

基于模糊粒子群算法的变压器故障自动识别^①

朱苏航¹ 吕干云² (1 金华职业技术学院 信息工程学院 浙江金华 321007;

2 浙江师范大学 信息学院 浙江 金华 321004)

摘要: 提出了一种新的基于模糊粒子群算法的电力变压器故障自动识别方法。首先对基于油中溶解气体分析得到五种关键气体含量数据进行特殊预处理,得到识别四种故障需要的六个关键特征。然后给出了一个新的模糊聚类目标函数,在此基础上,根据已有的故障样本利用粒子群算法得到各类故障的最优聚类中心;并由此计算出各测试样本到各个聚类中心之间的距离以及相应的隶属度,最后识别出样本的变压器故障类型。测试结果显示,该方法能有效诊断识别出变压器高能放电、过热、低能放电和正常状态,精度可达92%。

关键词: 故障诊断;模糊粒子群;最优聚类中心;电力变压器

Fault Diagnosis of Power Transformer Based on Fuzzy PSO Algorithm

ZHU Su-Hang¹, LV Gan-Yun²

(1. Department of Information Science and Engineering, Jinhua College of Profession and Technology, Jinhua 321004, China; 2. Department of Information, Zhejiang Normal University, Jinhua 321004, China)

Abstract: A fuzzy particle swarm optimization (PSO) method is applied to fault diagnosis of power transformer for the first time. Content of five diagnostic gases dissolved in oil obtained by dissolved gas analysis (DGA) is preprocessed through a special data processing, and six features are extracted for fuzzy PSO algorithm. Then a new objective function is proposed for fuzzy clustering algorithm. Based on the function, PSO algorithm is trained for get the optimized clustering centers of all fault type. With the optimized clustering centers, the distance of the testing sample to centers are calculated, and then the membership degree is obtained. Finally, the four fault types of transformer are identified. The algorithms perform well in the testing, and the correct ratios of fault diagnosis reach an average of 92%.

Keywords: fault diagnosis; fuzzy PSO algorithm; optimized clustering centers; power transformer

1 引言

变压器作为电力系统的枢纽设备,其故障检测不仅有利于保证其安全运行,而且对于提高电力系统的可靠性都将有重要的影响。对变压器故障进行识别,是合理制定维护措施的重要依据。

数十年的经验表明,变压器油中的溶解气体的组成比例、浓度、种类密切相关。电力变压器内部故障

是变压器油中气体含量突然增长的主要原因,绝大部分变压器的缺陷都是从色谱分析发现的。1997年正式修订实施的DL/T596-1996《电力设备预防性试验规程》中,把色谱分析列为电力变压器的首位实验项目,色谱分析法是检测变压器内部故障性质的重要方法,目前被广泛运用^[1]。结合变压器油中特征气体的含量信息,人工神经网络、专家系统、模糊技术、

^① 基金项目:浙江省自然科学基金(Y1090182)

收稿时间:2009-05-25;收到修改稿时间:2010-03-21

灰色理论、支持向量机和遗传算法等在变压器故障识别领域具有广泛的应用^[2-7]，也取得了较好的效果。

粒子群优化算法(PSO)^[8]是一种基于群体的具有全局搜索能力的随机优化算法，已在电力系统无功优化^[9]和电力市场^[10]等取得了良好应用。本文提出了一种基于模糊粒子群算法的变压器故障识别方法，首先利用粒子群算法得到各类变压器故障的最优聚类中心，然后计算出待测样本到各聚类中心距离和相应的模糊隶属度，最后诊断识别出样本的变压器故障类型。

2 粒子群算法

PSO算法于1995年由Kennedy和Eberhart提出^[10]，是一种基于群体的具有全局搜索能力的随机优化算法。通常的优化问题可表述为如下形式：

$$\min f(X) \quad (1)$$

$$\text{s.t. } g(X) \geq 0 \quad (2)$$

$$X=(x_1, x_2, \dots, x_m) \quad (3)$$

其中， $f(X)$ 为目标函数， $g(X) \geq 0$ 为约束条件。

PSO不能直接处理约束条件，采用罚函数形式将其加入到目标函数，则PSO优化的适应度函数为：

$$\min F(X)=f(X)+U\{\min[g(X),0]\}^2 \quad (4)$$

其中 U 为惩罚系数。则基本PSO算法流程如下：

①随机生成初始粒子位置 $X_i=(x_{i1}, x_{i2}, \dots, x_{im})$ 和初始速度 $V_i=(v_{i1}, v_{i2}, \dots, v_{im})$ ， $i=1, 2, \dots, n$ 。n为粒子数；

