Impact of bursty losses on TCP performance

Eitan Altman, Konstantin Avrachenkov, Chadi Barakat

INRIA, 2004 route des Lucioles, 06902 Sophia Antipolis, France
Email:{altman,kavratch,cbarakat}@sophia.inria.fr

Abstract

In this paper we analyze the performance of a TCP-like flow control mechanism in a lossy environment. The transmission rate in the control scheme that we consider has a linear growth rate; whenever a loss occurs, the transmission rate is halved. This approximates the performance of several versions of TCP that divide their congestion window by two when a loss is detected. We propose a mathematical model that allows to account for burstiness in the loss process. We compute the expected transmission rate and its moments at some potential loss instants, and provide useful implicit and explicit expressions for the Laplace Stieltjis Transform. This allows us to compute explicitly the time average of the transmission rate as well as its moments. We show that the time average of the transmission rate is indeed sensitive to the distribution of losses, and not just to the average loss rate: for a given average loss rate, we show that the time average of the transmission rate *increases* with the burstiness of losses.

1 Introduction

Flow control mechanisms in the Internet, particularly those of the Transmission Control Protocol (TCP) [19,24], use the loss of packets as an indication of network congestion. In general, the transmission rate of the controlled flow is linearly increased until a loss occurs. The network is supposed here to be congested and the transmission rate is multiplicatively decreased in order to alleviate this congestion. In TCP as an example, the transmission rate is controlled via a congestion window which is increased in absence of losses and decreased upon loss detection [24]. Another set of flow control mechanisms recently introduced into multimedia applications consist in measuring the loss rate of packets and in controlling the transmission rate in a way to be friendly with TCP transfers [15,16]. Explicit expressions for TCP throughput for a given loss rate (e.g.[23]) are used for this purpose.

A good understanding of the impact of a loss process on the performance of a flow control mechanism is required for a good network and protocol tuning.

Preprint submitted to Elsevier Preprint *30 May 2000*

Several previous works have addressed the problem of TCP performance as a function of data packet losses. The focus on TCP is due to the dominance of TCP traffic in today's Internet. Some of these works [14,20,21,23] have studied the impact of the intensity of losses (or the average loss rate) on the performance. TCP packets are assumed to be lost independently with the same probability. Explicit expressions for the throughput of the TCP connection are derived by simply dividing the average number of packets transmitted between losses to the average time between losses. No other parameter than the packet loss probability is used to characterize the distribution of loss instants over time. As we will see in this paper, this will cause a wrong estimation of the throughput when a certain burstiness of losses exists. Other works [6,10,11] have addressed the problem of burstiness of packet losses but in the wireless environment context. It is known that due to multiple phenomena such as multi-path fading [12], wireless links as those found in terrestrial wireless networks or satellite networks present a certain degree of transmission error burstiness. The impact of consecutive packet losses on the different versions of TCP is studied in these works [6,10,11]. They model losses with a two-state Markov chain where small bursts of losses appear in an independent and uniform manner. They study then TCP performance as a function of the average rate of bursts as well as the average burst size. But, the new versions of TCP (New Reno, SACK) [13] are able to resist to consecutive packet losses and to reduce their window once for all packet losses in the same Round Trip Time (RTT) (i.e. from the same window of data). This will result in these models becoming similar to the previous ones since they study the impact on TCP performance of only the average rate at which the window is reduced.

In this paper we propose a completely different model for a TCP-like flow control protocol that, in addition to the average rate of window reduction events, accounts for the burstiness of these events. Rather than looking at the packet level and considering the probability that a packet is lost, we look at the transmission rate level and consider the moments at which the transmission rate is reduced. We associate then a loss process to these moments. A loss event is equivalent to a transmission rate reduction event. This loss can be the result of a single packet loss or multiple consecutive packet losses from the same window. This depends on the version of TCP and its reaction to packet losses. Our aim is to study the impact of burstiness of this loss process on the throughput of the connection. We compute the expectation of the transmission rate as well as expressions for its Laplace Stieltjis Transform at some potential loss instants. This allows us to compute the time average of the transmission rate which we call the throughput of the transfer. We show that the throughput is indeed sensitive to the distribution of losses, and not just to the average loss rate: for a given average loss rate, we show that the throughput *increases* with the burstiness. Our results are compared to simulations done with the `ns` simulator developed at LBNL [22] and a good match is reported.

The structure of the paper is as follows. In the next section we present our model for losses and for the controlled rate. Section 3 contains our analysis of the performance of the transmission rate in presence of losses. At the end of this section, we give the general expression of the throughput. The throughput in the case of an independent loss process having the same average loss rate as a bursty loss process is defined. This latter throughput is then used as a reference to show the effect of burstiness. In section 4, we study the impact of the parameters of the loss process on the performance. The analytical results are compared to simulation ones. The paper is concluded in section 5.

2 The model

Consider a flow control mechanism where the transmission rate grows linearly at a rate α per unit of time. The growth continues until a loss occurs. The transmission rate is halved and the linear growth is then resumed. This model approximates the performance of several versions of TCP. Indeed, the transmission rate of a TCP connection at any instant is equal to the window size divided by the RTT. The window in turn increases by 1 packet for every window's worth of Acknowledgments (ACK) [19]. This results in a linear rate increase by a factor $\alpha = 1/RTT^2$ if ACKs are not delayed at the destination and $\alpha = 1/2RTT^2$ if ACKs are delayed [24] (the linear growth is known to hold for TCP connections in which the bandwidth-delay product is large in comparison with queueing delays, since in that case, RTT is almost constant. If queueing delays are large, however, the actual growth is sub-linear, see [1]). This linear window increase corresponds to the congestion avoidance mode of TCP. The slow start mode is neglected in this paper due its fast exponential window increase. The model can also approximate any additive-increase multiplicative-decrease flow control mechanism.

Let us propose a model which accounts for burstiness of losses. The Gilbert model is often used in this context [12]. The path between the source and the destination called channel in the wireless terminology is assumed to have two states: *Good* and *Bad*; losses are assumed to occur in the Bad state. The time during which the channel is in a Good or in a Bad state is taken to be geometrically (or exponentially) distributed. We propose an extension of this model in order to handle generally distributed periods of Good and Bad states. Our model is related to the MAP (Markovian Arrival Process) process [18]. We allow losses to occur both in the Good state as well as in the Bad state; the occurrence of losses in each of these states is different. To that end, we define a series of potential losses. Let T_n denote the time at which the nth potential loss may occur. Let $D_n, n = 1, 2, \ldots$ be the sequence of times between potential losses: $D_n = T_{n+1} - T_n$. D_n are assumed to be i.i.d. with expectation d, second moment $d^{(2)}$ and Laplace Stieltjis Transform $D^*(s)$. Let X_n be the transmission rate just prior to the instant of the nth potential loss.

