

# 一种新参数优化算法及其在流量预测中的应用

邓宗强<sup>1</sup>, 曾碧卿<sup>2</sup>

(1. 华南师范大学 计算机学院, 广东 广州 510631;

2. 华南师范大学南海校区 信息工程与技术系, 广东 佛山 528225)

**摘要:** 为了提高网络流量预测的精度, 提出先进行小波变换后利用 LS-SVM 的网络流量预测模型, 对于 LS-SVM 参数的优化, 提出一种基于模拟退火算法的自适应混沌量子粒子群算法 (AS-QPSO)。该算法在量子粒子群算法的基础上加入了自适应和混沌特性, 使算法具有动态自适应性, 改善算法的全局寻优能力, 再引入模拟退火算法避免陷入局部最优, 使算法具有更好的收敛性和稳定性。实验结果表明: 与其他算法优化的 LS-SVM 模型相比, 该模型具有较好的泛化能力、更高的预测精度以及很好的稳定性。

**关键词:** 量子粒子群算法; 参数优化; 小波变换; 最小二乘支持向量机; 流量预测

**中图分类号:** TP301.6

**文献标识码:** A

**文章编号:** 1673-629X(2013)10-0036-05

**doi:** 10.3969/j.issn.1673-629X.2013.10.009

## A New Parameter Optimization Algorithm and Its Application in Traffic Prediction

DENG Zong-qiang<sup>1</sup>, ZENG Bi-qing<sup>2</sup>

(1. College of Computer, South China Normal University, Guangzhou 510631, China;

2. Department of Information Engineering and Technology, Nanhai Campus of South China Normal University, Foshan 528225, China)

**Abstract:** For improving the prediction accuracy of network traffic, a new network traffic prediction model is proposed based on wavelet transform and optimized LS-SVM. To optimize the parameters of LS-SVM, a kind of adaptive chaos quantum-behaved particle swarm optimization based on simulated annealing algorithm (AS-QPSO) is proposed. The algorithm joins adaptive and chaotic characteristics based on QPSO, making it dynamic adaption, improving the capacity of the global optimization. Then the simulated annealing algorithm is introduced to avoid falling into local optimum, the algorithm has better convergence and stability. Experimental results show that compared with other algorithm optimized LS-SVM model, the proposed model is more efficient with higher precision, better generalization performance and stability.

**Key words:** quantum-behaved particle swarm optimization; parameters optimization; wavelet transformation; least squares support vector machines; traffic prediction

## 0 引言

随着 Internet 的迅速发展, 网络流量时间序列已经成为一个非线性、多时间尺度变化的动力系统。寻求新的预测模型已成为研究当前网络流量行为的一个热点。而 Suykens J. A. K 提出最小二乘支持向量机法<sup>[1]</sup>, 目前已有学者将其用于时间序列预测<sup>[2-3]</sup>。但他们只采用单一的 LS-SVM 模型还不足以完全描述网络流量的复杂特性, 同时在参数选取上都缺乏合适

的方法。在预测精度上还不能取得满意的结果。

参数优化选择对模型预测精度有决定性作用。然而到目前为止, 还没有指导最小二乘支持向量机参数选择的有效方法。针对参数选取问题, 一般采用交叉检验或网格搜索方法<sup>[4]</sup>, 这些方法搜索速度慢、运算量大、耗内存多, 具有一定的盲目性。张颖璐<sup>[5]</sup>提出基于遗传算法优化支持向量机的网络流量预测, 遗传算法比较容易陷入局部极值<sup>[6]</sup>。考虑到粒子群算法比较容

收稿日期: 2012-12-30

修回日期: 2013-04-03

网络出版时间: 2013-07-24

基金项目: 教育部科学研究青年基金项目 (10YJC870044); 广东省自然科学基金资助项目 (8151063101000040)

作者简介: 邓宗强 (1988-), 男, 湖南邵阳人, 硕士研究生, CCF 会员, 研究方向为认知网络、智能算法; 曾碧卿, 教授, 硕士生导师, 博士, 研究方向为网络计算与分布式处理、认知网络。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20130724.1005.044.html>

易实现的优点,姚全珠等提出了基于 PSO 的 LS-SVM 特征选择与参数优化算法<sup>[7]</sup>,Zhou 等<sup>[8]</sup>提出将量子粒子群算法(QPSO)应用到支持向量机的参数选择中,但是 QPSO 存在过早收敛缺点。

