

5-28-2024

A warning method for low-voltage electrical safety hazard based on multi-dimensional features and random forests

Xiangqi XIAO

State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Key Laboratory of Intelligent Electrical Measurement and Application Technology in Hunan Province, Changsha 410004, China, 84578610@qq.com

Junhua HU

State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Key Laboratory of Intelligent Electrical Measurement and Application Technology in Hunan Province, Changsha 410004, China

Zhi YE

State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Key Laboratory of Intelligent Electrical Measurement and Application Technology in Hunan Province, Changsha 410004, China

Wenxian WU

State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Key Laboratory of Intelligent Electrical Measurement and Application Technology in Hunan Province, Changsha 410004, China

Yujian LUO

State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Key Laboratory of Intelligent Electrical Measurement and Application Technology in Hunan Province, Changsha 410004, China
Follow this and additional works at: <https://jepst.researchcommons.org/journal>

See next page for additional authors
Recommended Citation

XIAO, Xiangqi; HU, Junhua; YE, Zhi; WU, Wenxian; LUO, Yujian; ZOU, Sheng; and LUO, Chen (2024) "A warning method for low-voltage electrical safety hazard based on multi-dimensional features and random forests," *Journal of Electric Power Science and Technology*. Vol. 39: Iss. 2, Article 16.

DOI: 10.19781/j.issn.1673-9140.2024.02.016

Available at: <https://jepst.researchcommons.org/journal/vol39/iss2/16>

This Smart grid is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology. For more information, please contact dlxb04@163.com.

A warning method for low-voltage electrical safety hazard based on multi-dimensional features and random forests

Authors

Xiangqi XIAO, Junhua HU, Zhi YE, Wenxian WU, Yujian LUO, Sheng ZOU, and Chen LUO

引用格式:肖湘奇,胡军华,叶志,等.基于多维特征与随机森林的低压用电安全隐患预警方法[J].电力科学与技术学报,2024,39(2):143-151.

Citation: XIAO Xiangqi, HU Junhua, YE Zhi, et al. A warning method for low-voltage electrical safety hazard based on multi-dimensional features and random forests[J]. Journal of Electric Power Science and Technology, 2024, 39(2): 143-151.

基于多维特征与随机森林的低压用电 安全隐患预警方法

肖湘奇^{1,2}, 胡军华^{1,2}, 叶志^{1,2}, 吴文娴^{1,2}, 罗宇剑^{1,2}, 邹晟^{1,2}, 罗晨³

(1. 国网湖南省电力有限公司, 湖南长沙 410004; 2. 智能电气量测与应用技术湖南省重点实验室, 湖南长沙 410004;

3. 长沙理工大学电气与信息工程学院, 长沙 410114)

摘要:故障电弧作为低压用电常见安全隐患,因存在隐蔽性和随机性,导致难以在故障前夕有效感知,现有保护方法通常于故障发生后采取措施,容易造成电气火灾发生。针对这些问题,提出基于多维特征与随机森林的低压用电安全隐患预警方法。首先,搭建真型实验平台复现多种负载故障电弧真实场景,采集实验数据进行去噪、归一化和样本化等预处理;接着分析故障前、后波形变化情况,进而实行特征有效性分析,提取多维与故障相关程度较高的特征值,提高特征普适性。然后,搭建随机森林模型并进行超参数优化,以最小化节点信息熵为目标完成模型训练,提升模型整体性能和学习效率。最后,通过实验验证表明,所提方法在多种不同负载运行工况下预测准确率可高达 99.4% 以上,且预测准确率高于 4 种主流分类预测模型。

关键词:低压用电安全;故障电弧;随机森林算法;特征工程

DOI: 10.19781/j.issn.1673-9140.2024.02.016 中图分类号: TM501; TP183 文章编号: 1673-9140(2024)02-0143-09

A warning method for low-voltage electrical safety hazard based on multi-dimensional features and random forests

XIAO Xiangqi^{1,2}, HU Junhua^{1,2}, YE Zhi^{1,2}, WU Wenxian^{1,2},

LUO Yujian^{1,2}, ZOU Sheng^{1,2}, LUO Chen³

(1. State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; 2. Key Laboratory of Intelligent Electrical Measurement and

Application Technology in Hunan Province, Changsha 410004, China; 3. School of Electrical & Information Engineering,

Changsha University of Science & Technology, Changsha 410114, China)

Abstract: As a common safety hazard in low-voltage electricity use, fault arcs are difficult to perceive effectively on the eve of failure due to their concealment and randomness. Existing protection methods usually take measures after the occurrence of faults, which can easily cause electrical fires. To address these issues, a warning method for low-voltage electrical safety hazard is proposed on basis of multi-dimensional features and random forests. Next, a random forest model is built and hyper-parameters are optimized, with the goal of minimizing node information entropy to complete model training, so that enhances the overall performance and learning efficiency of the model. Finally, experimental verification shows that the proposed method achieves a prediction accuracy of over 99.4% with different loads, and its prediction accuracy is higher than that of four traditional classification prediction models.

