

引用格式:肖湘奇,肖宇,黄瑞,等.基于随机森林算法的低压系统漏电检测技术研究[J].电力科学与技术学报,2024,39(3):38-47.

Citation: XIAO Xiangqi, XIAO Yu, HUANG Rui, et al. Research on leakage detection technology of low-voltage power system based on random forest algorithm[J]. Journal of Electric Power Science and Technology, 2024, 39(3): 38-47.

基于随机森林算法的低压系统漏电检测技术研究

肖湘奇^{1,2}, 肖宇^{1,2}, 黄瑞^{1,2}, 黄燕娇¹, 贺星^{1,2}, 刘谋海^{1,2}, 慕静茹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, 湖南长沙 410114)

摘要:随着低压配电系统规模及用户需求量迅速增加,用户线路与家庭电气设备漏电故障频发,极易发生人体触电及电气火灾事故。剩余电流保护器作为检测低压漏电故障的常用手段,近年来因线路(或设备)存在对地泄漏电流而频频误动,极大降低了保护设备投运率与可靠性。为此,本文提出一种基于随机森林(random forest, RF)算法的低压配电系统漏电检测技术,为最大程度贴近真实漏电故障场景,充分考虑实际故障场景存在的正常泄漏电流过大、故障邻近支路负荷投切频繁等干扰因素以获得贴近真实故障场景的原始剩余电流数据;通过对原始剩余电流数据进行数据预处理,分析掌握剩余电流的频域与时域特性并利用傅里叶变换算法提取时频域特征,完成低压系统漏电检测模型的建立与训练;在施加多种干扰因素情况下对漏电检测模型进行深度测试,其漏电故障的检测准确率可达99.97%,实现了多种干扰因素情况下的低压配电系统漏电故障检测;最后将支持向量机(support vector machine, SVM)算法、K最近邻(K-nearest neighbor, KNN)算法与本文基于RF算法的漏电故障检测模型的准确性进行对比,以此验证了基于RF算法的低压系统漏电检测模型的准确性与可行性。

关键词:低压系统;漏电检测;随机森林算法;剩余电流

DOI: 10.19781/j.issn.1673-9140.2024.03.005 中图分类号: TM588 文章编号: 1673-9140(2024)03-0038-10

Research on leakage detection technology of low-voltage power system based on random forest algorithm

XIAO Xiangqi^{1,2}, XIAO Yu^{1,2}, HUANG Rui^{1,2}, HUANG Yanjiao¹,
HE Xing^{1,2}, LIU Mouhai^{1,2}, MU Jingru³

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

Measurement and Application Technology, Changsha 410004, China; 3. School of Electrical & Information Engineering,

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

Abstract: With the rapid increase in the scale of low-voltage distribution system and user demand, leakage faults of user lines and household electrical equipment occur frequently, which increase the risk of electric shock and electrical fire accidents. Residual current protector is a common method to detect low-voltage leakage fault. In recent years, due to the existence of leakage current to ground of lines (or equipment), it also frequently false operates, which greatly reduces the operation rate and reliability of protective equipment. To overcome these issues, this paper proposes a leakage detection technology for low-voltage distribution system based on random forest (RF) algorithm. In order to closely simulate the real leakage fault scenario, the original residual current data close to the real fault scenario can be obtained by fully considering the interference factors such as excessive normal leakage current and frequent switching of load in the adjacent branch of the fault scenario. Through data preprocessing of the original residual current data, the

收稿日期: 2022-06-04; 修回日期: 2024-03-17

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等研究; E-mail: 84578610@qq.com

frequency domain and time domain characteristics of the residual current are analyzed, and the time-frequency characteristics are extracted using the Fourier transform algorithm to complete the establishment and training of the low-voltage system leakage detection model. The leakage detection model is tested under the condition of multiple interference factors, and the results show that the detection accuracy of the leakage fault can reach 99.98%, realizing the leakage fault detection of low-voltage distribution system under the condition of multiple interference factors. Finally, support vector machine (SVM) algorithm, K-nearest neighbor (KNN) algorithm and the leakage fault detection accuracy based on random forest algorithm are compared to verify the accuracy and feasibility of the proposed leakage fault detection model of low-voltage system.

Key words: low-voltage system; leakage detection; random forest algorithm; residual current;

在低压配电系统运行中,一旦电路或电气设备漏电,电流会与设备外壳、零线、保护接零线间形成闭合回路,导致漏电电流不断增大、电力资源损失,还会发生人体触电及电气火灾事故^[1-2]。

剩余电流保护装置在中国低压电网中广泛应用,对避免因漏电而引起的电气火灾事故具有非常重要的作用。传统意义上的漏电检测方式主要是通过人工定期巡检,这种方式存在诸多弊端,不仅耗时耗力且不能全方位、无死角、高精度检测漏电电流,仍存在很大的漏电隐患;近年来,对于漏电的检测主要利用剩余电流保护装置,但是剩余电流保护器由于存在正常泄漏电流等原因会频繁出现误动作、拒动作等现象;此外,目前在线运行的剩余电流保护装置,其动作电流的整定值通常是将检测到的剩余电流有效值的大小作为是否动作的唯一判据。理论研究和实际运行经验表明,这种判据无法精确检测到故障漏电电流信号,在动作原理上存在先天不足和缺陷,也经常导致误动或拒动现象的发生,大大降低了剩余电流保护装置的动作可靠性和正确投运率^[3-5]。

针对上述问题,文献[6]提出了一种考虑剩余电流幅值和相位的鉴幅鉴相的漏电保护方法,通过设定固定的动作电流整定值判别漏电与否,但电网中的漏电电流会随外部环境有较大的变化,因此固定动作阈值的漏电保护不能满足实际需要;文献[7]提出用2~5个周时间间隔的剩余电流变化量表示故障漏电,并采用剩余电流和剩余电流变化量双重判据,实现故障漏电的正确保护,但该方法要基于泄漏电流在时间间隔保持不变的前提下才具有可行性,而在实际场景下,漏电流会随外部环境的变化随时改变,变化规律难以分辨,因此该方法也具有一定的局限性;文献[8]对漏电流变化情况进行分析,提出基于浮动阈值的自适应漏电保护方法,但该方法的阈值整定过程较为复杂,花费时间较长。文献[9]通过提取故障信号的时频域特征,

