



LUND UNIVERSITY

A monotonic property of the optimal admission control to an M/M/1 queue under periodic observations with average cost criterion

Cao, Jianhua; Nyberg, Christian

Published in:

Seventeenth Nordic Teletraffic Seminar, NTS 17, Fornebu, Norway, 25-27 August 2004

2004

[Link to publication](#)

Citation for published version (APA):

Cao, J., & Nyberg, C. (2004). A monotonic property of the optimal admission control to an M/M/1 queue under periodic observations with average cost criterion. In *Seventeenth Nordic Teletraffic Seminar, NTS 17, Fornebu, Norway, 25-27 August 2004* Fornebu : Telenor.

Total number of authors:

2

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

This is an author produced version of a paper presented at the
17th Nordic Teletraffic Seminar (NTS 17), Fornebu, Norway,
25-27 August, 2004.

This paper may not include the final
publisher proof-corrections or pagination.

Citation for the published paper:

J. Cao and C. Nyberg, 2004,

"A Monotonic Property of the Optimal Admission Control to an M/M/1
Queue under Periodic Observations with Average Cost Criterion",
*Seventeenth Nordic Teletraffic Seminar, NTS 17, Fornebu, Norway,
25-27 August 2004.*

ISBN: 82-423-0595-1. Publisher: Fornebu : Telenor.

A Monotonic Property of the Optimal Admission Control to an M/M/1 Queue under Periodic Observations with Average Cost Criterion

Jianhua Cao and Christian Nyberg
Department of Communication Systems
Lund University, Sweden
{jcao, cn}@telecom.lth.se

Abstract

We consider the problem of admission control to an M/M/1 queue under periodic observations with average cost criterion. The admission controller receives the system state information every τ :th second and can accordingly adjust the acceptance probability for customers who arrive before the next state information update instance. For a period of τ seconds, the cost is a linear function of the time average of customer populations and the total number of served customers in that period. The objective is to find a stationary deterministic control policy that minimizes the long run average cost. The problem is formulated as a discrete time Markov decision process whose states are fully observable. By taking the control period τ to 0 or to ∞ , the model in question generalizes two classical queueing control problems: the open and the closed loop admission control to an M/M/1 queue. We show that the optimal policy is to admit customers with a non-increasing probability with respect to the observed number of customers in the system. Numerical examples are also given.

1 Introduction

We consider the problem of admission control to an M/M/1 queue under periodic observations with average cost criterion. The admission controller receives the system state information every τ :th second and can accordingly adjust the control policy which is the probability that an arrival will be accepted into the system. For a period of τ seconds, the cost for making a particular decision is a linear function of the time average of customer populations waiting in queue and the total number of served customers in that period. The objective is to find a stationary deterministic control policy that minimizes long run average cost.

Our objective function definition which rewards departures and penalizes a long queue resembles to that of Naor (1969) and many others, see Stidham Jr. (1985); Stidham Jr. (1988, 2002) and references therein. Along this line of objective function definition, the optimal admission control in the cases of complete observable and non-observable queueing system are well understood see, e.g., Hassin and Haviv (2003). In

a brief summary, when the queue length is completely observable, the optimal policy is of threshold type and when the queue length is not observable, the optimal policy accepts customers with a fixed probability.

Many models of queueing system admission control, including ours, are based on Markov decision theory, see (Stidham Jr. and Weber, 1993, Section 4) and references therein. One ramification to the existing admission control problems is reducing the amount of information used for decision. Much work has been done along this direction. Fukuda (1986) also considered a queueing system under periodic observations. There are two types of customers in his system with the same service time distribution and different arrivals processes. The control policy considered is call gapping control where the low priority customers are blocked for the next period whenever the observed queueing system state exceeds a predetermined limit. Kuri and Kumar (1995) studied the discrete time admission control and routing to a queueing system consisting of two parallel queues with delayed queue length information. Their objective is to minimize the total expected discounted cost in which a fixed reward is received for each admitted customers and linear holding cost is incurred for customers waiting in queue. The optimal policy in case of one period information delay is of threshold type. Lin and Ross (2003) considered a multiple-server loss model where the admission controller is informed when an admitted customer finds all servers are busy but not informed when customers depart. In this system a cost is incurred if a new arrival is blocked and an even larger cost induced if an admitted customer is blocked by servers. They proved that the threshold type policy that blocks for a certain amount of time after an admitted arrival is optimal in case of single servers.

