

DOI: 10.12158/j.2096-3203.2025.01.016

融合 MHSA 与 Boruta 的电力系统暂态功角稳定关键特征筛选

王曼¹, 周小雨¹, 陈凡^{1,2}, 赖业宁², 朱瑛³

(1. 南京工程学院电力工程学院, 江苏 南京 211167; 2. 智能电网保护和运行控制国家重点实验室, 江苏 南京 211106; 3. 河海大学能源与电气学院, 江苏 南京 211100)

摘要: 现有暂态稳定特征选择方法中初始特征的选定会限制后续寻找最佳特征组合的能力, 同时缺乏客观方法来确定关键特征的数量, 为此, 文中提出一种融合多头自注意力 (multi-head self-attention, MHSA) 与 Boruta 的暂态功角稳定关键特征筛选方法。首先, 构建深度神经网络模型, 并在输入侧添加 MHSA 模块进行暂态稳定评估。MHSA 直接面向输入的电网特征, 可在模型训练过程中自适应调整注意力权重, 聚焦关键特征。其次, 利用 Boruta 算法生成真假特征组合, 经过 MHSA 模型的训练, 选择高于最大虚假特征权重的真实特征, 由模型本身确定关键特征数量。最后, 在 IEEE 39 和 IEEE 118 节点系统上进行算例分析。算例结果表明, 所提方法可在保证评估精度的同时大幅减少输入特征的数量, 相比于传统方法, 可选出评估精度更高的关键特征。

关键词: 多头自注意力 (MHSA); Boruta 算法; 暂态稳定; 特征选择; 关键特征; 虚假特征

中图分类号: TM712

文献标志码: A

文章编号: 2096-3203(2025)01-0155-10

0 引言

近年来, 人工智能和数据挖掘等技术凭借其强大的数据处理和挖掘能力在电网领域迅速发展, 有效解决了电网稳定评估过程中面临的机理复杂、数据维度庞大等问题^[1-2]。基于人工智能的电力系统暂态稳定评估通常采用反映电网运行状态的电气量作为模型的输入特征^[3-5], 但其数量较多, 直接应用于实际电网可能导致监测点过多、通信困难, 同时会引入冗余特征或与电网稳定无关的噪声特征。此外, 直接使用这些特征进行暂态稳定评估也会导致分析速度慢、结果精度低等问题^[6]。因此, 对于数量众多的电网特征, 须采用合适的方法进行智能筛选, 减少特征数量, 提高评估效率, 使其更适用于实际应用场景。

目前已有学者针对暂态稳定评估模型的输入电气特征展开研究, 在备选电气量中利用特征选择方法筛选关键特征。文献[7]通过 Fisher Score 分析人工选取的所有潮流特征量, 仅得到电气特征的重要性排名; 文献[8]基于关键潮流断面特征, 计算备选特征与断面特征的条件互信息, 选出关键特征。上述过滤式方法仅针对电气特征进行分析, 缺乏模型结果的引导, 难以选出关键的特征组合, 且关键特征的数量也需要人为选取。文献[9]采用包裹式的蚱蜢优化算法搜索关键特征, 并结合互信息优先搜索优先级高的特征, 但难以应对大规模特

征; 文献[10-12]均采用两阶段式特征选择方法, 先利用过滤式方法降低备选电气特征数量, 再采用包裹式方法搜索关键特征, 虽然可以提高计算效率, 但在一定程度上会破坏原始特征之间的相关性, 并且其特征选择结果依赖于包裹式方法中的内嵌分类器效果以及选定的初始特征。

现有各领域的特征选择方法除了上述过滤式方法和包裹式方法, 还有嵌入式方法^[13]。过滤式方法通过度量每个特征与标签之间的关系为特征打分, 并据此进行特征排序从而选择关键特征。常见的度量准则有互信息^[14]、ReliefF 准则^[15]、Fisher 准则^[16-17]等。过滤式方法从基础的互信息最大化 (mutual information maximization, MIM) 方法^[18]开始发展, 逐渐考虑特征之间的相关性和冗余性, 进而提出最小冗余最大相关^[19]、条件互信息^[20]、联合互信息 (joint mutual information, JMI)^[21]等方法, 但该方法筛选关键特征时无分类器参与, 所选特征的实际表现可能不佳^[22]。包裹式方法则将机器学习模型嵌入特征选择过程中, 可以考虑模型反馈, 以模型结果为评价标准引导特征选择过程, 同时由于机器学习模型对特征组合进行训练, 该方法还能考虑特征之间的互补性。常见的包裹式方法有递归特征消除、前后向序列搜索等, 也有学者将选出的最佳特征组合转化为 NP 难问题^[23]。包裹式方法以模型结果为导向筛选关键特征, 具有较高的准确率, 但仍有不可避免的缺点: 随着备选特征数量增加, 计算时间会显著增长, 并且一旦选定特征, 之后调整特征集的灵活性就会受限。因此, 在进行特征

收稿日期: 2024-06-06; 修回日期: 2024-08-17

基金项目: 国家自然科学基金资助项目 (U1966205)

选择时,需要权衡准确率和计算效率,并考虑后续模型调整的灵活性。在此背景下,嵌入式方法在模型训练时通过评估特征权重选择重要特征,如文献[24]使用极端梯度提升(extreme gradient boosting, XGBoost)模型的树权重确定特征重要性;文献[25]依据神经网络结束训练后的特征权重选取特征。但嵌入式方法依赖底层模型的性能,因此为提高特征选择的性能,引入注意力机制以增强模型对数据的焦点,如文献[26]在深层神经网络中加入注意力机制,提高了模型的性能和表达能力;文献[27]在训练过程中引入自注意力机制,自动学习高层次特征表示,并在公共数据集上应用,获得了较高的准确率。

