航空液压泵磨损状况预测

葛 薇 王少萍

(北京航空航天大学 自动化科学与电气工程学院, 北京 100191)

摘 要: 磨损是航空液压泵典型的渐进性故障之一,因磨损量难以测量,对磨损状况进行准确的预测比较困难. 针对上述问题,提出了基于多尺度数据的支持向量机预测方法,该方法将支持向量机用于时间序列预测的基本理论和数据多尺度分解、相空间重构方法结合,能更有效地挖掘时间序列的内在联系及变化规律. 采用回油流量作为反映航空液压泵磨损状况的敏感信号,将其分解为趋势项和随机项,采用多尺度支持向量机作等维信息一步预测和多步预测,利用网格方法对预测模型参数寻优. 对比传统支持向量机算法分析其预测精度,结果表明:多尺度支持向量机模型预测精度更高,适于中长期预测.

关 键 词: 航空液压泵; 磨损预测; 多尺度支持向量机; 数据分解; 相空间重构

中图分类号: TP 271$^+$.31

文献标识码: A **文 章 编 号**: 1001-5965(2011)11-1410-05

Wear condition prediction of hydraulic pump

Ge Wei Wang Shaoping

(School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: Wear is a typical progressive failure of aero hydraulic pump. It is difficult to measure wear loss. To solve precision wear condition prediction problem, multi-dimensional support vector machine (SVM) prediction method was proposed, based on theoretical basis of SVM applied to time series prediction, multi-dimensional data decomposition and phase space reconstruction. The inner relationship of time series can be mined and reflected more effectively by this method. Oil-return flow was chosen to reflect the wear condition of hydraulic pump and was decomposed into trend data and random data. Multi-dimensional SVM was applied to predict oil-return flow of the aero hydraulic pump one-step ahead and multi-step ahead with grid search optimization method. The results show that multi-dimensional SVM model has higher prediction precision and is very suitable for long-term forecasting compared with the predicted results of traditional SVM.

Key words: aero hydraulic pump; wear prediction; multi-dimensional support vector machine; data decomposition; phase space reconstruction

液压泵作为飞机液压系统的动力元件,其性能直接影响液压系统的工作状况. 据统计,航空液压泵 80% 以上的故障是由于磨损污染造成的. 尤其是飞机液压系统向高压化方向发展,油液污染导致的液压泵磨损愈加显著,因此实时监控磨损特征量,及时得到剩余寿命预测值,可以在保障液压系统安全的情况下有效使用液压泵,对于飞机液压系统实现视情维修有重要意义.

目前国内对液压泵磨损状况的监控主要是利用光谱或铁谱数据分析油液成分和颗粒物[1]. 由于液压泵型号、材料以及工作油液不同,需要建立丰富的油样数据库才能实现准确的磨损状况预测,很难推广到机载液压系统上. 有些文献采用轮廓仪测定摩擦副各典型工作区域的微观形态,但

收稿日期: 2010-07-20; **网络出版时间**: 2011-11-14 17:20; **DOI**: CNKI:11-2625/V.20111114.1720.013
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20111114.1720.013.html
基金项目: 国家 863 计划资助项目(2009AA04Z412); 航空创新基金资助项目(08D51010)
作者简介: 葛 薇(1985 -),女,河北石家庄人,硕士生,gewei.buaa@gmail.com.

必须拆卸液压泵,非常不便.

液压泵在工作时,各配对摩擦副存在高速相对运动并由此产生磨损. 磨损随工作时间增长而逐渐累积,当积累到一定程度时,润滑油膜的建立趋于不稳定,泵内泄漏量逐渐增大,最终导致泵出口流量和容积效率降低,不能满足使用要求. 文献[2]通过实验证明,在相同工况下,回油流量和最深的划痕有近似线性的关系. 因此选用回油流量作为反映液压泵磨损状况的敏感信号.

常用的时间序列预测方法存在各自固有的缺陷,如:AR/ARMA 等传统方法精度不高,且容错性较差,只适合做短区间预报[3];神经网络网络结构选择以经验为主,存在过学习和欠学习问题,容易陷入局部极小[4].

支持向量机建立在统计学习理论、VC 维理论和结构风险最小原理基础上,得到的是全局最优解,具有较好的泛化性能[5]. 多尺度分解方法,如:经验模式分解[6]、小波分解[7]等,能从不同尺度反映数据的内部特征. 相空间重构能有效地挖掘时间序列内在联系. 将上述方法结合起来,用于航空柱塞泵回油流量预测,通过实验验证这种方法的可行性与优越性.

