OPTIMAL SERVICE POLICY IN MULTI-QUEUING SYSTEM

Tadayoshi SHIOYAMA

Department of Mechanical and System Engineering
Kyoto Institute of Technology
Matsugasaki, Sakyo-ku
Kyoto 606, Japan

This paper deals with a multi-queuing system with a single-server providing service for customers of K classes. An arriving customer enters the buffer for his class. A reward is received at the completion of service. A holding cost is incurred by a customer in the system. All rewards and holding costs are assumed to be continuously discounted. The problem is to decide the type of class of customer to be served next in order to attain the minimal expected present value of total costs incurred over an infinite planning horizon. The problem is formulated as the discounted semi-Markov decision process. The main purpose of the paper is to find the properties of the optimal service policy. The monotonicities of the optimal policy are shown.

1. INTRODUCTION

We deal with a multi-queuing system with a single-server providing service for customers of K classes. The queuing system studied here is motivated by a need to model the entrance node onto the digital network [1] providing transmission services for customer premises as a set of dissimilar data sources (data terminal, telephone set, video set, facsimile set, etc). Given the expected increasing need to transmit simultaneously a variety of data types on one channel, it becomes extremely important to efficiently provide transmission services for these traffic types onto the shared channel.

Customers of class k arrive according to a Poisson process with mean arrival rate \( \lambda_k \), \( k=1,...,K \). These arrival processes are assumed to be independent. An arriving customer enters into the infinite buffer for his class. Service times of customers of class k, \( k=1,...,K \), are independent and identically distributed as a random variable \( S_k \) with distribution function \( F_k(\cdot) \) and Laplace-Stieltjes transform (LST) \( \Psi_k(\phi) = \exp(-\phi S_k) \).

\[
\Psi_k(\phi) = \mathbb{E}[\exp(-\phi S_k)] = \int_0^{\infty} \exp(-\phi t) dF_k(t), \quad \text{Re}(\phi) > 0.
\]

We suppose that a reward of \( T_k \) is received at the completion of class k service, \( k=1,...,K \). Different rewards represent different importance or urgency of each class of customers. We also suppose that a holding cost is incurred for time that a customer resides within the system. All rewards and costs are assumed to be continuously discounted with interest rate \( \beta > 0 \). That is, one dollar received at time \( t > 0 \) has a present value of \( \exp(-\beta t) \). The problem is to decide the type of class of customer to be served next in order to attain the minimal expected present value of total costs incurred over an infinite planning horizon. The problem is formulated as the discounted semi-Markov decision process. Harrison [2,3] has dealt with a multi-queuing system with infinite buffers and studied the priority rule that maximizes the expected rewards. The main purpose of the present paper is to find the properties of the optimal service policy in the multi-queuing system. The monotonicities of the optimal policy are shown.

2. OPTIMALITY EQUATION

Since each arrival process is Poisson, the discrete-parameter process formed by observing a state of the system only at epochs of service completion is an imbedded Markov chain. Depending on the observed state, the type of class to be served next is decided. Let \( n=(n_1,...,n_K) \) denote the observed state where \( n_k, \ k=1,...,K \), represents the number of class k customers in the system (waiting or being served). We define \( q(n,k,\phi) \) as the immediate expected holding cost during the transition time (i.e., the service time). Denote by \( V(n) \) the expected present value of costs incurred over the infinite planning horizon, given that the initial state is \( n \). We assume that the server can not be idle in state \( n=0 \). Then we have the following optimality equation: for \( n \neq 0 \)

\[
V(n) = V(n_1,...,n_K) = \min_k \{-T_k \Psi_k(\phi) + q(n,k,\phi) \}
\]

\[
+ \int_0^{\infty} \exp(-\beta t) E[V(n-e^+A) | S_+ = t] dF_k(t), \quad (2)
\]

where \( A \) is a random variable representing the service time.
where $e_k^x$ denotes the $k$-th unit vector and 
$A = (A_1, \ldots, A_K)$, $A_k$ representing the number of 
class $k$ customers arriving during service time 
$S_k$. We assume that $F_k(\cdot) = F(\cdot)$ and $\Psi_k(\beta) = \Psi(\beta)$. 
$k=1, \ldots, K$. For simplicity, we use $q(n)$ 
instead of $q(n,k,\beta)$.

For notational convenience, we define the 
following:

$$V(n) = \sum_{n'} \exp(-\beta t) \sum_{t=0}^\infty \sum_{\epsilon_k} E[V(n+A)|S_k=t]dF(t)$$ (3)

and

$$V_{\epsilon_k}(n) = \sum_{n'} \exp(-\beta t) \sum_{t=0}^\infty \sum_{\epsilon_k} E[V(n+A)|S_k=t]dF(t)$$ (4)

Then, the optimality equation (2) is rewritten 
by

$$V(n) = \min_{\epsilon_k} [q(n) + V_{\epsilon_k}(n)]$$ (5)

For $n=0$, if the next observation point is the 
epoch of the first arrival, then we have

$$V(0) = \sum_{n'} \sum_{t=0}^\infty \sum_{\epsilon_k} \exp(-\beta t) \sum_{t=0}^\infty \sum_{\epsilon_k} Q(n', dt)$$

where $Q(n', t|0)$ denotes the probability that, 
given that the state 0 is observed, the next 
transition is to state $n'$ and takes no longer 
than $t$ time units, and is given by

$$Q(n', t|0) = (\lambda_1/\lambda) \{1 - \exp(-\beta t)\}$$

if $n'=e^i, i=1, \ldots, K$,

$$=0, \quad \text{otherwise}.$$

$x = \lambda_1 + \lambda_2 + \ldots + \lambda_K$, 
and 
$\lambda_1$ is the mean arrival rate for class 1.

