East Asian Math. J. **24** (2008), No. 1, pp. 45–55

OPTIMAL CONTROL OF A QUEUEING SYSTEM WITH P_λ^M-SERVICE POLICY

Sunggon Kim and Jongho Bae

Abstract. We consider an $M/G/1$ queue with P_λ^M-service policy, which is a two-stage service policy. The server starts to serve with rate 1 if a job arrives to the sever in idle state. If the workload of the system upcrosses λ, then the service rate is changed to M and this rate continues until the system is empty. It costs to change the service rate to M and maintaining the rate. When the expectation of the stationary workload is supposed to be less than a given value, we derive the optimal value of M.

1. Introduction

In this paper, an $M/G/1$ queueing system with P_λ^M- service policy is considered. The queueing system with the policy serves jobs with two different service rates. Server is initially idle and starts to work with service rate 1 when a job arrives. The jobs arrive according to a Poisson process of rate $\nu > 0$ and the amount of work of jobs are independent and identically distributed with distribution function G. The server increases his/her service rate to $M \geq 1$ instantaneously, if the workload(often called virtual waiting time) exceeds threshold $\lambda > 0$, and continues to serve at rate M until the system becomes idle. The server restarts to work with service rate 1 when another job arrives. For the generality of the policy, we consider the case of $M = 1$. In this case, the queueing process with the policy is identical to the ordinary $M/G/1$ queueing system.

P_λ^M- service policy has been proposed by Bae et al. [2]. It is a two-stage service policy, which is an application of the well-known P_λ^M-releasing policy [5, 8, 13] in dam theory to the $M/G/1$ queueing system. In contrast with the present model, the dam with P_λ^M-releasing policy does not release the water until the level of water upcrosses the threshold λ. Using the level crossing

Received August 23, 2007.

2000 *Mathematics Subject Classification.* 60K25; 90B05.

Key words and phrases. Two-stage service policy, P_λ^M-service policy, queueing system, optimal control.

This work was supported by grant No. R01-2006-000-10906-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

technique [3, 4] and decomposition method [8], Bae et al. obtained the stationary distribution of the workload process of the queueing system with the policy. Recently, Lee [9], Lee and Kim [10] extended P_λ^M-service policy to the queueing system with vacations or set-up times. Kim and Bae [7] applied the policy to $G/M/1$ queueing system and they obtained the stationary workload of the system.

Kim *et al.* [6] considered the problem of finding optimal service rate M under a cost structure. Four costs are considered by them: the operating cost per unit time when the service rate is M, the cost of increasing the service rate from 1 to M, the penalty cost per unit time while the server is idle, the holding cost of a unit workload per unit time. Under the cost structure, P_λ^M-service policy with the optimal service rate M controls the queueing system so that the system has low workload in average and has low proportion of idle periods while the operating cost and the cost of increasing service rate are low. However, in some situations, a specific quality of service (QoS) are required. For example, mean waiting time, variance of waiting time, and the probability that the workload exceeds a given amount are required to be less than fixed values. In this paper, we consider the case that the mean waiting time is required to be less than a fixed value. To this end, we define some costs and address the problem how to find the optimal service rate M.

2. Decomposition of the workload process

Let m be the mean amount of the work induced by a job. Then, the traffic intensity $\rho = \nu m$ is the mean amount of work induced by jobs during a unit time. For the stability of the system, we assume $M > \rho$. Note that we assume $M \geq 1$. Then, the workload process of the system is regenerative process. Let $\{X(t); t \geq 0\}$ be the workload process of $M/G/1$ queueing system with P_λ^M-service policy. The regeneration points of workload process $\{X(t); t \geq 0\}$ are the epochs when the server starts to work after an idle period. The process $\{X(t); t \geq 0\}$ is non-Markovian, which make it difficult to analyze the process. To overcome this, we decompose $\{X(t); t \geq 0\}$ into three processes $\{X_1(t); t \geq 0\}$, $\{X_2(t); t \geq 0\}$, and $\{X_3(t); t \geq 0\}$ as the authors have done in [2]. Process $\{X_1(t); t \geq 0\}$ is formed by separating the periods of service rate 1 from the original process and then connecting these together. In the similar manner, Process $\{X_2(t); t \geq 0\}$ is formed from the periods of service rate M. Process $\{X_3(t); t \geq 0\}$ is formed by connecting the rest of original process, that is, the idle periods. Then, $X_3(t) \equiv 0$ for all $t \geq 0$. See figures 1, 2, and 3.

