

# 基于 ReliefF 的时频联合特征及随机森林的 配电网电缆故障识别方法

王 辉<sup>1</sup>, 陈佳宁<sup>2</sup>, 金 雪<sup>1</sup>, 冯 双<sup>2</sup>

(1. 国电南瑞南京控制系统有限公司, 江苏 南京 211106)

(2. 东南大学电气工程学院, 江苏 南京 210096)

**[摘要]** 提出一种基于 ReliefF 算法的时频联合特征及随机森林的配电网电缆故障识别方法. 针对零序电压, 从时域和频域构造 23 个故障敏感特征, 采用 ReliefF 算法进行特征选择, 得到最具分类能力的特征子集. 将特征子集作为基于随机森林的输入进行训练, 得到最终的分类模型, 实现了电缆故障类型识别. 所提方法与基于单一特征的方法相比, 能够更加充分地挖掘数据潜力, 同时由于采用 ReliefF 算法筛选了无关特征, 提高了算法效率. 最后采用 Matlab 软件进行仿真, 并与决策树、KNN、SVM 等算法进行比较, 仿真结果验证了所提方法的可行性和高准确性.

**[关键词]** 配电网电缆, 短路故障识别, 时频联合特征, ReliefF 算法, 随机森林

**[中图分类号]** TM726 **[文献标志码]** A **[文章编号]** 1672-1292(2020)02-0044-08

## Fault Identification Method of Distribution Cable Based on Time-Frequency Domain Features Extracted by ReliefF and Random Forest Algorithm

Wang Hui<sup>1</sup>, Chen Jianing<sup>2</sup>, Jin Xue<sup>1</sup>, Feng Shuang<sup>2</sup>

(1. Nari Technology Co., Ltd., Nanjing 211106, China)

(2. School of Electrical Engineering, Southeast University, Nanjing 210096, China)

**Abstract:** This paper proposes a fault identification method of distribution cable based on time-frequency domain features extracted by ReliefF and random forest algorithm. 23 fault-sensitive features in time and frequency domain are constructed with bases of the zero sequence voltage, which are then selected with ReliefF method to obtain the subset of features with strong ability to identify the types of fault. The subset of features serves as the input of random forest to train the model identifying the type of short circuit fault. Compared with the method utilizing single feature, the proposed method exploits the potential of data and is more robust. Besides, because of the elimination of irrelative features, the efficiency of the method is improved. The simulation is carried out in Matlab and compared with decision tree, KNN and SVM algorithms. The result demonstrates the feasibility and high accuracy of the proposed method.

**Key words:** cable of distribution network, short-circuit fault identification, time-frequency domain feature, ReliefF, random forest

随着智能配电网的建设, 配电网的安全可靠性要求在逐步提高. 快速准确地识别配电网故障, 对于故障切除和减少故障停电时间、提高供电可靠性十分重要<sup>[1-2]</sup>. 电缆短路故障是配电网中最常见的故障, 但由于短路故障形式多样, 波形受到故障位置、过渡电阻、故障初始角等因素影响, 表现形式各异. 相比于输电网, 配电网的成分更加复杂, 分支较多, 因此其故障特征也更为模糊. 此外, 我国中低压配电网广泛采用小电流接地方式, 发生故障时, 尤其是单相接地故障时, 故障特征微弱. 这些均给配电网电缆故障识别带来了困难, 亟需研究快速准确的配电网电缆故障识别方法<sup>[3-4]</sup>.

