

Analysis of TCP Westwood+ in high speed networks

E. Altman, C. Barakat, S. Mascolo, N. Möller and J. Sun

Abstract

TCP Westwood+ is modelled and analyzed using stochastic recursive equations. It is shown that for links with Poisson losses and independent and varying delays, TCP Westwood converges to a stationary process with a finite average throughput. The resulting throughput is computed explicitly, and it is shown that it does not depend on the filtering coefficient α in the bandwidth filter of TCP Westwood+.

I. INTRODUCTION

The Van Jacobson TCP (Transmission Control Protocol) congestion control is the widely-used transport protocol of the Internet [10]. Since its first proposal, named Tahoe TCP, it evolved through Reno up to the New Reno TCP which is nowadays the most used TCP in the Internet. In recent years, several new proposals and implementations of TCP control algorithms have been developed, motivated by a growing heterogeneity of networks (such as ad-hoc and sensor networks, high speed long distance wireless networks, and in particular satellite links), for which the initial TCP versions (developed for wireline networks) was not adequate any more.

In this paper we study the performance of the Westwood+ TCP version [5], [9], [11] that revealed to be particularly useful in scenarios affected by losses due to unreliable links. We focus on a single connection traversing a bottleneck.

Westwood+ TCP is novel with respect to Westwood TCP because of a new, simpler and unbiased estimator of the available bandwidth. It behaves exactly as TCP New-Reno version in increasing its window when there are no packet losses. Once a loss occurs, the behavior is different: instead of employing the classic TCP by half window decrease, Westwood+ decreases the window size to a new value that exactly matches the bandwidth available at the time of congestion. In particular the window size is set equal to the available bandwidth times the smallest RTT it has been observed so far. The rationale of this choice is to keep full the "available pipe", where the available pipe is the available bandwidth times the minimum round trip time.

The main novelty of Westwood+ is to substitute the "blind" by half multiplicative decrease mechanism of classic TCP with an adaptive setting that takes into account the bandwidth that is available at time of congestion. This feature reveals to be particularly advantageous in environments affected by losses not only due to congestion but also to unreliable links such as in the case of wireless links. From the point of view of the implementation, it is worth noting that Westwood+ TCP requires modifications only at the sender side and is completely backward compatible. From a technical point of view, the issue of obtaining an end-to-end bandwidth estimation is the most complex one. This issue has been addressed and solved in the past [5], [9], [11] and is beyond the scope of this work. For sake of completeness we only sketch the main idea which consists of employing the stream of returning acknowledgment packets to obtain an estimate of the available bandwidth at the time of congestion. Thus, assuming this bandwidth estimate, this paper aims at investigating the performance of Westwood+.

In this paper, we wish to study the throughput of persistent TCP Westwood+ connections over long wireless links (such as satellite links). The latter are characterized by random losses which are due to the noise on the channel rather than due to congestion, and where the round trip time has large (random)

variability (e.g. due to link layer features such as the ARQ) which are again not directly related to congestion [8].

Our modeling approach are based on [2], which, unlike many other models, takes explicitly into account the delay variability on the TCP connection. This feature of the model is needed when considering Westwood+, since the window size after a loss event is set a function of the estimated bandwidth delay product.

II. TCP WESTWOOD+: BACKGROUND

The novelty of Westwood and Westwood+ TCP is to substitute the multiplicative window decrease behaviour of standard TCP with an adaptive setting aiming at exact matching of the bandwidth available along the TCP connection path. The bandwidth estimate is obtained by filtering and averaging the stream of returning ACK packets. In particular, when three DUPACKs are received, both the congestion window (cwnd) and the slow start threshold (sssthresh) are set equal to the estimated bandwidth (BWE) times the minimum measured round trip time (RTT_{min}); when a coarse timeout expires the sssthresh is set as before while the cwnd is set equal to one.

