练过程中所关注的失稳样本, 并减少注意力引导函数对于评估其他样本的影响。

接着定义一个矩阵  $\mathbf{g} \in \mathbb{R}^{N \times U}$ ，如式(12)所示。

$$\mathbf{g} = \begin{bmatrix} g_{1,G1} & g_{1,G2} & \cdots & g_{1,GU} \\ g_{2,G1} & g_{2,G2} & \cdots & g_{2,GU} \\ \vdots & \vdots & & \vdots \\ g_{N,G1} & g_{N,G2} & \cdots & g_{N,GU} \end{bmatrix} \quad (12)$$

式中， $g_{k,Gu}$  代表第  $k$  个样本中第  $u$  台机组所对应的值。若第  $k$  个样本的失稳方式为所关注的失稳模式，且第  $u$  台机组属于关键机组，则  $g_{k,Gu}$  为 1，其余情况取值为 0。对于第  $k$  个样本的  $\mathbf{g}_k$  可表示为

$$\mathbf{g}_k = [g_{k,G1}, g_{k,G2}, \cdots, g_{k,GU}]^T \quad (13)$$

$\mathbf{g}_k$  的作用是在不同的样本中标记需要修正注意力权重的机组。若  $g_{k,Gu}$  的取值为 1，则代表第  $k$  个样本的第  $u$  台机组需要调整其注意力；若  $g_{k,Gu}$  的取值为 0，则代表第  $k$  个样本的第  $u$  台机组无须调整其注意力。

根据式(8)定义第  $k$  个样本的注意力权重矩阵  $\mathbf{A}_{awk}$ ，如式(14)所示。

$$\mathbf{A}_{awk} = [a_{k,G1}, a_{k,G2}, \cdots, a_{k,GU}] \quad (14)$$

第  $k$  个样本的注意力引导函数项如式(15)所示。

$$\Psi_k(s_k, \mathbf{A}_{awk}, \mathbf{g}_k) = s_k e^{(-\alpha \times \mathbf{A}_{awk} \times \mathbf{g}_k)} \quad (15)$$

式中， $\alpha$  为放缩因子，把  $\mathbf{A}_{awk}$  中的数值扩大  $\alpha$  倍，是为了将所需要增加的注意力权重值映射到更大空间中，使注意力引导函数能够对注意力权重的变化更加敏感，以获得更好的修正效果。

整体的损失函数定义如式(16)所示。

$$\sum_{k=1}^N L'_k = L_k(y_k, \hat{y}_k) + \lambda \Psi_k(s_k, \mathbf{A}_{awk}, \mathbf{g}_k) \quad (16)$$

式中， $\lambda$  为引导系数，相当于拉格朗日乘子，使注意力引导函数项在损失函数中作为约束条件，起到

弱约束的作用。其取值是根据注意力引导函数项与交叉熵函数项的数值而定，使得两者的数值能处于同一数量级，保证在模型的微调过程中注意力引导函数能够起到一定的引导作用，增加对于失稳机组的关注。在实验过程中发现，引导系数  $\lambda$  无需太多的优化和整定，根据最后一次优化结果的损失函数值与注意力引导函数值的数量级进行赋值，保证其值在一个数量级下能获得较好的优化效果。本文的引导系数取值为 0.4。 $L_k$  为交叉熵函数，用来衡量模型的分类误差， $L_k(y_k, \hat{y}_k)$  定义为

$$L_k(y_k, \hat{y}_k) = -y_k \log \hat{y}_k - (1 - y_k) \log(1 - \hat{y}_k) \quad (17)$$

式中： $y_k$  代表第  $k$  个样本的真实标签； $\hat{y}_k$  表示第  $k$  个样本的预测标签。

注意力引导函数将用于指导模型的训练过程，若模型能够在训练过程中捕捉到正确的失稳模式，就能够将原先误判的样本评估正确。在训练的过程中，注意力引导函数将特定失稳模式下关键机组的注意力权重引入至损失函数中。随着损失函数的降低能提高模型对于失稳机组的关注度、加强评估模型挖掘特定样本的主导失稳模式的能力，使得模型能够更加准确地评估这些特定失稳模式的样本，进而提高模型的预测精度。为了降低更新计算成本，可以在已经训练好的模型基础上进行微调，即对原有的模型通过修正损失函数进行再训练。更新模型的过程为离线执行，并不会增加在线评估的时间，流程如图 2 所示。

在实际运用的过程中，某些故障场景频繁发生，需要重点监测。然而模型对于这些易发生的故障场景的评估性能不佳。此时就可以通过本文方法对模型进行更新以保证电力系统的安全运行。或是某些