Processes $\{X_1(t); t \geq 0\}$ and $\{X_2(t); t \geq 0\}$ are now Markovian regenerative processes. In both processes, we will call each separated segment a cycle. The starting levels of cycles in $\{X_1(t); t \geq 0\}$ are independent and have the same distribution function $G(x)/G(\lambda)$, $0 \leq x \leq \lambda$. The starting levels of cycles in $\{X_2(t); t \geq 0\}$ are also independent and have the same distributions as random variable S_a, where $S_a = \lambda + L$ and L is the first exceeding amount over λ

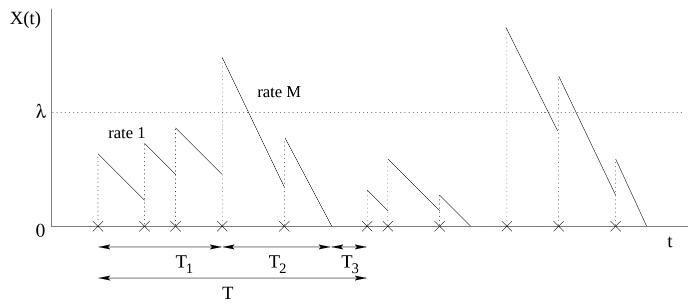

FIGURE 1. A sample path of workload process $\{X(t),\, t \geq 0\}$

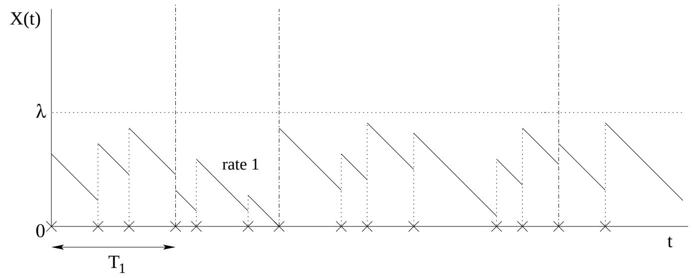

FIGURE 2. A sample path of $\{X_1(t),\, t \geq 0\}$

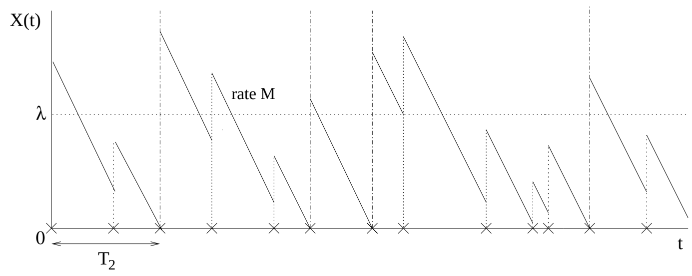

FIGURE 3. A sample path of $\{X_2(t),\, t \geq 0\}$

after the server restarts to work in the original process. Each idle period of the original process becomes a cycle in $\{X_3(t); t \geq 0\}$.

Let $F_i(x)$ be the stationary distribution function of $\{X_i(t); t \geq 0\}$, and T_i be the length of a cycle in $\{X_i(t); t \geq 0\}$, for $i = 1, 2, 3$. Let $F(x)$ be the stationary distribution function of $\{X(t); t \geq 0\}$ and T be the length of a regeneration cycle in $\{X(t); t \geq 0\}$. See figures 1, 2, and 3. Using the renewal reward theorem [11, p.133], Bae et al. [2] showed that for $x \geq 0$,

$$F(x) = \frac{\alpha E[T_1]}{E[T]} F_1(x) + \frac{\beta E[T_2]}{E[T]} F_2(x) + \frac{1/\nu}{E[T]}, \tag{1}$$

where α and β are the probabilities that there exists a period of service rate 1 and a period of service rate M in a cycle of $\{X(t); t \geq 0\}$, respectively. Then, $E[T] = \alpha E[T_1] + \beta E[T_2] + 1/\nu$.

