

基于优化相关向量机的 化工装备故障预测方法

郁鹏^{1*}, 吉卫喜^{1,2}, 钱陈豪^{1,2}, 张国祥¹

(1. 江南大学机械工程学院, 江苏无锡 214122;
2. 江苏省食品先进制造装备技术重点实验室, 江苏无锡 214122)

摘要:针对影响化工装备的主要因素—腐蚀速率,在采集少量压力容器状态数据的基础上,提出了基于自导式粒子群优化算法(SGPSO)的相关向量机(RVM)故障预测方法。首先判别影响最为显著的5个表征因素;其次,使用SGPSO算法得到RVM模型最优核函数参数;然后通过样本数据计算得到预测结果的均值与方差,同时计算得到预测结果的概率分布与置信区间。仿真结果表明,SGPSORVM方法具有更高的准确率以及可靠度,因此该方法能够更有效的预测腐蚀速率,及时对化工装备的故障进行维修。

关键词:化工装备;自导式粒子群优化算法;相关向量机;故障预测;腐蚀速率

中图分类号:TQ053.2

文献标志码:A

文章编号:0253-4320(2017)05-0189-04

DOI:10.16606/j.cnki.issn0253-4320.2017.05.045

The fault prediction method of chemical equipment based on improved relevance vector machine

YU Peng^{1*}, JI Wei-xi^{1,2}, QIAN Chen-hao^{1,2}, ZHANG Guo-xiang¹

(1. School of Mechanical Engineering, Jiangnan University, Wuxi 214122, China;

2. Jiangsu Key Laboratory of Advance Food Manufacturing Equipment & Technology, Wuxi 214122, China)

Abstract: This paper proposes a new fault predictive algorithm based on self-guided particle swarm optimization algorithm optimizing relevance vector machine(RVM) on the basis of collecting few status data from pressure vessels, aiming at corrosion rate—the major factor influencing chemical equipment. First, five characterized factors with the most significant influences are confirmed; Then, the optimized kernel function parameter of RVM model is obtained through SGPSO algorithm; Finally, the mean and variance of prediction results are worked out by using the sample data above and probability distribution and confidence interval of prediction results can also be obtained. The simulation results testify that SGPSORVM has high accuracy and reliability and can predict the corrosion rate more effectively, therefore the chemical equipment fault can be repaired on time.

Key words: chemical equipment; SGPSO; RVM; fault prediction; corrosion rate

压力容器、换热器、输送管道等是化工装备的重要组成部分。其使用条件比较苛刻,通常都承载着高压、高温、剧毒或腐蚀性介质。在近几年的实际检测工作中,发现压力容器故障也多是由于腐蚀而引发的壁厚减薄、凹坑腐蚀、腐蚀开裂等,腐蚀失效成为压力容器故障的重要类型之一^[1-2]。

文献[3-5]讨论了不同的统计分析方法在腐蚀研究中的应用,包括灰色理论、人工神经网络分析等方法,但是上述这些基于大样本的诊断方法难以适应化工装备故障这样只有少量故障样本的预测与诊断。其次,由于化工装备故障表征因素的复杂性、不确定性和不完备性,缺少主导失效特征,无法直接与腐蚀速率建立关系,因此,研究在已知少量监测数据条件下,如何有效利用影响显著的表征因素来获取高准确度和高置信度的预测结果就显得尤为重要。文献[6-7]讨论了支持向量机(support vector machine, SVM)方法在故障预测中的应用,该方法只