国内外学者对此展开了一些研究. 研究表明, 基于稳态信号对配电网故障进行分类存在适应性差的

问题,尤其是对中性点经消弧线圈接地的准确性差,因此目前大部分研究都基于暂态信号对故障进行分类<sup>[4]</sup>. 吴京<sup>[5]</sup>提出根据不同故障类型下小波能量谱的特征识别电缆故障类型. Das<sup>[6]</sup>分析了不同故障下三相电流在幅值和相位上的差异,通过建立隶属度函数识别故障类型. 随着机器学习的快速发展,一些基于机器学习算法的故障识别方法被提出. 李小薇<sup>[7]</sup>直接利用电缆正常状态和两种故障状态(相间短路和三相短路)下的电缆数据,构造 2 个支持向量机(support vector machine, SVM)进行识别. 杨春宇<sup>[8]</sup>和孙萌<sup>[9]</sup>分别以小波包分解的第四层能量谱和小波奇异熵作为特征向量,输入人工神经网络识别故障类型. 苏立<sup>[10]</sup>根据希尔伯特-黄变换提取地下电缆故障信号的特征分量,将提取的特征分量和地下电缆故障类型作为最小二乘支持向量机(least square support vector machine, LSSVM)的输入和输出,实现地下电缆故障类型的识别. 由此可见,现有方法一部分直接基于量测数据进行分析 and 构造判据,未深入挖掘量测数据的其他特征;另一部分基于经验提炼出量测数据的某些特征并结合机器学习进行故障类型识别,但未对所提特征的分类能力进行评价,分类能力不强的特征可能影响算法的效率<sup>[11]</sup>.

本文提出一种基于时频联合特征以及 ReliefF 算法的特征子集构造方法,利用电缆故障状态下零序电压分量构造时域及频域共 23 个特征,并采用 ReliefF 算法进行特征选择,得到最具分类能力的特征子集. 将特征子集作为随机森林的输入进行训练,得到最终的识别模型,实现电缆故障类型识别. 该方法相比现有算法有以下优势:首先,通过构造多个时频联合指标,充分挖掘故障数据的特征,相比基于单一判据的识别方法,具有更好的鲁棒性;其次,通过采用能够快速过滤冗余特征的 ReliefF 特征选择算法和组合多个训练集分类结果的随机森林算法,提升了识别算法的快速性、鲁棒性和准确性. 在 Matlab/Simulink 软件中,针对 10 种故障类型考虑多个影响因素进行仿真,并与普通决策树、K 最近邻(K-Nearest Neighbor)、SVM 算法比较,仿真结果验证了所提方法的可行性和高准确性.

## 1 基于 ReliefF 算法的配电网电缆故障的时频联合特征

### 1.1 配电网电缆故障的时频联合特征构造

电缆故障中最常见的 4 类短路故障为单相短路接地、两相短路、两相接地短路和三相短路. 根据故障相的不同,共可分为如图 1 所示的 10 种故障类型. 其中,XY 表示 X、Y 相间短路,XG 表示 X 相接地故障,X 和 Y 为 A、B 或 C. 本文的目标是通过电压暂态量测数据准确识别这 10 种故障类型.

以两馈线路为例,由于配电网馈线长度较短,可采用集中参数模型来表示. 当发生单相接地故障时,其等效电路如图 2 所示<sup>[4]</sup>. 图 2 中, $U$  为零序电压; $U_B$  为母线的零序电压; $R_f$  为故障点的故障电阻; $L$  为消弧线圈的电感值; $R_N$ 、 $L_N$ 、 $C_N$  分别为正常馈线的电阻、电感及电容; $R_F$ 、 $L_F$ 、 $C_F$  分别为故障馈线的电阻、电感及电容. 一般来说,消弧线圈为过补偿状态.

由等效电路图可计算正常馈线的阻抗表达式为:

$$Z_N = \frac{C_N R_N \omega - j + j C_N L_N \omega^2}{\omega C_N (2 - C_N L_N \omega^2 + j C_N R_N \omega)}. \quad (1)$$

正常馈线零序的电流表达式为:

$$I_N = \frac{U_B \omega C_N (2 - C_N L_N \omega^2 + j C_N R_N \omega)}{C_N R_N \omega - j + j C_N L_N \omega^2}. \quad (2)$$

由于中性点经过消弧线圈接地,所以消弧线圈的零序电流表达式为:

