

文章编号: 1001 - 9081 (2006) 10 - 2278 - 04

## 网络流量行为分析的小波加权混沌局域模型

雷霆<sup>1,2</sup>, 余镇危<sup>1</sup>

(1. 中国矿业大学(北京)机电与信息工程学院, 北京 100083; 2. 北京林业大学理学院, 北京 100083)  
(xleiting@126.com)

**摘要:** 结合小波变换和混沌局域模型的各自优势, 提出一种网络流量的预测模型。首先, 将网络流量时间序列进行小波分解得到高频信号序列和低频信号序列, 再用加权混沌局域模型对每一成分的信号序列分别进行预测, 对所有的预测分量进行小波重构就可以实现对网络流量的预测。用实际网络流量对该模型进行验证, 实验结果表明, 该模型具有较高的预测效果。

**关键词:** 网络流量; 混沌; 小波变换; 加权混沌局域模型

**中图分类号:** TP393.04; TP183 **文献标识码:** A

## Wavelet weighted chaos local-region model of network traffic behavior analysis

LEI Ting<sup>1,2</sup>, YU Zheng-wei<sup>1</sup>

(1. School of Mechanical Electronic & Information Engineering, China University of Mining & Technology at Beijing, Beijing 100083, China;  
2. School of Science, Beijing Forestry University, Beijing 100083, China)

**Abstract:** Integrating the advantage of wavelet transform with that of chaos local-region model, a new model of forecast network traffic was proposed. First, the network traffic time series was decomposed to the high frequency signal series and low frequency signal series and the weighted chaos local-region model was applied to predict these series respectively. Finally, forecasted traffic was achieved by wavelet reconstruction of all the forecasted components. The simulation results on real network traffic indicate that this model is more satisfactory than traditional methods in network traffic prediction.

**Key words:** network traffic; chaos; wavelet transform; weighted chaos local-region model

### 0 引言

网络流量行为的刻画对于实现数据可靠传输、合理分配网络资源具有重要的指导意义。研究表明, 随着计算机网络规模的扩展和业务种类的不断增长, 网络流量表现出混沌的行为特性, 使得传统的网络流量模型, 如 Markov 模型、Poisson 模型、回归模型等工具, 已无法有效地刻画真实的流量行为特性。因此, 寻求新的模型方法成为当前网络流量行为研究领域中的一个热点<sup>[1-3]</sup>。

小波变换由于其独特的多分辨分析能力, 在网络流量行为的研究领域日益受到广泛关注, 而混沌局域模型具有很强的非线性逼近功能而广泛应用于各学科领域。基于两者的各自优势, 本文将小波分析和混沌局域建模法结合起来, 提出了一种新的用于网络流量行为演化的分析模型, 即小波加权混沌局域模型, 利用该模型可以比较精确地预测出复杂网络流量的演化趋势。

### 1 小波变换

#### 1.1 小波变换定义

已知  $f(t) \in L^2(R)$ , 若满足条件  $\int_{-\infty}^{+\infty} |f(t)|^2 |t|^{-1} dt < +\infty$ , 或  $\int_{-\infty}^{+\infty} f(t) dt = 0$ , 其中  $\hat{f}(\omega)$  为  $f(t)$  的 Fourier 变换, 那么称  $f(t)$  为基本小波或母小波。将母小波函数做伸缩、

平移变换得:  $\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right)$ ,  $b \in R \setminus \{0\}$ , 称为小波函数, 简称小波。设网络流量时间序列为  $S(t)$ , 则它的小波变换定义为:

$$W_S(a, b) = \langle S, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} S(t) \overline{\psi_{a,b}(t-b)} dt$$

其中  $\overline{\psi_{a,b}(t-b)}$  表示  $\psi_{a,b}(t-b)$  的复共轭函数。

由定义可看出, 与 Fourier 变换只能做频域分析的情况不一样, 由于  $W_S(a, b)$  同时包含了时域和频域的信息, 所以小波变换可同时做时频分析。参数  $a$  的变化不仅改变小波的频谱结构, 而且改变其窗口的形状和大小, 而窗口位置的改变通过调节参数  $b$  来实现, 保证了小波函数的波形可窄可宽, 有利于观察信号特别是非平稳信号的细节行为, 比如瞬态或奇异点, 所以小波变换在时频域上都具有表征局部特征的行为和很强的分辨率。

