Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Control of a Random Walk with Noisy Delayed Information

E. Altman, G. Koole

Department of Operations Reasearch, Statistics, and System Theory

BS-R9316 1993

Control of a Random Walk with Noisy Delayed Information

Eitan Altman
INRIA
Centre Sophia Antipolis
06565 Valbonne Cedex, France

Ger Koole
CWI
P.O. Box 94079, 1090 GB Amsterdam
The Netherlands

Abstract

We consider the control of a random walk on the set of nonnegative integers. The controller has two actions. It makes decisions based on a noisy information on the current state but on full information on previous states and actions. We establish the optimality of a threshold policy, where the threshold depends on the last known action, and the noisy information on the current state. We apply the result to a problem of control of service and a problem of a control of flow in the presence of uncontrolled flows.

1991 Mathematics Subject Classification: 90B22, 90C40

Keywords & Phrases: Control of a random walk, noisy delayed information, service control, flow control

1 Introduction

Control problems with imperfect state information have received much attention in the literature. In such problems, one assumes that the decision maker has only access to an observation of the state (see e.g., Bertsekas [6], Hernández-Lerma [9]).

A special class of partially observable control problems is the one where information is delayed. In that case, the current state of the system becomes known to the controller only after some time T (possibly random). Such a model with T fixed, was recently analyzed by Altman and Nain [2]. They consider a discrete time model with N-Step Delayed State Information (N-SDSI) structure. By enlarging the state space to include the last observation of the state as well as all actions taken since that time, they transform the model into a standard fully observable MDP (Markov Decision

Process). They apply this transformation to obtain an optimal control policy for a flow control model with a unit information delay. A similar flow control problem as well as a routing problem, both with a unit time of information delay were solved by Kuri and Kumar [12]. The case of delay of N steps was considered in [4]. For some other recent research on optimal control of queueing networks with delayed information in continuous time, see Altman and Koole [1] who analyze a problem with two controllers with delayed information on both actions and on the state; and Koole [11] and Artiges [5] who study routing with delayed information.

Some other control problems with more involved information structure were considered by Grizzle and Marcus [8], Hsu and Marcus [10], Schoute [15] and Altman and Shimkin [3] who study decentralized control problems (i.e., there are several decision makers), using the Delayed Sharing of State Information (DSSI) pattern. In that case, each decision maker informs the others about his observation with a delay. Another feature of the models in [8] and [10] is that the controllers may possess some immediate noisy information about the current state of the system. After a unit of delay, however, the controllers obtain the exact information. In each case the imperfect state information (or partially observable) problem is seen to reduce to a perfect state information (or completely observable) problem by enlarging the state space.

In this paper we consider a control problem with a single controller with a noisy information structure, which is a special case of the information structure considered in Grizzle and Marcus [8], Hsu and Marcus [10]. We consider the control of a random walk on the set of nonnegative integers. A single controller has two actions a_0 and a_1. It makes decisions based on a noisy information on the current state but on full information on previous states and actions. We establish the optimality of a threshold policy, where the threshold $l(a, y)$ depends on the noisy information y on the current state, and on the last known action a. We then characterize the threshold, and show that $l(a_1, y) + 1 \geq l(a_0, y) \geq l(a_1, y)$, i.e., $l(a_0, y)$ is either $l(a_1, y)$ or $l(a_1, y) + 1$.

We apply the result to a problem of control of service and a problem of a control of flow in the presence of uncontrolled flows. The second application is a generalization of the model studied in [2] both in the information structure, and in the more general arrival structure.

The paper is organized as follows. In Section 2, we introduce the model, assumptions and notation. The main result is presented in Section 3, and finally, extensions and applications are presented in the last Section.

