

# 一种改进的高速 TCP 拥塞控制算法<sup>\*</sup>)

葛一楠 周驰岷 易勇

(成都大学 成都 610106)

**摘要** Sally Floyd 提出的高速 TCP 拥塞控制算法能提高 TCP 在高速网络中传送数据时的性能,但高速 TCP 与普通 TCP(以目前使用最广泛的 Reno TCP 为例)共存时存在公平性的问题。本文系统研究了高速 TCP 拥塞控制算法,分析了高速 TCP 与 Reno TCP 共存时的公平性与效率性,提出了一种改进的高速 TCP 拥塞控制算法。最后,通过仿真证明该算法有效地改进了高速 TCP 与 Reno TCP 共存时的性能。

**关键词** TCP 拥塞控制,高速 TCP,公平性,效率性

## An Enhanced High Speed TCP Congestion Control Algorithm

GE Yi-Nan ZHOU Chi-Min YI Yong

(Chengdu University, Chengdu 610106)

**Abstract** The high speed TCP congestion control algorithm brought forward by Sally Floyd. It can improve capability of TCP in high speed network, but there are some fairness problem between the high-speed TCP and common TCP (e. g. the most popular Reno TCP) when they are used in a single network. This issue studies the theory of high speed TCP congestion control algorithm and analyses the fairness and efficiency, and brings forward an enhanced high speed TCP congestion control algorithm.

**Keywords** TCP congestion control, High speed TCP, Fairness, Efficiency

### 1 引言

TCP 采用的拥塞控制算法已经被广泛地应用到互联网上。在实际应用中,当其应用于低带宽环境下的数据传输时,能够满足对吞吐量的需求。然而,当它应用于高带宽环境下的数据传输时(Gigabit 以上),由于传统 TCP 协议自身的问题,吞吐量往往不能满足实际需求。针对这种局限性,Sally Floyd 等人提出了针对高带宽、大时延环境下拥塞控制问题的高速 TCP(High Speed TCP)算法<sup>[1]</sup>。

### 2 高速 TCP 拥塞控制算法

高速 TCP 的基本思想是在拥塞避免阶段采用不同的窗口大小控制机制。高速 TCP 在拥塞窗口的增加上更加迅速,而在拥塞窗口的减少上更为缓慢,并最终保持一个比较大的拥塞窗口,以充分利用网络的带宽。表 1 为算法描述中使用的参数。

在慢启动阶段,高速 TCP 采用与普通 TCP(以目前最为流行的 Reno TCP 为例)相同的慢启动算法<sup>[3]</sup>;在拥塞避免阶段,采用基于加法增加/乘法减小的模型(AIMD)去改变拥塞窗口大小。但是,有别于 Reno TCP,它的递增递减参数是与拥塞窗口的大小有关的。当前拥塞窗口大小  $W$  小于  $W_{low}$  时,高速 TCP 使用与 Reno TCP 一样的递增递减参数;当前拥塞窗口大小  $W$  大于  $W_{low}$  时,高速 TCP 使用比 Reno TCP 较快的递增拥塞窗口、较慢的递减拥塞窗口,因而它维持了一个比 Reno TCP 更大的拥塞窗口。

当每个 ACK 应答到来的时候,高速 TCP 增加它的拥塞

窗口值

$$W \leftarrow W + \frac{\alpha(w)}{W} \quad (1)$$

当收到拥塞事件时,高速 TCP 减少它的拥塞窗口值

$$W \leftarrow (1 - \beta(w))W \quad (2)$$

表 1 高速 TCP 拥塞控制算法中的参数

参数名	含义
$W$	当前拥塞窗口大小
$W_{low}$	进入高速 TCP 的阈值
$W_{high}$	高带宽时的平均拥塞窗口
$\alpha(w)$	递增参数
$\beta(w)$	递减参数
$P_{low}$	对应于 $W_{low}$ 的丢包率
$P_{high}$	对应于 $W_{low}$ 的丢包率
$\beta_{high}$	高带宽时的递减系数

