2007 International Symposium on Nonlinear Theory and its Applications
NOLTA'07, Vancouver, Canada, September 16-19, 2007

Modelling of TCP/IP Congestion Control Mechanism with Queueing

Stefanos Politis and Paul Curran

School of Electrical, Electronic and Mechanical Engineering,
University College Dublin,
Belfield, Dublin 4, Ireland.
Email: stefanos.politis@ee.ucd.ie, paul.curran@ucd.ie

Abstract– In this paper, a model of a congestion control mechanism with queueing for a network running TCP is presented. The cases of one source and one link and two sources and one link are analyzed and the steady state behaviour is examined.

Key Words: Congestion Control, TCP, Nonlinear Systems.

1. Introduction

The Transmission Control Protocol (TCP), offers a reliable data transfer over the Internet. TCP provides an Additive Increase/Multiplicative Decrease (AIMD) mechanism according to which the source's window increases (exponentially at slow-start and linearly at congestion avoidance) until packets are discarded and congestion is detected. Then, the source reduces the window value [1,2,3].

A number of models for congestion control have been introduced. Kelly et al. [4] proposed a framework comprising a primal algorithm for TCP rate control and a dual algorithm for the AQM scheme. Hespanha [5] presented a new stochastic hybrid network model, where transitions between discrete modes are triggered by stochastic events. Recently, Shorten et al. [6] modelled communication networks drop-tail queueing and AIMD congestion control algorithms by employing the theory of nonnegative matrices.

In this paper, we present a preliminary analysis of a discrete time model of a congestion control mechanism with queueing for a network running TCP. We examine the cases of one source and one link and two sources and one link models. We analyze the network and examine the steady state behaviour.

2. One Source / One Link

The network consists of a source and sink connected via a link (fig.1). The source transmits data at the maximum allowed rate at all times. Assume the queue is initially empty. The detection of a congestion event by the source is via a timeout, after which the source's window is set to one. The link is managed by a Drop-Tail buffer, i.e. an arriving packet is discarded if the queue is full. The link is characterised by a parameter $N-1$, which represents the total number of packets that can be transmitted per unit of time without congestion occurring.

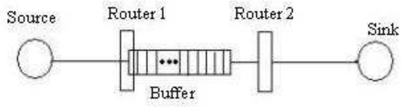

Figure 1: Network with 1 Source and 1 Link.

Let the parameter N be given by the general equation:

$$N = P + (Q_{MAX} - Q_k) + 1 \tag{1}$$

where P represents the number of packets which the link can process per time unit. The time unit corresponds to the time needed for the source to send a window of packets. Q_{MAX} represents the maximum number of packets, which the queue can hold and Q_k is the number of packets in the queue at the start of the k-th time interval.

If $W_k < N$ congestion does not occur and the queue empties unless the source times out early.

If $W_k \geq N$ congestion occurs and the queue empties unless the source times out early.

An early timeout takes place when, with the queue being full after one unit of time, the source does not wait long enough for the queue to empty, i.e. the source waits less than $1 + \left\lceil \frac{Q_{MAX}}{P} \right\rceil + RTT$

Where RTT represents the total time needed for a packet to reach the sink and the corresponding acknowledgement to reach the source.

By setting the retransmission timeout:

$$RTO \geq 1 + \left\lceil \frac{Q_{MAX}}{P} \right\rceil + RTT \tag{2}$$

early timeouts are not possible and

$$Q_k = 0 \text{ for all } k \tag{3}$$

So in the case of the one source and one one link, the parameter N is given by:

$$N = P + Q_{MAX} + 1 \tag{4}$$

3. Two sources / One Link

In the case of two sources and one link (fig.2), the same assumptions stand for the sources and link as in the case of one source and one link.

3.1. Synchronized Sources with no delay

When the two sources are synchronized with no delay, i.e. they start to transmit data at the same time, the congestion condition is given by:

$$W_{1k} + W_{2k} \geq P + Q_{MAX} + 1 \qquad (5)$$

3.2. Synchronized Sources with Unit delay

Assume that source one transmits W_{1k} packets in the time interval $[t_k, t_{k+1})$.