In this paper we establish that the average cost optimal policy is nonincreasing. This is an important property that can be exploited to accelerate numerical computations.

This paper is organized as follows: In Section 2 we define the Markov decision process model for our problem. We show that the average optimal solution is nonincreasing in Section 3. Numerical examples are give in Section 4. Finally, we conclude with some remarks.

2 The Markov Decision Process Model

2.1 Classes of Control Policies and Induced Markov Chains

We control the arrival rate to an M/M/1 queue using percentage blocking. For convenience, we assume that the service rate is 1 in the rest of this paper. The arrival rate to the system is λ which can be less or greater than 1. There is an admission controller that rejects arrivals in order to maintain a reasonable response time for admitted customers. The controller is characterized by the admission probability a , $0 \leq a \leq 1$. We are allowed to adjust the admission probability every τ :th seconds when the updated system state information becomes available. Fig. 1 illustrates the model.

Let $S := \{0, 1, 2, \dots\}$ be the set of system states. In our problem, the system state is the number of customers in the queue. A stationary deterministic control policy is an infinite sequence $(\pi(0), \pi(1), \dots) =: \pi$, with $0 \leq \pi(i) \leq 1, \forall i \in S$. It specifies that whenever the observed number of customers in the system is $i \in S$, the admission

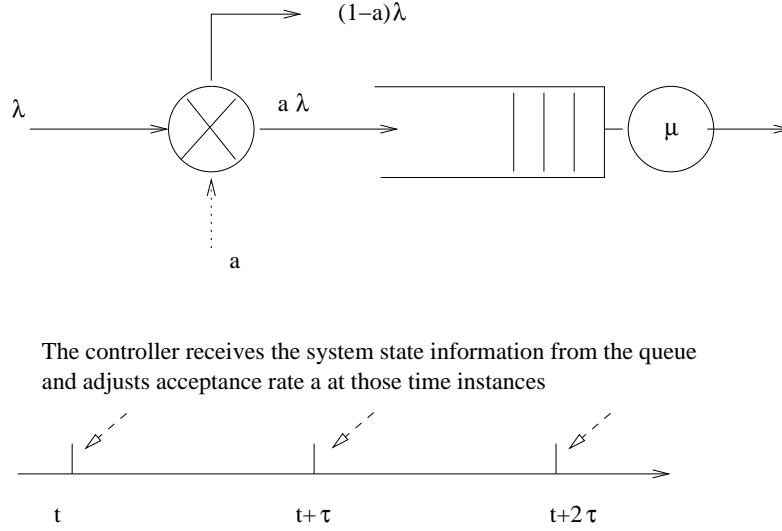


Figure 1: Arrival rate control to an M/M/1 queue with periodic observations.

probability should be adjusted to $\pi(i)$ and remain the same for the next τ seconds. Let Π be the set of stationary deterministic control policies. A more generic control class is, e.g., $\{\pi_k(i, t) \in [0, 1] \mid t \in [0, \tau), k \in \mathbb{N}, i \in S\}$, where $\pi_k(i, t)$ is the admission probability at time $k\tau + t$ when the observed number of customers at time $k\tau$ is i . We shall, however, content ourselves with the class of stationary deterministic control policies in this paper.

Let t_0, t_1, t_2, \dots be the time instances when the admission probability is adjusted and $t_0 = 0, t_{k+1} - t_k = \tau, \forall k \in \mathbb{N}$. Let $X(t) \in S$ be the system state at time t . Clearly $\{X(t_k)\}_{k=0}^{\infty}$ can be regarded as a discrete time Markov chain. For convenience denote $X_k := X(t_k)$.