$$\text{其中: } \alpha(w) = \frac{2w^2 p(w) \beta(w)}{2 - \beta(w)} \quad (3)$$

$$\beta(w) = \frac{\log W - \log W_{low}}{\log W_{high} - \log W_{low}} (\beta_{high} - 0.5) + 0.5 \quad (4)$$

$$p(w) = \exp \left[ \frac{\log W - \log W_{low}}{\log W_{high} - \log W_{low}} (\log p_{high} - \log p_{low}) + \log p_{low} \right] \quad (5)$$

### 3 改进的高速 TCP 拥塞控制算法

虽然高速 TCP 算法能提供比普通 TCP 更大的吞吐量,但其性能还有待改善提高。典型情况是:当 Reno TCP 和高

葛一楠 副教授,硕士,主要研究方向为自动控制;周驰岷 硕士,主要研究方向为计算机网络、嵌入式系统;易勇 教授,博士,主要研究方向为计算机网络与通信。

速 TCP 共享链路时,因为拥塞时 Reno TCP 的拥塞窗口增长速率慢于高速 TCP 的增长速率,所以经过若干次拥塞事件后,高速 TCP 会不断地挤占 Reno TCP 的可用带宽,这就是常说的公平性问题。

TCP 拥塞控制的另一个问题是效率 (efficiency) 问题。网络资源的使用效率是源端要求的总资源与网络资源的接近程度决定的。如果源端要求的总资源  $X(t)$  接近或等于网络所能提供的资源  $X_{eff}$ , 那么这种算法就是高效率的。超载  $X(t) > X_{eff}$ , 或负载不足  $X(t) < X_{eff}$ , 是效率不高的表现。效率只与总资源的利用率有关, 而与各个源端之间的资源利用无关。

### 3.1 基本思想

由于只有当带宽利用率饱和时高速 TCP 才会与 Reno TCP 发生带宽争夺的公平性问题, 因此分两种情况分别处理:

一是当信道带宽没有被充分利用时, 高速 TCP 可以按上述的递增递减参数来调整拥塞窗口。由于有足够的带宽可用, 高速 TCP 不会夺取本该由 Reno TCP 占有的带宽。

二是当信道带宽已被充分利用时, 因为高速 TCP 的递增参数比 Reno TCP 的大而递减参数比 Reno TCP 的小, 所以它会占用本应由 Reno TCP 占用的带宽。此时调整高速 TCP 算法的递增递减参数, 直至由于数据包丢失而使通道利用率低于 100% 时, 再转为运行高速 TCP 拥塞避免算法。

### 3.2 状态转换图

遵循上述思想, 修改现有高速 TCP 拥塞控制算法, 其状态转换图如图 1 所示。

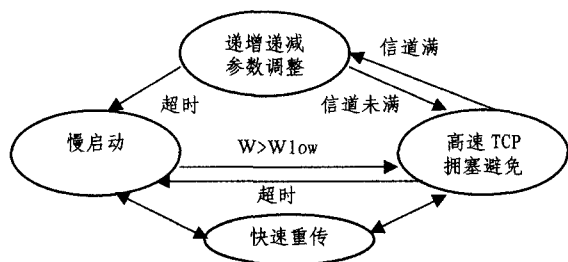


图 1 改进的高速 TCP 拥塞控制算法状态转换图

### 3.3 参数调整

当检测到信道带宽满后, 窗口大小减半进入拥塞避免阶段。此时, 每确认收到一个新的 ACK, 拥塞窗口大小增加  $\alpha(w)/W$ 。源端如果收到三个重复的 ACK, 则将窗口减少到  $(1 - \beta(w))W$ ; 如果发生超时, 则窗口降为 1。将递增参数  $\alpha(w)$  和递减参数  $\beta(w)$  作为变量加入到 TCP 的发送速率控制公式<sup>[7]</sup>:

$$B = \frac{1}{RTT \sqrt{\frac{2bp}{3}} + T_0 \min\left(1, 3\sqrt{\frac{3bp}{8}}\right) p(1 + 32p^2)} \quad (6)$$

