

DOI: 10.19666/j.rlfed.202402030

# 基于 Powershap 特征选择的电力系统 暂态稳定评估

陈超<sup>1,2</sup>, 余成波<sup>1,2</sup>, 左立昕<sup>1,2</sup>

(1.重庆理工大学电气与电子工程学院, 重庆 400054;

2.重庆市能源互联网工程技术研究中心, 重庆 400054)

**[摘要]** 为进一步提高暂态稳定评估 (transient stability assessment, TSA) 的精准度和可靠性, 提出一种基于统计学与 Shapley 值结合的特征选择方法 (Powershap), 并建立电力系统 TSA 模型。首先, 根据电力系统运行时的稳态分量构建输入特征集, 采用 Powershap 将数据集分为多个数据子集进行训练, 筛选出关键特征集; 其次, 利用关键特征集训练多个 CatBoost 模型并进行 TSA, 生成 TSA 模型; 最后, 在新英格兰 10 机 39 节点系统和加入新能源发电的新英格兰 54 机 118 节点系统上进行仿真实验, 并给出评估结果。实验得出: 在新英格兰 10 机 39 节点系统中采用基于 Powershap 特征选择的方法进行分类, 其准确率能够达到 99.79%; 在改进的新英格兰 54 机 118 节点系统上, 其准确率能够达到 99.49%, 说明该方法能够有效进行电力系统暂态稳定评估, 并且验证了所提 TSA 模型具有较好的鲁棒性与泛化能力。

**[关键词]** 电力系统; 暂态稳定评估; 特征选择; Powershap; CatBoost

**[引用本文格式]** 陈超, 余成波, 左立昕. 基于 Powershap 特征选择的电力系统暂态稳定评估[J]. 热力发电, 2024, 53(8): 143-151. CHEN Chao, YU Chengbo, ZUO Lixin. Power system transient stability assessment based on Powershap feature selection[J]. Thermal Power Generation, 2024, 53(8): 143-151.

## Power system transient stability assessment based on Powershap feature selection

CHEN Chao<sup>1,2</sup>, YU Chengbo<sup>1,2</sup>, ZUO Lixin<sup>1,2</sup>

(1.School of Electrical and Electronic Engineering, Chongqing University of Technology, Chongqing 400054, China;

2.Chongqing Energy Internet Engineering Technology Research Center, Chongqing 400054, China)

**Abstract:** To further improve the accuracy and reliability of transient stability assessment (TSA), a feature selection method (Powershap) based on the combination of statistics and Shapley values is proposed, and a power system transient stability assessment model is established. Firstly, the input feature set is constructed based on the steady-state components during the operation of the power system. Powershap is used to divide the dataset into multiple subsets for training, and key feature sets are selected. Then, multiple CatBoost models are trained using key feature sets and transient stability assessments are conducted to generate transient stability assessment models. Finally, simulation experiments are conducted on the New England 10-machine 39-node system and the New England 54-machine 118-node system with the addition of new energy generation, and evaluation results are provided. The experiments show that, in the 10-machine 39-node system in New England, using the Powershap feature selection method for classification can achieve an accuracy of 99.79%. On the improved New England 54-machine 118-node system, its accuracy can reach 99.49%, indicating that the method can effectively perform

收稿日期: 2024-02-26

基金项目: 重庆市自然科学基金创新发展联合基金 (2023CCZ082); 重庆市教委科研基金 (2023CYJH009)

Supported by: Chongqing Natural Science Foundation Innovation and Development Joint Fund (2023CCZ082); Research Fund of Chongqing Municipal Education Commission (2023CYJH009)

第一作者简介: 陈超 (1997), 男, 硕士研究生, 主要研究方向为电力系统智能运行与控制, R315CC@stu.cqut.edu.cn.

通信作者简介: 余成波 (1965), 男, 博士, 教授, 主要研究方向为信息获取与处理技术、远程测试与控制技术 (无线传感网络) 及电气设备物联网技术, yuchengbo@cqut.edu.cn.

transient stability assessment of power systems. It is verified that the proposed TSA model has good robustness and generalization ability.

**Key words:** power system; transient stability assessment; feature selection; Powershap; CatBoost

为实现高质量可持续发展,我国政府提出“双碳”发展的重大战略目标,以应对资源与环境压力。随着电力电子技术和新能源领域的发展,我国电网中接入大量电力电子与新能源设备的趋势日益普遍,这使得电网系统变得更加复杂<sup>[1]</sup>。因此,能够快速、精确地对电力系统暂态稳定进行评估,对于电力系统稳定运行具有重要意义。

电力系统暂态稳定评估(transient stability assessment, TSA)的常见方法有时域仿真法、直接法<sup>[2]</sup>和机器学习法<sup>[3]</sup>。其中,时域仿真法能够比较真实地模拟暂态响应过程,但计算复杂度较大,无法满足实时评估应用的需求。直接法计算速度较快,但对于我国目前大规模的电力网络系统,需要耗费大量计算资源并且难以构造精确的能量函数,因此该方法适用性较差。随着国内智能电网的发展和同步相量测量单元(phasor measurement unit, PMU)的广泛应用,机器学习法因其计算速度快、泛化能力强和评估精度高等优势,已然成为 TSA 研究中的热点领域。

