Advances in Applied Mathematics 应用数学进展, 2020, 9(5), 782-790
Published Online May 2020 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2020.95093

Hans汉斯

Research on Nonlinear Combined Network Traffic Prediction Based on ELM

Xuehua Yu[1], Xiangzhen Xiong[2], Zhigang Wang[1*]

[1]School of Science, Hainan University, Haikou Hainan
[2]School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang Jiangsu
Email: *wzhigang@hainanu.edu.cn

Received: May 1st, 2020; accepted: May 19th, 2020; published: May 26th, 2020

Abstract

In order to accurately and reliably predict the short-term trend of network traffic, this paper proposes a nonlinear combination algorithm based on ELM for network traffic prediction based on the chaotic and non-linear characteristics of network traffic. Firstly, the original data are preprocessed, phase space reconstruction, and then Volterra filter adaptive prediction model, BP neural network prediction model, RBF neural network prediction model, wavelet neural network prediction model, SVM prediction model and extreme learning machine prediction model are established. The traffic time series is predicted and analyzed, and then the prediction result is used as the input of the combined ELM to construct a nonlinear combined prediction model. The results show that the mean squared error of the combined forecast is 5.51%, and its evaluation indexes are better than the sub-models. The validity and superiority of the combined strategy are verified, which can provide a better reference for network traffic prediction.

Keywords

Network Traffic, Phase Space Reconstruction, Neural Networks, Combined Forecast, Extreme Learning Machine

基于ELM的非线性组合网络流量预测研究

俞雪华[1]，熊相真[2]，王志刚[1*]

[1]海南大学理学院，海南 海口
[2]江苏大学计算机科学与通信工程学院，江苏 镇江
Email: *wzhigang@hainanu.edu.cn

收稿日期：2020年5月1日；录用日期：2020年5月19日；发布日期：2020年5月26日

*通讯作者。

文章引用: 俞雪华, 熊相真, 王志刚. 基于ELM的非线性组合网络流量预测研究[J]. 应用数学进展, 2020, 9(5): 782-790.
DOI: 10.12677/aam.2020.95093

摘　要

为了精准预测网络流量的短期走向，本文针对网络流量具有混沌性以及非线性性的变化特征，提出了一种基于ELM的非线性组合算法用于网络流量的预测方法。首先对原始数据进行预处理，相空间重构，然后建立Volterra滤波器自适应预测模型，BP神经网络预测模型，RBF神经网络预测模型，小波神经网络预测模型，SVM预测模型以及极限学习机预测模型对流量时间序列进行预测分析，再将其预测结果作为组合ELM的输入，构建起非线性组合预测模型。结果显示，该组合预测均方百分误差为5.51%，其各项评价指标均优于各子模型，验证了该组合策略的有效性与优越性，可为网络流量的预测提供一个较好的参考。

关键词

网络流量，相空间重构，神经网络，组合预测，极限学习机

1. 引言

如今互联网飞速发展，网络规模不断扩大，网络结构更加复杂，用户数量不断增加，如何有效、合理地分配网络资源面临着巨大挑战。因此，对网络流量的预测就变得越来越重要。准确的网络流量预测对网络安全防护、拥塞控制、带宽资源分配等具有关键作用。

研究表明，网络流量数据具有混沌性，非线性性和自拟合性等特点[1]。对于混沌性，使用相空间重构法可有效的提高预测的准确率。而对于非线性性，基于传统的线性预测并不能对其进行准确预测，如线性回归分析，马尔科夫模型等无法适应现代大规模网络流量数据的建模与预测，这些线性回归模型无法有效地提取网络流量的特征，需要提高预测精度。近年来学者们使用基于自相似流量的预测模型，有Volterra 滤波器模型，BP 神经网络模型，RBF 神经网络模型，小波神经网络模型，支持向量机模型和极限学习机模型，得到了一些预测结果[2]-[14]。然而单一的预测往往无法将时间序列的特征最大限度的提取出来，存在一定的局限性。为了不丢失其中有用的信息片段，本文提出对网络流量进行非线性组合预测[15][16]。结果显示，该组合预测各项评价指标均优于各子模型，可为网络流量的预测提供一个较好的参考。

