

一种可用于绝缘子沿面放电状态识别的算法

贺博, 高乃奎, 万军, 张爱德, 孙健, 彭宗仁

(西安交通大学电力设备电气绝缘国家重点实验室, 710049, 西安)

摘要: 在深入分析污秽绝缘子沿面放电特性的基础上, 发展建立了一种新型的模式识别算法. 该算法借鉴 Bayes 分类原理来实现电流的判别, 借鉴模糊数学理论来实现湿度的判别, 然后通过仲裁机制在综合电流、湿度信息的基础上确定出状态类属. 检验结果表明, 该算法具有好的通用性, 利用一次试验数据训练出的算法在完成另外次试验数据的识别时具有很高的准确率, 并且误判均发生在相邻状态, 所以非常适合基于泄漏电流和湿度信号的放电状态识别.

关键词: 沿面放电; 模式识别; 模糊化; 交互验证

中图分类号: TM216.1 **文献标志码:** A **文章编号:** 0253-987X(2008)10-1269-06

Novel Algorithm for Discerning Surface Discharge State of Contaminated Insulators

HE Bo, GAO Naikui, WAN Jun, ZHANG Aide, SUN Jian, PENG Zongren

(State Key Laboratory of Electrical Insulation and Power Equipment, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: It is important for accomplishing online monitoring and risk prediction of flashover from contaminated insulators to develop efficient algorithm for identifying surface charge state of insulators. Following the analysis of surface discharge character, a novel discernment algorithm is proposed, where the discernment process contains three steps: the discharge states are discerned based on leakage current via Bayes discernment algorithm; the stages are discerned based on humidity through fuzzy theory; the results from the first and two steps are combined by an arbitration operation and the final result about the discharge state is achieved. The validity testing results show that the novel algorithm is endowed with a good universality and the algorithm trained out by one group data enables to identify the discharge state of the other artificial test with good accuracy and the mistake happens within the adjacent states. So the algorithm is suitable for discerning surface discharge state of HV insulators through LC and humidity signal.

Keywords: surface discharge; model discernment; fuzzy algorithm; cross-validity

发展建立高效的污秽闪络在线监测和风险预警机制, 对于预防污闪、增加电网的可靠性具有重要意义. 目前, 针对污闪在线监测的研究已成为热点, 提出了不同的监测方案, 但其中最具应用价值的是基于泄漏电流(LC)的方案, 主要原因是泄漏电流信号在采集传输上便于实现^[1]. 目前, 泄漏电流的采集和传输技术已经趋于成熟, 但在信号的综合分析和应用上, 还存在困难. 本文的目的是发展一种可以适用于泄漏电流信号分析的算法, 为实现沿面放电状态

识别和闪络发生风险预测提供参考.

1 沿面放电及泄漏电流特征分析

1.1 试验设计

本文泄漏电流数据来自预积污试品恒压、洁净雾条件下的人工污秽试验. 图1为试验接线图, 其中调压变压器规格为 220 V/(0~110) V, 500 kV·A; 变压器规格为 110 V/110 kV, 650 kV·A; 电压互感器规格为 150 kV/100 V, 试验电源可以提供

10 A 的短路电流. 试品为 XP-70 型普通悬式瓷绝缘子, 雾室尺寸(长×宽×高)为 1.8 m×1.8 m×2.0 m. 污秽度灰密值(NSDD)均取 1×10^{-3} g/cm², 盐密值(ESDD)取 0.1×10^{-3} g/cm². 使用慢受潮法使试品受潮, 整个试验过程中始终控制雾室的湿度和温度. 试品每组 7 片, 所加电压为 63.5 kV, 模拟 110 kV 输电线路配置.

