Flow Control in Communication Networks with Multiple Users and Action Delays: A Team-Theoretic Approach *

Eitan ALTMAN[†] , Tamer BAŞAR[‡] and Rayadurgam SRIKANT[§]

Abstract

Rate-based flow control has been chosen by the ATM Forum Technical Committee of Traffic Management as the main approach for controlling the traffic of the ABR (Available Bit Rate) transfer capacity. The behavior of the sources and destinations in ABR has been standardized, as well as the way in which information and control commands are conveyed by the ATM switches to the sources and destinations. Yet, the specific control mechanism has not been specified, and left to the designer. The purpose of this paper is to develop a methodology, based on control-theoretic tools, whereby optimal rate-based flow controllers can be obtained. An important issue that arises in this context is one of action delays: it takes time from the moment control information is sent to a source, until an action is taken by it, and until subsequently that action affects the state of the switch that had initiated that information. We first show that general control problems using different *action delays* can be transformed to equivalent team-control problems with different *information delays*. We then obtain an explicit solution to the flow control problem using decentralized LQG (Linear Quadratic Gaussian) team theory. Simulation examples illustrate the analytical results.

1 Introduction

Reactive flow control and congestion control are widely used today in the emerging telecommunications superhighways. They are designed to regulate the transmission of information from a source to a destination so as to avoid unacceptably long delays, high loss rates and other effects of congestion which are undesirable for both the individual user as well as the entire network. Examples of such controls are the TCP/IP congestion control in the Internet [13], and the control of ABR (Available Bit Rate) traffic in ATM (as was standardized by the ATM forum [1]). Such control mechanisms are especially used in regulating "best effort" type traffic, such as data transfer; this type of traffic is often

*Paper submitted for an invited session at the 26th IEEE Conference on Decision and Control, to be held in San Diego, CA, December 10-12, 1997. All correspondence on this submission should be sent to Prof. Tamer Başar, whose address appears below.

†INRIA B.P. 93, 06902 Sophia Antipolis Cedex, France. Tel: (33) 93 65 77 86; Fax: 93 65 77 65. Email: altman@sophia.inria.fr

‡Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA. Tel: (217) 333-3607; Fax: (217) 244-1653; Email: tbasar@decision.csl.uiuc.edu. Research supported by NSF Grants NSF ECS 92-20632, NSF ECS 93-12807, and a Grant from Scientific Systems, Mass.

§Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA. Email: rsrikant@uiuc.edu. Research supported by a Grant from Scientific Systems, Mass.

non real-time, and thus the rate of its transmission can be adapted to the state of the network. This is in contrast with guaranteed bandwidth services (which are typically required by real-time voice or video applications) which have the drawback that only a fixed number of sessions can be handled; when the entire bandwidth is used, new arriving sessions are blocked.

Designing congestion control architectures for Internet-type applications is quite hard since the control is performed by the individual users. We are thus faced with a non-cooperative game framework, where different decision makers control their traffic in a decentralized way, and where the performance measure of each user is affected by the decisions of all other users.

The situation we analyze in this paper is different. It is the one arising in the control of ABR traffic, for which the control actions are taken by the network itself, and not the users. More precisely, the rate at which a user is allowed to transmit is determined by the switches on the virtual circuit between the source and the destination. This gives rise to a simpler theoretical framework of a centralized control, as the control objectives and policies can be determined by a single entity – the network.

Unlike the situation of the Internet, where window-type flow control is used, the control mechanism that has been chosen by the ATM Forum Technical Committee of Traffic Management [1] for controlling ABR traffic is the rate control approach [12]. The behavior of the sources and destinations for ABR traffic has been standardized, as well as the way in which information and control commands are conveyed by the ATM switches to the sources and destinations. Yet, the specific control mechanism has not been specified and left to the designer. The purpose of this paper is to design optimal rate-based flow control for ABR type architectures using control-theoretic tools.

In doing so, we are faced with a situation of action delays: it takes time from the moment control information is sent to a source, until an action is taken by it, and until subsequently that action affects the state of the switch that had initiated that information. We first show that general control problems where a single controller controls several sources with different *action delays* can be transformed to equivalent team-control problems with different *information delays*. We then solve the flow control problem using the theory of decentralized LQG (Linear Quadratic Gaussian) teams. This approach uses an idealized linear queue dynamics, already introduced and justified in [2, 10, 11, 3, 14], and a quadratic cost structure, as was used in the context of a single user control design in [4, 5, 6, 15]. Even though the linear-quadratic structure is essential for the validity of the results obtained in this paper, Gaussiannes (of the statistics of interfering traffic) is by no means essential, and the result remains valid for an ARMA process driven by any finite variance independent sequence.

2 Mathematical model, and derivation of optimal decentralized flow controllers

The model we adopt in this paper is a natural extension of the one considered in [4, 5], from single user to the multiple (M) users case, where the presence of multiple users also brings in a much richer information structure. On the other hand, the uncertainties here are initially assumed to have Gaussian statistics,[1] whereas the ones in the models of [4] and [5] were taken to be general and unknown (which led to a *minimax* formulation). Another difference between the model of this paper and those of [4] and [5] is that here queue length information is noise free (but delayed), whereas in the latter it was acquired through a noisy channel.

