

DOI: 10.19783/j.cnki.pspc.201263

电压暂降源异质堆叠集成学习识别法

汪颖, 陈春林, 肖先勇

(四川大学电气工程学院, 四川 成都 610065)

摘要: 电压暂降源分类识别存在可获得信息不完备的问题。针对现有单一识别法弱学习特点和组合识别法一致性强的问题, 提出一种基于异质堆叠集成学习的暂降源识别方法, 提升识别模型的泛化能力和鲁棒性。把线路故障分为普通故障和雷击故障, 以10类单一电压暂降源的识别为目标, 选取9个表征特征差异的波形统计参数, 构建27维识别特征向量。引入堆叠集成算法, 以5种差异性强的单一识别法为基分类器, 用随机森林法作元分类器, 建立异质堆叠集成识别模型。通过PSCAD仿真数据和实测数据验证, 并与现有6种识别法比较, 结果表明, 该方法识别精度高, 噪声鲁棒性良好, 具有良好的工程实用性。

关键词: 电压暂降源识别; 异质性; 差异性; 堆叠集成; 识别精度; 噪声鲁棒性

Heterogeneous stacking integrated learning identification method for voltage sag sources

WANG Ying, CHEN Chunlin, XIAO Xianyong

(College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

Abstract: The classification and identification of voltage sag sources has the problem of having incomplete information. Given the weak learning characteristics of the existing single identification method and the strong consistency of the combined identification method, a sag source identification method based on heterogeneous stacking ensemble learning is proposed to improve the generalization ability and robustness of the recognition model. Line faults are subdivided into common faults and lightning faults, and ten types of single voltage sag sources are used as the identification target. Nine waveform statistical parameters that can characterize the differences are selected to construct a 27-dimension recognition vector. The stacking ensemble algorithm is introduced, five highly differentiated single recognition methods are used as base-classifiers, and random forest is selected as the meta-classifier to establish a heterogeneous stacking ensemble recognition model. Through the verification of a PSCAD simulation model and measured data, and comparison with six typical sag source identification methods, it is shown that the proposed method has high identification accuracy and good anti-noise performance with good engineering value.

This work is supported by the National Natural Science Foundation of China (No. 51807126).

Key words: voltage sag source identification; heterogeneity; difference; Stacking integration; identification accuracy; noise robustness

0 引言

电压暂降造成巨大损失^[1-4], 已成为社会热点。通过识别暂降源, 确定更有效的解决方案, 是解决电压暂降问题的必然要求。在仅有暂降幅值、频次等有限信息情况下, 暂降源识别面临信息不完备问题, 现有识别方法多为异质弱学习模型, 研究多异质集成强学习暂降源识别方法, 具有重要理论价值

和现实意义^[5-6]。

国内外对暂降源识别开展了大量研究^[7-18], 主要有单一法和组合法, 重点研究特征提取和分类识别算法^[7-8]。特征提取法包括小波变换^[9]、S变换^[10-11]、希尔伯特黄变换^[12-13]等。单一分类识别法有相似度匹配^[14]和分类器识别^[10-13, 15-18]两大类。前者通过距离矩阵, 按相似度进行识别; 后者采用支持向量机^[10-11, 15]、决策树^[12-13]、极限学习机^[16]、神经网络^[17-18]等算法, 通过样本训练识别模型。这些方法无论如何改进, 在信息不完备情况下, 均呈弱学习

基金项目: 国家自然科学基金项目资助(51807126)

特点, 通过集成学习可弥补此不足。国内外提出了自助法、自助聚合(Bagging)、随机森林、提升法(Boosting)、堆叠法(Stacking)等集成学习法, 考虑多弱学习模型的偏置、方差等, 建立有强学习特点的集成模型, 提升识别性能, 更符合暂降源识别的实际。

自助法、自助聚合(Bagging)、随机森林、提升法(Boosting)、堆叠法(Stacking)等集成学习法均为组合识别法, 现有暂降源单一识别法的不足是, 不同场景下识别误差大, 可能出现过拟合, 泛化能力不足^[19-20]等, 可通过异质堆叠集成学习解决这些问题。

本文以 27 维暂降波形特征为依据, 以 10 类单一暂降源为目标, 选取 5 种异质识别模型, 用随机森林(RF)法作元分类器, 提出一种异质堆叠(Stacking)集成暂降源识别法。通过仿真和实测数据验证本文方法, 并与现有 6 种方法比较, 结果表明, 本文方法具有识别准确度高、噪声鲁棒性好等优点, 工程应用前景良好。

1 暂降源异质堆叠集成识别法原理

1.1 异质堆叠集成原理

集成学习是将若干弱学习识别器集成为强学习识别器的方法, 分同质集成和异质集成。前者所用基础识别器相同, 误差相似度高; 后者采用不同类型基础识别器, 准确率高, 泛化能力强。