Keywords: low-voltage electricity safety; fault arc; random forest algorithm; characteristic engineering

收稿日期: 2022-09-01; 修回日期: 2022-12-06

基金项目: 国家电网有限公司总部科技项目(5700-202155204A-0-0-00)

通信作者: 肖湘奇(1989—), 男, 工程师, 主要从事电能计量及用电安全技术研究, E-mail: 84578610@qq.com

随着科技水平的提高,家用电器种类增多,普通家庭用户对电能的需求不断上升,大部分家居产品都需通过电能驱动,而电能带来诸多便利的同时也给人们带来了许多安全隐患^[1-2]。线路长期过载、不规范接线以及电线质量参差不齐等原因导致电弧故障频发,且因其存在随机性和隐蔽性而无法在故障萌芽状态有效感知。据有关统计数据表明,中国每年发生的电气火灾超10万起,70%的重、特大火灾是电气火灾,而电气火灾中80%以上是由故障电弧引起^[3-6],给人们的经济财产和人身安全带来了巨大威胁。

针对故障电弧隐患,目前国内外学者已开展了丰富的研究。检测电弧故障的方法可分为3种:基于信号、基于数学模型和基于数据分析的方法。基于信号的电弧检测方法主要应用于配电柜、变电站和中低压配电系统的开关柜等定点检测场景,具有检测范围受限、成本过高且信号容易被环境淹没等缺点,难以广泛使用于各种复杂的低压用电场景^[7-9];基于数学模型的检测方法则主要是建立电流、功率、电压和能量等其中两两或多者之间的时变关系方程,在认识电弧本质电气关系的基础上找到电弧规律,为电弧特征分析提供足够的样本,但该方法存在模型产生的数据与实际情况之间存在差距的问题^[10-12]。随着人工智能的发展,大多学者选择基于实验数据并结合人工智能算法实现故障电弧的检测。文献[13]通过样本化正常和故障电弧2种电流数据,训练深度长短时记忆(long short-term memory, LSTM)网络实现故障电弧的辨识检测;文献[14]通过经验模态分解和信号相关理论实现故障电弧特征的提取,并搭建概率神经网络(product-based neural network, PNN)检测模型实现故障电弧的检测。但以上方法均是针对当前时刻是否发生故障电弧进行研究,而故障电弧一旦发生即有可能导致电气火灾发生。因此,实现故障电弧提前预测更有利于对其进行正确处理,提高用电安全系数。

为解决以上问题,本文通过搭建真型实验平台复现多场景故障电弧,采集实验数据进行去噪、归一化和样本化等预处理,进而分析故障前、后电流信号特征的变化规律,并提取多维故障信号有效特征值,完成故障隐患特征数据集构建。同时搭建随机森林集成学习模型,通过局部寻优方法实现模型超参数优化,并以信息熵最小化为目标进行模型训练,完成基于随机森林的低压用电安全隐患预警模

型的搭建,实现故障隐患的提前精准预警,提高低压用户用电安全系数。

1 随机森林算法原理

单棵决策树的分类预测结果往往过于简单而使得分类预测准确率不高,为提高电弧故障分类预测效果,本文采用随机森林算法搭建隐患预测模型。随机森林^[15-16]作为一种采用“集思广益”思想的集成学习算法,结合多棵决策树进行集成学习,每棵决策树输出一个分类投票结果,通过集成所有决策树的分类投票结果,并将投票数量最多的分类结果作为最终输出结果,从而使分类预测结果与实际值更加接近。

1.1 决策树原理

决策树作为随机森林的基础学习器,是一种既可以处理回归问题,同时又可以处理分类问题的学习算法,结构原理如图1所示,其中包括节点和有向边,节点又分为叶子和内部节点,叶子节点代表一个类,内部节点则代表一种特征属性。采用从根节点开始自上而下不断拆分到叶节点的方式展开分支,并以信息增益、基尼指数或信息增益比等作为节点分裂评估指标,通过测量子集纯度不断进行分裂,直到满足预设条件终止。

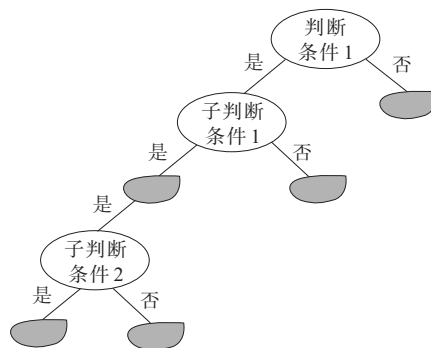


图1 决策树结构原理

Figure 1 Schematic diagram of decision tree

1.2 随机森林算法

随机森林算法通过引入2个随机量解决了决策树过拟合问题,提高了模型整体性能,可用于分析复杂且互不独立的高维特征,处理具有缺失值的数据时也能够具有较快的学习速度和较高的鲁棒性。而故障电弧发生前线路负载电流存在微小畸变,高频信号比例逐渐增加,但客观上故障特征仍然微弱,随机森林模型对此类微小现象具有高度敏感性,可精准感知电弧故障前信号的微小变化,实现

串联电弧故障快速精准预测。因此,为提高故障电弧隐患预测和辨识效果,本文分别采用随机特征选取和重复采样(Bootstrap)2个随机量,构建包含多棵决策树的随机森林模型,如图2所示,具体步骤如下。