3. Costs of the system

Kim *et al.* [6] considered the problem of finding optimal service rate M under a cost structure. Four costs are considered by them:

- $h(M)$: the operating cost per unit time when the service rate is M.
- $g(M)$: the cost of increasing the service rate from 1 to M.
- C_i: the penalty cost per unit time while the server is idle.
- C_h: the holding cost of a unit workload per unit time.

The cost function of a queueing system with P_λ^M-service policy is the sum of the costs in the above. The functions $h(M)$ and $g(M)$ are defined for $M \geq 1$ because it is assume that $M \geq 1$. When $M = 1$, there is no change of service rate, therefore, no cost occurred, which implies $h(1) = 0$ and $g(1) = 0$. They assumed that $h(M)$ and $g(M)$ are nondecreasing and twice differentiable convex functions including linear functions. These are usually assumed for the cost functions. They also assumed that C_i and C_h are positive. By doing this, P_λ^M-service policy with the service rate M of the minimum total cost controls the queueing system so that the system may has low workload in average and has low proportion of idle periods while the operating cost and the cost of increasing service rate are low. In this paper, we consider the case that the mean waiting time is required to be less than a fixed value. In other words, the time which a job spends in the system is expected to be less than a fixed value. Then, there is no reason to consider the penalty cost for the idle state of the server and the holding cost, in other words, $C_i = 0$ and $C_h = 0$. We assume that $h(M)$ and $g(M)$ are strictly increasing and twice differentiable convex function including linear functions. The cost during a cycle in the present model is $\beta g(M) + \beta E[T_2]h(M)$ in average. Since the mean cycle length is $E[T]$, the problem to solve is the following:

$$\begin{aligned} \textbf{minimize} \quad & C(M) = \frac{\beta}{E[T]} g(M) + \frac{\beta E[T_2]}{E[T]} h(M). \\ \textbf{subject to} \quad & E[W] \leq w_0, \end{aligned} \tag{2}$$

where W is the waiting time of a job and w_0 is a constant. To solve the nonlinear optimization problem in the above, we need to know the explicit form of $E[W]$. However, it is difficult to obtain the explicit form of $E[W]$. Let Y be the workload of the system at the arrival of a job. Then, it is clear that

$W \le Y$. Thus, $E[Y] \le w_0$ implies that $E[W] \le w_0$. Now, in this paper, we consider the following nonlinear programming:

$$(3) \qquad \begin{aligned} &\textbf{minimize} \quad C(M) = \frac{\beta}{E[T]} g(M) + \frac{\beta E[T_2]}{E[T]} h(M). \\ &\textbf{subject to} \quad E[Y] \le w_0, \end{aligned}$$

The optimal solution of the above problem is not the optimal solution of the problem (2). However, it is a feasible solution and can be used as an approximate optimal rate M of the problem (2).

To solve the nonlinear optimization problem in the above, we need to know the explicit form of $E[T_2]$, $E[T]$, $E[Y]$ and the values of α and β. $E[Y]$ will be obtained in the next section. Clearly, $\alpha = G(\lambda)$. Bae *et al.* [2] obtained that

$$\beta = \frac{H'(\lambda)}{\nu H(\lambda)}, \quad E[T_1] = \frac{1}{\nu G(\lambda)} \left(H(\lambda) - 1 - \frac{H'(\lambda)}{H(\lambda)} \int_0^\lambda H(x)dx \right),$$

