On scheduling and routing problems with a partially observed environment

Ger Koole
INRIA Sophia Antipolis
B.P. 93, 06902 Sophia Antipolis, France

Olaf Passchier
Leiden University
Postbus 9512, 2300 RA Leiden, The Netherlands

Technical Report TW-95-13, Leiden University, 1995

Abstract

We consider a special type of control problem with an independent environment that governs the type of transition of the controlled system. Assuming full observation on the system, but no observations on the state of the environment, we show using the theory of partially observed Markov decision processes that the equivalent fully observed control problem has the same structure with the state of the environment replaced by a distribution on the environment states.

We apply this result to Markov arrival processes (MAPs), and show that the estimation of an MAP is again an MAP. From this it immediately follows that many results in the literature hold also for partially observed arrival processes.

1 Introduction

In open scheduling and routing models one can often make a clear distinction between the arrival process and the system to be controlled. For some models the optimal policy is independent of the arrival process (e.g. shortest queue routing ([10]), the μc rule, LEPT ([9])). More often however the optimal policy can only be partially characterized (e.g., that the optimal policy is of threshold type), and in this case the optimal policy usually depends on the state of the arrival process (e.g., the threshold value is a function of the state of the arrival process). Examples can be found in [4, 5, 1, 6]. Furthermore, several results of this type which are derived for Poisson processes can easily be generalized to more general arrival streams like Markov arrival processes (MAPs), notably the results in [8, 3, 11]. This leads also to optimal policies depending on the arrival processes.

While it is not unusual to have full up-to-date state information on the system, it is of interest to study scheduling policies which are not allowed to depend on the current state of the environment that governs the arrival process. For example, in a communication network with incoming traffic generated by on-off sources it is not realistic to assume that the state of each source is known to the network controller; the only thing that often can be done is to estimate the number of busy sources based on current (and previous) traffic characteristics. And this information might be useful to the controller: in the case of admission control one might want to reject more customers if more arrivals are to be expected, due to a high number of active sources. It is the objective of this note to show that the structural results that hold in the case of full information still hold in the case of no information on the arrival process.

To do so, we formulate a general scheduling model consisting of an environment and a system which can be controlled. The environment changes state independently of the controlled system, and it influences the controlled system only through marks, where the mark indicates which type of transition should occur in the controlled system. For example, the mark could indicate whether and at which queue an arrival will occur. It is assumed that the state of the system and the marks are observed, but not the state of the environment itself. Using theory on partially observed stochastic control problems we show that the optimal control depends only on the estimation of the environment and the current state of the controlled system, and that the problem which needs to be solved to find this optimal control has the same structure. This means that results which hold for the fully observed model hold also for the model without observation of the environment.

An interesting application is the Markov arrival process (MAP). An MAP (both in continuous and in discrete time) together with the system at which the arrivals occur, is of the form sketched above. Moreover, the process estimating an MAP is again an MAP, and thus all the results in the literature for fully observed MAPs hold also for the case that the MAP is not observed. For the case of arrivals according to a Markov modulated Poisson process (MMPP, a poll of on-off processes is an example of this), it can be seen that the approximation is an MAP, but not an MMPP. This argues for the use of MAPs for the derivation of structural results.

2 The model

We consider a discrete time model, with state space $\mathcal{X} \times \mathcal{Y} \times \mathcal{I}$, respectively giving the states of the environment, the possible marks, and the possible states of the system. The action set (assumed the same for each state) is $\mathcal{A}$. If the process is in state (x, y, i), and action a is chosen, then state (x', y', i') is reached at the next epoch with probability $r((x', y', i')|(x, y, i), a)$. We assume that r has the following structure:

$$r((x', y', i')|(x, y, i), a) = r^A(x', y'|x)r^Q(i'|i, y', a) \qquad (1)$$

We see clearly that the interaction between the environment and the controlled system is through the mark y'. There are direct costs $c((x, y, i), a)$. In applications these costs often depend only on i and a.

Note also that r does not depend on y, and therefore neither does the optimal policy. Thus we could formulate an equivalent problem with state space $\mathcal{X} \times \mathcal{I}$ and transitions $\hat{r}((x', i')|(x, i), a) = \sum_{y'} r^A(x', y'|x)r^Q(i'|i, y', a)$. However, we believe that the analysis is clearer when we keep the mark in the states. In applications the use of the mark is often implicit.

We assume that in state (x, y, i) (y, i) is observed, meaning that there is no information on the state of the environment, but full information on the mark and on the state of the system. If the system is to be controlled at t, and $(y_1, i_1), \ldots, (y_t, i_t)$ are the observations up to t, then the control is allowed to depend on all these observations. In general, for a model with full observations, it can be shown that there exists an optimal policy which is a function only of the current state. This does not hold for the model with partial observations, the use of previous observations can improve the estimation of the current state. We take a closer look at this state approximation in the next section.

