Algorithms Seminar 2001–2002,
F. Chyzak (ed.), INRIA, (2003), pp. 69–72.

Available online at the URL
http://algo.inria.fr/seminars/.

Microscopic Behavior of TCP

Philippe Robert
INRIA Rocquencourt

February 11, 2002

Summary by Christine Fricker

Abstract

TCP (Transmission Control Protocol) is the ubiquitous data transfer protocol in communication networks. We focus on the control of the congestion of TCP. One long TCP connexion is studied when the loss rate of a packet tends to zero. It is shown that the Markov processes renormalized in a suitable way converge to a limit related to an auto-regressive Markov process. From a probalistic point of view, exponential functionals associated to compound Poisson processes play a key role. Analytically, the natural framework of this study turns out to be q-calculus. The talk presents a joint work with Vincent Dumas, Fabrice Guillemin and Bert Zwart (see [2] and [3]).

1. Introduction

In communication networks, sources send data to destinations via routers with limited capacity and TCP is a protocol which allows to transmit data with reliability in a loss network. The basic principles of TCP are due to Cerf and Kahn in 1973 and are based on acknowledgment: a source transmits at most W packets without response from destination. The control of congestion is due to Jacobson in 1987. Roughly speaking, if W packets are successfully transmitted, then the so-called congestion window size W is incremented by one; if a packet is lost, then W is divided by 2 (more generally multiplied by a factor δ). This is of course a simplification of the real algorithms involved, but the basic mechanism of reducing the congestion (called congestion avoidance) is captured by this model. Other algorithms (Slow Start, Fast Retransmit, and Fast Recovery) are also discussed (see [3]).

Consider the exchange between the source and the destination: each packet has some probability of being lost. The influence of the network is described in our model only through this loss process. We assume first that the packets are lost independently (a more general model where packets are lost by bursts will be considered in Section 4). Thus the sequence of the congestion window sizes is a Markov chain (W_n^α) on $\mathbb{N}$ with probability transitions

$$p\big(x, \min(x+1, w_{\max})\big) = e^{-\alpha x},$$
$$p\big(x, \lfloor \delta x \rfloor\big) = 1 - e^{-\alpha x},$$

where $w_{\max}$ is the maximum congestion size, $\delta \in (0,1)$ and $\alpha > 0$. The problems of special interest are estimations of the throughput (defined here as the mean congestion window size) and the stationary behavior. Asymptotic estimates will be presented when the loss rate α tends to zero.

Among other works, simulations are due to Floyd and Madhavi. Approximated models are investigated by Ott et al. [4] and Padhye et al., and analytical results are due to Adjih et al., Altman et al. [1], and Baccelli et al.

2. Convergence Results When the Loss Rate Tends to Zero

The main result of this section is that the congestion window size is of the order of $1/\sqrt{\alpha}$ when the loss rate or equivalently α tends to zero. For the sake of simplicity, assume that the maximum window size $w_{\max}$ is infinite.

Theorem 1. *If* $\lim_{\alpha\to 0}\sqrt{\alpha}\,W_0^\alpha = \overline{w}$ *and* $W^\alpha(t) = \sqrt{\alpha}\,W^\alpha_{\lfloor t/\sqrt{\alpha}\rfloor}$ *then* $\big(W^\alpha(t)\big)$ *converges in distribution to the Markov process* $\big(\overline{W}(t)\big)$ *given by* $\overline{W}(0)=\overline{w}$ *whose infinitesimal generator is*

$$\Omega(f)(x) = f'(x) + x\big(f(\delta x) - f(x)\big). \tag{1}$$

where f *is* $\mathcal{C}^1$ *on* $\mathbb{R}^+$.

A similar result is also valid for the embedded process (V_n^α) on $\mathbb{N}$ where V_n^α is the state of the Markov chain (W_n^α) just after the nth loss. It is clearly a Markov chain whose transitions are such that if $V_0^\alpha = x \geq 1$ then

$$V_1^\alpha = \big\lfloor \delta(x + G_x^\alpha)\big\rfloor$$

where $\mathbf{P}(G_x^\alpha \geq m) = \exp\big(-\alpha(mx + m(m-1)/2)\big)$. Indeed, $\sqrt{\alpha}$ is the right scaling for (G_x^α) and (V_n^α).