堆叠(Stacking)集成^[21]一般采用两层结构, 如图 1, 第一层为基分类器, 第二层为元分类器, 集成性能主要决定于基分类器效果和差异性。

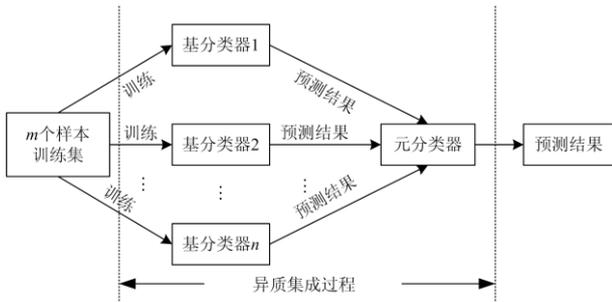


图 1 异质堆叠(Stacking)集成识别模型

Fig. 1 Heterogeneous Stacking integrated recognition model

1.2 基分类器选择

K 近邻算法(KNN)、支持向量机(SVM)、梯度提升决策树(GBDT)、极端梯度提升树(XGBOOST)、随机森林(RF)等分类器算法的优缺点如表 1。

如表 1 所示: KNN 鲁棒性好, 适合处理含噪声的电压暂降数据, 但计算复杂; 而 RF 训练速度快, 计算简单, 能弥补 KNN 的不足; 实际电网中, 变压器激磁、电机启动等电压暂降源数量少, 属于小

样本数据, SVM 非常适合处理这类数据; 但暂降特征维数一般不高, 引入擅长低维数据处理的 GBDT 算法可提升模型精度和效率; 采用 XGBoost 可有效防止识别模型过拟合。综上, 五种算法差异明显且各有优缺点, 符合暂降源异质堆叠集成识别要求。

表 1 典型算法对比

算法	特点	缺点
KNN	非参数监督, 异常值不敏感, 鲁棒性好	计算复杂, 易受样本影响
SVM	监督学习, 适合小样本、非线性及高维分类识别	参数和核函数难以确定
GBDT	Boosting 算法类, 适合低维数据, 可灵活处理各种类型数据	弱学习器之间存在依赖关系, 难以并行训练数据
XGBoost	Boosting 算法类, 泰勒二阶展开, 防止过拟合, 支持列抽样	弱学习器之间存在依赖关系, 难以并行训练数据
RF	Bagging 算法类, 抗过拟合能力、泛化能力强, 能处理高维数据, 训练速度快	可能存在很多相似的决策树, 掩盖了真实的结果

2 异质堆叠识别法实现步骤

电压暂降源识别目标包括识别单相接地(S₁)、两相接地(S₂)、相间短路(S₃)、三相故障(S₄)、雷击单相接地(S₅)、雷击两相接地(S₆)、雷击相间短路(S₇)、雷击三相(S₈)^[13,22-23]、变压器激磁(S₉)和电机启动(S₁₀)等。

2.1 特征提取

2.1.1 时域特征

选取 8 个能够突出暂降波形差异的时域特征^[18,24]: 均值 F_{Mean} 、方差 F_{Std} 、有效值 F_{Rms} 、斜度 F_{Skew} 、峭度 F_{Kur} 、峰值系数 F_{Crest} 、波形系数 F_{Ffa} 和对数能量 F_{Le} 。不同暂降源引起的电压暂降会在该 8 个时域特征中显示出不同的数据特点。线路故障和雷击故障引起的电压暂降, 其均值 F_{Mean} 会比较接近暂降幅值; 而变压器激磁和电机启动两种原因造成的电压暂降其均值会较大程度偏离暂降幅值。

2.1.2 雷击识别特征

为准确识别出雷击故障(S₅—S₈), 需结合雷击现象, 选取能辨别出雷击故障的特征。雷电冲击时刻, 由于存在雷电行波, 系统电压中掺杂着大量高频分量; 随着雷击结束, 雷电流入地, 行波消失, 高频分量逐渐衰减; 在电压暂降结束之时, 高频分量几乎完全衰减^[22-23]。参考文献[12], 对原始电压暂降波形经 HHT 变换处理, 选取 IMF1 分量的最大幅值 F_{1max} 作为雷击识别特征。

2.1.3 特征向量

基于选取的 9 个特征, 构造特征向量 \mathbf{F}_0 :

$$\mathbf{F}_0 = [F_{\text{Mean}}, F_{\text{Std}}, F_{\text{Rms}}, F_{\text{Skew}}, F_{\text{Kur}}, F_{\text{Crest}}, F_{\text{Ffa}}, F_{\text{Le}}, F_{\text{I}_{\text{max}}}] \quad (1)$$

考虑 ABC 三相, 构成 $9 \times 3 = 27$ 维特征向量 $\mathbf{F}_{0\text{-ABC}}$ 如式(2)。

$$\mathbf{F}_{0\text{-ABC}} = [F_0(A), F_0(B), F_0(C)] \quad (2)$$

为提升可比性, 27 维特征采用离差标准化 (Max-Min) 方法进行归一化, 如式(3)所示。

$$f = \frac{F - F_{\min}}{F_{\max} - F_{\min}} \quad (3)$$

式中, F_{\max} 和 F_{\min} 是指标 F 的最大值和最小值。

2.2 构建堆叠集成模型第一层

对 m 个暂降训练样本, 提取 27 维特征并归一化, 得特征向量 $\mathbf{x}_i = [f_{i-1}, f_{i-2}, \dots, f_{i-27}]$, 各样本对应的暂降源为 $\mathbf{y}_i \in \mathbf{S} = \{\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_{10}\}$, $i = 1, 2, \dots, m$ 。模型第一层输入为训练样本 \mathbf{x}_i 及其对应的暂降源标签 \mathbf{y}_i , 训练中采用 10 折交叉验证, 输出为验证数据 $\mathbf{x} = [f_1, f_2, \dots, f_{27}]$ 的 5 个基分类器预测结果。

