

Queue Control under Time-Variant Delays: A Discrete Time System Approach

Peter H. Bauer *and Mihail L. Sichitiu
Department of Electrical Engineering
University of Notre Dame,
Notre Dame, IN 46556
msichiti@bach.helios.nd.edu
pbauer@mars.ee.nd.edu

Kamal Premaratne *
Department of Electrical and
Computer Engineering
University of Miami
Coral Gables, FL 33124
kamal@miami.edu

Abstract

This paper introduces a discrete time model for time-variant delays and investigates the very nature of such a delay. It is shown that a linear system-delay interface is a system theoretic necessity for the construction of composite linear systems with time-variant delays. Based on this analysis, two interfaces of particular importance are presented and used to obtain new, simple to check stability results for queue control systems. The relevance of the presented modeling and stability results on queue control systems to QoS control in modern communication networks is illustrated via several examples.

1 Introduction

Due to the rapidly expanding field of networking, new applications for high speed networks in various disciplines are arising. Especially the areas of distributed control, integrated control-communication systems, teleoperation, variable bit-rate control, etc. all rely on the newly arising network technology to communicate control and measurement signals in real time. A large number of papers have focused on the problem of closing the feedback loop through communication networks (see for example [1]-[15]) and fundamental issues such as modeling, stability, performance, fault tolerance, synchronization, etc. have been addressed. The available literature focusing on fundamental aspects of this work can be viewed as belonging to one of two categories: Queue Control (control of the network itself) [1]-[5] and control of an external system using network connections [7]-[15]. In the second category, much of the available work was dedicated to integrated communication control systems, which basically use a LAN like Ethernet to connect the individual systems.

*This work was supported by NSF grants ANI 9726253, ANI 9726247

In both categories the network connection was modeled in the most simple case by a fixed and known delay and in the most elaborate models by uncertain time-variant delays [7]-[10],[14],[15]. However all the models introduced were tailored to the particular application under consideration. Also non-ideal effects such as packet loss in the network, bounds on queue size and bit-rates, quantization effects, etc. have not been jointly modeled in the context of a system theoretic analysis of the arising feedback system.

In this paper, we will take a fresh, new look at modeling feedback loops with network connections, including the effects of time-variant uncertain delays, packet loss and rate as well as buffer occupancy nonlinearities on the communication link. This will be done in section 2.1 by at first modeling the nature of a time-variant delay itself, i.e. outside the framework of any particular application. This is then followed in section 2.2 by the definition of interfaces between the network and the linear system, which arises as a system theoretic necessity rather than a hardware necessity. We will provide interfaces that model the connection between the link and the linear system for several major classes of data flow. These models will be sufficiently general to include both categories of control systems mentioned previously (queue control using variable bit-rates and general controller-plant pairs with network connections.) In section 2.3, properties of the developed network connection models are derived. Finally, in section 2.4 we provide complete models for the case of a queue control system, i.e. for the case of an integrator plant. This is an entirely new approach to modeling time-variant links and the arising feedback system and has a number of unique advantages as will be pointed out later.

In the second part of the paper (section 3) we will analyze the resulting models of queue control systems in terms of stability behavior. At first we will prove the existence and nonexistence of equilibria for the two different categories of feedback systems. In a second step we will derive conditions for the equilibria to be stable, if such equilibria exist.

2 Models

We can distinguish between two distinct components a time-variant delay connecting linear systems: the first one is the delay introduced by the communication link, (i.e. the nature of the delay itself) while the second one is the interface we use to connect the network to the receiving/sending system. Let us analyze the two components separately. As it will be seen in the following sections, this interface arises as a system theoretic necessity rather than being motivated by hardware.

2.1 The nature of time-variant communication delays in discrete time

In Figure 1 the first component of a time-variant delay communication link is shown : the delay $d(n)$. In what follows we will assume for simplicity that at the network ends we have single input single output (SISO) systems, i.e. the signals at the the network ends are scalar. The input to the time-variant delay $d(n)$ is a sequence of input sets $u(n)$ whereas the

output is given by a sequence of output sets $v(n)$. The sets $u(n)$ and $v(n)$ describe all input/output signals (scalar or otherwise) that are entering or leaving the block $d(n)$ within a sampling period, i.e., between time instant $n-1$ and time instant n . The output signal is a sequence of sets satisfying the following relation:

$$v(m) = \bigcup_{n+d(n)=m} u(n) \quad (1)$$

i.e. the set $v(m)$ is the union of all sets $u(n)$ which satisfy $n+d(n) = m$ with $n \geq 0$, $d(n) \geq 0$.

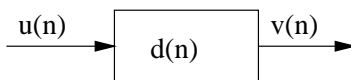


Figure 1: A discrete time-variant delay

We will use the following notations:

- $M(n)$ is the cardinality of the input $u(n)$ at time n .
- $N(n)$ is the cardinality of the output $v(n)$ at time instant n .
- $d(n)$ is a function $d : \mathcal{N} \rightarrow \mathcal{N}$ (where \mathcal{N} is the set of positive integers) satisfying the inequality:

$$0 \leq \underline{d} \leq d(n) \leq \bar{d} < \infty. \quad (2)$$

$d(n)$ represents the delay the communication link introduces at time instant n for the input $u(n)$.

The properties of the delay are completely described by the function $d(n)$. From (1) and the above introduced notation, the function $N(n)$ can be computed from the values of $d(n)$ with the following expression:

$$N(n) = \sum_{i \in U(n)} \text{card}(M(i)) \quad (3)$$

with $U(n)$:

$$U(n) = \{i \mid i + d(i) = n, i \geq 0\} \quad (4)$$

where by $\text{card}(A)$ we denote the cardinality of the set A .

Notice that at some time instances the set $v(n)$ may contain more than one value while at other time instances the set may be empty. This is due to the discrete nature of time in our system. Intuitively, between two ticks of the system clock there might be one packet arriving from the communication link, more than one packet or none.

Equations (1)-(4) represent the most general description of a time-variant discrete time delay. By nature these delay inputs/outputs must be sets (due to the finite and discrete time resolution). This model does not make any assumption on how the delay is interfaced with the linear

system (input/output buffer, register, FIFO queues, etc.) and describes solely the nature of the delay itself.

For the case of multi-input multi-output systems we can have two possible implementations:

- if the components of a multi-output signal are individually sent in different packets, then we will can treat each component as a scalar signal and use the model presented above.
- if all the components of a multi-output signal are packed and sent together then we can use the same approach with the difference that instead being sets of scalars $u(n)$ and $v(n)$ will be sets of vectors.

