

PERFORMANCE BOUNDS AND SUBOPTIMAL POLICIES FOR MULTI-CLASS QUEUE

A. Madankan, University of Zabol, Zabol, Iran, Amadankan@uoz.ac.ir

In this paper, we consider a general class of a queuing system with multiple job types and flexible service facility. We use a stochastic control policy to determine the performance loss in multi-class M/M/1 queue. The considered system is originally a Markov decision processes (MDP). The author showed how to compute performance bounds for the stochastic control policy of MDP with an average cost criteria. In practice, many authors used heuristic control policies due to some hardness in computing and running mathematically optimal policies. The authors found bounds on performance in order to an optimal policy where the goal of this job is to compute the difference of optimality and a specific policy. In other words, this study shows that, the optimal bounds of the average queue length for any non-idling policies can be found by a factor of service rates.

Keywords: queuing system; multiple job classes; stochastic control policy.

Introduction

We are interested to use an average cost per period (ACPP) in a multi-class job M/M/1 queue to determine the performance loss associated with using a control policy. We consider non-idling policies, which always mean serving jobs as long as there are jobs in the queue. The problem considered in this paper is to control a single server queue with multiple job class. [1,2] have shown that the optimal policy for this problem is known where the implementation of that optimal policy needs the exact information about the service rate for each class. However, note that it is a little difficult to analyse the performance of some policies, such as FIFO (see [1]).

Indeed it is known (see for example [3,4]) that the $c\mu$ -rule is the optimal control in two main settings: (i) generally distributed service requirements among all non-preemptive disciplines and (ii) exponentially distributed service requirements among all preemptive disciplines. In the preemptive case $c\mu$ -rule is only optimal if the service times are exponentially distributed. The queuing system that we considered to work on, is in the discrete-time case where the other case and its optimal control policy is studied before and is $c\mu$ rule. In the discrete-time case, optimality of $c\mu$ rule was established in [4,5]. We recall that $c\mu$ -rule is the discipline that gives a strict priority in descending order of $c_k\mu_k$, where c_k and μ_k refer to a cost and the inverse of the mean service requirement, respectively, of class k .

The problem that we considered is Markov decision processes with an objective of an average cost per period. The main job of this paper is to find a method to determine the performance loss associated with using an optimal control. We produce a systematic approach to reach that and to evaluate the difference between optimality and the costs from a specific policy. We use the presented methods to supply a relation between two costs, one by a specific sub-optimal policy and the other one by an optimal policy. We believe that our results open an interesting topic for the further research. For instance, well-known optimality results in a single-class queue like the optimality of the Shortest Service time discipline or the optimality of FCFS can all be derived as corollaries of the

queue. In order to get insights into the structure of the optimal policy in the multi-class case we consider several relevant cases where the service time distributions are exponential.

Finding an optimal control policy for a Markov decision process is one of the highlighted topics for infinite (or even very large) state space systems where computing it often is intractable [3]. It is not easy to run even the known general form of the optimal control policy due to some difficulties since at each time there is a slot in the discrete time process. In each time slot, we need high computational work to evaluate the costs of control action [6]. As a result, these are the reasons why we use sub-optimal heuristic control policies in practice.

In the present work, by using the general discussed methodology we present a factor which contains only the service rates for sub-optimality of the queue for any considered policy. The obtained bound establishes that there should not be much effects on the queue length when the service rates are approximately the same. Our contribution bounds on ACP for the problem of controlling multi-class queues. Many authors have already worked to find bounds in Markov chain problems and also in finding the optimal control and analysis of multi-class queues.

Due to the average cost of Markov decision processes for finite state Markov chains, as it is done in [7], bounds are used to provide convergence of an iteration algorithm. On the other hand for general state spaces that we considered here, the obtained bounds are associated with Lyapunov theorems for Markov chains where one can find a similar upper bound in [5]. For systems with positive unbounded costs, the standard Lyapunov theorems are used to produce upper bounds. The bounds can be considered universalization of the Lyapunov bounds and also the finite state bounds. These bounds have a feature that with unbounded costs one can provide an upper bound and lower bound.