目前,在暂态稳定的特征选择中大多采用过滤式和包裹式方法,鲜有对嵌入式方法的研究。文中旨在研究结合注意力机制的暂态稳定特征选择方法,增强模型对数据的敏感度和选择准确度。为增强暂态稳定特征选择的灵活性以及特征数量选择的客观性,融合多头自注意力(multi-head self-attention, MHA)机制与 Boruta 算法^[28]优化关键特征的识别过程。MHA 模型负责特征加权与动态选择, Boruta 算法则负责提供一种衡量真实及虚假特征权重的策略,从而确定关键特征的判定阈值。完成特征选择后,在 IEEE 39 和 IEEE 118 节点系统上利用 4 种机器学习分类器对所选特征性能进行算例分析,验证了所提方法的有效性和合理性。

1 暂态功角稳定关键特征筛选模型

文中提出的融合 MHA 与 Boruta 的暂态功角稳定特征选择模型如图 1 所示,主要由 Boruta 以及 MHA 神经网络组成。图 1 中, x_j 为一个样本输入的第 j 个真实特征, $j=1,2,\dots,n$, n 为真实或虚假特征的数量; x'_j 为输入的第 j 个虚假特征。

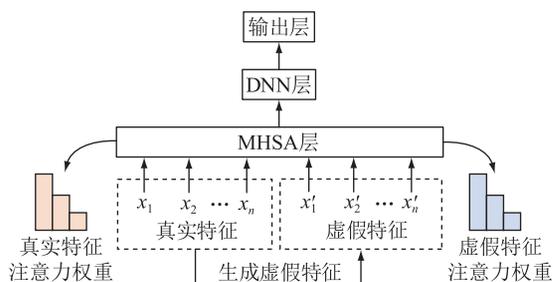


图 1 融合 MHA 与 Boruta 的暂态功角稳定特征选择模型
Fig.1 Transient power angle stability feature selection model integrating MHA and Boruta

Boruta 算法是机器学习中随机森林(random forest, RF)算法的扩展,其核心思想为构建“随机影子”虚假特征集进行特征选择,通过真实特征和虚

假特征之间的竞争,识别出对模型预测结果有重要影响的关键特征。基于此,利用 Boruta 算法根据备选电气特征生成一组虚假特征,与真实特征共同训练,形成竞争机制,从而筛选出关键特征。

MHA 神经网络包括 MHA 层和深度神经网络(deep neural network, DNN)层。MHA 机制常用于自然语言处理等领域,其优势在于能同时关注序列中的不同位置,更好地捕捉上下文信息。传统暂态稳定特征选择方法,尤其是两阶段式方法通常忽略特征之间的交互,从而导致割裂特征整体的关联性。文中通过 MHA 机制实现所有备选电气特征的全局关注,输入特征经过注意力层运算后形成新的加权特征,由 DNN 层负责学习该加权特征与暂态稳定性之间的非线性关系,从而更全面地进行稳定性判别。

2 算法原理介绍

2.1 Boruta 算法原理

图 2 为利用 Boruta 算法生成虚假特征的示意图,其中 s_1, s_2, s_3 为 3 个样本。随机排列每列真实特征的值生成一组具有相同分布的虚假特征, Boruta 算法利用这些特征训练模型,通过比较真实特征和虚假特征的重要性确定真正重要的特征。

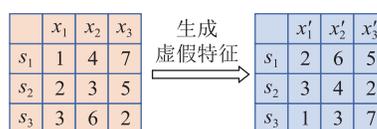


图 2 生成虚假特征示意

Fig.2 Schematic diagram of generating virtual features

选取电力系统中的稳态潮流特征,如各节点的电压幅值、相角和各输电线路的有功、无功功率特征作为文中的备选电气特征 X_0 。

$$X_0 = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix} \quad (1)$$

式中: $x_{i,j}$ 为第 i 个潮流样本的第 j 个真实特征, $i=1, 2, \dots, m$, m 为总潮流样本数量。

经过 Boruta 算法生成一组虚假电气特征,并与真实特征拼接,可表示为:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} & x'_{1,1} & x'_{1,2} & \cdots & x'_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} & x'_{2,1} & x'_{2,2} & \cdots & x'_{2,n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} & x'_{m,1} & x'_{m,2} & \cdots & x'_{m,n} \end{bmatrix} \quad (2)$$

式中: X 为拼接真假特征的模型输入特征; $x'_{i,j}$ 为第*i*个潮流样本的第*j*个真实特征对应的虚假特征。

2.2 MHSA 神经网络模型

文中提出的MHSA机制特征选择模型的基本架构如图3所示,主要由MHSA模块和DNN模块两部分组成。模型引入残差连接防止梯度退化,加速算法收敛,同时采用Mish函数作为激活函数。

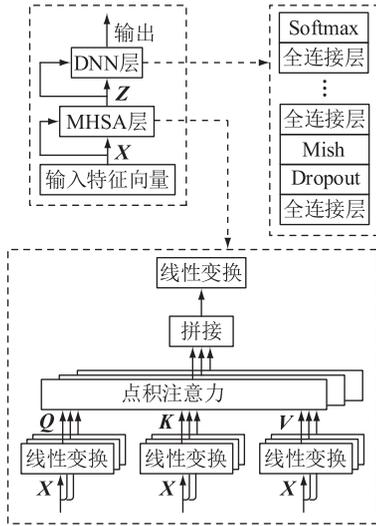


图3 MHSA神经网络模型架构

Fig.3 MHSA neural network model architecture

由图3可知,模型输入特征 X 经过不同的线性变换分别得到查询矩阵 Q 、键矩阵 K 和值矩阵 V ,计算公式^[29]为:

$$\begin{cases} Q = W^Q X \\ K = W^K X \\ V = W^V X \end{cases} \quad (3)$$

式中: W^Q 、 W^K 、 W^V 为3种权重矩阵,在训练过程中通过反向传播不断更新参数, W^Q 、 W^K 、 $W^V \in \mathbf{R}^{2n \times d_k}$, d_k 为训练过程设置的批处理大小。

