

基于 K - L 特征提取与支持向量机的油浸式 变压器故障诊断模型研究

方 飏

(四川省电力工业调整试验所, 四川 成都 610072)

摘 要: 针对变压器多故障问题, 提出了基于 Mercer 核函数的欧式距离查询策略算法, 并建立了基于 Karhunen - Loeve (K - L) 特征提取与支持向量机的变压器故障诊断模型, 利用 K - L 变换提取信号的特征值, 最后通过支持向量机学习算法完成对信号的选择与分类。通过实例应用表明: 所训练的 SVM 分类器较之直接任意选取训练样本作为训练集的传统方法具有更高的诊断率。

关键词: 油浸变压器; 故障诊断; 支持向量机; K - L 特征提取

Abstract: Aiming at the multi - fault problems of transformer, the Euclidian distance query algorithm based on Mercer kernel function is proposed and a model based on Karhunen - Loeve (K - L) feature extracting and support vector machine (SVM) is established for fault diagnosis of oil - immersed transformer, the eigenvalue of signal is extracted using K - L transform, and finally, the SVM learning algorithm is introduced to select and classify the training sample data. The result shows that the precision of the trained SVM classifier is better than that of the traditional method, and the reliability and effectiveness using the above mentioned method are satisfied in fault diagnosis.

Key words: oil - immersed transformer; fault diagnosis; support vector machine; K - L feature extracting

中图分类号: TM855 文献标志码: A 文章编号: 1003 - 6954(2012)06 - 0058 - 04

变压器是电力系统中的重要设备, 其运行状态直接影响到系统的安全性。油中溶解气体分析 (dissolved gas analysis, DGA) 是目前诊断充油电气设备故障的重要手段, 已作为变压器内部潜伏性故障的主要试验项目列于 DL/T 596 - 1996 变压器试验项目的首位^[1]。DGA 往往能够较准确、可靠地发现逐步发展的潜伏性故障, 防止由此引起重大事故。近年来的研究已将各种人工智能技术如神经网络、粗糙集、聚类算法、支持向量机等引入到变压器的故障诊断当中。

支持向量机算法^[2-3]是在统计学理论上发展起来的一种新的模式识别方法。它是由 Vapnik 等人于 1995 年提出的从统计学习理论发展出的一种模式识别方法, 是目前较为流行的适用于小样本训练的大边缘分类器。

K - L 变换是经典的传统方法, 适用于任何概率分布, 它能保留原样本中方差最大的数据分量, 所以 K - L 变换起了减小相关性、突出差异性的效果, 因此是在均方误差最小的意义下获得数据降维的最佳变换。

这里针对变压器的 5 种运行模式 (4 种故障模式与 1 种正常运行模式) 提出了建立基于支持向量机与 K - L 特征提取技术的变压器故障诊断模型。该模型的基本思想是在训练数据预选取阶段采用基于 Mercer 核函数的欧式距离的查询策略选取最能代表各类的样本数据, 从而能够更加准确地构造 SVM 分类器, 减少训练样本数和缩短训练时间。同时利用 K - L 变换的特征提取技术在系统分类精度不降低的情况下将样本数据由六维降到三维, 提取了样本数据集的特征值, 同时通过三维图表示了降为三维的样本数据, 从而为进一步提高分类正确率奠定了基础。

1 支持向量机分类原理

1.1 线性情况

当训练样本集线性可分时, 分类超平面的描述为

$$\omega \cdot x + b = 0 \quad (1)$$

式 (1) 中, 向量 ω 为分类超平面的权系数; b 是分类

阈值。最优分类超平面可通过解下面的凸二次优化问题获得。

$$\min \phi(w) = \|w\|^2 / 2 \quad (2)$$

$$y_i(w \cdot x_i) + b - 1 \geq 0, i = 1, 2, \dots, n \quad (3)$$

通过求解,得到最优分类超平面的分类判别函数为

$$f(x) = \text{sgn} \left[\sum_{i=1}^n \alpha_i y_i (x_i \cdot x) + b \right] \quad (4)$$

其中, α_i 为拉氏乘子,拉氏乘子不为 0 的解向量为支持向量。

1.2 非线性情况

对于非线性问题,SVMs 通过选择适当的非线性变换,将输入空间 X 中的训练样本映射到某个高维特征空间 F ,使得在目标高维空间中这些样本线性可分。根据泛函的有关理论,若核函数 $K(x, x_i)$ 满足 Mercer 条件,则对应某一变换空间中的内积 $\langle \phi(x_i) \cdot \phi(x) \rangle$,函数 $\phi: x \rightarrow F$ 是一个从非线性输入空间 X 到高维特征空间 F 的映射,所以求映射 $\phi: x \rightarrow F$ 只要知道如何由输入 x, x_i 计算内积 $\langle \phi(x_i) \cdot \phi(x) \rangle$ 即可,由

$$K(x_i, x) = \phi(x_i) \cdot \phi(x) \quad (5)$$

将式(4)重写,即可得到对应高维空间的分类函数为

$$f(x) = \text{sgn} \left[\sum_{i=1}^n \alpha_i y_i K(x_i \cdot x) + b \right] \quad (6)$$

