

基于平稳小波变换与随机森林的电缆早期故障识别方法

李胜辉¹, 白雪¹, 董鹤楠¹, 卢宏², 郭朝云²

(1. 国网辽宁省电力有限公司电力科学研究院, 辽宁 沈阳 110006;
2. 四川大学电气工程学院, 四川 成都 610065)

摘要: 分析电能质量监测装置获得的电能质量数据,为解决电力系统设备故障提供了一个新的方向。电缆作为电力系统中的重要设备,在出现永久性故障之前会出现一段时间的电缆早期故障,早期故障虽然不会引起保护装置动作,但可能是永久性故障的前兆,降低电力系统运行的安全性。本文提出基于平稳小波变换与随机森林相结合的电缆早期故障识别方法。首先利用平稳小波变换对故障相电流信号进行分解,根据获得的小波系数计算统计特征量;然后将提取的特征量作为随机森林分类器的输入,从多种过电流扰动中识别出电缆早期故障。将该方法与传统模式识别方法相比较,结果表明本文所提方法识别准确率更高,时间更短。

关键词: 平稳小波变换; 随机森林; 电缆早期故障; 过电流扰动

DOI: 10.12067/ATEEE1809022 **文章编号:** 1003-3076(2020)03-0040-09 **中图分类号:** TM726.4

1 引言

电缆早期故障是电缆永久性故障之前的一个阶段,常规的过电流检测装置不能检测到,但在一段时间内会重复发生,最后导致永久性故障^[1,2]。虽然电缆早期故障不能被保护装置检测到,但在电流波形中会出现一定的特征,通过对波形特征的分析,提出对应的检测算法,从而在故障发生之前对电缆进行检修或更换,有利于提高电力系统供电可靠性和避免发生永久性故障。

电缆早期故障的特征可总结为:①多发生在电压峰值时刻,因为此时电压应力最大,电缆绝缘更容易被击穿;②按照持续时间可分为半周波早期故障和多周波早期故障,持续时间分别为1/4~1/2个周波和1~4个周波;③具有重复性,多为单相接地故障;④不能引起电力系统保护装置动作^[3-5]。

目前,国内外对电缆早期故障的研究取得了一定的成果。文献[2]利用傅里叶变换和小波变换来分解信号,选取特征量,最后利用支持向量机和人工神经网络两种机器学习算法来完成分类识别。文献[5]分析了电缆早期故障的波形特征,利用小波变

换分解电缆电流故障信号,选取能量值和均方根值作为特征量来检测电缆中的过电流扰动,再根据电缆早期故障的持续时间和峰值对早期故障进行识别。文献[6]将小波变换与灰色关联分析方法相结合,利用小波变换选取时域特征量,初步检测过电流扰动,最后计算时域特征量与类别为电缆早期故障的参考样本之间的灰色关联度,判断关联度最大的为电缆早期故障。文献[7]中采用一种在线地下电缆监测系统记录故障数据,首先利用小波包理论分析电缆电流故障信号,将电流信号分解到不同的频率范围内,再对可能出现的过电流扰动进行识别;最后利用无监督学习的自组织映射,对不同的扰动进行分类识别。但在实际运行过程中,电缆的过电流扰动众多,误检测率高。文献[8]基于现场测量的数据,利用自组织映射方法构建数值模型,最后利用统计变化的检测算法确定故障时间。文献[9]利用卡尔曼滤波,首先基于故障电流计算的基频零序电流分量检测电缆早期故障,再通过计算的方差信号进一步确定是否为电缆早期故障。文献[10]假设电缆早期故障为电弧故障,利用电弧畸变的特点,构建仿真模型,计算故障电压总谐波畸变率,并与参考

收稿日期: 2018-09-11
作者简介: 李胜辉(1982-),男,辽宁籍,高级工程师,硕士,主要从事新能源与电能质量的研究;
白雪(1988-),女,辽宁籍,工程师,硕士,主要从事电能质量的研究。

值相比较,最后判断是否为电缆早期故障。目前的研究方法大多基于信号处理方法提取电缆早期故障的特征或者根据一定的假设条件建立仿真电路模型,在电力系统条件改变后,上述方法可能不适用。

现代智能电网中监测装置的广泛应用,使得大量实时电能质量数据能够被准确地记录^[11]。对电能质量数据进行分析,可以解决故障诊断、系统设备运行和负荷预测等问题。目前对电缆早期故障的研究未对电能质量数据进行深入分析,适用性不强。而机器学习算法能够深入地挖掘数据中隐藏的信息,目前已经广泛用于负荷预测^[12]、故障诊断^[13]等领域。

本文提出了一种平稳小波变换与随机森林(Random Forest, RF)相结合的电缆早期故障识别方法。利用平稳小波变换对原始电流信号进行分解,基于分解得到的系数计算特征量,最后利用随机森林进行分类识别。将本文方法与决策树(Decision Tree, DT)^[14]、K近邻(K-Nearest Neighbor, KNN)^[15]、朴素贝叶斯分类器(Naive Bayes, NB)^[13]和概率神经网络(Probabilistic Neural Network, PNN)^[16]相比较,结果表明,本文提出方法具有更高的准确率和更短的时间。