2.2.1 基分类器 KNN

采用欧氏距离度量:

$$d(\mathbf{x}, \mathbf{x}_i) = \sqrt{\sum_{n=1}^{27} (f_n - f_{i-n})^2} \quad (4)$$

找出与 \mathbf{x} 最近邻的 k 个点, 涵盖 k 个领域, 记为 $N_k(\mathbf{x})$, 采用多数规则, 决定类别 y_{KNN} 。

$$y_{\text{KNN}} = \arg \max_{S_j} \sum_{\mathbf{x}_i \in N_k(\mathbf{x})} I(\mathbf{y}_i = S_j) \quad (5)$$

式中, I 为指示函数。当 $\mathbf{y}_i = S_j$ 时, I 为 1, 否则 I 为 0 ($i = 1, 2, \dots, m, j = 1, 2, \dots, 10$)。

2.2.2 基分类器 SVM

设分类最优超平面方程为

$$\boldsymbol{\omega}^T \mathbf{x} + b = 0 \quad (6)$$

目标函数为

$$\min \frac{1}{2} \|\boldsymbol{\omega}\|^2, \text{ s.t. } \mathbf{y}_i (\boldsymbol{\omega}^T \mathbf{x}_i + b) \geq 1, i = 1, L, m \quad (7)$$

目标函数是二次的, 约束条件是线性的, 利用拉格朗日对偶性, 变换为对偶变量优化问题; 引入核函数, 推广到非线性分类, 用于暂降源识别。

训练集 $\mathbf{x}_i (i = 1, 2, \dots, m)$ 非线性变换后, 验证数据 \mathbf{x} 的预测结果为

$$y_{\text{VSM}} = \text{sgn} \left(\sum_{i=1}^m \alpha_i \mathbf{y}_i K(\mathbf{x}_i, \mathbf{x}) + b \right) \quad (8)$$

式中: α_i 为拉格朗日系数; \mathbf{y}_i 为暂降源标签; $K(\mathbf{x}_i, \mathbf{x})$ 为高斯核函数, 如式(9)所示。

$$K(\mathbf{x}_i, \mathbf{x}) = \exp \left(-\frac{\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2} \right) \quad (9)$$

2.2.3 基分类器 GBDT

训练 \mathbf{x}_i 和暂降源标签 \mathbf{y}_i 构成初始化分类器, 如式(10)所示。

$$f_0(\mathbf{x}) = \arg \min_{S_j} \sum_{i=1}^m L(\mathbf{y}_i, S_j) \quad (10)$$

式中, $L(\mathbf{y}_i, S_j)$ 为平方损失函数, 如式(11)所示。

$$L(\mathbf{y}_i, S_j) = \frac{1}{2} (\mathbf{y}_i - S_j)^2 \quad (11)$$

目标为最小化损失函数, GBDT 算法采用损失函数的负梯度来拟合损失函数。设迭代 K 轮生成 K 棵决策树, 第 k 次迭代的负梯度, 即残差为

$$r_{ik} = - \left[\frac{\partial L(\mathbf{y}_i, f(\mathbf{x}_i))}{\partial f(\mathbf{x}_i)} \right]_{f(\mathbf{x})=f_{k-1}(\mathbf{x})} \quad (12)$$

训练过程中以第 k 轮残差作暂降源新标签, 将 (\mathbf{x}_i, r_{ik}) 作为下一轮训练数据, 得新分类树 $f_{k+1}(\mathbf{x})$, 使损失函数逐渐减小, 对应的决策树叶子节点区域为 R_{jk} , $j = 1, 2, \dots, J$, J 为叶子节点数。

对叶子区域 $j = 1, 2, \dots, J$, 计算最佳拟合值, 如式(13)所示。

$$\gamma_{jk} = \arg \min_{\gamma} \sum_{\mathbf{x}_i \in R_{jk}} L(\mathbf{y}_i, f_{k-1}(\mathbf{x}_i) + \gamma) \quad (13)$$

根据拟合值更新分类器, 如式(14)。

$$f_k(\mathbf{x}) = f_{k-1}(\mathbf{x}) + \sum_{j=1}^J \gamma_{jk} I(\mathbf{x} \in R_{jk}) \quad (14)$$

GBDT 迭代完成后, 得验证数据 \mathbf{x} 的预测结果为

$$y_{\text{GBDT}} = f(\mathbf{x}) = f_0(\mathbf{x}) + \sum_{k=1}^K \sum_{j=1}^J \gamma_{jk} I(\mathbf{x} \in R_{jk}) \quad (15)$$

2.2.4 基分类器 XGBoost

XGBoost 是对 GBDT 的改进, 预测结果为

$$y_{\text{XGBoost}} = f(\mathbf{x}) = f_0(\mathbf{x}) + \sum_{k=1}^K \sum_{j=1}^J \gamma_{jk} I(\mathbf{x} \in R_{jk}) \quad (16)$$

与 GBDT 不同, XGBoost 在优化时, 同时采用了一阶和二阶导数信息, 目标为最小化损失函数, 如式(17)所示。

$$L^{(k)} = \sum_{i=1}^m l(\mathbf{y}_i^{(k-1)} + f_k(\mathbf{x}_i), \mathbf{y}_i) + \Omega(f_k) \quad (17)$$

式中: $L^{(k)}$ 为第 k 轮目标函数; $\mathbf{y}_i^{(k-1)}$ 为前 $k-1$ 棵树预测值; f_k 为第 k 棵树预测结果; l 为预测值与真实值间的训练误差函数; Ω 表示树模型复杂度。