TSA 研究中常见的机器学习模型有循环神经网络(RNN)<sup>[4-5]</sup>、支持向量机(SVM)<sup>[6-7]</sup>、随机森林(RF)<sup>[8]</sup>、决策树(DT)<sup>[9-10]</sup>等。文献[11]将聚类思想融入人工神经网络中,通过调整参数提升电力系统 TSA 的准确度,但该方法并不具有普适性。文献[12]采用核支持向量机进行电力系统 TSA,结果有更好的评估精度,对训练时长和内存占用较为友好。文献[13]选择所有发电机的功角数据作为输入特征数据集,实验结果表明在部分发电机数据缺失的情形下,所提方法能保持较高的准确度,但没有考虑节点上丢失数据的情况。文献[14]使用一种 TCN-GAT 的方法对电力系统暂态进行评估,对模型的损失函数进行了改进,减少了模型对失稳样本误判分类的数量,所得准确率较高。文献[15]提出使用 Fisher score 特征选择法,对电力系统暂态稳定进行评估,能够有效压缩数据特征集的规模,减少数据特征中包含的冗余信息和噪声干扰,但模型计算效率和准确度有待进一步提高。文献[16]提出一种 KTBoost 算法对特征关系进行探索,进而对电力系统暂态稳定裕度进行评估,其评估结果相较于部分传统模

型更为精确。

在现有研究中,特征选择方法主要有过滤法、包装法、嵌入法 3 种<sup>[17]</sup>。其中,过滤法计算效率高,可以快速准确筛选出与目标变量相关性较高的特征,但同时可能会忽略特征与目标变量之间的复杂关系;包装法考虑了特征与学习器之间的相互关系,因此能够更精准地评估特征的重要性,但容易产生过拟合且在特征维度较高的情况下计算速度较慢;嵌入法结合了过滤法和包装法的优点,计算速度快,但对于不同算法需要设计对应的嵌入方法,普适性较低。

Powershap 是一种基于统计学假设检验和 Shapley 值结合的包装器特征选择方法。Powershap 由解释组件和核心组件 2 个部分组成,其核心假设是与已知的随机特征相比,有效信息特征对预测结果的影响更大。根据 Powershap 的基准测试和仿真结果得知, Powershap 优于其他过滤法,预测性能与包装法相当,但其运行速度更快。因此 Powershap 在处理高维特征数据时能够快速直观地进行特征选择<sup>[18]</sup>。

基于此,本文提出一种基于 Powershap 特征选择的电力系统 TSA 方法。首先使用 Powershap 进行特征选择;然后使用多层 CatBoost<sup>[19]</sup>作为分类器,同时与其他分类器及特征选择方法进行了对比;最后,通过 PSASP 软件搭建 IEEE-39 节点和改进 IEEE-118 节点系统进行仿真实验,同时对模型进行鲁棒性测试和泛化能力测试,验证了该模型的精准度和可靠性。

## 1 电力系统暂态稳定评估模型

### 1.1 Powershap 算法原理

Powershap 算法由解释组件和核心组件 2 个部分组成。

1) 在解释组件中使用不同的随机种子,在不同的数据子集上训练多个模型且每个子集都由所有原始特征和 1 个随机特征组成。模型训练完成后,在样本外数据集上使用 Shapley 值解释特征的平均影响。这种做法可以减少特征集的规模,降低时间复杂度。使用 Shapley 值在样本外数据子集上量化每个特征对输出的影响,以评估真正的无偏影响。

最后, 取所有 Shapley 值的绝对值并求平均值, 以得到每个特征的总平均影响。

2) 在核心组件中, 把得到每个特征的总平均影响与核心组件中随机特征的影响进行比较量化, 从而筛选出所有含有信息量的特征。

$$P_{(v,x)} = \sum_i^n \frac{T(x > v_i)}{n} \quad (1)$$

式中:  $v$  为单个特征的 Shapley 平均值数组, 其迭代次数和长度相同;  $x$  为单个值;  $i$  为迭代次数;  $v_i$  为第  $i$  次迭代下, 单个特征的 Shapley 平均值数组;  $n$  为  $n$  组数据;  $T$  为指示函数;  $p$  为  $x$  值高于该次迭代的平均 Shap 值的百分比。

Powershap 算法需要设置  $\alpha$  和  $i$  2 个超参数,  $\alpha$  为  $p$  值阈值,  $i$  为迭代次数。测试的统计功效为  $1-\beta$ , 其中  $\beta$  是假阴性 FN 的概率。如果对测试样本的统计检验输出  $p$  值  $\alpha$ , 则表示在当前数据条件下, 测试样本可能被标记为显著的概率。如果统计检验中的数据较少,  $\alpha$  值可能很低, 但  $\beta$  值可能很大, 进而导致输出结果不可信。因此, 对于给定的  $\alpha$ , 统计功效应尽可能接近 1, 以避免假阴性 FN 出现。统计检验  $\alpha$  的功效可以使用基础检验分布  $H_1$  的累积分布函数  $F$  进行计算:

$$\alpha(x) = F_{H_0}(x) \quad (2)$$

$$\text{Power}(\alpha) = F_{H_1}(F_{H_0}^{-1}(\alpha)) \quad (3)$$

式中:  $H_0$  为随机特征影响分布;  $H_1$  为经测试的特征影响分布。

为了避免用户手动调整超参数, Powershap 提供了自动模式, 这种自动模式通过计算  $\alpha$  的统计功效, 自动确定和优化迭代超参数  $i$ 。本文使用 Powershap 自动模式来寻找最优参数, 该模式下默认首先执行 10 次迭代计算, 对于每个特征, Powershap 使用 student- $t$  分布进行功效检验, 进而计算效应量和统计功效。接着利用计算出的效应量, 计算出达到预定功效要求所需的迭代次数。 $\alpha$  值设定为 0.01, 如果所需的迭代次数大于已执行的迭代次数, Powershap 将继续执行额外的迭代次数。之后, Powershap 会重新计算所需的迭代次数, 并不断重新执行, 直至达到预定功效所需迭代次数。

