TOWARDS A STATE-SPACE APPROACH TO CONGESTION AND DELAY CONTROL IN COMMUNICATION NETWORKS

Henri-François Raynaud, Rim Hammi, Caroline Kulcsár
L2TI – Institut Galilée – Université Paris Nord – 99 avenue J.B. Clément – 93430 Villetaneuse – France
corresponding author: raynaud@l2ti.univ-paris13.fr

Abstract

This paper explores the application of state-space methods to congestion and delay control in communication networks. In the absence of buffer under- or overflow, the transmission delay for packets moving through a congested link turns out to be a non-linear output of a system with linear dynamics. As a consequence, the congestion/delay control problem can be addressed, in the presence of input saturation, by combining a disturbance feedforward plus state feedback control with an observer designed to cope with measurement delays. Both the feedback and observer designs are state-space versions of a Smith predictor.

1 Introduction

In the last two decades, the design and performance evaluation of efficient congestion control methods for packet-switching computer communication networks has emerged as a major engineering challenge. Most attempts to formulate congestion control problems in the language of control engineering have focused on an input-output approach. That line of research has led to innovative proposals for new or improved feedback control laws, notably in the technologically more flexible framework of ATM networks, and also to a better understanding of existing algorithms (see, *e.g.*, [2, 11, 9, 6, 12]). Other approaches used to control the input-output behavior of network elements include neural networks [10], fuzzy control [4] or utility functions [7].

In this contribution, we investigate the network congestion control problem using a state-space approach, in the simple case of a single router connected to a buffered link. This very simple system can be seen as an elementary building block with which more complex networks are constructed. Thus, a good understanding of its dynamics, and of the associated control problems, is a necessary first step towards the design of control methods of practical interest. Also, in any actual network, congestion control is only required upstream of links where demand exceeds available bandwidth, *i.e.* for a sub-network of reduced complexity. This explains why this model has been extensively studied from an input-output point of view [2, 11, 12].

Another important motivation for this work is the need to elaborate congestion control procedures able to meet the quality of service (QoS) requirements of real-time applications such as video transmission [8]. Most real-time applications use Real time Transport Control Protocol (RTCP) which provides only end-to-end aggregated feedback information such as mean loss rate and transmission delay, therefore imposing severe limitations on the structure of the rate control algorithm. It is hoped that the so-called active network technologies can provide a technological framework for implementing a wider range of alternative control procedures [5, 3].

A well-known merit of state-space methods is that they enable to separate the problem of designing a control action based on available measurements into a full information control problem and a state estimation problem, both of which can then be tackled using standard and constructive methods. In this case, the analysis of this system using standard state-space concepts leads to possible achievement of both congestion, delay and output flow control using an appropriate linear feedback plus feedforward scheme, associated if need be with a linear observer.

2 Elementary network model

Let us consider first a network made up of a router connected to a single link, with one source and one destination (figure 1). The packets arriving at the router are stored in a first-in, first-out (FIFO) buffer. The time unit is defined as the number of sampling periods. Assuming for the moment that i/ the time needed by the router to transfer incoming packets to the buffer and outgoing packets to the link is significantly smaller than the sampling period and that ii/ the source is capable of instantly adjusting its data rate according to the solicitations of the control procedure, the basic equation governing the buffer dynamics [2] is

$$c(t+1) = \mathrm{Sat}_{c_M}\left(c(t) + u(t) - b(t)\right) \quad , \tag{1}$$

where $c(t)$, c_M, $u(t)$ and $b(t)$ denote respectively the buffer congestion, maximum buffer capacity, source rate and available link bandwidth, while

$$\mathrm{Sat}_{c_M}(c) \triangleq \begin{cases} 0 & \text{if} \quad c < 0, \\ c & \text{if} \quad 0 \le c \le c_M, \\ c_M & \text{if} \quad c > c_M. \end{cases} \tag{2}$$

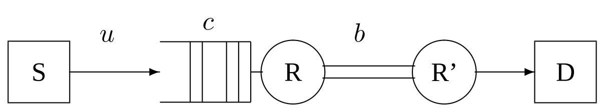

Figure 1: Elementary network configuration with one source, one buffer, one link

Let us further assume that the control u is constrained by $0 \leq u \leq u_M$. Clearly, if the link bandwidth remains superior to u_M, the system will then reach the (non-linear) stable equilibrium point $c = 0$ no matter what the source chooses to emit, and the appropriate control policy is therefore $u = u_M$. Consequently, we need only to consider here situations where b satisfies a constraint in the form $0 \leq b \leq b_M < u_M$..

