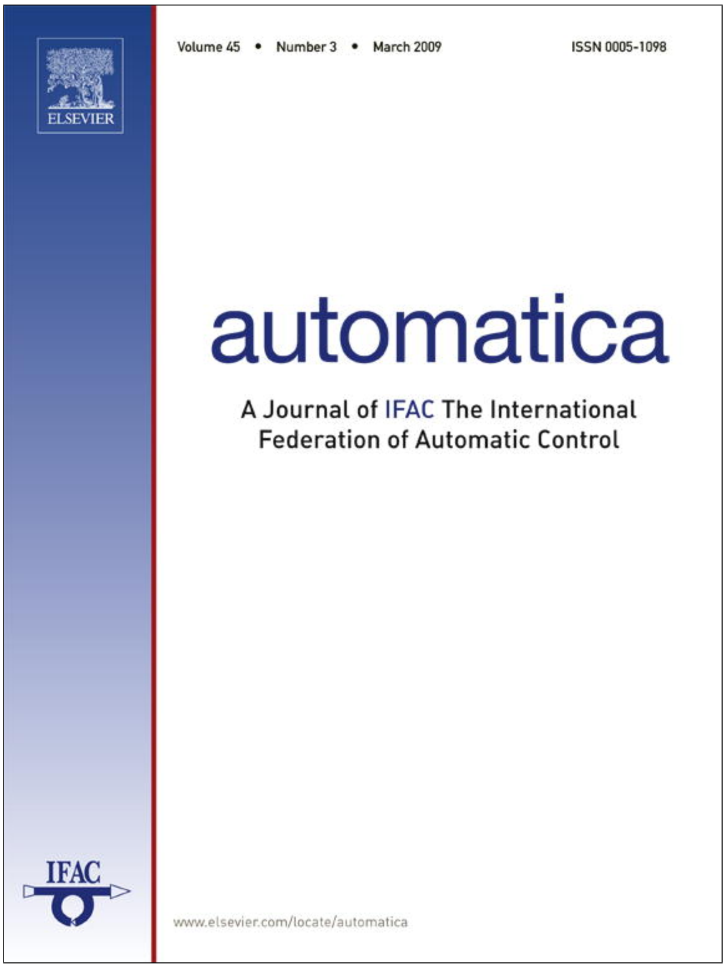

Volume 45 • Number 3 • March 2009
ISSN 0005-1098
ELSEVIER
automatica
A Journal of IFAC The International Federation of Automatic Control
IFAC
www.elsevier.com/locate/automatica

Automatica 45 (2009) 757–763

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

An H_∞ approach to the controller design of AQM routers supporting TCP flows[☆]

Feng Zheng [*], John Nelson

Department of Electronic and Computer Engineering, University of Limerick, Limerick, Ireland

ARTICLE INFO

Article history:
Received 9 August 2005
Received in revised form
8 February 2008
Accepted 2 October 2008
Available online 24 December 2008

Keywords:
Congestion control
AQM
Stability
Time delay
Disturbance rejection

ABSTRACT

In this paper, we aim at developing an H_∞ approach to the design of an active queue management (AQM) based congestion controller for the Internet Protocol. In the approach, the available link bandwidth is modelled as a nominal constant value, which is known to the link, plus a time-variant disturbance, which is unknown. The design objective is to minimize the ratio, denoted as γ, between the norm of the perturbed queue length and that of the disturbance. An important feature of the approach is that the system performances, including the disturbance rejection ratio γ and hence the stability of closed-loop systems, are guaranteed for all round-trip times that are less than a known value. The design is described by matrix inequalities, which can be efficiently solved by the linear matrix inequality toolbox in Matlab.

1. Introduction

The RED (random early detection) is a popular method for the congestion control of TCP networks, but recent analysis (Hollot, Misra, Towsley, & Gong, 2002; Low, Paganini, & Doyle, 2002) has shown that the scheme is prone to instabilities when network delays are taken into account. Some new design methods are reported in Chen and Yang (2005, 2007), Fan, Arcak, and Wen (2004), Hollot et al. (2002), Peng, Xu, and Lin (2006) and Quet and Özbay (2004) for the RED. In Fan et al. (2004) and Hollot et al. (2002), the stability of closed-loop congestion control system taking into account round-trip delays was addressed using the small gain theorem; in Quet and Özbay (2004), a robust controller for the AQM-based congestion control problem was presented based on the infinite-dimensional system theory; in Chen and Yang (2005) and Chen and Yang (2007), the H_∞ and μ analysis approaches for the RED congestion control algorithm were proposed, respectively; in Peng et al. (2006), the so-called α-stability criterion was used to design the marking probability function.

While great progress has been made in new congestion control schemes, some problems are still not sufficiently addressed. One important problem is the robustness of the congestion control algorithm against the disturbance on the available link bandwidth since it is often time-varying and cannot be exactly measured. In previous studies such as (Altman, Başar, & Srikant, 1999; Cavendish, Gerla, & Mascolo, 2004; Mascolo, 1999; Yan & Bitmead, 2005), all the available link bandwidths are modelled as a disturbance. In practice, it is more appropriate to model the available link bandwidth as a nominal constant value, which is known to the link, plus a disturbance, which is unknown and time-varying. The nominal part consists of those bandwidths occupied by long-lived connections such as FTPs, while the unknown disturbance consists of those bandwidths occupied by short-lived connections such as http. In this paper, we aim at developing an H_∞ approach to the design of the RED congestion algorithms by considering a new model of the available link bandwidth.

For the H_∞ design, the main difference between this paper and the previous studies (Chen & Yang, 2005, 2007; Quet & Özbay, 2004) is[1]: the approach proposed here uses a time-domain H_∞ design method, which can deal with the situation where the round-trip time varies with time, while the previous reports use frequency-domain design methods, which *theoretically* require that the system under consideration is time invariant.

☆ This work was supported by National Communications Network Research Centre, a Science Foundation Ireland Research Project, under Grant 03/IN3/1396. This paper was presented at the International Control Conference 2006, Glasgow, Scotland, 30 Aug.–1 Sept. 2006. This paper was recommended for publication in revised form by Associate Editor Xiaohong Guan under the direction of Editor Toshiharu Sugie.

