

# On Optimal Call Admission Control in a Resource-Sharing System

Eitan Altman, Tania Jiménez, and Ger Koole

**Abstract**—In this paper, we consider call admission control of multiple classes without waiting room. We use event-based dynamic programming for our model. We show that sometimes the customer classes can be ordered: if it is optimal to accept a class, then to accept a more profitable class is optimal too. We demonstrate submodularity of the minimum cost for the 2-classes problem and establish some properties of optimal policies. Then we formulate a fluid model that allows us to study the optimal control for the large-capacity case. We show that in the case of same service time distributions, the control problem can be reduced to a model with a one-dimensional (1-D) state space, and a trunk reservation policy is optimal. We present numerical examples that validate our results.

**Index Terms**—Fluid model, stochastic knapsack, submodularity, trunk reservation.

## I. INTRODUCTION

THE advent of ATM has renewed the interest in resource-sharing models with different resource requirements. For example, in the ATM context a video connection will demand more (effective) bandwidth (i.e., more resources) than a regular voice connection. Traditionally these models are called trunk reservation models, due to the most often used policy for admitting new customers, based on thresholds depending only on the currently available resources (where the resources are the trunk lines in a telecommunication network). In this paper, we study such a resource sharing model: we derive structural properties for the optimal policy for the actual model, and we fully determine the optimal policy for the approximating fluid model.

This model is defined as follows. Traffic from  $N$  classes share  $B$  resources. Traffic of class  $i$  arrives according to a Poisson process with parameter  $\lambda_i$ . Each customer demands  $b_i$  resources, with  $b_i$  integer. All the resources taken by a class  $i$  customer are released simultaneously after an exponentially distributed service time with parameter  $\mu_i$ . States are denoted with  $x = (x_1, \dots, x_N)$ , where  $x_i$  signifies the number of class  $i$  customers in the system. We assume that  $\sum_{i=1}^N b_i x_i \leq B$ , i.e., there is no waiting room, customers who do not find suf-

ficient resources are automatically blocked. There is a central controller who can reject arriving customers based on full state information. Each class  $i$  customer who enters the system gives rise to a reward of  $r_i$ . In this paper, we will study admission policies for this model for the discounted reward criteria.

This problem is also known as the *stochastic knapsack* (Ross and Tsang [1], Gaviou and Rosberg [2]). In [1], the optimal *coordinate convex* policy is determined. This class of policies (to which the optimal policy does not belong in general) is of particular interest as it leads to product-form solutions of the steady state probabilities (the influence of changing the parameters for coordinate convex policies are studied by Ross and Yao [3] and Nain [4]). In [2] a good policy is found based on an approximation of the value function. The special case where the service rates and the reward per acceptance do not depend on the customer's class was studied in Feinberg and Reiman [5] (under some constraints), and the optimality of a trunk reservation policy is established (see Miller [6], [7] for early results on trunk reservation).

We study the optimal policy for the unconstrained case. For the model with different  $\mu_i$  we derive several monotonicity results, the strongest for the model with  $N = 2$  and  $b_1 = b_2 = 1$ . We illustrate with numerical results the nonmonotonicity of the optimal policies (see also [8]).

Although these results give insight in the structure of the optimal policy, it does not help us much in finding a good policy. Therefore we considered also an approximation of the model for which we could find the optimal policy, namely the fluid model. We computed the optimal policy for several parameter values, both for the original model and its fluid approximation. For other related fluid models, see also [9]–[12].

## II. STRUCTURAL RESULTS WITH DYNAMIC PROGRAMMING

### A. Dynamic Programming

Our results are derived by inductively proving properties of the dynamic programming (DP) value function. To formulate the DP equation we will make use of *event-based DP*, as introduced in [13]. This greatly simplifies the proofs and also enables us to consider several systems at the same time.

To every possible event in the system corresponds a DP *event operator* which maps the set  $F$  of all real-valued functions of the state variable  $x$  into itself. For arrivals of a class  $i$  ( $i = 1, \dots, N$ ) customer we have (with  $e_i$  the  $i$ th unity vector) the corresponding arrival operator  $T_{A_i}$  defined by

$$T_{A_i} f(x) = \max\{r_i + f(x + e_i), f(x)\} \quad (1)$$

Paper approved by P. E. Rynes, the Editor for Switching Systems of the IEEE Communications Society. Manuscript received April 5, 1999; revised February 10, 2000 and February 13, 2001. This paper was presented in part at the INFORMS Tel Aviv Conference, June 29, 1998 and at the 37th IEEE CDC, Tampa, FL December 1998.

E. Altman is with INRIA, 06902 Sophia Antipolis, France (e-mail: Eitan.Altman@sophia.inria.fr).

T. Jiménez is with CESIMO—Centro de Simulación y Modelos, Facultad de Ingeniería, Universidad de Los Andes, Mérida, Venezuela (e-mail: tania@ing.ula.ve).

G. Koole is with Divisie Wiskunde & Informatica, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands (e-mail: koole@cs.vu.nl).

Publisher Item Identifier S 0090-6778(01)08169-7.

for  $f \in F$ . The function  $T_{A_i} f(x)$  may be interpreted as the optimal value function for a one-stage problem in which one must decide to accept or reject a class- $i$  arrival, after which a terminal state-dependent revenue is earned according to the function  $f$ .

For the departures the situation is more complicated. Here we have  $\lfloor B/b_i \rfloor$  event operators for class  $i$ , namely for  $k = 1, \dots, \lfloor B/b_i \rfloor$

$$T_{D_i}^{(k)} f(x) = \begin{cases} f(x - e_i), & \text{if } x_i \geq k \\ f(x), & \text{otherwise.} \end{cases} \quad (2)$$

Denote by  $C(x)$  the costs associated with state  $x$ . These costs are only used to prevent the system from leaving the state space, i.e., we take

$$C(x) = \begin{cases} 0, & \text{if } \sum_{i=1}^N b_i x_i \leq B \\ \infty, & \text{otherwise.} \end{cases}$$

Using the above the DP operator  $T$  for the continuous time model as defined in the introduction becomes (using the well-known uniformization technique, see for example Lippman [7])

$$Tf(x) = -C(x) + \beta \sum_{i=1}^N \lambda_i T_{A_i} f(x) + \beta \sum_{i=1}^N \mu_i \sum_{k=1}^{\lfloor B/b_i \rfloor} T_{D_i}^{(k)} f(x). \quad (3)$$

Here we have assumed without loss of generality that  $\sum_{i=1}^N \lambda_i + \sum_{i=1}^N \lfloor B/b_i \rfloor \mu_i = 1$ ;  $\beta \in (0, 1)$  is the discount factor (when analyzing the discounted reward criterion). One may also analyze the average reward criterion by taking  $\beta = 1$ . In [13] it is shown how to generalize this to general (but state-independent) arrival streams, by adding an environment state. We will not deal with that generalization here, but the results of this section hold also for general arrivals.

The operator  $T$  is a linear combination of event operators. This is typical for continuous time models; only one event can happen at a time. Using the same event operators we can also model discrete time models, by taking convolutions of event operators, representing events happening simultaneously. Note however that not all the results that follow are valid for discrete-time models. We discuss this issue when appropriate.

In the following we prove certain properties of the *value function*  $V_n$ , defined by  $V_{n+1}(x) = TV_n(x)$  and  $V_0(x) = -C(x)$ . As these properties hold for all  $n$ , they hold also for the limiting optimal policy (for discounted or average reward), as follows from standard results in Markov decision theory, see e.g., Puterman [14]. Note that no additional assumptions are needed because our model has a finite state space.

### B. Ordering Classes

Sometimes two or more of the customer classes can be ordered, in the sense that if it is optimal to accept the less profitable one in a certain state, then an arriving customer of the more profitable class should also be accepted.

This idea is formalized in the following lemmas. First we show that  $V_n$  is decreasing (we use increasing and decreasing in the nonstrict sense) in all of its components.

**Lemma II.1:** For all  $x, j$  and  $n$ , we have  $V_n(x) \geq V_n(x + e_j)$ .

*Proof:* It is easily seen that  $V_0$  (i.e.,  $-C$ ) satisfies the above inequality. We show that if  $f(x)$  satisfies the above

inequality, then so does  $T_{\bullet} f(x)$  for all event operators. As the inequality is maintained under linear combinations, the lemma follows directly by induction on  $n$ .

