

多路自相似流输入下基于 FBM 模型的丢包率分析

刘畅,周武旻

(中国科学技术大学无线网络通信实验室,安徽合肥 230027)

摘要:从分形布朗运动(FBM)模型出发,结合高斯输入下的丢包性能,推导了多路自相似流输入下丢包率的理论分析.结果显示,在多路自相似流输入情况下,若各条流的均值、方差近似相等,则 Hurst 参数最大的流对合成流的丢包率具有决定性的影响.该结论通过仿真得到了很好的验证.所提出的分析方法和得出的结论对于网络节点性能分析、缓冲区设计及资源调度等具有实际的指导意义.

关键词:多路输入;自相似流;丢包率

中图分类号:TP393.02 **文献标识码:**A **doi:**10.3969/j.issn.0253-2778.2010.01.014

Packet loss rate analysis based on FBM model with multiple self-similar traffic input

LIU Chang, ZHOU Wuyang

(Wireless Information Network Lab, USTC, Hefei 230027, China)

Abstract: Theoretical analysis of packet loss rate with multiple self-similar traffic input was proposed based on FBM model and the condition with Gaussian input. The results show that if the mean and variance of each traffic flow are approximately equal, the traffic flow with the largest Hurst parameter has a decisive influence on the packet loss rate of multiple traffic. This conclusion was verified by simulation and can be used for the analysis of network node performance, buffer design, and resource allocation.

Key words: multiple input; self-similar traffic; packet loss rate

0 引言

丢包率是网络服务质量(QoS)的重要指标之一,也是长期以来网络领域一直研究的重要课题.在网络中造成丢包的原因有很多,如信道质量下降或网络拥塞,然而在有线数据网络中,一般认为拥塞是造成丢包的主要原因.因此对拥塞造成的丢包性能分析可用于衡量网络的拥塞状况,评估网络对不同业务的承载能力,进一步可指导缓冲区的队列调度策略等.

大量研究表明,分组交换的网络中业务流普遍具有自相似性^[1-4].这些具有自相似性的流在不同的时间尺度上均呈现突发性,不会随着时间尺度的增大而被平滑掉.因此,传统的针对泊松过程的排队理论不再适用于分析自相似流的排队和丢包性能,现有的分析一般建立在业务流的自相似模型基础上,如分形布朗运动(FBM)模型^[5].FBM是自相似高斯过程, Kim等分析了有限缓存在高斯过程输入下的丢包率^[6],因此其结论可以用来分析网络中自相似业务流输入情况下的丢包性能.

收稿日期:2008-05-14;修回日期:2008-06-20

基金项目:国家重点基础研究发展(973)计划(2007CB310602),中国瑞典国际合作基金(2008DFA11950),中国高技术研究发展(863)计划(2009AA011506),IMT-Advanced 关键技术研究(2008BAH30B12),IMT-Advanced 试验平台(2009ZX03003-009)资助.

作者简介:刘畅,男,1983年生,博士生.研究方向:无线资源管理、业务建模. E-mail: changool@mail.ustc.edu.cn

通讯作者:周武旻,博士/教授. E-mail: wyzhou@ustc.edu.cn

然而,在实际网络中有时我们关心的是多路自相似流输入下的丢包率,比如核心路由器接收来自不同边缘路由器的多路自相似流,在网络优化或线路合并时某节点路由器可能会增加一路新的自相似流输入等,这些情况下在缓冲区设计及出口带宽规划时就需要考虑多路输入的丢包行为。

因此,我们分析了多路自相似流输入下的丢包特性,推导了相应的丢包率公式,分析得到了一些有意义的结论,通过仿真结果也进一步佐证了结论的合理性和正确性.本文的分析方法和结论对于网络节点性能分析、缓冲区设计及资源调度等具有实际的指导意义。

1 数学分析

1.1 FBM模型简介

标准分形布朗运动过程 $\{Z_H(t):t \geq 0\}$ 满足下面条件:

- ① $Z_H(0) = 0$;
- ② $\text{Var}\{Z_H(t)\} = t^{2H} (H \in [0.5, 1))$;
- ③ 具有如下协方差结构: $\text{Cov}(Z_H(t), Z_H(s)) =$

$\frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$. Norros^[5]给出了基于分形布朗运动的自相似业务流模型,定义如下:

$$A(t) = mt + \sqrt{an}Z_H(t), t \in (-\infty, +\infty) \quad (1)$$

其中, $A(t)$ 表示到时刻 t 为止到达业务流的总量,它由三个参数来刻画: m, a, H . $m > 0$ 为平均到达速率, $a > 0$ 为方差系数,它们有如下关系:

$$a = \sigma^2 / m \quad (2)$$

其中, σ^2 为一个时间单元内的业务流量的方差. $H \in [0.5, 1)$ 是 $Z_H(t)$ 的自相似参数(Hurst 参数).

1.2 单路输入的丢包率分析

设无限缓存系统中节点队列长度的分布为 $P(X > x)$, X 表示节点队列长度,以信息包(packet)为单位且包长相等,则有限缓存条件下丢包率 $P_L(x)$ 为^[6]

$$P_L(x) = \alpha P(X > x) \quad (3)$$

其中, x 为缓冲区大小,单位同样为 packet. α 是一个常量且有:

$$\alpha = P_L(0) / P(X > 0).$$

在高斯输入情况下,常量 α 可以通过下式算出

$$\alpha = \frac{1}{m \sqrt{2\pi\sigma}} \exp\left(-\frac{(c-m)^2}{2\sigma^2}\right) \times$$

$$\int_c^\infty (r-c) \exp\left(-\frac{(r-m)^2}{2\sigma^2}\right) dr \quad (4)$$

式中, c 为服务速率,可以认为是网络节点的输出速率。

Norros 在文献[5]中给出了基于 FBM 模型的自相似流输入下的无限缓存节点的队列长度分布,近似为

$$P(X > x) = \exp\left(-\frac{(c-m)^{2H}}{2\kappa(H)^2\sigma^2} x^{2-2H}\right) \quad (5)$$

其中, $\kappa(H) = H^H(1-H)^{1-H}$. 将式(5)代入式(3)就可以得到丢包率 $P_L(x)$

$$P_L(x) = \alpha \exp\left(-\frac{(c-m)^{2H}}{2\kappa(H)^2\sigma^2} x^{2-2H}\right) \quad (6)$$

1.3 多路输入的丢包率分析

由式(6)可得到单路自相似流输入下的丢包率,然而实际网络中输入某节点的自相似流有时不止一条,因此在缓冲区设计及出口带宽规划时,往往需要考察多路自相似流输入时节点的丢包特性。

假设节点的输入是 N 路自相似流,每一路均以 FBM 模型的三个参数 m, a, H 来描述,分别记为

$$X_i(m_i, a_i, H_i) (i = 1, \dots, N),$$

合成流记为

$$\mathbf{X}(m, a, H) = \sum_{i=1}^N X_i(m_i, a_i, H_i).$$

在实际网络中,来自不同链路的流可认为相互独立.故本文将 N 路自相似流看作 N 个独立的随机过程,可得

$$m = \sum_{i=1}^N m_i \quad (7)$$

$$a = \frac{\sum_{i=1}^N m_i a_i}{\sum_{i=1}^N m_i} \quad (8)$$

若互不相关的两个自相似过程 X_1 和 X_2 分别具有 Hurst 参数 H_1 和 H_2 ,则 $X_1 + X_2$ 为具有 Hurst 参数 $\max(H_1, H_2)$ 的自相似过程^[7]. 因此,合成后的 Hurst 参数为

$$H = H_{\max} = \max(H_1, H_2, \dots, H_N) \quad (9)$$

(9)式成立的条件是假设序列长度无限,对于实际的有限长队列合成后的 H 值只能近似等于其中的最大值. 本文的仿真结果表明合成流的 Hurst 值略小于理论值,相差不大,故式(9)可作为合成后 Hurst 参数的近似估计。