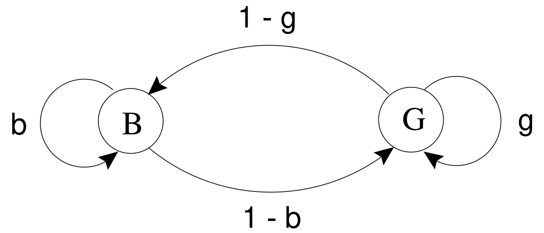


Fig. 1. The Markov chain associated to the path

Potential losses are transformed to real losses with a certain probability. This is similar to MAP processes in which at each state transition an arrival can occur with a probability that depends on the state. Let Y_n be the state of the path at the nth potential loss instant. We consider the states B (for Bad) in which a potential loss is transformed to a real loss with probability p_B, and G (for Good) in which it is transformed with a smaller probability p_G. We shall assume throughout that $p_G \leq p_B$ and that $p_B > 0$. We assume further that the sequences $\{Y_n\}$ and $\{D_n\}$ are independent.

The random process $\{Y_n\}$ is assumed to be a Markov chain with the following transition matrix (Figure 1)

$$P = \begin{bmatrix} g & \bar{g} \\ \bar{b} & b \end{bmatrix}.$$

State 1 (2) corresponds to the Good (Bad) state of the path. We shall assume throughout the paper that $g, b \in (0, 1)$. The Markov chain $\{Y_n\}_{n=1}^{+\infty}$ is then ergodic with stationary probabilities,

$$\pi_G = \frac{1-b}{2-b-g} = \frac{\bar{b}}{\bar{b}+\bar{g}}, \qquad \pi_B = \frac{1-g}{2-b-g} = \frac{\bar{g}}{\bar{b}+\bar{g}}.$$

The average loss rate is given by,

$$R = \frac{p_G \pi_G + p_B \pi_B}{d}.$$

This is equal to the average number of times the source reduces its rate per unit of time.

3 Performance Analysis

Define the two random variables U_n and V_n describing the behavior of the transmission rate when a potential loss occurs. They correspond to the two states of the path. A value one of these variables means that the potential loss causes really a reduction in the transmission rate. A value zero however means that X_n is not affected (i.e. a real loss didn't occur). We have,

$$P(U_n = 1) = p_G, \qquad P(U_n = 0) = 1 - p_G$$

$$P(V_n = 1) = p_B, \qquad P(V_n = 0) = 1 - p_B$$

The evolution of the transmission rate is the following,

$$\begin{aligned} X_{n+1} &= (1 - U_n)X_n 1\{Y_n = G\} + U_n \frac{X_n}{2} 1\{Y_n = G\} \qquad (1) \\ &+ (1 - V_n)X_n 1\{Y_n = B\} + V_n \frac{X_n}{2} 1\{Y_n = B\} + \alpha D_n \\ &= \left(1 - \frac{U_n}{2}\right) X_n 1\{Y_n = G\} + \left(1 - \frac{V_n}{2}\right) X_n 1\{Y_n = B\} + \alpha D_n \qquad (2) \end{aligned}$$

1{A} is the indicator function which is equal to 1 if expression A is true and to 0 otherwise.

Define the column vector

$$\mathbf{X}_n = (X_n 1\{Y_n = G\}, X_n 1\{Y_n = B\})^T.$$

Define the matrix

$$Q_n = \begin{pmatrix} (1 - U_n/2)1\{Y_{n+1} = G\} & (1 - V_n/2)1\{Y_{n+1} = G\} \\ (1 - U_n/2)1\{Y_{n+1} = B\} & (1 - V_n/2)1\{Y_{n+1} = B\} \end{pmatrix}.$$

Finally, define the column vector

$$\mathbf{D}_n = (D_n 1\{Y_n = G\}, D_n 1\{Y_n = B\})^T.$$

Then it follows from (2)

$$\mathbf{X}_{n+1} = Q_n \mathbf{X}_n + \alpha \mathbf{D}_n. \qquad (3)$$

We begin by studying the convergence and stability of the process $\mathbf{X}_n$. The next result follows from [8] or from Theorem 2A in [17]. (To show that the conditions of the Theorem indeed hold, one may follow the approach in the Appendix in [2].)

Theorem 1 *Assume that Y_n contains a single recurrent class and is initially in steady state. Consider an arbitrary initial state $\mathbf{X}_0$. Then,*

$$\mathbf{X}_n^* = \alpha \sum_{j=1}^{\infty} \left(\prod_{i=n-j}^{n-1} Q^i \right) \mathbf{D}_{n-j-1} \qquad (4)$$

is the only solution of (3) and is ergodic. The sum of the right hand of (4) converges absolutely almost surely. Furthermore, $|\mathbf{X_n} - \mathbf{X_n}^*| \to 0$ a.s. for all $\mathbf{X}_0$ on the same probability space as $\{(Q_n, \mathbf{D}_n)\}$. In particular, the distribution of $\mathbf{X}_n$ converges to that of $\mathbf{X}_n^*$ as $n \to +\infty$.

Next we study the existence of moments of X_n. We define for this purpose the following Laplace Stieltjis Transforms (LST),

$$Z_n(s,G) = E\left[e^{-sX_n}1\{Y_n = G\}\right], \qquad Z_n(s,B) = E\left[e^{-sX_n}1\{Y_n = B\}\right]$$

In addition, let us define

$$\mathbf{Z}_n(s) = \left[Z_n(s,G) \; Z_n(s,B) \right].$$

Theorem 2 *The Laplace Stieltjis Transform* $\mathbf{Z}_n(s)$ *satisfies the following recurrent equation,*

$$\mathbf{Z}_{n+1}(s) = D^*(\alpha s)\mathbf{Z}_n(s)P_1 + D^*(\alpha s)\mathbf{Z}_n(\frac{s}{2})P_2, \tag{5}$$

where

$$P_1 = \begin{bmatrix} g(1-p_G) & \bar{g}(1-p_G) \\ \bar{b}(1-p_B) & b(1-p_B) \end{bmatrix} \quad \textit{and} \quad P_2 = \begin{bmatrix} gp_G & \bar{g}p_G \\ \bar{b}p_B & bp_B \end{bmatrix}.$$

