<table>
<thead>
<tr>
<th>Title</th>
<th>Further Results on the N-Policy for the M/G/l Queue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>KIMURA, Toshikazu</td>
</tr>
<tr>
<td>Citation</td>
<td>HOKUDAI ECONOMIC PAPERS, 17: 71-73</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1987</td>
</tr>
<tr>
<td>Doc URL</td>
<td><a href="http://hdl.handle.net/2115/30737">http://hdl.handle.net/2115/30737</a></td>
</tr>
<tr>
<td>Type</td>
<td>bulletin</td>
</tr>
<tr>
<td>File Information</td>
<td>17_P71-73.pdf</td>
</tr>
</tbody>
</table>
Further Results on the N-Policy for the M/G/1 Queue

Toshikazu KIMURA
Associate Professor
Faculty of Economics
Hokkaido University

ABSTRACT

This note deals with the N-policy for the M/G/1 queue with the following linear cost structure: costs per unit time for keeping the server on or off, fixed costs for turning the server on or off, and a holding cost per customer in the system per unit time.

The N-policy is defined as the control policy which turns the server on when N or more customers are present and off when the system is empty. The average cost rate is used as a criterion for optimality. Some sufficient conditions under which the optimal policy falls into specific forms are provided.

This note deals with the N-policy for operating the M/G/1 queueing system with a linear cost structure. The average cost rate is used as a criterion for optimality. Some sufficient conditions under which the optimal operating policy falls into specific forms are provided.

The M/G/1 system is specified by the following assumptions and notation.

Customers individually arrive at the system according to a Poisson process at rate \( \lambda (>0) \). Their service times are nonnegative i.i.d. random variable with mean \( 1/\mu \) and finite variance \( \sigma^2 \). Let \( \rho \equiv \lambda/\mu \) be the traffic intensity and assume \( \rho < 1 \).

The N-policy is defined as the control policy which turns the server on whenever \( N \) or more customers are present and off only when the system is empty. We assume the same linear cost structure as in Heyman [2] and Bell [1]. That is, costs for keeping the server off (\( R_1 \)) or on (\( R_2 \)) are incurred per unit time, where it is assumed that \( r_1 \geq r_2 \). Fixed costs for turning the server on (\( R_1 \)) and turning him off (\( R_2 \)) are incurred. These four costs are assumed to be nonnegative and finite. Moreover, there is a cost per unit time as a penalty for the delay of customers. This cost (\( h \)) is proportional to the number of customers in the system and has finite positive value.

For the M/G/1 system under the N-policy, Heyman [2] showed that the average cost rate over an infinite horizon is given by

\[
C(0) = r_2 + hL(1),
\]

\[
C(N) = r_1 + \rho (r_2 - r_1) + hL(N) + \frac{2}{N} (1 - \rho) (R_1 + R_2), \quad \text{for } N \geq 1,
\]

where

\[
L(N) = \frac{1}{2} (N - 1) + \rho + \frac{\rho^2 + \lambda^2 \sigma^2}{2(1 - \rho)}, \quad N \geq 1,
\]

which denotes the mean number of customers in the M/G/1 system under the N-policy [3]. For \( N = 0 \), the corresponding 0-policy is defined as the policy which always
keeps the server on. A policy is called optimal if it minimizes the average cost rate $C(\ast)$. From (1) and (2), Heyman also showed that the optimal value of $N( = N^\ast)$ can be obtained by

$$C(N^\ast) = \min \{ C(0), C([\tilde{N}]), C([\tilde{N}]) \}$$

where

$$\tilde{N} = \sqrt{\frac{2\lambda(1 - \rho)(R_1 + R_2)}{h}}.$$  \hspace{1cm} (5)

In the following theorems, we shall investigate several effects of the costs and the traffic to the optimal policy.

Theorem 1. If $r_1 = r_2$, then the optimal policy is the 0–policy.

Proof: The proof is executed by distinguishing the following two cases ;

Case 1. $\tilde{N} \leq 1/2$.

For this case, it is clear that $\min_{N \geq 1} C(N) = C(1)$. Hence,

$$\min_{N \geq 1} C(N) - C(0) = C(1) - C(0) = \lambda (1 - \rho) (R_1 + R_2) > 0.$$  \hspace{1cm} (4)

That is, the optimal policy is the 0–policy.

Case 2. $\tilde{N} > 1/2$.