2 核距离与 K-L 变换

根据 SVM 分类原理,经向量化处理的变压器 DGA 数据可直接作为训练样本数据,但此方法使得 SVM 训练算法速度比较慢,这是因为训练样本的数量决定了二次规划问题目标函数中矩阵的维数,使得求解规划问题的速度与维数呈指数增长。为了提高训练速度,减少学习样本数,缩短训练时间,则采取主动选择对于确立分类器最重要的新样本进行训练,来进一步设计新的分类器,从而使得可以用尽可能少的标注样本数来实现较高的分类精度。

支持向量机学习算法包括两个独立的部分 c, f , c 是一个 SVM 多类分类器, f 是一个查询函数,决定应从候选集中选择哪些样本进行训练。由于检测变压器运行模式是一个多分类的模式识别问题,因此在这里 SVM 分类器采用的是 1-a-r 方法。这里通过引 λ 类质心和基于 Mercer 核函数的欧式距离等概念来定义查询函数 f 。其查询策略是通过调节基于 Mercer 核函数欧式距离的选择范围来选

择最能代表各类的样本数据。这种策略最大程度地缩减了训练样本数据,降低了算法的运算量。

2.1 基于 Mercer 核函数的欧式距离

Mercer 核函数的原理是将输入样本数据空间非线性映射到新的特征空间。

$$\phi: R^N \rightarrow H, x \mapsto \phi(x) \quad (7)$$

样本数据空间中的欧式距离在新的特征空间 H 里可以表示为

$$\begin{aligned} d_H(x, y) &= \sqrt{\|\phi(x) - \phi(y)\|^2} \\ &= \sqrt{\phi(x) \cdot \phi(x) - 2\phi(x) \cdot \phi(y) + \phi(y) \cdot \phi(y)} \end{aligned} \quad (8)$$

根据 Mercer 核映射 ϕ 的性质,输入空间的点积在新的特征空间 H 中可以用 Mercer 核函数来表示,即

$$K(x, y) = (\phi(x) \cdot \phi(y)) \quad (9)$$

由式(7)和式(8)可以表示为

$$d_H(x, y) = \sqrt{K(x, x) - 2K(x, y) + K(y, y)} \quad (10)$$

2.2 基于 K-L 变换的特征提取

K-L 特征提取^[8]是通过映射(或变换)的方法把高维特征向量变换为低维特征向量。通过特征提取获得的特征是原始特征集的某种组合,即 $A: Y \rightarrow X$,可见,新的特征中包含了原有全体特征的信息。特征提取的关键问题是求出最佳变换矩阵,使得变换后的 m 维模式空间中,类别可分性准则值最大。

设 x 为 n 维随机向量, x 可以用 n 个正交基向量的加权和来表示为

$$x = \sum_{i=1}^n \alpha_i \varphi_i \quad (11)$$

式中, α_i 为加权系数; φ_i 为正交基向量,满足

$$\varphi_i^T \varphi_j = \begin{cases} 1 & \text{当 } i=j \\ 0 & \text{当 } i \neq j \end{cases} \quad (12)$$

从 n 个特征向量中取出 m 个组成变换矩阵 A , 即 $A = (\varphi_1, \varphi_2, \dots, \varphi_m)$ ($m < n$), 这时, A 是一个 $n \times m$ 维矩阵, x 为 n 维向量, 经过 $A^T x$ 变换, 得到降维为 m 的新向量。通过选取 m 个特征向量构成变换矩阵 A , 使降维的新向量在最小均方误差准则下接近原来的向量 x 。

对于式(12)现在只取 m 项,对略去的项用预先选定的常数 b_j 来代替,这时对 x 的估计值为

$$\hat{x} = \sum_{j=1}^m \alpha_j \varphi_j + \sum_{j=m+1}^n b_j \varphi_j \quad (13)$$

由此产生的误差为

$$\Delta x = x - \hat{x} = \sum_{j=m+1}^n (\alpha_j - b_j) \varphi_j \quad (14)$$

均方误差为

$$\begin{aligned} \bar{\varepsilon}^2 &= \sum_{j=m+1}^n E [\alpha_j^2] = \sum_{j=m+1}^n E [\varphi_j^T x] (\varphi_j^T x)^T \\ &= \sum_{j=m+1}^n E [\varphi_j^T x x^T \varphi_j] \\ &= \sum_{j=m+1}^n \varphi_j^T R \varphi_j = \sum_{j=m+1}^n \lambda_j \quad (15) \end{aligned}$$