The transition probability between states i and j is denoted as $P(j \mid i, \pi(i))$. It is actually the probability that at time τ there are j customers in an M/M/1 queue with arrival rate $\pi(i)\lambda$ when there are i customers in the system initially. Once a control policy $\pi \in \Pi$ and the initial state are given, the associated Markov chain $\{X_k\}$ is fully specified. Notice that the transition probability $P(\cdot \mid \cdot, a)$ is continuous in a .

2.2 One Step Cost

Given a fixed arrival rate, one would like to accept as few customers as possible in order maintain a short mean response time. But, on the other hand, to have a high throughput, the arrival rate ought to be close to the service rate. Hence one has to balance low response time and high throughput. One way to resolve this dilemma is to formulate an constrained optimization problem where the throughput is the objective to be maximized while the response time is subject to a given constraint. Another way is to consider an unconstrained optimization problem whose objective is to minimize

a cost which is a function of response time and throughput. This paper discusses the second approach.

To make the definition of cost more precise, we introduce some auxiliary notation.

Let $\bar{N}(t, i, a)$, $\bar{N} : [0, \tau] \times S \times [0, 1] \mapsto [0, \infty)$, be the expected number of customers in an M/M/1 queue at time t when initially there are i customers and the admission probability is a .

The time average number of customers in the system between time 0 and τ is

$$\frac{1}{\tau} \int_0^\tau \bar{N}(t, i, a) dt.$$

The average number of severed customers(not including the rejected customers) between time 0 and τ is

$$\frac{1}{\tau} (i + a\lambda\tau - \bar{N}(\tau, i, a)).$$

Let C be the cost to maintain one customer in the system per time unit and R be the reward for a departure. To avoid the trivial case, we assume that $0 < C < R$. The one-step cost, i.e. the time average of net cost incurred in a period of τ seconds between two adjacent control parameter adjustment instances, is denoted as $c(i, a)$, and is defined as follows,

$$c(i, a) := \frac{C}{\tau} \int_0^\tau \bar{N}(t, i, a) dt - \frac{R}{\tau} (i + a\lambda\tau - \bar{N}(\tau, i, a)). \quad (1)$$

It is easy to convince oneself that the step cost $c(i, a)$ is not necessarily monotonic in i or a .

2.3 Average Cost and β -Discount Cost

The expected long-run average cost or average cost incurred by a policy $\pi \in \Pi$ is defined as

$$J(i, \pi) := \limsup_{N \rightarrow \infty} E_i^\pi \left[\frac{1}{N} \sum_{k=0}^{N-1} c(X_k, \pi(X_k)) \right],$$

where i is the initial number of customers in the system .

The optimal average cost for any initial state $i \in S$ is defined as

$$J^*(i) := \inf_{\pi \in \Pi} \{J(i, \pi)\}.$$

A policy π is average cost optimal, or average optimal, if $J(i, \pi) = J^*(i)$ for all $i \in S$.

While the average cost is of primal interest of this paper, β -discounted cost is helpful in the proof of the existence of the average optimal policy.

Given some $\beta \in (0, 1)$, the β -discount cost incurred by a policy $\pi \in \Pi$ is defined as

$$J_\beta(i, \pi) := E_i^\pi \left[\sum_{k=0}^{\infty} \beta^k c(X_k, \pi(X_k)) \right].$$

For any initial state $i \in S$, the optimal β -discount cost is defined as

$$J_\beta^*(i) := \inf_{\pi \in \Pi} \{J_\beta(i, \pi)\}$$

A policy π is β -discount cost optimal, or β -discount optimal, if $J_\beta(i, \pi) = J_\beta^*(i)$ for all $i \in S$.

2.4 The Existence of Average Cost Optimal Policy

There are at least three well developed methods to prove the existence of the average optimal policy when the state space is enumerable infinite and step cost is unbounded see Arapostathis et al. (1993, Section 5.2 and Section 5.3) for review. The first one is Hordijk's Lyapunov stability condition (Hordijk, 1974). Sennott's three necessary conditions (Sennott, 1989, 1998, 2000) can be counted as the second. The third one is Borkar's convex analytic approach (Borkar, 1988, 2000). Sennott's conditions are particularly well suited for queueing problems (Sennott, 1998) where the step cost often, if not always, grows as more customers accumulate in the queue. Our proof of the existence of the average optimal policy for our particular model is based on Sennott's conditions.

