Designing Stable ABR Flow Control with Rate Feedback and Open-Loop Control: First-Order Control Case

Song Chong
Dept. of Electronic Engineering
Sogang University
Seoul, Korea

Ramesh Nagarajan   Yung-Terng Wang
Performance Analysis Dept.
Bell Laboratories
Holmdel, NJ, USA

Abstract

In this paper we present a control-theoretic approach to design stable rate-based flow control for ATM ABR services. The flow control algorithm that we consider has the most simple form among all the queue-length-based flow control algorithms, and is referred to as *First-order Rate-based Flow Control* (FRFC) since the corresponding closed loop can be modeled as a first-order retarded differential equation. We analyze the equilibrium and the asymptotic stability of the closed loop for the case of multiple connections with diverse round-trip delays. We also characterize the asymptotic decay rate at which the stable closed loop tends to the equilibrium. The decay rate is shown to be a concave function of control gain with its maximum being the inverse of round-trip delay. We also consider an open loop control in which the queue control threshold is dynamically adjusted according to the changes in the available bandwidth and the number of connections. This open loop control is shown to be necessary and effective to prevent the closed loop from converging to an undesirable equilibrium point.

1 Introduction

Recently there has been a great interest in feedback-based flow control for high-speed wide-area ATM networking. In particular, a *rate-based approach* has been studied extensively [2]-[11] and adopted by the ATM Forum as the standard for the flow control of the Available Bit Rate (ABR) service [4, 12].

Rate-based flow control algorithms can be classified into two types. The first type is the queue-length-based rate control that we consider in this paper. In this type of control the rates are computed based on a certain function of the difference between the observed queue length and a queue threshold, and a certain fairness in rate allocation among users is accomplished as a consequence of the queue-length control. Examples of this type can be found in [2] [5]-[8] [11]. The other type of rate-based flow control [3, 10] is to compute directly rate allocations in a way that a certain fairness property is satisfied. Typically, in this latter approach, the queue length is not explicitly controlled.

There are important criteria in the design of high-performance ABR flow control algorithms. In the following, we summarize the criteria in consideration.

- Maximal link utilization and small cell loss, and consequently maximal throughput in steady-state.

- Stability (preferably asymptotic stability) of the steady-state solution for the case of multiple virtual circuits (VCs) with long and diverse round-trip delays.

- Fair bandwidth allocation among ABR streams; guarantee of standard fairness criteria such as *MAX-MIN fair share* [12].

- Fast and uniform convergence irrespective of number of active VCs.

- Adaptability to the changes in the operational environment, for instance, changes in available bandwidth and the number of active VCs.

- Simplicity in implementation.

We consider a control-theoretic rate-based flow control algorithm which we first proposed in [6] and refer to as *First-Order Rate-Based Flow Control* (FRFC). The FRFC is a queue-length-based flow control algorithm where the rate allocated to each ABR user is the difference between observed network queue length and queue threshold, multiplied by a control gain. The analysis will show that this simple form of control can readily achieve the above performance objectives.

2 Closed-Loop Control and Modeling

Consider a network with a single bottleneck link as in Figure 1. The assumptions employed for the analysis of the FRFC algorithm are as follows and are fairly standard [2] [5]-[8]:

A.1. The traffic is viewed as a deterministic fluid flow and the network queueing process and the feedback control is continuous in time.
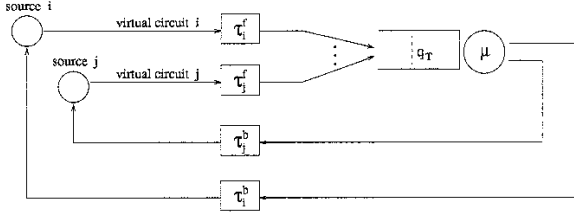
Figure 1: Network model with a single bottleneck link

A.2. The round-trip-time, τ_i, of virtual circuit i is the sum of forward-path delay, τ_i^f, and the backward-path delay, τ_i^b, which consists of propagation, queueing, transmission and processing times. We assume that τ_i is a constant.

A.3. The sources are *persistent* until the system reaches steady state.

A.4. There are no arrivals and departures of virtual circuits until the system reaches steady state.

A.5. The available bandwidth of the bottleneck link is constant until the system reaches steady state. Also, the buffer size at the bottleneck link is assumed infinite.

The assumptions **A.4-A.5** will be removed as necessary in Section 5 and in Section 6 to deal with dynamic environments.