作为支持向量机(support vector machine, SVM)分类模型的输入向量,进行故障类型识别,但该方法中的算法模型泛化能力及鲁棒性较差。文献[10]利用量子遗传计算的快速寻优性和神经计算的自适应性有机结合构建故障检测系统,但该方法中的数据样本数仅有30组,模型泛化能力弱,欠缺说服力。此外,数据的真实性是一切基于智能算法模型的漏电检测的重要前提与关键,上述基于算法的漏电检测方法在数据采集阶段并未加入真实漏电故障场景下存在的干扰因素,如线路或设备的正常泄漏电流过大、故障邻近支路的负荷投切频繁、感知距离改变等对漏电检测的影响。

鉴于此,本文充分使用智能信息处理技术,如傅里叶变换、小波变换、随机森林(random forest, RF)算法等^[11-17],利用课题组低压用电安全实验室构建漏电故障真型物理实验平台,首次考虑实际配用电场景存在的线路(或设备)正常泄漏电流过大、故障邻近支路的负荷投切频繁与故障感知距离过大等干扰因素,获取大量贴近真实漏电故障场景的原始剩余电流数据;分析掌握剩余电流的频域与时域特性并利用傅里叶变换算法提取时频域特征即可完成低压系统漏电检测模型的建立与训练;在各类干扰因素情况下对漏电检测模型进行深度测试,实现多种干扰因素情况下的低压系统漏电故障检测;最后,将SVM算法、K最近邻(K-nearest neighbor, KNN)算法与本文的基于RF算法的漏电故障检测准确性进行对比,验证基于RF算法的低压系统漏电检测模型的准确性与可行性。

1 算法原理

1.1 决策树算法

决策树是RF算法的基础部分,RF是决策树的集合。决策树是一种经典的机器学习算法,由3类节点组成的树形结构分类器包括:根节点、内部节点和终端节点。根节点(决策节点)将整个种群或

样本分为2个或多个同类集;内部节点(机会节点)确定决策树中的某个属性节点;终端节点(叶节点)表示分类后获得的分类标记。决策树结构是以分支关系定义的层次结构,按照根节点—内部节点—叶节点的决策规则从根节点不断拆分到达叶节点展开分支的过程^[18-19]。

决策树生成算法流程图如图1所示。

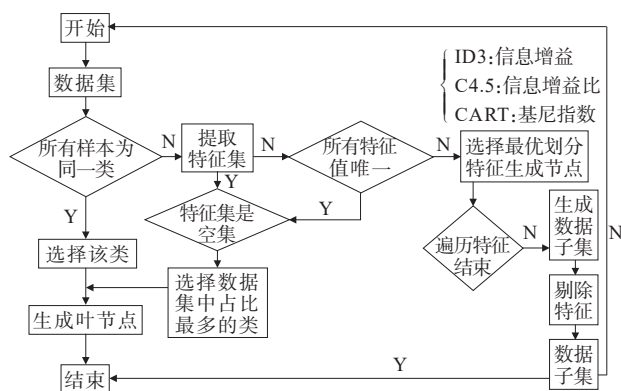


图1 决策树生成算法流程

Figure 1 Decision tree generation algorithm flowchart

决策树划分属性原则主要有基尼指数信息增益、信息增益比,依据其划分原则构建不同的决策树算法,主要有ID3、C4.5、分类回归树(classification and regression tree, CART)算法。

由于CART决策树不依赖特定类型的输入数据,同时也不受数据集中异常值的影响^[20],因此本文使用CART决策树构建算法来创建单个决策树。

CART算法通过递归地构造二叉树,分类树用基尼指数最小化准则,进行特征选择,生成决策树。CART决策树构建步骤如下。

1) 输入训练集、基尼系数的阈值、切分的最少样本个数阈值。

2) 当前节点的数据集为 D ,如果样本个数小于阈值或没有特征,则返回决策子树,当前节点停止递归。

3) 计算样本集 D 的基尼系数,如果基尼系数小于阈值,则返回决策子树,当前节点停止递归。

对某个样本 D ,其样本个数为 N ,根据特征 A 是否取某一可能值 a ,把样本 D 分为左右两部分 D_1 和 D_2 。在属性 A 的条件下,样本 D 的基尼系数计算公式如下:

$$G(N|A=a) = \frac{|D_1|}{N} G(D_1) + \frac{|D_2|}{N} G(D_2) \quad (1)$$

4) 在计算出来的各个特征的各个值的基尼系数中,选择基尼系数最小的特征 A 及其对应的取值 a 作为最优特征和最优切分点。

$$\begin{cases} D_1 = (x, y) \in D | A(x) = a \\ D_2 = D - D_1 \end{cases} \quad (2)$$

数据集 D 的基尼指数的计算公式如下:

$$G(D) = \sum_{i=1}^n p(x_i)(1-p(x_i)) \quad (3)$$

式中, $p(x_i)$ 是分类 x_i 出现的概率; n 为分类的数目。 $G(D)$ 反映了从数据集 D 中随机抽取2个样本,其类别标记不一致的概率。因此, $G(D)$ 越小,则数据集 D 的纯度越高。

5) 对左右的子节点递归调用步骤1)~4),生成CART。

6) 输出CART决策树。

1.2 自助法抽样

在统计学中,自助法(bootstrap)是一种从给定训练集中有放回的均匀抽样方法。RF通过自助法重采样技术,从原始训练样本集 N 中有放回地重复随机抽取 k 个样本(k 一般与 N 相同)生成新的训练样本集,然后根据自助样本集生成 n 个分类树组成RF^[21]。

自助法有放回抽样步骤如下。

1) 在一个含有 N 个样本的原始数据集中,随机抽取一个样本,并标记次数 i 。

2) 将步骤1)随机抽取到的样本重新放回原始数据集中,此时原始数据集仍含有 N 个样本。

3) 重复步骤1)、2),直至抽取到的样本个数 i 与原始数据集中的样本个数 N 相等,即可得到1个与原始数据集数量相同的自助集 $customfeature_k$,用于第 k 个子树的训练。

4) 重复步骤1)、2)、3),直至 k 与RF的决策树数量 n (由参数 $n_estimators$ 决定)相等时,完成训练集的生成;并将原始数据集中未被抽取到的样本作为测试集。自助法有放回抽样示意图如图2所示。