Proof: We write,

$$\begin{aligned}
&E\left[e^{-sX_{n+1}}1\{Y_{n+1} = G\}\right] \\
&= gE\left[e^{-s((1-\frac{U_n}{2})X_n+\alpha D_n)}1\{Y_n = G\}\right] + \bar{b}E\left[e^{-s((1-\frac{V_n}{2})X_n+\alpha D_n)}1\{Y_n = B\}\right] \\
&E\left[e^{-sX_{n+1}}1\{Y_{n+1} = B\}\right] \\
&= \bar{g}E\left[e^{-s((1-\frac{U_n}{2})X_n+\alpha D_n)}1\{Y_n = G\}\right] + bE\left[e^{-s((1-\frac{V_n}{2})X_n+\alpha D_n)}1\{Y_n = B\}\right]
\end{aligned}$$

Using the fact that,

$$\begin{aligned}
E\left[e^{-s(1-\frac{U_n}{2})X_n}1\{Y_n = G\}\right] &= (1-p_G)Z_n(s,G) + p_G Z_n(s/2,G), \\
E\left[e^{-s(1-\frac{V_n}{2})X_n}1\{Y_n = B\}\right] &= (1-p_B)Z_n(s,B) + p_B Z_n(s/2,B)
\end{aligned}$$

we obtain the required relations. ◇

Now with the help of recurrent equation (5) we can investigate the convergence of moments $E[X_n^k]$ for an arbitrary initial state X_0. First we define

$$\mathbf{x}^{(k)} = \left[x_G^{(k)} \; x_B^{(k)} \right]$$

where

$$x_G^{(k)} = \lim_{n\to\infty} E[X_n^k 1\{Y_n = G\}], \quad x_B^{(k)} = \lim_{n\to\infty} E[X_n^k 1\{Y_n = B\}]. \tag{6}$$

Note that the convergence of the above moments implies the convergence of $E[X_n^k]$. Furthermore,

$$x^{(k)} = \lim_{n\to\infty} E[X_n^k] = x_G^{(k)} + x_B^{(k)}.$$

In the next theorem we formulate conditions for existence of the limits in (6).

Theorem 3 *Let first k moments of D_n exist and p_G or p_B be positive. Then the moments $x_G^{(k)}$ and $x_B^{(k)}$ exist and can be calculated from the following recurrent relation*

$$\mathbf{x}^{(k)} = \sum_{i=1}^{k} C_k^i \alpha^i d^{(i)} \mathbf{x}^{(k-i)} [P_1 + \frac{1}{2^{k-i}} P_2][I - P_1 - \frac{1}{2^k} P_2]^{-1}, \tag{7}$$

where $\mathbf{x}^{(0)} = [\pi_G \ \pi_B]$. Finally, $x^{(k)} = x_G^{(k)} + x_B^{(k)}$ are also moments of the process X_n in the stationary regime.

Proof: We first differentiate k times the recurrent relation (5) with respect to s

$$\mathbf{Z}_{n+1}^{(k)}(s) = \sum_{i=0}^{k} C_k^i \alpha^i D^{*(i)}(\alpha s) \mathbf{Z}_n^{(k-i)}(s) P_1 + \sum_{i=0}^{k} C_k^i \alpha^i D^{*(i)}(\alpha s) \frac{1}{2^{k-i}} \mathbf{Z}_n^{(k-i)}(\frac{s}{2}) P_2$$

where C_k^i is the binomial coefficient. Then, we take $s = 0$ to get

$$\mathbf{Z}_{n+1}^{(k)}(0) = \sum_{i=0}^{k} C_k^i \alpha^i D^{*(i)}(0) \mathbf{Z}_n^{(k-i)}(0) [P_1 + \frac{1}{2^{k-i}} P_2]. \tag{8}$$

We recall that

$$\mathbf{Z}_n^{(k)}(0) = (-1)^k \left[E[X_n^k 1\{Y_n = G\}] \ E[X_n^k 1\{Y_n = B\}] \right].$$

Let us introduce the following augmented vector

$$\Xi_n^{(k)} = \left[\mathbf{Z}_n^{(0)}(0) \ \mathbf{Z}_n^{(1)}(0) \ \cdots \ \mathbf{Z}_n^{(k)}(0) \right].$$

Then recursions (8) can be written in the following matrix form

$$\Xi_{n+1}^{(k)} = \Xi_n^{(k)} \Pi$$

where

$$\Pi = \begin{bmatrix} P & \alpha d^{(1)}P & \alpha^2 d^{(2)}P & \cdots & \alpha^k d^{(k)}P \\ 0 & P_1+\frac{1}{2}P_2 & 2\alpha d^{(1)}[P_1+\frac{1}{2}P_2] & \cdots & \alpha^{k-1}d^{(k-1)}[P_1+\frac{1}{2}P_2] \\ 0 & 0 & P_1+\frac{1}{4}P_2 & \cdots & \alpha^{k-2}d^{(k-2)}[P_1+\frac{1}{4}P_2] \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & & \cdots & P_1+\frac{1}{2^k}P_2 \end{bmatrix}$$

Note that if $p_G > 0$ or $p_B > 0$, then all matrices $[P_1 + \frac{1}{2^k}P_2], k \geq 1$ are sub-stochastic and hence they have eigenvalues with modulus less than one. P is the transition matrix of ergodic Markov chain. Therefore it has only one eigenvalue equal to one. Since the spectrum of Π is the union of spectrums of diagonal sub-matrices $[P_1 + \frac{1}{2^k}P_2]$ and P, we conclude that Π has only one eigenvalue equal to one and the other eigenvalues with modulus less than one. The latter implies that the powers of Π converge to the eigenprojection corresponding to the eigenvalue one. Consequently, the moments $E[X_n^k 1\{Y_n = G\}]$ and $E[X_n^k 1\{Y_n = B\}]$ are also convergent.

Once the convergence is proven, we can let n go to infinity in (8) to obtain (7). The last statement of the theorem follows immediately from the existence of limits (6) and Theorem 1. ◇

Corollary 4 *Let $E[D_n] < \infty$ and p_G or p_B be positive. Then*

$$x_G = x_G^{(1)} = \alpha d \frac{\gamma_B(\pi_B - b) + \pi_G}{1 - \gamma_B b - \gamma_G g + \gamma_B \gamma_G (g + b - 1)}$$

$$x_B = x_B^{(1)} = \alpha d \frac{\gamma_G(\pi_G - g) + \pi_B}{1 - \gamma_B b - \gamma_G g + \gamma_B \gamma_G (g + b - 1)}$$

with $\gamma_G = 1 - \frac{p_G}{2}$, $\gamma_B = 1 - \frac{p_B}{2}$.