Let $r_i(t)$, $i = 1, \cdots, n$, denote the rate allocation to virtual circuit i, which is computed by the switch at time t. Also, let $q(t)$, $\dot{q}(t)$ and μ respectively denote the queue length, its derivative at time t, and the available bandwidth at the bottleneck link. The rate-based flow control algorithm that we consider in this paper is a switch algorithm of the following simple form:

$$r_i(t) = (\ -\frac{K_i}{n}(\ q(t) - q_T\ )\ )^+, \quad K_i > 0 \quad (1)$$

where K_i is the control gain, q_T is the queue length threshold for the flow control, and the symbol $(\ \cdot\ )^+$ denotes $max\{\ \cdot,\ 0\ \}$. We refer to this algorithm as *First-Order Rate-Based Flow Control* (FRFC) since as we will see, the behavior of the closed-loop system with this form of algorithm is governed by a first-order differential equation. In contrast, most other existing algorithms found in [2, 5, 7, 8, 11] can be viewed as a second-order flow control since the rate is modulated via its derivative and thus the behavior of the closed-loop system is governed by a second-order differential equation. For instance, the algorithm in [7, 8] has the following form:

$$\dot{r}_i(t) = -a_i r_i(t) - b_i(\ q(t) - q_T\ ) \quad (2)$$

where a_i and b_i are two positive control constants. A noticeable feature of the FRFC algorithm is the control

gain K_i scaled by the number of VCs n. It will be shown later that such a scaling can yield a uniform convergence rate of the bottleneck queue irrespective of the number of VCs.

According to the above assumptions, the queueing process at the bottleneck link is given as

$$\dot{q}(t) = \begin{cases} \sum_{i=1}^{n} r_i(t - \tau_i) - \mu, & q(t) > 0 \\ (\ \sum_{i=1}^{n} r_i(t - \tau_i) - \mu\ )^+, & q(t) = 0. \end{cases} \quad (3)$$

In the next section we investigate the equilibrium points and the asymptotic stability of the equation (3) when the control (1) is applied.

3 Equilibrium and Asymptotic Stability

Let q_∞ and $r_{i\infty}$ respectively denote the steady-state solution of $q(t)$ and $r_i(t)$ in the closed-loop system given by (1) and (3). By taking the limit as $t \to \infty$ on both sides of (1) and (3) with $\lim_{t\to\infty} \dot{q}(t) = 0$, we can easily get the following result on the existence of equilibrium [6].

Proposition 3.1 *For the FRFC, there exist two steady-state solutions (equilibrium points):*
if $q_T > \frac{n\mu}{\sum_{i=1}^{n} K_i}$,

$$q_\infty = q_T - \frac{n\mu}{\sum_{i=1}^{n} K_i}; \quad r_{i\infty} = \frac{K_i}{\sum_{i=1}^{n} K_i}\mu, \quad \forall\ i, \quad (4)$$

otherwise,

$$q_\infty = 0; \quad r_{i\infty} = \frac{K_i}{n}q_T \leq \frac{K_i}{\sum_{i=1}^{n} K_i}\mu, \quad \forall\ i. \quad (5)$$

Note that q_∞ cannot be greater than or equal to the control threshold q_T. The network queue can be stabilized at either zero or a certain value smaller than q_T. For given $n\mu$, the choice of q_T and K_i, $i = 1, \cdots, n$, determines where for the system to converge. If q_T and K_i, $i = 1, \cdots, n$ are chosen such that $q_T > \frac{n\mu}{\sum_{i=1}^{n} K_i}$, the system has the equilibrium in (4) where the bottleneck link is fully utilized and any desired sharing of the bottleneck bandwidth can be accomplished through a proper selection of control gains. For instance, by taking identical gains, one can achieve MAX-MIN fair bandwidth sharing. On the other hand, if q_T and K_i, $i = 1, \cdots, n$ are chosen such that $q_T \leq \frac{n\mu}{\sum_{i=1}^{n} K_i}$, the system has the equilibrium in (5) where the bandwidth sharing can be still fair but the available bandwidth cannot be fully utilized.

Now we investigate the asymptotic stability of the equilibrium point (4) at which full link-utilization is achieved. For understanding the asymptotic stability,

we ignore the non-linearities introduced by the buffer floor and the control. Define

$$e(t) = q(t) - q_\infty = q(t) - q_T + \frac{n\mu}{\sum_{i=1}^n K_i}. \tag{6}$$

By combining (6) and the linear parts of (3) and (1), we obtain the following closed-loop equation

$$\dot{e}(t) + \frac{1}{n} \sum_{i=1}^n K_i e(t - \tau_i) = 0 \tag{7}$$

which is a first-order retarded differential equation [1, 13]. The characteristic equation of the closed-loop equation (7), denoted by $D(s)$, is

$$D(s) = s + \frac{1}{n} \sum_{i=1}^n K_i e^{-s\tau_i} = 0 \tag{8}$$

which is an exponential polynomial of s. For *asymptotic stability* of the closed-loop equation (7), all the roots of the characteristic equation (8) must have negative real parts [1, 13].