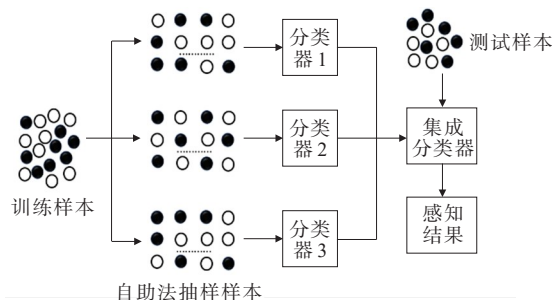


图2 自助法有放回抽样示意图

Figure 2 Bootstrap has a put-back sample

2 建立低压系统漏电检测模型

基于 RF 算法的低压系统漏电检测模型的建立包括漏电检测实验(获取原始剩余电流数据)、原始

剩余电流数据的预处理、特征提取、模型建立和模型评估与测试 5 个部分,每部分之间层层递进,互相依托,共同完成了基于 RF 算法的低压系统漏电检测模型的建立。模型建立框图如图 3 所示。

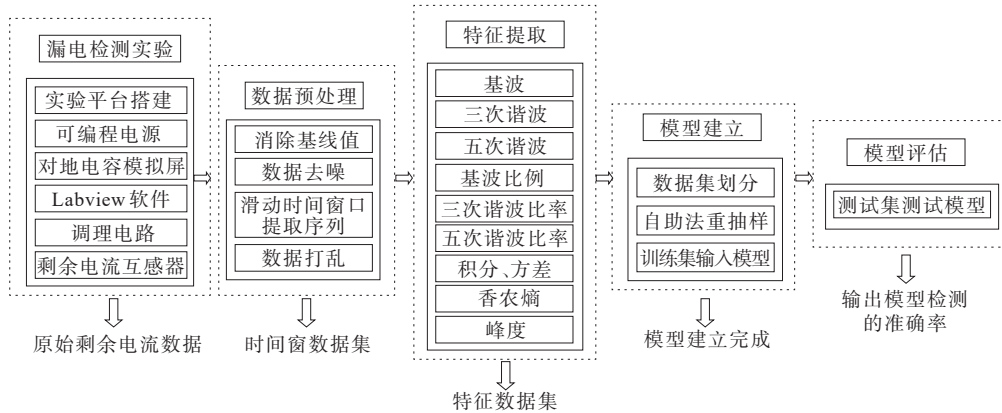


图 3 基于 RF 的低压系统漏电检测模型的建立过程

Figure 3 The process of establishing a leakage detection model for low-voltage system based on random forest

2.1 实验数据来源

实验的原始数据样本来源于课题组低压用电安全实验室搭建的真型漏电物理实验平台,低压系统漏电实验电路示意图如图 4 所示。

漏电检测实验的实验设备包括:可编程电源、Labview 数据采集设备、对地电容模拟屏、小电流互感器、调理电路、家庭用电负载(阻性负载、非线性电力电子型负载)。

剩余电流信号由剩余电流互感器测得,测得的剩余电流经过调理电路将模拟电流信号转化为可用于数据采集、控制过程、执行计算显示读出的数字信号;可编程电源为线路提供相电压为 220 V 的交流电;火线与零线一同穿过剩余电流互感器,当发生漏电或人身触电故障时,火线电流、零线电流与故障电流的关系表达式为

$$I_L + I_N = I_{\Delta r} \quad (4)$$

式中, I_L 为流过相线的相电流; I_N 为零线电流。当发生漏电或人身触电故障时,剩余电流保护器所检测到的电流不为 0,而是 $I_{\Delta r}$ 。通过信号线连接至 Labview 数据采集设备,即可记录线路的正常工作状态和发生漏电故障时的总剩余电流数据。

由于线路在正常工作状态下存在的正常漏电流主要是由线路对地电容产生的,且正常漏电流对剩余电流的大小存在较大的影响,因此在本次真型低压系统漏电故障实验中,利用了课题组对地电容模拟屏模拟了 2 km 线路对地电容,加入了线路对地泄漏电流对低压用户漏电故障检测的干扰因素;根据漏电故障检测的实际感知距离需求,在本实验中将剩余电流传感器与漏电故障发生处之间的感知距离设为 25 m,以此考虑增大故障感知距离

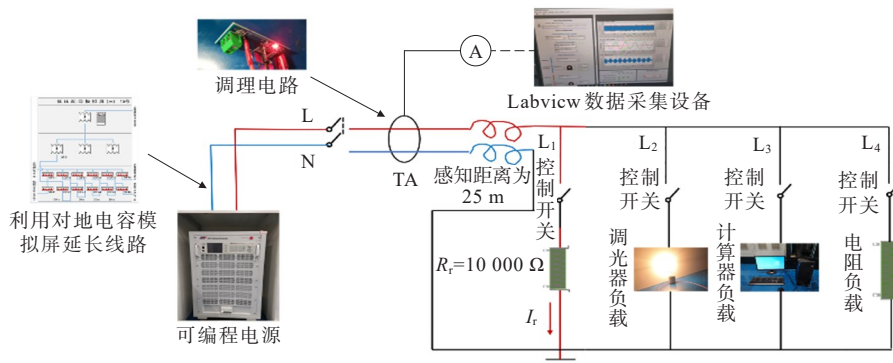


图 4 漏电检测实验电路原理

Figure 4 Schematic diagram of the leakage detection experimental circuit

对漏电故障检测的干扰影响,真型线路局部连接示意图如图5所示。

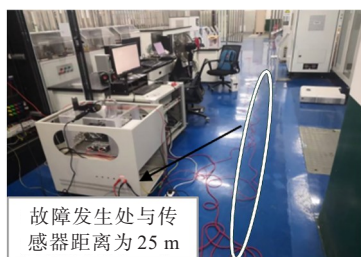


图5 真型线路局部连接示意图

Figure 5 Schematic diagram of the local connection of the true line

支路 L_1 为漏电故障发生支路,可利用漏电控制开关操控漏电故障发生时间,故障电阻设置为 $5\ 000\ \Omega$ 与 $10\ \text{k}\Omega$;支路 L_2 、支路 L_3 与支路 L_4 均为负载支路,分别代表非线性电力电子负载与阻性负载类型,真型实验场景连接局部示意图如图6所示;在采集线路正常与故障剩余电流数据时开合负荷投切开关,加入故障邻近支路的负荷投切对漏电故障检测的干扰因素。