## 1.2 CatBoost 模型

CatBoost 模型基于 GBDT 模型改进而来, 是梯度提升框架的一个高效实现形式, 其采用对称树作为基学习器, 基本原理如下。

在给定一个含有  $m$  个样本的训练集中,

$x_i = (x_i^1, x_i^2, x_i^3 \cdots x_i^n)$  代表含有  $n$  维特征的第  $i$  个数据;  $y_i$  代表第  $i$  个数据映射的标签值, 其中, 标签值为 0 表示系统稳定, 标签值为 1 表示系统不稳定。

使用 CART 初始化构建一个弱学习器  $f_1(x_i)$ , 并定义迭代数  $s \in [2, S]$ ,  $S$  表示总迭代数。在每一轮中, 通过最小化损失函数来训练从 CART 集合空间  $H$  中选择的最优弱学习器, 该损失函数基于前一轮输出的强学习器  $L(y_i, f_{s-1}(x_i))$ 。

$$h_s = \underset{H}{\operatorname{argmin}} \sum_{i=1}^m L(y_i, f_{s-1}(x_i) + h_s(x_i)) \quad (4)$$

式中:  $h_s(x_i)$  为输入为  $x_i$  时, 弱学习器  $f_s(x_i)$  的输出值。

第  $s$  次迭代的第  $i$  个样本损失函数的负梯度为:

$$g_s^i = -\frac{\partial L(y_i, f_{s-1}(x_i))}{\partial f_{s-1}(x_i)} \quad (5)$$

本轮弱学习器经过数据集  $(x_i, g_s^i)$  拟合得出, 它也表示第  $s$  棵 CART。对于任意叶节点  $Q_s^j (j=1, 2, 3 \cdots J)$  的数据样本, 其中  $J$  表示叶节点数, 经拟合后, 叶节点最佳输出值为  $z_s^j$ :

$$z_s^j = \underset{z}{\operatorname{argmin}} \sum_{x_i \in Q_s^j} L(y_i, f_{s-1}(x_i) + z) \quad (6)$$

全部叶节点的最佳输出可由式(6)得出,  $z$  表示叶节点的输出值, 进一步得出经历  $s$  次迭代后, CART 拟合函数为:

$$h_s(x_i) = \sum_{j=1}^J z_s^j T \quad (7)$$

式中:  $T$  表示指示函数;  $x_i \in Q_s^j$ 。因此, 第  $s$  次迭代完成后的强学习器为:

$$f_s(x_i) = f_{s-1}(x_i) + \sum_{j=1}^J c_s^j T \quad (8)$$

CatBoost 能够自动处理类别特征, 无需进行编码或转换。此外, CatBoost 通过引入随机性和正则化技术来降低过拟合的风险, 并使用随机负采样来平衡类别样本, 并利用对称树结构和对特征的随机排序来减少过拟合, 解决了预测偏移和梯度偏差的问题, 进而保证了 TSA 模型的可靠性与时效性。

## 1.3 电力系统 TSA 模型建立

本文采用基于 Powershap 特征选择方法并使用 CatBoost 作为分类器, 进一步提出 TSA 模型: Powershap 将数据集分成多个数据子集进行训练, 进而得出各个特征的 Shapely 值和特征贡献度排名。下一步将各个数据子集输入多层 CatBoost 模型进行训练, 最后取平均值作为评估结果, 其模

型如图 1 所示。

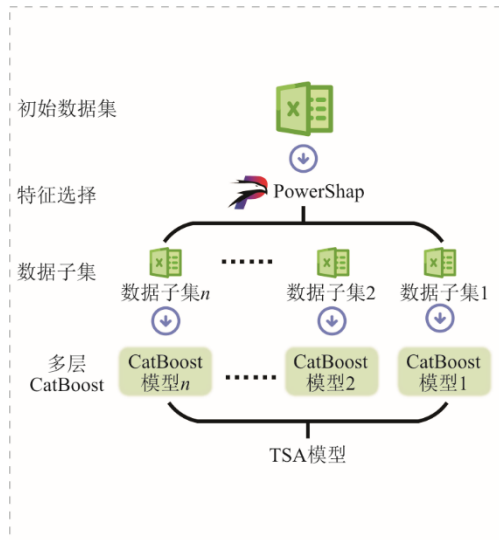


图 1 TSA 模型  
Fig.1 The TSA model

## 2 暂态稳定评估方案

### 2.1 暂态稳定评估的问题描述

电力系统 TSA 实质上为二分类问题,通过系统受扰动后,各发电机功角的暂态稳定指数 (transient stability index, TSI) 对数据进行标签:

$$I_{TSI} = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|} \quad (9)$$

式中:  $\Delta\delta_{\max}$  为仿真结束后任意 2 台发电机工作时的最大功角差。若  $I_{TSI} > 0$ , 则判定系统处于稳定状态; 若  $I_{TSI} \leq 0$ , 则判定系统失稳。样本集数据为模型训练的输入, 其中可分为  $X$  和  $Y$  两大类数据:

$$X = [x_1, x_2 \cdots x_m] \quad (10)$$

$$Y = [y_1, y_2 \cdots y_m]^T \quad (11)$$

式中:  $X$  为一组包含  $m$  组  $n$  维的数据集合;  $n$  为特征数;  $Y$  为一组对应  $X$  的标签数据集, 当  $Y=1$  时, 标记样本为稳定;  $Y=0$  时, 标记样本为失稳。

### 2.2 暂态稳定评估的问题描述

对于机器学习方法的 TSA, 其输入特征选取对模型评估性能的优劣有较大影响, 因此选取一组合适的输入特征集非常关键。本文选取的特征量均可由 PMU 直接量测得来<sup>[20]</sup>, 进而可以保证实际中数据来源的准确性与实时性。为保证数据较为完整, 数据集由故障前、故障时、故障后的样本所得, 选取特征见表 1。

