

文章编号:1000-582X(2008)02-0175-04

IP 网络流量产生器的研究与实现

陈 晨,裴昌幸,陈 南,易运晖

(西安电子科技大学 ISN 国家重点实验室,西安 710071)

摘 要:分析了造成 IP 网络流量自相似性的根本原因以及 ON/OFF 源汇聚程度对自相似度的影响。针对上述分析结论,实现了一个能够单点或分布式产生 IP 网络自相似业务量的产生器。实际环境中测试显示,该产生器能够较好地反应文件大小,协议类型,ON/OFF 逗留时间以及源汇聚程度与自相似度之间的关系。

关键词:IP 网络;ON/OFF 模型;业务量产生器;自相似

中图分类号:TN915

文献标志码:A

Study and Implementation of Traffic Generator in IP Network

CHEN Chen, PEI Chang-xing, CNEN Nan, YI Yun-hui

(National Key Lab of Integrated Service Networks, XiDian University, Xi'an 710071, P. R. China)

Abstract: According to the simulation results, the factors causing the forming of self-similarity in IP network is presented. Through analysis based on ON/OFF model, we show the influence of traffic source integration level on self-similarity degree. Using the above results, an IP network traffic generator is implemented. Tests in real environment show that the generator can better reflect the relationship of file sizes, protocol types, ON/OFF residence time, traffic source integration levels and self-similarity degrees.

Key words: IP network; ON/OFF model; traffic generator; self-similarity

许多分组交换网络中的信息流具有长相关性/自相似性^[1-7],为了对自相似业务流的排队特性进行研究,需要生成自相似随机序列。到目前为止,人们提出了许多产生自相似随机序列的方法。RMD (random midpoint displacement)^[8]算法是其中的一种方法,此种方法实现简单、快速,但准确性较差;另一种自相似业务流的生成方法(以下称 Hosking^[9]方法)除可以生成 FGN(fractional Gaussian noise)外,还可以生成具有渐近自相似特性的业务流,这主要由相关函数 $r(k)$ 来决定。Hosking 方法为在线生成自相似业务流的方法,需要的存储空间较多,计算量也

较大;还有就是 ON/OFF 源汇聚模型^[10],作为一种基于用户行为模拟的产生方法,其不仅具有明确的物理意义,而且通过简单的改变分布和汇聚参数就可以达到调整自相似度的目的,因此非常适合于实际环境中产生自相似业务量。

首先对 IP 网络中造成流量自相似性的因素进行了分析;然后利用分析的结果,对 ON/OFF 源汇聚模型展开讨论,并通过仿真,分析 ON/OFF 源汇聚模型的性能以及分布参数与网络实体参数之间的关系;最后给出一个较为完善的 ON/OFF 自相似业务量产生器的实现方案。

收稿日期:2007-09-20

基金项目:国家自然科学基金重点项目(60132030);国家自然科学基金重点项目(60572147);西安市工业科技攻关项目(GG04018)

作者简介:陈 晨(1977-),男,西安电子科技大学博士生,主要从事无线网络测量,接入控制方面的研究。裴昌幸(联系人),教授,博士生导师,(Tel)029-88204486;(E-mail)chxpei@xidian.edu.cn。

1 IP 网络中业务量自相似原因分析

1.1 问题提出背景

网络业务流呈现出自相似的特性,引起了网络业务分析的新热潮。不管是网络业务流模型还是网络性能分析,不得不重新考虑这种新的特性。当然,对于产生这种业务流的原因更是一个迫切需要解决的问题。文中通过大量仿真,对 IP 网络中业务量自相似性的成因做出了详尽的分析,为实现一个能够真实反映 IP 网络业务量特性的业务量产生器提供了理论依据。