At $t = t_k$ either source one had timed out or it had received the last outstanding acknowledgement. Then:

$$Q_k' \equiv Q(t_k + 1) =$$

$$\begin{cases} Q_k + W_{1k} - P & \text{if } 0 \leq Q_k + W_{1k} - P \leq Q_{MAX} \\ 0 & \text{if } Q_k + W_{1k} - P \leq 0 \\ Q_{MAX} & \text{if } Q_k + W_{1k} - P > Q_{MAX} \end{cases} \qquad (6)$$

W_{2k} packets are transmitted in the time interval $[t_k + j_k, t_k + j_k + 1)$. Then:

$$Q_k'' \equiv Q(t_k + j_k + 1) =$$

$$\begin{cases} Q_k' + W_{2k} - j_k P & \text{if } Q_{MAX} \geq Q_k' \geq (j_k - 1)P \\ & \text{and } 0 \leq Q_k' + W_{2k} - j_k P \leq Q_{MAX} \\ 0 & \text{if } Q_{MAX} \geq Q_k' \geq (j_k - 1)P \\ & \text{and } Q_k' + W_{2k} - j_k P \leq 0 \\ Q_{MAX} & \text{if } Q_{MAX} \geq Q_k' \geq (j_k - 1)P \\ & \text{and } Q_k' + W_{2k} - j_k P > Q_{MAX} \\ W_{2k} - P & \text{if } Q_k' < (j_k - 1)P \\ & \text{and } 0 \leq W_{2k} - P \leq Q_{MAX} \\ 0 & \text{if } Q_k' < (j_k - 1)P \\ & \text{and } W_{2k} - P < 0 \\ Q_{MAX} & \text{if } Q_k' < (j_k - 1)P \\ & \text{and } W_{2k} - P > Q_{MAX} \end{cases} \qquad (7)$$

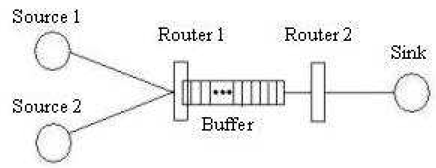

Figure 2: Network with 2 Sources and 1 Link.

Case 1

If sources one and two congest:

$$Q_k' = Q_{MAX} \qquad (8)$$

and $Q_k'' =$

$$\begin{cases} Q_{MAX} + W_{2k} - j_k P & \text{if } 0 \leq Q_{MAX} + W_{2k} - j_k P \leq Q_{MAX} \\ 0 & \text{if } Q_{MAX} + W_{2k} - j_k P \leq 0 \\ Q_{MAX} & \text{if } W_{2k} - j_k P > 0 \end{cases} \qquad (9)$$

But since source two congests, $Q_k'' = Q_{MAX}$ and

$$Q_{k+1} \equiv Q_k'' - (h_k - 1)P =$$

$$\begin{cases} Q_{MAX} - (h_k - 1)P & \text{if } Q_{MAX} \geq (h_k - 1)P \\ 0 & \text{if } Q_{MAX} < (h_k - 1)P \end{cases} \qquad (10)$$

Where h_k $(h_k \geq 1)$ represents the additional units of time required, from $t = t_k + j_k$, for source one to receive the final acknowledgement.

Case 2

If source one does not congest and source two congests then the expression for Q_{k+1} is again (10).

Case 3

If source one congests and source two does not congest, by combining (7) and (8) we get

$$Q_{k+1} =$$

$$\begin{cases} Q_{MAX} + W_{2k} - (j_k + h_k - 1)P & \\ \quad \text{if } (h_k - 1)P \leq Q_{MAX} + W_{2k} - j_k P \leq Q_{MAX} & \\ 0 \quad \text{if } Q_{MAX} + W_{2k} - j_k P < (h_k - 1)P & \end{cases} \qquad (11)$$

Case 4

If neither source one not source two congest, sub case two of (6) gives:

$$Q_k' = 0 \quad \text{if} \quad Q_k + W_{1k} - P < 0 \qquad (12)$$

But then

$$Q(t_k + j_k + 1) = \begin{cases} W_{2k} - P \ \ if \ \ W_{2k} \geq P \\ \ \ 0 \quad if \ \ W_{2k} < P \end{cases} \tag{13}$$