②验证粒子位置、速度的等式约束及变量上下限约束；

③将各粒子位置 X_i 代入优化适应度函数 $F(X)$ ，计算各粒子适应度值；

④将粒子 i 适应度值与其个体最优适应度值 P_i 比较，若比 P_i 小(最小化)，则用当前值替换 P_i ，并用当前粒子位置 X_i 更新自身最优粒子位置 $X_{ibest}=(x_{pi1}, x_{pi2}, \dots, x_{pim})$ ；

⑤将各粒子适应度值与全局最优适应度值 P_g 进行比较，如比 P_g 小(最小化)，用此值替换 P_g ，并用此粒子位置 X_i 更新粒子群全局最优粒子位置 $X_{gbest}=(x_{g1}, x_{g2}, \dots, x_{gm})$ ；

⑥用式(5)和式(6)更新各粒子速度和位置：

$$v_{id}^{k+1} = wv_{id}^k + c_1r_1(xp_{id} - x_{id}^k) + c_2r_2(x_{gd} - x_{id}^k) \quad (5)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (i=1, 2, \dots, n, \quad d=1, 2, \dots, m) \quad (6)$$

其中， w 为速度惯性权重，通常取0.4~0.9， c_1 为认知权重， c_2 为社会学习权重， c_1, c_2 通常取2， r_1, r_2 为(0,1)间随机数，上标表示迭代次数， n 为粒子数， m 为粒子维数，也就是待优化参数的个数。 v_{id}, x_{id} 和 xp_{id} 分别表示第 i 个粒子的第 d 维的速度、当前坐标和最优坐标， x_{gd} 为全局最优粒子的第 d 维坐标；

⑦返回②，直到满足一定的收敛判断条件。

3 模糊粒子群的变压器故障识别

设变压器故障样本数据集 $X=\{x_1, x_2, \dots, x_n\} \subset R^m$ 是 m 维向量空间中的一个子集， $x_k=\{x_{k1}, x_{k2}, \dots, x_{km}\} \subset R^m$ 为样本 $x_k(k=1, 2, \dots, n)$ 的特征向量， x_{kj} 为特征向量 x_k 在第 j 维特征上的值。对于样本集 X 的模糊聚类分析就是要产生 X 的 c 划分。本文模糊聚类采用目标函数形式为：

$$\text{Min } J_q(V) = \sum d_{ik}^2 - \alpha \sum d_{ik}^2 \quad (7)$$

其中， $V=\{v_1, v_2, \dots, v_c\}$ ，代表 c 个聚类中心向量， $v_i=(v_{i1}, v_{i2}, \dots, v_{im}) \in R^m$ ， $d_{ik}=\|x_k - v_i\|$ 为样本 x_k 与第 i 个聚类中心向量 v_i 之间的距离，式(7)前一项代表本类样本到该聚类中心向量 v_i 之间的距离最短，后一项代表其它类的样本到该聚类中心向量 v_i 之间的距离尽量大， α 为小于1的权系数。

在上述目标函数中，样本 x_k 与聚类中心 v_i 之间的距离通常采用欧氏距离，即

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{kj} - v_{ij})^2} \quad (8)$$

基于模糊粒子群的变压器故障识别步骤如下：

①采用特殊方法进行数据预处理，将变压器故障样本的特征提取到特征向量中；

②初始化聚类类别数 $c(c=4)$ ，训练样本个数 n ，测试样本个数 l ，样本特征值维数 m ，权系数 α ，初始化聚类中心矩阵 V ，设置PSO算法参数；