2. 基于 ELM 的非线性组合预测模型构建

设网络流量时间序列为 $x_i, i=1,2,\cdots,n$，采用 m 种预测模型对其进行预测，得到预测结果 $y_j, j=1,2,\cdots,m$，其平均值为 $\bar{y}$，然后根据某种组合规则确定一个非线性映射函数 $\Phi(\cdot)$，则组合预测结果为：

$$y=\Phi\left(y_1,y_2,\cdots,y_m,\bar{y}\right) \tag{1}$$

不同的非线性映射函数所能提取的子模型信息具有很大的差异，因此选择一个能够综合利用各子模型有效信息的非线性映射函数对于网络流量预测具有重要意义，由于极限学习机具有优异的泛化性能，且能在保证学习精度的前提下比神经网络速度更快，所以本文使用极限学习机来确定该非线性映射函数，构建的非线性组合预测模型如图 1 所示。

极限学习机(ELM)是一种基于前馈神经网络的快速学习算法[17]，其主要特点是输入层与内隐层之间的连接权重和内隐层的阈值可以随机生成，而且在训练过程中不用对其进行调整，隐藏层和输出层之间的连接权重不需要迭代更新，而是通过解方程组方式一次性确定。对于一个具有 L 个内隐层节点的单层隐层神经网络，存在 β_i, W_i 和 b_i，使得输出误差最小，即

$$\sum_{j=1}^{L}\beta_i g\left(W_i\cdot X_j+b_j\right)=t_j,\ j=1,2,\cdots,N \tag{2}$$

式中 $g(x)$ 为激活函数，本文选用 sigmoid 函数，W_i 为输入权重，β_i 为输出权重，b_i 为第 i 个隐层单元的偏差。该算法中，一旦随机确定了隐藏层的输入权重 W_i 和偏差 b_i，就相应地确定了隐藏层的输出矩阵，输出权重 β 即可通过求解一个方程 $H\beta=T$ 得到，因此具有较快的学习速度。其具体步骤为：

Step 1：对于一列原始网络流量时间序列，本文先对其进行数据预处理，利用 Q-Q 图及标准化残差找出并替换掉其中的异常值。

Step 2：然后对其进行归一化处理，利用 C-C 法求出嵌入维数 m 和延迟时间 τ，并在相位空间中对原始数据进行了重构，得到各子模型的训练样本集和测试样本集。

Step 3：然后分别使用 Volterra 自适应滤波器预测模型、BP 神经网络预测模型、RBF 神经网络预测模型、小波神经网络预测模型、支持向量机预测模型以及极限学习机预测模型进行独立预测，反复训练网络 10 次，并取其平均值，得到 6 个子模型预测结果。

Step 4：将以上 6 个子模型预测结果及其平均值作为组合极限学习机的输入，反复训练网络 10 次，并取平均值得到最终的预测结果。

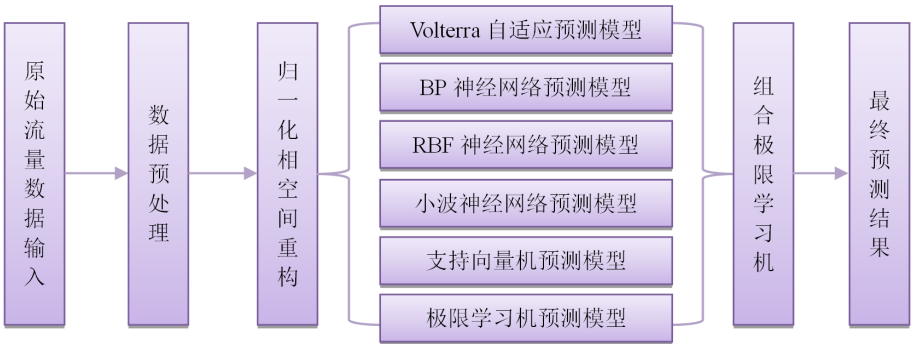

Figure 1. Nonlinear combined forecasting model
图 1. 非线性组合预测模型

3. 相空间重构

现实生活中所测得的网络流量往往只是一列随时间变化的具有某种意义的观测值，它是众多因素综合作用的结果，具有一定的混沌性，相空间重构理论指出[1] [18]：对于混沌系统，任何组件的演化都是由与之交互的其他变量决定的，因此，每个分量的发展过程中都包含了该系统相关分量的所有信息。对于一个一维网络流量时间序列 $s(t_i)=\{s(t_1),s(t_2),\cdots,s(t_n)\}$，$n$ 为时间序列样本数，通过延迟坐标法可以构造 $M=n-(m-1)\tau$ 个 m 维相空间矢量：