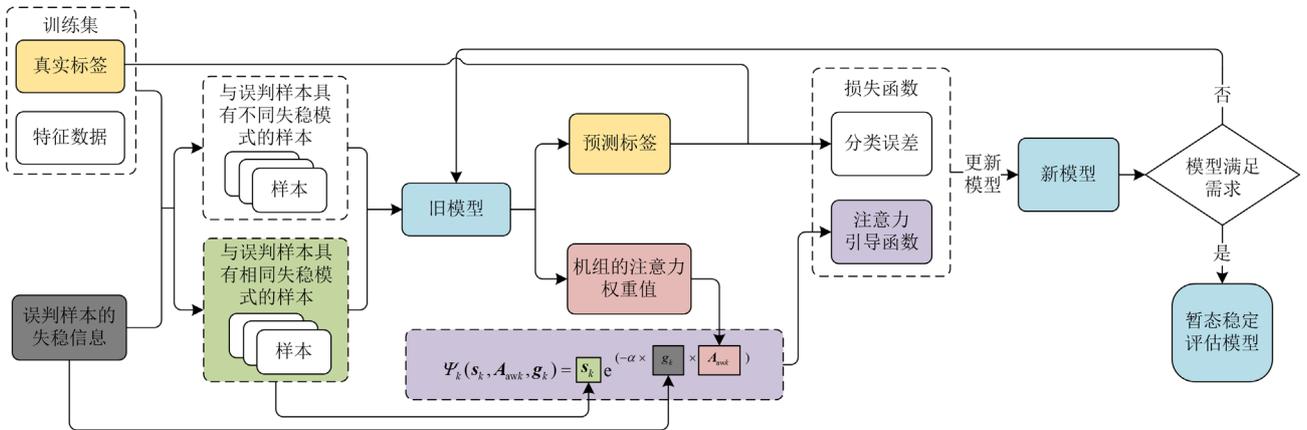


图 2 微调模型的流程

Fig. 2 Process of fine-tuning

故障场景会对电力系统造成严重的危害, 因此要求模型对于这些故障场景有更高的评估精度, 此时也可以采用本文提出的更新方法对模型进行进一步的优化, 以便及时采取合适的控制策略。

### 3 基于 Transformer 的暂态稳定评估模型

#### 3.1 评估模型设计

将 Transformer 编码器与全连接层和逻辑分类器相结合, 以建立暂态稳定评估模型。

特征量输入 Transformer 编码器后得到  $\mathbf{x}_{\text{output}}$ 。

$\mathbf{x}_{\text{output}}$  作为输入通过全连接层并采用激活函数 Relu 激活, 如式(18)所示。

$$\mathbf{x}_z = \max(0, \mathbf{W}_F \mathbf{x}_{\text{output}} + \mathbf{b}_F) \quad (18)$$

式中:  $\mathbf{x}_z$  为全连接层的输出;  $\mathbf{W}_F$  为是全连接层的权重矩阵;  $\mathbf{b}_F$  为偏置矩阵。

全连接层的输出将输入到逻辑分类器中, 如式(19)所示。

$$\hat{y} = \frac{1}{1 + e^{-\mathbf{W}_L \mathbf{x}_z + \mathbf{b}_L}} \quad (19)$$

式中:  $\mathbf{W}_L$  为逻辑分类器的权重矩阵;  $\mathbf{b}_L$  为偏置矩阵;  $\hat{y}$  为模型的评估结果。

#### 3.2 稳定性指标

通过时域仿真法生成原始数据集后, 每个样本的真实标签用式(20)<sup>[22]</sup>标记。

$$\gamma = 360^\circ - |\Delta\delta_{\max}| \quad (20)$$

式中,  $\Delta\delta_{\max}$  为仿真时刻任意两个发电机之间功角差的最大值。如果  $\gamma > 0$ , 则该样本为稳定样本, 其标签为 1, 否则为不稳定样本, 其标签为 0。

#### 3.3 输入特征选取

电力系统暂态稳定评估模型的准确性与输入特征量的选择有一定的关系。本文选取故障清除后的发电机功角和角速度数据作为模型的输入特征量。

发电机转子运动方程如式(21)所示。根据转子运动方程可知, 发电机的功角数据最直接地反映系统的稳定性。而角速度为功角的一阶导数, 可以反映功角的变化趋势。文献[33]选择发电机的功角和角速度数据作为模型的输入, 并取得了良好的评估效果。因此本文选择发电机的功角和角速度数据来训练评估模型。

$$\begin{cases} T_J \times \frac{d^2\delta}{dt^2} = M_T - M_E \\ \frac{d\delta}{dt} = \omega - \omega_0 \end{cases} \quad (21)$$

式中:  $T_J$  为发电机的惯性时间常数;  $\omega_0$  为系统的同步转速;  $\delta$  为发电机的功角;  $\omega$  为发电机的转子角速度;  $M_T$  为发电机的原动机转矩;  $M_E$  为发电机的电磁转矩。

同步相量测量单元(phasor measurement unit, PMU)采集到的数据从时间上可分为故障发生前、故障清除时刻和故障清除后。故障前的数据属于稳态信息, 而故障期间的 PMU 测量容易受到干扰, 导致数据不可靠, 因此本文选择故障清除后的数据。文献[21]也表明故障清除后的数据对于暂态稳定评估问题具有更高的价值。

#### 3.4 数据处理

在数据送入评估模型之前需要用 Z-Score 进行预处理。将不同量级的数据统一转化为同一个量级, 以消除单位的影响和保证数据之间的可比性, 如式(22)所示。

$$x' = \frac{x - \bar{X}}{\sigma_x} \quad (22)$$

式中:  $x'$  为 Z-Score 标准化后的数据;  $x$  为原始特征量的值;  $\bar{X}$  为原始特征量的均值;  $\sigma_x$  为原始特征量的标准差。

#### 3.5 模型性能的评价指标

仅使用准确率难以全面评估模型性能, 本文将采用准确率  $P_{\text{ACC}}$ 、漏报率  $P_{\text{FD}}$ 、误警率  $P_{\text{FA}}$  作为衡量模型性能的评价指标, 各指标的定义如式(23)—式(25)<sup>[21]</sup>所示。  $T_P$ 、 $T_N$ 、 $F_P$ 、 $F_N$  的具体含义如表 1 所示。

$$P_{\text{ACC}} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \times 100\% \quad (23)$$

$$P_{\text{FD}} = \frac{F_P}{F_P + T_N} \times 100\% \quad (24)$$

$$P_{\text{FA}} = \frac{F_N}{F_N + T_P} \times 100\% \quad (25)$$

表 1 暂态稳定评估结果

Table 1 Results of transient stability assessment

| 真实标签 | 预测标签  |       |
|------|-------|-------|
|      | 稳定    | 不稳定   |
| 稳定   | $T_P$ | $F_N$ |
| 不稳定  | $F_P$ | $T_N$ |