1. Problem Definition

The model that we want to consider is a multi-classes queuing system where each job belongs to a variety distinct classes. On the assumption that the system deals with one class of the job, then the order of serving jobs is not important and also does not impress the quantities such as average queue length, the results from any simple control policies (such as first come-first service (FCFS)) and more complex control policies are the same. While jobs depend on service time then the order of serving impress on quantities such average queue length. There are many policies which minimizes an average queue length for finite multi-class M/M/1 queuing system where one can find in [3]. There are some differences between multi-classes queue rather than single class queue which, for instance, are (1) the frequently of the job arrival of some classes is more than some other classes, (2) the service time of some classes is longer than other classes.

Due to the optimal policy for the considered model, we give priority to classes according to an average service time which means job classes with shorter average service times have a higher priority than job classes with longer average service times. Also, we consider preempting in the service where it means we temporarily stop serving a low priority job when a job of higher priority arrives. This means we consider the difference between classes, need statistically information of serving all classes, and allow to preempt jobs in the service if it is needed. Due to sub-optimal control policy, let consider the set of policies that server serves all jobs in the queue without idling and taking rest. We call these theme non-idling

policies. We denote that there is a factor which contains only the service rates when it bounds the queue lengths. The obtained bound, for almost the same service rates for all classes, produces quantifying the inherent sense where the control policy affects just a little bit on the queue length.

1.1. Notation

Let consider discrete-time Markov decision processes where for such a system we consider \mathcal{Y} be a measurable general state space with respect to some given σ -field $\mathcal{B}(\mathcal{Y})$, a finite set of available actions \mathcal{A} at each time slot, and measurable cost function $z : \mathcal{Y} \times \mathcal{A} \rightarrow R$. At state $y \in \mathcal{Y}$, action $a \in \mathcal{A}$ is the cause of specified cost.

Let consider stochastic kernel p to evolve the state. The kernel is $p : \mathcal{Y} \times \mathcal{Y} \times \mathcal{A} \rightarrow [0, 1]$ and its definition is presented by

$$p(B, y, a) = Pr \left\{ Y_{t+1} \in B \mid Y_t = y, A_t = a \right\}, \forall y \in \mathcal{Y}, a \in \mathcal{A}, B \in \mathcal{B}(B).$$

In addition, $p(B, \cdot, \cdot) : \mathcal{Y} \times U \rightarrow [0, 1]$ is a measurable function for each $B \in \mathcal{B}(B)$. In the supposed system, we considered the performance of systems subject to the static state-feedback policies. We consider $\rho : \mathcal{Y} \rightarrow \mathcal{A}$ to be a measurable function that depends on the running system state, it chooses the action in each time slot. Let define a set of all measurable policies ρ by set $\mathcal{Q} = \{\rho : \mathcal{Y} \rightarrow \mathcal{A} \mid \rho \text{ is measurable}\}$.

The state evolution is a random process which is based on policy $\rho \in \mathcal{Q}$, is time-homogeneous Markov chain (Y_0, Y_1, \dots) . Stochastic kernel $p(B, y, \rho(y))$ specifies the transition probability for the Markov chain.

To convenience and to ease of use of notation we define $EC = \frac{1}{t} \sum_{k=0}^{t-1} E \left[c(Y_k, \rho(Y_k)) \mid Y_0 = y \right]$ and $Ec = \frac{1}{t} \sum_{k=0}^{t-1} E \left[c(Y_k) \mid Y_0 = y \right]$ and use Eg to show the expectation of $g(Y_{t+1})$ on condition $Y_t = y$. Now we consider the following performance function for a system under specific policy $\rho \in \mathcal{Q}$ in terms of ACPP

$$J(\rho) = \limsup_{t \rightarrow \infty} EC,$$

where we can define

$$J_{opt} = \inf_{\rho \in \mathcal{Q}} \left\{ \liminf_{t \rightarrow \infty} EC \right\}.$$

Finding and developing tools to determine policies which achieve ACPP $J(\rho)$ close to J_{opt} , is our goal.