TCP Westwood differs with respect to TCP Westwood+ mainly for the way the available bandwidth is computed. In details, Westwood+ computes one sample of available bandwidth every round trip time [11] using all data acked in one round trip time, whereas Westwood [5] computes one sample every received ack. The latter way has been shown to provide "aliased" samples, that is, the available bandwidth is overestimated up to several orders of magnitude. Samples are then filtered using a low-pass filter. It has been shown that using different type of low-pass filters does not affect the performance of TCP in a significant way [7] so that currently Westwood+ implements a standard exponential filter such as the one used by TCP to average round trip time samples [11]. The pseudo code of the Westwood+ algorithm is reported below:

```

a) On ACK reception:
   cwnd is increased accordingly to the Reno algorithm;
   the end-to-end bandwidth estimate BWE is computed;
b) When 3 DUPACKs are received:
   sssthresh =max(2, (BWE* RTTmin) / seg_size);
   cwnd = sssthresh;
c) When coarse timeout expires:
   sssthresh = max(2, (BWE* RTTmin) / seg_size);
   cwnd = 1;

```

In other words, Westwood+ additively increases the cwnd as standard New Reno, when ACKs are received. On the other hand, when a congestion episode happens, Westwood employs an adaptive setting of cwnd and sssthresh so that it can be said that Westwood+ follows an Additive-Increase/Additive-Decrease paradigm instead of the standard multiplicative decrease one.

III. MODELING WESTWOOD+

A. Modeling with linear stochastic recursive equations

The TCP window W_n of Westwood+ at the beginning of the n th RTT evolves according to:

$$W_{n+1} = \begin{cases} W_n + \beta & \text{if no loss occurred during } n\text{th RTT} \\ B_n \cdot RTT_{\min} & \text{if there is at least one loss during } n\text{th RTT} \end{cases} \quad (1)$$

where β is the additive increase factor, RTT_{\min} is the minimum value of RTT measured so far, and B_n is the bandwidth estimation obtained as the output of the following low pass filter [11, eq. 1]:

$$B_{n+1} = \begin{cases} \alpha B_n + \bar{\alpha} \frac{W_n}{RTT_n} & \text{if no loss occurred during } n\text{th RTT} \\ B_n & \text{if there is at least one loss during } n\text{th RTT} \end{cases} \quad (2)$$

where $\bar{\alpha} := 1 - \alpha$, and where α is a constant set equal to 0.9 [11, p. 27]. RTT_n is the duration of the n th RTT.

Remark 3.1: An alternative filter is used in [9, eq. 3]. In both cases the update of the estimated value is done every RTT, where as in the original Westwood version, the filter used to be updated upon arrivals of ACKs. This caused problems in the case of ACK compressions which made the Westwood TCP too aggressive and motivated the change to the Westwood+ version of the filter, see details in [9].

Let Z_n be the indicator that equals 1 if there has been at least one loss during the n th RTT and is otherwise 0. Combining (1) and (2) we obtain the vector recursive equation:

$$\begin{pmatrix} W_{n+1} \\ B_{n+1} \end{pmatrix} = A_n^w \begin{pmatrix} W_n \\ B_n \end{pmatrix} + C_n^w \text{ where } A_n^w = \begin{pmatrix} \bar{Z}_n & RTT_{\min} Z_n \\ \frac{\bar{\alpha} \bar{Z}_n}{RTT_n} & \alpha \bar{Z}_n + Z_n \end{pmatrix}, \quad C_n^w = \begin{pmatrix} \beta \bar{Z}_n \\ 0 \end{pmatrix}. \quad (3)$$

In order to work with the same units we denote $X_n = W_n / RTT_{\min}$. Then (3) becomes

$$\begin{pmatrix} X_{n+1} \\ B_{n+1} \end{pmatrix} = A_n \begin{pmatrix} X_n \\ B_n \end{pmatrix} + C_n \text{ where } A_n = \begin{pmatrix} \bar{Z}_n, & Z_n \\ \bar{\alpha} \bar{Z}_n \frac{RTT_{\min}}{RTT_n}, & \alpha \bar{Z}_n + Z_n \end{pmatrix}, \quad C_n = \begin{pmatrix} \frac{\beta \bar{Z}_n}{RTT_{\min}} \\ 0 \end{pmatrix}. \quad (4)$$

Eq. (4) is a linear SRE (Stochastic Recursive Equation), see e.g. [4], [6], which frequently arises in the analysis of TCP (see e.g. [1], [3], [13], [14]). Moreover, all elements of A_n and C_n are nonnegative. Matrices that have these properties frequently arise in modeling TCP, see [14].

