

利用改进遗传算法与 LS-SVM 进行变压器故障诊断

张凯¹, 黄华平², 杨海涛³, 谢庆²

(1. 石家庄供电公司, 河北省 石家庄市 050000; 2. 电力系统保护与动态安全监控教育部重点实验室(华北电力大学), 河北省 保定市 071003; 3. 广安电业局调通中心, 四川省 广安市 638000)

A Transformer Fault Diagnosis Method Integrating Improved Genetic Algorithm With Least Square Support Vector Machine

ZHANG Kai¹, HUANG Hua-ping², YANG Hai-tao³, XIE Qing²

(1. Shijiazhuang Power Supply Corporation, Shijiazhuang 050000, Hebei Province, China;

2. Key Laboratory of Power System Protection and Dynamic Security Monitoring and Control

(North China Electric Power University), Ministry of Education, Baoding 071003, Hebei Province, China;

3. Communication and Dispatching Center of Guang'an Electric Power Bureau, Guang'an 638000, Sichuan Province, China)

ABSTRACT: Least square support vector machines (LS-SVM) can solve small sample nonlinear multi-classification problem well, so it is applicable to the power transformer fault diagnosis by dissolved gas analysis (DGA), however the selection of parameters c and σ^2 greatly impacts the diagnosis result, thus the optimized selection of these parameters is necessary. In this paper, the improved genetic algorithm (IGA) is applied to the optimized selection of c and σ^2 . The initial population of IGA is randomly generated by coding mechanism, in this way, the search space can be quickly expanded and the diversity of individuals in the populations can be stabilized, thus both global search ability and convergence speed can be effectively improved. The LS-SVM optimized by IGA is applied to fault diagnosis by multi-sets of chromatographic data of transformer oil and the result shows that optimizing LS-SVM by IGA, the optimized selection of c and σ^2 can be implemented effectively and the accuracy of power transformer fault diagnosis can be improved.

KEY WORDS: power transformer; fault diagnosis; improved genetic algorithm (IGA); least square support vector machines (LS-SVM); dissolved gas analysis (DGA); parameter optimization

摘要: 最小二乘支持向量机(least square support vector machines, LS-SVM)能较好地解决小样本、非线性数据特征的多分类问题,适用于电力变压器油色谱故障诊断,但参数

c 与 σ^2 的选取对诊断结果影响较大,因此有必要对其进行优化选择。文中利用改进遗传算法(improved genetic algorithm, IGA)对 c 与 σ^2 参数进行寻优。IGA采用了编码机制随机产生初始种群,这样可快速扩大搜索空间,稳定群体中个体多样性,有效提高全局搜索能力和收敛速度。文中采用IGA优化后的LS-SVM对多组变压器油色谱数据进行故障诊断分析。结果表明,IGA可以有效实现对LS-SVM算法中 c 与 σ^2 的优化选取,提高变压器故障诊断的准确率。

关键词: 变压器; 故障诊断; 改进遗传算法; 最小二乘支持向量机; 溶解气体分析; 参数优化

0 引言

电力变压器是电力系统的核心设备,其故障诊断一直以来受到了学术界和工程界的广泛关注^[1-5]。油中溶解气体分析(dissolved gas analysis, DGA)被公认为是一种探测变压器初期故障和进行绝缘寿命评估的有效手段^[6-8]。它为间接了解变压器内部的一般隐患提供了重要依据。

近年来利用DGA数据诊断变压器不同故障的方法很多,最基本的是采用IEC/IEEE推荐的三比值法。另外还有多种智能技术,如模糊推理^[9]、人工神经网络^[10]、物元理论^[11]等已被引入变压器的油色谱故障诊断中。但由于实际所获取的变压器油色谱故障样本较少,模糊推理在故障数据样本的分析计算上会出现失真现象,不同气体成分对故障的灵敏度存在较大差异。人工神经网络诊断方法由于隐层单元数目难以确定,网络最终权值受初值影响较

基金项目: 长江学者和创新团队发展计划资助项目(IRT0515)。

Program Supported by Changjiang Scholars and Innovative Research Team in University (IRT0515).

