

文章编号:1005-0523(2015)06-0121-07

PSO-SVM 算法在智能建筑环境监控系统中的应用

傅军栋, 邹欢, 康水华

(华东交通大学电气与电子工程学院, 江西 南昌 330013)

摘要:针对智能建筑环境监控系统中,多个传感器独立工作可能会造成系统误判的问题,提出了一种基于粒子群优化支持向量机(PSO-SVM)的环境质量综合评价模型,即利用粒子群算法快速优化支持向量机的惩罚参数和核函数参数,然后采用 ZigBee 无线传感网络采集的环境数据对 PSO-SVM 分类模型进行训练和测试。实验结果表明 PSO-SVM 分类器对环境质量判断的平均精度达到 94.44%,且分类结果稳定。将这种方法应用于智能建筑环境监控系统中,可以增加系统监测数据的准确性,提高系统工作的可靠性。

关键词:环境监控;多传感器;粒子群优化;支持向量机

中图分类号:TM715

文献标志码:A

DOI:10.16749/j.cnki.jecjtu.2016.01.018

智能建筑环境监控系统需要实时监测建筑内的各种环境变量,并控制相应的电气设备对环境进行调节,使用户的生活更加安全、舒适和智能。与传统的建筑环境监控系统相比,智能建筑环境监控系统在安装、维护以及节能方面都具有明显的优势^[1-2]。

由于传感器相对独立地工作,采用单个传感器的数据来评价环境质量具有一定的不确定性,因此如何快速准确地对环境信息进行融合是智能建筑环境监控系统的一项重要任务。王超采用了模糊综合评判的算法对多传感器信息进行加权融合处理,需设计各项影响因素对环境质量等级的隶属函数,这种方法能够减少单一传感器的局限性^[3];周廷刚采用了遗传算法训练神经网络权重的方法建立了 GA-BP 模型,克服了单纯的 BP 算法容易陷入局部极小的缺点^[4];陈祖云采用了支持向量机对环境质量进行评价,该模型具有较高的识别能力和鲁棒性,且算法简单,训练样本少^[5]。支持向量机通过结构风险最小化原理来提高泛化能力,可以较好地解决小样本、非线性、高维数、局部极小点等实际问题,可用于模式分类和非线性回归,适用于环境质量评价中。

由于支持向量机核函数的构造及其参数的选择对分类效果有较大的影响,而传统的参数选择一般是通过交叉验证的方法,需要大量的训练样本,优化过程非常耗时,且很难精确找到最优参数^[6],粒子群优化算法简便易行,收敛快速,可以减少试算的盲目性和提高预测精度,因此本文采用粒子群优化算法优化支持向量机的惩罚参数及核函数参数^[7],使优化后的支持向量机能够更好地进行环境质量的判断。

1 智能建筑环境监控系统构成

所设计的智能建筑环境监控系统如图 1 所示,采用 Zigbee 技术将建筑内环境及各控制设备组成无线网络系统,以 ARM 嵌入式网关^[8]为信息控制中心,对传感器获得的环境参数进行算法处理得出环境质量的等级,再对相应的电气设备进行控制从而达到环境自动调节的作用。系统包括 3 大模块:环境信息的采集和调控模块、信息管理中心模块以及上位机模块。

收稿日期:2015-08-12

作者简介:傅军栋(1972—),男,副教授,研究方向为电力系统、建筑电气及智能化。

其中,家庭网关作为整个系统的核心,对环境数据进行处理的过程如图2所示,读取环境数据后进行环境质量等级的判断,2002年国家环保部颁布了我国《室内空气质量标准》^[9],对影响室内空气的各污染物程度作了等级的划分:一级表示室内环境良好,二级表示室内环境可以保证大众健康,三级表示能保护员工健康,基本能居住或办公的室内环境。本系统根据此划分标准把环境质量简化为3个等级:质量在二级以上的为“安全”:代表环境适宜人类生活,人员舒适性良好;质量在二级到三级之间的为“一般”:代表环境不会对人的健康造成伤害,但会影响人体的舒适性;质量三级以下为“危险”:代表环境参数超过了健康标准的要求,将对人的健康造成伤害。当判断环境等级为“安全”时,处理器驱动指示灯显示为绿灯;当判断环境等级为“一般”时,显示为黄灯,并根据用户设置判断是否进行智能调整;当判断环境等级为“危险”时,显示为红灯,服务器将自动启动消防报警设施,包括自动灭火设施、防烟排烟设施、自动报警设施、应急照明设施等。