Denote

$$Y_n = \begin{pmatrix} X_n \\ B_n \end{pmatrix}.$$

IV. STEADY STATE BEHAVIOR

We provide in this Section some general characteristics of the Y_n process. In future sections we shall use these to compute the throughput.

A. Solving iteratively recursion (4)

Iterating equation (4), we obtain:

$$\begin{aligned} Y_n &= A_{n-1} Y_{n-1} + C_{n-1} = A_{n-1} A_{n-2} Y_{n-2} + A_{n-1} C_{n-2} + C_{n-1} \\ &= \dots = \sum_{j=0}^{n-1} \left(\prod_{i=n-j}^{n-1} A_i \right) C_{n-j-1} + \left(\prod_{i=0}^{n-1} A_i \right) Y_0 \end{aligned} \quad (5)$$

where for $n > j$, we use the convention $\prod_{i=j}^{n-1} A_i = A_{n-1} A_{n-2} \dots A_j$ and $\prod_{i=n}^j A_i = I$, where I is the identity matrix.

Consider a probability space generated by the sequence $\{(A_n, C_n)\}$. The process Y_n is defined on this probability space simultaneously for all initial conditions Y_0 using the recursion (4). Under suitable conditions,

$$Y_n^* := \sum_{j=0}^{\infty} \left(\prod_{i=n-j}^{n-1} A_i \right) C_{n-j-1} \quad (6)$$

is well defined, it is the unique solution of (4) and is stationary ergodic. Moreover $|Y_n - Y_n^*| \rightarrow 0$ a.s. for all Y_0 on the same probability space as $\{(A_n, C_n)\}$.

B. The eigenvalues of A_n

For each n , one or the other of the off-diagonal elements of A_n is zero, so that A_n is either upper or lower triangular. Then the eigenvalues of A_n are the elements on the diagonal, \bar{Z}_n and $\alpha\bar{Z}_n + Z_n$.

A standard sufficient condition for the convergence of Y_n to a finite stationary regime Y_n^* is that A_n is a contracting matrix. But A_n is not contracting, since for each n one of its eigenvalues equals 1. Yet, one can show the convergence of Y_n to Y_n^* as above using arguments as in [3, p. 8].

V. THROUGHPUT ANALYSIS UNDER INDEPENDENCE CONDITIONS

We shall assume throughout this section the following assumption (for which we provide later sufficient conditions).

Assumption A1: A_n and Y_n are independent and $\{Y_n\}$, defined as the column vector $Y_n = (X_n, B_n)^T$, converges to a stationary ergodic process $Y_n^* = (X_n^*, B_n^*)^T$.

A. First moments of the limit process

Taking expectations in Eq. (4) yields the following

$$E[Y_{n+1}] = E[A_n]E[Y_n] + E[C_n], \quad \text{implying } E[Y_0^*] = E[A_0]E[Y_0^*] + E[C_0].$$

Hence we get at steady state the following expression for the first moments (provided that $I - E[A_0]$ is invertible):

$$E[Y_0^*] = (I - E[A_0])^{-1}E[C_0] \quad (7)$$

B. Throughput

In eq. (7) we obtained the first moments of Y^* at special points in time: those at round-trip time boundaries (this is known as the expectation with respect to the Palm measure). To compute the actual throughput we shall use the following formula:

$$\text{Thp} = \frac{E[S_0]}{E[RTT_0]}.$$

S_0 is the number of packets transmitted during RTT_0 , which is clearly equal to $W_0 = X_0 \cdot RTT_{\min}$ (we suppose that TCP implements the Nagle algorithm [12] and does not transmit partially filled packets).

