中山大学学报（自然科学版）
第39卷 第6(A)期 ACTA SCIENTIARUM NATURALIUM Vol.39 No.6(A)
2000年 11月 UNIVERSITATIS SUNYATSENI Nov. 2000

文章编号：0529-6579（2000）06-0161-03

# 网络流量的分形研究方法*

许 勇，张 凌，董守斌
（华南理工大学信息网络工程研究中心，广东 广州 510640）

**摘 要：** 介绍了自相似分析方法、分形插值方法、分形预测方法和重分形分析方法在网络流量研究中的应用.

**关键词：** 网络；流量；分形；自相似性；重分形

**中图分类号：** TN915.01 **文献标识码：** A

近年来，网络测量技术的发展使得网络研究人员测量和分析大量的业务数据成为可能. 90年代初期，由美国贝尔实验室和 Boston 大学研究人员所组成的小组采集了大量关于以太网 Ethernet[1]和可变比特率 VBR（variable bit rate）视频服务[2]的业务数据，并利用分形（fractal）的思想对这些数据进行了分析. 研究结果证明：实际网络流量普遍存在统计上的自相似性，该特性与流量发生的时间、地点或信元编码方式无关.

分形理论[3]总结了自然现象中不规则形态的自相似规律，描述了自然界中很广泛一类物质的基本特性：自相似性. 网络流量在数学上可表征为一个随机过程，它所蕴涵的自相似性是统计意义上的. 自相似性反映了网络流量在各个不同的时间标度上的结构相似性. 对局部性态也非常复杂的业务流，应用重分形理论和方法[3,4]可得到其重分形谱函数的描述.

## 1 自相似分析方法

若一个连续时间过程 $Y = \{Y(t), t \in T\}$ 满足

$$Y(t)^d = a^{-H}Y(at), t \in T, \forall a > 0, 0 \leqslant H < 1 \tag{1}$$

则称 $Y$ 为自相似过程[1]，等式的含义是指其有限维分布相同，$H$ 称为 Hurst 系数.

我们一般处理离散形式的网络业务流序列，它的定义如下，设 $X = \{X(i), i \geqslant 1\}$ 为平稳序列，其 $m$ 阶平滑过程定义为

$$X^{(m)}(k) = \frac{1}{m}\sum_{i=(k-1)m+1}^{km} X(i), k = 1,2,\cdots \tag{2}$$

如果序列 $X$ 是一满足式(1)过程 $Y$ 的差分过程，即 $X(i) = Y(i+1) - Y(i)$，则对所有的 $m$，我们有

$$X^d = m^{1-H}X^{(m)} \tag{3}$$

称满足式（3）的平稳序列为严格自相似的.

* **基金项目：** 国家863计划资助项目（863-317-01-03-99）；华南理工大学自然科学基金资助项目.
**收稿日期：** 2000-09-18；**作者简介：** 许 勇（1971～），男，博士后；**研究方向：** 分形几何理论及其应用、网络行为分析、数学模型.

## 2 分形插值方法和分形预测方法

### 2.1 分形插值方法

一般情况下，我们测量得到的是离散的时间－流量坐标数据．线性分形插值方法[5]是利用已知坐标数据模拟整体函数图象的一种有效方法．

设观察数据坐标为 $\{(x_i,y_i)\}$，$i=(0,1,\cdots,N)$，$x_i\in[x_0,x_N]$，$y_i\in[a,b]$，我们要找一个分形插值函数 $f(x)$，使得 $f(x):[x_0,x_N]\to[a,b]$ 是连续函数，且 $f(x_i)=y_i$，$i=(0,1,\cdots,N)$．对 $i=(1,2,\cdots,N)$，定义

$$\begin{cases} L_i(x)=x_{i-1}+(x_i-x_{i-1})(x-x_0)/(x_N-x_0) \\ F_i(x,y)=b_ix+Hy+k_i \end{cases} \tag{4}$$

其中 $H$ 为 Hurst 系数，

$$\begin{cases} b_i=[y_i-y_{i-1}-H(y_N-y_0)]/(x_N-x_0) \\ k_i=y_{i-1}-Hy_0-b_ix_0,\ i\in\{1,2,\cdots,N\} \end{cases} \tag{5}$$

令

$$T_i(x,y)=(L_i(x),F_i(x,y)),\ i=1,2,\cdots,N \tag{6}$$

则 $\{T_i(x,y)\}$，$i=(1,2,\cdots,N)$ 构成一个迭代函数系，其吸引子是我们所需要的分形插值函数 f (x) 的图象．有关迭代函数系及其吸引子的具体描述可参见文献 [3]．