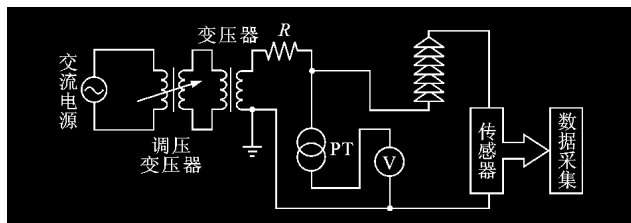


图1 人工污秽试验接线图

按 5 000 次/s 的速率采样泄漏电流, 每采集 10 个点即求取平均值并保存, 每 0.02 s 求得所保存的 10 个历史平均值中的最大值作为有用数据. 这样, 每秒可以得到 50 个数据点, 均反映了泄漏电流最大值的变化. 在数据处理过程中有求均值的过程, 使得噪声得到了很大的抑制, 同时需保存、传输的数据量也大大减少, 降低了系统对硬件的要求, 所得到实际泄漏电流的包络线表明了沿面放电的强度.

1.2 泄漏电流特征分析

图 2 为实测的泄漏电流波形, 可见随着湿度的增加和沿面放电的发展, 泄漏电流包络线总体上呈增大的趋势, 结合试验观察并参考相关研究结果, 整个试验过程中沿面放电可以划分为 8 个状态^[2], 具体如表 1 所示.

按照污秽闪络的机理, 影响沿面放电的两个主要因素是环境湿度和表面受污情况, 而泄漏电流只是放电的外在表现. 泄漏电流的大小并不能完全表征放电状态和污闪风险, 因为较低的污秽水平在较大的湿度影响下可以具有和较大污秽水平在较低湿度下相同的电流幅值, 但相同的泄漏电流却来自于不同的放电阶段, 前者的放电阶段等级通常要高于后者^[3].

对于表 1 所示的放电阶段来说, 阶段 1 为试验之初, 试品表面保持干燥的情况, 反映的是正常绝缘水平下的状态, 此时表面无放电, 电流幅值恒定, 无冲击起伏, 状态易于判断, 在此不加以讨论. 阶段 7 和 8 为临闪状态, 此时相对湿度通常很大($\geq 91\%$), 电流的幅值和冲击程度也非常大, 对于这一阶段下的闪络预警, 可以采用极值风险函数的方法^[3], 不在本文讨论的范围内. 本文只讨论阶段 2~6 下的放电

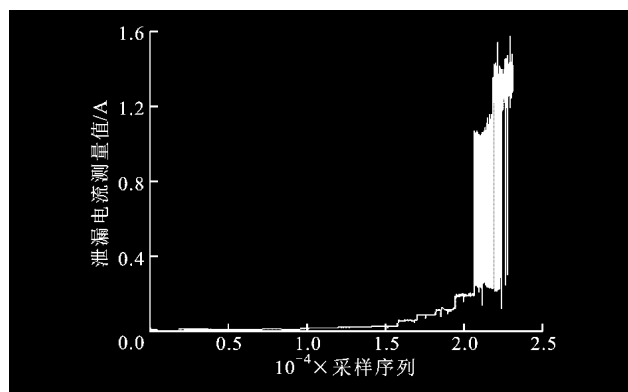


图2 泄漏电流实测波形(采样频率为 50 112)

模式识别方法.

表1 XP7 绝缘子污秽试验过程放电状态

放电阶段	放电特征
1	无明显放电现象
2	紫火星, 淡紫丝状放电, 微声
3	紫色刷状放电, 黄、白小火星
4	桔黄色短电弧, 脉冲密集, 放电持续, 声大
5	脉冲频率降低, 1/3 泄距
6	脉冲间隔加长, 明亮桔红色
7	主电弧, 沟槽内密集小电弧
8	强烈放电, 几乎贯通泄距, 仍能耐受
	红色电弧, 随时可能发生闪络

2 算法设计

污秽闪络机理复杂, 影响因素众多, 导致了人工污秽试验随机性很强, 重复性很差. 就利用泄漏电流实现沿面放电监测和闪络预警来讲, 难点和重点是借助人工污秽试验电流特征, 建立起有效的识别算法, 算法应该具有良好的通用性, 可以适用于多组试验数据的判别. 经上述分析, 制定算法的实质是挖掘数据中蕴涵的统计规律, 所以本文考虑用统计模式识别理论来处理问题.

