

一类支持向量机的设备状态自适应报警方法

张庆^{1,2}, 徐光华^{1,2}, 华成¹, 张熠卓¹

(1. 西安交通大学机械工程学院, 710049, 西安; 2. 西安交通大学机械制造系统工程国家重点实验室, 710049, 西安)

摘要: 为了提高对异常状态识别的适应性和有效性, 提出了一种基于一类支持向量机的设备状态自适应报警方法. 该方法使用一类支持向量机的在线算法, 动态估计监测参数在高维特征空间中的最优分布区域, 将新数据与上一时刻分布区域的相对距离作为异常指标, 描述监测参数的统计特征变化, 辨识出设备的异常状态. 通过对仿真数据的报警效果分析, 以及将该方法应用于对加热炉风机的振动监测中, 得到的异常报警结果能够满足实际监测的需要, 证明该方法具有异常的识别敏感性、缓慢劣化包容性和状态迁移适应性的特点.

关键词: 一类支持向量机; 自适应报警; 异常状态

中图分类号: TH165.3 **文献标志码:** A **文章编号:** 0253-987X(2009)11-0061-05

Self-Adaptive Alarm Method for Equipment Condition Based on One-Class Support Vector Machine

ZHANG Qing^{1,2}, XU Guanghua^{1,2}, HUA Cheng¹, ZHANG Yizhuo¹

(1. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China; 2. State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To improve the adaptability and effectiveness of recognition on abnormal condition, a self-adaptive alarm method for equipment condition based on one-class support vector machine (OC-SVM) is proposed. The optimum distribution area of monitoring parameters in high-dimensional feature space is dynamically estimated with on-line algorithm of OC-SVM. The abnormal index, determined by the relative distance between the new data and the distribution area at the previous moment, describes the statistical feature variation of monitoring parameters and identifies the abnormal condition of the equipments. The characteristics, such as the sensitivity in abnormal condition recognition, the toleration in slow deterioration and the adaptability in condition alternation, are verified by simulation data. The present method is further applied to the vibration monitoring of heating furnace fan. The alarms under the actual abnormal condition meet the demand of equipment monitoring.

Keywords: one-class support vector machine; self-adaptive alarm; abnormal condition

目前,最常用的设备异常报警标准是国际标准化组织制定的 ISO2372 和 ISO10816,它们规定了不同类型设备不同品质级别的振动强度范围,而温度、压力、油液等监测参数也具有相应的异常阈值^[1]. 这些报警标准的共同之处是建立在试验统计基础上的固定阈值,没有充分考虑实际运行条件和现场工况,

因此在应用中会频繁出现误报警和漏报警问题. 随着研究工作的深入,模糊逻辑、神经网络等模型方法被用来实现监测报警^[2],决定其性能的关键是充足的先验知识,而获得先验知识要付出高昂的代价,同时,当设备的工况改变或者状态迁移后,原有的模型参数和结构可能已不适用,因此在适应性方面还有

待提高。

一类支持向量机^[3]是在统计学习的理论基础上提出的一种无监督学习方法,其目的是将目标类样本与其他样本进行有效区分,具有无需先验知识和结构风险最小化等优点,因此被广泛应用于异常检测^[4]和故障诊断^[5-6]中。本文设计了一类支持向量机的在线算法,动态估计监测参数的分布区域,在高维特征空间中建立异常指标并设置阈值,实现了设备状态的自适应监测报警。在高维动态特征空间中对异常指标设置的固定阈值,等价于原始数据空间对监测参数本身设置的非线性自适应阈值,因此报警问题的复杂性被大大降低。

1 理论基础

1.1 一类支持向量机

一类支持向量机的基本原理如下:假设有源于同一分布的 n 维向量样本 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l \in \mathbf{R}^n$, 通过非线性函数 ϕ 将样本映射到高维的特征空间 F 中,在 F 中寻找一个体积小而包含样本点尽可能多的区域,使其与非本类样本以最大间隔分开。以二次型方程来表示该优化问题

$$\begin{aligned} \min: & \frac{1}{2} \|\boldsymbol{\omega}\|^2 + \frac{1}{\nu l} \sum_i \xi_i - \rho \\ \text{s. t. } & (\boldsymbol{\omega} \cdot \phi(\mathbf{x}_i)) \geq \rho - \xi_i \end{aligned} \quad (1)$$