$$S_j=\left(s(t_j),s(t_j+\tau),\cdots,s(t_j+(m-1)\tau)\right),(j=1,2,\cdots,M) \tag{3}$$

于是重构的相空间为：

$$S=[S_1,S_2,\cdots,S_n]^{\mathrm{T}} \tag{4}$$

式中 m 表示网络流量的嵌入维数，τ 表示网络流量的延迟时间，S 为 $M\times m$ 维矩阵。

依据 Takens 嵌入定理[19]，对于一个维数为 D 的吸引子，当嵌入维数 $m\geq 2D+1$ 时，重构的吸引子能保持原来吸引子的拓扑特性，即重构的相空间与原始网络流量时间序列是拓扑等价的。因此，在相空间重构中，选择适合的嵌入维数 m 和延迟时间 τ 是非常重要的。当嵌入维数 $m<2D+1$ 时，重构的相空间将不能恢复原始奇异吸引子的特性，过大时，重构的相空间将引起过多的噪声，从而使得预测精度下降；当延迟时间 τ 过小时，相邻的两点相关性太强而产生大量的冗余，使得预测精度下降，过大时，重构的相空间将丢失大量原始时间序列的特性。同时本文认为嵌入维数和延迟时间是相关的，因此采用 C-C 方法，通过计算嵌入窗宽 $t_w=(m-1)\tau$ 来求得嵌入维数和延迟时间。

4. 预测子模型构建

4.1. Volterra 自适应滤波器预测模型

Volterra 是一种非线性自适应 FIR 滤波器[20]。由于网络流量的混沌和非线性，Volterra 非线性系统函数能够利用本身的高阶矩信息，实现对网络流量的准确预测，本文利用 Volterra 级数展开公式，建立了网络流量的非线性预测模型，同时由于无穷级数在实际中的实现较为困难，因此本文采用二次截断 m 次求和，其预测模型为：

$$x(n+1)=h_0+\sum_{i_1=0}^{m-1}h_1(i_1)x(n-i_1\tau)+\sum_{i_1=0}^{m-1}\sum_{i_2=0}^{m-1}h_k(i_1,i_2)x(n-i_1\tau)x(n-i_2\tau) \tag{5}$$

式中 $h_k(i_1),h_k(i_1,i_2)$ 为 k 阶 Volterra 核，其可用归一化最小均方自适应算法求解。

4.2. BP 神经网络预测模型

BP 神经网络是一种反向传递并且能够修正误差的多层映射函数，它解决了多层网络中隐含单元学习权的学习问题，能够以任意精度拟合任何连续非线性函数。本文采用三层网络结构：输入层、隐藏层和输出层。其算法主要思想分为信息正向传递和误差反向传播两大过程。本文使用的 BP 神经网络预测模型[21]为：

$$s_{i+1}=\frac{1}{1+\exp\left(-\sum_{j=1}^{p}v_jb_j+\gamma\right)} \tag{6}$$

式中 s_{i+1} 为预测值，p 为隐藏层节点数，v_j 为隐藏层至输出层的连接权重，γ 为输出层的阈值。

4.3. RBF 神经网络预测模型

RBF 神经网络是一种具有多层前向网络类似结构且带有单隐藏层的三层前馈网络[13] [22]，它可以任意精度地逼近任意连续函数。构建 RBF 网络的关键在于构造从输入层到隐藏层 $X\to\phi_i(X)$ 的非线性映射：

$$\phi_i(X)=\exp\left(\frac{\|X-c_i\|^2}{-2\sigma_i^2}\right),(i=1,2,\cdots,m) \tag{7}$$