2.1 传统识别算法及其不足

传统模式识别算法主要有 Bayes 分类算法和 Fisher 分类算法.

2.1.1 Bayes 分类算法的特点和适用性分析
Bayes 分类算法的基础是 Bayes 公式, 分类的原则是保证所有样本对于所属类具有最大的附属概率

$$P(\omega_i | X) > P(\omega_j | X) \Rightarrow X \in \omega_i \quad (1)$$

求解式(1)是 Bayes 分类算法实现的难点, 因为理想的求解结果应该是将 p 维空间 Ω^p 进行完全互斥划分, 划分成 R 个子空间, 重叠的部分将导致判别错

误^[4].所以, Bayes 识别算法目前针对两类问题:各模式均服从正态分布,分布密度已知或者可以利用样本加以估计,并且样本方差齐性,不同的模式分布具有相同的方差;样本各分量值限于 0、1 的问题.

传统 Bayes 算法难以适用于污闪识别,因为我们要根据电流和湿度来实现判别,而 2~6 各阶段的电流显然不具有相同的方差,再者湿度也不是服从正态分布的随机变量,而是受人为控制的受控量.

2.1.2 Fisher 分类算法的特点和适用性分析

Fisher 判别是基于坐标变换的思想,通过选择合适的坐标将待判别数据作投影,并使得在新坐标下各观测数据在不同组内差异较大,在相同组内方差较小,从而实现分组判别.详细算法可参考文献[4].

理论上讲, Fisher 判别不需要正态分布的限制,也不要求方差齐性,但由于涉及到主因子的求取,对于人工污秽试验来讲,沿面放电通常发生于比较大的湿度水平下,并且随着湿度的增加,放电幅值不断增加.所以,在利用 Fisher 判别时,算法训练过程容易受单一量的影响(由于电流的信息在训练中权值较小,训练结果过分依靠湿度),导致了训练结果通用性很差,训练出的方法只能用于该次试验数据的判别,而不能适用于其他次试验的准确判别.

2.2 新型识别算法结构及流程

图 3 为 2~6 阶段的数据特征示意图,反映出了沿面放电发展的基本特性:随着湿度的增加,电流幅值和方差不断增加,湿度和电流共同来表征不同放电状态.

本文利用以下思路来设计识别算法:先分别计算电流和湿度属于不同类属(放电阶段)的概率,取附属概率最大的类为判断结果,如果两者的判断结果相同,则将数据归于此类,如果类属判断结果不相同,则综合利用湿度、电流的判断结果,利用合适的仲裁机制来决定数据的类属.算法具体步骤如下.

(1) 泄漏电流数据预处理. 首先拟合出电流和湿度的关系 $i = f(h)$, 进而求出电流的总残差, 然后通过各类属所包含的湿度跨度, 将总残差按照湿度跨度比例分配到各类属中去, 并认为在同一类属中, 方差是齐性的. 这样, 组内方差的不均匀性将大大减小. 总方差的求取为