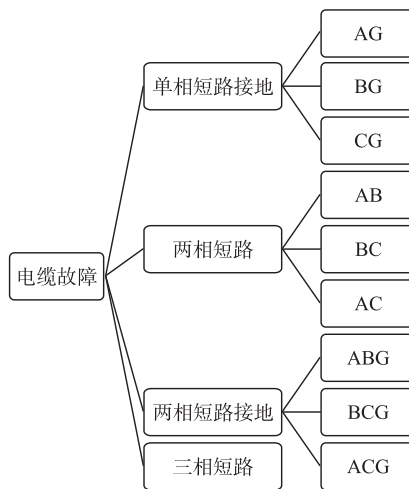


图 1 电缆故障类型

Fig. 1 Types of cable fault

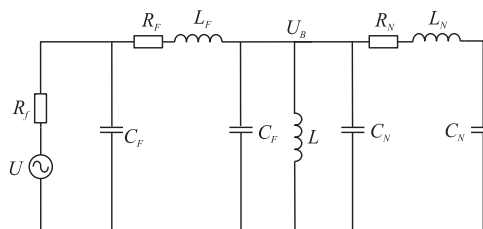


图 2 两馈线配网等效电路图

Fig. 2 Equivalent circuit of two-feeder distribution system

$$I_L = \frac{U_B}{Z_L} = \frac{U_B}{j\omega L} \quad (3)$$

故障馈线的零序电流为正常馈线零序电流和经过消弧线圈的电流之和,可得故障馈线零序电流表达式为:

$$I_F = -I_N - I_L = -\frac{U_B(2L\omega^2 C_N + \omega^2 C_N L_N - L\omega^4 C_N^2 L_N - 1 - j\omega C_N R_N + jL\omega^3 C_N R_N)}{L\omega(C_N R_N \omega - j + jC_N L_N \omega^2)}, \quad (4)$$

即母线的零序电压为:

$$U_B = \frac{I_F L\omega(C_N R_N \omega - j + jC_N L_N \omega^2)}{2L\omega^2 C_N + \omega^2 C_N L_N - L\omega^4 C_N^2 L_N - 1 - j\omega C_N R_N + jL\omega^3 C_N R_N}. \quad (5)$$

其他故障情况下,零序电压的表达式也可同理推出. 但配电网分支较多,故障种类也较多,难以根据推导的表达式来进行故障的分类. 因此,本文提出一种基于数据分析的方法进行故障的识别.

当电力系统发生以上 10 种不同类型的故障时,首先要对故障特征进行提取,提取的故障特征对故障识别的准确性非常重要. 本文从时域和频域两个方面,分别提取故障时零序电压信号的 23 个特征量. 时域特征量包括均值、标准差、方根幅值、均方根值、峰值、歪度、峭度、峰值系数、裕度、波形指标、脉冲指标共 11 个指标. 频域特征量包括中心频率、方差、偏斜度、峰度、频度中心、频率标准差、均方根频率、波形稳定系数、变异系数、歪度、峭度和均方根比率共 12 个指标. 具体数学表达式如表 1 所示.

表 1 时频联合特征

Table 1 Time and frequency domain features

序号	时域特征		序号	频域特征	
	特征名称	数学表达式		特征名称	数学表达式
1	均值	$F_1 = \frac{\sum_{n=1}^N x(n)}{N}$	12	中心频率	$F_{12} = \frac{\sum_{k=1}^K s(k)}{K}$
2	标准差	$F_2 = \sqrt{\frac{\sum_{n=1}^N (x(n) - T_1)^2}{N - 1}}$	13	方差	$F_{13} = \frac{\sum_{k=1}^K (s(k) - F_1)^2}{K - 1}$
3	方根幅值	$F_3 = \left( \frac{\sum_{n=1}^N \sqrt{ x(n) }}{N} \right)^2$	14	偏斜度	$F_{14} = \frac{\sum_{k=1}^K (s(k) - F_1)^3}{K(\sqrt{F_2})^3}$
4	均方根值	$F_4 = \sqrt{\frac{\sum_{n=1}^N x(n)^2}{N}}$	15	峰度	$F_{15} = \frac{\sum_{k=1}^K (s(k) - F_1)^4}{KF_2^2}$
5	峰值	$F_5 = \max  x(n) $	16	频度中心	$F_{16} = \frac{\sum_{k=1}^K f_k s(k)}{\sum_{k=1}^K s(k)}$
6	歪度	$F_6 = \frac{\sum_{n=1}^N (x(n) - T_1)^3}{(N - 1)T_2^3}$	17	频率标准差	$F_{17} = \sqrt{\frac{\sum_{k=1}^K (f_k - F_5)^2 s(k)}{K}}$
7	峭度	$F_7 = \frac{\sum_{n=1}^N (x(n) - T_1)^4}{(N - 1)T_2^4}$	18	均方根频率	$F_{18} = \sqrt{\frac{\sum_{k=1}^K f_k^2 s(k)}{\sum_{k=1}^K s(k)}}$
8	峰值系数	$F_8 = \frac{T_5}{T_4}$	19	波形稳定系数	$F_{19} = \frac{\sum_{k=1}^K f_k^2 s(k)}{\sqrt{\sum_{k=1}^K s(k) \sum_{k=1}^K f_k^4 s(k)}}$
9	裕度	$F_9 = \frac{T_5}{T_3}$	20	变异系数	$F_{20} = \frac{F_6}{F_5}$

