文章编号: 1001-0920(2009)09-1367-04

支持向量机和一类模糊推理系统的等效性及其应用

孙 坚[1], 郑恩辉[1], 邹 超[1], 刘长东[2]
(1. 中国计量学院 机电工程学院, 杭州 310018; 2. 浙江天地环保工程有限公司, 杭州 310012)

摘 要: 支持向量机(SVM)和模糊推理系统(FIS)分别源于统计学习理论(SLT)和认知学两个不同的领域. 在一定约束条件下,提出并证明了 SVM 和一类基于规则的 FIS 的函数等效性定理. 在此基础上,提出基于 SVM 学习过程的 FIS(MBFIS)的设计方法. MBFIS 继承了 SVM 良好的泛化能力和对"维数灾难"的避免能力,也继承了基于规则的 FIS 的显式推理能力. Benchmark 数据实验表明,MBFIS 具有良好的分类性能.

关键词: 支持向量机; 模糊推理系统; 平移不变核
中图分类号: TP18　　**文献标识码**: A

Functional equivalence and its application of support vector machines and rule-based fuzzy inference system

SUN Jian[1], *ZHENG En-hui*[1], *ZOU Chao*[1], *LIU Chang-dong*[2]
(1. College of Mechatronics Engineering, China Jiliang University, Hangzhou 310018, China; 2. Zhejiang Tiandi Environmental Protection Engineering Company Limited, Hangzhou 310012, China. Correspondent: ZHENG En-hui, E-mail: ehzheng@cjlu.edu.cn)

Abstract: Support vector machines (SVM) and fuzzy inference system (FIS) are motivated by statistical learning theorem (SLT) and cognizing theorem respectively. Under some restrictions, the functional equivalence between SVM and rule-based FIS is proposed. Based on the learning mechanism of SVM, the algorithm of designing a rule-based FIS, Mercer binary FIS(MBFIS), is presented. The MBFIS algorithm has the good generalization ability and can avoid the "curse of dimension", and has the transparent inference ability. Experimental results based on a few Benchmark data sets show that the MBFIS algorithm has better classification performance.

Key words: Support vector machines; Fuzzy inference system; Translation invariant kernel

1 引 言

模糊集合理论和模糊逻辑是处理复杂系统不确定性的强有力工具,是对人类认知过程的模拟,一些模糊模型已被证明是通用模拟器[1,2]. 模糊推理系统(FIS)的核心是模糊规则的提取,以此作为决策的依据. 因其显式推理能力增加了决策的透明度和可信度,FIS 已成功应用于控制系统辨识、信号和图像处理、模式分类、信息提取等众多领域[3-5]. 然而,现有的 FIS 易遭遇"维数灾难"和"过学习"等问题,模糊集的划分和规则数的确定缺乏良好的理论依据.

Vapnik 等[6]提出的支持向量机(SVM)是一类强大的机器学习算法. SVM 基于结构风险最小化(SRM)准则,折衷结构风险(模型复杂度)和经验风险,求取最优分类超平面. 由于具有坚实的理论基础和良好的泛化性能,SVM 广泛应用于故障诊断、医疗诊断、入侵检测和欺诈检测等领域. 但是,SVM 获得的知识隐藏在大量的 Lagrange 系数和支持向量中,决策过程无法为用户所理解. 与人工神经网络相似,SVM 生成一个"黑箱"模型,决策过程缺乏透明度,不具有显示推理能力,因而在安全性要求高的领域的应用受到一定的限制[7]. 如在医疗诊断方面,不仅要知道 SVM 的决策结果,而且要知道其决策过程. 文献[8,9]将 SVM 用于模糊推理系统的建模,提出了基于 SVM 建立模糊推理系统的方法. 本文受其启发,提出并证明了 SVM 和一类 FIS 的函数等效性定理.

在介绍 FIS 和 SVM 基本知识的基础上,本文

收稿日期: 2008-10-11; **修回日期**: 2009-01-21.
基金项目: 浙江省自然科学基金项目(Y1080950); 国家公益行业专项项目(2007GYJ016).
作者简介: 孙坚(1964—),男,杭州人,副教授,从事在线动态检测、计量校准等研究; 郑恩辉(1975—),男,辽宁新民人,讲师,博士,从事(代价敏感)数据挖掘、复杂系统建模与控制等研究.

基于一定约束条件,提出并证明 SVM 和一类 FIS 的函数等效性定理,进而给出基于 SVM 学习过程的 FIS 设计方法. 该 FIS 基于 SRM 准则,具有良好的泛化能力,自动确定规则数,能够避免“维数灾难”.