Due to the independence assumption A1, RTT_0 is independent of X_0 . If we denote by e_1 the row vector $(1, 0)$, the TCP throughput becomes equal to,

$$\text{Thp} = \frac{E[X_0^*]RTT_{\min}}{E[RTT_0]} = \frac{e_1 E[Y_0^*]RTT_{\min}}{E[RTT_0]} = e_1 (I - E[A_0])^{-1} E[C_0] \frac{RTT_{\min}}{E[RTT_0]} \quad (8)$$

Let's find the explicit expression of the throughput for the following particular loss process.

Assumption A2: Losses occur according to a Poisson process with intensity ν . The RTT_n sequence is i.i.d., independent of this loss process.

Remark: Assumption **A2** implies that the number of losses that occur during RTT_n only depends on RTT_n and not on the number of losses during RTT_k for $k \neq n$. This implies that X_n is independent of Z_n , of \bar{Z}_n and of RTT_n . Thus, the fact that R_n are i.i.d. implies that X_n is independent of A_n . The same arguments also hold for showing that B_n is independent of A_n . We conclude therefore that Assumption **A2** implies Assumption **A1**.

Assumption **A2** is natural in the context of long wireless links, such as satellite communications. In such links, RTT_n may be quite variable due to link layer retransmissions (ARQ). The loss process itself can be caused by external factors as transmission errors caused by noise, equipment failures, etc. In view of **A2**, the conditional probability of at least one loss during RTT_n is $1 - \exp(-\nu RTT_n)$. Then the (unconditional) probability of no loss event during RTT_n is given by $\bar{p} := E[1 - \exp(-\nu RTT_n)]$. Also define $q := E[\exp(-\nu RTT_n)/RTT_n]$, $\bar{q} = 1 - q$ and $\bar{p} := 1 - p$.

Since TCP reduces its window once (or ideally should) for any number of losses during a round-trip time, we have in the view of assumption **A2**, $E[Z_n] = p$. Hence, $E[A_n]$ has the following form:

$$A := E[A_n] = \begin{pmatrix} \bar{p} & p \\ \bar{\alpha}qRTT_{\min} & \alpha + \bar{\alpha}p \end{pmatrix} \quad (9)$$

One can now form the matrix $(I - E[A_n])$, invert it and compute the throughput of TCP using (8). But before doing that, let's first check whether the inversion of the matrix $(I - E[A_n])$ is possible. This is equivalent to studying the stability conditions of our system. If for any particular parameter setting the inversion is not possible, this will mean that the protocol is not stable under this setting and that the throughput of TCP will infinitely grow (if the parameter setting remains the same).

Note that in reality one cannot have this explosion of the throughput. At some moment, the network starts to congest and limits the TCP throughput, which certainly leads to a change in the network setting otherwise the throughput would continue growing.

C. Stability and performance analysis

Consider the row sums of (9). On the first row, the sum is clearly 1. For the second row, we have $\bar{\alpha}(qRTT_{\min} + p) + \alpha$. This is at most 1, since $RTT_{\min} \leq RTT_n$ and

$$p + qRTT_{\min} = E[1 - \exp(-\nu RTT_n)(1 - RTT_{\min}/RTT_n)] \leq 1 \quad (10)$$

Hence, the eigenvalues of A are of magnitude at most 1.

Let's first consider the border cases, which correspond to an unstable system, and an A that is a stochastic matrix. If $\alpha = 1$, then the A matrix is stochastic for all round-trip time processes, and the system is unstable. If $\alpha < 1$, then A is stochastic iff we have equality in equation (10), which is equivalent to the condition that $RTT_n = RTT_{\min}$ a.s., or $E[RTT_k] = RTT_{\min}$.

On the other hand, if $\alpha < 1$ and $E[RTT_n] > RTT_{\min}$, then we have a strict inequality in (10), and it follows that A is a sub-stochastic matrix. Its eigenvalue a , which has the largest norm, is a real number smaller than 1. This implies that $I - A$ is invertible and hence an explicit expression for the throughput exists. When forming the inverse $(I - E[A])^{-1}$, it turns out that the first column is independent of α . Hence, $E[Y_0]$ and the average throughput is independent of the parameter α .