So

$$Q_{k+1} = \begin{cases} W_{2k} - h_k P \ \ if \ \ W_{2k} \geq h_k P \\ \ \ 0 \qquad if \ \ W_{2k} < h_k P \end{cases} \tag{14}$$

Whereas from sub case one of (6):

If $0 \leq Q_k + W_{1k} - P < j_k P$ (15)

Then

$$Q(t_k + j_k) = 0$$

Which implies that

$$Q_{k+1} = \begin{cases} W_{2k} - h_k P \ \ if \ \ W_{2k} \geq h_k P \\ \ \ 0 \qquad if \ \ W_{2k} < h_k P \end{cases}$$

If $j_k P \leq Q_k + W_{1k} - P < Q_{MAX}$ (16)

Then

$$Q(t_k + j_k) = Q_k + W_k - j_k P \tag{17}$$

But

$$Q_k'' =$$

$$\begin{cases} Q_k + W_{1k} + W_{2k} - (j_k + 1)P \\ \qquad\qquad if \ \ Q_k + W_{1k} + W_{2k} \geq (j_k + 1)P \\ \ \ 0 \qquad if \ \ Q_k + W_{1k} + W_{2k} < (j_k + 1)P \end{cases} \tag{18}$$

Which implies that

$$Q_{k+1} =$$

$$\begin{cases} Q_k + W_{1k} + W_{2k} - (j_k + h_k)P \\ \qquad\qquad if \ \ Q_k + W_{1k} + W_{2k} \geq (j_k + h_k)P \\ \ \ 0 \qquad if \ \ Q_k + W_{1k} + W_{2k} < (j_k + h_k)P \end{cases} \tag{19}$$

So, the network can be described by a 5-dimentional state with the window sizes $\mathrm{W}_{1k}, \mathrm{W}_{2k}$, the threshold values $\mathrm{Th}_{1k}, \mathrm{Th}_{2k}$ and the queue length Q_{k+1} as the five components.

The network also depends on four parameters, j_k, RTT, Q_{MAX} and P. Assume that j_k and RTT are constant, where

$$RTT = j_k + h_k \tag{20}$$

Also assume

$$Q_{MAX} \geq (j_k - 1)P \tag{21}$$

so the queue does not empty every time.

Q_{k+1} evolves according to (10), (11), (14), (19), whereas the window and threshold evolution follows the next set of equations:

Mode 1 – Slow Start

If source i does not congest and $W_{ik} \leq Th_{ik}$

$$\begin{bmatrix} W_{i,k+1} \\ Th_{i,k+1} \end{bmatrix} = \begin{bmatrix} W_{ik} + \left\lfloor \dfrac{W_{ik} + 1}{2} \right\rfloor \\ Th_{i,k} \end{bmatrix} \tag{22}$$

Mode 2 – Congestion Avoidance

If source i does not congest and $W_{ik} > Th_{ik}$

$$\begin{bmatrix} W_{i,k+1} \\ Th_{i,k+1} \end{bmatrix} = \begin{bmatrix} s_{r_i(k)} \\ Th_{i,k} \end{bmatrix} \tag{23}$$

where $s_{j+1} = s_j + \dfrac{1}{s_j}, \ \ j = 0,1,..., r_i(k-1)$ (24)

$$s_0 = W_{ik} \tag{25}$$

and $r_i(k) = \left\lfloor \dfrac{W_{ik} + 1}{2} \right\rfloor$ (26)

Mode 3 – Congestion Detection

If source i does experience congestion

$$\begin{bmatrix} W_{i,k+1} \\ Th_{i,k+1} \end{bmatrix} = \begin{bmatrix} 1 \\ Th_{i,k} \end{bmatrix} \tag{27}$$

In the specific network:
No congestion occurs when

$$W_{1k} + W_{2k} < Q_{MAX} + P + 1$$

Source one congests when

$$W_{1k} + W_{2k} \geq Q_{MAX} + P + 1$$

$$W_{1k} > \frac{1}{2}(Q_{MAX} + P) \ \ and \ \ W_{2k} < \frac{1}{2}(Q_{MAX} + P) \tag{28}$$

Source two congests when

$$W_{1k} + W_{2k} \geq Q_{MAX} + P + 1$$

$$W_{1k} < \frac{1}{2}(Q_{MAX} + P) \ \ and \ \ W_{2k} > \frac{1}{2}(Q_{MAX} + P) \tag{29}$$