式中: $\boldsymbol{\omega}$ 为超平面法向量; ρ 为边界系数; $\xi_i \geq 0$ 为松弛因子; l 为样本点数; $0 < \nu \leq 1$ 为折衷系数,它决定分布区域的体积与包含样本点多少的折衷。该方程从空间几何的角度来看,是将样本投影到一个超球体中,寻找一个与 $\boldsymbol{\omega}$ 正交、距离球心 $\rho / \|\boldsymbol{\omega}\|$ 的超平面,使大多数样本处于被超平面分割并远离超球中心的区域,为保证该区域的结构紧凑,少数样本被排除在分布区域之外,与边界的距离为 ξ_i 。

构造拉格朗日函数求解式(1),化简后得到

$$\begin{aligned} \min: & \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s. t. } & 0 \leq \alpha_i \leq \frac{1}{\nu l} \quad \sum_{i=1}^l \alpha_i = 1 \end{aligned} \quad (2)$$

式中: α_i 为拉格朗日系数; K 为核函数; $K(\mathbf{x}_i, \mathbf{x}_j)$ 为 $\phi(\mathbf{x}_i)$ 和 $\phi(\mathbf{x}_j)$ 的内积运算,并记做 $K_{i,j}$ 。式(2)是典型的二次规划问题,求解得到的 α_i 多数等于 0,只有少部分 α_i 大于 0,将其对应的样本 $\mathbf{x}_i$ 称为支持向量。定义决策函数

$$f(\mathbf{x}) = \sum_{i=1}^l \alpha_i K(\mathbf{x}_i, \mathbf{x}) - \rho \quad (3)$$

当 $f(\mathbf{x}) > 0$ 时, $\mathbf{x}$ 位于样本的分布区域之内, $f(\mathbf{x}) = 0$ 时, $\mathbf{x}$ 位于分布区域边界上, $f(\mathbf{x}) < 0$ 时, $\mathbf{x}$ 位于样本分布区域之外。由式(3)可知, $\alpha_i = 0$ 的样本在计算中将被忽略,因此全体目标样本在特征空间中的分布仅取决于支持向量。

1.2 一类支持向量机的在线算法

支持向量机的核心运算是二次规划问题的求解,计算量为 $O(l^3)$,按照常规算法,当增加新样本或删除旧样本时,就必须重新求解二次规划问题,因计算量大,所以不适用于处理连续更新的样本集。为此, Gert 和 Tomaso^[7] 模拟物理学的绝热过程,提出了支持向量机的增量减量学习方法。本文在该方法的基础上,研究了一类支持向量机的在线算法。

决策函数将样本集分为非支持向量集 N 、边界支持向量集 M 和非边界支持向量集 V 。定义 M 中样本的下标集合为 S ,即 $\mathbf{x}_i \in M$ 时, $i \in S$ 。这 3 个向量集中的样本存在以下等式关系:当 $\mathbf{x}_i \in N$ 时, $\alpha_i = 0$;当 $\mathbf{x}_i \in V$ 时, $\alpha_i = 1/\nu l$;当 $\mathbf{x}_i \in M$ 时, $f(\mathbf{x}_i) = 0$ 。增量减量学习方法的基本原理是:当新样本 $\mathbf{x}_c$ 被加入或旧样本被删除时,通过调节拉格朗日系数 α_c 来保证决策方程和约束条件对所有样本仍然有效,并使上述等式关系保持成立。

在假定 3 个向量集之间不发生样本交换的前提下, α_c 与 M 中样本的拉格朗日系数有如下变化关系

$$\begin{bmatrix} -\Delta\rho \\ \Delta\alpha_M \end{bmatrix} = -\mathbf{Q}_M^{-1} \begin{bmatrix} 1 \\ \mathbf{K}_{M,c} \end{bmatrix} \Delta\alpha_c \quad \mathbf{Q}_M = \begin{bmatrix} 0 & \mathbf{E}' \\ \mathbf{E} & \mathbf{K}_M \end{bmatrix} \quad (4)$$

式中: $\mathbf{K}_M$ 是 M 的核函数矩阵; $\mathbf{K}_{M,c}$ 是 M 与 $\mathbf{x}_c$ 的内积矢量; $\mathbf{E}$ 是全为 1 的列向量。对于 V 和 N 中的样本, α_i 分别保持 0 和 $1/\nu l$ 不变,即 $\Delta\alpha_i = 0$,仅 $f(\mathbf{x}_i)$ 变化,其变化量与 $\Delta\alpha_c$ 的关系如下

$$\Delta f(\mathbf{x}_i) = \gamma_i \Delta\alpha_c \quad \gamma_i = \mathbf{K}_{i,c} + \sum_{j \in S} \mathbf{K}_{i,j} \beta_j + \beta_\rho \quad (5)$$