续表 1

Table 1 continued

序号	时域特征		序号	频域特征	
	特征名称	数学表达式		特征名称	数学表达式
10	波形指标	$F_{10} = \frac{T_4}{\frac{1}{N} \sum_{n=1}^N  x(n) }$	21	歪度	$F_{21} = \frac{\sum_{k=1}^K (f_k - F_5)^3 s(k)}{KF_6^3}$
11	脉冲指标	$F_{11} = \frac{T_5}{\frac{1}{N} \sum_{n=1}^N  x(n) }$	22	峭度	$F_{22} = \frac{\sum_{k=1}^K (f_k - F_5)^4 s(k)}{KF_6^4}$
			23	均方根比率	$F_{23} = \frac{\sum_{k=1}^K (f_k - F_5)^{1/2} s(k)}{K \sqrt{F_6}}$

注:  $x(n)$  为故障时采集的零序电压时域信号,  $n=1, 2, \dots, N$ ,  $N$  为总的采集点数;  $s(k)$  为故障时零序电压的频率分量,  $k=1, 2, \dots, K$ ,  $K$  为谱线数量;  $f_k$  为第  $k$  条谱线的值。

## 1.2 基于 ReliefF 的特征子集选择

上节共构造了 23 个时频联合特征,若采用所有特征对分类器进行训练会大大增加计算量。此外,当特征维度超过一定数量时,分类器的性能可能下降。因此,本文采用 ReliefF 算法从 23 个时频联合特征中选择适用于电缆故障识别的特征子集。

Relief 算法是一种经典的多变量过滤式特征选择算法,但只能用于两分类问题。ReliefF 算法在此基础上进行了改进<sup>[12-13]</sup>,通过将多类别问题分解为多个两分类问题,实现多分类问题的特征选择。其基本原理是根据每个特征和类别的相关性计算得到相应特征的权重,与类别相关性越低的特征权重越小,最后去除低于某一权重的低相关性特征。

采用 ReliefF 算法进行电缆故障类型识别的特征子集选择的具体方法为:

- (1) 选择任意一个样本  $X$ , 其对应的故障类型为  $c$ ;
- (2) 分别从与样本  $X$  同类和不同类的样本集中找出  $K$  个最近邻样本  $H_j$  和  $M_j$  ( $j=1, 2, \dots, K$ );
- (3) 根据式(6)更新每个特征  $F_p$  ( $p=1, 2, \dots, P$ ) 的权重  $W(F_p)$ :

$$W(F_p) = W(F_p) - \frac{1}{mK} \sum_{j=1}^K d(F_p, X, H_j) + \frac{1}{mK} \frac{P(c)}{1-P(c)} \sum_{j=1}^K d(F_p, X, M_j), \quad (6)$$