If X_n is in the stationary regime, then the LST $\mathbf{Z}(s)$ of its distribution function satisfy the following implicit equation

$$\mathbf{Z}(s) = D^*(\alpha s)\mathbf{Z}(s)P_1 + D^*(\alpha s)\mathbf{Z}(\frac{s}{2})P_2$$

By repeated iterations, it is possible to write $\mathbf{Z}(s)$ in an explicit form (for s such that $Re(s) \geq 0$). To that end, for any sequence A_n of square matrices of the same size, we shall use the following notation:

$$\prod_{i=0}^{n} A_i := A_n \times ... \times A_1 \times A_0.$$

Assuming that we start initially at the stationary regime, we have

$$
\begin{aligned}
\mathbf{Z}(s) = &\mathbf{Z}_0(s)\\
&= \mathbf{Z}_1(s) = D^*(\alpha s)\mathbf{Z}(s/2)P_2[I - D^*(\alpha s)P_1]^{-1}\\
&= \mathbf{Z}_2(s)\\
&= D^*(\alpha s)D^*(\alpha s/2)\mathbf{Z}(s/4)P_2[I - D^*(\alpha s/2)P_1]^{-1}P_2[I - D^*(\alpha s)P_1]^{-1}\\
&= \ldots\\
&= \mathbf{Z}_{n+1}(s) = \left(\prod_{i=0}^{n} D^*(\alpha s/2^i)\right)\mathbf{Z}(s/2^n)\prod_{i=0}^{n}(P_2[I - D^*(\alpha s/2^i)P_1]^{-i})
\end{aligned}
$$

Since this holds for any n, we conclude that the following limit exists

$$
\lim_{n\to\infty}\mathbf{Z}(s/2^n)\prod_{i=0}^{n}\left(D^*(\alpha s/2^i)P_2[I - D^*(\alpha s/2^i)P_1]^{-i}\right) = \mathbf{Z}(s)
$$

Since $\lim_{n\to\infty}\mathbf{Z}(s/2^n) = (\pi_G, \pi_B)$, and since the product above is uniformly bounded in n (for $Re(s) \geq 0$), we further conclude that

$$
\mathbf{Z}(s) = \lim_{n\to\infty}(\pi_G, \pi_B)\prod_{i=0}^{n}\left(D^*(\alpha s/2^i)P_2[I - D^*(\alpha s/2^i)P_1]^{-i}\right),
$$

where the above limit is well defined.

3.1 Calculation of the throughput

Denote by $\bar{x}$ the throughput of the transmission..

Theorem 5 *Let $E[D_n] < \infty$ and $E[D_n^2] < \infty$. Then, the throughput can be expressed as*

$$
\bar{x} = \lim_{t\to+\infty}\frac{1}{t}\int_0^t X(t)dt = \gamma_G x_G + \gamma_B x_B + \frac{1}{2}\alpha\frac{d^{(2)}}{d}
$$

where x_G and x_B are given in Corollary 4.

Proof: From the last statement of Theorem 1 we conclude that $\{X_n\}$ is ergodic Markov chain. Hence, $\{T_n, X_n\}$ is an ergodic marked point process. From [9, Ch. 4] it follows that the associated continuous time process of the transmission rate evolution $X(t)$ is ergodic as well. The latter fact implies that the throughput, that is, the time average transmission rate, is a.s. equal to $E[X(t)]$, the expectation of the transmission rate at arbitrary time moment. This expectation can be calculated by using the following inversion formula of the Palm theory (see e.g. [7, Ch.1, Sec.4])

$$
E[X(t)] = \frac{1}{d}E^0[\int_0^{T_1} X(t)dt] \tag{9}
$$

where $E^0[.]$ is an expectation associated with Palm distribution. In particular, $P^0(T_1 = 0) = 1$. Using (2) and (9), we write

$$
\begin{aligned}
\bar{x} &= \frac{1}{d} E^0 [\int_0^{T_1} ((1-U_0/2)X_0 1\{Y_0 = G\} + (1-V_0/2)X_0 1\{Y_0 = B\} + \alpha t)dt] \\
&= \frac{1}{d} E^0 [(1-U_0/2)X_0 1\{Y_0 = G\} D_0 + (1-V_0/2)X_0 1\{Y_0 = B\} D_0 + \frac{1}{2}\alpha D_0^2] \\
&= E^0[(1-U_0/2)X_0 1\{Y_0 = G\}] + E^0[(1-V_0/2)X_0 1\{Y_0 = B\}] + \frac{1}{2}\alpha \frac{d^{(2)}}{d} \\
&= \gamma_G x_G + \gamma_B x_B + \frac{1}{2}\alpha \frac{d^{(2)}}{d}
\end{aligned}
$$

◇

3.2 The reference throughput

To study the effect of burstiness, we change in what follows the parameters of the Markov chain (b and g) while keeping the average loss rate unchanged. The throughput in the bursty case is then compared to the throughput when the path is subjected to a non-bursty loss process having the same average loss rate. We denote this latter throughput $\bar{x}_r$ and we use it as a reference to evaluate the impact of burstiness. A non-bursty loss process is obtained when we have the same loss probabilities in the two states. We call this probability p. To get the same average loss rate as in the bursty case, p must be equal to,

$$p = dR = p_G \pi_G + p_B \pi_B.$$

Lemma 6 *On a non-bursty path, the source achieves a throughput of*

$$\bar{x}_r = \frac{2-p}{p}\alpha d + \frac{1}{2}\alpha \frac{d^{(2)}}{d} \tag{10}$$

Proof: This expression of $\bar{x}_r$ can be easily obtained by substituting in the expression of $\bar{x}$ (Theorem 5), γ_G and γ_B by their values as a function of p, the loss probability in the two states. We have,

$$\gamma = \gamma_G = \gamma_B = 1 - \frac{p}{2}.$$

The parameters of the Markov chain disappear and we get an expression of the reference throughput as a function of p and the distribution of potential losses. ◇

3.3 Comparison with previous works

Consider a particular case where $p = 1$. In this case, all potential losses cause a reduction in the transmission rate. This forms a loss process similar (even more general) to the one used in many previous works [14,21,23]. These works suppose that in the stationary regime, TCP rate varies in a cyclic manner between two values X and $2X$. They found that the time average transmission rate is about $3X/2$. Our model shows well that in the presence of a non-bursty path with $p = 1$, the expectation of the transmission rate just prior to a loss is equal to $2\alpha d$. The expectation of the transmission rate just after a real loss is simply αd. Thus, αd in our model corresponds to their X. However, our model does not give the same throughput they found. In our expression for the throughput, we see the appearance of the second moment of the time between losses $d^{(2)}$. To get their result, the second moment of the time between losses must be equal to the square of its average. This is only the case for a constant inter-loss time equal to d. Although they use a probabilistic loss model, their works transform the loss process into a deterministic one which results in the disappearance of the term $d^{(2)}$ in their analysis. The second moments of X_n and D_n are taken equal to the square of their average rates. The difference between these works and ours is that the packet loss probability is used for the calculation of these quantities. This deterministic evolution of the window can be seen as a normal result of the mutual independence that they assume between processes $\{X_n\}$ and $\{D_n\}$. Our model however, in addition to the consideration of the burstiness, proposes the exact expression of the throughput. It shows that the average time between losses as well as the second moment of this time must be considered otherwise the throughput will be underestimated. As an example, in the case of exponentially distributed inter-loss times, $d^{(2)}$ is equal to $2d^2$ and the throughput is simply equal to the average transmission rate just prior to losses $E[X_n]$.