A complete proof of the existence of average optimal policy for a Markov decision process satisfies Sennott's three conditions relies on the vanishing discounted approach. The general idea is that one treats the average cost case as the limit of the discounted cost problem. The vanishing discounted approach was also used by Blackwell (1962) to prove the existence of average optimal policy for a model with finite state space and action set, and Derman (1966) for countable state space, finite action set and bounded costs.

The existence β -discount optimal policy of our problem is easy to check, see Arapostathis et al. (1993, Lemma 2.1). We shall verify Sennott's three conditions in our particular problem in Lemma 1 which guarantees existence of average optimal policy using vanishing discount approach see e.g. Arapostathis et al. (1993, Theorem 5.9)

Lemma 1. (a) For every $i \in S$ and every $\beta \in (0, 1)$, $J_\beta^*(i) < \infty$.

(b) There exists a nonnegative integer L such that

$$h_\beta(i) := J_\beta^*(i) - J_\beta^*(0) \geq -L$$

(c) There exists a function $M : S \mapsto \mathbb{R}$ such that $h_\beta(i) \leq M(i)$ for all $i \in S$ and $\beta \in (0, 1)$ and for every $i \in S$ and $a \in [0, 1]$ such that $\sum_j P(j | i, a) M(j) < \infty$.

Proof. (a). Clearly there exists a stationary deterministic policy, e.g. $\pi := (0, 0, \dots)$ such that the induced MC is ergodic and $J(i, \pi) = 0$ for all $i \in S$. By the Tauberian Theorem, we have

$$\liminf_{\beta \uparrow 1} (1 - \beta) J_\beta^*(i) \leq \limsup_{\beta \uparrow 1} (1 - \beta) J_\beta^*(i) \leq J(i, \pi)$$

Hence $J_\beta^*(i)$ is finite for all $i \in S$ and $\beta \in (0, 1)$.

(b) Consider a queueing system with two M/M/1 queues, one queue with i customers initially but no arrivals, another queue starts empty and the β -discount optimal policy

is employed. The discounted cost for the whole system is therefore $J_\beta^*(0) + J_\beta(i, \mathbf{0})$, where $\mathbf{0} = (0, 0, \dots)$ is a policy that rejects all arrivals. Clearly $J_\beta^*(0) + J_\beta(i, \mathbf{0}) \leq J_\beta^*(i)$, hence we have

$$J_\beta^*(i) - J_\beta^*(0) \geq J_\beta(i, \mathbf{0})$$

Let $n := \lfloor R/C \rfloor$. For all $i \in S$, $J_\beta(i, \mathbf{0}) \geq J_\beta(n, \mathbf{0}) \geq -nR$. Therefore $h_\beta(i) \geq -\lfloor R/C \rfloor R$.

(c) Let Π^* be the class of policies inducing irreducible, ergodic MC and $c_{i,0}(\pi)$ the expected total cost of a first passage from i to 0. Clearly Π^* is not empty. Let $M(0) = 0$ and for $i \geq 1$, $M(i) = \inf_{\pi \in \Pi^*} c_{i,0}(\pi)$. We then have $J_\beta^*(i) \leq M(i) + J_\beta^*(0)$ and for all $i \in S$ and $a \in [0, 1]$, and some $\pi \in \Pi^*$,

$$\sum_{j \in S} P(j | i, a) M(j) \leq \sum_j P(j | i, a) c_{i,0}(\pi) = c_{i,0}(\pi) < \infty.$$

□

3 The Monotonic Property of the Average Cost Optimal Policy

Intuitively when more customers are observed in one control instance, less should be admitted into the system in the next control period of τ seconds. Hence it is reasonable to conjecture that the average optimal policy φ is nonincreasing in i , i.e. $\varphi(i) \geq \varphi(i+1), \forall i \in S$. This fact can be easily verified in the cases of closed loop and open loop admission control. In this section we shall prove that this is true in general when $0 \leq \tau < \infty$. Let $N(t, i, a)$ be the number of customers in a queue (with arrival rate $a\lambda$ and service rate 1) at time t with initial condition $N(0, i, a) = i$. Notice that given t, i, a , $N(t, i, a)$ is a random variable.