3 The result

We use the theory on partially observed Markov decision processes, as it can be found in for example [2, 7]. There it is shown that a problem with partial information can be solved by solving a second problem with full information, where the states are replaced by probability distributions on the states. It is readily seen (for example, by (6.2) on p. 83 of [7]) that in the current model these distributions have the form (p, y, i), where p is a distribution on $\mathcal{X}$. When we denote the transitions of the second problem with $\tilde{r}$, then it follows from the theory (for example, by theorem 7.1 on p. 85 of [7]) that (write $p(x)$ for the mass that p puts on x)

$$\begin{aligned}\tilde{r}((p', y', i')|(p, y, i), a) &= \sum_{x,x'} p(x) r((x', y', i')|(x, y, i), a) \\ &= \sum_{x,x'} p(x) r^A(x', y'|x) r^Q(i'|i, y', a)\end{aligned}$$

for p' such that

$$p'(x') = \frac{\sum_x p(x) r^A(x', y'|x)}{\sum_{x,x'} p(x) r^A(x', y'|x)}.$$

Note that the distribution p' depends only on p and y': this means that to estimate the state of the environment it suffices to know all the marks (and the initial distribution of the environment). Write $T(p, y')$ for this distribution. Now we can rewrite $\tilde{r}((p', y', i')|(p, y, i), a)$ as follows (with $\mathrm{I}\{\cdot\}$ the indicator function):

$$\tilde{r}((p', y', i')|(p, y, i), a) = \mathrm{I}\{p' = T(p, y')\} \sum_{x,x'} p(x) r^A(x', y'|x) r^Q(i'|i, y', a),$$

which is in the form of (1), as we can take $\tilde{r}^A(p', y'|p) = \mathrm{I}\{p' = T(p, y')\} \sum_{x,x'} p(x) r^A(x', y'|x)$ and $\tilde{r}^Q(i'|i, y', a) = r^Q(i'|i, y', a)$.

As direct costs we have $\tilde{c}((p, y, i), a) = \sum_x p(x) c((p, y, i), a)$.

Remark We could generalize our result by allowing partial observations of the environment instead of no information at all. Technically this is straightforward to do, but it would add significantly to the notation. Therefore we decided to confine ourselves to the current model.

4 Markov arrival processes

An important application of the above result are Markov arrival processes. An MAP consists of a Markovian environment that can generate customers at each transition. The customers arrive at this system, whose evolution is otherwise stochastically independent of the MAP. The MAP can be considered both in continuous and discrete time. Most applications are formulated in continuous time, making it the most interesting case to consider. We start however with the simpler discrete time model.

The crucial difference between discrete and continuous time models is that in discrete time models multiple events can happen at the same time: for example in a discrete time queue customers can arrive at and depart from the system at the same time.

Discrete time model Each transition function for the environment r^A, as given in (1), can be written as $r^A(x', y'|x) = \lambda(x'|x)q(y'|x, x')$, which is the typical form of an MAP in discrete time: λ gives the transition probabilities of the arrival process (note that $\lambda(x|x)$ can be non-zero), and q the generation of marked arrivals, given the particular transition that took place (we assume that no arrival is represented by mark 0, making q a probability distribution for each x and x'). Under an MAP the type of transition depends on the mark of the arrival, giving indeed the form of (1). As every distribution r^A (and thus also its estimation $\tilde{r}^A$) can be written in the form λq, it follows that the estimation of an MAP is again an MAP. This gives the discrete time result.

Continuous time model Now we consider the continuous time model. The usual way to deal with a continuous time Markovian control problem is to *uniformize* it. An early paper applying this method to a control model is [8]. It amounts to considering the model at intervals with an exponentially distributed length. The crucial idea is that the length of this interval does not depend on the state of the system. This can be achieved by inserting fictitious transitions, if we assume that the transition rates in each state are uniformly bounded. Thus we assume that the rates of the MAP, $\lambda(x'|x)$, are bounded by Λ for each x: $\sum_{x'} \lambda(x'|x) \le \Lambda$. The probability of an arrival with mark y is $q(y|x, x')$, and if such an arrival occurs, the system changes state with probabilities $r^Q(i'|i, y, a)$. We denote the rates of the internal changes with $s(i'|i, a)$, and they are bounded by Γ: $\sum_{i'} s(i'|i, a) \le \Gamma$ for all i. We add fictitious transitions (from x to x) to the MAP, with the inactive mark (say) 0 which has $r^Q(i|i, 0, a) = 1$ for all i and a, as to assure that $\sum_{x'} \lambda(x'|x) = \Lambda$ for each x, and we add transitions from i to i to the system. Finally assume (by scaling) that $\Lambda + \Gamma = 1$. This finishes the uniformization: what results is a discrete time model with the transitions occurring at the points of a Poisson process with rate 1. This discrete time model has as transitions:

$$r((x', i')|(x, i), a) = \sum_y \lambda(x'|x)q(y|x, x')r^Q(i'|i, y, a) + \mathrm{I}\{x = x'\}s(i'|i, a). \qquad (2)$$

To model this in the form (1), we introduce an extra mark "$*$", denoting the transition internal to the controlled system, where $q(*|x, x')$ can only be positive if $x = x'$. Furthermore we take $r^Q(i'|i, *, a) = s(i'|i, a)/\Gamma$.