损失函数在 $\mathbf{y}_i^{(k-1)}$ 处泰勒展开, 去掉常数项, 第 k 轮损失函数为

$$L^{(k)} = \sum_{i=1}^m [g_i f_k(\mathbf{x}_i) + \frac{1}{2} h_i f_k^2(\mathbf{x}_i)] + \Omega(f_k) \quad (18)$$

2.2.5 基分类器 RF

RF 是多决策树组成的 Bagging 集成算法, 多棵

决策树并行训练, 训练速度快。输入为暂降训练数据 x_i 和暂降源标签 $y_i \in S = \{S_1, S_2, \dots, S_{10}\}$, $i = 1, 2, \dots, m$ 。

假设训练共产生 K 棵分类决策树, 得决策树序列 $\{f_1(x), f_2(x), \dots, f_K(x)\}$, 构成多分类模型系统, 验证数据 x 的预测结果采用多票表决得:

$$y_{RF} = \arg \max_{S_j} \sum_{k=1}^K I(f_k(x) = S_j) \quad (19)$$

2.3 构建堆叠集成模型第二层

第二层元分类器采用训练速度快, 原理简单的 RF。数据 x 经第一层处理后所得输出 y_p 为

$$y_p = [y_{KNN}, y_{SVM}, y_{GBDT}, y_{XGBoost}, y_{RF}] \quad (20)$$

训练样本 x_i 经第一层基分类器处理后的输出为 y_{pi} , 把 (y_{pi}, y_i) 如多元分类器 RF 进一步训练, $i = 1, 2, \dots, m$, y_{pi} 为 5 维特征。

假设第二层训练产生 H 棵决策树, 得树序列 $\{g_1(x), g_2(x), \dots, g_H(x)\}$, 验证数据 x 经堆叠两层结构的最终预测结果 $y_{x\text{-stacking}}$ 为

$$y_{x\text{-stacking}} = \arg \max_{S_j} \sum_{h=1}^H I(g_h(x) = S_j) \quad (21)$$

暂降源识别流程, 如图 2。

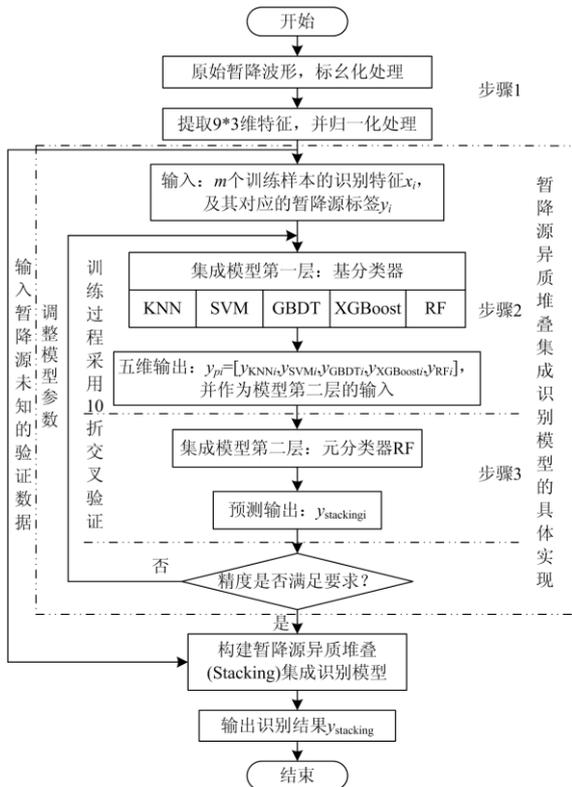


图 2 暂降源异质堆叠集成识别流程

Fig. 2 Flow chart of voltage sag source identification based on stacking heterogeneous integration

3 方法验证

3.1 仿真模型

参考[5]和[18], 在 PASCAD 平台搭建线路故障、变压器投运和电机启动三种暂降源模型, 如图 3。

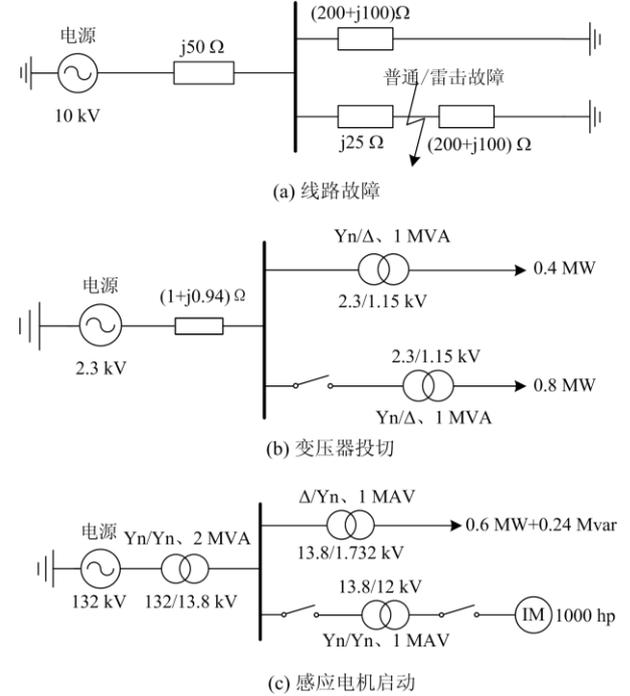


图 3 电压暂降源仿真模型

Fig. 3 Simulation model of voltage sag source

系统频率 50 Hz, 采样频率 6.4 kHz, 通过改变模型参数, 如故障类型、持续时间(0.1~0.8 s)、短路阻抗(0~50 Ohms)、变压器/电机额定容量(1~10 MVA)、线路负荷(0.2~1 MVA)等, 生成 $S_i (i = 1, 2, \dots, 10)$ 类暂降样本各 200 个。

