Analysis of Two Competing TCP/IP Connections

E. Altman
INRIA, 2004 route des Lucioles, 06902 Sophia Antipolis, France
CESIMO, Universidad de Los Andes, Facultad de Ingeneria, Merida, Venezuela

T. Jiménez
CESIMO, Universidad de Los Andes, Facultad de Ingeneria, Merida, Venezuela

R. Núñez-Queija
CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Eindhoven University of Technology, Dept. Math. & Comp. Sc., The Netherlands

ABSTRACT

Many mathematical models exist for describing the behavior of TCP/IP under an exogenous loss process that does not depend on the window size. The goal of this paper is to present a mathematical analysis of two asymmetric competing TCP connections where loss probabilities are directly related to their instantaneous window size, and occur when the sum of throughputs attains a given level. We obtain bounds for the stationary throughput of each connection, as well as an exact expression for symmetric connections. This allows us to further study the fairness as a function of the different round trip times. We avoid the simplifying artificial synchronization assumption that has frequently been used in the past to study similar problems, according to which whenever one connection looses a packet, the other one looses a packet as well.

2000 Mathematics Subject Classification: 60K25, 68M20, 90B18, 90B22.
Keywords & Phrases: congestion control in data networks, transmission control protocol (TCP).

Note: R. Núñez-Queija is a member of PNA Theme 2.1.

1. Introduction

The mathematical analysis of the performance of TCP has been a major research area in networking. Different types of approaches have been suggested and validated. On one hand, there have been models focusing on a single connection that is subject to some exogenous loss process (which does not depend on that connection), see e.g. [1, 2, 3, 13]. This approach is appealing when there is a large amount of traffic, so that we can neglect the effect of the single connection on events that cause losses. An alternative approach is necessary when the window increase of a connection is itself a central cause for losses. This occurs typically when a small number of connections compete over bandwidth, say, at a bottleneck link. A main mathematical approach for studying this situation has been to study several connections sharing a bottleneck, and then make the simplifying assumption that all connections reduce their windows simultaneously upon congestion [9, 12]. With this approach, it has been shown

[12] that the throughput achieved by a TCP connection is inversely proportional to RTT^{α} with $1 < \alpha < 2$, where RTT is the two-way propagation delay of the connection. However, it turns out that in practice this assumption does not hold, except for drop tail buffers and connections with similar Round Trip Times (RTTs) [15].

In two recent papers, [4, 7], a Markovian model with continuous state space has been proposed to study the performance of TCP when flows are not synchronized. Instead of synchronization, it is assumed that the connections reduce their windows upon congestion with a probability that equals to their share of the bandwidth upon the congestion. Exact formulas for TCP throughput in a similar setting are obtained in [6], assuming that when congestion occurs, the probability that a given connection suffers from it does not depend on the window sizes. However, simulation studies [4] indicate that this probability does depend on the connection's bandwidth share. The performance measures of TCP were obtained in [4, 7], by discretizing this Markov chain and computing numerically the steady state distribution of the discretized model. It was then concluded in [4] (and validated through simulations) that in that scenario, TCP is more fair than predicted by models that assume synchronization of losses: the throughput share of each connection was observed to be approximately proportional to $RTT^{-0.85}$. This as opposed to $RTT^{-\alpha}$ with $\alpha > 1$ in the synchronized case.

The first goal of the current paper is to go one step further than the numerical results in [4] and provide analytical expressions and provable bounds on the throughput of competing TCP connections. Our second objective is to substantiate the qualitative conclusions that were obtained in [4] through numerical studies and simulations on the bandwidth sharing between the connections, and prove their validity beyond the parameter values chosen there. Indeed, we show that TCP is more fair than predicted in the synchronized setting, a connection's share being inversely proportional to its RTT when the latter tends to infinity. Note that this matches the prediction of TCP's throughput by models for a single connection [1, 13]. For moderate RTTs we provide an approximation that matches the numerical results for the discretized Markov chain in [4] (and explains the observed approximate proportionality between the throughput and $RTT^{-0.85}$).

The remainder of the paper is organized as follows. In Section 2 we describe the mathematical model for the throughput of two concurring TCP connections. A recursive expression for the moments of the throughput at loss instants is derived in Section 3. When both connections have the same RTT, this allows to compute all moments of the throughput, but in the unsymmetric case the mean (from which all higher moments can be determined) remains unknown. Section 4 shows how the recursion leads to an implicit equation for the *distribution* of the throughput at loss instants. Determining bounds for the mean throughput at loss instants (when the RTTs are different) is the subject of Section 5. This leads to bounds for the time average throughput when the RTTs are unequal in Section 6 (in the symmetric case the time average throughput of both connections can be determined exactly). In Section 7 we show that the obtained bounds match the exact order of magnitude of the throughput when the ratio of the RTTs tends to infinity. A surprisingly accurate approximation is derived in Section 8. Finally, Section 9 summarizes the paper.