③利用训练样本，将目标函数(7)设为PSO适应度函数，采用PSO计算最优聚类中心矩阵 V ；

④由中心矩阵 V ，计算测试样本 x_k 与各聚类中心 v_i 之间的距离 d_{ik} ，并根据下式计算出模糊隶属度矩阵 U ：

$$\mu_{ik} = \left(\frac{1}{d_{ik}} \right) / \left(\sum_{i=1}^c \frac{1}{d_{ik}} \right) \quad (i=1, 2, \dots, c; k=1, 2, \dots, l) \quad (9)$$

⑤根据隶属度矩阵 U ，判别出测试样本的故障类型。

模糊聚类算法一定程度上克服了对数据内在形状分布的依赖,能对不同形状分布数据正确聚类,计算速度快,而且克服了对噪声及野值数据的敏感,增强了算法鲁棒性。但其聚类性能具有依赖于聚类中心初始值的缺点。针对模糊聚类算法对初始值敏感、易陷入局部最优的缺点,本文提出了一种基于粒子群优化的模糊聚类方法。该法根据聚类准则设计适应度函数,利用粒子群优化算法对聚类中心进行优化,后面仿真实验证明了算法的可行性和有效性。

4 案例分析

本文所讨论的电力变压器故障对象为高能放电、低能放电、过热和正常这四种变压器的状态。由于数据来源有限,算例分析采用的数据样本共 75 组,前 50 组为训练样本,后 25 为测试样本。50 组训练样本中,高能放电和过热故障样本各 15 组,低能放电故障和正常样本各 10 组。25 组测试样本中,过热故障样本 10 组,高能放电、低能放电故障和正常样本各 5 组。

4.1 变压器故障识别的特征提取

故障诊断的数据基础是通过色谱分析得到的变压器油中各种特征气体 (H₂、CH₄、C₂H₆、C₂H₄、C₂H₂) 的含量信息,这些气体的含量信息反映了电力变压器的高能放电、低能放电、过热和正常这四种状态。但是,每组测试数据所包括的上述 5 种特征气体含量信息都是指绝对含量,这些数据都是绝对值数据,为了更好地进行故障诊断识别,对这些绝对值数据进行数据处理,产生识别需要的特征量,包括:

(1)5 种特征气体的相对含量。它是通过各组测试数据中各种特征气体的含量数据除以本组含量最大的气体含量数据,共 5 个特征量,记为 y(1)、y(2)、y(3)、y(4)和 y(5);

$$y(k) = p(k) / \max_{i=1}^5 p(i) \quad (10)$$

其中 k=1,2,..., 5, p(k)为 5 种特征气体的绝对含量;

(2)反映各组测试中最大气体含量数据之间横向比较的一个特征量。由式(10)得到

$$y(6) = \lg 10 \left(\max_{i=1}^5 p(i) \right) \quad (11)$$

6 个特征量组成特征向量 y [y(1), y(2), y(3),

y(4), y(5), y(6)]^T, 它反映了识别变压器高能放电、低能放电、过热和正常这 4 种状态需要的各种特征。

4.2 最优聚类中心的计算

利用 50 组训练样本,根据目标函数(7)利用 PSO 算法进行最优聚类中心的计算。在本算例中,根据经验粒子群算法参数设置如下:进化代数设为 140,种群规模设为 60,粒子最大速度为 0.035,速度惯性权重 w 取 0~1 的随机数,认知学习因子 c₁ 和社会学习因子为 c₂ 都取 2。训练样本各类故障的特征气体含量典型数据如表 1 所示:

表 1 各类故障的特征气体含量典型数据

变压器状态	特征气体含量(H ₂ , CH ₄ , C ₂ H ₆ , C ₂ H ₄ , C ₂ H ₂)
高能放电	260 130 29 84 92
低能放电	550 53 34 20 0
过热故障	228 380 82 1012 19
正常状态	10 4 3 33 6

利用 PSO 算法得 4 种故障状态的最优聚类中心

Center_{高能放电}=[0.831 0.282 0.104 0.275 0.722 2.162];

Center_{过热故障}=[0.213 0.632 0.167 1.090 0.047 2.952];

Center_{低能放电}=[1.121 0.150 0.032 0.053 0.024 2.557];