More specifically, we consider here a discrete-time model, where a time unit corresponds to a round trip delay. Let q_n denote the queue length at a bottleneck link, and μ_n denote the effective service rate available for traffic of the given source in that link at the beginning of the nth time slot. Let r_{mn} denote the effective source rate for user m ($m = 1, \ldots, M$) during the nth time slot. This effective source rate may actually be the outcome of an action taken by user m several time steps earlier, say at $n - d_m$, where d_m is a positive integer that stands for *action delay*. In our formulation, we will not recognize this delay explicitly, but instead (see below) include a delay factor in the information available to each user for the construction of transmission rates for that user. As shown in *Appendix 1*, these two formulations are equivalent, that is, to any problem that incorporates action delays there corresponds one with information delays, and *vice versa*.

Now, in terms of the notation introduced, the queue length evolves according to

$$q_{n+1} \quad = \quad q_n + \sum_{m=1}^{M} r_{mn} - \mu_n \,. \tag{1}$$

Since several sources with varying transmission rates may share the same bottleneck link, the service rate μ_n available to the M users may change over time in an unpredictable way. The other users (e.g., CBR and VBR traffic) may then be represented as some interference, which we model as generated by a p-dimensional stable ARMA process:

$$\mu_n \quad = \quad \mu + \xi_n \tag{2}$$

$$\xi_n \quad = \quad \sum_{i=1}^{p} \alpha_i \, \xi_{n-i} + \phi_{n-1}, \quad \phi_n \sim N(0, k^2)\,, \tag{3}$$

where μ is the constant nominal service rate (known to all users), and α_i, $i = 1, \ldots, p$, and k are also known parameters. $\{\phi_n\}_{n \geq 1}$ is a zero-mean *i.i.d.* Gaussian sequence with variance k^2.

[1] As discussed in Appendix 2, the results obtained in this paper are valid even if this assumption does not hold, as long as the noise sequence driving the ARMA process (to be introduced shortly) is independent and has finite variance.

The objective function to minimize (by all M users) is

$$J = \limsup_{N\to\infty} \frac{1}{N} E\left\{ \sum_{n=1}^{N} \left[(q_n - Q)^2 + \sum_{m=1}^{M} \frac{1}{c_m^2} (r_{mn} - a_m \mu_n)^2 \right] \right\} \tag{4}$$

where Q is the target queue length, c_m's are some positive constants, and

$$\sum_{m=1}^{M} a_m = 1 .$$

The *first additive term* above represents a penalty for deviating from a desirable queue length. Setting some desired level of queue reflects the fact that we do not wish to have a large queue, so as to avoid losses, and on the other hand, we do not wish to have a low level of queue, since if the queue is empty and the input rate is lower than the service rate then there is a waste of potential throughput. We have taken the weighting on this term to be 1, as any other positive value for the positive weight can be absorbed into the other variables by an appropriate normalization. The *second additive term*, on the other hand, is a measure of the quality with which the input rate for each user tracks a given fraction of the available service rate, where the c_m's are weighting terms that serve to prioratize relative importances of these individual terms (among different users as well as collectively with respect to the first additive term). For example, if we desire "fair" sharing of the available bandwidth, we would choose

$$a_1 = a_2 = \cdots = a_M = \frac{1}{M} ,$$

assuming that everything else is also symmetric for the users. If the feedback delay is different for different users, then the c_m's will have to be picked unequal accordingly. These tradeoffs will be illustrated through examples in Section 3. The information available to user m at time n is I_{n-d_m}, where

$$I_n := \{q_1, q_2, \ldots, q_n\} \tag{5}$$

and d_m's denote delays in the acquisition of queue length information. Note that a crucial assumption here is that the delay for each user is independent of n. In view of this, and without any loss of generality, we can take d_m's to be ordered in the following way:

$$0 \le d_1 \le d_2 \le \cdots \le d_M .$$

Hence,

$$r_{mn} = \tilde{g}_{mn}(I_{n-d_m}),\ n = 1, 2, \ldots,; \quad m = 1, 2, \ldots, M,$$

where $\tilde{g}_{mn}$ is some measurable function, with respect to which J will be minimized.

We now introduce the new (appropriately shifted) variables

$$x_n \quad := \quad q_n - Q \tag{6}$$

$$u_{mn} \quad := \quad r_{mn} - a_m \mu \tag{7}$$

which will serve as the state and the control, respectively. The queue dynamics (1) can be re-written in terms of these quantities as:

$$x_{n+1} \quad = \quad x_n + \sum_{m=1}^{M} u_{mn} - \xi_n \tag{8}$$

$$\xi_{n+1} \quad = \quad \sum_{i=1}^{p} \alpha_i \, \xi_{n+1-i} + \phi_n \tag{9}$$

and the cost function J as:

$$J \quad = \quad \limsup_{N \to \infty} \frac{1}{N} E \left\{ \sum_{n=1}^{N} \underbrace{\left[(x_n)^2 + \sum_{m=1}^{M} \frac{1}{c_m^2} (u_{mn} - a_m \, \xi_n)^2 \right]}_{\ell_n(x_n, u_{mn})} \right\} \tag{10}$$

where ℓ_n denotes the instantaneous cost. Note further that (9) can equivalently be written as

$$\xi_n \quad = \quad (0, 0, \ldots, 0, 1)_{1 \times p} \; \eta_n \equiv D' \, \eta_n \tag{11}$$

$$\eta_{n+1} \quad = \quad C \, \eta_n + D \, \phi_n \tag{12}$$

where

$$C := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ \alpha_p & \alpha_{p-1} & \cdot & \cdots & \cdot & \alpha_1 \end{pmatrix}_{p \times p} , \qquad D = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix}_{p \times 1} . \tag{13}$$

Clearly,

$$\eta_n = (\xi_{n-p+1}, \cdots, \xi_n) \; . \tag{14}$$

Since the original ARMA process was assumed to be stable, the $p \times p$ matrix C above is a stable matrix (*i.e.* all its eigenvalues are in the unit circle).