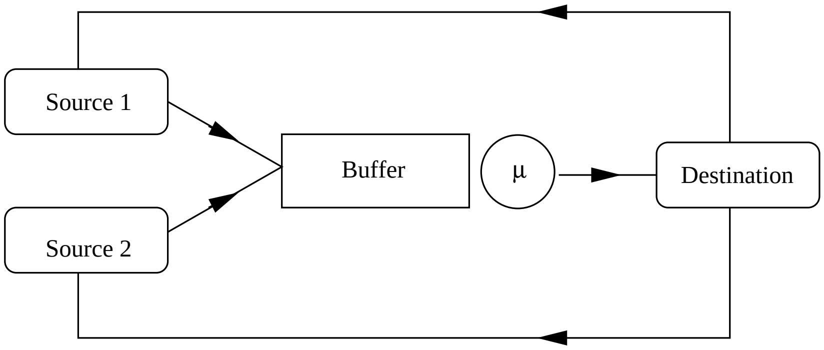

Figure 1: Two TCP connections on a path of bandwidth μ.

2. The mathematical model

Suppose that two TCP sources (1 and 2) share a path of bandwidth μ as depicted in Figure 1. The two sources are assumed to have the same packet length. Denote the RTTs of these connections by T_1 and T_2. Denote also by $W_1(t)$ and $W_2(t)$ the window sizes of the two connections at time t. The rate of a connection at an instant t can be written as,

$$X_k(t) = W_k(t)/T_k \qquad \text{with } k = 1, 2.$$

We assume that the two sources run a TCP version able to recover from losses without resorting to timeout and slow start. A SACK version or a New-Reno version can be used [10]. Upon detection of congestion, the TCP source divides its window by two, recovers from losses, and then resumes increasing its window. We assume that the transfers are permanent, i.e., the sources always have packets to send. We further assume that the queueing delay is small with respect to the propagation delay so that the RTT is approximately constant. This is reasonable with active buffers where the queue length is maintained at small values [11]. We consider the case when the window of TCP increases by one packet every RTT (i.e., delay ACK mechanism [14] is disabled; the analysis of the case of delay ACKs can be handled in exactly the same way our analysis below). The window and the rate of each source grow then linearly as a function of time as shown in [12] where a fluid model for the window evolution is used. We write for $k = 1, 2$

$$\frac{dW_k(t)}{dt} = \frac{dW_k(t)}{dack_k} \times \frac{dack_k}{dt} = \frac{1}{W_k(t)} \times \frac{W_k(t)}{T_k} = \frac{1}{T_k}.$$

This linear growth continues until a congestion occurs. Due to our assumption that queueing time is small, it is possible to consider that congestion occurs when the sum of the rates of the two connections reach a threshold bandwidth μ (which could correspond to the capacity of a bottleneck). We assume that a congestion event causes losses to one connection and only this connection divides its window by two. The window growth of the other connection is not affected. Given the probabilistic drop of packets at the onset of congestion, the probability that a specific connection is affected can be approximated by its share of the bottleneck bandwidth upon congestion.

Definition 2.1. *Denote by t_n the instant at which the nth congestion event occurs. Let $W_1(t_n)$ (resp. $W_2(t_n)$) be the window size of source 1 (resp. 2) just prior to this event. We* *assume that instants t_n are given by:*

$$X_1(t_n) + X_2(t_n) = W_1(t_n)/T_1 + W_2(t_n)/T_2 = \mu. \tag{2.1}$$

The probability that a source k (k=1,2) reduces its window at instant t_n is equal to:

$$p_k = X_k(t_n)/\mu = W_k(t_n)/(\mu T_k).$$

Our aim is to determine how fair the two connections share the bottleneck bandwidth μ, how well they utilize this bandwidth, and how much the network parameters affect the overall performance.

The preceding description gives rise to the Markovian model from [4] which we describe now. Given that the two processes $W_1(t_n)$ and $W_2(t_n)$ are related to each other by equation (2.1), we can transform the problem from a two-dimensional problem to a one dimensional-one. The study of one of the two processes is sufficient to describe the other. In the sequel we focus on the study of connection 1. We start by calculating the relationship between $W_1(t_n)$ and $W_1(t_{n+1})$ as well as the time $(t_{n+1} - t_n)$ between two consecutive congestion events. The window variation as a function of time and the sum of the rates at instants t_n and t_{n+1} are used. First we state the main results.