2 基于平稳小波变换的特征量提取

由于电缆早期故障信号存在突变点,为了在小波变换后不丢失突变点的信息,采用平稳小波变换^[17]对电缆早期故障信号进行分解。基于分解得到的小波系数计算初始特征量,将这些初始特征量作为 RF 的输入,提高 RF 的效率。

2.1 小波变换

假设 $\psi(t)$ 满足容许条件,如式(1)所示,则 $\psi(t)$ 可称为一个母小波。

$$C_\psi = \int_R \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty \quad (1)$$

式中, $\Psi(\omega)$ 表示 $\psi(t)$ 的傅里叶变换。

则对于一个连续信号 $g(t)$,其对应的连续小波变换(Continue Wavelet Transform, CWT)^[18]为:

$$CWT(a, b) = \frac{1}{\sqrt{a}} \int_R g(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (2)$$

式中, $\overline{\psi(t)}$ 表示 $\psi(t)$ 的复共轭函数; a 表示与频率有关小波函数的尺度因子; b 表示与时间有关的位移因子。

在实际运用中,需要将连续小波离散化。假设

$a = a_0^j, b = kb_0, j, k \in Z$, 于是可以得到对应的离散小波变换(Discrete Wavelet Transform, DWT)为:

$$DWT(j, k) = a_0^{-j/2} \psi(a_0^{-j}t - kb_0) \quad (3)$$

2.2 平稳小波变换

平稳小波变换在每次分解时不再进行向下采样,这样就可以使得变换后的高频细节系数和低频近似系数的长度与原始信号长度相等,保存了原始信号大部分有价值的信息^[17]。

平稳小波变换的分解过程是采用不同尺度的正交小波基将信号分解到不同频段上,获得对应频段的高频细节系数和低频近似系数,其中高频细节系数表示暂态现象,低频近似系数表示基频分量^[19]。

信号 $g(t) \in L^2(R)$ 对应的分解公式为:

$$g(t) = \sum_{i=1}^I c_{ij} \phi_{ij}(t) + \sum_{i=1}^I \sum_{j=1}^J d_{ij} \psi_{ij}(t) \quad (4)$$

$$c_{i,j+1} = \sum_{m=1}^M H(m-2j) c_{ij} \quad (5)$$

$$d_{i,j+1} = \sum_{m=1}^M G(m-2j) c_{ij} \quad (6)$$

式中, $\phi_{ij}(t)$ 表示尺度函数, $\phi_{ij}(t) = 2^{-j/2} \phi(2^{-j}t - i)$; m 表示当前采样窗口的序号值; M 表示采样窗口的总采样点数; i 表示小波函数的分解程度; j 表示小波函数的离散程度; $c_{i,j+1}$ 和 $d_{i,j+1}$ 表示近似系数和细节系数; $H(j)$ 和 $G(j)$ 表示第 j 层的低通滤波器和高通滤波器。对应的小波分解过程如图 1 所示。

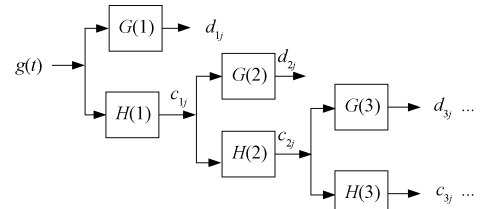


图 1 平稳小波变换分解图

Fig.1 Graph of stationary wavelet decomposition

2.3 特征量的选取

特征量计算公式如表 1 所示。本节利用平稳小波分解电缆故障相电流信号,分解层数为 8 层,得到 8 个细节系数和 1 个近似系数,再按照表 1 中列出的特征计算公式,计算对应的特征量,最后组成包含 81 个特征的特征向量 \mathbf{X} ,作为下一步分类识别的输入。本文研究的信号包括半周波早期故障、多周波早期故障、电容投切、变压器激磁涌流、恒定阻抗故障和正常情况。对以上不同故障类型进行 3 层小波变换,分解结果如图 2 所示,其中 c_3 表示第 3 层近似系数, $d_1 \sim d_3$ 分别表示第 1~3 层细节系数。