1　多尺度支持向量机理论分析

1.1　支持向量机

支持向量机用于时间序列预测的数学描述如下[8]:给定训练数据样本集合$\{(\boldsymbol{x}_1,y_1),\cdots,(\boldsymbol{x}_l,y_l)\}$,其中$\boldsymbol{x}_i\in\mathbf{R}^n$,$y_i\in\mathbf{R}$,$i=1,2,\cdots,l$,利用非线性映射$\boldsymbol{\Phi}$将数据$\boldsymbol{x}$映射到高维特征空间$H$,在$H$空间内寻找一个线性函数$f(\boldsymbol{x})$来逼近数据:

$$f(\boldsymbol{x})=(\boldsymbol{\omega},\boldsymbol{\Phi}(\boldsymbol{x}))+b \tag{1}$$

式中,$\boldsymbol{\Phi}:\mathbf{R}^n\to H$;$\boldsymbol{\omega}\in H$;$b$为阈值.

采用线性ε不敏感损失函数,引入松弛因子ξ_i和ξ_i^*,则回归问题等价于求解以下优化问题:

$$\left.\begin{aligned}&\min\ R(\boldsymbol{\omega},\xi_i,\xi_i^*)=\\&\qquad\frac{1}{2}\boldsymbol{\omega}\cdot\boldsymbol{\omega}+c\sum_{i=1}^{n}(\xi_i+\xi_i^*)\\&\text{s.t.}\ \ y_i-(\boldsymbol{\omega},\boldsymbol{\Phi}(\boldsymbol{x}))-b\leqslant\varepsilon+\xi_i\\&\qquad(\boldsymbol{\omega},\boldsymbol{\Phi}(\boldsymbol{x}))+b-y_i\leqslant\varepsilon+\xi_i^*\\&\qquad\xi_i\geqslant 0,\xi_i^*\geqslant 0\end{aligned}\right\} \tag{2}$$

式中,ε为回归逼近误差管道大小,控制支持向量的个数和泛化能力;c为惩罚函数,控制对超出误差ε的样本的惩罚程度.

根据 Mercer 核定理,引入核函数$K(\boldsymbol{x}_i\cdot\boldsymbol{x}_j)$,将求解式(2)的最优问题转化为它的对偶问题:

$$\left.\begin{aligned}&\max\ W(\alpha,\alpha^*)=-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i-\alpha_i^*)\cdot\\&\qquad(\alpha_j-\alpha_j^*)K(\boldsymbol{x}_i\cdot\boldsymbol{x}_j)+\\&\qquad\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)y_i-\sum_{i=1}^{l}(\alpha_i+\alpha_i^*)\varepsilon\\&\text{s.t.}\ \ \boldsymbol{\omega}=\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)\boldsymbol{\Phi}(\boldsymbol{x}_i)\\&\qquad 0\leqslant\alpha_i,\alpha_i^*\leqslant c\end{aligned}\right\} \tag{3}$$

此时回归函数$f(\boldsymbol{x})$可以直接表示为

$$f(\boldsymbol{x})=\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)K(\boldsymbol{x}_i\cdot\boldsymbol{x}_j)+b \tag{4}$$

支持向量机通过核函数实现非线性映射,将样本空间映射到高维乃至于无限维的特征空间,把寻找最优线性回归超平面的算法转化为求解凸约束条件下的凸规划问题. 理论上,凸规划问题一定有全局最优解.

1.2　多尺度支持向量机

1.2.1　时间序列多尺度分解

工程问题中的时间序列一般是比较复杂的、广义平稳或者非平稳的,是许多因素共同作用的结果,不容易直接分析其变化规律,需将其分解进行逐项分析,再综合最终结果[9]. 一般将时间序列Y按不同性质分解为以下 4 部分的组合[10]:

$$Y=f(T,C,S,I) \tag{5}$$

式中,T为趋势分量,描述由于某种根本性因素影响,时间序列的中长期变化趋势;C为循环分量,描述趋势曲线在长期时间内呈现摆动的现象;S为季节分量,描述时间序列由于自然条件影响,随季节转变的周期性变动;I为不规则分量,描述随机因素引起的序列变动.