To simplify the expressions, we introduce some further notation: Let

$$a := (T_1)^2, b := (T_2)^2, c := (a^{-1} + b^{-1})^{-1}, \text{ and } r := \frac{b}{a}.$$

Let $S_{n+1} := t_{n+1} - t_n$. We define the instantaneous throughput of connection 1 by $X_n = W_1(t_n)/T_1 = X_1(t_n)$. The instantaneous throughput of connection 2 at t_n is then equal to $\mu - X_n$. The process X_n evolves as follows (see [4, Thm. 1]):

Theorem 2.1. *If connection 1 is hurt by congestion at instant t_n, the next congestion will occur after time*

$$S_{n+1} = \frac{c}{2} X_n,$$

and the instantaneous throughput of connection 1 prior to that event will be equal to,

$$X_{n+1} = \frac{1+2r}{2(1+r)} X_n. \tag{2.2}$$

If connection 2 is hurt by the congestion, connection 1 continues to increase its window without reduction until the next congestion event which occurs after a time,

$$S_{n+1} = \frac{c(\mu - X_n)}{2}.$$

In this case, the instantaneous throughput of connection 1 prior to the next congestion event will be equal to,

$$X_{n+1} = \frac{\mu r}{2(1+r)} + \frac{2+r}{2(1+r)} X_n. \tag{2.3}$$

The distribution of X_{n+1}, which can be viewed as the state of connection 1 at instant t_{n+1} is only a function of the state at time t_n. Thus, the stochastic process X_n forms a Markov chain. By Theorem 2.1 we have

$$X_{n+1} = \begin{cases} \frac{1+2r}{2(1+r)} X_n, & \text{w.p. } \frac{X_n}{\mu}, \\ \mu - \frac{2+r}{2(1+r)}(\mu - X_n), & \text{w.p. } 1 - \frac{X_n}{\mu}. \end{cases} \tag{2.4}$$

In the sequel, we shall study this Markov chain assuming it has reached stationarity.

3. Recursion for the moments of the throughput at loss instants

In this section we derive recursive equations for the moments of the throughput of both connections. These will be the basis for the subsequent analysis, when we obtain implicit equations for the probability distribution and its Laplace-Stieltjes transform, as well as analytic bounds for the mean in case of different round trip times and an exact expression for the symmetrical case.

Assuming the process X_n is in steady state (in which case we may omit the subscript n), we obtain

$$\begin{aligned} E[X^k] &= E\Big[E[(X_{n+1})^k | X_n]\Big] \\ &= E\left[\frac{X_n}{\mu}\left[\frac{1+2r}{1+r}\cdot\frac{X_n}{2}\right]^k + \frac{\mu - X_n}{\mu}\left[\frac{\mu r + (2+r)X_n}{2(1+r)}\right]^k\right] \\ &= Z_1(k) + Z_2(k) \end{aligned} \tag{3.1}$$

where

$$Z_1(k) = \frac{2(1+r)}{\mu(1+2r)}\left(\frac{1+2r}{2(1+r)}\right)^{k+1} E[X^{k+1}],$$

$$\begin{aligned} Z_2(k) &= -E\left[\frac{2(1+r)}{\mu(2+r)}\left(\frac{(2+r)(X_n-\mu)}{2(1+r)}\right)\left(\frac{\mu r + X_n(2+r)}{2(1+r)}\right)^k\right] \\ &= \frac{2(1+r)}{2+r}E\left[\left(\frac{\mu r + (2+r)X_n}{2(1+r)}\right)^k\right] \\ &\quad - \frac{2(1+r)}{\mu(2+r)}E\left[\left(\frac{\mu r + (2+r)X_n}{2(1+r)}\right)^{k+1}\right] \end{aligned}$$

Note that (3.1) enables us to compute recursively all moments of the distribution of X, once we know $E[X]$. In particular, equation (3.1) with $k=1$, $k=2$ and $k=3$ gives

$$(1-r)E[X^2] = \mu r\,(\mu - 2E[X]), \tag{3.2}$$

$$(1-r)E[X^3] = \frac{\mu r\left(\mu^2 r + \mu(4+r)E[X] - (8+5r)E[X^2]\right)}{3(1+r)}, \tag{3.3}$$

$$\begin{aligned} (1-r)E[X^4] &= \frac{\mu r}{7r^2+13r+7} \times \Big(-2(5r^2+15r+12r)E[X^3] \\ &\quad + 6\mu(2+r)E[X^2] + 2\mu^2 r(3+r)E[X] + \mu^3 r^2\Big) \end{aligned} \tag{3.4}$$

If $r=1$ (the symmetric case) then the coefficient of $E[X^{k+1}]$ in (3.1) vanishes. In particular, we directly get $E[X] = \mu/2$ from (3.2), $E[X^2] = 7\mu^2/26$ from (3.3) and $E[X^3] = 2\mu^3/13$ from (3.4).