$$S = \sum_{j=1}^n (i_j - f(h_j))^2 \quad (2)$$

各组内方差为

$$S_i = \frac{d_{hi} S}{\sum_{j=1}^5 d_{hj}} \quad (3)$$

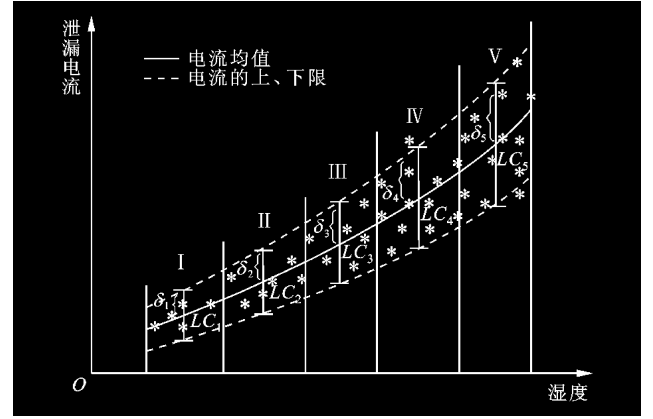


图3 人工污秽试验数据趋势示意图

式中: d_{hi} 为第 i 组的湿度范围, 通过对试验数据作距离聚类而获得, 同时还可以求出各组内数据个数 n_i 和电流中心 $\bar{i}_i$.

利用组内方差来实现组内标准差的估计

$$\delta_i = \left(\frac{S_i}{n_i - 1} \right)^{1/2} \quad (4)$$

先验概率的估计为

$$p(\omega_i) = \frac{n_i}{\sum_{i=1}^5 n_i} \quad (5)$$

(2) 电流归属的 Bayes 判别. 由于 $p(X|\omega) \approx N\left(\frac{i_x - \mu}{\delta}\right)$, 其中 μ, δ 为各组的均值和方差, 所以式(1)等价于

$$N\left(\frac{i_x - \mu_i}{\delta_i}\right) P(\omega_i) > N\left(\frac{i_x - \mu_j}{\delta_j}\right) P(\omega_j) \Rightarrow X \in \omega_i \quad (6)$$

用式(6)来实现电流归属的判别.

(3) 利用待判断电流对应的湿度来决定湿度判据下数据的归属, 将数据判入湿度分量所属的类空间, 如

$$h_x \in R_{hi} \Rightarrow X \in \omega_i \quad (7)$$

式中: R_{hi} 为第 i 类对应的湿度区间.

(4) 如果(2)、(3)步的判断结果相同, 则完成数据的类属判别, 程序返回, 否则重新计算湿度属于各类属的隶属度, 并经由仲裁机制来决定数据类属, 如第(5)步所示.

(5) 湿度隶属度计算. 选用高斯形函数来计算隶属度

$$g(h) = \exp\left(-\frac{(h - h_i)^2}{2\sigma^2}\right) \quad (8)$$

式中: h_i 为待判断湿度大小; σ 为参变量, 文中取为 15 (对应湿度范围 35%~99%), 式(8)函数值位于

$(0,1]$ 之间,并且在 h_i 处具有峰值 1. 利用

$$A = \int_{h_{\min}}^{h_{\max}} g(h) d(h) \quad (9)$$

计算高斯函数的面积,其中积分上、下限 $h_{\min}$ 、 $h_{\max}$ 分别表示试验进行中湿度的上、下限,文中取 30%、99%,利用

$$g_1(h) = \frac{g(h)}{A} \quad (10)$$

来修正 $g(h)$,得到隶属度密度函数 $g_1(h)$,该函数在待判断数据处具有峰值,并在整个湿度范围内积分为 1. 利用

$$P(h_i \in \omega_j) = \int_{\omega_{j1}}^{\omega_{j2}} g_1(h) d(h) \quad (11)$$

计算湿度属于类属 ω_j 的概率 $P(h_i \in \omega_j)$, ω_{j1} 、 ω_{j2} 为 j 类湿度的上、下限.

(6) 仲裁机制. 将式(6)和式(11)得到的电流和湿度分属不同类的概率求和,得到和值为不同类属的概率,取和概率最大的类属为类属判断结果,如下

$$R_i = P_i + P'_i \quad (12)$$

$$R_j > R_i \Rightarrow X \in \omega_i \quad (13)$$

在判断过程中,第一次判断时,电流先验概率、均值和方差均由聚类结果初始化,此后一旦数据归属被确定,立即更新组内电流的先验概率、均值和方差,对于湿度,则保持聚类分析的结果不变,不参与更新.

