On the Dynamics of TCP's Higher Moments

Robert N. Shorten, Christopher M. Kellett, and Douglas J. Leith

***Abstract*— In this paper we extend a recently derived positive systems model for TCP to capture the dynamics of TCP's higher order moments in drop-tail environments. Experimental results are given to illustrate the accuracy of our model.**

***Index Terms*— Additive-increase multiplicative-decrease (AIMD), network congestion control.**

I. Introduction

THE TCP transport protocol is used to carry the great majority of traffic in the Internet, including web, email, file transfers, music and video downloads *etc*. However, while the behavior of communication networks employing TCP has been the subject of much study, the dynamics of interacting/competing TCP flows remains relatively poorly understood, particularly in environments where drop-tail queueing is used. Recently, [7] presented a new positive systems model of the dynamics of interacting TCP flows in drop-tail environments. The analysis in [7] is confined to the first-moment (or mean value) of the flow peak congestion windows. In this paper we extend this analysis to encompass the dynamics of the second-moments (related to the variance) of the flow peak congestion windows.

Understanding the dynamics of the higher moments of the network congestion windows is important for a number of reasons. One fundamental reason is that the first moment provides no information regarding the fluctuations in the peak congestion window about its mean value. The fluctuations are, however, of considerable interest in unsynchronized network conditions i.e. where not every flow sees a packet loss at every congestion event, which is perhaps the norm in real networks. As has been observed by other authors, and as we shall see, in unsynchronized conditions the flow congestion windows can fluctuate in a complex manner, including exhibiting fractal behavior (e.g. see Fig. 1). Modelling of higher moments is essential for gaining a better understanding of such behavior. In particular, the *second moments* are known to be important for characterizing transient network unfairness and fluctuations over short time scales [1], [2], [6]. It is important to stress that interactions between flows as they compete for bandwidth play a fundamental role here and capturing the *dynamics* of interactions between flows is a prerequisite to understanding how networks behaves over short time scales.

II. A dynamic model of AIMD networks

Recently, in [7], a linear algebraic model of the AIMD algorithm was proposed to model a set of N AIMD sources, with multiplicative backoff factors $\beta_i \in [0,1)$, and additive increase factor $\alpha_i > 0$, that compete for bandwidth at a single bottleneck router. We denote the nonnegative integers by $\mathbb{Z}_{\geq 0}$ and let $w_i(k)$, $k \in \mathbb{Z}_{\geq 0}$ denote the value of the window size for the i^{th} flow just before the k^{th} network congestion event. It was shown that the evolution of each flow between successive congestion events can modeled as

$$w_i(k+1) = \beta_i(k) w_i(k) + \frac{\alpha_i}{\sum_\ell \alpha_\ell} \sum_\ell \left(1 - \beta_\ell(k)\right) w_\ell(k) \quad (1)$$

for all $k \in \mathbb{Z}_{\geq 0}$ where $\ell \in \{1, \ldots, N\}$ and $\beta_i(k)$ is a random variable that equals β_i when the i^{th} flow sees packet loss at the k^{th} network congestion event and is 1 otherwise.

The network dynamics can be conveniently captured in matrix form,

$$W(k+1) = A(k)W(k), A(k) \in \{A_1, .., A_m\}, \quad (2)$$

where $W(k) = [w_1(k)...w_N(k)]^T$ is the vector of flow congestion windows, $m = 2^N - 1$ and each A_i is a column stochastic matrix. Each matrix captures the various ways in which flows can be informed of congestion, and the probability[1] that A_i is invoked at the k^{th} event is denoted ρ_i. Note that this equation is a dynamic equation that models the evolution of the network in terms of an infinite product of random matrices.

This approach captures the evolution of the network by modeling the network state at congestion events. Geometrically, the state $W(k)$ evolves in a stochastic manner on a hyperplane that lies in the positive orthant in N dimensions. From a mathematical perspective, $W(k)$ is completely characterized by a probability density, the support of which lies on this hyperplane. An example of the support of such a density (the set of W values that can be reached asymptotically) is depicted in Figure 1. The complex nature of the support is evident, and it is also clear that the mean of $W(k)$ provides only limited information.

The fidelity of this model has been verified in extensive empirical studies (see [7]), and a number of mathematical properties of the model have been established in [8], [9]. For example, ergodicity (the existence of a limiting distribution) has been established under general conditions; in particular, under the assumption that the ρ_i are constant, or under the assumption that the ρ_i are functions of W (as is normal in networks with routers operating AQM's). It is also empirically shown in [7] that the assumption of constant ρ_i in networks employing drop-tail buffers appears to be valid. In this case, the network first moments converge to the Perron eigenvector of $\sum_{i=1}^m \rho_i A_i$ (see [7, Theorem 2]) for which an explicit

Manuscript received September 12, 2006. The associate editor coordinating the review of this letter and approving it for publication was Nikos Nikolaou.