2.2 Interfacing linear systems with time-variant delays

As we saw in the previous section the output signal set of a time-variant delay is of varying cardinality thus making it impossible to represent it as a scalar time function. In order to “connect” the communication link to a linear time-invariant (LTI) system we must define an interface that would make the output signal proper, i.e. of constant dimension. For reasons of notational simplicity we choose scalar signals, i.e. SISO systems are assumed. Depending on the type of application different interfaces must be defined. Essetially, an interface defines an operation on the output set $v(m)$ that returns only a single scalar value.

2.2.1 The Hold Freshest Sample Interface

In the case of a regular plant/controller pair which exchanges samples through a packet switched network, a common solution for the case when a new sample did not arrive on time is to use the “hold freshest sample” interface (see Figure 2). This interface holds the most recently received sample both at the plant and controller input until a more recent sample arrives. If more than one sample arrives in any one sampling interval, the most recent sample is chosen. If an old sample arrives after a newer one was received it is simply discarded, being considered outdated. The latter might happen if the packets can arrive out of order (for example on a network using UDP/IP with dynamical routing). Of course, an implementation of this interface relies on the existence of a time-stamp that will indicate the “age” of the samples. (If the communication channel is a FIFO structure, then a time stamp is not necessary to implement an HFS interface.)

The time-variant delay/HFS interface combination was first introduced in [6] as the so called “input variable delay”. In this work the input variable delay was viewed as a linear time-variant system rather than a time-variant delay/interface combination.

We will provide here a formal description of the HFS interface:

At any time instant n , denoting with $f(n)$ the time-stamp of the freshest sample received by time instant n and using the notations defined in equations (3),(4) we can have one of the following three cases:

- (a) $N(n) = 0$

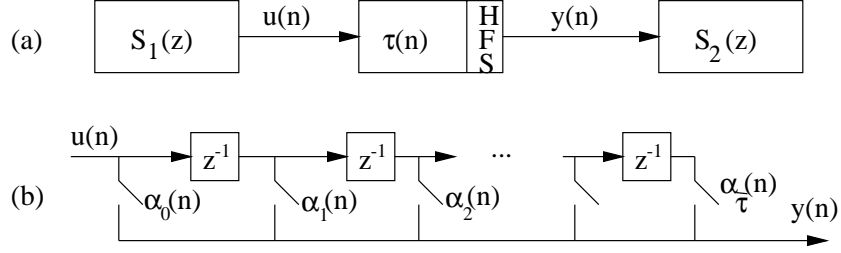


Figure 2: A time-variant delay with a hold freshest sample interface. (2a) Block diagram of the HFS delay and interface with linear system. (2b) Detailed block diagram using a cascade of unit delays.

that is there is no signal at the output of the time-variant delay block. In this case we simply hold the previous sample, thus we have:

$$y(n) = y(n-1)$$

(b) $\max(U(n)) < f(n-1)$

that is all the signals that arrived at time n are older than the freshest sample at time $n-1$. Also in this situation we discard the newly arrived (but older) signals and we hold the previous sample:

$$y(n) = y(n-1)$$

(This situation cannot occur in FIFO structures.)

(c) $\max(U(n)) > f(n-1)$

that is at least one more recent sample arrived at time instant n and we choose the most recent one:

$$y(n) = u(f(n))$$

where we update $f(n)$ by:

$$f(n) = \max(U(n))$$

We can combine the three cases presented above by writing:

$$\begin{aligned} y(n) &= u(f(n)) \quad \text{where} \\ f(n) &= \max\{U(n) \cup \{f(n-1)\}\}. \end{aligned}$$

Thus we obtain a scalar output from a scalar input, i.e. we avoid output signal sets at the delay block. Furthermore we can write:

$$y(n) = u(n - \tau(n)) \tag{5}$$

with

$$\tau(n) = n - f(n). \tag{6}$$

i.e. $y(n)$ is now defined for all n . Observe that without defining a proper interface between the time-variant delay and the linear system, equation (5) does not describe the time-variant delay, since there might be time instances n for which no $y(n)$ exists!

With the help of equations (5),(6), one can now analyze the resulting system [7], [14], [15] using the available linear systems tools.

Notice that if $d(n)$ is bounded as in equation (2) then $\tau(n)$ is also bounded:

$$0 \leq \underline{\tau} \leq \tau(n) \leq \bar{\tau} < \infty \quad (7)$$

Using notation similar to the one in [10] and using Figure 2b, we can write equation (5) as:

$$y(n) = \alpha_{\underline{\tau}}(n)u(n - \underline{\tau}) + \alpha_{\underline{\tau}+1}(n)u(n - \underline{\tau} - 1) + \dots + \alpha_{\bar{\tau}}(n)u(n - \bar{\tau})$$

where

$$\alpha_i(n) = \begin{cases} 1 & \text{if } \tau(n) = i \\ 0 & \text{otherwise} \end{cases} \quad i = \underline{\tau}, \dots, \bar{\tau}, \quad \forall n \geq 0 \quad (8)$$

Furthermore $\sum_{i=\underline{\tau}}^{\bar{\tau}} \alpha_i(n) = 1 \quad \forall n \geq 0$.

The corresponding state space model is given by the system matrices:

$$A = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}}_{\bar{\tau}} \quad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (9)$$

$$C(n) = (\alpha_{\bar{\tau}}(n) \dots \alpha_1(n)); \quad D(n) = (\alpha_0(n));$$

where if $\underline{\tau} > 0$ we have $\alpha_j(n) = 0 \quad \forall j = 0 \dots \underline{\tau} - 1 \quad \forall n \geq 0$.

Notice that by HFS definition the coefficients $\alpha_i(n)$ cannot vary arbitrarily from one time instant to another: the delay $\tau(n)$ is restricted by $\tau(n+1) \leq \tau(n) + 1$, and hence we have:

$$\alpha_i(n) = 1 \Rightarrow \alpha_j(n+1) = 0 \quad \forall j > i + 1$$

In other words the used sample from the time-variant delay output ages with time, but not faster. Depending on the particular applications (CANs and ICCS, control over ATM networks, etc.) additional constraints on $\tau(n)$ can be formulated.

The delays encountered on the communication link $d(n)$ are in general not equal to the delays perceived at the output of the HFS interface $\tau(n)$. Given the values of $d(n)$ one can compute the values of $\tau(n)$ but not the other way around. The bounds on the values of $d(n)$ and $\tau(n)$ are however equal: considering equations (2) and (7), we have that $\underline{d} = \underline{\tau}$ and $\bar{d} = \bar{\tau}$.

Finally a word on the effect of packet loss: For an HFS time-variant delay interface, packet loss simply introduces an additional delay. In fact it

is simple to show that if p consecutive data samples are lost, the maximum delay bound $\bar{\tau}$ will have to be substituted by $\bar{\tau} + p$. In other words, the HFS interface has to hold the most recent data sample for p more time instances compared with the case where no data loss is allowed.