4. Throughput distribution at loss instants

Before focusing on the mean throughput in the remainder of the paper, we first employ (3.1) to find an implicit expression for the distribution function of X. Define

$$\beta := \frac{1+2r}{2(1+r)}, \quad u := \frac{\mu r}{2(1+r)}, \quad v := \frac{2+r}{2(1+r)}.$$

Then we can rewrite

$$Z_1(k) = \frac{E[(\beta X)^{k+1}]}{\beta\mu},$$

$$Z_2(k) = \frac{1}{v}\Big(E[(u+vX)^k] - \mu^{-1}E[(u+vX)^{k+1}]\Big).$$

Define the Laplace Stieltjes transform (LST)

$$F(s) := E[\exp(-sX)] = \sum_{k=0}^{\infty} \frac{(-s)^k E[X^k]}{k!}.$$

Note that since X is bounded, $F(s)$ is well defined for all s. Then substituting (3.1), we obtain:

$$\begin{aligned}
F(s) &= \frac{1}{\beta\mu}\sum_{k=0}^{\infty} \frac{(k+1)(-s)^k E[(\beta X)^{k+1}]}{(k+1)!} \\
&\quad + \frac{1}{v}\sum_{k=0}^{\infty} \frac{(-s)^k E[(u+vX)^k]}{k!} \\
&\quad - \frac{1}{v\mu}\sum_{k=0}^{\infty} \frac{(k+1)(-s)^k E[(u+vX)^{k+1}]}{(k+1)!} \\
&= -\frac{1}{\mu}F'(\beta s) + \frac{1}{v}\exp(-us)F(vs) \\
&\quad + \frac{1}{\mu}\exp(-us)F'(vs) - \frac{u}{v\mu}e^{-us}F(vs).
\end{aligned}$$

To invert the above transforms, we use the relations between a probability distribution density $f(x)$ and its LST $F(s) = \int_0^\infty \exp(-sx)f(x)dx$:

$$\mathcal{L}^{-1}F(s/\alpha) = \alpha f(\alpha x), \qquad \mathcal{L}^{-1}(\exp(-sr)F(s)) = f(x-r),$$

$$\mathcal{L}^{-1}F(\alpha s)' = -\frac{x}{\alpha^2} f\Big(\frac{x}{\alpha}\Big).$$

We get the following relation for the probability density function $f(x)$ of X:

$$\begin{aligned}
f(x) &= \frac{1}{\beta^2\mu} x f\left(\frac{x}{\beta}\right) + \frac{1}{v^2} f\left(\frac{x-u}{v}\right) \\
&\quad - \frac{1}{\mu v^2}(x-u) f\left(\frac{x-u}{v}\right) - \frac{u}{v^2\mu} f\Big(\frac{x-u}{v}\Big).
\end{aligned}$$

5. Mean throughput at loss instants

We already observed in Section 3 that if $r = 1$ then all moments of X, i.e., the rate of connection 1 at instants of losses, can be obtained, starting with the mean $E[X] = \mu/2$. We now derive bounds on $E[X]$ when $r \neq 1$. Equation (3.2) may alternatively be written as

$$E[X^2] = r\, E[(\mu - X)^2]. \tag{5.1}$$

If $r \leq 1$, i.e., connection 1 is the one with the largest round trip time, then

$$\begin{aligned} E[X]^2 &= (1-r)E[X]^2 + rE[X]^2 \\ &\leq (1-r)E[X^2] + rE[X]^2 \\ &= rE[(\mu - X)^2 - X^2] + rE[X]^2 \\ &= r\,(\mu - E[X])^2\,. \end{aligned}$$

Therefore, we have the following bound on the ratio of the expected throughputs of the connections just before loss instants, when $r \leq 1$:

$$\frac{E[X]}{\mu - E[X]} \leq \sqrt{r} = \frac{RTT_2}{RTT_1}.$$

By symmetry we may write when $r \geq 1$

$$\frac{\mu - E[X]}{E[X]} \leq \frac{1}{\sqrt{r}}.$$

Together this gives a direct bound on the expected throughput of a connection just before a loss instant

$$E[X] \begin{cases} \leq \frac{\mu\sqrt{r}}{1+\sqrt{r}} & \text{if } r \leq 1, \\ \geq \frac{\mu\sqrt{r}}{1+\sqrt{r}} & \text{if } r \geq 1. \end{cases} \tag{5.2}$$

(Note that indeed $E[X] = \mu - E[X] = \mu/2$ when $r = 1$.)