表 1 输入特征集

Tab.1 The input feature set

序号	输入特征	序号	输入特征
1	母线电压幅值	11	负荷端电流
2	发电机功角	12	交流传输线 $i$ 侧电压
3	发电机端电压	13	交流传输线 $j$ 侧电压
4	发电机端电流	14	交流传输线 $i$ 侧相角
5	发电机角速度	15	交流传输线 $j$ 侧相角
6	发电机有功功率	16	交流线传输线 $i-j$ 侧有功功率
7	发电机无功功率	17	交流线传输线 $j-i$ 侧有功功率
8	负荷有功功率	18	交流线传输线 $i-j$ 侧无功功率
9	负荷无功功率	19	交流线传输线 $j-i$ 侧无功功率
10	负荷端电压		

### 2.3 暂态稳定评估过程

#### 2.3.1 离线训练

本文提出的基于 Powershap 特征选择的 TSA 综合模型的离线训练具体步骤如下。

1) 通过 PSASP 仿真软件对搭建的 IEEE-39 节点、改进 IEEE-118 节点系统进行时域仿真, 进行潮流计算后在系统中设置三相短路故障, 计算出电力系统暂态稳定数据。

2) 将仿真数据进行整理, 并把数据打好稳定与不稳定标签; 使用五折交叉验证的方式, 将数据输入 Powershap 进行特征选择, 筛选出关键特征集, 并按照特征贡献度进行排序。

3) 将 Powershap 特征选择后的数据集按照 4:1 比例分为训练集和测试集, 训练多层 CatBoost 模型, 完成 TSA 综合模型的构建, 输出评价指标结果。

#### 2.3.2 在线评估

通过读取 PMU 量测的实时数据, 结合 Powershap 进行特征选取, 将结果输入进 TSA 模型, 并输出评估结果。如果判定结果为不稳定, 系统将提示工作人员对系统作出进一步操作, 避免发生大规模故障。

#### 2.3.3 模型更新

在实际情况下, 电网运行会受到各种因素影响, 如检修计划、紧急事故等情况会导致电网拓扑结构发生变化。随着天气季节变化, 负荷端需求也会随之变化。另外大规模新能源设备和分布式设备接入也会造成电能质量波动。因此, 进行 TSA 模型更新是非常必要的, 当电力系统运行工况发生显著变化时, PMU 将上传实时电力系统运行数据, 进一步重新选择特征, 最后训练模型并进行评估, TSA 流程框架如图 2 所示。

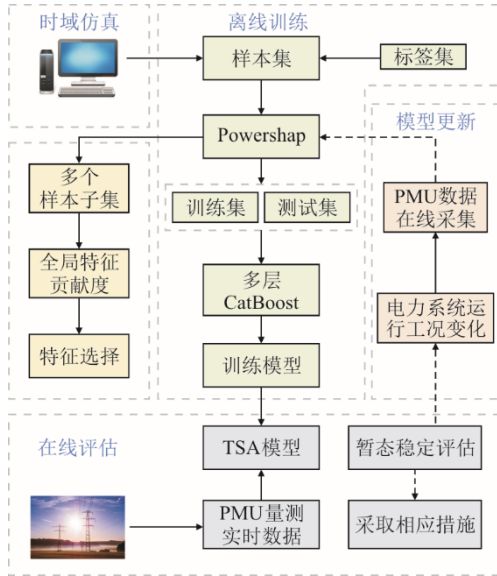


图 2 TSA 流程框架  
Fig.2 Framework of the TSA flow

### 2.4 模型评价指标

为做到 TSA 的全面性, 本文引入混淆矩阵, 采用准确率和  $F_1$  分数进行评价。准确率 ( $A_c$ ) 能够反映模型全局的正确率; 查准率 (precision) 代表预测失稳样本的比例; 召回率 (recall) 代表被正确预测的失稳样本占真实失稳样本的比例;  $F_1$  分数能更好地反映模型对失稳样本的预测能力。各类指标定义见表 2。

表 2 TSA 混淆矩阵  
Tab.2 Confusion matrix of TSA

评估模型	真实结果	
	稳定	失稳
稳定	TP	FP
失稳	FN	TN

$$A_c = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

$$\delta_{Precision} = \frac{TN}{TN + FN} \quad (13)$$

$$\delta_{Recall} = \frac{TN}{TN + FP} \quad (14)$$

$$F_1 = \frac{2 \times \delta_{Precision} \times \delta_{Recall}}{\delta_{Precision} + \delta_{Recall}} \quad (15)$$

## 3 算例分析

### 3.1 算例系统介绍

通过 PSASP 软件搭建新英格兰 10 机 39 节点系统。该系统中基准功率设置为 100 MV·A, 系统额

定频率为 50 Hz。系统中包括 10 台发电机、39 条母线以及 46 条线路。在仿真过程中, 模拟负荷以 5% 为步长, 在 70%~120% 之间增长变化, 共计 11 种负荷水平。为保证母线电压维持在 0.95~1.05 p.u. 范围内, 根据负荷水平相应调整发电机出力情况。在输电电路上设置三相短路故障, 故障位置设置在离输电线路首端的 1%、25%、50%、75%、99% 处。故障发生时刻为 1 s 时, 仿真的故障持续时间分别为 0.1、0.2、0.3 s, 最终生成 5 610 个样本, 其中稳定样本 3 448 个, 不稳定样本 2 162 个。