对此文中提出一种先进行小波变换后利用 LS-SVM 的网络流量预测模型,同时提出基于模拟退火算法的自适应混沌量子粒子群算法(AS-QPSO),并将提出的 AS-QPSO 算法应用于 LS-SVM 参数的优化上。文中对实际的网络流量数据进行实验,与一些现有的优化参数算法比较,证明了 AS-QPSO 算法具有动态自适应性,避免陷入局部最优,具有更好的收敛性和稳定性,同时提高了预测的精确性。为解决参数选取问题提供了一条有效途径。

## 1 最小二乘支持向量机的原理

最小二乘支持向量机利用等式约束代替标准支持向量机的不等式约束,且它们的损失函数是不同的,如下是其推导过程:给定  $M$  训练样本  $\{x_i, y_i\}_{i=1}^M$ , 式中,  $x_i \in R^m$  为  $M$  维的训练样本输入;  $y_i \in R$  为训练样本输出。

LS-SVM 算法的损失函数为:

$$\min \{ J(W, \xi) = \frac{1}{2} W^T W + \frac{C}{2} \sum_{i=1}^m \xi_i^2 \} \quad (1)$$

约束条件:

$$y_i = W^T g(x_i) + b + \xi_i, i = 1, 2, \dots, M \quad (2)$$

式中,  $J$  为以  $W$  为自变量的优化函数;  $C$  为惩罚系数,来调节训练误差和模型复杂度,使之具有较好的泛化能力,引入拉格朗函数对其进行求解。

$$L(W, b, a, \xi) = J(W, \xi) - \sum_{i=1}^m a_i [W^T g(x_i) + b + \xi_i - y_i] \quad (3)$$

根据 Karush-Kuhn-Tucker(KKT) 最优化条件,拉格朗日函数分别对  $W, \xi, b, a$  求导可得

$$\begin{cases} \frac{\partial L}{\partial W} = 0 \\ \frac{\partial L}{\partial b} = 0 \\ \frac{\partial L}{\partial d} = 0 \\ \frac{\partial L}{\partial \xi} = 0 \end{cases} \Rightarrow \begin{cases} W = \sum_{i=1}^m a_i g(x_i) \\ \sum_{i=1}^m a_i = 0 \\ a_i = C \xi_i \\ W^T g(x_i) + b + \xi_i - y_i = 0 \end{cases} \quad (4)$$

通过消去式(4)中的  $W, \xi$ , 把优化问题转化为求解如下线性方程组:

$$\begin{pmatrix} 0 & I_m^T \\ I_m & \Omega + C^{-1} I_m \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

通过求解式(4)可得  $b = \frac{I_m^T (\Omega + C^{-1} I_m)^{-1} y}{I_m^T (\Omega + C^{-1} I_m)^{-1} I_m}, a =$

$$(\Omega + C^{-1} I_m)^{-1} (y - I_m b)$$

用于函数估计的 LS-SVM 为:

$$y(x) = \sum_{i=1}^m a_i K(x, x_i) + b \quad (5)$$

文中采用 RBF 核函数:

$$K(x_i, x_j) = \exp \left[ -\frac{(x_i - x_j)^2}{2\sigma^2} \right]$$

RBF 核函数是具有全局收敛特性的线性学习算法前馈网络,它学习速度快,并且是应用最广泛的核函数。另外,采用 RBF 核函数的 LS-SVM 需确定核宽度  $\sigma$  和正则化参数  $C$  两个参数,它们是 LS-SVM 模型的 2 个重要参数,它们的选取直接影响着模型的泛化能力和学习能力<sup>[9-10]</sup>。对此,文中提出 AS-QPSO 算法,优化选择 LS-SVM 模型中的  $\sigma$  和  $C$  参数,以提高模型的泛化性能。

## 2 新型改进的 QPSO 算法:AS-QPSO 算法

### 2.1 LS-SVM 参数和 QPSO 算法

LS-SVM 的参数对预测结果好坏起着决定性的作用,参数  $C, \sigma$  的确定十分重要。文献[8, 11] 分别利用 GA 和 PSO 对 LS-SVM 参数进行优化,减少了算法的计算量,但 GA 和 PSO 算法需设定的参数比较多,且容易陷入局部最优。QPSO<sup>[12]</sup> 具有强大全局搜索能力但也会陷入局部最优点,文中采用改进的量子粒子群算法(AS-QPSO)对  $C, \sigma$  进行寻优。