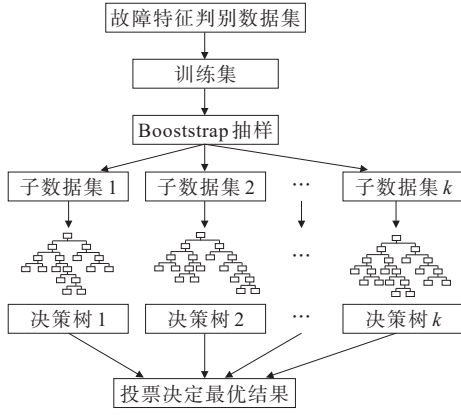


图2 随机森林算法结构原理

Figure 2 Schematic diagram of random forest algorithm

1) Bootstrap 抽样。将基于具有时间序列的正常和故障电弧数据生成的故障电弧隐患预测特征数据集记为 $F = ([x_{11}, x_{12}, \dots, x_{1N}, y_1], \dots, [x_{i1}, x_{i2}, \dots, x_{iN}, y_i])^T$, 通过 Bootstrap 抽样法在数据集 F 中随机抽取 k 次, 生成 k 个新的训练样本子集 $F_k (k=1, 2, \dots, k)$ 。其中, 每个训练子集 F_k 为一个只能取有限个样本的离散数据集, 概率分布和信息熵可表示为

$$\begin{cases} P(F_k) = p_j (j=1, 2, \dots, n) \\ H(F_k) = -\sum_{j=1}^n p_j \log_2 p_j \end{cases} \quad (1)$$

式中, $P(F_k)$ 为概率分布; $H(F_k)$ 为信息熵。

因此在训练子集 F_k 中, 将选择的特征因子 B 对 F_k 的信息增益定义为

$$G(F_k, B) = H(F_k) - H_B(F_k) \quad (2)$$

式中, $H_B(F_k)$ 为特征因子 B 在给定条件下数据集 F_k 的信息熵, 即

$$H_B(F_k) = -\sum_{|F_k^v|} p_j \log_2 \frac{|F_k^v|}{|F_k|} \quad (3)$$

其中, V 表示特征因子 B 包含的取值个数, 对应条件下的信息增益率可表示为

$$G_R(F_k, B) = \frac{G(F_k, B)}{H(F_k)} \quad (4)$$

2) 将 k 个子训练集不做剪枝处理生成最大深度决策树, 再根据式(4)计算出随机选取的部分特征

因子的信息增益率。

3) 设定最小信息增益率阈值, 并对信息增益率最大的特征因子进行分裂, 重新计算信息增益率, 并进行重复分裂操作, 当最终信息增益率小于设定阈值时停止分裂, 即选出最优分裂节点。

4) 将 k 个具有时间序列的正常和故障电弧数据子集训练生成的多个决策树集成, 建立随机森林集成学习模型, 并根据所有决策树中出现最多的分类结果, 作为随机森林集成学习模型的最终预测值。

1.3 GridSearch 超参数优化法

超参数优化可有效减少模型训练计算时间并提高预测精度, 使故障预警模型达到局部最优效果。在随机森林集成学习算法中, 需要调节的重要超参数包括: 弱评估器的规模 ($n_estimators$), 规模越大则模型准确率越高, 但同时也会增加算法计算成本; 每棵决策树的最大深度 (max_depth), 其对模型效果和过拟合都能够起到很好地调节作用; 每个叶子节点所需包含的训练样本数量 ($min_samples_leaf$), 用于抑制模型过拟合; 可以使用的最大特征数量 ($max_features$), 用于“决策树”的精修, 可直接限制所使用的特征数量而停止树的分裂。模型具有超参数数量少、各参数都能直观表达其所代表的含义等优点。

GridSearch 作为一种超参数优化方法, 通过超参数搜索域内交叉组合的方式形成超参数样本向量, 对域内候选参数进行轮番取样, 并构建临时验证模型, 基于验证集数据对超参数样本向量进行精确度评估, 其优化原理如图3所示。

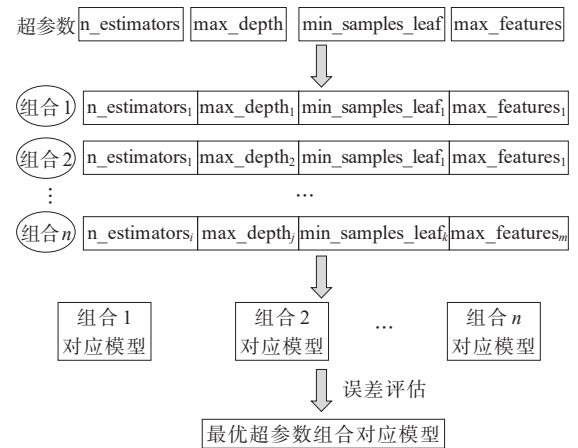


图3 模型超参数优化原理示意

Figure 3 Schematic diagram of model hyperparameter optimization

由图3可以看出,评估结果与其对应的超参数样本向量依次匹配并逐个排序,最终可得出超参数搜索域中局部最优的超参数样本组合,确保模型运算性能。该方法对于小参数深度学习模型具有很好的适用性,能够简单且有效地寻找到模型最佳超参数组合。因此,本文选取 GridSearch 超参数优化方法对模型超参数进行最优组合。

2 数据获取与波形分析

2.1 真型实验模拟

故障电弧是一种气体游离放电现象。当线路长期过载运行导致线路内芯发热,或是线路长期处于潮湿、含腐蚀性物质的空气中时,会导致线路绝缘层逐渐老化、破损而发生空气击穿现象,即故障电弧现象,也是生活用电中频繁发生的一种现象。为模拟实际生活场景中的真实故障电弧隐患,在实验室已有电动电弧发生装置条件下,依照国家标准搭建真型实验平台,其接线原理如图4所示。