To find the necessary and sufficient condition for $D(s) = 0$ to have stable roots, one can appeal to Pontryagin's criterion [1, 8] assuming discrete delays of rational ratios. For more general case with continuous delays or discrete delays of irrational ratios, Stépán's criterion [13] provides a way to construct the necessary and sufficient condition. However, constructing such a condition in an explicit form is extremely complicated for the system with multiple connections of diverse round-trip delays.

Instead of finding the necessary and sufficient condition, we have derived a useful sufficient condition based on Stépán's criterion, which can be found in our earlier work [6]. We re-state the result below for convenience.

Proposition 3.2
The closed-loop system (7) is asymptotically stable if $\frac{1}{n}\sum_{i=1}^n K_i \tau_i < 1$.

For the case with $K_i = K$, $\forall i$, which is of our particular interest, the above stability condition implies that $K < \bar{\tau}^{-1}$ where $\bar{\tau}$ is the average round-trip delay of all the active VCs, i.e., $\bar{\tau} = \frac{\sum_{i=1}^n \tau_i}{n}$. Note that the condition is easy to use and requires only the estimate of average round-trip delay.

Next we derive the necessary and sufficient condition for asymptotic stability for a special case that all the round-trip delays are identical. Let $\tau_i = \tau$, $\forall i$. Then, the closed-loop equation (7) becomes

$$\dot{e}(t) + \frac{1}{n} \sum_{i=1}^n K_i e(t - \tau) = 0. \tag{9}$$

This equation may be normalized so that the time lag τ is unity. Let $t = \tau\xi$. In terms of the new variable ξ, (9) becomes

$$\dot{e}(\xi) + Pe(\xi - 1) = 0 \tag{10}$$

where

$$P = \frac{\tau}{n} \sum_{i=1}^n K_i. \tag{11}$$

The characteristic equation of (10), denoted by $H(z)$, is

$$H(z) = ze^z + P = 0. \tag{12}$$

For asymptotic stability of the closed-loop equation (10), all the roots of (12) must have negative real parts. To find the necessary and sufficient condition for $H(z) = 0$ to have such roots, one can appeal to Pontryagin's criterion [1, 8].

Proposition 3.3
The closed-loop system (9) is asymptotically stable if and only if $\frac{\tau}{n}\sum_{i=1}^n K_i < \frac{\pi}{2}$.

The proof is omitted due to the limited space. The stability condition is sufficient and necessary for the special case with homogeneous delays, whereas the condition given in Proposition 3.2 is only a sufficient condition for the general case with heterogeneous delays. Consider the case that $\tau_i = \tau$, $K_i = K$, $\forall i$. In this case, these two conditions imply that $K < \tau^{-1}$ and $K < \frac{\pi}{2}\tau^{-1}$ respectively, i.e., the actual range of stable K is larger than the range given by the sufficient condition by the factor $\pi/2$. Therefore we can argue that the stability condition in Proposition 3.2 is fairly tight in spite of being only a sufficient condition.

4 Asymptotic Decay Rate and Dependence on Control Gain

In this section we determine the rate at which the stable closed-loop system approaches steady state. Consider the principal root z^*, which is the root of (12) having the largest real part. Let $\alpha = -\mathrm{Re}(z^*) > 0$. Then, the asymptotic solution of the closed-loop equation (10) satisfies the following inequality:

$$\lim_{\xi \to \infty} |e(\xi)| \leq Ce^{-\alpha\xi} \tag{13}$$

where C is a properly chosen positive real constant taking into account initial conditions. In terms of the original variable $t(= \tau\xi)$, (13) can be rewritten by

$$\lim_{t \to \infty} |e(t)| \leq Ce^{-\frac{\alpha}{\tau}t}. \tag{14}$$

Note that $\frac{\alpha}{\tau}$ is the asymptotic decay rate at which the system tends to the equilibrium point. Hence the inverse of it, $\frac{\tau}{\alpha}$, is the time constant of the closed-loop system, i.e., the time it takes for a small perturbation around the equilibrium point to decrease by a factor of e^{-1}. Consider the case that $\tau_i = \tau$, $\forall i$. The change of variable $z = \psi - \sigma\tau$ transforms the characteristic equation (12) to