Center_{正常状态}=[0.822 0.392 0.207 0.605 0.219 0.865];

需要注意的是,最优聚类中心是在特征向量空间。

4.3 故障识别测试

利用 PSO 算法得到最优聚类中心,对 25 组测试样本进行故障识别测试。其特征气体含量数据如表 2 所示:

表 2 25 组测试样本的 5 种特征气体含量数据

测试样本	气体含量(H ₂ CH ₄ C ₂ H ₆ C ₂ H ₄ C ₂ H ₂)
测试样本 1(高能放电)	23 12 12 10 61
测试样本 2(高能放电)	42 62 5 63 73
测试样本 3(高能放电)	162 35 5.6 30 44

测试样本 4(高能放电)	334	39.9	47.8	5.35	247.6
测试样本 5(高能放电)	200	48	14	117	131
测试样本 6(过热故障)	336	419	105	1047	21
测试样本 7(过热故障)	5	217	69	523	6
测试样本 8(过热故障)	766	993	116	665	4
测试样本 9(过热故障)	16	237	92	470	0
测试样本 10(过热故障)	117	357	92	468	4
测试样本 11(过热故障)	80	153	42	276	18
测试样本 12(过热故障)	15	125	29	574	7
测试样本 13(过热故障)	8	631	254	2020	39
测试样本 14(过热故障)	73	520	140	1200	6
测试样本 15(过热故障)	42	97	157	600	0
测试样本 16(低能放电)	565	93	34	47	0
测试样本 17(低能放电)	980	73	58	12	0
测试样本 18(低能放电)	160	90	27	17	5
测试样本 19(低能放电)	150	53	34	20	0
测试样本 20(低能放电)	35	25	0	23	22
测试样本 21(正常)	5.8	2.7	1.8	0.8	0
测试样本 22(正常)	14.7	3.8	10.5	2.7	0.2
测试样本 23(正常)	6.7	10	11	71	3.9
测试样本 24(正常)	5.8	2.7	1.8	0.8	0
测试样本 25(正常)	0.33	0.26	0.04	0.27	0

根据 4 类故障对应的最优聚类中心,由(8)可计算得到 25 组测试样本到各中心的距离,然后根据式(9)计算得到测试样本与 4 类故障(分别为高能放电、过热、低能放电和正常)之间隶属度,结果数据如表 3 所示:

表 3 25 组测试样本的 5 种特征气体含量原始数据

测试样本	气体含量(H ₂ CH ₄ C ₂ H ₆ C ₂ H ₄ C ₂ H ₂)				
测试样本 1	0.6342	0.0815	0.1320	0.1523	
测试样本 2	0.4740	0.1693	0.1394	0.2173	
测试样本 3	0.4537	0.0599	0.4016	0.0848	
测试样本 4	0.6144	0.0551	0.2855	0.0450	
测试样本 5	0.7626	0.0590	0.1288	0.0496	
测试样本 6	0.0390	0.9065	0.0390	0.0155	
测试样本 7	0.0651	0.8502	0.0539	0.0307	
测试样本 8	0.1796	0.5364	0.2159	0.0681	
测试样本 9	0.0594	0.8633	0.0483	0.0289	

测试样本 10	0.0517	0.8813	0.0423	0.0246
测试样本 11	0.1439	0.6870	0.1012	0.0678
测试样本 12	0.1010	0.7656	0.0874	0.0460
测试样本 13	0.0768	0.8110	0.0783	0.0338
测试样本 14	0.0339	0.9186	0.0326	0.0149
测试样本 15	0.1069	0.7501	0.0943	0.0487
测试样本 16	0.0677	0.0338	0.8824	0.0161
测试样本 17	0.1220	0.0781	0.7674	0.0325
测试样本 18	0.3007	0.0862	0.5262	0.0869
测试样本 19	0.2696	0.0716	0.5801	0.0786
测试样本 20	0.3874	0.0925	0.1560	0.3642
测试样本 21	0.0981	0.0401	0.0741	0.7877
测试样本 22	0.1916	0.0719	0.1470	0.5896
测试样本 23	0.2748	0.3019	0.1674	0.2559
测试样本 24	0.0981	0.0401	0.0741	0.7877
测试样本 25	0.1594	0.1031	0.1238	0.6137