## 4 算例分析

本文采用 PSD-BPA 仿真软件生成原始数据集, 测试系统为 IEEE39 节点系统<sup>[15]</sup>和 华东电网系统。

### 4.1 IEEE39 节点系统

#### 4.1.1 样本集的构造

IEEE39 节点系统仿真细节如表 2 所示,共生成 15 362 个样本,其中稳定样本有 9296 个,不稳定样本有 6066 个。选择故障清除时刻和故障清除后 4 个周波内的所有发电机的功角与角速度偏差作为模型的输入特征。训练集和测试集的比例为 1:1,其中不稳定样本均为 3033 个,稳定样本均为 4648 个。

表 2 IEEE39 节点系统的仿真细节

Table 2 Simulation details based on the IEEE39-bus system

| 仿真参数     | 参数细节                |
|----------|---------------------|
| 故障线路     | 33 条输电线路            |
| 故障类型     | 永久性三相短路             |
| 故障位置     | 每条线路的 10%、50%和 90%处 |
| 负荷水平     | 80%~115%            |
| 故障持续时间   | 6~18 个周波            |
| PMU 采样频率 | 120 次/s             |
| 仿真时长     | 4 s                 |

#### 4.1.2 不同算法的比较

为了测试本文暂态稳定评估模型的性能,与常用的机器学习方法 DT、SVM、CNN、LSTM 和融合注意力机制(attention mechanism)的 LSTM 模型(ALSTM)<sup>[34]</sup>进行了比较。

Transformer 中的  $h$  为 8,子模块数为 2。LSTM 的隐藏层数为 2,隐藏层的维度为 60。ALSTM 的结构可参考文献[34],在 LSTM 模型的基础上融合了注意力机制,通过注意力机制处理 LSTM 模型的隐藏层输出以获取更好的评估性能。CNN 模型有 3 层卷积层,每层卷积层后连接  $2 \times 2$  的最大池化层,最后连接 2 个全连接层,激活函数采用 Relu。采用分类回归树(classification and regression tree, CART)算法搭建基于 DT 的评估模型。SVM 采用径向基函数,超参数  $C=10$ ,  $\gamma=0.1$ 。

为了保证公平性,所有试验结果均为 10 次测试后的平均性能,比较结果如表 3 所示。Transformer 的优势在于利用多头注意力机制来捕获全局信息,从而提取更关键的特征,具有更强的数据挖掘能力。从表 3 可以看出,基于 Transformer 的暂态稳定评估模型与其他的评估模型相比,在准确率  $P_{ACC}$ 、漏报率  $P_{FD}$ 、误警率  $P_{FA}$  指标上均具有更好的表现。

#### 4.1.3 样本的失稳模式识别

为了验证基于 Transformer 评估模型的可解释性,将从样本的机组注意力权重和其失稳模式的对应关系对模型的可解释性进行分析。

失稳样本 1 的失稳信息为:IEEE39 节点系统的负载水平为 110%时,母线 5-6 之间发生永久性三相短路故障,故障位置位于线路的 90%处,故障持续

时间为 13 个周波。失稳样本 1 的功角趋势如图 3 所示,各机组注意力分布如图 4 所示。

表 3 不同分类模型的评估结果

Table 3 Evaluation results of different classification models

| 评估模型        | 评价指标         |             |             |
|-------------|--------------|-------------|-------------|
|             | $P_{ACC}/\%$ | $P_{FD}/\%$ | $P_{FA}/\%$ |
| Transformer | 98.80        | 1.68        | 0.88        |
| LSTM        | 98.02        | 3.20        | 1.18        |
| ALSTM       | 98.22        | 2.67        | 1.20        |
| CNN         | 98.03        | 3.10        | 1.23        |
| DT          | 95.38        | 7.35        | 2.84        |
| SVM         | 96.80        | 5.61        | 1.64        |