### 3.2 评估模型性能对比

为了选择一个性能更优的分类器, 将数据集分别输入 KNN<sup>[21]</sup>、DT、SVM、LightGBM<sup>[22]</sup> 模型进行评估, 并与 CatBoost 模型进行对比。其中 KNN 采用网格搜索确定  $k$  值为 10; SVM 采用网格搜索法设定最优惩罚程度  $C=10$ ; DT 与 LightGBM 同样采用网格搜索法寻找最优参数, DT 算法中设置最大深度为 5, 最大叶节点数为 10, 最小样本数为 2; LightGBM 的最优学习率为 0.1, 最大深度为 5, 分类器数目为 200; CatBoost 模型采用早停法确定最优学习率为 0.4, 最大深度为 6, 分类器数目为 50。上述模型均采用同一数据集且采用 5 折交叉验证, 评估结果如图 3 所示。同时将 Powershap 与其他特征选择方法进行了对比, 分类器均使用 CatBoost, 评估结果见表 3。

对于 KNN 模型, 算法虽然较为简单且易理解, 但在处理大规模数据集或高维数据时容易出现样本之间过于稀疏, 导致结果较差; 对于 SVM 模型, 在小规模系统上表现较好, 但应用到复杂的电力系统中将使计算开销变大; 对于 DT 模型, 当树的深度和分支较大时, 容易造成模型过拟合; 对于 LightGBM 模型, 该模型使用直方图进行特征分割, 在处理高维稀疏数据时直方图很难有效捕捉特征之间的关联。本文中使用的 CatBoost 作为分类器能够有效解决梯度偏差问题且有效性高。采用 Fisher 特征选择方法时所需时间虽然短, 但其对数据分布要求较高, 否则容易导致结果失真和特征选择错误; MIC<sup>[23]</sup> 特征选择方法在处理高维特征时会增加计算复杂度, 导致计算时间较长; 采用递归特征消除法<sup>[24]</sup> 的评估结果虽然较好, 但由于需要多次训练模型, 因此在特征数量较多时会导致计算时间长; LR 嵌入法<sup>[25]</sup> 易受样本不平衡影响, 会导致结果出现偏

差;采用基于 Powershap 进行特征选择的 TSA 模型兼具着特征选择时间较短和评估精度高的优点,因而有着更加优异的评估性能,能够适用于大规模电力系统。表 4 为经过 Powershap 特征选择后输出的贡献度排名前 10 的特征,贡献度越高表示该特征对系统稳定性影响越大。

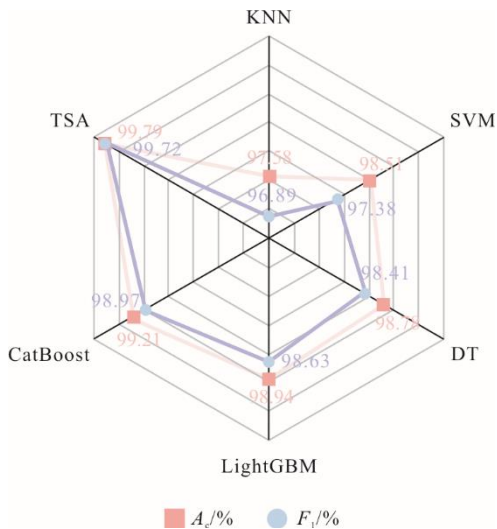


图 3 模型评估结果对比

Fig.3 Comparison of model evaluation results

表 3 特征选择方法对比

Tab.3 Comparison of feature selection methods

方法	Ac/%	F1/%	特征选择时间/s
Fisher	99.49	99.34	0.22
MIC	99.41	99.28	557.84
递归特征消除	99.64	99.53	4 133.88
LR 嵌入法	99.55	99.42	7.31
Powershap	99.79	99.72	70.69

表 4 贡献度前 10 排名

Tab.4 Top 10 contribution rankings

排名	特征名	贡献度
1	发电机_BUS-31_δ(deg.)	1.098 392 248
2	发电机_BUS-38_It(p.u.)	0.742 980 361
3	发电机_BUS-38_δ(deg.)	0.568 831 861
4	支路_AC2_Pij(p.u.)	0.541 546 583
5	支路_AC1_Pij(p.u.)	0.513 051 391
6	支路_AC1_Pji(p.u.)	0.501 971 364
7	支路_AC2_Pji(p.u.)	0.496 336 877
8	发电机_BUS-31_It(p.u.)	0.337 646 335
9	支路_AC14_Pij(p.u.)	0.252 183 586
10	发电机_BUS-38_ω(p.u.)	0.225 379 854

3.3 鲁棒性测试

在实际电网运行中,PMU 量测装置采集数据时可能会出现数据缺失或受到噪声干扰等情况,本文

考虑了 2 种方式对系统进行测试。结果表明 TSA 模型在受到不同噪声和丢失不同比例的特征数据的情况下均能保持较高的评估精度,表明本文所提方法相较于其他模型有着更高的可靠性和有效性,表现出较好的鲁棒性,这是由于采用基于 Powershap 的特征选择方法能够筛选出影响度较大的信息特征且过滤掉无关特征,进而降低了结果对特征数据量的依赖性,因此 TSA 模型能够适应电网运行的特殊情况。