基于上述得到的3种权重矩阵,充分考虑输入潮流特征之间的关联性,对输入特征计算注意力权重,形成新的加权特征,计算公式为:

$$Z_0 = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (4)$$

式中: Z_0 为新的加权特征; $1/\sqrt{d_k}$ 为缩放因子,可以避免点乘的结果过大造成归一化指数函数Softmax梯度过低,提高训练稳定性。

在自注意力机制的基础上,生成*h*组相等大小的子空间 Q_l 、 K_l 、 V_l , $l=1,2,\dots,h$ 。对每组子空间分别进行注意力计算,最后将每组的输出结果进行拼接,计算公式为:

$$H_l = \text{Softmax} \left(\frac{Q_l(K_l)^T}{\sqrt{d_k}} \right) V_l \quad (5)$$

$$Z = [H_1 \ H_2 \ \dots \ H_h] W^o \quad (6)$$

式中: H_l 为第*l*组子空间的注意力头计算结果; Z 为MHSA层计算后的加权特征; W^o 为可学习的参数矩阵,对其进行线性变换可获得最终注意力层输出。

其中,用于衡量输入电气特征 X 对暂态稳定性贡献程度的注意力权重为:

$$A = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) \quad (7)$$

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,2n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,2n} \\ \vdots & \vdots & & \vdots \\ a_{2n,1} & a_{2n,2} & \dots & a_{2n,2n} \end{bmatrix} \quad (8)$$

式中: A 为注意力权重矩阵; $a_{e,f}$ 为第*e*个特征分配给注意力层的第*f*个输出特征的权重, $e, f = 1, 2, \dots, 2n$ 。

而获取模型对输入特征的注意力权重则采用自注意力权重,即取权重矩阵的对角元素,MHSA则需要计算所有子空间注意力权重平均值。最终输入特征的自注意力权重矩阵可表示为:

$$W = \frac{1}{h} \sum_{l=1}^h \text{diag}(A_l) \quad (9)$$

式中: W 为MHSA下的权重矩阵,具有 $2n$ 维,其中前*n*个为真实特征的权重,后*n*个为虚假特征的权重; A_l 为第*l*组子空间的注意力权重矩阵。

权重矩阵 W 可直观展示各输入特征对暂态稳定评估模型的贡献程度,其中真实特征权重子矩阵为 W_{rel} ,虚假特征权重子矩阵为 W_{fake} 。将 W_{fake} 的最大值作为判定阈值,即可筛选出可靠的关键特征。

3 融合MHSA与Boruta的暂态稳定关键特征筛选

3.1 原始输入特征的选取

用于暂态稳定评估的输入特征大致可以分为2类:第一类输入特征由故障前、故障期间以及故障后的电气量构成;第二类则由故障前的稳态运行特征构成。第一类输入特征通常会考虑故障后电气量的时序特性来提高暂态稳定评估模型的精度,但此类特征只能在故障发生后才能监测到。由于电力系统的暂态过程十分短暂,给予故障后紧急控制措施的反应时间有限,因此文中采用故障前的稳态运行特征,如各节点的电压幅值、相角和各输电线路的有功、无功功率作为待选特征集合。

3.2 稳定判据与模型输出

文中采用扰动后发电机转子功角的暂态稳定指数(transient stability index, TSI)作为对样本稳定

的判据,表达式为:

$$I_{TSI} = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|} \quad (10)$$

式中: I_{TSI} 为 TSI 计算结果; $\Delta\delta_{\max}$ 为仿真期间任意 2 台发电机功角差的最大值。当 $I_{TSI} > 0$ 时,判定为稳定样本;反之,判定为失稳样本。

3.3 暂态稳定关键特征筛选流程

图 4 为融合 MHSa 与 Boruta 的暂态稳定关键特征筛选框图,主要分为暂态稳定样本生成、特征预处理、暂态稳定关键特征筛选、关键特征评价四部分,各部分详细步骤如下。

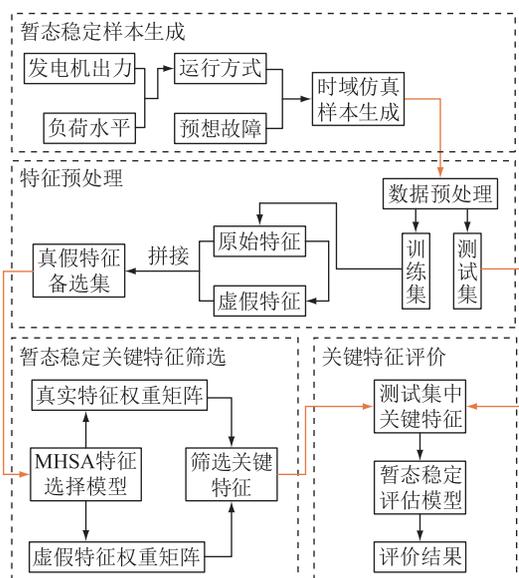


图 4 融合 MHSa 与 Boruta 的暂态稳定关键特征筛选

Fig.4 Transient stability key feature selection integrating MHSa and Boruta

(1) 暂态稳定样本生成:采用仿真工具改变各发电机出力和负荷水平以获取不同运行工况下的电力系统运行方式,结合预想故障通过时域仿真获取暂态稳定样本,利用 3.2 节所述暂态稳定判据为每个样本附加稳定与否的标签。

(2) 特征预处理:针对时域仿真所获得的暂态稳定样本,根据原始输入特征获取数据集。将标准化后的数据集随机分为训练集和测试集,训练集用于关键特征的筛选,测试集用于评价所选特征的优劣。其中训练集基于 Boruta 算法的思想,将真实特征与虚假特征送入模型中进行联合训练。

(3) 暂态稳定关键特征筛选: MHSa 模型由 MHSa 层和 DNN 层组成,训练完成后可获取真实特征权重矩阵以及虚假特征权重矩阵,将虚假特征权重矩阵的最大值作为阈值筛选出关键特征。

(4) 关键特征评价:筛选出关键特征后,在测试集中剔除非关键特征并采用常见的机器学习分类

器进行暂态稳定评价。为避免单一模型对数据的过拟合,采用 RF、轻量梯度提升机(light gradient boosting machine, LGBM)、袋装分类器(bagging classifier, BG)、决策树(decision tree, DT)4 种机器学习分类器,利用不同性能的模型全面评价关键特征,确保所选特征的可信度。

4 算例分析

文中采用 IEEE 39 和 IEEE 118 节点系统作为仿真算例,验证所提融合 MHSa 与 Boruta 的暂态稳定关键特征筛选方法的有效性,并与当前经典特征选择方法进行比较。