A function $f: \mathbb{N} \rightarrow \mathbb{R}$ is discrete convex if $f(i+2) - f(i+1) \geq f(i+1) - f(i)$ for all $i \in \mathbb{N}$. Let \mathfrak{F} be the space of discrete convex functions that map from S to \mathbb{R} . The following lemma follows.

Lemma 2. *Given $t > 0$, for all $f \in \mathfrak{F}$, $E[f(N(t, i, a))]$ is supermodular in (i, a) .*

Proof. We have to show that for all $f \in \mathfrak{F}$, $i \in S$, $0 \leq b < a \leq 1$,

$$E[f(N(t, i+1, a))] + E[f(N(t, i, b))] \geq E[f(N(t, i, a))] + E[f(N(t, i+1, b))]$$

The following stochastic order relations can be verified by the normalization technique: $N(t, i+1, a) \geq_{st} N(t, i, a)$, $N(t, i+1, a) \geq_{st} N(t, i+1, b)$, $N(t, i+1, b) \geq_{st} N(t, i, b)$, $N(t, i+1, a) - N(t, i, a) \geq_{st} N(t, i+1, b) - N(t, i, a)$. Since f is discrete convex, we have $f(N(t, i+1, a)) - f(N(t, i, a)) \geq_{st} f(N(t, i+1, b)) - f(N(t, i, a))$ which leads to the claim after taking expectation on both sides. □

Lemma 3. *Given $t > 0$, for all $f \in \mathfrak{F}$. $i \in S$, $0 \leq b < a \leq 1$,*

$$E[f(N(t, i, a))] + E[f(N(t, i + 2, b))] \geq 2E\left[f\left(N\left(t, i + 1, \frac{a+b}{2}\right)\right)\right]$$

Proof. Since f is discrete convex, it is enough to show that

$$N(t, i, a) + N(t, i + 2, b) \geq_{st} 2N\left(t, i + 1, \frac{a+b}{2}\right). \quad (2)$$

Consider a system consisting of two identical but independent M/M/1 queues. The total arrival rate is $(a + b)\lambda$ and initially there are $2i + 2$ customers in total in two queues. Suppose we are allowed to distribute $2i + 2$ customers into two queues at time 0 and split the incoming traffic by random routing. Total number of customers in the whole system at time t can be minimized (in stochastic order sense) by placing $i + 1$ customers in each queue and splitting traffic evenly. Thus (2) holds. \square

Lemma 4. Define $g : \mathfrak{F} \times S \times [0, 1] \mapsto \mathbb{R}$ as follows $g(f)(i, a) := c(i, a) + E[f(N(\tau, i, a))]$, then $g(f)(i, a)$ is supermodular in (i, a) and $\min_a g(f)(i, a)$ is convex in i .

Proof. (Supermodularity) By Lemma 2, $E[N(t, i, a)]$ is supermodular in (i, a) . Recall the definition of the step cost

$$c(i, a) = \frac{C}{\tau} \int_0^\tau E[N(t, i, a)] dt - \frac{R}{\tau} (i + a\lambda\tau - E[N(\tau, i, a)]).$$

Hence $c(i, a)$ is supermodular in (i, a) too. Also from Lemma 2, $E[f(N(\tau, i, a))]$ is supermodular. Thus $g(f)(i, a)$ is supermodular in (i, a) .