2 Model and Assumptions

We will consider the control of a random walk in discrete time, defined by

$$X_{n+1} = (X_n + g(\eta_n, A_n))^+, \tag{1}$$

where X_n denotes the state and A_n the action at time n. The state space $\mathbf{X}$ is the set of nonnegative integers $\mathbb{N}$, and there are two actions a_0 and a_1 available in each state. The function g is integer valued, and η_n is a sequence of $\mathbb{R}^K$-valued i.i.d. random variables. Further assume that the actions are ordered, i.e. $a_0 < a_1$. Note that the displacement is, in some sense, independent of X_n: For $f : \mathcal{Z} \to \mathbb{R}$, an arbitrary function such that $f(x) = f(0)$, for all $x < 0$, we have

$$E(f(X_2 - 1)|X_1 = x + 1, A_1 = a) = E(f(X_2)|X_1 = x, A_1 = a) \tag{2}$$

The applications in section 4 consider a special case, where η_n consists of two independent components η_n' and η_n'', respectively governing the arrivals and the departures at a queue.

In the general model, the controller does have full information on the previous states, but not on the present state. Instead, it has, at time $n + 1$, noisy information Y_n (taking values in some Borel space $\mathbf{Y}$) on η_n:

$$Y_n = h(\eta_n, \zeta_n),$$

where ζ_n is a sequence of i.i.d. random variables generating the noise. This is a special case of the information structure studied in [8] and [10]. Note that Y_n is independent of X_n and A_n. This implies, by using (2), that the posterior transition probabilities are also shift invariant: let $f : \mathcal{Z} \to \mathbb{R}$ be an arbitrary function such that $f(x) = f(0)$, for all $x < 0$. Then

$$E(f(X_2 - 1)|X_1 = x + 1, A_1 = a, Y_1 = y) = E(f(X_2)|X_1 = x, A_1 = a, Y_1 = y). \tag{3}$$

We can model our random walk as a Markov Decision Process (MDP) with partial state observation. By enlarging the state space from $\mathbf{X}$ to $\mathbf{X} \times \mathbf{A} \times \mathbf{Y}$ we obtain an equivalent fully observed

MDP. (This is the standard way of enlarging the state space, see [8] or [10].) Z_{n+1}, the state at time $n+1$, is given by $Z_{n+1} = (X_n, A_n, Y_n)$, the state at time $n+1$.

Let Y_n have probability mass function F_2, and let F_1 be the probability mass function of X_{n+1}, given X_n, A_n and Y_n. Thus $F_1(x|z) = P(X_n = x|Z_n = z)$ and $F_2(B) = P(Y_n \in B)$.

The transition probabilities of the MDP, for state z_n and action a_n, are

$$P(Z_{n+1} \in (\{x'\}, \{a''\}, B)|Z_n = (x, a, y), A_n = a') = 1(a'' = a')F_1(x'|(x, a, y))F_2(B)$$

Note that (3) can be written as

$$\sum_{x' \in \mathbf{X}} F_1(x'|x+1, a, y)f(x'-1) = \sum_{x' \in \mathbf{X}} F_1(x'|x, a, y)f(x') \tag{4}$$

We shall assume that

$$\mathbf{C0}: \qquad F_1(\cdot|z^2) \geq_{st} F_1(\cdot|z^1)$$

for $z^i = (x^i, a^i, y^i)$, $i = 1, 2$, $y^1 = y^2$ when either $a^2 \geq a^1$ and $x^2 = x^1$, or $x^2 > x^1$. The relation $\geq_{st}$ is the stochastic ordering, see e.g. Ross [14].

To obtain some insight in condition C0, integrate F_1 over y to get the transitions of the underlying unobserved model:

$$F(x'|(x, a)) = \int_y F_1(x'|(x, a, y))F_2(dy)$$

It is easily seen that if C0 holds, then also $F(\cdot|(x^2, a^2)) \geq_{st} F(\cdot|(x^1, a^1))$ holds, for $x^1 = x^2$ and $a^2 \geq a^1$, or $x^2 > x^1$. Now choose $(x^1, a^1) = (x, a_0)$ and $(x^2, a^2) = (x, a_1)$. This shows us that under a_1 the displacement is stochastically bigger than under a_0. Choosing $(x^1, a^1) = (x, a_1)$ and $(x^2, a^2) = (x+1, a_0)$ shows that the difference in displacement, after actions a_0 and a_1, is stochastically smaller than 1. This gives the following total ordering of the states, just by looking at the state to be reached next:

$$(0, a_0), (0, a_1), (1, a_0), \ldots, (x, a_0), (x, a_1), (x+1, a_0), \ldots. \tag{5}$$

The description of the model is complete after specifying the costs. We consider an immediate cost function $c : \mathbf{X} \times \mathbf{A} \to \mathbb{R}$ and assume that c satisfies the property C1 introduced below.