需少量样本就能得到较好的预测结果,但随着样本数的增加,模型参数逐渐增多、模型结构会更加复杂,而且需要保证核函数满足 Mercer 条件。

针对以上不足,本文中提出了一种基于自导式粒子群优化算法^[8](self-government particles swarm optimization, SGPSO)的相关向量机(relevance vector machine, RVM)故障预测方法。该方法克服 SVM 规则化系数难确定、核函数受 Mercer 条件限制等局限,还可以概率形式输出结果。目前, RVM 已经成功应用到网络热点预测^[9]、故障诊断^[10]和智能预测^[11]等领域。而 SGPSO 算法是在标准的 PSO 算法寻优基础上,将之前很多次个体自我寻优的信息应用在粒子群此次全局寻优的过程中,因而提高搜索全局最优的能力。本文中基于 SGPSO 进行核参数的优化选取并进行仿真分析,最终通过某企业压力容器监测的显著表征因素数据验证了上述预测方法的可行性。

收稿日期:2016-09-19;修回日期:2017-03-07

基金项目:江苏省产学研联合创新资金(BY2014023-30);江南大学自主科研计划-青年基金理工类(JUSRPI16027)

作者简介:郁鹏(1991-),男,硕士生,主要研究方向为先进制造、数字化管理与全生命周期 MRO,通讯联系人, yupeng666891@163.com。

1 压力容器腐蚀状况的监测

1.1 设备腐蚀情况

通过对某企业的一批在用压力容器进行检修观察,收集相关样本数据。采用声发射技术^[12]对几台设备进行腐蚀监测,发现有 3 台蒸压釜釜体底部均有明显的腐蚀凹坑,壁厚最大减薄达 4 mm,成溃疡状分布。因此,使用相关向量机理论建立腐蚀速率预测模型,对腐蚀速率与相关表征因素的关系进行了分析,对未来可能的壁厚进行了预测。

1.2 腐蚀速率表征因素分析

根据压力容器腐蚀机理及专家经验,主要有 9 种主要的表征因素,如存储介质(Cl⁻、Ca(OH)₂、O₂浓度等)、pH、使用温度等。按照文献[13]的方法划分 9 种表征因素等级,最终得出影响较为显著的 5 种表征因素:含氧量(x₁)、使用温度(x₂)、铁离子变化(x₃)、pH(x₄)、CO₂含量(x₅)。采用声发射检测出壁厚差,计算得到腐蚀速率(x₀)。由于采样过程较为漫长,每个月监测 1 次,此次总共收集了 14 组数据,具体见表 1。

表 1 5 种表征因素及腐蚀速率数据

样本 序号	表征因素					腐蚀速率
	x ₁ / (mg·L ⁻¹)	x ₂ / °C	x ₃ / (mg·L ⁻¹)	x ₄	x ₅ / (mg·L ⁻¹)	x ₀ / (μm·a ⁻¹)
1	10.58	194.9	17.8	7.7	13.85	68.2
2	13.25	194.4	9.7	7.0	18.47	85.3
3	12.33	195.1	8.3	8.0	14.32	87.5
4	9.86	195.0	16.4	7.4	16.30	86.4
5	10.54	193.9	8.5	7.3	12.92	101.2
6	14.20	194.3	9.6	8.2	21.12	88.7
7	15.35	194.6	13.7	7.9	15.30	86.7
8	11.80	194.3	2.8	8.4	17.68	84.3
9	10.25	193.8	9.3	6.5	16.22	78.2
10	13.80	194.7	10.1	7.6	14.50	82.3
11	14.30	193.5	30.2	8.3	13.90	80.3
12	12.77	194.6	27.3	7.9	17.32	74.8
13	12.86	194.4	13.7	9.3	15.68	80.5
14	13.45	194.6	10.3	8.0	15.92	81.2

2 RVM 回归预测建模及优化

2.1 RVM 回归预测建模

假设训练样本集为 {x_i, t_i}_{i=1}^N, 目标函数值 t = [t₁, t₂, ..., t_N]^T 独立且自带噪声 ε_i, 则:

$$t_i = f(x_i, w) + \varepsilon_i \quad (1)$$

式中, ε_i 服从均值为 0, 方差为 σ² 的高斯分布, 则 RVM 学习监督模型就可以表示为:

$$\begin{cases} f(x_i, W) = \sum_{i=1}^N w_i k(x, x_i) + w_0 \\ k(x, x_i) = \exp[-(\|x - x_i\|^2)/\sigma^2] \end{cases} \quad (2)$$

