

# 自相似流量对网络性能的影响

乐红兵<sup>1</sup>, 马书南<sup>2</sup>, 帅训波<sup>3</sup>

(1. 江南大学 网络中心, 江苏 无锡 214121; 2. 江南大学 信息工程学院, 江苏无锡214122;  
3. 中石油勘探开发研究院廊坊分院 地球物理与信息研究所, 河北廊坊 065007)

**摘要:** 阐述了自相似流量对网络性能诸因素的影响, 并通过实验研究了丢包率与自相似性的关系。实验采用 ON/OFF 模型, 用 NS 平台进行模拟。

**关键词:** 自相似性 网络性能 丢包率 延迟

Leland 等<sup>[1]</sup>通过对局域网的测试, 以及 Paxson 等<sup>[2]</sup>对广域网的测试, 均发现网络流量具有自相似性(self-similarity), 并且流量的特性参数在很长的时间尺度上是相关联的, 即具有长相关性(long-range-dependence)。

自相似性的发现, 引起了网络性能的改变, 使传统意义上的网络特性分析变得复杂。当网络性能用吞吐量、丢包率和包传输率表示时, 随着重尾程度的增加, 网络性能逐渐降低。另外, 增加网络资源如链路带宽或缓冲区容量时, 能明显提高网络性能。当传输较大文件时, 对大缓冲区容量来说, 吞吐量的改进和队列延迟的增加是同时发生的。如果保持缓冲区的利用率不变, 吞吐量增大是以队列延迟的提高为代价的。

## 1 自相似性

自相似性指局部以某种方式与整体相似。从统计意义上讲, 局部适当放大后, 与整体具有相同的统计分布。下面给出常用时间序列的二阶自相似性的数学定义:

令  $X = \{X_t, t \in N\}$  是广义平稳随机过程, 均值为  $\mu$ , 方差为  $\sigma^2$ , 且自相关函数为  $r(k), k \geq 0$ 。假设  $X$  的自相关函数形式表示为:

当  $k \rightarrow \infty$  时,  $r(k) \sim k^{-\beta} L(k)$

其中,  $0 < \beta < 1$  且  $L$  缓慢变化至无穷, 即: 对所有的  $x > 0$  来说,  $\lim_{t \rightarrow \infty} L(tx)/L(t) = 1$ 。对每个  $m = 1, 2, 3 \dots$ , 令  $X^{(m)} = (X_k^{(m)} : k = 1, 2, 3 \dots)$  表示新的协方差稳态时间序列。对每一个  $m = 1, 2, 3 \dots$ ,  $X^{(m)}$  由  $X_k^{(m)} = 1/m(X_{km-m+1} + \dots + X_{km}), k \geq 1$  得出。过程  $X$  称为精确二阶自相似序列<sup>[1]</sup>。

自相似序列具有以下特性:

(1) 过程  $X^{(m)}$  随着  $m \rightarrow \infty$ , 自相关结构不变, 或者说非退化的, 即对于不同的  $m$  值,  $X^{(m)}$  直观上相似, 即它们有基本一致的自相关函数。

(2) 长相关性。其自相关函数随时滞的增加呈双曲线衰减。自相似是长相关的简单模型。因此,  $\sum r(k)$  是不可和的, 即  $\sum_{k=-\infty}^{\infty} r(k) = \infty$

(3) 慢衰减方差。自相似过程方差满足:  $\text{var}(X^{(m)}) \sim cm^{-\beta}, m \rightarrow \infty$ 。其中,  $0 < \beta < 1, c$  是与  $m$  无关的正常数。

## 2 自相似流量对网络性能的影响

当缓冲区的容量到达一定程度后, 缓冲区容量对丢失率的影响越来越小<sup>[7]</sup>。这表明不能用增加缓冲区的方法来减少分组丢失率。出现这种情况的原因是自相似流量具有长相关性, 使得输入流的突发性更大。许多研究表明, 长相关性能引起网络性能下降, 这是因为它比常规队列分析需要更大的缓冲区, 并增加了队列延迟和包丢失率。当吞吐量随着自相似程度的提高而逐渐衰减时, 队列延迟的增加更为明显。当流量的自相似程度很高时, 队列延迟几乎成比例增加, 这种现象对网络服务质量来说有负面影响。网络用户希望有服务质量的保障, 如提供较大的带宽、较低的丢包率等。系统根据服务所需的系统资源进行分配, 并满足各种约束条件, 如延迟时间约束、缓冲区的空间约束, 以及保证数据传输吞吐量的带宽约束。为了更好地利用网络资源, 需要根据网络业务流的特点, 确定所需的各种资源。