4 A case study

In the sequel we consider the special case where,

$$p_G = 0, \gamma_G = 1, \qquad p_B = 1, \gamma_B = 1/2.$$

In other words we suppose that if the path is in the Bad state, each potential loss is transformed into a real loss, and if it is in the Good state no real losses occur. This model is sufficiently general to allow to vary both the average loss rate as well as the burstiness. Substituting in the expressions of x_G and x_B (Corollary 4), we get

$$x_B = 2\alpha d, \qquad x_G = \alpha d \frac{\bar{b} + \pi_G}{\bar{g}} \tag{11}$$

The throughput is given by,

$$\bar{x} = x_G + \frac{1}{2} x_B + \frac{1}{2} \alpha \frac{d^{(2)}}{d} \tag{12}$$

Remark : *It may seem remarkable that x_B does not depend on the transition probabilities of the Markov chain. This can easily be explained using the following argument. The mean time between losses is clearly $1/R = d/\pi_B$, so the mean increase in the transmission rate between two consecutive losses is*

$\alpha d/\pi_B$. Since we assume that we are in the stationary regime, the mean decrease in the transmission rate between losses should thus equal to the mean increase. But the mean decrease in the rate is half its mean value at loss. Thus,

$$E[X_n|Y_n = B] = 2\alpha d/\pi_B.$$

We conclude that indeed,

$$x_B = E[X_n 1\{Y_n = B\}] = E[X_n|Y_n = B]P(Y_n = B) = 2\alpha d.$$

4.1 The deviation of $\bar{x}$ from $\bar{x}_r$

The non-bursty path that has the same average loss rate is obtained when taking a loss probability p equal to π_B in the two states. The reference throughput in the non-bursty case is then,

$$\bar{x}_r = \frac{2\alpha}{R} - \alpha d + \frac{1}{2}\alpha \frac{d^{(2)}}{d}.$$

Given a certain average loss rate, we increase the burstiness by increasing b and g in such a way that their ratio remains unchanged. This guarantees that π_B and π_G, and therefore the average loss rate R, remain the same. To study the deviation of the throughput from the non-bursty case, we express $\bar{x}$ as a function of $\bar{x}_r$ and the parameters of the Markov chain. We get,

$$\bar{x} = \bar{x}_r + \alpha d \pi_G \left[\frac{1}{\bar{g}} - \frac{1}{\pi_B}. \right] \tag{13}$$

It is clear from this expression of $\bar{x}$, that the non-bursty case is obtained when $\bar{g} = b = \pi_B$. In our particular case, $\bar{g}$ is the probability that the next potential loss causes a real loss given that we are in the Good state. b is that the probability that it causes a loss given that we are in the Bad state. In the non-bursty case, these two probabilities must be equal. At the same time, they must be equal to π_B, the probability that the next potential loss causes a real loss independently of the current state.

4.2 Effect of the intensity and burstiness of losses

In this section, we study how the throughput varies as a function of R and the burstiness (via d, b and g). We shall show in particular that for a fixed loss rate R, the throughput increases when the burstiness increases. To facilitate the analysis, we suppose that the time between potential losses is exponentially distributed.

First, we study the effect of an increase in R on the performance. An increase in R can be caused by an increase in the number of potential losses per unit of time $(1/d)$ or by an increase in π_B. To study these two cases, we write $\bar{x}$ as,

$$\bar{x} = \alpha d \left[2 + \frac{\bar{b}}{\bar{g}} + \frac{\bar{b}}{\bar{g}(\bar{b}+\bar{g})}. \right] \quad (14)$$

It is clear that when d decreases, the throughput deteriorates. The increase in π_B can be caused by an increase in $\bar{g}$ or a decrease in $\bar{b}$. The two cases result also in throughput deterioration.

Suppose now that d is fixed as well as π_B and π_G. We increase b and g in order to increase the burstiness. The reference throughput remains constant given that it is only a function of the average loss rate. Equation (13) shows well that the average transmission rate improves when losses start to appear in bursts.

4.3 Impact of transmission rate limitation

Consider the case of TCP flow control. In the absence of losses on the link, the transmission rate increases until reaching a maximum value given by the window advertised by the receiver [24]. Once this window is reached, the transmission rate remains constant until the next loss occurs. Our model does not account for this limitation. It works well when losses are frequent so that the maximum window is rarely reached. We write first some conditions on the loss process to define the region where our previous model works properly. Then, we present a simple approximate calculation to account for this window limitation.

Suppose that the transmission rate is bounded by X_{max}. The point where the transmission rate is most likely to reach the maximal value corresponds to the Good state and it is just before the first potential loss in a Bad state. This is the first reduction in the transmission rate after getting out of a Good state. For our previous model to be correct, the expectation of the transmission rate at this point must be much smaller than the upper bound. This condition can be written as,

$$E[X_n|Y_n = B, Y_{n-1} = G] << X_{max}.$$

Taking into account that

$$\begin{aligned} E[X_n|Y_n = B, Y_{n-1} = G] &= E[X_n|Y_{n-1} = G] \\ &= E[X_{n-1}|Y_{n-1} = G] + \alpha d = x_G/\pi_G + \alpha d \end{aligned}$$

we get the following condition $x_G/\pi_G + \alpha d << X_{max}$. The larger the average loss rate and the lower the burstiness are, the more likely is that this condition holds. For a given loss rate, the increase in burstiness stretches the duration of the Good state and makes it more likely that the transmission rate reaches the upper bound.