Now introduce the notation:

$$z_n := \begin{pmatrix} x_n \\ \eta_n \end{pmatrix}, \quad u_n := \begin{pmatrix} u_{1n} \\ \vdots \\ u_{mn} \end{pmatrix}, \quad \tilde{u}_n := \begin{pmatrix} u_{1n} - a_1 \xi_n \\ \vdots \\ u_{mn} - a_m \xi_n \end{pmatrix}.$$

Then, z_n is generated by

$$\begin{aligned} z_{n+1} &= Az_n + Bu_n + D\phi_n \\ &= \widetilde{A}z_n + B\tilde{u}_n + D\phi_n \end{aligned} \tag{15}$$

where A and $\widetilde{A}$ are $(p+1) \times (p+1)$ matrices, and B is a $(p+1) \times M$ matrix, defined as follows:

$$A := \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & & & & \\ \cdot & & C & & \\ 0 & & & & \end{pmatrix}, \quad \tilde{A} := \begin{pmatrix} 1 & 0 & \cdots & 0 & -\bar{a} \\ 0 & & & & \\ \cdot & & C & & \\ 0 & & & & \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

with

$$\bar{a} := 1 + \sum_{i=1}^{M} a_i \, .$$

Furthermore, the instantaneous cost ℓ_n can be expressed, in terms of this new notation, as:

$$\ell_n = |z_n|_P^2 + |\tilde{u}|_R^2 \tag{16}$$

where $|\cdot|$ denotes the Euclidean norm in an appropriate finite-dimensional vector space, and

$$P := \text{ diag } (1, 0, \cdots, 0) \ , \quad R := \text{ diag } \left(\frac{1}{c_1^2}, \cdots, \frac{1}{c_M^2} \right) . \tag{17}$$

Note that $(\widetilde{A}, B)$ is a stabilizable pair (as C, whose p eigenvalues are also the eigenvalues of $\widetilde{A}$,is a stable matrix and the other eigenvalue of $\widetilde{A}$, 1, is controllable), and $(\widetilde{A}, P)$ is a detectable pair. Hence, if z_n were available at time n, then the optimum minimizing solution for $\tilde{u}$ would be (from standard discrete-time linear regulator theory [8]):

$$\tilde{u}_n = -[R + B'SB]^{-1}B'S\widetilde{A}\, z_n \tag{18}$$

where $S > 0$ is the unique solution of

$$S = P + \widetilde{A}'S[I - B(R + B'SB)^{-1}B'S]\widetilde{A} \tag{19}$$

in the class of nonnegative-definite matrices. Furthermore, this control renders the closed-loop system

$$z_{n+1} = [I - B[R + B'SB]^{-1}B'S]\widetilde{A}\, z_n + D\, \phi_n$$

asymptotically stable.

Note that the corresponding choice for u_n is

$$u_n = \tilde{u}_n + \begin{pmatrix} 0 & 0 & \cdots & 0 & a_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_M \end{pmatrix} z_n \equiv -K z_n \tag{20}$$

where

$$K := \begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_M \end{bmatrix} = [R + B'SB]^{-1} B'S\widetilde{A} - \begin{pmatrix} 0 & 0 & \cdots & 0 & a_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_M \end{pmatrix}. \tag{21}$$

For the individual user, m, we can write down the corresponding controller as

$$u_{mn} = -K_m \, z_n. \tag{22}$$

The state z_n is however not available to u_{mn}, but only I_{n-d_m} is. What we have here is a decentralized optimal control problem, but since the information is *nested* i.e., what is available to user m at time n is available to all users $m' > m$), a separation result holds [9], allowing us to replace z_n in (22) by the conditional mean of z_n given I_{n-d_m}. Hence, the optimum minimizing solution under the original information is

$$\begin{aligned} u^*_{mn} &= g^*_{mn}(q_{n-d_m}, q_{n-d_m-1}, \ldots) \\ &= -K_m \, \hat{z}^m_{n|n-d_m} \end{aligned} \tag{23}$$

where

$$\hat{z}^m_{n|n-d_m} = \zeta^m_{n,d_m} \tag{24}$$

$$\zeta^m_{n,\ell+1} = A\zeta^m_{n,\ell} + \begin{bmatrix} \sum_{i=1}^M \hat{u}^m_{i,n+\ell-d_m} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; \zeta^m_{n,0} = \hat{z}_{n-d_m|n-d_m} \tag{25}$$

$$\ell = 0, 1, \ldots, d_m$$

$$\hat{u}^m_{i,n+\ell-d_m} := \begin{cases} -K_i \zeta^m_{n,\ell} & \text{if } \ell > d_i \\ u^*_{i,n+\ell-d_m} & \text{if } \ell \leq d_i \end{cases} \tag{26}$$

$$\hat{z}_{k|k} = \left(\hat{x}_{k|k}, \hat{\eta}'_{k|k}\right)', \tag{27}$$

where

$$\hat{x}_{k+1|k+1} = x_{k+1}$$

$$
\begin{aligned}
\hat{\eta}_{k|k} &= \left(\xi_{k+1-p|k}, \xi_{k+2-p|k}, \ldots, \xi_{k|k}\right) \\
\hat{\xi}_{k+1|k+1} &= -\sum_{i=1}^{p} \alpha_i \left[x_{k+2-i} - x_{k+1-i} - \sum_{m=1}^{M} u^*_{m,k+1-i}\right] \\
\hat{\xi}_{k+1+j-p|k+1} &= -\left(x_{x+2+j-p} - x_{k+1+j-p} - \sum_{m=1}^{M} u^*_{m,k+1+j-p}\right), \quad j = 1, \ldots, p-1. \qquad (28)
\end{aligned}
$$