* Corresponding address: Institute of Communications Technology, Faculty of Electrical Engineering and Computer Science, Leibniz University Hannover, 30167 Hannover, Germany. Tel.: +49 5117622848; fax: +49 5117623030.

E-mail addresses: feng.zheng@ikt.uni-hannover.de, feng.zheng@ul.ie (F. Zheng), john.nelson@ul.ie (J. Nelson).

1 The following comments refer to the limitation in the relevant approaches of the controller *design*. Note that the time-varying delay is considered in all the studies in their simulations.

0005-1098/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2008.10.014

About the controller design for time delay systems, readers are referred to Richard (2003) for a comprehensive survey, where the results for various kinds of time delay systems and design approaches are reviewed. Notice that the TCP congestion control model involves time delay not only in the state variable but also in the control input. Due to this fact, some design methods proposed recently, e.g., Lee, Kwon, and Park (2007), Lee, Moon, Kwon, and Park (2004) and Yang, Wang, Hung, and Gani (2006), cannot apply to the problem considered here. For the effect of the delay in the control input on the controller design, readers may be referred to Zheng, Cheng, and Gao (1994).

The starting point of the paper is the fluid model of TCP congestion-avoidance algorithm as proposed in Misra, Gong, and Towsley (2000) (see also Hollot et al. (2002)). A theoretical justification of how this model fits for the practical process is shown in Misra et al. (2000).

Notation. $\mathbb{R}^n$ denotes n-dimensional Euclidean space, $\mathbb{L}_2^n[0,\infty)$ stands for the space of functions taking values in $\mathbb{R}^n$ and that are square integrable over $[0,\infty)$, and I is an identity matrix whose dimension is implied from context. The notation $X > Y$ means that $X - Y$ is positive definite.

2. Problem formulation

In Misra et al. (2000), a dynamic model of TCP behavior was developed using fluid-flow and stochastic differential equation analysis. Similar to Hollot et al. (2002), here a simplified version of that model is used which neglects the TCP timeout mechanism. This model is described by the following coupled and nonlinear delay-differential equations:

$$\begin{cases} \dot{W}(t) = \dfrac{1}{\tau(t)} - \dfrac{W(t)}{2}\dfrac{W(t-\tau(t))}{\tau(t-\tau(t))}p(t-\tau(t)), \\ \dot{q}(t) = \begin{cases} -C(t) + \dfrac{N(t)}{\tau(t)}W(t), & q(t) > 0, \\ \max\left\{0, -C(t) + \dfrac{N(t)}{\tau(t)}W(t)\right\}, & q(t) = 0, \end{cases} \\ \tau(t) = \dfrac{q(t)}{C(t)} + T_p, \end{cases} \tag{1}$$

where W is the TCP window size (in packets), q the queue length in the router (in packets), τ the round-trip time (in secs), C the available link capacity (in packets/s), T_p propagation delay (in secs), N the number of TCP sessions, and p the probability of packet mark.

The queue length and window-size are positive and bounded quantities, i.e., $q \in [0, \bar{q}]$ and $W \in [0, \bar{W}]$, where $\bar{q}$ and $\bar{W}$ denote buffer capacity and maximum window size, respectively. The marking probability p belongs to the interval $[0, 1]$. In practical networks, the available link capacity changes with time and it is difficult to measure. Therefore it is taken as a disturbance in a lot of studies, see e.g. Cavendish et al. (2004), Fan et al. (2004) and Mascolo (1999). In this paper, it is supposed that the nominal value, say C_0, of $C(t)$ is known, while $\delta C(t) \triangleq C(t) - C_0$ is unknown and considered as a disturbance to the system.

Take (W, q) as the state and p as the input. For a given triplet of network parameters (N, C_0, T_p), any triplet (W_0, q_0, p_0) that is in the set $\Omega = \{(W_0, q_0, p_0) : W_0 \in [0, \bar{W}], q_0 \in [0, \bar{q}], p_0 \in [0, 1], \tau_0 = \frac{q_0}{C_0} + T_p, W_0 = \frac{\tau_0 C_0}{N}, p_0 = \frac{2}{W_0^2}\}$ is a possible operating point. Now define

$$\delta W = W - W_0, \qquad \delta q = q - q_0, \qquad \delta p = p - p_0, \qquad \delta C = C - C_0.$$

Following the same line as in Hollot et al. (2002) and taking into account the similar simplifying consideration on the nesting function for the round-trip time as in Low et al. (2002), we can obtain the linearized version of (1) as follows

$$\begin{cases} \delta\dot{W} = -\dfrac{N}{\tau_0^2 C_0}[\delta W(t) + \delta W(t-\tau_0)] \\ \qquad - \dfrac{1}{\tau_0^2 C_0}[\delta q(t) - \delta q(t-\tau_0)] \\ \qquad - \dfrac{\tau_0 C_0^2}{2N^2}\delta p(t-\tau_0) + \dfrac{\tau_0 - T_p}{\tau_0^2 C_0}[\delta C(t) - \delta C(t-\tau_0)], \\ \delta\dot{q} = \dfrac{N}{\tau_0}\delta W(t) - \dfrac{1}{\tau_0}\delta q(t) - \dfrac{T_p}{\tau_0}\delta C(t). \end{cases} \tag{2}$$

The objective of this paper is to develop an H_∞ design approach for the problem of AQM-based congestion control based on the dynamic model (2), which guarantees the ratio between the norm of some desired variables and that of the disturbance being less than some specified value. Furthermore, this specified value for the ratio can be minimized for a given group of network parameters. To this end, we will first study the H_∞ control of general linear time delay systems and then apply the result to the above system.

We would like to point out that our design belongs a type of "re-design", which means that the basic congestion control algorithm or structure is the same as the RED. The only difference lies in how to calculate the marking probability, which will be addressed in Theorem 1.

3. H_∞ control of linear time delay systems

Now consider the following system

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-\tau(t)) + B_0 u(t) + B_1 u(t-\tau(t)) + Dv(t), \quad z(t) = Hx(t), \tag{3}$$

where $x(t) \in \mathbb{R}^n$ is the system state, $u(t) \in \mathbb{R}^{n_u}$ the control input, $v \in \mathbb{R}^{n_v}$ the exogenous disturbance, $z \in \mathbb{R}^{n_z}$ the controlled output, and τ the time-delay involved. Suppose that τ is upper-bounded by τ_m: $0 \le \tau \le \tau_m$. All matrices are of appropriate dimensions. Throughout this section, it is defined that $A = A_0 + A_1$ and $B = B_0 + B_1$.