In such a situation, it is possible to choose u such that $0 \leq c \leq c_M$. Thus, we can, at least initially, reduce the dynamics to the linear regime

$$c(t+1) = c(t) + u(t) - b(t) \quad . \tag{3}$$

Obviously, this linear regime is unstable. Should one desire to drive the congestion to some desired value c^r, an appropriate choice of u would be

$$u(t) = b(t) - k\left(c(t) - c^r(t)\right) \quad , \tag{4}$$

with $0 < k \leq 1$. Combining (3) and (4) yields the stable closed loop dynamics:

$$c(t+1) = (1-k)\,c(t) + kc^r(t) \quad , \tag{5}$$

which guarantees that the tracking error $\Delta c \triangleq c - c^r$ converges monotonously towards zero, while u converges towards b. In the next section, a sensible choice of c^r will be deduced from a study of the congestion-induced delay.

3 Control of congestion-induced delay

The congestion-induced transmission delay $d(t)$ experienced by the destination at time t is the difference between t and the time at which the packets presently exiting from the buffer have been emitted. Keeping this delay constant is, in some important respects, a more relevant control objective than maintaining the congestion at a prespecified level. To begin with, it means something for the user, while buffer congestion does not. Secondly, in real-time applications such as video transmission, an important QoS criterion is delay fluctuation (jitter).

In the absence of fixed delays, and assuming that no buffer overflow has occurred at least in the recent past, the delay incurred by packets percolating through a FIFO buffer is

$$d \triangleq \min\left\{ d > 0 \;\text{ such that }\; \sum_{s=1}^{d} u(t-s) \geq c(t) \right\}. \tag{6}$$

Assuming with no loss of generality that $d \leq d_M$, the adequate choice of state for this system is the vector with $n \triangleq d_M + 1$ coordinates

$$x(t) \triangleq \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \vdots \\ x_n(t) \end{pmatrix} \triangleq \begin{pmatrix} c(t) \\ u(t-1) \\ u(t-2) \\ \vdots \\ u(t+1-n) \end{pmatrix} \quad . \tag{7}$$

Equations (3) and (6) can be rewritten as a *linear state transition* equation associated with a *non-linear observation* equation:

$$x(t+1) = Ax(t) + Bu(t) + \Gamma b(t) \quad , \tag{8}$$

$$d(t) = h(x(t)) \quad , \tag{9}$$

with

$$A \triangleq \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \; B \triangleq \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; \Gamma \triangleq \begin{pmatrix} -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{10}$$

$$h(x) \triangleq \min\left\{ 0 < d < d_M \text{ such that } \sum_{s=1}^{d} x_{s+1}(t) \geq x_1(t) \right\}. \tag{11}$$

Using these formula, and for any constant desired delay d^r, one can construct a reference trajectory (u^r, x^r) for the state transition equation (8) which will result in $d = d^r$. A suitable choice is

$$u^r(t) \triangleq \frac{1}{2}\left(b(t+d^r-1) + b(t+d^r)\right) \quad , \tag{12}$$

$$x^r(t) \triangleq \begin{pmatrix} c^r(t) \\ u^r(t-1) \\ u^r(t-2) \\ \vdots \\ u^r(t+1-n) \end{pmatrix} \quad , \tag{13}$$

with $$c^r(t) \triangleq \frac{1}{2} b(t+d^r-1) + \sum_{s=0}^{d^r-2} b(t+s) \quad . \tag{14}$$

From the definitions of x^r and h, the identity $h(x^r(t)) = d^r$ is equivalent to

$$\sum_{s=1}^{d^r-1} u^r(t-s) < c^r(t) \leq \sum_{s=1}^{d^r} u^r(t-s) \quad . \tag{15}$$

Using this characterization, it is easily checked that if $x(0) = x^r(0)$ and $u = u^r$, then the identities $x(t) = x^r(t)$ and $d(t) = h(x^r(t)) = d^r$ will hold for all $t \geq 0$.