These are the recursive equations generating the estimators for the users. Details of this derivation are included in *Appendix 2*. We should note that the quantity $\hat{z}_{k|k}$ above stands for the (*centralized*) best estimate (in the least squares sense – see *Appendix 2*) of z_k based on the measurements $x_{k+1}, x_k, \ldots$ common to all users. Now, to summarize:

Theorem 1 *For the decentralized multi-user communication problem with delayed and partially shared queue length information, there exists a unique set of controllers that minimizes the performance index J defined by (4), which is given for user m by*

$$
\begin{aligned}
u^*_{mn} &= g^*_{mn}(q_{n-d_m}, q_{n-d_m-1}, \ldots) \\
&= -K_m \, \hat{z}^m_{n|n-d_m}
\end{aligned}
$$

where K_m is defined by (21), and $\hat{z}^m_{(\cdot|\cdot)}$ is generated by (24), (25), and (28). $\diamond$

3 Simulation Results

To illustrate the performance of the controller derived in the previous section, we present simulation results for the case where there are three ABR users who are attempting to fairly share the available bandwidth. The examples study the effect of the delays d_m, and the choice of the weights c_m, on the *centered* queue length x_n and the centered bandwidth used by user m, $\tilde{u}_{mn}$ $\forall$ $m = 1, 2, 3$. In the examples considered in this section, we will assume $a_1 = a_2 = a_3 = 1/3$, and the noise variance $k^2 = 1$.

Example 1: We choose $d_1 = d_2 = d_3 = 0$, and $c_1 = c_2 = c_3 = 1$. Thus, this is a case with no delay which can be used later to compare with cases where at least one of the delays is non-zero. Figures 1 and 2 plot x_n, $\tilde{u}_{1n}$ and $\tilde{u}_{2n}$ as a function of n. Since the plots for $\tilde{u}_{1n}$, $\tilde{u}_{2n}$ and $\tilde{u}_{3n}$ are identical, the ones for $\tilde{u}_{2n}$ and $\tilde{u}_{3n}$ are not shown here.

Example 2: We choose $d_1 = d_2 = d_3 = 10$, and $c_1 = c_2 = c_3 = 1$. Comparing this with Example 1, we allow the feedback delay for both users to be non-zero, however we assume that the delay is equal for both users. Figures 3 and 4 show that the variability around zero is more for these plots as compared to the corresponding plots for Example 1. Thus, the presence of delay has made it harder to *regulate* the queue at the desired operating point.

Figure 1: x_n vs. n, with $d_1 = d_2 = d_3 = 1$, $c_1 = c_2 = c_3 = 1$.

Figure 2: $\tilde{u}_{1n}$ vs. n, with $d_1 = d_2 = d_3 = 1$, $c_1 = c_2 = d_3 = 1$.

Figure 3: x_n vs. n, with $d_1 = d_2 = d_3 = 10$, $c_1 = c_2 = c_3 = 1$.

Figure 4: $\tilde{u}_{1n}$ vs. n, with $d_1 = d_2 = d_3 = 10$, $c_1 = c_2 = c_3 = 1$.

Figure 5: x_n vs. n, with $d_1 = 0$, $d_2 = 5$, and $d_3 = 10$, $c_1 = c_2 = c_3 = 1$.

Example 3: Here we consider the case where $d_1 = 0$, $d_2 = 5$, and $d_3 = 10$, and $c_1 = c_2 = c_3 = 1$. This is the case that is most pertinent to the main results of this paper. The plots of $\tilde{u}_{1n}$, $\tilde{u}_{2n}$ and $\tilde{u}_{3n}$ in Figures 6, 7 and 8 are no longer identical. However, they both oscillate in a fairly symmetric fashion around zero, indicating that the long-run average of $\tilde{u}_{1n}$, $\tilde{u}_{2n}$ and $\tilde{u}_{3n}$ are zero. Recalling that $\tilde{u}_{mn}$ denote the fluctuations around one-third of the available bandwidth (since $a_1 = a_2 = a_3 = 1/3$), this indicates that the bandwidth sharing between the users is still fair in the long-term even when the delays are unequal. More surprising is the effect of unequal delay on the variability of $\tilde{u}_{1n}$ and $\tilde{u}_{2n}$ around zero. Roughly, all the users have similar variability around zeoo! Even though user 1 has a smaller delay, since this is a team problem, it also "works harder" to keep the queue length at the desired level. Hence, its variability is actually slightly larger compared to the other two users despite the fact that its delay is smaller. Finally, note that, in Figure 5, the variability in x_n is larger than what was observed in Example 1 but smaller than in Example 2.