图 3(a)中, 雷击故障采用叠加标准双指数雷电流模拟^[12]。因 75%~90%的雷电流为负极性, 因此仿真中采用负极性雷电冲击波。

3.2 仿真结果分析

3.2.1 模型评价指标

为验证提出方法的性能, 基于混淆矩阵, 采用准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall)、特异度 (Specificity) 和 F_1 值 5 个指标评价所提方法的性能。以 $S_i (i = 1, 2, \dots, 10)$ 类样本为例, 5 个指标如表 2。

表 2 中: T_P 为分类结果为 S_i 且实际是 S_i 的样本数; F_N 是实际为 S_i 类但分类结果非 S_i 类的样本数; F_P 为分类结果为 S_i 但实际非 S_i 类的样本数; T_N 为分类结果非 S_i 且实际非 S_i 类的样本数。

表 2 模型评价指标表达式

Table 2 Expressions of model performance evaluation index

指标	描述	表达式
准确率	识别类型和实际类型一致的样本占总样本的比例	$Accuracy = \frac{T_p + T_N}{T_p + F_N + F_p + T_N}$
精确率	预测为 S_i 类的样本中, 实际类别为 S_i 的样本所占的比例	$Precision = \frac{T_p}{T_p + F_p}$
召回率	实际为 S_i 类的样本中, 被识别为 S_i 类样本的比例	$Recall = \frac{T_p}{T_p + F_N}$
特异度	非 S_i 类样本的识别准确率	$Specificity = \frac{T_N}{T_N + F_p}$
F_1 值	精确率和召回率的调和均值	$F_1 = \frac{2T_p}{2T_p + F_p + F_N}$

3.2.2 无噪声数据分析

对仿真数据(训练:测试比例为 7:3)进行识别, 测试样本(各 60 组)识别结果对应的混淆矩阵如图 4。



图 4 异质堆叠集成识别模型分类的混淆矩阵

Fig. 4 Confusion matrix for classification of heterogeneous integrated recognition models

混淆矩阵的行为实际类别, 列为预测类别。图 4 中, 1 组 S_5 误分为 S_1 , 1 组 S_7 误分为 S_3 , 1 组 S_9 误分为 S_{10} ; S_1 — S_4 、 S_6 、 S_8 、 S_{10} 准确, 指标如表 3。

表 3 识别模型性能评价指标

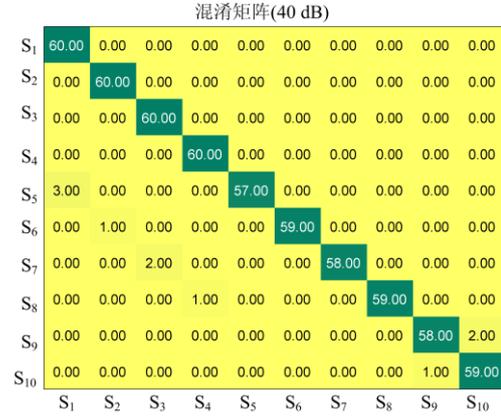
Table 3 Identification model performance evaluation index

暂降类型	Accuracy	Precision	Recall	Specificity	F_1
S_1	99.8%	98.4%	100%	99.8%	99.2%
S_2	100%	100%	100%	100%	100%
S_3	99.8%	98.4%	100%	99.8%	99.2%
S_4	100%	100%	100%	100%	100%
S_5	99.8%	100%	98.3%	100%	99.2%
S_6	100%	100%	100%	100%	100%
S_7	99.8%	100%	98.3%	100%	99.2%
S_8	100%	100%	100%	100%	100%
S_9	99.8%	100%	98.3%	100%	99.2%
S_{10}	99.8%	98.4%	100%	99.8%	99.2%
最小值	99.8%	98.4%	98.3%	99.8%	99.2%

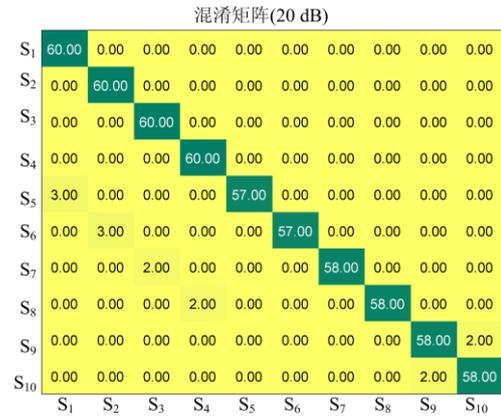
由表 3 可知, 指标最小值分别为 99.8%、98.4%、98.3%、99.8%、99.2%, 说明本文方法识别准确率高, 满足无噪声数据识别要求。

3.2.3 噪声数据鲁棒性分析

在仿真数据中添加不同程度的高斯白噪声, 验证方法的鲁棒性, 识别混淆矩阵如图 5。



(a) 40 dB 噪声数据结果



(b) 20 dB 噪声数据结果



(c) 10 dB 噪声数据结果

图 5 识别结果对应的混淆矩阵(噪声数据)