综上所述,多路输入情况下网络节点的丢包率可表示为

$$P_L(x) = \alpha P(\mathbf{X} > x), \mathbf{X} = \sum_{i=1}^N X_i \quad (10)$$

其中,

$$P(\mathbf{X} > x) = \exp \left[- \frac{(c - \sum_{i=1}^N m_i)^{2H_{\max}}}{2\kappa(H_{\max})^2 \sum_{i=1}^N \sigma_i^2} x^{2-2H_{\max}} \right] \quad (11)$$

$$\alpha = \frac{1}{(\sum_{i=1}^N m_i) \sqrt{2\pi \sum_{i=1}^N \sigma_i^2}} \exp \left[- \frac{(c - \sum_{i=1}^N m_i)^2}{2 \sum_{i=1}^N \sigma_i^2} \right] \times \int_c^\infty (r - c) \exp \left[- \frac{(r - \sum_{i=1}^N m_i)^2}{2 \sum_{i=1}^N \sigma_i^2} \right] dr \quad (12)$$

2 公式解析

由式(10)~(12)可知,丢包率与多个条件有关,包括各路自相似流的均值、方差、Hurst 值以及节点服务速率、缓冲区大小等. 本文主要考察 Hurst 参数对合成流丢包率的影响,为简化分析,后续的公式解析及仿真中均假设组成合成流的各单路输入流具有相等的均值和方差. 虽然实际网络中的各单路流可能具有不同的均值、方差及 Hurst 值,但上述公式中并没有对各单路流之间参数的关系作出约束,因此仍然可以利用本文的公式及分析方法进行分析,得到合成流的丢包率特性.

不失一般性,本节分别针对单路输入、Hurst 参数相同的两路输入、Hurst 参数不同的两路输入情况用 Matlab 进行丢包率公式解析,得出结论:①在保证链路利用率不变的情况下,若两条流的均值、方差相近,自相似性较小的数据流在与自相似性较大的流混合输入时,丢包率会增大;②两条流混合输入下,平均丢包率与 Hurst 参数较小的流的 Hurst 值无关.

2.1 单路输入的丢包率公式解析

假设系统稳定的前提下,平均到达速率 m 不能大于节点服务速率 c . 记 $\rho = m/c$ 为链路利用率,有 $\rho \leq 1$. 本文后面的分析和仿真中设缓冲区大小为 B , B 的单位为 packet, m 和 c 的单位均为 packet/s.

下面我们先考虑丢包率和缓冲区大小及 Hurst 参数间的关系. 取 $m=300, \sigma^2=100, c=310$, 不同 H 下丢包率随 B 的变化曲线如图 1 所示.

从图 1 中可以看出,对于不同的 H , 丢包率随

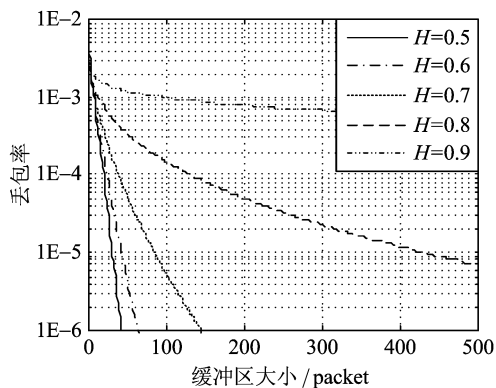


图 1 丢包率随缓冲区大小变化曲线

Fig. 1 Variation of packet loss rate with buffer size

缓冲区增大而下降的速度各不相同, H 越大下降速度越慢. 当 $H=0.5$ 时, $B=50$ 的丢包率就小于 10^{-6} , 而 $H=0.9$ 时, $B=500$ 的丢包率仍大于 10^{-4} .

