An LMI Approach to Robust Congestion Control of ATM Networks

Jun Lin, Lihua Xie*, and Huanshui Zhang

Abstract: In this paper, ATM network congestion control with explicit rate feedback is considered. In ATM networks, delays commonly appear in data transmission and have to be considered in congestion control design. In this paper, a bounded single round delay on the return path is considered. Our objective is to design an explicit rate feedback control that achieves a robust optimal H_2 performance regardless of the bounded time-varying delays. An optimization approach in terms of linear matrix inequalities (LMIs) is given. Saturation in source rate and queue buffer is also taken into consideration in the proposed design. Simulations for the cases of single source and multiple sources are presented to demonstrate the effectiveness of the design.

Keywords: Congestion control, network, delay systems, robust H_2 control, LMI.

1. INTRODUCTION

Asynchronous transfer mode (ATM) is a key technology for integrating broadband multimedia services (B-ISDN) in heterogeneous networks. The concept of virtual circuit (VC) deployed combines the advantages of circuit-switched technology (high speed) and packet-switched technology (flexibility, tolerance). Congestion exists in both ATM and non-ATM networks and congestion control is critical in order to improve network utilization. It is the most important mechanism to prevent networks from collapse as well as to ensure quality of service. Among the several services offered by ATM, the available bit rate (ABR) service plays a central role in regulating the network traffic, as it guarantees zero cell loss rate if the source obeys the dynamically varying traffic management signals from the network.

A congestion control scheme is required to efficiently allocate the unused bandwidth of the link to the ABR traffic in order to improve network utilization. Two different congestion control mechanisms are provided in the ATM standard [1]. The first mechanism allows a switch to communicate its congestion status with the sources by using a single bit in the resource management (RM) cells while the other allows the switches to explicitly designate the cell transmission rate by modifying the explicit rate (ER) value of the RM cell. The advantage of the single-bit approach is the implementation simplicity, although it has been shown to exhibit oscillatory behavior [16]. ER switch is rather complex but can improve performance and the effectiveness of an ER switch is highly dependent on the determination method of the ER value.

Based on simplifying assumptions, in existing literature ER feedback congestion control of ABR class of traffic in ATM networks is formulated as linear time-invariant systems without delay [8], linear time-invariant systems with constant delay [3,13] and linear time-invariant systems with uncertain constant delay [6]. In [8], a counter-based congestion avoidance scheme is used. Linear-quadratic-Gaussian (LQG) approach with state-augmentation is used in [3,13]. Bound of delay is used in solving the LQG problem in [8]. We note that there have been many recent interests in control of discrete-time systems with delays; see, for example, [7,11,21] and the references therein.

In this paper, we will focus on the delay which RM cells encounter on the return path. The congestion problem is formulated as a state feedback control of systems with bounded time varying input delays. In the application of congestion control, we are more concerned with the best average performance of the network over a long period of time. Hence, the H_2 performance measure is very appropriate. Therefore, our objective is to design a congestion control that would give rise to a robust optimal H_2 performance regardless of the time-varying delay on the return path.

Manuscript received May 10, 2005; revised October 13, 2005; accepted October 23, 2005. Recommended by Editor Keum-Shik Hong.

Jun Lin and Lihua Xie are with the School of Electrical and Electronic Engineering, Nanyang Technological University, BLK S2, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mails: junlin@pmail.ntu.edu.sg, elhxie@ntu.edu.sg).

Huanshui Zhang is with Information and Control Research Center, Shenzhen Graduate School of Harbin Institute of Technology, HIT Campus, Shenzhen University Town, Xili, Shenzhen 518055, P. R. China (e-mail: h_s_zhang@hit.edu.cn).

* Corresponding author.

As the time-varying delay is assumed to be bounded with known upper bound, the state augmentation approach can be applied to transfer the system into a linear system without delay but with parameter uncertainty. An LMI approach is adopted to design a robust H_2 controller. Saturation in source rate and queue buffer is also taken into consideration in the design of congestion control.

The rest of the paper is arranged as follows. In Section 2, we will introduce a time-varying uncertain delay model for the congestion control of ABR traffic in ATM networks. Section 3 converts the system into a delay free uncertain system via state augmentation and derives a robust H_2 controller with the help of LMI. Simulation results are presented in Section 4 and some conclusion is drawn in Section 5.

2. SYSTEM DESCRIPTION

The mathematical model of congestion control is taken from [2,3]. The ABR source is the only traffic class which responds to feedback information of a node for rate adjustment to prevent network congestion and to maintain quality of service (QoS) to all connections. The feedback information is the available transmission capacity (bandwidth) and queue level at the bottleneck node. Since the available node capacity for the ABR source changes over time in an unpredictable way due to the higher priority sources, the CBR (constant bit rate) and VBR (variable bit rate) source rates are represented as interferences. Let ζ_k denote the higher priority source (interference) which is modelled as a stable ARMA process [4]. Such a formulation allows for long-range correlated traffic. Let q_k be the queue length at the bottleneck and μ_k the effective service rate available for the traffic of the given source in that link at time instant k. Let r_k denote the effective source rate measured at the congestion switch. Without loss of generality, we consider the case of single connection. In this case, the queue length equation is given by