For convenience, we shall extend the definition of c to $\mathcal{Z} \times \mathbf{A}$ (where $\mathcal{Z}$ are the set of all integer numbers), such that $c(y,p) = c(0,p)$, for all $y < 0$ and $p = a_0, a_1$.

A function $f : \mathcal{Z} \times \mathbf{A} \to \mathbb{R}$ with the property that

$$f(y,p) = f(0,p), \qquad \forall y < 0, p = a_0, a_1 \tag{6}$$

is said to satisfy property **C1** if

$$\begin{cases} (i) \;\; f(x,p) - f(x,q) \text{ is nondecreasing in } x \text{ for any actions } p \geq q, \\ (ii) \;\; f(x+1,p) - f(x,q) \text{ is nondecreasing in } x \text{ for any actions } p \text{ and } q. \end{cases}$$

Note that (ii) implies the convexity and monotone increasingness of $f(x,p)$ in x, for all p. More precisely, it can be seen that a function $g : \mathbf{X} \times \mathbf{A} \to \mathbb{R}$ satisfies (i), (ii) and monotonicity in x (i.e., $g(x,p)$ is nondecreasing in x for any fixed p) if and only if its extension through (6) satisfies C1.

We denote the problem of minimizing the expected cost for a horizon of n and initial state z by $\mathbf{Q}_n$.

3 The optimality of a threshold policy

Let $J^n(z)$ denote the cost for a horizon of n steps, and let $J^0(z) = 0$. Define

$$V^n(z,a) = E\left\{ c(X_1,a) + J^n(Z_2) \middle| Z_1 = z, A_1 = a \right\}$$

Then

$$J^{n+1}(z) = \min_a V^n(z,a) \tag{7}$$

The Markov policy u that minimizes $V^n(z,a)$ in the (enlarged) state z when there are n steps to go, $n = 0, 1, ...$ is known to be optimal.

Denote $\hat{J}^n(x,a) = \int J^n(x,a,y) F_2(dy)$. We shall understand below $\hat{J}^n(x,a) = \hat{J}^n(0,a)$ for $x < 0$ (we thus consider the extension (6) of $\hat{J}^n$). Then

$$V^n(z,a') = \sum_{x'} F_1(x'|z) \left[c(x',a') + \hat{J}^n(x',a') \right] \tag{8}$$

and (7) yields

$$\hat{J}^{n+1}(x,a) = \int F_2(dy) \left\{ \min_{a'} \sum_{x'} F_1(x'|(x,a,y)) \left[c(x',a') + \hat{J}^n(x',a') \right] \right\}$$

Theorem 3.1 *Assume C0 and that c satisfies C1. Then*

(i) There exists an optimal policy u^* for $\mathbf{Q}_n$ which is of a (time-dependent) threshold type, such that if at time n it is optimal to use a_0 at state (x,a,y), then for any $x' > x$ it is also optimal to use a_0 at states (x',a,y).

(ii) For all $n \geq 1$, $\hat{J}^n$ satisfies C1.

Proof. A sufficient condition for the existence of an optimal threshold policy at stage n (when there are n steps to go) is that for $z = (x,a,y)$,

$$V^n(z,a_1) - V^n(z,a_0) \text{ is nondecreasing in } x. \tag{9}$$

Since

$$V^n(z,a_1) - V(z,a_0) = \sum_{x' \in \mathbf{X}} F_1(x'|z) \left[c(x',a_1) - c(x',a_0) + \hat{J}^n(x',a_1) - \hat{J}^n(x',a_0) \right] \tag{10}$$

it follows that a sufficient condition for (9) to hold is that both $\hat{J}^n$ and c satisfy C1(i), and C0 holds. Indeed, this implies that the term in the square brackets of (10) is nondecreasing in x', and (9) then follows from C0.