式中 c_i 为第 i 个隐藏层高斯函数的中心矢量，其与输入向量 X 维数相同，σ_i 为第 i 个隐藏层高斯函数的宽度，$\|\cdot\|$ 为欧式范数，m 为隐藏层中的节点数。以及从隐藏层到输出层 $\phi_i(X)\to Y$ 的线性映射：

$$Y=\sum_{i=1}^{m}\omega_i\phi_i(X) \tag{8}$$

式中 ω_i 为隐藏层与输出层之间的连接权重，Y 即为预测值。

4.4. 小波神经网络预测模型

小波神经网络是小波分析和 BP 神经网络相结合的产物[14]，它既包含了神经网络良好的训练能力和泛化能力，也具有小波函数良好的时频特性。其网络结构与 BP 神经网络结构相一致，由输入层、隐层和输出层组成，只是将 BP 神经网络中隐藏层的激励函数 Sigmoid 换成了小波函数，本文所选择的小波函数为 Morlet 函数：

$$f(x)=\exp\left(-\frac{x^2}{2}\right)\times\cos(1.75x) \tag{9}$$

其中隐藏层每个节点的输出为：

$$\phi(x_i)=f_i\left(\frac{\sum_{i=1}^{m}\omega_{ij}x_i-b_j}{a_j}\right),\ j=1,2,\cdots,p \tag{10}$$

式中 a_j 为小波函数的缩放因子，b_j 为小波函数的平移因子，p 为隐藏层的节点数。其预测模型为：

$$y=\sum_{i=1}^{p}\omega_i\phi(x_i) \tag{11}$$

式中 ω_i 为隐藏层到输出层的连接权重。

4.5. 支持向量机预测模型

用支持向量机来进行预测，与回归函数估计问题类似，不同的是支持向量机要找一个超平面 $x\omega+b=0$ 使得支持向量到超平面的距离尽可能大，为解决线性不可分问题，需构造一个从低维空间到高维空间的映射，从而使得样本点在高维空间中线性可分，同时引入核函数，可以将高维空间中的复杂点积运算转化为低维空间的核函数计算。本文选择 RBF 核函数

$$K(x,y)=\exp\left(-\gamma\|x-y\|^2\right) \tag{12}$$

最后所得的回归函数为：

$$f(x)=\sum_{i=1}^{nsv}\left(\alpha_i-\alpha_i^*\right)K(x_i,y_j)+b \tag{13}$$

式中 nsv 为支持向量的个数，$\left(\alpha_i-\alpha_i^*\right)$ 为对应支持向量的系数，可以通过引入 Lagrange 函数求解一个二次规划问题求出，b 为阈值，可以通过其中两个支持向量求出。SVM 的输出是中心节点的线性组合，每个中心节点对应一个支持向量。

5. 模型评价标准

本文选用绝对误差 Err，平均绝对百分误差 $MAPE$ 和相对误差 $Perr$ 作为预测精度的评测标准，其定义分别为：

$$Err=x(n)-x_p(n) \tag{14}$$

$$MAPE=\frac{1}{N_p}\sum_{n=1}^{N_p}\left|\frac{x(n)-x_p(n)}{x(n)}\right| \tag{15}$$

$$Perr=\frac{\sum_{n=1}^{N_p}\left(x(n)-x_p(n)\right)^2}{\sum_{n=1}^{N_p}x^2(n)} \tag{16}$$

式中，N_p 表示训练数据集个数，$x(n)$ 表示真实值，$x_p(n)$ 表示预测值。

6. 实证分析

6.1. 实验环境及数据来源

本文的实验数据如图 2 所示，其来源于 kaggle 的网络流量预测比赛，该数据表示维基百科某文章在连续的 803 天内的页面点击量。本文的实验环境为 spss v21.0 以及 matlab 2013a，先使用 spss 对网络流量进行预分析，找出并替换其中的异常值，然后采用 matlab 语言编写算法计算程序，并应用 matlab 构建了 6 个子预测模型，分别是 Volterra 滤波器模型、BP 神经网络模型、RBF 神经网络模型、小波神经网络模型、支持向量机模型以及极限学习机，将其预测结果作为组合 ELM 的输入构建组合预测模型，并对该网络流量时间序列进行预测对比实验。