Now, the closer $E[X_n|Y_n = B, Y_{n-1} = G]$ is to X_{max}, the more important is the impact of the receiver window. At the beginning, the receiver window starts to impact the transmission rate evolution only during the Good state of the path. We can assume that during the Bad state, the transmission rate still have the same evolution as that predicted by our previous model. The receiver window starts to impact the two states once the expectation of the transmission rate just prior to losses in the Bad state becomes larger than X_{max}. This latter condition can be written as,

$$E[X_n|Y_n = B, Y_{n-1} = B] >> X_{max}, \text{ i.e. } \qquad \frac{1}{2}\frac{x_B}{\pi_B} + \alpha d >> X_{max}.$$

Once the upper bound starts to impact a state, we make the assumption that during that state the transmission rate always reaches its maximal value. This is the kind of assumption made in [23]. Using the above two conditions, we separate first the space into three regions. In the first region, the transmission rate is not affected by X_{max}. In the second region, the Good state is affected. In the third region, both states are affected. We use then the above assumption to calculate the throughput of the transfer during each state of the path. Let

$$\bar{x}_G = E[X(t)|Y(t) = G], \qquad \bar{x}_B = E[X(t)|Y(t) = B]$$

where the expectation is with respect to the stationary probability. Thus, the throughput is simply equal to,

$$\bar{x} = \pi_G \bar{x}_G + \pi_B \bar{x}_B. \tag{15}$$

Let us define the three regions and calculate $\bar{x}_G$ and $\bar{x}_B$ for each of them:

$\mathbf{E[X_n|Y_n = B, Y_{n-1} = G] < X_{max}}$**:** The transmission rate limitation in this case has no influence and the throughput given by equation (12) can be considered.

$\mathbf{E[X_n|Y_n = B, Y_{n-1} = G] > X_{max}}$ **but** $\mathbf{E[X_n|Y_n = B, Y_{n-1} = B] < X_{max}}$**:** During the Bad state, the transmission rate limitation has no impact and $\bar{x}_B$ can be simply approximated by taking $p = 1$ in equation (10). This is the throughput obtained when the transmission rate is reduced at every potential loss, which is the case for the Bad state. Thus,

$$\bar{x}_B = \alpha d + \alpha \frac{d^{(2)}}{d}.$$

During the Good state, however, another throughput is to be considered. In average, the transmission rate at the beginning of the Good state is equal to,

$$x_0 = E[X_n|Y_n = G, Y_{n-1} = B] = \frac{1}{2}\frac{x_B}{\pi_B} + \alpha d.$$

The average duration of the Good state is $d/\bar{g}$. Using our assumption that the transmission rate during the Good state always reaches X_{max}, we consider that the transmission rate grows first from x_0 to X_{max}, then stays at X_{max} until the beginning of the Bad state. We can find the following expression for $\bar{x}_G$,

$$\begin{aligned}\bar{x}_G &= \frac{\bar{g}}{d}\left(\int_0^{(X_{max}-x_0)/\alpha}(x_0+\alpha t)dt + \int_{(X_{max}-x_0)/\alpha}^{d/\bar{g}} X_{max}dt\right)\\ &= \frac{\bar{g}}{d}\left(\frac{X_{max}^2 - X_0^2}{2\alpha} + X_{max}\left(\frac{d}{\bar{g}} - \frac{X_{max}-X_0}{\alpha}\right)\right).\end{aligned}$$

Given $\bar{x}_G$ and $\bar{x}_B$, the throughput can be calculated using equation (15).

$\mathbf{E[X_n|Y_n = B, Y_{n-1} = B] > X_{max}}$: In this case we assume that the transmission rate always reaches X_{max}. The transmission rate just before the occurrence of a real loss can be taken equal to X_{max}, that is, we can now take $x_0 = X_{max}$. Hence,

$$\begin{aligned}\bar{x}_G &= X_{max}\\ \bar{x}_B &= \frac{1}{d}\left(\int_0^{(X_{max}/2\alpha)}(X_{max/2}+\alpha t)dt + \int_{(X_{max})/2\alpha}^{d} X_{max}dt\right) = X_{max} - \frac{X_{max}^2}{8\alpha d}.\end{aligned}$$

And the total throughput is equal to

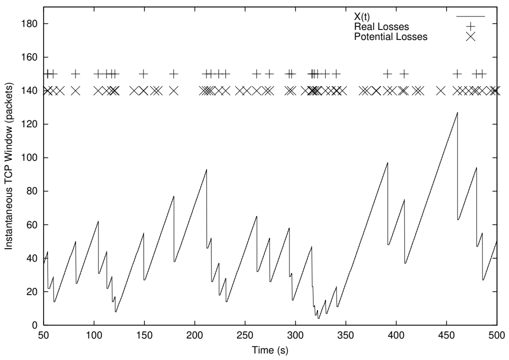
$$\bar{x} = \pi_G \bar{x}_G + \pi_B \bar{x}_B = X_{max} - \frac{X_{max}^2 \pi_B}{8\alpha d}.$$

The difference between our calculation here and the calculation in [23] is that we benefit from the use of a Markov chain, so we can introduce two refined conditions instead of one as in [23]. It was assumed in [23] that only when $E[X_n]$ exceeds X_{max}, the transmission rate limitation starts to impact the throughput.

4.4 Validation of the model

4.4.1 The simulation scenario

We validate our model using the TCP implementation in the `ns` simulator [22]. We consider long TCP transfers to eliminate the impact of the transient behavior at the beginning of the connection. Also, we use the SACK version [13] of TCP since it is able to recover from losses quickly and with a low probability of TimeOut and slow start. We suppose that the receiver acknowledges every data packet, so the window increases by one packet every RTT. We suppose also that the receiver window is very large so that it does not affect the transmission rate. Later, we will show that the estimations derived in Section 4.4



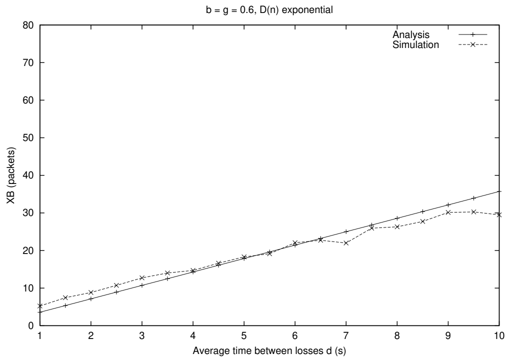


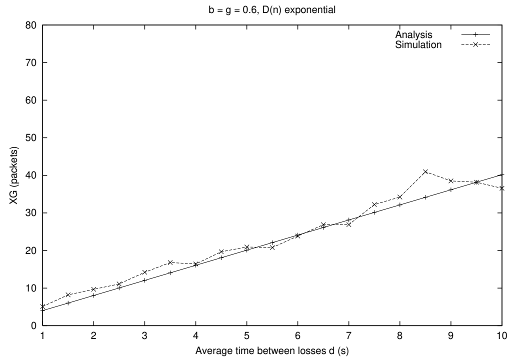
Fig. 2. The variation of $X(t)$ vs. time

Fig. 3. The variation of x_B vs. d

well agree with simulation results obtained in the case when the window evolution is limited by the receiver window. We consider the TCP window size in packets as the transmission rate in our mathematical model since this window varies linearly as a function of time and is divided by two upon loss detection. The different rates in our model are then expressed in terms of packets and need to be divided by the RTT in order to get the real rates.