$$q_{k+1} = q_k + r_k - \mu_k. \tag{1}$$

The effective service rate is modelled as

$$\mu_k = \mu + \zeta_k, \tag{2}$$

$$\zeta_{k+1} = \sum_{i=1}^{p} l_i \zeta_{k+1-i} + \rho w_k, \tag{3}$$

where μ is the constant nominal service rate and $\{l_i\}_{i=1,2,\cdots,p}$ are known parameters. $\{w_k\}_{k\geq 1}$ is a zero-mean i.i.d. Gaussian sequence with unit variance and ρ is a known constant. We assume that there is no cell loss and let u_k denote explicit cell rate (ER) calculated by switch. The delay between u_k and r_k is d_k which is the round trip delay. d_k consists of two path delays, one is return path delay and the other is forward path delay. On the return path RM cells travel from the switch to the source. On the forward path the user data travels from the source through the congested switch. In ATM network, packet delays, transmission delays, processing delays and queuing delays exist in transmission on the both paths and the queuing delay is dominant [18]. Actually the two paths are one single communication link and hence the round trip delay d_k can be considered. The relationship between u_k and r_k can be expressed as

$$r_k = u_{k-d_k}, \tag{4}$$

where d_k is known to be bounded with upper bound m.

In the case when d_k is known, i.e. the explicit ER is time-stamped, we formulate the congestion control problem as a LQG control problem [2], that is, we seek a feedback control $u_k = u_k(q_k, \mu_k)$ that minimizes the cost

$$J = \lim_{N\to\infty} \frac{1}{N} E\left\{\sum_{k=1}^{N} [(q_k - q_d)^2 + \lambda^2 (r_k - \mu_k)^2]\right\}, \tag{5}$$

where q_d is the target queue length and λ is a weighting factor. It is clear that the objective is to make the queue buffer close to the desired level while the difference between the source rate and the service rate should not be too large.

The above criterion combines the performance of queue length and accumulation of the difference between switch input and output. We should note that round trip delay in transmission is one reason of the disagreement between the switch input and output.

In the case when the delay d_k is not known, we shall be concerned with designing a controller that minimizes an upper bound of (5) for all possible time-varying uncertainties within the bound of m.

Observe that the above criterion does not consider the saturation in queue buffer and service rate. We shall address this issue later.

Remark 1: The choice of weighting λ_i depends on the importance of keeping the queue length close to the desired level in relation to that of reducing the difference between the source and service rates.

Remark 2: In [12], an H_∞ control approach is adopted where w_k is assumed to an energy bounded deterministic signal. While guaranteeing the worst-case performance, the H_∞ approach is generally conservative.

Remark 3: In the multi-source case, the control objective becomes finding inputs $r_{i,k}$ so as to minimize

$$J = \lim_{N\to\infty} \frac{1}{N} E\left\{\sum_{k=1}^{N}[(q_k - q_d)^2 + \sum_{i=1}^{M}\lambda_i^2(r_{i,k} - a_m\mu_k)^2]\right\}, \tag{6}$$

where $M, \lambda_i, r_{i,k}$ and a_m are respectively the number of sources, weighting, effective source rate for source i and share of input bandwidth of the i-th source ($\sum_{i=1}^{M} a_i = 1$). If we consider fair sharing of available bandwidth, we can simply define $a_i = 1/M, i = 1,2,\cdots,M$. There will be jitter in queue length when new sources add in the congested switch. The reason is that RM cells come back to different sources with different delays. See the examples in Section 4. In spite of the jitter, the fair sharing of bandwidth is still a good scheme for its simplicity in application.

Denote

$$x(k) = \begin{pmatrix} q_k - q_d \\ \zeta_{k+1-p} \\ \vdots \\ \zeta_k \end{pmatrix},$$

$$v(k - d_k) = u_{k-d_k} - \mu = r_k - \mu,$$

$$z(k) = \begin{bmatrix} q_k - q_d, \\ \lambda(r_k - \mu_k) \end{bmatrix}.$$

Then, the system (1)-(4) can be rewritten as

$$x(k+1) = Ax(k) + Bv(k - d_k) + B_1 w(k), \tag{7}$$

$$z(k) = Cx(k) + Dv(k - d_k), \tag{8}$$

where $0 < d_k \le m$, $w(k)$ is a white noise with unit variance, and

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & l_p & l_{p-1} & l_{p-2} & \cdots & l_1 \end{bmatrix},$$

$$B = [1 \quad 0 \quad \cdots \quad 0]^T,$$

$$B_1 = [0 \quad 0 \quad \cdots \quad 0 \quad \rho]^T,$$

$$C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & -\lambda \end{bmatrix},$$

$$D = [0 \quad \lambda]^T.$$

Obviously, (5) can be written as

$$J = \lim_{N\to\infty} \frac{1}{N} E\left\{\sum_{k=1}^{N} z^T(k) z(k)\right\}. \tag{9}$$

We assume that each ABR source is both compliant and greedy, which implies that a command received from the congested switch is immediately executed.

In the multi-source case, effective service rate μ_k and effective source rate r_k are vectors. However, in order to simplify the problem, we assume that d_k, r_k, μ_k are scalars. In this case, we only consider the total queue length, and after the computation, the output of the controller is multiplied with the weight w_i which depends on the number of flows [18].