式中,  $m$  为算法重复次数;  $P(c)$  为第  $c$  类故障样本在总样本中的概率;  $d(F_p, X, H_j)$  表示样本  $X$  在时频特征  $F_p$  上与样本  $H_j$  的距离;  $d(F_p, X, M_j)$  表示样本  $X$  在时频特征  $F_p$  上与样本  $M_j$  的距离, 其表达式为:

$$d(F_p, X, H_j) = \frac{|V(F_p, X) - V(F_p, H_j)|}{\max(F_p) - \min(F_p)}, \quad (7)$$

$$d(F_p, X, M_j) = \frac{|V(F_p, X) - V(F_p, M_j)|}{\max(F_p) - \min(F_p)}, \quad (8)$$

式中,  $V(F_p, A)$  表示样本  $A$  的特征  $F_p$  的值,  $A$  为  $X, H_j$  或  $M_j$ ;

- (4) 重复以上过程  $m$  次, 得到最终所有  $P$  个特征的权重  $W(F_p)$ ;
- (5) 设定权重阈值  $\alpha$ , 选择权重大于  $\alpha$  的特征构成特征子集。

## 2 随机森林算法原理及构建

决策树是一种通过多变量输入建立估计模型的有监督机器学习算法, 实现简单、过程清晰。但决策树是一种弱分类器, 分类能力有限。随机森林算法采用随机抽样的方法训练多棵决策树, 对目标问题进行投票表决来构成一个性能更好的模型, 获得模型的准确度高、泛化能力较强, 且不易产生过拟合现象, 因而在计算机视觉、文本分类、金融市场风险预测等领域获得了广泛关注和应用<sup>[14-15]</sup>。

### 2.1 随机森林基分类器——决策树

决策树是一种通过训练样本构建二叉树来分类的算法, 模型中包含 3 种节点: 根节点、决策节点和叶

节点. 其中,根节点包含所有的样本,决策节点表示一个特征的判定,叶节点表示最终的一个类别. 采用决策树进行分类的过程就是从根节点开始,根据决策节点中对某个特征的判别,将样本分配到相应的分支. 重复该过程,直至到达叶节点,获得样本最终所属类别.

构造决策树最关键的环节为节点分裂算法,本文采用分类与回归树(classification and regression tree, CART)算法. 该算法基于基尼指数最小进行分裂,基尼指数的大小代表数据的不纯度,一个数据集  $D$  的基尼指数可表示为:

$$\text{Gini}(D)=1-\sum_{i=1}^M P_i^2,$$

(9)

式中, $M$  为数据类别个数; $P_i$  为第  $i$  类样本在数据集  $D$  中所占的比例. 当数据集中只有一种类型的样本时,基尼指数最小,等于 0. 基尼指数越大,说明数据集  $D$  中的数据纯度越低,越分散;基尼指数越小,说明数据集  $D$  中的数据纯度越高,越集中.

若特征  $A$  为分裂特征,则分裂之后该数据集的基尼指数可表示为:

$$\text{Gini}_A(D)=\sum_{j=1}^K \frac{D_j}{D}\text{Gini}(D_j),$$

(10)

式中, $K$  表示数据集  $D$  被分成的份数; $D_j$  为分裂后的  $K$  个数据集其中的第  $j$  个. 分裂特征的选择标准为使得数据集  $D$  的基尼指数最小.

2.2 随机森林的构建

以决策树为基决策器,随机森林构建的过程主要包括 3 部分:Bootstrap 重抽样,训练多棵决策树,多棵决策树投票表决. 算法具体步骤为:

- (1)假设原数据集包含  $T$  个样本,采用 Bootstrap 重抽样从原样本集中抽样  $T$  次构成新的样本集,重复进行  $V$  次,得到  $V$  个相互独立的样本集;
- (2)针对每个样本集分别构建决策树,在每棵决策树构建过程中的节点分裂时,都随机从所有  $P$  个特征中抽取  $p$  个特征,按照 CART 方法进行生长,得到  $V$  棵决策树;
- (3)将  $V$  棵决策树进行组合,采用投票表决的方式得到最终的分类结果.