大, 准确率不高; 物元理论在应用中对样本的一致性和样本量的要求严格, 对小样本数据仍存在一定不足。

支持向量机是在统计学理论上发展起来的一种新的机器学习方法。支持向量机^[12-15]在变压器的故障诊断中, 利用其置信范围最小化以控制诊断中风险问题; 并利用分层决策解决大类别分类问题的同时, 将 K-邻近搜索聚类应用于分层决策, 修正支持向量分类器参数难于选择带来的误差。该模型可有效地求取变压器油中气体特征空间中的最优解, 提高了故障诊断的正确性。最小二乘支持向量机(least square support vector machines, LS-SVM)^[14-15]是标准支持向量机的一种扩展, 优化指标采用平方项, 并用等式约束代替标准支持向量机的不等式约束, 即将二次规划问题转化为线性方程组求解, 降低了计算的复杂性, 提高了求解速度。

LS-SVM 优化模型参数的选取对算法性能有很大影响。参数选择的恰当与否会对样本数据的惩罚和训练误差带来很大的影响, 因此很有必要对其进行优化。文献[14]使用遗传支持向量机来解决电力变压器 DGA 故障问题, 具有一定的现实意义。但是算法在参数寻优过程中会出现因遗传优秀个体的破坏而导致优化过程不完全收敛。因此本文利用改进遗传算法(improved genetic algorithm, IGA)进行 LS-SVM 参数的选取。IGA 应用编码机制随机产生初始种群, 快速扩大搜索空间, 实现对 LS-SVM 参数的寻优。

1 改进遗传算法

1.1 遗传算法

遗传算法是一种基于自然选择和自然遗传的全局优化算法。它可以有效地利用已有的信息来处理搜索那些有希望改善解质量的串。它对算法所产生的每个染色体进行评价, 并基于适应度值来改变染色体, 使适应性好的染色体比适应性差的染色体有更多繁殖机会。这样群体中的个体一代一代优化, 并逐渐逼近最优解^[16]。

构成基本遗传算法的要素主要有: 染色体编码、个体适应度评价、遗传算子(选择算子、交叉算子、变异算子)以及遗传参数设置等。

遗传算子主要有以下 3 种:

1) 选择算子。选择操作建立在对个体的适应度进行评价的基础上。选择操作的主要目的是避免

基因缺失、提高全局收敛性和计算效率。

2) 交叉算子。随机从中间群体中抽取 2 个个体, 并按照某种交叉策略使 2 个个体互相交换部分染色体码串, 从而形成 2 个新的个体。

3) 变异算子。通常按照一定的概率(一般比较小)改变染色体中某些基因的值。变异是遗传算法中保持物种多样性的一个重要途径。通过选择与变异能够进行生物进化。

1.2 改进遗传算法及其参数选取

1.2.1 简介

本文提出的改进遗传算从提高全局搜索能力和收敛速度出发, 采用了编码机制技术。为保持种群的多样性, 提出了随机地产生初始种群的方法。遗传操作中采用变异概率随适应度自动改变的自适应遗传算法, 达到了快速扩大搜索空间, 稳定群体中个体多样性的目的。在进化若干代后, 将较差解空间删除。在进化过程中以同样的群体大小只在较好解空间中进行繁殖, 加大了较好解空间的个体密度, 有效地保证了性能优良个体的产生, 使演化过程能跳出局部最优解的误区, 在形成新种群进程中避免了局部早熟现象, 提高了全局最优解的搜索能力和收敛速度。

1.2.2 选择算子

本文选择运算使用比例选择算子, 种群个体数为 N , 个体 i 的适应度为 f_i , 则个体被选取的概率为

$$P_i = f_i / \sum_{k=1}^N f_k。$$

1.2.3 交叉算子

利用单点交叉算子对种群进行交叉运算。单点交叉是指在个体编码中只随机设置 1 个交叉点, 然后在该点相互交换 2 个配对个体的部分染色体。图 1 为单点交叉运算示意图。