where $H(x) = \sum_{n=0}^\infty \rho^n G_e^{*n}(x)$, $G_e(x) = (1/m)\int_0^x (1-G(u))du$, the equilibrium distribution function of G, and $*n$ is the n-fold Stieltjes convolution with G_e^{*0} being Heaviside function. It is shown in Asmussen [1, p.113] that $H(x)$ is well-defined for all nonnegative ρ and x. Since the process $\{X_2(t); t \ge 0\}$ in a cycle is identical to the workload process of an $M/G/1$ queueing system with initial workload S_a in a busy period, it follows from Wolff [12, p.393] that the expectation of T_2 is given by

$$E[T_2] = \frac{E[S_a]}{M - \rho}.$$

The probability density function of S_a is also obtained by Bae et al. [2]. However, the explicit form is complicated. Kim et al. [6] obtained a simpler form of $\Pr\{S_a > s\}$, which is given by, for $s \ge \lambda$,

$$\Pr\{S_a > s\} = \frac{\nu H(\lambda)}{H'(\lambda)} \int_{0-}^\lambda (1 - G(s-u))\, dH(u) - \nu \int_0^\lambda (1 - G(s-u)) H(u)\, du.$$

By integrating the above, we derive

$$E[S_a] = \frac{H(\lambda)}{H'(\lambda)} - (1-\rho)\frac{H(\lambda)^2}{H'(\lambda)} + (1-\rho)\int_0^\lambda H(u)\, du.$$

4. Analysis of the cost function

From the results of the pervious section, we can see that

$$C(M) = \frac{\beta g(M) + \beta E[S_a] h(M)/(M-\rho)}{\alpha E[T_1] + \beta E[S_a]/(M-\rho) + 1/\nu}.$$

Let $A = \alpha E[T_1] + 1/\nu$. Then, the above equation is rewritten as

$$(4) \qquad C(M) = \frac{(M-\rho)\beta g(M) + \beta E[S_a] h(M)}{A(M-\rho) + \beta E[S_a]}.$$

The behavior of the function $C(M)$ is given by the following proposition.

Proposition 1. *If $\rho \geq 1$ and $\beta E[S_a]h'(\rho) - Ah(\rho) + \beta g(\rho) < 0$, then $C(M)$ is strictly decreasing for $M < m_0$ and strictly increasing for $M > m_0$, where m_0 is the solution of the equation $\beta E[S_a]h'(M) - Ah(M) + \beta g(M) = 0$. Otherwise, $C(M)$ is strictly increasing.*

Proof. By differentiating Eq. (4) with respect to M, we have

$$C'(M) = \frac{\beta E[S_a]N(M)}{(A(M-\rho) + \beta E[S_a])^2},$$

where

$$N(M) = \left(\frac{A(M-\rho)}{E[S_a]} + \beta\right)(E[S_a]h'(M) + (M-\rho)g'(M)) + \beta g(M) - Ah(M). \tag{5}$$

The derivative of $N(M)$ is computed as

$$N'(M) = \left(\frac{A(M-\rho)}{E[S_a]} + \beta\right)(E[S_a]h''(M) + (M-\rho)g''(M) + 2g'(M)). \tag{6}$$

Since $h(M)$ and $g(M)$ are strictly increasing and twice differentiable convex functions for $M > 1$, we have that $N'(M) > 0$ for $M > 1$. Thus, $N(M)$ is strictly increasing function for $M > 1$. Since $g(M)$ is strictly increasing and convex, $g'(M) > \epsilon$ for $M > 1 + \delta$, where δ and ϵ are positive numbers. Then, Eqn. (6) says that $N(M)$ goes to ∞ as M goes to ∞.

Note that M is assumed to be larger than $\max\{1, \rho\}$. Thus, $N(M)$ has its infimum value at $\max\{1, \rho\}$. When $\rho < 1$, we obtain by substituting 1 to M in Eq. (5) that for $M > 1$,

$$N(M) > \left(\frac{A(1-\rho)}{E[S_a]} + \beta\right)(E[S_a]h'(1) + (1-\rho)g'(1)),$$

which is nonnegative. When $\rho \geq 1$, we obtain by substituting ρ to M in Eq. (5) that for $M > \rho$,

$$N(M) > \beta E[S_a]h'(\rho) - Ah(\rho) + \beta g(\rho),$$

whose value can be negative or positive. $C'(M)$ has the same sign as $N(M)$, which completes the proof.