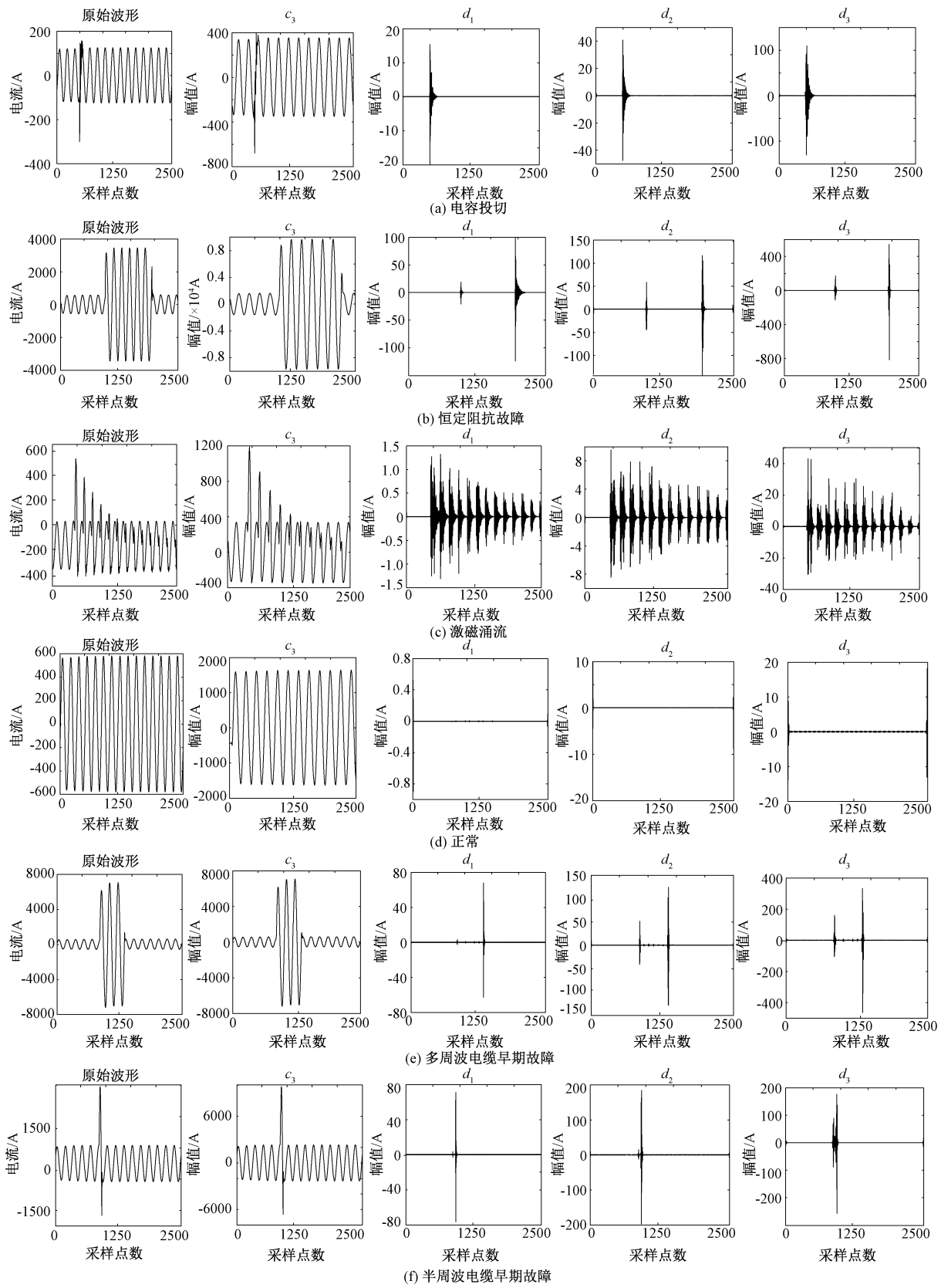


图 2 不同类型的平稳小波分解图

Fig.2 Stationary wavelet decomposition diagram of different fault types

表 1 特征量的计算公式

特征量	计算公式
能量 (En)	$En_i = \sum_{j=1}^N \{c_{ij}^2, d_{ij}^2\}$
平均值 (μ)	$\mu_i = \frac{1}{N} \sum_{j=1}^N \{c_{ij}, d_{ij}\}$
方差 (Std)	$Std_i = \sqrt{\frac{1}{N-1} \sum_{j=1}^N \{(c_{ij} - \mu_i)^2, (d_{ij} - \mu_i)^2\}}$
均方根值 (RMS)	$RMS_i = \sqrt{\frac{1}{N} \sum_{i=1}^N \{c_{ij}^2, d_{ij}^2\}}$
峰值 (Cre)	$Cre_i = \left\{ \frac{\text{Max } c_{ij}}{RMS_i}, \frac{\text{Max } d_{ij}}{RMS_i} \right\}$
信息熵 ($InfEnt$)	$InfEnt_i = - \sum_{j=1}^N \{c_{ij}^2 \log_2 c_{ij}^2, d_{ij}^2 \log_2 d_{ij}^2\}$
斜度 ($Skew$)	$Skew_i = \left\{ \frac{E(c_{ij} - \mu_i)^3}{Std_i^3}, \frac{E(d_{ij} - \mu_i)^3}{Std_i^3} \right\}$
峭度 (Kur)	$Kur_i = \left\{ \frac{E(c_{ij} - \mu_i)^4}{Std_i^4}, \frac{E(d_{ij} - \mu_i)^4}{Std_i^4} \right\}$
能量熵 (Ent)	$Ent_i = -En_i \log_2 \frac{En_i}{En}$