For a prescribed scalar $\gamma > 0$, define the performance index as

$$J(\gamma) = \int_0^\infty (z^T(t)z(t) - \gamma^2 v^T(t)v(t))\mathrm{d}t. \tag{4}$$

The objective is to find a control law of the type $u(t) = Kx(t)$ such that the closed-loop system satisfies $J(\gamma) < 0$ for any $v \in L_2^{n_v}[0,\infty)$. Furthermore, minimize γ if possible. Note that the requirement $J(\gamma) < 0$ means that

$$\frac{\|z\|}{\|v\|} \triangleq \frac{\sqrt{\int_0^\infty z^T(t)z(t)\mathrm{d}t}}{\sqrt{\int_0^\infty v^T(t)v(t)\mathrm{d}t}} < \gamma, \tag{5}$$

where $\|\cdot\|$ refers to the 2-norm in the space $L_2^{n_v}[0,\infty)$. Eq. (5) says that the ratio between the norm of the controlled output and that of the disturbance is less than a specified number γ.

To solve the above problem, the bounded real lemma (BRL) for time delay systems is needed. Up to now, several versions of BRL have been reported, see, e.g., Fridman and Shaked (2001) and Shaked, Yaesh, and de Souza (1998). Among these, the result recently presented in Fridman and Shaked (2001) seems most tight. Therefore, the version in Fridman and Shaked (2001) will be adopted, tailored to the system considered here.

Lemma 1 ((Fridman & Shaked, 2001)). *Consider system (3) with $u(t) \equiv 0$. If there exist matrices $R > 0$ and $P \triangleq \begin{bmatrix} P_1 & 0 \\ P_2 & P_3 \end{bmatrix}$ with $P_1 > 0$ such that the following linear matrix inequality (LMI)*

$$\begin{bmatrix} P^{\mathrm{T}}\begin{bmatrix} 0 & I \\ A & -I \end{bmatrix} + \begin{bmatrix} 0 & A^{\mathrm{T}} \\ I & -I \end{bmatrix} P + \begin{bmatrix} H^{\mathrm{T}}H & 0 \\ 0 & \tau_m R \end{bmatrix} & P^{\mathrm{T}}\begin{bmatrix} 0 \\ D \end{bmatrix} & \tau_m P^{\mathrm{T}}\begin{bmatrix} 0 \\ A_1 \end{bmatrix} \\ \begin{bmatrix} 0 & D^{\mathrm{T}} \end{bmatrix} P & -\gamma^2 I & 0 \\ \tau_m \begin{bmatrix} 0 & A_1^{\mathrm{T}} \end{bmatrix} P & 0 & -\tau_m R \end{bmatrix} < 0 \tag{6}$$

holds, then system (3) achieves $J(\gamma) < 0$.

Based upon Lemma 1, the following theorem can be established.

Theorem 1. *Consider system (3). If there exist matrices $Q_1 > 0$, Q_2, Q_3, Y and positive number ε such that the matrix inequality in Box I holds, then the closed-loop system achieves $J(\gamma) < 0$ with the controller*

$$u(t) = Kx(t), \quad K = YQ_1^{-1}. \tag{7}$$

Proof. Substituting Eq. (7) into (3) and denoting

$$\bar{A}_0 = A_0 + B_0 K, \quad \bar{A}_1 = A_1 + B_1 K, \quad \bar{A} = \bar{A}_0 + \bar{A}_1,$$

according to Lemma 1, it can be concluded that the closed-loop system (3) achieves $J(\gamma) < 0$ if matrix inequality (6) holds when A and A_1 in (6) are substituted by $\bar{A}$ and $\bar{A}_1$, respectively. But matrix inequality (6) is difficult to solve, so it is converted into an easily solvable matrix inequality. From

$$P^{\mathrm{T}}\begin{bmatrix} 0 & I \\ \bar{A} & -I \end{bmatrix} + \begin{bmatrix} 0 & \bar{A}^{\mathrm{T}} \\ I & -I \end{bmatrix} P + \begin{bmatrix} H^{\mathrm{T}}H & 0 \\ 0 & \tau_m R \end{bmatrix} < 0$$

it can be seen that $-(P_3 + P_3^{\mathrm{T}})$ must be negative definite. Thus P is nonsingular. Let

$$P^{-1} \triangleq Q \triangleq \begin{bmatrix} Q_1 & 0 \\ Q_2 & Q_3 \end{bmatrix}.$$

Define $\Delta = \mathtt{diag}\{Q, I, Q_1\}$, where diag denotes a block diagonal matrix with its diagonal elements being specified by its arguments, and denote

$$\Psi_1 = \begin{bmatrix} 0 & I \\ \bar{A} & -I \end{bmatrix} Q + Q^{\mathrm{T}} \begin{bmatrix} 0 & \bar{A}^{\mathrm{T}} \\ I & -I \end{bmatrix}.$$

Multiplying the left-hand side of (6) by Δ^{T} and Δ on the left and right, respectively, the matrix inequality (8) can be obtained

$$\begin{bmatrix} \Psi_1 + Q^{\mathrm{T}}\begin{bmatrix} H^{\mathrm{T}} & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & \tau_m R \end{bmatrix}\begin{bmatrix} H & 0 \\ 0 & I \end{bmatrix} Q & \begin{bmatrix} 0 \\ D \end{bmatrix} & \tau_m \begin{bmatrix} 0 \\ \bar{A}_1 \end{bmatrix} Q_1 \\ \begin{bmatrix} 0 & D^{\mathrm{T}} \end{bmatrix} & -\gamma^2 I & 0 \\ \tau_m Q_1 \begin{bmatrix} 0 & \bar{A}_1^{\mathrm{T}} \end{bmatrix} & 0 & -\tau_m Q_1 R Q_1 \end{bmatrix} < 0. \tag{8}$$