图 1 单点交叉运算示意图

Fig. 1 Diagram of one point crossover operation schematic

1.2.4 变异算子

采用如下式的自适应变异概率算子 P_m 。当种群中各个体适应度趋于一致或者趋于局部最优时, 使 P_m 增加; 当群体适应度比较分散时, 使该解得以保护, 进入下一代。而低于平均适应度值的个体, 相对应于较高的 P_m 将被淘汰掉。所以自适应的 P_m 能

够提供相对于某个解的最佳 P_m ，在保持群体多样性的同时，保证了遗传算法的收敛性^[16]。

$$P_m = \begin{cases} P_{m1} - \frac{(P_{m1} - P_{m2})(f_{\max} - f)}{f_{\max} - f_{\text{avg}}}, & f \geq f_{\text{avg}} \\ P_{m1}, & f < f_{\text{avg}} \end{cases}$$

其中： f 为要变异的个体适应度值； f_{avg} 为每代种群的平均适应度值； f_{\max} 为群体中最大的适应度值。

本文中 $P_{m1}=0.1$ ， $P_{m2}=0.01$ 。

1.2.5 最优保存策略

在改进遗传算法中，使父代中一定数量的优秀个体直接进入下一代，这样可防止复制、交叉或变异中的偶然因素破坏优秀个体，使优化过程在概率性上全局收敛。这是增强算法稳定性和收敛性的有效方法。一般认为，简单遗传算法不是全局收敛的，而保留最优个体的遗传算法是全局收敛的。

2 基于改进遗传算法的 LS-SVM

2.1 LS-SVM 多分类算法

$T_i = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ 为训练样本集，其中 $x_n \in \mathbf{R}^d$ ， $y_n \in \{1, 2, \dots, k\}$ ， $i = 1, 2, \dots, n$ 。在 LS-SVM 二分类的基础上，采用一对多分类算法，建立多类分类器^[12]。对 k ($k \geq 2$) 类分类问题，取所有 $y = i$ 和 $y \neq i$ 的训练样本，构造 k 个训练子集 T_{ii} 。在构造 k 个分类器中的第 m 分类器时 ($m \in y_i$)，将 m 类训练样本作第一类，类别号 $y_i^m = +1$ ，其余 $m-1$ 类作一类，类别号 $y_i^m = -1$ 。

对测试样本分类时，测试样本先输入给分类器 1，若判别函数输出为 1，则判断类别为第 1 类故障；否则自动输入给下一个分类器 2，若判别函数输出为 1，则判断类别为第 2 类，否则自动输入给下一个分类器 3；以此类推，直到分类器 k 。若判断函数输出为 1，则判断类别为第 k 类故障；若输出为 -1，测试样本不在 k 种类别之内，属于其他类别。多类分类流程如图 2 所示。

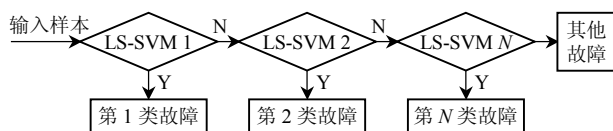


图 2 多类分类流程

Fig. 2 Flow chart of multi-classification

2.2 核函数及参数选择

LS-SVM 将优化问题转化为求解一线性方程组，使基于 LS-SVM 的分类模型具有很好的泛化能力。另外 LS-SVM 中的参数对算法性能有很大影

响。如 c 取得小，对样本数据的惩罚就小，使训练误差变大，算法泛化能力变好；反之，相应的 $\omega^T \omega / 2$ 的权重就小，算法的泛化能力变差。采用高斯核函数时， σ^2 的值太大或太小会对样本数据造成过学习或者欠学习的现象。本文利用改进遗传算法进行 LS-SVM 参数的选取，核函数选用高斯径向基核函数。