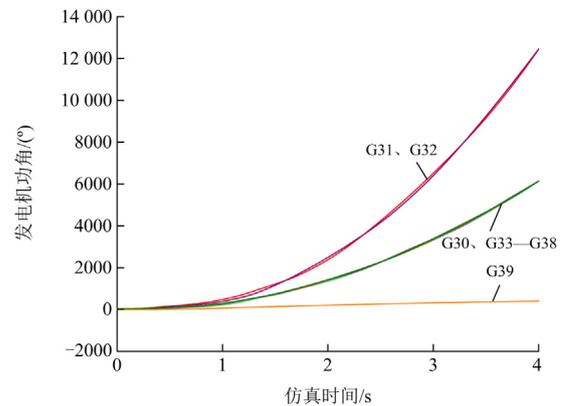


图 3 失稳样本 1 中各台机组的功角趋势

Fig. 3 Power angle trend of each generator in unstable sample 1

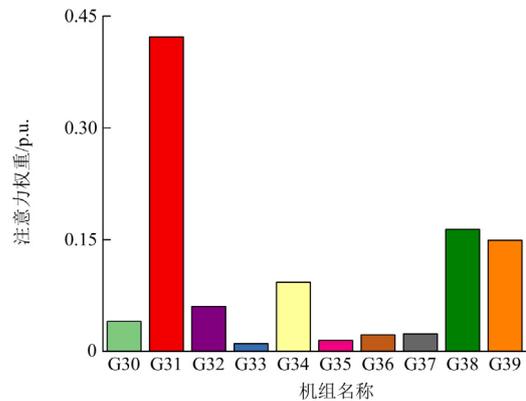


图 4 失稳样本 1 中各台机组的注意力权重分布

Fig. 4 Attention weight distribution of each generator in unstable sample 1

从图 3 可以看出,当系统失稳时,机组分为 3 群:第一群为 G31 和 G32;第二群为 G30 和 G33—G38;第三群为 G39。

从图 4 可以看出,评估模型对机组 G31、G38 和 G39 分配了较高的权重。在该失稳模式下,这三台机组即可很好地表征整个系统的失稳模式。基于

Transformer 的评估模型主要关注于 G31、G38 和 G39, 表明了该模型正确地捕捉到系统的主导失稳模式。G31 的数据变化最明显, 其被分配的注意力权重远高于其他机组, 说明数据变化趋势最明显的机组对模型的决策起到决定性的作用。

在评估样本时, 模型总是在每个机群中关注一台机组, 并依靠这几台机组作出评估。同时模型对数据变化趋势最明显的机组分配的关注度最高且远高于其他机组, 说明该机组对评估结果最为关键。注意力权重分布与失稳模式相联系, 在一定程度上提升了模型的可解释性。专业人员可以根据机组的注意力权重分布判断模型是否准确地捕捉到样本的主导模式, 分析决策结果是否可靠。

#### 4.1.4 注意力引导函数的测试

根据机组的注意力权重分布, 可以解释模型的评估规则, 并且用于指导模型的优化。如果在评估过程中, 模型没有很好地识别失稳模式, 对于关键机组关注不足, 可以通过提升关键机组的注意力权重改善该类样本的准确率。

本节为了验证注意力引导函数对于模型优化的有效性, 以某一失稳模式为例, 通过注意力引导函数提高模型在评估该失稳模式的样本时对于关键机组的关注度, 使得模型能够更准确地挖掘到系统的主导模式, 进而提高模型的准确率。

使用 4.1.1 节构建的训练集对模型进行多次训练, 随机选取一次训练后的模型, 其在测试集上的准确率为 97.96%, 共有 156 个样本评估错误。在评估错误的样本中选择一个失稳样本, 其失稳信息如下: IEEE39 总线系统的负载水平为 85% 时, 母线 17—27 之间发生永久性三相短路故障, 故障位置位于线路的 90% 处, 故障持续时间为 15.5 个周波。失稳样本 2 中各台机组的功角趋势如图 5 所示, 其失稳模式如表 4 所示。

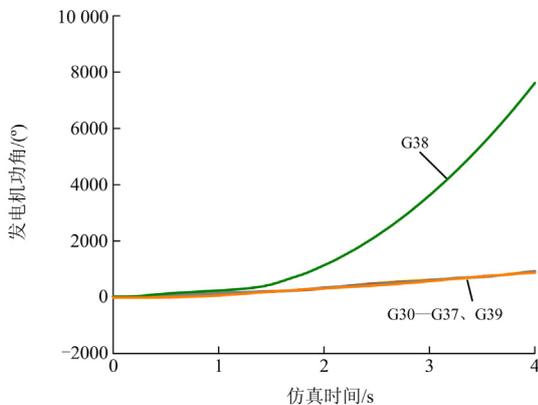


图 5 失稳样本 2 中各台机组的功角趋势

Fig. 5 Power angle trend of each generator in unstable sample 2

表 4 失稳样本 2 的失稳模式

Table 4 Instability pattern of unstable sample 2

| 机群编号 | 机组编号        |
|------|-------------|
| 机群 1 | G38         |
| 机群 2 | G30—G37、G39 |

表 4 中, 电力系统失稳的方式为两群失稳, 超前机群中仅有机组 G38, 滞后机群由其余机组组成。微调前的模型在评估失稳样本 2 时注意力权重分布如图 6 所示。

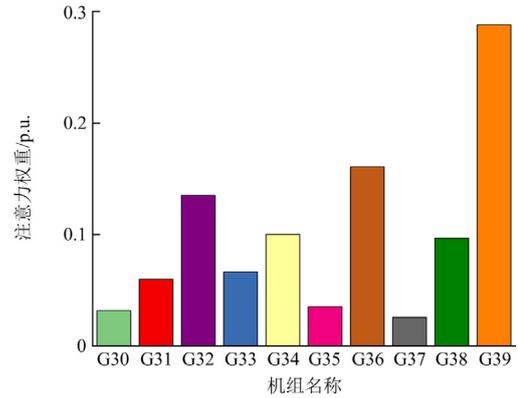


图 6 微调前失稳样本 2 的注意力权重分布

Fig. 6 Attention weight distribution of unstable sample 2 before fine-tuning

由表 4 和图 6 可知, 微调前的 Transformer 模型在评估失稳样本 2 时, 对属于超前机群的 G38 机组关注不足, 其注意力权重值仅有 9.67%, 表明模型在评估时没有很好地识别其主导模式, 导致评估错误。

