文章编号:1000 - 582X(2004)05 - 0028 - 03

基于最大 Lyapunov 指数的网络业务流量预测*

罗　燕[1,2],汪纪锋[2],曹长修[1]
(1. 重庆大学 自动化学院,重庆　400030; 2. 重庆邮电学院,重庆　400065)

摘　要:在高速网络资源分配与拥塞控制研究中,网络业务流量的预报是一个具有重要意义的课题。基于准确的业务预报,网络管理和控制方案更易于适应业务流量的动态变化,从而达到优化网络性能的目的。而高速网络中大量存在着以自相似性为特征的多种业务流量。已有研究表明,这种自相似特性与混沌现象的吸引子有着紧密的联系。笔者利用混沌时间序列的重构相空间方法,对高速网络中自相似信源的速率做出了预测,并给出了最大可预报时间。该方法的预测模式简单,仿真结果表明,预测的精度也比较高。

关键词:Lyapunov 指数; 自相似性;预测;重构相空间
中图分类号:TP393　　　　**文献标识码**:A

对高速网络中的业务流量进行预测是防止网络拥塞现象发生的一种有效的手段。对业务流量的精确预测,能为各种网络管理方案如互联网中资源预留协议 RSVP 以及 ATM 网中带宽分配方案提供可靠的数据依据,进而提高网络的利用率,并减少拥塞的发生。传统的预测手段有数理统计方法,有基于自回归分数整合滑动平均(FARIMA)模型的方法[1],近来有使用神经网络预测业务量和利用模糊判决规则来预测网络流量的方法[2-3]。

大量的理论分析和仿真实验表明,高速网络中广泛存在着以自相似性为特征的业务流,如视频业务等。最近,人们又发现自相似现象与混沌现象之间存在着紧密的联系,它们的某些特征量具有相同的值。笔者利用 Takens 关于混沌时间序列的重构相空间理论,以及最大 Lyapunov 指数的定义,对仿真自相似业务源的速率做出了预测,并给出了最大可预报时间。仿真表明,预测效果较好。

1　相空间重构

对原始时间序列 x_1, x_2, Λ, x_N,Takens 证明了可以找到一个合适的嵌入维 $m \geq 2d+1$,d 是混沌吸引子的关联维数,以及时间延迟 τ 来重构相空间 R^m 如下:

$$(Y(i) = (x(i), x(i+\tau), x(i+2\tau), \cdots, x(i+(m-1)\tau)), i = 1, 2, \cdots, N \qquad (1)$$

$Y(i)$ 是相空间中第 i 个相点,总的相点数 $M = N - (m-1)\tau$。

如此得到的相空间在拓扑等价意义下与原混沌序列是微分同胚的。相点 $Y(i)$ 的轨迹保持了原混沌序列(系统)的特征。这里的特征是指混沌系统的不变量,一般指 3 个量:关联维数 d,Lyapunov 指数和 Kolmogorov 熵。

在重构相空间中,m 和 τ 的选取十分重要。有多种的计算 d 和 τ 或同时计算出 d 和 τ 的方法。笔者选取标准的 $G-P$ 算法,即利用式(1)的向量 $Y(i)$ 计算关联积分:

$$C_m(r) = \frac{1}{M^2}\sum_{i=1}^{M-1}\sum_{j=1}^{M} H[r - d(Y(i), Y(j))] \qquad (2)$$

式中 M 为相点总数,r 为阀值,H 为阶跃函数,$d(Y(i), Y(j))$ 为向量 $Y(i)$ 和 $Y(j)$ 的距离。

由于(2)式当 $r \to 0$ 时存在关系:$\lim_{r\to 0} C_m(r) \propto r^d$,则关联维数定义为

$$d = \lim_{m\to\infty}\lim_{r\to 0}\frac{\ln C_m(r)}{\ln r} \qquad (3)$$

在实际计算中,先给出一个较小的 m_0(如 4,5 等)对应

* 收稿日期:2003 - 12 - 28
基金项目:受重庆市科委攻关项目"通信网网络管理智能预警模型研究"(2002 - 3303)的资助。
作者简介:罗燕(1966 -),女,重庆大学博士研究生,重庆邮电学院讲师,主要研究方向为网络建模及流量控制。