3 随机森林

随机森林^[12]是 Leo Breiman 提出的一种分类器组合算法,集成多个 DT 模型来提高预测精度,利用多个 DT 模型的决策结果,通过投票机制得出最后的识别结果。电缆早期故障的波形数据中存在一定的隐藏信息,通过传统信号特征提取方法只能提取浅层特征。RF 具有强大的数据训练和学习能力,可以准确发现数据中隐藏的信息,有利于从多种过电流扰动中准确识别出电缆早期故障。

RF 的结构如图 3 所示。假设有 n 个 DT 组成的一个 RF 分类器,每个 DT 都是一个分类模型。RF 的具体过程为:首先利用 Bootstrap 重抽样方法随机从原始样本中抽取与原始样本容量相同的样本,再对抽取的样本建立 DT 模型,通过 DT 的训练算法进行训练,最后组合多个 DT 的预测结果,通过决策树的投票机制,将获得票数多的 DT 模型分类结果作为最后结果^[20],即

$$H(X) = \arg \max_Z \sum_{i=1}^n I(h_i(X) = Z) \quad (7)$$

式中, n 表示总的 DT 个数; $h_i(X)$ 表示第 i 个 DT 模型; Z 表示 RF 的分类目标; I 表示示性函数,指事件发生与否与 0,1 两值函数的对应关系。

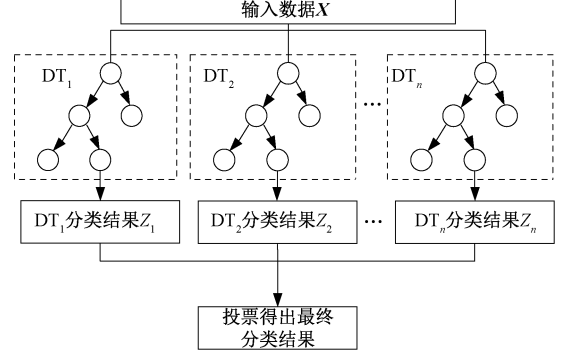


图 3 随机森林示意图

Fig.3 Flowchart of random forest

RF 满足一定的收敛特性,定义余量函数 mr 如式(8)所示。式(8)反映了输入数据 X 对应的正确分类结果 Z 超过其他错误分类结果的程度。余量函数的值越大,分类性能越好。

$$mr(X, Z) = \alpha v_n I(h_n(X) = Z) - \max_{k \neq Z} \alpha v_n I(h_n(X) = k) \quad (8)$$

式中, α, v 为求解余量函数时的系数。

定义 RF 的泛化误差 PE^* 及其收敛表达式分别为:

$$PE^* = P_{X,Z}(mr(X, Z) < 0) \quad (9)$$

$$\lim_{n \rightarrow \infty} PE^* = P_{X,Z}(P_{\theta}(h_n(X) = Z)) - \max_{k \neq Z} P_{\theta}(h_n(X) = k) < 0) \quad (10)$$

因此,RF 不会随着 DT 数量的增加而产生过拟合问题,提高了分类结果的正确率。

由于电缆早期故障与其他过电流扰动在波形上存在一定的相似,直接利用原始数据作为 RF 的输入,可能不能准确识别出电缆早期故障,而且所花费的时间更长。为了提高算法的分析效率和识别准确率,对不同类型的数据进行小波变换,将经过小波变换提取的初始特征量作为 RF 模型的输入数据,通过 RF 对输入数据的学习,得出分类规则。本文研究的是三分类问题,即从正常情况、其他过电流扰动和电缆早期故障三种类型中准确地识别出电缆早期故障类型。

4 电缆早期故障的识别方法

本文利用平稳小波变换分析电力系统故障信号时的优点,对电缆早期故障信号进行平稳小波变换,

并且考虑了其他 3 种过电流扰动(变压器激磁涌流、电容投切、恒定阻抗故障),为了提高算法的效率,基于小波变换系数提取初始特征量。由于 RF 分类器在分类识别中具有良好的性能,于是将提取到的初始特征量组合成时频特征向量,最后利用 RF 从众多过电流扰动中准确识别出电缆早期故障。算法流程图如图 4 所示。

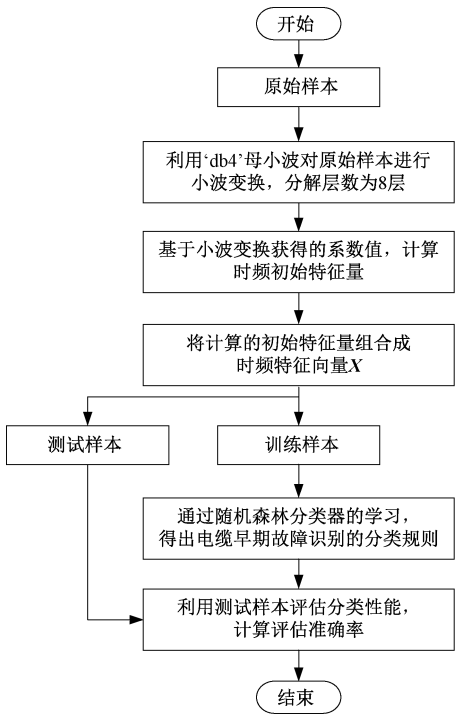


图 4 电缆早期故障识别算法流程图
Fig.4 Flowchart of algorithm for cable incipient fault identification