式中, λ_j 是 x 的自相关矩阵 R 的第 j 个特征值; φ_j 是与 λ_j 对应的特征向量。显然所选的 λ_j 值越小, 均方误差也越小。

综上所述, 基于 K-L 变换的特征提取的步骤如下^[4]: (1) 平移坐标系, 将模式总体的均值向量作为新坐标系的原点; (2) 求出自相关矩阵 R ; (3) 求出 R 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 及其对应的特征向量 $\varphi_1, \varphi_2, \dots, \varphi_n$; (4) 将特征值从大到小排序, 如 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq \dots \geq \lambda_n$, 取前 m 个大的特征值所对应的特征向量构成变换矩阵 $A = (\varphi_1, \varphi_2, \dots, \varphi_m)$; (5) 将 n 维的原向量变换成 m 维的新向量 $y = A^T x$ 。

利用 K-L 变换方法完成了对输入样本的特征值提取, 降低了输入样本维数, 使 SVM 能够快速分类。

3 故障诊断流程

(1) 计算整个训练数据集的平均值 $v = \frac{1}{N} \sum_{X \in U} X$ 作为整个数据集的质心。

(2) 计算各类的数据 $X_j^i (i = 1, \dots, 5)$ 与这个质心的基于 Mercer 核函数欧式距离 $d_{ij} = \|\phi(X_j^i) - \phi(v)\|$ 。

(3) 对得到的每类基于核函数的欧拉距离按一定序列排序。如低温过热故障所得的基于核函数的欧拉距离采用降序排列; 而高温过热故障所得的基于核函数的欧拉距离采用升序排列。

(4) 根据各类的欧式距离 d_{ij} 分别为每一类设

表 1 两种方法诊断的结果对比

运行模式	测试数据	传统方法	检测精度 /%	所提方法	检测精度 /%
正常	30	30/28	100/93	30/29	100/97
低能放电	10	8/9	80/90	9/8	90/80
低温过热	10	6/4	60/40	7/5	70/50
高温过热	10	9/4	90/50	10/10	100/100
高能放电	10	10/10	100/100	10/9	100/90
总计	70	63/55	90.00/78.57	66/61	94.28/87.14

注: a/b , a 表示基于最优参数的故障诊断率; b 表示基于经验选取的参数的故障诊断率。

定一个上限阈值 α_i^+ 和一个下限阈值 α_i^- , 进而根据各类的阈值分别选取相应类的训练集。

(5) 得到的 5 类数据集即可作为训练集。

(6) 通过 K-L 变换对所选取的训练集进行特征提取, 将其降为三维以降低运算量。

(7) 将测试数据代入训练好的分类器中进行试验, 如果测试误差较大, 返回步骤(3)。

(8) 达到精度即可作为故障分类器。

4 变压器故障诊断及分析

根据模拟试验和大量的现场试验, 将故障类型分为低能放电、低温过热、高温过热和高能放电。与正常运行情况共组成 5 类运行模式。既充分考虑到 DGA 数据自身的可分性, 又可使故障的诊断结论可从故障现象明确确定。同时每类故障样本分别由 6 种气体值(H_2 、 CH_4 、 C_2H_6 、 C_2H_4 、 C_2H_2 和 CO) 组成。所收集到故障明确且测量、诊断结果无误的变压器油中溶解气体的含量数据共 161 个样本。各种类型故障样本数目比例基本符合对充油电气设备故障类型的统计结果, 较好地反映了变压器的故障模式。在计算基于 Mercer 核函数的欧式距离时选择了线性核函数($K(x, y) = (x \cdot y + 1)^{p_1}, p_1 = 2$); 选择高斯核函数($K(x, y) = \exp(-\|x - y\|^2 / 2p_1^2)$)作为 SVM 训练的核函数。通过使用网格搜索法确定最优参数(核宽 $P_1 = 1$, 惩罚系数 $C = 1000$), 结束训练条件 $\text{eps} = 0.001$ 。

通过所提学习方法选取 29 例正常模式、12 例低能放电模式、8 例低温过热模式、9 例高温过热模式和 20 例高能放电模式训练样本组成训练集。同时与通过传统学习方法随机选择等量的样本数据作为训练集进行比较, 测试数据如表 1 所示。

图 1 是训练集的相对位置关系图。图中实线表

示正常运行模式,虚线表示高能放电模式,星号表示低能放电模式,加号表示高温过热模式,菱形表示低温过热模式。从图中可知,所得的训练集的三维图正确地反映了油浸式变压器实际的运行模式。高能放电模式样本数据与低能放电模式样本数据位置紧密但较易区分;而高温过热模式样本数据与低温过热模式样本数据混淆严重,由此推断,这两种故障模式较难区分。表 1 显示了两种方法训练的分类器对测试数据正确诊断率的仿真结果。同时为检验参数寻优对分类结果的影响,对基于经验选取参数($p_1 = 0.1, C = 100$)的多类 SVM 模型与基于最优参数的多类 SVM 模型的分类结果进行了对比。