The above proposition says that $C(M)$ has its minimum at $M = m_0$ if $\rho \geq 1$ and $\beta E[S_a]h'(\rho) - Ah(\rho) + \beta g(\rho) < 0$. Otherwise, it has its minimum at $M = \max\{1, \rho\}$.

5. Expected workload seen by a job and its behavior

In this section, we obtain $E[Y]$, the expected workload seen by a job. Since the arrival process is Poisson, PASTA [12] says that the workload seen by a job has the same distribution as the stationary workload, i.e. $F(x)$. From Eq. (1), it follows that

$$E[Y] = \frac{\alpha E[T_1]}{E[T]}E[Y_1] + \frac{\beta E[T_2]}{E[T]}E[Y_2], \tag{7}$$

where Y_1 and Y_2 are the stationary workload of the processes $\{X_1(t), t \geq 0\}$ and $\{X_2(t), t \geq 0\}$, respectively.

Proposition 2. *Let Y be the workload seen by a job. Then,*

$$E[Y] = \frac{B + (\beta E[S_a^2]/2)/(M-\rho) + (\nu\beta E[S_a]E[S^2]/2)/(M-\rho)^2}{A + \beta E[S_a]/(M-\rho)},$$

where

$$B = \frac{1}{\nu}\left(\int_0^\lambda x\, dH(x) + \frac{H'(\lambda)}{2H(\lambda)}\int_0^\lambda x^2\, dH(x) - \frac{\lambda^2}{2}H'(\lambda)\right).$$

Proof. Let $f_1(x)$ and $f_2(x)$ be the derivatives of $F_1(x)$ and $F_2(x)$, respectively. Then, Bae et al. [2] obtained that

$$f_1(x) = \frac{1}{\nu\alpha E[T_1]}\left(H'(x) - \frac{H(x)H'(\lambda)}{H(\lambda)}\right),$$

and, for $0 < x \leq \lambda$,

$$f_2(x) = \frac{H_M(x)}{E[T_2]},$$

for $x \geq \lambda$,

$$f_2(x) = \frac{H_M(x) - \int_\lambda^x H_M(x-y) f_{S_a}(y)dy}{E[T_2]},$$

where $H_M(x) = \sum_{n=0}^\infty (\rho/M)^n G_e^{*n}(x)$ and $f_{S_a}(x)$ is the probability density function of S_a. From the above equations, we can evaluate

$$E[Y_1] = \frac{1}{\nu\alpha E[T_1]}\left(\int_0^\lambda x\, dH(x) + \frac{H'(\lambda)}{2H(\lambda)}\int_0^\lambda x^2\, dH(x) - \frac{\lambda^2}{2}H'(\lambda)\right)$$

and

$$E[Y_2] = \frac{1}{2E[T_2]}\left(\frac{E[S_a^2]}{(M-\rho)} + \frac{\nu E[S_a]E[S^2]}{(M-\rho)^2}\right).$$

Another approach to derive $E[Y_2]$ is given by Kim *et al.* [6]. By applying the above two equations to Eq. (7), we obtain the desired result.

Let $x = A + \beta E[S_a]/(M-\rho)$. Then, the above equation is rewritten as

$$E[Y] = \frac{E[S_a^2]}{2E[S_a]} - \frac{\nu A E[S^2]}{2\beta E[S_a]} + \frac{\nu E[S^2]}{2\beta E[S_a]}x + \left(B - \frac{AE[S_a^2]}{2E[S_a]} + \frac{\nu A^2 E[S^2]}{2\beta E[S_a]}\right)\frac{1}{x}$$

and its derivative with respect to x is given by

$$\frac{\partial E[Y]}{\partial x} = \frac{\nu E[S^2]}{2\beta E[S_a]} - \left(B - \frac{AE[S_a^2]}{2E[S_a]} + \frac{\nu A^2 E[S^2]}{2\beta E[S_a]}\right)\frac{1}{x^2}. \tag{8}$$

When x is increasing, the behavior of $E[Y]$ is determined according to whether the coefficient of $1/x^2$ of the above equation is positive or not.