Consider first  $T_{A_i}$  for some  $i$ . Suppose that  $f$  is decreasing in all components. Then  $\max\{r_i + f(x + e_i), f(x)\} \geq \max\{r_i + f(x + e_i + e_j), f(x + e_j)\}$  for all  $x, i$ , and  $j$ . This shows that  $T_{A_i} f$  is decreasing in all components for all  $i$ .

The terms concerning the  $T_{D_i}^{(k)}$  are even easier. Note that  $T_{D_i}^{(x_j+1)} f(x) = T_{D_i}^{(x_j+1)} f(x + e_j)$ . ■

For ease of notation we consider classes 1 and 2; as we can renumber, this does not restrict generality.

**Lemma II.2:** Consider the  $N$ -class problem with  $N \geq 2$ . If  $b_1 \leq b_2$  and  $\mu_1 \geq \mu_2$ , then for all  $x$  and  $n$  we have  $V_n(x + e_1) \geq V_n(x + e_2)$ .

*Proof:* The cost function  $-C$  and thus  $V_0$  satisfy the above inequality because  $b_1 \leq b_2$ .

We show that if  $f(x)$  is decreasing in all its components and satisfies the above inequality, then so does  $T_{\bullet} f(x)$  for all event operators. As the inequality is closed under linear combinations, the lemma follows directly by induction on  $n$  (using Lemma II.1).

Suppose that  $f$  satisfies  $f(x + e_1) \geq f(x + e_2)$  and is monotone decreasing.

Consider first  $T_{A_i}$  for some  $i$ . Then clearly  $\max\{r_i + f(x + e_1 + e_i), f(x + e_1)\} \geq \max\{r_i + f(x + e_2 + e_i), f(x + e_2)\}$ , so  $T_{A_i} f$  satisfies the inequality if  $f$  does.

The terms concerning the  $T_{D_i}^{(k)}$  are easy, except  $T_{D_1}^{(x_1+1)}$  and  $T_{D_2}^{(x_2+1)}$ . For the value function (3) we need that

$$\begin{aligned} \mu_1 T_{D_1}^{(x_1+1)} f(x + e_1) + \mu_2 T_{D_2}^{(x_2+1)} f(x + e_1) \\ \geq \mu_1 T_{D_1}^{(x_1+1)} f(x + e_2) + \mu_2 T_{D_2}^{(x_2+1)} f(x + e_2) \end{aligned}$$

which is equivalent to  $\mu_1 f(x) + \mu_2 f(x + e_1) \geq \mu_1 f(x + e_2) + \mu_2 f(x)$ . This holds because  $\mu_1 \geq \mu_2$ ,  $f(x + e_1) \geq f(x + e_2)$  and  $f(x) \geq f(x + e_2)$ . The Lemma is now established via induction. ■

Note that we used explicitly the form of (3) in the proof. Now we formulate the first result of this section concerning the form of an optimal policy.

**Theorem II.3:** If  $b_1 \leq b_2$ ,  $\mu_1 \geq \mu_2$  and  $r_1 \geq r_2$ , and in a state  $x$  it is optimal to accept class 2 customers, then it is also optimal to accept class 1 customers.

*Proof:* By Lemma II.2 and  $r_1 \geq r_2$  we have that  $r_1 + V_n(x + e_1) \geq r_2 + V_n(x + e_2)$ . Thus if  $r_2 + V_n(x + e_2) \geq V_n(x)$ , i.e., admission of a class 2 customer is optimal, then  $r_1 + V_n(x + e_1) \geq V_n(x)$ , i.e., admission of a class 1 customer is also optimal. ■

### C. Submodularity for $N = 2$

Assume that  $N = 2$ .

**Lemma II.4:** For all  $n$  and  $x \geq 0$

$$V_n(x + e_1) + V_n(x + e_2) \geq V_n(x) + V_n(x + e_1 + e_2). \quad (4)$$

*Proof:* It is clear that  $V_0$  (i.e.,  $-C$ ) satisfies the above inequality. We show that if  $f(x)$  satisfies the above inequality, then so does  $T_{\bullet} f(x)$  for all event operators. As the inequality is closed under linear combinations, the lemma follows directly.

The proof for the operators  $T_{D_i}$  is easily given, using the inductive assumption; let us therefore consider in detail the operator  $T_{A_1}$  ( $T_{A_2}$  is similar). Denote by  $a_1$  ( $a_2$ ) the maximizing action in  $T_{A_1}f(y)$  for  $y = x$  ( $x + e_1 + e_2$ ), where action 0 (1) refers to rejecting (accepting) a customer. If  $a_1 = a_2$ , then the result follows easily. If  $a_1 = 1$  and  $a_2 = 0$ , then

$$\begin{aligned} & T_{A_1}f(x + e_1) + T_{A_1}f(x + e_2) \\ & \geq f(x + e_1) + r_1 + f(x + e_1 + e_2) \\ & = T_{A_1}f(x) + T_{A_1}f(x + e_1 + e_2). \end{aligned}$$

If  $a_1 = 0$  and  $a_2 = 1$ , then

$$\begin{aligned} & T_{A_1}f(x + e_1) + T_{A_1}f(x + e_2) \\ & \geq r_1 + f(x + 2e_1) + f(x + e_2) \\ & \geq f(x) + r_1 + f(x + 2e_1 + e_2), \end{aligned}$$

the second inequality follows by using (4) twice.  $\blacksquare$

Inequality (4) is known as *submodularity*. The following monotonicity result is a direct consequence of it (see, e.g., Altman and Koole [15]):

*Theorem II.5:* For every fixed value of  $x_i$ ,  $i = 1, 2$ , there is a threshold level  $L^i(x_i)$  such that in state  $x = (x_1, x_2)$  a customer of type  $i$  is admitted if and only if  $x_j < L^i(x_i)$  where  $j \neq i$ .

In [16] and [15], a theory is developed around submodularity. This was done in the setting of costs instead of rewards; a general theory around supermodularity could not be developed. Translating the current model to costs would change sub- to supermodularity. This makes our model interesting from a theoretical point of view, and explains as well the restriction to  $N = 2$ .

Unfortunately, for every fixed value of  $x_1$  there does not exist in general a threshold level  $L'(x_1)$  such that in state  $x = (x_1, x_2)$  a customer of type 2 is admitted if and only if  $x_2 < L'(x_1)$ . To prove that such a threshold exists typically requires that concavity be established, which, in general, does not hold in our case. In fact, it is easy to establish counter examples for the case that  $b_1 \neq b_2$ , where the monotonicity is reversed due to the boundaries, see Section IV.

#### D. The One-Dimensional (1-D) Case

So far we have derived two partial results: one for the case that customer classes can be ordered in some special way, and one for  $N = 2$ . A special case for which the optimal policy is fully characterized is the one where all  $\mu_i$  are the same (equal to  $\mu$ , say), and  $b_i = 1$  for all  $i$ .  $V_n(x)$  then depends only on the total number of customers present, so we assume here that the argument of  $V_n$  is 1-D.

*Theorem II.6:* (Miller [6], Lippman [7]):

- i) For all  $x \leq B - 2$  and  $n$  we have that  $2V_n(x + 1) \geq V_n(x) + V_n(x + 2)$ , i.e.,  $V_n$  is concave.
- ii) For each class the optimal policy is of threshold form, i.e., for each customer class there is a critical level above which no customers are admitted.

This theorem shows that the optimal policy uses the idea of trunk reservation: the thresholds assure that some of the servers (called trunk lines in the telephone network) are kept free (are

reserved) for other traffic classes. Of course capacity is only kept free for traffic classes that give a higher profit. This can be shown as follows. Let  $T_i$  be the threshold for class  $i$  traffic, and assume that  $r_1 \geq \dots \geq r_N$ . Using Theorem II.3 we conclude that also  $T_1 \geq \dots \geq T_N$ . Thus we can say that  $B - T_N$  trunks are reserved for traffic of the classes  $1, \dots, N - 1$ , of which  $B - T_{N-1}$  trunks are reserved for traffic of the classes  $1, \dots, N - 2$ , etc.