1.2 仿真分析

首先给出重尾分布的定义:

$$P(X > x) \sim x^{-\alpha} \text{ 当 } x \rightarrow \infty, \quad (1)$$

其中 $0 < \alpha < 2$,也就是说如果忽略一些值比较小的随机变量的分布特性,大部分的随机变量分布都近似于幂律分布的话,就称这个随机变量服从重尾分布。常用的一种简单地重尾分布是 Pareto 分布,它的 pdf 如下:

$$P(x) = \alpha k^\alpha x^{-\alpha-1}, \quad (2)$$

其中 $\alpha, k > 0$ 且 $x \geq k$ 。它的 cdf 如下:

$$F(x) = P[X \leq x] = 1 - (k/x)^\alpha, \quad (3)$$

其中 k 称为截止参数,代表随机变量的最小可能取值。

重尾分布与常规分布例如正态分布,指数分布相比有许多特殊性质。如果 $\alpha < 2$,重尾分布具有无限方差;假如 $\alpha \leq 1$,重尾分布还具有无限均值。因此,随着 α 减小,大部分的概率质量集中在分布曲线的尾部。另外,用 $\text{pareto}(k, \alpha)$ 来完全确定一个 pareto 分布。其中, k 为截止参数, α 为形状参数。

下面利用 opnet 来作为实验仿真平台,并利用 R/S 估计法来评估 Hurst 参数。

实验 1:文件大小(size)服从 pareto 重尾分布, ON 和 OFF 逗留时间服从指数分布,仿真时间 5 min。仿真结果见表 1。

表 1 ON,OFF 逗留时间服从指数分布,文件大小服从 pareto 分布时的 Hurst 参数

ON 分布	OFF 分布	文件大小分布	Hurst
Exp(0.1)	Exp(0.1)	Pareto(1 024,1.8)	0.611
Exp(0.1)	Exp(0.1)	Pareto(1 024,1.5)	0.781
Exp(0.05)	Exp(0.1)	Pareto(1 024,1.5)	0.792

实验 2:文件大小(size)服从指数分布, ON 和 OFF 逗留时间服从 pareto 分布,分别采用 udp 协议和 tcp 协议,仿真时间 5 min。仿真结果见表 2~3。

表 2 udp 协议下 Hurst 参数仿真结果

ON 分布	OFF 分布	文件大小分布	Hurst
Pareto(0.1,1.6)	Pareto(0.1,1.6)	Exp(1 024)	0.689
Pareto(0.1,1.4)	Pareto(0.1,1.4)	Exp(1 024)	0.797
Pareto(0.1,1.2)	Pareto(0.1,1.2)	Exp(1 024)	0.868

表 3 tcp 协议下 Hurst 参数仿真结果

ON 分布	OFF 分布	文件大小分布	Hurst
Pareto(0.1,1.6)	Pareto(0.1,1.6)	Exp(1 024)	0.734
Pareto(0.1,1.4)	Pareto(0.1,1.4)	Exp(1 024)	0.829
Pareto(0.1,1.2)	Pareto(0.1,1.2)	Exp(1 024)	0.945

实验 3:文件大小服从 pareto 重尾分布, ON 和 OFF 逗留时间服从 pareto 分布,仿真时间 5 min。仿真结果见表 4~5。

表 4 ON/OFF 逗留时间,文件大小均服从 pareto 分布时 Hurst 参数

ON 分布	OFF 分布	文件大小分布	Hurst
Pareto(0.1,1.6)	Pareto(0.1,1.6)	Pareto(1 024,1.6)	0.663
Pareto(0.1,1.4)	Pareto(0.1,1.4)	Pareto(1 024,1.6)	0.676
Pareto(0.1,1.2)	Pareto(0.1,1.2)	Pareto(1 024,1.6)	0.721

表 5 ON/OFF 逗留时间,文件大小均服从 pareto 分布时 Hurst 参数

ON 分布	OFF 分布	文件大小分布	Hurst
Pareto(0.1,1.6)	Pareto(0.1,1.6)	Pareto(1 024,1.4)	0.788
Pareto(0.1,1.6)	Pareto(0.1,1.6)	Pareto(1 024,1.2)	0.889
Pareto(0.1,1.6)	Pareto(0.1,1.6)	Pareto(1 024,1.1)	0.936