6.2. 数据预处理

从图 2 中可以明显看出，有一些点已经远离数据主体，这些离群点可能会成为强影响点，使得模型训练结果会偏向这些数据点，因此为了使得该预测更加准确可行，应该找出这些异常值并删除或替换。根据正态分布的 3σ 原则，当标准化残差 Z 大于 3 或小于−3 时，说明该值为异常值，本文用该异常值左右最近的两个非异常值的平均值替代该异常值。

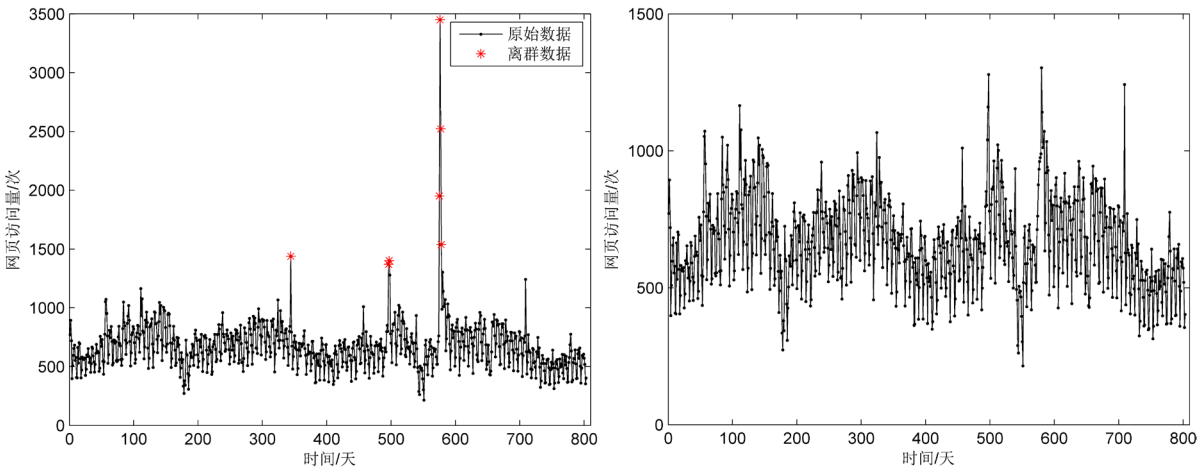

Figure 2. Network traffic data (comparison of original data before processing and data after processing)
图 2. 网络流量数据(处理前原始数据和处理后数据对比)

6.3. 预测结果及分析

本文首先将这些数据归一化处理为均值为 0，方差为 1 的时间序列，其具体操作为：

$$x'(i)=\frac{x(i)-\min\left(x(i)\right)}{\max\left(x(i)\right)-\min\left(x(i)\right)} \tag{17}$$

式中 $x(i)$ 表示原始值，$x'(n)$ 表示归一化值。然后利用 C-C 法确定了最优的延迟时间 τ 为 4，嵌入维数 m

为 2，对网络流量进行相空间重构，建立起预测模型的训练样本和实验样本。

6.3.1. 组合预测结果

本文的预测方式分为单一预测，子预测以及组合预测 3 类。首先将数据集的前 750 个数据作为 6 个子模型的训练集，后 53 个作为测试集，经各子模型训练得到各模型的单一预测结果，然后将数据集的前 560 个数据作为 6 个子模型的训练集，后 243 个作为测试集，经各子模型训练得到 6 个子预测结果，将这 6 个子预测结果及其平均值作为组合 ELM 的输入，前 190 个用作训练集，最后 53 个用作测试集，真实值作为组合 ELM 的输出，经组合 ELM 训练得到组合预测结果，然后对比分析组合预测以及 6 个子模型单一预测。同时由于各预测模型网络训练时的随机性，使得每次训练结果有所差异，为使得预测结果更有说服力，因此本文将各模型训练 10 次，10 次预测结果的平均值作为各模型的最终预测结果。其中图 3 表示各单一预测和组合预测的结果对比，图 4 表示组合预测的结果及误差。