Applying Schur complements, it can be seen that (8) is equivalent to

$$\begin{bmatrix} \begin{bmatrix} 0 & I \\ \bar{A} & -I \end{bmatrix} Q + Q^{\mathrm{T}}\begin{bmatrix} 0 & \bar{A}^{\mathrm{T}} \\ I & -I \end{bmatrix} & \begin{bmatrix} 0 \\ D \end{bmatrix} & \tau_m \begin{bmatrix} 0 \\ \bar{A}_1 \end{bmatrix} Q_1 & Q^{\mathrm{T}}\begin{bmatrix} H^{\mathrm{T}} & 0 \\ 0 & I \end{bmatrix} \\ \begin{bmatrix} 0 & D^{\mathrm{T}} \end{bmatrix} & -\gamma^2 I & 0 & 0 \\ \tau_m Q_1 \begin{bmatrix} 0 & \bar{A}_1^{\mathrm{T}} \end{bmatrix} & 0 & -\tau_m Q_1 R Q_1 & 0 \\ \begin{bmatrix} H & 0 \\ 0 & I \end{bmatrix} Q & 0 & 0 & -\begin{bmatrix} I & 0 \\ 0 & (\tau_m R)^{-1} \end{bmatrix} \end{bmatrix} < 0. \tag{9}$$

Now suppose $R = \varepsilon Q_1^{-1}$ and let $KQ_1 = Y$. Expanding the block matrices in (9) shows that (9) is equivalent to the matrix inequality in Box I. The proof is completed. ∎

Remark 1. From Theorem 1, one can see an interesting feature of the approach: the system performances, including the disturbance rejection ratio γ and the implied stability of the closed-loop system, are guaranteed for all time delay that is less than τ_m. This feature is especially important for the congestion control problem since the round-trip time is actually state-dependent and hence time-varying, whereas its upper bound can be roughly estimated.

In congestion control, one important problem is to find how large the time delay is allowed to be such that the network can still be stabilized or an H_∞ performance index can still be guaranteed. This problem can be easily dealt with based on the following corollary.

Corollary 1. *Consider system (3). For a given γ, if there exist matrices $Q_1 > 0$, $\bar{Q}_1 > 0$, Q_2, Q_3, Y and positive number ε such that the following matrix inequalities*

$$\begin{bmatrix} \Psi_2 & \eta_1 & \eta_2 \\ \eta_1^{\mathrm{T}} & -\varepsilon \bar{Q}_1 & 0 \\ \eta_2^{\mathrm{T}} & 0 & -\frac{1}{\varepsilon}\bar{Q}_1 \end{bmatrix} < 0 \tag{10}$$

$$\bar{Q}_1 < \frac{1}{\tau_m} Q_1 \tag{11}$$

hold, where

$$\Psi_2 = \begin{bmatrix} Q_2 + Q_2^{\mathrm{T}} & Q_1 A^{\mathrm{T}} + Y^{\mathrm{T}} B^{\mathrm{T}} - Q_2^{\mathrm{T}} + Q_3 & 0 & Q_1 H^{\mathrm{T}} \\ AQ_1 + BY - Q_2 + Q_3^{\mathrm{T}} & -Q_3 - Q_3^{\mathrm{T}} & D & 0 \\ 0 & D^{\mathrm{T}} & -\gamma^2 I & 0 \\ HQ_1 & 0 & 0 & -I \end{bmatrix},$$

$$\eta_1 = \begin{bmatrix} 0 & (A_1 Q_1 + B_1 Y)^{\mathrm{T}} & 0 & 0 \end{bmatrix}^{\mathrm{T}}, \quad \eta_2 = \begin{bmatrix} Q_2 & Q_3 & 0 & 0 \end{bmatrix}^{\mathrm{T}},$$

then the closed-loop system achieves $J(\gamma) < 0$ with the controller

$$u(t) = Kx(t), \quad K = YQ_1^{-1}.$$

Further more, τ_m can be maximized by solving the generalized eigenvalue problem defined by (10) and (11).

Note that the inequality in Box I is not an LMI. For a given ε, it becomes an LMI. Therefore, to solve it, a sweeping process (sweeping all ε in the interval $(0, \infty)$) is likely to be involved. Another way is to use the method presented in Zheng, Wang, and Lee (2002) to provide some clue for such ε. Using the approach in Zheng et al. (2002), it can be shown that if other variables except ε are fixed, the most possible ε which satisfies the matrix inequality in Box I is given by

$$\varepsilon^* = \sqrt{\frac{\mathtt{tr}\left(\eta_1 Q_1^{-1} \eta_1^{\mathrm{T}}\right)}{\mathtt{tr}\left(\eta_2 Q_1^{-1} \eta_2^{\mathrm{T}}\right)}}, \tag{12}$$

where tr denotes the trajectory of a square matrix. Based on (12), a recursive process may be employed to solve the inequality in Box I.

4. Application to RED-based congestion control problem

In this section, the result obtained in the preceding section will be applied to the RED-based congestion control problem.

A single bottleneck router running N TCP flows is considered. The simulations are conducted in Matlab. The dynamic model for the window size and router queue length is

$$\dot{W}_i(t) = \frac{1}{\tau_i(t)} - \frac{W_i(t)}{2} \frac{W_i(t - \tau_i(t))}{\tau_i(t - \tau_i(t))} p(t - \tau_i(t)), \quad i = 1, \ldots, N, \tag{13}$$

$$\begin{bmatrix} Q_2+Q_2^{\mathrm{T}} & Q_1A^{\mathrm{T}}+Y^{\mathrm{T}}B^{\mathrm{T}}-Q_2^{\mathrm{T}}+Q_3 & 0 & 0 & Q_1H^{\mathrm{T}} & Q_2^{\mathrm{T}} \\ AQ_1+BY-Q_2+Q_3^{\mathrm{T}} & -Q_3-Q_3^{\mathrm{T}} & D & \tau_m(A_1Q_1+B_1Y) & 0 & Q_3^{\mathrm{T}} \\ 0 & D^{\mathrm{T}} & -\gamma^2 I & 0 & 0 & 0 \\ 0 & \tau_m(Q_1A_1^{\mathrm{T}}+Y^{\mathrm{T}}B_1^{\mathrm{T}}) & 0 & -\tau_m\varepsilon Q_1 & 0 & 0 \\ HQ_1 & 0 & 0 & 0 & -I & 0 \\ Q_2 & Q_3 & 0 & 0 & 0 & -\frac{1}{\tau_m\varepsilon}Q_1 \end{bmatrix} < 0$$

Box I.

$$\dot{q}(t) = \begin{cases} -C(t) + \sum_{i=1}^{N} \frac{W_i(t)}{\tau_i(t)} & \text{when } q(t) > 0, \\ \max\left\{0, -C(t) + \sum_{i=1}^{N} \frac{W_i(t)}{\tau_i(t)}\right\} & \text{when } q(t) = 0, \end{cases} \quad (14)$$

where $\tau_i(t) = \frac{q(t)}{C(t)} + T_{pi}$. To simulate the scenario that different connections may have different round-trip time, we allow T_{pi} to depend on i randomly in some range.