做为母小波的函数, 必须具有良好的时、频局部特性, 以更好地刻画信号的突变信息, 因此怎样恰当地选择小波函数, 对于较客观全面地刻画信号的行为特性是很重要的问题, 常用的小波函数有 Haar 小波、Shannon 小波、Daubechies 小波等。

对于离散的信号序列, 进行小波变换时所用的是二进制离散形式:

$$W_S(2^{-j}, k) = \langle S, \psi_{2^{-j}, k} \rangle = \sum_{n=-\infty}^{+\infty} S(n) \overline{\psi_{2^{-j}, k}(2^{-j}n - k)} \quad (1)$$

收稿日期: 2006 - 04 - 17; 修订日期: 2006 - 06 - 13 基金项目: 国家教育部博士点基金资助项目 (20030290003)

作者简介: 雷霆 (1972 - ), 男, 广西百色人, 讲师, 博士研究生, 主要研究方向: 通信与信息系统、网络行为学; 余镇危 (1942 - ), 男, 上海人, 教授, 博士生导师, 主要研究方向: 计算机网络、下一代网络体系结构。

整数  $j$  和  $k$  分别表示频域分辨率和时间平移量。

### 1.2 小波分解与小波重构

设时间序列为  $\{S(t), t = 1, 2, \dots\}$ 。选择一组恰当的小波函数基  $\{\psi_{jk}\}_{jk \in \mathbb{Z}}$  与尺度函数基  $\{\phi_{jk}\}_{jk \in \mathbb{Z}}$ ，分别计算小波系数  $D_j(t) = \langle S, \psi_{jk} \rangle_{jk}$  与尺度系数  $A_j(t) = \langle S, \phi_{jk} \rangle_{jk}$ ，其中计算内积  $\langle S, \psi_{jk} \rangle$  的方法和计算  $\langle S, \phi_{jk} \rangle$  类似，只需把 (1) 式中的积分核  $\psi$  改为尺度函数的共轭即可。将  $S(k)$  进行  $L$  层小波分解，其结果可形式地表示为：

$$S(t) = \sum_{j=1}^L D_j(t) + A_L(t) \quad (2)$$

(2) 式右边的第一部分表示序列在小波子空间上的投影，为高频信号，所显示的是信号的细节部分，描述信号突变、瞬时的特性； $A_L(t)$  部分是序列在尺度空间的投影，是低频信号，所显示的是信号的近似部分，描述了信号的长期演化趋势、周期性等较为确定的性质。 $L$  为小波分解的层数， $L$  越大，表明在该层时间的分辨率越小，所能观察到信号的细节就越多。图 1 所表示的是  $L = 3$  时小波变换的三层分解过程。

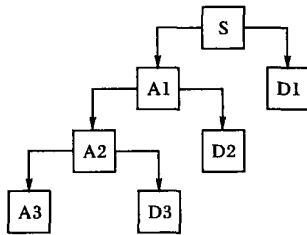


图 1 三层小波分解的过程

将信号做小波分解以方便观察不同尺度内信号的某些重要特性是小波分析的一个重要作用，但是这并不是小波分析的全部应用。小波分析另一个重要的应用是小波重构，也就是原始信号在经过分析、比较或者滤波、消噪、加密、预测等处理后得到高频信号和低频信号，然后再按照 (2) 式来合成、构造所需要的信号，这个过程实际上是小波变换的逆变换，也称为小波合成。

## 2 加权混沌局域模型

### 2.1 混沌的诊断

所谓混沌，是一种貌似无规则的运动，是在确定性复杂系统中不需要附加任何外在因素就能出现类似随机的行为。混沌系统的最大特点在于系统的演化对初始条件十分敏感，这使得对复杂系统的长期预测成为不可能。

对于一个复杂系统，判断它是随机还是混沌是一个很重要的问题。目前用得比较多的方法是关联维数法<sup>[4]</sup>和 Lyapunov 指数法。关联维数是描述动力学系统奇怪吸引子复杂程度的重要参数，具有非整数吸引子的维数，表征了动力学系统的运动轨道在相空间中随时间演变的混沌行为，关联维数越高，意味着吸引子的复杂程度越高。