Consider now the tracking errors defined by $\Delta u \triangleq u - u^r$, $\Delta x \triangleq x - x^r$, $\Delta d \triangleq d - d^r$. If we apply a standard state-feedback control in the form $\Delta u = -K\Delta x$, or equivalently $u = u^r - K(x - x^r)$, where the control gain K is chosen so that $A - BK$ is a stability matrix, *i.e.* has all its eigenvalues of modulus strictly less than one, then the state tracking error Δx shall be driven towards zero according to the closed-loop dynamics $\Delta x(t+1) = (A - BK)\,\Delta x(t)$.

Because the linear open-loop dynamics are the series connection of an integrator and a delay line, the congestion feedback described in section 2 suffices to stabilize the closed-loop. In addition, the following proposition holds:

Proposition 1 *Assume that c is governed by the linear regime (3), that the link bandwidth at all $t \geq 0$ is constrained by*

$$0 < b_m \leq b(t) \leq b_M < u_M \quad , \tag{16}$$

and that the source rate control is

$$u^l(t) = u^r(t) - k(c(t) - c^r(t)) \quad , \tag{17}$$

$$u(t) = Sat_{u_M}\left(u^l(t)\right) \quad , \tag{18}$$

where u^r, x^r are defined by (12)-(14) and $0 < k \leq 1$. Then for any initial congestion $c(0)$, both Δx and Δd converge towards zero. Furthermore, the convergence of Δx is exponential with guaranteed rate $1 - k_m$, where

$$k_m \triangleq \min\left\{k, \frac{b_m}{|\Delta c(0)|}, \frac{u_M - b_M}{|\Delta c(0)|}\right\} \quad , \tag{19}$$

whereas Δd converges to zero in finite time.

Proof. Should u^l remain at all times in the interval $[0, u_M]$, the closed-loop dynamics of c would be governed by

$$\Delta x_1(t) = \Delta c(t) = (1-k)^t \Delta c(0) \quad , \tag{20}$$

$$\Delta x_j(t) = \Delta u(t-j) = -k(1-k)^{t-j} \Delta c(0) \tag{21}$$

for $j > 1$. In order to derive similar relations for the saturated control, it suffices to note that (17)-(18) can be rewritten as $\Delta u(t) = -k_S(t)\Delta c(t)$ with

$$k_S(t) = k \text{ if } -u^r(t) \leq -k\Delta c(t) \leq u_M - u^r(t), \tag{22}$$

$$k_S(t) = \frac{u_M - u^r(t)}{-\Delta c(t)} \text{ if } -k\Delta c(t) > u_M - u^r(t), \tag{23}$$

$$k_S(t) = \frac{u^r(t)}{\Delta c(t)} \text{ if } k\Delta c(t) > u^r(t). \tag{24}$$