R. N. Shorten and D. J. Leith are with The Hamilton Institute, NUI Maynooth, Ireland (email: robert.shorten@nuim.ie).

C. M. Kellett is with the School of Electrical Engineering and Computer Science, University of Newcastle, Australia. This work was done while he was with The Hamilton Institute.

Digital Object Identifier 10.1109/LCOMM.2007.061477.

[1]We note that the synchronization rate of flow j (the proportion of network congestion events at which flow j sees a packet loss), denoted λ_j, is related to the matrix probabilities ρ_i by $\lambda_j = \sum \rho_i$, where the summation is taken over all those matrices for which the j^{th} flow is notified of congestion.

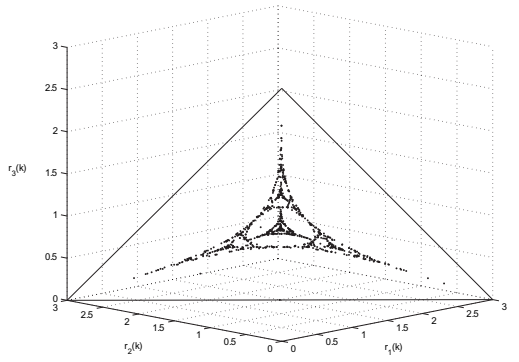

Fig. 1. Approximation of asymptotic image set for a network of three flows. $\alpha_i = 1, \beta_i = 0.25$ for all i.

expression is available in terms of the drop probabilities and the parameters α_i and β_i. Furthermore, a bound on the rate of convergence is given by the second largest eigenvalue of $\sum_{i=1}^{m} \rho_i A_i$.

III. A DYNAMIC MODEL OF TCP'S SECOND MOMENT

To the best of our knowledge, a general model that captures the dynamics of TCP's higher moments has not been developed in the literature. Such a model follows from the evolution of $W(k)$ described in the previous section. Specifically, the second moments are governed by

$$E\{W(k+1)W(k+1)^T\} = E\{A(k)W(k)W(k)^T A(k)^T\},$$

where $E\{\cdot\}$ denotes the expectation operator. Writing $Y(k) = Vec[W(k)W(k)^T]$, $Z(k) = E\{Y(k)\}$, and using the Kronecker product, we may write the above as $Z(k+1) = E\{Y(k+1)\} = E\{A(k) \bigotimes A(k) Y(k)\}$. Basic properties such as ergodicity of the random variable $Z(k+1)$ may be established as in [8]. Here, we proceed by assuming that the ρ_i are constant, as has been found to be the case in networks with drop-tail buffers. Then, we have that

$$\begin{aligned} Z(k+1) &= E\left\{A(k) \bigotimes A(k)\right\} Z(k) \\ &= \left(\sum_{i=1}^{m} \rho_i \left(A_i \bigotimes A_i\right)\right) Z(k) = \hat{M} Z(k). \end{aligned}$$

since we have assumed $A(k)$ to be independent of $W(k)$. We note that $\hat{M}$ is a positive, column stochastic matrix. Therefore, the Perron eigenvector describes the asymptotic value of the second moments of the network [5, Theorem 8.2.11], and the second largest eigenvalue of $\hat{M}$ describes the rate at which the moments converge. In the remainder of this section we give explicit formulae for this asymptote, and for the rate of convergence of the network second moment.

A. Convergence rate of second moment

To obtain an explicit expression for the second moment convergence rate we make use of the following linear algebraic result for Kronecker products of stochastic matrices of our form.

Theorem 1 ([3]): Consider the matrices $\hat{M} = \sum_{i=1}^{n} \rho_i \left(A_i \bigotimes A_i\right)$ and $M = \sum_{i=1}^{n} \rho_i A_i$. Then: (i) the eigenvalues of M are eigenvalues of $\hat{M}$; (ii) all eigenvalues of M and $\hat{M}$ that coincide, and which are different from 1, have multiplicity at least two; and (iii) the second eigenvalue of M matches the second eigenvalue of $\hat{M}$.

In our context, the above theorem states that the first and second moment of the network converge at exactly the same rate. Since an explicit formula for the second largest eigenvalue of M is given in [7], then this can be used to quantify the second eigenvalue of $\hat{M}$. More specifically, if we denote $E\{\beta_i\}$ as the expectation of the random variable $\beta_i(k)$, then the results given in [7] state that the second largest eigenvalue of M is bounded above by the largest $E\{\beta_i\}$, and below by the second largest $E\{\beta_i\}$, for $i \in \{1, ..., m\}$.