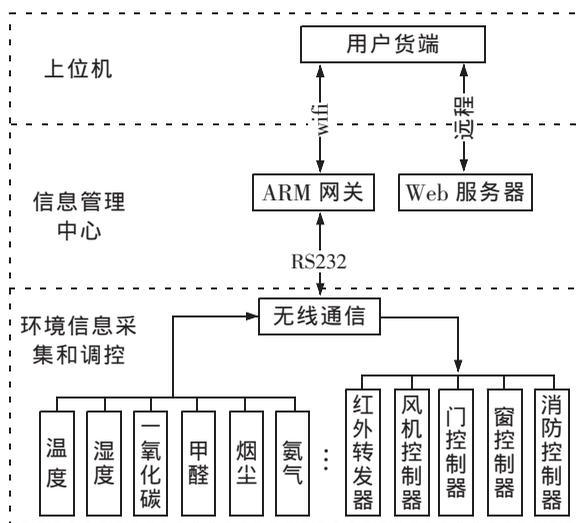


图1 智能建筑环境监控系统构成

Fig.1 Structure of intelligent building environment monitoring system

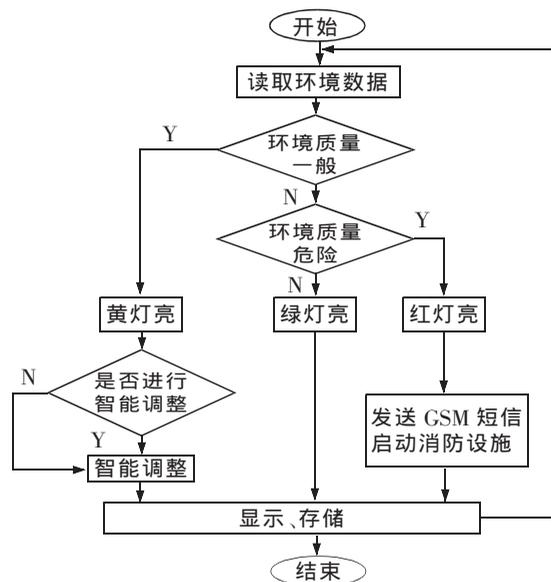


图2 处理环境数据流程图

Fig.2 Flow diagram of environment data processing

智能建筑环境监控系统的实现需要依靠多个传感器协同工作,因此多个传感器之间的信息融合对用户端的判决起到了关键性的作用。本文结合环境数据非线性、小样本、多维等特点,充分利用支持向量机的结构风险最小化与粒子群算法快速全局优化的特点,建立 PSO-SVM 的多传感器信息融合的数学模型。

2 基于 PSO-SVM 的多传感器信息融合模型

2.1 SVM 分类器

支持向量机(support vector machines, SVM)^[10-11]基本思想是通过非线性映射函数把数据样本映射到高维特征空间,再在高维特征空间中求得最优分类面来分离训练样本点,使得训练样本点与最优分离面距离最大化。基本原理如下:给定样本集 $T=\{(x_i, y_i) | i=1, 2, 3, \dots, m\}$, 其中 x_i 为输入矢量, y_i 为对应的期望输出, m 为样本数。引入松弛变量 ξ_i , 用于衡量实际指示值与支持向量机输出之间的距离。则最优分类超平面可以描述为

$$\left. \begin{aligned} \min & \frac{\|w\|^2}{2} + c \sum_{i=1}^m \xi_i, (i=1, 2, 3, \dots, m) \\ \text{s.t.} & \left\{ \begin{aligned} y_i(w \cdot x_i + b) & \geq 1 - \xi_i \\ \xi_i & \geq 0 \end{aligned} \right. \end{aligned} \right\} \circ \quad (1)$$