Note that the model (13)–(14) is much more complicated than the model (1), and it is difficult to design the congestion controllers based directly on the model (13)–(14). Instead, we will use the model (1) to design the controllers. This idea for the model simplification in the controller design stage has been implicitly used in Hollot et al. (2002) and Quet and Özbay (2004).

The short-lived http flows are introduced into the router and modelled with δC as a birth-and-death process. Specifically, construct δC as:

$$\delta C(t) = B_{av}\kappa(t),$$

where $\kappa(t)$ is a birth-and-death process with the birth and death rates being λ and μ, respectively, and B_{av} is the average transmission rate of http flows. According to Bertsekas and Gallager (1992), it is appropriate to use a birth-and-death process to model http flows. Note that the value of the available link capacity at the equilibrium may be also over-estimated, so δC might take negative values too. Considering this fact, $\kappa(t)$ is allowed to take negative values. It is natural to assume that the birth rate and death rates are equal. Thus such a process is null-recurrent, i.e., the process does not keep visiting any state frequently. Therefore, $\kappa(t)$ will diverge inevitably. Obviously, this does not match the practical situation. To remedy it, we place lower and upper bounds for $\kappa(t)$, namely $-\kappa_{\max} \le \kappa(t) \le \kappa_{\max}$, where $\kappa_{\max}$ is a positive number. Thus $\kappa(t)$ can be viewed as a modified birth-and-death process with lower and upper barriers. This is realized in simulation by simply removing the constraint $\kappa \ge 0$ in the original model of the birth–death process and placing the new constraint $-\kappa_{\max} \le \kappa(t) \le \kappa_{\max}$ on it.

Note that in model (2), the delayed version of the exogenous disturbance also appears in the dynamics of the system. To take into account of this fact, define

$$\bar{v}(t) = [\delta C(t) \quad \delta C(t-\tau_0)]^{\mathrm{T}}$$

and change the objective function for H_∞ control design as

$$\bar{J}(\gamma) = \int_0^\infty (z^{\mathrm{T}}(t)z(t) - \gamma^2 \bar{v}^{\mathrm{T}}(t)\bar{v}(t))\mathrm{d}t.$$

It is clear that the performance $J(\gamma) < 0$ is satisfied if the closed-loop system achieves $\bar{J}(\gamma) < 0$.

Associating model (2) with the general system model studied in Section 3, we have

$$x(t) = \begin{bmatrix} \delta W(t) \\ \delta q(t) \end{bmatrix}, \qquad A_0 = \begin{bmatrix} -\frac{N}{\tau_0^2 C_0} & -\frac{1}{\tau_0^2 C_0} \\ \frac{N}{\tau_0} & -\frac{1}{\tau_0} \end{bmatrix},$$

$$A_1 = \begin{bmatrix} -\frac{N}{\tau_0^2 C_0} & \frac{1}{\tau_0^2 C_0} \\ 0 & 0 \end{bmatrix},$$

$$B_0 = 0, \qquad B_1 = \begin{bmatrix} -\frac{\tau_0 C_0^2}{2N^2} \\ 0 \end{bmatrix}, \qquad D = \begin{bmatrix} \frac{\tau_0 - T_p}{\tau_0^2 C_0} & -\frac{\tau_0 - T_p}{\tau_0^2 C_0} \\ -\frac{T_p}{\tau_0} & 0 \end{bmatrix}.$$

The matrix H is chosen as $H = [0\ 1]$ for all cases to be studied.

The following two cases are investigated: Case (i) - high speed network and Case (ii) - low speed network. In both cases, we set $T_{pi} = (200 + 40\xi)$ ms, where ξ is a random variable depending on i, $i = 1, \ldots, N$, and uniformly distributed over the interval $[-0.5, 0.5]$.[2] Other parameters are chosen as $\bar{W} = 20$, $q_0 = 500$, $\bar{q} = 800$, $B_{av} = 32$ packets/sec, which is equivalent to a bandwidth of 128 Kb/s with a packet size of 500 Bytes, and the birth and death rates are chosen to be $\lambda = 200$ (1/s) and $\mu = 200$ (1/s), respectively, which means that about every 5 ms there is a http flow being connected to or leaving the network.

In Case (i), we set $C_0 = 25\,000$ packets/s, $N = 1200$, and $\kappa_{\max} = 50$. Such a setting means that the link bandwidth is 100 Mb/s if the packet size is 500 Bytes and the maximal link bandwidth disturbance is about 6.4% of its nominal value.

In Case (ii), we set $C_0 = 2500$ packets/s, $N = 120$, and $\kappa_{\max} = 25$. Such a setting means that the link bandwidth is 10 Mb/s if the packet size is 500 Bytes and the maximal link bandwidth disturbance is about 32% of its nominal value.

With the above settings, the equilibriums are

Case (i): $W_0 = 4.5833$, $q_0 = 500$, $p_0 = 0.0952$;
Case (ii): $W_0 = 8.3333$, $q_0 = 500$, $p_0 = 0.0288$.