2.2.2 The Variable Bit Rate Interface

Modeling a variable bit-rate channel requires to quantify the number of bits $u(n)$ sent into the channel at any given time n . Hence, the input $u(n)$ signifies the amount of data sent per discrete time unit rather than being the actual number symbol that is transmitted like in the HFS interface. (A specific example of such a variable bit-rate delay interface would be in congestion control, where buffer occupancy levels need to be controlled by varying the source rates.)

Let us consider the link between the controller and the plant as it is shown in Figure 3. Essentially, the introduced interface uses the addition operators for all members of the set $v(m)$ to produce a single output.

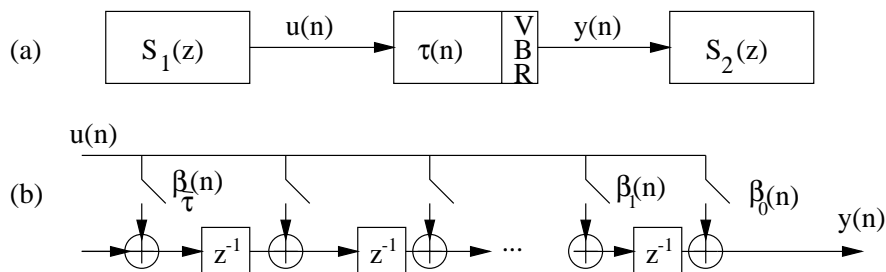


Figure 3: A time-variant delay with a VBR interface. $u(n)$ denotes the amount of data entering the delay during a clock cycle. $y(n)$ is the amount of data exiting the delay. (3a) Time-variant delay - linear system block diagram. (3b) Detailed block diagram using unit delay cascade.

In what follows, we assume at first that there are no packet/cell losses on the communication channel, but we do not require the FIFO property. (If necessary, such a constraint could be easily incorporated in the model.) Also, $S_2(z)$ is a buffer or a queue which can be modeled as a digital integrator.

In [6], the time-variant delay/linear system interface combo shown in Figure (3b) was originally introduced as a linear time-variant SISO system called “output variable delay”.

At any time instant n , using the notations defined in equations (3),(4) we can have one of the following two situations:

(a) $N(n) = 0$

that is there is no signal at the output of the time-variant delay (i.e. no packet was received). In this case, the received number of bits is zero and

$$y(n) = 0$$

(b) $N(n) > 0$

that is at least one signal is exiting the time-variant delay (i.e. at least one packet was received). In this case all the values that arrived during this time period are added, (i.e. representing the total number of received bits) since (by an initial assumption) the communication link cannot lose data:

$$y(n) = \sum_{i \in U(n)} u(i)$$

The corresponding state space model (see Figure 3b) is:

$$A = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}}_{\bar{\tau}} \quad B(n) = \begin{pmatrix} \beta_1(n) \\ \beta_2(n) \\ \beta_3(n) \\ \vdots \\ \beta_{\bar{\tau}}(n) \end{pmatrix} \quad (10)$$

$$C = (1 \ 0 \ \dots \ 0 \ 0) \quad D(n) = (\beta_0(n))$$

where

$$\beta_i(n) = \begin{cases} 1 & \text{if } d(n) = i \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

with $i = 0, \dots, \bar{d}$.

Furthermore the condition $\sum_{i=\underline{d}}^{\bar{d}} \beta_i(n) = 1$ needs to be imposed since each sample encounters exactly one delay value.

Using equation (10) we can also write for the VBR interface:

$$y(n) = \sum_{i=\underline{\tau}}^{\bar{\tau}} \beta_i(n-i) u(n-i) \quad (12)$$

In case the communication channel is a FIFO structure, additional constraints similar to those presented in section 2.2.1 must be imposed: if $\beta_i(n) = 1$, then $\beta_j(n+1) = 0 \ \forall j = \underline{\tau}, \dots, i-1$.

Finally, packet loss can be modeled by a disturbance at $u(n)$ (the input side of the time-variant delay). The disturbance signal will then act like a negative input with a maximum amplitude resulting from the amount of lost bits, i.e. the number of lost bits in the packet.

2.2.3 Maximum Delay

Of course it is well known, that it is always possible to reduce the problem of time-variant delays to a time-invariant problem if we introduce additional delays such that the sum of the delays experienced by a signal on the communication links and the artificial delays is always constant (i.e. time-invariant). Thus if the delays encountered on the communication

links are bounded as in equation (2) we can make the total delay equal to \bar{d} . We will call this way of handling the delays on the communication links the “maximum delay interface”. Similarly to the case of the hold freshest sample interface the implementation of the maximum delay interface is dependent on the existence of a time-stamp on the communication packets. This type of delay interface is commonly used for streaming audio/video over the Internet.

2.3 Properties of time-variant delays with interfaces

In this section we will explore some of the properties of the time-variant delays and their associated interfaces.

2.3.1 Commutativity of TVDs

The question asked in this section is whether or not we can interchange two different time-variant delays and obtain the same equivalent system in both cases.

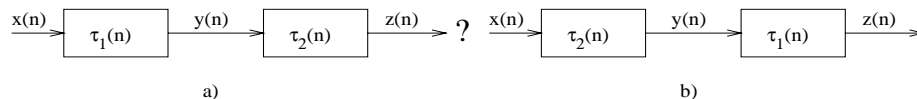


Figure 4: Commutativity of two time-variant delays

With the notations in Figure 4 we have for case (a):

$$\begin{aligned}
 y(k) &= \bigcup_{i \in U(k)} x(i) = \bigcup_{i + \tau_1(i) = k} x(i) \\
 z(n) &= \bigcup_{k + \tau_2(k) = n} y(k) = \bigcup_{i + \tau_1(i) + \tau_2(i + \tau_1(i)) = n} x(i)
 \end{aligned} \tag{13}$$

Using a similar deduction we obtain for case (b):

$$z(n) = \bigcup_{i + \tau_2(i) + \tau_1(i + \tau_2(i)) = n} x(i) \tag{14}$$

Comparing equations (13) and (14) we can state that in general commutativity does not hold. One notable exception is the case when the two delays are constant. The addition of a hold freshest sample interface or of a variable bit-rate interface does not change the situation, i.e. the two time-variant systems remain non-commutative.

2.3.2 Commutativity of a TVD with a LTI System

The question asked in this section is whether or not we can interchange a time-variant delay and a linear time-invariant system and obtain the same equivalent system in both cases.

A positive answer to this question would enable us to simplify a closed loop system involving communication links (see for example Figure 6) by allowing the grouping of the time-variant delays (which can be further concatenated as we will see in section 2.3.3) and by grouping the controller and the plant into one linear time-invariant system.