2. Bounds of Markov Chain without Control

Y. Wang and S. Boyd in [8] showed how to compute performance bounds for finite-horizon stochastic control problems with linear system dynamics and arbitrary constraints, objective, noise distribution, and with asymmetric costs and constraint sets. Here we present a methodology to find an approach to determine bounds on ACPP. In this order and before presenting a lower bound on ACPP, first we try to find the bounds for Markov chain without control (or with a given state-feedback control) and then we develop this approach to present a lower bound on ACPP for any policy and then we investigate the

difference between the optimality and the given policy with using an upper bound and a lower bound on the cost by a given policy and by any policy, represently. Here we consider

Theorem 1. *Suppose $g : \mathcal{Y} \rightarrow R$ is a measurable function. We suppose $C(y) = c(y) + Eg - g(y)$ and define*

$$\alpha_u = \sup_{y \in \mathcal{Y}} \left\{ C(y) \right\}, \quad \alpha_l = \inf_{y \in \mathcal{Y}} \left\{ C(y) \right\}.$$

If there exists an $\varepsilon > 0$ such that

$$\sup_{y \in \mathcal{Y}} \left\{ E \left[|g(Y_{t+1})|^{1+\varepsilon} \mid Y_t = y \right] - |g(y)|^{1+\varepsilon} \right\} < \infty,$$

then for all $y \in \mathcal{Y}$,

$$\alpha_u \geq \limsup_{t \rightarrow \infty} Ec, \quad \alpha_l \leq \liminf_{t \rightarrow \infty} Ec.$$

This theorem is the main result of this paper and we use it to determine upper and lower bounds on the average cost incurred by Markov chains with general measurable state spaces. But before we prove Theorem 1, we express the existence of all expectations that we need.

Lemma 1. *Suppose $g : \mathcal{Y} \rightarrow R$ is a measurable function, consider*

$$\sup_{y \in \mathcal{Y}} C(y) < \infty, \quad \inf_{y \in \mathcal{Y}} C(y) > -\infty,$$

and there exists an $\varepsilon > 0$ such that

$$\sup_{y \in \mathcal{Y}} \left\{ E \left[|g(Y_{t+1})|^{1+\varepsilon} \mid Y_t = y \right] - |g(y)|^{1+\varepsilon} \right\} < \infty, \tag{1}$$

then for all $y \in \mathcal{Y}$

$$E \left[|g(Y_{t+1})|^{1+\varepsilon} \mid Y_0 = y \right] < \infty, \tag{2}$$

$$\left| E \left[c(Y_k) \mid Y_0 = y \right] \right| < \infty. \tag{3}$$

Proof. Suppose that there is M such that (1) equals to M , so the immediate result is

$$E \left[|g(Y_{t+1})|^{1+\varepsilon} \mid Y_t = y \right] \leq M + |g(y)|^{1+\varepsilon}, \quad \forall y \in \mathcal{Y}.$$

With clarity $E[|g(Y_0)|^{1+\varepsilon} | Y_0 = y] < \infty$. Also, using induction, if for some t , $E[|g(Y_t)|^{1+\varepsilon} | Y_0 = y] < \infty$, then

$$E \left[|g(Y_{t+1})|^{1+\varepsilon} \mid Y_0 = y \right] \leq M + E \left[|g(Y_t)|^{1+\varepsilon} \mid Y_0 = y \right] < \infty.$$

and as a result,

$$E \left[|g(Y_t)|^{1+\varepsilon} \mid Y_0 = y \right] < \infty.$$

Also from the fact that $g(y) \leq 1 + |g(y)|^{1+\varepsilon}$,

$$E \left[g(Y_t) \mid Y_0 = y \right] \leq 1 + E \left[|g(Y_t)|^{1+\varepsilon} \mid Y_0 = y \right] < \infty, \forall y \in \mathcal{Y}.$$

And also, since $-g(y) \geq 1 + |g(y)|^{1+\varepsilon}$,

$$-\infty < -1 - E \left[|g(Y_t)|^{1+\varepsilon} \mid Y_0 = y \right] \leq E \left[g(Y_t) \mid Y_0 = y \right], \forall y \in \mathcal{Y}.$$