2.3 改进遗传 LS-SVM 流程

本文所用的改进遗传 LS-SVM 模型，采用 IGA 来优化 LS-SVM 中的参数，具体过程如下：

- 1) 系统初始化，包括参数和初始种群。
- 2) 计算初始目标函数值。
- 3) 判断是否达到结束条件，若达到，则输出结果；否则转步骤 4)。
- 4) 进行适应度值的计算、选择、交叉和变异。
- 5) 计算子代的目标函数值，并重插子代到种群，转入步骤 3)。

2.4 参数选择

LS-SVM 有 c 与 σ^2 2 个参数需要优化。建立一个具有 50 个个体，每个个体有 2 个变量的随机种群作为原始种群。同时为避免 LS-SVM 的过学习现象，将训练样本划分为正式训练和辅助训练样本，目标函数为正式训练样本集和辅助训练样本集的正判率之和。在遗传搜索的初始阶段，正式和辅助训练样本集的正判率会同时增加；随着搜索的进行，正式训练样本集正判率的增加会变慢；当正式训练样本集的正判率进一步提高，辅助训练样本集的正判率减小时，就要停止一串搜索^[9-10]。

本文采用真值编码方法可以在较大空间遗传搜索，改善 IGA 的计算复杂性，提高运算效率。

3 基于 IGA 与 LS-SVM 的变压器故障诊断

3.1 故障特征量的选择与故障分类

变压器绝缘材料热分解所产生的气体达 20 种左右。为有利于变压器内部故障诊断，选定必要的气体作为分析对象很重要。目前国内外所分析的气体对象不统一，根据国内当前情况同时考虑经济性及判断效果我们选择 5 种气体(即 H_2 、 CH_4 、 C_2H_6 、 C_2H_4 和 C_2H_2)作为研究对象。为降低各种气体之间由于量值差异造成的影响，处理中将 DGA 原始数据进行了归一化。

根据 IEC60599 规定，故障类型可分为 6 种：局部放电(PD)、低能放电(D₁)、高能放电(D₂)、低温过热(T₁)、中温过热(T₂)和高温过热(T₃)，再加上正

常状态(NC), 诊断状态总共有7种。

3.2 故障诊断流程

基于改进遗传 LS-SVM 的变压器故障诊断步骤如下:

- 1) 对各样本的5种气体浓度数据进行归一化处理, 形成样本集合。
- 2) 用IGA进行参数优化。
- 3) 将优化得到的参数带入LS-SVM的多分类算法, 得到相应的分类模型。
- 4) 利用建好的分类模型对样本进行诊断。

3.3 实例分析

本文所用的样本库是从各个资料上^[17-18]选取的有明确结论的137组样本, 选择其中的87组作为训练集, 其余50组作为验证集。变压器的各种状态在样本集的分布情况见表1。

表1 各种状态在样本集中的分布
Tab. 1 The distribution of variety state in samples

故障类型	NC	PD	D ₁	D ₂	T ₁	T ₂	T ₃
训练集	8	13	9	16	14	10	17
验证集	4	9	4	10	8	6	9

将 c 与 σ^2 按实数编码方式进行编码, 寻优区间分别为(0,200)、(0,50)。改进遗传算法的种群规模为30, 交叉概率 P_c 为0.9, 变异概率 P_m 为0.033, 进化代数为1000, 优化模型的正确率为91.7%。表2为一些具体的实际样本用改进遗传LS-SVM诊断的结果。可以看出, 改进遗传LS-SVM进行的变压器故障诊断的正确率是较高的。

表2 实际样本诊断结果
Tab. 2 Diagnosis results of the actual samples

序号	H ₂	CH ₄	C ₂ H ₆	C ₂ H ₄	C ₂ H ₂	实际故障类型	本文算法
1	56.0	61.00	75.00	32.00	31.00	PD	D ₁
2	33.1	26.20	6.20	5.30	0.27	NC	NC
3	176.2	205.90	47.70	75.60	69.70	D ₁	D ₁
4	70.2	69.50	28.90	241.20	10.87	T ₃	T ₃
5	162.0	34.90	5.60	30.10	44.00	D ₂	D ₂
6	345.0	112.25	27.40	51.50	58.75	D ₁	D ₁
7	172.9	334.10	172.00	812.50	37.70	T ₃	T ₃
8	2587.8	7.90	4.79	1.49	0.00	PD	PD
9	1677.0	653.90	81.70	1004.90	419.10	D ₂	D ₂