QPSO 算法只需用位置向量来描述粒子的状态,按照下面三个公式更新位置信息

$$m_{\text{best}}(k+1) = \frac{1}{M} \sum_{i=1}^M P_i(k) = \left( \frac{1}{M} \sum_{i=1}^M P_{i1}(k), \dots, \frac{1}{M} \sum_{i=1}^M P_{id}(k) \right) \quad (6)$$

$$P_{ij}(k+1) = (\varphi_1 * P_{ij}(k) + \varphi_2 * P_{gj}(k)) / (\varphi_1 + \varphi_2) \quad (7)$$

$$X_{ij}(k+1) = \begin{cases} P_{ij}(k+1) + \beta(K) * |m_{\text{best}}(k+1) - X_{ij}(k)| * \ln\left(\frac{1}{u}\right), \text{rand}(0, 1) \leq 0.5 \\ P_{ij}(k+1) - \beta(K) * |m_{\text{best}}(k+1) - X_{ij}(k)| * \ln\left(\frac{1}{u}\right), \text{rand}(0, 1) > 0.5 \end{cases} \quad (8)$$

$$\beta(k) = 1.0 - 0.5 * (k / \text{MAXITER}) \quad (9)$$

其中, MAXITER 为最大迭代次数;  $m_{\text{best}}$  为粒子群平均最优位置;  $P_{ij}$  为第  $i$  个个体在第  $j$  维上的最佳位置;  $M$  为粒子群规模;  $\varphi_1, \varphi_2, u$  均为  $(0, 1)$  之间的随机数;  $\beta$  为 QPSO 算法的惯性权值,用它来控制算法收敛的快慢,一般  $\beta(k)$  按式(9) 进行线性变化,粒子搜索

范围则会在进化后期变得很小,甚至停滞。

## 2.2 $\beta(k)$ 更新公式的改进-动态自适应型

为了解决上面提到的这一问题,文中首先对  $\beta(k)$  更新公式进行改进。提出一种自适应更新公式,算法中的  $\beta(k)$  更新公式如下:

$$\beta = \beta_0 - \beta_1 * V_d + \beta_2 * J_d \quad (10)$$

在粒子进化过程中,算法中的收缩扩张系数  $\beta$  取值是随着进化次数的增加而线性减少的,但是真正的进化过程中,它却不是线性变化的而是复杂的。仅仅用原先策略来控制  $\beta$  值会使算法出现早熟,陷入局优。如果按照式(9), $\beta$  是不能表现粒子真正进化的过程的,会使算法收敛表现过慢。而在算法中引入了量子粒子群进化速度因子  $V_d$  和聚集度因子  $J_d$ ,同时将收缩扩张系数  $\beta$  表示为式(10)的话,则算法可以动态的调整  $\beta$  值<sup>[11]</sup>,这样在每次迭代过程中具有动态的适应性了。

$V_d$  为当前的全局最佳适应度值和上一次迭代的全局最佳适应度值的比值,  $0 < V_d < 1$ , 这样的参数  $V_d$  既与上次的全局适应度值有关,考虑在迭代进化中的粒子群的进化历史,同时它又与当前的全局最佳适应度值有关,表示了算法的进化速度的快慢。 $V_d$  值的大小关系到进化速度的快慢,它的值越小表示进化速度越快,当达到 1 时,则算法寻找到了最优解;进化速度因子  $V_d$  值变大,则进化速度变慢了,即粒子的搜索范围要随之缩小, $\beta$  值要变小。

$$V_d = \text{fit\_gen\_asacqps}(i) / \text{fit\_gen\_asacqps}(i-1) \quad (11)$$

$J_d$  的大小关系到算法中的粒子群的聚集性和多样性。它的值越大表示算法中的粒子群的聚集性越大,算法中的粒子群多样性越小,当达到 1 时,算法中所有粒子完全具有同一聚集性,则非常容易陷入局部最优解。 $\beta = \beta_0 - \beta_1 * V_d + \beta_2 * J_d$  ( $\beta_0$  为  $\beta$  的初始值,  $\beta_1$  为  $V_d$  作用下的权重,  $\beta_2$  为  $J_d$  作用下的权重)。QPSO 算法中的惯性权重  $\beta$  直接关系到整个算法的收敛性,它的值大,收敛的速度快,且能利于全局搜索,却难得到精确的解;它的值小,则能得到精确的解,且能利于局部搜索,但是收敛速度就比较慢。