3. H_2 CONTROLLER DESIGN

3.1. State augmentation

We consider a state feedback control $v(k) = Fx(k)$ for the system (7)-(8). Then, we have

$$\begin{cases} x(k+1) & = Ax(k) + BFx(k - d_k) + B_1 w(k), \\ z(k) & = Cx(k) + DFx(k - d_k). \end{cases} \tag{10}$$

Since there always exists a round delay on the forward and return paths, we can assume that the delay is at least one unit. Hence, by considering that $d_k \in \{1,2,\cdots,m\}$, we transfer the system (10) into an equivalent system as follows by state augmentation:

$$(\Sigma): \xi(k+1) = \bar{A}_k \xi(k) + \bar{B} w(k),$$
$$z(k) = \bar{C}_k \xi(k),$$

where

$$\xi(k) = \begin{pmatrix} x(k) \\ \xi_1(k) \\ \vdots \\ \xi_m(k) \end{pmatrix},$$

$$\bar{A}_k = \begin{pmatrix} A & 0 & \cdots & \overbrace{B}^{(d_k+1)-th\,block} & \cdots & 0 & 0 \\ F & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & 0 \end{pmatrix},$$

$$\bar{B} = \begin{pmatrix} B_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$\bar{C}_k = \begin{pmatrix} C & 0 & \cdots & 0 & \overbrace{DF}^{(d_k+1)-th\,block} & 0 & \cdots & 0 \end{pmatrix}.$$

Here, B and DF are at the $(d_k+1)-th$ block column of $\bar{A}_k$ and $\bar{C}_k$, respectively, and the dimensions of $\xi(k)$, $\bar{A}_k, \bar{B}, \bar{C}_k$ are $(p+1+m)\times 1$, $(p+1+m)\times(p+1+m)$, $(p+1+m)\times 1$ and $2\times(p+1+m)$, respectively.

Under the state feedback $v(k)=Fx(k)$, for the given time delay d_k, it is well known that the cost of (9) is in fact the square of the H_2 norm of the system (Σ). Hence, the congestion control problem formulated in the last section becomes the problem of designing the state feedback control gain F such that the closed-loop system (Σ) is stable and its H_2 norm is minimized.

If delay d_k is a constant, the system (Σ) will be time-invariant ($\bar{A}_k=\bar{A}$, $\bar{C}_k=\bar{C}$), the H_2 norm square of the system can be computed as [22]

$$\|G(z)\|_2^2 = trace(\bar{B}^T L_o \bar{B}) = trace(\bar{C} L_c \bar{C}^T), \tag{11}$$

where L_c and L_o are the reachability and observability Gramians

$$\bar{A}^T L_o \bar{A} - L_o + \bar{C}^T\bar{C} = 0, \tag{12}$$

$$\bar{A} L_c \bar{A}^T - L_c + \bar{B}\bar{B}^T = 0. \tag{13}$$

However, in ATM networks the delay d_k is generally unknown and time-varying. In this situation, (Σ) is an uncertain system with uncertainty in $\bar{A}_k$ and $\bar{C}_k$. Hence, the congestion control is to design the feedback gain F to minimize an upper bound of the cost (9) or the H_2 norm square of (Σ).

Theorem 1: Given the system (Σ) with time-varying $\bar{A}_k$ and $\bar{C}_k$, the following hold [5]:

(a) If the system (Σ) is exponentially stable, then

$$\|\Sigma\|_2^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} trace[\bar{C}_k L_c(k)\bar{C}_k^T], \tag{14}$$

where L_c satisfies the following Lyapunov difference equation

$$L_c(k+1) = \bar{A}_k L_c(k)\bar{A}_k^T + \bar{B}\bar{B}^T. \tag{15}$$

(b) If there exist bounded matrices P, Q, W and a scalar γ such that

$$\begin{bmatrix} P-\bar{B}\bar{B}^T & \tilde{A}_i \\ \tilde{A}_i^T & Q \end{bmatrix} > 0, \tag{16}$$

$$\begin{bmatrix} W & \tilde{C}_i \\ \tilde{C}_i^T & Q \end{bmatrix} > 0, \quad i=1,2,\cdots,m, \tag{17}$$

$$trace(W) < \gamma^2, \tag{18}$$

$$PQ = I, \tag{19}$$

where

$$\tilde{A}_i = \begin{pmatrix} A & 0 & \cdots & \overbrace{B}^{(i+1)-thblock} & \cdots & 0 & 0 \\ F & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & 0 \end{pmatrix}, \tag{20}$$

$$\tilde{C}_i = \begin{pmatrix} C & 0 & \cdots & 0 & \overbrace{DF}^{(i+1)-thblock} & 0 & \cdots & 0 \end{pmatrix}, \tag{21}$$

then the system (Σ) is exponentially stable and

$$\|\Sigma\|_2 < \gamma. \tag{22}$$

In this situation, a suitable state feedback controller is

$$v(k) = Fx(k). \tag{23}$$

Proof: (a) In view of [5], define $L_c(k) \triangleq E[\xi(k)\xi^T(k)]$.

$$\begin{aligned} L_c(k+1) &= E[\xi(k+1)\xi^T(k+1)] \\ &= E\{[\bar{A}_k\xi(k)+\bar{B}w(k)][\bar{A}_k\xi(k)+\bar{B}w(k)]^T\} \\ &= E[\bar{A}_k\xi(k)\xi^T(k)\bar{A}_k^T] + \bar{B}\bar{B}^T \\ &= \bar{A}_k L_c(k)\bar{A}_k^T + \bar{B}\bar{B}^T \end{aligned} \tag{24}$$

