

水电机组中改进支持向量机算法的故障检测应用研究

水电机组故障数据通常为小样本的精密数据，而传统的支持向量机故障检测方法通常对大数量样本进行分类分析才有效，对小数据故障样本的分析存在局限性。对支持向量机方法进行了改进，提出最小二乘支持向量机方法并应用到水电机组故障检测中，对改进过程进行了详细的介绍，塑造故障检测多元分类模型，实现水电机组故障的准确检测。实验对采集的13个水电机组故障样本类型进行仿真分析，结果说明，优化支持向量机具有较高的小样本数据学习和预测性能，将其用于水电机组故障检测是有效、可行的。

肖华中/广西水利电力职业技术学院

黄华标/广西方元电力股份有限公司桥巩水电站



肖华中/副教授

关键词/Keywords

支持向量机·
小样本·
水电机组·
故障检测·

故障诊断技术对于确保工业系统安全运行具有重要意义^[1-2]，但是，在实际问题中，处理故障诊断时往往只能利用有限的样本数据集，其样本数目是非常有限的。水电机组故障数据通常为小样本的精密数据，而传统支持向量机通常对大数量样本数据进行分类分析，存在局限性^[3-5]，实际检测水电机组故障数据存在明显偏差，影响了方法在水电机组故障检测过程中的实际使用效果^[6-8]。本文提出了适用于小样本模式识别的优化支持向量机方法，也就是最小二乘支持向量机，分析了优化最小二乘支持向量机参数选择的方法，塑造水电机组故障检测多元分类模型，提高支持向量机模型的故障分类检测质量，实现水电机组故障的准确检测。实验对采集的13个水电机组故障样本类型进行仿真分析，结果说明，优化支持向量机具有较高的小

样本数据学习和预测性能，将其用于水电机组故障检测是有效、可行的。

水电机组故障检测原理

1. 水电机组故障检测相关理论

一个水电机组故障检测系统往往有两个组成部分：传感器、安全隐患分析程序。传感器用来获取转子表面承受的压力、尾水管表面裂缝程度等相关参数；安全隐患分析程序则用于计算安全隐患区域的空间位置，并叠加处理各安全隐患局部区域，最终实现水电机组故障检测。其原理如图1所示。

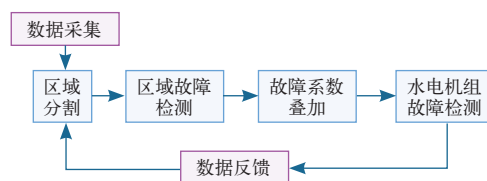


图1 故障检测原理框图

经上述检测分析后，还可以获得不同区域的相关性参数值，由这些参数值可以描述出整个水电机组不同区域的故障相关性。

2. 水电机组故障检测方法缺陷

对于水电机组中不同区域的安全隐患，其重要程度在整个故障检测过程中往往是存在差异的。有时会因局部区域危险程度的突变而造成检测系统发生虚报警现象，这主要是由于水电机组故障数据通常为小样本的精密数据，而传统的支持向量机故障检测方法通常对大数量样本进行分类分析才有效。局部区域危险系数在传统算法中往往会随着采

新世纪广西高等教育教改工程项目（2010JGZ072）。

集的水电机组图像像素数目的增加而降低,这就将造成不同区域安全隐患的重要程度差异系数增加,从而造成其相关性参数值的增加。

为此提出一种优化的支持向量机方法用于水电机组故障检测中,提高了水电机组故障检测的准确度。

支持向量机算法

1. 基本的支持向量机算法

支持向量机是在线性可分条件下的一种最佳分类方法,对于非线性问题,应进行非线性转换成高维空间内的线性问题,从变换空间中获取最佳分类面。

设有 n 个训练样本的数据集 $(x_i, y_i) (i=1, 2, \dots, n)$, $x_i \in R^n$, $y_i \in R$ 是第 i 个输入数据的输入类型,若获取函数关系为线形函数,则是线形回归,用 $f(x) = w^T x + b$ 表示,其中 w 是权值变量, b 表示偏差;否则是非线形回归。该种情况下,采用一个非线性映射 $\varphi(x)$ 将数据影射到高维特征空间中,在该空间中进行线性回归,则有 $f(x) = w^T \varphi(x) + b$ 。按照结构风险最小化原理,融入结构风险函数