$$J_d = \text{global\_fitness} / \text{avgfitness\_gen\_asacqps}(i) \quad (12)$$

## 2.3 融入模拟退火算法-避免陷入局部最优

提出的 AS-QPSO 算法引入模拟退火算法不仅能够克服基本粒子群算法陷入局部寻优的缺点,而且算法的收敛性和稳定性都很好。在 AS-QPSO 算法中,SA 算法用来处理每个粒子如下:  $\Delta = \text{Fitness}(P_{\text{best}}) - \text{Fitness}(G_{\text{best}})$ , 如果  $\Delta < 0$ , 以概率赋  $G_{\text{best}} = P_{\text{best}}$ ; 如果  $\Delta \geq 0$ , 以概率  $\text{prob} = \exp(-\Delta/\text{temp})$  赋  $G_{\text{best}} = P_{\text{best}}$ ,  $t_n =$

$\alpha * t_{n-1}$ 。  $P_{\text{best}}$  是粒子  $i$  最好的位置,  $G_{\text{best}}$  是所有粒子最好的位置,  $\text{temp}$  是当前的温度。随着温度的降低,赋值的概率最差逐渐趋近于零,SA 使可能跳出局部最优搜索全局最优解,当终止条件满足时,该算法将停止。初始温度  $t_0 = 3$ , 终止温度  $t_f = 0.01$ ,  $\alpha = 0.9$ 。

## 2.4 $\beta$ 加入混沌特性-改善算法的全局寻优

网络流量存在混沌特性,利用混沌变量的随机性、遍历性和规律性优化搜索,可保持群体多样性,使算法跳出局部最优,以改善算法的全局寻优能力。通过混沌映射迭代产生混沌变量,再将混沌变量值域放大到原粒子取值空间,从而产生新的变量,继续迭代优化。随机性使搜索避免陷入局部最优,如果它的遍历性能能够控制得当,最终能以任意精度逼近真实的最优解。文中用到混沌方程是 Logistic 方程,  $\mu$  为控制参量,当  $\mu = 4$ ,  $0 \leq R_i(t) \leq 1$ , Logistic 完全处于混沌状态。

$$\begin{cases} R_i = \text{rand}(1, \text{asacqps\_option.dimension}); i = 1, 2 \\ R_j = 4 * R_i * (1 - R_i); j = 1, 2 \\ p = (R_1 * \text{local\_x}(j, :) + R_2 * \text{global\_x}) / (R_1 + R_2) \end{cases} \quad (13)$$

利用混沌特性更新  $R$ , 再计算  $P$ , 根据混沌运动遍历的特性,将混沌计算、模拟退火算法引入量子粒子群算法中,同时动态地调整 QPSO 的收缩扩张系数  $\beta$ , 使该算法具有动态自适应性。AS-QPSO 算法融合了 QPSO 算法的计算简便和混沌的遍历性等特点。较大地提高了 QPSO 算法的寻优能力和解的精确性。

## 2.5 AS-QPSO 算法流程图

图 1 为 AS-QPSO 算法流程图。

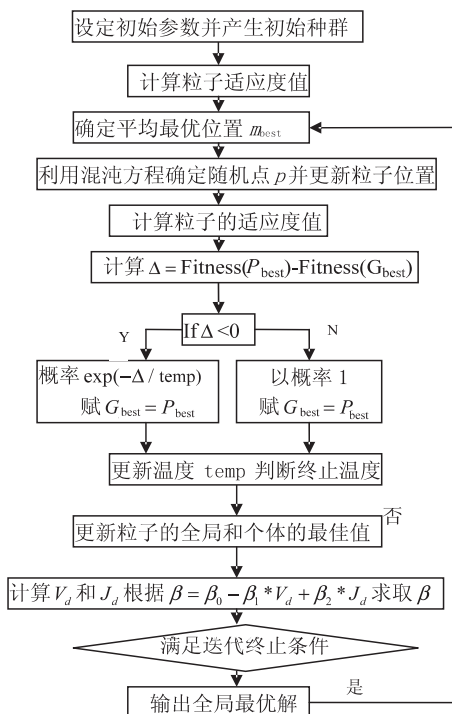


图 1 AS-QPSO 算法流程图



### 3 基于小波变换和 AS-QPSO 算法优化 LS-SVM 的预测方法

网络流量组合预测模型进行步骤如下:

1) 首先利用小波对流量数据进行分解。

对流量数据的分解层数和要用的小波基的选择对预测结果的精度是有较大影响的。因需选择合适的分解层数和利用到的小波基。文中利用 wavedec 函数对测试数据进行 2 层分解,用 wrcoef 进行单支重构。小波分解对时间序列进行了平滑,这样用小波变换可以将复杂的网络流量转化为用传统的预测方法来对分解后的时间序列进行预测。

2) 对流量数据进行归一化处理。不同时间段的流量数据是差别比较大的,所有首先要对小波分解后的信号利用如下公式进行归一化预处理,来提高预测精度。

$$x = \frac{x - \min(x_i)}{\max(x_i) - \min(x_j)}$$

3) 模型初始化。选取训练集和测试集,初始化各参数。

4) 模型组合优化。经过小波变换处理过的信号在高频部分是近似平稳的,可以考虑用 Linear 核 LS-SVM 来预测;低频部分是显著非平稳和非线性的,可以采用非线性的 RBF 核 LS-SVM 来预测<sup>[12]</sup>。利用 AS-QPSO 算法训练 LS-SVM 的参数  $C$  和  $\sigma$ 。

5) 先对预测模型利用优化出的参数进行训练,再对测试集分别进行预测,得到各预测结果。

6) 结果重构。先对各个预测结果进行反归一化处理,再小波重构得到最终预测结果。

## 4 实验结果与分析

### 4.1 仿真实验

实验所采用的原始流量样本数据来自 <http://ita.ee.lbl.gov/html/contrib/LBL-TCP-3.htm> 中的实际流量数据 lbl-tcp-3.tcp,仿真平台为 matlabR2011b。采用 1 s 为间隔进行重采样,得到长度为 7 199 的流量序列,并对此序列的前 100 个数据作为训练集,后 50 个数据作为测试集。首先将原始流量数据进行预处理,再利用 db2 小波对归一化处理后的数据进行 3 层分解。然后将分解后得到的各高频信息和低频信息分别单支重构到原级别上,用不同核函数 LS-SVM 预测,分别用 PSO/QPSO/AS-QPSO 算法优化各参数。最后再叠加各个子序列的预测结果,最终得到网络流量数据。为了检验模型在网络流量预测中的有效性,做了 4 个实验,实验 1~3:分别基于 PSO/QPSO/AS-QPSO 算法训练小波变换的 LS-SVM 模型进行预测;实验 4:直接用 LS-SVM 模型进行流量预测。文中选择 RMSE(根

均方误差)作为评价模型的预测效果:  $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^m (y_i - \hat{y}_i)^2}$ ,  $y_i$  为预测真实值,而  $\hat{y}_i$  为  $y_i$  的预测值。如果 RMSE 值越小,则表示模型的预测效果就越好。

### 4.2 结果分析

表 1 为三种模型预测的最优适应度值和 MSE 误差比较。

表 1 三种模型预测的最优适应度值和 MSE 误差比较

	Cvaccuracy	MSE
PSO-LS-SVM	0.013 968 578 434 87	70.652 864 941 162
QPSO-LS-SVM	0.007 969 432 885 50	62.723 183 463 996
AS-QPSO-LS-SVM	0.005 380 526 556 8	54.387 152 479 055