Since $0 \leq k_S(t) \leq k$ for all $t \geq 0$, we get $|\Delta c(t+1)| = (1 - k_S(t))|\Delta c(t)| \leq |\Delta c(0)|$. Using this uniform boundedness of $|\Delta c(t)|$, the definition of u^r in (12) and equation (16), we can guarantee that for all $t \geq 0$, $k_S(t) \geq k_m$, so that (20)-(21) can be replaced by $|\Delta x_1(t)| \leq (1-k_m)^t |\Delta c(0)|$ and $|\Delta x_j(t)| \leq k(1-k_m)^{t+1-j}|\Delta c(0)|$ for $j > 1$. This establishes that Δx converges exponentially towards zero with guaranteed rate $1 - k_m$. To complete the proof, one needs only to confirm that $\Delta x \to 0$ implies $\Delta d \to 0$. From (15), we deduce that $h(x^r(t) + \Delta x(t)) = d^r$ if and only if the two inequalities $\frac{1}{2}b(t) + \Delta x_1(t) - \sum_{s=1}^{d^r-1}\Delta x_{s+1}(t) \geq 0$ and $\frac{1}{2}b(t-1) - \Delta x_1(t) + \sum_{s=1}^{d^r}\Delta x_{s+1}(t) > 0$ are satisfied. It is immediately checked that this will be true, in the worst possible scenario, as soon as $\|\Delta x(t)\| < b_m/2$. ■

4 Taking into account fixed delays

Assume now that the linear buffer regime (3) is replaced by

$$c(t+1) = c(t) + u(t - T_c) - b(t) \quad , \tag{25}$$

where the total control delay T_c is the sum of the time T_s needed for the source to adjust its rate and of an incompressible processing time T_p incurred by packets before they enter the router buffer. The total delay experienced by the packets is the sum of T_p and of the congestion delay

$$d(t) = \min\{d > 0 \text{ such that } \sum_{s=1}^{d} u(t - T_c - s) \geq c(t)\}. \tag{26}$$

To construct a state-space representation for this system, one can retain the state vector x defined by (7), albeit with $n \triangleq d_M + T_c + 1$. Clearly, this representation can be written in the form (8)-(9) for a suitable choice of the matrices A, B, Γ, and with the non-linear function h modified according to (26).

It is immediately checked that in order to obtain a reference trajectory corresponding to a constant congestion delay d^r, the only modification to (12)-(14) should be

$$u^r(t) \triangleq \frac{1}{2}(b(t + d^r + T_c - 1) + b(t + d^r + T_c)) \quad . \tag{27}$$

The following proposition states that for this choice of reference trajectory, asymptotic convergence of both Δx and Δd towards zero can be guaranteed in the presence of input saturation for a suitably modified version of the control law defined in proposition 1:

Proposition 2 *Assume that c is governed by the linear regime (25), that the link bandwidth at all $t \geq 0$ is constrained by (16), and that the source rate control is*

$$u^l(t) = u^r(t) - k(c(t) - c^r(t)) - k\sum_{s=1}^{T_c}(u(t-s) - u^r(t-s)), \tag{28}$$

$$u(t) = Sat_{u_M}\left(u^l(t)\right) \quad , \tag{29}$$

where u^r, x^r are defined by (27), (13)-(14), and where $0 < k \leq 1$. Then for any initial congestion $c(0)$, both Δx and Δd converge towards zero. Furthermore, the convergence of Δx is exponential with guaranteed rate $1 - k_m$, where

$$k_m \triangleq \min\left\{k, \frac{b_m}{\delta}, \frac{u_M - b_M}{\delta}\right\} \quad , \tag{30}$$

$$\delta \triangleq \left|\Delta c(0) + \textstyle\sum_{j=1}^{T_c} \Delta u(-j)\right| \quad , \tag{31}$$

whereas Δd converges to zero in finite time.

Since it involves a few somewhat tedious algebraic manipulations, the proof of this proposition is presented in Appendix A. However, the general conception underlying this particular choice of control feedback is completely straightforward. In fact, (28) is the discrete-time version of a Smith predictor corresponding to the proportional feedback $\Delta u = -k\Delta c$. Thus, in the absence of input saturation, and for the special choice of initial conditions $\Delta u(-j) = 0$ for $1 < j \leq T_c$, this control would result in the closed-loop dynamics

$$\Delta c(t) = (1-k)^{t-T_c} \Delta c(0) \quad . \tag{32}$$

5 Estimation of c from delayed measurements

Let us now assume that measurements of c have to travel back from the buffer to the source using another network route, resulting in a measurement delay, so that the congestion measurement available at time t is

$$m(t) \triangleq c(t - T_m) \quad , \tag{33}$$

where the measurement delay T_m is (at least for the moment) assumed to be both constant and known. Note that in practice, this last assumption can be achieved using the standard Network Time Protocol (NTP).