3 结果合理性验证

表2、表3分别为经典 Bayes 分类和 Fisher 分类方法下的交互验证结果,交互验证的方法为从污秽试验中人工挑选出一批数据,这些数据具有明显的物理特征,可以准确判断出其所属的污闪阶段(金牌数据),按比例取出一定数量的数据进行训练,在训练完成后,利用剩余的数据来检验训练效果,用于训练的数据不再用于检验.

3.1 经典算法识别结果

由交互验证结果可知,对于经典 Bayes 和 Fisher 分类方法来说,在实现人工污秽数据的分类上均表现出了比较好的正确率,其中 Fisher 方法由于没有正态分布和方差齐性的要求,正确率更高. Bayes 法尽管要求试验数据来自于多维正态总体,但由于在试验过程中,湿度指标的变化水平比较少,使得不同污闪发展阶段湿度差异较大,造成了仅按照湿度即可以分辨出大部分数据类别的情况. 这在某种程度上掩盖了数据来源不服从正态和方差不齐的缺陷,所以 Bayes 法也表现出了较好的正确率.

为了进一步检验分类方法的通用性,本文利用以上经过训练的分类器对另外一组试验数据进行判别. 训练数据和检验数据来自于不同的试验过程,但该两组试验是平行试验,在试品特征、参数和试验安排上完全相同. 由于试验的随机性,在相同污闪阶段内两组数据存在差异.

表4、表5组间检验结果表明,这两种方法在执行组间判别时,判别准确率大大降低,通用性表现很差. 所以,利用经典 Bayes 或 Fisher 判据在同一批数据内实现训练和判别可以得到比较满意的准确率,但这种训练后的判别方法缺乏通用性,如果利用其去判断另外一组试验数据,误判率将非常大,对于随机性比较强的人工污秽试验数据来讲,意义不大.

3.2 改进后算法识别结果

采用和4.1部分相同的方法来检验改进后算法的判断结果,分组内验证和组间验证两部分. 其中组内验证采用交互验证法,训练和检验数据来自于同一组试验,但在用途上加以区分,不能雷同. 组间检验则是采用两组平行试验的数据,其中一组用于算法训练,另外一组用于检验,主要用来检验算法的通用性.

表6、表7为组内检验和组间检验结果,与经典方法识别结果相对比,改进后的分类算法不仅对于同组试验数据作交互检验表现出了很高的正确率,

表2 交互验证结果 I (Bayes 判别)

数据来源	模式1		模式2		模式3		模式4		模式5		合计	
	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%
1	673	93.9	27	3.8	14	2.0	3	0.4	0	0.0	717	100
2	16	6.5	218	88.3	13	5.3	0	0.0	0	0.0	247	100
3	1	0.3	15	5.2	263	91.6	8	2.8	0	0.0	287	100
4	0	0.0	6	4.5	10	7.5	118	88.1	0	0.0	134	100
5	0	0.0	0	0.0	9	6.3	13	9.2	120	84.5	142	100

表 3 交互验证结果 II (Fisher 判别)

数据来源	模式 1		模式 2		模式 3		模式 4		模式 5		合计	
	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%
1	690	96.2	18	2.5	9	1.3	0	0.0	0	0.0	717	100
2	9	3.6	232	93.9	6	2.4	0	0.0	0	0.0	247	100
3	0	0.0	13	4.5	269	93.7	5	1.7	0	0.0	287	100
4	0	0.0	0	0.0	7	5.2	127	94.8	0	0.0	134	100
5	0	0.0	0	0.0	6	4.2	11	7.7	125	88.0	142	100

表 4 组间验证结果 I (Bayes 判别)

数据来源	模式 1		模式 2		模式 3		模式 4		模式 5		合计	
	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%
1	56	56.0	34	34.0	9	9.0	1	1.0	0	0.0	100	100.0
2	18	18.0	59	59.0	22	22.0	1	1.0	0	0.0	100	100.0
3	0	0.0	23	23.0	48	48.0	29	29.0	0	0.0	100	100.0
4	0	0.0	2	2.0	17	17.0	65	65.0	16	16.0	100	100.0
5	0	0.0	1	1.0	14	14.0	22	22.0	63	63.0	100	100.0