These observations can be summarized as follows.

Theorem: Under Assumption **A2**, Westwood+ TCP converges to a stationary process, with a finite average throughput, iff $\alpha < 1$ and $E[RTT_n] > RTT_{\min}$. If either condition is violated, the system is unstable and the throughput increases without limit. Furthermore, the average throughput does not depend on α .

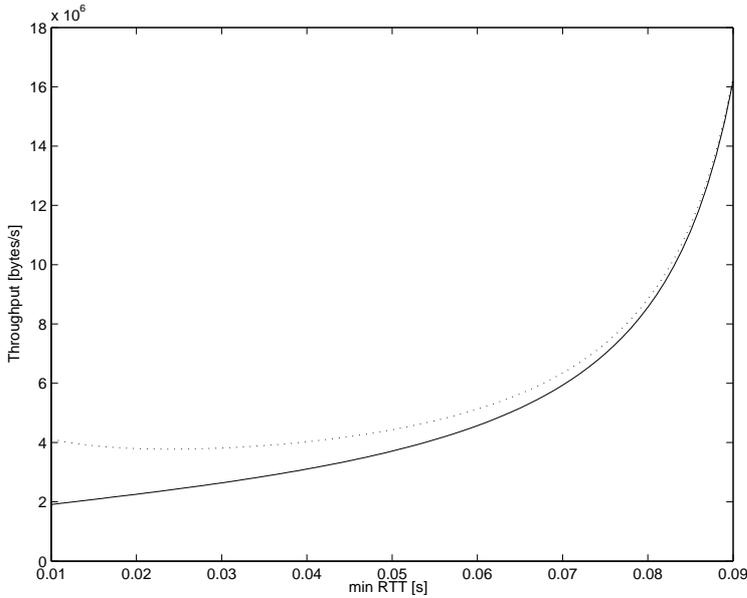


Fig. 1. Throughput for varying RTT_{\min} . In this scenario, the average RTT is 0.1 s, and it is divided into a constant component RTT_{\min} and a stochastic, exponentially distributed, component. The solid curve is the resulting average throughput, and the dotted curve is the average of the bandwidth estimate B .

In reality, the case $\alpha = 1$ corresponds to a bandwidth estimation that does not change with time. So the protocol will get a throughput function of the initial value of B_n . One needs a value of α strictly less than one to update the bandwidth estimation and to make the protocol throughput converge to a value independent of the initial value of B_n .

In summary, when TCP Westwood+ operates over a high speed link of constant round-trip time with random errors, it fully utilizes the available resources and drives the network into congestion. If the round-trip time varies for any reason, as for example retransmissions at the link-level or mobility, TCP Westwood+ can not saturate arbitrarily high capacity networks. One has to compute the expression of the throughput in (8), and if it found to be less than the available bandwidth, this will mean that TCP Westwood+ will not drive the network into congestion under this setting. One can always remove the congestion events by reducing the value to which TCP Westwood+ sets its window at the onset of a loss. This can be done for example by artificially lowering the minimum round-trip time of the connection.

VI. NUMERICAL EVALUATION

Assume, as in the previous section, that the RTT_n sequence is i.i.d., and that the loss process is Poisson with intensity ν , independent of the RTT_n process. For a concrete example, assume that $RTT_n = RTT_{\min} + d_n$, where d_n is exponentially distributed with average $1/\lambda$. Then we can compute

$$p = E[1 - \exp(-\nu(RTT_n))] = 1 - \frac{\lambda}{\nu + \lambda} \exp(-\nu RTT_{\min}) \quad (11)$$

$$q = E[\exp(-\nu RTT_n)/RTT_n] = \lambda \exp(\lambda RTT_{\min}) \int_{(\nu+\lambda)RTT_{\min}}^{\infty} \frac{\exp(-t)dt}{t} \quad (12)$$

A. Influence of RTT_{\min}

Consider the scenario $E[RTT_n] = 0.1$, with RTT_{\min} varying and λ chosen so that the average total delay is constant. Also set $\nu = 0.1$, $\beta = 1500$, and $\alpha = 0.9$. Figure 1 displays the throughput as RTT_{\min} varies.