$$R = \frac{1}{2} \|w\|^2 + CR_{\text{cmp}}(f) \quad (1)$$

式中, C 为惩罚函数参数,调整模型复杂度参数 $\|w\|^2$ 和经验风险 $R_{\text{cmp}}(f)$ 用于提高模型的泛化性能。设置经验风险为

$$R_{\text{cmp}}(f) = \frac{1}{n} \sum_{i=1}^n [y_i - f(x_i)]^2 \quad (2)$$

则在权 w 空间中的支持向量机分类问题能够优化成

$$\min_{w,b,e} H(w,b,e) = \frac{1}{2} \|w\|^2 + \frac{1}{2} \sum_{i=1}^n \xi_i^2 \quad (3)$$

限制规范为: $y_i[w^T \varphi(x_i) + b] = 1 - \xi_i$, γ 为可调整的正则化参数。

通过拉格朗日函数,将优化问题转换到相应的对偶空间中,则有

$$L(w,b,\xi,\alpha) = H(w,b,e) - \sum_{i=1}^n \alpha_i \{y_i[w^T \varphi(x_i) + b] + \xi_i - 1\} \quad (4)$$

式中,拉格朗日乘子 $\alpha_i \in R$,对上式进行优化有

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^n \alpha_i y_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^n \alpha_i y_i = 0 \\ \frac{\partial L}{\partial \xi_i} = 0 \Rightarrow \alpha_i = \gamma \xi_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \Rightarrow y_i[w^T \varphi(x_i) + b] + \xi_i - 1 = 0 \end{cases} \quad (5)$$

过滤变量 w, ξ ,获取如下矩阵方程

$$\begin{pmatrix} 0 & -Y \\ Y & XX^T + \gamma^{-1}I \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (6)$$

式中, $Y = (y_1, \dots, y_n)^T$; $I = [1, \dots, 1]$; $x_i = [\varphi(x_i)^T y_1, \dots, \varphi(x_i)^T y_n]$; $\alpha = [\alpha_1, \dots, \alpha_n]$ 。

按照Mercer条件,存在映射函数 φ 和核函数 $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$,令 $\Omega = XX^T$,则有

$$\Omega_{kl} = y_k y_l / \varphi(x_k)^T \varphi(x_l) = y_k y_l K(x_k, x_l) \quad (7)$$

LS-SVM的分类决策函数为

$$y(x) = \text{sgn} \left[\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right] \quad (8)$$

式中, α_i, b 由式(6)获取,按照泛函分析可得,符合Mercer定理规范的函数则为核函数 $K(x_i, y_i)$ 。核函数 $K(x_i, x)$ 存在如下几种形式:

(1) d 阶多项式核函数

$$K(x_i, x) = (1 + x_i^T x)^d \quad (9)$$

(2) RBF核函数

$$K(x_i, x) = \exp(-\|x - x_i\| / 2\sigma^2) \quad (10)$$

(3) 两层感知神经网络核函数

$$K(x_i, x) = \tan[k(x_i^T x) + \theta] \quad (11)$$

式中, σ, k 和 θ 都为可调常数。

2. 改进支持向量机算法

为了使得支持向量机的算法适应水电机组小样本故障数据分析的需要,引入最小二乘法对其改进,利用最小二乘法的小数据学习能力对水电机组故障数据进行优化,改进后算法主要参数包括正则化参数 γ 以及核函数参数。通常将径向基核(RBF核函数)作为核函数,因此可将分类问题看作是找到一个水电机组故障数据超参数 γ 与径向基核参数 σ ,使得最小二乘支持向量机的学习能力与

检测能力达到最优。

具体实现步骤为：

1) 确定水电机组故障数据参数 γ 与 σ^2 的数值取值区间，基于最小二乘支持向量机规则，通过经验可知， $\gamma \in [0.1, 5\ 000]$ ， $\sigma^2 \in [0.01, 10\ 000]$ 。

2) 在由 γ 与 σ^2 建立的 $m \times n$ 二维平面中，塑造水电机组故障数据最大取值范围参数对 (γ_i, σ_j^2) ， $(i=1, 2, \dots, m; j=1, 2, \dots, n)$ 。