(Convexity) We have to show that $\forall i \in S$

$$\min_a g(f)(i + 2, a) + \min_a g(f)(i, a) \geq 2 \min_a g(f)(i + 1, a).$$

We shall show that there exist some $w \in [0, 1]$ such that

$$\min_a g(f)(i + 2, a) + \min_a g(f)(i, a) \geq 2g(f)(i, w) \quad (3)$$

Let $u := \arg \min_a g(f)(i, a)$, $v := \arg \min_a g(f)(i + 2, a)$. By the supermodularity of g , $v \leq u$. If $u = v$, then let $w = u = v$, the inequality above is evident. If $v < u$ then let $w = \frac{u+v}{2}$. By Lemma 3,

$$E[N(t, i + 2, v)] + E[N(t, i, u)] \geq 2E[N(t, i + 1, w)]$$

and

$$E[f(N(t, i + 2, v))] + E[f(N(t, i, u))] \geq 2E[f(N(t, i + 1, w))].$$

Therefore (3) is valid. \square

Theorem 1. If $\varphi \in \Pi$ is average optimal, then $\varphi(i) \geq \varphi(i + 1)$ for all $i \in S$.

Proof. Consider the value iteration. The value function $V_k : S \mapsto \mathbb{R}$ at the k :th iteration is defined as follows

$$V_{k+1}(i) = \min_a \left\{ c(i, a) + \sum_j V_k(j) P(j | i, a) \right\}$$

with $V_0(j) = 0, \forall j \in S$. The average optimal policy is defined as follows

$$\varphi(i) = \lim_{k \rightarrow \infty} \arg \min_a c(i, a) + \sum_j V_k(j) P(j | i, a), \forall i \in S$$

By induction, we have that for all k , $V_k(i)$ is convex in i and $c(i, a) + \sum_j V_k(j) P(j | i, a)$ is supermodular in (i, a) . Hence $\varphi(i)$ is nonincreasing in i . \square

4 Numerical Examples

In our model, the state space is countable infinite. For numerical calculations, however, the state space must be truncated (cutting off the tail). Rigorous analysis of the effects of state space truncation for dynamic programming in general can be found in Fox (1971); Whitt (1978, 1979). The approximating sequence method described in Sennott (1998, Chapter 8) is also very illuminating. In our particular case, we note that the difference in the optimal cost between the original model and the state space truncated model decreases as the probability of unexpected blocking diminishes. Once the state space is truncated, both the value iteration algorithm and the policy iteration algorithm (Puterman, 1994, Section 8.5 and 8.6) can be used to calculate the optimal policy and the optimal average cost. The policy iteration algorithm which is used in our numerical examples is described briefly as follows.

Step 1 Initialization: An arbitrary stationary policy π is chosen.

Step 2 Value determination: For the current policy π , compute the unique solution $\{J, h_i\}$ for the following system of linear equations

$$\begin{aligned} h_i &= c(i, \pi(i)) - J + \sum_{j \in S'} h_j P(j | i, \pi(i)), \quad \forall i \in S' \\ h_0 &= 0 \end{aligned}$$

where S' is the truncated state space.

Step 3 Policy improvement: The new policy π' is obtained as follows

$$\pi'(i) = \arg \min_{0 \leq a \leq 1} c(i, a) - J + \sum_{j \in S'} h_j P(j | i, a) \quad \forall i \in S'$$

Step 4 Convergence test. Let $d(\pi', \pi)$ be a distance measure of two policies such as follows $d(\pi', \pi) = \sum_{i \in S'} |\pi'(i) - \pi(i)|$. If the new policy π' is within the give distance ϵ of the old policy π , i.e. $d(\pi', \pi) < \epsilon$, the algorithm stops, otherwise goto step 2 with π replaced by π' .