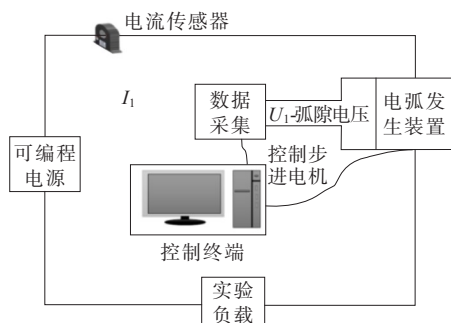


图4 实验平台接线原理

Figure 4 Wiring diagram of experimental platform

实验平台主要由可编程电源、电弧发生装置、数据采集、实验负载等模块组成。可编程电源和故障电弧发生装置均由电脑控制,可编程电源根据需要输出不同电压等级与频率的交流电压;电弧发生装置主要由机箱风扇、步进电机和碳铜电极等组成,碳棒为固定电极,铜棒为移动电极。实验前将碳棒与铜棒紧靠构成正常供电回路,采集正常状态电流,然后缓慢移动铜棒发生电弧采集故障电弧电流,其拉弧距离可由电脑设定自动控制,避免产生人为误差。实验负载模块可分别接入需要模拟的负载进行实验,获取不同实验负载正常和故障电弧电流数据。现场实物如图5所示。

针对 220 V/50 Hz 低压家庭用电场景,选取不同类型生活常用电器作为实验负载,如表1所示。

通过负载内部元件将负载属性分为阻性、感性及非线性负载等类型。由于吹风机发热工作时电阻丝是吹风机内部主要耗电部件,因此将吹风机归类为阻性负载;可控硅负载和计算机因含有电力电子元件,归类为非线性负载;手持电钻由于内部含有电动机,归类为感性负载。

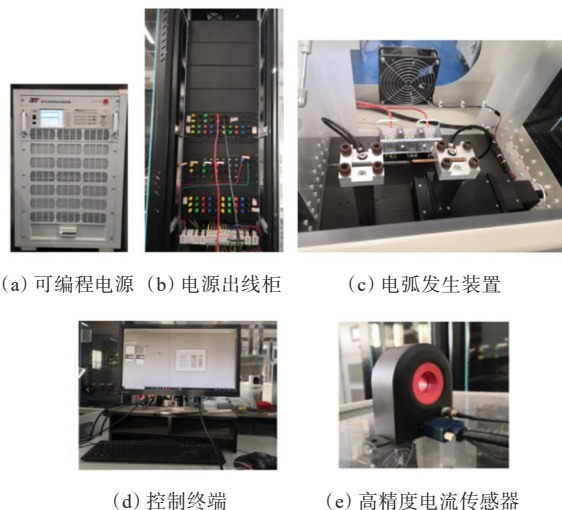


图5 实验原理与装置实物

Figure 5 Diagrams of experimental devices

表1 实验负载分类

Table 1 Classification of experimental load

负载属性	负载名称	电气参数
阻性	吹风机	220 V, 1 000 W
非线性	电子调节器	220 V, 300 W
	计算机	220 V, 225 W
感性	电钻	220 V, 700 W

将实验采样频率设置为 10 kHz,通过高频专业数据采集装置采集具有时间序列的正常和故障电弧 2 种状态实验数据,并进行有效数据选取后得到负载最终样本个数,如表2所示。

表2 样本个数

Table 2 Number of samples

状态	吹风机	电子调节器	计算机	电钻
正常	14 603	22 402	2 870	14 603
故障	15 001	22 399	2 230	15 001

2.2 实验数据处理与分析

为消除环境噪声和不同量纲的影响,对数据进行去噪、归一化等预处理,并利用 Python 中的 Matplotlib 绘图库完成预处理后的数据样本波形绘制,分析波形变化规律。

1) 数据去噪。

在测量小功率负载电流信号时,环境噪声信号容易对其产生较大影响,为避免噪声影响,选择sym8小波函数对电流信号进行6层分解,并利用MATLAB自带的基于heursure的启发式阈值选取函数,自动选取阈值,去除小波系数小于阈值的信号,并对保留信号进行信号重构,进而实现信号去噪^[17-18]。以阻性负载为例,去噪前、后波形如图6所示,可以看出,去噪后波形明显变得更加平滑,波形毛刺显著减少。

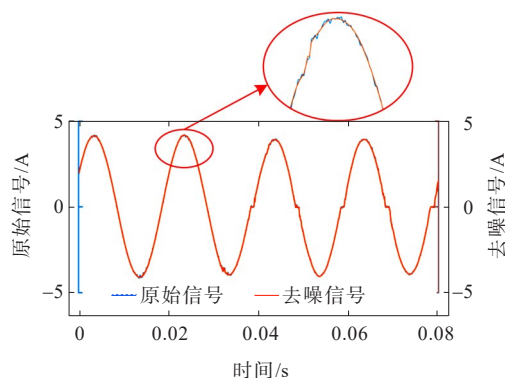


图6 信号去噪前、后波形对比

Figure 6 Comparison waveforms before and after signal denoising

2) 归一化。

为消除不同负载和特征之间不同量纲的影响,使模型能够更快收敛,将各负载电流样本数据经线性变换归一化为 $[-1, 1]$,具体公式如下:

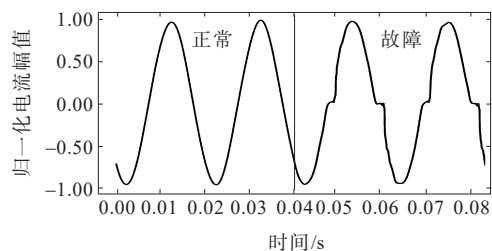
$$x^* = \frac{2 \times \left(x - \frac{x_{\max} + x_{\min}}{2} \right)}{x_{\max} - x_{\min}} \quad (5)$$

式中, x 、 x^* 分别为归一化前、后的对应样本值; $x_{\max}$ 、 $x_{\min}$ 分别为样本数据的最大、最小值。

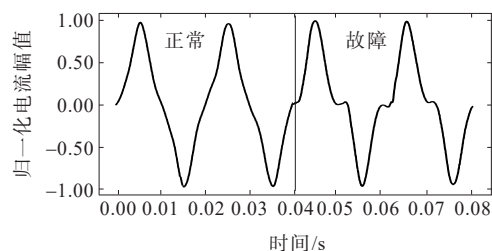
3) 波形分析。

通过对多次实验现象的观察,电弧燃烧剧烈程度受负载功率影响,负载功率越大燃烧越剧烈。几种负载正常和故障电弧波形如图7所示,可以看出,正常运行工况下,吹风机电流波形呈现为平滑正弦波,电钻波形则类似三角波,故障电弧发生后两者波形在过零点处均伴随有“零休”现象出现。计算机正常运行时波形呈现类似正弦波轮廓,故障电弧运行工况下波形出现许多毛刺,有大量尖峰脉冲出现。开关类电子调节器负载故障电弧前、后波形变化不大,且在电流过零时均有“零休”现象出现。因此,利用电流过零时间或电流过零率等单一时域特

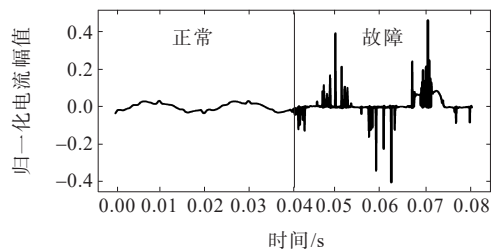
征难以实现故障电弧的准确辨识,容易造成误判和漏判现象发生。



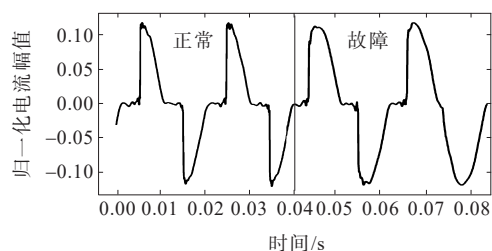
(a) 吹风机负载



(b) 电钻负载



(c) 计算机负载



(d) 电子调节器负载

图7 正常和故障电弧波形分析

Figure 7 Waveforms of normal and fault arc

针对以上问题,本文采用时频域相结合的特征分析方法,对线路电流多个维度特征进行分析,并提取重要特征构建故障电弧隐患预测特征数据集。

3 多维特征分析与提取

为提高特征泛化能力,解决单一特征量作为故障特征时容易失效的问题,本文选取基波、三次谐波、五次谐波3个频域特征,过零指标(一个采

样周期内电流在过零区间的样本点个数,可以反映“零休”现象的持续时间)、峰态2个时域特征以及香农熵等共6个特征值形成特征向量,构建故障电弧隐患预测特征数据集。同时,为证明所提特征的有效性,对以上多维特征进行分析,各类负载故障前、后特征值如表3所示,在故障前、后,4类负载基波值均减小,三次谐波和五次谐波值均明显增大。吹风机、电钻和计算机3类负载在故

障电弧发生后电流过零指标明显增加;计算机在故障发生后峰态值显著增大,其他3类负载峰态值在故障前、后也有明显的增大趋势;4类负载故障前、后香农熵均发生明显变化。可见6个特征指标与故障电弧是否发生存在较强关联性,可作为故障电弧隐患预测特征指标,从而说明本文选取的特征指标具有有效性,可用于构建故障电弧隐患预测特征数据集。