式中: w 为权值向量; $\|w\|^2$ 为 w 的欧拉范数; c 为惩罚因子,用于实现对错分样本惩罚程度的控制; b 为偏差; $w x_i$ 表示内积。引入拉格朗日系数 a_i 将上述优化问题转化为二次规划的优化问题。由于环境数据是非线性的,因此要将原输入空间的样本映射到高维的特征空间中,再构造最优分类超平面。引入满足 Mercer 条件的核函数 $k(x_i, x_j)=\varphi(x_i)\varphi(x_j)$ 来代替内积运算,最终得到分类决策函数:

$$f(x)=\text{sign}\left(\sum_{i=1}^m a_i y_i k(x_i, x_j)+b\right) \quad (2)$$

本文选用具有全局收敛性且适应性强的径向基核函数(其中 g 指核宽度):

$$k(x_i, x_j)=\exp\left(-\frac{\|x_i-x_j\|}{2g^2}\right) \quad (3)$$

2.2 粒子群优化算法

粒子群优化算法(particle swarm optimization, PSO)^[12]思想来源对鸟群觅食行为的模仿,通过群体中个体之间的协作和信息共享来寻找最优解,原理如下:设在一个 D 维的搜索空间内,由 n 个粒子组成的种群 $X=(X_1, X_2, \dots, X_n)$,根据目标函数可计算出每个粒子位置 X_i 对应的适应度值。每次迭代过程中,粒子通过个体极值和群体极值来更新自身的速度和位置:

$$V_{id}^{k+1}=\rho V_{id}^k+c_1 r_1(P_{id}^k-X_{id}^k)+c_2 r_2(P_{gd}^k-X_{gd}^k) \quad (4)$$

$$X_{id}^{k+1}=X_{id}^k+V_{id}^{k+1} \quad (5)$$

式中: ρ 为惯性权重; $d=1, 2, \dots, D$; $i=1, 2, \dots, n$; k 为当前迭代次数; V_{id} 为粒子的速度; c_1 和 c_2 为加速度因子; r_1 和 r_2 是分布在 $[0, 1]$ 区间的随机数。

2.3 基于 PSO 的 SVM 参数优化

PSO 优化过程中,自适应粒子根据实际情况,通过动态平衡其全局搜索与局部搜索能力,能快速准确地找到支持向量机的最优参数,从而有效进行识别。通过优化得到 SVM 误差最小的惩罚参数 c_{best} 和核函数半径参数 g_{best} ,使优化后的 SVM 能够更好地进行分类判断^[13]。过程如图 3 所示。

步骤 1:初始化,随机产生一组粒子的初始位置和速度,即确定 SVM 参数,选用一定输入样本建立 SVM 模型;

步骤 2:根据 SVM 分类决策函数确定其函数值,进行粒子适应度值的计算及评估;

步骤 3:根据适应度值更新个体极值 c 和全局极值 g ,按式(4)和(5)更新粒子的速度和位置;

步骤 4:寻找全局最优参数,若满足终止条件,结束寻优;若不满足终止条件,则迭代搜索更新粒子的速度和位置,转至步骤 2;