From the definition of the H_2 norm,

$$\begin{aligned} \|\Sigma\|_2^2 &= \lim_{N\to\infty} E\left(\frac{1}{N}\sum_{k=1}^{N} z^T(k)z(k)\right) \\ &= \lim_{N\to\infty} E\frac{1}{N}\sum_{k=1}^{N} trace(z(k)z^T(k)) \\ &= \lim_{N\to\infty} E\frac{1}{N}\sum_{k=1}^{N} trace(\bar{C}_k\xi(k)\xi^T(k)\bar{C}_k^T) \\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} trace(\bar{C}_k L_c(k)\bar{C}_k^T). \end{aligned} \tag{25}$$

(b) It can be easily known that if there exist constant matrices P and W such that

$$\begin{pmatrix} P-\bar{B}\bar{B}^T & \bar{A}_k P \\ P\bar{A}_k^T & P \end{pmatrix} > 0, \tag{26}$$

$$\begin{pmatrix} W & \bar{C}_k P \\ P\bar{C}_k^T & P \end{pmatrix} > 0, \quad k=1,2,\cdots, \tag{27}$$

then $L_c(k) < P$ and

$$\|\Sigma\|_2^2 < \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} trace(W). \quad (28)$$

Moreover, $\bar{A}(k)\in\{\tilde{A}_i\},\ \bar{C}(k)\in\{\tilde{C}_i\},\ i=1,2,\cdots,m.$ From LMIs (16) and (17), we have that for all k,

$$\begin{pmatrix} P-\bar{B}\bar{B}^T & \bar{A}_kP \\ P\bar{A}_k^T & P \end{pmatrix} > 0, \quad (29)$$

$$\begin{pmatrix} W & \bar{C}_kP \\ P\bar{C}_k^T & P \end{pmatrix} > 0. \quad (30)$$

Then

$$\begin{aligned}\|\Sigma\|_2^2 &= \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} trace\left(\bar{C}_kL_c(k)\bar{C}_k^T\right)\\ &\le \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} trace\left(\bar{C}_kP\bar{C}_k^T\right)\\ &\le trace(W).\end{aligned}$$

□

Remark 4: In the multi-source case, if we use the controllers $v_j(k)=f_jFx(k),\ j=1,2,\cdots,M$, where M is the number of sources, F is the same as in Theorem 1 and $\sum_{j=1}^{M} f_j = 1$, $f_j\ge 0,\ j=1,2,\cdots,M$, then the system

$$x(k+1) = Ax(k) + B\sum_{j=1}^{M} v_j(k-d_k^j) \quad (31)$$

can be rewritten as

$$x(k+1) = Ax(k) + B\sum_{j=1}^{M} f_jFx(k-d_k^j), \quad (32)$$

where d_k^j is the round trip delay of source j. After state augmentation, we get

$$\begin{bmatrix} P-\bar{B}\bar{B}^T & \bar{A}_k \\ \bar{A}_k^T & Q \end{bmatrix} = \sum_{j=1}^{M} f_j \begin{bmatrix} P-\bar{B}\bar{B}^T & \tilde{A}_k^j \\ \tilde{A}_k^{jT} & Q \end{bmatrix} > 0, \quad (33)$$

$$\begin{bmatrix} W & \bar{C}_k \\ \bar{C}_k^T & Q \end{bmatrix} = \sum_{j=1}^{M} f_j \begin{bmatrix} W & \tilde{C}_k^j \\ \tilde{C}_k^{jT} & Q \end{bmatrix} > 0. \quad (34)$$

Hence, we can conclude that in the multi-source case, state feedback controller $v_j(k)=f_jFx(k),\ j=1,2,\cdots,M$, can still guarantee the H_2 performance for the system. Note that $f_j,\ j=1,2,\cdots,M$, can be treated as the weights of the sources and in the later simulation we take $f_j = 1/M,\ j=1,2,\cdots,M$.

Note, however, that (16)-(19) do not constitute a convex optimization due to the equality constraint (19) which is bilinear in P and Q. Note that although no global solution is available for the bilinear problem, some useful algorithms have been proposed; see e.g. [9,14]. In [9], the above constrained non-convex optimization is formulated as:

$$\min\ trace(QP)\ \ subject\ to \quad (35)$$

$$LMIs\ (16)-(18)\ and \quad (36)$$

$$\begin{pmatrix} P & I \\ I & Q \end{pmatrix} \ge 0. \quad (37)$$

The above optimization solves the bilinear problem (16)-(19) if and only if the minimum solution of $trace(QP)$ is $p+1+m$.

Note from (20)-(21) that F is affine in $\tilde{A}_i$ and $\tilde{C}_i$.

To find an optimal H_2 controller, minimization of γ^2 is to be carried out. From the description above, we need to minimize both $trace(W)$ and $trace(QP)$. By extending the sequential linear programming matrix method (SLPMM) in [14], we propose the following procedure for the optimization.

Algorithm 1:

1. Given a positive scalar γ, solve the semidefinite programming problem (SDP) (16)-(18) and (37) for an initial (P^0,Q^0,W^0,F^0) and set $k=0$. Note that if the SDP does not admit a solution, increase γ until a feasible solution subject to LMIs (16)-(18) exists.
2. Solve the following LMI problem for the variables P,Q,W,F:

   $\min trace(P^kQ + PQ^k)$

   $subject\ to\ LMIs\ (16)-(18)$

   Set $P^{k+1}=P,\ Q^{k+1}=Q$.
3. If a stopping criterion is satisfied, say $k=k_{max}$, here k_{max} is the largest number of iterations we choose, then exit. Otherwise, set $k=k+1$ and go to step 2.