上述 4 部分分量若相互独立,则为相加关系;若相互关联,则为相乘关系. 以加法模型为例,数据分解的步骤如下所示:

1) 通过移动平均得到趋势循环量$(T+C)$,则季节变动和不规则变动$(S+I)=Y-(T+C)$;

2) 利用曲线拟合等方法求出趋势分量T,则循环分量$C=(T+C)-T$;

3) 计算各周期中相同期的值的平均数,并进行调整后,作为季节分量S的值,则不规则分量$I=(S+I)-S$.

1.2.2　相空间重构

在现实系统的研究中,由于某些客观条件限制,只能观测到系统某个单一指标的时间序列. 为了更有效地进行预测模型建模,需要深入挖掘数据间的关联关系. 选择适当的嵌入维数,将原始的

一维时间序列转化为多维相空间[11].

建立映射关系:$R^k \rightarrow R$. 经过映射,预测模型的输入向量 $\boldsymbol{X}$ 和输出向量 $\boldsymbol{Y}$ 为

$$\left.\begin{aligned}\boldsymbol{X}&=\begin{bmatrix}x_1 & x_2 & \cdots & x_k\\ x_2 & x_3 & \cdots & x_{k+1}\\ \vdots & \vdots & \cdots & \vdots\\ x_{n-m} & x_{n-m+1} & \cdots & x_{n-1}\end{bmatrix}\\ \boldsymbol{Y}&=\begin{bmatrix}x_{k+1}\\ x_{k+2}\\ \vdots\\ x_n\end{bmatrix}\end{aligned}\right\} \tag{6}$$

式中,k 为最佳嵌入维数,它的选取以最终预报误差(FPE, Final Prediction Error)最小为准则. FPE 准则用模型一步预报误差的方差 $e_{\mathrm{FPE}(k)}$ 来判定预测模型阶数是否适用. $e_{\mathrm{FPE}(k)}$ 越小,认为模型拟合得越好.

$$e_{\mathrm{FPE}(k)}=\frac{n+k}{(n-k)^2}\sum_{t=k+1}^{n}(x_t-\hat{x}_t)^2 \tag{7}$$

式中,x_t 为实际值;$\hat{x}_t$ 为预测值. $e_{\mathrm{FPE}(k)}$ 为 k 的函数,因此总可以找到一个最优值 k_{opt} 使 $e_{\mathrm{FPE}(k)}$ 最小,该值即为最佳嵌入维数.

1.2.3 等维信息预测

为提高预测模型中长期预测精度,将相继得到的实验数据不断考虑进去,更新预测模型输入向量. 每补充一个新信息的同时去掉一个最老的信息,这样既考虑了随着时间推移,老信息功能不断退化和新信息功能不断增强,又使系统能适应不断发展变化的实际情况. 模型维数始终保持不变,因此称为等维信息预测.

等维信息 1 步预测以实验数据为基础,每次只预测下一个数据点. 提前 1 步的预测值 $\hat{x}_{n+1}$ 为

$$\hat{x}_{n+1}=f(x_{n-(k-1)},x_{n-(k-2)},\cdots,x_n) \tag{8}$$

实际工程应用中,有的维修备件准备需要较长时间,提前 1 步预测不能及时地将故障征兆反馈给维修人员,因此需要提前多步预测. 等维信息多步预测利用已有的实验数据和当前已经预测得到的数据,动态更新输入,保持数据集合维数不变. 提前 m 步的预测值 $\hat{x}_{n+m}$ 为

$$\hat{x}_{n+m}=f(x_{n-(k-m)},\cdots,\hat{x}_{n+1},\cdots,\hat{x}_{n+m-1}) \tag{9}$$

1.2.4 建立多尺度支持向量机

结合支持向量机、多尺度数据分解和相空间重构理论,将时间序列 $\boldsymbol{y}$ 分解成 s 个分量:

$$\boldsymbol{y}=\boldsymbol{x}_1+\boldsymbol{x}_2+\cdots+\boldsymbol{x}_s \tag{10}$$

对于分解得到的 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s$,采用 FPE 准则寻找最佳嵌入维数 $k_1,k_2,\cdots,k_s$,利用支持向量机建立预测模型,得到预测函数 $f_1,f_2,\cdots,f_s$,则最终预测值 p_y 为

$$p_y=\sum_{i=1}^{s}f_i(\boldsymbol{x}_i) \tag{11}$$

选用径向基核函数进行非线性映射:

$$K(\boldsymbol{x}\cdot\boldsymbol{x}_i)=\exp\{-\gamma|\boldsymbol{x}-\boldsymbol{x}_i|^2\} \tag{12}$$

各个分量的预测模型结构如图 1 所示.