B. The asymptotic second moment

We now obtain an explicit formula for the Perron eigenvector of $\hat{M}$. To aid exposition we make the following definitions: $\mu_i(k)$ denotes the value of $E\{w_i(k)\}$ and $\mu^T = [\mu_1, \ldots, \mu_n]$; σ_{ij} denotes the asymptotic value of $E\{w_i(k)w_j(k)\}$, $\nu^T = [\sigma_{11}, \ldots, \sigma_{nn}]$ and ψ is the vector of remaining $E\{w_i(k)w_j(k)\}$, $i \neq j$, ordered $\psi^T = [\sigma_{12}, \ldots \sigma_{1n}, \sigma_{21}, \sigma_{23}, \ldots, \sigma_{n-1,n}]$. In the following we also assume, without any loss of generality, that $w_1(k) + \cdots + w_n(k) = 1$.

First, observe that $\mu_i(k) = E\{w_i(k)\} = E\{w_i(k)(w_1(k) + \cdots + w_n(k)\} \quad \forall k \in \mathbb{Z}_{\geq 0}$. Then,

$$\sigma_{ii} + \sum_{j=1,\, j\neq i}^{n} \sigma_{ij} = \mu_i. \tag{3}$$

Consequently, it follows that the Perron vector of $\hat{M}$ must satisfy $I\nu + G\psi = \mu$, where I is the $n \times n$ identity matrix, and G is an appropriately dimensioned matrix.

We now use the following fact. The identity

$$\begin{aligned} &(1 - E\{\beta_i(k)\beta_j(k)\})\sigma_{ij} = \\ &\quad \frac{\alpha_j}{2\alpha_i}(1 - E\{\beta_i^2(k)\})\sigma_{ii} + \frac{\alpha_i}{2\alpha_j}(1 - E\{\beta_j^2(k)\})\sigma_{jj} \end{aligned} \tag{4}$$

holds for all distinct pairs of network sources i and j, $i \neq j$.

Then, by considering all distinct pairs we obtain the matrix equation $\psi = F\nu$ where F is an appropriately dimensioned matrix. By combining these observations it follows that $(I + GF)\nu = \mu$. Note that all entries of these matrices can be easily calculated from the network parameters and from the ρ_i, and consequently, the network second moments ν can be explicitly computed.

Example 1: The case where all flows share the same AIMD parameters, drop probabilities, and round-trip times, has been studied in [1], where an explicit expression for the moments is given. We observe that the expression obtained from our model is significantly simpler than that in [1, Theorem 1].

Assume first that $\alpha_i = \alpha$, $\beta_i = \beta$ for all $i \in \{1, \ldots, N\}$ and that all flows are equally likely to experience a drop at

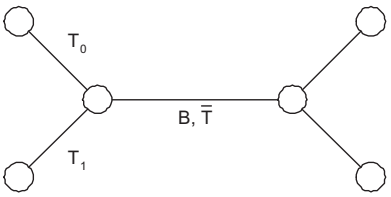

Fig. 2. Dumbbell topology.

congestion. Then the asymptotic value of the second moments are given by

$$\sigma_{ii} = \frac{\frac{1}{N}}{1 + (N-1)\frac{1-E\{\beta^2\}}{1-E\{\beta_i(k)\beta_j(k)\}}}, \quad \forall i \in \{1, \ldots, N\}.$$

This expression follows from the assumption that the window sizes sum to 1 (i.e., the link capacity is normalized to 1), that $\mu_i = \frac{1}{N}$ for all $i \in \{1, \ldots, N\}$ and that, with the assumption that all parameters are the same, it follows that

$$\frac{\sigma_{ij}}{\sigma_{ii}} = \frac{1 - E\{\beta_i^2(k)\}}{1 - E\{\beta_i(k)\beta_j(k)\}}$$

for all distinct sources i and j. Using these basic observations and equations (3) and (4), the result follows. ■

IV. Model Validation

The objective of this section is to present a brief empirical study to validate some of the predictions of the previous section.

We investigate the behavior of two TCP flows in the dumbbell topology shown in Figure 2. We use the standard AIMD parameters (i.e., $\alpha = 1$ and $\beta = \frac{1}{2}$) and round-trip times of 22ms and 122ms for flows 1 and 2 respectively. It is well known that networks of TCP flows with drop-tail queues can exhibit a wide range of deterministic drop behaviors [4]. However, most real networks carry at least a small amount of web traffic and our simulations incorporate a small level of this traffic to disrupt the coherent structure associated with phase effects and other complex phenomena previously observed in simulations of unsynchronized networks [4]. By performing repeated packet-level simulations with different random seed values for the web traffic generator, the ensemble average congestion window can be estimated.