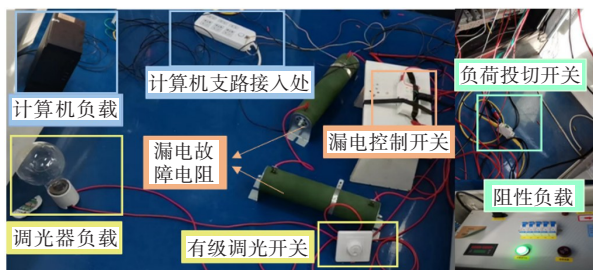


图6 真型实验场景连接局部示意图

Figure 6 Schematic diagram of the connection of the true experimental scene

在已增大故障感知距离与线路对地电容的前提下,操控漏电控制开关与负荷投切开关,原始剩余电流数据采集方法如下(以故障电阻为 $5\ 000\ \Omega$ 时为例)。

1) 控制可编程电源输出 $220\ \text{V}$ 交流电—运行NI数据采集系统—打开调光器负载正常工作—加入干扰(开合电阻支路的负荷投切开关制造干扰)—动作故障漏电支路的漏电控制开关(发生漏电故障)—加入干扰(开合电阻支路的负荷投切开关制造干扰)。

2) 控制可编程电源输出 $220\ \text{V}$ 交流电—运行NI数据采集系统—打开计算机负载正常工作—加入干扰(开合电阻支路的负荷投切开关制造干

扰)—动作故障漏电支路的漏电控制开关(制造漏电故障)—加入干扰(开合电阻支路的负荷投切开关制造干扰)。

通过模拟低压供电系统的漏电故障状态和正常运行状态,数据采集系统从模拟实验中采集总剩余电流数据 I_{Di} 。数据采集系统的采样频率设置为 $10\ \text{kHz}$,频率为 $50\ \text{Hz}$,故每个工频周期包括 200 个采样点,形成原始数据集矩阵 O_{D} 。该矩阵结构为 M 行 3 列,其中第 i 行代表第 i 个采样点时刻的数据,第 1 列表示采样点编号,第 2 列表示总剩余电流数据,第 3 列表示标签值(正常运行状态标签为“ 0 ”、漏电状态标签为“ 1 ”)。

2.2 数据预处理

1) 消除调理电路的基线值。

漏电实验过程中,在实验电路加入直流分量,虽可防止烧毁二极管,确保调理电路不反向,但也致使电流互感器的测量基线值升高。因此,若在保证总剩余电流检测的准确性,需先切断负载,空采线路中的直流数据,对电流互感器进行校准。本次实验的空采时间约为 $5.7\ \text{s}$,空采数据为 $57\ 047$,取全部空采数据的平均值;将所有电流数据减去空采数据平均值即可消除调理电路的基线值,得到准确的总剩余电流数据,具体计算方法如下:

$$x_{\text{T}} = x_i - \frac{x_1 + x_2 + \dots + x_n}{n} \quad (5)$$

式中, x_i 为某一时刻数据样本; n 为空采样本个数; $x_1 + x_2 + \dots + x_n$ 为 n 个空采样本数据。

2) 对原始剩余电流数据进行去噪处理。

漏电实验中采集到的原始剩余电流数据存在噪声,剩余电流波形毛刺较多,如图7(a)所示。利用MATLAB中的小波信号降噪软件对原始剩余电流进行去噪^[22],去噪后的剩余电流波形如图7(b)所示,可见经去噪处理之后的总剩余电流波形更光滑,毛刺较去噪之前更少。

3) 滑动时间窗口提取序列。

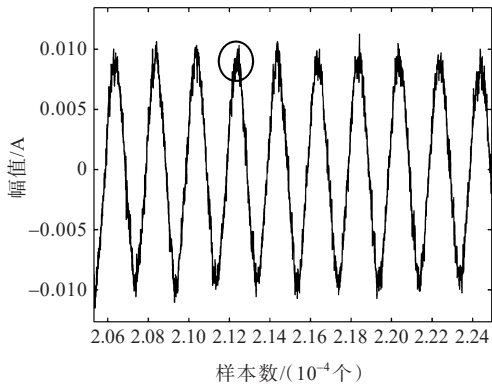
在实际应用中,系统在工作时只与工作点附近的数据有较强的相关性,而与远离工作点区域数据的相关性并不强。因此,本文建立一个滑动时间窗口模型,并保持数据区间长度为 200 个采样点。当有一个新数据加入时,最早的一个数据相应地从上区间滑出,随着数据区间不断更新,模型也相应地由新区间的数据不断更新。

4) 数据乱序、归一化处理。

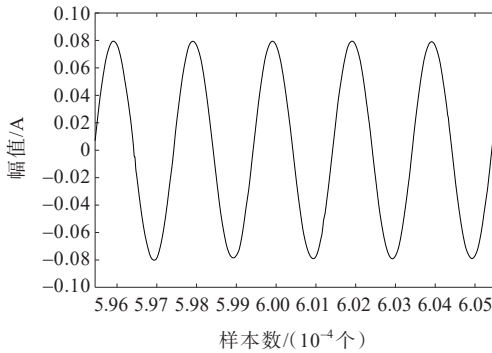
为提高低压系统漏电检测模型的泛化能力和精度,需对时间窗数据进行打乱;使用Excel的rand()函数对各个时刻的数据特征赋予随机数并对其进行排序,以此起到打乱时间窗数据集(时序性)的作用;为解决数据指标之间的可比性问题,采用最大-最小标准化法对数据进行归一化处理:

$$x' = \frac{2 \cdot \left(x_i - \frac{x_{\max} + x_{\min}}{2} \right)}{x_{\max} - x_{\min}} \quad (6)$$