AS-QPSO 算法寻找最优参数的运动轨迹(见图 2)。

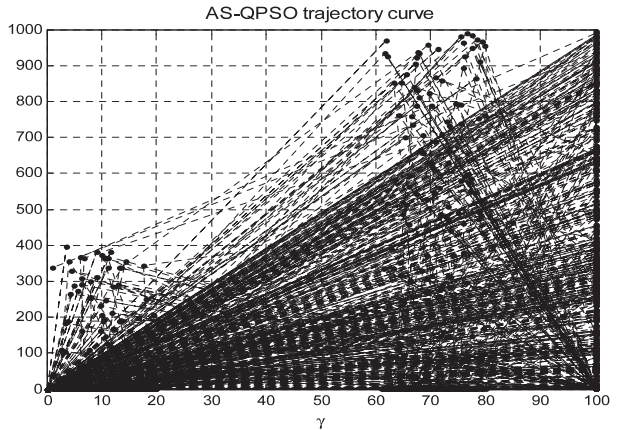


图 2 AS-QPSO 运动轨迹图

粒子群算法 PSO、QPSO、AS-QPSO 适应度函数收敛曲线如图 3。

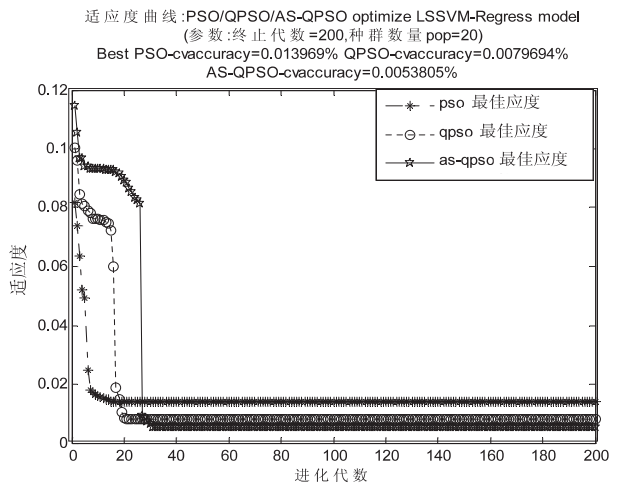


图 3 适应度函数收敛曲线

由表 1 和图 3 可知,因随机性,初始种群的优劣是无法控制的,在迭代次数相同时,标准 PSO 算法在初始种群较优的条件下只能搜索到较差精度,而 QPSO 算法和基于退火算法的 AS-QPSO 算法的初始种群优

劣程度是相当的,综合多种优点的 AS-QPSO 算法在初始种群最差的条件搜索到了最高的精度。三种算法运行过程中发现,粒子群规模越大,在迭代次数相同时运行时间越长,预测精度则越高,这是由于大种群分布范围广,有更高概率寻找最优解。同时小种群也能在增加迭代次数的条件下找到指定精度的最优解。AS-QPSO 算法不仅预测精度很高,误差值最小,而且预测性能稳定。