We also obtain complementary bounds for $E[X]$, i.e., a lower bound when $r \leq 1$ and an upperbound when $r \geq 1$. Using $E[X^2] < \mu^2$ we have from (5.1)

$$\frac{\mu/2 - E[X]}{1-r} = \frac{E[X^2]}{2\mu r} \leq \frac{\mu}{2r}.$$

We may write this as

$$\left|\frac{\mu}{2} - E[X]\right| \leq |1-r|\,\frac{\mu}{2r}.$$

Combining this with (5.2) gives for $r < 1$

$$\frac{\mu}{2(1+\sqrt{r})^2} < \frac{\mu/2 - E[X]}{1-r} \leq \frac{\mu}{2r}$$

and the symmetric expression for $r > 1$.

6. Time average throughput

We denote the mean rate of connection $k = 1, 2$, at arbitrary time instants by $\overline{X}_k$. We shall now employ (5.2) to derive bounds for $\overline{X}_k$ and for $\overline{X}_1/\overline{X}_2$, as well as exact expressions for $\overline{X}_k$ for the symmetric case $r = 1$. This ratio is a measure for the fairness of bandwidth sharing between the two connections. By Theorem 2.1 we have, using the inversion formula [5, Ch. 1, Sec. 4] $\overline{X}_1 = E\left[\int_{t_{n-1}}^{t_n} X_1(t)dt\right]/E\left[S_n\right]$,

$$\begin{aligned}\overline{X}_1 &= \frac{E\left[\frac{cX^2}{2\mu}\left(\frac{X}{2}+\frac{cX}{4a}\right)+\frac{c(\mu-X)^2}{2\mu}\left(X+\frac{c(\mu-X)}{4a}\right)\right]}{E\left[\frac{cX^2}{2\mu}+\frac{c(\mu-X)^2}{2\mu}\right]} \\ &= \frac{E\left[cX^2\left(\frac{X}{2}+\frac{cX}{4a}\right)+(\mu-X)^2\left(X+\frac{c(\mu-X)}{4a}\right)\right]}{E\left[X^2+(\mu-X)^2\right]} \\ &= \frac{E\left[\frac{3}{2}X^3-\mu(2+\frac{3r}{4(1+r)})X^2+\mu^2(1+\frac{3r}{4(1+r)})X+\mu^3\frac{r}{4(1+r)}\right]}{E\left[2X^2-2\mu X+\mu^2\right]} \qquad (6.1)\end{aligned}$$

Substituting $E[X] = \mu/2$, $E[X^2] = 7\mu/26$ and $E[X^3] = 2\mu/13$ into (6.1) we obtain for $r = 1$

$$\overline{X}_1 = \overline{X}_2 = \frac{3}{7}\mu,$$

implying a utilization of 86% in the symmetric case. We note that this gives an excellent correspondence with previously obtained numerical results: the difference between our computed throughput and the one obtained numerically in [4, Fig. 2] is 0.4%.

For $r \neq 1$, we can write

$$\overline{X_1} = \mu\, h_1(E[X]), \qquad (6.2)$$

with

$$h_1(x) := \frac{(1+r)(4+9r)x-\mu r(7+6r)}{4(1-r)(1+r)(\mu-2x)}.$$

Note that if $r \neq 1$ then $h_1(x)$ is increasing in x for all values of x except at the point $x = \frac{1}{2}\mu$ where $h_1(x)$ switches from $+\infty$ to $-\infty$. Substituting (5.2) into (6.2) gives

$$\overline{X}_1 \begin{cases} \leq & \mu\sqrt{r}\frac{4-3\sqrt{r}+3r-3r\sqrt{r}}{4(1-r)(1+r)}, \quad \text{if } r<1, \\ \geq & \mu\sqrt{r}\frac{4-3\sqrt{r}+3r-3r\sqrt{r}}{4(1-r)(1+r)}, \quad \text{if } r>1. \end{cases} \qquad (6.3)$$

By symmetry, replacing $E[X]$ and r in (6.2) with $\mu - E[X]$ and $1/r$, resp., we also have

$$\overline{X_2} = \mu\frac{(1+r)(4r+9)(\mu-E[X])-\mu(7r+6)}{4(1-r)(1+r)(\mu-2E[X])}. \qquad (6.4)$$

We note that the above bounds do not perform well for r in the neighborhood of 1.