4.1 仿真设置与样本生成

采用电力系统分析软件包(power system analysis toolbox, PSAT)进行仿真获取暂态稳定评估所需要的数据集。为获取足够的样本进行研究,文中选取暂态稳定特征选择及控制策略研究^[30]中具有代表性的故障,设置预想故障为:1.0 s 时,IEEE 39 节点系统线路 bus16—bus21 的首端发生三相短路故障,故障持续 0.1 s 后切除该故障线路,仿真时长为 5 s;在初始标准潮流下,发电机出力和负荷水平在 80%~120% 范围内随机变化,以获取不同的运行场景。为避免样本不平衡带来的影响,按照 1:1 的比例选取 5 000 个稳定样本和 5 000 个失稳样本,并按 7:3 随机分为训练集和测试集^[31]。原始输入特征如表 1 所示,共选取 170 个特征,每个特征按顺序对应表中的序号。

表 1 原始输入特征

Table 1 Raw input features

特征序号	电气物理量
1~39	各节点电压相角
40~78	各节点电压幅值
79~124	各输电线路有功功率
125~170	各输电线路无功功率

4.2 模型评价指标

电力系统暂态稳定评估属于二分类问题,为全面评价特征选择的结果,采用表 2 所示暂态稳定评估混淆矩阵进行评价。其中, n_{TP} 、 n_{TN} 为预测正确的样本数量; n_{FP} 、 n_{FN} 为预测错误的样本数量。

准确率(accuracy, Acc)可以反映模型的全局评估准确性。暂态稳定评估更注重对失稳样本的判别,误判率(false positive rate, FPR)可以反映稳定评估中将失稳样本错判为稳定样本的情况。电力系统暂态稳定评估中 FPR 应尽量低。其中 Acc 和 FPR 的计算公式为:

表 2 暂态稳定评估混淆矩阵定义

Table 2 Definition of transient stability assessment confusion matrix

样本状态	样本数量	
	稳定(模型预测)	失稳(模型预测)
稳定	n_{TP}	n_{FP}
失稳	n_{FN}	n_{TN}

$$A_{Acc} = \frac{n_{TP} + n_{TN}}{n_{TP} + n_{TN} + n_{FP} + n_{FN}} \quad (11)$$

$$F_{FPR} = \frac{n_{FN}}{n_{TN} + n_{FN}} \quad (12)$$

式中: A_{Acc} 为 Acc 计算结果; F_{FPR} 为 FPR 计算结果。

4.3 仿真模型训练过程与结果

文中所建立的 MHSa 特征选择模型具体参数设置如下:与输入特征直接相连的注意力层共有 340 个神经元,与真假特征的总个数相等,使得每个神经元可以专注处理特定的输入特征,注意力层输入、输出的维度一致;综合考虑模型性能与计算时间,设置注意力头数量为 2;全连接层之间添加 0.2 的 Dropout 防止模型过拟合;优化器采用 Adam,初始学习率为 0.005。

4.3.1 暂态稳定模型训练及特征选择过程可视化

文中模型在训练集和测试集上的损失函数值以及准确率结果如图 5 所示。

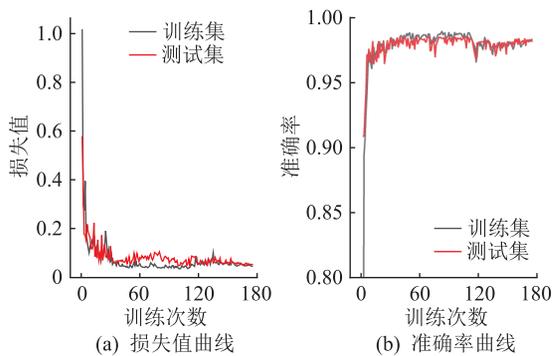


图 5 暂态稳定评估模型训练结果

Fig.5 Transient stability assessment model training results

由图 5 可知,随着训练的进行,训练集和测试集的损失值逐渐下降并趋于稳定,同时,模型在训练集和测试集上的准确率也逐渐稳定,最佳训练次数约为 170~180。结果表明,模型在经过一定训练后,在新的、未见过的数据集上能够保持一定的性能。

文中的暂态稳定特征选择方法由模型确定关键特征的数量,为此,图 6 展示了模型训练过程中所选特征数量的变化。每次训练结束后,选出的关键特征经过 RF 分类器进行评价,并以图形化方式呈现。通过图 6 可以直观地了解每次训练中选出的关

键特征的质量。

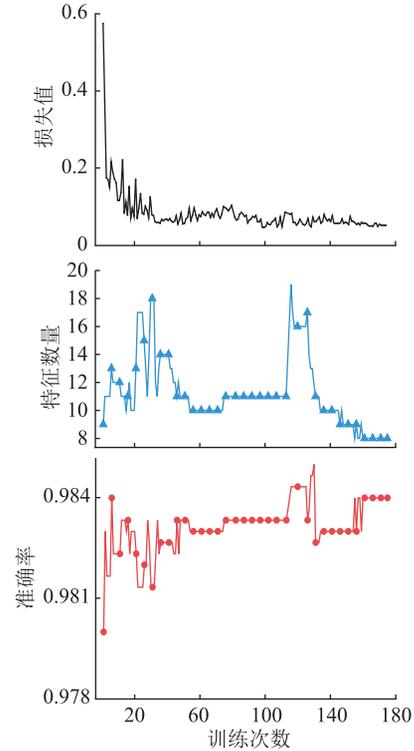


图 6 特征选择模型训练过程

Fig.6 Training process of feature selection model

由图 6 中的损失值曲线可知,随着训练次数增加,模型的损失值逐渐下降,模型的拟合程度逐渐提高。特征数量曲线代表关键特征的数量随着训练的进程而不断更新,模型在训练过程中动态调整特征的权重,并相应更新选取的关键特征。准确率曲线展示了 RF 对关键特征的评价结果,在模型训练后期,关键特征子集的评价结果较高,同时特征数量也减少,从而获得最佳的特征组合。综上所述可知,文中方法在训练过程中能够逐步提升模型的拟合能力,并根据特征的重要性动态筛选关键特征。通过这种方式,模型能够在保持较高评价结果的同时减少特征数量,得到最佳的特征组合。