1.3 仿真结果分析

由表 1 可见,当 ON 和 OFF 的逗留时间服从指数分布的时候,文件大小的重尾分布基本上完全决定了生成业务量的自相似性。由表 1 的 2,3 行可见,随着文件大小分布的 α 值减小,Hurst 参数逐渐增加;而通过对表 1 的 3,4 行数据对比发现,在文件大小的分布参数不变的情况下,ON 和 OFF 逗留时间分布的改变对最终输出的业务量的自相似性的影响很小。

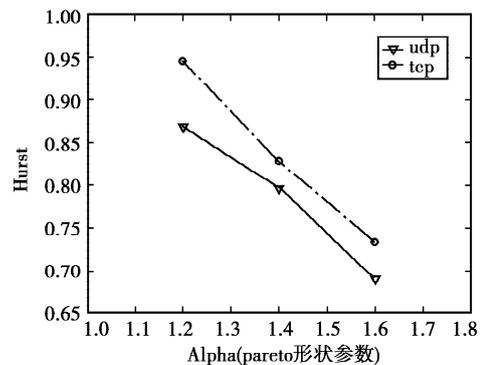


图 1 不同协议对自相似度的影响

由表 2,3 可见,在相同的分布参数的情况下,使用了 tcp 协议的数据源要比使用 udp 协议的数据源具有更大的 Hurst 值,说明了传输层的流量控制,整形以及保证重传机制是造成自相似性的一个重要原因。图 1

比较了不同的协议对业务量自相似度的影响。

由表 4,5 可见,当 ON 和 OFF 的逗留时间以及文件大小都服从重尾分布的时候,ON 和 OFF 的分布参数的改变对最终业务量的自相似度有局部性的影响;而当文件大小的重尾分布程度较为明显的时候,即使 ON 和 OFF 的逗留时间也有较为明显的重尾分布,对最终业务量的自相似度也起不到决定作用。因此,Hurst 参数的值主要还是由文件大小的分布来决定。另外,如果文件大小的分布只是轻度的重尾分布,那么在 ON 过程中文件发送间隔时间的分布的重尾性会加重最终生成的业务量的自相似性;但是,如果文件大小的分布已经是明显的重尾分布,那么间隔时间的分布的重尾性几乎对最终的自相性的程度没有影响。最后,传输层所采用协议的类型也会影响自相似性的存在。例如,如果传输层使用 UDP 协议,那么下行流量的自相似性将被破坏,换句话说,业务量的突发度将只能在短程相关尺度(SRD)下存在;不过如果传输层采用 tcp 协议,例如 Tcp Reno, Tahoe 或者 Vegas,那么业务量的长程相关性(LRD)将被保留下来并且被传送到 MAC 层。因此,传输层协议是否采用流控和保证重传的机制将会影响自相似性的存在或者存在程度。

2 多 ON/OFF 源汇聚的业务量自相似性分析

2.1 仿真分析

本节将通过叠加多个 ON/OFF 数据源来研究业务量的自相似输出特性。仿真参数如表 1:

表 6 仿真参数表

仿真参数	数值
仿真时间/s	300
协议	Tcp
ON/OFF 分布	Pareto(0.01,1.5)
文件大小分布	Pareto(1 024,1.5)
Qos	Best Effort
信道带宽/bps	10 M

2.2 结果分析

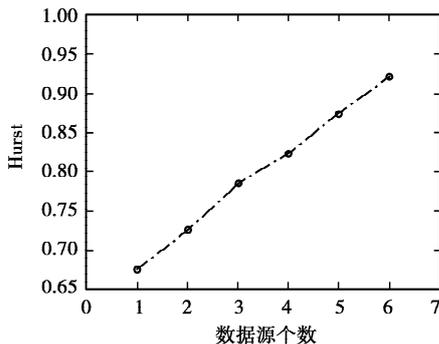


图 2 数据源汇聚程度与自相似度的关系

由图 2 可见,随着数据源汇聚程度的增加,在相同的 ON/OFF 逗留时间和文件大小分布的情况下,输出业务量的自相似程度也逐渐增加。这说明由于数据源叠加所造成的链路可用带宽减小,发送队列队长增加,以及接受队列缓存溢出导致丢包所造成的重传等因素都能造成业务量的自相似。