A similar model, with two types of traffic and in the context of cellular communications, has been studied in [17].

### III. FLUID MODEL

In this section, we present a fluid model that approximates the original model when the number of resources  $B$  as well as the rate of call arrivals  $\lambda$  are large. Explicit solutions are obtained for the fluid model, which can be used to construct almost-optimal policies for the original problem.

Let us make some general observations before specifying the model. A fluid model is characterized by the absence of randomness in the system. This makes a fluid model usually perform better than the originating queueing system. (In [18] this general observation is proven for a specific model.) Indeed, in a queueing system with two classes, trunk reservation will always be optimal if the reward of customer class 1 is high enough, even if there is enough capacity to accommodate the average total traffic load. This is because capacity is reserved in case, due to the random behavior, a burst of class 1 customers arrives. This phenomenon does not play a role in the fluid model. But still, as we will see, trunk reservation can be optimal, even if the total traffic load does not exceed the service capacity.

#### A. The Model

- *The states:* The discrete state space is replaced with a continuous one  $\mathbf{X} = \{x \in \mathbb{R}^N : x \geq 0, \sum_{i=1}^N b_i x_i \leq B\}$ , representing the amount of sessions in the system.
- *The arrivals:* We replace the discrete arrivals by a continuous stream; the amount of class  $i$  traffic that arrives per time unit is assumed to be deterministic with rate  $\lambda_i$ .
- *The departure rate:* The departure rate of one unit of session of type  $i$  is  $\mu_i$ . Hence, if there are  $x_i$  sessions of type  $i$  in the system, then the departures of type  $i$  sessions is  $x_i \mu_i$ .
- *Control and policies:* The arrivals are controlled;  $u_i$ , the admission control for class  $i$ , represents the fraction of the arrivals that are accepted to the system. The evolution of the system is given by

$$\frac{dx(t)}{dt} = f(x(t), u(t)), \text{ where } f_i(x, u) = u_i \lambda_i - \mu_i x_i. \quad (5)$$

The control actions are admission vectors which are elements of  $U = [0, 1]^N$ . We consider policies that are measurable functions of the time of the form  $\pi : \mathbb{R} \rightarrow U$ , with the following state constraints. If at time  $t$  the state is  $x$  and action  $u$  is used, then

$$\sum_{i=1}^N (u_i \lambda_i - \mu_i x_i) \leq 0 \text{ if } \sum_{i=1}^N b_i x_i = B. \quad (6)$$

A policy satisfying (6) is called feasible. Define  $D(x)$  to be the set of actions that can be used at state  $x$ .

- *The reward:*  $r_i > 0$  is a reward per unit of type  $i$  sessions admitted to the system. With  $r(u) = \sum_{i=1}^N r_i \lambda_i u_i$  being the immediate reward, we consider the total discounted cost  $J_\beta(x, \pi) = \int_0^\infty e^{-\beta t} r(u_t) dt$ , where  $u_t$  is the action at time  $t$ , obtained when a policy  $\pi$  is used. (Note that, although the immediate reward is related only to the action and not the state of the system, the state of the system plays an important role since the feasible admission actions are constrained once the boundary  $\sum_{i=1}^N b_i x_i = B$  is attained.)

Our aim is to find a policy  $\nu$  that achieves  $J_\beta(x, \nu) = J_\beta(x) \stackrel{\text{def}}{=} \sup_\pi J_\beta(x, \pi)$ .

*Remark III.1:* The constant  $\beta$  used in  $J_\beta$  is not the same as the one used in the previous section for the discretized model. Writing the discounting as an exponent is usual in continuous time models. However, there is an easy formula relating both forms of discounting (see, e.g., [19] for this and other details concerning uniformization).

The above model can be interpreted as a limit of a sequence MDP ( $n$ ) of MDPs of the form of the original one, in which the parameters are chosen appropriately. This is done by scaling both the intensities and the states by the same factor. We shall add below the superscript  $c$  to denote the parameters of the continuous queueing model, i.e., the one appearing in (5).

- 1)  $\lambda_i(n) = n\lambda_i^c = n\lambda_i(1)$ ,  $\mu_i(n) = \mu_i^c = \mu_i(1)$  (due to the next item, this will mean that all rates are  $n$  times faster in the  $n$ th model).
- 2)  $x_i(n) = x_i(1)/n$ ; this is the rescaling in space by  $n$ . The state space now contains fractions instead of integers, and the basic unit corresponding to a single session in the  $n$ th model is  $1/n$ . The arrival and departure operators  $A_i$  and  $D_i$  have to be redefined accordingly in (1) and (2) by letting  $e_i$  correspond to  $1/n$  times the (original)  $i$ th unity vector. Thus, keeping the initial state  $x(n, 0)$  constant in  $n$  means that we multiply the number of ongoing sessions by  $n$ .
- 3)  $B$  and the  $b_i$ s are kept unchanged; due to the transformation of the state space, this means that the constraint on the total number of calls, weighted by the  $b_i$ s, increases linearly with  $n$ .
- 4) The discount factor  $\beta(n)$  of the  $n$ th model is given as  $\beta(n) = (\beta(1))^{1/n} = e^{-\beta^c/n}$ .
- 5) The reward:  $r_i(n) = r_i(1)/n = r_i^c/n$ .

*Remark III.2:* The above scaling allows for the convergence of not only the rewards, but also of the state process. In practice, if we only want to compute the rewards using DP, we may use a simpler scaling. i) We need not rescale the state, so we may take  $x_i(n) = x_i(1)$  in step 2) above. In that case, the vector  $e_i$  used in (1) and (2) is just as originally defined: the  $i$ th unit vector. But then, ii) we need to rescale  $B$  to  $nB$  in step 3). The other parameters are the same as in the previous scaling.

The motivation for the above scaling is the following. The number of arrivals in a time unit  $\Delta$  in the  $n$ th model is of the order of  $n\lambda_i^c \Delta$ . Assume that  $\Delta$  is small,  $n$  large, and that the policy  $u$  is constant during  $\Delta$ . Then we can approximate the

incurred reward during a time  $\Delta$  as  $\lambda_i(n)\Delta r_i(n)$  (where  $r_i(n)$  depends on  $u$ ), since a reward of  $r_i(n)$  is earned by each arrival and there are approximately  $n\lambda_i(n)\Delta$  arrivals. The discounting in the original model was related to events: each time an event occurred we multiplied by the discount factor  $\beta$ . The total number of events that occur during the period  $\Delta$  in the  $n$ th model is of the order of  $n\Delta$  of which the number of arrivals is of order of  $n\lambda\Delta$  (recall that in the original model, the sum of arrival rates plus maximum service rate was taken to be one unit). The total discounting during the period of  $\Delta$  is thus  $(\beta(n))^{n\Delta}$ . The overall reward for class  $i$  is thus approximated by

$$\int_0^\infty e^{-\beta^c t} E r^c(u_t) dt \sim \sum_{m=0}^\infty (\beta(n)^{n\Delta})^m \lambda_i(n) E r_i(n) \Delta.$$

This gives us the above scalings. We do not include a mathematical study of the convergence of the discrete model to the fluid model, this is beyond the scope of this paper.

Next we derive certain useful properties of our fluid model. First note that the model is indeed totally deterministic. This means that the trajectory can be calculated for each policy and initial state. These trajectories are the solutions of the linear differential (5), which are given by

*Lemma III.3: (Trajectories):* The state trajectories under any policy  $u \in U$  are given by

$$x_i(t) = e^{-\mu_i t} x_i(0) + \lambda_i e^{-\mu_i t} \int_0^t e^{\mu_i s} u_i(s) ds \quad (7)$$

where  $u_i(t)$  is the action used at time  $t$ .

We will also need the equivalent of the DP value function of the previous section for the current continuous time and state space model. It is the following Hamilton-Jacobi-Bellman (HJB) equation:

$$\beta v(x) = \max_{u \in D(x)} \{r(x) + f(x, u) \cdot \nabla v(x)\} \quad (8)$$

where

$$f(x, u) \cdot \nabla v(x) \stackrel{\text{def}}{=} \sum_{i=1}^N f_i(x, u) \frac{\partial v(x)}{\partial x_i}.$$

Define  $\mathcal{D}$  to be the set of functions  $\psi: \mathbf{X} \rightarrow \mathbb{R}$  which are uniformly continuous over  $\mathbf{X}$ . We have the following ([20] Theorem. 3.3, see also [21] Sections III.8–9, [22] and [23]).