在可靠传输中, 当自相似程度增加时, 网络的吞吐量、丢包率和重传率也平缓增长, 而平均队列长度和响应时间快速增加, 表示需要更多的缓冲区和带宽。如果增加网络资源, 如网络带宽, 将极大地提高网络性能。若采用增大缓冲区的办法, 虽然能降低丢包率, 但也将导致较长的队列等待时间, 增加延迟时间, 不符合时间约束较高的应用要求。因此, 在特定的网络带宽下, 需要平衡队列等待的延迟与数据包丢失率之间的关系。

## 3 自相似性与丢包率的关系

网络性能通常用队列延迟、丢包率等因素表示, 下面选取丢包率作为性能的评价因素, 研究自相似性与丢包率之间的关系。

### 3.1 丢包率和重传机制

当自相似的程度发生变化时, 需对网络性能进行评估。在应用层, 流量的尺度不变, 突发性仅受重尾的影响, 它由 Pareto 分布的形参  $\alpha$  获得, 并作为控制变量<sup>[3]</sup>。

首先基于传输层采用 UDP 不可靠传输协议, 在不同重尾分布  $\alpha$  值下, 网络流量中的分组丢弃率随  $\alpha$  值的增大而降低。Pareto 分布的重尾变化结果如图 1 和图 2 所示, 对两个缓冲区容量来说, 它们分别表示丢包率与  $\alpha$  的函数关系, 以及包传输率与  $\alpha$  的函数关系。随着缓冲区容量的增加, 丢包率呈下降趋势。当瓶颈缓冲区非常小时,  $\alpha$  值的改变几乎不引起网络性能的变化, 此时性能曲线是平滑的, 其原因是突发性在  $\alpha$  值接近 2 时引发了高丢包率。当缓冲区的容量较大时, 如图 1 中缓冲区容量为 128KB 时, 高突发性呈现在  $\alpha=1$  至  $\alpha=1.1$  之间的某值处。当  $\alpha$  接近 1 时, 可以发现在丢包率与重传机制之间逐渐发生了衰减。可以发现, 丢包率和重传机制与  $\alpha$  呈现线性关系, 而且该特性是通过拥塞控制的行为获得的。由图 1 可以得出, 当瓶颈带宽发生变化时, 对于固定长度的缓冲区来说, 自相似性与丢包率是相关的。

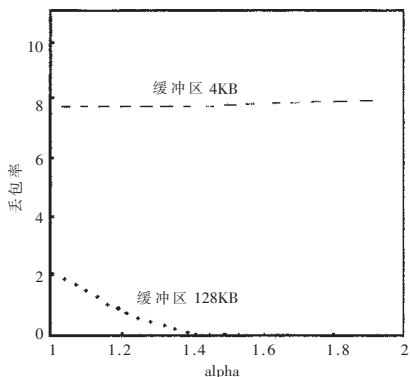


图 1 丢包率与  $\alpha$  的关系

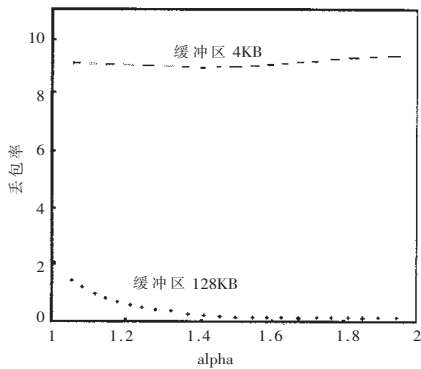


图 2 包传输率与  $\alpha$  的关系

### 3.2 实验过程与结果分析

下面对采用 TCP 可靠通信协议、基于 ON/OFF 的叠加模型进行自相似性与丢包率间的关系分析。当丢失率较高时, 通常会导致超时, 超时期间没有数据包发送, OFF 持续期间重尾分布加重, 造成 TCP 流的突发现象, 自相似程度较高; 当丢失率很低时, 在 ON 期间, 源端可以发送最大的分组数, 此时没有超时发生, 自相似程度较低。本文通过实验研究了当多个 TCP 连接通过单个瓶颈链路时, 自相似性与丢包率的关系, 以及丢包率对网络流量自相似特性程度的影响; 通过设计一个丢失模