$$(\psi - \sigma\tau)e^{\psi} + Pe^{\sigma\tau} = 0. \tag{15}$$

If we choose σ to be the supremum of positive real number for which the transformed characteristic equation (15) has all roots in the left half plane, then $\alpha = \sigma\tau$, and the asymptotic decay rate and the time constant are σ and σ^{-1} respectively. The sufficient and necessary conditions that all roots of a characteristic equation in the form of (15) be in the left half plane were given by earler works in [14]. We re-state these conditions as a set of inequalities:

$$\sigma\tau < 1 \tag{16}$$

$$\sigma\tau e^{-\sigma\tau} < P \tag{17}$$

$$P < \frac{u_1}{\sin u_1} e^{-\sigma\tau} \tag{18}$$

where u_1 is the root of the equation

$$u = \sigma\tau\tan u \tag{19}$$

in the interval $(0, \pi)$. We now choose σ to be the supremum of positive real numbers satisfying (16)-(18). The inequality (18) is equivalent to

$$P < \frac{u_1}{\sin u_1} e^{-\frac{u_1}{\tan u_1}}. \tag{20}$$

The right side of (20) is monotone increasing on $0 < u_1 < \pi$. It follows from (19) that u_1 decreases when σ increases. Thus an increase in σ produces a decrease in the right side of (20), and the maximum σ satisfying this inequality is obtained by solving the corresponding equality. By an analogous argument, since an increase in σ on $\sigma\tau < 1$ produces an increase in the left side of (17), the maximum satisfying this inequality is also obtained by solving the corresponding equality. If any of the three equalities corresponding to (16), (17), (20) has no solution for σ, the the corresponding inequality places no restriction on σ. We summarize our results as follows:

Proposition 4.1 *Consider the stable closed-loop system (9). Let σ_1, σ_2, σ_3 be the positive solutions of*

$$\sigma\tau = 1 \tag{21}$$

$$\sigma\tau e^{-\sigma\tau} = \frac{\tau}{n}\sum_{i=1}^{n} K_i, \quad \sigma\tau \leq 1 \tag{22}$$

$$\frac{\tau}{n}\sum_{i=1}^{n} K_i = \frac{u_1}{\sin u_1} e^{-\frac{u_1}{\tan u_1}}, \quad u_1 = \sigma\tau\tan u_1, \quad u_1 \in (0, \pi) \tag{23}$$

respectively, with the understanding that $\sigma_i = \infty$ if the corresponding equation has no solution. Then, the asymptotic decay rate σ of the closed-loop solution is given by

$$\sigma = \min\{\sigma_1, \ \sigma_2, \ \sigma_3\} \tag{24}$$

and $\sigma \leq \tau^{-1}$.

Note that the asymptotic decay rate is a function of only two system parameters, round-trip delay τ and the average of control gains, $\bar{K} = \frac{1}{n}\sum_{i=1}^{n} K_i$. The decay rate is indenpendent of the number of active VCs, which is a highly desirable property in practice since it guarantees an identical rate of convergence no matter how many VCs are active. In contrast, it can be shown that if the gain is not scaled by the number of VCs as in the control (1), the asymptotic decay rate becomes no longer independent of the number of VCs. What matters in this type of control is that one needs to estimate the maximum number of possible VCs, say n_{max}, and the gains should be selected for this extreme case such that $\bar{K} < \frac{\pi}{2n_{max}\tau}$. As a consequence, in nominal cases with n much smaller than n_{max}, the system would converge unnecessarily slowly due to the small value of the chosen gains. Note also that the asymptotic time constant σ^{-1} cannot be smaller than the round-trip delay, which means that the time it takes for a small perturbation around the equilibrium point to decrease by a factor of e^{-1} is at least τ.

Next we investigate the dependence of σ on $\bar{K}$ for a given τ. The function in the left side of (22), $\sigma\tau e^{-\sigma\tau}$, is monotone increasing on $0 < \sigma\tau < 1$ with minimum value 0 at $\sigma\tau = 0$ and maximum value e^{-1} at $\sigma\tau = 1$. The function in the right side of the first equation in (23)

$$g(u) \equiv \frac{u}{\sin u} e^{-\frac{u}{\tan u}} \tag{25}$$

is monotone increasing on $0 < u < \tau$ with $g(0) = e^{-1}$, $\lim_{u \to \pi} g(u) = \infty$. Also, the function in the second equation in (23), $\frac{u}{\tan u}$, is monotone deceasing on $0 < u < \pi$ with $\lim_{u \to 0} = 1$. For a given τ, define

$$K^* = e^{-1}\tau^{-1}, \quad K^{**} = g(\frac{\pi}{2})\tau^{-1} = \frac{\pi}{2}\tau^{-1}. \tag{26}$$