Unfortunately it can be shown that both for the HFS interface and for the VBR interface the commutativity with an LTI system does not hold in general. The proof is beyond the scope of this paper.

2.3.3 Reducing Two TVDs to One Equivalent TVD

The question answered in this section is: can we reduce two time-variant delays to one equivalent time-variant delay?

An affirmative answer will enable us to model multiple communication links with time-variant delays with only one equivalent communication link with time-variant delays.

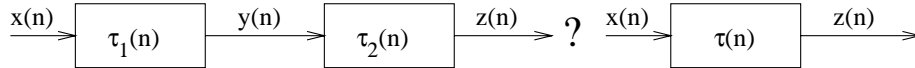


Figure 5: Reducing two delays to one equivalent delay

Following the deduction presented in section 2.3.1 we obtain:

$$z(n) = \bigcup_i x(i) \quad \text{such that } i + \tau_1(i) + \tau_2(i + \tau_1(i)) = n$$

Thus given the values of $\tau_1(i)$ and $\tau_2(i)$ we can construct $\tau(i)$ as:

$$\tau(i) = \tau_1(i) + \tau_2(i + \tau_1(i)) \quad (15)$$

Thus we can reduce the two time-variant delays to one equivalent time-variant delay. Moreover, if we have $\underline{\tau}_1 \leq \tau_1(n) \leq \bar{\tau}_1$ and $\underline{\tau}_2 \leq \tau_2(n) \leq \bar{\tau}_2$ then we have:

$$\underline{\tau}_1 + \underline{\tau}_2 \leq \tau(n) \leq \bar{\tau}_1 + \bar{\tau}_2.$$

Notice that reducing the two pairs of time-variant delays in section 2.3.1 yields two different time-variant delays having the same delay bounds.

2.4 Closed loop models

We will present in this section the closed loop models for a typical queue control system. At first we will present an idealized model of the system which will be further enhanced by introducing the queue saturation

nonlinearity and the possibility of the sources to disconnect at arbitrary times.

2.4.1 An Initial Idealized Model

A congestion control system can usually be decomposed into two components: the rate control and the queue control. The rate control aims to ensure that the sum of the incoming rates is equal to the output rate from the queue. The queue control's objective is to adjust the occupancy level of the queue. The two control components can be decoupled and analyzed separately.

Figure 6 depicts a typical congestion control mechanism where m sources are connected to one common queue. The quantity y_{out} is the constant amount of data by which the queue is depleted at each time instant. (For example y_{out} could be the data transmitted on an outgoing network link or the amount of data processed per time unit.) The quantities $\frac{y_{out}}{m}$ represent the rate control component and we assume that each source receives a fair, equal share of the available rate (i.e. amount of data per time unit). The input $y_0(n)$ is the desired queue occupancy level.

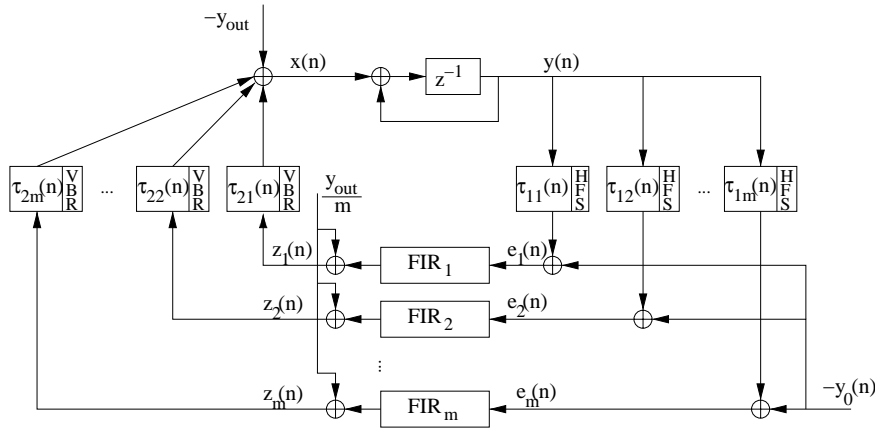


Figure 6: Linear system model

The Queue:

The dynamic behavior of the queue is described by the linear time-invariant difference equation:

$$y(n+1) = y(n) + x(n). \quad (16)$$

The quantity $x(n)$ represents the net input into the buffer and it can be positive or negative.

The Sources:

If general controllers are allowed, the overall system cannot be described using a single difference equation anymore.

For simplicity, we assume that the sources have FIR controllers with real coefficients. The output is composed by the rate control component $\frac{y_{out}}{m}$ and the queue control component - the output of the FIR filter:

$$z_\nu(n) = \frac{y_{out}}{m} + \sum_{i=0}^k c_{i\nu} e_\nu(n-i), \quad \nu = 1, \dots, m. \quad (17)$$

We assume all FIR filters have the same order k which is not a restriction considering that we may have zero coefficients. Also we do not consider the trivial case where all coefficients of an FIR filter are zero.

The Delays on the Return Paths:

On the return path the queue informs the sources about its current occupancy level $y(n)$. The integer delays $\tau_{1\nu}(n)$ describes the sum of all delays the signal $y(n)$ encounters before it arrives at each input of the FIR filters. In this case an HFS interface is appropriate to model the behavior of such a source:

$$e_\nu(n) = y(n - \tau_{1\nu}(n)) - y_0(n), \quad \nu = 1, \dots, m \quad (18)$$

We assume boundedness of the delay, i.e.

$$0 \leq \tau_{1\nu} \leq \tau_{1\nu}(n) \leq \bar{\tau}_1 \quad \nu = 1, \dots, m \quad (19)$$

and HFS restrictions on $\tau_{1,\nu}(n)$ $\nu = 1, \dots, m$.

The Delays on the Forward Paths:

The integer delays $\tau_2(n)$ describe the sum of all delays the data on the ν th forward path $z_\nu(n)$ encounters before it arrives at the queue input. In this case a VBR interface must be used to model the data flow on the communication links:

$$x(n) = \sum_{\nu=1}^m \sum_{j=\tau_2}^{\bar{\tau}_2} \beta_{j\nu}(n-j) z_\nu(n-j) - y_{out} \quad (20)$$

where the coefficients $\beta_{j\nu}$ depend on the delays encountered in the forward paths $\tau_{2\nu}(n)$ as defined in equation (11)

Similarly as in the case of the delays on the return path, boundedness of the delays is assumed:

$$0 \leq \tau_2 \leq \tau_{2\nu}(n) \leq \bar{\tau}_2 \quad \nu = 1, \dots, m \quad (21)$$

and VBR restrictions are imposed on $\tau_{2\nu}(n)$ $\nu = 1, \dots, m$.