4 Conclusions

We have studied rate-control in high-speed networks using a team-theoretic framework. Specifically, we have addressed the problem where a single bottleneck node issues rate control messages to multiple users. The action delay (the amount of elapsed time from the instant a control action is initiated by the node to the instant at which it receives bits at the appropriate rate from the user) is different for the different users. Our main contribution is to obtain an optimal rate controller for the case with different delays, which we believe is new. Using a first-order ARMA process driven by a Gaussian noise as our model for the available bandwidth at the node, we have derived the optimal rate controls using the fact that the information is nested. An appealing feature of the controller is the fact that

Figure 6: $\tilde{u}_{1n}$ vs. n, with $d_1 = 0$, $d_2 = 5$, $d_3 = 10$, and $c_1 = c_2 = 1$.

Figure 7: $\tilde{u}_{2n}$ vs. n, with $d_1 = 0$, $d_2 = 5$, $d_3 = 10$, and $c_1 = c_2 = c_3 = 1$.

Figure 8: $\tilde{u}_{3n}$ vs. n, with $d_1 = 0$, $d_2 = 5$, $d_3 = 10$, and $c_1 = c_2 = c_3 = 1$.

it does *not* depend upon the variance of the noise that derives the ARMA process.

We have restricted our analysis here to the case where perfect queue length information (q_n) is used, but no information about the available bandwidth (μ_n). The results presented can however be extended in a straightforward manner to the cases where only imperfect information is available about q_n and either perfect or imperfect information is available about μ_n, with the "imperfectness" caused in each case by the presence of independent additive Gaussian sequence(s). While more information on μ_n would improve the performance of the system, it would not significantly change the structure of the controller.

We have presented simulation results to show that team theory can be used to design optimal rate controllers that are fair even when the delays experienced by the users are different. There are several avenues for future research. We do not have expressions relating the weights c_i to queue length or the variability in the bandwidth utilized by the different users. Obtaining this would be an important area of future work since this would allow one to immediately relate c_i to the delay d_i. Secondly, we have assumed that the delays are known perfectly. While our controller can be used even when the delay is not precisely known, by basing it on some rough estimate, it would be important to analyze the sensitivity the performance measures of interest when the delay is variable. A third area of research would be to treat the case where different sources may have different bottleneck nodes. We may then assume that each switch has a table of those connections for which that switch is the bottleneck; this can be done by using the CCR (Current Cell Rate) information on RM cells, where the users indicate their current transmission rate, and comparing these rates to the ACR. If the CCR of source m is smaller than the ACR issued to source m, d_m time units earlier, then the switch would conclude that it is not a bottleneck node for that source. The control would then be computed only for those sources for which the switch is the bottleneck. It then remains to check by simulations whether this approach would give good performance, and if not, explore ways of improving it.

5 Appendix 1: Equivalence of action and information delays

We prove in this *Appendix* the equivalence of dynamic team problems with two types of delays, namely:
(i) *information delay*: at time t the controller knows about the state of the system only until time $t - N_1$, where $N_1 > 0$ is the corresponding delay.
(ii) *action delay*: the action taken by the controller at time t influences the state of the system only at time $t + N_2$where $N_2 > 0$ is the corresponding delay.

The equivalence we will establish is in the following sense: A team problem with action delays between a central controller and several controlled units can be transformed into an equivalent team problem with no action delays but with different information delays.

Discrete-time control problem with action delays.

Considering a class of team problems more general than the one formulated in the main body of the paper, let there be M units (e.g. users) controlled by a central controller. Define a vector $\theta = (\theta_1, ..., \theta_M)$ of action-delays: the impact of a control decision taken by the central controller on the mth unit is affected only θ_m time units later, as will become clear in (29) below.

Let $\mathbf{X}$ be some state-space, $\mathbf{A} = (\mathbf{A}_1 \times \mathbf{A}_2 \times \ldots \times \mathbf{A}_M)$ a product action space, and $\mathcal{Y}$ the space of measurements. Let x_n be the state at time n, and u_{mn} the action taken by unit m at time n. We define the vector of "relevant actions" at time n as $w_n = (w_{1n}, ..., w_{Mn}) := (u_{1,n-\theta_1}, u_{2,n-\theta_2}, \ldots, u_{M,n-\theta_M})$. The state of the system at time $n+1$ will be determined by the state at time n, the relevant actions w_n at time n, and some other quantity (such as random noise), ϵ_n, taking values in some space $\mathcal{E}$:

$$x_{n+1} = f_n(x_n, w_n, \epsilon_n), \qquad n \in I\!N := \{1, 2, \ldots\}, \tag{29}$$

where f_n is a given function $f_n : \mathbf{X} \times \mathbf{A} \times \mathcal{E} \to \mathbf{X}$.

The measurement y_n at time n is a function of the current system state, x_n, and some other quantity (such as random noise), v_n, taking values in some space $\mathcal{V}$; that is,

$$y_n = h(x_n, v_n), \qquad n \in I\!N, \tag{30}$$

for some functions $h_n : \mathbf{X} \times \mathcal{V} \to \mathcal{Y}$.

The composite control is a sequence of vectors $\{u_{mn}, n \in I\!N, m = 1, ..., M\}$, where u_{mn} is the command sent to unit m at time n, which is a function of the available information:

$$u_{mn} = g_{mn}(\{y_j\}, 1 \le j \le n).$$

A discrete-time team control problem with delayed information.
Now we consider the situation of M different controllers that are faced with a team problem. The information for controller m on the system state is acquired after a delay of θ_m time units.