Consider the case $0 < \bar{K} < K^*$. Since $\bar{K} < K^*$ implies that $\frac{\tau}{n}\sum_{i=1}^{n} K_i < e^{-1}$ and $g(u_1) > e^{-1}$ for $u_1 \in (0, \pi)$, (23) has no solution, i.e., $\sigma_3 = \infty$. On the other hand, since the function $\sigma\tau e^{-\sigma\tau}$ is monotone increasing on $0 < \sigma\tau < 1$ with minimum value 0 at $\sigma\tau = 0$ and maximum value e^{-1} at $\sigma\tau = 1$, (22) has a unique solution σ_2 which is smaller than τ^{-1}. Hence, $\sigma = \sigma_2 < \tau^{-1}$ in this case. Differentiation of (22) with respect to $\bar{K}$ gives

$$\frac{d\sigma}{d\bar{K}} = \frac{1}{e^{-\sigma\tau}(1 - \sigma\tau)} > 0 \quad \text{when } \sigma\tau < 1. \tag{27}$$

Therefore, we conclude that σ is monotone increasing with respect to $\bar{K}$.

Consider the case $K^* < \bar{K} < K^{**}$. Since this case implies that $e^{-1} < \frac{\tau}{n}\sum_{i=1}^{n} K_i < \frac{\pi}{2}$ and the value of the function $\sigma\tau e^{-\sigma\tau}$ is smaller than e^{-1} for $\sigma\tau < 1$, (22) has no solution, i.e., $\sigma_2 = \infty$. Since $g(u_1)$ is monotone increasing and the image of $g(u_1)$ is in $(\frac{1}{e}, \frac{\pi}{2})$ for $u_1 \in (0, \frac{\pi}{2})$, a unique solution u_1 of the first equation in (23)
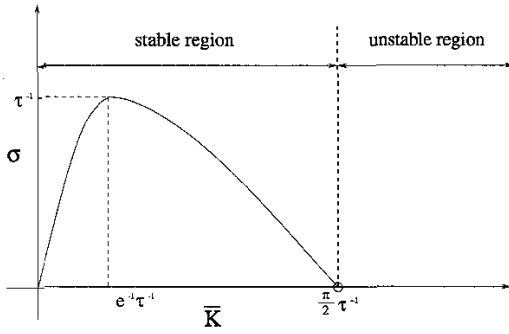
Figure 2: Dependence of asymptotic decay rate σ on control gain $\bar{K}$ for given round-trip delay τ

exists in the interval $(0, \frac{\pi}{2})$ and is monotone increasing as $\bar{K}$ increases. Furthermore, since the function $\frac{u_1}{\tan u_1}$ is monotone deceasing with respect to u_1 and the image of it, $\sigma\tau$, is in $(0, 1)$ for $u_1 \in (0, \frac{\pi}{2})$, a unique solution σ of the second equation in (23) exists in the interval $(0, \tau^{-1})$ and is monotone decreasing as $\bar{K}$ increases. It is obvious that $\sigma = \sigma_3$ in this case.

If $\bar{K} \geq K^{**}$, $\frac{\tau}{n} \sum_{i=1}^{n} K_i \geq \frac{\pi}{2}$ which implies that the closed-loop system is unstable (see Porposition 3.3). Finally, if $\bar{K} = K^*$, $\frac{\tau}{n} \sum_{i=1}^{n} K_i = e^{-1}$ and consequently $\sigma = \sigma_1 = \sigma_2 = \sigma_3 = \tau^{-1}$. We summarize our results as follows:

Proposition 4.2 *Consider the stable closed-loop system (9). The asymptotic decay rate σ is a monotone increasing function of $\bar{K}$ given by (22) with its value being smaller than τ^{-1} for $0 < \bar{K} < K^*$, and a monotone decreasing function of $\bar{K}$ given by (23) with its value being smaller than τ^{-1} for $K^* < \bar{K} < K^{**}$. Moreover, σ has its maximum τ^{-1} at $\bar{K} = K^*$.*

Figure 2 depicts the results of Proposition 4.2. Interestingly, the asymptotic decay rate is a concave function with respect to the average control gain, i.e., an increase in $\bar{K}$ up to $e^{-1}\tau^{-1}$ produces a monotone increase in σ but once $\bar{K}$ exceeds a certain value, $e^{-1}\tau^{-1}$, the increase in $\bar{K}$ rather produces a monotone decrease in σ. This is very different from the characteristics of an ordinary first-order system with no delay where the decay rate can be increased infinitely.