Proposition 3. *If* $B - AE[S_a^2]/(2E[S_a]) + \nu A^2 E[S^2]/(2\beta E[S_a]) \leq 0$, *then* $E[Y]$ *is strictly decreasing with respect to* M. *Otherwise,* $E[Y]$ *is strictly increasing for* $M < m_1$ *and strictly decreasing for* $M > m_1$, *where*

$$m_1 = \rho + \frac{\beta E[S_a]}{\sqrt{\frac{2\beta E[S_a]}{\nu E[S^2]} B - \frac{\beta E[S_a^2]}{\nu E[S^2]} A + A^2} - A}.$$

Proof. If the coefficient of $1/x^2$ in Eq. (8) is less than or equal to 0, then $E[Y]$ is strictly increasing with respect to x. Since x is strictly decreasing function of M, we can see that $E[Y]$ is strictly decreasing with respect to M. For the other case, Eq. (8) says that $E[Y]$ is strictly decreasing function of x for $x < x^*$ and $E[Y]$ is strictly increasing function of x for $x > x^*$, where x^* is the solution of

$$\frac{\nu E[S^2]}{2\beta E[S_a]} - \left(B - \frac{AE[S_a^2]}{2E[S_a]} + \frac{\nu A^2 E[S^2]}{2\beta E[S_a]} \right) \frac{1}{x^2} = 0.$$

Then, the remaining part of the proof immediately follows if we recall that x is strictly decreasing function of M and its form is $A + \beta E[S_a]/(M - \rho)$.

Usually, for a queueing system having uniform service rate, increasing service rate results in decreasing of the stationary workload. However, we can see from the above proposition that $E[Y]$ may increase when the service rate M increases. This is due to the fact that the queueing system considered in this paper has two service rates and high value of M makes increasing the proportion of time when the system serves with rate 1.

6. Finding the optimal service rate M

In this section, we present how to find the optimal service rate M for the optimization problem (3). To this end, we first observe the some values of $E[Y]$ given by Table 1.

M	$E[Y]$
1 when $\rho < 1$	$E[S] + \nu E[S^2]/(2(1-\rho))$
ρ when $\rho \geq 1$	∞
∞	B/A

TABLE 1. The expected stationary workload for some values of M

Propositions 1 and 3 enable us to divide all cases into six. Let $C = B - AE[S_a^2]/(2E[S_a]) + \nu A^2 E[S^2]/(2\beta E[S_a])$. Then, the six cases are the followings:

Case 1: $C \leq 0$ and $\rho < 1$.
Case 2: $C \leq 0$, $\rho \geq 1$ and $\beta E[S_a]h'(\rho) - Ah(\rho) + \beta g(\rho) \geq 0$.
Case 3: $C \leq 0$, $\rho \geq 1$ and $\beta E[S_a]h'(\rho) - Ah(\rho) + \beta g(\rho) < 0$.