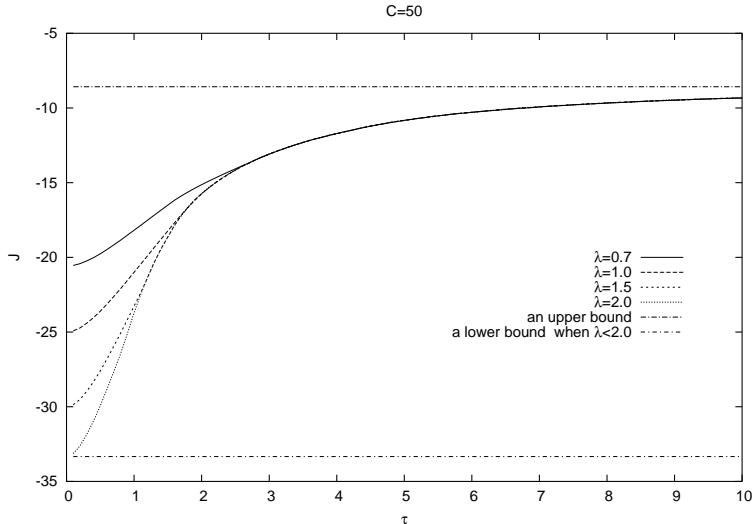


Figure 2: Average optimal cost versus the control interval τ when $C = 50$.

Note that due to the state space truncation, the expression for the average number of customers in the system $\bar{N}(i, t, a)$ and the transition probability $P(j | i, a)$ must be swapped from the M/M/1 model to the M/M/1/K model where K is the truncated queue size.

Using the same techniques in proving Lemma 4 and Theorem 1, one can show that the function to be minimized in the policy improvement step is convex and supermodular if the policy chosen in the initialization step is nonincreasing. As a result many convex programming techniques can be used to find the minimum in step 3 and further, the search region for $\pi'(i)$ can be narrowed from $[0, 1]$ to $[0, \pi'(i-1)]$.

In the following discussions, we set the reward coefficient to a constant $R = 100$.

In Fig. 2, we show how the optimal average cost changes as the control interval τ increases for different combinations of cost coefficient $C = 50$ and arrival rates $\lambda = 0.7, 1.0, 1.5, 2.0$. Notice that when the offered traffic intensity and the cost of waiting are high, e.g. $C = 50, \lambda = 2.0$, the optimal average cost for models with short control intervals is sensitive to τ . We also plot the theoretical lower and the upper bound of the optimal average cost in these figures.

We have shown in Section 3 that the average optimal policy is nonincreasing. In Fig. 3, we give examples of the average optimal control policies in cases of $\tau = 0.1, 5.0, 10.0$ when $\lambda = 1.0$ and $C = 30$. The optimal control policy in the case of open loop control $\tau \rightarrow \infty$ is also shown. If the closed loop control is used (i.e. $\tau = 0.0$) when $\lambda = 1.0, C = 30$, the average optimal policy is of threshold type and the optimal threshold is 1.

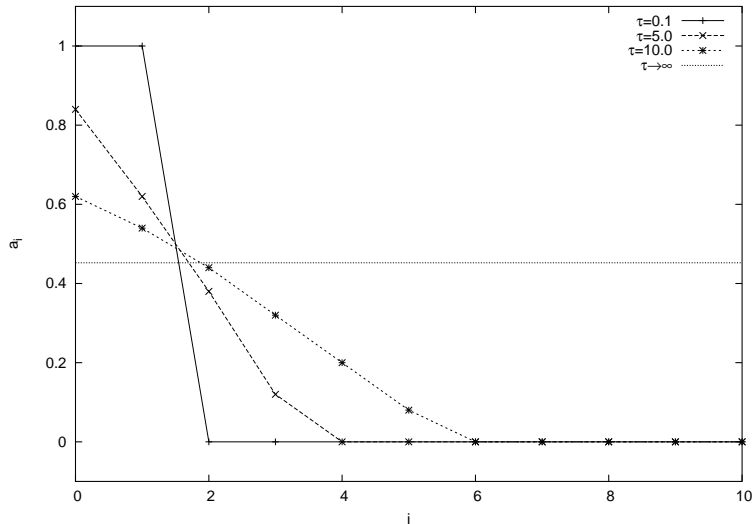


Figure 3: The average optimal control policy when $\lambda = 1.0$ and $C = 30$.