再考虑丢包率随着链路利用率增加的情况. 取 $\sigma^2=100, c=310, H=0.8$, 图 2 给出了缓冲区大小不相同, 丢包率随着平均到达率的提高(即链路利用率的提高)的变化曲线. 可见,随着链路负载加重,缓冲区大小对丢包率的影响逐渐减小,当 m 趋于 c 时丢包率趋于相等.

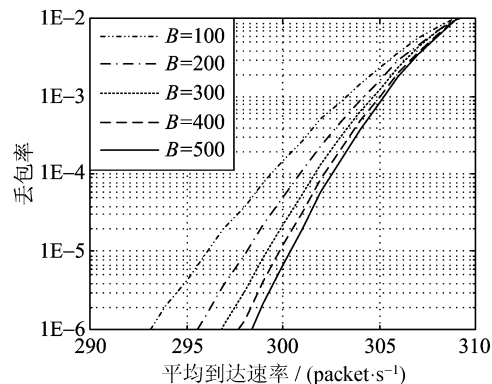


图 2 丢包率随平均到达率变化曲线

Fig. 2 Variation of packet loss rate with mean input rate

2.2 多路输入的丢包率公式解析

2.2.1 Hurst 相同的两路流叠加

(I) 取 $m_1=m_2=300, \sigma_1^2=\sigma_2^2=100, c_1=c_2=310, H_1=H_2=0.5$. 假设两路输入下的链路利用率不变,即合成流的输出速率 $c=c_1+c_2$. 合成流的丢包率与单一流的对比曲线如图 3(a)所示.

(II) 取 $m_1=m_2=300, \sigma_1^2=\sigma_2^2=100, c_1=c_2=310, H_1=H_2=0.8$. 同样假设两路输入下的链路利用率不变,合成流的丢包率与单一流的对比曲线如图 3(b)所示.

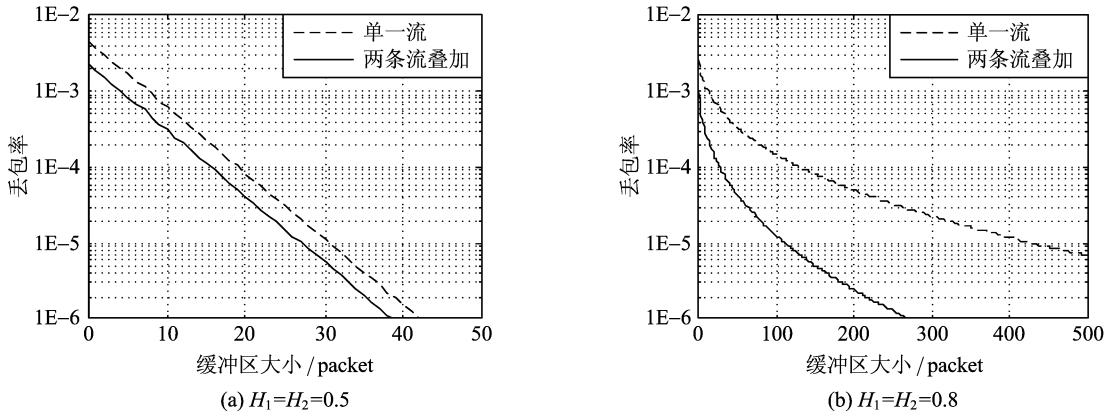


图 3 Hurst 相同的流叠加的丢包率解析

Fig. 3 Packet loss rate analysis of traffic flows with the same Hurst

由图 3 可以看出,均值、方差及 Hurst 参数相等的两条流叠加输入,在保证链路利用率不变的情况下,合成流的平均丢包率相对于单一流均有所降低,没有自相似性的两条流(即 $H=0.5$)合成后降低的幅度较小,而 $H=0.8$ 时合成流的丢包率比单一流降低得多。

2.2.2 Hurst 不同的两路流叠加

取 $m_1=m_2=300, \sigma_1^2=\sigma_2^2=100, c_1=c_2=310, H_1=0.5, H_2=0.8$, 两路输入情况下的链路利用率不变。单条流和叠加流的丢包率对比曲线如图 4 所示。