2 支持向量机与核策略

针对两类非线性可分问题,假定已知观测样本集

$$(x_1, y_1), \cdots, (x_i, y_i), \cdots, (x_n, y_n),$$
$$x_i \in R^l,\ y_i \in \{+1, -1\},\ i = 1, 2, \cdots, n \quad (1)$$

能被超平面 $(wx) - b = 0$ 分类. 学习问题为

$$\min R(w, \xi) = \frac{1}{2} \| w \|^2 + C\Big(\sum_{i=1}^{n} \xi_i\Big);$$
$$\text{s.t. } y_i(x_i w + b) \geq 1 - \xi_i,$$
$$\xi_i \geq 0,\ i = 1, 2, \cdots, n. \quad (2)$$

其中: $\| w \|^2 / 2$ 模拟模型的复杂度(VC 维),为结构风险; ξ_i 为松弛变量, $\sum_{i=1}^{n} \xi_i$ 模拟模型在训练样本集上的误差,为经验风险; C 为松弛因子,折衷结构风险和经验风险. 引入 Lagrange 乘子 α_i 构造目标函数的对偶形式,得到 SVM 的决策函数为

$$f(x) = \operatorname{sgn}(w^* x + b) = \operatorname{sgn}\Big(\sum_{i=1}^{s} \alpha_i^* y_i x_i x + b^*\Big). \quad (3)$$

其中: $w^* = \sum_{i=1}^{l} \alpha_i^* y_i x_i$, α_i^* 表示不为零的 Lagrange 乘子,其对应的样本称为支持向量(SV), s 表示 SV 的个数.

以上为线性情况. 对于非线性情况,依据 Cover 定理[6],输入空间 R^l 通过非线性映射 $\Psi: R^l \to H$ 映射到高维特征空间 H,即将非线性可分问题转化为线性可分问题进行求解. 此时决策函数为

$$f(x) = \operatorname{sgn}(w^* x + b) = \operatorname{sgn}\Big(\sum_{i=1}^{l} \alpha_i^* y_i K(x_i, x) + b^*\Big), \quad (4)$$

其中 $K(x_i, x) = \psi(x_i)\psi(x)$ 为核函数,在 SVM 算法中[6],要求核函数满足 Mercer 定理.

3 SVM 和一类 FIS 的函数等效性

3.1 BFIS 和 SVM 函数等效的充要条件

若模糊模型第 j 条规则形式为

$$\text{if } A_j^1 \text{ and } \cdots \text{and } A_j^n, \text{ then } b_j. \quad (5)$$

其中: A_j^k 为模糊集,其隶属度函数为 $\alpha_j^k: R \to [0, 1]$; $j = 1, 2, \cdots, m$ 为规则数; $k = 1, 2, \cdots, n$ 为输入向量的维数; $b_j \in R$. 若选取“积”为模糊结合操作,选取“加”为模糊规则聚集操作,选取面积中心法(COA)去模糊化,则该模糊模型为 TS(Takagi-Sugeno) 模糊模型的特殊形式[8]. 模型的输入输出映射 $F: R^n \to R$ 定义为

$$F(x) = \frac{\sum_{j=1}^{m} b_j \prod_{k=1}^{n} \alpha_j^k(x_k)}{\sum_{j=1}^{m} \prod_{k=1}^{n} \alpha_j^k(x_k)}. \quad (6)$$

假如式(6) 的分母为零,则增加一条规则为

$$\text{if } A_0^1 \text{ and } \cdots \text{and } A_0^n \text{ then } b_0. \quad (7)$$

其中: $b_0 \in R$; 隶属度函数 $\alpha_0^k(x_k) \equiv 1$, $j = 1, 2, \cdots, n$; $x_k \in R$. 此时输入输出映射为

$$F(x) = \frac{b_0 + \sum_{j=1}^{m} b_j \prod_{k=1}^{n} \alpha_j^k(x_k)}{1 + \sum_{j=1}^{m} \prod_{k=1}^{n} \alpha_j^k(x_k)}, \quad (8)$$

其中 $\prod_{k=1}^{n} \alpha_j^k(x_k)$ 称为规则 j 的触发强度,即决策权重,规则 0 的决策权重为 1.