*Lemma III.4:* There is a unique viscosity solution  $v \in \mathcal{D}$  of (8) and  $v(x) = J_\beta(x)$ .

### B. Optimality Results

In this subsection we study the general model. We derive some results for the special case of light load. We obtain the structure for arbitrary load in the following subsection, restricting to the special case of equal service rates.

*Theorem III.5:* Assume that  $\sum_{i=1}^N b_i \lambda_i / \mu_i \leq B$ . Then

- Any optimal policy starting from any initial state  $x \in \mathbf{X}$  results in a trajectory that converges to the point  $x^*$ , given by  $x_i^* = \lambda_i / \mu_i$  for all  $i$ ;
- There exist  $x' \in \mathbf{X}$  such that

$$\lambda_i \leq x'_i \mu_i, \quad \text{for all } i \quad (9)$$

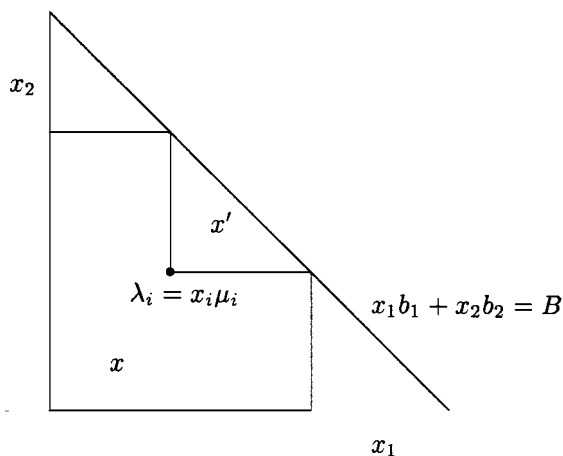


Fig. 1. Illustration of Theorem III.5.

and for any initial state  $x$  with  $x \leq x'$  componentwise it is optimal for all time  $t \geq 0$  to accept all arriving customers.

*Proof:* Note that the point  $x^*$  exists if and only if  $\sum_{i=1}^N b_i \lambda_i / \mu_i \leq B$ . We start with proving ii). The set of possible  $x'$  consists of all  $x' \in \mathbf{X}$  with  $x' \geq x^*$ . It is easy to see that it is optimal to accept all customers if, by doing so, the boundary  $B$  is not crossed. Any policy not doing so is sub-optimal, as  $r_i > 0$  for all  $i$ . We prove that it is optimal to accept all traffic for  $x'$  as specified in the theorem. Because the trajectories for  $x$  are upperbounded by those for  $x'$ , the result for  $x$  follows then directly. Indeed, from (7) we have

$$\begin{aligned} x'_i(t) &= e^{-\mu_i t} x'_i(0) + e^{-\mu_i t} \lambda_i \int_0^t e^{\mu_i s} ds \\ &= x'_i(0) - (1 - e^{-\mu_i t}) \left( x'_i(0) - \frac{\lambda_i}{\mu_i} \right) \leq x'_i(0) \end{aligned} \quad (10)$$

and  $x'(0) \in \mathbf{X}$  by assumption.

The convergence of the trajectory for any  $x \leq x'$  to  $x^*$  follows easily from (10).

Now consider a trajectory due to an optimal policy starting at an arbitrary initial state. Then the trajectory satisfies

$$x_i(t) \leq x_i(0) - (1 - e^{-\mu_i t}) \left( x_i(0) - \frac{\lambda_i}{\mu_i} \right).$$

This means that after a finite time the trajectory is within a region, as in ii), for which accepting forever all sessions is optimal, for which we already showed that the trajectory converges to  $x^*$ . ■

In Fig. 1 the set of states for which  $u_i = 1$  is optimal because of Theorem III.5 is drawn for  $N = 2$ . The middle triangle gives the possible values for  $x'$  (in which (9) holds), the region below it gives the possible values for  $x$ . In fact, for one of the two remaining triangles (the one corresponding to an excess of faster customers)  $u_i = 1$  is also optimal.

Thus we see that trunk reservation is not used within the set of states  $\{x | x \leq x', \lambda_i \leq x'_i \mu_i\}$ . This is for the simple reason that we can show that all traffic can be accepted, without any overflow. This is not the case for initial states outside of this set.

It can well be the case that the trajectory belonging to  $u_i = 1$  for all  $t$  and  $i$  goes outside of  $\mathbf{X}$ . We study this case in the next subsection.

### C. A Simpler Equivalent Model

We make in this subsection the following assumption:

*Assumption III.6:* (A1) The service rates do not depend on  $i$ ,  $\mu_i = \mu$ .

As we will see below this assumption suffices to treat the system as a 1-D system. To do the same in the discrete model we needed the  $b_i$  to be equal as well.

Define

$$y(x) \stackrel{\text{def}}{=} \sum_{i=1}^N x_i b_i, \quad w(u) \stackrel{\text{def}}{=} \sum_{i=1}^N u_i \lambda_i b_i. \quad (11)$$

*Lemma III.7:* Consider two models with initial states  $x(0)$  and  $x'(0)$ , respectively, such that

$$\sum_{i=1}^N b_i x_i(0) = \sum_{i=1}^N b_i x'_i(0) = y(x(0)).$$

Then

- i) Any policy that uses  $u(t)$  at time  $t$  that is feasible for initial state  $x(0)$  is also feasible for initial state  $x'(0)$ , and  $r_t \stackrel{\text{def}}{=} \sum_{i=1}^N r_i \lambda_i u_i(t)$  in the two cases;
- ii) For any two policies  $u(t)$  and  $u'(t)$  satisfying  $w(u(t)) = w(u'(t))$  for all  $t$  the trajectories  $y(t) = y(x(t))$  and  $y'(t) = y(x'(t))$  are the same, and if  $u(t)$  is a feasible policy, then so is  $u'(t)$ .

*Proof:* It follows from (5) that  $y_t$  satisfies

$$\frac{dy}{dt} = w(u) - \mu y \quad (12)$$

with the constraint  $dy/dt \leq 0$  if  $y = B$ . The evolution of  $y_t$  is then the same in the two cases. Since the constraints on the policies are functions of  $y$ , the same constraints are satisfied, and thus  $u(t)$  are indeed feasible for both cases. Since the discounted cost is only a function of  $u$ , which is the same in both cases, it follows that the immediate rewards  $r_t$  are the same. This completes the proof of i).

Part ii) follows directly from (12) and part i). ■

The above Lemma hints that we may aggregate the state and action spaces. Indeed, we note that at any time  $t$ , an action  $u(t)$  may be changed to  $u'(t)$  without affecting the future evolution of the  $y$ s, and hence the future constraints on  $u$ . On the other hand, the rewards do not depend on  $x$  directly, but only on  $u$ . This motivates the following definition of an improvement of a policy.

*Definition III.8:* Consider a feasible policy  $\pi$ . Define the improvement  $\pi'$  of  $\pi$  as follows. Denote

$$g(w) \stackrel{\text{def}}{=} \max_{u'} \sum_{i=1}^N r_i \lambda_i u'_i, \quad \text{s.t.} \quad \sum_{i=1}^N u'_i \lambda_i b_i = w, \quad 0 \leq u'_i \leq 1. \quad (13)$$

Define  $\hat{\pi}(w) = g(w)$ .

If  $\pi$  uses  $u$  at time  $t$ , then  $\pi'$  is defined to be the policy that uses at time  $t$  the action  $\hat{\pi}(w(u))$ .

*Lemma III.9:* For all  $x$  we have  $J_\beta(x, \pi') \geq J_\beta(x, \pi)$ . The reward corresponding to  $\pi'$  is given by  $J_\beta(x, \pi') = \int_0^\infty e^{-\beta t} g(w_t) dt$ , where  $w_t = w(u(t))$ .