1) 通过向初始数据集中加入 5 种信噪比的高斯白噪声,模拟实际数据受到干扰的情形。实验结果如图 4 所示。

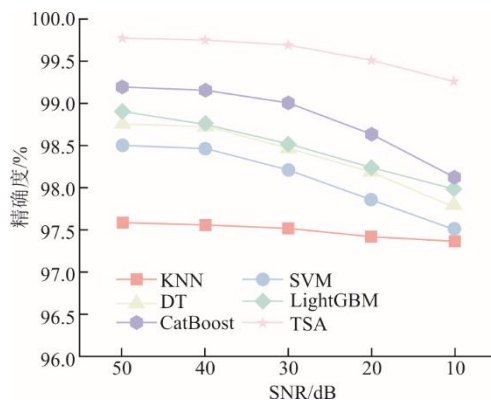


图 4 不同模型受噪后的精确度

Fig.4 Accuracy of different models after noise exposure

2) 通过 Powershap 特征选择并输出特征贡献度排名,从中删除影响度较小的 5 种不同比例的特征来模拟实际数据缺失的情况,实验结果如图 5。

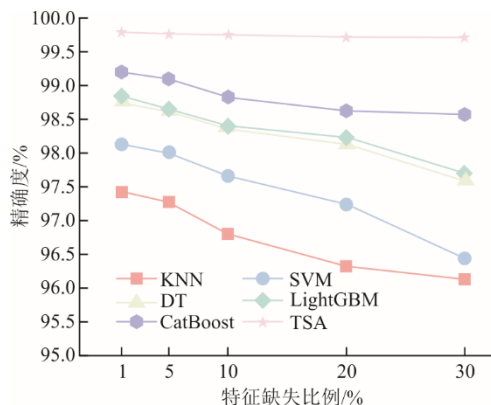


图 5 不同模型缺失特征的精确度

Fig.5 Accuracy of missing features for different models

3.4 泛化能力测试

在实际电网运行中,常常会出现电力系统故障以及日常检修维护等情形。为验证基于 Powershap

的 TSA 模型的泛化能力,对下面 6 种新拓扑场景进行测试,每一种拓扑结构的数据都根据第 3.1 节所述的方法,分别生成 500 组数据样本,并将对应数据样本分别输入 TSA 模型进行评估。表 5 为不同拓扑的评估结果。

表 5 不同拓扑的评估结果  
Tab.5 Evaluation results for different topologies

拓扑	拓扑结构	$A_c/\%$	$F_1/\%$
1	线路 9-39 断开	98.78	97.79
2	3 号发电机退出运行	99.60	98.89
3	线路 6-7 断开; 线路 25-26 断开	98.42	97.38
4	4 号发电机退出运行; 线路 23-24 断开	98.65	97.67
5	线路 2-3、线路 16-19、线路 28-29 断开	97.41	96.36
6	9 号发电机退出运行; 线路 4-5、线路 17-18 断开	97.63	96.55

由表 5 可知,不同拓扑结构一定程度上会影响模型的精确度  $A_c$  和  $F_1$  分数。在拓扑 5 与拓扑 6 结构中,模型精确度虽然下滑较大,但均能保持在可接受的评估结果内,进而说明本文 TSA 模型具有较好的泛化能力。

### 3.5 改进的新英格兰 118 节点系统

为进一步验证本文所提 TSA 方法在较大电网系统下的有效性,采用改进的新英格兰 118 节点系统进行测试。系统中基准功率设置为 100 MV·A,系统额定频率为 50 Hz,包含 118 条母线、179 条线路、9 台变压器、54 台发电机(其中节点 89、100、110 处的发电机用单台 1.5 MW 风力发电机组成的风电场代替,节点 66、104 处的发电机采用单台 1.5 MW 光伏阵列组成的光伏电厂代替)和 91 个负荷。在仿真过程中,模拟负荷以 5%为步长,在 80%~120%之间增长变化,共计 9 种负荷水平。为保证母线电压维持在 0.95~1.05 p.u.范围内,根据负荷水平相应调整发电机出力情况。在输电线路路上设置三相短路故障,故障位置设置在离输电线路首端的 1%、50%、99%处。故障发生时刻为 1 s 时,仿真的故障持续时间分别为 0.1、0.2、0.3 s,生成 9 558 个样本,其中稳定样本 7 236 个,不稳定样本 2 322 个。

#### 3.5.1 评估模型对比

采用 3.2 节所述方法确定参数,并对各个模型进行对比,给出相应的评估结果见表 6 与表 7。结果表明:在较大测试系统下,采用基于 Powershap 的 TSA 模型进行暂态稳定评估相比其他模型同样

能获得较高的精确度。

表 6 模型评估结果对比  
Tab.6 Comparison of model evaluation results

方法	$A_c/\%$	$F_1/\%$
KNN	97.35	95.53
SVM	97.72	95.81
DT	97.91	96.17
LGBM	98.13	96.82
CatBoost	98.56	96.99
Powershap	99.49	98.96

表 7 特征选择方法对比  
Tab.7 Comparison of feature selection methods

方法	$A_c/\%$	$F_1/\%$	特征选择时间/s
Fisher	98.10	97.12	0.35
MIC	98.05	97.01	1 640
递归特征消除	99.28	99.15	12 721
LR 嵌入法	98.24	97.33	19.86
Powershap	99.49	98.96	225.84

#### 3.5.2 鲁棒性测试

为进一步测试 TSA 模型在大系统下的鲁棒性,采用 3.3 节所述方法进行测试,图 6 与图 7 分别为在 118 节点系统所得数据集中添加 5 种不同噪声干扰和随机删除 5 种不同比例特征数据时的结果对比。由图 6 和图 7 可见,本文所提方法在大系统下,相较其他方法同样有着优异的评估结果。

#### 3.5.3 泛化能力测试

为进一步测试 TSA 模型在大系统下的泛化能力,对 6 种新拓扑场景进行测试,每种拓扑结构生成 500 组数据样本,随后将数据输入模型进行测试,表 8 为 118 节点系统的新拓扑结构测试结果。由表 8 可见,本文所提模型在较大系统拓扑结构改变时,其评估精度不会出现较大幅度的下降,表明本文所提模型具有较好的泛化能力。