3 基于 ReliefF 算法的时频联合特征及随机森林的配电网电缆故障识别流程

根据基于 ReliefF 算法的时频联合特征子集获取方法和随机森林的构造方法,配电网电缆故障识别方法的流程如图 3 所示. 该方法分为 3 个部分,首先是训练样本的获得,然后根据训练样本提取时频特征并

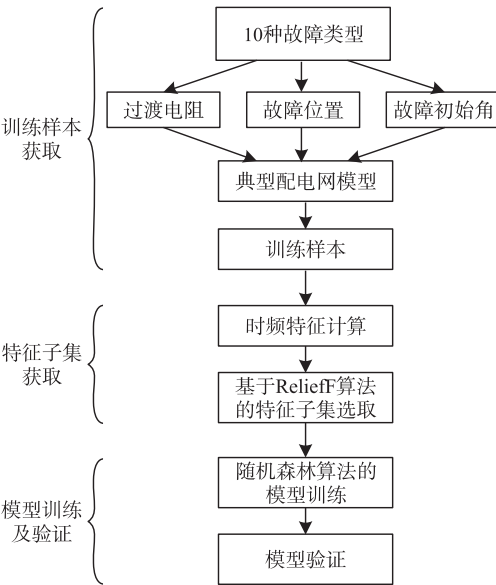


图 3 基于 ReliefF 算法的时频联合特征及随机森林的配电网电缆故障识别方法流程示意图

Fig. 3 Flow chart of the fault identification method of distribution cable based on time-frequency domain features extracted by ReliefF and random forest algorithm



基于 Relieff 算法选择特征子集,最后进行随机森林模型的训练和模型验证,得到最终的识别模型. 当故障发生时,对采集数据提取特征子集中的特征,输入识别模型,即可得到故障类型.

在训练样本的获得中,分别设置如图 1 所示的 10 种故障. 此外,还考虑以下 3 种因素对故障特征的影响:

- (1)过渡电阻:系统发生短路故障时短路点存在过渡电阻,主要包括电弧电阻和接地电阻;
- (2)故障位置:配电网电缆具有一定长度,不同位置都有可能发生故障,对故障信号有显著影响;
- (3)故障初始角:当系统在某些故障初始角情况下发生短路时,故障特征可能十分微弱,如电压过零点或电压幅值相等方向相反时.

4 算例分析

本文在 Matlab/Simulink 中建立典型的配电网模型,仿真模型如图 4 所示.

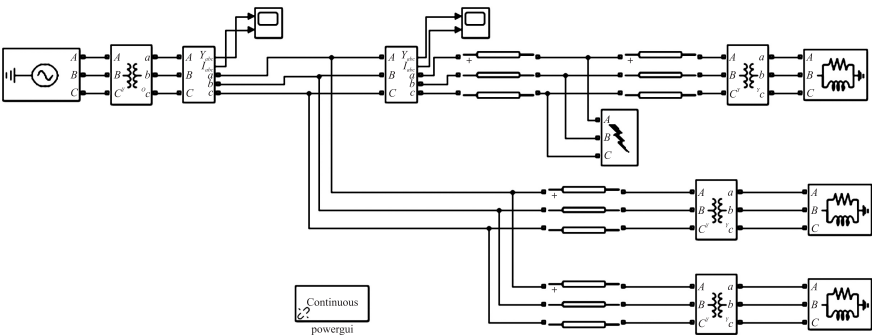


图 4 典型配电网仿真模型

Fig. 4 Simulation model of typical distribution network

设置图 1 中的 10 种故障,针对每种故障,考虑过渡电阻在 5~500 Ω 之间变化,故障位置在 2~8 km 变化,故障初始角在 0~360°之间变化,每种变化情况下分别得到 1 000 组样本. 样本的采样频率为 100 kHz,采样时间为一个周波.

首先对获得的样本提取表 1 中的 23 个时频联合特征形成特征集  $S$ ,并采用 Relieff 算法计算每个特征的权重,结果如图 5 所示. 由图 5 可知,与类型相关性越大的特征权重越大,因此选择权重大于 0.2 的 8 个特征得到最终的特征子集  $S_1$ , $S_1 = \{F_1, F_5, F_6, F_{13}, F_{17}, F_{20}, F_{21}, F_{23}\}$ .