We show by induction that $\hat{J}^n$ satisfies C1, hence establishing both claims of the Theorem. Assume that $\hat{J}^n$ satisfies C1. We shall show that for every y, $J^{n+1}(\cdot,y,\cdot)$ also satisfies C1, from which the inductive claim is established.

We begin by establishing C1(i). Let p be the action that achieves the minimum in

$$\min_a \sum_{x' \in \mathbf{X}} F_1(x'|x+1,a_1,y)[c(x',a) + \hat{J}^n(x',a)];$$

let q be the action that achieves the minimum in

$$\min_a \sum_{x' \in \mathbf{X}} F_1(x'|x,a_0,y)[c(x',a) + \hat{J}^n(x',a)].$$

Then

$$
\begin{aligned}
& J^{n+1}(x+1,a_1,y) - J^{n+1}(x+1,a_0,y) - [J^{n+1}(x,a_1,y) - J^{n+1}(x,a_0,y)] \\
= \ & \min_a \sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_1,y)[c(x',a) + \hat{J}^n(x',a)] \\
& \qquad - \min_a \sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_0,y)[c(x',a) + \hat{J}^n(x',a)] \\
& - \left(\min_a \sum_{x'\in\mathbf{X}} F_1(x'|x,a_1,y)[c(x',a) + \hat{J}^n(x',a)] \right. \\
& \qquad \left. - \min_a \sum_{x'\in\mathbf{X}} F_1(x'|x,a_0,y)[c(x',a) + \hat{J}^n(x',a)] \right) \\
\geq \ & \sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_1,y)[c(x',p) + \hat{J}^n(x',p)] \\
& \qquad - \sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_0,y)[c(x',p) + \hat{J}^n(x',p)] \\
& - \left[\sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_1,y)[c(x'-1,q) + \hat{J}^n(x'-1,q)] \right. \\
& \qquad \left. - \sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_0,y)[c(x'-1,q) + \hat{J}^n(x'-1,q)] \right] \\
= \ & \sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_1,y)[c(x',p) - c(x'-1,q) + \hat{J}^n(x',p) - \hat{J}^n(x'-1,q)] \\
& - \sum_{x'\in\mathbf{X}} F_1(x'|x+1,a_0,y)[c(x',p) - c(x'-1,q) + \hat{J}^n(x',p) - \hat{J}^n(x'-1,q)] \geq 0
\end{aligned}
$$

The first inequality follows from (4) and the definition of p and q. The last inequality follows from C0 and the fact that the term in square brackets is nondecreasing in x' due to C1.

It remains to establish C1(ii). Let $\hat{p}$ be the action that achieves the minimum in

$$
\min_a \sum_{x'\in\mathbf{X}} F_1(x'|x+2,y,p)[c(x',a) + \hat{J}^n(x',a)]
$$

and let $\hat{q}$ be the action that achieves the minimum in

$$\min_{a} \sum_{x' \in \mathbf{X}} F_1(x'|x,y,q)[c(x',a) + \hat{J}^n(x',a)].$$

It follows from the inductive assumption that (9) holds for n and therefore that $\hat{p} \leq \hat{q}$.