4.3.2 注意力权重系数及关键特征

利用文中所提方法进行暂态稳定关键特征子集的筛选,图 7 为不同训练次数下真假特征注意力权重系数的计算结果,虚线左侧为真实特征,右侧为虚假特征。通过观察曲线的高度可以判断权重系数较高的特征主要在真实特征中,这些真实特征在暂态稳定性判别中起到了关键作用。尽管引入虚假特征可能会增加噪声特征,但由于注意力机制的存在,模型仍能关注每个特征与其他特征之间的相对重要性。因此,即使存在噪声特征的干扰,文中模型也可以从众多特征中找出关键特征。

将低于虚假特征阈值的真实特征视为无关特

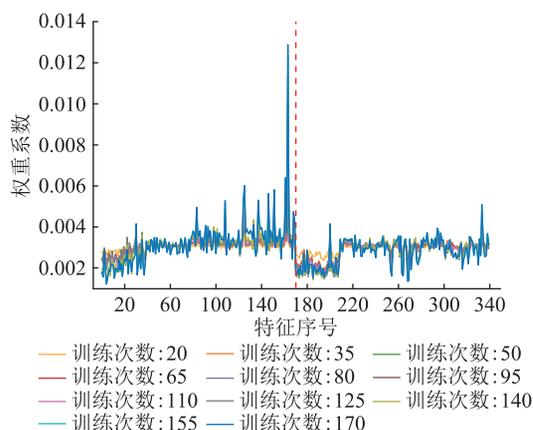


图7 真、假特征的注意力权重系数

Fig.7 Attention weight coefficient for real and virtual features

特征,能够过滤对暂态稳定性评估没有实际意义的特征,使模型更好区分真实特征的重要性,更准确地筛选关键特征,为暂态稳定性评估提供更有意义的信息。

根据图7中曲线高度最终可筛选出8个关键特征,其中虚假特征为第333号特征,对应真实特征Q_16_21。将此8个特征权重单独进行归一化,如表3所示,各个特征在系统中的分布位置如图8所示。

表3 暂态稳定关键特征名称及权重

Table 3 Names and weights of transient stability key features

序号	关键特征名称	特征权重
1	Q_16_21	0.338 9
2	Q_25_26	0.129 4
3	Q_1_39	0.112 6
4	Q_13_14	0.103 3
5	P_3_4	0.090 6
6	Q_6_31	0.079 5
7	P_17_27	0.074 1
8	Q_1_2	0.071 7

特征权重反映了所选关键特征在暂态稳定性中的相对重要性。表3中,Q_16_21的权重最高,为故障发生所在线路,表明该线路在故障时对系统稳定性产生最大影响。在系统结构不变的情况下,当故障确定时,影响暂态稳定性的自变量主要是发电机出力以及负荷水平,机组出力和负荷变化对线路功率具有最直接的影响,因此文中所选的关键特征多为线路上的功率。

分别采用RF、LGBM、BG、DT评估选定的关键特征,特征选择前后的Acc和总计算时间如表4所示。经过特征选择后,RF的Acc保持不变,DT的

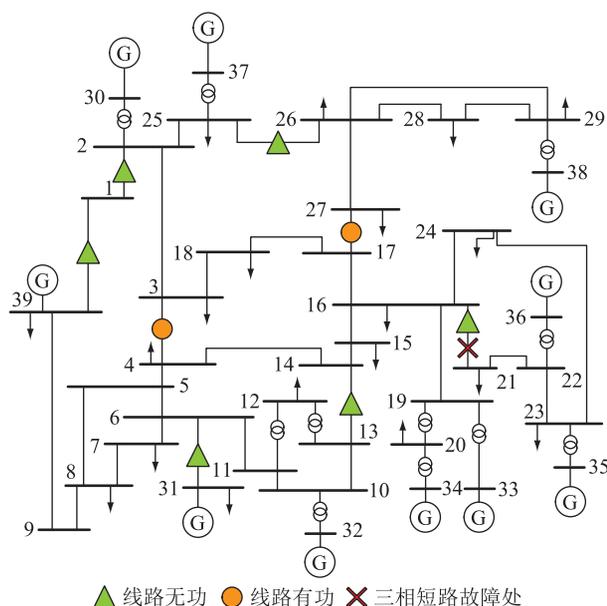


图8 关键特征分布位置

Fig.8 Distribution location of key features

Acc有所提高,其他分类器的Acc略微下降。由于特征数量减少,模型的计算压力减轻,4种模型的计算时间显著缩短。特征选择既可以降低模型计算成本,又能保持相对较高的评估准确性。

表4 特征选择前后Acc和计算时间对比

Table 4 Comparison of Acc and calculation time before and after feature selection

分类器	原始特征		关键特征	
	A_{Acc}	t/s	A_{Acc}	t/s
RF	0.984 0	2.99	0.984 0	0.93
LGBM	0.982 7	1.19	0.981 3	0.27
BG	0.983 0	3.20	0.981 0	0.26
DT	0.973 7	0.62	0.976 7	0.06

4.4 特征选择方法效果分析

为验证文中方法的有效性,分别与MIM、JMI、前向搜索策略(forward search strategy, FSS)特征选择方法进行对比。其中,MIM、JMI属于过滤式方法;FSS属于包裹式方法,其内嵌的分类器采用RF(下文表示为RF-FSS)。不同特征选择方法结果对比如表5所示。文中方法所选关键特征的数量由模型自适应确定,且最终选出了8个关键特征,为方便对比,各方法均选择8个关键特征。采用RF、LGBM、BG和DT作为分类器,比较各方法所选特征在不同分类器上判断结果的精度。由表5可知,文中所选特征子集在各种分类器上的Acc和FPR均优于其他方法。