We now proceed to construct an estimate $\widehat{c}(t)$ of $c(t)$ based on the knowledge of past and present values of m, b and u. We denote as η the partial state

$$\eta(t) \triangleq \begin{pmatrix} c(t) \\ c(t-1) \\ \vdots \\ c(t+1-T_m) \end{pmatrix} \quad . \tag{34}$$

Using this notation, and assuming that the dynamics of c are described by the linear regime (25), we obtain the linear state-space representation

$$\eta(t) = A_m \eta(t-1) + B_m u(t - T_c - 1) + \Gamma_m b(t-1) \, , \tag{35}$$

$$m(t) = C_m \eta(t-1) \quad , \tag{36}$$

with

$$A_m \triangleq \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad , \tag{37}$$

$$B_m \triangleq \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \Gamma_m \triangleq \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad , \tag{38}$$

$$C_m \triangleq \begin{pmatrix} 0 & \cdots & 0 & 1 \end{pmatrix} \quad . \tag{39}$$

The estimation problem is solved by using an observer:

$$\begin{aligned} \widehat{\eta}(t) &= A_m \widehat{\eta}(t-1) + B_m u(t - T_c - 1) + \Gamma_m b(t-1) \\ &\quad + L\left(m(t) - \widehat{m}(t)\right) \quad , \end{aligned} \tag{40}$$

where L is an observer gain and

$$\widehat{m}(t) \triangleq C_m \widehat{\eta}(t-1) \quad . \tag{41}$$

The dynamics of the estimation error $\widetilde{\eta} \triangleq \eta - \widehat{\eta}$ then become

$$\widetilde{\eta}(t) = (A_m - LC_m)\, \widetilde{\eta}(t-1) \quad . \tag{42}$$

Thus, an appropriate choice of L will yield a convergent estimate of c in the form $\widehat{c}(t) = \widehat{\eta}_1(t)$. A suitable solution is to use the dual version of the feedback gain selection approach in section 4:

Proposition 3 *For the linear congestion regime (25) and the measurement equation (33), the observer (40)-(41), with the gain*

$$L = \begin{pmatrix} g \\ \vdots \\ g \end{pmatrix} \quad , \tag{43}$$

where $0 < g \leq 1$, and with an initial condition in the form

$$\widehat{\eta}(0) = \begin{pmatrix} \widehat{c}(0) \\ \widehat{c}(0) + b(-1) - u(-1) \\ \vdots \\ \widehat{c}(0) + \sum_{j=1}^{T_m - 1} \left(b(t-j) - u(t-j)\right) \end{pmatrix} \quad , \tag{44}$$

guarantees that for any trajectory of the control u, the identification error $\widetilde{\eta}$ shall converge exponentially towards zero with rate $1-g$, and that, for $\widehat{c}(t) = \widehat{\eta}_1(t)$,

$$\widetilde{c}(t) \triangleq c(t) - \widehat{c}(t) = (1-g)^t\, \widetilde{c}(0) \quad . \tag{45}$$

The proof is a straightforward adaptation of the proof of proposition 2.

It is interesting to note that this observer design can be adapted almost effortlessly to the case of a variable measurement delay $T_m(t)$. In order to achieve this, one needs only to define η so as to accommodate the maximum anticipated value of $T_m(t)$, and to turn both C_m and L into time-varying vectors.