From (6.3) we have that, when $r \to 0$, then $\overline{X}_1/(\mu\sqrt{r}) \leq 1$. In Section 7 we show that the order of this upper bound is exact, i.e., $\overline{X}_1$ tends to zero as $\sqrt{r}$ (and not faster) when $r \to 0$.

Note also that $\overline{X}_2 \to \frac{3}{4}\mu$ when $r \to 0$, which indeed corresponds to the throughput of a single connection.

For $r \neq 1$, (6.2) and (6.4) give

$$\frac{\overline{X}_1}{\overline{X}_2} = h(E[X]) \tag{6.5}$$

where

$$h(x) := \frac{(1+r)(4+9r)x - \mu r(7+6r)}{(1+r)(4r+9)(\mu - x) - \mu(7r+6)}.$$

If $r \neq 1$, then h is increasing in x for all values of x except at

$$x = \mu \frac{3+6r+4r^2}{(1+r)(4r+9)},$$

at which point $h(x)$ switches from $+\infty$ to $-\infty$. Substituting (5.2) into (6.5) gives

$$\frac{\overline{X}_1}{\overline{X}_2} \begin{cases} \leq \quad \sqrt{r}\left(\frac{4-3\sqrt{r}+3r-3r\sqrt{r}}{3-3\sqrt{r}+3r-4r\sqrt{r}}\right), & \text{if } r < r_0, \\ \geq \quad \sqrt{r}\left(\frac{4-3\sqrt{r}+3r-3r\sqrt{r}}{3-3\sqrt{r}+3r-4r\sqrt{r}}\right), & \text{if } r > 1/r_0, \end{cases} \tag{6.6}$$

where $x = r_0 \approx 0.32$ is the unique root in $(0,1)$ of $-3 + 7\sqrt{x} - 6x + 7x\sqrt{x} - 3x^2$. We shall show that these bounds give the right order of magnitude when $r \to 0$ (or, by symmetry, when $r \to \infty$). More specifically, we show that

$$\liminf_{r \to 0} \frac{1}{\sqrt{r}} \cdot \frac{\overline{X}_1}{\overline{X}_2} \geq \frac{2}{3}. \tag{6.7}$$

Note that from (6.6) we also have

$$\limsup_{r \to 0} \frac{1}{\sqrt{r}} \cdot \frac{\overline{X}_1}{\overline{X}_2} \leq \frac{4}{3}. \tag{6.8}$$

This proves that the bounds have the correct order of magnitude in the regime $r \to 0$: $\overline{X}_1 \sim \sqrt{r} = RTT_2/RTT_1$. We shall prove (6.7) by showing the equivalent asymptotic behavior for $E[X]$ (the mean rate of connection 1 at loss instants). This is done in Section 7. Note that (6.6) is only informative for $r < r_0 < 1$ and $r > 1/r_0 > 1$. However, in Section 8 we derive an approximation for $\overline{X}_1/\overline{X}_2$ which performs well for all values of r (and also has the right asymptotics for small and large values of r).

7. Asymptotic bound at loss instants

From (2.4) it is straightforward to see that if $X_n = x$ then the drift is positive if $x < x_0$, negative if $x > x_0$ and zero if $x = x_0$, where $x_0 := \mu\sqrt{r}/(1+\sqrt{r})$. This feeds the intuition that the process X_n "tends to be in the neighborhood of x_0". Note that for $r \to 0$ we have

$x_0 \sim \mu\sqrt{r}$. We shall now construct a process Y_n that mimics this behavior of X_n, but is easier to analyze. More specifically, the Markov chain Y_n evolves according to

$$Y_{n+1} = \begin{cases} \frac{(1+2r)Y_n}{2(1+r)}, & \text{w.p. } \begin{cases} t/\mu, & \text{if } Y_n \le t, \\ 1, & \text{if } Y_n > t, \end{cases} \\ \frac{(2+r)Y_n + r\mu}{2(1+r)}, & \text{w.p. } \begin{cases} \mu - t/\mu, & \text{if } Y_n \le t, \\ 0, & \text{if } Y_n > t, \end{cases} \end{cases} \tag{7.1}$$

where $t \in (0, \mu)$ is an arbitrary threshold. (We shall later choose $t = 2\mu\sqrt{r}$ to obtain the sharpest bound possible using this construction.) By a simple coupling argument it can be shown that if $P(X_0 \le x) \le P(Y_0 \le x)$ for all $0 \le x \le \mu$, then also $P(X_n \le x) \le P(Y_n \le x)$ for all $0 \le x \le \mu$ and $n \ge 1$. Therefore $E[X] \ge E[Y]$ where Y has the stationary distribution of the process Y_n.