表3 特征数据集
Table 3 Feature data sets

负载类型及状态	基波	三次谐波	五次谐波	过零指标	峰态	香农熵
吹风机正常	96.729 10	1.558 96	0.160 13	1	1.528 56	17.837 07
	...	...	...	...	...	...
	96.140 53	1.885 27	0.094 15	2	1.537 08	17.909 20
	96.137 27	1.889 63	0.089 89	2	1.537 16	17.909 66
吹风机故障电弧	90.116 18	5.444 60	2.756 17	7	1.609 93	18.357 83
	...	...	...	...	...	...
	90.373 41	5.777 67	2.588 77	7	1.617 30	18.158 87
	90.370 92	5.781 88	2.588 82	7	1.617 41	18.158 05
电钻正常	75.008 07	15.399 40	2.576 11	2	2.039 70	17.173 75
	...	...	...	...	...	...
	75.018 25	15.355 50	2.651 74	1	2.039 49	17.184 48
	75.018 30	15.354 40	2.654 28	1	2.039 51	17.183 93
电钻故障电弧	69.447 16	21.471 30	1.645 38	7	2.295 00	14.655 58
	...	...	...	...	...	...
	69.098 03	22.258 40	1.632 47	8	2.305 55	14.420 18
	68.955 29	22.551 80	1.800 04	9	2.310 73	14.327 96
计算机正常	3.180 40	0.506 95	0.418 54	0	1.577 49	0.342 40
	...	...	...	...	...	...
	3.456 06	0.531 95	0.329 82	0	1.519 92	0.394 72
	3.454 37	0.526 21	0.332 11	0	1.523 15	0.394 18
计算机故障电弧	2.508 92	0.121 58	0.529 50	80	29.633 06	0.731 75
	...	...	...	...	...	...
	2.423 06	0.096 20	0.626 13	83	30.714 57	0.786 63
	2.422 83	0.095 27	0.626 80	83	30.715 07	0.786 6
电子调节器正常	10.093 22	0.225 42	0.131 53	2	1.533 31	2.192 44
	...	...	...	...	...	...
	10.110 59	0.231 31	0.132 98	2	1.535 74	2.197 95
	10.111 40	0.231 25	0.133 65	2	1.535 81	2.198 22
电子调节器故障	9.933 23	0.498 83	0.207 19	3	1.597 91	1.858 90
	...	...	...	...	...	...
	9.935 19	0.494 93	0.204 83	3	1.597 02	1.859 03
	9.935 02	0.495 33	0.205 35	3	1.597 11	1.859 45

4 基于随机森林算法的用电安全隐患预警模型

安全隐患预警是一种在故障发生之前进行预警信息反馈的机制。本文构建故障电弧隐患预警模型,通过当前时刻电流数据信息对未来0.02 s是否发生故障电弧作出预测,有利于在故障电弧发生前进行正确处理,将故障电弧消灭在萌芽时刻。采用训练速度快、能够在大数据集上运行且不易过拟合的随机森林算法进行模型搭建,如图8所示,具体步骤如下。

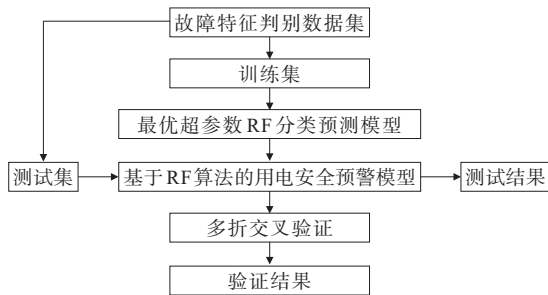


图8 用电安全隐患预警模型构建流程

Figure 8 Flowchart of warning model construction for electrical safety hazards

1) 数据获取。

通过真型实验采集具有时间序列的正常和故障电弧数据,并对数据进行去噪和归一化等预处理,并通过时间窗对实验数据进行样本化处理。

2) 特征工程。

通过对实验数据进行特征分析,提取与故障电弧相关程度较高的特征指标,并构建多维故障电弧隐患预测特征数据集。本文选取特征为电流基波、三次谐波、五次谐波3个频域特征,过零指标、峰态2个时域特征以及香农熵。

3) 模型超参数设置。

随机森林超参数相对较少,根据网格搜索具有操作简单且局部范围内可找到最优超参数等优点,本文选择网格搜索局部寻优方法,对模型超参数进行优化。

4) 模型数据拟合。

得到具有时间序列的样本化特征数据集后,按4:1的比例将故障电弧隐患预测特征数据集划分为训练集和测试集,并将训练集读入模型,使模型在最优超参数环境下完成训练集数据拟合,并得到学习好的故障电弧隐患预警模型。

5) 模型性能评估。

通过测试集对已经训练好的隐患预警模型进行预测准确率及训练时间评估,同时,使用多折交叉验证并完成模型泛化能力的测试,判断是否进行过拟合评估。

5 实验验证

5.1 整体预测结果分析

对实验数据进行预处理后构建故障电弧隐患预测数据集,并按照4:1的比例划分为训练数据集和测试数据集。模型超参数设置:决策树个数 $n_estimators$ 为10,树的最大深度 max_depth 为8,最大特征数 $max_features$ 为3,叶子节点最少样本数 $min_samples_leaf$ 为1。利用信息熵(entropy)来选择最合适的节点,采用真实实验获取的数据样本对模型行进行训练和测试,测试结果如表4所示。

表4 测试结果

Table 4 Test results

负载类型	样本个数			准确率/ %
	测试	预测正确	预测错误	
吹风机	5 920	5 890	30	99.49
电钻	5 920	5 902	18	99.70
计算机	1 020	1 017	3	99.71
电子调节器	8 960	8 924	36	99.59

由表4可知,对于通过时域波形较难辨识是否发生故障电弧的电子调节器负载,本文所提方法预测准确率可达99.59%,且在不同负载类型和样本个数条件下,预测准确率均能达到99.40%以上,说明模型具有有效性。为进一步对模型进行泛化能力测试,在4种不同负载下分别对模型进行多折交叉验证,结果如图9所示,可以看出,交叉验证准确率均超过99.55%,说明模型训练未过拟合,且具有较强泛化能力。

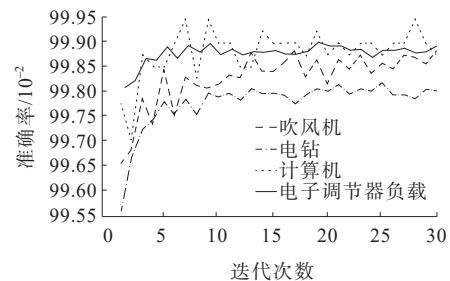


图9 交叉验证曲线

Figure 9 Cross-validation curves

5.2 模型对比分析

为进一步体现模型整体优越性,在同一训练和测

试样本集下,选取门控循环神经网络(gated recurrent neural network, GRU)、循环神经网络(recurrent neural network, RNN)、决策树(decision tree, DT)、BP(back propagation)神经网络等4个主流模型进行对比分析,测试结果如表5所示。