定义 1[9] 设一个模糊系统包含 $m+1$ 条模糊规则,其中规则 0 有式(7) 的形式,规则 $j(j = 1, 2, \cdots, m)$ 有式(6) 的形式. 如果模糊推理系统使用“积”进行模糊结合,使用“加”进行规则聚集,使用 COA 解模糊化,则该系统定义一个二元 FIS(BFIS),其决策函数为

$$f(x) = \operatorname{sign}(F(x) + t). \quad (9)$$

其中: $F(x)$ 由式(8) 定义, $t \in R$ 为阈值.

令式(9) 中 $t = 0$,又因为式(8) 分母为正,不失一般性,式(9) 的 BFIS 重写为

$$f(x) = \operatorname{sign}\Big(\sum_{j=1}^{m} b_j \prod_{k=1}^{n} \alpha_j^k(x_k) + b_0\Big). \quad (10)$$

而 SVM 的决策函数为

$$f(x) = \operatorname{sign}\Big(\sum_{i=1}^{s} \alpha_i^* y_i K(x_i, x) + b^*\Big). \quad (11)$$

BFIS 的决策函数(10) 和 SVM 的决策函数(11) 具有形式上的等价性,均具有求和形式,且求和符号后面的表达式都具有 $az + b$ 的形式. 对于 BFIS 和 SVM, a 分别为 b_j 和 $\alpha_i^* y_i$, z 分别为 $\prod_{k=1}^{n} \alpha_j^k(x_k)$ 和 $K(x_i, x)$, b 分别为 b_0 和 b^*.

定理 1 对于给定的样本集合,如果 BFIS 和 SVM 的决策函数均具有式(10) 和(11) 的形式,则 BFIS 和 SVM 函数等效的充要条件是以下 4 个条件同时成立:1) SV 的集合和模糊规则基的集合相等; 2) 对应的系数相等,即 $\alpha_j^* y_j = b_j (j = 1, 2, \cdots, m)$; 3) SVM 的核函数和 BFIS 的触发强度相等,即 $K(x_j, x) = \prod_{k=1}^{n} \alpha_j^k(x_k)$, $j = 1, 2, \cdots, m$; 4) 对应的阈值相等,即 $b^* = b_0$.

定理 1 提供了基于 SVM 学习 BFIS 的思路:首先基于 FIS 触发强度的计算方法构造一类核函数,该类核函数的计算自动完成模糊结合操作;然后证明该类核函数满足 Mercer 定理,即该类核可用于建模 SVM 分类器;最后将基于 SVM 学习到的参数根据定理 2 引入 BFIS,即完成了基于 SVM 学习 BFIS 的过程.

3.2　核二元模糊推理系统

定义 2[9]　如果对于第 $k(k \in \{1,2,\cdots,n\})$ 个输入矢量,隶属度函数 $a_j^k: R \to [0,1], j = 1,2,\cdots,m$,由隶属度函数 a^k 通过平移变换产生,即对位置参数 $z_j^k \in R$,有

$$a_j^k(x_k) = a^k(|x_k - z_j^k|),\ a^k(0) = 1$$

成立,则该 BFIS 定义为标准 BFIS(SBFIS).

SBFIS 的隶属度函数具有平移不变性,即隶属函数的值取决于其到位置参数的距离,而与位置参数的具体数值无关. 如果将触发强度 $\prod_{k=1}^{n} a_j^k(x_k - z_j^k)$ 定义为一个核函数

$$K(x,z_j) = \prod_{k=1}^{n} a^k(x_k - z_j^k),$$

则 $K(x,z_j) = K(x - z_j)$ 成立,即该核函数具有平移不变性,称之为平移不变核.

定义 3　如果在 SBFIS 中,触发强度根据平移不变核

$$K(x,z_j) = \prod_{k=1}^{n} a^k(x_k - z_j^k)$$

计算,则该 SBFIS 定义为核 BFIS(KBFIS),其决策规则为

$$f(x) = \mathrm{sign}\Big(\sum_{j=1}^{m} b_j K(x,z_j) + b_0\Big). \qquad (12)$$

3.3　MBFIS 和 SVM 函数等效的充要条件

Mercer 核在机器学习领域得到大量研究,是一种有效的将线性机器扩展为非线性机器的方法,SVM 中要求核函数满足 Mercer 定理.

引理 1[6]　一个平移不变核 $K(x,z) = K(x - z)$ 是 Mercer 核,当且仅当其如下傅立叶变换非负:

$$F[K](w) = \frac{1}{(2\pi)^{n/2}} \int_{R^n} K(x) \mathrm{e}^{-i\langle w,x\rangle} \mathrm{d}x.$$

定义 4　若 SBFIS 中的平移不变核是一个 Mercer 核,则相应的 KBFIS 定义为 Mercer BFIS(MBFIS).