式中 β_j 和 β_ρ 的计算方法为

$$\begin{bmatrix} -\beta_\rho \\ \beta_{j \in S} \end{bmatrix} = -\mathbf{Q}_M^{-1} \begin{bmatrix} 1 \\ \mathbf{K}_{M,c} \end{bmatrix}$$

式(4)与式(5)确定了样本不在 3 个向量集间交换的条件下,调整 α_c 会引起 ρ, α_i 以及 $f(\mathbf{x}_i)$ 的变化。在实际情况, $\Delta\alpha_c$ 有可能导致数据集之间的样本交换,例如当 $\Delta\alpha_c = -f(\mathbf{x}_i)/\gamma_i$ 时,对于 V 中的样本 $\mathbf{x}_i$,由式(5)可知 $f(\mathbf{x}_i) + \Delta f(\mathbf{x}_i) = 0$,则该样本进入 M 。表 1 列出了样本在各向量集间交换所对应的 $\Delta\alpha_c$ 值,为了减少计算的复杂性, α_c 需要根据交换过程逐步变化,以保证每次只有一个样本发生交换,直

到 $\mathbf{x}_c$ 被确定属于 3 个向量集之一,同时决策方程和约束条件对所有样本仍然有效。

表 1 $\mathbf{x}_i$ 在向量集间交换时所对应的值

退出向量集	进入向量集	$\Delta\alpha_c$
$\mathbf{M}$	$\mathbf{V}$	$\frac{1/vl-\alpha_i}{\beta_i}$
$\mathbf{M}$	$\mathbf{N}$	$-\frac{\alpha_i}{\beta_i}$
$\mathbf{V}$	$\mathbf{M}$	$-\frac{f(\mathbf{x}_i)}{\gamma_i}$
$\mathbf{N}$	$\mathbf{M}$	$-\frac{f(\mathbf{x}_i)}{\gamma_i}$
$\mathbf{x}_c$ (增量)	$\mathbf{M}$	$-\frac{f(\mathbf{x}_i)}{\gamma_i}$
$\mathbf{x}_c$ (增量)	$\mathbf{V}$	$\frac{1}{vl}-\alpha_c$
$\mathbf{x}_c$ (减量)	$\mathbf{N}$	$-\alpha_c$

新样本 $\mathbf{x}_c$ 加入的增量学习过程如下:①由式(3)计算 $f(\mathbf{x}_c)$,如果 $f(\mathbf{x}_c)>0$,则 $\mathbf{x}_c \in \mathbf{N}$,整个过程结束,否则转到②;②根据表 1,计算所有样本在向量集间发生交换以及 $\mathbf{x}_c$ 进入 $\mathbf{M}$ 或 $\mathbf{V}$ 时的 $\Delta\alpha_c$,寻找最小的 $\Delta\alpha_c$,根据式(4)或式(5)进行各参数的调整,并进行对应样本的交换;③如果 $\mathbf{x}_c$ 进入 $\mathbf{M}$ 或 $\mathbf{V}$,则停止,否则转到②。

当旧样本 $\mathbf{x}_c$ 从样本集中被删除时,如果 $\mathbf{x}_c \in \mathbf{N}$,则可以被直接删除,否则就需要逐步调整 α_c ,直到 $\mathbf{x}_c$ 进入 $\mathbf{N}$,其减量学习过程的步骤与增量学习过程类似。学习过程中,如果 $\mathbf{M}$ 发生变化,式(4)中的一 Q_M^{-1} 也随之变化,为了避免重复的矩阵求逆,可以根据 Woodbury 公式进行递归计算^[7],以提高算法的实时性。在 3 个向量集样本交换的过程中,有可能出现 $\mathbf{M}$ 为空的情况,此时令 $\gamma_i=1$,根据式(5)仅更新 $f(\mathbf{x}_i)$ 。在线算法通过局部调整,在高维空间中获得全局最优的当前样本集分布区域,避免了二次规划求解,因此适合在线处理动态监测的数据。