*Proof:* Directly from Lemma III.7, part ii). ■

We are thus ready to introduce an equivalent simplified 1-D model with

- **State space:**  $\mathbf{Y} = [0, B]$ ,
- **Action space:**  $W = \left[0, \sum_{i=1}^N \lambda_i b_i\right]$ ,
- **Policies:** These are measurable functions from  $\mathbb{R}^+$  to  $W$ , satisfying the following state constraints for all  $t: w(t) \leq \mu y$  (this is equivalent to  $dy/dt \leq 0$ ) if  $y = B$ . We denote by  $D(y)$  the actions available at state  $y$ .
- **Rewards:** The immediate reward is  $g(y, w) = g(w)$  given in (13). The total discounted reward is  $\bar{J}_\beta(y, \mathbf{w}) = \int_0^\infty e^{-\beta t} g(w_t) dt$  where  $\mathbf{w}$  is the policy that uses  $w(t)$  at time  $t$ . The optimal reward is  $\bar{J}_\beta(y) \stackrel{\text{def}}{=} \sup_{\mathbf{w}} \bar{J}_\beta(y, \mathbf{w})$ .

Introduce the HJB equation

$$\beta v(y) = \max_{w \in D(y)} \left[ g(w) + (w - \mu y) \frac{dv(y)}{dy} \right]. \quad (14)$$

As with Lemma III.4, we have

*Lemma III.10:* There is a unique solution  $v \in \mathcal{D}$  (defined in Lemma III.3) of (14), and  $v(x) = \bar{J}_\beta(y)$ .

As for the general case we make the distinction between models that have enough capacity to handle all offered traffic and models that do not have this capacity.

*Theorem III.11:* If  $\sum_{i=1}^N \lambda_i b_i \leq \mu B$  then it is optimal always to accept all arriving calls.

*Proof:* This follows directly because the policy that always accepts all calls has a negative drift on the boundary  $y = B$ . ■

We shall thus consider from now on only the nontrivial case  $\sum_{i=1}^N \lambda_i b_i > \mu B$ . To simplify notation we assume (without restricting generality) that  $\mu = 1$ . We also assume that  $r_1/b_1 \geq r_2/b_2 \geq \dots \geq r_N/b_N$ . Recall that a policy is called *stationary* if it uses at any  $t_1$  and  $t_2$  the same actions, whenever the states at  $t_1$  and  $t_2$  are the same. We characterize the optimal policy.

*Theorem III.12:* Assume  $\sum_{i=1}^N \lambda_i b_i > \mu B$ . The set  $\Pi$  consists of all policies  $\pi = u(t)$  with  $u(t)$  of the following form: for some  $l$   $u_1(t) = \dots = u_l(t) = 1$ ,  $u_{l+1}(t) \in [0, 1]$ , and  $u_{l+2}(t) = \dots = u_N(t) = 0$ . Then

- i) there is an optimal policy  $\pi^* \in \Pi$ , (characterized by  $w^*(t) := w(u^*(t))$ );
- ii) with  $y^*(t)$  the state process generated by  $\pi^*$ , we have that  $y^*(t)$  is increasing to  $\sup_t y^*(t) = B$ .

*Proof:* By Lemma III.9 an arbitrary policy  $\pi$  can be improved by applying  $\pi'$ , that uses  $g(w(t))$  at all  $t$ . We show that  $\pi' \in \Pi$ . By a simple change of variables we see that (13) is equivalent to

$$g(w) = \max_{h_i} \sum_{i=1}^N \frac{r_i}{b_i} h_i \quad \text{s.t.} \quad \sum_{i=1}^N h_i = w, \quad 0 \leq h_i \leq \lambda_i b_i.$$

Now it is obvious that this maximum is obtained by a policy in  $\Pi$ . This shows part i).

Next we show that  $\sup_t y^*(t) = B$ . Suppose that  $\sup_t y^*(t) = b < B$ . Not all traffic is admitted, as this would lead to a trajectory with  $\sup_t y^*(t) > B$ . Thus we can add additional traffic consuming up to  $B - b$  units of service capacity to the system at all times without crossing the boundary  $B$ . This increases the reward, and therefore  $\sup_t y^*(t) < B$  cannot be optimal. A similar argument can be used to show that  $\lim_{t \rightarrow \infty} y^*(t) = B$ .

We continue with showing that  $y^*(t)$  is increasing. Suppose that  $y^*(t)$  is decreasing on some interval  $[t_1, t_2]$ . As  $\lim_{t \rightarrow \infty} y^*(t) = B$ ,  $y^*(t)$  has to increase after  $t_2$  to regain  $y^*(t_1)$ . If  $y^*(t)$  is decreasing (increasing) in  $t$  then the drift is negative. Note also that the optimal policy, the solution to the HJB equation, is stationary. But we saw that the drift in states  $[y^*(t_1), y^*(t_2)]$  can be negative or positive, depending on the time. This is in contradiction with the stationarity.

A simple more heuristic argument is as follows: due to the discounting it is advantageous to move rewards forward in time. Earlier arrivals mean also earlier departures, leaving room for more admissions later on. This shows that it can never be optimal to have a negative drift in some state. ■

In the Appendix we sketch a numerical approach for solving (14). We first note that the term in brackets in (14) is concave and piecewise linear in  $w$  for every  $y$  for which  $dv(y)/dy$  exists. Therefore, for such  $y$ 's, it attains its maximum on one of the points of the form  $\sum_{i=1}^l \lambda_i b_i$ , except for the boundary  $y = B$ , where the argmax need not be achieved at an extreme point (and may typically be an interior point). We now use (14) to obtain an optimal stationary policy.

*Theorem III.13:*

- i) The solution  $v$  of (14) is concave decreasing.
- ii) Let  $w^*$  be a stationary policy that selects an action that achieves the argmax of the brackets in (14) at any  $y \in [0, B]$  for which  $dv(y)/dy$  exists. Then such a  $w^*$  can be chosen to be piecewise constant in  $y$ , and decreasing in  $y$ ; any such choice is an optimal policy.

The proof of the theorem is delayed to the Appendix.

We conclude that there is a set  $\hat{Y} = \{y_0, y_1, \dots\}$  containing at most  $N$  elements, such that  $y_n \in \mathbf{Y}$  is a decreasing sequence and  $w_n = \sum_{j \leq n} \lambda_j b_j$  achieves the argmax in (14) on the interval  $I_n \stackrel{\text{def}}{=} (y_n, y_{n+1})$ .

It follows from the definition of  $I_n$  that the solution  $v$  of (14) satisfies  $\beta v(y) = g(w_n) + (w_n - \mu y) dv(y)/dy$ . This differential equation allows us to compute  $v$  on the whole interval. We recall that the differential equation of the form  $dv/dy - f(y)v(y) = \hat{g}(y)$  has the solution

$$v(y) = \exp\left(\int_y^r f(s) ds\right) \times \left[ v(y') + \int_{y'}^r \exp\left(-\int_{y'}^r f(s) ds\right) \hat{g}(r) dr \right].$$

This yields for  $y \in I_n$

$$v(y) = \left( \frac{y_n \mu - w_n}{y \mu - w_n} \right)^{\beta/\mu} \left( v(y_n) - \frac{g(w_n)}{\beta} \right) + \frac{g(w_n)}{\beta}.$$

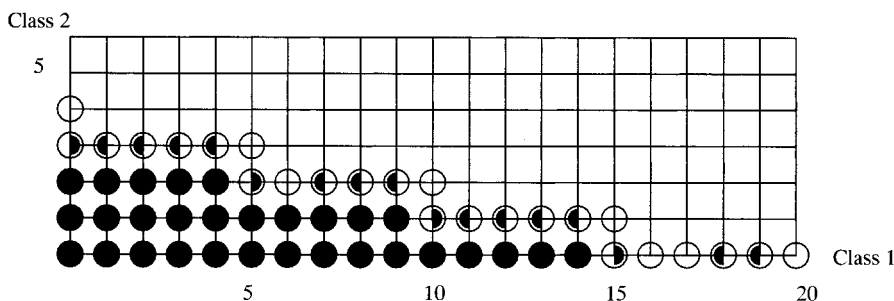


Fig. 2. Example 1.

To initiate the computation, we use the fact that at  $y_0 = B$ , it is optimal to use  $w = \mu B$  which is the policy  $w$  that results in always staying on the boundary. This follows directly from (14) and Lemma III.10. Thus  $v(B) = v(y_0) = g(\mu B)/\beta$ .