The above analysis considered the case with homogeneous delays. Some more analytical work is needed to understand the quantitative asymptotic behavior in the case with heterogeneous delays.

5 Open-Loop Control: Getting to Desired Equilibrium

So far we have studied the equilibrium and the asymptotic stability of the ABR closed loop in a static environment where the bottleneck bandwidth μ and the set of active VCs are assumed to be unchanged.

In reality, however, these assumptions are no longer true. The available bandwidth at the bottleneck link is time-varying since it depends on the instantaneous aggregate traffic of higher-priority services such as CBR and VBR. Also, the set of active VCs keeps changing due to the frequent arrivals and departures of VCs. One of the major problem in such a dynamic environment is that the quantity $\frac{n\mu}{\sum_{i=1}^{n} K_i}$ changes and hence, as shown in Proposition 3.1, the equilibrium point (i.e., the steady-state solution) of the system varies. More specifically, if $\frac{n\mu}{\sum_{i=1}^{n} K_i}$ grows and exceeds q_T due to the changes in μ and the set of active VCs, the link would become under-utilized and the queue would converge to zero. On the other hand, if we choose q_T large enough to to avoid such a under-utilization of link bandwidth, q_∞ would grow and so the likelihood of cell loss would increase for given buffer budget.

What is desirable in such a dynamic environment is the capability to keep $q_T - \frac{n\mu}{\sum_{i=1}^{n} K_i}$, i.e., q_∞, constant and positive. To accomplish this, we need an open-loop control in which the switch adaptively changes either q_T or the control gains whenever the changes in μ and the set of active VCs are detected. Considering the large number of VCs in a high speed link, we choose the former option. The dynamic queue threshold (DQT) algorithm that we propose in this paper is to change the queue control threshold in the following manner:

$$q_T(t) = \frac{|I(t)|\mu(t)}{\sum_{i \in I(t)} K_i} + \epsilon, \quad \epsilon > 0 \tag{28}$$

where $\mu(t)$ and $I(t)$ respectively denote time-varying bottleneck bandwidth and the set of active VCs with cardinality $|I(t)|$, and ϵ is the desired value of the ABR queue length in the steady state. For the case with $K_i = K, \forall i$, the above DQT algorithm is simplified to

$$q_T(t) = \frac{\mu(t)}{K} + \epsilon, \quad \epsilon > 0. \tag{29}$$

Consider the closed-loop system behavior with no buffer floor when the FRFC with DQT is applied. For simplicity we assume that only μ is time-varying while the number of VCs is fixed at n. Then, by combining (1), (3) and (28), we get the following closed-loop equation for the case with DQT

$$\dot{q}(t) + \frac{1}{n}\sum_{i=1}^{n} K_i q(t - \tau_i) = \frac{\epsilon}{n}\sum_{i=1}^{n} K_i - \mu(t) \tag{30}$$

$$+ \sum_{i=1}^{n} \frac{K_i}{\sum_{j=1}^{n} K_j}\mu(t - \tau_i).$$

In contrast, the closed-loop equation (7) for the case with static q_T can be rewritten as

$$\dot{q}(t) + \frac{1}{n}\sum_{i=1}^{n} K_i q(t - \tau_i) = \frac{q_T}{n}\sum_{i=1}^{n} K_i - \mu(t) \tag{31}$$
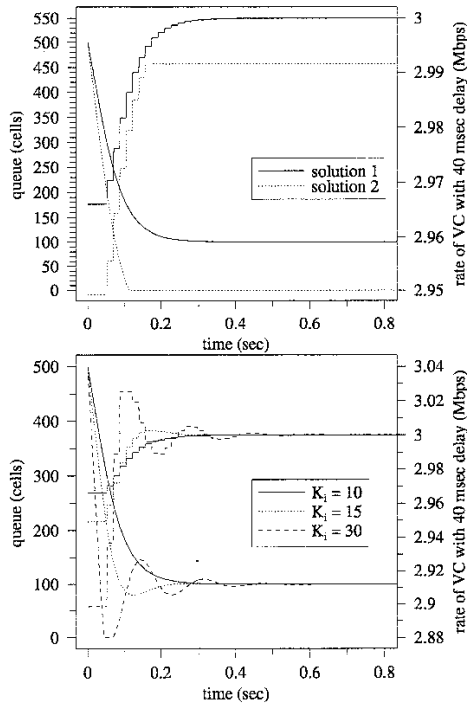
Figure 3: Control performance in static environment ($\mu = 150$ Mbps, # of VCs = 50, $\tau_i \in [10, 40]$ msec, $B = 5{,}000$ cells): (a) two steady-state solutions with $K_i = 10.0, \forall i$, (b) effect of control gains K_i.