2 自适应报警方法

监测参数的变化与设备的损伤或劣化相比,是一个相对快变的过程,这是由于在干扰因素的持续作用下,不论设备处于稳态还是非稳态,其监测参数都会持续出现随机的微小波动。对于运行状态稳定的设备,尽管监测参数波动细节的产生原因无法被

准确推知,但分布特征在概率意义上则保持稳定,设备状态一旦发生变化,监测参数的分布特征也随之改变,否则所选参数不能胜任监测工作,设备处于不可监控状态。根据监测参数分布特征与状态变化的同步性,建立了如图 1 所示的一类支持向量机的自适应报警方法,具体步骤如下。

步骤 1 使用长度为 m 的滑动窗对连续的监测参数进行分割,在 $k-1$ 时刻得到目标样本集 $\{\mathbf{x}_{k-m}, \dots, \mathbf{x}_{k-2}, \mathbf{x}_{k-1}\}$ 。

步骤 2 使用一类支持向量机的在线算法计算目标样本集的空间分布区域。

步骤 3 在 k 时刻得到新的监测参数 $\mathbf{x}_k$,根据下式计算异常指标

$$I_k = -\lg\left[\sum_{i=k-m}^{k-1} \alpha_i K(\mathbf{x}_i, \mathbf{x}_k) / \rho_{k-1}\right] \quad (6)$$

步骤 4 如果 I_k 大于报警阈值 η ,则报警。

步骤 5 $\mathbf{x}_k$ 加入目标样本集,并将 $\mathbf{x}_{k-m}$ 从样本集中删除,在线计算新的空间分布区域。

步骤 6 $k-1 \leftarrow k$,重复步骤 3~5,实现连续的在线监测报警。

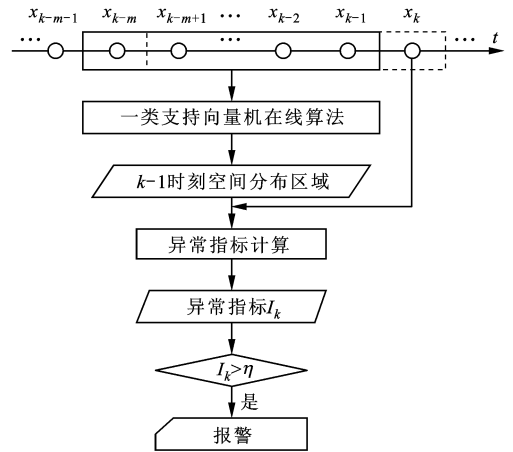


图 1 自适应报警方法的流程框图

异常指标是高维特征空间中的相对距离指标,由式(3)所示的决策函数变形得到,对其设置阈值 η ,以决定新的监测参数距离上一时刻分布区域达到何种相对程度时才需要报警。假设 $k-1$ 时刻的目标样本集满足正态分布 $N(\boldsymbol{\mu}_{k-1}, \boldsymbol{\sigma}_{k-1})$, $\boldsymbol{\mu}_{k-1}$ 和 $\boldsymbol{\sigma}_{k-1}$ 分别为正态分布的均值与方差,则 $I_k \geq \eta$ 等价于

$$\frac{\|\mathbf{x}_k - \boldsymbol{\mu}_{k-1}\|}{\boldsymbol{\sigma}_{k-1}} \geq C(K, v) \quad (7)$$

式中: $C(K, v)$ 是与一类支持向量机计算系数相关的函数。由此可知,自适应报警方法与传统的统计分析方法在本质上是相同的,其优势在于对分布区域估

计的结构风险最小化和全局最优性,因此更适于处理动态的小样本问题。

3 仿真验证与结果分析

为分析自适应报警方法对异常状态的辨识效果,使用正态分布的随机信号进行仿真验证,通过加入野点或调节正态分布的均值和方差,模拟分布稳定、稀疏异常点、剧烈突变和微弱缓变监测参数的变化情况,如图2所示。