4 结论

最小二乘支持向量机能较好地解决小样本、非线性数据特征的多分类问题, 适用于电力变压器油色谱故障诊断。但 c 与 σ^2 参数的选取对诊断结果影响较大, 文中利用一种新的改进遗传算法对 c 与 σ^2 参数进行寻优。IGA算法采用了编码机制, 随机产

生初始种群, 这样可快速扩大搜索空间, 稳定群体中个体多样性, 有效提高全局搜索能力和收敛速度。变压器故障诊断实例分析结果表明, IGA可以有效实现对LS-SVM算法中 c 与 σ^2 的优化选取, 提高了变压器故障诊断的准确率。

参考文献

- [1] Castro A R G, Miranda V. Knowledge discovery in neural networks with application to transformer failure diagnosis[J]. IEEE Trans on Power Systems, 2005, 20(2): 717-724.
- [2] 钱政, 严璋, 罗承沐. 电力变压器绝缘故障综合诊断方法的研究[J]. 电网技术, 2002, 26(2): 32-36.
Qian Zheng, Yan Zhang, Luo Chengmu. A study on synthetic diagnosis for insulation fault of power transformers[J]. Power System Technology, 2002, 26(2): 32-36(in Chinese).
- [3] 胡文平, 尹项根, 张哲, 等. 变压器故障诊断与多传感器信息融合[J]. 高电压技术, 2002, 28(2): 32-36.
Hu Wenping, Yin Xianggen, Zhang Zhe, et al. Transformer faults diagnosis and multi-sensor information fusion[J]. High Voltage Engineering, 2002, 28(2): 32-36(in Chinese).
- [4] 张国云, 章兢. 基于模糊支持向量机的多级二叉树分类器的水轮机调速系统故障诊断[J]. 中国电机工程学报, 2005, 25(8): 100-104.
Zhang Guoyun, Zhang Jing. Fuzzy SVM-based multilevel binary tree classifier for fault diagnosis of hydroturbine speed regulating system[J]. Proceedings of the CSEE, 2005, 25(8): 100-104(in Chinese).
- [5] Hsu C W, Lin C J. A comparison of methods for multi class support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2): 415-425.
- [6] 李清泉, 王伟, 王晓龙. 利用DGA-NN诊断油浸式电力变压器故障[J]. 高电压技术, 2007, 33(8): 48-51.
Li Qingquan, Wang Wei, Wang Xiaolong. Fault diagnosis of oil-immersed power transformer by DGA-NN[J]. High Voltage Engineering, 2007, 33(8): 48-51(in Chinese).
- [7] 焦卫东, 钱苏翔, 杨世锡, 等. 基于非线性主分量分析: 多层感知器网络的变压器油中溶解气体数据分析[J]. 中国电机工程学报, 2007, 27(6): 72-76.
Jiao Weidong, Qian Suxiang, Yang Shixi, et al. NPCA-MLP network for analysis and diagnosis of DGA data from power transformer[J]. Proceedings of the CSEE, 2007, 27(6): 72-76(in Chinese).
- [8] 王南兰, 邱德润. 油中溶解气体分析的变压器故障诊断新方法[J]. 高电压技术, 2006, 32(6): 35-37.
Wang Nanlan, Qiu Derun. New method of power transformer fault diagnosis by dissolved gas-in-oil analysis[J]. High Voltage Engineering, 2006, 32(6): 35-37(in Chinese).
- [9] 王子建, 何俊佳, 尹小根. 基于改进模糊ISODATA算法的变压器故障诊断[J]. 高压电器, 2006, 42(1): 11-13, 17.
Wang Zijian, He Junjia, Yin Xiaogen. Fault diagnosis of transformer based on revised fuzzy ISODATA algorithm[J]. High Voltage Apparatus, 2006, 42(1): 11-13, 17(in Chinese).
- [10] 徐志钮, 律方成. Nguyen-Widrow法用于变压器油气分析故障诊断[J]. 高电压技术, 2006, 32(8): 46-48.
Xu Zhiniu, Lü Fangcheng. Application of Nguyen-Widrow method in transformer dissolved gas-in-oil analysis fault diagnosis[J]. High Voltage Engineering, 2006, 32(8): 46-48(in Chinese).
- [11] 黄文涛, 赵学增, 王伟杰, 等. 基于物元模型的电力变压器故障的可拓诊断方法[J]. 电力系统自动化, 2004, 28(13): 45-49.