Lyapunov 指数表示系统演化轨道收敛或发散的比率，用来度量在相空间中初始条件不同的两条相邻轨道随时间按指数速度吸引或分离的程度。对于一个  $n$  维混沌系统，存在着一个 Lyapunov 指数谱，若所有的指数按从大到小的顺序排列结果为  $\lambda_1, \lambda_2, \dots, \lambda_n$ ，则称  $\lambda_1$  为最大 Lyapunov 指数， $\lambda_1$  表征了混沌系统状态演化过程。若两相邻轨道  $Y(i)$  与  $Y(j)$  经过  $n$  次迭代后分别得到相点  $Y(i+n)$  与  $Y(j+n)$ ，两者距离的分离速度为指数级，即：

$$|Y(i+n) - Y(j+n)| = d(0) e^{\lambda_1 n}$$

其中  $d(0)$  是它们的初始距离。 $\lambda_1 = 0$  对应于稳定的边界，初始误差保持原样； $\lambda_1 < 0$  表明相体积收缩，轨道在局部稳定，对初始条件不敏感，对应于周期运动；当  $\lambda_1 > 0$ ，运动轨道在每个局部不稳定，相邻轨道按指数速度分离，轨道在整体性的稳定因素作用下反复折叠，形成奇异吸引子。 $\lambda_1$  同时也表征了相轨道对初始信息的遗忘的程度，给出系统最大的可预测的时间尺度  $T = 1/\lambda_1$ ，即当提前预测的时间小于  $T$  时，所预测的对象才属于可预测的范围。

### 2.2 相空间重构理论

描述网络流量行为的传统模型，往往把流量发生突发、振荡等复杂这些现象看成是系统外部的干扰所引起的随机噪声，而采用统计的方法将其噪声过滤掉。实际上，网络流量行为是混沌的，它所表现出来的无规则复杂行为是流量本身所固有的属性，是一种内随机性，不能把它忽略。所以有必要考虑建立起能刻画网络流量混沌特性的数学模型，以实现通过网络流量行为的演化趋势进行有效地预测和控制。

实际上建立这样的模型是非常困难的，为此，人们往往用非线性逼近的方式来对它的复杂行为特性进行分析，相空间重构理论就是在这样的背景下被提出来的。所谓相空间重构就是通过已知的系统少量信息来重新构造一个与原系统等价的混沌吸引子，用来研究系统动力学行为的方法。混沌理论认为，系统是一个整体，其中的任一分量的动力学行为都是由它与交互其他分量所决定的，它的演化包含了系统丰富的信息，基于这些结论，Takens 提出了相空间重构的理论基础，即 Takens 定理<sup>[5]</sup>。

Takens 定理表明，已知混沌时间序列  $\{S(t), t = 1, 2, \dots, n\}$ ，对于任意滞后时间  $m$  和嵌入空间维  $m \geq 2D_2 + 1$  ( $D_2$  为关联维数)，在吸引子上一定可找到某个光滑映射  $F: R^m \rightarrow R^m$ ，满足  $Y(t+1) = F[Y(t)]$ ，其中  $Y(t)$  是  $m$  维向量，即  $Y(t) = (S(t), S(t+1), \dots, S(t+(m-1)))$ ， $t = 1, 2, \dots, n - (m - 1)$ 。

这个结论不仅给我们提供了相空间重构的理论基础，实际上也给出了由一个混沌时间序列进行相空间重构的方法。由定理可以知道，重构相空间中的轨道在微分同胚意义下与原系统是等价的，原系统的动力学性质在这个重构的相空间保持不变，这样通过重构起来的这个相空间，能使我们在无法全面了解系统的各变量之间的相互关系情况下，也可通过目前所掌握的少数分量的信息来对整个系统的动力学行为进行研究，了解整个系统的动力学特性。

在重构相空间过程中，必须适当地选择滞后时间  $m$  和重构相空间的嵌入维数  $m$ ，否则会引入不必要的噪声，从而无法保证原系统的真实信息在重构相空间表现出来。