$$\begin{aligned}
&\hat{J}^{n+1}(x+2,y,p) - \hat{J}^{n+1}(x+1,y,p) - [\hat{J}^{n+1}(x+1,y,q) - \hat{J}^{n+1}(x,y,q)] \\
&= \min_{a} \sum_{x' \in \mathbf{X}} F_1(x'|x+2,y,p)[c(x',a) + \hat{J}^n(x',a)] \\
&\qquad - \min_{a} \sum_{x' \in \mathbf{X}} F_1(x'|x+1,y,p)[c(x',a) + \hat{J}^n(x',a)] \\
&\quad - \left(\min_{a} \sum_{x' \in \mathbf{X}} F_1(x'|x+1,y,q)[c(x',a) + \hat{J}^n(x',a)] \right. \\
&\qquad \left. - \min_{a} \sum_{x' \in \mathbf{X}} F_1(x'|x,y,q)[c(x',a) + \hat{J}^n(x',a)] \right) \\
&\geq \sum_{x' \in \mathbf{X}} F_1(x'|x+2,y,p)[c(x',\hat{p}) + \hat{J}^n(x',\hat{p})] \\
&\qquad - \sum_{x' \in \mathbf{X}} F_1(x'|x+2,y,p)[c(x'-1,\hat{q}) + \hat{J}^n(x'-1,\hat{q})] \\
&\quad - \left(\sum_{x' \in \mathbf{X}} F_1(x'|x+1,y,q)[c(x',\hat{p}) + \hat{J}^n(x',\hat{p})] \right. \\
&\qquad \left. - \sum_{x' \in \mathbf{X}} F_1(x'|x+1,y,q)[c(x'-1,\hat{q}) + \hat{J}^n(x'-1,\hat{q})] \right) \\
&= + \sum_{x' \in \mathbf{X}} F_1(x'|x+2,y,p)[c(x',\hat{p}) - c(x'-1,\hat{q}) + \hat{J}^n(x',\hat{p}) - \hat{J}^n(x'-1,\hat{q})] \\
&\quad - \sum_{x' \in \mathbf{X}} F_1(x'|x+1,y,q)[c(x',\hat{p}) - c(x'-1,\hat{q}) + \hat{J}^n(x',\hat{p}) - \hat{J}^n(x'-1,\hat{q})] \geq 0
\end{aligned}$$

The last inequality follows from C0, C1 and the fact that by the inductive assumption, $\hat{J}^n$ satisfies C1. This establishes the proof. ∎

According to Theorem 3.1, for any $a \in \mathbf{A}$ and $y \in \mathbf{Y}$, there exists some threshold $l(a,y)$, such that if $x > l(a,y)$, it is optimal to use a_0, and otherwise it is optimal to use a_1. Next we give some characterization of the function l.

Theorem 3.2 *For any $y \in \mathbf{Y}$,*

$$l(a_1, y) + 1 \geq l(a_0, y) \geq l(a_1, y). \tag{11}$$

Proof. Fix n. It follows from (10) that

$$\begin{aligned} & V^n((x,a_0,y),a_0) - V^n((x,a_0,y),a_1) - [V^n((x,a_1,y),a_0) - V^n((x,a_1,y),a_1)] \\ &= \sum_{x' \in \mathbf{X}} F_1(x'|(x,a_0,y)) \left[c(x',a_0) - c(x',a_1) + \hat{J}^n(x',a_0) - \hat{J}^n(x',a_1)\right] \\ &\quad - \sum_{x' \in \mathbf{X}} F_1(x'|(x,a_1,y)) \left[c(x',a_0) - c(x',a_1) + \hat{J}^n(x',a_0) - \hat{J}^n(x',a_1)\right] \geq 0 \end{aligned} \tag{12}$$

The last inequality follows from C0, and from the fact that both c and $\hat{J}^n$ satisfy C1; hence the term in the square brackets is nonincreasing in x'.

Assume that in state $z = (x, a_0, y)$, a_0 is optimal, i.e., $V^n(z, a_1) \geq V^n(z, a_0)$. It follows from (12) that a_0 is optimal also in state (x, a_1, y), since $V^n((x,a_1,y),a_1) \geq V^n((x,a_1,y),a_0)$. This implies the second inequality in (11). To obtain the first inequality, we note that for any $p, q \in \mathbf{A}$,

$$\begin{aligned} & V^n((x,y,p),a_0) - V^n((x,y,p),a_1) - [V^n((x+1,y,q),a_0) - V^n((x+1,y,q),a_1)] \\ &= \sum_{x' \in \mathbf{X}} F_1(x'|(x,y,p)) \left[c(x',a_0) - c(x',a_1) + \hat{J}^n(x',a_0) - \hat{J}^n(x',a_1)\right] \\ &\quad - \sum_{x' \in \mathbf{X}} F_1(x'|(x+1,y,q)) \left[c(x',a_0) - c(x',a_1) + \hat{J}^n(x',a_0) - \hat{J}^n(x',a_1)\right] \geq 0 \end{aligned} \tag{13}$$

The last inequality follows from C0, and from the fact that both c and $\hat{J}^n$ satisfy C1; hence the term in the square brackets is nonincreasing in x.