块, 研究丢包率对自相似程度的影响。

实验采用网络拓扑结构, 在 NS 平台上进行模拟。实验中共有 8 个源端点, 8 个目的端点, 每个源端点  $k$  和目的端点  $(8+k)$  组成一个连接。所有的源端点和所有的目的端点均与链路上的丢失模块相连。所有源端点与丢失模块之间的链路带宽设置为 3MB, 所有目的端点与丢失模块之间的链路带宽设置为 6MB。丢失模块所在链路上的带宽设置为 100MB。源端点与丢失模块之间的链路延迟初始值设置为 0.3ms, 目的端点与丢失模块之间的链路延迟初始值设置为 0.2ms, 丢失模块所在链路上的延迟初始值设置为 0.2ms。包的大小设为 500KB, 形参数设为 1.5。当 TCP 连接数为  $N=8$ , 丢失率分别为  $P=0.1$ 、0.2 时, 对数尺度图如图 3 所示。

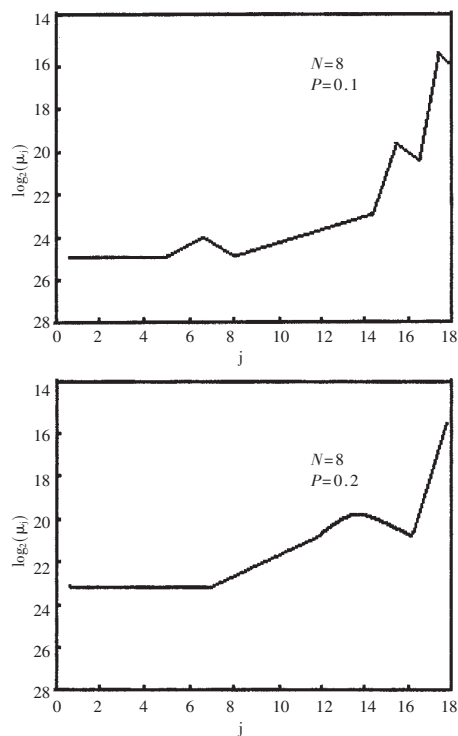


图 3 丢包率不同的对数尺度图

通过对链路上的聚合流量进行自相似分析, 以验证丢失率对流量自相似的影响。流量的自相似性随丢失率的增加而增大, 当多个 TCP 争用带宽时, 会产生较高的突发性, 自相似程度增加。另外, 随着丢包率的提高, 导致自相似性的程度也逐渐增加。

自相似流量对网络性能有不利的影响。当自相似程度提高时, 为了使吞吐量或丢包率处于平衡状态, 需要一个极大的缓冲区容量, 但缓冲区的增加会导致更严重的队列延迟。在自相似程度较高时, 增加网络带宽能够明显改善丢包率。流量的自相似性随着丢包率的提高而增加, 反之, 当自相似程度增加时, 网络的丢包率也相应提高。

#### 参考文献

[1] LELAND W, TAQQU M, WILLINGER W, et al. On the self-similar nature of ethernet traffic(extended version)[J],

- IEEE/ACM Trans.Networking, 1994 ,(2):1-15.
- [2] PAXSON V, FLOYD S. Wide area traffic: The failure of poisson modeling [J]. IEEE/ACM Transactions on Networking, June 1995,3(3):226-244.
- [3] PARK K. Self-similar network traffic and its control. In: Proc.25<sup>th</sup> Allerton Conference on Communication,Control and Computing, 1997,(10).
- [4] NORROS I. A storage model with self-similar input[J], Queueing Systems,1994:387-396.
- [5] ADAS A, MUKHERJEE A. On resource management and QoS guarantees for long range dependent traffic[J]. IEEE Infocom 1995:779-787.
- [6] GROSSGLAUSER M. On the relevance of long-range dependence in network traffic[J]. ACM Sigcomm,1996: 15-24.
- [7] 吴泽民,郑少仁. 自相似流量及其对网络性能的影响. 解放军理工大学学报,2000,5(1):29-34.
- [8] PARK K. On the effect and control of self-similar network traffic: A simulation perspective. Simulation Conference, Dec 1997. (收稿日期:2006-10-06)