Proposition 1. *For* $x \in \mathbb{R}^+$, *as* α *tends to zero, the sequence* $\big(\sqrt{\alpha}\,G^\alpha_{\lfloor x/\sqrt{\alpha}\rfloor}\big)$ *converges in distribution to a random variable* $\overline{G}_x$ *with the property that for* $y \geq 0$,

$$\mathbf{P}\big(\overline{G}_x \geq y\big) = \exp^{-(xy+y^2/2)}. \tag{2}$$

Theorem 2. *If* $\lim_{\alpha\to 0}\sqrt{\alpha}\,V_0^\alpha = \overline{v}$ *then* $(\sqrt{\alpha}\,V_n^\alpha)$ *converges in distribution to the Markov chain* $\big(\overline{V}_n\big)$ *with* $\overline{V}_0 = \overline{v}$ *and transitions*

$$\overline{V}_{n+1} = \delta\big(\overline{V}_n + \overline{G}_{\overline{V}_n}\big).$$

3. The Equilibrium

Up to now, a closed form expression for the invariant probabilities of the Markov chains (W_n^α) and (V_n^α) is not known, but only bounds in some special cases. Nevertheless these invariant probability measures converge in distribution when α tends to 0 respectively to the distribution of $\overline{W}_\infty$, a random variable with distribution the invariant distribution of $\big(\overline{W}(t)\big)$ and of $\overline{V}_\infty$, a random variable with distribution the invariant distribution of $\big(\overline{V}_n\big)$. These limiting probabilities have rather simple closed form expressions. The key argument is the following result.

Lemma 1. *For* $x > 0$, *if* $\overline{G}_x$ *is defined by (2) then*

$$\big(x + \overline{G}_x\big)^2 \stackrel{dist.}{=} 2E_1 + x^2$$

where E_1 *is an exponentially distributed random variable wih parameter 1.*

It implies the important fact that the square of the limiting embedded Markov chain $\big(\overline{V}_n^2\big)$ is an AR (auto-regressive) process. By definition a process (X_n) is AR if and only if $X_{n+1} = A_nX_n + B_n$ where (A_n) and (B_n) are i.i.d. In Altman [1] the AR property is assumed, a priori, for the Markov chain $\big(\overline{V}_n\big)$ itself. The following result presents this property which leads to a close form expression for the distribution of $\overline{V}_\infty$ and its density function.

Proposition 2. *The sequence* $\left(\overline{V}_n^2\right)$ *is AR. More precisely for* $n \in \mathbb{N}$,

$$\overline{V}_{n+1}^2 = \delta^2\left(\overline{V}_n^2 + 2E_n\right)$$

where (E_n) *is an i.i.d. sequence of exponentially random variables with parameter 1. The distribution of* $\overline{V}_\infty$ *is thus given by*

$$\overline{V}_\infty \overset{dist.}{=} \sqrt{2\sum_{n=1}^{+\infty} \delta^{2n} E_n} \overset{dist.}{=} \sqrt{2\int_0^{+\infty} \delta^{2N(s)}\, ds}$$

where N *is a Poisson process with parameter 1. The density function* h_δ *of* $\overline{V}_\infty$ *is given by*

$$h_\delta(x) = \frac{1}{\prod_{n=1}^{+\infty}(1-\delta^{2n})} \sum_{n=1}^{+\infty} \frac{1}{\prod_{k=1}^{n-1}(1-\delta^{-2k})} \delta^{-2n} x e^{-\delta^{-2n}x^2/2} \qquad (x \geq 0).$$

The throughput of the TCP model is defined in the literature by $\rho^\alpha(\delta) = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^n W_k^\alpha$. The ergodic theorem for the Markov chain (W_n^α) gives that $\rho^\alpha(\delta) = \mathbf{E}(W_\infty^\alpha)$. Using the embedded chain $\left(\overline{V}_n\right)$ and defining the asymptotic throughput as $\overline{\rho}(\delta) = \lim_{\alpha\to 0} \sqrt{\alpha}\, \rho^\alpha(\delta)$, the following result can be deduced from Proposition 2.

Corollary 1. *The asymptotic throughput of the TCP model when* α *tends to 0 is given by*

$$\overline{\rho}(\delta) = \frac{\delta}{(1-\delta)\,\mathbf{E}(\overline{V}_\infty)} = \sqrt{\frac{2}{\pi}} \prod_{n=1}^{+\infty} \frac{1-\delta^{2n}}{1-\delta^{2n-1}}.$$

Remark. For the case of TCP, δ is set to 2 and the throughput $\overline{\rho}(1/2)$ is approximately 1.3098, which is the value observed in earlier simulations and experiments by Floyd, Padhye, and Madhavi.

4. **A More General Model**

A model with correlated losses generalizes the previous one. The evolution of W_n^α, the congestion window size over the nth RTT (Round Trip Time) interval, i.e., the maximum number of packets that can be transmitted without receiving acknowledgement from destination, is given as previously by the AIMD (additive increase, multiplicative decrease) algorithm: $W_{n+1}^\alpha = W_n^\alpha + 1$ when none of the W_n^α packets is lost and $W_{n+1}^\alpha = \max\left(\lfloor W_n^\alpha \rfloor, 1\right)$ otherwise. Nevertheless packet losses occur by clumps: if a packet is lost then several packets are lost during the following RTT intervals. These "clumps" are i.i.d. (see [3] for a complete definition). In particular if X_n is the number of losses in the nth clump then (X_n) is i.i.d. Though (W_n^α) is not a Markov chain, the embedded chain at the end of the consecutive clumps (V_n^α) is still Markov. Thus convergence results of Section 2 when the loss rate α tends to zero are valid with the infinitesimal generator in Theorem 1

$$\Omega(f)(x) = f'(x) + x\int_{\mathbb{R}^+} \left(f(\delta^u x) - f(x)\right) \mathbf{P}_{X_1}(du)$$

where the distribution of X_1 is denoted by $\mathbf{P}_{X_1}$ and δ replaced by δ^{X_1} in Theorem 2. As to Section 3, Proposition 2 is replaced by the following.