We consider the same spaces $\mathbf{X}, \mathbf{A}$ and $\mathcal{Y}$ of states, actions and measurements, respectively, as well as the spaces of driving sequences, $\mathcal{E}$ and $\mathcal{V}$. Let $\overline{x}_n$ be the state at time n, and $\overline{u}_{mn}$ the action taken by controller m at time n. Define the vector of actions at time n by $\overline{w}_n = (\overline{u}_{1n}, \overline{u}_{2n}, \ldots, \overline{u}_{Mn})$. The state of the system at time $n+1$ will be determined by the state and actions at time n and the state of the environment, ϵ_n:

$$\overline{x}_{n+1} = f_n(\overline{x}_n, \overline{w}_n, \epsilon_n), \qquad n \in I\!N, \tag{31}$$

where f_n and ϵ_n are defined as in (29). The measurement sequence $\{\overline{y}_n\}$ and the noise sequence $\{v_n\}$ are related to each other as in (30):

$$\overline{y}_n = h(\overline{x}_n, v_n), \qquad n \in I\!N. \tag{32}$$

The composite control is a sequence of vectors $\{u_{mn}, n \in I\!N, m = 1, ..., M\}$, where u_{mn} is the command taken by unit m at time n, which is a function of the available information:

$$\overline{u}_{mn} = \overline{g}_{mn}(\left\{\overline{y}_j\right\}, 1 \le j \le n - \theta_m),$$

with the convention adopted being that for $n < \theta_m$, $\overline{u}_{mn}$ is just a constant element of $\mathbf{A}_m$

The equivalence between the two models

Theorem 2 *Suppose that*

$$\overline{x}_n = x_n, \quad n = 1, 2, ..., n_0$$

where $n_0 := \max_i \theta_i$, *and*

$$w_{mn} = \overline{u}_{mn}, \quad 1 \le n < \theta_m,$$

and let $\overline{g}_{mn} = g_{mn}$ *for all* m *and* $n \ge n_0$. *Then the same state, measurement and action trajectories are obtained in the two systems:*

$$x_n = \overline{x}_n, \quad w_n = \overline{u}_n, \quad y_n = \overline{y}_n, \quad n = 1, 2, ...$$

Proof: The assumptions of the Theorem imply that $y_j = \overline{y}_j$ for $1 \le j \le n_0$. This, together with the fact that $\overline{g}_{mn} = g_{mn}$ implies that

$$w_{mn} = \overline{u}_{mn}, \quad m = 1, ..., M \tag{33}$$

for $n = \theta_m$, which in turn implies that

$$x_n = \overline{x}_n, \qquad \text{and} \qquad y_n = \overline{y}_n \tag{34}$$

for $n = n_0 + 1$. Continuing by induction, (33) and (34) are then seen to hold for all n. ∎

When optimal control problems are considered, Theorem 2 can be used to transform a centralized control problem with different action delays into a team problem with different information delays, provided that the corresponding costs are defined in a compatible way; more precisely, if the cost to be minimized in the first problem is a function of the trajectories $\{x_n, w_n, n \in I\!N\}$, and the cost in the second problem is the same function of the trajectories $\{\overline{x}_n, \overline{w}_n, n \in I\!N\}$, then any optimal policy for one yields (through Theorem 2) an optimal policy for the other problem.

6 Appendix 2: Proof of Theorem 1

In this *Appendix*, we provide the details of the derivation that led to Theorem 1 in Section 2. Specifically, we want to show that the least squares estimate of z_n for user m, given his d_m unit delayed information on the queue length, I_{n-d_m}, that is

$$\hat{z}^m_{n|n-d_m} = E[z_n | I_{n-d_m}] \tag{35}$$

is generated by (24)–(27).

Let us first consider the construction of a central estimator, which uses nondelayed queue length information:

$$\hat{z}_{k|k} := E[z_k | q_k, q_{k-1}, \ldots] \tag{36}$$

where we take $u_{m\ell}$, $m = 1, \ldots, M$, $\ell \leq k$, to be fixed as a measurable function of $\{q_\ell, q_{\ell+1}, \ldots\}$. Now, (36) is a *singular* estimation problem since the first component of z_k is measured perfectly. A way to get around the difficulty caused by this nonstandard nature of the estimation pronblem is first to introduce a noisy measurement process:

$$\begin{aligned} y_k &= c' z_k + v_k \\ c' &= (1\,,\,0\,,\,\ldots\,,\,0)\,, \quad v_k \sim N(0, \epsilon) \end{aligned}$$

where $\{v_k\}$ is an *i.i.d.* Gaussian sequence with a small variance $\epsilon > 0$, then apply the standard Kalman filtering theory to the resulting nonsingular estimation problem, and finally let $\epsilon \to 0$. An alternative method would be to write $\hat{z}_{k|k}$ in (36) as

$$\hat{z}_{k|k} = h_k(q_k, q_{k-1}, \ldots)$$

where h_k is a 2-dimensional measurable function of its arguments, and to minimize the following quadratic error with respect to h_k:

$$E\left\{ |h_k(q_k, q_{k-1}, \ldots) - z_k|^2 \right\} \tag{37}$$

where $|\cdot|$ denotes the Euclidean norm. Let h_k^i denote the $i'th$ component of h_k. Then, the minimum of (37) is