3 业务量产生器的算法描述以及实现细节

本节根据前几节研究分析的结果,设计并实现了一个 IP 网络的自相似业务量生成器。通过利用 V-T(variance time plot)图测量业务量生成器的输出业务量,发现所提出的业务量生成算法能够较好地反映文件大小,传输协议以及汇聚程度与自相似度之间的关系以及真实业务量的统计特性。

3.1 设计思想描述

流量产生器分为单点式和分布式两种产生模式。单点式产生器可以在一台计算机上通过不同的发送端口模拟多台计算机的数据源叠加效果,并且每一个被模拟的数据源都可以分别设置自己的分布参数和发送速率。在接收端(linux 下实现),接收父进程通过 fork 子进程来模拟多台数据源的并发连接,并将结果收集并显示。分布式产生模式采用服务器/客户机模式。客户机端负责数据源的产生和连接服务器;服务器端负责处理发送的流量数据并通过图形显示。服务器和客户机通过 CORBA ORB 连接。系统结构的示意图如下:

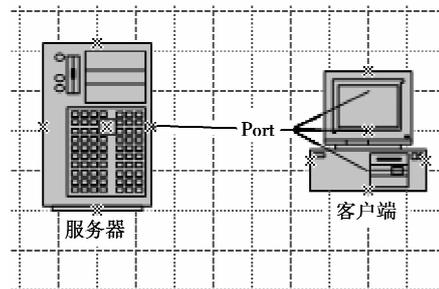


图 3 单点式产生器示意图

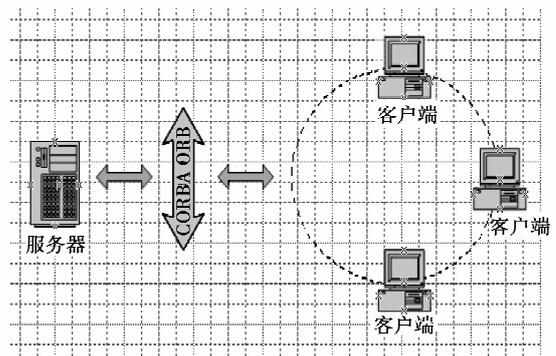


图 4 分布式产生器示意图

下面给出单点式流量产生器的算法流程图如图 5~6 所示。

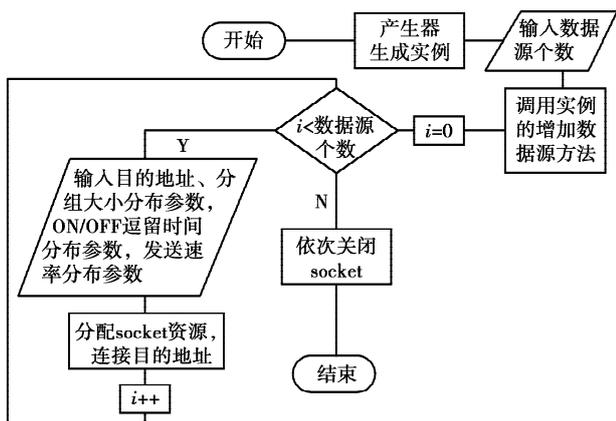


图 5 流量产生器发送端流程

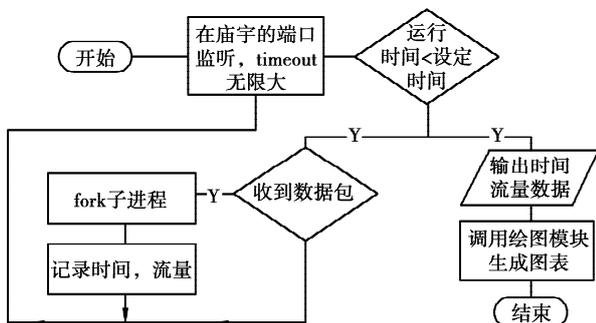


图 6 流量产生器接收端流程

3.2 系统架构

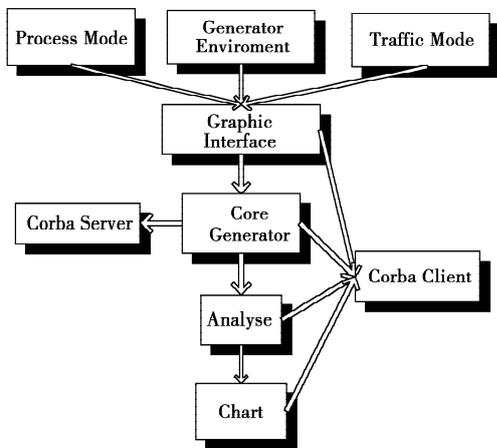


图 7 系统结构图

系统的结构图如图 7 所示, 相应的模块说明如下。

CoreGenerator: 流量产生器核心模块, 包含 easyGenerator 和 complexGenerator 模块。其中 easyGenerator 模块负责产生简单的基于 winsocket 的流量, 而 complexGenerator 继承了 CorbaServer 的功