式中, x_i 为某一时刻数据样本; x_{\max} 、 x_{\min} 分别为所有样本中的最大样本数据与最小样本数据。



(a) 未经去噪处理



(b) 经去噪处理

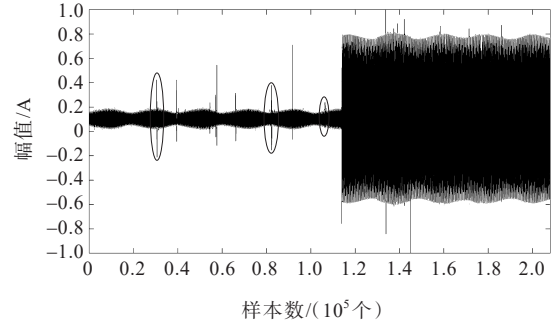
图7 原始剩余电流去噪处理

Figure 7 Raw residual current denoising processing

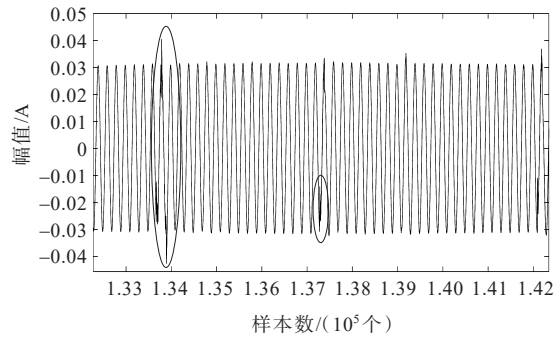
2.3 剩余电流波形特征

利用MATLAB对原始漏电电流数据进行分析。正常状态—漏电状态的剩余电流整体发展过程如图8(a)所示,正常状态—漏电状态的剩余电流变化特点细节如图8(b)所示。由图8可知,由于在实验过程中施加外部干扰因素,因此剩余电流波形中出现幅值突变与无规律波动现象,当漏电故障阻值较大,且存在负荷投切等干扰的情况下,正常情况下的突变幅值甚至可超过故障状态下的剩余电

流幅值,因此仅凭幅值等单一特征无法精准检测漏电故障。



(a) 整体发展过程



(b) 变化特点细节图

图8 正常至漏电故障的剩余电流变化特征

Figure 8 Characteristic of the residual current change from normal to leakage

2.4 特征提取

为增加模型检测的效率,降低计算机的运行内存,对原始数据进行降维处理,即特征提取,可用少数、更具有代表性的特征数据反映整体原始数据的效果。利用傅里叶变换的方法从原始时间窗口数据表中提取不同特征,包括基波、3次谐波、5次谐波、基波比率、3次谐波比率、5次谐波比率、积分、香农熵、峰度、方差、对特征提取后的数据集按照7:1:2的比例进行训练集、验证集和测试集的划分。选取208 890组数据作为训练集,146 223组数据作为验证集,剩余的41 778组数据作为测试集。

上述特征在正常运行状态和漏电故障状态下的10种特征分布如图9所示。其中,[0,1]为正常状态区间;[1,2]为漏电故障状态区间。基于验证集对特征分布进行可视化处理后,可看出单维特征值分布区域具有重叠部分(以图9(a)为例,图中圈出标注部分即为故障特征分布重叠部分),无法精确检测漏电故障发生与否,因此本文将构建多维故障特征向量用于漏电故障检测模型的建立。

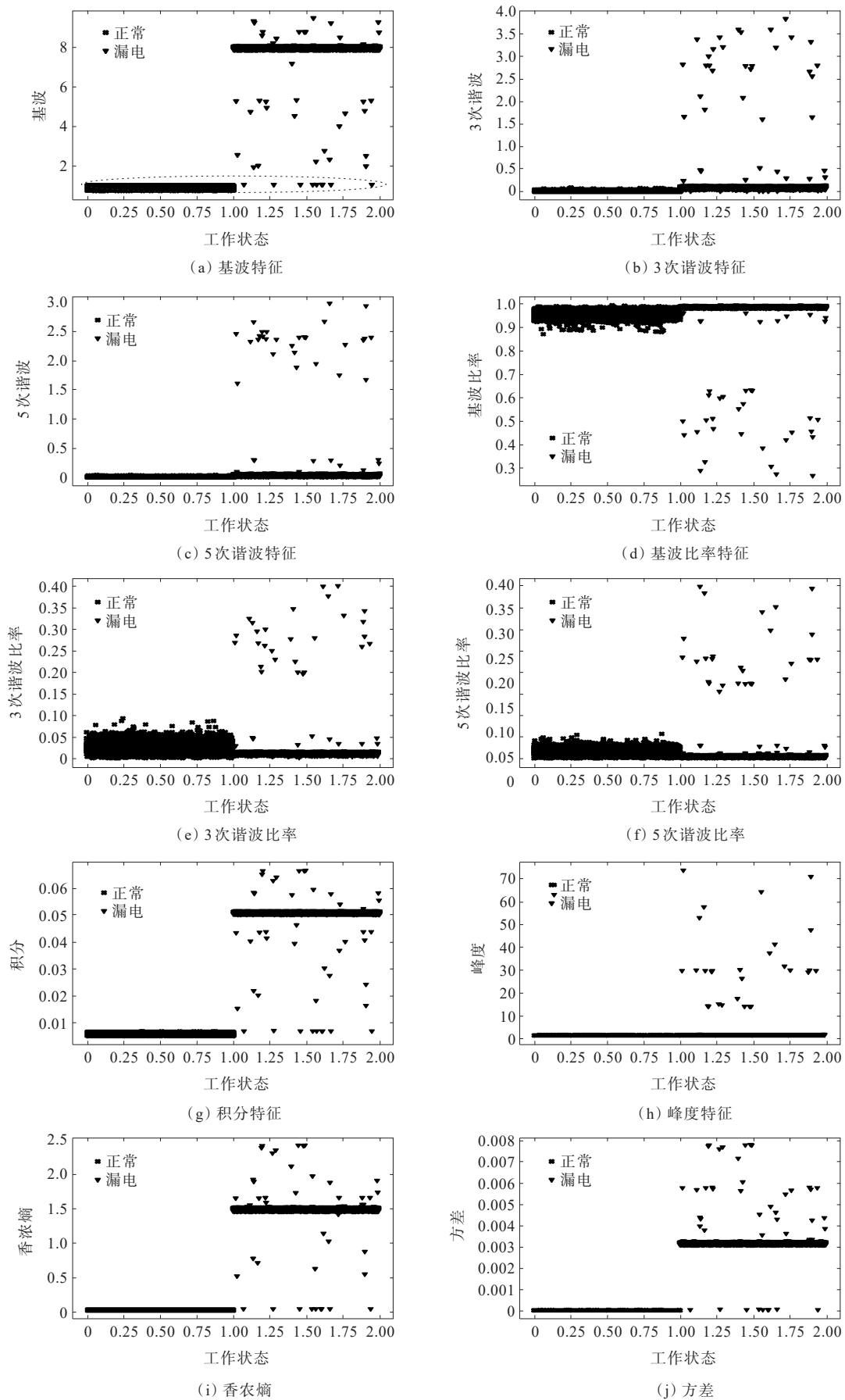


图9 特征分布可视化图示

Figure 9 Visualization of feature distributions

2.5 建立基于 RF 的漏电故障检测模型

利用 RF 算法展开建模,建模过程如图 10 所示。首先,将特征数据集分为训练集与测试集;其次,采用自助法重抽样从训练样本集中随机抽取多个训练样本子集,对每个子集分别进行决策树建模,综合多棵树的决策结果,通过投票最终确定检测结果,并完成低压系统漏电检测模型的建立。