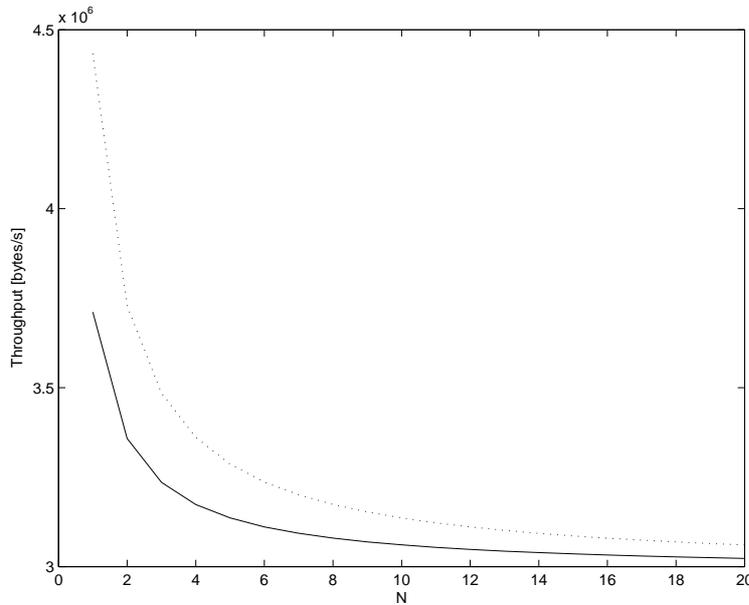


Fig. 2. Throughput for varying RTT variance. In this scenario, the average RTT is 0.1 s, and it is divided into a constant component $RTT_{\min} = 0.05$ and a stochastic $\Gamma(N, 20N)$ -distributed component. The standard deviation of the RTT is $0.05/\sqrt{N}$, decreasing to the right in the figure. The solid curve is the resulting average throughput, and the dotted curve is the average of the bandwidth estimate B .

B. Influence of RTT variance

To see how the variance of the RTT influences the throughput, we consider a scenario where both the minimum and the average RTT are fixed, but the variance varies. To do this, assume that d_n follows a $\Gamma(N, \lambda)$ -distribution, i.e., it is the sum of N exponentially distributed delays. Then $E[d_n] = N/\lambda$ and $\sigma^2(d_n) = N/\lambda^2$. For this distribution we get

$$p = 1 - \left(\frac{\lambda}{\nu + \lambda} \right)^N \exp(-\nu RTT_{\min}) \quad (13)$$

$$q = \lambda \exp(\lambda RTT_{\min}) \frac{1}{\Gamma(N)} \left(\frac{\lambda}{\nu + \lambda} \right)^{N-1} \int_{(\nu+\lambda)RTT_{\min}}^{\infty} \frac{\exp(-t)}{t} (t - (\nu + \lambda)RTT_{\min})^{N-1} dt \quad (14)$$

Now put $RTT_{\min} = 0.05$, $\lambda = 20N$, then for all N , the average total RTT is 0.1, and the variance decreases as N is increased. The resulting throughput is shown in Figure 2.

C. Parameter limits

In Figure 1, we see that the throughput tends to infinity as RTT_{\min} approaches $E[RTT_n]$. This confirms the theoretical results, since in this case the system becomes unstable.

On the other hand, in Figure 2, the throughput approaches a finite value when $N \rightarrow \infty$. In the limit, RTT_n approaches the constant value 0.1 s. However, RTT_{\min} is kept artificially fixed at a smaller value. This smaller RTT_{\min} is the reason the system remains stable.

D. Estimation bias

For both considered scenarios, one can also observe that the bandwidth estimate B_n exhibits a small bias (the difference between the solid and the dotted curves), and that this bias increases with the variance of RTT_n .