Now if

$$\sup_{y \in \mathcal{Y}} \left\{ C(y) \right\} \leq \alpha_u < \infty,$$

then

$$c(y) \leq \alpha_u + g(y) - E \left[g(Y_{t+1}) \mid Y_t = y \right], \forall y \in \mathcal{Y},$$

where $|E[g(Y_{t+1})|Y_t = y]| = |E[g(Y_1)|Y_0 = y]| < \infty$. Accordingly,

$$E \left[c(Y_t) \mid Y_0 = y \right] < \infty, \forall t \text{ and } \forall y \in \mathcal{Y}. \tag{4}$$

In the similar way, if

$$\inf_{y \in \mathcal{Y}} \left\{ C(y) \right\} \geq \alpha_l > -\infty,$$

then

$$c(y) \geq \alpha_l + g(y) - E \left[g(Y_{t+1}) \mid Y_t = y \right], \forall y \in \mathcal{Y}.$$

Accordingly

$$-\infty < E \left[c(Y_t) \mid Y_0 = y \right], \forall t \text{ and } \forall y \in \mathcal{Y}. \tag{5}$$

Now it is clear that (3) is the result of combining of (4), (5). □

Now we can prove theorem 1.

Proof. [Proof of Theorem 1] By the definition of α_u we have

$$\alpha_u \geq \frac{1}{t} E \left[\sum_{k=0}^{t-1} C(Y_k) \mid Y_0 = y \right] = \frac{1}{t} \sum_{k=0}^{t-1} E \left[c(Y_k) \mid Y_0 = y \right] + \frac{1}{t} \left(E \left[g(Y_t) \mid Y_0 = y \right] - g(y) \right),$$

which means

$$\frac{1}{t} \sum_{k=0}^{t-1} E \left[c(y_k) \mid Y_0 = y \right] \leq \alpha_u + \frac{1}{t} \left(g(y) - E[g(Y_t) \mid Y_0 = y] \right).$$

In the similar way,

$$\frac{1}{t} \sum_{k=0}^{t-1} E \left[c(Y_k) \mid Y_0 = y \right] \geq \alpha_l + \frac{1}{t} \left(g(y) - E[g(Y_t) \mid Y_0 = y] \right).$$

Now if we show that $\lim_{t \rightarrow \infty} \frac{1}{t} E \left[g(Y_t) \mid Y_0 = y \right] = 0$, for all $y \in \mathcal{Y}$, then we have completed the proof.

Suppose that the supreme of $E \left[|g(Y_{t+1})|^{1+\varepsilon} \mid Y_t = y \right] - |g(y)|^{1+\varepsilon}$ for all $y \in \mathcal{Y}$ is equal M where $M < \infty$. So

$$\begin{aligned} M &\geq \frac{1}{t} E \left[\sum_{k=0}^{t-1} \left(E[|g(Y_{k+1})|^{1+\varepsilon} \mid Y_k] - |g(Y_k)|^{1+\varepsilon} \right) \mid Y_0 = y \right] = \\ &= \frac{1}{t} \left(E \left[|g(Y_t)|^{1+\varepsilon} \mid Y_0 = y \right] - |g(y)|^{1+\varepsilon} \right) \end{aligned} \quad (6)$$

with some algebra we have that.

Also we know that

$$\left(|g(y)|^{1+\varepsilon} + tM \right)^{\frac{1}{1+\varepsilon}} \leq |g(y)| + (tM)^{\frac{1}{1+\varepsilon}}, \quad (7)$$

therefore (6) and (7) imply that

$$E \left[|g(Y_t)| \mid Y_0 = y \right] \leq |g(y)| + (tM)^{\frac{1}{1+\varepsilon}}. \quad (8)$$

By taking lim sup as t approaches ∞ , from both sides of (8) we have

$$\limsup_{t \rightarrow \infty} \frac{1}{t} E \left[|g(Y_t)| \mid Y_0 = y \right] \leq \lim_{t \rightarrow \infty} \frac{1}{t} \left(|g(y)| + (tM)^{\frac{1}{1+\varepsilon}} \right) = 0,$$

and so

$$\lim_{t \rightarrow \infty} \frac{1}{t} E \left[g(Y_t) \mid Y_0 = y \right] = 0, \quad \forall y \in \mathcal{Y}.$$

□

3. Bounds with Control

We used Theorem 1 to provide bounds on ACP that acquired by Markov chains. These bounds are established in the general measurable state spaces of the Markov chains. In this section we extend the result to establish a lower bound on ACP acquired by any policy to bound the difference between $J(\rho)$ and J_{opt} for some specific ρ .