A comparison of the predictions of the model (2) against the output of a packet-level NS simulation is depicted in Figure 3. Similarly to [7], the synchronization rates used in the model are measured from the simulations.

We make the following observations:

1) The predictions of [7, Theorem 2] for the ensemble average are in close agreement with the observed simulation results. The disparity in expected window size for the two flows arises from the difference in round-trip times.
2) The first and second moment converge at the same rate as predicted by Theorem 1.
3) The predictions of the model for the evolution of the second moment are in remarkable agreement with the observed simulation results.

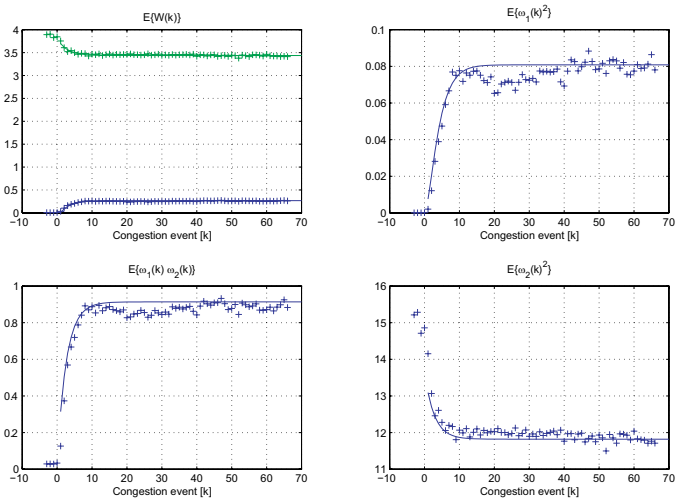

Fig. 3. First and second moments of window size by congestion epoch for dumbbell topology of Figure 2. Key: +*NS* simulation result (average over 200 runs); solid line - predictions of model. Network parameters: B=50Mb, q_{max}=50 packets, $\bar{T}$=20ms, T_0=102ms, T_1=2ms; approximately 0.5% bidirectional background web traffic.

V. Conclusions

We extend a recently derived positive systems model for networks of interacting AIMD flows to include the higher order moments of the flow peak congestion windows. A key point to note is that our model accurately captures the dynamic behavior of the higher moments. This is useful in evaluating the short-term behavior of networks in which TCP is deployed, e.g. evaluation of the short-term unfairness between network flows. Furthermore, we derived a simple, explicit expression for the asymptote of the second moments in terms of the network parameters for a homogeneous network. An explicit expression in the more general case of nonhomogeneous networks is an open problem.

Acknowledgements

This work was supported by SFI grants 04/IN3/I460 and 04/BR/M006. RS gratefully acknowledges discussions with Rade Stanojevic relating to material presented in this paper.

References

[1] F. Baccelli and D. Hong, "AIMD, fairness, and fractal scaling of TCP traffic," in *Proc. IEEE INFOCOM 2002*, pp. 229–238.
[2] F. Baccelli and D. Hong, "Interaction of TCP flows as billiards," *IEEE/ACM Trans. Networking*, vol. 13, no. 4, pp. 841–853, Aug. 2005.
[3] A. Berman, T. Laffey, A. Leizarowitz, and R. Shorten, "On the second eigenvalue of matrices associated with TCP," *Linear Algebra and its Applications*, vol. 416, pp. 175–183, 2006.
[4] S. Floyd and V. Jacobson, "On traffic phase effects in packet-switched gateways," *Internetworking: Research and Experience*, vol. 3, no. 3, pp. 115–156, Sept. 1992.
[5] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
[6] J. Lee, S. Bohacek, J. P. Hespanha, and K. Obrarzka, "Modeling communication networks with hybrid systems," accepted for publication in *IEEE Trans. Networking*, 2006.
[7] R. Shorten, F. Wirth, and D. Leith, "A positive systems model of TCP-like congestion control: asymptotic results," *IEEE/ACM Trans. Networking*, vol. 14, no. 3, pp. 616–629, June 2006.
[8] R. N. Shorten, C. King, F. Wirth, and D. Leith, "Modeling TCP congestion control dynamics in drop-tail environments," accepted for publication in *Automatica*, 2006.
[9] F. Wirth, R. Stanojevic, R. N. Shorten, and D. Leith, "Stochastic equilibria of AIMD communication networks," accepted for publication in *SIAM Matrix Analysis*, 2006.