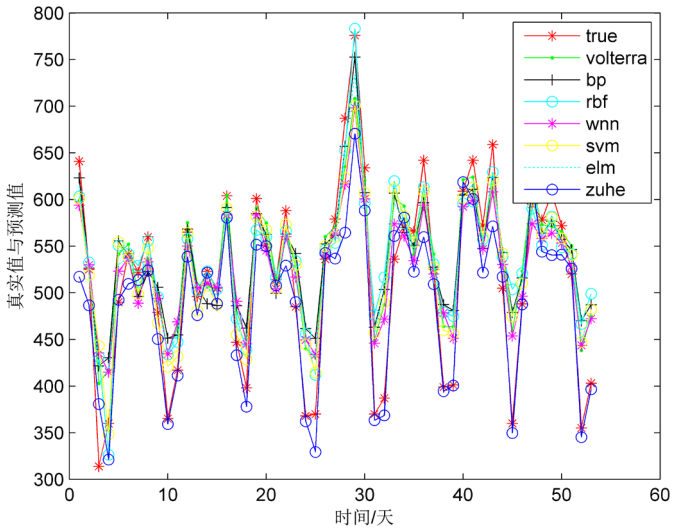

Figure 3. Comparison between single forecast and combined forecast
图 3. 单一预测与组合预测对比

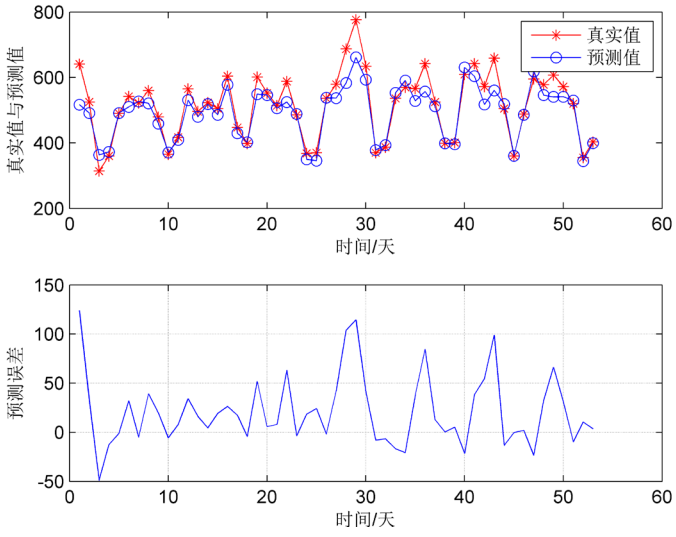

Figure 4. Combined prediction results and errors
图 4. 组合预测结果及误差

6.3.2. 结果分析

通过对上述仿真结果的分析，可以得出以下结论：

1) 从图 3 可以看出，6 种子模型都具有较高的预测精度，均有预测值与真实值相等的情况存在，能够很好地反映网络流量变化的趋势和规律，而且预测误差也在比较小的误差范围内，从表 1 可以看出，6 种子模型的均方百分误差分别为 8.31%，9.46%，8.37%，8.42%，7.86%，7.88%都满足小于 10%的预测精度，说明了子模型对网络流量时间序列信息的捕捉比较全面，具有很好的非线性拟合能力，同时进一步说明了这些子模型的有效性与优越性，这也为组合预测打好了良好的基础。

2) 从图 4 可以看出，组合预测结果已经基本与真实值走势吻合，其预测均方百分误差为 5.51%，小于各子模型的误差，说明了本文的组合策略能够较好的综合各子模型所捕捉的序列信息，使得各子模型优势互补，因而得到一个更为精确的预测结果，也说明了本文的组合策略的有效性与可行性，可以为网络流量的预测提供一个较好的参考。

Table 1. Error comparison of prediction results
表 1. 预测结果误差对比

	Volterra	BP	RBF	WNN	SVM	ELM	组合 ELM
MAPE	8.31%	9.46%	8.37%	8.42%	7.86%	7.88%	5.51%
Perr	0.0076	0.0101	0.0099	0.0081	0.0080	0.0075	0.0066

基金项目

海南省自然科学基金高层次人才项目(2019RC168)和海南省科协青年科技英才创新计划(QCXM201806)资助。

参考文献

[1] 顾兆军, 李冰, 刘涛. 基于 PSO-Elman 模型的网络流量预测[J]. 现代电子技术, 2019, 42(1): 82-86.