$y = B$  is the only point at which we need to consider  $w$  which does not belong to the set of points of the form  $\sum_{i=1}^t \lambda_i b_i$ ; outside the boundary, we do not need anymore to constrain the actions, so that one of the  $w_n$ 's will be optimal in a neighborhood of  $y = B$ .

#### IV. EXPERIMENT RESULTS

In this section we test the DP with some realistic data. In particular, we study numerically the structure of the optimal policies validating the structural results obtained in Section II-A. Our numerical study shows that the optimal policies are, in general, not monotone in all components. We also compare numerically the original problem with its fluid approximation.

The examples were coded in C and run on a SUN Sparc Station V.

##### A. Example 1

We study the model of Section II-A with two classes of customers ( $N = 2$ ), 20 units of total resources ( $B = 20$ ) and the following parameters for each class: Class 1:  $\lambda_1 = 10, \mu_1 = 1, b_1 = 1, r_1 = 1$ , Class 2:  $\lambda_2 = 2, \mu_2 = 1, b_2 = 5, r_2 = 10$ .

The computer program ends when the variation of the value function ( $V_n$ ) between two iterations is less than  $\epsilon = 0.00000001$ . The number of iterations was 928 for this example. The discount factor for this example is:  $\beta = 0.99$ .

Fig. 2 represents the *optimal policy* obtained. A black circle means acceptance of customers from any one of the two classes. A circle whose only left (right, respectively) half is black means that we accept customers of class one (of class two, respectively) and reject those of the other class. A white circle means that we reject both types of arrivals. In the  $x$  ( $y$ , respectively) axis we represent the customers of class 1 (2, respectively).

Fig. 3 shows the threshold level  $L(x)$  given by the Theorem II.5.

##### B. Example 2

In this example we still use two classes of customers, but class one has a much larger reward. The total resources are still twenty units ( $B = 20$ ). Class 1:  $\lambda_1 = 10, \mu_1 = 1, b_1 = 1, r_1 = 1$ , Class 2:  $\lambda_2 = 1, \mu_2 = 0.5, b_2 = 5, r_2 = 1000$ .  $\epsilon = 0.000001, \beta = 0.95$ , N. of iterations = 276.

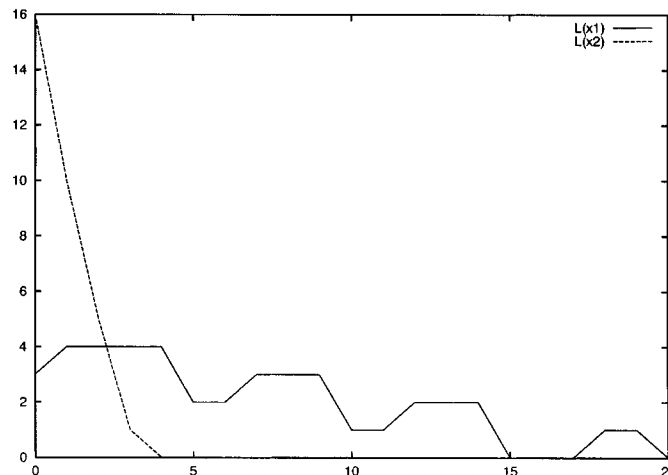


Fig. 3.  $L(x)$ .

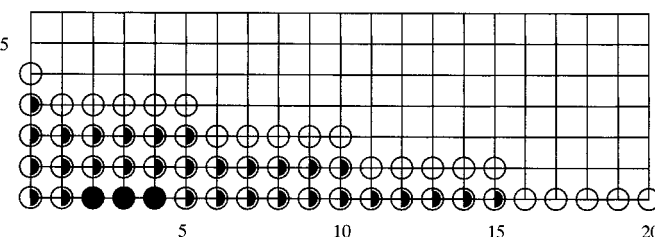


Fig. 4. Example 2.

Fig. 4 shows that only when there are 0 customers of class 1 and 2, 3 or 4 of class 2, then it is optimal to accept customers from both classes, otherwise customers of class 1 should be rejected, and it is optimal to accept class 2 customers as long as we do not reach the boundary.

Fig. 5 shows the threshold level  $L(x)$  given by the Theorem II.5.

##### C. Example 3

We use in this example three classes of customers  $N = 3$ , the total amount of resources is still  $B = 20$ , Class 1:  $\lambda_1 = 10, \mu_1 = 5, b_1 = 2, r_1 = 2$ , Class 2:  $\lambda_2 = 12, \mu_2 = 1, b_2 = 2, r_2 = 3$ , Class 3:  $\lambda_3 = 5, \mu_3 = 1, b_3 = 2, r_3 = 4, \epsilon = 0.0000001, \beta = 0.99$ .

We obtained a convex acceptance region for the optimal policy, which we describe graphically by specifying the acceptance region for each class.

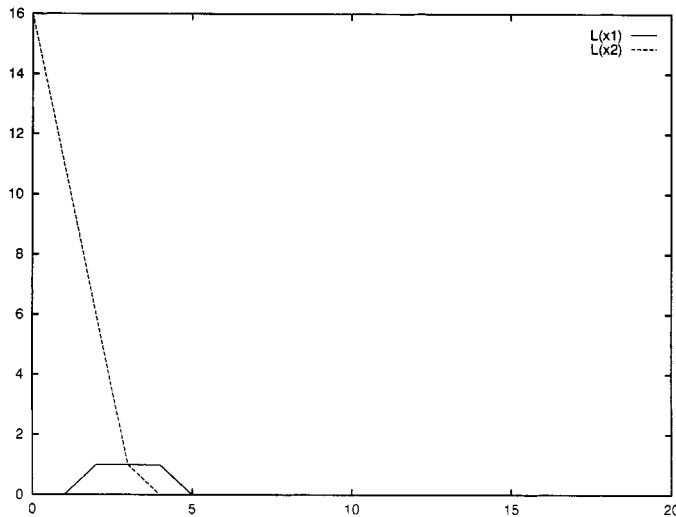
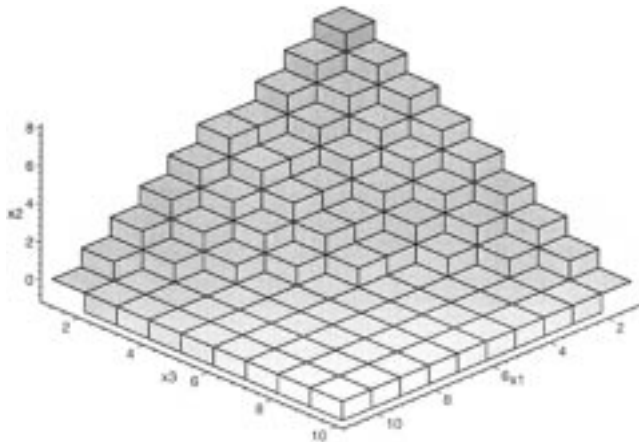
Fig. 5.  $L(x)$ .

Fig. 6. Example 3.

The acceptance region for class 2, for example, is a function of the number of customers of classes 1 and 3. In order to plot it, we define the acceptance function  $g^2 : \mathbf{X} \rightarrow R$ ; an arrival of class 2 is accepted at state  $x = (x_1, x_2, x_3)$  if and only if  $x_2 \leq g^2(x_1, x_3)$ . We present the function  $g^2$  in Fig. 6.

We did the same for all classes and the numerical results validate the structural results obtained in Section II-A. However, in both examples 1 and 2, the optimal policy is not monotone in all coordinates.

#### D. Example 4

We considered the following example for comparing the MDP with the fluid model.

The parameters are: Class 1:  $\lambda_1 = 20$ ,  $\mu_1 = 1$ ,  $b_1 = 1$ ,  $r_1 = 10$ , Class 2:  $\lambda_2 = 10$ ,  $\mu_2 = 1$ ,  $b_2 = 5$ ,  $r_2 = 100$ .  $B = 20$ ,  $\beta = 0.99$ ,  $\epsilon = 0.000\,000\,1$ .