定理 2　如果对于给定的样本集合 MBFIS 和 SVM,其决策函数具有式(12) 和(11) 的形式,则 MBFIS 和 SVM 函数等效的充要条件是以下 4 个条件同时成立:1) SV 的集合和模糊规则前件位置参数的集合相等;2) 与每个 SV 对应的 Langrage 乘子和类标号的积与每个规则的后件相等,即 $a_i^* y_j = b_j (j = 1,2,\cdots,m)$;3) 触发强度计算采用的平移不变核满足 Mercer 定理,且 SVM 使用该核函数;4) 对应的阀值相等,即 $b^* = b_0$.

定理 1 与定理 2 的区别在 3),它给出了函数等效性的具体实现方法. 研究发现,根据表 1 中的隶属度函数定义的平移不变核满足 Mercer 定理. 例如,根据等腰三角形模糊隶属度函数定义的平移不变核为

$$K(x,z) = \prod_{k=1}^{n} a_j^k(x_k) = \prod_{k=1}^{n} \max(1 - |b(x_k - z_j^k)|, 0). \qquad (13)$$

令 $z_j^k = 0$,不失一般性,$K(x,z)$ 的傅立叶变换为

$$F[K](w) = \frac{1}{(2\pi)^{n/2}} \int_{R^n} K(x) \mathrm{e}^{-i\langle w,x\rangle} \mathrm{d}x = \frac{1}{(2\pi)^{n/2}} \int_{R^n} \prod_{k=1}^{n} a_j^k(x_k) \mathrm{e}^{-i w_k x_k} \mathrm{d}x = \prod_{k=1}^{n} \frac{1}{(2\pi)^{1/2}} \frac{\sin^2(w/2b)}{(w/2b)^2} > 0. \qquad (14)$$

表 1　可用于定义 Mercer 的隶属度函数

名称	表达式
等腰三角形	$\mu(x) = \max(1 - d\|x\|, 0, d > 0$
高斯	$\mu(x) = \mathrm{e}^{-dx^2},\ d > 0$
柯西	$\mu(x) = 1/(1 + dx^2),\ d > 0$
拉普拉斯	$\mu(x) = \mathrm{e}^{-d\|x\|},\ d > 0$

根据引理 1,根据等腰三角形模糊隶属度函数定义的平移不变核是一个 Mercer 核. 同理可证,根据表 1 中的其他隶属度函数定义的平移不变核也是 Mercer 核.

4　基于 SVM 学习建立 MBFIS

现有的 FIS 基于 ERM 准则,模型复杂度的确定没有较好的理论依据. 定理 2 基于一定约束给出了 MBFIS 和 SVM 函数等效性的充要条件,为建立基于 SRM 准则的 FIS 提供了一个新方法. 若给定样本集,建立 MBFIS 步骤如下:首先从定理 2 条件 3) 出发,定义基于某些隶属度函数的平移不变核,该核为 Mercer 核,其计算完成各维的模糊结合操作;然后将该核引入 SVM,基于给定的样本集学习得到 SVM 的决策函数;最后基于定理 2 的条件 1),2),3) 将 SVM 的学习结果应用到 FIS 中,得到基于 SRM 准则的 MBFIS.

基于上述步骤建立的 MBFIS 具有以下良好特性:

1) 需要的先验信息只有参数 C 和隶属度函数的选择.

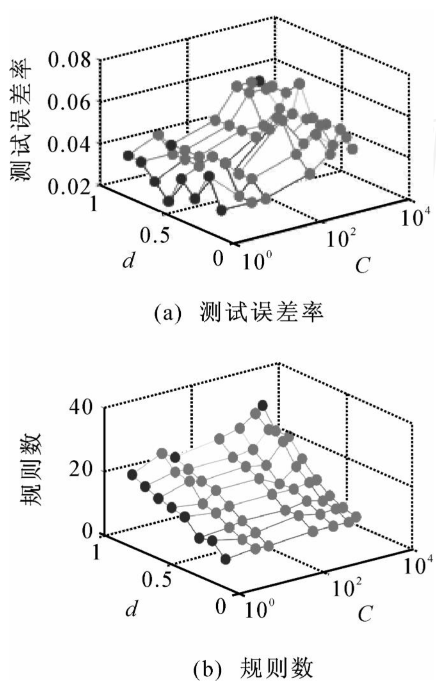
2) 自动产生模糊规则集,且模糊规则的数量与输入空间的维数无关,等于 SV 的数量,在这个意义上,避免了“维数灾难”. 另外,由于 SV 的稀疏性,模糊规则数少于训练样本数.