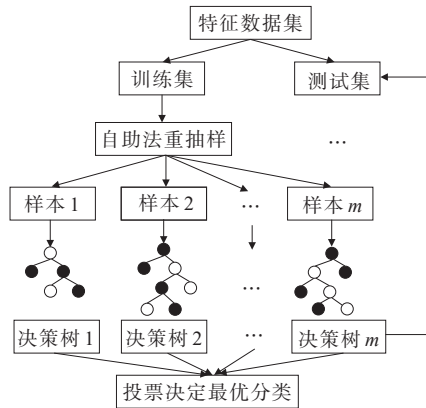


图 10 基于 RF 算法的低压漏电检测模型的建立

Figure 10 Establishment of low-voltage leakage detection model based on random forest algorithm

将训练集作为 RF 模型的输入,即可训练得到低压系统漏电检测模型,最后利用测试集测试低压系统漏电检测模型的漏电检测性能。经测试,低压供电系统的漏电检测模型准确率可达 99.98%。

利用 RF 回归进行模型的构建, $n_estimators$ 为决策树的个数; $random_state$ 为每一个树利用 Bagging 策略中的自助法进行重抽样(即有放回的袋外随机抽样)时,随机选取样本的随机数种子; fit 函数进行模型的训练, $predict$ 函数进行模型的预测。

1) 导入相关包。

从 `sklearn_ensemble` 中导入 RF 分类器;从 `sklearn.datasets` 中使用 `load()` 函数导入数据集;导入 `pandas` 包中的 `read_excel` 方法用于读取数据;导入 `matplotlib.pyplot` 模块可用于基分类器与对应的分类准确率变换过程可视化。

2) 导入并读取数据集,设置相关参数。

导入线路在正常运行和发生漏电故障状态下的数据集,即 `data=load_OD()`;使用导入的 `pandas` 包中的 `read_excel` 方法读取 OD 数据集。

相关参数设置如表 1 所示。

3) RF 模型的训练。

参数设置完毕后,使用 `fit` 函数进行模型的训

练,以训练数据为参数,进入训练阶段,即可输出训练完毕的 RF 模型。

表 1 RF 模型参数设置

Table 1 Table of random forest model parameter settings

参数名称	参数设置
<code>n_estimators</code>	9
<code>oob_score</code>	false
<code>max_depth</code>	12
<code>max_features</code>	3
<code>bootstrap</code>	true
<code>criterion</code>	entropy
<code>random_state</code>	1

2.6 模型的评估与使用

使用 `score` 函数随机在测试集中对森林模型的性能进行评估,输出相应评估值为 99.98%,将其作为模型分类能力的评估指标,该评估值表示基于 RF 算法的低压系统漏电检测模型能力合格。

将测试特征数据集作为 RF 模型的输入,模型中每个决策树将会输出一个工作状态判别结果,遵循“少数服从多数”的原则对所有决策树的识别结果进行统计,从而将结果占比最大的感知状态作为低压系统漏电检测模型的最终结果。使用 `predict` 函数以测试集为模型的输入,查看输出,进而得到对应的类别(正常运行标签为“0”、漏电状态标签为“1”)。

3 模型深度测试与对比

为证明本文建立的基于 RF 模型的低压漏电故障检测模型的优势与模型的稳定性,进行多次真实漏电实验,选取有效数据样本对已经搭建好的漏电检测模型进行多次深度测试。具体方法如下:选取上万个与训练数据样本没有交叉的实验数据,并对其数据进行乱序处理,加入干扰样本(分别在正常状态与漏电故障状态工作操控故障邻近支路的负荷投切开关,即为负荷投切频繁型干扰;增大对地电容模拟屏的线路长度以增大线路对地泄漏电流,即为正常泄漏电流过大型干扰;增大漏电故障感知距离,即为感知距离增大型干扰)。

此外,将基于 RF 算法的漏电检测模型(本文所用方法)与基于邻近算法和基于支持向量机的漏电检测模型的检测准确率进行对比^[23-25]。其中,检测模型准确率=(正确预测样本数/总预测样本数)×100%。

将测试集标签和3种模型测试集输出标签构成的二维矩阵分别作为混淆矩阵算法^[26]模型的输入,输出3种模型的混淆矩阵图示,如图11所示。

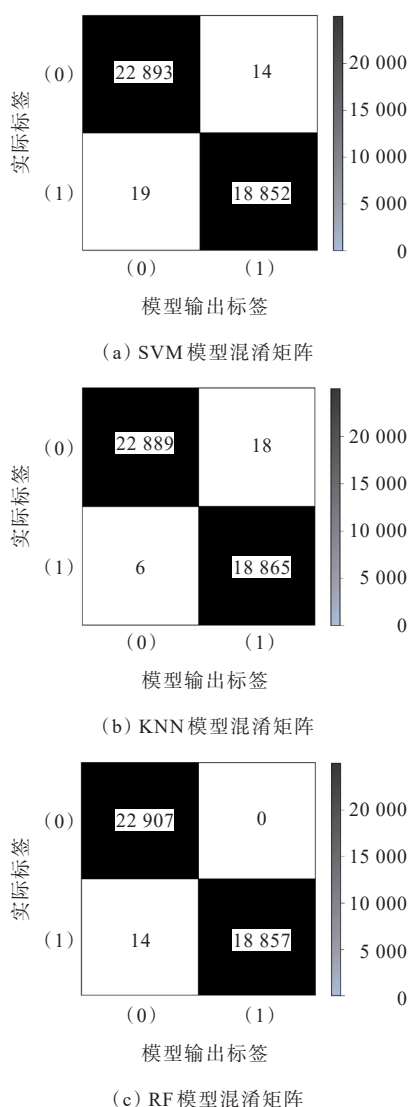


图 11 3种模型的混淆矩阵图示

Figure 11 Confusion matrix for three models

由图11的3种模型的混淆矩阵可得如下结论。

1) SVM模型的错误识别次数为33次,其中漏电状态被识别为正常状态19次,正常状态被识别为故障状态14次。其在3个模型中识别准确率最低,SVM漏电检测模型的检测准确率为99.92%。