3) 将10个水电机组故障数据网格节点参数对 (γ_i, σ_j^2) 发送到最小二乘支持向量机内，通过学习样本完成学习，同时对得到的准确分类率进行输出。将与最大准确分类率相应的节点值 (γ_i, σ_j^2) 看作是最佳水电机组故障数据参数对，同时将其作为新的中心点。

4) 将水电机组故障数据最佳参数值的0.01~10倍作为扩展网格的宽，也就是参数改变的大小，塑造水电机组故障数据新参数对 (γ_i, σ_j^2) 的 $m \times n$ 二维网格平面。

5) 反复进行步骤2)，直至达到要求的学习准确度。

经大量研究表明，对水电机组故障数据新参数进行确定的过程完全是自动进行的，简化了算法的计算，使得操作更加简单。通过上述过程对二乘支持向量机水电机组故障数据参数不断进行改进，将得到的最终结果作为所需确定的水电机组故障数据参数对。

实验与分析

本文以某水电厂机组数据为样本数据完成故障诊断试验。当前常用的水电机组故障诊断方法是变工况试验以及频谱分析，通过变工况试验能够准确获取水、电和机三种振源的影响大小，而依据振动频谱，该变量能够有效体现出故障的特征，从而对不同的特征进行判断。所以本文采用 $0.17 \sim 0.5f$ ， $2f$ ， $\geq 3f$ ，50 Hz频率，100 Hz频率，将振动和负荷、励磁电流的关系看作是条件属性，依次用 $c_1 \sim c_6$ 进行描述。将上述五种故障看作是决策属性，在决策表中依次用1，2，3，4，5进行描述。详细描述如表1所示。取全部样本数

据（13个）作为检验样本集，取前10个样本数据作为训练样本集，通过本文方法完成对支持向量机参数的改进。假设其种群规模是40，最大循环次数是300，SVM参数可在0.1~1 000之间进行搜索， σ 的取值范围为0.01~1 000。

表1 振动频谱征兆域训练样本

故障样本	条件属性						决策属性
	c_1	c_2	c_3	c_4	c_5	c_6	
u_1	0.02	0.04	0.79	0.97	0.93	0.07	2
u_2	0.02	0.21	0.86	0.93	0.92	0.06	2
u_3	0.02	0.03	0.91	0.99	0.99	0.03	2
u_4	0.07	0.06	0.93	0.63	0.63	0.04	3
u_5	0.07	0.09	0.99	0.61	0.61	0.04	3
u_6	0.02	0.23	0.94	0.03	0.03	0.08	4
u_7	0.02	0.03	0.96	0.06	0.06	0.24	4
u_8	0.06	0.97	0.08	0.03	0.05	0.13	5
u_9	0.09	0.07	0.21	0.08	0.08	0.98	5
u_{10}	0.05	0.03	0.99	0.07	0.07	0.02	6
u_{11}	0.02	0.03	0.99	0.06	0.06	0.08	4
u_{12}	0.04	0.98	0.08	0.03	0.03	0.03	5
u_{13}	0.06	0.03	1	0.07	0.07	0.03	3