步骤 5:得到使 SVM 误差最小的惩罚参数 c 和核函数半径参数 g ,用于 SVM 训练和分类预测。

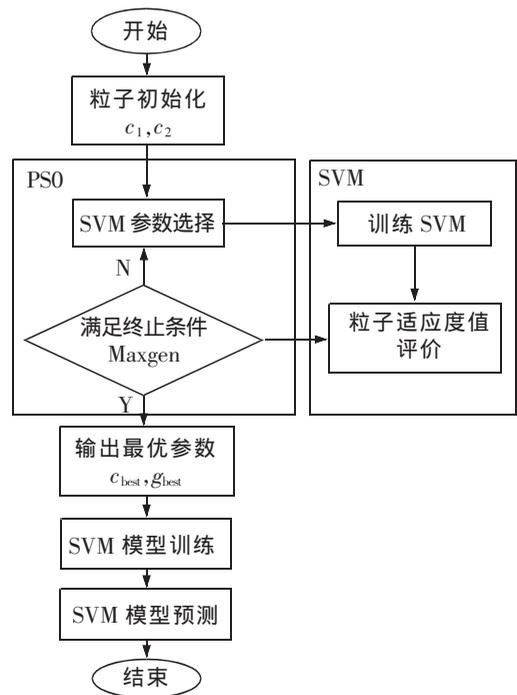


图 3 基于 PSO 的 SVM 参数优化过程
Fig.3 Process of optimizing the SVM parameters with PSO

3 实验及结果分析

3.1 数据来源及特征提取

考虑影响人体舒适度和健康的主要环境因素,包括温度、湿度、CO、烟尘、甲醛、氨气 6 种环境因素,为了增加对比性,30 组实验监测数据引自文献[14],如表 1 所示,其中环境质量等级为“1”表示安全,环境质量

等级为“2”表示一般,环境质量等级为“3”表示危险。

表1 环境数据

Tab.1 Environment data

序号	温度/°C	相对湿度/%	CO/(mg·m ⁻³)	烟尘/(mg·m ⁻³)	甲醛/(mg·m ⁻³)	氨气/(mg·m ⁻³)	环境质量等级
1	22.3	55.6	10	0.17	0.19	0.11	3
2	23.1	38.7	8.8	0.14	0.09	0.21	1
3	16.5	58.1	9.4	0.16	0.08	0.22	2
4	21.2	30.8	8.9	0.22	0.2	0.15	2
5	22.4	60.3	9.1	0.16	0.12	0.16	1
6	26.5	72.1	11.2	0.17	0.16	0.31	3
7	16.7	49.2	7.7	0.19	0.07	0.22	2
8	22.8	59.2	9.2	0.18	0.16	0.39	2
9	21.8	61.5	9.3	0.19	0.25	0.19	2
10	23.9	48.7	7.5	0.22	0.15	0.13	1
11	16.8	62.4	9.4	0.24	0.23	0.1	3
12	18.4	61.2	9.6	0.23	0.19	0.23	1
13	21.5	49.6	8.7	0.21	0.16	0.24	1
14	28.8	34.6	8.7	0.21	0.22	0.31	3
15	19.7	47.6	8.7	0.26	0.04	0.22	1
16	27.6	59.5	10.1	0.26	0.22	0.14	3
17	17.9	61.4	10.5	0.34	0.05	0.18	3
18	25.7	66.7	9.5	0.28	0.07	0.31	3
19	18.6	56.2	10.3	0.32	0.18	0.11	2
20	18.7	48.6	8.2	0.27	0.16	0.35	3
21	20.4	62.5	10.4	0.38	0.07	0.12	3
22	15.4	54.3	9.8	0.37	0.21	0.24	3
23	29.2	36.5	11.3	0.31	0.18	0.21	3
24	29.4	34.6	10.9	0.37	0.25	0.35	3
25	19.3	39.5	10.1	0.44	0.11	0.24	3
26	18.2	66.6	10.3	0.41	0.22	0.33	3
27	21.5	38.7	8.9	0.39	0.13	0.16	2
28	23.4	58.7	8.8	0.37	0.23	0.29	3
29	24.6	53.1	8.7	0.41	0.18	0.22	3
30	17.5	62.4	8.4	0.31	0.2	0.23	3

由于每种环境因素的影响大小不同,为了更好地进行训练和识别,需要进行归一化处理^[15],采用线性函数转换:

$$P'=(P-P_{\min})/(P_{\max}-P_{\min})。 \quad (6)$$

式中: P 为原始数据; P' 为归一化后的数据。

3.2 构造 PSO-SVM 分类器

将得到的特征样本数据按照一定比例,随机分为训练样本集和测试样本集,为了使训练的 SVM 分类器有更好的分类性能,训练样本和测试样本数据需要进行合理的分配^[16]。分析表 1 中的数据,可知在 30 组数据中,环境质量等级为 1 的有 6 组,等级为 2 的有 7 组,等级为 3 的有 17 组,选择训练样本时 3 种等级的数据都应该包括,保证训练样本的规模要求,兼顾分类器的泛化能力,还需考虑参数优化及分类识别时的计算成