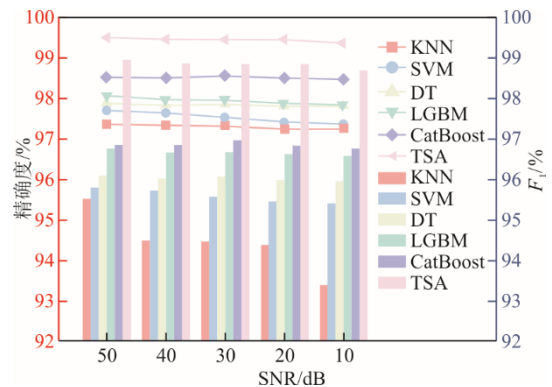


图 6 各模型受噪的精确度和  $F_1$  值  
Fig.6 The accuracy and  $F_1$  value of noise for each model

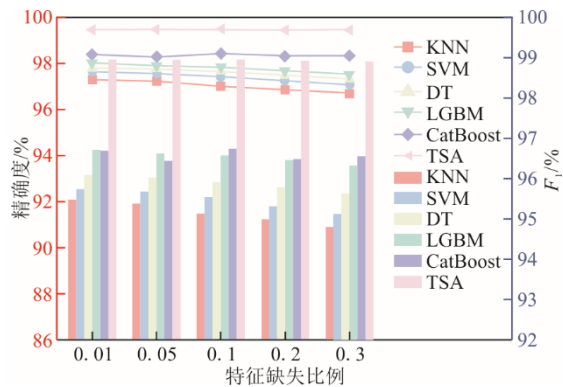


图7 缺失不同比例特征数据下的精确度和  $F_1$  值对比  
Fig.7 Comparison of accuracy and  $F_1$  value under missing feature data with different proportions

表 8 118 节点系统的新拓扑结构  
Tab.8 New topology structure of the 118-node system

拓扑	拓扑结构	$A_c$ /%	$F_1$ /%
1	线路 8-30、49-51 断开	98.27	97.14
2	7 号发电机退出运行；线路 15-33 断开	98.45	97.39
3	14 号 32 号发电机退出运行	97.16	96.12
4	线路 4-11、40-42、85-89 断开	97.24	96.15
5	29 号发电机退出运行；线路 3-12、49-54 断开	97.27	96.21
6	10 号、26 号发电机退出运行；线路 59-61 断开	96.78	95.63

## 4 结 论

针对电力系统暂态稳定性问题,本文提出了一种基于 Powershap 特征选择的电力系统 TSA,并在 IEEE-39 节点和改进的 IEEE-118 节点系统上进行了仿真验证,实验结果表明:

1) 基于 Powershap 特征选择方法能够有效挖掘数据集中各个特征之间的关系,并且可以依据 Shapley 值得出特征贡献度,方便后续能够有效地进行可解释性分析研究。将其与 Fisher、MIC、递归特征消除法以及 LR 嵌入法进行比较,本文所提特征选择方法评估结果更优。

2) 通过比较 KNN、SVM、DT、LightGBM、CatBoost 的评估结果,在经 Powershap 特征选择后使用 CatBoost 分类器能够使 TSA 模型更加优越。

3) 针对实际电网运行中可能遇到的噪声干扰、数据缺失以及拓扑结构改变的问题,本文所提 TSA 模型具有较好的鲁棒性和泛化能力。

本文所对比的模型部分是传统模型,不能完全体现出所提模型的优越性,如何采用改进模型加以对比以及基于 Powershap 的可解释性是下一步需要

研究的内容。

## [参 考 文 献]

- [1] 杜冬梅,曹冬惠,何青.“双碳”目标下我国电力行业低碳转型的思路探讨[J].热力发电,2022,51(10):1-9.  
DU Dongmei, CAO Donghui, HE Qing. Discussion on the low-carbon transformation of China's power industry under the "dual carbon" goal[J]. Thermal Power Generation, 2022, 51(10): 1-9.
- [2] 杨博,陈义军,姚伟,等.基于新一代人工智能技术的电力系统稳定评估与决策综述[J].电力系统自动化,2022,46(22):200-223.  
YANG Bo, CHEN Yijun, YAO Wei, et al. Review on stability assessment and decision for power systems based on new-generation artificial intelligence technology[J]. Automation of Electric Power Systems, 2022, 46(22): 200-223.
- [3] 赵冬梅,谢家康,王闯,等.基于 Bagging 集成学习的电力系统暂态稳定在线评估[J].电力系统保护与控制,2022,50(8):1-10.  
ZHAO Dongmei, XIE Jiakang, WANG Chuang, et al. On-line transient stability assessment of a power system based on Bagging ensemble learning[J]. Power System Protection and Control, 2022, 50(8): 1-10.
- [4] 张雅婷,刘颂凯,张磊,等.针对数据缺失的电力系统暂态稳定评估方法[J].电力系统及其自动化学报,2023,35(3):59-68.  
ZHANG Yating, LIU Songkai, ZHANG Lei, et al. Assessment method for power system transient stability with missing data[J]. Proceedings of the CSU-EPSA, 2023, 35(3): 59-68.
- [5] 谷广超,轩克辉.基于局部编码和多头注意力模型的电力系统暂态稳定性评估[J].计算机应用与软件,2023,40(8):114-120.  
GU Guangchao, XUAN Kehui. Transient stability assessment of power system based on partial encoding and multi-head attention[J]. Computer Applications and Software, 2023, 40(8): 114-120.
- [6] 田芳,周孝信,于之虹.基于支持向量机综合分类模型和关键样本集的电力系统暂态稳定评估[J].电力系统保护与控制,2017,45(22):1-8.  
TIAN Fang, ZHOU Xiaoxin, YU Zhihong. Power system transient stability assessment based on comprehensive SVM classification model and key sample set[J]. Power System Protection and Control, 2017, 45(22): 1-8.
- [7] 戴远航,陈磊,张玮灵,等.基于多支持向量机综合的电力系统暂态稳定评估[J].中国电机工程学报,2016,36(5):1173-1180.  
DAI Yuanhang, CHEN Lei, ZHANG Weiling, et al. Power system transient stability assessment based on multi-support vector machines[J]. Proceedings of the CSEE, 2016, 36(5): 1173-1180.
- [8] 武宇翔,韩肖清,牛哲文,等.基于变权重随机森林的暂态稳定评估方法及其可解释性分析[J].电力系统自动化,2023,47(14):93-104.  
WU Yuxiang, HAN Xiaoping, NIU Zhewen, et al. Transient stability assessment method based on variable weight random forest and its interpretability analysis[J]. Automation of Electric Power Systems, 2023, 47(14): 93-104.
- [9] WANG D, RUEDA J L, RAKHSHANI E, et al.