能, 可以作为一个 Corba 服务以供 CorbaClient 分布式的调用。两者都提供了数据源的产生功能;

GraphicInterface: 提供了流量产生器所需要的界面功能, 包括初始化界面, 主操作界面, 弹出对话框的设计等等;

Analyse: 该模块负责生成数据的分析, 包括最小二乘法拟和, 时间方差法估计 Hurst 参数等等;

ProcessMode: 此模块包含了分组大小、ON/OFF 逗留时间等随机变量的分布生成函数;

GeneratorEnvironment: 此模块是 coreGenerator 的一个 interface, 负责与 corbaServer 和 corbaclient 通信;

TrafficMode: 此模块扩展了流量产生器所能处理的流量模型, 包括 ON/OFF 模型, Fractal ARIMA 模型等等;

Chart: 此模块负责最终生成流量的收集, 图形显示和处理;

ExceptionPackage: 此模块处理其它模块所抛出的异常并作相应处理;

CorbaClient: 此模块用于分布式流量产生器时生成 corba 客户端;

CorbaServer: 此模块用于分布式流量产生器时生成 corba 服务器。

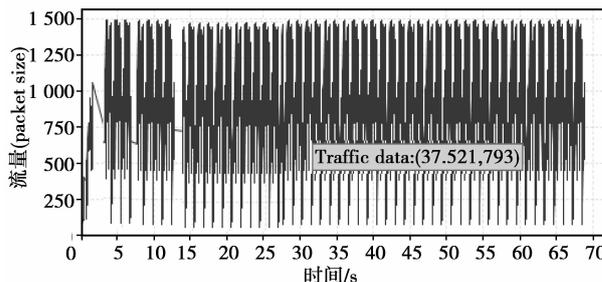


图 8 流量产生器产生的自相似流量图

3.3 结果分析

业务量生成器的生成参数为: 40 个数据源, ON 状态的逗留分布服从 pareto 分布, 形状参数为 1.2, OFF 状态的逗留分布服从 pareto 分布, 形状参数为 1.4, packet size 服从 pareto 分布, 形状参数为 1.2, 网络负载为 0.3, 网络连接由于采用 socket, 具有 tcp 协议的所有功能。图 8 为在接收端收到的流量所生成的图形。将流量数据用 V-T (variance time plot) 图分析后, 得到生成流量的 HURST 参数大概为 0.88 左右, 基本上符合预设的 0.9 的目标。

(下转第 190 页)