为了能够正确评估失稳样本 2, 在原有的损失函数中引入注意力引导函数, 并对模型进行微调更新。微调前后模型的评估结果如表 5 所示。

表 5 模型微调前后的评估结果

Table 5 Evaluation results before and after fine-tuning

|       | $P_{ACC}/\%$ | $P_{FD}/\%$ | $P_{FA}/\%$ |
|-------|--------------|-------------|-------------|
| 模型微调前 | 97.96        | 1.12        | 2.65        |
| 模型微调后 | 98.10        | 0.92        | 2.54        |

微调后的模型在评估失稳样本 2 时, 各台机组的注意力权重分布如图 7 所示。由图 7 可知, 通过注意力引导函数能有效增加关键机组的注意力权重, 模型对 G38 的注意力权重值提升至 40.5%。提高关键机组的权重值后能正确评估失稳样本 2。

由表 5 可得出微调后的模型, 整体的误判率降低了 7%, 对于失稳样本的误判率降低了 17.65%。对失稳样本误判的降低除了因为示例的失稳样本 2 能够评估正确外, 还因为其他误判样本的失稳模式

与所关注的失稳模式相同，在评估这些样本时 G38 的关注度也被提高，使模型可以评估正确。

结合漏报率  $P_{FD}$  和误警率  $P_{FA}$  还可以得出，在微调模型前后对于其他失稳模式的样本以及稳定样本的评估结果前后基本一致，这也表明了注意力引导函数对其他样本的影响很小。

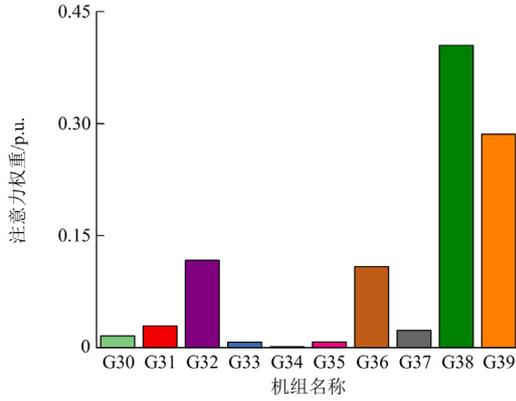


图 7 微调后失稳样本 2 的注意力权重分布

Fig. 7 Attention weight distribution of unstable sample 2 after fine-tuning

接下来讨论不同的引导系数  $\lambda$  对模型更新过程产生的影响。针对失稳样本 2 设置不同的引导系数  $\lambda$  对模型进行更新，模型更新后的评估结果如表 6 所示。

由表 6 可以看出，引导系数  $\lambda$  的取值在一定范围内时能有效地对模型进行更新优化。而如果取值过大或过小，则可能会对更新过程有一定的影响。引导系数  $\lambda$  取值太小会导致注意力引导函数在更新过程中发挥不了效果，更新后的模型依然无法正确评估选定的失稳模式。引导系数  $\lambda$  取值太大，更新后的模型虽能正确评估选定的失稳模式，但可能会导致有少部分其他的样本在更新后被评估错误。因此在更新过程中应当注意引导系数的取值。

表 6 在不同的引导系数  $\lambda$  下更新后模型的评估结果

|                  | $P_{ACC}$ | $P_{FD}$  | $P_{FA}$   | 正确评估失稳样本 2 |
|------------------|-----------|-----------|------------|------------|
| 原始模型             | 97.96%    | 1.12%(34) | 2.65%(122) | 否          |
| $\lambda = 0.04$ | 97.99%    | 1.09%(33) | 2.60%(121) | 否          |
| $\lambda = 0.20$ | 98.07%    | 0.96%(29) | 2.56%(119) | 是          |
| $\lambda = 0.40$ | 98.10%    | 0.92%(28) | 2.54%(118) | 是          |
| $\lambda = 0.80$ | 98.05%    | 0.92%(28) | 2.62%(122) | 是          |
| $\lambda = 4.00$ | 97.90%    | 0.99%(30) | 2.82%(131) | 是          |

#### 4.2 华东电网系统

进一步在华东电网系统中测试本文所提优化

模型评估规则方法的有效性。本文将 597 条电压等级为 500 kV 或 1000 kV 的输电线路作为测试线路。故障类型为  $N-1$  三相短路、 $N-2$  三相短路和两相接地短路，其他仿真细节如表 7 所示。PSD-BPA 共生成了 35 840 个样本，其中不稳定样本有 7719 个，稳定样本有 28 121 个。训练样本数为 23 040 (包含 4969 个不稳定样本，18 071 个稳定样本)，测试样本数为 12 800 (包含 2750 个不稳定样本，10 050 个稳定样本)。选择故障清除时刻和故障清除后 3 个周波时刻的所有发电机功角与角速度偏差数据作为评估模型的输入特征量。

表 7 华东电网系统的仿真细节

Table 7 Simulation details based on East China power grid

| 仿真参数     | 参数细节                |
|----------|---------------------|
| 故障线路     | 597 条输电线路           |
| 故障位置     | 每条线路的 10%、50%和 90%处 |
| 负荷水平     | 100%                |
| 故障持续时间   | 5~20 个周波            |
| PMU 采样频率 | 100 次/s             |

训练好的模型在测试集上的准确度为 97.85%。为了验证本文提出的方法在实际系统中的有效性，在评估错误的样本中选择 4 个具有不同失稳模式的样本，这 4 个样本的失稳模式如表 8 所示。

表 8 失稳样本 3—6 的失稳模式

Table 8 Instability pattern of unstable sample 3—6

| 样本编号   | 机群 1    | 机群 2            |
|--------|---------|-----------------|
| 失稳样本 3 | G31     | G1—G30, G32—426 |
| 失稳样本 4 | G36、G37 | G1—G35, G38—426 |
| 失稳样本 5 | G22、G23 | G1—G21, G24—426 |
| 失稳样本 6 | G20、G21 | G1—G19, G22—426 |