式中, k(x, x_i) 为核函数; σ 为高斯核宽度; w_i 为权值系数。引入 β = 1/σ², 则对应似然函数为:

$$P(t | W, \beta) = (\beta/2\pi)^{N/2} \exp(-\beta/2 \|t - \varphi W\|^2) \quad (3)$$

若直接采用极大似然法估计 W 和 σ², 容易导致过拟合, 因此, 采用稀疏贝叶斯理论, 假定 w_i 服从均值为 0, 方差为 α_i⁻¹ 的高斯分布, 得到超参数 α 的先验概率分布为:

$$P(W, \alpha) = \prod_{i=1}^N N(w_i | 0, \alpha_i^{-1}) \quad (4)$$

根据贝叶斯公式求得所有未知权值向量 W 的后验概率分布:

$$P(W | t, \alpha, \beta) = [P(t | W, \beta)P(W, \alpha)] / [P(t | \alpha, \beta)] = N(W | \mu, \Sigma) \quad (5)$$

后验协方差矩阵 Σ 和均值向量 μ 分别为:

$$\begin{cases} \Sigma = (\beta\varphi^T\varphi + A)^{-1} \\ \mu = \beta \Sigma \varphi^T t \end{cases} \quad (6)$$

式中, A = diag(α₀, α₁, ..., α_N)。通过最大化边缘似然函数 P(t | α, β) 实现超参数 α_j 和 β 的优化:

$$P(t | \alpha, \beta) = \int P(t | W, \beta)P(W | \alpha) dW = (2\pi)^{-N/2} |I/\beta + \varphi\varphi^T/A|^{-1/2} \cdot \exp[-1/2t^T(I/\beta + \varphi\varphi^T/A)^{-1}t] \quad (7)$$

然后不断迭代整理得:

$$\begin{cases} \alpha_j = \lambda_j / \mu_j^2 \\ \beta = (N - \sum_j \lambda_j) / (\|t - \varphi\mu\|^2) \end{cases} \quad (8)$$

式中, j = 0, 1, 2, ..., N; μ_j 是权值后验均值向量 μ 的第 j 个元素; λ_j = 1 - α_j Σ_{jj}, Σ_{jj} 是后验协方差矩阵 Σ 的第 j 个对角元素。

RVM 通过不断迭代更新, 直到参数都收敛或达到最大训练次数为止, 得到超参数 α_{MP} 和噪声方差 σ_{MP}²。此时, 给定一个输入值 x*, 对应其估计值 t* 的预测分布为:

$$P(t_* | t, \alpha_{MP}, \sigma_{MP}^2) = \int P(t_* | W, \sigma_{MP}^2)P(W | t, \alpha_{MP}, \sigma_{MP}^2) dW \quad (9)$$

式中, 预测均值和方差分别是:

$$\begin{cases} t_* = \mu^T \varphi(x_*) \\ \sigma_*^2 = \sigma_*^2 + \varphi^T(x_*) \Sigma \varphi(x_*) \end{cases} \quad (10)$$

这样, RVM 对新观测值的预测输出即为 f(x*, μ)。根据预测均值和方差可以得到预测值的概率密度和分布:

$$\begin{cases} p(t_*) = (1/\sqrt{2\pi\sigma_*}) \exp[-(t_* - f_*)/2\sigma_*^2] \\ F(t_*) = (1/\sqrt{2\pi\sigma_*}) \int_{-\infty}^{t_*} -[(q - d_*)/2\sigma_*^2] dq \end{cases} \quad (11)$$