The above process can be repeated for a smaller γ in order to optimize the H_2 performance.

Alternatively, we can set the objective function to be minimized as:

$$\min_{(P,Q,W,F)} \gamma, s.t. (16)-(19). \quad (38)$$

We can define multi-objective programming problem:

$$\min_{(P,Q,W,F)} trace(PQ) + trace(W),$$
$$s.t. (16),\ (17)\ and\ (37)$$

and apply the following algorithm.

Algorithm 2 (SLPMM[14])**:**

1. Determine (P^0, Q^0, W^0, F^0) satisfies (16), (17) and (37).
   For $k = 1, 2, ...$ do
2. Determine (U^k, V^k, Z^k, H^k) as the solution of
   $$\min_{(P,Q,W,F)} trace(PQ^k + P^kQ) + trace(W),$$
   $$s.t. (16),\ (17)\ and\ (37)$$
3. If $trace(U^kQ^k + P^kV^k) + trace(Z^k) = 2trace(P^kQ^k) + trace(W^k)$, then stop.
4. Compute $\beta \in [0,1]$ by solving
   $$\min_{\beta\in[0,1]} trace[(P^k + \beta(U^k - P^k))(Q^k + \beta(V^k - Q^k)) + (W^k + \beta(Z^k - W^k))].$$
5. Set $P^{k+1} = (1-\beta)P^k + \beta U^k$, $Q^{k+1} = (1-\beta)Q^k + \beta V^k$, $W^{k+1} = (1-\beta)W^k + \beta Z^k$, $F^{k+1} = H^k$.

In the above, we have discussed the design of a linear time-invariant controller which is independent of delays. Note that in the case when the controller has an access to the delays, i.e. the service rate is time-stamped, we can have the following less conservative design.

Theorem 2: If the delay d_k is known at time instant k and there exist matrices P_i, Q_i, W_i $(i = 1, 2, \cdots, m)$ and scalar γ such that

$$\begin{bmatrix} P_j - \bar{B}\bar{B}^T & \tilde{A}_i \\ \tilde{A}_i^T & Q_i \end{bmatrix} > 0, \tag{39}$$

$$\begin{bmatrix} W_i & \tilde{C}_i \\ \tilde{C}_i^T & Q_i \end{bmatrix} > 0, \tag{40}$$

$$trace(W_i) < \gamma^2, \tag{41}$$

$$P_iQ_i = I, \quad i, j = 1, 2, \cdots, m, \tag{42}$$

where $\tilde{A}_i, \tilde{C}_i$ are respectively defined in (20) and (21), then the system (Σ) is exponentially stable and

$$\|\Sigma\|_2 < \gamma. \tag{43}$$

In this situation, a suitable controller can be

$$v(k) = F_i x(k). \tag{44}$$

Proof: In view of (42), it follows from (39)-(40) that

$$\begin{pmatrix} P_j - \bar{B}\bar{B}^T & \tilde{A}_iP_i \\ P_i\tilde{A}_i^T & P_i \end{pmatrix} > 0, \tag{45}$$

$$\begin{pmatrix} W_i & \tilde{C}_iP_i \\ P_i\tilde{C}_i^T & P_i \end{pmatrix} > 0, \quad i = 1, 2, \cdots, m. \tag{46}$$

Hence, for any k, there exist bounded matrix sequences $P(k)$, $W(k)$ such that

$$\begin{bmatrix} P(k+1) - \bar{B}\bar{B}^T & \bar{A}_kP(k) \\ P(k)\bar{A}_k^T & P(k) \end{bmatrix} > 0, \tag{47}$$

$$\begin{bmatrix} W(k) & \bar{C}_kP(k) \\ P(k)\bar{C}_k^T & P(k) \end{bmatrix} > 0, \tag{48}$$

where $\bar{A}_k = \tilde{A}_{d_k} \in \{\tilde{A}_i\}$, $\bar{C}_k = \tilde{C}_{d_k} \in \{\tilde{C}_i\}$ and $P(k) = P_{d_k}$, $d_k \in \{1, 2, \cdots, m\}$. According to [5], we know that the system Σ is exponentially stable and

$$\|\Sigma\|_2^2 < \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} trace[W(k)] < \gamma^2.$$

Remark 5: Theorem 2 is obviously less conservative than Theorem 1 as when setting $P_i = P$ and $W_i = W$, $i = 1, 2, \cdots, m$, the former reduces to the latter.

Note that [18] addresses the design of congestion control to ensure the stability of the system. In Theorems 1 and 2 above, we present an approach of designing a congestion control that achieves the best possible average performance in terms of the H_2 norm minimization. In the present paper, a model of effective service rate is used, which can give rise to better performance of congestion control. It is worth highlighting that since the sampling period of controller is larger than the RM cell spacing, there is sufficient time to perform the required computation [18].

On the other hand, saturation exists in sources and switch because of limited available bandwidth and queue buffer. Saturation can affect the efficiency of congestion control, so it should be taken into consideration. However, since the system we are dealing with is stochastic with white Gaussian noise inputs, it makes sense to discuss the saturation problem in terms of probability. It is well known that for a Gaussion random variable η with zero mean and variance σ^2,

$$Pr(\eta < 3\sigma) = 0.99865,$$

where Pr means possibility.