We next define MDP (10) in which we rescale by 10 the previous MDP: The parameters for the MDP (10) are: Class 1:  $\lambda_1 = 200$ ,  $\mu_1 = 1$ ,  $b_1 = 1$ ,  $r_1 = 1$ , Class 2:  $\lambda_2 = 100$ ,  $\mu_2 = 1$ ,  $b_2 = 5$ ,  $r_2 = 10$ .  $B = 200$ ,  $\beta = 0.998\,995\,47$ ,  $\epsilon = 0.000\,000\,1$ .

The Figs. 7 and 8 show the optimal policy for the original MDP and the rescaled one, respectively.

We next illustrate the fact that under assumption A1, introduced in Section III-C, the value function  $V(x)$  of the limit fluid model (or in fact, already the rescaled model MDP (10)) depends on the state  $x$  only through  $y(x)$  given in (11) (so the limit problem is indeed 1-D).

Figs. 9 and 10 show the  $V(x)$  averaged over all values of  $x$ 's that give the same  $y$ . For each  $y$ , Fig. 9 presents in addition to the average, also the discrepancy (i.e., the variations) from the average, and Fig. 10 presents the standard deviations from that average, both for the MDP (10) model. In both figures,  $V(x)$  is given in the vertical axis and  $y$  (taking values between 0 and 200) is given in the horizontal one.

The maximal discrepancy is always inferior to 1.34% of the average of the  $V(x)$ .

The standard deviation is always inferior to 1.4% of the average of  $V(x)$ .

We conclude that the MDP (10) model is already a good approximation of the one dimensional limit fluid model.

## V. CONCLUSION

We considered a discrete model for optimal call admission control with multiple classes of customers without waiting room. We have used DP techniques to solve this model, and we obtained structural properties of the optimal policies. In particular, in the case of two classes we showed that for each of the two classes there exist a threshold such that a customer of the other class is admitted only if the number of them in the system is under the threshold. We further illustrated by numerical examples that optimal policies are in general not monotone.

We then considered an approximation of the discrete model by a fluid model. We solved the new model and showed that, under fairly general assumptions, the control problem can be reduced to a 1-D state space model. We presented numerical experimentation that validate our results.

## APPENDIX

We use the approach and notation of [24] to solve (14). First, time is discretized to an interval of a small length  $h$ . This gives rise to a discrete time MDP whose value,  $\tilde{V}^h$  satisfies [24, p. 81]

$$\tilde{V}^h(y) = \max_{w \in D(y)} \left[ hg(w) + e^{-\beta h} \tilde{V}^h(y + h(w - \mu y)) \right].$$

We then discretize the state and obtain an MDP with the new finite state space:  $Y^h = \{0, h, 2h, \dots, B\}$ . Denote the corresponding value by  $V^h$ . To obtain  $V^h$  at  $y \in Y^h$ , we replace the argument in  $\tilde{V}^h(y + h(w - \mu y))$  by the convex combination of the values at points in  $Y^h$ . This results in the following DP equation [24, p. 92]:  $V^h(y) = TV^h(y)$  for all  $y \in Y^h$ , where

$$T^h f(y) \stackrel{\text{def}}{=} \max_{w \in D(y)} \left[ hg(w) + e^{-\beta h} (\gamma w f(z + h) + \gamma \mu y f(z - h) + (1 - \gamma(w + \mu y)) f(z)) \right] \quad (15)$$

where  $\gamma \stackrel{\text{def}}{=} \left( \sum_{i=1}^N \lambda_i b_i + \mu B \right)^{-1}$ . Equation (15) can be interpreted as a DP corresponding to a *stochastic* control problem,



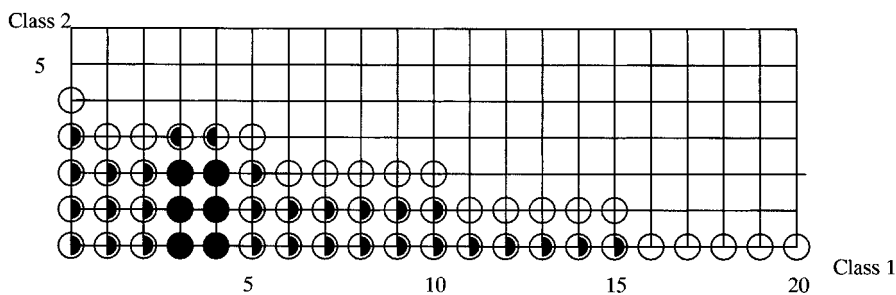


Fig. 7. Optimal policy for the MDP.

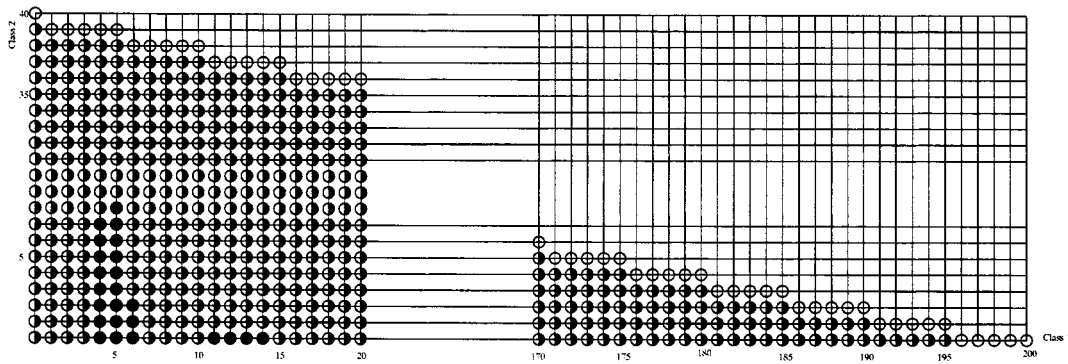


Fig. 8. Optimal policy for the MDP (10).

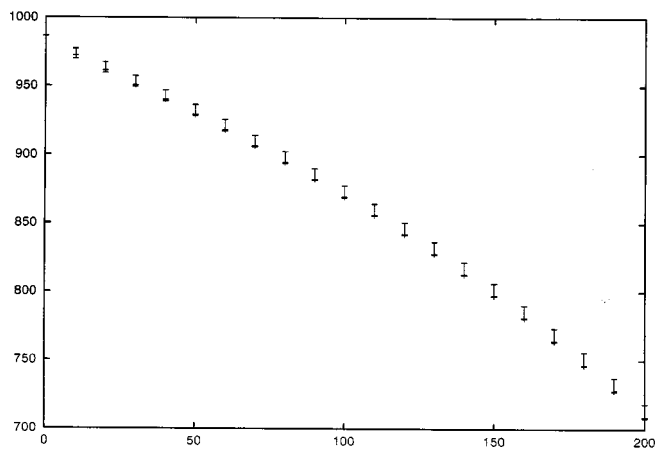


Fig. 9. Average value of  $V(x)$  and the discrepancy.

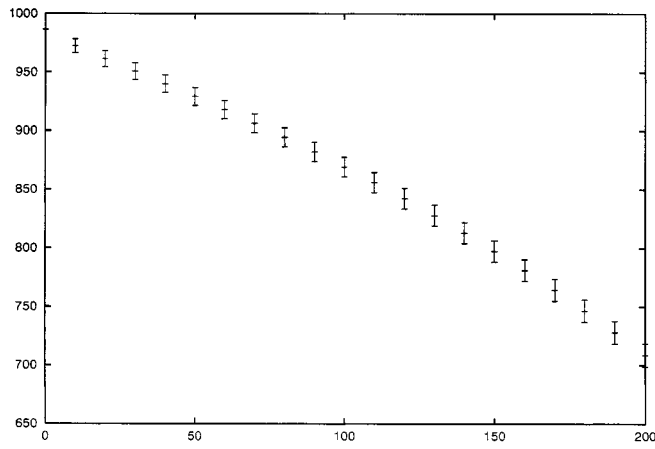


Fig. 10. Average value of  $V(x)$  and the standard deviation.

and can thus be solved using standard value iteration. The convergence of  $V^h(y)$  to  $v(y)$  is established in [24, p. 83] (a linear interpolation can be used to approximate  $v(y)$  at  $y \notin Y^h$ ). Note however that in this reference  $D(y)$  is taken to be independent of  $y$ . The proof for our case, in which  $D(y)$  is different at the boundary  $y = B$  is handled in the same way using the two discretization steps, see [25], and the convergence follows from Theorems 2.7 and 4.2 in [25] (or the Appendix of [22]).