with time-varying $\mu(t)$. As we see in (30) and (31), the major difference between the DQT case and the static queue threshold case is the third term in the right side of the DQT case (30). The role of this term is to nullify the effect of time-varying $\mu(t)$ with delays. In particular, if $\mu(t)$ varies slowly or is piecewise constant with reasonably long intervals, the term $-\mu(t) + \sum_{i=1}^{n} \frac{K_i}{\sum_{j=1}^{n} K_j} \mu(t - \tau_i)$ remains small in magnitude or as a superposition of impulses so that the effect of $\mu(t)$ becomes nearly nullified as the system approaches steady-state. In contrast, with a static queue threshold, the effect of $\mu(t)$ remains governing the dynamics of $q(t)$ as you see in (31). This difference will result in superior performances of the FRFC with DQT in dynamic environments, as will be shown in the next section through simulations.

6 Simulation Results

In this section, we simulate the network model to examine the performance of the FRFC algorithm. First we consider a static scenario where the assumptions A.4 and A.5 hold. The bottleneck bandwidth μ and the buffer size B are respectively set to 150 Mbps and 5,000 cells, and there are 50 active VCs sourcing the traffic into the link. The round-trip delay τ_i of VCs is

chosen uniformly in the range [10, 40] (msec) to represent long propagation delays. To take into account the discrete-time effect of control, the FRFC is applied in the sample-and-hold manner with intervals defined by the rate of VC. We choose this interval aggressively long as if RM (resource management) cells [12] are issued every 128 data cells by the ABR sources. Figure 3 a shows the two steady-state solutions, (4) and (5), with $K_i = 10.0, \forall i$, and $q_0 = 500$ cells. If we choose q_T at 35,477 cells, the queue $q(t)$ approaches 100 cells and the user rate $r_i(t)$ converges to the fair allocation (= 3 Mbps) as time goes. For the illustration, the rate trajectory of a VC with 40 msec round-trip delay is plotted in the figure. On the other hand, if we choose q_T at 35,277 cells, $q(t)$ converges to 0 and $r_i(t)$ approaches 2.9915 Mbps as computed in (5). Notice that it is not necessary for q_T to be smaller than the buffer size B.

Figure 3 b shows the effect of control gains on the queue and user rates with $K_i = 10.0$, 15.0 and 30.0. In the above simulation scenario, the choice of $K_i = 10.0$, 15.0 and 30.0 satisfies the stability condition in Proposition 3.2 since $\frac{1}{\bar{\tau}} \approx 40$. While changing the gain, we kept $\epsilon \ (= q_T - \frac{n\mu}{\sum_{i=1}^{n} K_i})$ positive and constant at 100 cells by changing q_T correspondingly. For the larger K_i, the system suffers from poor transient behavior such as overshoots at the risk of link under-utilization and cell loss, but still remains asymptotically stable. This example tells that the sufficient stability condition that we derive can serve as a practically good gain selection criterion.