6 Observer-based output feedback delay control

The feedback delay control law proposed in section 4 may be combined with the observer in section 5 into a standard observer-based control design. Indeed, the exponential stability of the closed-loop system with input saturation and direct state feedback will be preserved for the observer-based feedback. Thus, as a special case of the main theorem in [1], we can state the following property:

Proposition 4 *Let the source rate control be defined as*

$$\begin{aligned} u^l(t) &= u^r(t) - k(\widehat{c}(t) - c^r(t)) \\ &\quad - k \sum_{s=1}^{T_c} (u(t-s) - u^r(t-s)), \end{aligned} \tag{46}$$

$$u(t) = \mathit{Sat}_{u_M}\left(u^l(t)\right) \quad , \tag{47}$$

where $\widehat{c}(t) = \widehat{\eta}_1(t)$ is obtained using the observer in proposition 3. Then under the assumptions of propositions 2 and 3, for any initial congestion $c(0)$, both Δx and Δd converge towards zero, respectively exponentially and in finite time.

7 Extension to N sources sharing a single link

The results presented above can be extended to the scenario where N source-destination pairs share a single buffer/link (figure 2). For the sake of clarity, we shall deal here only with the case $N = 2$, the extension to larger values of N being self-evident.

The N sources scenario differs from the single source case mainly in one important qualitative respect: the data flows received by the different sources now depend not only on the link rate $b(t)$ but also on the internal state itself. If $N = 2$, the linear congestion regime is now described by

$$c(t+1) = c(t) + u_1(t - T_{c1}) + u_2(t - T_{c2}) - b(t), \tag{48}$$

where u_1, u_2 and T_{c1}, T_{c2} are the rates and control delays for the two sources. The appropriate state vector should be

$$x(t) \triangleq \begin{pmatrix} c(t) \\ u_1(t-1) \\ \vdots \\ u_1(t - d_M - T_{c1}) \\ u_2(t-1) \\ \vdots \\ u_2(t - d_M - T_{c2}) \end{pmatrix}, \tag{49}$$

where u_1, u_2 are the rates for the two sources and d_M is the maximum value of the congestion delay. The state dynamics are linear, in the form (8), while the congestion delay is given by

$$d(t) = \min \left\{ \begin{array}{c} d > 0 \text{ such that } \sum_{s=1}^{d} u_1(t - T_{c1} - s) \\ + \sum_{s=1}^{d} u_2(t - T_{c2} - s) \geq c(t) \end{array} \right\}. \tag{50}$$

Let now $y_1(t)$, $y_2(t)$ be the data rates received by the two sources, or more precisely the total number of packets received at each of the two destinations during the time slot $(t-1, t]$. These rates are difficult to predict with absolute accuracy, since they depend on the way the packets emitted by the sources are inserted into the buffer. However, if in any given slice of the buffer the two classes of packets are randomly mixed and small enough compared with the total number of packets emitted during one sampling interval, one can confidently assume that $y_1(t)/y_2(t)$ is close to the corresponding ratio of the input rates at time $t - d(t)$. Then, one can use the approximation

$$y_j(t) = b(t) \times \frac{u_j(t - d(t))}{u_1(t - d(t)) + u_2(t - d(t))} \quad . \tag{51}$$

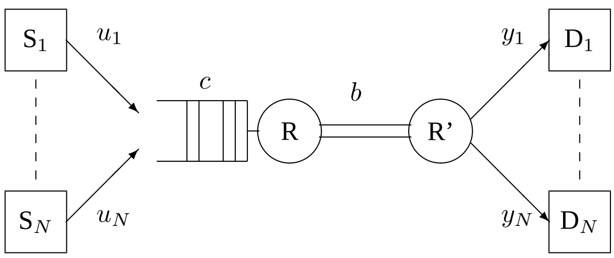

Figure 2: N sources sharing one buffer and one link

Thus, $y_1(t)$ and $y_2(t)$ are non-linear, somewhat complicated but nevertheless piecemeal continuous, functions of $x(t)$ and $b(t)$. As a consequence, any control procedure ensuring that d and u_1, u_2 converge towards prespecified values will also enable to control the output rates y_1, y_2. Keeping in mind this important remark, we now proceed to adapt the control procedures of section 4 to the multi-source case. This can be achieved by defining a suitable scalar feedforward plus feedback control, then splitting it between the various sources. To begin with, we define a scalar reference control as

$$u_0^r(t) \triangleq \frac{1}{2}\left(b(t + d^r - 1) + b(t + d^r)\right) . \tag{52}$$