Assume that the upper bound of source rate is $\bar{r}$, that is

$$r_k = sat(u_{k-d_k}), \tag{49}$$

$$sat(u_{k-d_k}) = \begin{cases} 0, & if \ \ u_{k-d_k} < 0 \\ u_{k-d_k}, & if \ \ 0 \le u_{k-d_k} \le \bar{r} \\ \bar{r}, & if \ \ \bar{r} < u_{k-d_k}. \end{cases} \tag{50}$$

Observe that

$$v(k) = u_k - \mu, \tag{51}$$

$$v(k) = Fx(k) = F\begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix}\xi(k), \tag{52}$$

where $\xi(k)$ is the state of the system (Σ) whose covariance $L_c(k) = E[\xi(k)\xi^T(k)]$ satisfies $L_c(k) < P$. Hence, if

$$F\begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} P \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} F^T \le (\frac{\bar{r}-\mu}{3})^2, \tag{53}$$

then the probability that

$$u_{k-d_k} \le \bar{r} \tag{54}$$

is 0.99865.

Note that (53) can also expressed as

$$\begin{pmatrix} (\bar{r}-\mu)^2/9 & F\begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} \\ \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} F^T & Q \end{pmatrix} \ge 0, \tag{55}$$

where $PQ = I$.

Similarly, assume that the upper bound of queue buffer is $\bar{q}$. We have that

$$q_k - q_d = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} x(k) \tag{56}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}\begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix}\xi(k)$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}\xi(k). \tag{57}$$

Then, if

$$\begin{pmatrix} (\bar{q}-q_d)^2/9 & \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & Q \end{pmatrix} \ge 0, \tag{58}$$

then the probability that

$$q_k < \bar{q} \tag{59}$$

is 0.99865.

4. EXAMPLE

Consider a congested switch with only one source. We adopt the similar parameters as given by [18].

The bandwidth available for ABR traffic $b_0 = 1500$ cells/s.
The maximum rate $R_{1,max} = 2b_0 = 3000$ cells/s.
The buffer length $\bar{q} = 10000$ cells/s.
The buffer set point $q_0 = (1/2)y_{max} = 5000$ cells.
The controller cycle time $T = 1$ ms.
The maximum delay on the path m = 10 ms (including delay in the forward and backward paths).

Here, the delay is unknown and time-varying. We assume that the link capacity is a 2nd order auto-regressive (AR) process with parameters $l_1 = l_2 = 0.4$ and the driving zero-mean Gaussian white noise process has a variance equal to 1 [18]. By using the LMI toolbox of the MATLAB package, the following feedback controller is obtained with $\gamma = 3.25$:

$$F = [-0.0114 \quad 0.1210 \quad 0.1379].$$

Fig. 1 shows the buffer occupancy trajectory for time-variant return path delays between 0 and 9 ms, while keeping VBR delays in the forward path fixed at 1 ms. The desired equilibrium at $q_d = 5000$ cells is clearly maintained.

Further, we will consider a more realistic example, with 100 sources feeding into the same congested switch and the system starts at the equilibrium. We use the same parameters as in the previous example, and fix the delay in the forward path to be 1 ms. Since

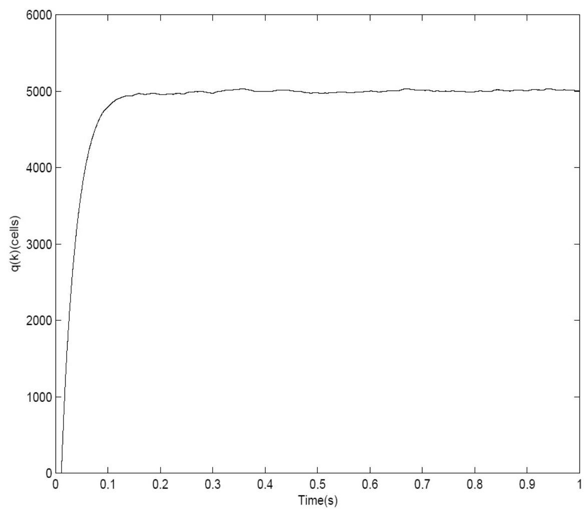

Fig. 1. Performance of buffer-level vs time.

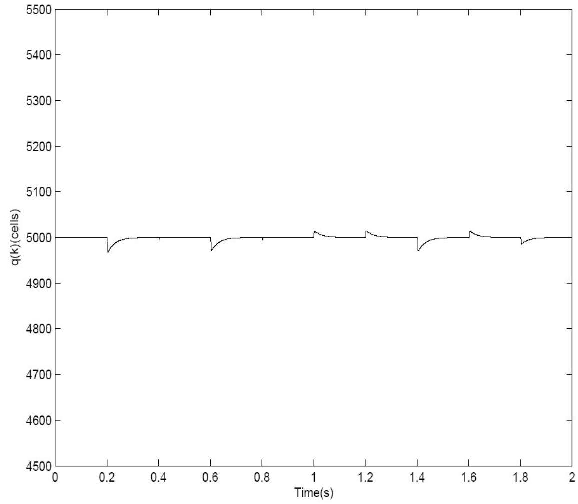

Fig. 2. Performance of buffer-level vs time.