$$
\begin{aligned}
\min_{h_k^1,\ldots,h_k^{p+1}} \quad & E\left\{|h_k^1(q_k,q_{k-1},\ldots)-x_k|^2+\sum_{i=1}^{p}|h_k^{i+1}(q_k,q_{k-1},\ldots)-\xi_{k-p+i}|^2\right\} \\
= & \min_{h_k^2,\ldots,h_k^{p+1}} E\left\{\sum_{i=1}^{p-1}|h_k^{i+1}(q_k,q_{k-1},\ldots)-\xi_{k-p+i}|^2+|h_k^{p+1}(q_k,q_{k-1},\ldots)-\sum_{i=1}^{p}\alpha_i\xi_{k-i}-\phi_{k-1}|^2\right\} \\
= & \min_{h_k^{p+1}} E\left\{|h_k^{p+1}(q_k,q_{k-1},\ldots)-\sum_{i=1}^{p}\alpha_i\xi_{k-i}|^2\right\}+\mathrm{var}\ (\phi_{k-1}) \\
= & \ \mathrm{var}\ (\phi_{k-1})
\end{aligned}
$$

where the second-to-last line has followed because $\{\phi_k\}$ is an independent sequence, and in particular ϕ_{k-1} is independent of $q_k, q_{k-1}, \ldots$; furthermore, we have picked here

$$
h_k^{i+1}(q_k,q_{k-1},\ldots)=-\left(x_{k+1+i-p}-x_{k+i-p}-\sum_{m=1}^{M}u_{m,k+i-p}^*\right)\ , \quad i=1,\ldots,p-1\,,
$$

to make the first term zero. The last line has followed because of the fact that the specific choice

$$
h_k^{p+1}(q_k,q_{k-1},\ldots)=-\sum_{i=1}^{p}\alpha_i\left(x_{k+1-i}-x_{k-i}-\sum_{m=1}^{M}u_{m,k-i}^*\right)
$$

makes the first term in the second-to-last line *zero.* Likewise, in going from the first to the second line above, we picked

$$
h_k^1(q_k,q_{k-1},\ldots)=x_k=q_k-Q\,.
$$

Hence,

$$
\hat{z}_{k|k}=\begin{pmatrix} x_k \\ \hat{\eta}_{k|k}\end{pmatrix} \tag{38}
$$

where $\hat{\eta}_{k|k}$ is generated by (28). We should note that an ϵ-noise perturbation argument would have led to the same result.

We now turn to the delayed measurement case. Let us focus on a generic user m, at time instant n, with the policies prior to n determined according to (23). Then, letting

$$
\hat{z}_{n|n-d_m}^m=E[z_n|q_{n-d_m},q_{n-d_m-1},\ldots]=:\begin{pmatrix}\hat{x}_{n|n-d_m}^m \\ \hat{\eta}_{n|n-d_m}^m\end{pmatrix}\,,
$$

we have, from (8) and (9),

$$\hat{x}^m_{n|n-d_m} = \hat{x}^m_{n-1|n-d_m} - \sum_{i=1}^{m-1} K_i E[\hat{z}^i_{n-1|n-d_i-1}|I_{n-d_m}] + \sum_{i=m}^{M} u^*_{i,n-1} - \hat{\xi}^m_{n-1|n-d_m} \tag{39}$$

$$\hat{\xi}^m_{n|n-d_m} = \sum_{i=1}^{p} \alpha_i \, \hat{\xi}^m_{n-i|n-d_m} \tag{40}$$

where the last relationship can be re-written in terms of η as:

$$\hat{\eta}^m_{n|n-d_m} = C\, \eta^m_{n-1|n-d_m}\,; \quad \hat{\xi}^m_{n|n-d_m} = D' \hat{\eta}^m_{n|n-d_m}\,. \tag{41}$$

Here, we have made use of the initial hypothesis that the information is nested, that is user m has access to all the information available to users i, $i > m$, but not necessarily to the information of users j, $j < m$. Now, a recursive application of (41) leads to:

$$\hat{\eta}^m_{n|n-d_m} = C^{d_m} \hat{\eta}^m_{n-d_m|n-d_m}$$

Furthermore, since for $i < m$, we have

$$\begin{aligned} K_i E[\hat{z}^i_{n-1|n-d_i-1}|I_{n-d_m}] &= K_i E[\, E[z_{n-1}|I_{n-1-d_i}]|I_{n-d_m}] \\ &= \left\{ \begin{array}{ll} K_i E[z_{n-1}|I_{n-1-d_i}] = -u^*_{i,n-1} & \text{if } d_i = d_m \\ K_i E[z_{n-1}|I_{n-d_m}] & \text{if } d_i < d_m \end{array} \right. =: -\hat{u}^m_{i,n-1}\,, \end{aligned}$$

this results in the following equivalent expression for (39):

$$\hat{x}^m_{n|n-d_m} = \hat{x}^m_{n-1|n-d_m} + \sum_{i=1}^{m-1} \hat{u}^m_{i,n-1} + \sum_{i=m}^{M} u^*_{i,n-1} - \hat{\xi}^m_{n-1|n-d_m}\,.$$