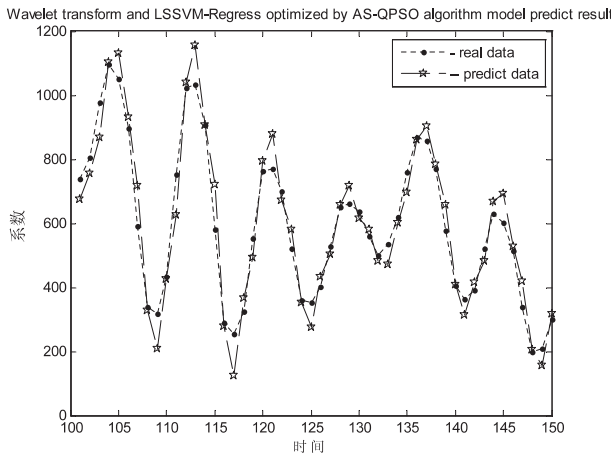


图4 基于 AS-QPSO 算法优化小波变换的 LS-SVM 模型的网络流量预测

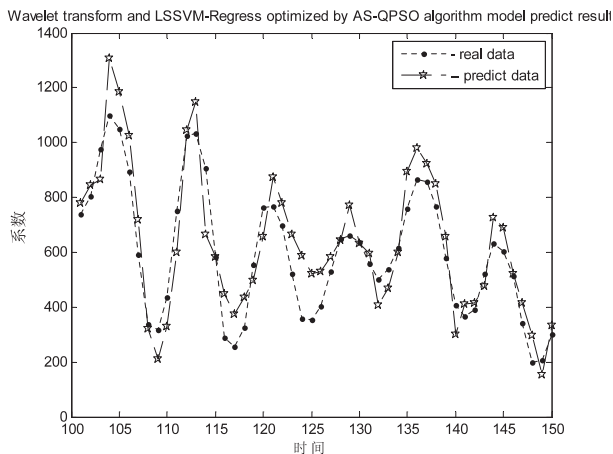


图5 基于 LS-SVM 模型的网络流量预测

从图4和图5可见,基于 AS-QPSO 算法优化的小波变换的 LS-SVM 的网络流量预测模型的预测性能明显高于其他两种模型,预测值与实际值基本相符。从表1和图4、图5可以看出基于 AS-QPSO 算法优化的经小波变换的 LS-SVM 进行流量预测时,不但预测性能很稳定,而且预测精度也很高。从实验结果中可以看出,在未经小波变换而直接对流量数据进行 LS-SVM 的实验中,没有很好地反映出流量数据的特性,并且预测精度较差。综上所述,采用基于 AS-QPSO 算法优化的小波变换的 LS-SVM 模型能达到所要求的精度,为网络流量预测提供一种方法。

## 5 结束语

文中提出了基于小波变换 LS-SVM 的网络流量预测模型,利用文中提出来的 AS-QPSO 算法对 LS-SVM 参数进行自动优化,有效提高了模型的泛化能力。实验结果表明文中模型的预精度要优于现有的一些模型,对于复杂的网络流量能取得较好的预测效果,具有鲁棒性好、泛化能力强和预测精度高等特点。将模拟退火算法、混沌机制、自适应机制引入到量子粒子群算法,充分利用了混沌的遍历特性和量子粒子群算法计算效率高等优点,具有较高的搜索效率,且有广泛的应用前景,可为应用提供科学依据。

## 参考文献:

- [1] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers [J]. Neural Processing Letters, 1999, 9 (3):293-300.
- [2] 李明迅,孟相如,袁荣坤,等.融合提升小波降噪和 LSSVM 的网络流量在线预测[J]. 计算机应用,2012,32(2):340-342.
- [3] 尹华,吴虹.最小二乘支持向量机在混沌时间序列中的应用[J]. 计算机仿真,2011,28(2):225-227.
- [4] An Senjian, Lu Wanquan, Venkatesh S. Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression [J]. Pattern Recognition, 2007, 40(8):2154-2163.
- [5] 张颖璐.基于遗传算法优化支持向量机的网络流量预测[J]. 计算机科学,2008,35(5):177-179.
- [6] 吴景龙,杨淑霞,刘承水.基于遗传算法优化参数的支持向量机短期负荷预测方法[J]. 中南大学学报(自然科学版),2009,40(1):180-184.
- [7] 姚全珠,蔡婕.基于 PSO 的 LS-SVM 特征选择与参数优化算法[J]. 计算机工程与应用,2010,46(1):134-136.
- [8] Zhou Lincheng, Yang Huizhong, Liu Chunbo. QPSO-based hyper-parameters selection for LS\_SVM regression [C]//Proceedings of Fourth International Conference on Natural Computation. Jinan, China; IEEE, 2008:130-133.
- [9] 崔海青,刘希玉.基于粒子群算法的 RBF 网络参数优化算法[J]. 计算机技术与发展,2009,19(12):117-119.
- [10] 胡旺,李志蜀.一种更简化而高效的粒子群优化算法[J]. 软件学报,2007,18(4):861-868.
- [11] Osowski S, Siroic R, Markiewicz Z, et al. Application of support vector machine and genetic algorithm for improved blood cell recognition [J]. Instrumentation and Measurement, 2009, 58 (7):2159-2168.
- [12] 曾勃炜,徐知海,付爱英,等.融合蚁群算法与支持向量机的网络流量预测[J]. 南昌大学学报(理科版),2011,35 (4):406-408.

一种新参数优化算法及其在流量预测中的应用

作者：邓宗强, 曾碧卿, DENG Zong-qiang, ZENG Bi-qing

作者单位：邓宗强, DENG Zong-qiang(华南师范大学 计算机学院, 广东 广州, 510631), 曾碧卿, ZENG Bi-qing(华南师范大学南海校区 信息工程与技术系, 广东 佛山, 528225)

刊名：计算机技术与发展

ISTIC

英文刊名：Computer Technology and Development

年, 卷(期): 2013(10)

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_wjtz201310009.aspx](http://d.g.wanfangdata.com.cn/Periodical_wjtz201310009.aspx)