The simulation scenario consists of a TCP connection crossing a 2 Mbps link. The RTT of the connection is taken equal to 560 ms. TCP packets are of total size 1000 bytes. we add our loss model to the simulator and we associate it to the 2 Mbps link. We account only for losses on the link and we study their impact on the throughput. We chose the parameters of the simulation in a way to not get losses in the other parts of the network. This clearly requires that losses are frequent so that the buffers in network routers do not overflow. The first condition of Section 4.3 is always satisfied with X_{max} equal to the network capacity. The purpose of the present paper is to create and to validate the model which takes into account the burstiness of losses. Currently, we are working on how to infer the parameters of our model from a real TCP trace [4].

The time between potential losses is taken to be exponentially distributed. Figure 2 shows a typical variation of the congestion window of the TCP connection. We see well how potential losses are transformed into real losses and how real losses cause a reduction of the window by a factor of two. In what follows, we run the connection for one hour and then we calculate the values of x_G, x_B and $\bar{x}$. These simulation results are then compared to those given by our analysis. When simulating, x_G (resp. x_B) is calculated by summing the window sizes when a potential loss occurs and the link in the Good state (resp. in the Bad state), then by dividing this sum by the total number of potential losses. $\bar{x}$ is calculated as the throughput of the connection over one hour expressed in terms of packets/s times the RTT. This gives the time average of the congestion window. First, we fix the parameters of the Markov chain of the link and we vary the time between potential losses. Second, we vary the burstiness while fixing the average rate of losses.



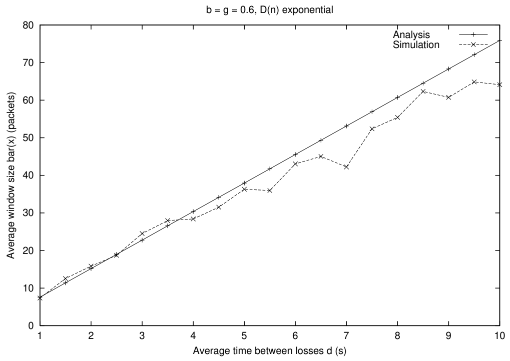


Fig. 4. The variation of x_G vs. d

Fig. 5. The variation of $\bar{x}$ vs. d

4.4.2 Impact of the average loss rate

In our first set of simulations, d is varied between 1 and 10 seconds. b and g are however taken equal to 0.6. Our analysis predicts a linear variation of the three quantities x_G, x_B and $\bar{x}$ (equations (11) and (14)). Figures 3, 4 and 5 show well the match between simulation and analytical results.

We shall give some details about the way the losses are generated and then explain the small deviations from the analytical results. We see that the slope of the line given by simulation is slightly smaller than the one given by analysis. The simulated model consists of individual packets that are sent in bursts on the link. The lossy link may not be carrying TCP packets when a potential loss has to be transformed into a real loss. At small d, losses are frequent and the window is most of the time of small size. When the window is very small and an event of loss is simulated, there might not be an actual packet on the link to which this loss corresponds. This results in many real losses considered by the analytical model but not considered by the simulation. Now, when d increases, the window becomes larger and the probability that the link is not carrying TCP packets when a potential loss occurs becomes smaller. Thus, the simulation line becomes closer to the analytical line.

To overcome the above problem, we simulate a loss as an event that causes the loss of all the packets that cross the lossy link during a certain time interval (100 ms in our simulations). By taking a large time interval to represent potential losses, we solve the problem of small windows. However, large windows see a large number of lost packets which causes sometimes a TimeOut and a slow start. For this reason, we see that the simulation results fall below the analytical ones at large d.

4.4.3 The impact of burstiness

We fix here the average time between potential losses to 5 seconds and we change the transition probabilities b and g while keeping $b = g$. This results in $\pi_G = \pi_B = 0.5$ which guarantees that the average loss rate remains con-

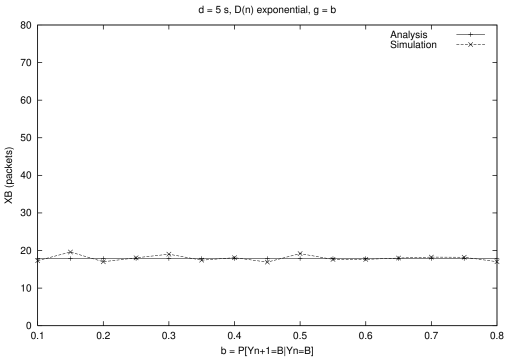


Fig. 6. The variation of x_B vs. b

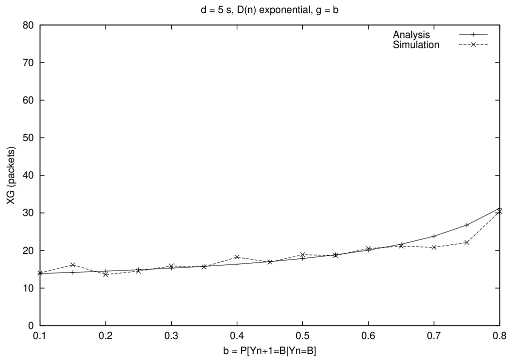


Fig. 7. The variation of x_G vs. b

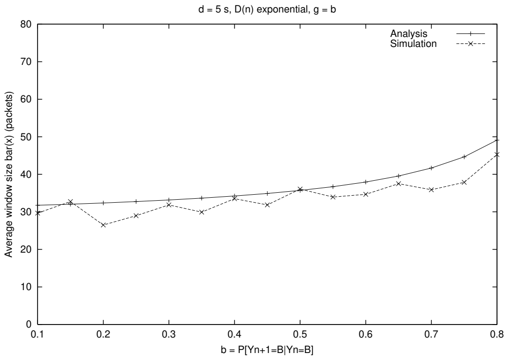


Fig. 8. The variation of $\bar{x}$ vs. b

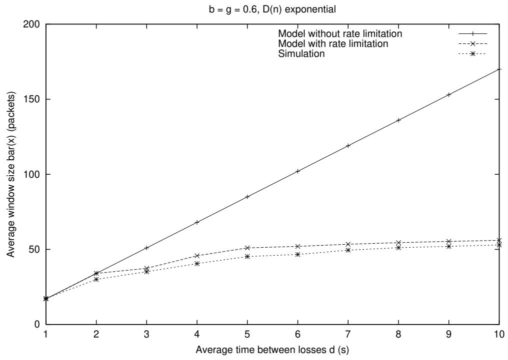


Fig. 9. $\bar{x}$ vs. d with a limitation on $X(t)$

stant. Our analysis shows that x_B must not change (equation (11)). x_G and $\bar{x}$ must however increase as a result of the increase in burstiness (equations (11) and (14)). Figures 6, 7 and 8 validate our analytical results. In particular, it is clear from Figure 8 that by increasing b from 0.1 to 0.8, the average throughput increases by around 60% even though the average loss rate remains unchanged. This confirms our result concerning the improvement in performance when losses become clustered.