一个相空间(1)。然后观察由式(2)计算出来的当 r 在某个适当范围内变化时 $\ln C_{m_0}(r)$ 与 $\ln r$ 的比例系数 d 是否常数。再增加嵌入维 $m_1 > m_0$，重复上面的计算过程，直到估计值 d 不再随 m 的增长而变化。此时算出的 d 就是序列 x_1, x_2, x_N 的关联维数。另一方面，如果 d 随 m 的增长不收敛于一个稳定的值，则表明所考虑的序列是一个纯随机序列，而非混沌序列。

对 τ 的选取也有多种算法。在此选择一种简单的算法。考虑到 τ 的含义，可选择原始序列的平均峰值时间(mtbp)作为时间窗 $\tau_w = (m-1)\tau$[5]，τ_W 代表了原始序列的不相关性，由 m 和 τ_w 的值就可以算出 τ 的值。

如上选取 $m \geqslant 2d+1$ 和 τ 就得出了满足条件的重构相空间 $\{Y(i)\}$。

混沌系统的特征量 Lyapunov 指数是从整体上反映动力系统的混沌水平的一个量。对于一个 n 维混沌系统，它存在着指数谱，按照从大到小的顺序排列为 $\lambda_1, \lambda_2, \cdots, \lambda_n$。Wolf 等人基于相轨道线，相平面，相体积等的演化来估计 $\lambda_1, \lambda_2, \cdots, \lambda_n$，提出了 Wolf 数值计算方法。

在指数谱 $\lambda_i (1 \leqslant i \leqslant n)$ 中，λ_1 定量地表现了相空间中两相邻轨道线的分离问题，即两相邻相点 $Y(i)$ 与 $Y(j)$ 的 n 次迭代后的相点 $Y(i+n)$ 与 $Y(j+n)$ 距离分离的平均指数量，如下：

$$|Y(i+n) - Y(j+n)| = d(0) \cdot e^{n\lambda_1} \quad (4)$$

$d(0)$ 是它们的初始距离，采用 Wolf 的轨道跟踪法来估计 λ_1。

此外，如果将式(4)变换为：

$$\frac{|Y(i+n) - Y(j+n)|}{d(0)} = e^{n\lambda_1} \quad (5)$$

设式(5)超过某一临界值 C 时，可以认为轨道发散到使运动不可预言了，这时所经历的时间就是临界时间 t_0，即：

$$C = \frac{|Y(i+n) - Y(j+n)|}{d(0)} = e^{\lambda_1 t_0}$$

从而有：$t_0 = \frac{1}{\lambda_1}\ln C$。通常取 $C = e$ 或更小，则得到最大可预报时间：$t_0 = \frac{1}{\lambda_1}$

2　预测模式

根据 Wolf 计算 λ_1 的方法，可以得出基于最大 Lyapunov 指数的预测模式。首先，在重构相空间中选择预报中心点 $Y(P)$，它的选择应该使得其下一次迭代 $Y(P+1)$ 中含有待预测的量 x_{N+1}，然后寻找它的最近距离的邻点 $Y(K)$、$K < P$，设它们的距离为 $d_P(0)$。

$$d_P(0) = \min_{1 \leqslant j \leqslant M} |Y(P) - Y(j)| = |Y(P) - Y(K)|$$

，这里 $|\cdot|$ 为距离。

根据 λ_1 的物理含义，两相点 $Y(P)$ 与 $Y(K)$ 各自迭代一次后距离分离为：

$$|Y(P+1) - Y(K+1)| = |Y(P) - Y(k)| \cdot e^{\lambda_1} \quad (6)$$

在(6)式中，只有 $Y(P+1)$ 向量中的最后一个分量 x_{N+1} 是未知的，从而可以由此式计算出 x_{N+1}，即原时间序列的一步预测值 x_{N+1}。同样地，选取不同的中心点与不同的迭代次数，还可以做出有限步预测值。

3　仿真实验及讨论

3.1　实验概况

采用 Hosking 算法来实现一个自相似过程。根据 Hosking 算法产生的前 2 000 个值作为样本点 $x_1, x_2, \cdots, x_{2000}$，分别计算出 $m = 13, \tau = 5, \lambda_1 = 0.002\,7$，则 $t_0 = \frac{1}{\lambda_1} = 370(\mathrm{ms})$，再根据预测模式分别做出了后续 20 个点的预测值，并与 Hosking 算法所得的值相比较，如图 1 所示，误差的计算如图 2。