cases	M^* (the optimal value of M)
Case 1	if $w_0 \le B/A$, then there is no feasible solution. if $B/A < w_0 < E[S] + \nu E[S^2]/(2(1-\rho))$, then $M^* = m^*$. if $E[S] + \nu E[S^2]/(2(1-\rho)) \le w_0$, then $M^* = 1$.
Case 2	if $w_0 \le B/A$, then there is no feasible solution. if $B/A < w_0$, then $M^* = m^*$.
Case 3	if $w_0 \le B/A$, then there is no feasible solution. if $B/A < w_0$ and $m^* \ge m_0$, then $M^* = m^*$. if $B/A < w_0$ and $m^* < m_0$, then $M^* = m_0$.
Case 4	if $m_1 \le 1$ and $w_0 \le B/A$, then there is no feasible solution. if $m_1 \le 1$ and $B/A < w_0 < E[S] + \nu E[S^2]/(2(1-\rho))$, then $M^* = m^*$. if $m_1 \le 1$ and $E[S] + \nu E[S^2]/(2(1-\rho)) \le w_0$, then $M^* = 1$. if $m_1 > 1$ and $w_0 \le \min\{B/A, E[S] + \nu E[S^2]/(2(1-\rho))\}$, then there is no feasible solution. if $1 < m^* < m_1$ and $\min\{B/A, E[S] + \nu E[S^2]/(2(1-\rho))\} < w_0$ $< \max\{B/A, E[S] + \nu E[S^2]/(2(1-\rho))\}$, then $M^* = 1$. if $1 < m_1 < m^*$ and $\min\{B/A, E[S] + \nu E[S^2]/(2(1-\rho))\} < w_0$ $< \max\{B/A, E[S] + \nu E[S^2]/(2(1-\rho))\}$, then $M^* = m^*$. $1 < m_1$ and $w_0 \ge \max\{B/A, E[S] + \nu E[S^2]/(2(1-\rho))\}$, then $M^* = 1$.
Case 5	if $w_0 \le B/A$, then there is no feasible solution. if $B/A < w_0$, then $M^* = m^*$.
Case 6	if $w_0 \le B/A$, then there is no feasible solution. if $B/A < w_0$ and $m^* \ge m_0$, then $M^* = m^*$. if $B/A < w_0$ and $m^* < m_0$, then $M^* = m_0$.

TABLE 2. The optimal solution of problem (3). In the table, m^* is the solution of $E[Y] = w_0$. If there are more than one solution, then m^* is the smallest one.

Case 4: $C > 0$ and $\rho < 1$.
Case 5: $C > 0$, $\rho \ge 1$ and $\beta E[S_a]h'(\rho) - Ah(\rho) + \beta g(\rho) \ge 0$.
Case 6: $C > 0$, $\rho \ge 1$ and $\beta E[S_a]h'(\rho) - Ah(\rho) + \beta g(\rho) < 0$.

In cases 1 and 2, Propositions 1 and 3 say that $E[Y]$ is strictly decreasing and $C(M)$ is strictly increasing. It is the optimal M such that $E[Y] = w_0$ if exists. From Table 1, we can see that the equation $E[Y] = w_0$ has solution when $B/A < w_0 \le E[S] + \nu E[S^2]/(2(1-\rho))$ for $\rho < 1$ and $B/A < w_0$ for $\rho \ge 1$. When $w_0 \le B/A$, there is no M satisfying $E[Y] \le w_0$. When $w_0 > E[S] + \nu E[S^2]/(2(1-\rho))$ and $\rho < 1$, then $M = 1$ is the optimal service rate.

In case 3, Propositions 1 and 3 say that $E[Y]$ is strictly decreasing and that $C(M)$ is decreasing for $M < m_0$ and increasing for $M > m_0$. Let m^* be the unique solution of $E[Y] = w_0$ if exists. Then, it is clear that $M = m^*$ is the optimal service rate if $m^* \ge m_0$. Otherwise, $M = m_0$ is the optimal service rate. Since the existence of such M with which $E[Y] = w_0$ depends only on the value of $E[Y]$, the condition of the existence is the same the above. If $w_0 \le B/A$, there is no M satisfying $E[Y] \le w_0$.

In case 4, Propositions 1 and 3 say that $C(M)$ is strictly increasing and $E[Y]$ is strictly increasing for $M < m_1$ and strictly decreasing for $M > m_1$. If $m_1 \leq 1$, then $E[Y]$ is strictly decreasing for $M \in [1, \infty)$. Then, $C(M)$ and $E[Y]$ show the same behavior as in case 1 for $M \in [1, \infty)$ because we assume $M \geq 1$. Thus, we can find the optimal M in the same manner as in case 1. Now, we consider the case that $m_1 > 1$. Let m^* be the smallest M with which $E[Y] = w_0$ if exists. If $m^* \leq m_1$, then 1 is the optimal solution because $E[Y]$ is an increasing function of M for $M \in [1, m^*)$. If $m^* > m_1$, then m^* is the optimal solution because $E[Y]$ is a decreasing function of M for $M > m^*$. The condition of the existence of M with which $E[Y] = w_0$ can be deduced from Table 1. When w_0 has a value between B/A and $E[S] + \nu E[S^2]/(2(1-\rho))$, the equation $E[Y] = w_0$ has one solution. If $w_0 > \max(E[S] + \nu E[S^2]/(2(1-\rho)), B/A)$, then 1 is the optimal solution. If $w_0 < \min(E[S] + \nu E[S^2]/(2(1-\rho)), B/A)$, then there is no M such that $E[Y] \leq w_0$.