- field in the inhomogeneous goaf with arbitrary shape[J]. *Mechanics in Engineering*, 2007, 29(2): 28-30.
- [5] 李宗翔, 王晓冬, 王波. 采空区场流数值模拟程序(G3)实现与应用[J]. *湖南科技大学学报*, 2005, 20(3): 16-20
LI ZONG-XIANG, WANG XIAO-DONG, WANG BO. Realization and application of numerical simulating program (G3) for field flow of goaf[J]. *Journal of Hunan University of Science & Technology*, 2005, 20(3): 16-20.
- [6] 李宗翔, 李海洋, 贾进章. Y形通风采空区开区注氮防火的数值模拟[J]. *煤炭学报*, 2005, 30(5): 593-597.
LI ZONG-XIANG, LI HAI-YANG, JIA JIN-ZHANG. Numerical simulation of prevent spontaneous combustion by nitrogen injection in goaf of Y-type ventilation face [J]. *Journal of China Coal Society*, 2005, 30(5): 593-597.
- [7] 李宗翔. 采空区遗煤自燃过程及其规律的数值模拟研究[J]. *中国安全科学学报*, 2005, 15(6): 15-19.
LI ZONG-XIANG. Numerical simulation of spontaneous combustion process of residual coal in goaf [J]. *China Safety Science Journal*, 2005, 15(6): 15-19.
- [8] 李宗翔, 海国治, 秦书玉. 采空区风流移动规律的数值模拟与可视化显示[J]. *煤炭学报*, 2001, 26(1): 76-80.
LI ZONG-XIANG, HAI GUO-ZHI, QIN SHU-YU. Numerical simulation and visual display of air flow pattern in goaf [J]. *Journal of China Coal Society*, 2001, 26(1): 76-80.
- [9] 李宗翔, 孙广义, 王继波. 回采采空区非均质渗流场风流移动规律的数值模拟[J]. *岩石力学与工程学报*, 2001, 2(增刊): 1578-1581.
LI ZONG-XIANG, SUN GUANG-YI, WANG JI-BO. Numerical simulation of air flow pattern in the inhomogeneous seepage flow field of goaf [J]. *Chinese Journal of Rock Mechanics and Engineering*, 2001, 2(S): 1578-1581.
- [10] 李宗翔, 孟宪臣, 赵国忱. 任意形冒落非均质采空区流场流态数值模拟[J]. *力学与实践*, 2005, 27(6): 26-28.
LI ZONG-XIANG, MENG XIAN-CHEN, ZHAO GUO-CHEN. Numerical simulation of flow pattern of caving flow field in the inhomogeneous goaf with arbitrary shape [J]. *Mechanics in Engineering*, 2001, 27(6): 26-28.
- [11] 邢玉忠, 郭勇义, 吴世跃. 采空区紊流漏风相关系数的研究[J]. *煤炭学报*, 2001, 26(5): 525-529.
XING YU-ZHONG, GUO YONG-YI, WU SHI-YUE. Study on the relative coefficient of air leakage in the gob [J]. *Journal of China Coal Society*, 2001, 26(5): 525-529.

(编辑 张小强)

(上接第 178 页)

参考文献:

- [1] PAXON V, FLOYD S. Wide area traffic: the failure of poisson modeling [J]. *IEEE/ACM Transactions on Networking*, 1995, 3(1): 226-244.
- [2] 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.
- [3] BERAN J, SHERMAN R, TAQQ M S, et al. Long-range dependence in variable-bit-rate video traffic [J]. *IEEE Transaction on Communications*, 1995, 43(2/3/4): 1566-1579.
- [4] CROVELLA M E, BESTAVROS A. Self-similarity in world wide web traffic-evidence and possible causes [J]. *IEEE/ACM Transactions on Networking*, 1997, 5(6): 835-846.
- [5] ABRY P, PARANIUK R, FLANDRIN P, et al. Multi-scale nature of network traffic [J]. *IEEE Signal Processing Magazine*, 2002, 19(3): 28-46.
- [6] OLIVEIRA C, JAIME B K, SUDA T. Long-range dependence in IEEE 802.11b wireless LAN traffic: an empirical study* [C]. *Proceedings of the 18th Annual IEEE Computer Communications Workshop (CCW 2003)*, California, USA, 2003.
- [7] TAKETSUGU J, HARA S. Self-similarity in cellular network traffic: another cause of long-range dependence [C]. *Wireless Personal Multimedia Communication 2001, WPMC'01*, Aalborg, Denmark, September 9 to 12, 2001.
- [8] MANDELBROT B, NESS J. Fractional brownian motion, ractional noise and applications [J]. *SIAM Review*, 1968, 10(4): 422-437.
- [9] HOSKING J. Fractional differencing [J]. *Biometrika*, 1981, 68(1): 165-176.
- [10] ALEXANDER R, BROWNLEE N, ZIEDINS I. Modeling self-similar network traffic [R]. *Springfield: University of Auckland*, 1995.

(编辑 吕建斌)