从图 4 可以看出,叠加流对应的平均丢包率小

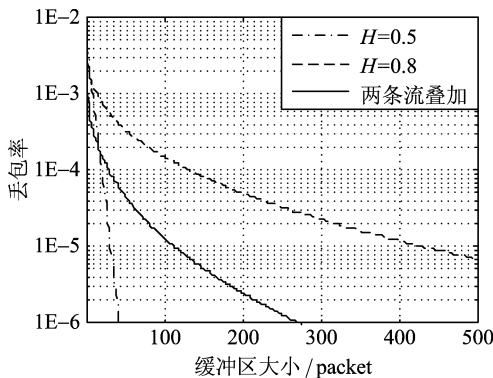


图 4 Hurst 不同的流叠加的丢包率解析

Fig. 4 Packet loss rate analysis of traffic flows with different Hurst

于 $H=0.8$ 的单一流的丢包率,但大于 $H=0.5$ 的单一流,即在保证链路利用率不变的情况下,若两条流的均值、方差相近,自相似性较小的数据流在与自相似性较大的流混合输入时,丢包率会增大。

更进一步,综合图 4 与图 3(b)可以看出其叠加流的丢包率曲线相同,结合式(9)可以得出:两条流混合输入下,平均丢包率与其中 Hurst 参数较小的流的 Hurst 值无关。

3 仿真实验

下面我们对本文提出的多路输入下的丢包率公式及其解析所得的结论进行仿真验证,仿真利用分形高斯噪声(FGN)^[8]产生不同 m, σ^2 和 H 的自相似流,各条流彼此独立,以此来模拟实际网络中的数据流.分别计算它们的均值、方差,并用周期图法^[1]估计 Hurst 参数.之后计算这些数据流在不同混合情况下的均值、方差和 Hurst 参数,并与式(7)~(9)进行比较。

由表 1 可以看出,不同自相似流合成后的均值、方差、Hurst 基本满足式(7)~(9),即合成流的均值以及方差等于各独立流的均值之和以及方差之和,合成流的 Hurst 等于各独立流 Hurst 的最大值.从仿真结果来看,其结果与公式的理论值之间存在一定误差,这是由于仿真中的数据不可能趋近于无限

表 1 不同数据流的混合特性

Tab. 1 Parameters of heterogeneous traffics

	组 1			组 2			组 3		
	均值	方差	Hurst	均值	方差	Hurst	均值	方差	Hurst
流 1	300	100	0.604 1	400	200	0.509 2	100	400	0.697 5
流 2	300	100	0.806 9	500	300	0.873 8	250	300	0.604 9
合成流	600	198.8	0.788 4	900	494.7	0.864 9	350	696.3	0.689 4

所致,不过误差并不大,因此 $H = \max(H_1, H_2, \dots)$ 仍然可以作为 Hurst 参数的近似估计。

同样的,为不失一般性,本文对 Hurst 参数相同和不同两种情况下的丢包率进行了仿真验证。

(I) Hurst 参数相同

设输入的两条流有 $m_1 = m_2 = 300$, $\sigma_1^2 = \sigma_2^2 = 100$, $H_1 = 0.8069$, $H_2 = 0.8026$, 合成流的输出速率 $c = 620$ 。图 5 给出了根据式(10)~(12)计算出的理论曲线和仿真得到的实际曲线对比,结果显示二者基本接近。