式中, q 为 $F(t_*)$ 的积分变量, $q \in (-\infty, t_*)$; d_* 为 t_* 的决策值。

2.2 基于自导式粒子群优化算法的 RVM 参数寻优

由于选用高斯函数作为 RVM 的核函数, 显然, 核参数(高斯核宽度 σ) 即为影响 RVM 性能的主要因素^[14]。在实际应用中一般都是主观判断选取核参数, 存在一定的盲目性, 因此需要采用合理的方法对核参数进行科学地求取。基本粒子群算法(particle swarm optimization, PSO) 是根据每个粒子个体的历史最优位置和粒子群的历史全局最优位置来更新粒子的速度, 但在寻优过程中, 粒子群往往会疏漏一些潜在的有用信息, 陷入过早收敛的局面。SGPSO 算法是在基本粒子群算法 PSO 的基础上提出的一种优化智能算法, 其原理为: 利用每个粒子自身在历次实验中所获得的局部最优位置集合 $P_{ad}(t)$, 来引导粒子在解空间中的移动。 $P_{ad}(t)$ 是指算法运行 t 次后所搜索到的每个粒子局部最优位置的集合, 由前 $t-1$ 次实验后的局部最优粒子所组成。优化后粒子位置更新公式如下:

$$\begin{aligned} v_{id}(k+1) &= wv_{id}(k) + c_1r_1[p_{id}(k) - x_{id}(k)] + \\ & c_2r_2[p_{gd}(k) - x_{gd}(k)] \quad (12) \\ x_{id}(k+1) &= \end{aligned}$$

$$\alpha x_{id}(k) + (1 - \alpha)[p_{ad}^i(t) - x_{id}(k)] + v_{id}(k+1) \quad (13)$$

$$P_{ad}^i(t) = P_{ad}^i(1) \cup P_{ad}^i(2) \cup \dots \cup P_{ad}^i(t-1) \quad (14)$$

其中, w 为惯性权值; c_1 和 c_2 为加速系数; r_1 和 r_2 是 2 个在 $[0 \sim 1]$ 范围内变化的随机数, 用来保持粒子群的多样性; k 为迭代次数; α 是 $[0 \sim 1]$ 内的固定值, 表示共享继承先前实验中局部最优信息的水平, 当 SGPSO 算法第一次运行时, 由于没有局部最优信息可继承, 因此 α 取 1。公式(13)表示第 t 个粒子第 k 次实验时, 将运行 SGPSO 算法 $t-1$ 次所搜索的局部最优赋予 $P_{ad}(t)$ 。

为了提升模型预测性能, 本文中 RVM 模型训练采用 k 折交叉验证的方法, 以误差最小为目标, 采用 SGPSO 算法对 σ 进行优化, 最终选取最优值, 具体优化步骤如下:

步骤 1: 初始化参数, 微粒个数 m 、惯性权值 w 、加速系数 c_1 和 c_2 、随机数 r_1 和 r_2 、最大允许迭代次数 k 、各微粒的初始位置 P_1 和初始速度 v_1 等。

步骤 2: 对 RVM 进行测试, 建立预测模型。测试结果满足允许误差, 则继续下一步, 否则, 修改

参数。

步骤 3: 初置全局最优位置, 使用 k 折交叉验证, 计算每个粒子最优位置 P_i' 的适应值 F_i , 取最大 F_i 粒子所在位置作为全局最优位置 P_g' 。

步骤 4: 根据公式(12)、(13), 跟新粒子位置和速度, 计算粒子当前的适应值 F_i , 和 $P_{ad}(t)$ 集合中值进行比较, 选取最大 F_i 对应位置为最优位置 P_i' , 并更新 $P_{ad}(t)$ 集合。

步骤 5: 将更新后每个粒子的 P_i' 与全局最优位置 P_g' 对比, 若表现更优, 则将此粒子 P_i' 替代 P_g' , 否则不变。

步骤 6: 如满足设定终止条件, 则停止迭代寻优, 将 P_g' 作为最终结果; 如不满足, 则令 $i = i + 1$, 跳转至步骤 4 继续迭代。

2.3 SGPSORVM 故障预测模型

SGPSORVM 故障预测模型的基本原理为: 首先对样本数据进行相空间重构, 得到 RVM 的训练样本及测试样本; 接着, 采用 k 折交叉验证法和 SGPSO 算法优化 RVM 模型核参数; 最后建立 SGPSORVM 的预测模型, 对测试样本进行计算得到腐蚀速率的预测均值 Y_* 与预测方差 σ_*^2 , 进一步得到预测结果的概率分布 $p(t_*)$ 与概率密度函数 $F(t_*)$ 。预测模型流程见图 1。