Next we consider dynamic environments where μ is piecewise constant with 1 sec intervals. For the FRFC with static threshold, we set q_T at 35,877 cells aiming at $q_\infty = 500$ cells with $\mu(0)=150$ Mbps. This design implies that if μ does not vary, $q(t)$ will converge to 500 cells. For the FRFC with DQT in (28), ϵ was fixed at 500 cells. The trajectory of $\mu(t)$ is plotted in Figure 4 a as a solid curve. Also, in Figure 4 a, the user rate $r_i(t)$ of a VC with longest round-trip delay (=40 msec) is compared for the two cases. With DQT $r_i(t)$ tracks well $\mu(t)$ with a time lag, whereas without DQT $r_i(t)$ suffers from loss of bandwidth as observed during the time interval [1, 2] (sec). Figure 4 b explains why such a loss of bandwidth occurs without DQT. As explained in (31), with static threshold $q(t)$ essentially tracks the dynamics of $-\mu(t)$, consequently hitting both buffer floor and ceiling (see Figure 4 b). As also shown in Figure 4 b, the bandwidth utilization drops while $q(t)$ hits buffer floor, and hence the user suffers from the loss of bandwidth. On the other hand, if DQT is applied, $q(t)$ remains in the neighborhood of ϵ (= 500 cells), maintaining full utilization of bandwidth and no loss. This is because the time-varying dynamics of $\mu(t)$ does not directly affect the dynamics of $q(t)$. Rather,
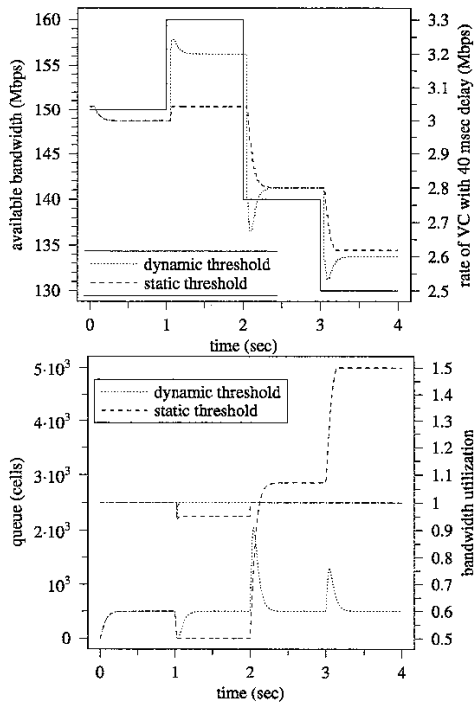
Figure 4: Control performance when $\mu(t)$ is piecewise constant (# of VCs = 50, $\tau_i \in [10, 40]$ msec, $B = 5,000$ cells, $K_i = 10.0$, $\forall i$): (a) trajectories of $\mu(t)$ (solid curve) and $r_i(t)$ of a VC with 40 msec round-trip delay, (b) trajectories of $q(t)$ and bandwidth utilization.

the jumps in $\mu(t)$ affect $q(t)$ as impulses so that the effect of jumps vanish after a certain transient period, as explained in (30).

More simulation results can be found in [6] and a comprehensive cell-based simulation study for multi-hop configurations are under way and the results will be reported in a separate paper.

References

[1] R. Bellman and K. L. Cooke, *Differential-Difference Equations* (Academic Press, New York, 1963).

[2] L. Benmohamed and S. M. Meerkov, Feedback Control of Congestion in Packet Switching Networks: The Case of Single Congested Node, *IEEE/ACM Trans. on Networking* **1**(6)(1993) 693-708.

[3] D. Bertsekas and R. Gallager, *Data Networks* (Prentice Hall, 1992).

[4] F. Bonomi and K. W. Fendick, The Rate-Based Flow Control Framework for the Available Bit Rate ATM Service, *IEEE Network* **9**(2)(1995) 25-39.

[5] F. Bonomi, D. Mitra and J. B. Seery, Adaptive Algorithms for Feedback-Based Flow Control in High-Speed, Wide-Area ATM Networks, *IEEE J. Select. Areas on Communications* **13**(7)(1995) 1267-1283.

[6] S. Chong, R. Nagarajan and Y. T. Wang, First-Order Rate-Based Flow Control with Dynamic Queue Threshold for High-Speed Wide-Area ATM Networks, *Proc. of SPIE Conference on Performance and Control of Network Systems* **3231** (1997) 259-270.

[7] S. Chong, Second-Order Rate-Based Flow Control with Dynamic Queue Threshold for High-Speed Wide-Area ATM Networks, preprint 1997.

[8] A. Elwalid, Analysis of Adaptive Rate-Based Congestion Control for High-Speed Wide-Area Networks, *Proc. IEEE ICC'95*, pp. 1948-1953.

[9] E. Hernandez-Valencia et al., Rate Control Algorithms for the ATM ABR Service, *European Transactions on Telecommunications* **8**(1)(1997) 7-20.

[10] L. Kalampoukas, A. Varma and K. K. Ramakrishnan, An Efficient Rate Allocation Algorithm for ATM Networks Providing Max-Min Fairness, Technical Report UCSC-CRL-95-29, Computer Engineering Dept., University of California, Santa Cruz, June 1995.

[11] K. K. Ramakrishnan and R. Jain, A Binary Feedback Scheme for Congestion Avoidance in Computer Networks with a Connectionless Network Layer, *Proc. ACM SIGCOMM'88*, pp. 303-313.

[12] S. Sathaye, *ATM Forum Traffic Management Specification*, Version 4.0, Feb. 1996.

[13] G. Stépán, *Retarded Dynamical Systems: Stability and Characteristic Functions* (Longman Scientific & Technical, 1989).

[14] N. D. Hayes, Roots of the Transcendental Equation Associated with a Certain Differential-Difference Equation, *J. London Math. Soc.*, **25** (1950) 226-232.