Having rewritten (2) in the form of (1) shows that the partial information model has again the form of (1), due to the results of the previous section. What we should show however, is that the estimation has again the form of (2). To do that, it suffices to show that $\tilde{r}((p', *, i')|(p, y, i), a)$ has the form $\mathrm{I}\{p = p'\}r^Q(i'|i, *, a)$, that is, that $T(p, *) = p$. This basically means that the observation of $*$ does not change the estimation of the environment. But this is easily verified due to the fact that $*$ is observed in every environment state with equal probability.

Example Consider a model with arrivals according to an MAP, occurring in a single class, thus $q(0|x, x') + q(1|x, x') = 1$, with mark 1 denoting the arrival. Furthermore, take $\mathcal{I} = \mathbb{N}^m$ and $\mathcal{A} = \{1, \ldots, m\}$. The n-dimensional state space models n parallel queues. Take $r^Q(i|i, 0, a) = 1$ and $r^Q(i + e_a|i, 0, a) = 1$ for all i and a, with e_a the ath unity vector. Thus action a corresponds to sending an arrival to queue a. The rates s correspond to the (possible) departures: $s(i - e_j|i, a) = \mu_j$ for i such that $i_j > 0$. We add dummy transitions as to assure that $\sum_j s((i - e_j)^+|i, a) = \sum_j \mu_j = \Gamma$. This gives exactly the routing model of [4], as defined on p. 499 and with value function given by (2.1). Thus our results show that the monotonicity results obtained for the optimal policy hold as well for the model with a partially observed MAP.

This is just one of the models for which the results can be generalized to partially observed environments; also the results in [4, 5, 1, 6] are obtained for MAPs and hold also for partially observed MAPs. Furthermore, there are many results in the literature (notably those in [8, 3, 11]) obtained for Poisson processes which hold also for MAPs, observed or not.

Remark The Markov modulated Poisson process (MMPP) is often used to model bursty traffic. An MMPP is an MAP with $q(y|x, x') = 0$ if $x \neq x'$. It is clear that, in the partial observation case, the arrival of a customer gives information on the state of the MMPP: if the MMPP models a single on-off source, then the arrival of a customers assures us that the source is in the on state. Thus with the arrival of a customer the estimation changes, making the estimation an MAP, but not an MMPP. This argues for the use of MAPs instead of MMPPs when deriving structural results.

Remark In [1] a unified framework is presented to deal at the same time with continuous and discrete time models. There it is used to prove properties related to submodularity, but it is also indicated how to apply it to other types of scheduling models. In this note it has been shown how to deal with fully and partially observed environments at the same time. Combining all shows that a single proofs can suffice for dealing at the same time with all four different types of models.

5 Conclusion

We considered a special type of control problem with an environment that governs the transitions in the system. Assuming full observation on the system, but no observations on the state of the environment, we showed using the theory of partially observed Markov decision processes that the equivalent control problem has the same structure where the state of the environment is replaced by a distribution.

We applied this result to Markov arrival processes, and showed that the estimation of an MAP is again an MAP. From this it immediately follows that many results in the literature hold also for partially observed environments.

References

[1] E. Altman and G.M. Koole. On submodular value functions of dynamic programming. Technical Report 2658, INRIA Sophia Antipolis, 1995.

[2] D.P. Bertsekas. *Dynamic Programming.* Prentice-Hall, 1987.

[3] B. Hajek. Optimal control of two interacting service stations. *IEEE Transactions on Automatic Control*, 29:491–499, 1984.

[4] A. Hordijk and G.M. Koole. On the assignment of customers to parallel queues. *Probability in the Engineering and Informational Sciences*, 6:495–511, 1992.

[5] A. Hordijk and G.M. Koole. On the optimality of LEPT and μc rules for parallel processors and dependent arrival processes. *Advances in Applied Probability*, 25:979–996, 1993.

[6] G.M. Koole and Z. Liu. Nonconservative service for minimizing cell loss in ATM networks. In *Proceedings of the 1st Workshop on ATM Traffic Management*, 1995. To appear.

[7] P.R. Kumar and P. Varaiya. *Stochastic Systems.* Prentice-Hall, 1986.

[8] S.A. Lippman. Applying a new device in the optimization of exponential queueing systems. *Operations Research*, 23:687–710, 1975.

[9] R. Righter. Scheduling. In M. Shaked and J.G. Shanthikumar, editors, *Stochastic Orders and their Applications*, pages 381–432. Academic Press, 1994.

[10] J. Walrand. *An Introduction to Queueing Networks.* Prentice-Hall, 1988.

[11] R.R. Weber and S. Stidham, Jr. Optimal control of service rates in networks of queues. *Advances in Applied Probability*, 19:202–218, 1987.