Although these additional some restrictions on how the two delays can change if y and z are sent via a data communication network, in our analysis that is to follow, we make no further assumptions on $\tau_{1,\nu}(n)$, $\tau_{1,\nu}(n)$ $\nu = 1, \dots, m$.

The resulting difference equation:

With (16)-(18) and (20) we obtain for the difference equation of the overall system:

$$\sum_{\nu=1}^m s_{\nu}(n) \neq 0 \quad \forall n \geq 0$$

i.e. at least one controller has to be connected.

Also define:

$$sat_Q[x] = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq q_{MAX} \\ q_{MAX} & \text{if } x > q_{MAX} \end{cases}$$

with $q_{MAX} > 0$. The input x to the nonlinearity is the quantity that describes the queue input.

Following the same steps as in section 2.4.1, the resulting difference equation arises:

$$\begin{aligned} y(n+1) = & sat_Q \left[y(n) + \frac{y_{out}}{\sum_{\nu=1}^m s_{\nu}(n)} \sum_{\nu=1}^m s_{\nu}(n) \sum_{j=\tau_2}^{\bar{\tau}_2} \beta_{j\nu}(n-j) + \right. \\ & + \sum_{\nu=1}^m s_{\nu}(n) \sum_{j=\tau_2}^{\bar{\tau}_2} \sum_{i=0}^k \beta_{j\nu}(n-j) c_{i\nu} y(n - T_{i,j,\nu}(n)) \\ & \left. + \sum_{\nu=1}^m s_{\nu}(n) \sum_{j=\tau_2}^{\bar{\tau}_2} \sum_{i=0}^k \beta_{j\nu}(n-j) c_{i\nu} y_0(n-i-j) - y_{out} \right] \quad (24) \end{aligned}$$

3 Equilibria and Stability

In this section we will study the existence of equilibria for the closed loop systems introduced in the previous section. If such an equilibrium exists, we will provide sufficient, easy to check conditions for asymptotic stability of those systems.

3.1 Non existence of equilibria in systems with VBR interfaces

At first we will introduce a Lemma, that essentially shows that the overall system cannot stay at an equilibrium if at least one of the delays $\tau_{2,\nu}(n)$ in the forward path is non-constant.

Lemma 1 *The system (22) cannot have a stable, non-zero equilibrium if the delay trajectories on the forward path are time-variant i.e. $\tau_{2,\nu}(n) \neq const$.*

Proof: Consider the ν^{th} feedback loop, with the VBR part shown in Figure 3. An equilibrium at $y(n) = y_0 \quad \forall n \geq 0$ implies $e_{\nu}(n) = 0 \quad \forall \nu = 1, \dots, m$ and hence that $z_{\nu}(n) = \frac{y_{out}}{m} = const, \quad \forall n \geq 0$. It can now be seen from Figure 3 or from the equations (12), (11) that despite the

constant input, the output is not constant if the delays on the forward path $\tau_{2\nu}$ are time variant.

Only for the degenerate (and not meaningful) case of $y_{out} = 0 \Rightarrow z_\nu(n) = 0$ will there be an equilibrium corresponding to $y(n) = y_0$, $z_\nu(n) = 0$.

Since the ν^{th} feedback path does not have a non-zero equilibrium if the delays $\tau_{2\nu}(n)$ are non-constant, the overall system in (22) will not have a non-zero equilibrium if the delays $\tau_{2,\nu}$ are non-constant.

Comments:

- If the delays on the VBR side are constant i.e. $\tau_{2,\nu} = const$, an equilibrium in the ν^{th} loop is achieved regardless of $\tau_{1,\nu}(n)$, if the reference input rate $y_0(n) = y_0$ is constant. This assumes $s_\nu(n) = const$.
- A similar result (as in Lemma 1) can be formulated for the case when $s_\nu = s_\nu(n)$ if the sources cannot communicate with one another instantly. Assume the system is at equilibrium and one source disconnects itself. If the other sources do not adjust the output rate accordingly, the queue will be depleted by more than the quantity sent by the active sources and the equilibrium will be disturbed.
- A similar argument like in Lemma 1 can be made for the system in equation (24).

3.2 Stability of systems with HFS interfaces

Since the system with time-variant delays in the forward path does not have an equilibrium, and the delays in the return path are usually more critical, we will address the case of time-invariant uncertain delays in the forward path, time-variant uncertain delays in the feedback path and time-invariant connection signal. Typically the sources are on or off for an extended period of time, which justifies such an analysis.

There are a number of results in the literature [16]-[18] which could be applied to the problem at hand, after it is translated into a state space representation. The resulting tests, however, are of prohibitively high complexity, especially if the delay interval is large. In this paper, we pursue another avenue. We will derive a sufficient condition for global asymptotic stability of the systems (22) and (24) (with constant delays on the forward paths), which can easily be tested and is not very conservative.

3.2.1 Stability of the linear model

If an equilibrium exists, the equilibrium point of the linear system depicted in Figure 6 is achieved at $y(n) = y_0$ and $z_\nu(n) = \frac{y_{out}}{m}$. Subtracting the equilibrium point from equation (22), the resulting zero input system with the equilibrium point shifted to the origin is:

$$\tilde{y}(n+1) = \tilde{y}(n) + \sum_{\nu=1}^m \sum_{j=\tau_2}^{\bar{\tau}_2} \sum_{i=0}^k \beta_{j\nu}(n-j) c_{i\nu} \tilde{y}(n - T_{i,j,\nu}(n)) +$$

$$+ \frac{y_{out}}{m} \sum_{\nu=1}^m \sum_{j=\tau_2}^{\bar{\tau}_2} \beta_{j\nu}(n-j) - y_{out} \quad (25)$$

When the delays on the forward paths are fixed ($\tau_{2\nu}(n) = \tau_{2\nu}$) we have exactly one coefficient $\beta_{j\nu} = 1$ for $j = \tau_{2\nu}$ and zero otherwise and equation (25) becomes:

$$\tilde{y}(n+1) = \tilde{y}(n) + \sum_{\nu=1}^m \sum_{i=0}^k c_{i\nu} \tilde{y}(n - T_{i,\tau_{2\nu},\nu}(n)) \quad (26)$$

Theorem 2:

The system (26) is globally asymptotically stable if the following three conditions are satisfied:

(a) $c_{i\nu} \leq 0 \quad \forall i = 0, \dots, k, \quad \forall \nu = 1, \dots, m$

(b) The transfer function

$$H_\epsilon(z) = \frac{1}{1 + (\epsilon-1)z^{-1} + \epsilon z^{-2} + \dots + \epsilon z^{-(L+1)}} = \mathcal{Z}\{h_\epsilon(n)\},$$

$L = k + \bar{\tau}_1 + \bar{\tau}_2$ is stable for

$$\epsilon = \sum_{\nu=1}^m \sum_{i=0}^k |c_i| \text{ with } \epsilon > 0 \text{ and } L \geq 2$$

(c) $\sum_{n=0}^{\infty} |h_\epsilon(n)| \epsilon < \frac{1}{L}$

Proof:

Considering equations (19) and (21), $T_{i\nu}$ defined in equation (23) satisfies the inequality:

$$\tau_1 + \tau_2 + i \leq T_{i,j,\nu}(n) \leq \bar{\tau}_1 + \bar{\tau}_2 + i$$

$\forall i = 0, \dots, k$ and $\forall \nu = 1, \dots, m$. Define $L = \bar{\tau}_1 + \bar{\tau}_2 + k$.