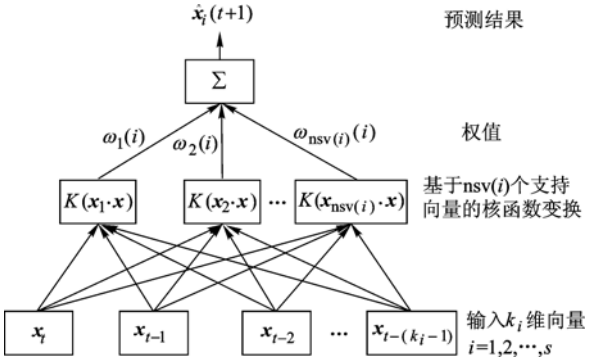

图 1 支持向量机预测模型结构

该预测模型中,如式(2)所示,c 控制对超出误差的样本的惩罚程度,γ 控制着支持向量对输入量变化的敏感程度,这两个参数直接影响着预测模型的预测精度. 采用网格寻优方法,令 c 和 γ 在一定范围内取离散值,取使得最终训练集预测精度最高的 c 和 γ 作为最后模型的参数. 在得到支持向量机最优参数后,即可以进行预测.

2 实例分析

2.1 数据来源

对某型飞机液压泵进行长试试验,在试验台泵出口处设置涡轮流量传感器监测泵出口流量. 在全流量工况下,每隔 5 h 记录一次回油流量,当回油流量超出 2.8 L/min 时,判定液压泵失效,停止采集.

图 2 为液压泵在整个试验过程中回油流量定时采集得到的时间序列,在采集到第 238 个点时,也就是工作到约 1 190 h,回油流量超过2.8 L/min,停止试验.

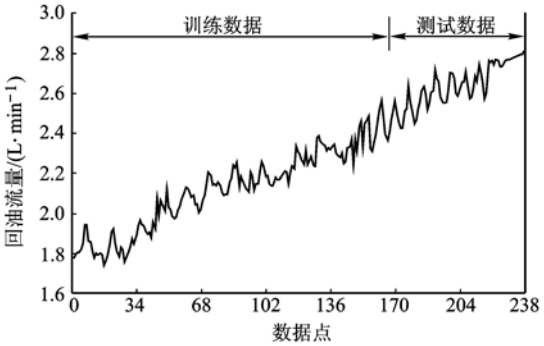

图 2 飞机液压泵长试试验回油流量变化的时间序列

2.2　多尺度支持向量机预测模型建立

通过对多台航空液压泵试验结果分析,液压泵磨损状况是一个逐渐严重的过程,其回油流量呈现逐渐上升的趋势. 图 2 所示试验数据也表明了这一趋势. 因此认为回油流量序列是一个非平稳的、带有近似线性趋势项存在的序列$\{x_i\}$. 将回油流量数据分解为趋势项x_{i1}和随机项x_{i2},分别建立支持向量机预测模型,流程如图 3 所示.

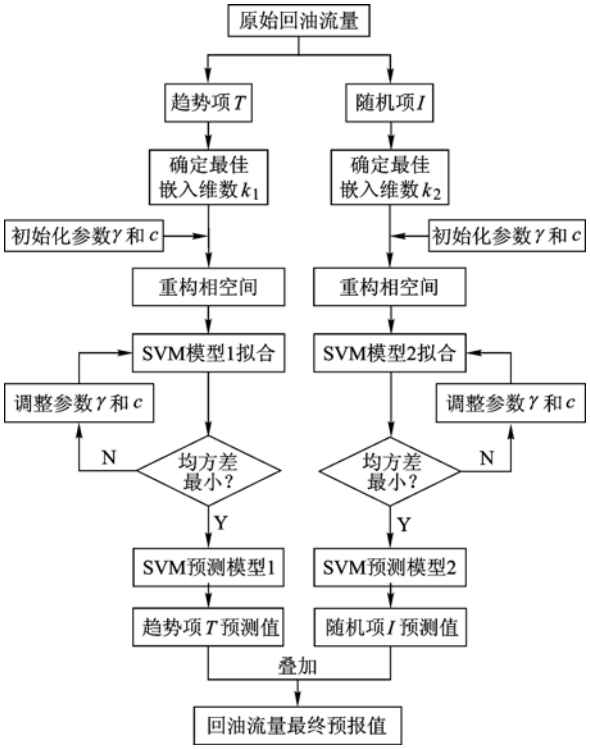

图 3　多尺度支持向量机回油流量预测流程

采用如下公式进行数据分解:

$$x_{i1} = \begin{cases} \dfrac{1}{i}\sum_{j=0}^{i} x_i & i \leqslant l \\ \dfrac{1}{l}\sum_{j=i-l}^{i} x_i & l < i \leqslant n \end{cases} \qquad (13)$$

$$x_{i2} = x_i - x_{i1}$$

式中,l 为光滑系数,它越大,时间序列越光滑. 选取 $l=30$ 进行数据分解.

采用 FPE 准则选取最佳嵌入维数 k_1 和 k_2 构建训练样本空间,嵌入维数和最终预报误差的关系如图 4 所示.

图 4 显示,趋势项 k_1 为 3 时,最终预报误差最小,$e_{\mathrm{FPE}(k)}=1.1538\times10^{-4}$,因此 $k_1=3$;随机项 k_2 为 10 时,最终预报误差最小,$e_{\mathrm{FPE}(k)}=1.6\times10^{-3}$,因此 $k_2=10$.

利用网格方法寻优(c,γ),趋势项预测模型的最佳参数为:$c=32$,$\gamma=0.0884$;随机项预测模型的最佳参数为:$c=0.0442$,$\gamma=22.6274$.

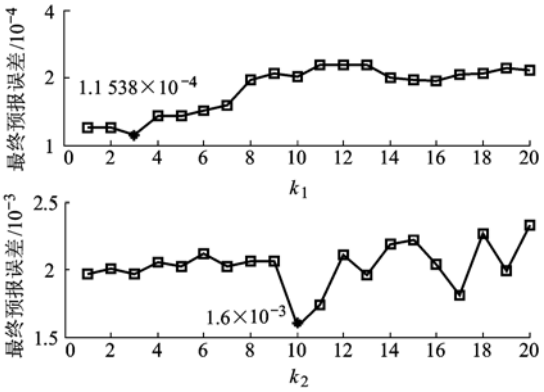

图 4　嵌入维数和最终预报误差的关系

2.3　预测模型的测试及结果评价

将回油流量数据分为两组,前 166 个数据作为训练数据集合,用于训练支持向量机预测模型,后 72 个数据作为测试数据集合,用于检验模型预测精度. 对回油流量分别作提前 1 步、提前 2 步和提前 3 步预测,预测结果如图 5 ~ 图 7 所示.

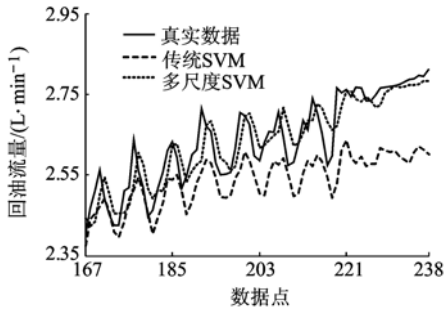

图 5　提前 1 步预测

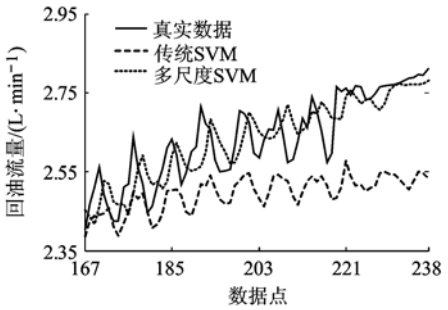

图 6　提前 2 步预测

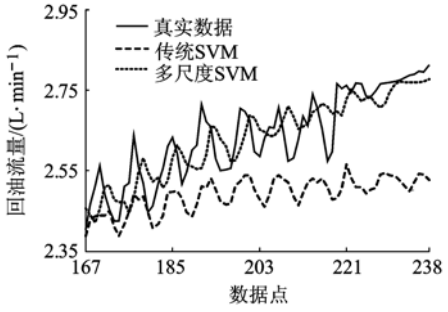

图 7　提前 3 步预测

选取平均绝对误差 σ_{MAE} 和均方误差 σ_{MSE} 为评价标准,其表达式为

$$\sigma_{\mathrm{MAE}} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i| \tag{14}$$

$$\sigma_{\mathrm{MSE}} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2} \tag{15}$$

将传统支持向量机和多尺度支持向量机预测结果对比,如表 1 所示.