Lemma 2. *Suppose $h : \mathcal{Y} \rightarrow R$ is a measurable function, consider α_l be the infimum of the $\{c(Y, a) + E[g(Y_{t+1}) \mid Y_t = y, A_t = a] - g(y)\}$ for all $y \in \mathcal{Y}$ and $a \in \mathcal{A}$. And for any static state-feedback control policy $\rho : \mathcal{Y} \rightarrow \mathcal{A}$ such that*

$$\sup_{y \in \mathcal{Y}} \left\{ E[|g(Y_{t+1})|^{1+\varepsilon} \mid Y_t = y, A_t = \rho(y)] - |g(y)|^{1+\varepsilon} \right\} < \infty,$$

has ACP where satisfies

$$\alpha_l \leq \liminf_{t \rightarrow \infty} \frac{1}{t} \sum_{k=0}^{t-1} E[c(Y_k, \rho(Y_k)) \mid Y_0 = y], \quad \forall y \in \mathcal{Y}.$$

Proof. For any state feedback policy $\rho \in P$

$$\begin{aligned} \alpha_l &= \inf_{y \in \mathcal{Y}, a \in \mathcal{A}} \{c(y, a) + E[g(Y_{t+1})|Y_t = y, A_t = a] - g(y)\} \leq \\ &\leq \inf_{y \in \mathcal{Y}} \{c(y, \rho(y)) + E[g(Y_{t+1})|Y_t = y, A_t = \rho(y)] - g(y)\}. \end{aligned}$$

Now by applying Theorem 1 the proof is done. □

4. Control Policy of Multi-Class Queue

The considered queuing system is a discrete time model. Let us consider in the each time slot t , we have $W_t^i \in \{0, 1\}$ arrival jobs in class i to the queue. Also let's put restriction to have at most one arrival in each time slot (we can consider a small time interval to have it), which means $\sum_i W_t^i \leq 1$. Let consider $W_t = (W_t^1, \dots, W_t^N)$ denotes the arrival vector.

Also consider that any vectors W_t and $W_{t'}$ are i.i.d for any individual time slots, $t' \neq t$. We consider $\lambda_i = E[W_t^i]$, where it is independent from t . Let Y_t^i denote the number of jobs in class i at given time t , and consider $Y_t = (Y_t^1, \dots, Y_t^N)$ as a state vector. Also we characterize control sequence U_i in each time slot t such that

$$A_t^i = \begin{cases} 1, & \text{if a job of class } i \text{ is being serviced,} \\ 0, & \text{otherwise.} \end{cases}$$

A service, of a job of class i that is being served in a given time slot, will be completed with probability θ_i where this θ_i is independent from the service history. Also, let's define the number of departure jobs of class i in time slot t (those are served successfully) as random variable $D_t^i = A_t^i I(Y_t^i) B^i$. In this random variable, parameter B^i briefs a Bernoulli random variable where $E[B^i] = \theta_i$ and I is the indicator function,

$$I(y) = \begin{cases} 0, & \text{if } y = 0, \\ 1, & \text{otherwise.} \end{cases}$$

We have the form of the queue length dynamic according to $Y_{t+1}^i = Y_t^i + W_t^i - D_t^i$ for each $i \in \{1, \dots, N\}$.

This problem is a Markov decision process with ACPP criteria since we aimed to choose how to serve different job classes to minimize an average queue length. The state space, for the considered model, is $\mathcal{Y} = Z_+^N$, where the actions are chosen from the $\mathcal{A} = \left\{ a \in \{0, 1\}^N \mid \sum_{i=1}^N a_i = 1 \right\}$.