滞后时间  $m$  的计算有很多方法，因为自相关函数可以用来刻画时间序列与其滞后序列的线性独立程度，所以可用它来确定。时间序列  $\{S(t), t = 1, 2, \dots, n\}$  的自相关函数定义为  $ACF(\tau) = \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} (S_i - \bar{S})(S_{i+\tau} - \bar{S}) / \sigma^2$ ， $\sigma^2$  为时间序列的方差。从直观上看，当选取  $ACF(\tau) = 0$  时所对应的滞后  $\tau$  可以保证嵌入空间各个分量的线性独立。但是实践证明，由此所得到的  $\tau$  往往偏大，实际中将  $ACF(\tau)$  下降到  $1 - \frac{1}{e}$  即约 0.632 时对应的  $\tau$  作为重构相空间的时间滞后会取到更好的效果<sup>[6]</sup>。如何确定  $m$  同样有许多算法，用得比较多的是 GP 算

法<sup>[4]</sup>,该算法可通过编程来实现,用起来比较方便。

### 2.3 加权混沌局域模型的建立

在 Takens定理,除了  $m$  外,还有一个关键的问题没解决,那就是如何确定映射  $F$ 。由于定理没有指明映射  $F$  的具体形式,所以怎样确定  $F$  是较为困难的问题。

构造  $F$  的方法很多,目前常用的方法大多数是采用基于相空间重构理论的各种拟合模型来逼近  $F$ 。根据所使用数据的范围,这些模型大体上可分为混沌全域模型和混沌局域模型两类。混沌全域模型计算量大,且仅适用于光滑函数,对混沌系统不太适合。混沌局域模型的基本思想是在相空间中选择预测中心邻域内的若干点而不是所有的点,建立一个线性或非线性的拟合函数来近似表示  $F$ 。该模型应用比较广泛,下面介绍混沌局域建模的方法。

假设流量时间序列为  $\{S(t), t = 1, 2, \dots, n\}$  为已知,为了把该序列的未来的行为状态即  $S(n+1), S(n+2), \dots$  预测出来,可以采用先对该时间序列重构相空间,然后建立局域模型来预测的方法。选择适当的滞后时间  $\tau$  与嵌入相空间维数  $m$ ,则可得到该序列的重构相空间相应的  $N = n - (m - 1)\tau$  个相点:  $Y(1) = (S(1), S(1 + \tau), \dots, S(1 + (m - 1)\tau))$ ,  $Y(2) = (S(2), S(2 + \tau), \dots, S(2 + (m - 1)\tau))$ ,  $\dots$ ,  $Y(N) = (S(N), S(N + \tau), \dots, S(N + (m - 1)\tau))$ 。这些相点的行为状态都已知,我们将预测时间序列的下一步演化行为,即  $Y(N + 1) = (S(N + 1), S(N + 1 + \tau), \dots, S(N + 1 + (m - 1)\tau))$ 。显然,如果  $Y(N)$  下一步的状态  $Y(N + 1)$  能被预测出来,则时间序列  $\{S(t), t = 1, 2, \dots, n\}$  下一步的行为状态  $S(n + 1)$  也就被预测出来了,因为  $S(n + 2)$  是  $m$  维向量  $Y(N + 1)$  的最后一个分量。

由于  $Y(N + 1)$  是  $Y(N)$  演化结果,为此,可考虑以  $Y(N)$  为预测中心,把与  $Y(N)$  空间距离最近的一些点  $Y(i_1), Y(i_2), \dots, Y(i_M)$  (其中  $M < N$ ) 找出来,因为在相空间中,这些点对  $Y(N)$  的动力学行为影响最大,它们的平均演化行为可以看成是  $Y(N)$  的演化近似,所以,如果设  $Y(i_1), Y(i_2), \dots, Y(i_M)$  的下一步演化行为分别是  $Y(i_1 + 1), Y(i_2 + 2), \dots, Y(i_M + 1)$ ,在  $M < N$  条件下,  $Y(i_1 + 1), Y(i_2 + 2), \dots, Y(i_M + 1)$  显然都是已知的。这样,  $Y(N + 1)$  预测值可以用下面的公式来计算,即<sup>[7]</sup>:

$$Y(N + 1) = \frac{1}{M} \sum_{k=1}^M Y(i_k + 1) \quad (3)$$

这是一个常用的混沌预测公式,叫零阶局域模型。当  $M = 1$  时,称为最近点模型,而当  $M = N$ ,则意味着拟合过程用到了相空间的所有点,是一种全域模型。(3) 式所表示的模型一个明显不足是没有考虑到各个邻近点与中心点的空间距离对预测结果的影响,而实际上相空间各点与中心点的空间距离是一个非常重要的参数,预测的精度,往往取决于和中心点最邻近的那些点。因此,将距离作为一个拟合参数引入到演化方程中,在一定程度上可提高预测的精度,而且有一定的去噪能力<sup>[8]</sup>。所以,文献 [9] 对该模型做了改进,得到加权混沌局域模型:

$$Y(N + 1) = \frac{\sum_{k=1}^M Y(i_k) \exp[-A(d_k - d_{\min})]}{\sum_{k=1}^M \exp[-A(d_k - d_{\min})]} \quad (4)$$