其中  $RTT$  为  $RTT$  时间,  $T_0$  为超时时间,  $P$  即为  $P_{low}$ ,  $b$  为每个 ACK 确认的数据包个数。得到

$$B(\alpha, \beta) = \frac{1}{RTT \sqrt{\frac{2bp(1-\beta)}{\alpha(1+\beta)}} + T_0 \min\left(1, 3\sqrt{\frac{bp(1-\beta^2)}{2\alpha}}\right) p(1 + 32p^2)} \quad (7)$$

如果所有的拥塞都不发生定时器超时, 则分母中没有  $T_0$  部分, 表明  $B(\alpha, \beta)$  此时不依赖于  $T_0$ 。当拥塞有超时发生时,

虽然有  $T_0$  部分, 但  $B(\alpha, \beta)$  也不直接依赖于  $T_0$ 。比较这两部分可知: 当  $P$  很小时,  $RTT$  部分占主导地位, 此时  $B(\alpha, \beta)$  主要由其决定; 而当  $P$  增加时,  $T_0$  部分逐渐变大, 最后增加到 1。

从式(7)可知, 通过控制  $(\alpha, \beta)$  的取值可得到期望的发送速率  $B(\alpha, \beta)$ , 使 TCP 的发送速率  $B(\alpha, \beta) = B(\alpha = 1, \beta = 0.5)$ 。通过选取合适的  $\alpha$  值, 使在较大的丢包率范围内  $B(\alpha, \beta)$  和 TCP 发送速率的差值最小。定义误差方程:

$$E_p(\alpha) = \int_0^1 w(p) \left| 1 - \frac{B(\alpha, \beta)}{B(\alpha = 1, \beta = 0.5)} \right| dp \quad (8)$$

当  $\beta = 0.5, T_0 = 3RTT$  时,  $\alpha$  在  $(0.5, 0.8)$  的范围内取值, 误差方程得到的差值较小。本改进算法取中值  $\alpha = 0.65$  作为参数调整阶段的取值。

## 4 公平性与效率性分析

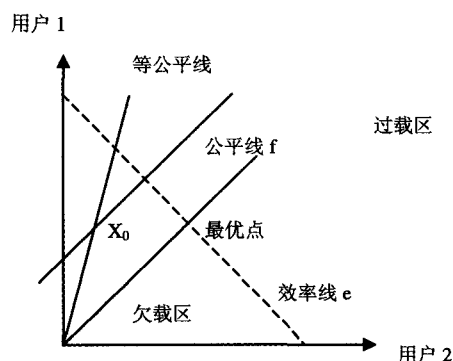


图 2 两个用户的向量表示

对于高速 TCP 和 Reno TCP 都采用的加法增加/乘法减小的模型 (AIMD), 我们常常用  $n$  维向量空间来描述系统的状态。如图 2 所示, 以两个用户为例: 两个用户的资源分配  $\{X_1(t), X_2(t)\}$  可以用 2 维空间的点  $(X_1, X_2)$  来表示。纵坐标表示用户 1 的资源分配, 横坐标表示用户 2 的资源分配。在所有的资源分配策略中,  $X_1 + X_2 = X_{eff}$  是有效率的分配策略, 这与图中的效率直线  $e$  相对应; 在所有的分配策略中,  $X_1 = X_2$  是公平的分配策略, 这与图中的公平直线  $f$  相对应。两条直线的交点  $(X_1/2, X_2/2)$  就是最优点, 分配策略的目的就是尽量将系统状态转换到此点上来。