Proposition 3.

$$\overline{V}_\infty^2 = \delta^{2X_1}\left(\overline{V}_\infty^2 + 2E_1\right)$$

where X_1, E_1, *and* $\overline{V}_\infty$ *are independent random variables,* E_1 *being a random variable with an exponential distribution with parameter 1.*

Let I be a solution to

$$I \stackrel{dist.}{=} \beta^{X_1} I + E_1$$

where $\beta = \delta^2$ and E_1, I, and X_1 are independent. By definition, it turns out that I is the exponential functional associated to the Lévy process $Y(t) = \log \frac{1}{\beta} \sum_{k=1}^{N(t)} X_k$, N being a Poisson process with parameter 1. This functional occurs in mathematical finance (Asian options) where the Lévy process is generally a brownian motion with drift. In this setting, Bertoin, Carmona, Monthus, Petit, Yor, and many others (see for example Yor [5] for a survey) proved that the density of I is the solution of an integro-differential equation and that the moments of I are known. We present here an expression of the distribution of I for some special cases ($X_1 = 1$, X_1 with exponential distribution, and X_1 having a rational generating function). The Laplace transform of I can be expressed as a q-hypergeometric function (see [3] for details). The following proposition gives its fractional moments, in particular $E(\sqrt{I})$.

Proposition 4. *For each real s, if $-s$ is not in $\mathbb{N}^*$, $\mathbf{E}\big(\beta^{(s+1)X_1}\big) < \infty$ and $\mathbf{E}\big(\frac{1}{1-\beta^{X_1}}\big) < \infty$ then*

$$\mathbf{E}(I^s) = \Gamma(s+1) \prod_{k=1}^{+\infty} \frac{1 - \mathbf{E}\big(\beta^{(s+k)X_1}\big)}{1 - \mathbf{E}\big(\beta^{kX_1}\big)}.$$

As a sketch of the proof, to obtain the fractional moments, if $\psi(\lambda) = \mathbf{E}(e^{-\lambda} I)$ then, from the definition of I, we derive

$$\psi(\lambda) = \frac{1}{1+\lambda} \mathbf{E}\big(\psi(\lambda \beta^{X_1})\big),$$

which gives a simple recurrence relation on the Mellin transform $\psi^*(s) = \int_0^{+\infty} \psi(\lambda)\lambda^{s-1}\, d\lambda$ of ψ. Then, using the fact that $\psi^*(s) = \mathbf{E}(I^{-s})\Gamma(s)$ for $\Re(s) > 0$, one proves the result.

As in the independent losses model, asymptotic throughput when α tends to zero can be derived.

Theorem 3. *The asymptotic throughput for the correlated model when α tends to zero is given by*

$$\overline{\rho}_{X_1}(\delta) = \lim_{\alpha \to 0} \sqrt{\alpha\, \mathbf{E}(X_1)}\, \rho^\alpha = \sqrt{\frac{2\,\mathbf{E}(X_1)}{\pi}} \prod_{n=1}^{+\infty} \frac{1 - \mathbf{E}\big(\delta^{2nX_1}\big)}{1 - \mathbf{E}\big(\delta^{(2n-1)X_1}\big)}.$$

To conclude it is possible to compare throughputs for different distributions of X_1. In particular, the throughput for the independent losses model is a lower bound for the throughput of a correlated losses model (see [3] for details).

Bibliography

[1] Altman (E.), Avrachenkov (K.), and Barakat (C.). – A stochastic model of TCP/IP with stationary random losses. In *Proceedings ACM-SIGCOMM'00*, vol. 4, pp. 231–242. – Stockholm, 2000.

[2] Dumas (Vincent), Guillemin (Fabrice), and Robert (Philippe). – A Markovian analysis of additive-increase multiplicative-decrease algorithms. *Advances in Applied Probability*, vol. 34, n° 1, 2002, pp. 85–111.

[3] Guillemin (Fabrice), Robert (Philippe), and Zwart (Bert). – AIMD algorithms and exponential functionals. – April 2002. Preprint.

[4] Ott (Teunis J.), Kemperman (J. H. B.), and Mathis (Matt). – Window size behavior in TCP/IP with constant loss probability. – Available online at `http://www.psc.edu/networking/papers/papers.html#teunis`, November 1996.

[5] Yor (Marc). – *Exponential functionals of Brownian motion and related processes.* – Springer-Verlag, Berlin, 2001, x+205p.