其中  $d_k$  分别是各邻近点  $Y(i_k)$  到中心点  $Y(N)$  的距离,  $d_{\min}$  是所有  $d_k$  中的最小值;参数  $A$  可根据实际情况选取,通过修正  $A$ ,可使模型更好匹配于不同的动力学系统。(4) 式所表

示的模型虽然考虑到了不同的邻近点对预测中心点动力学行为影响的差异,但对于怎样选取参数  $A$  存在一定的盲目性和主观性,用不同的  $A$  值会得到不同结果,容易使模型产生不稳定的振荡。针对这些不足,本文拟提出以下新的加权混沌局域模型。

假设预测中心  $Y(N)$  的邻近点  $Y(i_1), Y(i_2), \dots, Y(i_M)$  与  $Y(N)$  的距离分别是  $R_1, R_2, \dots, R_M$ , 令  $R = \frac{1}{\sum_{k=1}^M \frac{1}{R_k}}$ ,  $w_j = \frac{1/R_j}{R}$  ( $j = 1, 2, \dots, M$ )。则  $Y(N + 1)$  的值可由以下的加权零阶模型预测出来,即:

$$Y(N + 1) = \sum_{k=1}^M w_k Y(i_k + 1) \quad (5)$$

注意到  $Y(N + 1)$  是  $Y(i_1 + 1), Y(i_2 + 2), \dots, Y(i_M + 1)$  的一种加权平均。由 (5) 式可以看出,邻近点距离  $Y(N)$  越小,对预测结果影响就越大,所以加权平均刻画了相空间中不同邻近点在  $Y(N)$  的动力学演化过程中所起作用的差异。同时,每个邻近点的权重是确定的,克服了上述 (5) 式所表示的加权模型振荡不稳定的缺点。

完成了一步预测以后,确定  $Y(N + 1)$  为新的预测中心,重复以上步骤,经过一定次数的迭代,可做出时间序列演化的有限步预测值。

## 3 小波加权混沌局域模型

结合小波变换和加权混沌局域模型的各自优点,下面建立一种用于网络流量行为演化分析的模型,该模型称为加权小波混沌局域模型,建模的步骤叙述如下:

1) 将所采集到的网络流量序列  $\{S(t), t = 1, 2, \dots, n\}$  进行  $L$  层小波分解,得到高频信号序列  $\{D_j(t), t = 1, 2, \dots, n\}$  ( $j = 1, 2, \dots, L$ , 下同) 和低频信号序列  $\{A_L(t), t = 1, 2, \dots, n\}$ , 总共  $L + 1$  个信号序列。

2) 对于每个信号成分  $\{D_j(t), t = 1, 2, \dots, n\}$  和  $\{A_L(t), t = 1, 2, \dots, n\}$ , 分别选择适当的滞后时间  $\tau_D$  和  $\tau_A$ , 以及适当的嵌入维数  $m_D^j$  和  $m_A^L$ , 按各自的序列进行相空间重构,得到相应重构相空间:  $\{Y_D^j(t) = (D_j(t), D_j(t + \tau_D), \dots, D_j(t + (m_D^j - 1)\tau_D))\}$ ,  $t = 1, 2, \dots, N = n - (m_D^j - 1)\tau_D$  和  $\{Y_A^L(t) = (A_L(t), A_L(t + \tau_A), \dots, A_L(t + (m_A^L - 1)\tau_A))\}$ ,  $t = 1, 2, \dots, N = n - (m_A^L - 1)\tau_A$ 。