(1)在 PSCAD/EMTDC 中搭建电缆早期故障电路模型,获得原始信号样本,包括半周波早期故障、多周波早期故障、电容投切、变压器激磁涌流、恒定阻抗故障和正常情况,如图 5 所示。其中电缆线路用 PI 线路模型,采样率为 10kHz,仿真时间 0.5s;(1)~(4)分别表示接入故障类型

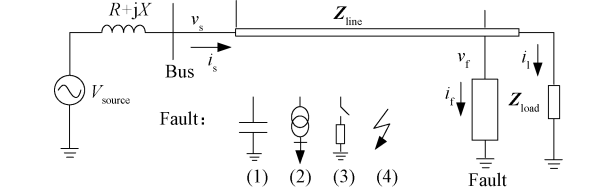


图 5 电缆早期故障仿真电路模型

(2)利用‘db4’母小波对收集到的信号进行平稳小波变换,分解层数为 8 层,共获得 8 个细节系数和 1 个近似系数,基于表 1 中的公式计算对应的特征量,最后将 81 个特征量组成时频特征量。

(3)将组成的时频特征量归一化在[0,1],作为 RF 的输入,通过 RF 组织多个 DT 对输入数据进行学习,最后经过多个 DT 的投票得出最后结果。

(4)利用测试样本测试 RF 的性能,通过输出类别和实际类别的比较,得出分类识别准确率。

5 算例分析

5.1 样本情况描述

本文研究的是三分类问题,包括电缆早期故障、其他过电流扰动和正常电流波形情况。图 5 中,电容投切用恒定电容 C 模拟,激磁涌流用容量为 S 的变压器模拟,恒定阻抗故障用电阻 Z 模拟。电缆早期故障用 Kizilcay 电弧模型^[21]来模拟,电弧模型公式为:

dg/dt = 1/tau * (|i_f|/(u_0 + r_0 * |i_f|) - g) (11)

g = i_f/v_f (12)

式中,g 表示电弧电导;tau 表示电弧时间常数;u_0 和 r_0 分别表示特征电弧电压和电阻;v_f 和 i_f 分别表示故障点的电压和电流。

为了能够获得不同类型的样本数据,设置了不同的仿真参数。其中 C、S、Z 的取值各 70 个,tau、u_0、r_0 的取值分别为 2、7、5 个,故障位置 40 处。所设置的参数符合真实的扰动情况,提高了仿真样本的真实性和特征的典型性,如表 2 所示。

表 2 电缆早期故障仿真电路的模型参数
Tab.2 Parameters of cable incipient fault simulation circuit model

参数	数值
V _{source} /kV	25
R+jX(pu)	0.35 + j0.68
Z _{line} (pu)	0.65 + j0.86
Z _{load} (pu)	6.56 + j3.03
故障位置	40 个(500~20000m,每 500m 递增)
C/mF	1~7000(每 100mF 递增)
S/(kV·A)	1~7000(每 100kV·A 递增)
Z/Ω	1~7000(每 100Ω 递增)
tau/ms	0.2~0.4(每 0.2ms 递增)
u ₀ /V	300~1500(每 200V 递增)
r ₀ /Ω	0.001~0.013(每 0.003Ω 递增)

Fig.5 Simulation circuit model for cable incipient fault

通过表 2 的仿真参数设置,改变设置不同参数组合情况,最后可以获得每种类型各 2800 个样本,其中每个样本包含 2560 个采样点。最后通过随机选取样本的原则,选择每种类型中的 2100 个样本为训练样本,剩下的 700 个样本作为测试样本来测试分类器的性能。

图 6 为各种类型的波形图。从图 6 可以得出,除正常电流波形情况外,其余 5 种类型都会造成电流的增大,表现出过电流扰动。多周波电缆早期故障与恒定阻抗故障类似,半周波电缆早期故障与电容投切故障类似。因此直接利用传统信号处理方法提取波形特征,不能准确地识别出电缆早期故障。