### 2.2 分形预测方法

应用上面线性分形插值函数的思想，可以构造线性分形预测方法，有效地实现对未来流量数据的预测．

如果已经知道前 $(n+1)$ 个时间—流量坐标数据，因网络业务流具有自相似结构，可得到此流量数据对应的 Hurst 参数 $H$，同时也可以估计出在第 $(n+2)$ 时刻网络流量数据值．由式 (1)，可得

$$y_{n+1}^i=\left(\frac{x_i}{x_{n+1}}\right)^{-H}y_i,\ i=0,1,\cdots,n \tag{7}$$

令

$$y_{n+1}=\frac{1}{n+1}\sum_{i=0}^{n}y_{n+1}^i \tag{8}$$

即取 $y_{n+1}$ 为 $\{y_{n+1}^i,\ i=(0,1,\cdots,n)$ 的统计平均值，于是得到坐标数据 $(x_{n+1},y_{n+1})$．这样，通过插值可得到定义域为 $[x_n,x_{n+1}]$ 上的流量数据．

## 3 重分形理论和分析方法

设 $n\in N$，$\mu$ 是定义在单位区间 $[0,1]$ 上的一个概率分布，将 $[0,1]$ 等分为 $2^n$ 个长度均为 $2^{-n}$ 的小区间．定义

$$I_{n,k}=[k2^{-n},(k+1)2^{-n}],\ k=0,1,\cdots,2^n-1 \tag{9}$$

若 $x\in[0,1]$，不妨设 $x\in I_{n,k}$，记为 $I_n(x)$．则点 $x$ 的 Hölder 指数定义为

$$\alpha_h(x)=\lim_{n\to\infty}\frac{\lg\mu(I_n(x))}{\lg|I_n(x)|}=-\lim_{n\to\infty}\frac{\log_2\mu(I_n(x))}{n} \tag{10}$$

从定义可以看出，当 $n$ 足够大时，我们有 $\mu(I_n(x))\approx|I_n(x)|^{\alpha_h(x)}$，故 Hölder 指数 $\alpha_h(x)$ 表示了一点(或局部区间)的特征．

对 $q \in R$,定义

$$\tau_h(q) = \frac{\lg \sum_{0 \le j < 2^n} \mu(I_{n,j})^q}{\lg |I_n(x)|} = \frac{\log_2 \sum_{0 \le j < 2^n} \mu(I_{n,j})^q}{n} \tag{11}$$

令 $\tau_h(q) = \lim_{n \to \infty} \inf \tau_h(q)$.对非负实数 $\alpha$,定义

$$E_\alpha = \{x : \alpha_h(x) = \alpha\}, f(\alpha) = \dim E_\alpha \tag{12}$$

其中 $\dim E_\alpha$ 表示 $E_\alpha$ 的 Hausdorff 维数.函数 $f$ 称为概率分布 $\mu$ 的重分形谱函数. 从定义可以看出,重分形谱函数 $f$ 全面地刻划出了概率分布 $\mu$ 的整体特征.应用重分形理论可以证明[3],$f$ 是 $\tau$ 的勒让德变换,即

$$f(\alpha) = \inf_{q \in R}(q\alpha - \tau(q)) \tag{13}$$

**参考文献:**


[1] LELAND W. On the self-similar nature of Ethernet traffic (extended version)[J]. IEEE/ACM Transactions on Networking, 1994,2(1):1.

[2] BERAN J. Long-range dependence in variable bit rate video traffic[J]. IEEE Transactions on Communications, 1995,43(2/3/4):1566.

[3] FALCONER K J. Fractal geometry: mathematical foundation and applications[M]. New York: John Wiley and Sons,1990.

[4] RIEDI R H. An improved multifractal formalism and self-similar measures[J]. J Math Anal Appl, 1995,189: 462.

[5] BARNSLEY M F. Fractal functions and interpolation[J]. Const Approx,1986,2:303.


## Fractal Methods in the Study of Network Traffic


*XU Yong* *, *ZHANG Ling*, *DONG Shou-bin*



**Abstract**: With the high development of computer network, the study of the network traffic is deeper. Now fractal theory is one of the most important methods in the research of the traffic. This paper introduces the applications of self - similar analysis method, the fractal interpolation method, the fractal prediction method and multifractal theory in the study of the traffic.



**Keywords**: Network; traffic; fractal; self-similar; multifractal



* Information Network Engineering and Research Center, South China University of Technology, Guangzhou 510640, China