表5 模型对比分析

Table 5 Comparison results of models %

模型类别	预测准确率			
	吹风机	电钻	计算机	电子调节器
GRU	97.22	96.97	88.17	76.76
RNN	96.86	96.60	86.71	78.86
DT	97.13	97.73	91.04	92.15
BP	99.30	99.23	98.13	97.15
本文方法	99.49	99.70	99.71	99.59

由表5中数据可知,在吹风机和电钻2种负载运行工况下,5种模型均能达到较好的预测效果,而在计算机和电子调节器等非线性负载运行工况下,GRU、RNN、DT、BP这4个模型预测准确率均有明显下降,泛化能力较差。本文所提方法在4种不同类型负载运行工况下均表现出较强优势,预测准确率均为最高,进一步体现了模型的优越性。

5.3 故障临界区测试结果分析

故障临界区是体现模型预测效果的重要区域。为验证模型在故障临界区的预测效果,提取故障发生点前3个周期与后2个周期的数据样本单独制作样本子集,并随机打乱顺序后将样本子集按比例划分为训练集和测试集,在不改变模型超参数的条件下进行模型训练与测试,测试结果如图10所示。

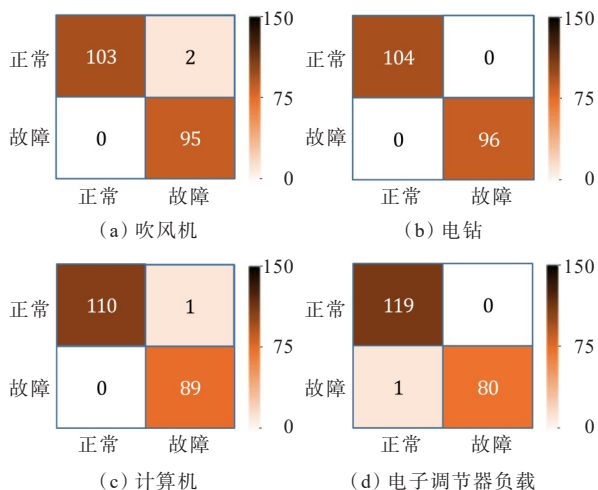


图10 故障临界区测试结果混淆矩阵

Figure 10 Confusion matrix of test results in fault critical region

由图10可以看出,在故障临界区域,电钻预测结果没有出现错误,计算机和电子调节器负载均只出现1个错误预测结果,4种负载预测准确率均达到99.00%以上,说明本文所提方法能够在故障前夕准确预测故障电弧是否发生。

6 结语

故障电弧作为低压用电中的常见安全隐患,存在隐蔽性、随机性和危险性等特点,导致难以在故障前夕有效感知。针对这些问题,搭建真型实验平台,重现多种负载运行工况下的故障电弧场景,并采集实验数据进行数据分析。从故障前、后波形变化规律出发进行特征工程,构建了完整的故障电弧隐患预测多维特征数据集,解决了通过单一特征量进行故障预测容易失效的问题。训练随机森林集成学习模型,并通过局部寻优方法进行超参数优化,以最小化节点信息熵为目标驱动模型参数调整,完成用电安全隐患预警模型搭建。通过实验验证表明,本文所提方法能够实现低压用电安全隐患精准预测,在多种负载运行工况下,本文方法预测准确率高于GRU、RNN、DT、BP等4个主流分类预测模型,为提高居民用电安全系数,减少电气火灾的发生提供了新的方法。

参考文献:

- [1] 高锴. 电气火灾监控系统和电气火灾预防[J]. 建筑电气, 2013, 32(11): 54-56.
GAO Kai. Alarm and monitoring system for electric fire protection and electric fire prevention[J]. Building Electricity, 2013, 32(11): 54-56.
- [2] 刘任鹏, 朱阳, 何畅. 考虑材料不确定性的特高压GIS瓷套管地震易损性研究[J]. 高压电器, 2022, 58(8): 91-98.
LIU Renpeng, ZHU Yang, HE Chang. Study on seismic vulnerability of porcelain bushing for ultra-high-voltage gis considering material uncertainty[J]. High Voltage Apparatus, 2022, 58(8): 91-98.
- [3] 徐云聪, 张逸, 林才华, 等. 基于电气特性-物理参数耦合的交流电弧炉模型[J]. 电工技术学报, 2024, 39(6): 1643-1655.
XU Yuncong, ZHANG Yi, LIN Caihua, et al. AC arc furnace model based on coupling of electrical and physical parameters[J]. Transactions of China Electrotechnical Society, 2024, 39(6): 1643-1655.
- [4] 刘旭晖, 朱振宇, 郭玥, 等. 一种电弧引燃电缆的电气火源模拟方法[J]. 电力工程技术, 2022, 41(2): 136-142.
LIU Xuhui, ZHU Zhenyu, GUO Yue, et al. Electrical fire