将获取的特征子集  $S_1$  作为随机森林的输入,对配电网电缆故障识别模型进行训练. 同时将未经过 Relieff 算法筛选的特征集  $S$  作为随机森林的输入,同样进行模型训练. 采用经典的十折交叉验证方法对模型的准确率进行验证,即将所有样本分为 10 个子集,每个子集均做一次测试集,其他 9 个作为训练集,每次测试均得到相应的正确率,10 次结果正确率的平均值作为对算法精度的估计. 一般认为,采用十折交叉验证得到的准确率更接近模型的真实情况. 两种输入情况下训练时间和模型准确率如表 2 所示.

从表 2 可见,经过 Relieff 筛选后训练时间减少了 41.9%,而准确率则相同,说明采用 Relieff 算法删除了无关特征,提高了算法的效率,同时保证了识别的准确率. 采用本文方法得到的故障辨识模型的混淆矩阵如图 6 所示.

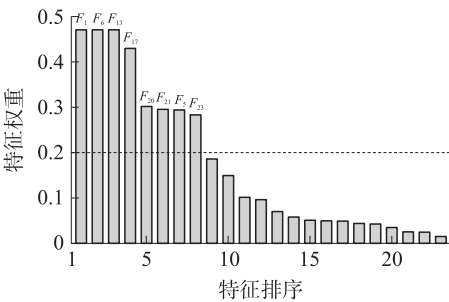


图 5 Relieff 算法得到的特征权重

Fig. 5 Predictor importance weight obtained with Relieff algorithm

表 2 不同特征下模型训练时间及准确率

Table 2 Training time and accuracy with different feature sets

特征集	训练时间/s	准确率/%
全部特征集 $S$	88.4	99.3
特征子集 $S_1$	51.3	99.3

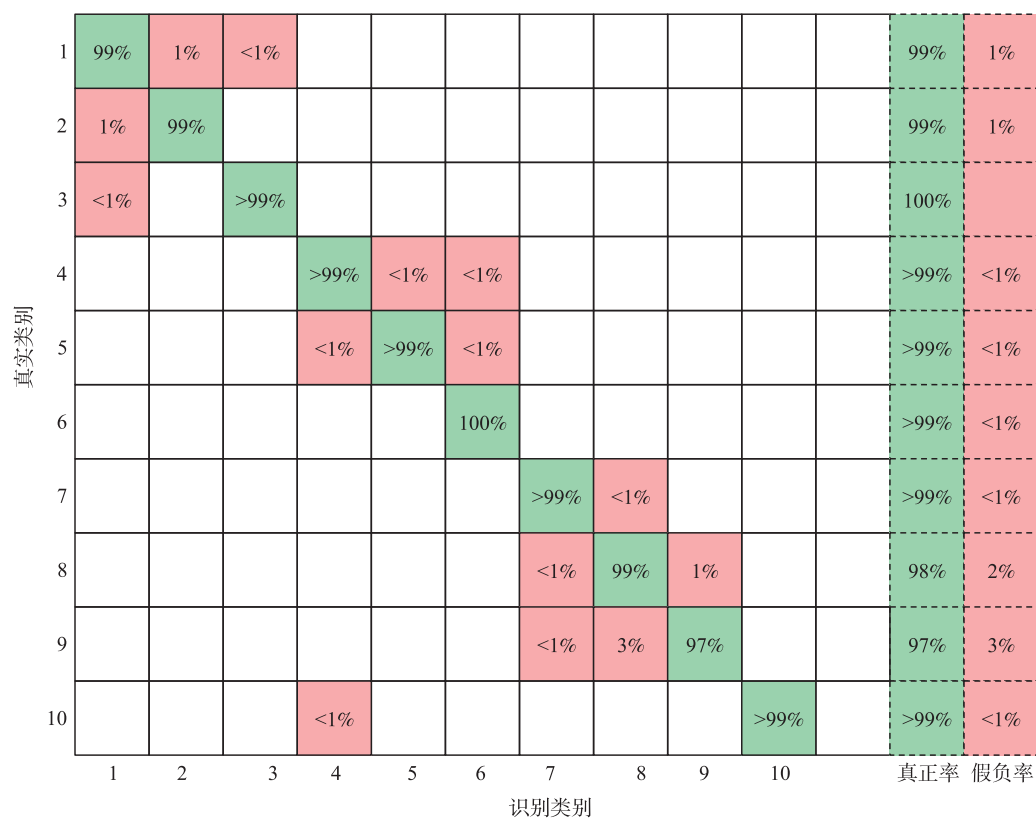


图 6 故障辨识模型的混淆矩阵

Fig. 6 Confusion matrix of the fault identification model

由图 6 可见,对 10 种故障类型的识别准确率均达到 97% 以上. 以同样的特征子集作为输入,分别采用决策树、SVM 和 KNN 算法进行训练,并用十折交叉验证方法验证模型的准确率,所得结果如表 3 所示. 由表 3 可知,随机森林算法所得模型的准确率最高,因此最适用于配电网电缆故障的识别.

5 结论

针对配电网电缆常见的 10 种短路故障,本文基于 ReliefF 算法的时频联合特征及随机森林提出了一种高准确度的识别方法,得出结论如下:

- (1) 根据配电网电缆短路故障特点,利用零序电压从时域和频域构造了 23 个故障敏感特征,充分挖掘了数据中的故障信息;
- (2) 进一步采用 ReliefF 算法筛选得到最具分类能力的特征子集,避免了无关特征对算法效率的影响;
- (3) 采用随机森林算法训练故障识别模型,仿真结果表明针对电缆短路故障识别问题,随机森林相比决策树、SVM 和 KNN 等其他机器学习算法具有更高的识别准确率和更高的计算效率.

[参考文献](References)

[1] 王成山,李鹏,于浩. 智能配电网的新形态及其灵活性特征分析与应用[J]. 电力系统自动化,2018,42(10):13-21.