Assuming that $0 \leq u_1 \leq u_{M1}$ and $0 \leq u_2 \leq u_{M2}$, the reference trajectory x^r is defined through (49), using (14) and

$$u_j^r(t) \triangleq \frac{u_{Mj}}{u_{M1} + u_{M2}} \times u_0^r(t) \quad . \tag{53}$$

Routine calculations show that this defines a proper reference trajectory. The next step is to adapt to the multi-source case the feedback gain design in section 4. It turns out that the only (mild) difficulty is to properly adapt the change of coordinates associated with the control delays T_{c1} and T_{c2}. Finally, one obtains the control

$$\begin{aligned} u_0^l(t) &= u_0^r(t) - k\left(c(t) - c^r(t)\right) \\ &\quad - k \left(\begin{array}{c} \sum_{s=1}^{T_{c1}} \left(u_1(t-s) - u_1^r(t-s)\right) \\ + \sum_{s=1}^{T_{c2}} \left(u_2(t-s) - u_2^r(t-s)\right) \end{array} \right) \end{aligned} \tag{54}$$

$$u_j(t) \triangleq \text{Sat}_{u_{Mj}} \left(\frac{u_{Mj}}{u_{M1} + u_{M2}} \times u_0^l(t) \right) \quad . \tag{55}$$

Because u_1, u_2 necessarily enter and leave their respective lower and upper saturations simultaneously, this control retains all the convergence properties laid out in proposition 2. Thus, if $b_M < u_{M1} + u_{M2}$, it can be guaranteed that as t increases, $\Delta x \to 0$ and $\Delta d \to 0$, so that

$$y_j(t) \to \frac{u_{Mj}}{u_{M1} + u_{M2}} \times u_0^r(t) \quad . \tag{56}$$

8 Perspectives

One open issue is on-line identification of the link bandwidth and competing non-controlled traffic. We should also mention the ability of the control scheme to reject additional disturbances, for example those resulting from the packetized nature of the data flows.

Another important concern is robustness *vis-à-vis* uncertainties/variations in the system's parameters, and especially the control and measurement delays. As usual, one would expect that such robustness concerns lead one to impose limitations to the feedback and observer gains, and thus to limit closed-loop performance. Finally, there is the non trivial problem of evaluating the impact of buffer under- or overflow on the transient responses of the controlled system.

Besides the obvious issue of implementation and performance evaluation of this algorithm over a real network, the perspectives for further developments include “scaling up” this state-space approach to more complex network combinations involving several routers. As an ultimate, yet perhaps unrealistic, goal, one can dream of a comprehensive set of rules which would permit to combine and coordinate elementary “control agents” operating at different locations so as to be able to control any possible network configuration.

References

[1] Bao X, Zongli L, Sontag ED. Finite gain stabilization of discrete-time linear systems subject to actuator saturation. Automatica 2000; 36:269-277.

[2] Benmohamed L, Meerkov SM. Feedback control of congestion in packet switching networks: the case of a single congested mode. IEEE/ACM Transactions on Networking 1993; 6(1): 693-708.

[3] Calvert K. Architectural framework for active networks, version 1.0., 1999: http://www.cc.gatech.edu/projects/-canes/papers/arch-1-0.ps.gz.

[4] Cheng RG, Chang CJ. Design of a fuzzy traffic controller for ATM networks. IEEE/ACM Transactions on Networking 1996; 4(3): 460-469.

[5] Clark D, Tennehouse DL. Architectural considerations for new generation of protocols. ACM symposium on communication architectures & protocols, 1999; 2000-2008.

[6] Delli Priscoli F, Pietrabissa A. Resource management for ATM-based geostationary satellite networks with onboard processing. Computer Networks 2002; 39: 43-60.

[7] Gibbens RJ, Kelly FP. Resource pricing and the evolution of congestion control. Automatica 1999; 35: 1969-1985.