- source simulation for igniting cable[J]. Electric Power Engineering Technology, 2022, 41(2): 136-142.
- [5] 孙宇晗,程永锋,卢智成,等.特高压换流站高压换流变及防火墙上设备减隔震分析研究[J].高压电器, 2022, 58(8): 204-212.
- SUN Yuhang, CHENG Yongfeng, LU Zhicheng, et al. Analysis and research on vibration control and isolation of high voltage converter transformer and equipment on firewall in UHV converter station[J]. High Voltage Apparatus, 2022, 58(8): 204-212.
- [6] BAO G H, GAO X Q, JIANG R, et al. A novel differential high-frequency current transformer sensor for series arc fault detection[J]. Sensors, 2019, 19(17): 3649.
- [7] 郭志伟,徐子涛,刘波,等.基于最优特征量选取的开关柜故障判别方法研究[J].电测与仪表, 2023, 60(8): 85-91.
- GUO Zhiwei, XU Zitao, LIU Bo, et al. Research on switchgear fault judgment method based on optimal feature quantity selection[J]. Electrical Measurement & Instrumentation, 2023, 60(8): 85-91.
- [8] 林奕夫,叶兆平,刘冬晨,等.高压开关柜局部放电 TEV 信号传播衰减特性[J].电力工程技术, 2023, 42(5): 224-231.
- LING Yifu, YE Zhaoping, LIU Dongchen, et al. Propagation and attenuation of TEV signals by partial discharge in high voltage switchgear[J]. Electric Power Engineering Technology, 2023, 42(5): 224-231.
- [9] 丁锐,陈羽,孙伶雁,等.基于邻波电流差与随机性的交流串联电弧故障识别[J].电力系统保护与控制, 2023, 51(8): 169-178.
- DING Rui, CHEN Yu, SUN Lingyan, et al. Series arc fault detection in low-voltage AC power lines based on absolute difference of the neighboring waveform of the current and randomness[J]. Power System Protection and Control, 2023, 51(8): 169-178.
- [10] 刁晓虹,董昊,侯磊,等.基于 CEEMDAN-HT 和 SAE 的直流故障电弧诊断方法研究[J].电网与清洁能源, 2023, 39(10): 1-8+18.
- DIAO Xiaohong, DONG Hao, HOU Lei, et al. Research on the DC fault arc detection method based on CEEMDAN-HT and SAE[J]. Power System and Clean Energy, 2023, 39(9): 1-8+18.
- [11] 巩泉役,彭克,陈羽,等.基于电弧随机性和卷积网络的交流串联电弧故障识别方法[J].电力系统自动化, 2022, 46(24): 162-169.
- GONG Quanyi, PENG Ke, CHEN Yu, et al. Identification method of AC series arc fault based on randomness of arc and convolutional network[J]. Automation of Electric Power Systems, 2022, 46(24): 162-169.
- [12] 李响,张炜祺,赵瑞锋,等.小电阻接地系统弧光接地故障建模与特征计算方法[J].供用电, 2022, 39(6): 27-32+39.
- LI Xiang, ZHANG Weiqi, ZHAO Ruifeng, et al. Modeling and characteristic calculation method of arc ground fault in low resistance grounding system[J]. Distribution & Utilization, 2022, 39(6): 27-32+39.
- [13] 余琼芳,路文浩,杨艺.基于深度长短时记忆网络的多支路串联故障电弧检测方法[J].计算机应用, 2021, 41(S1): 321-326.
- YU Qiongfang, LU Wenhao, YANG Yi. Multi-branch series fault arc detection method based on deep long short-term memory network[J]. Journal of Computer Applications, 2021, 41(S1): 321-326.
- [14] 苏晶晶,许志红.基于 EMD 和 PNN 的故障电弧多变量判据诊断方法[J].电力自动化设备, 2019, 39(4): 106-113.
- SU Jingjing, XU Zhihong. Diagnosis method of multi-variable criterion based on EMD and PNN for arc fault diagnosis[J]. Electric Power Automation Equipment, 2019, 39(4): 106-113.
- [15] 封钰,宋佑斌,金晟,等.基于随机森林算法和粗糙集理论的改进型深度学习短期负荷预测模型[J].发电技术, 2023, 44(6): 889-895.
- FENG Yu, SONG Youbin, JIN Sheng, et al. Improved deep learning model for forecasting short-term load based on random forest algorithm and rough set theory[J]. Power Generation Technology, 2023, 44(6): 889-895.
- [16] 王健,易姝慧,刘俊杰,等.基于随机森林算法和稳态波形的非介入式工业负荷辨识[J].中国电力, 2022, 55(2): 82-89.
- WANG Jian, YI Shuhui, LIU Junjie, et al. Non-intrusive industrial load identification based on random forest algorithm and steady-state waveform[J]. Electric Power, 2022, 55(2): 82-89.
- [17] 张静,陶彬彬.基于 MATLAB 的小波去噪的研究[J].辽宁工业大学学报(自然科学版), 2022, 42(3): 177-182.
- ZHANG Jing, TAO Binbin. Research on wavelet denoising based on MATLAB[J]. Journal of Liaoning University of Technology(Natural Science Edition), 2022, 42(3): 177-182.
- [18] 徐明林.基于小波降噪和经验模态分解的滚动轴承故障诊断[D].哈尔滨:哈尔滨工业大学, 2013.
- XU Minglin. Fault diagnosis of rolling bearings based on wavelet noise reduction and empirical modal decomposition[D]. Harbin: Harbin Institute of Technology, 2013.