Since $C(\tilde{N})$ must be less than or equal to $\min_{N \geq 1} C(N)$, it follows from (1), (2), (3) and (5) that

$$\min_{N \geq 1} C(N) - C(0) \leq C(\tilde{N}) - C(0)$$

$$= h\{ L(\tilde{N}) - L(1) \} + \frac{1}{2} h\tilde{N}$$

$$= h(\tilde{N} - \frac{1}{2}) > 0.$$  \hspace{1cm} (4)

That is, the optimal policy is the 0–policy.\]

From the result of Theorem 1, suppose hereafter that $r_2 > r_1$.

Theorem 2. Suppose that $R_1 + R_2 > 0$.

(i) If $r_2 - r_1 < \mu(R_1 + R_2)$, then there exists a unique $\lambda^\ast \in (0, \mu)$ such that for any $\lambda \in [\lambda^\ast, \mu)$ the optimal policy is the 0–policy.

(ii) Otherwise there exists a unique $\lambda^\ast \in (0, \mu)$ such that for any $\lambda \in [\lambda^\ast, \mu)$ the optimal policy is the 1–policy.

Proof : From (5), we have

$$\tilde{N}^2 = \frac{2(R_1 + R_2)}{\mu h} \lambda(\mu - \lambda).$$  \hspace{1cm} (6)

That is, $\tilde{N}^2$ can be regarded as a quadratic function of $\lambda$. Therefore $\tilde{N}^2$ is
Further Results on the N-Policy for the M/G/1 Queue

monotonically decreasing for \( \lambda > \mu/2 \) and gets less than unity as \( \lambda \) increases, i.e.,
\[
2(R_1 + R_1) \lambda (\mu - \lambda) / \mu h \leq 1.
\]
This inequality can be deduced to
\[
(\lambda - \mu/2)^2 \geq \frac{\mu^2}{4} \left(1 - \frac{2h}{\mu(R_1 + R_2)}\right).
\]

Hence, if \( \mu(R_1 + R_2) \leq 2h \), then (7) always holds, that is, \( \bar{N} \leq 1 \) for any \( \lambda \in [0, \mu) \). On the other hand, if \( \mu(R_1 + R_2) > 2h \), then \( \bar{N} \leq 1 \) for \( \lambda \geq \lambda_1 \) with
\[
\lambda_1 = \frac{\mu}{2} \left(1 + \sqrt{1 - \frac{2h}{\mu(R_1 + R_2)}}\right).
\]

It is clear that \( \lambda_1 \in (\mu/2, \mu) \). Consequently, a sufficient condition for \( \bar{N} \leq 1 \), which is independent of the costs, is \( \lambda \geq \lambda_1 \). Assume here that \( \lambda \) is sufficiently large so that \( \bar{N} \leq 1 \). Then,
\[
\min_{N \geq 1} C(N) - C(0) = C(1) - C(0)
= (1 - \rho)(\lambda (R_1 + R_2) - (r_2 - r_1)).
\]

Hence, if \( \lambda \geq \lambda_2 \equiv (r_2 - r_1) / (R_1 + R_2) \) and \( 0 < \lambda_2 < \mu \), it follows that \( \min_{N \geq 1} C(N) \geq C(0) \) and the optimal policy is the 0-policy. In other words, if \( r_2 - r_1 < \mu (R_1 + R_2) \), then the optimal policy is the 0-policy for \( \lambda \geq \max(\lambda_1, \lambda_2) \in (0, \mu) \). For the case that \( \lambda_2 \geq \mu \), i.e., \( r_2 - r_1 \geq \mu (R_1 + R_2) \), it follows for any \( \lambda \) with \( \bar{N} \leq 1 \) that
\[
\min_{N \geq 1} C(N) = C(1) < C(0),
\]
and so that the optimal policy is the 1-policy for \( \lambda \) satisfying \( \bar{N} \leq 1 \). Such \( \lambda \) is given, for example, by \( \lambda \in [\lambda_1, \mu) \). Thus the proof is completed. \( \square \)

Corollary 2.1. If \( r_2 - r_1 < \mu(R_1 + R_2) \leq 2h \), then the optimal policy is

the 1-policy for \( 0 < \lambda < (r_2 - r_1) / (R_1 + R_2) \), and

the 0-policy for \( (r_2 - r_1) / (R_1 + R_2) \leq \lambda < \mu \).

Corollary 2.2. If \( r_2 - r_1 \geq \mu(R_1 + R_2) \) and \( 2h \geq \mu(R_1 + R_1) \), then the optimal policy is the 1-policy.

The proofs of these corollaries obviously follow that of Theorem 2, and hence are omitted here.

Theorem 2 and its corollaries imply that the optimal policy eventually falls into the 0- or 1-policy as the traffic becomes heavy.

References