3) 由  $\{Y_D^j(t), t = 1, 2, \dots, N = n - (m_D^j - 1)\tau_D\}$  和  $\{Y_A^L(t), t = 1, 2, \dots, N = n - (m_A^L - 1)\tau_A\}$ , 分别按 (3) 式所表示的加权混沌局域模型,计算出  $Y_D^j(N + 1)$  和  $Y_A^L(N + 1)$  的预测值。从而得到  $A_j(n)$  与  $D_j(n)$  下一步演化的预测值  $D_j(n + 1)$  和  $A_L(n + 1)$ , 它们分别是  $Y_D^j(N + 1)$  和  $Y_A^L(N + 1)$  的最后一个分量。

4) 利用小波重构技术得到信号,即可求出网络流量在  $n + 1$  时刻的预测值  $S(n + 1)$ , 即:

$$S(n + 1) = \sum_{j=1}^L D_j(n + 1) + A_L(n + 1) \quad (6)$$

5) 完成一步预测以后,以  $Y_D^j(N + 1)$  和  $Y_A^L(N + 1)$  为新的预测中心,然后重复以上的预测步骤,则经过  $T$  次迭代后,可得到网络流量在经过  $T$  时刻后的预测值  $S(n + T)$ 。

由整个建模过程不难看出,小波变换将复杂的网络流量时间序列分解成若干不同频率成分的低频信号序列和低频信

号序列,因为网络流量不同的频率成分都有自身的制约因素与发展演化规律,在流量行为演化过程中的作用机制、影响程度完全不一样,所以小波分解与重构正好抓住了网络流量多尺度特征及其各频率成分的发展趋势和对原流量时间序列的不同作用机制,刻画了流量时间序列中的近似部分和细节部分的相互依赖、相互制约的非线性关系。而加权混沌局域模型则考虑到了与预测中心点距离不同的状态点,对中心点动力学行为演化的影响的差异性,在某种程度上克服了传统局域模型是以平等对待的方式来刻画各个邻近状态点与预测中心点动力学行为的相互关系这方面的局限。因此,将小波变换与加权混沌局域模型各自优点结合起来的加权小波混沌局域模型具有更好的非线性逼近功能,从而具有准确拟合复杂网络流量的动力学行为的功能。

#### 4 模型仿真

本文数据来源于新闻组网络系统 USENET 上某个 NNTP (新闻传送协议)服务器从 2005 年 6 月 10 日到 2005 年 8 月 27 日共 79 天的访问流量日志<sup>[10]</sup>,该服务器的 NNTP 协议由一种高效的 innfeed 程序来实现。服务器流量的采样时间粒度为 h,每天 24h,得到 1896h 的流量记录,从而形成了一个网络流量时间序列  $\{s(t), t = 1, 2, \dots, 1896\}$ 。这里将由前 1866 个数据建立小波加权混沌局域模型,其余 30 个的数据用于检验模型的预测效果。流量时间序列  $\{s(t), t = 1, 2, \dots, 1866\}$  的自相关函数图像如图 2 所示,可以看出,该流量相关函数是拖尾的,下降过程比较慢,这个结果在一定程度上说明了网络流量表现出比较复杂长程相关,所以很自然的想到它可能会表现出混沌的行为特性。通过编程计算得知最大 Lyapunov 指数  $\lambda = 0.02632 > 0$ ,这个结果印证了我们的猜想,所以可考虑用本文的模型对其行为未来的演化趋势进行预测。可预测的范围是未来  $T = 1/\lambda = 38\text{h}$  内的网络流量发展趋势,为此可将  $\{s(t), t = 1, 2, \dots, 1866\}$  后续 30 步的行为状态预测出来。