MIM仅计算单个特征对暂态稳定预测结果的贡献程度,未考虑特征之间的冗余和互补。JMI则

表5 不同特征选择方法结果对比

Table 5 Comparison of results for different feature selection methods

方法	RF		LGBM	
	A_{Acc}	F_{FPR}	A_{Acc}	F_{FPR}
MIM	0.947 7	0.048 0	0.951 0	0.045 3
JMI	0.980 3	0.027 3	0.979 7	0.028 7
RF-FSS	0.980 7	0.026 7	0.979 0	0.024 7
文中方法	0.984 0	0.024 7	0.981 3	0.024 0
方法	BG		DT	
	A_{Acc}	F_{FPR}	A_{Acc}	F_{FPR}
MIM	0.952 0	0.055 3	0.942 7	0.042 7
JMI	0.978 0	0.027 3	0.972 7	0.031 3
RF-FSS	0.979 0	0.024 0	0.975 3	0.022 0
文中方法	0.981 0	0.023 3	0.976 7	0.021 7

通过最大化特征组合与暂态稳定标签之间的联合互信息,配合搜索算法进行特征选择,其考虑特征之间的关联,所选特征在多种分类器上的表现均明显优于MIM。RF-FSS采用具体模型的评价结果引导特征选择过程,在4种分类器中除了LGBM的准确率略低于JMI,其他分类器的结果均有所提升,但该方法的计算时间较长。JMI和RF-FSS均从空集开始逐步添加特征,一旦某个特征被选中添加到组合中,就无法更改,这可能导致无法找到更优的特征组合方案。

文中通过注意力层对全部输入特征进行注意力运算,并结合DNN自适应地调整特征的权重,确保选取的关键特征具有最高的代表性,因此文中方法不存在已选特征无法更改的情况。同时,文中方法具有较高的数据拟合能力,选出的关键特征在4种分类器上具有最高的Acc和最低的FPR,最大程度地提升了模型的性能。

4.5 关键特征抗噪声能力分析

实际工程中,采用同步相量测量单元(phasor measurement unit, PMU)对系统特征进行监测。由于采样过程中可能存在误差和干扰,根据IEEE电力系统同步相量数据传输标准,PMU测量误差应小于1%^[32]。为观察所选择的关键特征在测量误差情况下的表现,在测试集中添加1%的随机误差,根据表5中的结果,使用准确率相对较高的RF分类器进行评价,特征子集评价结果如表6所示。

在1%噪声情况下,使用各种特征选择方法筛选的关键特征子集通过RF分类器进行分类,分类结果相较于无噪声时更差。各方法的Acc分别降低了0.008 0、0.007 6、0.007 7、0.003 3,而FPR则上升了0.006 7、0.005 4、0.003 3、0.003 3。其中,文中方

表6 噪声情况特征子集评价结果

Table 6 Evaluation results of feature subsets in the case of noise

方法	无噪声		1%噪声	
	A_{Acc}	F_{FPR}	A_{Acc}	F_{FPR}
MIM	0.947 7	0.048 0	0.939 7	0.054 7
JMI	0.980 3	0.027 3	0.972 7	0.032 7
RF-FSS	0.980 7	0.026 7	0.973 0	0.030 0
文中方法	0.984 0	0.024 7	0.980 7	0.028 0

法所选的关键特征子集在该噪声环境下的Acc只有轻微变化,仅降低了0.003 3,这表明文中方法具有最稳定的分类准确性,即使某些关键特征存在误差,其他特征仍能弥补并保持分类的准确率,证明了文中方法的有效性和鲁棒性。

4.6 IEEE 118 节点系统

为进一步验证文中方法在大规模系统上的有效性,文中利用IEEE 118节点系统对文中所提融合MHSA与Boruta的暂态功角稳定特征选择方法进行验证。

首先基于4.1节中的仿真设置利用PSAT工具箱进行仿真获取样本。故障设置为线路bus64—bus63首段,同时获取系统各节点电压幅值、相角和线路有功、无功功率特征作为潮流特征量,共构成608维备选特征。

采用文中方法进行训练并筛选关键特征,经过200次训练,最终筛选出35个关键特征。参考4.4节评价所筛选的特征的质量,并与传统特征选择方法进行对比,各特征选择方法的对比结果见表7。

表7 IEEE 118节点系统下不同特征选择方法结果对比

Table 7 Comparison of results for different feature selection methods on the IEEE 118-node system

方法	RF		LGBM	
	A_{Acc}	F_{FPR}	A_{Acc}	F_{FPR}
MIM	0.896 1	0.104 4	0.893 9	0.105 6
JMI	0.897 8	0.098 9	0.895 6	0.100 0
RF-FSS	0.899 4	0.104 4	0.896 7	0.104 4
文中方法	0.911 1	0.085 6	0.901 1	0.102 2
方法	BG		DT	
	A_{Acc}	F_{FPR}	A_{Acc}	F_{FPR}
MIM	0.895 0	0.110 0	0.873 9	0.133 3
JMI	0.895 6	0.104 4	0.877 2	0.121 1
RF-FSS	0.900 0	0.103 3	0.887 8	0.108 9
文中方法	0.910 0	0.094 4	0.877 8	0.117 8

从表7可以观察到,在RF、LGBM和BG模型上,文中方法在评价结果上相较于其他特征选择方法表现更出色。在DT模型中,虽然包裹式方法在

特征选择中以评价结果为标准,具有较高的准确率,但与 RF-FSS 相比,其关键特征数量高达 608 维,因此计算时间较长。总体而言,文中提出的方法在多数模型上展现出较好的评价结果。

5 结论

针对电力系统暂态稳定关键特征筛选中初始特征的选定限制寻找最佳关键特征组合,以及关键特征数量的确定方式存在主观性的问题,文中提出了融合 MHSA 与 Boruta 的暂态稳定关键特征筛选方法,将注意力机制和输入特征直接相连,配合 Boruta 算法可动态、客观地选取关键特征。以 IEEE 39 和 IEEE 118 节点系统为例进行算例分析,得出以下结论:

(1) 文中提出的 MHSA 特征选择模型可以自适应调整特征权重,相比于过滤式和包裹式方法,能够更加灵活地选出与暂态稳定相关的关键特征。

(2) 融合 Boruta 算法后,在训练过程中可以根据虚假特征权重的最大值自适应确定关键特征的数量。此外,虚假特征是没有实际意义的特征,经过虚假特征权重阈值的筛选,可以选出真实特征中对暂态稳定性起关键作用的特征。

(3) 文中方法可以选出合理的关键特征组合。由于特征之间的互补性,在数据存在测量误差的情况下,其暂态稳定判别结果仍能保持较高的水平。

文中选取的暂态稳定关键特征建立于一种预想故障的情况,即一种故障位置对应一组关键特征,实际应用时需要大量计算获取每种预想故障下的特征。因此,为选出适应不同故障下的特征,还须进一步研究筛选可应对电网拓扑变化的关键特征。

致 谢

本文得到智能电网保护和运行控制国家重点实验室项目(SGTYHT/20-JS-221),江苏省研究生科研与实践创新计划项目(SJCX23_1212)资助,谨此致谢!