表 1 预测指标对比

预测方法	预测步数	σ_{MAE}	σ_{MSE}	平均相对误差/%	误差限/%
传统支持向量机	提前 1 步	0.0995	0.0136	3.70	8.55
	提前 2 步	0.1412	0.0190	5.25	10.20
	提前 3 步	0.1457	0.0195	5.42	10.32
多尺度支持向量机	提前 1 步	0.0380	0.0056	1.46	4.75
	提前 2 步	0.0484	0.0071	1.86	5.76
	提前 3 步	0.0494	0.0073	1.90	5.91

从表 1 的预测指标可以看出,传统支持向量机随着时间推进,预测误差越来越大. 多尺度支持向量机对不同尺度数据进行分析,并在预测时动态更新输入向量,因此有更高的预测精度.

3 结 论

将多尺度支持向量机预测结果和传统支持向量机预测结果比较,实验数据和分析结果表明:

1) 多尺度分解和相空间重构可以有效发掘时间序列的内在规律;

2) 多尺度支持向量机预测具有较高的预测精度,即使提前 3 步预测,预测误差仍控制在 6% 以内,较好地跟随了回油流量的变化趋势. 对于提前 3 步预测,相当于提前 15 h 预测,有充分的时间通知维修人员进行备件准备.

参考文献 (References)

[1] 赵大庆. 液压泵污染磨损与控制[M]. 北京:北京煤炭工业出版社,1993:1 - 50
Zhao Daqing. Pollution wear and control of pump [M]. Beijing: China Coal Industry Publishing House, 1993: 1 - 50 (in Chinese)

[2] 张文超. 机载液压泵源健康管理系统研究[D]. 北京:北京航空航天大学自动化科学与电气工程学院,2008
Zhang Wenchao. Research on on-board PHM of hydraulic power system [D]. Beijing: School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, 2008 (in Chinese)

[3] Wang Wenchuan, Chau Kwok-Wing, Cheng Chuntian, et al. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series [J]. Journal of Hydrology, 2009, 374(3 - 4): 294 - 306

[4] Hippert H S, Pedreira C E, Souza R C. Neural networks for short-term load forecasting: a review and evaluation [J]. Power Systems, 2001, 16(1): 44 - 55

[5] Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning method [M]. Beijing: China Machine Press, 2005: 1 - 124

[6] Peng Xinjun, Wang Yifei. A normal least squares support vector machine (NLS-SVM) and its learning algorithm [J]. Neurocomputing, 2009, 72(16 - 18): 3734 - 3741

[7] Li Lin, Ji Hongbing. Signal feature extraction based on an improved EMD method [J]. Measurement, 2009, 42(5): 796 - 803

[8] Osowski S, Garanty K. Forecasting of the daily meteorological pollution using wavelets and support vector machine [J]. Engineering Applications of Artificial Intelligence, 2007, 20(6): 745 - 755

[9] 康卓,黄竞伟,李艳,等. 复杂系统数据挖掘的多尺度混合算法[J]. 软件学报,2003,14(7):1229 - 1237
Kang Zhuo, Huang Jingwei, Li Yan, et al. A multi-scale mixed algorithm for data mining of complex system [J]. Journal of Software, 2003, 14(7): 1229 - 1237 (in Chinese)

[10] 杨叔子,吴雅,轩建平,等. 时间序列分析的工程应用[M]. 2 版. 武汉:华中科技大学出版社,2007:1 - 20,175 - 222
Yang Shuzi, Wu Ya, Xuan Jianping, et al. Time series analysis in engineering application [M]. 2nd. Wuhan: Huazhong University of Science & Technology Press, 2007: 1 - 20, 175 - 222 (in Chinese)

[11] Li Hongxing, Xu Li D. Feature space theory: a mathematical foundation for data mining [J]. Knowledge-Based Systems, 2001, 14(5 - 6): 253 - 257

(编 辑:刘登敏)