4.4.4 Case of a limitation on the transmission rate

We now consider a case where the receiver window is set to a finite value so that it limits the evolution of the congestion window. We set b and g to 0.6 and we take an exponential time between potential losses of average 5 s. We reduce the RTT of the connection to 250 ms and we set the receiver window to the bandwidth- delay product. We change d from 1 to 10. By simple calculation, we see that in this setting, we cross the three regions we defined in Section 4.3 while introducing the limitation on the transmission rate for our model. Figure 9 shows how our approximation correctly estimates the real throughput. The model without limitation on the transmission rate leads to a clear overestimation of the real throughput in this scenario.

5 Conclusions

In this paper we evaluate the impact of burstiness of losses on the performance of a TCP-like flow control protocol. We define a model for losses using potential losses and a two-state Markov chain. We then calculate the throughput and the moments of the transmission rate at some potential loss instants. The throughput is compared to the one achieved when operating over a non-bursty path having the same average loss rate. Our main result is that for a given loss rate, the performance improves when losses tend to appear in bursts. We conduct a set of simulations with `ns` to validate the analytical results. A good match between simulation and analysis has been noticed.

Our current and future work on TCP modeling will be devoted to extend the Markov model of the path to more than two states [4] and to study non Markovian models as well [3]. We shall focus also on the identification of the parameters of the Markov model of the path [4] and try to compute the the exact expression of the throughput in case of limitation on the transmission rate (which renders the model for the transmission rate increase non-linear) [5]. We need also to extend our model to account for losses detected via a conservative TimeOut mechanism [3]. In case of these losses, the source stays idle for a certain time between the occurrence of the loss (i.e. the stop of the ACK clock) and the resumption of the transmission. Our assumption that the transmission rate resumes its linear increase directly after the loss will not be valid. These type of losses have been shown to be frequent on long delay connections [23]. Our experiments [3] have validated this conclusion. However, on short delay connections, we noticed that the TimeOut phenomenon is so rare. We believe that the introduction of TimeOut is worthy on connections where the window sizeăis of a small number of packets and where the TimeOut mechanism is quite is conservative.

Acknowledgement
The work of the second author was financed by a grant from CNET France-Telecom on flow control in High Speed Networks. The work of the third author was financed by an RNRT (French National Research Network in Telecommunications) "Constellations" project on satellite communications.

References

[1] E. Altman, F. Boccara, J. Bolot, P. Nain, P. Brown, D. Collange, and C. Fenzy "Analysis of TCP/IP flow control mechanism in hight- speed wide-area networks", *IEEE Conference on Decision and Control*, Dec 95.

[2] E. Altman, K. Avrachenkov, and C. Barakat, "TCP in presence of bursty losses", *ACM SIGMETRICS*, Jun 2000.

[3] E. Altman, K. Avrachenkov, and C. Barakat, "A stochastic model of TCP/IP with stationary random losses", *ACM SIGCOMM*, Sep 2000. A more detailed

version: INRIA Research Report RR-3824, Nov 1999.

[4] E. Altman, K. Avrachenkov, C. Barakat, and P. Dube, “TCP over a multi-state Markovian path”, working paper.

[5] E. Altman, K. Avrachenkov and C. Barakat, and R. Núñez-Queija, “State-dependent MG1 type queueing analysis for congestion control”, working paper.

[6] F. Anjum and L. Tassiulas, “On the behavior of different TCP algorithms over a wireless channel with correlated packet losses”, *ACM SIGMETRICS*, Mar 1999.

[7] F. Baccelli and P. Bremaud, “Elements of queueing theory: Palm-Martingale calculus and stochastic recurrences”, Springer-Verlag, 1994.

[8] A. Brandt, “The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients”, *Advances in Applied Probability*, Vol. 18, 1986.

[9] A. Brandt. P. Franken, and B. Lisek, “Stationary stochasitic models”, Wiley, 1990.

[10] H. Chaskar, T. V. Lakshman, and U. Madhow, “On the design of interfaces for TCP/IP over wireless”, *IEEE MILCOM*, 1996.

[11] A. Chockalingam, M. Zorzi, and R.R. Rao, “Performance of TCP on wireless fading links with memory”, *IEEE ICC*, Jun 1998.

[12] E.N. Gilbert, “Capacity of a burst-noise channel”, *Bell Systems Technical Journal*, Sep 1960.

[13] K. Fall and S. Floyd, “Simulation-based comparisons of Tahoe, Reno, and SACK TCP”, *ACM Computer Communication Review*, Jul 1996.

[14] S. Floyd, “Connections with multiple congested gateways in packet-switched networks part 1: one-way traffic”, *ACM Computer Communication Review*, Oct 1991.

[15] S. Floyd and K. Fall, "Promoting the use of end-to-end congestion control in the Internet", *IEEE/ACM Transactions in Networking*, Aug 1999.

[16] S. Floyd, M. Handley and J. Padhye, “Equation-based congestion control for unicast applications: the extended version”, *ACM SIGCOMM*, Sep 2000.

[17] P. Glasserman and D. D. Yao, “Stochastic vector difference equations with stationary coefficients”, *Journal of Applied Probability*, Vol. 32, 1995.

[18] G. Koole and S. Asmussen, “Marked point processes as limits of Markovian arrival streams”, *Journal of Applied Probability*, 1993.

[19] V. Jacobson, “Congestion avoidance and control”, *ACM SIGCOMM*, Aug 1988.

[20] A. Kumar, “Comparative performance analysis of versions of TCP in a local network with a lossy link”, *IEEE/ACM Transactions on Networking*, Aug 1998.

[21] 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.

[22] The LBNL Network Simulator, *ns*, http://www-nrg.ee.lbl.gov/ns.

[23] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, “Modeling TCP throughput: a simple model and its empirical validation”, *ACM SIGCOMM*, Sep 1998.

[24] W. Stevens, “TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery algorithms”, *RFC 2001*, Jan 1997.