表 5 组间验证结果 II (Fisher 判别)

数据来源	模式 1		模式 2		模式 3		模式 4		模式 5		合计	
	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%
1	66	66.0	26	26.0	8	8.0	0	0.0	0	0.0	100	100.0
2	12	12.0	71	71.0	17	17.0	0	0.0	0	0.0	100	100.0
3	0	0.0	14	14.0	64	64.0	22	22.0	0	0.0	100	100.0
4	0	0.0	2	2.0	13	13.0	73	73.0	12	12.0	100	100.0
5	0	0.0	0	0.0	9	9.0	24	24.0	67	67.0	100	100.0

表 6 改进后算法交互验证结果

数据来源	模式 1		模式 2		模式 3		模式 4		模式 5		合计	
	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%
1	705	98.3	12	1.7	0	0.0	0	0.0	0	0.0	717	100.0
2	3	1.2	240	97.2	4	1.6	0	0.0	0	0.0	247	100.0
3	0	0.0	3	1.1	279	97.2	5	1.7	0	0.0	287	100.0
4	0	0.0	0	0.0	2	1.5	127	94.8	5	3.7	134	100.0
5	0	0.0	0	0.0	0	0.0	4	2.8	138	97.2	142	100.0

表 7 改进后算法组间验证结果

数据来源	模式 1		模式 2		模式 3		模式 4		模式 5		合计	
	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%	数量	比例/%
1	91	91.0	0	9.0	0	0.0	0	0.0	0	0.0	100	100.0
2	3	3.0	93	93.0	4	4.0	0	0.0	0	0.0	100	100.0
3	0	0.0	2	2.0	93	93.0	5	5.0	0	0.0	100	100.0
4	0	0.0	0	0.0	3	3.0	95	95.0	2	2.0	100	100.0
5	0	0.0	0	0.0	0	0.0	9	9.0	91	91.0	100	100.0

对于来自不同次试验的数据同样表现出了很高的判别正确率。

对比表 6 和表 7 可知,对于组内判别,其精度要高于组间检验,但误判数据均集中在相邻的组内,没

有发生大的误判,所以,对于人工污秽试验数据判断来讲,改进算法具有比较好的兼容性,即通过一次试验训练出判别规则后,可以以较高的精度实现其他试验数据的判断,并且误判结果均限于相邻状态.

4 结 论

本文在分析绝缘子沿面放电具体数据特征和基本性质的基础上,发展了一种新型的识别算法,用于实现2~6放电阶段的识别.检验结果表明,改进算法具有良好的通用性,利用一次试验数据训练出的算法可以准确完成另外组试验数据的检验,并且误判只限于相邻状态,即便误判也不会出现大的偏差.所以,该算法非常适用于高压绝缘子沿面放电状态监测.

参考文献:

- [1] 贺博,林辉. 绝缘子污秽闪络的研究现状及思索[J]. 电瓷避雷器,2006,44(2):7-11.
- HE Bo, LIN Hui. Present status and future prospects for flashover research of contaminated insulators [J]. Insulator And Surge Arresters, 2006, 44(2): 7-11.
- [2] 李治,王永江,王黎明,等. 绝缘子泄漏电流和放电现象初步研究[J]. 高电压技术,2002, 28(120):1-4.
- LI Zhi, WANG Yongjiang, WANG Liming et al. Research of leakage current and discharge phenomena of Insulators [J]. High Voltage Engineering, 2002, 28 (120): 1-4.
- [3] AMARH F. Electric transmission line flashover prediction system[D]. Tempe, USA: Arizona State University, 2001:30-100.
- [4] 秦超英,赵选民,师义民,等. 数理统计[M]. 西安:西北工业大学出版社,1999.

(编辑 杜秀杰)