We solve the optimality equation by the 
following iteration scheme:

$$V(n) = V(n)$$ (6)

for $n=0$,

$$U(n) = \sum_{n'} \sum_{t=0}^\infty \sum_{\epsilon_k} \exp(-\beta t) \sum_{t=0}^\infty \sum_{\epsilon_k} V(n')Q(n', dt|0).$$ (7)

$$U_{\epsilon_k}(n) = \sum_{n'} \sum_{t=0}^\infty \sum_{\epsilon_k} \exp(-\beta t) \sum_{t=0}^\infty \sum_{\epsilon_k} F_{\epsilon_k}(n')Q(n', dt|0).$$ (8)

$$V(n) = \min_{\epsilon_k} [q(n) + U_{\epsilon_k}(n)], \quad \text{and}$$ (9)

$$V(n) = \sum_{n'} \sum_{t=0}^\infty \sum_{\epsilon_k} \exp(-\beta t) \sum_{t=0}^\infty \sum_{\epsilon_k} V(n')Q(n', dt|0).$$ (10)

The iteration scheme converges to the optimal 
solution [4].

3. MONOTONICITY OF OPTIMAL POLICY

Denote by $I_k$ the set $\{n; a(n) = k\}$, where $a(n)$ is 
the optimal action (decision) in state $n$. The set $I_k$, $k=1, \ldots, K$, is called an increasing set 
when it satisfies the condition that if $n \in I_k$, 
then $n+e_k \in I_k$. An optimal policy is called 
monotone if the set $I_k$ is an increasing set 
under the optimal policy. In the following, 
the sufficient condition for the set $I_k$ to be 
an increasing set, is shown. Denote by $I_k(n)$, 
k=1, \ldots, K, the set $I_k$ at the $n$-th iteration.

**Lemma 1.** If $U^{(n)}(n)$ satisfies the following condition: for $k, k'=1, \ldots, K$ and $k' \leq k$,

$$U^{(n)}(n-e_k' + e_k) - U^{(n)}(n-e_k'),$$ (10)

then $I_k(n)$ is an increasing set.

**Proof.** It suffices to show that if $U^{(n)}(n) - U^{(n)}(n-k)$, $k'=k$, then $U^{(n)}(n-k') - U^{(n)}(n-k)$. 
Since it holds that

$$U^{(n)}(n-k') - U^{(n)}(n-k) = (T_{k'} + T_k)\Psi(n-k) - U^{(n)}(n-k'),$$

we obtain using the condition (10) the 
following relation.

$$2(T_{k'} + T_k)\Psi(n-k) - U^{(n)}(n-k) = 2U^{(n)}(n-k') - U^{(n)}(n-k') - U^{(n)}(n-k) \geq 0.$$

The proof is completed.

In order to find the condition for $I_k(n)$ to 
be an increasing set for any $n$, we introduce 
the condition:

$$V^{(n)}(n) - 2e_k' - e_k - e_k = U^{(n)}(n-k) - U^{(n)}(n-k') - U^{(n)}(n-k) \geq 0.$$ (11)

In the following we assume that the number $K$ 
of classes is two. Then, we have the following 
lemmas concerning the properties of 
$\min_{\epsilon_k} U^{(n)}(n)$.

**Lemma 2.** If $U^{(n)}(n)$ satisfies the conditions 
(10) and (11), then $\min_{\epsilon_k} U^{(n)}(n)$ satisfies the condition (10).

**Proof.**

(a) In the case of $\min_{\epsilon_k} U^{(n)}(n-e_k') = U^{(n)}(n-e_k')$ and $\min_{\epsilon_k} U^{(n)}(n-e_k' + e_k') = U^{(n)}(n-e_k' + e_k')$, 
for $k'=k$, we have from (10) the following.

$$\min_{\epsilon_k} U^{(n)}(n-e_k' + e_k') - \min_{\epsilon_k} U^{(n)}(n-e_k') = U^{(n)}(n-e_k' + e_k') - U^{(n)}(n-e_k') \geq 0.$$

(b) In the case of $\min_{\epsilon_k} U^{(n)}(n-e_k') = U^{(n)}(n-e_k')$ and $\min_{\epsilon_k} U^{(n)}(n-e_k' + e_k') = U^{(n)}(n-e_k' + e_k')$, 
we have the following by the way similar to (a)
In the case of \( \min U_a(n)(n-e^k) = U_k c(n)(n-e^k) \) and \( \min U_a(n)(n-e^k) = U_k c(n)(n-e^k) \), we have from (10) and (11) the following relation.

\[
\begin{align*}
\min U_a(n)(n-e^k) &= U_k c(n)(n-e^k) \\
\min U_a(n)(n-e^k) &= U_k c(n)(n-e^k) \\
\min U_a(n)(n-e^k) &= U_k c(n)(n-e^k) \\
\min U_a(n)(n-e^k) &= U_k c(n)(n-e^k) \\
\end{align*}
\]

Hence, \( \min U_a(n)(n) \) satisfies the condition (10) for all cases. The proof is completed.