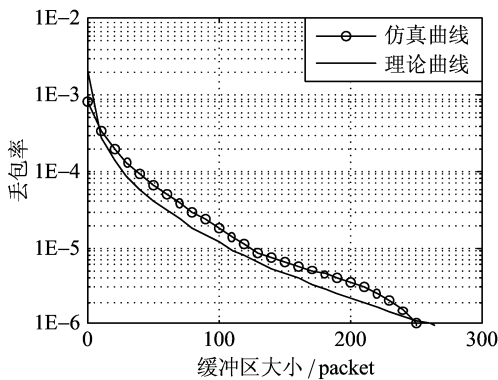


图 5 相同 Hurst 参数的仿真结果

Fig. 5 Simulation result with the same Hurst

(II) Hurst 参数不同

针对两条不同 Hurst 参数的自相似流混合,同样取 $m_1 = m_2 = 300$, $\sigma_1^2 = \sigma_2^2 = 100$, 合成流的输出速率 $c = 620$, $H_1 = 0.8069$, H_2 分别等于 0.6126 和 0.7071。丢包率曲线对比如图 6 所示,结果显示理论曲线和仿真得到的实际曲线基本接近。

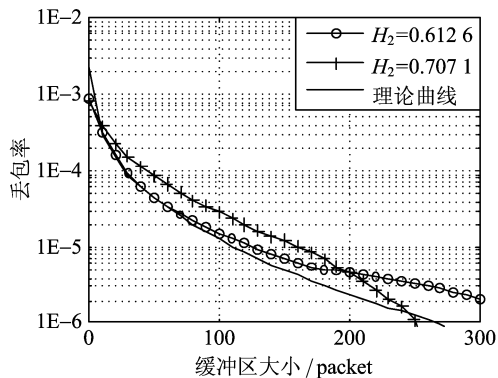


图 6 不同 Hurst 参数的仿真结果

Fig. 6 Simulation result with different Hurst

综合图 5 和图 6 可以看出, $H = 0.8069$ 的流和 H 为 0.8026、0.6126、0.7107 的流分别混合输入,丢包率与两条 $H = 0.8069$ 的流混合的丢包率接近,验证了本文提出的多路自相似流输入下节点丢

包率公式的准确性和所得结论的正确性。

4 结论

本文基于网络自相似流的 FBM 模型,在高斯输入下丢包率公式的基础上,推导并分析了多路自相似流输入下的丢包率公式,得出一个有意义的结论,即多路自相似流输入下,若各独立流的均值、方差近似相等,则 Hurst 参数最大的流对合成后的丢包率具有决定性的影响。分析和仿真结果也表明了公式的准确性和结论的正确性。本文的公式及分析方法可以用来分析实际网络中多路自相似流混合输入下节点的丢包性能,进而对网络规划、网络中多路汇聚流量分析、网络节点缓冲区设计以及资源调度等提供有益的借鉴和指导。

参考文献(References)

- [1] Leland W E, Taqqu M S, Willinger W, et al. On the self-similar nature of Ethernet traffic (extended version) [J]. IEEE/ACM Transactions on Networking, 1994, 2(1):1-15.
- [2] Beran J, Sherman R, Taqqu M S, et al. Long-range dependence in variable-bit-rate video traffic[J]. IEEE Transactions on Communications, 1995, 43 (234): 1 566-1 579.
- [3] Field A J, Harder U, Harrison P G. Measurement and modelling of self-similar traffic in computer networks [J]. IEE Proceedings- Communications, 2004, 151 (4):355-363.
- [4] Karagiannis T, Molle M, Faloutsos M. Long-range dependence ten years of Internet traffic modeling[J]. IEEE Internet Computing, 2004, 8(5):57-64.
- [5] Norros I. On the use of fractional Brownian motion in the theory of connectionless networks [J]. IEEE Journal on Selected Areas in Communications, 1995, 13(6):953-962.
- [6] Kim H S, Shroff N B. Loss probability calculations and asymptotic analysis for finite buffer multiplexers [J]. IEEE/ACM Transactions on Networking, 2001, 9(6):755-768.
- [7] Xiao Junfang, Zeng Guihua. Structural modeling of heterogeneous traffics [C]// Proc. of Wireless Communications, Networking and Mobile Computing. Wuhan, 2005, 2:1 028-1 032.
- [8] Paxson P. Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic[J]. Computer Communication Review, 1997, 27(5):5-18.