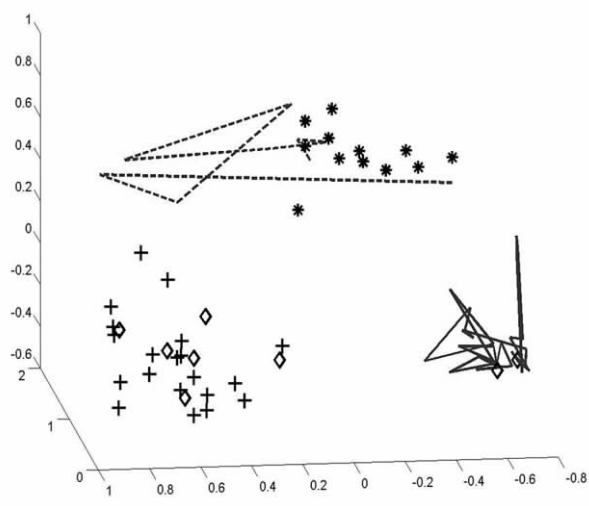


图 1 训练集三维图

从表 1 对测试样本的检测率可知,通过所提方法获得的训练集构造的分类器比传统方法构造的分类器更精确。因此故障诊断正确率更高。同时可知通过寻优所得的基于最优参数的故障诊断模型比基于经验的参数模型具有更高的检测正确率。

通过对变压器油中气体容量($\times 10^{-6}$)的现场测试,具体值为: $\phi(H_2) = 127$, $\phi(CH_4) = 107$, $\phi(C_2H_2) = 244$, $\phi(C_2H_4) = 154$, $\phi(C_2H_6) = 11$, $\phi(CO) = 174$ 。通过所提方法建立的诊断模型,诊断结果为高能放电。通过实际的检测,故障是由于次级线圈放电而造成整个线圈的损坏。

5 结 论

提出了将支持向量机学习算法与基于 K-L 特征提取技术应用于变压器故障诊断模型的新算法。通过主动选择各类数据集中能正确代表各类相对关

系的样本数据来训练分类器,因此提出的主动学习方法能够在缩减训练样本和计算量的情况下仍能达到传统方法的检测精度。同时引入了基于 K-L 变换的特征提取技术,使能够实现样本数据从六维降到三维,并通过 Matlab 作图将其显示在三维图中。实验表明,采用所提方法相对于传统方法具有更高的故障诊断率。在仿真实验中,分类的准确度与核函数类型的选择和相应参数的选取密切相关,因此下一步工作是从 SVM 的分类机理出发对模型的参数选择与精确度的提高可以展开进一步的研究。

参考文献

- [1] DL/T 596 - 1996, 电力设备预防性试验规程 [S].
- [2] Warmuth M K, Liao J, Ratsch G et al. Active Learning with Support Vector Machines in the Drug Discovery Process [J]. Journal of Chemical Information and Computer Sciences, 2003, 43(2): 667 - 673.
- [3] Hsu C W, Chang C C, Lin C J. A Practical Guide to Support Vector Classification [Z]. 2003 - 07. <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.
- [4] Wang Liguang, Zhang Ye, GU Yanfeng. The Research of Simplification of Support of Structure of Multi - class Classifier Vector Machine [J]. Journal of Image and Graphics, 2005, 10(5): 572 - 574.
- [5] Cortes C, Vapnik V. Support Vector Networks [J]. Machine Learning, 1995, 20(1): 273 - 297.
- [6] 肖建华. 智能模式识别方法 [M]. 广州: 华南理工大学出版社, 2006.
- [7] Wilson D R, Tony R M. Improved Heterogeneous Distance Functions [J]. Journal of Artificial Intelligence Research, 1997, 6(1): 1 - 34.
- [8] 段丹青, 陈松乔, 杨卫平. 基于 SVM 主动学习的入侵检测系统 [J]. 计算机工程, 2007, 33(1): 153 - 155.
- [9] Vapnik V. An Overview of Statistical Learning Theory [J]. IEEE Transactions on Neural Networks, 1999, 10(5): 988 - 999.
- [10] Cui Jiang, Wang You - ren. Analog Circuit Faults Diagnosis Based on Clustering Preprocess and SVM [J]. Computer Applications, 2006, 26(8): 1977 - 1979.

作者简介:

方 飏(1973), 男, 工程师, 四川省电力工业调整试验所电气室主任, 长期从事高电压技术相关试验工作及研究。

(收稿日期: 2012 - 07 - 25)