By use of Theorem 1, we find that the optimal disturbance rejection index is $\gamma_{opt} = 0.1719$ and 0.1783 for Cases (i) and (ii), respectively, and the corresponding controllers are

Case (i): $K_{H_\infty} = [-0.0206\ 1.7345 \times 10^{-5}]$;

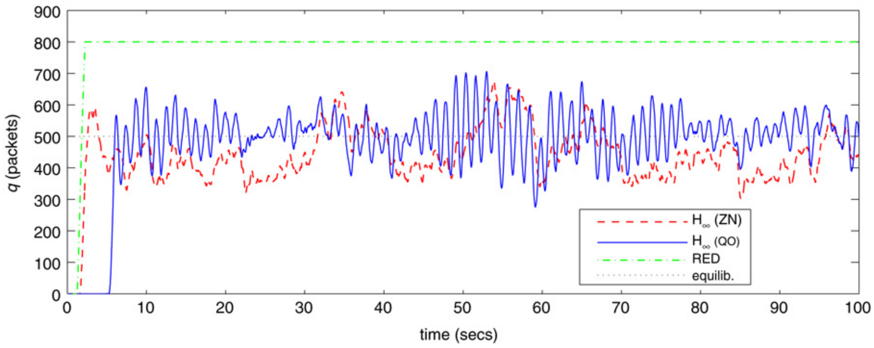
Case (ii): $K_{H_\infty} = [0.0032\ 3.4502 \times 10^{-5}]$.

The performance of the controller developed here and that of RED will be compared. The controller structure of RED is illustrated in Fig. 10 of Hollot et al. (2002). With the above parameter settings, it is reasonable to choose

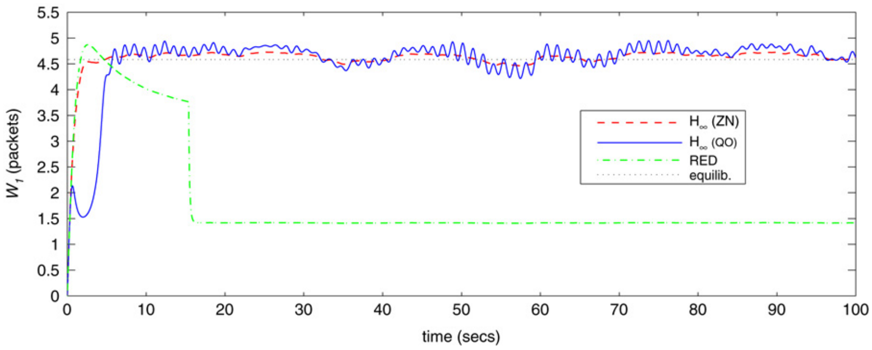
$$q_{RED,\min} = 100, \qquad q_{RED,\max} = 700,$$

where $q_{RED,\min}$ is the minimal queue length under which the sending packets will not be marked/dropped, $q_{RED,\max}$ is the maximal queue length above (but excluding) which all the sending

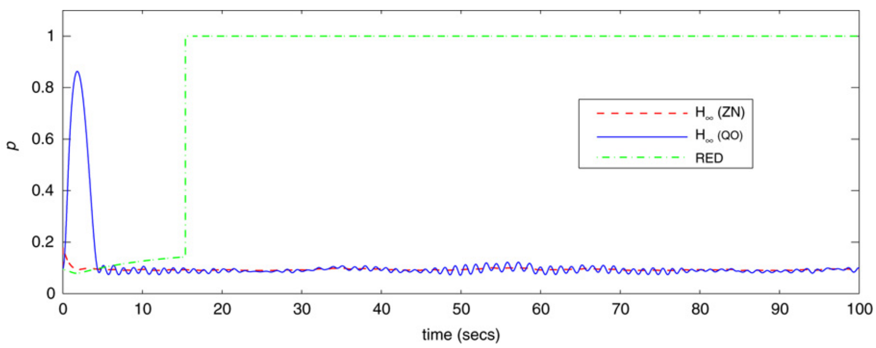
[2] Note that T_{pi} is a random variable instead of a random process. This means that once it is chosen for each i, it will be fixed in the whole time of one simulation. So the randomness of T_{pi} over i can describe the random nature of the locations of different sources.



(a) Queue length.



(b) Window size.



(c) Mark probability.

Fig. 1. The responses and control inputs using H_∞ and RED controllers for Case (i), where the shorthand notation ZN denotes the approach developed in this paper, and QO the approach in Quet and Özbay (2004).

packets will be marked/dropped. The parameter $p_{\max}$ is chosen such that the marking probability is p_0 when $q = q_0$. This constraint yields that

Case (i): $p_{\max} = 0.1428$, $\quad L_{RED} = 2.3802 \times 10^{-4}$;
Case (ii): $p_{\max} = 0.0432$, $\quad L_{RED} = 7.2000 \times 10^{-5}$.

With the above settings and using the approach developed in Hollot et al. (2002), the constant of the RED low-pass filter can be chosen as

Case (i): $K_{RED} = 0.0972$; $\quad$ Case (ii): $K_{RED} = 0.2667$;

which guarantees the stability of the closed-loop system with the phase margin being 90° and 139° for Cases (i) and (ii), respectively. These phase margins are reasonably large for a typical control system.

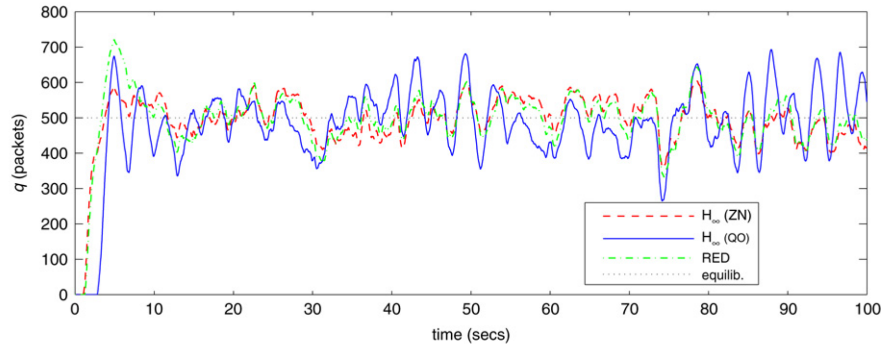
The approach proposed here is also compared with the H_∞ design approach proposed in Quet and Özbay (2004). The controller structure of Quet and Özbay (2004) and the involved parameters are illustrated in Fig. 1 and Eq. (21)–(25) of Quet and Özbay (2004). With the above settings, the controller parameters can be calculated as the following

Case (i): $a_2 = 0.0017$, $\quad b_2 = 33.3346$; $\quad c_2 = 22.9091$;
$c = 1.3791$; $\quad \gamma = 31.6693$.
Case (ii): $a_2 = 5.9802 \times 10^{-4}$, $\quad b_2 = 10.4539$;
$c_2 = 4.0300$; $\quad c = 17.5330$; $\quad \gamma = 70.7657$.