本,如从环境质量等级为 1 的 6 组数据中随机选取 5 组作为训练样本,其余组作为测试样本。最终选择结果如表 2 所示,共选取了 21 组数据进行样本训练,其余 9 组数据作为样本测试,训练集用于对 SVM 分类器进行训练,测试集则用于输入 SVM 分类器进行分类测试,得到环境质量的判断结果。若结果与测试目标一致,说明判断正确,否则判断错误。

表 2 样本训练集和测试集的选取

Tab.2 Sample selection of training set and testing set

环境等级	对应的样本组数	训练组数	测试组数
1	6	5	1
2	7	5	2
3	17	11	6
总和	30	21	9

构造 PSO-SVM 分类器时还需要设置 PSO 算法的初始参数^[17],惯性权重的选择非常关键,较大的惯性权重有较好的全局搜索能力,而较小的惯性权重有较强的局部搜索能力,因此随着迭代次数的增加,惯性权重应不断减少,从而使得粒子群算法在初期具有较强的全局收敛能力,而晚期具有较强的局部收敛能力,本文惯性权重采用 $p_i=0.9-0.5i/\text{Maxnum}$ 的计算方法,随迭代次数由 0.9 线性递减到 0.4;学习因子 c_1 较大时,会使粒子过多地在局部范围徘徊,而学习因子 c_2 较大时会促使粒子过早收敛到局部最小值,为了平衡随机因素的作用,学习因子设置为 $c_1=c_2=2$;粒子规模设置为 20,最大迭代次数初设为 100;为防止粒子的盲目搜索,将其位置和速度限制在一定的区间内, c 的取值范围为[0.1, 1 000], g 的取值范围为[0.1, 100]。采用 MATLAB 编写 PSO-SVM 非线性拟合程序。

3.3 结果与讨论

把选好的 21 组训练数据输入模型进行训练,按照图 3 的过程进行迭代,当 PSO 满足迭代终止条件,优化过程结束,得到惩罚参数和核函数半径参数的最优值 c_{best} 和 g_{best} 。如图 4 所示,随着迭代次数的增加,实际适应度(准确率)逐渐趋近于最佳参数的适应度,当进化到 22 代左右时就可以找到 SVM 的最佳参数,输出的最优参数为 $c_{\text{best}}=17.208 8, g_{\text{best}}=3.906 4$ 。为了结果更加准确,进行 10 次测试,SVM 参数 c, g 和准确率如表 3 所示,可以看出训练集分类准确率平均达到 94.74%,寻找最佳参数的准确率高。

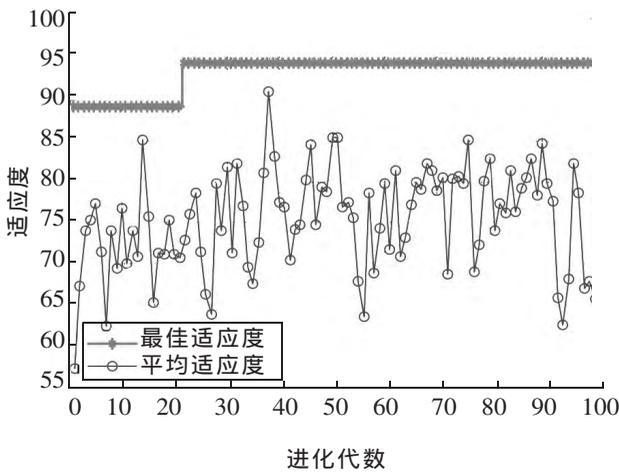


图 4 SVM 训练集分类精度

Fig.4 Classification accuracy of SVM training set

表 3 PSO-SVM 参数优化及其训练集分类精度

Tab.3 Training set classification accuracy and optimal parameters of PSO-SVM classifiers

次数	最优参数 c	最优参数 g	分类精度/%
1	3.363	3.007 3	94.74
2	3.575 8	3.582 2	94.74
3	12.256 7	4.350 2	94.74
4	100	4.342 1	94.74
5	5.177 1	3.125 1	94.74
6	85.814 7	3.953 2	94.74
7	14.143 9	5.426 5	89.47
8	10.686	1.830 8	94.74
9	2.078	4.462 1	94.74
10	90.113 5	3.235 2	94.74