所有在效率线以下的点表示欠载, 分配策略会要求用户增加负载量。如点  $X_0$ , 加法增加策略使两个用户的负载量都沿着  $45^\circ$  直线以特定递增系数增加。

所有在效率线以上的点都是过载点, 分配策略会使用户的负载量减少。乘法减少沿着连接原点和该点的直线减少。任意一点的公平度为

$$F(x) = \frac{(x_1 + x_2)^2}{2(x_1^2 + x_2^2)} \quad (9)$$

在乘法分配策略中, 系数  $b$  不会影响公平度, 即  $(X_1, X_2)$  和  $(NX_1, NX_2)$  具有相同的公平度。那么所有位于连接原点直线上的点公平度相同, 因此把通过原点的直线叫做等公平线。

图 3 表示了系统采用加法增加/乘法减少算法 (AIMD) 时从初始位置  $X_0$  变化的过程。由于初始点位于效率线以下, 因此系统要求用户增加负载量, 采用加法算法沿着  $45^\circ$  直线增加, 使其负载状态点运动到  $X_1$ 。此时该点位于效率线之上, 因此采用乘法减少算法减少负载量, 使得状态沿着连接点  $X_1$  和原点的直线向原点运动。此时状态点运动到  $X_2$ , 又位

于效率线之下,那么系统重复上述算法过程。每一次的重复算法,使得公平度提高,直到最后系统收敛到理想状态点,围绕目标点振动。

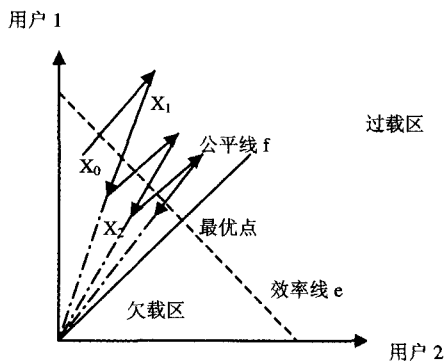


图 3 AIMD算法的收敛

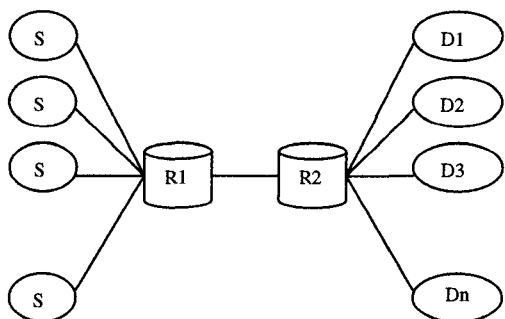


图 4 仿真网络拓扑图

基于以上分析,我们知道 AIMD 算法总是能收敛到公平线上并围绕着效率线振动。因此可以说,这种改进的高速 TCP 拥塞控制算法能够使各数据流最终达到公平性和效率性要求。

## 5 算法仿真

仿真网络拓扑如图 4 所示,图中链路带宽均为 1Gbps,并且假设无论远端发射多少数据,接收端都来得及接收。

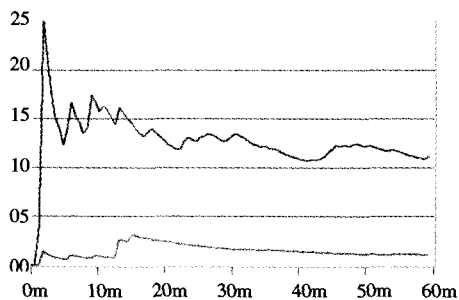


图 5 高速 TCP 与 Reno TCP 共存

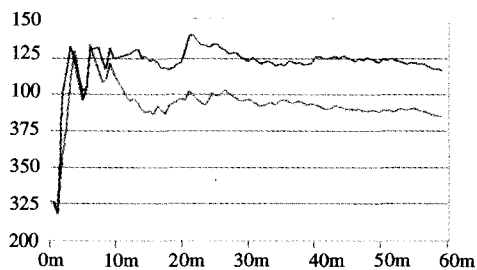


图 6 改进的高速 TCP 与 Reno TCP 共存

图 5 是一条高速 TCP 连接与一条 Reno TCP 连接共存于同一信道进行 60 分钟仿真的情况。图 6 是一条改进高速 TCP 连接与一条 Reno TCP 连接共存于同一信道进行 60 分钟仿真的情况。