2) KNN模型的错误识别次数为24次,其中漏电状态被识别为正常状态6次,正常状态被识别为漏电状态18次,其漏电检测模型的检测准确率为99.94%。

3) RF模型的错误识别次数为14次,其中漏电状态被识别为正常状态14次,正常状态被识别为漏电状态0次,RF漏电检测模型的检测准确率为

99.97%。

4) 通过将RF模型与KNN、SVM模型相比,RF模型的漏电检测准确率最高,可达99.97%,满足故障检测合格标准,识别错误样本数最少。

4 结语

本文基于低压漏电检测真型物理实验平台,充分考虑实际用电场景存在的正常泄漏电流过大、故障邻近支路负荷投切频繁等干扰因素,最大程度采集贴近真实故障场景的原始剩余电流数据;通过将总原始剩余电流数据进行数据预处理(消除基线值、数据去噪、滑动事件窗口提取序列、数据乱序处理、数据集划分),利用傅里叶变换算法对数据集进行时频域特征提取,完成低压系统漏电检测模型的建立与训练;最后利用包含干扰样本的实时数据对漏电检测模型进行深度测试,得出以下结论:

1) 本文所建立的基于RF算法漏电检测模型适用于包含多种干扰因素的真实漏电故障场景,且漏电故障检测效果较好;

2) 通过对单个特征进行特征分布可视化处理,可看出单维特征向量无法清晰准确地区分正常状态与故障状态,本文通过构建多维特征向量,可达到很好的漏电故障检测效果;

3) 通过将本文的基于RF算法的低压系统漏电检测模型与基于SVM、基于KNN算法的漏电检测模型对比,基于RF算法的低压系统漏电检测模型的检测准确率均高于其余2种模型。

参考文献:

- [1] 朱领军. 漏电检测的研究与分析[D]. 杭州:浙江大学, 2018.
ZHU Lingjun. Research and analysis of leakage detection [D]. Hangzhou: Zhejiang University, 2018.
- [2] 汪倩, 陈思磊, 刘家欣, 等. 低压直流系统剩余电流检测技术研究综述[J]. 高压电器, 2023, 59(6): 1-12.
WANG Qian, CHEN Silei, LIU Jiabin, et al. Review on research of residual current detection technology in low voltage DC systems[J]. High Voltage Apparatus, 2023, 59(6): 1-12.
- [3] 韩晓慧. 低压配电网剩余电流保护优化算法研究[D]. 北京: 中国农业大学, 2017.
HAN Xiaohui. Research on the optimization algorithm of residual current protection for low-voltage distribution network[D]. Beijing: China Agricultural University, 2017.
- [4] 余磊, 贾科, 温志文, 等. 计及量测数据丢失的主动配电网电流保护自适应整定方法[J]. 电力系统自动化, 2022,

- 46(15):145-152.
YU Lei, JIA Ke, WEN Zhiwen, et al. Adaptive setting method of over current protection in active distribution network considering measurement data missing[J]. Automation of Electric Power Systems, 2022, 46(15): 145-152.
- [5] 吴行健, 石晶, 陈红坤, 等. 基于短路电流小半波特征的短路故障快速检测方法[J]. 电力系统保护与控制, 2022, 50(16): 86-94.
WU Xingjian, SHI Jing, CHEN Hongkun, et al. Fast short-circuit fault detection method based on a small half-wave feature[J]. Power System Protection and Control, 2022, 50(16): 86-94.
- [6] 张福良, 蔡志远. 鉴幅鉴相式剩余电流保护方法的研究[J]. 低压电器, 2014(8): 17-19.
ZHANG Fuliang, CAI Zhiyuan. Research on amplitude and phase discrimination method of residual current protection[J]. Low Voltage Apparatus, 2014(8): 17-19.
- [7] 武一, 李奎, 岳大为, 等. 消除剩余电流保护动作死区的理论与方法[J]. 电工技术学报, 2008, 23(6): 44-49.
WU Yi, LI Kui, YUE Dawei, et al. Theory and method of eliminating operating dead zone of residual current protection[J]. Transactions of China Electrotechnical Society, 2008, 23(6): 44-49.
- [8] 李奎, 郝运金, 赵成晨, 等. 基于马尔可夫过程的漏电保护器服役状态及其可靠性分析[J]. 电工技术学报, 2023, 38(18): 5061-5076.
LI Kui, HAO Yunqian, ZHAO Chengchen, et al. Service states and reliability analysis of residual current device based on Markov process[J]. Transactions of China Electrotechnical Society, 2023, 38(18): 5061-5076.
- [9] 韩晓慧, 杜松怀, 苏娟, 等. 触电信号暂态特征提取及故障类型识别方法[J]. 电网技术, 2016, 40(11): 3591-3596.
HAN Xiaohui, DU Songhuai, SU Juan, et al. Fault transient feature extraction and fault type identification for electrical shock signals[J]. Power System Technology, 2016, 40(11): 3591-3596.
- [10] 王金丽, 刘永梅, 杜松怀, 等. 基于剩余电流固有模态能量特征的生物触电故障诊断模型[J]. 农业工程学报, 2016, 32(21): 202-208.
WANG Jinli, LIU Yongmei, DU Songhuai, et al. Fault diagnosis model for biological electric shock based on residual current intrinsic mode function energy features[J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(21): 202-208.
- [11] 赵凤展, 杨仁刚. 基于短时傅里叶变换的电压暂降扰动检测[J]. 中国电机工程学报, 2007, 27(10): 28-34+109.
ZHAO Fengzhan, YANG Rengang. Voltage sag disturbance detection based on short time Fourier transform[J]. Proceedings of the CSEE, 2007, 27(10): 28-34+109.
- [12] 邓志祥, 潘建兵, 黄灿英, 等. 基于傅里叶变换与最大关联距离相结合的配电网故障选线新方法[J]. 电网与清洁能源, 2022, 38(4): 95-103.
DENG Zhixiang, PAN Jianbing, HUANG Canying, et al. A novel fault line selection method of distribution network based on FFT and maximum correlation distance[J]. Power System and Clean Energy, 2022, 38(4): 95-103.
- [13] 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.
- [14] 岳首志, 洪海生, 邓祺, 等. 基于曲线特征聚类与信息聚合的电力负荷预测[J]. 长沙理工大学学报(自然科学版), 2023, 20(6): 128-139.
YUE Shouzhi, HONG Haisheng, DENG Qi, et al. Power load forecasting based on curve feature clustering and information aggregation[J]. Journal of Changsha University of Science & Technology (Natural Science), 2023, 20(6): 128-139.
- [15] JURADO F, SAENZ J R. Comparison between discrete STFT and wavelets for the analysis of power quality events[J]. Electric Power Systems Research, 2002, 62(3): 183-190.
- [16] 邵能灵, 侯志俭, 李涛, 等. 基于小波分析的电力系统短期负荷预测方法[J]. 中国电机工程学报, 2003, 23(1): 45-50.
TAI Nengling, HOU Zhijian, LI Tao, et al. New principle based on wavelet transform for power system short-term load forecasting[J]. Proceedings of the CSEE, 2003, 23(1): 45-50.
- [17] CHUI C K. Wavelets: a tutorial in theory and applications [M]. Boston: Academic Press, 1992.
- [18] 袁晓龙, 梅雪, 黄嘉爽, 等. 基于随机森林算法的特征选择在 fMRI 数据中的应用[J]. 微电子学与计算机, 2014, 31(8): 132-135.
YUAN Xiaolong, MEI Xue, HUANG Jiashuang, et al. Feature selection based on random forest algorithm and the applications in fMRI data[J]. Microelectronics & Computer, 2014, 31(8): 132-135.
- [19] 王晓佩. 基于随机森林的冷水机组典型故障诊断方法研究[D]. 广州: 华南理工大学, 2020.
WANG Xiaopei. Study on the fault diagnosis method of chiller based on random forest[D]. Guangzhou: South China University of Technology, 2020.
- [20] 李志强, 高大兵, 苏盛, 等. 基于大数据的智能电表入侵检测方法[J]. 电力科学与技术学报, 2016, 31(1): 121-126.

