

基于遗传算法优化支持向量机的交通流量预测

李晓斌

(华南理工大学 自动化科学与工程学院, 广东 广州 510641)

摘 要: 为了得到性能优越的 SVM 预测模型, 实现城市交通流量的准确预测, 文中提出基于遗传算法优化支持向量机(GA-SVM)的城市交通流量预测方法. 其中通过遗传算法对 SVM 中的训练参数进行优化处理, 以得到优化的 SVM 预测模型. 实验结果表明: 用 GA-SVM 对城市交通流量预测, 预测精度远优于人工神经网络.

关键词: 支持向量机; 遗传算法; 城市交通流量; 预测模型

中图分类号: U491.1

文献标识码: A

文章编号: 1000-7180(2010)10-0186-03

Forecasting Urban Traffic Flow Based on Support Vector Machine Optimized by Genetic Algorithm

LI Xiao-bin

(Institute of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China)

Abstract: In order to gain the excellent SVM forecasting model and realize the accurate forecasting of urban traffic flow, support vector machine optimized by genetic algorithm is presented to forecast urban traffic flow. Genetic algorithm (GA) is introduced to optimize the parameters of support vector machine in this model, which can gain optimized SVM forecasting model. The experimental results indicate that the proposed GA-SVM model has better forecasting accuracy than artificial neural networks.

Key words: support vector machine; genetic algorithm; urban traffic flow; forecasting model

1 引言

交通流量预测作为智能交通系统的基础, 其准确预测是实现科学规范的交通规划、交通诱导和交通控制的重要基础^[1]. 由于交通流量存在随机性、复杂性和不确定性, 难以建立合适的数学模型, 因而人工智能方法在交通流量预测越来越受重视^[2]. 目前人工神经网络在交通流量预测应用较多^[3-4], 如 BP 神经网络、RBF 神经网络. BP 神经网络存在训练数据需求大、收敛速度慢等问题, 而 RBF 神经网络克服了 BP 神经网络训练数据需求大、收敛速度太慢等缺点, 日渐取代 BP 神经网络成为一种新的交通流量预测方法. 然而, 由于人工神经网络基于经验最小优化原理, 容易陷入局部极值, 从而影响了其泛化能力.

近年来, 支持向量机(SVM)在时序预测中应用越来越广泛, 它采用结构风险最小化原则, 能有效解决小样本、非线性等回归问题, 泛化能力强, 并能找到全局最优解, 克服了神经网络局部极值的难题^[5-8]. 然而, 支持向量机中训练参数的选择对其预测性能有很大影响. 鉴于遗传算法具有隐含的并行性和强大全局搜索能力, 文中提出遗传算法优化支持向量机(GA-SVM)城市交通流量预测方法, 其中将遗传算法(GA)用于确定 SVM 中的训练参数.

2 SVM 回归原理

给定训练样本集 $T = \{x_i, y_i\}_{i=1}^m$, 其中 m 为训练样本数目, x_i 为输入矢量, y_i 为输出值. 在 SVM 中, 利用非线性函数 φ 将输入空间数据 x 映射到高维特征空间并在该空间用函数 $f(x) = (w \cdot \varphi(x))$

+ b 进行线性回归, 其中, w 为权值矢量, b 为偏差^[7].

引入非负松弛变量 ξ 和 ξ^* , 用于度量不敏感带外的训练样本的偏离程度 ϵ , 则 SVM 优化的目标函数为

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) \quad (1) \\ \text{s. t.} \quad & \begin{cases} y_i - w \cdot \varphi(x) - b \leq \epsilon + \xi_i, & \xi_i \geq 0 \\ w \cdot \varphi(x) + b - y_i \leq \epsilon + \xi_i^*, & \xi_i^* \geq 0 \end{cases} \end{aligned}$$

式中, ϵ 为不敏感损失函数参数, C 为惩罚参数, 用于控制对超出误差的样本的惩罚程度. 引入拉格朗日乘子 α_i 和 α_i^* , 式(1) 优化问题转化为对偶问题

$$\begin{aligned} \max \quad & \sum_{i=1}^m y_i (\alpha_i - \alpha_i^*) - \epsilon \sum_{i=1}^m (\alpha_i + \alpha_i^*) \quad (2) \\ & - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(x_i, x_j) \\ \text{s. t.} \quad & \sum_{i=1}^m (\alpha_i - \alpha_i^*) = 0, C \geq \alpha_i, \alpha_i^* \geq 0 \end{aligned}$$