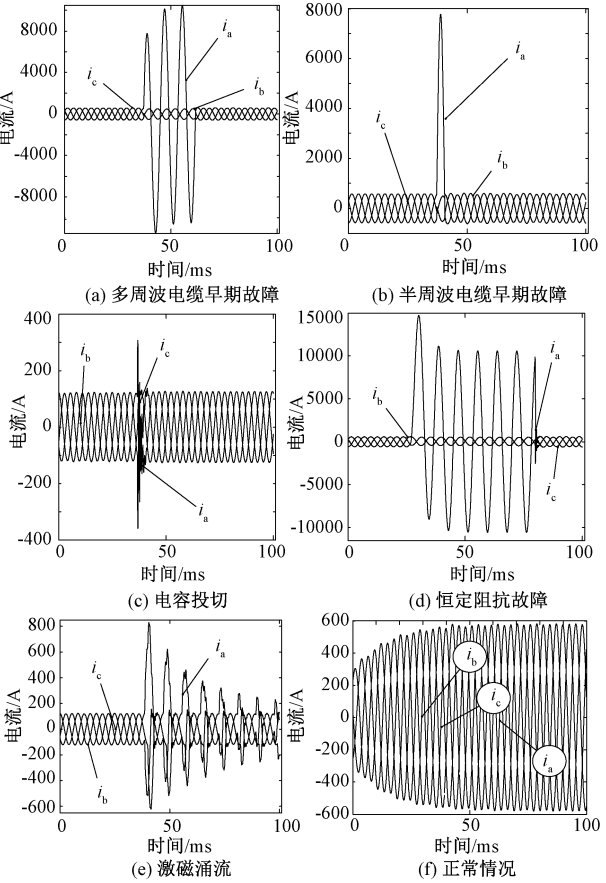


图 6 不同干扰类型的馈线电流波形图
Fig.6 Feeder current waveforms under different interference

RF 分类器能够准确地学习数据中隐藏特征,有利于对故障类型的分类识别。因此利用‘db4’母小波对不同类型的样本数据进行分解,按照表 1 的特征量计算公式提取初始特征量,最后将提取到的初始特征量归一化到[0,1],作为 RF 分类器的输入数据。

5.2 信噪比的影响

由于实际电力系统中的信号存在一定的噪声,

为了更好地符合实际信号,对测试数据分别添加不同信噪比(Signal Noise Ratio, SNR)的噪声来进行分析,所添加的噪声类型为高斯白噪声。对添加了噪声的数据分别进行 10 次测试,测试结果如表 3 所示。

表 3 不同信噪比对识别结果的影响

Tab.3 Effect of different SNR on identification results	
SNR	平均识别准确率 (%)
0	99.2
10	79.3
20	86.8
30	93.7
40	97.2
50	94.6

从表 3 可以得出,在添加了不同信噪比的噪声后,平均测试准确率略有下降,但仍能够准确识别出电缆早期故障。

5.3 所提方法与传统方法的比较

将本文方法与 DT、KNN、NB 和 PNN 相比较,所有的训练和测试样本完全一致。为了防止实验过程中的偶然性和特殊性,每种方法都进行了 10 次实验,得到最终的平均测试准确率,结果如图 7 所示。

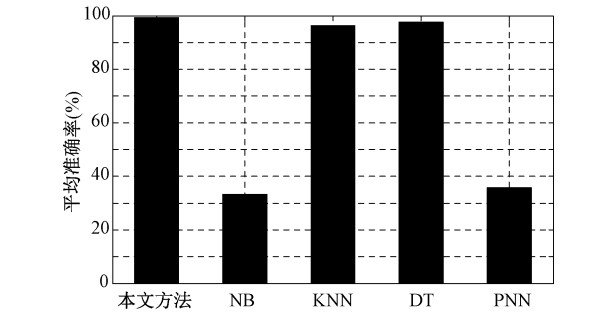


图 7 不同方法进行 10 次实验的平均准确率(SNR=0)
Fig.7 Mean identification accuracy of 10 trials using different methods (SNR=0)

为了进一步与以上方法进行比较,本文还采用了用于多分类性能的指标。

- (1)精确率(Precision)
- $$Precision = \frac{TP}{TP + FP} \times 100\%$$
- (2)召回率(Recall)
- $$Recall = \frac{TP}{TP + FN} \times 100\%$$
- (3) F_1 值:表示精确率和召回率的调和均值
- $$\frac{2}{F_1} = \frac{1}{Precision} + \frac{1}{Recall}$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \times 100\%$$

(16)

式中, TP 表示真阳性的样本数量; FP 表示假阳性的样本数量; FN 表示假阴性的样本数量。

详细的识别结果如表 4 所示。从表 4 可以得出, 本文提出方法的平均准确率、 $Precision$ 、 $Recall$ 和 F_1 值最高。其中平均准确率比 KNN 和 DT 分别高 2.96% 和 1.58%, 而且所用的仿真时间也更低, 更利于电缆早期故障的在线监测; NB 和 PNN 的平均准确率相比本文提出的方法、KNN 和 DT 则略有差距。

表 4 不同方法的识别结果 (SNR=0)

Tab.4 Identification result of different methods (SNR=0)

方法	平均识别 准确率 (%)	$Precision$	$Recall$	F_1	时间/s
本文方法	99.2	99.2	98.9	98.7	5.882
NB	32.57	32.89	0	0	1.490
KNN	96.24	95.2	89	93.2	119.500
DT	97.62	94.7	96.15	98.04	252.689
PNN	33.33	33.33	0	0	39.511

为了进一步考虑噪声对以上方法的影响, 采用 SNR 为 40dB 的数据进行测试, 测试结果如表 5 所示。从表 5 可以看出, 在添加噪声后, 本文所提方法的平均准确率、 $Precision$ 、 $Recall$ 和 F_1 值也最高, 而且相对其他方法更稳定。

表 5 不同方法的识别结果 (SNR=40dB)

Tab.5 Identification result of different methods (SNR=40dB)