5 Conclusions

We have presented an admission control model for an M/M/1 queue under periodic observations with average cost criterion. This model degenerates into two well known queueing control problems when the observation interval becomes zero and infinite. The corresponding discrete time Markov decision process is obtained via embedding the state transition instances at the state information update epochs. The state transition probability is related to the transient solution of state probability distribution of an M/M/1 queue. The step cost is linearly proportional to the time average number of customers in the queue and the time average number of departures per time unit between two adjacent control instances. With help of value iteration algorithm and induction, we proved that the average optimal policy is nonincreasing. Several numerical examples are also provided in a separate section.

Many claims/arguments in the paper are not specific to the M/M/1 queue, e.g. Lemma 1 used in the proof of the existence of average optimal policy; Lemma 2 and 3 in the proof of the monotonic property of optimal policy. However one must proceed with caution when the approach in defining the discrete time Markov decision process for the M/M/1 model in this paper is extended to an M/G/1 or an G/M/1 queue.

References

- Arapostathis, A., Borkar, V., Fernández-Gaucherand, E., Ghosh, M. K., and Marcus, S. I. (1993). Discrete-time controlled Markov processes with average cost criterion: a survey. *SIAM J. Control and Optimization*, 31(2):282–344.

- Blackwell, D. (1962). Discrete dynamic programming. *The Annals of Mathematical Statistics*, 33(2):719–726.
- Borkar, V. S. (1988). *Stochastic Differential Systems, Stochastic Control Theory and Applications*, volume 10 of *The IMA Volumes in Mathematics and Its Applications*, chapter Control of Markov chains with long-run average cost criterion, pages 57–77. Springer-Verlag.
- Borkar, V. S. (2000). *Markov Decision Processes: Model, Methods and Open Problems*, chapter Convex analytic methods in Markov decision processes. Kluwer 2000.
- Derman, C. (1966). Denumerable state Markov decision processes, average cost criterion. *Ann. Math. Statist.*, 37:1545–1553.
- Fox, B. L. (1971). Finite-State Approximations to Denumerable-State Dynamic Programs. *J. Math. Anal. Appl.*, 34(665–670).
- Fukuda, A. (1986). Input regulation control based on periodical monitoring using call gapping control. *Electronics Comm. Japan, Part 1*, 69(11):84–92.
- Hassin, R. and Haviv, M. (2003). *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*. Kluwer Academic Publishers.
- Hordijk, A. (1974). Dynamic Programming and Markov Potential Theory. Technical report, Math. Centre Tract No. 51, Mathematisch Centrum, Amsterdam.
- Kuri, J. and Kumar, A. (1995). Optimal control of arrivals to queues with delayed queue length information. *IEEE Tran. Automatic Control*, 40(8):1444–1450.
- Lin, K. Y. and Ross, S. M. (2003). Admission control with incomplete information of a queueing system. *Operations Research*, 51(4):645–654.
- Naor, P. (1969). The regulation of queue size by levying tolls. *Econometrica*, 37(1).
- Puterman, M. L. (1994). *Markov Decision Processes, Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc.
- Sennott, L. I. (1989). Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operations Research*, 37:626–633.
- Sennott, L. I. (1998). *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley & Sons.
- Sennott, L. I. (2000). *Markov Decision Processes: Model, Methods and Open Problems*, chapter Average reward optimization theory for denumerable state spaces. Kluwer 2000.
- Stidham Jr., S. (1985). Optimal control of admission to a queueing system. *IEEE Tran. on Automatic Control*, 30(8).

- Stidham Jr., S. (1988). *Stochastic differential systems: stochastic control theory and applications*, chapter Scheduling, Routing, and Flow Control in Stochastic Networks, pages 529–561. Springer-Verlag.
- Stidham Jr., S. (2002). Analysis, design and control of queueing systems. *Operations Research*, 50(1):197–216.
- Stidham Jr., S. and Weber, R. (1993). A survey of Markov decision models for control of networks of queues. *Queueing Systems*, 13:291–314.
- Whitt, W. (1978). Approximations of dynamic programs I. *Mathematics of Operations Research*, 3(3):231–243.
- Whitt, W. (1979). Approximations of dynamic programs, II. *Mathematics of Operations Research*, 4(2):179–185.