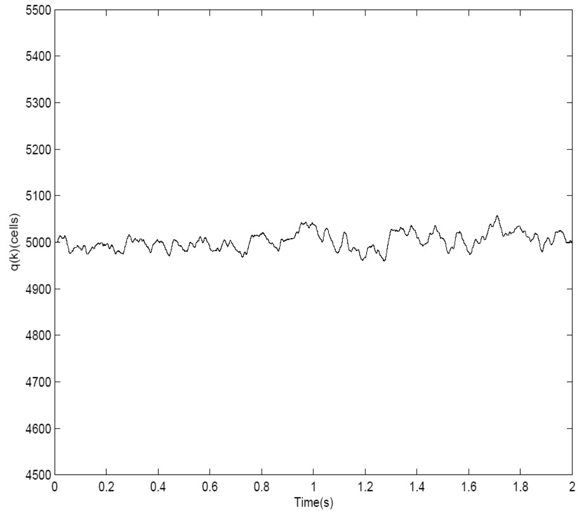

Fig. 3. Performance of buffer-level vs time.

the same controller that stabilizes one source stabilizes multiple sources, the same controller $F=[-0.0114 \quad 0.1210 \quad 0.1379]$ will be used. Ten additional sources join sequentially at 200-ms intervals. The results of the simulation are shown in Figs. 2 and 3. Fig. 2 shows the result under the condition of constant bandwidth. As each source joins the switch, a small glitch can be observed. This is the effect of delays in updating the weights. When a new source is connected, the switch will compute new weights and send them to the sources. However, the updated weights and the reaction of the sources to those updated weights are delayed and therefore, for a brief period of time, the sum of the weights may not be equal to one. After all sources are updated, the control scheme takes effect and brings the buffer to the desired set point. After considering the random nature of the bandwidth available, that is

$$\zeta_{k+1} = 0.4\zeta_k + 0.4\zeta_{k-1} + w_k ,$$

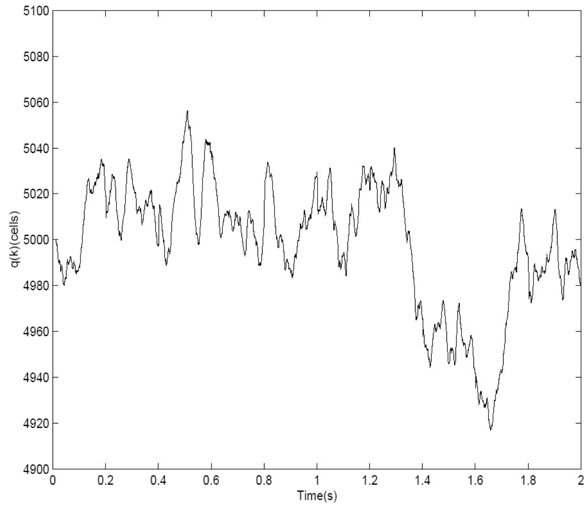

Fig. 4. Performance of buffer-level vs time.

where w_k is white gaussian noise, the result is shown in Fig. 3.

When buffer length is near to buffer set point, saturation may happen during transmission. We assume the buffer length $\overline{q} = 5050 cells/s$ and we want the probability of the queue length saturation is below 0.00135 . By incorporating (58) in the optimization, the optimal state feedback gain is

$$F = \begin{bmatrix} -0.0117 & 0.1198 & 0.1382 \end{bmatrix}.$$

We carry out 50 simulations and compute the total number of instants where queue buffer is saturated. Fig. 4 is the result of one simulation, solid line is the result of the design where the upper bound of queue length is taken into consideration. The result shows that the number of saturation for the case when buffer saturation is not considered in congestion control design is 1926. The number is reduced to 1758 when the buffer saturation limit is considered and the percentage of unsaturated control inputs

$$\frac{number\ of\ unsaturated\ control\ inputs}{total\ number\ of\ time\ steps} = 98.242\%$$

is close to the theoretical result 99.865%.

5. CONCLUSIONS

The paper studies the congestion control problem where a time-variant delay model for the ABR option of ATM networks is considered. The problem has been formulated as a robust H_2 control problem for linear systems with delays. State augmentation was applied to convert the delay problem into a robust control problem for systems with parameter uncertainties. An optimization approach for solving bilinear matrix inequalities was proposed to obtain an optimal H_2 congestion controller. Simulations were given to demonstrate the effectiveness of the controller.

REFERENCES

[1] *ATM Forum Traffic Management Specifications* v. 4.0, April 1996.

[2] E. Altman, T. Basar, and R. Srikant "Congestion control as a stochastic control problem with action delays," *Proc. of 34th IEEE Conference on Decision and Control*, New Orleans, LA., pp. 1389-1394, 1995.

[3] E. Altman and T. Basar "Optimal rate control for high speed telecommunication networks," *Automatica*, vol. 35, pp. 1937-1950, 1999.

[4] E. Altman, T. Basar, and R. Srikant "Robust rate control for ABR sources," *INFOCOM '98, Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies, Proceedings. IEEE*, vol. 1, pp. 166-173, 1998.

[5] K. A. Barbosa and C. E. de Souza "Robust H_2 filtering for discrete-time uncertain linear systems using parameter-depedent Lyapunov functions," *Proc. of 41st IEEE Conf. Decision and Control*, pp. 3224-3229, 2002.

[6] L. Benmohamed and S. M. Meerkov "Feedback control of congestion in packet switching network: The case of a single congested node," *Proc. of IEEE/ACM Trans, Networking*, vol. 1, pp. 693-708, December 1993.