In each time slot t , the cost gives the number of jobs of all classes where it represents by $c(Y_t) = \sum_{i=1}^N Y_t^i$. Let $\rho : \mathcal{Y} \rightarrow \mathcal{A}$ be a policy, and consider sequence of (Y_0, Y_1, \dots) for the queue length process supporting by this policy. Then the average queue length supporting by policy ρ is

$$J(\rho) = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{k=0}^{t-1} E[c(Y_k) | Y_0 = 0].$$

The control policy affects on the average queue length where it is the problem that we want to consider. Stability is a measure that under, the bounded queue lengths exists for

any non-idling policy. We call a policy is non-idling if $y \neq 0$, then we have $y_i > 0$ when $\rho(y) = a_i$. Lemma 3 is for the cases that the system is not stabilized under a non-idling policy and since the result is standard (see, [16]), we omitted the proof.

Lemma 3. *On condition that $\sum_{i=1}^N \frac{\lambda_i}{\theta_i} > 1$, we do not have any non-idling policy with bounded average queue length.*

Theorem 2. *Consider that \mathcal{Q}_{NI} be the set of all policies that are non-idling one and let $\rho \in \mathcal{Q}_{NI}$ be an arbitrary one, and also let*

$$J_{opt} = \inf_{\rho \in \mathcal{Q}_{NI}} \left\{ \liminf_{t \rightarrow \infty} EC \right\}.$$

If $\sum_{i=1}^N \frac{\lambda_i}{\theta_i} < 1$, then $J(\rho) < \infty$ and we have

$$\frac{J(\rho)}{J_{opt}} \leq \frac{\max_i \{\theta_i\}}{\min_i \{\theta_i\}}.$$

Proof. In order to have a lower bound for J_{opt} , we define

$$g_l(y) = K_1 \left(\left(\sum_{i=1}^N \frac{y_i}{\theta_i} \right)^2 + k_2 \sum_{i=1}^N \frac{y_i}{\theta_i} \right),$$

where the coefficients K_1 and K_2 are

$$K_1 = \frac{\min_i \{\theta_i\}}{2 \left(1 - \sum_{i=1}^N \frac{\lambda_i}{\theta_i} \right)}, K_2 = \left(1 - 2 \sum_{i=1}^N \frac{\lambda_i}{\theta_i} \right).$$

Let $\Delta_l(y, a) = E[g_l(Y_{t+1}) | Y_t = y, A_t = a] - g_l(y)$. For all $y \neq 0$, action a which minimizes $z(y) + \Delta_l(y, a)$ has

$$E \left[\sum_{i=1}^N \frac{D_t^i}{\theta_i} | Y_t = y, A_t = a \right] = 1.$$

Therefore,

$$\min_{a \in \mathcal{A}} \{z(y) + \Delta_l(y, a)\} = \sum_{i=1}^N \left(1 - \frac{\min_j \{\theta_j\}}{\theta_i} \right) y_i + \alpha \min_j \{\theta_j\}.$$

In order to have an upper bound on J_ρ , we use following function,

$$g_u(y) = K_3 \left(\left(\sum_{i=1}^N \frac{y_i}{\theta_i} \right)^2 + k_2 \sum_{i=1}^N \frac{y_i}{\theta_i} \right),$$

where constant K_3 is

$$K_3 = \frac{\max_i \{\theta_i\}}{2 \left(1 - \sum_{i=1}^N \frac{\lambda_i}{\theta_i}\right)}.$$

Let consider $\Delta_u(y) = E[g_u(Y_{t+1})|Y_t = y] - g_u(y)$. With using the dynamic of queue length $Y_{t+1}^i = Y_t^i + W_t^i - D_t^i$ we have

$$z(y) + \Delta_u(y) = \sum_{i=1}^N \left(1 - \frac{\max_j \{\theta_j\}}{\theta_i}\right) y_i + \alpha \max_j \{\theta_j\},$$

where

$$\alpha = \frac{\sum_{i=1}^N \left(1 + \frac{1}{\theta_i}\right) \frac{\lambda_i}{\theta_i} - 2 \left(\sum_{i=1}^N \frac{\lambda_i}{\theta_i}\right)^2}{2 \left(1 - \sum_{i=1}^N \frac{\lambda_i}{\theta_i}\right)}.$$