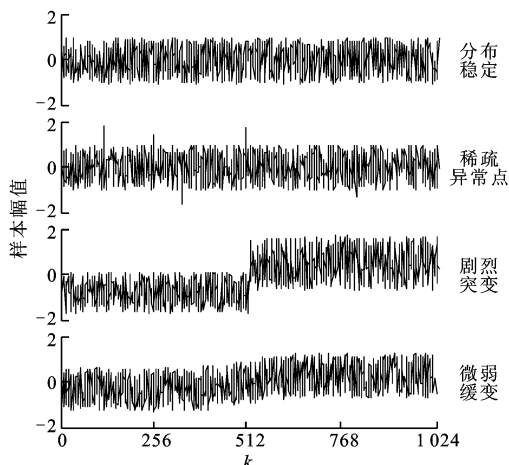


图2 仿真验证数据

取 $m=50$, $v=0.1$, 选择高斯函数为核函数, 平滑因子 $\sigma=0.6$, 使用一类支持向量机的在线算法估计滑动窗内数据的空间分布, 根据式(6)计算得到的异常指标序列如图3所示。

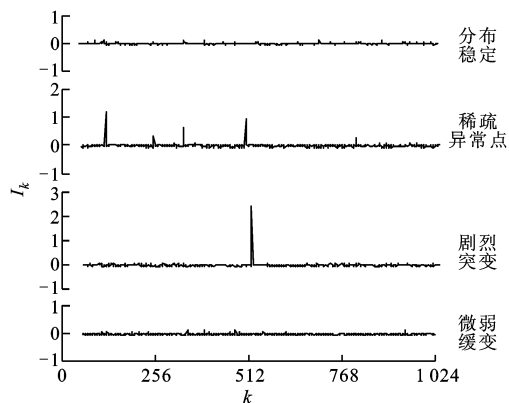


图3 仿真数据的异常指标

当数据分布稳定时, I_k 在 0 值附近微小波动, 当数据中存在稀疏的异常点时, 由于距离上一个时刻的分布区域较远, 因此 I_k 明显增大. 因为后续时刻没有新的异常点出现, 所以异常点持续停留在 V 中, 直到从滑动窗中被删除. 稀疏的异常点只会引起

对应时刻的 I_k 变化, 对后续数据点的异常识别没有影响. 图2中剧烈突变的仿真数据使临界点处的 I_k 迅速增大, 随着后续数据点持续落在旧分布区域之外, V 中的数据转移到 M 中, 分布区域随之发生改变, 最终趋于平稳并对应于新的状态, 使 I_k 重新回到 0 值附近的小范围内. 该结果表明, 自适应报警方法只在状态变化临界时产生异常报警, 状态完成迁移后则自动消除报警, 在实际应用中, 能够避免由设备工况变化或维修后引起的长期误报警现象. 分布微弱缓变的数据用来仿真设备的缓慢劣化, 由于数据分布区域的变化缓慢, 部分数据虽然超出上一时刻的分布, 但偏离程度都非常小, 同时后续数据持续改变滑动窗内目标样本的分布区域, 因此 I_k 一直保持在 0 值附近.

自适应报警方法使用一类支持向量机, 来寻求分布区域的紧凑程度与包含样本数目的最优折衷, 一部分数据点被自动归于 V , 并排除在目标类之外. 正是这种折衷的作用, 保证了该方法在对异常点和剧烈突变准确识别的前提下, 鲁棒跟踪数据的变化. 针对不同的监测对象, 通过调节 v 、 σ 和 η , 自适应报警方法能够实现对异常状态的敏感识别、对缓慢劣化的持续包容以及对状态迁移的自适应.

4 应用实例

某钢铁集团公司的中板工序加热炉 3# 风机是由电机、连轴器和鼓风机构成的, 根据旋转电机振动测定方法的国家标准, 选择振动速度信号的有效值——均方根值 R_{ms} 作为监测参数, 对电机自由端和输出端的轴承座进行了垂直方向、水平方向和轴向的定期监测. 从 2005 年 6 月 4 日起至 2006 年 12 月 5 日期间, 共采集信号 215 次, 由于维修以及检测人员的安排问题, 使信号的采集间隔长短不一, 中板工序每年 9 月为停机检修时间, 而 2005 年 9 月电机曾返厂维修, 2006 年 11 月底该设备发生事故, 检修解体后发现, 电机轴承保持架断裂, 滚动体破碎. 因此, 以电机自由端垂直方向的监测参数为例, 使用本文方法进行分析, 其监测参数和报警效果如图4所示.

选择 $m=24$, $v=0.1$, $\sigma=1$, 设置异常指标的报警阈值 $\eta=0.7$, 在表2中列出了报警时间、监测参数以及异常指标值. 对报警结果进行分析后发现, 该风机在 2005 年 9 月以前运行平稳, 振动小, 但经过返厂维修后, 出现第 1 次报警, 此后振动幅度达到一个新的水平, 于 2006 年 3 月 7 日达到顶峰, 出现了

第 2 次报警.

