Convergence results for synchronised communication networks

A. Berman and R. Shorten and D. Leith

***Abstract*— We present a simplified model of a network of TCP-like sources that compete for a shared bandwidth. We show that: (i) networks of communicating devices operating AIMD congestion control algorithms may be modelled as a positive linear system; (ii) that such networks possess a unique stationary point; and (iii) that this stationary point is stable. Using these results we establish conditions for the fair co-existence of traffic in networks employing heterogeneous AIMD algorithms, and present convergence rates to equilibria for such networks.**

Key Words : Second eigenvalue of a positive matrix; Network congestion control; Communication networks

I. Introduction

In this paper we present recently derived results concerning the dynamic properties of synchronised communication networks [1], [2]. We consider the problem of developing a systems theoretic framework that is suitable for the design and analysis of congestion control systems for communication networks. This problem has become topical in the context of internet congestion control [3], [4], [5], [6], [7], [8], but is also of relevance in a variety of problems where a number of devices compete for a shared resource. In this note we present preliminary results in this direction. More specifically, we present a theory for the design and analysis of synchronised networks that operate TCP-like congestion control algorithms and that operate drop-tail queues. We show that tools from the theory of positive linear systems may be employed to resolve convergence, stability, and performance issues in such networks. Finally, we present bounds on the rate of convergence to the network equilibrium.

This paper is structured as follows. We use positive linear systems theory to model TCP-based communications network in Section 2. In Section 3 we analyse the positive linear system obtained in Section 2 and present stability and convergence results. In Section 4 we present experimental results that confirm the analysis presented in Section 3. The implications of our results for the design of communication networks are discussed in Section 5.

This work was supported by Science Foundation Ireland grant 00/PI.1/C067 and by the Enterprise Ireland grant SC/2000/084/Y.

A. Berman is with the Department of Mathematics, The Technion, Haifa, Israel: berman@techunix.technion.ac.il

R. Shorten is with the Hamulton Institute, NUI, Maynooth, Ireland: robert.shorten@may.ie

D. Leith is with the Hamulton Institute, NUI, Maynooth, Ireland: doug.leith@may.ie

II. A model of TCP congestion control

We consider a network of n-sources competing for shared bandwidth. A communication network consists of a number of sources and sinks connected together via links and routers. We assume that these links can be modelled as a constant propagation delay together with a queue to buffer bursty traffic, and that all of the sources are operating a TCP-like congestion control algorithm.

TCP operates a window based congestion control algorithm. The TCP standard defines a variable $cwnd$ called the congestion window. Each source uses this variable to track the number of sent unacknowledged packets that can be in transit at any time, i.e. the number of packets in the 'pipe' formed by the links and buffers in a transmission path. When the window size is exhausted, the source must wait for an acknowledgement (ACK) before sending a new packet. Congestion control is achieved by dynamically adapting the window size according to an additive-increase multiplicative-decrease (AIMD) law. The basic idea is for a source to gently probe the network for spare capacity and rapidly back-off its send rate when congestion is detected. A typical window evolution is depicted in Figure 2 ($cwnd_i$ at the time of detecting congestion is denoted by w_i in this figure). Over the

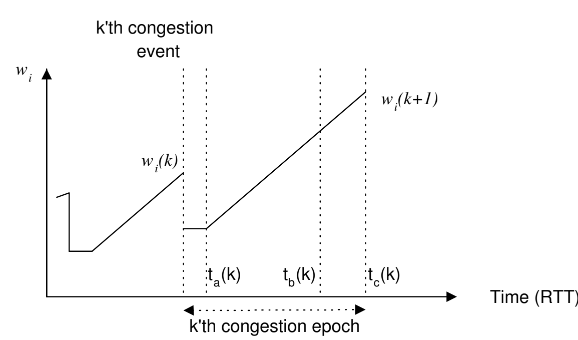

Fig. 1: Evolution of window size

kth congestion epoch three important events can be discerned: $t_a(k), t_b(k)$ and $t_c(k)$ in Figure 2. The time $t_a(k)$ is the time at which the number of unacknowledged packets in the pipe equals $\beta_i w_i(k)$; $t_b(k)$ is the time at which the pipe is full; and $t_c(k)$ is the time at which packet drop is detected by the sources. Note that we measure time in units of round-trip time (RTT)[1].