Now consider the following time-variant system with uncertain, time variant parameters $\Delta a_i(n)$, $i = 0, \dots, L$, $L \geq 1$:

$$\begin{aligned} y(n+1) = y(n) & - (\epsilon + \Delta a_0(n))y(n) \\ & - (\epsilon + \Delta a_1(n))y(n-1) \\ & - \dots \\ & - (\epsilon + \Delta a_L(n))y(n-L) \end{aligned} \quad (27)$$

In what follows, we will formulate conditions on ϵ and $\Delta a_i(n)$ such that the set of systems described by (27) contains the system of equation (26), i.e. the time variant system (26) is an element of the set of uncertain time variant systems (27).

At first we choose:

$$\epsilon = - \sum_{\nu=1}^m \sum_{i=0}^k c_{i\nu} \quad (28)$$

Imposing the condition

$$\sum_{i=0}^L \Delta a_i = -L\epsilon \quad (29)$$

on the uncertainties and letting

$$-\epsilon \leq \Delta a_i \leq 0, \quad i = 0 \dots L \quad (30)$$

ensures that system (26) can be represented by (27). Notice that this guarantees that each coefficient from equation 27 $-(\epsilon + \Delta a_i(n)) \in [-\epsilon, 0]$, $\forall i = 0 \dots L$ and that the sum of these coefficients is exactly $-\epsilon = \sum_{\nu=1}^m \sum_{i=0}^k c_{i\nu}$.

From (29) and (30) we have:

$$\sum_{i=0}^L |\Delta a_i| = L\epsilon \quad (31)$$

From Theorem 1 in [19] it is known that the system (27) is asymptotically stable if

$$\sum_{n=0}^{\infty} |h_{\epsilon}(n)| \cdot \gamma < 1 \quad (32)$$

$$\gamma > \sum_{i=0}^L |\Delta a_i(n)| \quad (33)$$

Since by equation (31), in our case we have

$$\gamma = L\epsilon + \mu \quad (34)$$

with μ arbitrarily small and positive real. From (32), (33) we obtain for stability:

$$\sum_{n=0}^{\infty} |h_{\epsilon}(n)| \cdot [L\epsilon + \mu] < 1$$

and hence

$$\sum_{n=0}^{\infty} |h_{\epsilon}(n)| \epsilon < \frac{1}{L}$$

which is condition (c) in the theorem. Hence condition (c) together with (a) and (b) guarantees global asymptotic stability of the system.

q.e.d.

Comments:

- Since $\sum_{n=0}^{\infty} |h_{\epsilon}(n)| > 1$, an upper bound for the values of ϵ which satisfy condition (c) is $\epsilon < \frac{1}{L}$, i.e. ϵ will have to decrease with increasing maximum delays $\bar{\tau}_1, \bar{\tau}_2$.

- The theorem provides important guidelines for controller design by providing bounds on $\sum_{\nu=1}^m \sum_{i=0}^k |c_{i\nu}|$, the 1 norm on the controller coefficients.

Now assume that ϵ is sufficiently small such that $h_\epsilon(n) \geq 0$ for all $n \geq 0$.

In this case

$$\sum_{n=0}^{\infty} |h_\epsilon(n)| = \sum_{n=0}^{\infty} h_\epsilon(n) = H_\epsilon(z)|_{z=1} = \frac{1}{(L+1)\epsilon}$$

and condition (c) is always satisfied. The existence of such an ϵ is guaranteed by the following lemma:

Lemma 3:

There exists $\epsilon_0 > 0$ such that $\forall \epsilon \ 0 < \epsilon < \epsilon_0$ the impulse response of the system with the following Z-transform:

$$H_\epsilon(z) = \frac{1}{1+(\epsilon-1)z^{-1}+\epsilon z^{-2}+\dots+\epsilon z^{-(L+1)}} = \mathcal{Z}\{h_\epsilon(n)\}$$

is strictly positive.

The proof can be found in [15].

3.2.2 Stability of the generalized nonlinear model

In the case of the generalized system the equilibrium point is $y(n) = y_0$, $z_\nu(n) = \frac{y_0 s_\nu}{\sum_{\nu=1}^m s_\nu(n)}$. If we subtract the equilibrium point from equation (24) we obtain:

$$\tilde{y}(n+1) = \text{sat}_q \left[\tilde{y}(n) + \sum_{\nu=1}^m s_\nu(n) \sum_{i=0}^k c_{i\nu} \tilde{y}(n - T_{i,\tau_{2\nu},\nu}(n)) \right] \quad (35)$$

where $\text{sat}_q(x) = \text{sat}_Q(x + y_0) - y_0$. Define $q_{\min} = -y_0$ and $q_{\max} = y_{MAX} - y_0$. Then the shifted saturation nonlinearity is:

$$\text{sat}_q[x] = \begin{cases} q_{\min} & \text{if } x < q_{\min} \\ x & \text{if } q_{\min} \leq x \leq q_{\max} \\ q_{\max} & \text{if } x > q_{\max} \end{cases}$$

with $q_{\min} < 0$ and $q_{\max} > 0$ (see Figure 8).