Lemma 3. If \( U^{(n)}(n) \) satisfies the conditions (10) and (11), then \( \min U_a(n)(n) \) satisfies the condition (11).

Proof.

(a) In the case of \( \min U_a(n)(n-e^k) = U_k c(n)(n-e^k) \) and \( \min U_a(n)(n-2e^k) = U_k c(n)(n-2e^k) \),

\[
\begin{align*}
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\end{align*}
\]

(b) In the case of \( \min U_a(n)(n-e^k) = U_k c(n)(n-e^k) \) and \( \min U_a(n)(n-2e^k) = U_k c(n)(n-2e^k) \),

\[
\begin{align*}
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\end{align*}
\]

(c) In the case of \( \min U_a(n)(n-e^k) = U_k c(n)(n-e^k) \) and \( \min U_a(n)(n-2e^k) = U_k c(n)(n-2e^k) \),

\[
\begin{align*}
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\min U_a(n)(n-2e^k) &= U_k c(n)(n-2e^k) \\
\end{align*}
\]

The proof is completed.

The equation (7) is rewritten by

\[
U^{(n)}(n) = \sum \exp(-\beta t) E[V^{(n-1)}(n+i) | S_i(t)] dF(t)
\]

where \( P(A) \) denotes the probability that \( A_i, i=1,\ldots,K \), customers of class \( i \) arrive during the service time \( t \). Hence, we have the following.

\[
\begin{align*}
U^{(n)}(n-e^k) &= U^{(n)}(n)(n-e^k) \\
U^{(n)}(n-e^k) &= U^{(n)}(n)(n-e^k) \\
U^{(n)}(n-e^k) &= U^{(n)}(n)(n-e^k) \\
U^{(n)}(n-e^k) &= U^{(n)}(n)(n-e^k) \\
\end{align*}
\]
\(+V^{(n-1)}(n-e^kA)\)dF(t)

\[=\int \exp(-\beta t) P(A)(q(n-e^k\cdot e^k+A)-q(n+A))
\]

\[-q(n-e^k\cdot A)+q(n-e^kA)
\]

\[+\min_u U_u^{(n-1)}(n-e^k\cdot e^k+A)-\min_u U_u^{(n-1)}(n+A)
\]

\[-\min_u U_u^{(n-1)}(n-e^k\cdot A)
\]

\[+\min_u U_u^{(n-1)}(n-e^kA)\]dF(t) \hfill (13)

From (7) through (9), (12) through (14) and Lemmas 2 and 3, we obtain the following lemma concerning \(U_u(n)\) and \(V_v(n)\).

**Lemma 4.** Suppose \(q(n)\) satisfies the conditions (10) and (11). If \(V^{(n-1)}(n)\) or \(U^{(n-1)}(n)\) satisfies the conditions (10) and (11), then \(V^{(n)}(n)\) and \(U^{(n)}(n)\) also satisfy the conditions (10) and (11). 

From iteration scheme (6) through (9) and Lemmas 1 and 4, we obtain the following proposition concerning the monotone property of the optimal policy.

**Proposition 1.** Suppose \(q(n)\) satisfies the conditions (10) and (11). For any \(n\), \(V^{(n)}(n)\) and \(U^{(n)}(n)\) satisfy the conditions (10) and (11). Therefore, for any \(n\), \(I_k^{(n)}\), \(k=1,\ldots,K\), is an increasing set and \(\lim I_k^{(n)}=I_k\), \(k=1,\ldots,K\), is also an increasing set.

Proposition 1 implies that when the immediate expected holding cost \(q(n)\) satisfies the conditions (10) and (11), if the optimal action \(a(n)\) is \(k\) (i.e., the type \(k\) is optimally decided to be served next) in state \(n\), then the action \(a(n+e^k)-k\) is also optimal in state \(n+e^k\).

**4. Conclusions**

We have discussed the monotone property of the optimal service policy in a multi-queuing system. For simplicity, we have proved the monotonicity for the case where the number \(K\) of classes is two. It is conjectured that the optimal policy has the monotone property for the case of any \(K\). In the large scale system, it is difficult in practical to obtain the exact optimal solution for the Markov decision process because of time consuming. The heuristic solution becomes practical for the large scale system. The monotone property of the optimal policy is considered useful to find the heuristic solution because using the monotonicity, the problem of finding the optimal policy is reduced to the simple one in which it is sufficient to find only the boundaries of the increasing sets \(I_k\), \(k=1,\ldots,K\) (control limits).

**References**