We consider a network of sources operating AIMD congestion control algorithms. Each source is parameterized by an additive increase parameter and a multiplicative decrease factor, denoted α_i and β_i respectively. These parameters satisfy $\alpha_i \geq 1$ and $0 < \beta_i < 1 \ \forall i \in \{1, ..., n\}$. We assume that the event times t_a, t_b and t_c indicated in Figure 2 are the same for every source, i.e. that the sources are synchronised.

Let $w_i(k)$ denote congestion window size of source i immediately before the kth network congestion event is detected by the sources; see Figure 2. It follows from the definition of the AIMD algorithm that the window evolution is completely defined over all time instants by knowledge of the $w_i(k)$ and the event times $t_a(k), t_b(k)$ and $t_c(k)$ of each congestion epoch. We therefore only need to investigate the behaviour of these quantities.

We have that $t_c(k) - t_b(k)$ = 1; namely, each source is informed of congestion exactly one RTT after the first dropped packet was transmitted. Also,

$$w_i(k) \geq 0, \sum_{i=1}^{n} w_i(k) = P + \sum_{i=1}^{n} \alpha_i, \ \forall k > 0, \qquad (1)$$

where P is the maximum number of packets which can be held in the 'pipe'; this is usually equal to $q_{max} + BT$ where q_{max} is the maximum queue length of the congested link, B is the service rate in packets per second and T is the round-trip time. At the $(k+1)$th congestion event

$$w_i(k+1) \quad = \quad \beta_i w_i(k) + \alpha_i [t_c(k) - t_a(k)]. \qquad (2)$$

and

$$\begin{aligned} t_c(k) - t_a(k) \quad &= \quad \frac{1}{\sum_{i=1}^{n} \alpha_i} [P - \sum_{i=1}^{n} \beta_i w_i(k)] \\ &+ \quad 1. \end{aligned} \qquad (3)$$

Hence, it follows that

$$\begin{aligned} w_i(k+1) \quad &= \quad \beta_i w_i(k) \\ &+ \quad \frac{\alpha_i}{\sum_{j=1}^{n} \alpha_i} [\sum_{i=1}^{n} (1-\beta_i) w_i(k)], \end{aligned} \qquad (4)$$

and that the dynamics of the entire network can be written in matrix form as

$$W(k+1) = AW(k), \qquad (5)$$

[1]RTT is the time taken between a source sending a packet and receiving the corresponding acknowledgement, assuming no packet drop

where $W^T(k) = [w_1(k), \cdots, w_n(k)]$, and

$$\begin{aligned} A \quad &= \quad \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} \\ &+ \quad \frac{1}{\sum_{j=1}^{n} \alpha_i} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} 1-\beta_1 \cdots 1-\beta_n \end{bmatrix}. \end{aligned} \qquad (6)$$

In the sequel it is convenient to write A in the form

$$A \quad = \quad I - Y + xy^T, \qquad (7)$$

where $Y = diag([1-\beta_1, ..., 1-\beta_n])$, $x^T = \frac{1}{\sum_{i=1}^{n} \alpha_i}[\alpha_1, ..., \alpha_n]$ and $y^T = [1-\beta_1, ..., 1-\beta_n]$, and where we assume that the entries of $I - Y$ have been ordered:

$$\begin{gathered} \beta_1 = \cdots = \beta_{k_1} = \gamma_1, \\ \beta_{k_1+1} = \cdots = \beta_{k_1+k_2} = \gamma_2, \\ \vdots \\ \beta_{k_1+k_2+....+k_{s-1}+1} = \cdots = \beta_{k_1+k_2+....+k_s} = \gamma_s, \end{gathered}$$

with $k_1 + k_2 + ... + k_s = n$ and $\gamma_1 < \gamma_2 < ... < \gamma_s$.

Comment 1: The matrix A is strictly positive and it follows that the synchronised network (5) is a positive linear system.

Comment 2: The vector x is a probability vector and the matrix A is a column stochastic matrix. Thus, the Perron eigenvalue $\rho(A) = 1$ and the all the eigenvectors of A, except the Perron vector, are orthogonal to $e^T = [1, ..., 1]$.

III. The Spectrum of the Network Matrix

We now present the main mathematical results of the paper. The results are derived in [2], [1] and establish basic properties of the communication networks under study in this paper.

Theorem 3.1: Let A be defined as in Equation (6). Then, a Perron eigenvector of A is given by $x_p^T = [\frac{\alpha_1}{1-\beta_1}, ..., \frac{\alpha_n}{1-\beta_n}]$.

The following corollary follows from Theorem 3.1 and properties of non-negative matrices [9], [10].