Using a sector description for the saturation nonlinearity provides:

$$\text{sat}_q(x) = \alpha x, \quad \alpha \in (0, 1] \quad (36)$$

Assuming that condition (a) in Theorem 2 is satisfied, the maximum argument in the saturation function of (35) is then given by:

$$\begin{aligned} & \max \left\{ y(n) + \sum_{\nu=1}^m s_\nu \sum_{i=0}^k c_{i\nu} y(n - T_{i,\tau_{2\nu},\nu}(n)) \right\} = \\ & = q_{\max} + \sum_{\nu=1}^m \sum_{i=0}^k c_{i\nu} q_{\min} \end{aligned}$$

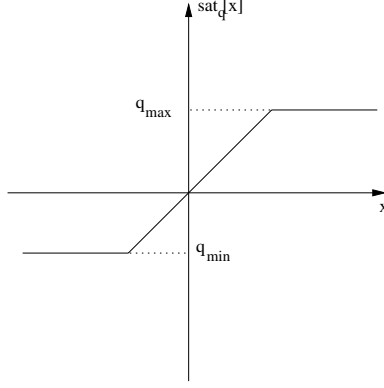


Figure 8: The shifted saturation non-linearity

$$= q_{\max} + |q_{\min}| \sum_{\nu=1}^m \sum_{i=0}^k |c_{i\nu}|$$

Similarly, the minimum argument in the saturation function of (35) is given by:

$$\begin{aligned} \min & \left\{ y(n) + \sum_{\nu=1}^m s_{\nu} \sum_{i=0}^k c_{i\nu} y(n - T_{i,\tau_{2,\nu},\nu}(n)) \right\} = \\ & = q_{\min} + \sum_{\nu=1}^m \sum_{i=0}^k c_{i\nu} q_{\max} \\ & = -|q_{\min}| - q_{\max} \sum_{\nu=1}^m \sum_{i=0}^k |c_{i\nu}| \end{aligned}$$

Hence the lower sector bound can be refined:

$$\min \left\{ \frac{q_{\max}}{q_{\max} + |q_{\min}| \sum_{\nu=1}^m \sum_{i=0}^k |c_{i\nu}|}, \frac{|q_{\min}|}{|q_{\min}| + q_{\max} \sum_{\nu=1}^m \sum_{i=0}^k |c_{i\nu}|} \right\} = \underline{\alpha} \quad (37)$$

For the case where $q_{\max} = -q_{\min}$ we have:

$$\underline{\alpha} = \frac{1}{1 + \sum_{\nu=1}^m \sum_{i=0}^k |c_{i\nu}|} = \frac{1}{1 + \epsilon};$$

Therefore

$$\text{sat}_q(x) = \alpha x, \quad \alpha \in [\underline{\alpha}, 1], \underline{\alpha} > 0$$

Theorem 4:

The system (35) is globally asymptotically stable if the following three conditions are satisfied:

- (a) $c_{i\nu} \leq 0 \quad \forall i = 0, \dots, k, \quad \forall \nu = 1, \dots, m$
- (b) The transfer function

$$H_{\underline{\alpha}}(z) = \frac{1}{1 + (\epsilon - \underline{\alpha})z^{-1} + \epsilon z^{-2} + \dots + \epsilon z^{-(L+1)}} = \mathcal{Z}\{h_{\underline{\alpha}}(n)\},$$

$$L = k + \tau_{1\max} + \tau_{2\max}$$
 is stable for

$$\epsilon = \sum_{\nu=1}^m \sum_{i=0}^k |c_{i\nu}|$$
 with $\epsilon > 0$ and $L \geq 1$
- (c) $\sum_{n=0}^{\infty} |h_{\underline{\alpha}}(n)| < \frac{1}{(L+1)\epsilon + (1-\underline{\alpha}) - \delta}$
 with $\delta = \min_{\nu} \sum_{i=0}^k |c_{i\nu}|$.

Proof:

Equation (35) now becomes with (36) and (37):

$$y(n+1) = \alpha y(n) + \sum_{\nu=1}^m s_{\nu} \sum_{i=0}^k \alpha c_{i\nu} y(n - T_{i, \tau_{2\nu}, \nu}(n))$$

with $\alpha \in [\underline{\alpha}, 1]$

For the new nominal system, we choose:

$$y(n+1) = (\underline{\alpha} - \epsilon)y(n) - \epsilon y(n-1) - \dots - \epsilon y(n-L)$$

with ϵ defined as in equation (28). The uncertainty is now bounded by:

$$\gamma > (1 - \underline{\alpha}) + (L+1)\epsilon - \min_{\nu} \sum_{i=0}^k |c_{i\nu}|$$

Denote $\delta = \min_{\nu} \sum_{i=0}^k |c_{i\nu}|$. The stability condition of Theorem 1 in [19] becomes:

$$\sum_{n=0}^{\infty} |h_{\underline{\alpha}}(n)| [(L+1)\epsilon + (1-\underline{\alpha}) - \delta] < 1$$

or equivalently:

$$\sum_{n=0}^{\infty} |h_{\underline{\alpha}}(n)| < \frac{1}{(L+1)\epsilon + (1-\underline{\alpha}) - \delta} \quad (38)$$

which is satisfied as condition (c) in the hypothesis.

q.e.d.

If ϵ is sufficiently small such that $h_{\underline{\alpha}}(n) \geq 0$ for all $n \geq 0$, then

$$\sum_{n=0}^{\infty} |h_{\underline{\alpha}}(n)| = \sum_{n=0}^{\infty} h_{\underline{\alpha}}(n) = \frac{1}{1 - \underline{\alpha} + (L+1)\epsilon}.$$

We then satisfy (38) in all cases.

4 Examples

4.1 The effect of the time-variant delays on the forward path

In this section we will illustrate via a very simple example the destabilizing effects of time-variant delays on the forward paths. To isolate the phenomenon we construct a simple example: assume we have only one loop ($m = 1$), no non-linearities and no disconnecting sources, no delays on the return path ($\tau_{1,1}(n) = 0 \quad \forall n \geq 0$), the buffer set point $y_0 = y_{\max}/2$ (50%) and the depletion rate $y_{out} = y_{\max}/10$ per time step (i.e. 10% of the buffer is depleted per time instant) (This corresponds to Figure 6 for $m=1$). The delays on the forward path can vary $\tau_{2,1}(n) \in \{0, 1, 2, 3, 4, 5\}$, i.e. $\bar{\tau}_2 = 5$. We will use a proportional controller. The controller gain is chosen such that the system is stable for all the fixed delay systems arising from $\tau_{2,1}(n) = \tau_2 \quad \tau_2 \in \{0, 1, 2, 3, 4, 5\}$.

We start the simulation with the system at the equilibrium point ($y(n) = y_0$) and with $\tau_2(n) = 0$. After 10 sampling periods we start to randomly change the delays in the forward path. In order to illustrate the fact that even slow changes will perturb the equilibrium, we will only change the delays by one unit at a time. In Figure 9 (a) the time trace of the delays on the forward path is shown. In Figure 9 (b) the effect of those changes on the buffer occupancy level is shown. It is clear that, as predicted, the equilibrium has been perturbed. If the same delay trace is applied on the return path (HFS interface) while keeping the delays on the VBR side constant, the system has an equilibrium point and the equilibrium point is stable (see Figure 9 (c)).