这 4 个样本均呈两群失稳模式，其中机群 1 为超前机群。针对不同的失稳模式分别对模型进行微调，提高对不同关键机组的注意力权重，其在测试集上的评估结果如表 9 所示。其中模型 A 代表微调前的模型，模型 B—E 分别代表针对样本 3—6 的失稳模式进行微调后的模型。模型 F 表示同时针对样本 3—6 的失稳模式进行微调后的模型。

表 9 不同模型的评估结果

Table 9 Evaluation results of different models

| 模型编号 | $P_{ACC}$ | $P_{FD}$   | $P_{FA}$    |
|------|-----------|------------|-------------|
| 模型 A | 97.85%    | 3.20% (88) | 1.86% (187) |
| 模型 B | 97.88%    | 2.98% (82) | 1.89% (190) |
| 模型 C | 97.96%    | 2.55% (70) | 1.90% (191) |
| 模型 D | 98.13%    | 2.84% (78) | 1.60% (161) |
| 模型 E | 98.00%    | 2.69% (74) | 1.80% (181) |
| 模型 F | 97.95%    | 2.33% (64) | 1.97% (198) |

模型 B—E 能准确地评估相应的样本 3—6, 即微调后的模型能够学习到对应的评估规则。模型 F 能够正确评估样本 3—5, 而依然无法正确评估样本 6, 这也说明了同时针对多种失稳模式进行微调, 多种失稳模式的评估规则之间可能会相互影响。

由表 9 可知, 无论是针对单一的失稳模式还是同时针对多个失稳模式进行更新微调, 微调更新后的模型在整体的准确度上都能有所提升。从漏报率  $P_{FD}$  可以看出, 除了能够正确评估示例的样本, 对于具有相同失稳模式的样本模型也能够正确评估。同时针对 4 种失稳模式进行微调的模型 F 在漏报率  $P_{FD}$  的指标上取得了最好的效果, 对于失稳样本的误判降低了 27.27%。

值得注意的是, 模型若是仅针对一个失稳模式进行更新微调, 那么对于其他失稳模式的样本和稳定样本的影响是非常微弱的。然而若同时关注多个失稳模式时, 微调后的模型虽然能够有效提升对这几个失稳模式的样本的评估精度, 但会导致少部分其他失稳模式的样本和稳定样本在微调前后的评估结果不一致。由表 9 可以看出, 模型 F 的误警率  $P_{FA}$  相比微调之前提高了 0.11%, 说明模型同时关注了多个失稳模式导致对其他样本的评估规则产生了微小影响, 有少部分之前评估正确的稳定样本在重新训练后被评估错误。

#### 4.3 模型更新效率与在线评估时间的讨论

本文考虑到更新计算成本问题, 因此在已经训练好的模型基础上进行微调, 即对原有的模型通过修正损失函数进行再训练。在不同规模的电力系统下, 模型更新所需时间如表 10 所示。

表 10 不同电力系统下更新所需的时间

Table 10 Update time in different power systems

| 测试系统        | 更新时间/s |
|-------------|--------|
| IEEE39 节点系统 | 30.03  |
| 华东电网系统      | 367.63 |

由表 10 可以看出, 在不同电力系统中本文方法的更新计算成本都不高。通过本文方法能够在短时间内迅速调整模型, 使得模型不断满足实际需求。

暂态稳定评估对评估模型的时效性和快速性具有较高要求。在不同电力系统下单个样本的计算时间如表 11 所示。

表 11 在不同电力系统下单个样本的计算时间

Table 11 Calculation time for each sample in different power systems

| 测试系统        | 样本评估时间/ms |
|-------------|-----------|
| IEEE39 节点系统 | 2.30      |
| 华东电网系统      | 4.89      |

表 11 表明基于 Transformer 的暂态稳定评估模型完全可以满足在线评估的需求。

## 5 结论

本文以 Transformer 编码器为核心搭建了暂态稳定评估模型, 结合机组的注意力权重分布与样本的失稳信息对模型评估结果的可解释性进行了分析。在此基础上, 对评估错误的失稳模式, 引入注意力引导函数, 辅助模型挖掘到正确的主导模式。通过 IEEE39 节点系统和华东电网系统的仿真, 得到如下结论:

1) 基于 Transformer 的暂态稳定评估模型识别了样本的主导模式, 并依据代表性机组作出评估。机组注意力权重的分布与样本的失稳信息相互印证, 提升了模型的可解释性。

2) 对于评估错误的样本, 其主要原因在于模型无法正确识别系统的主导模式。通过注意力引导函数微调后的模型, 能有效提高对指定失稳模式的挖掘以及模型的预测精度。

3) 注意力引导函数针对单一失稳模式对模型进行微调时, 微调后的模型在尽可能不影响其他样本评估准确率的前提下, 减少对指定失稳模式样本的误判。针对多个失稳模式进行微调时, 微调后的模型虽然会对其他样本的评估造成一定的影响, 但能够同时提高这几个指定失稳模式的准确率。

模型的更新微调过程可能会导致少部分其他的样本在微调前后的评估结果不一致, 为保持模型对其他样本的评估能力, 下一步工作需要探索如何减少更新过程对其他样本产生的影响。