- 都:电子科技大学,2017.
- LI Qi. Micro-grid operation optimization strategy based on evaluation of economic benefit[D]. Chengdu: University of Electronic Science and Technology of China,2017.
- [21] 郁琛,李尚轩,谢云云,等.考虑交通网与配电网信息融合的台风后配电网抢修策略优化[J].电力系统自动化,2022,46(4):15-24.
- YU Chen,LI Shangxuan,XIE Yunyun,et al.Optimization of post-typhoon rush repair strategy for distribution network considering information integration of traffic network and distribution network[J]. Automation of Electric Power Systems,2022,46(4):15-24.
- [22] 李咸善,周晓岚,姚俊伟,等.考虑车主多模式需求响应模糊意愿的优化调度策略[J].电力系统保护与控制,2023,51(2):89-101.
- LI Xianshan,ZHOU Xiaolan,YAO Junwei,et al.Optimal dispatch strategy considering fuzzy intention of multi-mode demand response of vehicle owners[J]. Power System Protection and Control,2023,51(2):89-101.
- [23] SUN X C,CHEN J,ZHAO H R,et al.Sequential disaster recovery strategy for resilient distribution network based on cyber-physical collaborative optimization[J]. IEEE Transactions on Smart Grid,2023,14(2):1173-1187.
- [24] KENNEDY J, EBERHART R. Particle swarm optimization[C]//Proceedings of ICNN'95-International Conference on Neural Networks. Perth, WA, Australia. IEEE,1995:1942-1948.
- [25] CHEN W N,ZHANG J,CHUNG H S H,et al.A novel set-based particle swarm optimization method for discrete optimization problems[J]. IEEE Transactions on Evolutionary Computation,2010,14(2):278-300.
- [26] 任郡枝.光伏配电网中储能系统规划方法研究[D].济南:山东大学,2020.
- REN Junzhi. Research on planning method of energy storage system in photovoltaic distribution network[D]. Jinan:Shandong University,2020.
- [27] 陈健,林咨良,赵浩然,等.考虑信息耦合的电一气综合能源系统韧性优化方法[J].中国电机工程学报,2020,40(21):6854-6864.
- CHEN Jian, LIN Ziliang, ZHAO Haoran, et al. Optimization method for resilience of integrated electric-gas system with consideration of cyber coupling[J]. Proceedings of the CSEE,2020,40(21):6854-6864.
-
- (上接第47页 Continued from page 47)
- LI Zhiqiang,GAO Dabing,SU Sheng,et al.Big data based intrusion detection method of smart meters[J]. Journal of Electric Power Science and Technology,2016,31(1):121-126.
- [21] 乔黎伟,王静怡,郭炜,等.基于随机森林算法的中短期用电量预测[J].电力科学与技术学报,2020,35(2):150-156.
- QIAO Liwei,WANG Jingyi,GUO Wei,et al.Medium and short-term electricity demand prediction based on random forests algorithm[J]. Journal of Electric Power Science and Technology,2020,35(2):150-156..
- [22] 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 element bearing based on wavelet denoising and empirical mode decomposition[D]. Harbin: Harbin Institute of Technology,2013.
- [23] 李静,徐路路.基于机器学习算法的研究热点趋势预测模型对比与分析:BP神经网络、支持向量机与LSTM模型[J].现代情报,2019,39(4):23-33.
- LI Jing, XU Lulu. Comparison and analysis of research trend prediction models based on machine learning algorithm—BP neural network, support vector machine and LSTM model[J]. Journal of Modern Information, 2019,39(4):23-33.
- [24] 耿丽娟,李星毅.用于大数据分类的KNN算法研究[J].计算机应用研究,2014,31(5):1342-1344+1373.
- GENG Lijuan, LI Xingyi. Improvements of KNN algorithm for big data classification[J]. Application Research of Computers,2014,31(5):1342-1344+1373.
- [25] 范昕炜.支持向量机算法的研究及其应用[D].杭州:浙江大学,2003.
- FAN Xinwei.Support vector machine and its applications [D].Hangzhou:Zhejiang University,2003.
- [26] 孔英会,景美丽.基于混淆矩阵和集成学习的分类方法研究[J].计算机工程与科学,2012,34(6):111-117.
- KONG Yinghui, JING Meili. Research of the classification method based on confusion matrixes and ensemble learning[J]. Computer Engineering & Science, 2012,34(6):111-117.