VII. CONCLUSIONS

This paper considers a model for TCP Westwood+ using the framework of stochastic recursive equations. In the case of independent delay and a Poisson loss process, it is shown that if the round-trip time is constant, throughput is unbounded. This means that Westwood+ TCP can achieve full utilization for arbitrarily high link capacities. On the other hand, if the round-trip time varies for any reason, as for example retransmissions at the link-level or mobility, TCP Westwood+ converges to a stationary process with finite average throughput. By computing the resulting throughput explicitly, it is shown that the throughput is independent of the α parameter in Westwood's the bandwidth filter.

The results are illustrated by numerical computation of the average throughput, for delays with exponential or gamma distribution.

REFERENCES

- [1] E. Altman, C. Barakat and K. Avratchenkov, "A stochastic Model of TCP/IP with Stationary Ergodic Random Losses", ACM-Sigcomm, Aug. 28 - Sept. 1, Stockholm, Sweden, 2000. See also INRIA Research Report RR-3824 . available at <http://www-sop.inria.fr/mistral/personnel/Eitan.Altman/fl-cont.html> at entry [2.11]
- [2] E. Altman, C. Barakat and V. Ramos-Ramos, "Analysis of AIMD protocols over paths with variable delay", Proceedings of IEEE Infocom, Hong-Kong, March 2004. see also a later journal version at *Computer Networks* Vol 48/6, pp 972-989, 2005,
- [3] F. Baccelli and D. Hong. AIMD, Fairness and Fractal Scaling of TCP Traffic. In *Proc. of IEEE Infocom 2002*, New York, NY, <http://www.di.ens.fr/~trec/aimd>, 2002.
- [4] A. Brandt, "The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients", *Advances in Applied Probability*, Vol. 18, 1986.
- [5] C. Casetti, M. Gerla, S. Mascolo, M. Sanadidi, R. Wang, "TCP Westwood: end-to-end bandwidth estimation for enhanced transport over wireless links", *ACM Wireless Networks*, (Special issue with Selected papers from Mobicom 2001), vol. 8, no. 5, pp.467-479, Sept. 2002. Kluwer Academic Publisher.
- [6] P. Glasserman and D. D. Yao, "Stochastic vector difference equations with stationary coefficients", *J. Appl. Prob.*, Vol 32, pp 851-866, 1995.
- [7] L.Grieco, S. Mascolo "End-to-end bandwidth estimation for congestion control in packet networks", 2nd international workshop on QoS in Multiservice IP Networks (QoS-IP 2003), Milan, Feb. 24-26, 2003.
- [8] M. Choon Chan and R. Ramjee, "TCP/IP performance over 3G wireless links with rate and delay variation", *MOBICOM*, Sept. 23-28, Atlanta, Georgia, USA, pp. 71-82, 2002.
- [9] S. Mascolo, L. A. Grieco, R. ferorelli, P. Camarda , G. Piscitelli, "Performance evaluation of Westwood+ TCP congestion control", *Performance Evaluation*, 55 (2004), pp. 93-111, Special Issue with Selected papers from Globecom 02, Elsevier, North-Holland, January 2004.
- [10] V. Jacobson, "Congestion avoidance and control", *ACM SIGCOMM*, August 1988.
- [11] L. Grieco, S. Mascolo, "Performance evaluation and comparison of Westwood+, New Reno, and Vegas TCP congestion control", *ACM SIGCOMM Computer Communication Review* archive Volume 34 , Issue 2, pp. 25-38, April, 2004.
- [12] J. Nagle, "Congestion control in IP/TCP internetworks", *RFC 896*, Jan. 1984.
- [13] R. N. Shorten, D. J. Leith, J. Foy and R. Kilduff, "Analysis and design of congestion control in synchronized communication networks", *Automatica*, 2004.
- [14] R. Shorten, F. Wirth and D. Leith, "A positive systems model of TCP-like congestion control: Asymptotic results", manuscript, 2005.
- [15] A. Berman, T. Laffey, A. Leizarowitz and R. Shorten, "On the second eigenvalues of matrices associated with TCP", manuscript, 2005.