## 参考文献

- [1] 董新洲, 汤涌, 卜广全, 等. 大型交直流混联电网安全运行面临的问题与挑战[J]. 中国电机工程学报, 2019, 39(11): 3107-3119.  
DONG Xinzhou, TANG Yong, BU Guangquan, et al. Confronting problem and challenge of large scale AC-DC hybrid power grid operation[J]. Proceedings of the CSEE, 2019, 39(11): 3107-3119.
- [2] 江涵, 岳程燕, 严兴煜, 等. 高比例可再生能源系统惯量约束对灵活性分析的影响研究[J]. 电力系统保护与控制, 2021, 49(18): 44-51.  
JIANG Han, YUE Chengyan, YAN Xingyu, et al. Influence of system inertia on flexibility resource analysis for an interconnection system with a high proportion of intermittent renewable energy[J]. Power System Protection and Control, 2021, 49(18): 44-51.
- [3] 徐博涵, 向月, 潘力, 等. 基于深度强化学习的含高比

- 例可再生能源配电网就地分散式电压管控方法[J]. 电力系统保护与控制, 2022, 50(22): 100-109.
- XU Bohan, XIANG Yue, PAN Li, et al. Local decentralized voltage management of a distribution network with a high proportion of renewable energy based on deep reinforcement learning[J]. Power System Protection and Control, 2022, 50(22): 100-109.
- [4] CHAN K W, CHEUNG C H, SU H T, et al. Time domain simulation based transient stability assessment and control[C] // International Conference on Power System Technology, October 13-17, 2002, Kunming, China.
- [5] LIU Y, SUN K, YAO R, et al. Power system time domain simulation using a differential transformation method[J]. IEEE Transactions on Power Systems, 2019, 34(5): 3739-3748.
- [6] ZOU Y, YIN M H, CHIANG H D. Theoretical foundation of the controlling UEP method for direct transient-stability analysis of network-preserving power system models[J]. IEEE Transactions on Circuits and Systems I-Fundamental Theory and Applications, 2003, 50(10): 1324-1336.
- [7] XUE Y, CUTSEM T V, RIBBENSPAVELLA M. A simple direct method for fast transient stability assessment of large power-systems[J]. IEEE Transactions on Power Systems, 1988, 3(2): 400-412.
- [8] 张晓英, 史冬雪, 张琏, 等. 基于 CPSO-BP 神经网络的风电并网暂态电压稳定评估[J]. 智慧电力, 2021, 49(10): 38-44.
- ZHANG Xiaoying, SHI Dongxue, ZHANG Jin, et al. Transient voltage stability assessment of power system integrated with wind power based on CPSO-BP neural network[J]. Smart Power, 2021, 49(10): 38-44.
- [9] 朱轶伦, 罗焯锋, 高强, 等. 基于 LSTM 的电力暂态稳定在线评估及预测研究[J]. 电网与清洁能源, 2021, 37(3): 38-46.
- ZHU Yilun, LUO Yefeng, GAO Qiang, et al. Research on online assessment and prediction of power system transient stability based on LSTM[J]. Power System and Clean Energy, 2021, 37(3): 38-46.
- [10] 赵慧敏, 帅智康, 沈阳, 等. 基于深度学习的多虚拟同步机微电网在线暂态稳定评估方法[J]. 电力系统自动化, 2022, 46(9): 109-117.
- ZHAO Huimin, SHUAI Zhikang, SHEN Yang, et al. Online transient stability assessment method for microgrid with multiple virtual synchronous generators based on deep learning[J]. Automation of Electric Power Systems, 2022, 46(9): 109-117.
- [11] 杨跃, 刘友波, 刘俊勇, 等. 基于神经网络预测校核的暂态稳定预防控制[J]. 电网技术, 2018, 42(12): 4076-4084.
- YANG Yue, LIU Youbo, LIU Junyong, et al. Preventive transient stability control based on neural network security predictor[J]. Power System Technology, 2018, 42(12): 4076-4084.
- [12] 陈厚合, 王长江, 姜涛, 等. 基于投影能量函数和 Pin-SVM 的电力系统暂态稳定评估[J]. 电工技术学报, 2017, 32(11): 67-76.
- CHEN Houhe, WANG Changjiang, JIANG Tao, et al. Transient stability assessment in bulk power grid using projection energy function and support vector machine with pinball loss[J]. Transactions of China Electrotechnical Society, 2017, 32(11): 67-76.
- [13] OYAMA T, YOKOYAMA A, SAKAEDA S, et al. Online transient stability assessment employing decision tree and maximal Lyapunov exponent[C] // International Conference on Smart Grids and Energy Systems (SGES), November 23-26, 2020, Perth, Australia.
- [14] LIU C X, SUN K, RATHER Z H, et al. A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees[J]. IEEE Transactions on Power Systems, 2014, 29(2): 717-730.
- [15] 朱乔木, 党杰, 陈金富, 等. 基于深度置信网络的电力系统暂态稳定评估方法[J]. 中国电机工程学报, 2018, 38(3): 735-743.
- ZHU Qiaomu, DANG Jie, CHEN Jinfu, et al. A method for power system transient stability assessment based on deep belief networks[J]. Proceedings of the CSEE, 2018, 38(3): 735-743.
- [16] 邵美阳, 吴俊勇, 李宝琴, 等. 基于两阶段集成深度置信网络的电力系统暂态稳定评估[J]. 电网技术, 2020, 44(5): 1776-1787.
- SHAO Meiyang, WU Junyong, LI Baoqin, et al. Transient stability assessment of power system based on two-stage ensemble deep belief network[J]. Power System Technology, 2020, 44(5): 1776-1787.
- [17] 朱乔木, 陈金富, 李弘毅, 等. 基于堆叠自动编码器的电力系统暂态稳定评估[J]. 中国电机工程学报, 2018, 38(10): 2937-2946, 3144.
- ZHU Qiaomu, CHEN Jinfu, LI Hongyi, et al. Transient stability assessment based on stacked autoencoder[J]. Proceedings of the CSEE, 2018, 38(10): 2937-2946, 3144.
- [18] 田芳, 周孝信, 史东宇, 等. 基于卷积神经网络的电力系统暂态稳定预防控制方法[J]. 电力系统保护与控制, 2020, 48(18): 1-8.
- TIAN Fang, ZHOU Xiaoxin, SHI Dongyu, et al. A preventive control method of power system transient stability based on a convolutional neural network[J]. Power