由表 3 分析可发现,测试样本 1~5 与高能放电故障的隶属度最高,可诊断为变压器高能放电;测试样本 6~15 与过热故障的隶属度最高,为变压器过热故障;测试样本 16~19 与低能放电故障的隶属度最高,出为变压器低能放电;测试样本 21、22、24、25 与正常状态的隶属度最高,为变压器正常。同样,样本 20 可诊断为变压器高能放电,样本 23 可诊断为过热故障。25 个测试样本中 23 个正确诊断识别,2 个错误识别(测试样本 20, 23),识别正确率为 92%。在 23 个正确识别样本中,大部分样本(21 个)与对应类型故障的隶属度在 0.5 以上,而与其它对应类型故障的隶属度都较小,即能较可靠识别。2 个误识别样本 20、23,四种故障类型中的最高隶属度小于 0.4,且与其它故障类型的隶属度较接近,在此情况下,为了提高诊断正确率,可采用变压器故障诊断的其它方法(如 IEC 三比值法)进行复诊。

通过调查资料得到如下结果,国家标准改良三比值法诊断的准确率为 90%,BP 神经网络诊断正确率约为 79%,小波神经网络的准确率可达到 90%左右,IEC 三比值法根据实际经验准确率约为 70%,专家系统效果与专家经验关联较大,正确率约为 89%,粗糙集诊断方法正确率约 81%,灰色理论故障的正确率约为

88%。比较起来,该模糊粒子群算法的变压器故障识别方法计算简单,识别正确率为92%,在普通PC机上整个训练和测试花费的时间约30秒左右,能适用于变压器故障的在线识别。

5 结论

本文分析了一种基于模糊粒子群算法的变压器故障识别方法。首先对基于油中溶解气体分析得到五种关键气体含量数据进行特殊预处理后,得到六个关键特征。然后给出了一个模糊聚类目标函数,在此基础上,根据已有的故障样本利用粒子群算法得到各类故障的最优聚类中心,计算出各测试样本到各聚类中心的距离,及相应隶属度,最后诊断识别出被测样本故障类型。粒子群优化算法是有效随机优化算法,能很好地获取识别需要的最优聚类中心。利用模糊粒子群算法实现变压器故障识别,原理清晰、算法简单,效果较好,测试结果显示,该法能有效诊断识别出变压器高能放电、过热、低能放电和正常状态,总体精度可达92%,是变压器故障识别的一种很好选择。

参考文献

- 1 孙才新,郭俊峰等.基于行为的变压器油色谱分析模糊诊断专家系统研究.电工技术学报,2001,16(3):49-53.
- 2 徐文,王大忠.人工神经网络在变压器特征气体法故障诊断中的应用.高电压技术,1996,22(2):27-30.
- 3 彭宁云,文习山,舒翔.模糊神经网络在变压器故障诊断中的应用.高电压技术,2004,33(5):14-17.
- 4 潘翀,陈伟根.基于遗传算法进化小波神经网络的电力变压器故障诊断.电力系统自动化,2007,31(13):88-92.
- 5 束洪春,孙向飞,司大军.电力变压器故障诊断专家系统知识库建立和维护的粗糙集方法.中国电机工程学报,2002,22(2):31-35.
- 6 吕千云,程浩忠.基于改进灰色关联分析的变压器故障识别.中国电机工程学报,2004.10,24(10):121-126.
- 7 LV GY, CHENG HZ. Fault Diagnosis of Power Transformer Based on Multi-layer SVM Classifier. Electric Power Systems Research, 2005,74(1):1-8.
- 8 Kennedy J, Eberhart R. Particle Swarm Optimization. Proceedings of IEEE International Conference on Neural Networks, 1995,(4):1942-1948.
- 9 沈茂亚,丁晓群.自适应免疫粒子群算法在动态无功优化中应用.电力自动化设备,2007,27(1):31-35.
- 10 马豫超,侯志俭,蒋传文等.基于粒子群算法求解电力市场发电商最优供给函数模型.电力系统自动化,2006,30(2):45-50.