式中, $K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$ 为核函数, 核函数的引入使得高维空间中的内积运算可以通过输入空间中的函数来实现, 避免了维数灾难^[8], 常用的核函数有径向基核函数, 多项式核函数等, 鉴于径向基核函数只有一个变量需确定, 相比其它核函数需确定的变量较少, 因此径向基核函数用于构造文中的 SVM 回归模型. 最后获得的 SVM 回归函数为

$$f(x) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (3)$$

3 基于 GA - SVM 交通流量预测建模

3.1 建立训练样本集

利用训练样本建立 SVM 回归模型, 其描述为 $y_t = f(a_t, a_{t-1}, \dots, a_{t-k})$, 其中 $x_t = a_t, a_{t-1}, \dots, a_{t-k}$ 为输入矢量, $y_t = a_{t+1}$ 为输出值, k 为输入矢量的维数.

3.2 SVM 参数优化

惩罚参数 C 、RBF 的核参数 σ 、不敏感损失参数 ϵ 对 SVM 回归模型的预测效果影响很大, 为了获取较高预测性能的 SVM 回归模型, 则需对 C 、 σ 、 ϵ 进行组合优化. 优化过程如图 1 所示.

由于遗传算法具有隐含的并行性和强大全局搜索能力, 可以在很短的时间内搜索到全局最优点. 因此, 文中利用遗传算法对这些参数进行优化, 具体步骤为:

(1) SVM 参数编码与 GA 参数设置. 将 SVM

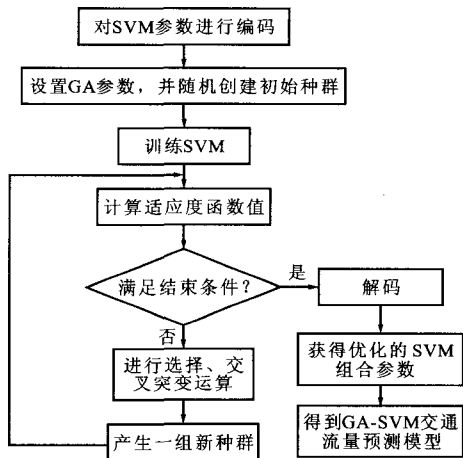


图 1 GA - SVM 交通流量预测模型优化过程

训练参数 σ 、 C 、 ϵ 以进行染色体基因编码. 进行 GA 参数设置, 如交叉率, 变异率等. 随机产生一组表示 SVM 参数值的染色体.

(2) 计算种群中各个个体的适应度. 采用平均相对误差作为适应度函数来衡量选取参数的优劣.

(3) 根据个体适应度, 通过轮盘赌法, 将具有较高适应度值的染色体被复制; 并通过单点交叉法将两条染色体中的基因进行随机的交换, 交叉率为 0.8; 变异是指将染色体中的基因编码“0”与“1”进行互换, 变异率设定为 0.02.

(4) 重复步骤(2)~(4), 使初始确定的一组支持向量机参数不断进化, 直到满足结束条件.

(5) 对得到的最优解译码, 得到优化的参数 σ 、 C 、 ϵ .

4 应用实例

以文献[9]中某交叉路口 12:45~14:40 中记录的 24 组交通流量数据(每隔 5 分钟记录一次)作为实验数据, 12:45~14:20 期间记录的数据为训练数据, 14:25~14:40 期间记录的数据为测试数据, 如表 1 所示. 归一化试验数据, 取输入矢量的维数为 5, 构建训练样本集, 如式(4)所示. 并利用遗传算法优化 SVM 参数, 遗传算法的计算参数为: 交叉率和变异率分别为 0.8 和 0.02, 种群数为 20, σ 、 C 、 ϵ 取值范围分别为: $[0, 1]$ 、 $[1, 100]$ 、 $[0.0001, 0.01]$. 获得优化的 SVM 参数: $\sigma = 0.479$, $C = 10.06$, $\epsilon = 0.0039$. 分别采用 BP 神经网络、RBF 神经网络与 GA - SVM 进行对比分析, 其中 BP 神经网络由 5 个输入节点、10 个隐节点和 1 个输出节点组成. 采用平均相对误差 (MAPE) 评价城市交通流量预测精

度.这三种预测方法的拟合与预测的曲线如图 2 所示.