Fig. 5 Confusion matrix corresponding to the sag source identification result (noise data)

加入噪声后, S_1 — S_4 仍识别准确, S_5 — S_{10} 有少量误判。 S_5 误判为 S_1 , S_6 误判为 S_2 , S_7 误判为 S_3 , S_8 误判为 S_4 。这是由于 S_1 — S_4 为非雷击暂降, 而 S_5 — S_8 为雷击暂降, 因两者暂降波形在起始时刻的差异减少导致少量误判。

用表 2 指标来评价, 所提方法的性能指标最小值如表 4。

表 4 噪声数据的性能评价指标(最小值)

Table 4 Noise data performance evaluation index (minimum value)

噪声	Accuracy	Precision	Recall	Specificity	F_1
无噪声	99.8%	98.4%	98.3%	99.8%	99.2%
40 dB	99.5%	95.2%	95.0%	99.4%	97.4%
20 dB	99.3%	95.2%	95.0%	99.4%	96.7%
10 dB	98.8%	93.8%	93.3%	99.3%	94.1%

可见, 添加噪声后, 性能指标均有所降低, 10 dB 噪声下效果最差, 但准确率仍高达 98.8%, F_1 指标不低于 94.1%, 证明本文方法有良好的鲁棒性。

3.2.4 与现有方法对比

与现有 6 种方法比较的结果如表 5。

表 5 与现有方法对比结果(无噪声数据)

Table 5 Comparison with existing methods (no noise data)

暂降源	识别精度/%						
	本文方法	方法一 ^[9]	方法二 ^[18]	方法三 ^[12]	方法四 ^[16]	方法五 ^[10]	方法六 ^[5]
S_1	99.8	100	100		100		
S_2	100	92.7	100		/		
S_3	99.8	93.8	97.92	100	/		
S_4	100	100	97.92		/	98.75	100
S_5	99.8	/	/	/	/		
S_6	100	/	/	100	/		
S_7	99.8	/	/		/		
S_8	100	/	/		/		
S_9	99.8	97	95.83	94	98.3	100	100
S_{10}	100	97.2	100	97	99.4	98.75	98.33

注: 表格中‘/’表示未考虑该项暂降源; ‘合并’表示未严格区分, 将几类暂降源合并为一类研究。

表 5 中: 方法一、二将短路细分为单相接地、两相接地、相间短路和三相短路导致的 4 类暂降源, 未考虑雷击暂降; 方法三考虑了雷击暂降, 但未细分故障类型; 方法四仅考虑了单相接地; 方法五、六将所有故障合并为一类暂降源。本文把雷击和非雷击故障分为 8 种障类型(S_1 — S_8), 识别精度不低于 99.8%; 误判发生在同类故障间, 如 S_5 误判为 S_1 。对比表明, 本文方法对线路故障、变压器激磁、电