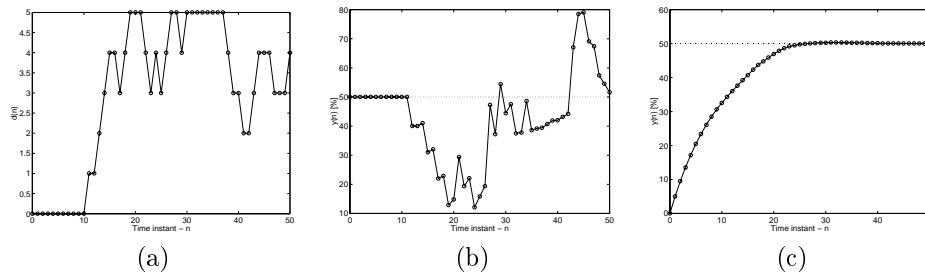


Figure 9: The effect of a time-variant delay on the buffer occupancy. (a) Time variant delay in the forward path (b) Corresponding buffer occupancy when the return path delay is kept constant. (c) Buffer occupancy when the delays on the forward path are kept constant and the delay in the return path is time-variant.

4.2 Computing stable controller gains when an equilibrium exists

In this section we will illustrate the application of Theorem 2 in computing stable controller gains when an equilibrium exists.

We will consider the case of $m = 6$ sources, each equipped with a proportional controller $c_1 - c_6$ (see Figure 6 for $m = 6$). We assume that the sources do not disconnect and we do not model the buffer nonlinearity. On the forward path (the VBR side) we assume constant delays $\tau_{2\nu}(n) = 2 \forall n \geq 0$. On the return path (HFS side) we consider time variant delays bounded by $\bar{\tau}_1 = 20$. Then L (as defined in Theorem 2) equals to 22. The condition (c) in Theorem 2 results in the following condition:

$$\sum_{\nu=1}^6 |c_\nu| < 0.001634$$

A time-invariant analysis of the system (i.e. checking stability of all frozen time systems) results in the following condition:

$$\sum_{\nu=1}^6 |c_\nu| < 0.0698$$

which is known not to guarantee stability of the time-variant system.

5 Conclusion

This paper introduced a few significant and new insights regarding buffer occupancy control in data networks:

- For the first time, the nature of a time-variant delay itself and the need for a linear system interface is introduced and analyzed.
- For the first time, it was shown that a variable bit-rate control scheme with time-variant communication delays cannot have an equilibrium point if the VBR link is time-variant.
- Simple, easy to check, sufficient stability conditions for the queue control systems were derived.

Current and future work addresses the problem of finding less conservative stability tests without having to pay the price of using NP hard tests [18].

References

- [1] C. E. Rohrs, R. A. Berry, and S. J. O'Halek, "Control engineer's look at ATM congestion avoidance," *Computer Communication*, vol. 19, pp. 226–234, Mar. 1996.
- [2] C. E. Rohrs and R. A. Berry, "A linear control approach to explicit rate feedback in ATM networks," in *Proc. IEEE Infocom*, pp. 277–282, 1997.
- [3] L. Benmohamed and S. M. Meerkov, "Feedback control of congestion in packet switching networks: The case of a single congested node," *IEEE/ACM Transactions on Networking*, vol. 1, Dec. 1993.

- [4] L. Benmohamed and S. M. Meerkov, "Feedback control of congestion in packet switching networks: The case of multiple congested nodes," *International Journal of Communication Systems*, vol. 10, pp. 227–246, Sept. 1997.
- [5] P. Narváez and K. Siu, "Optimal feedback control for ABR service in ATM," Tech. Rep. 1, d'Arbeloff Laboratory for Information Systems and Technology, MIT, 1997.
- [6] M. M. Ekanayake, *Robust Stability of Discrete Time Nonlinear Systems*. PhD thesis, University of Miami, 1999.
- [7] R. Krtolica, U. Özgüner, H. Chan, H. Goktas, J. Winkelman, and M. Liubakka, "Stability of linear feedback systems with random communication delays," *International Journal of Control*, vol. 59, pp. 925–953, Apr. 1994.
- [8] H. Chan and U. Özgüner, "Closed loop control of systems over a communication network with queues," *International Journal of Control*, vol. 62, pp. 493–510, Sept. 1995.
- [9] H. Chan and Ü. Özgüner, "Control of interconnected systems over a communication network with queues," in *Proc. 33rd Conference on Decision and Control*, pp. 4104–4109, 1994.
- [10] H. Chan and U. Özgüner, "Optimal control of systems over a communication network with queues via a jump system approach," in *Proc. 1995 IEEE Conference on Control Applications*, pp. 1148–1153, 1995.
- [11] Y. Halevi and A. Ray, "Integrated communication and control systems: Part I - analysis," *Journal of Dynamic Systems, Measurement and Control*, vol. 110, pp. 367–373, Dec. 1988.
- [12] Y. Halevi and A. Ray, "Integrated communication and control systems: Part II - design considerations," *Journal of Dynamic Systems, Measurement and Control*, vol. 110, pp. 374–381, Dec. 1988.
- [13] A. Bemporad, "Predictive control of teleoperated constrained systems with unbounded communication delays," in *Proc. of the 37th IEEE Conference on Decision and Control*, pp. 2133–2138, Dec. 1998.
- [14] P. H. Bauer, M. L. Sichertiu, and K. Premaratne, "Controlling an integrator through data networks: Stability in the presence of unknown time-variant delays," in *Proc. 1999 IEEE International Symposium on Circuits and Systems*, vol. 5, pp. 491–494, 1999.
- [15] P. H. Bauer, M. L. Sichertiu, and K. Premaratne, "Closing the loop through communication networks: The case of an integrator plant and multiple controllers," in *Proc. 38th IEEE Conference on Decision and Control*, pp. 2180–2185, Dec. 1999.
- [16] P. H. Bauer, K. Premaratne, and J. Durán, "A necessary and sufficient condition for robust asymptotic stability of time-variant discrete systems," *IEEE Transactions on Automatic Control*, vol. 38, pp. 1427–1430, Sept. 1993.
- [17] R. Brayton and C. Tong, "Constructive stability and asymptotic stability of dynamical systems," *IEEE Trans. on Circuits and Systems*, vol. 27, pp. 1121–1130, 1980.

- [18] A. Bhaya and F. Mota, "Equivalence of stability concepts for discrete time-varying systems," *International Journal of Robust and Nonlinear Control*, vol. 4, pp. 725–740, 1994.
- [19] S. A. Yost and P. H. Bauer, "Asymptotic stability of linear shift-variant difference equations with diamond shaped uncertainties," in *IEEE International Symposium on Circuits and Systems*, pp. 785–788, May 1995.