[2] 姜楠,王琦,王恩荣,等. 分布式光伏电源接入对配电网可靠性的影响研究[J]. 南京师范大学学报(工程技术版), 2016,16(2):1-9.

[3] 方毅,薛永端,宋华茂,等. 谐振接地系统高阻接地故障暂态能量分析与选线[J]. 中国电机工程学报,2018,38(19):

- 5636-5645,5921.
- [4] 张姝. 配电网弱故障接地保护与定位方法研究[D]. 成都:西南交通大学,2018.
  - [5] 吴京. 基于模分量小波能量谱的电缆故障识别方法研究[D]. 西安:西安科技大学,2014.
  - [6] DAS B. Fuzzy logic-based fault-type identification in unbalanced radial power distribution system[J]. IEEE Transactions on Power Delivery,2006,21(1):278-285.
  - [7] 李小微. 基于 IPSO-SVM 的电缆故障识别[D]. 西安:西安科技大学,2014.
  - [8] 杨春宇. 电力电缆故障分析与诊断技术的研究[D]. 大连:大连理工大学,2013.
  - [9] 孙萌. 基于小波奇异熵的配电网电缆接地故障研究[D]. 西安:西安理工大学,2017.
  - [10] 苏立. 基于 HHT 变换和 FOA\_LSSVM 的电缆故障诊断[J]. 计算机与现代化,2017(9):96-101,105.
  - [11] 何清,李宁,罗文娟,等. 大数据下的机器学习算法综述[J]. 模式识别与人工智能,2014,27(4):327-336.
  - [12] JENKE R,PEER A,BUSS M. Feature extraction and selection for emotion recognition from EEG[J]. IEEE Transactions on Affective Computing,2014,5(3):327-339.
  - [13] 何涛,胡洁,夏鹏,等. 基于 ReliefF 算法与遗传算法的肌电信号特征选择[J]. 上海交通大学学报,2016,50(2):204-208.
  - [14] SABAHI F,AHMAD M O,SWAMY M N S. Perceptual image hashing using random forest for content-based image retrieval[C]// 2018 16th IEEE International New Circuits and Systems Conference(NEWCAS). Montreal,Canada:IEEE,2018:348-351.
  - [15] 彭微. 基于随机森林的文本分类并行化研究[D]. 湘潭:湘潭大学,2018.

[责任编辑:严海琳]