机启动导致的暂降源识别精度均较高, 且可进一步区分 8 种故障类型的暂降源。

4 实测数据验证

用南方某省 2019 年监测数据进一步验证, 部分实测波形如图 6。

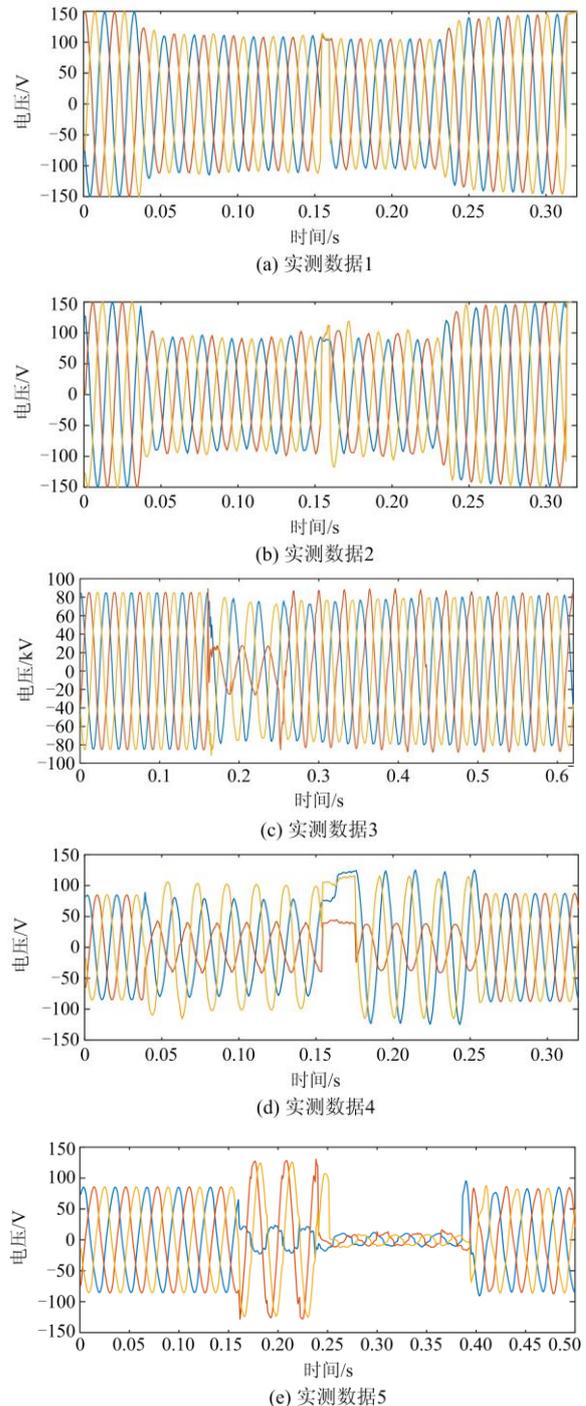


图 6 实测数据波形

Fig. 6 Measured data waveform

图 6(a)为 2019 年 4 月 19 日某 10 kV 馈线三相短路故障实测数据, 图 6(b)为 2019 年 4 月 22 日某 10 kV 馈线三相短路故障实测数据, 图 6(c)为 2019 年 5 月 27 日某 10 kV 母线单相接地故障实测数据, 图 6(d)为 2019 年 6 月 05 日某 10 kV 馈线单相接地故障实测数据, 图 6(e)为 2019 年 10 月 11 日某 10 kV 馈线单相接地故障后发展为三相故障的实测数据。

应用本文方法, 所得识别结果如表 6。

表 6 实测数据识别结果

Table 6 Recognition results of measured data

实测数据	实际暂降源	识别暂降源
1	三相短路	三相短路(S ₄)
2	三相短路	三相短路(S ₄)
3	单相接地故障	单相接地故障(S ₁)
4	单相接地故障	单相接地故障(S ₁)
5	单相故障后发展为三相故障	三相短路(S ₄)

结果表明, 本文方法能正确识别 1~4 组实测数据。数据 5 出现了一点偏差, 原因在于本文识别目标未考虑多级暂降, 这也是有待下一步研究的问题。

5 结论

本文基于 5 种异质识别模型, 提出一种电压暂降源异质堆叠集成识别法, 得到以下结论。

1) 所提暂降源异质堆叠集成识别法, 准确率不低于 98.8%, 噪声鲁棒性好。

2) 对于故障引起的暂降, 所提方法可准确识别雷击和非雷击引起的暂降。

3) 实测数据验证表明, 本文方法能正确识别单一暂降源, 多级暂降源识别准确率还有待提升。

因故障类型变化导致的多级暂降源、复合暂降源等的识别, 还有待进一步研究。

参考文献

- [1] 谢伟伦, 薛峰, 黄志威. 基于网络传播特性的配电网电压暂降随机评估方法[J]. 电力系统保护与控制, 2020, 48(8): 163-171.
XIE Weilun, XUE Feng, HUANG Zhiwei. Stochastic estimation method of voltage sags for a distribution network based on network propagation property[J]. Power System Protection and Control, 2020, 48(8): 163-171.
- [2] XU Y, WU Y, ZHANG M, et al. Sensitivity of programmable logic controllers to voltage sags[J]. IEEE Transactions on Power Delivery, 2019, 34(1): 2-10.
- [3] ZHAO W, SHANG L, SUN J. Power quality disturbance classification based on time-frequency domain multi-feature and decision tree[J]. Protection and Control of Modern Power Systems, 2019, 27(4). DOI: 10.1186/s41601-019-0139-z.
- [4] XIAO X, HE H, WANG Y. Analytical model of ac contactors for studying response mechanism to multi-dimensional voltage sag characteristics and its novel applications[J]. IET Generation, Transmission & Distribution, 2019, 13(17): 3910-3920.
- [5] 李晨懿, 杨家莉, 徐永海, 等. 模糊综合评价在电压暂降源识别中的应用[J]. 电网技术, 2017, 41(3): 1022-1028.
LI Chenyi, YANG Jiali, XU Yonghai, et al. Application of comprehensive fuzzy evaluation method on recognition of voltage sag disturbance sources[J]. Power System Technology, 2017, 41(3): 1022-1028.
- [6] ZHENG Z, QI L, WANG H, et al. Recognition method of voltage sag causes based on two-dimensional transform and deep learning hybrid model[J]. IET Power Electronics, 2020, 13(1): 168-177.
- [7] 夏杰, 张建文, 王金川. 基于 FFT 与形态加权动态时间规整的电压暂降源辨识新方法[J]. 电测与仪表, 2019, 56(1): 28-34.
XIA Jie, ZHANG Jianwen, WANG Jinchuan. A novel method of voltage sag source identification based on FFT and shape-based weighted dynamic time warping[J]. Electrical Measurement & Instrumentation, 2019, 56(1): 28-34.
- [8] WEI P, XU Y, WU Y, et al. Research on classification of voltage sag sources based on recorded events[J]. CIRED-Open Access Proceedings Journal, 2017, 1: 846-850.
- [9] 褚佳伟, 袁晓冬, 陈兵, 等. 结合小波分析和改进型 DTW 距离的配电网电压暂降源辨识方法[J]. 电网技术, 2018, 42(2): 637-643.
CHU Jiawei, YUAN Xiaodong, CHEN Bing, et al. A method for distribution network voltage sag source identification combining wavelet analysis and modified DTW distance[J]. Power System Technology, 2018, 42(2): 637-643.
- [10] 杨秀, 张彤瑶, 潘爱强, 等. 基于 S 变换与多维多形的电压暂降源特征识别[J/OL]. 电网技术: 1-8. DOI: 10.13335/j.1000-3673.pst.2019.2002
YANG Xiu, ZHANG Tongyao, PAN Aiqiang, et al. Voltage sag source feature identification with S transform and multidimensional fractal[J/OL]. Power System Technology: 1-8. DOI: 10.13335/j.1000-3673.pst.2019.2002.
- [11] TANG Y, WEI R, CHEN K, et al. Voltage sag source identification based on the sign of internal resistance in a