To finish our job in this proof, we need to show that for any policy $\rho \in Q_{NI}$ which is non-idling we have,

$$\sup_{y \in Y} \{E[g_u(Y_{t+1})^2|Y_t = y, A_t = \rho(y)] - g_u(y)^2\} < \infty,$$

and

$$\sup_{y \in Y} \{E[g_l(Y_{t+1})^2|Y_t = y, A_t = \rho(y)] - g_l(y)^2\} < \infty.$$

Functions g_u and g_l have the following form

$$g(x) = K(x^2 + K_2x),$$

where in the above equation, variable $x = \sum_{i=1}^N \frac{y_i}{\theta_i}$. Now by squaring g we have

$$g(x)^2 = K^2(x^4 + 2k_2x^3 + k_2^2x^2).$$

The expected drift for any non-idling policy, $E[g(Y_{t+1})^2|Y_t = y, A_t = \rho(y)] - g(y)^2$ is a polynomial of degree third where its third order term is equal to

$$4k^2 \left(\sum_{i=1}^N \frac{\lambda_i}{\theta_i} - 1\right) \left(\sum_{i=1}^N \frac{y_i}{\theta_i}\right)^3.$$

It is obviously negative for all $y \neq 0$ when the system is stable, where it means, the expected drift in g^2 has upper bound for all policies in Q_{NI} . Hence, bounds α_u and α_l are well-founded and satisfy

$$\frac{J(\rho)}{J_{opt}} \leq \frac{\alpha_u}{\alpha_l} = \frac{\max_i \{\theta_i\}}{\min_i \{\theta_i\}}.$$

□

The point of this job is that we did not consider simply the maximum and minimum service rates of queues to find the lower and upper bound of the average queue length of the multi-class queue, respectively. Although these type approaches, that consider the minimum and maximum service rates of queues give the bounds but they can be made a large difference between these bounds for the given service rates. In fact, the bounds by provided in Theorem 2 is tighter than the bound obtaining from considering minimum and maximum service rate.

Conclusion

In this paper, we considered a queuing problem which is Markov decision process in the general state space and we described the method for computing bounds on the costs in such processes with average cost per period. Our method naturally yields the factor that can be used at the problem of controlling of a multi-class queue to find a bound on. The bound that we found is a relation between the average queue length acquired by any policies and acquired by an optimal policy where this bound is totally different with the bound that is obtained by applying minimum and maximum service rate in queues which serve multi-class jobs.

References

1. Atar R., Mandelbaum A., Reiman M.I. Scheduling a Multi-Class Queue with Many Exponential Servers: Asymptotic Optimality in Heavy Traffic. *The Annals of Applied Probability*, 2004, vol. 14, no. 3, pp. 1084–1134. DOI: 10.1214/10505160400000233
2. Regan K., Boutilier C. Robust Policy Computation in Reward-Uncertain MDPs Using Nondominated Policies. *Twenty-Fourth AAAI Conference on Artificial Intelligence*, Atlanta, July, 2010, pp. 1127–1133.
3. Kebarighotbi A., Cassandras C.G. Optimal Scheduling of Parallel Queues with Stochastic Flow Models: The $c\mu$ -rule Revisited. *IFAC Proceedings Volumes*, 2011, no. 44, pp. 8223–8228.
4. Shanthikumar J., Yao D. Multiclass Queuing Systems: Polymatroidal Structure and Optimal Scheduling Control. *Operations Research*, 1992, vol. 40, no. 2, pp. 293–299. DOI: 10.1287/opre.40.3.S293
5. Meyn S., Tweedie R. *Markov Chains and Stochastic Stability*. London, Springer, 1993. DOI: 10.1007/978-1-4471-3267-7
6. Puterman M.L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New Jersey, John Wiley and Sons, 2009.
7. Schweitzer P.J., Seidmann A. Generalized Polynomial Approximations in Markovian Decision Processes. *Journal of Mathematical Analysis and Applications*, 1985, vol. 110, pp. 568–582. DOI: 10.1016/0022-247X(85)90317-8
8. Yang Wang, Boyd S. Performance Bounds and Sub-Optimal Policies for Linear Stochastic Control via LMIs. *International Journal of Robust and Nonlinear Control*, 2011, no. 21, pp. 1710–1728. DOI: 10.1002/rnc.1665
9. Osipova N., Ayesta U., Avrachenkov K. Optimal Policy for Multi-Class Scheduling in a Single Server Queue. *21st International Teletraffic Congress*, 2009, p. 10951139.
10. Jia Li, Zhang H.M. Bounding Queuing System Performance with Variational Theory. *Transportation Research Procedia*, 2015, no. 7, pp. 519–535.
11. Senderovich A., Weidlich M., Gal A., Mandelbaum A. Queue Mining for Delay Prediction in Multi-Class Service Processes. *Information Systems*, 2015, no. 53, pp. 278–295. DOI: 10.1016/j.is.2015.03.010