参考文献:

- [1] 张诗熠,王怀远,李剑,等. 基于失稳模式自适应捕捉的暂态稳定评估方法[J]. 电力系统保护与控制,2024,52(18):35-44.
ZHANG Shiyi,WANG Huaiyuan,LI Jian,et al. Transient stability assessment method based on adaptive capture of instability patterns[J]. Power System Protection and Control,2024,52(18):35-44.
- [2] 张津铭,赵健,唐维溢,等. 动态拓扑下基于终端数据挖掘的配电网异常节点检测[J]. 智慧电力,2024,52(7):1-9.

- ZHANG Jinming,ZHAO Jian,TANG Weiyi,et al. Abnormal node detection in distribution network based on terminal data mining in dynamic topology[J]. Smart Power,2024,52(7):1-9.
- [3] 郭剑,朱炳铨,徐泰山,等. 基于历史数据聚类分析的暂态功角稳定故障筛选[J]. 电力工程技术,2020,39(2):75-80.
GUO Jian,ZHU Bingquan,XU Taishan,et al. Transient power angle stability contingency screening based on clustering analysis of historical data[J]. Electric Power Engineering Technology,2020,39(2):75-80.
- [4] 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.
- [5] 刘颂凯,胡竞哲,杨超,等. 基于改进 CDBN 的电力系统暂态稳定评估[J]. 智慧电力,2023,51(6):8-14,92.
LIU Songkai,HU Jingzhe,YANG Chao,et al. Transient stability assessment of power systems based on improved CDBN[J]. Smart Power,2023,51(6):8-14,92.
- [6] 高晗,蔡国伟,杨德友,等. 基于累积贡献率和可解释人工智能的静态电压稳定裕度估计特征量筛选方法[J]. 电力自动化设备,2023,43(4):168-176.
GAO Han,CAI Guowei,YANG Deyou,et al. Feature selection approach based on FCC-eAI in static voltage stability margin estimation[J]. Electric Power Automation Equipment,2023,43(4):168-176.
- [7] 李鹏,董鑫剑,孟庆伟,等. 基于 Fisher Score 特征选择的电力系统暂态稳定评估方法[J]. 电力自动化设备,2023,43(7):117-123.
LI Peng,DONG Xinjian,MENG Qingwei,et al. Transient stability assessment method for power system based on Fisher Score feature selection[J]. Electric Power Automation Equipment,2023,43(7):117-123.
- [8] 胡润滋,马晓忱,孙博,等. 基于特征选择的暂态安全评估方法及其可解释性研究[J]. 电网技术,2023,47(2):755-763.
HU Runzi,MA Xiaochen,SUN Bo,et al. Transient safety assessment and its interpretability based on feature selection[J]. Power System Technology,2023,47(2):755-763.
- [9] 姜新雄,徐箭,廖思阳,等. 基于增强进化计算的电力系统暂态稳定关键特征智能选择方法[J]. 中国电机工程学报,2023,43(14):5358-5372.
JIANG Xinxiong,XU Jian,LIAO Siyang,et al. An intelligent selection method for power system transient security key features based on enhanced evolutionary computation[J]. Proceedings of the CSEE,2023,43(14):5358-5372.
- [10] 吴双,胡伟,张林,等. 基于 AI 技术的电网关键稳定特征智能选择方法[J]. 中国电机工程学报,2019,39(1):14-21,316.