Corollary 3.1: For a network of synchronised time-invariant AIMD sources: (i) the network has a Perron eigenvector $x_p^T = [\frac{\alpha_1}{1-\beta_1}, ..., \frac{\alpha_n}{1-\beta_n}]$; and (ii) the Perron eigenvalue is $\rho(A) = 1$. It follows that all other eigenvalues of A satisfy $|\lambda_i(A)| < \rho(A)$. The network possesses a unique stationary point $W_{ss} = \Theta x_p$, where Θ is a positive constant such that the constraint (1) is satisfied; $\lim_{k \to \infty} W(k) = \Theta x_p$, and the rate of convergence of the network to W_{ss} depends upon the second largest eigenvalue of A ($max|\lambda|, \lambda \neq 1 \in spec(A)$).

It follows from the corollary that the second largest eigenvalue of the matrix A determines the convergence properties of the entire network. It is therefore important to determine this eigenvalue. Theorem 3.2, which is the main result of the paper and whose proof is given in [2], provides a characterisation of all the eigenvalues of the matrix A. It shows that all the eigenvalues of A are real and positive and lie in the interval $[\beta_1, 1]$. In particular, the second largest eigenvalue is bounded above by β_n.

Theorem 3.2: Consider the matrix (7). Then, the following statements are true.

(a) The matrix A is diagonally similar to a (real) positive diagonal matrix.
(b) Except for the Perron eigenvalue, all of the eigenvalues of A lie in the interval $[\beta_1, \beta_n]$.
(c) More specifically, if $k_i > 1$, then γ_i is an eigenvalue of A of multiplicity $k_i - 1$, and the remaining eigenvalues are simple, and with the exception of 1 lie in the intervals, (γ_1, γ_2), $(\gamma_2, \gamma_3), ..., (\gamma_{s-1}, \gamma_s)$.
(d) In particular, if all the β's are distinct, then $\beta_1 < \lambda_1 < \beta_2 < ...... < \beta_{n-1} < \lambda_{n-1} < \beta_n < \lambda_n = 1$.

IV. Experimental results

In this section we present a short example to illustrate our results. Consider a network with a simple dumb-bell topology having a bottleneck link of bandwidth 10Mbs, link propagation delay 50ms and buffer size of 40 packets (operating a drop-tail queueing policy). Consider two TCP sources operating AIMD congestion control strategies, source 1 having increase parameter $\alpha_1 = 1$ and back-off factor $\beta_1 = 0.5$ and source 2 having $\alpha_2 = 1.5$, $\beta_2 = 0.25$. Results from an ns-2 simulation are shown in Figure IV. Note that this is a detailed nonlinear packet-level simulation. Observe that $\alpha_1/(1-\beta_1) = \alpha_2/(1-\beta_2)$ and hence our analysis predicts that the equilibrium solution is fair (Perron eigenvector is $[1\ \ 1]^T$), in agreement with detailed simulation results. Note that fairness is achieved despite the sources having different AIMD parameters i.e. there exists an 'ecosystem' of AIMD strategies which co-exist fairly. Moreover, the second largest eigenvalue is bounded above by $\beta_1 = 0.5$. Thus, the rate of convergence to a fixed point is bounded by β_1^k where k is the congestion epoch and using this it follows that the 95% rise time is predicted to be $\log 0.05/\log \beta_1$, i.e. a rise time of 4.3 congestion epochs, again in excellent agreement with detailed simulation results.

V. Discussion and concluding remarks

In this paper we present a simple model of a network of sources competing for a shared bandwidth.

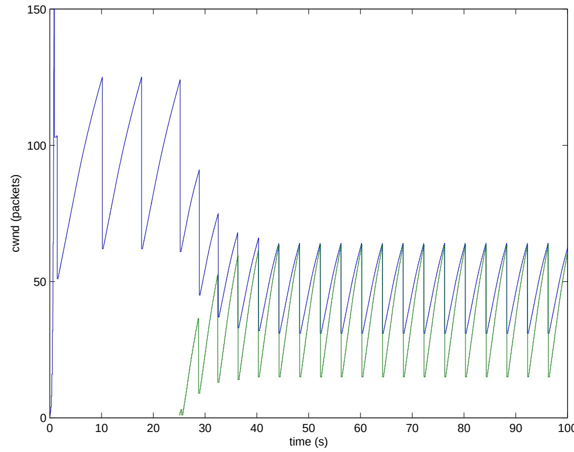

Fig. 2: ns-2 packet-level simulation ($\alpha_1 = 1$, $\beta_1 = 0.5$, $\alpha_2 = 1.5$, $\beta_2 = 0.25$, dumb-bell with 10Mbs bottleneck bandwidth, 100ms propagation delay, 40 packet queue).