利用本文提出的最小二乘支持向量机分析方法，将水电机组的故障数据从之前的9维减至2维，获取的分类效果如图2~图6所示。

分析图1可知本文方法，可对水电机组各类故

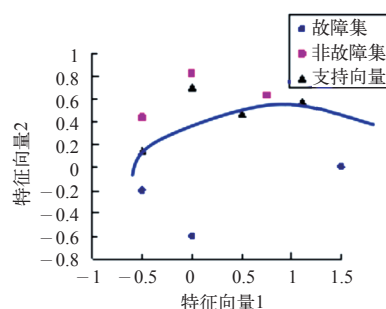


图2 转子不对中故障诊断结果

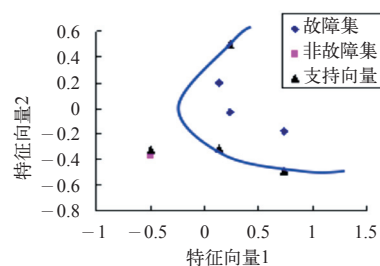


图3 动静碰摩故障诊断结果

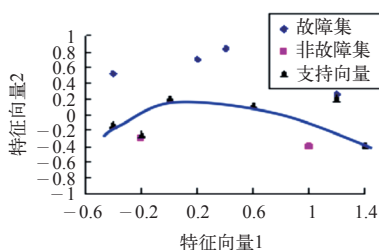


图4 转子质量不平衡故障诊断结果

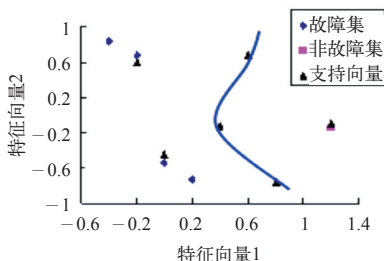


图5 尾水管偏心涡带故障诊断结果

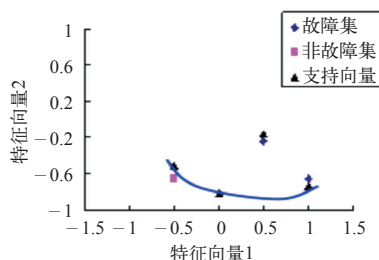


图6 磁极不均匀故障诊断结果

障进行准确判断。

最后，将本文方法的测试结果与传统的支持向量机方法进行对比，其分类结果如表2所示。结果说明，本文方法具有较高的分类性能，分类准确率达到100%，对应的超参数 γ 和径向基核参数 σ 分别为75.012 7以及0.205 5，而传统方法的分类准确率是60.427%，表明采用本文方法的寻优性能，能够很好地得到全局最佳支持向量机，实现水电机组故障的准确分类检测。采用本文方法改进的径向基参数 σ 值高于传统方法，说明本文方法可获取更多的最佳样本分类参数对，其样本总体分类准确率高于传统方法，验证了本文方法的有效性。

表2 优化方法比较结果

方法类型	分类正确率 (%)	超参数 γ	径向基参数 σ
传统方法	60.42	76.225 1	0.010 5
本文方法	100	75.021 7	0.205 5

结束语

水电机组故障数据通常为小样本的精密数据，而传统支持向量机通常对大数量样本数据进行分类分析，存在局限性，实际检测水电机组故障数据存在明显偏差，提出了适用于小样本模式识别的优化支持向量机方法，也就是最小二乘支持向量机，分析了优化最小二乘支持向量机参数选择的方法，塑造水电机组故障检测多元分类模型，提高支持向量机模型的故障分类检测质量，实现水电机组故障的准确检测。实验对采集的13个水电机组故障样本类型进行仿真分析，结果说明，优化支持向量机具有较高的小样本数据学习和预测性能，将其用于水电机组故障检测是有效、可行的。

参考文献

- [1] 周东华, 叶银忠. 现代故障诊断与容错控制 [M]. 北京: 清华大学出版社, 2000.
- [2] 黄丽华. 冷冻机状态监测与故障诊断专家系统的开发与研究 [D]. 上海: 华东理工大学, 2000.
- [3] 吴晓辉, 刘炯, 梁永春. 支持向量机在电力变压器故障诊断中的应用 [J]. 西安交通大学学报, 2007, 41 (6): 722-726.
- [4] 钟秉林, 黄仁. 机械故障诊断学 [M]. 北京: 机械工业出版社, 1997.
- [5] 饶泓, 虞国全, 胡倩如. 基于支持向量机的径向基网络结构优化 [J]. 计算机工程与应用, 2008, 44 (5): 67-69.
- [6] 黄戈, 崔建武, 陈晓芸, 等. 基于粒子群优化神经网络的水轮机振动故障诊断 [J]. 电网与清洁能源, 2009, 25 (4): 54-57.
- [7] 贾嵘, 张云, 洪刚. 基于改进PSO的LSSVM参数优化在变压器故障诊断中的应用 [J]. 电力系统保护与控制, 2010, 38 (17): 120-124.
- [8] 雷明. 基于单冲程频谱特征故障智能检测软件设计 [J]. 科技通报, 2014, 2 (30): 44-46.

(收稿日期: 2014-11-10) **EA**