Assume that in state $z = (x, y, p)$, a_0 is optimal, i.e., $V^n(z, a_1) \geq V^n(z, a_0)$. It follows from (13) that a_0 is optimal also in state (x, y, q), since $V^n((x+1,y,q),a_1) \geq V^n((x+1,y,q),a_0)$. This implies the first inequality in (11) (by setting $p = a_0$ and $q = a_1$). ∎

4 Applications and Extensions

The results can easily be extended to the case where ζ_n and η_n are independent but not i.i.d. Another straight forward extension is to the case where the random walk is on the set of all integers (not just the nonnegative). In that case the dynamics are given by:

$$X_{n+1} = X_n + g(\eta_n, A_n)$$

We now present two possible applications of Theorem 3.1 to the control of queues with noisy delayed information. We then discuss situations that yield different type of noisy delayed information.

Service Control Consider a discrete time queue with an infinite buffer. At the beginning of time n there are $\eta'_n \geq 0$ arrivals. Let X_n be the number of customer just before the arrivals occur. Let $\eta''_n \in \{0,1\}$ be a sequence of i.i.d. Bernoulli random variables with parameter α, representing potential service completions. The action A_n corresponds to the decision whether to enable the service at time n. Let $\mathbf{A} = \{-1, 0\}$, i.e., $a_0 = -1$ and $a_1 = 0$; take $g(\eta_n, A_n) = \eta'_n + A_n \eta''_n$. If the queue is non-empty and $A_n = a_0$ then with probability α a customer will leave the system at the end of the slot. Consider the immediate cost $c(x,a) = f(x) + \gamma a$, where f, representing a holding cost, is increasing and convex, and $\gamma \geq 0$ is a cost for deciding to serve. This structure of c ensures that C1 holds. It is easily seen that also C0 holds.

Flow Control Consider a discrete time queue with an infinite buffer. Consider K streams of arrivals. Here $\eta_n^{(l)}$, $l = 1, \ldots, K$, represents the arrival streams and $\eta_n^{(K+1)}$ represents the departures. At the beginning of time n there are $\eta_n^{(l)} \geq 0$ arrivals from sources $l = 1, \ldots, K-1$. These are uncontrolled arrivals. Let $\eta_n^{(K)} \in \{0,1\}$ be a sequence of i.i.d. Bernoulli random variables with parameter α, representing potential arrivals from source K at the beginning of slot n. The action A_n denotes the decision of whether to enable or not the potential arrival at time n. Let $\mathbf{A} = \{0,1\}$, i.e., $a_0 = 0$ and $a_1 = 1$; if $A_n = a_1$ then with probability α a customer will arrive from stream K. At the end each slot, if the queue is non-empty, then service succeeds with probability β and a customer leaves the system. Let $\eta_n^{(K+1)} = -1$ if service succeeds, otherwise $\eta_n^{(K+1)} = 0$. Take $g(\eta_n, A_n) = \sum_{l=1}^{K-1} \eta_n^{(l)} + A_n \eta_n^{(K)} + \eta_n^{(K+1)}$. The immediate cost is $c(x,a) = f(x) + \gamma a$, where f, representing a holding cost, is increasing and convex, and $\gamma \leq 0$ is interpreted as a reward for accepting customers, and hence a reward for increasing the throughput. This structure of c ensures that C1 holds. Again it is easily seen that C0 holds as well.

For both models we may consider the following type of informations:

(1) No information on the current state. This means that the state information as well as information about arrivals and service are known with a unit delay. In that case the threshold policy obtained by Theorem 3.1 is a function of the last known action. The flow control model corresponding to this case was studied in [2] (for the case $k = 1$, i.e., only one controlled arrival stream).

(2) Partial delayed information. Consider a situation where the controller (service controller or flow controller) gets in time the information from the beginning of the last slot only yet the information about events occurring in the end of the slot do not arrive in time for the decision making. Then the controller has all the information about the arrivals in the last slot but not on service completions. It thus has more information than in the previous case, but less than full information on the current state. In the case of service control this could mean that $Y_n = (\eta_n')$, in the case of flow control $Y_n = (\eta_n^{(1)}, \cdots, \eta_n^{(K)})$.