3) 每条模糊规则被参数化为一个训练样本 (x_j, y_j) 和相应的非零 Lagrange 乘子 α_j,其中:x_j 为前件隶属函数的位置,$y_j\alpha_j$ 为后件常数.

4) 基于 SRM 准则而不是 ERM 准则,可控制模型的复杂度和经验风险的折衷,具有良好的泛化性能.

5 实验结果与分析

基于一些 Benchmark 数据集[10] 实验研究 MBFIS 的分类性能,实验结果取 20 次实验的平均值. 图 1 给出了 MBFIS 在 Iris 数据集上训练误差率、测试误差率和规则数随高斯参考函数参数 d 与 SVM 参数 C 的变化. 可见,随着 C 和 d 的增加,训练误差率减小,测试误差率增加,规则数减少.



(a) 测试误差率

(b) 规则数

图 1 C 和 d 对 MBFIS 在 Iris 数据集上分类性能的影响

表 2 不同参考函数的 MBFIS 在 Iris 上的测试误差率

	高斯	柯西	拉普拉斯	等腰三角形
规则数	18	16	20	23
测试误差率	0.0267	0.0332	0.0316	0.0370

表 2 给出了 MBFIS 使用不同隶属度函数时在 Iris 数据集上的规则数和测试误差率. 可见,高斯隶属度函数的 MBFIS 获得了最好的性能.

表 3 给出了 MBFIS 使用高斯参考函数时在一些 Benchmark 数据集上测试误差率和几种经典分类算法测试误差率的比较,在其中 4 个数据集上获得了最好测试性能,说明 MBFIS 具有良好的分类性能.

表 3 MBFIS 和几种经典分类算法测试误差率比较

数据集	C4.5	k-N	Bayes	BP-N	RBF	MBFIS
Iris	0.081	0.064	0.067	0.033	0.041	**0.027**
Waveform	**0.102**	0.137	0.183	0.163	0.147	0.134
Adult	**0.089**	0.091	0.104	0.133	0.197	0.102
DNA	0.076	0.146	0.068	0.088	**0.041**	**0.041**
Australian	0.099	--	0.136	**0.087**	0.107	0.108
Diabetes	0.131	0.324	0.239	0.198	0.218	**0.127**
Segment	0.040	0.077	0.265	**0.054**	0.069	0.072
Satimage	0.150	0.094	0.287	0.139	0.121	**0.091**
Shuttle	**0.100**	0.440	0.450	0.430	0.140	0.118
Vehicle	0.266	0.275	0.558	**0.207**	0.307	0.343

6 结 论

基于一定约束,本文提出并证明了 SVM 和一类模糊推理系统(MBFIS)具有函数等效性,进而提出了利用 SVM 的学习过程建立 MBFIS 的方法,并基于一些 Benchmark 数据集进行了实验研究. 结论如下:

1) 当 FIS 的隶属度函数取表 1 中的函数时,MBFIS 系统触发强度的计算定义了一个平移不变核,该核是 Mercer 核,可用于 SVM 学习;

2) 在一定约束条件下,SVM 和一类模糊推理系统(MBFIS)具有函数等效性,可基于 SVM 的学习过程建立 MBFIS;

3) 提出的 MBFIS 基于 SRM 准则,折衷结构风险和经验风险,具有良好的泛化能力;

4) MBFIS 自动确定规则数,规则数等于 SVM 学习中 SV 个数,而与样本的维数无关,因而避免了“维数灾难”.

本文是针对分类问题完成了 SVM 和一类模糊推理系统的函数等效性研究,而针对回归问题进行类似研究,并将研究结果应用于温度场的回归建模是进一步要研究的内容.

参考文献(References)

[1] Mitra S, De R K, Pal S K. Knowledge-based fuzzy MLP for classification and rule generation[J]. IEEE Trans on Neural Networks, 1997, 8(6): 1338-1350.

[2] Yen J, Wang L, Gillepsie C W. Improving the Interpretability of TSK fuzzy models by combining global learning and local learning[J]. IEEE Trans on Fuzzy Systems, 1998, 6(4): 530-537.