- Huang Wentao, Zhao Xuezheng, Wang Weijie, et al. Extension diagnosis method of power transformer faults based on matter-element model[J]. Automation of Electric Power Systems, 2004, 28(13): 45-49(in Chinese).
- [12] 董明, 孟源源, 徐长响, 等. 基于支持向量机及油中溶解气体分析的大型电力变压器故障诊断模型研究[J]. 中国电机工程学报, 2003, 23(7): 88-92.
- Dong Ming, Meng Yuanyuan, Xu Changxiang, et al. Fault diagnosis model for power transformer based on support vector machine and dissolved gas analysis[J]. Proceedings of the CSEE, 2003, 23(7): 88-92(in Chinese).
- [13] 贾嵘, 徐其惠, 李辉, 等. 最小二乘支持向量机多分类法的变压器故障诊断[J]. 高电压技术, 2007, 33(6): 110-113.
- Jia Rong, Xu Qihui, Li Hui, et al. Fault diagnosis of transformer using multi-class least squares support vector machine[J]. High Voltage Engineering, 2007, 33(6): 110-113(in Chinese).
- [14] 肖燕彩, 陈秀海, 朱衡君. 遗传支持向量机在电力变压器故障诊断中的应用[J]. 上海交通大学学报, 2007, 41(11): 1878-1886.
- Xiao Yancai, Chen Xiuhai, Zhu Hengjun. The application of genetic support vector machine in power transformer fault diagnosis[J]. Journal of Shanghai Jiaotong University, 2007, 41(11): 1878-1886(in Chinese).
- [15] 邓宏贵, 曹建, 罗安. 一种新的遗传算法及其在变压器故障诊断中的应用[J]. 中南大学学报: 自然科学版, 2005, 36(3): 481-485.
- Deng Honggui, Cao Jian, Luo An. A novel genetic algorithm and its application to transformer fault diagnosis[J]. Journal of Central South University of Technology: Natural Science Edition, 2005, 36(3): 481-485(in Chinese).
- [16] 余楠. 一种改进的遗传算法及其在旅行商问题中的应用[J]. 电脑开发与应用, 2009, 22(1): 35-36.
- Yu Nan. An improved genetic algorithm and application for traveling salesman problem[J]. Computer Development & Applications, 2009, 22(1): 35-36(in Chinese).
- [17] 彭宁云. 基于 DGA 技术的变压器故障智能诊断系统研究[D]. 武汉: 武汉大学, 2004.
- [18] 肖燕彩, 陈秀海, 朱衡君. 基于最小二乘支持向量机的变压器油中气体浓度预测[J]. 电网技术, 2006, 30(11): 91-94.
- Xiao Yancai, Chen Xiuhai, Zhu Hengjun. Forecasting of gas concentration in power transformer oil based on least square support vector machine[J]. Proceedings of the CSEE, 2006, 30(11): 91-94(in Chinese).



张凯

收稿日期: 2009-05-26。

作者简介:

张凯(1980—), 男, 硕士, 主要从事变电运行技术管理工作, E-mail: zhangkai0603@163.com。

(责任编辑 王晔)