12. Huang Qing, Chakravarthy S.R. Analytical and Simulation Modeling of a Multi-Server Queue with Markovian Arrivals and Priority Services. *Simulation Modelling Practice and Theory*, 2012, no. 28, pp. 12–26. DOI: 10.1016/j.simpat.2012.05.010
13. Casale G., Sansottera A., Cremonesi P. Compact Markov-Modulated Models for Multiclass Trace Fitting. *European Journal of Operational Research*, 2016, vol. 255, no. 3, pp. 822–833. DOI: 10.1016/j.ejor.2016.06.005
14. Lefeber E., Lammer S., Rooda J.E. Optimal Control of a Deterministic Multiclass Queuing System For Which Several Queues Can Be Served Simultaneously. *Systems and Control Letters*, 2011, vol. 60, no. 7, pp. 524–529. DOI: 10.1016/j.sysconle.2011.04.010
15. Walraevens J., Bruneel H., Fiems D., Wittevrongel S. Delay Analysis of Multiclass Queues with Correlated Train Arrivals and a Hybrid Priority/Fifo Scheduling Discipline. *Applied Mathematical Modelling*, 2017, no. 45, pp. 823–839. DOI: 10.1016/j.apm.2017.01.044
16. Kleinrock L. *Queueing Systems. Volume II: Computer Applications*. New Jersey, John Wiley and Sons, 1976.
17. Ching-Tarng Hsieh, Lam S.S. Two Classes of Performance Bounds for Closed Queueing Networks. *Performance Evaluation*, 1987, vol. 7, no. 1, pp. 3–30. DOI: 10.1016/0166-5316(87)90054-X
18. Koukopoulos D., Mavronicolas M., Spirakis P. Performance and Stability Bounds for Dynamic Networks. *Journal of Parallel and Distributed Computing*, 2007, vol. 67, no. 4, pp. 386–399. DOI: 10.1016/j.jpdc.2006.11.005

Received June 28, 2018

УДК 519.872**DOI: 10.14529/mmp190104****ГРАНИЦЫ ПРОИЗВОДИТЕЛЬНОСТИ И СУБОПТИМАЛЬНЫЕ СТРАТЕГИИ ДЛЯ МНОГОКЛАССОВОЙ ОЧЕРЕДИ**

A. Madankan, Университет Заболы, г. Забола, Иран, Amadankan@uoz.ac.ir

В этой статье рассматривается общий класс системы массового обслуживания с несколькими типами заданий и гибкими возможностями обслуживания. Используется стохастическая стратегия управления для определения потери производительности в многоклассовой очереди М/М/1. Рассматриваемая система изначально представляется собой марковский процесс принятия решений. В работе показано, как рассчитать границы производительности для стратегии стохастического управления марковского процесса принятия решения с критериями средней стоимости. На практике многие исследователи использовали эвристические стратегии управления из-за некоторой сложности в вычислениях и использовании математически оптимальных стратегий. Цель данной работы заключается в расчете разницы между оптимальной и конкретной стратегий, а также в нахождении границы производительности для оптимальной стратегии. Другими словами, это исследование показывает, что оптимальные границы средней длины очереди для любых стратегий без простоя можно найти с помощью коэффициента скорости обслуживания.

Ключевые слова: система массового обслуживания; многоклассовые задачи; стратегия стохастического контроля.

Али Маданкан, кафедра «Информатика», Университет Заболы (г. Забола, Иран), Amadankan@uoz.ac.ir.

Поступила в редакцию 28 июня 2018 г.