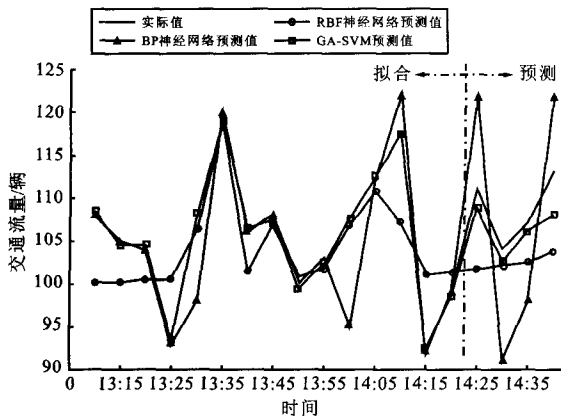


图 2 GA-SVM、RBF 神经网络、BP 神经网络的预测值与实际值对比曲线图

如表 2 所示,GA-SVM、RBF 神经网络、BP 神经网络这三种预测模型的预测精度分别为:2.13%、5.63%、9.54%,可以看出在交通流量预测中,GA-

SVM 的预测精度远高于 RBF 神经网络、BP 神经网络.

$$X = \begin{bmatrix} a_{1245} & a_{1250} & a_{1255} & a_{1300} & a_{1305} \\ a_{1250} & a_{1255} & a_{1300} & a_{1305} & a_{1310} \\ & & \dots & & \\ a_{1355} & a_{1400} & a_{1405} & a_{1410} & a_{1415} \end{bmatrix}$$
$$Y = \begin{bmatrix} a_{1310} \\ a_{1315} \\ \dots \\ a_{1420} \end{bmatrix} \tag{4}$$

式中 X 中为输入矢量,Y 为输出值.

5 结束语

文中提出遗传算法优化支持向量机城市交通流量预测方法,其中将遗传算法(GA)用于确定 SVM 中的训练参数.实验结果表明:用 GA-SVM 对城市交通流量预测,预测精度远优于人工神经网络.

表 1 国内某交叉路口 12:45~14:40 中记录的 24 组交通流量数据 (单位:辆)

时间	交通流量	备注	时间	交通流量	备注	时间	交通流量	备注
12:45	107	训练	13:25	93	训练	14:05	112	训练
12:50	71	训练	13:30	98	训练	14:10	122	训练
12:55	86	训练	13:35	120	训练	14:15	92	训练
13:00	80	训练	13:40	106	训练	14:20	99	训练
13:05	98	训练	13:45	108	训练	14:25	111	预测
13:10	108	训练	13:50	100	训练	14:30	104	预测
13:15	105	训练	13:55	103	训练	14:35	107	预测
13:20	104	训练	14:00	95	训练	14:40	113	预测

表 2 GA-SVM 与 RBF 神经网络、BP 神经网络预测结果比较

时间	14:25	14:30	14:35	14:40	MAPE/%
实际值	111	104	107	113	
GA-SVM 预测值	108.8252	102.5703	106.0631	108.1262	2.13
相对误差/%	1.96	1.38	0.87	4.31	
RBF 神经网络预测值	101.7480	102.1116	102.5166	103.7537	5.63
相对误差/%	8.33	1.82	4.18	8.18	
BP 神经网络预测值	121.5340	91.0498	97.9172	121.7084	9.54
相对误差/%	9.49	12.46	8.48	7.71	

参考文献:

[1] 犹勇. 组合预测方法在城市交通流量预测中的应用[J]. 重庆工商大学学报:自然科学版, 2007, 24(2): 122 - 125.

[2] 朱文兴,龙艳萍,贾磊. 基于 RBF 神经网络的交通流量

预测算法[J]. 山东大学学报:工学版, 2007, 37(4): 23 - 27.

[3] 申慧,刘知贵,李春菊. 基于 BP 神经网络的交通流量预测设计[J]. 西南科技大学学报, 2008, 23(2): 72 - 75.