- "Thevenin's equivalent circuit"[J]. International Transactions on Electrical Energy Systems, 2017, 27(12): 1-18.
- [12] 李新, 武利会, 范心明, 等. 基于希尔伯特黄变换和决策树的电压暂降源识别方法[J]. 科学技术与工程, 2019, 34(19): 198-205.
LI Xin, WU Lihui, FAN Xinming, et al. Voltage sag sources identification method based on Hilbert-Huang transform and decision tree[J]. Science Technology and Engineering, 2019, 34(19): 198-205.
- [13] 崔灿, 肖先勇, 吴奎华, 等. 基于 HHT 和特征离散化的电压暂降扰动源分类特征提取[J]. 电力系统保护与控制, 2018, 46(24): 8-15.
CUI Can, XIAO Xianyong, WU Kuihua, et al. Feature extraction in classification of voltage sag sources based on HHT and feature discretization[J]. Power System Protection and Control, 2018, 46(24): 8-15.
- [14] 邱海峰, 陈兵, 袁晓冬, 等. 基于动态时间弯曲距离的电压暂降源辨识方法[J]. 电力系统保护与控制, 2017, 45(13): 7-13.
QIU Haifeng, CHEN Bing, YUAN Xiaodong, et al. Identification of voltage sag sources based on DTW[J]. Power System Protection and Control, 2017, 45(13): 7-13.
- [15] 李陶然, 张宸宇, 史明明, 等. 计及雷击情况的基于 PDT-SVM 暂降源辨识方法研究[J]. 电力工程技术, 2019, 38(5): 2-7.
LI Taoran, ZHANG Chenyu, SHI Mingming, et al. PDT-SVM-based sag source identification considering lightning strike[J]. Electric Power Engineering Technology, 2019, 38(5): 2-7.
- [16] 郑智聪, 王红, 齐林海. 基于深度学习模型融合的电压暂降源识别方法[J]. 中国电机工程学报, 2019, 39(1): 9-104, 324.
ZHENG Zhicong, WANG Hong, QI Linhai. Recognition method of voltage sag sources based on deep learning models' fusion[J]. Proceedings of the CSEE, 2019, 39(1): 9-104, 324.
- [17] SUN H, YI Hao, YANG G, et al. Voltage sag source identification based on few-shot learning[J]. IEEE Access, 2019, 7: 164398-164406.
- [18] 汪颖, 王欢, 张姝. 基于优化极限学习机的电压暂降源识别方法[J]. 电力系统自动化, 2020, 44(9): 135-145.
WANG Ying, WANG Huan, ZHANG Shu. Recognition method of voltage sag source based on optimized extreme learning machine[J]. Automation of Electric Power Systems, 2020, 44(9): 135-145.
- [19] 余东行, 张保明, 赵传, 等. 联合卷积神经网络与集成学习的遥感影像场景分类[J]. 遥感学报, 2020, 24(6): 717-727.
YU Donghang, ZHANG Baoming, ZHAO Chuan, et al. Scene classification of remote sensing image using ensemble convolutional neural network[J]. Journal of Remote Sensing, 2020, 24(6): 717-727.
- [20] 刘佳翰, 陈克绪, 马建, 等. 基于卷积神经网络和随机森林的三相电压暂降分类[J]. 电力系统保护与控制, 2019, 47(20): 112-118.
LIU Jiahao, CHEN Kexu, MA Jian, et al. Classification of three-phase voltage dips based on CNN and random forest[J]. Power System Protection and Control, 2019, 47(20): 112-118.
- [21] KLAUS N. Ensemble methods: foundations and algorithms by Zhi-Hua Zhou[J]. International Statistical Review, 2013, 81(3).
- [22] 肖先勇, 李逢, 邓武军. 雷击与普通故障引起的电压凹陷特征[J]. 高电压技术, 2009, 35(2): 309-314.
XIAO Xianyong, LI Feng, DENG Wujun. Characteristic identification of voltage sag caused by lightning stroke and short-circuit faults[J]. High Voltage Engineering, 2009, 35(2): 309-314.
- [23] 肖先勇, 李逢, 邓武军. 雷击与短路故障的 S 变换特征量识别方法[J]. 高电压技术, 2009, 35(4): 817-822.
XIAO Xianyong, LI Feng, DENG Wujun. Identification method of lightning stroke and short-circuit faults using S-transform characteristics[J]. High Voltage Engineering, 2009, 35(4): 817-822.
- [24] 汪颖, 孙建风, 肖先勇, 等. 基于卷积神经网络的电缆早期故障分类识别[J]. 电力系统保护与控制, 2020, 48(7): 10-18.
WANG Ying, SUN Jianfeng, XIAO Xianyong, et al. Cable incipient fault classification and identification based on optimized convolution neural network[J]. Power System Protection and Control, 2020, 48(7): 10-18.

收稿日期: 2020-10-20; 修回日期: 2020-12-16

作者简介:

汪颖(1981—), 女, 教授, 博士生导师, 研究方向为电能质量与优质供电; E-mail: 769429505@qq.com

陈春林(1996—), 男, 硕士研究生, 主要研究方向为电力市场与电能质量; E-mail: 2195250867@qq.com

肖先勇(1968—), 男, 通信作者, 教授, 博士生导师, 研究方向为电能质量与优质供电。E-mail: xiaoxianyong@163.com

(编辑 葛艳娜)