In this paper we have analysed the properties of a certain type of communication network. Specifically, we analyse the dynamic behaviour of synchronised communication networks where each source operates an AIMD congestion control algorithm and where each of the sources share the same RTT. While these assumptions do not apply to general communication networks, they are valid for important network types; in particular, for long-distance high-speed networks [11], [12], [1]. Our model incorporates the hybrid nature of TCP algorithm, time-varying delays on links, and drop-tail queueing (all features of networks of TCP sources).

A basic problem in the design of these networks is to ensure fairness of the equilibrium condition, good throughput of data, and to ensure rapid convergence to the equilibrium condition in the presence of network disturbances. Our results show that: (i) fairness[2] is ensured when A is symmetric, i.e., by choosing the $\frac{\alpha_i}{1-\beta_i}$ to be constant for all sources in the network (Theorem 3.1); and (ii) the rate of convergence of the network to its equilibrium state is determined by β_n in the case when β_n is an eigenvalue of $I - Y$ with multiplicity greater than one, and is bounded above and below by β_n and β_{n-1} respectively when β_n has multiplicity 1 (Theorem 3.2). Importantly, our results indicate that good data throughput, which is normally achieved by ensuring that the β_i's are set to large values, cannot be achieved without adversely affecting the network convergence properties. In particular, knowledge of the eigenvalue locations of the matrix A (Theorem 3.2) , and the fact that the non-Perron eigenvectors of this matrix satisfy $e^T z = 0$

[2] When defined to be an equal share of the network 'pipe' for all sources.

(Comment 2), provides a basis for understanding the transient response of such networks, and may provide a network for identification of network parameters from measured data. In this context our results may even provide a basis for network operators to detect of malicious network attacks (users setting their value of β to large values). Finally, we note that our results are also likely to provide valuable insights into the design of adaptive congestion control algorithms and this is currently an active area of research. We believe that the results presented here represent a small, but nevertheless important step in the mathematical design of communication networks. Future work will involve extending our results to the case of non-synchronised networks.

ACKNOWLEDGEMENTS

This work was supported by Science Foundation Ireland grant 00/PI.1/C067. This work was also partially supported by the European Union funded research training network *Multi-Agent Control*, HPRN-CT-1999-00107[3] and by the Enterprise Ireland grant SC/2000/084/Y. Neither the European Union or Enterprise Ireland is responsible for any use of data appearing in this publication.

REFERENCES

[1] R. Shorten, D. Leith, J. Foy, and R. Kilduff, "Analysis and design of synchronised communication networks." Submitted to Automatica, 2003.

[2] A. Berman, R. Shorten, and D. Leith, "Positive matrices associated with synchronised communication networks." Submitted to Linear Algebra and its Applications, 2003.

[3] V. Jacobson, "Congestion avoidance and control," in *Proceedings of SIGCOMM*, 1988.

[4] S. Floyd and K. Fall, "Promoting the use of end-to-end congestion control in the internet," *IEEE/ACM Transactions on Networking*, vol. 7, no. 4, pp. 458–472, 1999.

[5] Various authors, "Special issue on TCP performance in future networking environments," *IEEE Commuications magazine*, vol. 39, no. 4, pp. –, 2001.

[6] S. Low, F. Paganini, and J. Doyle, "Internet congestion control," *IEEE Control Systems Magazine*, vol. 32, no. 1, pp. 28–43, 2002.

[7] Various authors, "Special issue on internet technology and convergence of communication services," *Proceedings of the IEEE*, vol. 90, no. 9, pp. –, 2002.

[8] J. Hespanha, S. Hohacek, K. Obrarzka, and J. Lee, "Hybrid model of TCP congestion control," in *Hybrid Systems: Computation and Control*, pp. 291–304, 2001.

[9] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[10] A. Berman and R. Plemmons, *Nonnegative matrices in the mathematical sciences*. SIAM, 1979.

[11] S. Floyd, "High speed TCP for large congestion windows," tech. rep., Internet draft draft-floyd-tcp-highspeed-02.txt, work in progres, February 2003.

[12] T. Kelly, "On engineering a stable and scalable TCP variant," tech. rep., Cambridge University Engineering Department Technical Report CUED/F-INFENG/TR.435, 2002.

[3]This work is the sole responsibility of the authors and does not reflect the European Union's opinion