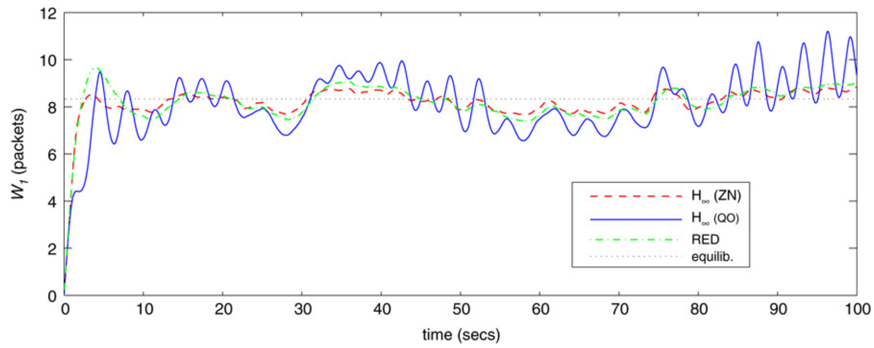
In all the simulations, the initial window size of every source and the initial queue length of the router are set to be zero. For each case, the same disturbance profile on the available link bandwidth is used for both H_∞ and RED controllers.

The results are illustrated in Figs. 1 and 2 for Cases (i)–(ii), respectively. The disturbances on the available link bandwidth for Cases (i) and (ii) are depicted in Fig. 3.

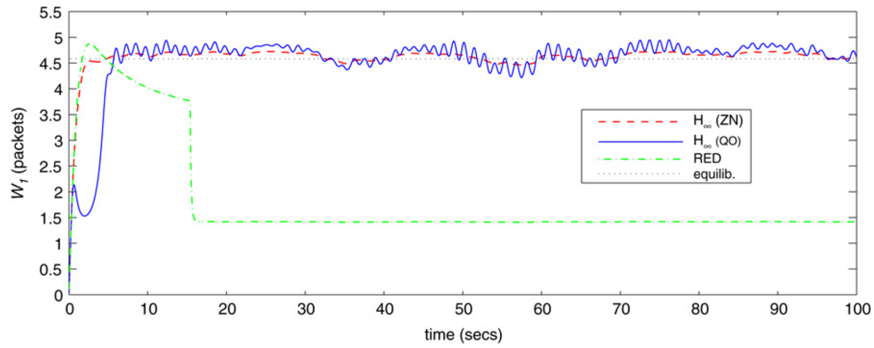
From Figs. 1 and 2, we can see that, in the case of high speed networks, by using the H_∞ congestion controller, a stable operating condition can be built up and maintained even in the situations where the available link bandwidth is subjected to persistent disturbances and the round-trip time varies with different TCP



(a) Queue length.

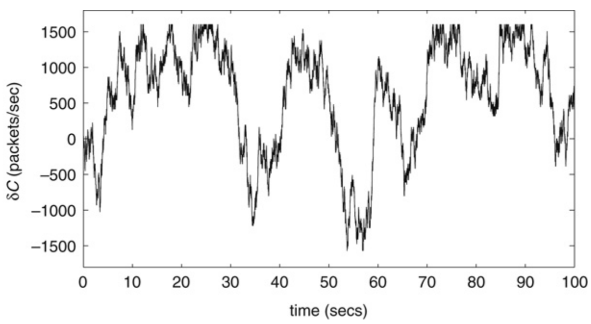


(b) Window size.

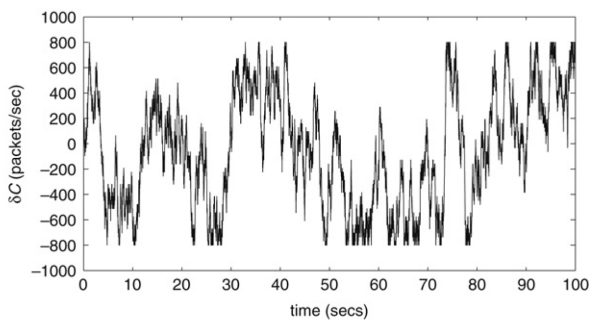


(c) Mark probability.

Fig. 2. The responses and control inputs using H_∞ and RED controllers for Case (ii).



(a) Case (i).



(b) Case (ii).

Fig. 3. The disturbance profile for Cases (i) and (ii). The same profile is used for the situations of both H_∞ and RED controllers.

sessions, while the RED congestion controller fails to do so. This is due to the lag of the response of the conventional RED controller to the sudden change of the network operating condition. From Fig. 2, one can see that both the H_∞ controller developed here and the RED congestion control algorithm work well for low speed networks, while the H_∞ controller developed in Quet and Özbay (2004) yields a more oscillatory response than the other two controllers do. This might be due to the fact that a weighting transfer function $1/s$ is placed on the H_∞ design in Quet and Özbay (2004) (see Eq. (20) therein). While this weighting function can improve the tracking performance of step-like reference inputs, it may weaken the rapidness of the system's response to the rapid fluctuation of available link bandwidth.

Observe the responses of the queue size for the duration of time from 0 to 5 s in Figs. 1 and 2 respectively. The following results can be seen: for both cases, the H_∞ (ZN) (the approach developed here) controller yields lower overshoot than the RED, and yields almost the same rise time as the RED and smaller rise time than the H_∞ (QO) (the approach developed in Quet and Özbay (2004)) controller.

From Fig. 2, it is observed that both the RED and H_∞ (ZN) controller produce similar results. This is due to the principle that the H_∞ design is a worst-case design. So one cannot expect that the H_∞ design always outperforms other approaches. We should admit that the RED works well in normal situations.

5. Concluding remarks

In this paper, a new design method for the H_∞ congestion controller for the Internet protocol TCP has been developed based on the LMI technique. Two events, i.e., the establishment of a fluid model for the TCP congestion-avoidance problem Misra et al. (2000) and the recent advances in time-delay system theory (see, e.g., Fridman and Shaked (2001)), make the effort possible. In the approach, the available link bandwidth is modelled as a nominal constant value, which is known to the link, plus a time-variant disturbance, which is unknown.