- MVMO-based identification of key input variables and design of decision trees for transient stability assessment in power systems with high penetration levels of wind power[J]. *Frontiers in Energy Research*, 2020, 8(41): 1-10.
- [10] GUEDES J D S, FERREIRA D D, BARBOSA B H G. A non-intrusive approach to classify electrical appliances based on higher-order statistics and genetic algorithm: a smart grid perspective[J]. *Electric Power Systems Research*, 2016, 140: 65-69.
- [11] 戴仁昶, 张伯明. 基于人工神经网络的暂态稳定性分析[J]. *电力系统自动化*, 2000, 24(12): 1-3.  
DAI Renchang, ZHANG Boming. Transient stability analysis based on ANN[J]. *Automation of Electric Power Systems*, 2000, 24(12): 1-3.
- [12] WANG B, FANG B, WANG Y, et al. Power system transient stability assessment based on big data and the core vector machine[J]. *IEEE Transactions on Smart Grid*, 2016, 7(5): 2561-2570.
- [13] 李宝琴, 吴俊勇, 邵美阳, 等. 基于集成深度置信网络的精细化电力系统暂态稳定评估[J]. *电力系统自动化*, 2020, 44(6): 17-26.  
LI Baoqin, WU Junyong, SHAO Meiyang, et al. Refined transient stability evaluation for power system based on ensemble deep belief network[J]. *Automation of Electric Power Systems*, 2020, 44(6): 17-26.
- [14] 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.
- [15] 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.
- [16] 王强, 刘炼, 陈浩. 基于关系探索和 KTBoost 的暂态稳定裕度评估[J]. *电力系统及其自动化学报*, 2022, 34(8): 35-43.  
WANG Qiang, LIU Lian, CHEN Hao. Transient stability margin assessment based on relationship exploration and KTBoost[J]. *Proceedings of the CSU-EPSA*, 2022, 34(8): 35-43.
- [17] 李郅琴, 杜建强, 聂斌, 等. 特征选择方法综述[J]. *计算机工程与应用*, 2019, 55(24): 10-19.  
LI Zhiqin, DU Jianqiang, NIE Bin, et al. Summary of feature selection methods[J]. *Computer Engineering and Applications*, 2019, 55(24): 10-19.
- [18] VERHAEGHE J, DONCKT J V D, ONGENAE F, et al. Powershap: a power-full shapley feature selection method[C]//European Conference, ECML PKDD 2022, Grenoble, France, September 19-23, 2022.
- [19] 王强, 陈浩, 刘炼. 基于多层 CatBoost 的电力系统暂态稳定评估[J]. *科学技术与工程*, 2022, 22(4): 1456-1464.  
WANG Qiang, CHEN Hao, LIU Lian. Transient stability assessment of power system based on multi-layer CatBoost[J]. *Science Technology and Engineering*, 2022, 22(4): 1456-1464.
- [20] 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.
- [21] 叶圣永, 王晓茹, 刘志刚, 等. 基于 Stacking 元学习策略的电力系统暂态稳定评估[J]. *电力系统保护与控制*, 2011, 39(6): 12-16.  
YE Shengyong, WANG Xiaoru, LIU Zhigang, et al. Power system transient stability assessment based on Stacking meta-learning strategy[J]. *Power System Protection and Control*, 2011, 39(6): 12-16.
- [22] 周挺, 杨军, 周强明, 等. 基于改进 LightGBM 的电力系统暂态稳定评估方法[J]. *电网技术*, 2019, 43(6): 1931-1940.  
ZHOU Ting, YANG Jun, ZHOU Qiangming, et al. Power system transient stability assessment method based on modified LightGBM[J]. *Power System Technology*, 2019, 43(6): 1931-1940.
- [23] 周惠怡, 刘颂凯, 张磊, 等. 考虑误分类约束的电力系统暂态稳定评估[J]. *电力系统及其自动化学报*, 2022, 34(6): 71-78.  
ZHOU Huiyi, LIU Songkai, ZHANG Lei, et al. Transient stability assessment for power system considering misclassification constraints[J]. *Proceedings of the CSU-EPSA*, 2022, 34(6): 71-78.
- [24] 张林林, 胡熊伟, 李鹏, 等. 基于极限学习机的电力系统暂态稳定评估方法[J]. *上海交通大学学报*, 2019, 53(6): 749-756.  
ZHANG Linlin, HU Xiongwei, LI Peng, et al. Power system transient stability assessment based on extreme learning machine[J]. *Journal of Shanghai Jiaotong University*, 2019, 53(6): 749-756.
- [25] 高书宇, 刘友波, 刘挺坚, 等. 嵌入轻量梯度提升机评估模型的暂态稳定预防-紧急协调控制[J]. *电力工程技术*, 2023, 42(5): 167-176.  
GAO Shuyu, LIU Youbo, LIU Tingjian, et al. Transient stability prevention-emergency coordinated control embedded with assessment model of light gradient boosting machine[J]. *Electric Power Engineering Technology*, 2023, 42(5): 167-176.

(责任编辑 杜亚勤)