方法	平均识别 准确率 (%)	$Precision$	$Recall$	F_1	时间/s
本文方法	97.2	96.8	96.1	96.45	6.871
NB	28.73	27.63	0	0	1.869
KNN	94.28	94.9	87.63	91.12	128.633
DT	83.97	83.5	84.25	83.87	259.375
PNN	32.67	31.68	0	0	43.625

本文提出方法具有较高准确率和更少的仿真时间。从故障电流的样本数据提取特征后, 将提取的特征作为 RF 分类器的输入, RF 采用 Bootstrap 采样的方法从训练样本中抽取样本, 然后建立 DT, 基于多个 DT 进行决策, 综合各个 DT 的结果进行分析, 最后通过投票机制得出结果, 能在多种类别中准确区分出电缆早期故障类型。

5.4 不同训练集和测试集的测试效果

为了体现本文所提方法的可信性和总体识别效

果, 采用不同的训练集和测试集对所提方法进行多次测试。将每种类型的 2800 个样本分成 28 组, 每组 100 个样本, 分别选择不同的组号, 最后组成不同的训练集和测试集, 测试结果如表 6 所示。

表 6 不同训练集和测试集的识别结果

Tab.6 Identification results for different train sets and test sets

组别	训练集	测试集	平均识别准确率(%)
1	1~21	22~28	97.6
2	5~25	1~4, 26~28	98.3
3	6~26	1~5, 27, 28	98.9
4	7~27	1~6, 28	98.6
5	8~28	1~7	99.1

从表 6 可以得出, 在不同的训练集和测试集下, 本文所提方法的平均识别准确率都超过 97%, 说明本文所提方法的总体识别效果很好, 能够准确地识别出电缆早期故障。

5.5 平稳小波变换提取特征的优点

本文对信号进行平稳小波变换, 计算对应的用统计参数表示的特征量, 再将特征向量作为 RF 的输入, 最后完成分类识别工作。将本文提出方法与原始数据作为 RF 的输入进行比较, 结果如表 7 所示。

表 7 RF 不同输入数据的识别结果

Tab.7 Identification result of different input data of RF

方法	平均识别准确率(%)	时间/s
本文方法	99.2	5.882
原始数据	88.7	128.57

从表 7 可以得出, 在经过平稳小波变换提取对应的统计特征后, 再将提取到的特征作为 RF 的输入, 更适用于电缆早期故障的识别, 在时间和准确率上都更优于直接将原始数据作为 RF 的输入, 证明了本文提出方法的可行性。

6 结论

本文将平稳小波变换与随机森林相结合, 提出了识别电缆早期故障的方法。并且在 PSCAD/EMTDC 中搭建了相应的电缆早期故障模型, 通过对模型收集到的数据进行分析, 得到结论如下:

(1) 提出了一种电缆早期故障特征提取算法, 通过对电缆早期故障信号进行平稳小波分解, 再基于得到的小波系数计算 81 个统计特征, 最后组成特征向量, 减少了电力系统中信号数据量大的问题, 提高了运算效率。

(2)将机器学习算法引入对电缆早期故障识别中,通过 RF 模型的分类,最后可以得到识别平均准确率为 99.2% (无噪声) 和 97.2% (信噪比为 40dB),相比其他传统模式识别方法具有一定的优势。

(3)将机器学习算法应用于电力系统中的故障诊断,为以后电力系统的故障诊断提供了一个新的思路,值得进一步研究。

参考文献 (References):