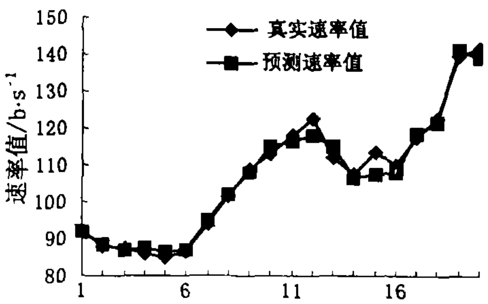

图 1　信源速率值与预测值的比较

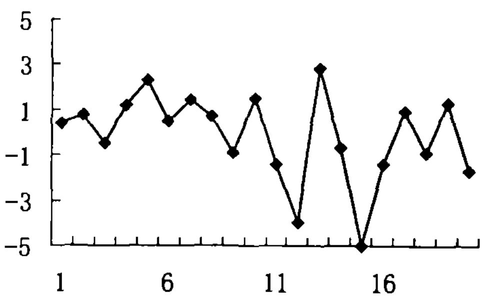

图 2　误差图

3.2　算法讨论

① 预测效果比较好，相对误差均在 -5% ~ 5% 以内。

②算法的复杂度为 $O(M^2)$。算法中主要是关联维数 d 和 λ_1 的计算。对于 d，是求关联积分 $C_m(r)$，即在 $[M \times M]$ 表格上逐点填写 $H = 1$ 的数值。这里的运算量一是来自于

向量距离的运算,二是各距离与不同阀值的比较。对于λ_1,需要计算各个相点间的距离并做出比较。

③d和λ_1的值不受距离的1,2,∞ - 范数表示的影响。为保证预报的唯一性,取欧氏距离作为相点间的距离,这样就增加了计算量。如果用$\| Y(i) - Y(j) \|_\infty$,即最大分量间距来定义距离,虽然减少了计算量,但同时有可能得出预测值x_{N+1}的某个区间,而不是一个值。

④由于计算的复杂度,要做到在线预测x_{N+1}比较难。这有待于改进上述的算法。

4 结 论

由于自相似现象与混沌现象之间的本质联系,从混沌时间序列的角度来研究自相似业务流速率的方法是可行的。基于最大Lyapunov指数的预测方法是直接根据数据序列本身所蕴含的客观规律(如不变量λ_1等)来进行预测,不需要事先建立主观的分析模型,它具有精度高,可信度大的优点。在当今大型网络中广泛分布着自相似的业务流,笔者的工作为如何管理和控制这些业务流量提供了一个可信的数据依据。但由于目前对混沌现象本质的基础研究没有取得突破性的进展,现有的研究和应用主要是基于数值计算来完成的[6]。所以,提高m,τ及λ_1的数值计算方法的可靠性、精度及速度是影响本预测方法应用的一个重要方面。

参考文献:

[1] 刘嘉昆,金志刚,薛飞,等.基于FARIMA过程的网络业务预报与应用[J].电子与信息学报,2001,23(4):403-407.

[2] 王萍,王同胜.ATM网络中使用神经网络的业务预测[J].天津大学学报,1998,31(1):115-119.

[3] 戚文芽,程时昕.一个利用模糊预测的ATM广域网流量控制算法[J].电子学报,1999,27(1):109-111.

[4] AYEDEMIR M, BOTTOMLEY L, COFFIN M, et al. Two Tools for Network Traffic Analysis[J]. Computer Networks, 2001, 36: 169-179.

[5] 杨绍清.一种最大李雅普诺夫指数估计的稳健算法[J].物理学报,2000, 49(4):636-640.

[6] 周越,杨杰.求解关联维数的快速算法研究[J].电子学报,2002,30(10):1 526-1 529.

A Prediction of Network Traffic Flow Based on Lyapunov Exponent

LUO Yan[1,2], *WANG Ji-feng*[2], *CAO Chang-xiu*[1]

(1. College of Automation Chongqing University, Chongqing 400065, China;

2. Chongqing University of Posts and Telecommunications, Chongqing 400030, China)

Abstract: The prediction of network traffic flow is a problem of great significance in the research work of resource allocation and congestion control. Based on this accurate prediction, the scheme of resource allocation and control can easily adapt to dynamic variations of incoming traffic flow. So the goal of optimal network performance is achieved. There are many sorts of traffic flow of self-similarity characteristics in high-speed network, and some research work has showed this self-similarity keeps in close contact with the attractor of chaos system. A rate prediction of self-similar traffic sources in high-speed network is proposed as well as the maximum of predictable time by applying the technology of phase space reconstruction about chaotic time series. This method has a simple prediction mode, and the result of simulations indicates it also has highly accurate results.

Key words: lyapunov exponent; network traffic; prediction; phase space reconstruction

(编辑 吕赛英)