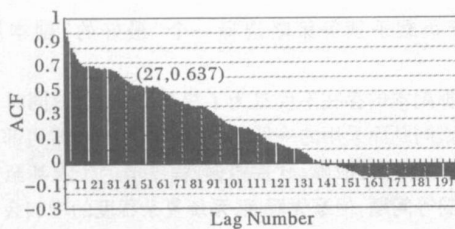


图 2 流量序列的自相关函数

用 Haar 小波函数对流量时间序列  $\{s(t), t = 1, 2, \dots, 1866\}$  进行小波分解,得分解的高频信号序列  $\{D_j(t), t = 1, 2, \dots, 1866\}$  ( $j = 1, 2, 3$ ) 和  $\{A_3(t), t = 1, 2, \dots, 1866\}$ 。显然,小波分解后所得到的低频信号部分  $\{A_3(k)\}$  相对高频信号  $\{D_i(k)\}$  来说是比较平稳的,但计算后发现其最大 Lyapunov 指数为  $0.001041$ ,说明  $\{A_3(k)\}$  仍然具有混沌特性,这与原流量的混沌程度以及小波分解层数较低等因素有关。所以,和对高频信号部分的预测一样,  $\{A_3(k)\}$  也可用加权混沌局域模型进行预测。分别对每个信号成分进行相空间重构,并且用加权混沌局域模型进行预测,最后对每个预测分量进行小波重构,就可把网络流量在未来时刻的预测值确定下来,图 3 列出了小波混沌局域模型对网络流量的预测效果。为了比较小波加权混沌局域模型的预测性能,本文同时采用了传统不加权的混沌预测模型,并且直接以原始流量数据作为模型输入

对网络流量进行预测。

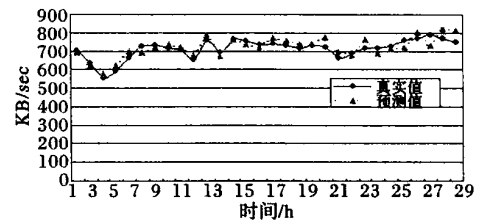


图 3 小波加权混沌局域模型的预测效果

两个模型的相对误差如图 4 所示,平均相对误差分别为 6.630% 和 9.40%,这些结果说明了加权小波混沌模型中的小波变换紧紧抓住了网络流量多尺度特征及其各频率成分的发展趋势和对原流量时间序列的作用机制,而加权混沌局域模型则考虑了和预测中心距离不同的状态点的动力学行为对中心点动力学行为演化的影响的差异性,在一定程度上弥补了传统局域模型的缺陷。

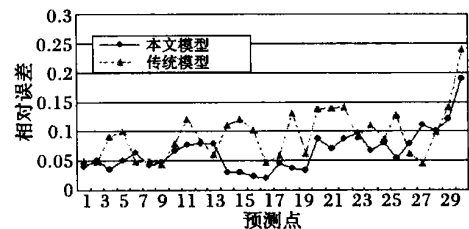


图 4 小波加权混沌局域模型与传统模型预测相对误差比较

#### 5 结语

本文提出了一种新的网络流量预测的小波加权混沌局域模型。小波分析对复杂网络流量具有多尺度分析功能,而加权局部模型具有很强的非线性逼近功能。所以基于两者优势的小波加权混沌局域模型能够比较准确预测复杂网络流量动力学行为的演化趋势。

#### 参考文献:

- [1] FELDMANN A, GLBERT AC, W LLNGER W, *et al* Looking behind and beyond self-similarity: Scaling phenomena in measured WAN traffic [A]. Proc. 35th Annual Allerton Conf. On Comm., Control, and Computing [C]. June 1997. 269 - 280.
- [2] LELAND WE, TAQUU MS, W LLNGER W, *et al* On the self-similar nature of ethernet traffic [J]. IEEE/ACM Transaction on Networking, 1994, 2 (1): 1 - 15.
- [3] W LLNGER W, TAQUU MS, SHERMAN R, *et al* Self-Similarity Through High-Variability: Statistical analysis of ethernet LAN traffic at the source level [A]. Proceedings of the ACM SIGCOMM '95 [C].
- [4] GRASSBERGER P, PROCACCIA J. Dimensions and entropies of strange attractors from a fluctuating dynamics approach [J]. Physics D, 1984, 13: 2146 - 2152.
- [5] TAKENS F. Determining strange attractors in turbulence [J]. Lecture notes in Math, 1981, 898: 339 - 359.
- [6] ABARBANEL HDI. The analysis of observed chaotic data in physical system [J]. Reviews of modern physics, 1993, 65 (4): 1331 - 1392.
- [7] 刘洪,李必强. 基于混沌吸引子的时间序列预测 [J]. 系统工程与电子技术, 1997, (2): 3 - 28.
- [8] 陈森发. 复杂系统建模理论与方法 [M]. 南京:东南大学出版社, 2005. 2: 153 - 174.
- [9] 吴耿锋,周佩玲,储阅春,等. 基于相空间重构的预测方法及其在天气预报中的应用 [J]. 自然杂志, 1999, 21 (2): 107 - 110.
- [10] Usenet daily report [EB/OL]. <http://newsfeed.ntcu.net/~news/>, 2005.