A similar reasoning leads to (provided that $d_m \geq 2$):

$$\hat{x}^m_{n-1|n-d_m} = \hat{x}^m_{n-2|n-d_m} + \sum_{i=1}^{m-1} \hat{u}^m_{i,n-2} + \sum_{i=m}^{M} u^*_{i,n-2} - \hat{\xi}^m_{n-2|n-d_m}\,,$$

where, by a slight abuse of notation,

$$\hat{u}^m_{i,n-2} := \left\{ \begin{array}{ll} u^*_{i,n-2} & \text{if } d_i \geq d_m - 1 \\ -K_i E[z_{n-2}|I_{n-d_m}] & \text{if } d_i < d_m - 1 \end{array} \right.$$

For a general $\ell' \geq 0$ such that $\ell' \leq d_m$, defining

$$\hat{u}^m_{i,n-\ell'} := \left\{ \begin{array}{ll} u^*_{i,n-\ell'} & \text{if } d_i \geq d_m + 1 - \ell' \\ -K_i E[z_{n-\ell'}|I_{n-d_m}] & \text{if } d_i < d_m + 1 - \ell' \end{array} \right. ,$$

we have

$$\hat{x}^m_{n-\ell'+1|n-d_m} = \hat{x}^m_{n-\ell'|n-d_m} + \sum_{i=1}^{m-1} \hat{u}^m_{i,n-\ell'} + \sum_{i=m}^{M} u^*_{i,n-\ell'} - \hat{\xi}^m_{n-\ell'|n-d_m} .$$

Now let

$$\zeta^m_{n,\ell} := \begin{pmatrix} \hat{x}^m_{n-\ell'|n-d_m} \\ \hat{\eta}^m_{n-\ell'|n-d_m} \end{pmatrix}, \quad \ell = d_m - \ell' .$$

Then, $\zeta^m_{n,\ell}$ is generated by (25), with the initial condition generated by (27). This completes the proof of the theorem.

Note that the solution is independent of the variance of $\{\phi_k\}$, and in the derivation above we have not made any use of the Gaussian statistics of this process, other than that the second moment is finite. The only place where Gaussianness was used in Section 2 was when certainty-equivalence was invoked in going from (22) to (23), which was necessitated by the existing theory on linear-quadratic stochastic teams. In our present set-up, however, because of the special structures of the system dynamics and the measurement process (that is, queue length information is available noise free, but with delay, and the information structure is nested), a similar result can be obtained for an arbitrary finite variance independent random sequence $\{\phi_k\}$. Details will not be given here.

References

[1] The ATM Forum Technical Committee, *Traffic Management Specification*, Version 4.0, 95-0013R8, April 1996.

[2] O. Ait-Hellal, E. Altman and T. Başar, "Rate based flow control with bandwidth information", *Proc. 35th IEEE Conference on Decision and Control*, pp. 2911-2916, Kobe, Japan, Dec. 1996.

[3] E. Altman, F. Baccelli and J. C. Bolot, "Discrete-time analysis of adaptive rate control mechanisms", *High Speed Networks and their performance*, H.G. Perros and Y. Viniotis Eds., North Holland, 121-140, 1994.

[4] E. Altman and T. Başar, "Optimal Rate Control for High Speed Telecommunication Networks", *Proc. 34th IEEE Conference on Decision and Control*, pp. 1389–1394, New Orleans, Louisiana, Dec. 1995.

[5] E. Altman and T. Başar, "Optimal Rate Control for High Speed Telecommunication Networks: the case of delayed information", First Workshop on ATM Traffic Management, WATM'95, IFIP, WG.6.2 Broadband Communication, Paris, pp. 115-122, December 1995.

[6] E. Altman, T. Başar, Z. Pan, "Admission and flow control in telecommunication networks as a hybrid control problem" *Proc. 1996 Conference on Information Sciences and Systems*, vol: II, pp. 705-710, Princeton, NJ, 1996.

[7] E. Altman and S. Stidham, "Optimality of monotonic policies for two-action Markovian decision processes, with applications to control of queues with delayed information", invited paper, *QUESTA* **21**, special issue on optimization of queueing systems, 267-291, 1995.

[8] B.D.O. Anderson and J.B. Moore, *Optimal Control: Linear Quadratic Methods*, Prentice-Hall, N.J., 1990.

[9] T. Başar and J. B. Cruz, Jr. Concepts and methods in multiperson coordination and control. In S. G. Tzafestas, editor, *Optimization and Control of Dynamic Operational Research Models*, chapter 11, pp. 351–387. North Holland, 1982.

[10] J-C. Bolot, "A self-tuning regulator for overload control in communication networks," *Proc. 31st Conf. on Decision and Control*, Tucson, AZ, pp. 1022-1023, Dec. 1992.

[11] J-C. Bolot and A. U. Shankar, "Analysis of a fluid approximation to flow control dynamics," *IEEE INFOCOM '92*, Florence, Italy, May 1992, pp. 2398-2407.

[12] F. Bonomi and K. W. Fendick, "The rate-base flow control framework for the Available Bit Rate ATM service", *IEEE Network*, pp. 25-39, March/April 1995.

[13] V. Jacobson, "Congestion avoidance and control", *ACM SIGCOMM '88*, Stanford, CA, 314-329, August 1988.

[14] S. Keshav, "A control-theoretic approach to flow control", *ACM SIGCOMM '91*, Zurich, Switzerland, September 1991.

[15] Z. Pan, E. Altman and T. Başar, "Robust Adaptive Flow Control in High Speed Telecommunication Networks", *Proc. 35th IEEE Conference on Decision and Control*, pp. 1341-1346, Kobe, Japan, Dec. 1996.