And both sources congest when

$$W_{1k} + W_{2k} \geq Q_{MAX} + P + 1$$
$$W_{1k} > \frac{1}{2}(Q_{MAX} + P) \ \ and \ \ W_{2k} > \frac{1}{2}(Q_{MAX} + P) \quad (30)$$

3.3. Simulation Results

MATLAB simulation for various combinations of initial window and threshold values and also for various values of RTT and j_k establishes that the network possesses various steady state behaviours. There are cases where the network converges to one or multiple fixed points, but also to 2-cycles, 3-cycles, up to 7-cycles.

Plotting Th_1 and Th_2 as functions of j_k (fig.3a, fig.3b), while considering for the first the upper bound of possible threshold values and for the later the lower bound, enabled us to realize that Th_1 is monotonically decreasing in j_k whereas Th_2 is monotonically increasing.

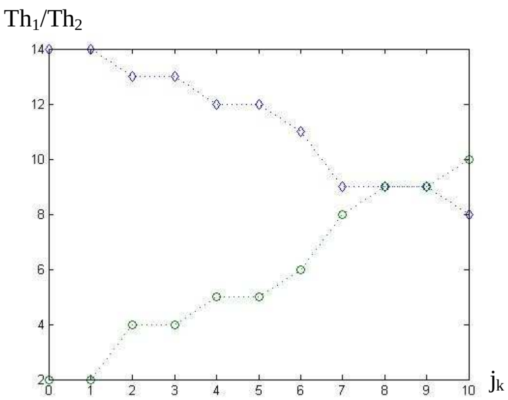

Figure 3a: Evolution of Thresholds ($W_1(0)$=1, $W_2(0)$=2, $Th_1(0)=Th_2(0)$=41, Q=32, P=2, RTT=15, j_k=[0,10])

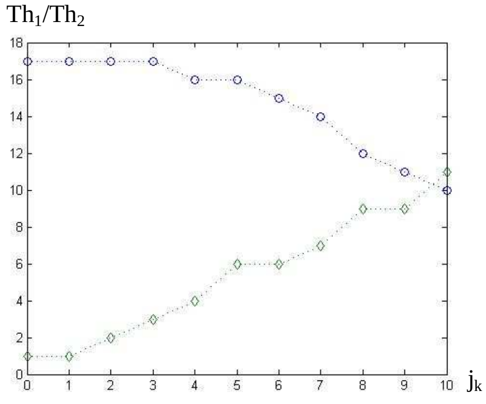

Figure 3b: Evolution of Thresholds ($W_1(0)$=1, $W_2(0)$=3, $Th_1(0)=Th_2(0)$=31, Q=32, P=2, RTT=20, j_k=[0,10])

4. Conclusions

In the case of two sources and one link, the network, converges not only to fixed points, but also 2-cycle, 3-cycle, up to 7-cycle. The thresholds, as a function of j_k, possesses the property of monotonicity. There are also indications that the bifurcation points could be simple fractions of RTT.

Future work will involve further analysis of the simulation results in order to determine the bifurcation points and the cycle-order. Theoretical proof of those properties will be attempted. Furthermore, the extension of the analysis to a m-sources and n-links network is the ultimate goal.

Acknowledgments

This research was supported by Science Foundation Ireland.

References

[1] V. Jacobson, "Congestion avoidance and control," in *Proc. ACM SIGCOMM*, Stanford, CA, pp. 314–329, Aug. 1988.

[2] M. Allman, V. Paxson, and W. Stevens, "TCP congestion control," RFC 2581, Apr. 1999.

[3] S. Floyd, "High Speed TCP for large congestion windows," RFC 3649, 2003.

[4] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, pp. 237–252, 1998.

[5] J.P. Hespanha, "Stochastic Hybrid Systems: Application to communication networks," *Hybrid Systems: Computation and Control, Proceedings*, 2993: 387-401, 2004.

[6] R. Shorten, F. Wirth and D. Leith, "A positive systems model of TCP-like congestion control: Asymptotic results" IEEE/ACM transactions on networking, vol. 14, no.3, June 2006.