[7] E. K. Boukas and Z. K. Liu, "Robust H_∞ control of discrete-time Markovian jump linear systems with mode-dependent time-delays," *IEEE Trans. on Automatic Control*, vol. 46, pp. 1918-1924, 2001.

[8] H. M. Choi, K. E. Nygard, and R. J. Vetter "Counter-based congestion avoidance scheme for ABR traffic in ATM networks," *Proc. of IEEE Int. Perforamnce, Computing and Communication Conf.*, pp. 284-290, February 1998.

[9] L. El Ghaoui, F. Oustry and M. Ait Rami "A cone complementary linearization algorithms for static output-deedback and related problems," *IEEE Trans. on Automatic Control*, vol. 42, pp. 1171-1176, 1997.

[10] G. Kesidis, *ATM Network Performance*, 2nd ed., Kluwer Academic Publishers, 1999.

[11] H. H. Kim and H. B. Park, "H_∞ control state feedback control for generalized continuous/ discrete time-delay system," *Automatica*, vol. 35, pp. 1443-1451, 1999.

[12] J. Hu, J. Lin, and L. Xie "Robust congestion control for high speed data networks with uncertain time-variant delays: An LMI control approach," *Proc. of the 30th IEEE Conf. on Local Computer Networks*, Sydney, Australia, pp. 474-475, November 2005.

[13] O. C. Imer and T. Basar "Optimal solution to a team problem with information delays: An application in flow control for communication networks," *Proc. of the 38th Conference on Decision and Control*, pp. 2697-2702, 1999.

[14] F. Leibfritz "An LMI-Based algorithm for designing suboptimal static H_2/H_∞ output feedback controllers," *SIAM J. Control Optim.*, vol. 39, no. 6, pp. 1711-1735, 2001.

[15] S. Mascolo "Smith's principle for congestion control in high speed ATM networks," *Proc. of the 36th Conference on Decision and Control*, pp. 4595-4600, 1997.

[16] C. E. Rohrs, R. A. Berry, and S. J. O'Halek "Control engineer's look at ATM congestion avoidance," *Comput. Commun.*, vol. 19, pp. 226-234, March 1999.

[17] C. Scherer, P. Gahinet, and M. Chilali "Multiobjective output-feedback control via LMI optimization," *IEEE Trans. on Automatic Control*, vol. 42, no. 7, pp. 896-911, 1997.

[18] M. L. Sichitiu, P. H. Bauer, and K. Premaratne "The effect of uncertain time-variant delays in ATM networks with explicit rate feedback: A control theoretic approach," *IEEE/ACM Trans. on Networking*, vol. 11, no. 4, pp. 628-637, 2003.

[19] L. Xie, "Output feedback H_∞ control of systems with parameter uncertainty," *Int. J. Control*, vol. 63, no. 4, pp. 741-750, 1996.

[20] L. Xie, Y. C. Soh, and C. Du "Robust H_2 estimation and control," *Proc. of Int. Conf. on Control, Automation, Robotics and Vision*, Singapore, December 2002.

[21] S. Xu and T. Chen, "Robust H_∞ control for uncertain discrete-time systems with time-varying delays via exponential output feedback controllers," *Systems and Control Letters*, vol. 51, pp. 171-183, 2004.

[22] K. Zhou, J. Doyle, and K. Golver, *Robust and Optimal Control*, Prentice-Hall, 1995.

Jun Lin was born in Fujian, China. He received the B.S degree in automatic control from University of Science and Technology of China, Hefei, China, in 2002. He is currently pursuing the Ph.D. degree at Nanyang Technological University, Singapore. His research interests include time delay systems, network congestion control and robust control.

Lihua Xie received the B.E. and M.E. degrees in electrical engineering from Nanjing University of Science and Technology in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Newcastle, Australia, in 1992.

Dr. Xie is currently a Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

He held teaching appointments in the Department of Automatic Control, Nanjing University of Science and Technology from 1986 to 1989. He also held visiting appointments with the University of Melbourne. His current research interests include robust control, networked control systems, time delay systems, and control of disk drive systems and smart structures. In these areas, he has published several papers and co-authored (with C. Du) the monograph *H-infinity Control and Filtering of Two-dimensional Systems* (Springer, 2002).

Dr. Xie is currently an associate editor of the IEEE Transactions on Automatic Control, International Journal of Control, Automation and Systems, and Journal of Control Theory and Applications. He is also a member of the Editorial Board of IEE Proceedings on Control Theory and Applications. He served as an associate editor of the Conference Editorial Board, IEEE Control Systems Society from 2000 to 2004.

Huanshui Zhang graduated in mathematics from the Qufu Normal University in 1986 and received his MSc and PhD degrees in control theory and signal processing from the Heilongjiang University, P.R. China, and Northeastern University, P.R. China, in 1991 and 1997, respectively. He worked as a postdoctoral fellow at the Nanyang Technological University from 1998 to 2001 and Research Fellow at Hong Kong Polytechnic University from 2001 to 2003.

He joined Shandong Taishan College in 1986 as an Assistant Professor and became an Associate Professor in 1994. He joined Shandong University in 1999 as a Professor. Currently he is a professor of Shenzhen Graduate School of Harbin Institute of Technology.

His interests include optimal estimation, robust filtering and control, time delay systems, singular systems, wireless communication and signal processing.