[3] Tomoharu Nakashimaa, Gerald Schaeferb, Yasuyuki Yokotaa. A weighted fuzzy classifier and its application to image processing tasks[J]. Fuzzy Sets and Systems, 2007, 158 (3): 284-294.

(下转第 1376 页)

方法的有效性.

参考文献(References)

[1] Rodrigues M A, Odloak D. MPC for stable linear systems with model uncertainty[J]. Automatica, 2003, 39(3): 569-583.

[2] Kothare M V, Balakrishnan V, Morari M. Robust constrained model predictive control using linear matrix inequalities[J]. Automatica, 1996, 32(10): 1361-1379.

[3] Kouvaritakis B, Rossiter J A, Schuurmans J. Efficient robust predictive control[J]. IEEE Trans on Automatic Control, 2000, 45(8): 1545-1549.

[4] Lee Y I, Kouvaritakis B. Constrained robust model predictive control based on periodic invariance [J]. Automatica, 2006, 42(12): 2175-2181.

[5] Feng L, Wang L J, Poh E K. Improved robust model predictive control with structured uncertainty[J]. J of Process Control, 2007, 17(7): 683-688.

[6] Andrey Y, Rolf F, Christian E, et al. Model predictive control of linear continuous time singular systems subject to input constraints[C]. IEEE Conf on Decision and Control. Atlantls, 2004, 2047-2051.

[7] Zhang L Q, Huang B. Robust model predictive control of singular systems [J]. IEEE Trans on Automatic Control, 2004, 49(6): 1000-1006.

[8] Ch Perez T, Goodwin G C. Stochastic output feedback model predictive control [C]. Proc of the American Control Conf. Piscataway: IEEE Press, 2001: 2412-2417.

[9] Xu Lee S M, Park Ju H. Output feedback model predictive control for LPV systems using parameter dependent Lyapunov function[J]. Applied Mathematics and Computation, 2007, 190(2): 671-676.

[10] 韩春艳, 刘晓华. 基于输出反馈的不确定连续系统的鲁棒预测控制[J]. 信息与控制, 2006, 35(6): 721-725.
(Han C Y, Liu X H. Robust predictive control for uncertain continuous-time system based on output feedback[J]. Information and Control, 2006, 35(6): 721-725.)

[11] 俞立. 鲁棒控制——线性矩阵不等式处理方法[M]. 北京: 清华大学出版社, 2002.
(Yu L. Robust control — Using linear matrix inequalities[M]. Beijing: Tsinghua University Press, 2002.)

[12] Esfahani S H, Petersen I R. An LMI approach to the output-feedback guaranteed cost control for uncertain time-delay systems [C]. Proc of the IEEE Conf on Decision and Control. Piscataway: IEEE Press, 1998: 1358-1363.

[13] Xu S Y, Paul V D, Stefan R, et al. Robust stabilization for singular systems with state delay and parameter uncertainty[J]. IEEE Trans on Automatic Control, 2002, 47(7): 1122-1128.

(上接第 1370 页)

[4] Lina C J, Chungb I F, Chenc C H. An entropy-based quantum neuro-fuzzy inference system for classification applications[J]. Neurocomputing, 2007, 70 (13/15): 2502-2516.

[5] 吕红丽, 贾磊, 王雷, 等. 基于模糊线性化预测模型的 HVAC 系统温度控制[J]. 控制与决策, 2006, 21(12): 1412-1416.
(Lv H L, Jia L, Wang L, et al. Model predictive control based on fuzzy linearization technique for temperature control of HVAC systems[J]. Control and Decision, 2006, 21(12): 1412-1416.)

[6] Burges C J C. A tutorial on support vector machines for pattern recognition[J]. Knowledge Discovery and Data Mining, 1998, 2(2): 121-167.

[7] Zhang Ying, Su Hong-ye, Chu Jian. Rules extraction from trained support vector machines [J]. Lecture Notes in Artificial Intelligence, 2005, 3518: 61-70.

[8] Chiang J H, Hao P Y. Support vector learning mechanism for fuzzy rule-based modeling: A new approach[J]. IEEE Trans on Fuzzy Systems, 2004, 12(1): 1-12.

[9] Chen Y X, Wang J Z. Support vector learning for fuzzy rule-based classification systems[J]. IEEE Trans on Fuzzy Systems, 2003, 11(6): 716-728.

[10] Merz C J, Murphy P M. UCI repository for machine learning data-Bases[EB/OL]. http://www.ics.uci.edu/ ,~mlearn/MLRepository.html, 1998.