通过 PSO-SVM 得到的模型对 9 个预测检验样本的预测值与实际值对比,经过 10 次测试发现准确率大多能达到 100%和 88.889%,图 5 是准确率为 100%时的测试效果图,图 6 是准确率为 88.889%时的测试效果图,可以看出准确率为 88.889%时,只有一个分类错误,环境质量等级 2 预测成等级 1,对人员安全不会产生很大的影响,总体来说,系统的可靠性比较高。

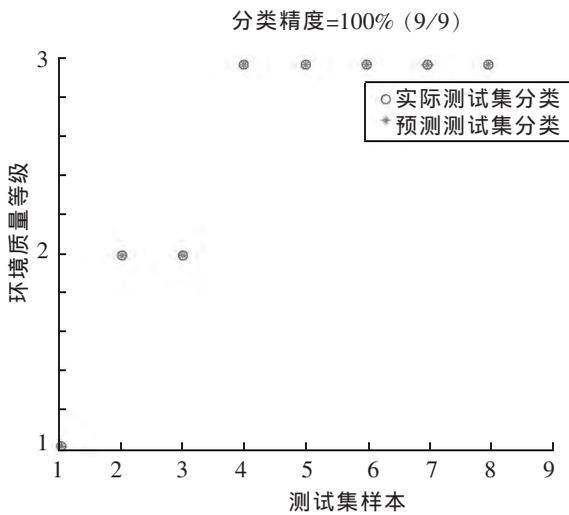


图5 测试集准确率为100%时的分类效果图

Fig.5 Effect diagram when testing set accuracy is 100%

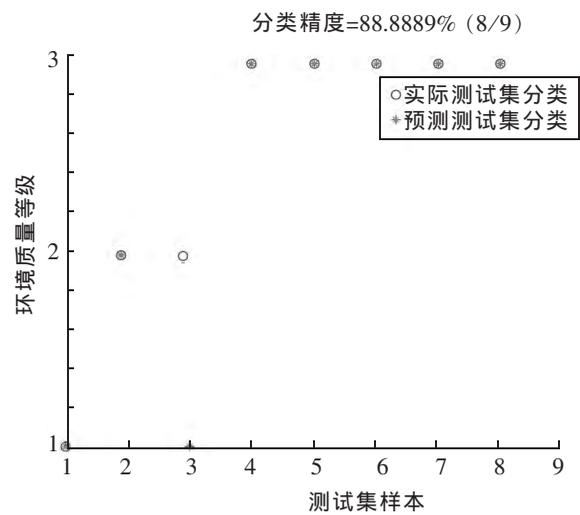


图6 测试集准确率为88.889%时的分类效果图

Fig.6 Effect diagram when testing set accuracy is 88.889%

为了进一步说明 PSO-SVM 模型的优越性,和未进行参数优化的 SVM 以及常用的增量型极限学习机 (increased extreme learning machine, I-ELM)、BP 神经元算法的分类结果进行对比,分别统计各模型运行 10 次的准确率,统计结果如表 4 所示,可以看出 PSO-SVM 算法平均分类精度最高,平均达到 94.44%,并且在 10 次运行中的分类准确率比较稳定,未经参数优化的 SVM 模型准确率相对偏低,I-ELM 和 BP 算法的模型平均分类精度分别为 83.33%和 80.00%,且波动较大。因此 PSO-SVM 模型应用于智能建筑环境监控系统中,在准确度和稳定性方面有一定的优势。

表 4 测试集分类准确率比较

Tab.4 Comparison of testing set classification accuracy

算法次数	PSO-SVM	SVM	I-ELM	BP
1	100	88.89	88.89	77.78
2	100	100	88.89	88.89
3	88.89	88.89	66.67	88.89
4	100	88.89	77.78	77.78
5	100	77.78	88.89	88.89
6	88.89	88.89	66.67	88.89
7	77.78	88.89	88.89	77.78
8	100	100	77.78	55.56
9	100	88.89	100	88.89
10	88.89	77.78	88.89	66.67
平均准确率	94.44	88.89	83.33	80.00