In case 5, $C(M)$ and $E[Y]$ show the same behavior as in Case 4. From Table 1, we can see that $E[Y]$ goes to ∞ as M goes to ρ, which implies that $m_1 < \rho$. Then, $E[Y]$ is strictly decreasing for $M \in (\rho, \infty)$. Then, $C(M)$ and $E[Y]$ show the same behavior as in case 2 for $M \in (\rho, \infty)$. Thus, we can find the optimal M in the same manner as in case 2 because we assume $M > \rho$.

In case 6, Propositions 1 and 3 say that $E[Y]$ is strictly decreasing for $M < m_1$ and strictly increasing for $M > m_1$, and say that $C(M)$ is strictly decreasing for $M < m_0$ and strictly increasing for $M > m_0$. In case 6, we assume $\rho > 1$. Then, by the same reason as in case 5, $E[Y]$ is strictly decreasing for $M \in (\rho, \infty)$. Then, $C(M)$ and $E[Y]$ show the same behavior as in case 3 for $M \in (\rho, \infty)$. Thus, we can find the optimal M in the same manner as in case 3 because we assume $M > \rho$. Now, the optimal values in all case are given in Table 2.

References

[1] S. Asmussen, *Applied Probability and Queues*, John Wiley & Sons, 1987.

[2] J. Bae, S. Kim and E. Y. Lee, *A P_λ^M policy for an $M/G/1$ queueing system,* Appl. Math. Modell. **26** (2002), 929–939.

[3] P. H. Brill and M. J. M. Posner, *Level Crossings in Point Process Applied to Queues: Single-server Case,* Oper. Res. **25** (1977), 662–674.

[4] J. W. Cohen, *On Up- and Downcrossings,* J. Appl. Probab. **14** (1977), 405–410.

[5] M. J. Faddy, *Optimal Control of Finite Dams: Discrete(2-stage) Output Procedure,* J. Appl. Probab. **11** (1974), 111—121.

[6] J. Kim, J. Bae and E. Y. Lee, An optimal P_λ^M-service policy for an $M/G/1$ queueing system, Appl. Math. Modell. **30** (2006), 38–48.

[7] S. Kim and J. Bae, *Stationary workload in a G/M/1 queueing system with P_λ^M-service policy,* Oper. Res. Lett. **36** (2008), 201–204

[8] E. Y. Lee and S. K. Ahn, *P_λ^M-Policy for a Dam with Input Formed by a Compound Poisson Process,* J. Appl. Probab. **35** (1998), 482–488.

[9] J. Lee, *An $M/G/1$ vacation queue under the P_λ^M-service policy,* J. Korean Statist. Soc. **36** (2007), 285–297.

[10] J. Lee and J. Kim, *Workload analysis of an $M/G/1$ queue under P_λ^M-service policy with a set-up time,* Appl. Math. Modell. **31** (2007), 236–244.
[11] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, New York, 1983.
[12] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*, Prentice Hall, 1989.
[13] D. Zuckerman, *Two-stage Output Procedure of a Finite Dam,* J. Appl. Probab. **14** (1977), 421–425.

Sunggon Kim: Department of Information Statistics and RINS, Gyeongsang National University, Jinju 660-701, Republic of Korea
E-mail address: sgkim@gnu.ac.kr

Jongho Bae: Department of Statistics, Chungnam National University, Daejeon 305-764, Republic of Korea
E-mail address: bae-jongho@cnu.ac.kr