A. Proof of Theorem III.13

i) If a function  $f : Y^h \rightarrow \mathbb{R}$  is concave decreasing then so is  $T^h f$ . This follows from Lemma 3.1 in [13] (which is formulated in terms of minimizing costs, thus the rewards should be

multiplied by  $-1$  and max should be changed into min before applying this lemma) by noting that

$$T^h f(y) = e^{-\beta h}(\gamma \mu y f(z-h) + (1 - \gamma \mu y)f(z)) + \max_{w \in D(y)} \left[ hg(w) + e^{-\beta h}(\gamma w(f(z+h) - f(z))) \right]. \quad (16)$$

Using this for  $f = 0$  together with value iteration we conclude that  $V^h = \lim_{n \rightarrow \infty} (T^h)^n f$  is also concave decreasing, and hence also the limit as  $h \rightarrow 0, v$ . This implies the required monotonicity of  $w_i^*(y)$ .

ii) That  $w^*$  is decreasing follows from the approximation and i). To prove that  $w^*$  is piecewise constant we first note that the function  $g(\cdot)$  is concave increasing and piecewise linear. For all

$y$  the maximum is obtained in one of the finite set of points  $\sum_{i=1}^I \lambda_i b_i$ , or in the point for which the drift in  $B$  is 0. Thus  $w^*$  is decreasing and there is only a finite set of possible values: therefore  $w^*$  is piecewise constant.

The last statement would follow directly from [21, Thm III.9.1] if the value were continuously differentiable. Since this is not guaranteed we modify that proof. Note that only at points in  $\hat{Y}$  (other than  $y_0$  and  $y_{|\hat{Y}|}$ ),  $dv(y)/dy$  need not exist. One can thus proceed as follows to generate  $w^*$ . (a) At  $y \notin \hat{Y}$ , select an action that achieves the argmax of the brackets in (14). (b) Select optimally the actions at  $y_n, 0 < n < |\hat{Y}|$ . The proof of [21, Thm III.9.1] can still be used for each interval  $I_n$  to show that the  $w^*$  is optimal. The question then remains of how to perform (b). Since  $\hat{Y}$  is finite, it follows from Lemma III.12 that either it is optimal to accept arrivals at all states, or that any choice of  $w^*(y)$  (which has already been determined for  $y \in \hat{Y}$ ) that guarantees that  $y(t)$  is monotone strictly increasing (for  $y < B$ ) will guarantee that  $w^*$  is optimal; indeed, all such policies give the same reward. One can choose in particular  $w^*(y_n) = \lim_{y \downarrow} w^*(y)$ , which results in a monotone decreasing policy. ■

#### REFERENCES

- [1] K. W. Ross and D. H. K. Tsang, "The stochastic knapsack problem," *IEEE Trans. Commun.*, vol. 37, pp. 740–747, 1989.
- [2] M. Gavious and Z. Rosberg, "A restricted complete sharing policy for a stochastic knapsack problem in B-ISDN," *IEEE Trans. Commun.*, vol. 42, pp. 2375–2379, July 1994.
- [3] K. W. Ross and D. D. Yao, "Monotonicity properties for the stochastic knapsack," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1173–1179, 1990.
- [4] P. Nain, "Qualitative properties of the Erlang blocking model with heterogeneous user requirements," *Queueing Syst.*, vol. 6, pp. 189–206, 1990.
- [5] E. A. Feinberg and M. I. Reiman, "Optimality of randomized trunk reservation," *Probab. Eng. Inform. Sci.*, vol. 8, pp. 463–489, 1994.
- [6] B. Miller, "A queueing reward system with several customer classes," *Management Sci.*, vol. 16, pp. 234–245, 1969.
- [7] S. A. Lippman, "Applying a new device in the optimization of exponential queueing systems," *Operat. Res.*, vol. 23, pp. 687–710, 1975.
- [8] J. D. Papatavrou, S. Rajagopalan, and A. J. Kleywegt, "The dynamic and stochastic knapsack problem with deadlines," *Management Sci.*, vol. 42, pp. 1706–1718, 1996.
- [9] S. Meyn, "The policy improvement algorithm for markov decision processes with general state space," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 191–196, 1997.
- [10] R.-R. Chen and S. Meyn, "Value iteration and optimization of multiclass queueing networks," *Queueing Syst.*, vol. 32, pp. 65–97, 1999.
- [11] J. Dai, "On the positive harris recurrence for multiclass queueing networks: A unified approach via fluid limit models," *Ann. Appl. Probab.*, vol. 5, pp. 49–77, 1995.
- [12] D. Bertsimas and X. Luo, "Separated continuous linear programming: Theory, algorithms, and applications," *SIAM J. Optimiz.*, to be published.
- [13] G. M. Koole, "Structural results for the control of queueing systems using event-based dynamic programming," *Queueing Syst.*, vol. 30, pp. 323–339, 1998.
- [14] M. L. Puterman, *Markov Decision Processes*. New York: Wiley, 1994.
- [15] E. Altman and G. M. Koole, "On submodular value functions and complex dynamic programming," *Stochastic Models*, vol. 14, pp. 1051–1072, 1998.
- [16] P. Glasserman and D. D. Yao, "Monotone optimal control of permutable GSMPs," *Math. Operat. Res.*, vol. 19, pp. 449–476, 1994.
- [17] R. Nagarajan, R. Ramjee, and D. Towsley, "On optimal call admission control in cellular networks," in *IEEE Infocom '96*, 1996, pp. 43–50.
- [18] E. Altman, T. Jiménez, and G. M. Koole, "On the comparison of queueing systems with their fluid limits," *Probability in the Engineering and Informational Sciences*, vol. 15, pp. 165–178, 2001.
- [19] R. F. Serfozo, "An equivalence between continuous and discrete time markov decision processes," *Operat. Res.*, vol. 27, pp. 616–620, 1979.
- [20] H. M. Soner, "Optimal control problem with state-space constraint," *SIAM J. Contr. Optimiz.*, vol. 24, pp. 552–562, 1986.
- [21] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*. New York: Springer-Verlag, 1993.
- [22] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*. San Francisco, CA: Birkhäuser, 1997.
- [23] H. Frankowska and S. Plaskacz, "Hamilton–Jacobi equations for infinite horizon control problems with state constraints," *Cahiers du Centre de Recherche Viabilité*, Univ. Paris Dauphine, Jeux. Contrôle, 9808, 1998.
- [24] H. J. Kushner and P. G. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*. New York: Springer, 1992.
- [25] F. Camilli and M. Falcone, "Approximation of optimal control problems with state constraints: estimates and applications," in *Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control*, B. S. Mordukhovich and H. J. Susmann, Eds. New York: Springer, 1996.



**Eitan Altman** received the B.Sc. degree in electrical engineering, the B.A. degree in physics, and the Ph.D. degree in electrical engineering, all from the Technion-Israel Institute, Haifa, Israel, in 1984, 1984, and 1990, respectively. In 1990, he received the B.Mus. degree in music composition from Tel-Aviv University.

Since 1990, he has been with INRIA (National Research Institute in Informatics and Control), Sophia-Antipolis, France. His current research interests include performance evaluation and control of telecommunication networks, stochastic control, and dynamic games. In recent years, he has applied control theoretical techniques in several joint projects with the French telecommunications company—France Télécom.



**Tania Jiménez** received the Ph.D. degree in parallel simulations from the University of Nice, Sophia-Antipolis, France, in 2000.

Her research interests include simulation as well as optimization and control of telecommunication networks. She is currently a teaching and research assistant with the Universidad de Los Andes, Merida, Venezuela, teaching computer science and simulation courses.



**Ger Koole** received the Ph.D. degree in 1992 from Leiden University in the area of stochastic operations research.

He held postdoctoral positions with CWI, Amsterdam, and INRIA, Sophia-Antipolis, France. Currently, he is a Professor with the Mathematics and Computer Science Department of the Vrije Universiteit Amsterdam. His research interests include queueing, stochastic control theory, and stochastic operations research and their applications to telecommunications and logistics. He is an

Associate Editor of *Operations Research Letters*.