(下转第 192 页)

4 结束语

基于 $0.18\mu\text{m}$ SiGe BiCMOS 工艺,采用传输线构造谐振回路,文中在调谐范围、功耗、相位噪声等方面做了适当的折中,设计了一个交叉耦合差分 VCO,调谐范围为 $58.377\text{GHz}\sim 60.365\text{GHz}$. 仿真结果表明该 VCO 拥有较好的性能,在工作电压为 2.2V ,偏置电流为 11mA 的条件下,在 60.365GHz 和 10MHz 频偏处的相位噪声为 -99.77dBc/Hz , FOM 为 -161.5 .

参考文献:

- [1] Guofu Niu, Zhenrong Jin, John D Cressler. Transistor noise in SiGe HBT RF technology[J]. IEEE JSSC, SEPTEMBER, 2001, 36(9):1424-1427.
- [2] Nakamura T, Hitachi, Kokubunji. A 59GHz push-push VCO with 13.9GHz tuning range using loop-ground transmission line for a full-band 60GHz transceiver[C]// ISSCC, SESSION 29 mm-WAVE CIRCUITS. Beijing, 2009.
- [3] Zhao Yan, Wang Zhigong, Li Wei, et al. Design and measurement of a 53 GHz balanced colpitts oscillator[J]. Journal of semiconductor, 2009, 30(1):1-4.
- [4] Rael J J, Abidi A A. Physical process of phase noise in differential LC oscillator[C]// IEEE 2000 CICC. USA: Florida, 2000:569-572.
- [5] 徐跃, 杨英强. 输入输出调谐的低相位噪声 CMOS 压控振荡器[J]. 微电子学与计算机, 2009, 26(6):50-53.
- [6] 杨丰林, 闵昊, 沈绪榜. CMOS 电感电容压控振荡器中对称噪声滤波技术的研究[J]. 微电子学与计算机, 2007, 24(5):41-44.
- [7] Veenstra H, Notten M G M. 60GHz VCOs with transmission line resonator in a $0.25\mu\text{m}$ SiGe BiCMOS technology[C]// IEEE Radio Frequency Integrated Circuits Symposium. Atlanta, 2008:119-122.
- [8] Ja-Yol Lee, Sang-Heung Lee, Haecheon Kim, et al. A 45-to-60-GHz two-band SiGe:C VCO for millimeter-wave applications[C]// IEEE Radio Frequency Integrated Circuits Symposium. Hawaii, Honolulu, 2007:709-712.
- [9] Jacques Verdier, José Cruz Nune Pérez, Christian Gontrand. Analysis and design of 20GHz VCOs using cross-coupled differential pair and balanced colpitts topologies in SiGe:C BiCMOS technology[C]// International Conference on Design & Technology of Integrated Systems. Nanoscale Era, 2008.

作者简介:

王 东 男, (1984-), 硕士研究生. 研究方向为射频模拟集成电路设计.

吴玉平 男, (1968-), 研究员, 硕士生导师. 研究方向为混合信号集成电路设计技术与方法学.

陈 岚 女, (1968-), 研究员, 博士生导师. 研究方向为超深亚微米芯片设计方法学及计算机体系结构.

(上接第 188 页)

- [4] 张九跃, 焦玉栋. 基于 RBF 神经网络的短时交通流量预测[J]. 山东交通学院学报, 2008, 16(3):32-36.
- [5] Johan Colliez, Franck Dufrenois, Denis Hamad. Optic flow estimation by support vector regression[J]. Engineering Applications of Artificial Intelligence, 2006, 19(7):761-768.
- [6] Kuan-Yu Chen. Forecasting systems reliability based on support vector regression with genetic algorithms[J]. Reliability Engineering & System Safety, 2007, 92(4):423-432.
- [7] 甄彤, 范艳峰. 基于支持向量机的企业信用风险评估研

究[J]. 微电子学与计算机, 2003, 23(5):136-139.

- [8] 单莘, 朱永宣, 郭军. 基于支持向量机的网络告警预测知识发现[J]. 微电子学与计算机, 2007, 24(6):35-37.

- [9] 胡佩峰. 交通流量短时预测方法研究[D]. 北京: 北京交通大学, 2006.

作者简介:

李晓斌 男, (1971-), 博士研究生. 研究方向为智能交通、交通诱导.