(3) Information with a random delay. In some cases the delay of information has random duration. In packet switching telecommunication networks, information is often obtained through acknowledgements from the destination that are piggy-packed on packets on the opposite direction. Hence the amount of delay in the information depends on the (random) amount of congestion on the reverse way of packets back to the source. In our simple model, we could assume that information on the service in the last slot does not come in time for the decision making, yet with some positive probability, the information about the arrivals that occurred in the beginning of the last slot did arrive in time. In the case of the control of service this could be modeled by

$$Y_n = \eta_n' + \zeta_n \tag{14}$$

where $\zeta_n = (\zeta_1, ..., \zeta_K)$ is a vector whose components may take values 0 or $-\infty$, and are independent; $\zeta_i = 0$ means that the information came in time, otherwise $\zeta_i = -\infty$.

(4) Noisy delayed information. We may assume that due to some unreliable medium, the information we get is noisy; it becomes reliable after some error correction which makes the information accurate after a unit delay.

(5) Full information. The case where the controller always has full information on the current state is a special case of the information structure that we consider. Note that even for this case, our results generalize some previous results for fully observable control models. The optimality of a threshold policy for flow control models are known, see e.g. [13] and [16]; the novelty of our model is that we consider the control of one flow in presence of other uncontrolled ones. (Note, however, that a flow control problem with another type of criteria was analyzed by [7], which involves also a second flow of uncontrolled traffic).

References

[1] E. Altman and G. Koole, "Stochastic Scheduling Games with Markov Decision Arrival Processes", *Journal Computers and Mathematics with Appl.*, 3rd special issue on "Differential Games", pp. 141-148, 1993.

[2] E. Altman and P. Nain, "Closed-Loop control with delayed information", *Performance Evaluation Review* **20**, pp. 193-204, 1992.

[3] E. Altman and N. Shimkin, "Decentralized flow control with delayed sharing information", in preparation.

[4] E. Altman and S. Stidham, "Monotonicity of Optimal Policies in a Two-Action Markov Decision Process, with Applications to Networks of Queues", in preparation.

[5] D. Artiges, "Optimal routing into two heterogeneous service stations with delayed information", submitted to IEEE-AC.

[6] D.P. Bertsekas, *Dynamic Programming. Deterministic and Stochastic Models*. Prentice Hall, Inc. Englewoods Cliff, 1987.

[7] H. Blanc, P. R. de Waal, P. Nain and D. Towsley, "Optimal Control of Admission to a Multiserver Queue with Two Arrival Streams", *IEEE Trans. Aut. Contr.*, Vol. AC-37, No. 6, pp. 785-797, 1992.

[8] J.W. Grizzle and S. I. Marcus, "A decentralized control strategy for multiaccess broadcast networks," *Large Scale Systems*, Vol. 3, pp. 75-88, 1982.

[9] O. Hernández-Lerma, *Adaptive Markov Control Processes*. Springer-Verlag, New York, 1989.

[10] K. Hsu and S. I. Marcus, "Decentralized control of finite state Markov processes," *IEEE Trans. Aut. Contr.*, Vol. AC-27, No. 2, pp. 426-431, Apr. 1982.

[11] G. Koole, "Optimal repairman assignment in two maintenance models which are equivalent to routing models with early decisions", Report BS-R9301 CWI, Amsterdam, January 1993.

[12] J. Kuri and A. Kumar, "Optimal Control of Arrivals to Queues with Delayed Queue Length Information", *Proceeding of the 31st IEEE Conference on Decision and Control*, pp. 997-998, Arizona, USA, December 1992.

[13] D.-J. Ma and A.M. Makowski, "Optimality results for a simple flow control problem," in *Proc. 26th Conf. on Decision and Control*, pp. 1852-1857, Los Angeles, CA, USA, Dec. 1987.

[14] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, New York, 1983.

[15] F.C. Schoute, "Decentralized control in packet switched satellite communication", *IEEE Trans. Aut. Contr.*, Vol. AC-23, No. 2, pp. 362-371, Apr. 1978.

[16] S. Stidham, Jr., "Optimal Control of Admission to a Queueing System," *IEEE Trans. Aut. Contr.*, Vol. AC-30, No. 8, pp. 705-713, Aug. 1985.