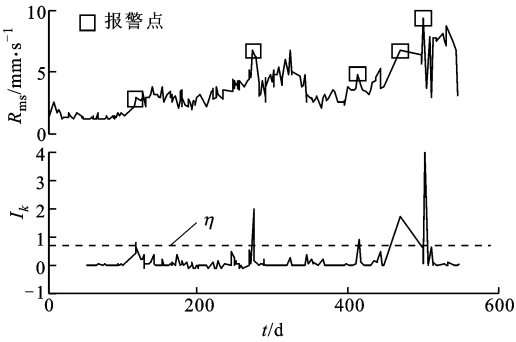


图 4 加热炉风机振动监测参数的自适应报警

表 2 风机的振动监测报警结果

报警日期	2005-09-29	2006-03-07	2006-07-24	2006-09-18	2006-10-20
$R_{ms}/\text{mm} \cdot \text{s}^{-1}$	2.8	6.8	4.8	6.7	9.3
I_k	0.779 3	1.974 1	0.899 6	1.769 5	3.975 3

该风机经过 2 次维修调整后,于 2006 年 5 月底到 7 月初其振幅下降,7 月 24 日其振动又突然增大,出现了第 3 次报警. 2006 年 9 月全工序停机检修后,设备振动再次异常,出现第 4 次报警,10 月 20 日振动异常剧烈,出现第 5 次报警,由于未进行及时处理,11 月底该风机因故障停机. 通过对电机自由端水平方向和输出端垂直、水平方向的监测参数进行自适应报警分析,得到了相同的结果. 由此判断,该风机故障的根本原因是维修不当所致. 因为电机的返厂维修破坏了该风机原始的平稳状态,随后的定期维修均无法恢复它的工作性能,并导致设备持续劣化,所以最终发生了电机轴承被损毁的严重事故. 本文方法在小样本条件下,动态跟踪风机状态的多次迁移,计算得到的报警点符合实际情况,能够满足现场监测的报警需要.

5 结 论

为了实现对设备异常状态的有效识别,本文提出了一种基于一类支持向量机的设备状态自适应报

警方法. 该方法利用在线一类支持向量机,连续求解监测参数在特征空间中的最优分布区域,通过新数据与分布区域的相对距离指标,判断异常状态的出现,实现了监测报警功能. 经仿真和试验数据验证,本文方法能够及时有效地识别异常,并且随设备状态迁移进行自适应调整,包容微弱缓慢的劣化过程,从而满足了工业现场的实际监测需求.

参考文献:

[1] 徐敏. 设备故障诊断手册 [M]. 西安:西安交通大学出版社,1998:1045-1060.

[2] FRANK P M. Residual evaluation for fault diagnosis based on adaptive fuzzy thresholds [C]// IEE Colloquium on Qualitative and Quantitative Modeling Methods for Fault Diagnosis. London, UK: IEE Press, 1995: 1-11.

[3] BERNHARD S, JOHN C P, JOHN S T, et al. Estimating the support of a high-dimensional distribution [J]. Neural Computation, 2001, 13(7): 1443-1471.

[4] DELPJINE P, PHILIPPE V, EMMANUEL D, et al. An abrupt change detection algorithm for buried landmines localization [J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(2):260-272.

[5] 胡桥,何正嘉,訾艳阳,等. 基于模糊支持矢量数据描述的早期故障智能监测诊断 [J]. 机械工程学报,2005, 41(12):145-150.

HU Qiao, HE Zhengjia, ZI Yanyang, et al. Incipient fault intelligent monitoring and diagnosis based on fuzzy support vector data description [J]. Chinese Journal of Mechanical Engineering, 2005, 41(12): 145-150.

[6] 钟清流,蔡自兴. 基于 OCSVM-CPSO 的自适应故障诊断 [J]. 计算机工程与应用,2007,43(8):18-20.

ZHONG Qinliu, CAI Zixing. Self-adaptive fault-diagnoses based on OCSVM-CPSO [J]. Computer Engineering and Applications, 2007, 43(8): 18-20.

[7] GERT C, TOMASO P. Incremental and decremental support vector machine learning [M]. Cambridge, MA,UK: MIT Press, 2001:409-415.

(编辑 管咏梅)