- System Protection and Control, 2020, 48(18): 1-8.
- [19] 李楠, 朱娜, 崔莹. 考虑代价敏感的 AC-LSTM 暂态稳定评估[J]. 电力系统保护与控制, 2022, 50(22): 160-169.  
LI Nan, ZHU Yuan, CUI Ying. AC-LSTM transient stability assessment considering cost-sensitivity[J]. Power System Protection and Control, 2022, 50(22): 160-169.
- [20] 吴思婕, 王怀远. 基于集成学习的时间自适应电力系统暂态稳定评估方法[J]. 电力系统保护与控制, 2022, 50(24): 112-119.  
WU Sijie, WANG Huaiyuan. Transient stability assessment of power system with time-adaptive method based on ensemble learning[J]. Power System Protection and Control, 2022, 50(24): 112-119.
- [21] 赵恺, 石立宝. 基于改进一维卷积神经网络的电力系统暂态稳定评估[J]. 电网技术, 2021, 45(8): 2945-2957.  
ZHAO Kai, SHI Libao. Transient stability assessment of power system based on improved one-dimensional convolutional neural network[J]. Power System Technology, 2021, 45(8): 2945-2957.
- [22] WANG H Y, WANG Q Y. Adaptive cost-sensitive assignment method for power system transient stability assessment[J]. International Journal of Electrical Power & Energy Systems, 2022, 135.
- [23] 张若愚, 吴俊勇, 李宝琴, 等. 基于迁移学习的电力系统暂态稳定自适应预测[J]. 电网技术, 2020, 44(6): 2196-2205.  
ZHANG Ruoyu, WU Junyong, LI Baoqin, et al. Self-adaptive power system transient stability prediction based on transfer learning[J]. Power System Technology, 2020, 44(6): 2196-2205.
- [24] GUIDOTTI R, MONREALE A, RUGGIERI S, et al. A survey of methods for explaining black box models[J]. ACM Computer Survey, 2019, 51(5).
- [25] ZHANG Y C, XU Y, BU S Q, et al. Online power system dynamic security assessment with incomplete PMU measurements: a robust white-box model[J]. IET Generation Transmission & Distribution, 2019, 13(5): 662-668.
- [26] 时纯, 刘君, 梁卓航, 等. 基于 GAN 和多通道 CNN 的电力系统暂态稳定评估[J]. 电网技术, 2022, 46(8): 3191-3202.  
SHI Chun, LIU Jun, LIANG Zhuohang, et al. Transient stability assessment of power system based on GAN and multi-channel CNN[J]. Power System Technology, 2022, 46(8): 3191-3202.
- [27] REN J Y, LI B Y, ZHAO M, et al. Optimization for data-driven preventive control using model interpretation and augmented dataset[J]. Energies, 2021, 14(12).
- [28] SINGH P, SHARMA A. Interpretation and classification of Arrhythmia using deep convolutional network[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71.
- [29] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C] // 31st Annual Conference on Neural Information Processing Systems (NIPS), 2017, Long Beach, CA, USA.
- [30] LIANG J K, TANG W Y. Ultra-short-term spatiotemporal forecasting of renewable resources: an attention temporal convolutional network-based approach[J]. IEEE Transactions on Smart Grid, 2022, 13(5): 3798-3812.
- [31] TOUBEAU J F, BOTTIEAU J, WANG Y, et al. Interpretable probabilistic forecasting of imbalances in renewable-dominated electricity systems[J]. IEEE Transactions on Sustainable Energy, 2022, 13(2): 1267-1277.
- [32] LIM B, ARIK S O, LOEFF N, et al. Temporal fusion transformers for interpretable multi-horizon time series forecasting[J]. International Journal of Forecasting, 2021, 37(4): 1748-1764.
- [33] 安军, 艾士琪, 刘道伟, 等. 基于短时受扰轨迹的电力系统暂态稳定评估方法[J]. 电网技术, 2019, 43(5): 1690-1697.  
AN Jun, AI Shiqi, LIU Daowei, et al. A power system transient stability assessment method based on short-time disturbed trajectories[J]. Power System Technology, 2019, 43(5): 1690-1697.
- [34] ZHOU H X, ZHANG Y J, YANG L F, et al. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism[J]. IEEE Access, 2019, 7: 78063-78074.

收稿日期: 2023-02-20; 修回日期: 2023-05-15

作者简介:

高发骏(1999—), 男, 硕士研究生, 研究方向为机器学习在电力系统中的应用; E-mail: 215057180@qq.com

王怀远(1987—), 男, 通信作者, 博士, 副教授, 研究方向为电力系统运行与控制; E-mail: 79749544@qq.com

党然(1991—), 男, 本科, 研究方向为人工智能算法研究与应用。E-mail: 584310203@qq.com

(编辑 姜新丽)