- [1] 徐铭铭,肖立业,王海风,等 (Xu Mingming, Xiao Liye, Wang Haifeng, et al.). 一种基于 Prony 算法的直流配电网电缆故障定位方法 (Novel method of locating cable fault in DC distribution based on Prony algorithm) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2015, 34 (4): 1-5.
- [2] 戴铭 (Dai Ming). 10kV 地下电缆早期故障检测与识别方法探讨 (Discussion methods of incipient fault detection and identification in 10 kV underground cables) [D]. 成都: 西南交通大学 (Chengdu: Southwest Jiaotong University), 2012.
- [3] Saurabh Kulkarni, Surya Santoso, Thomas A Short. Incipient fault location algorithm for underground cables [J]. IEEE Transactions on Smart Grid, 2014, 5 (3): 1165-1174.
- [4] Rohit Moghe, Mirrasoul J Mousavi, James Stoupis, et al. Field investigation and analysis of incipient faults leading to a catastrophic failure in an underground distribution feeder [A]. IEEE/PES Power Systems Conference and Exposition [C]. 2009. 1-6.
- [5] Sidhu T S, Xu Zhihan. Detection of incipient faults in distribution underground cables [J]. IEEE Transactions on Power Delivery, 2010, 25 (3): 1363-1371.
- [6] 姚海燕,张静,留毅,等 (Yao Haiyan, Zhang Jing, Liu Yi, et al.). 基于多尺度小波判据和时频特征关联的电缆早期故障检测和识别方法 (A method of incipient cable fault detection and identification based on multi-scale wavelet criterion and time-frequency feature correlation) [J]. 电力系统保护与控制 (Power System Protection and Control), 2015, 43 (9): 115-123.
- [7] Butler-Purpy Karen L, Cardoso J. Characterization of underground cable incipient behavior using time-frequency multi-resolution analysis and artificial neural networks [A]. 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century [C]. 2008. 1-11.
- [8] Mousavi Mirrasoul J, Butler-Purpy Karen L. Detecting incipient faults via numerical modeling and statistical change detection [J]. IEEE Transactions on Power Delivery, 2010, 25 (3): 1275-1283.
- [9] Teymoor Ghanbari. Kalman filter based incipient fault detection method for underground cables [J]. IET Generation, Transmission & Distribution, 2015, 9 (14): 1988-1997.
- [10] Zhang Wenhai, Xiao Xianrong, Zhou Kai, et al. Multi-cycle incipient fault detection and location for medium voltage underground cable [J]. IEEE Transactions on Power Delivery, 2017, 32 (3): 1450-1459.
- [11] 陈敬德,盛戈皞,吴继健,等 (Chen Jingde, Sheng Gehao, Wu Jijian, et al.). 大数据技术在智能电网中的应用现状及展望 (Application status and prospect of big data technology in smart grid) [J]. 高压电器 (High Voltage Apparatus), 2018, 54 (1): 35-43.
- [12] 黄青平,李玉娇,刘松,等 (Huang Qingping, Li Yujiao, Liu Song, et al.). 基于模糊聚类与随机森林的短期负荷预测 (Short-time load forecasting based on fuzzy clustering and random forest) [J]. 电测与仪表 (Electrical Measurement & Instrumentation), 2017, 54 (23): 41-46.
- [13] 陈俊尧,王志新 (Chen Junyao, Wang Zhixin). 基于贝叶斯网络 Noisy-Or 模型的低压漏电保护器机械组件故障诊断应用 (Fault diagnosis for mechanical components of leakage protection devices based on Bayesian networks Noisy-Or model) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2017, 36 (8): 74-79.
- [14] 徐立娟,吴春华,王元章 (Xu Lijuan, Wu Chunhua, Wang Yuanzhang, et al.). 基于决策树的光伏组件故障诊断方法研究 (Survey of fault diagnosis for PV modules based on decision tree) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2017, 36 (6): 83-88.
- [15] 朱念芳,林善明 (Zhu Nianfang, Lin Shanming). 基于邻域 KNN 算法的风电功率短期预测模型 (Short-term wind power prediction model based on KNN algorithm considering neighbors' density) [J]. 电测与仪表 (Electrical Measurement & Instrumentation), 2017, 54 (16): 20-24.
- [16] Suhail Khokhar, Abdullah Asuhaimi Mohd Zin, Aslam Pervez Memonb, et al. A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network [J]. Measurement, 2017, 95: 246-259.
- [17] Nason G P, Silverman B W. The stationary wavelet transform and some statistical application in wavelet and statis-

tics [M]. Berlin: Springer Verlag, 1995.

[18] 张士文, 张峰, 王子骏, 等 (Zhang Shiwen, Zhang Feng, Wang Zijun, et al.). 一种基于小波变换能量与神经网络结合的串联型故障电弧辨识方法 (Series arc fault identification method based on energy produced by wavelet transformation and neural network) [J]. 电工技术学报 (Transactions of China Electrotechnical Society), 2014, 29 (6): 290-295.

[19] Mallat S G. A theory for multi-resolution signal decomposition: The wavelet representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11 (7): 674-693.

[20] 方匡南, 吴见彬, 朱建平, 等 (Fang Kuangnan, Wu Jianbin, Zhu Jianping, et al.). 随机森林方法研究综述 (A review of technologies on random forests) [J]. 统计与信息论坛 (Statistics & Information Forum), 2011, 26 (3): 32-38.

[21] Kizilcay M, Pniok T. Digital simulation of fault arcs in power systems [J]. European Transactions on Electrical Power, 1991, 1 (1): 55-60.

Cable incipient fault identification based on stationary wavelet transform and random forest

LI Sheng-hui¹, BAI Xue¹, DONG He-nan¹, LU Hong², GUO Chao-yun²

(1. Electric Power Research Institute, State Grid Liaoning Power Supply Co., Ltd., Shenyang 110006, China;
2. College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Analyzing a large amount of power quality data acquired from the power quality monitor provides a new direction to solve faults of device in the power system. Cables are important devices, and its incipient fault may occur for a period of time before a permanent fault occurs in the cable. Although the incipient fault will not cause the protection device to operate, it may cause a permanent fault and reduce the safety of the power system operation. In this paper, a cable incipient fault identification method based on stationary wavelet transform and random forest is proposed. First, the fault phase current signal is decomposed by using stationary wavelet-transform, and the statistical features are computed from wavelet coefficients. Finally, the extracted features are taken as the input of the random forest classifier, and the cable incipient fault is identified from various over-current disturbances. This method is compared with traditional pattern recognition method. The results show that the proposed method has higher recognition accuracy and takes shorter time.

Key words: stationary wavelet transform; random forest; cable incipient fault; over-current disturbance