[8] Hammi R, Chen K, Blin JP, Loras F. A framework for responsive control of video communication using active networks. 18th World Telecommunication Congress, 2002.

[9] Imer OÇ, Compans S, Başar T, Skikant R. ABR congestion control in ATM networks. IEEE Control Systems Magazine 2001; 21(1): 38-56.

[10] Jagannathan S, Talluri J. Predictive congestion control of ATM networks: multiple sources/single buffer scenario. Automatica 2002; 38: 815-820.

[11] Mascolo S. Congestion control in high-speed communication networks using the Smith principle. Automatica 1999; 35:1921-1935.

[12] Mounier H, Bastin G. Compartmental modelling for traffic control in communication networks. Submitted to IEEE Transactions on Automatic Control.

A Proof of proposition 2

For the sake of clarity, we shall only deal here with the case $T_c = 3$, the adaptations to the general situation being obvious. Denoting as x^c the vector of the $T_c + 1 = 4$ first coordinates of x, the dynamics of this subsystem are

$$x^c(t+1) = A_c x^c(t) + B_c u(t) + \Gamma_c b(t) \quad , \tag{57}$$

where

$$A_c \triangleq \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad B_c \triangleq \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \Gamma_c \triangleq \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} .$$

Let $\Delta x^c \triangleq x^c - x^{cr}$, where x^{cr} is the vector of the $T_c + 1 = 4$ first coordinates of x^r. Then (57) can be rewritten as

$$\Delta x^c(t+1) = A_c \Delta x^c(t) + B_c \Delta u(t) \quad . \tag{58}$$

The matrix A_c has two eigenvalues: $z = 1$ (simple) and $z = 0$ (multiplicity $T_c = 3$). Since the modes associated to $z = 0$ are stable, we can apply feedback gain design strategy suggested in [1]: map that state transition equation (57) in a system of coordinates corresponding to the Jordan blocs decomposition of A_c, and apply a stabilizing feedback action based on the unstable mode(s), taking advantage of the parallel structure of the modal decomposition. In this case, the appropriate change of coordinates is $\gamma \triangleq P^{-1}\Delta x^c$, where

$$P \triangleq \begin{pmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

Let $\gamma^{[2]}$ be the vector of the $T_c = 3$ last coordinates of γ. In the new coordinates, the open-loop equation (58) becomes

$$\gamma_1(t+1) = \gamma_1(t) + \Delta u(t) \quad , \tag{59}$$

$$\gamma^{[2]}(t+1) = \widetilde{A}_2 \gamma^{[2]}(t) + \widetilde{B}_2 \Delta u(t) \quad , \tag{60}$$

$$\text{with} \quad \widetilde{A}_2 \triangleq \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \widetilde{B}_2 \triangleq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} . \tag{61}$$

We now choose $\Delta u(t) = -k\gamma_1(t)$ as the non-saturated control, which translates in the initial coordinates as $\Delta u(t) = -K_c \Delta x^c(t)$, with $K_c = (k\ \ 0\ \ 0\ \ 0)$, $P^{-1} = (k\ \ k\ \ k\ \ k)$. Applying this control would result in the stable closed-loop dynamics

$$\gamma_1(t+1) = (1-k)\gamma_1(t) \quad , \tag{62}$$

$$\gamma^{[2]}(t+1) = \widetilde{A}_2 \gamma^{[2]}(t) - \widetilde{B}_2 k \gamma_1(t) \quad . \tag{63}$$

The saturation can now be dealt with as in the proof of proposition 1, noting that here too the saturated input can always be expressed as $\Delta u(t) = -k_S(t)\gamma_1(t)$, and that $\Delta c(0)$ should be replaced by $\gamma_1(0) = \Delta c(0) + \sum_{j=1}^{T_c} \Delta u(-j)$.

Finally, since the coordinates of $\Delta x(t)$ not included in Δx^c are in the form $\Delta u(t-j)$, the exponential convergence of Δx^c towards zero implies the exponential convergence of Δx with the same rate, and therefore that $\Delta d \to 0$ in finite time.