From (7.1) we get

$$E[Y] = \frac{(1 - t/\mu)\left(r\mu - (1-r)\, E[Y\, \mathbf{1}(Y > t)]\right)}{r + (1-r)t/\mu}, \tag{7.2}$$

where $\mathbf{1}(\cdot)$ is the indicator function. Note that

$$\begin{aligned} E[Y\, \mathbf{1}(Y > t)] &\le \left(\frac{(r+2)t}{2(r+1)} + \frac{r\mu}{2(r+1)}\right) P(Y > t) \\ &= \left(t + \frac{r(\mu - t)}{2(r+1)}\right) P(Y > t). \end{aligned} \tag{7.3}$$

Next, we shall bound the probability $P(Y > t)$ from above. To this end we first write $P(Y > t) = 1/(1 + \tau_t)$, where τ_t is the return time to the set $\{Y > t\}$. This is justified because if $Y_n > t$ then $Y_{n+1} < t$. Of course, such a return time could start from different points above t, but when the starting point makes a difference we shall take the "worst case" so as to obtain a lower bound for τ_t and, hence, an upper bound for $P(Y > t)$. We shall prove (for later explanation these formulas have not been compressed)

$$\begin{aligned} \tau_t \;&\ge\; \hat{\tau}_t := K + \frac{t}{\mu}\sum_{k=1}^{K} a_k, \qquad (7.4) \\ K &:= \frac{t - \frac{2r+1}{2(r+1)}\left(t + \frac{r(\mu-t)}{2(r+1)}\right)}{r\mu/(2(r+1))}, \\ a_k &:= \left(\frac{2r+1}{2(r+1)}\left(t + \frac{r(\mu - t)}{2(r+1)}\right) + \frac{kr\mu}{2(r+1)}\right) \times \frac{(2r+1)/(2(r+1))}{r\mu/(2(r+1))}. \end{aligned}$$

K is the minimum number of steps for the process Y_n to get back above level t after it has dropped below. To see this, note that after dropping below t the process is surely below level $\frac{2r+1}{2(r+1)}\left(t + \frac{r(\mu-t)}{2(r+1)}\right)$. Since

$$E[Y_{n+1} - Y_n] \le r\mu/(2(r+1)),$$

it takes at least K steps to be above level t again.

At each of these K steps a new reduction can take place with probability t/μ. If a reduction takes place at the k-th step, it takes at least a_k additional steps to "recover" from this reduction (more if there are new intermediate reductions, but we shall neglect those). The expression for a_k is constructed in the same way as that for K.

Since $E[X] \geq E[Y]$ we have from (7.2), (7.3) and (7.4), for any choice of t

$$E[X] \geq \frac{r(\mu - t)}{r\mu + (1-r)t}\left(\mu - \frac{r\mu + (1-r)\left(t + \frac{r(\mu-t)}{2(r+1)}\right)}{r(1+\hat{\tau}_t)}\right).$$

Now choose $t = t(r) = c\sqrt{r}$ for some constant $c > 0$ independent of r. Note that

$$\lim_{r\to 0} \frac{r\hat{\tau}_{t(r)}}{t(r)} = \frac{1}{\mu}\left(1 + \frac{c^2}{4\mu^2}\right),$$

and, hence,

$$\liminf_{r\to 0} \frac{1}{\sqrt{r}} E[X] \geq \frac{c\mu^2}{4\mu^2 + c^2} = \frac{1}{2}\mu,$$

where we set $c = 2\mu$, for which the bound is the sharpest. Together with (6.2) this proves (6.7).

8. Approximation

We can derive a surprisingly accurate approximation for $\overline{X}_1/\overline{X}_2$ if we approximate $\overline{X}_1$ and $\overline{X_2}$ by the average throughputs in between two consecutive losses:

$$\begin{aligned}\overline{X}_1 &\approx \frac{1}{2}E[X] + \frac{1}{2}\left(E[X] - \frac{1}{2\mu}E[X^2]\right)\\ &= \frac{2-r}{2(1-r)}E[X] - \frac{r\mu}{4(1-r)},\end{aligned}$$

and the symmetrical expression for $\overline{X}_2$. This gives

$$\frac{\overline{X}_1}{\overline{X}_2} \approx \frac{(2-r)\,E[X] - \frac{1}{2}\mu r}{(2r-1)\,E[X] + \frac{1}{2}\mu\,(3-4r)}. \tag{8.1}$$