The developed approach can theoretically guarantee the system performances, including the disturbance rejection ratio γ and the implied stability of the closed-loop system for all round-trip times that are less than a known value, even though they are time-varying. This property is useful for robust congestion controller design. Finally, it is pointed out that the effectiveness of the proposed approach has been verified only via simulations in Matlab. Further verification via packet-based simulation tools such as NS2 or via experimental studies is needed.

Acknowledgments

The authors thank Dr. P.-F. Quet for providing them with the controller design package in Quet and Özbay (2004). The authors also wish to thank one of the anonymous reviewers for the detailed comments and for bringing to their attention the references Chen and Yang (2005, 2007).

References

Altman, E., Başar, T., & Srikant, R. (1999). Congestion control as a stochastic control problem with action delays. *Automatica*, *35*, 1937–1950.

Bertsekas, D., & Gallager, R. G. (1992). *Data networks*. Upper Saddle River, NJ: Prentice-Hall.

Cavendish, D., Gerla, M., & Mascolo, S. (2004). A control theoretical approach to congestion control in packet networks. *IEEE/ACM Transactions on Networking*, *12*, 893–906.

Chen, Q., & Yang, O. W. W. (2005). Design of AQM controller for IP routers based on H_∞ S/U MSP. In *2005 IEEE intern. conf. on communications* (pp. 340–344).

Chen, Q., & Yang, O. W. W. (2007). Robust controller design for AQM router. *IEEE Transactions on Automatic Control*, *52*, 938–943.

Fan, X., Arcak, M., & Wen, J. T. (2004). Robustness of network flow control against disturbances and time-delay. *Systems Control Letters*, *53*, 13–29.

Fridman, E., & Shaked, U. (2001). New bounded real lemma representations for time-delay systems and their applications. *IEEE Transactions on Automatic Control*, *46*, 1973–1979.

Hollot, C. V., Misra, V., Towsley, D., & Gong, W. B. (2002). Analysis and design of controllers for AQM routers supporting TCP flows. *IEEE Transactions on Automatic Control*, *47*, 945–959.

Lee, Y. S., Kwon, W. H., & Park, P. G. (2007). Authors reply: Comments on delay-dependent robust H_∞ control for uncertain systems with a state-delay. *Automatica*, *43*, 572–573.

Lee, Y. S., Moon, Y. S., Kwon, W. H., & Park, P. G. (2004). Delay-dependent robust H_∞ control for uncertain systems with a state-delay. *Automatica*, *40*, 65–72.

Low, S. H., Paganini, F., & Doyle, J. C. (2002). Internet congestion control. *IEEE Control Systems Magazine*, *22*(Feb.), 28–43.

Mascolo, S. (1999). Congestion control in high-speed communication networks using the Smith principle. *Automatica*, *35*, 1921–1935.

Misra, V., Gong, W.-B., & Towsley, D. (2000). Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED. In *Proc. ACM/SIGCOMM*, (pp. 151–160).

Peng, D., Xu, B., & Lin, Q. (2006). Congestion control research based on time-delay system robust analysis. In *Proc. 6th world congress on intelligent control and automation* (pp. 611–614).

Quet, P.-F., & Özbay, H. (2004). On the design of AQM supporting TCP flows using robust control theory. *IEEE Transactions on Automatic Control*, *49*, 1031–1036.

Richard, J. P. (2003). Time-delay systems: An overview of some recent advances and open problems. *Automatica*, *39*, 1667–1694.

Shaked, U., Yaesh, I., & de Souza, C. (1998). Bounded real criteria for linear time-delay systems. *IEEE Transactions on Automatic Control*, *43*, 1016–1022.

Yan, J., & Bitmead, R. R. (2005). Incorporating state estimation into model predictive control and its application to network traffic control. *Automatica*, *41*, 595–604.

Yang, F., Wang, Z., Hung, Y. S., & Gani, M. (2006). H_∞ control for networked systems with random communication delays. *IEEE Transactions on Automatic Control*, *51*, 511–518.

Zheng, F., Cheng, M., & Gao, W. B. (1994). Feedback stabilization of linear systems with distributed delays in state and control variables. *IEEE Transactions on Automatic Control*, *39*, 1714–1718.

Zheng, F., Wang, Q.-G., & Lee, T. H. (2002). A heuristic approach to solving a class of bilinear matrix inequality problems. *Systems Control Letters*, *47*, 111–119.

Feng Zheng received the B.Sc. and M.Sc. degrees in 1984 and 1987, respectively, both in electrical engineering from Xidian University, Xi'an, China, and the Ph.D. degree in automatic control in 1993 from Beijing University of Aeronautics and Astronautics, Beijing, China. In the past years, he held Alexander-von-Humboldt Research Fellowship at the University of Duisburg and research positions at Tsinghua University, National University of Singapore, and University of Limerick respectively. From 1995 to 1998 he was with the Center for Space Science and Applied Research, Chinese Academy of Sciences as an Associate Professor. Since July 2007 he has been with the Institute of Communications Technology, Leibniz University of Hannover as a Wissenschaftlicher Beirat. He is a co-recipient of several awards, including the National Natural Science Award in 1999 from the Chinese government, the Science and Technology Achievement Award in 1997 from the State Education Commission of China, and the SICE Best Paper Award in 1994 from the Society of Instrument and Control Engineering of Japan at the 33rd SICE Annual Conference, Tokyo, Japan. His research interests are in the areas of UWB-MIMO wireless communications, signal processing, and systems and control theory.

John Nelson is one of the cluster investigators in the Science Foundation Ireland funded National Communications Network Research Centre (NCNRC) at the University of Limerick. Previously, he was an Engineering Director in the Wireless Division at Parthus/Ceva Inc. At Ceva Inc, he undertook major system level projects in the areas of Bluetooth, GPS and 802.11. His core competency is a broad knowledge of wireless and fixed networks, and the skills to build them. Currently he leads UL's involvement in the EU FP6 CAALYX project (http://caalyx.eu), to improve care of the elderly through ambient assisting living.