[2] 殷礼胜, 何怡刚, 董学平, 等. 交通流量 VNNTF 神经网络模型多步预测研究[J]. 自动化学报, 2014(9): 2067-2073.

[3] 任师涛, 史志才, 吴飞, 等. 基于改进 BP 神经网络的路由器流量预测方法[J]. 传感器与微系统, 2018, 37(8): 49-50.

[4] 刘杰, 黄亚楼. 基于 BP 神经网络的非线性网络流量预测[J]. 计算机应用, 2007(7): 1770-1772.

[5] 王俊松, 高志伟. 基于 RBF 神经网络的网络流量建模及预测[J]. 计算机工程与应用, 2008, 44(13): 6-7.

[6] 张玉梅, 曲仕茹, 温凯歌. 基于混沌和 RBF 神经网络的短时交通流量预测[J]. 系统工程, 2007(11): 26-30.

[7] 李欣, 孙珊珊. 基于小波神经网络的网络流量预测研究[J]. 现代电子技术, 2016, 39(23): 98-99.

[8] 雷霆, 余镇危. 一种网络流量预测的小波神经网络模型[J]. 计算机应用, 2006(3): 526-528.

[9] 熊凡. 遗传算法优化支持向量机的网络流量混沌预测[J]. 现代电子技术, 2018, 41(18): 166-169.

[10] 袁开银, 魏彬. 相空间重构和极限学习机的网络流量预测模型[J]. 控制工程, 2018, 25(11): 2087-2091.

[11] 王思哲. 葡萄酒品尝评分的可信性度量方法研究[J]. 海南热带海洋学院学报, 2018, 25(5): 74-79.

[12] 王思哲, 王志刚, 何勇. 基于数据挖掘技术的葡萄酒评价体系研究[J]. 应用数学进展, 2015, 4(4): 376-384.

[13] Liu, J. and Guo, Z.H. (2014) Network Traffic Prediction Using Radial Basis Function Neural Network Optimized by Ant Colony Algorithm. *Sensors & Transducers*, **172**, 224-228.

[14] Tan, X.L., Fang, W.J. and Qu, Y. (2013) Network Traffic Prediction Algorithm Based on Wavelet Transform. *International Journal of Advancements in Computing Technology*, **5**, 183-190.

[15] 凌立文, 张大斌. 组合预测模型构建方法及其应用研究综述[J]. 统计与决策, 2019, 35(1): 18-23.

[16] Ahmad, Z. and Zhang, J. (2008) Selective Combination of Multiple Neural Networks for Improving Model Prediction in Nonlinear Systems Modelling through Forward Selection and Backward Elimination. *Neurocomputing*, **72**, 1198-1204. https://doi.org/10.1016/j.neucom.2008.02.005

[17] Yang, Y.M. and Jonathan Wu, Q.M. (2016) Extreme Learning Machine with Subnetwork Hidden Nodes for Regression and Classification. *IEEE Transactions on Cybernetics*, **46**, 2885-2898. https://doi.org/10.1109/TCYB.2015.2492468

[18] 陈铿, 韩伯棠. 混沌时间序列分析中的相空间重构技术综述[J]. 计算机科学, 2005(4): 67-70.

[19] 雷绍兰, 孙才新, 周湶, 张晓星. 电力短期负荷的多变量时间序列线性回归预测方法研究[J]. 中国电机工程学报, 2006(2): 25-29.

[20] 王兰, 李华强, 吴星等. 基于改进局域 Volterra 自适应滤波器的风电功率混沌时间序列预测模型[J]. 电力自动化设备, 2016, 36(8): 40-44.

[21] 刘春. 遗传算法优化 BP 神经网络的网络流量预测[J]. 信息安全与技术, 2014, 5(6): 82-86.

[22] 李立华. 基于径向基函数(RBF)神经网络模型的金融混沌预警研究[J]. 海南金融, 2012(6): 32-35.