Using (5.2) as an approximation for $E[X]$ gives

$$\frac{\overline{X}_1}{\overline{X}_2} \approx \frac{\sqrt{r}\,(4 + 3\sqrt{r})}{3 + 4\sqrt{r}}. \tag{8.2}$$

Comparing this with the numerical results from [4] shows that it provides a good approximation of the true ratio (the error being in the order of a few percent). This approximation being close to the true ratio, it also "explains" the observation in [4] that the ratio of the throughputs is well approximated by $(RTT_2/RTT_1)^{0.85} = (\sqrt{r})^{0.85}$: for moderate values of r this function is close to the approximation (for $r = 0.1$, 0.5, 0.8 the relative error is 8%, 2% and 0.6%, resp.). Moreover, the approximation matches the correct order of magnitude when $r \to 0$ (or when $r \to \infty$).

9. Conclusions

We have studied the throughput of two TCP connections that share a bandwidth μ without assuming that losses occur simultaneously to both connections. The transmission rates at loss instants were modeled as a Markov chain. By deriving a recursion on the moments of the transmission rates, we showed that all moments and, hence, the distribution can be determined once the means are known. For the mean transmission rates and their ratio we obtained bounds that proved to be very close to the numerical computations in [4]. Regarding fairness in bandwidth sharing we obtained the approximation (8.2), where $\sqrt{r} = R := RTT_2/RTT_1$ is the ratio of the round trip times. We proved that for $R \to 0$ this approximation has the right order of magnitude, i.e., $\overline{X}_1/\overline{X}_2$ is of the order R when $r \to 0$; this is formalized in (6.7) and (6.8). This contradicts the order R^α with $1 < \alpha \leq 2$ predicted by models assuming synchronized losses. The same order, however, is predicted by models for many competing TCP connections where the loss process for one connection is independent of its transmission rate. This suggests that the order of magnitude R is valid throughout the whole spectrum: for many and for few competing connections. However, it remains to verify this assertion for few, but more than two, connections.

References

1. E. Altman, K. Avratchenkov and C. Barakat, "A stochastic model of TCP/IP with stationary random losses", *ACM SIGCOMM 2000.*
2. E. Altman, K. Avratchenkov, C.Barakat, and R. Núñez-Queija, "State-dependent M/G/1 type queueing analysis for congestion control in data networks", *IEEE INFOCOM 2001.*
3. E. Altman, K. Avratchenkov, C. Barakat and R. Nunez Queija, "TCP modeling in the presence of nonlinear window growth", *Proceedings of the 17th International Teletraffic Congress*, Salvador da Bahia, Brazil, September 24-28, 2001.
4. E. Altman, C. Barakat, E. Laborde, P. Brown and D. Collange, "Fairness Analysis of TCP/IP", *Proceedings of IEEE Conference on Decision and Control (CDC'00)*, Sydney, Australia, December 2000.
5. F. Baccelli, P. Brémaud. *Palm probabilities and stationary queues.* Lecture Notes in Statistics, 41. Springer-Verlag, 1987.
6. F. Baccelli and D. Hong, "A.I.M.D, Fairness and Fractal Scaling of TCP Traffic" Technical Report, April 2001, RR-4155, INRIA Rocquencourt, France, 2001.
7. C. Barakat and E. Altman, "A Markovian Model for TCP Analysis in a Differentiated Services Network", *Workshop on Quality of future Internet Services*, Sep 2000.
8. B. Braden, et al.,"Recommendations on Queue Management and Congestion Avoidance in the Internet", *RFC 2309*, Apr 1998.
9. P. Brown, "Resource sharing of TCP connections with different round trip times", *IEEE Infocom*, Mar 2000.
10. K. Fall and S. Floyd, "Simulation-based Comparisons of Tahoe, Reno, and SACK TCP", *Computer Communication Review*, Jul 1996.
11. S. Floyd and V. Jacobson, "Random Early Detection gateways for Congestion Avoidance", *IEEE/ACM Transactions on Networking*, Aug 1993.
12. T.V. Lakshman and U. Madhow, "The performance of TCP/IP for networks with high

bandwidth-delay products and random loss", *IEEE/ACM Transactions on Networking*, Jun 1997.

13. J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP Throughput: a Simple Model and its Empirical Validation", *ACM SIGCOMM*, Aug. 1998.
14. W. Stevens, "TCP Slow-Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms", *RFC 2001*, Jan 1997.
15. L. Zhang, S. Shenker, and D.D. Clark, "Observations on the Dynamics of a Congestion Control Algorithm: The Effects of Two-Way Traffic", *ACM SIGCOMM*, Sep 1991.