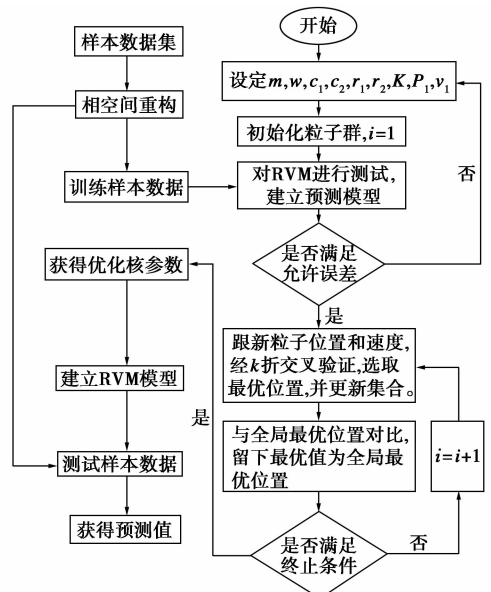


图 1 SGPSORVM 预测框架

3 实例应用分析

3.1 数据处理

为了保证模型训练的顺畅性, 将对数据进行归一化处理, 使得样本数据取值在 $[0 \sim 1]$ 之间。初步

设定嵌入维数 $d = 2$, 则经过相空间重构后剩余 12 组样本数据, 取前 7 组作为训练样本, 后 5 组作为测试样本。采用 SGPSO 算法对模型参数进行优化, 设置种群规模 10、迭代次数 100, 惯性权值 w 范围设为 $[0.4 \sim 0.9]$, 个体范围 $[0 \sim 400]$, 速度范围 $[-1 \sim 1]$ 。文献[8]指出, 当 c_1, c_2 取较大值 $[1.2 \sim 2.0]$, α 取偏小值 $[0.4 \sim 0.8]$ 时, SGPSO 算法优化结果较好。作为对比, 分别使用人工选取核参数的 RVM 以及 PSORVM 进行对比。

3.2 预测结果及分析

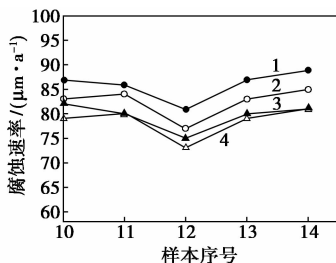
对优化后的模型进行预测, 将 7 组训练样本构成的特征向量 X 代入 SGPSORVM 模型中进行训练, 参照式(10)和(12)进行迭代计算, 得到超参数 α_{MP} 和噪声方差 σ_{MP}^2 , 参照式(14)计算得到预测均值, 其对应核参数取值与预测结果如表 2 所示, 对应误差率见表 3。由表 3 可以看出, SGPSORVM 的误差率要明显低于其他 2 种方法, 相对于 RVM 和 PSORVM 的预测方法, SGPSORVM 预测方法得到的预测结果表现较好。

表 2 模型核参数取值及预测值

样本 序号	RVM		PSORVM		SGPSORVM	
	σ 值	预测值	σ 值	预测值	σ 值	预测值
10	0.3	73.3	1.2	73.3	0.9	81.3
11	0.9	75.3	1.7	76.2	1.4	76.3
12	1.2	68.8	0.6	67.9	0.7	72.8
13	1.6	73.5	0.9	76.5	0.6	77.5
14	2.4	73.1	1.9	77.2	1.5	77.2