4 结语

结合支持向量机的结构风险最小化与粒子群算法快速全局优化的特点,建立基于 PSO-SVM 算法的环境质量评价模型。试验结果表明,通过 PSO 优化的 SVM 分类器具有较高的分类精度,PSO-SVM 模型相比其他常用算法在决策准确度和模型稳定性方面都具有一定的优势,应用于智能建筑环境监控系统中,有效地解决了多传感器信息融合的问题,提高了系统的可靠性。

参考文献:

- [1] 辛海亮,钟佩思,朱绍琦,等. 基于 ZigBee 的物联网智能家居控制系统[J]. 电子技术应用, 2013, 39(12): 79-82.
- [2] 朱伟兴,戴陈云,黄鹏. 基于物联网的保育猪舍环境监控系统[J]. 农业工程学报, 2012(11): 101-103.
- [3] 王超,肖童心,陈楠,等. 基于模糊数据融合理论智能家居环境监测系统研究[J]. 安全, 2014(7): 5-8.
- [4] 周廷刚. 基于遗传神经网络模型的大气环境质量评价方法[J]. 四川环境, 2003(3): 73-75.
- [5] 陈祖云,金波,邬长福. 支持向量机在环境空气质量评价中的应用[J]. 环境科学与技术, 2012(S1): 395-398.
- [6] 陆荣秀. 混合核函数支持向量机在系统建模中的应用[J]. 华东交通大学学报, 2010, 27(2): 63-66.
- [7] MAALI Y, Al-JUMAILY A. A novel partially connected cooperative parallel PSO-SVM algorithm: study based on sleep apnea detection[C]//Proceedings of the 2012 IEEE Congress on Evolutionary Computation (CEC), Brisbane, Australia, 2012: 1-8.
- [8] 李坚强,何穗强,明仲. 基于智能网关的数字家庭系统研究与设计[J]. 深圳大学学报:理工版, 2014(6): 630-637.
- [9] 中国疾病预防控制中心.GB/T 18883-2002 室内空气质量标准[S].北京: 中国标准出版社, 2003.
- [10] 彭春华,刘刚,孙惠娟. 基于小波分解和微分进化支持向量机的风电场风速预测[J]. 电力自动化设备, 2012, 32(1): 9-13.
- [11] 衷路生,齐叶鹏,杨辉,等. PSO 优化的 LS-SVM 在列车弓网系统的建模研究[J]. 华东交通大学学报, 2012, 29(3): 1-6.
- [12] KENNEDY J, EBERHART R C. Particle swarm optimization[C]//Proceedings of the IEEE Conference on Neural Network, Perth, Australia, 1995, 1942-1948.
- [13] 兰秀菊,张丽霞,鲁建厦,等. 基于小波分析和 PSO-SVM 的控制图混合模式识别[J]. 浙江工业大学学报, 2012, 40(5): 532-536.
- [14] 李檀. 建筑内环境监控仪设计与研究[D]. 沈阳: 沈阳建筑大学, 2011.
- [15] 费胜巍,苗玉彬,刘成良,等. 基于粒子群优化支持向量机的变压器故障诊断[J]. 高电压技术, 2009(3): 509-513.
- [16] 赵敏华,苏夔,徐立先. 基于 PSO-SVM 的 PMV 指标预测系统研究[J]. 建筑热能通风空调, 2015(3): 73-76.

Application of PSO-SVM Algorithm in Environmental Monitoring System of Intelligent Building

Fu Jundong, Zou Huan, Kang Shuihua

(School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: In environmental monitoring system of intelligent buildings, independent work of multiple sensors may cause misjudgment. Aiming at this problem, this paper proposes an environmental quality comprehensive evaluation model for optimizing the Support Vector Machine (SVM) by using the Particle Swarm Optimization (PSO) algorithm, in which the SVM parameters of the punishment and the kernel function are optimized by PSO and the established SVM classifiers are trained and tested by using the sample data collected by ZigBee wireless sensor networks. The experimental results show that the average recognition rate of the PSO-SVM classifier is up to 94.44% in evaluating environmental quality, and the classification results are stable. It suggests that the proposed method increase the accuracy of monitoring data and improve working reliability for the environmental monitoring system of intelligent buildings.

Key words: environmental monitoring; multiple sensor; particle swarm optimization (PSO); support vector machine(SVM)

(责任编辑 刘棉玲)