**结论** 图 5 中蓝色曲线表示高速 TCP 连接的数据流,红色曲线表示 Reno TCP 连接的数据流。仿真进行了 60 分钟。由仿真数据可知,高速 TCP 数据流占用了信道中的大部分带宽,而 Reno TCP 数据流所占带宽很小。

图 6 中蓝色曲线表示改进高速 TCP 连接的数据流,红色曲线表示 Reno TCP 连接的数据流。与图 5 相比,它们对信道带宽的占用情况得到极大改善。因此,通过以上的数学分析和仿真实验,本文中提出的改进的高速 TCP 拥塞控制算法改进了高速 TCP 数据流与 Reno TCP 数据流共存时的性能,满足了其共存时的公平性要求。

## 参考文献

- 1 Floyd S. RFC 3649, High Speed TCP for Large Congestion Windows. <http://www.ietf.org/rfc/rfc3649.txt?number=3649>
- 2 Floyd S. RFC 3742, Limited Slow-Start for TCP with Large Congestion Windows. <ftp://ftp.rfc-editor.org/in-notes/rfc3742.txt>
- 3 Floyd S, Ratnasamy S, Shenker S. Modifying TCP's Congestion Control for High Speeds. <http://www.icir.org/floyd/papers/hstcp.ps>. Nov01
- 4 Jain M, Dovrolis C. End to End Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput. [http://portal.acm.org/ft\\_gateway.cfm?id=633054&type=pdf&coll=GUIDE&dl=GUIDE,ACM&CFID=11111111&CFTOKEN=22222222](http://portal.acm.org/ft_gateway.cfm?id=633054&type=pdf&coll=GUIDE&dl=GUIDE,ACM&CFID=11111111&CFTOKEN=22222222)
- 5 Jain M, Dovrolis C. Pathload: A measurement tool for end-to-end available bandwidth. <http://citeseer.ist.psu.edu/cache/papers/cs/26060/http://zSzzSzwww.cis.udel.edu/zSzzSz~dovrolis/zSzzPaper-szSpam02.pdf/jain02pathload.pdf>
- 6 Stevens W R. TCP/IP Illustrate
- 7 Padhye J, Firoiu V, Towsley Y D, et al. Modeling TCP Throughput: A Simple Model and its Empirical Validation. <http://www.ecse.rpi.edu/Homeworkpages/shivkuma/teaching/sp2001/readings/Padhye98.pdf>
- 8 <http://www.isi.edu/nsnam/ns/>
- 9 de Souza E, Agarwal D. A High Speed TCP Study Characteristics and Deployment Issues
- 10 Tan Kun, Song Jingmin, Zhang Qian, et al. A Scalable and TCP-Friendly Congestion Control for High-speed Networks
- 11 Lee Junsoo, Bohacek S, Hespanha J P, et al. A Study of TCP Fairness in High-Speed Networks
- 12 Marquezl R, Altman E, Sole-Alvarezl S. Modeling TCP and High Speed TCP A Nonlinear Extension to AIMD Mechanisms
- 13 Tokuda K, Hasegawa G, Murata M. Performance Analysis of High Speed TCP and its Improvement for High Throughput and Fairness against TCP Reno Connections
- 14 Weigle M C, Sharma P, Freeman J R IV. Performance of Competing High-Speed TCP Flows
- 15 Leith D, Shorten R. TCP for high-speed and long-distance networks
- 16 Gu Yunhong, Hong Xinwei, Mazzucco M, et al. Rate-based Congestion Control over High BandwidthDelay Links