表 3 误差率对比表

样本 序号	RVM 误差率/%	PSORVM 误差率/%	PSORVM 误差率/%
10	10.96	8.53	1.22
11	6.25	5.11	5.00
12	8.00	9.33	2.67
13	8.75	4.91	3.75
14	9.88	4.94	4.94
平均误差	8.77	6.54	3.52



1—置信度为 95% 的上界; 2—实际值; 3—SGPSORVM;

4—置信度为 95% 的下界

图 2 腐蚀速率置信区

根据方差计算得到均值的置信区间, 以测试样本的 5 个数据集, 其预测值、实际值及置信度为 95% 的上下界计算结果如图 2 所示。从图 2 中可以看出, 腐蚀速率的实际值都包含在置信度 95% 的置信区间内, 有效地证明了 SGPSORVM 故障预测模型具有很好的可靠性, 在实际应用中是行之有效的。

4 结论

提出了一种基于自导式粒子群优化算法的相关向量机模型 SGPSORVM。

(1) 该模型利用自导式粒子群优化算法较强的全局寻优能力, 以较快的收敛速度寻优得到最优的核参数, 有效地避免了人工选取核参数的不确定性问题, 通过仿真对比, 得到较优的预测结果。

(2) 利用 RVM 概率输出的优势, 得到预测值的同时也获取了其置信区间, 对腐蚀速率的预测更为可靠, 为压力容器的故障预判提供了一种精准有效可靠的方法。

参考文献

- [1] 戴树和. 化工设备事故预测技术的研究[J]. 化工学报, 1998, 49(5): 624-631.
- [2] 黄坤利. 谈压力容器安装和维修的管理[J]. 设备管理与维修, 2014, (S2): 35-36.
- [3] 刘静, 李晓禄, 朱崇伟, 等. 利用人工神经网络技术预测气田环境下 316L 不锈钢临界点蚀温度[J]. 中国腐蚀与防护学报, 2016, 36(3): 205-211.
- [4] 李启锐, 黎贤娇, 李剑彬, 等. 炼厂装置设备腐蚀速率预测算法的研究与应用[J]. 现代计算机: 专业版, 2015, (20): 24-27, 32.
- [5] 杜京义, 韩娟, 寇水潮, 等. 基于模糊可拓层次分析的接地网腐蚀速率预测[J]. 计算机应用与软件, 2014, 31(6): 170-173, 197.
- [6] 湛璐, 贺兴时. 改进的支持向量机算法在人脸识别上的应用[J]. 纺织高校基础科学学报, 2015, 28(1): 108-115.
- [7] 寇永强, 贾江鸿, 程远方. 基于灰色支持向量机的套管腐蚀速率预测研究[J]. 石油机械, 2011, 39(4): 23-26.
- [8] 林蔚天. 改进的粒子群优化算法研究及其若干应用[D]. 上海: 华东理工大学, 2014.
- [9] 李睿, 赵保华, 王朋. 基于遗传优化加权相关向量机的网络热点预测模型[J]. 计算机应用与软件, 2015, 32(6): 312-315, 319.
- [10] 王波, 刘树林, 张宏利, 等. 相关向量机及其在机械故障诊断中的应用研究进展[J]. 振动与冲击, 2015, 34(5): 145-153, 167.
- [11] 范庚, 马登武, 张继军, 等. 基于自适应 RVM 的电子系统缓变故障预测方法[J]. 北京航空航天大学学报, 2013, 39(10): 1319-1324.
- [12] 侯素霞, 罗积军, 徐军, 等. 基于声发射技术的压力容器应力腐蚀检测研究[J]. 腐蚀科学与防护技术, 2006, 18(3): 220-221.
- [13] 张颖, 陈荣刚, 戴光, 等. 储罐底板腐蚀状况的贝叶斯判别预测方法[J]. 压力容器, 2010, (1): 31-34.
- [14] 王正方, 王勇, 刘秀华. 基于灰色系统理论的常压蒸馏装置腐蚀预测[J]. 中国石油大学学报: 自然科学版, 2010, 34(2): 114-118. ■