- WU Shuang, HU Wei, ZHANG Lin, et al. An intelligent key feature selection method of power grid based on artificial intelligence technology [J]. *Proceedings of the CSEE*, 2019, 39(1):14-21, 316.
- [11] 徐遐龄, 胡伟, 王春明, 等. 考虑特征组合效应的电网关键稳定特征筛选方法研究 [J]. *中国电机工程学报*, 2018, 38(8):2232-2238, 2533.
- XU Xialing, HU Wei, WANG Chunming, et al. Research on power systems key feature selection based on combination effect considering the stability rule [J]. *Proceedings of the CSEE*, 2018, 38(8):2232-2238, 2533.
- [12] 甄永赞, 阮程. 基于强化学习的混合元启发式暂态电压稳定特征选择方法及可解释性研究 [J]. *电网技术*, 2024, 48(4):1519-1532.
- ZHEN Yongzan, RUAN Cheng. Reinforcement learning-based hybrid element heuristic transient voltage stability feature selection and its interpretability [J]. *Power System Technology*, 2024, 48(4):1519-1532.
- [13] KHAIRE U M, DHANALAKSHMI R. Stability of feature selection algorithm: a review [J]. *Journal of King Saud University-Computer and Information Sciences*, 2022, 34(4):1060-1073.
- [14] 邓威, 郭钺秀, 李勇, 等. 基于特征选择和 Stacking 集成学习的配电网网损预测 [J]. *电力系统保护与控制*, 2020, 48(15):108-115.
- DENG Wei, GUO Yixiu, LI Yong, et al. Power losses prediction based on feature selection and Stacking integrated learning [J]. *Power System Protection and Control*, 2020, 48(15):108-115.
- [15] ZHANG B S, LI Y Y, CHAI Z. A novel random multi-subspace based ReliefF for feature selection [J]. *Knowledge-Based Systems*, 2022, 252:109400.
- [16] SUN L H, FU S, WANG F. Decision tree SVM model with Fisher feature selection for speech emotion recognition [J]. *EURASIP Journal on Audio, Speech, and Music Processing*, 2019, 2019(1):2.
- [17] 栾开宁, 杨世海, 黄艺璇, 等. 基于 Fisher-SVM 特征选择的负荷辨识研究 [J]. *电力科学与技术学报*, 2023, 38(4):230-239, 264.
- LUAN Kaining, YANG Shihai, HUANG Yixuan, et al. Research on Fisher-SVM feature selection based load identification [J]. *Journal of Electric Power Science and Technology*, 2023, 38(4):230-239, 264.
- [18] 毛莺池, 曹海, 平萍, 等. 基于最大联合条件互信息的特征选择 [J]. *计算机应用*, 2019, 39(3):734-741.
- MAO Yingchi, CAO Hai, PING Ping, et al. Feature selection based on maximum conditional and joint mutual information [J]. *Journal of Computer Applications*, 2019, 39(3):734-741.
- [19] LIU J, SUN H W, LI Y T, et al. An improved power system transient stability prediction model based on mRMR feature selection and WTA ensemble learning [J]. *Applied Sciences*, 2020, 10(7):2255.
- [20] BERMEJO P, GÁMEZ J A, PUERTA J M. Adapting the CMIM algorithm for multilabel feature selection. A comparison with existing methods [J]. *Expert Systems*, 2018, 35(1):e12230.
- [21] SALEM O A M, LIU F, CHEN Y P P, et al. Feature selection and threshold method based on fuzzy joint mutual information [J]. *International Journal of Approximate Reasoning*, 2021, 132:107-126.
- [22] CHEN K, ZHOU F Y, YUAN X F. Hybrid particle swarm optimization with spiral-shaped mechanism for feature selection [J]. *Expert Systems with Applications*, 2019, 128:140-156.
- [23] 谢彦祥, 刘天琪, 苏学能. Hadoop 架构下基于分布式粒子群算法的暂态稳定评估特征量选择 [J]. *电网技术*, 2018, 42(12):4107-4115.
- XIE Yanxiang, LIU Tianqi, SU Xueneng. Feature selection for transient stability assessment applying distributed PSO algorithm based on Hadoop architecture [J]. *Power System Technology*, 2018, 42(12):4107-4115.
- [24] TIAN J L, JIANG Y C, ZHANG J S, et al. High-performance fault classification based on feature importance ranking-XgBoost approach with feature selection of redundant sensor data [J]. *Current Chinese Science*, 2022, 2(3):243-251.
- [25] ZHANG H Q, WANG J, SUN Z Q, et al. Feature selection for neural networks using group lasso regularization [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2020, 32(4):659-673.
- [26] 梁露, 张智晟. 基于多尺度特征增强 DHTCN 的电力系统短期负荷预测研究 [J]. *电力系统保护与控制*, 2023, 51(10):172-179.
- LIANG Lu, ZHANG Zhisheng. Short-term load forecasting of a power system based on multi-scale feature enhanced DHTCN [J]. *Power System Protection and Control*, 2023, 51(10):172-179.
- [27] YAN C R, CHEN Y Z, WAN Y Q, et al. Modeling low- and high-order feature interactions with FM and self-attention network [J]. *Applied Intelligence*, 2021, 51(6):3189-3201.
- [28] SUBBIAH S, ANBANANTHEN K S M, THANGARAJ S, et al. Intrusion detection technique in wireless sensor network using grid search random forest with Boruta feature selection algorithm [J]. *Journal of Communications and Networks*, 2022, 24(2):264-273.
- [29] 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.
- [30] 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.

[31] 周挺,杨军,詹祥澎,等. 一种数据驱动的暂态电压稳定评估方法及其可解释性研究[J]. 电网技术,2021,45(11):4416-4425.

ZHOU Ting, YANG Jun, ZHAN Xiangpeng, et al. Data-driven method and interpretability analysis for transient voltage stability assessment[J]. Power System Technology, 2021, 45(11): 4416-4425.

[32] 张亮,安军,周毅博. 基于时间卷积和图注意力网络的电力系统暂态稳定评估[J]. 电力系统自动化,2023,47(7):114-122.

ZHANG Liang, AN Jun, ZHOU Yibo. Transient stability assessment of power system based on temporal convolution and graph

attention network[J]. Automation of Electric Power Systems, 2023,47(7):114-122.

作者简介:



王曼

王曼(1977),女,硕士,副教授,研究方向为电力系统运行与规划、人工智能在电力系统中的应用等(E-mail:njwm@njit.edu.cn);

周小雨(1996),男,硕士,研究方向为人工智能在电力系统中的应用;

陈凡(1981),女,博士,教授,研究方向为电力系统规划与可靠性分析、新能源接入等。

Fusion of MHSA and Boruta for key feature selection in power system transient angle stability

WANG Man¹, ZHOU Xiaoyu¹, CHEN Fan^{1,2}, LAI Yening², ZHU Ying³

(1. School of Electric Power Engineering, Nanjing Institute of Technology, Nanjing 211167, China;

2. State Key Laboratory of Smart Grid Protection and Control, Nanjing 211106, China;

3. College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China)

Abstract: In response to challenges posed by existing transient stability feature selection methods, which often encounter limitations in searching for the optimum combination of critical features and lack an objective criterion for determining the optimal number of key features, this paper introduces a novel approach. A transient power angle stability key feature selection method that seamlessly integrates multi-head self-attention (MHSA) and the Boruta algorithm. A deep neural network (DNN) with an MHSA model is initially constructed to execute transient stability assessments directly on the input grid features. The model dynamically adjusts attention weights during training, focusing on key features. Subsequently, the Boruta algorithm is employed to determine the number of key features. It generates a combination of real and virtual features, which the MHSA model trains to select the actual features that are higher than the maximum virtual feature weight, and the model autonomously determines the optimal number of key features. An analysis is conducted on the IEEE 39-node and 118-node systems to validate the proposed method. The results demonstrate that this approach ensures evaluation accuracy while significantly reducing the number of input features. Moreover, the key features identified exhibit higher evaluation accuracy than traditional methods.

Keywords: multi-head self-attention (MHSA); Boruta algorithm; transient stability; feature selection; key features; virtual features

(编辑 陆海霞)