

## **A SIMPLE APPROACH FOR THE SLOWER SERVER PROBLEM**

M. O. ABOU-EL-ATA and A. I. SHAWKY

*Faculty of Engineering at Shoubra, Cairo, Egypt.*

(Received April. 30,1998; Revised Feb 4, 1999; Accepted Feb. 15, 1999)

### **ABSTRACT**

In this paper we aim to study when and how to discard the slower server from a heterogeneous two-channels queueing system with a general discipline. Our approach depends basically on the delay probability  $P_0$ . Also we discuss if it is possible to discard any server in the homogeneous case. Then we deduce some special cases of our work.

### **1. INTRODUCTION**

The problem of discarding the slower is of great importance in practice. Rubinovitch [1,2] studied the problem of a heterogeneous two- channels queueing systems. In his first paper he discussed three simple models and gave the condition when to discard the slower server depending on the expected number of units in the system  $L$ . In the second paper he studied a queueing model with a stalling concept.

In our work we aim to introduce a simpler approach to find the condition when to discard the slower server in a heterogeneous two- channels queue. Our approach depends on the delay probability  $P_0$  which is more simpler than that depending on the expected number of units in the system as in Rubinowitch [1]. But we shall discuss first the homogeneous case to look for the condition when and how to discard a server of the multi-channels  $M/M/C$  queue if it is possible. Then we deduce some special cases and compare our results with that of Rubinovitch [1].

### **2. THE HOMOGENEOUS MODEL ANALYSIS**

Let us here look for the condition when and how to discard a server of the homogeneous multi-channels queue  $M/M/C$ . As found in most texts, the delay probability  $P_0$  is

$$P_0 = P_0(C) = \left[ \sum_{n=0}^{C-1} \frac{\rho^n}{n!} + \frac{\rho^C}{(C-1)!(C-\rho)} \right]^{-1}, \quad 0 \leq \rho = \frac{\lambda}{\mu} < C, \quad (1)$$

which is an increasing function of  $C$  when  $\rho$  is assumed to be constant. To answer the first question "when?" we know that:

$$P_0(C+1) > P_0(C),$$

but when it happens that for some values of  $\rho$ :  $P_0(C+1) < P_0(C)$  then we can discard any one of the  $C$  servers.

To answer the second question "how?" we try to find a condition on the values of the utilization factor  $\rho$  that could be greater or equal zero. This could be understood from the following lemma (1).

**Lemma 1.** For the homogeneous multi-channels queue:  $M/M/C$ , it is impossible to find any condition on  $\rho$  in order to discard a server.

The proof is obvious.

### 3. THE HETEROGENEOUS MODEL ANALYSIS

We study a heterogeneous two-channels  $M/M/2$  queue with a general discipline, to find the condition when to discard the slower server. Assume the units arrive to the system according to the Poisson distribution with rate  $\lambda$  and exponentially served with rates:  $\mu_1$  and  $\mu_2$  ( $\mu_1 \geq \mu_2$ ). The discipline considered is that of Krishnamoorthi [3] which is as follows:

- i) if the system is empty, the head unit joins the first server with probability  $\pi_1$  or to the second slower server with probability  $\pi_2$ ,  $\pi_1 + \pi_2 = 1$ .
- ii) if one server is empty, the head of the queue goes to it.
- iii) if both servers are busy, the head unit waits until one server becomes free.

Let us define the following probabilities:

$P_n$  = The steady-state probability that there are  $n$  units ( $n \geq 1$ ) in the system,

$P_0$  = prob. {there is no unit in the system}

$P_{10}$  = prob. {a unit being served in the first channel},

$P_{01}$  = prob. {a unit being served in the second channel}, and  $P_1 = P_{10} + P_{01}$ .

As usual the steady-state probability difference equations can be easily deduced in the following form

$$-\lambda P_0 + \mu_1 P_{10} + \mu_2 P_{01} = 0, \quad n = 0, \quad (2)$$

$$\left. \begin{aligned} -(\lambda + \mu_1) P_{10} + \lambda \pi_1 P_0 + \mu_2 P_2 &= 0, \\ -(\lambda + \mu_2) P_{01} + \lambda \pi_2 P_0 + \mu_1 P_2 &= 0, \end{aligned} \right\} \quad n = 1, \quad (3)$$

$$-(\lambda + \mu) P_n + \lambda_1 P_{n-1} + \mu P_{n+1} = 0, \quad \mu = \mu_1 + \mu_2, \quad n \geq 2. \quad (4)$$

Solving equations (2) – (4) for  $P_n$  we can get

$$P_0 = \frac{1 - \rho}{1 - \rho + \lambda / \Delta}, \quad (5)$$

where

$$\Delta^{-1} = \frac{\lambda + \pi_2 \mu_1 + \pi_1 \mu_2}{(2\rho + 1) \mu_1 \mu_2}, \quad \rho = \frac{\lambda}{\mu_1 + \mu_2} = \frac{\lambda}{\mu}, \quad (6)$$

$$P_1 = P_{10} + P_{01} = (\lambda / \Delta) P_0, \quad (7)$$

and

$$P_n = \rho^{n-1} P_1, \quad n \geq 1. \quad (8)$$

Now let us discuss the problem when to discard the slower server in the heterogeneous case. The condition depends on the delay probability  $P_0$  as we shall see in the following Lemma 2.

**Lemma 2.** The slower could be discarded in the heterogeneous two-channels queue M/M/2 for those values of  $\rho \leq \rho_c$  satisfying the following equation:

$$(\mu_1^2 + \mu_2^2) \rho_c^2 [\pi_1 (\mu_2^2 + \mu_1^2) + \mu_1 \mu_2] \rho_c + \mu_1 (\mu_2 - \mu_1) \pi_2 = 0.$$

**Proof.** As in Lemma 1, consider on the contrary that  $P_0$  for the queue M/M/2 is less than  $P_0$  for the queue M/M/1,

i.e.  $P_0(C = 2) < P_0(C = 1)$

$$\begin{aligned} \Rightarrow \frac{1 - \rho}{1 - \rho + (\lambda / \Delta)} &< 1 - \frac{\lambda}{\mu_1} \\ 1 - \rho &< (1 - \frac{\lambda}{\mu_1})(1 - \rho + \frac{\lambda}{\Delta}) \\ \frac{\lambda}{\Delta \mu_1} [(\mu - \Delta) \rho + \Delta - \mu_1] &< 0 \end{aligned}$$

i.e.

$$\left[ \mu - \frac{(2\rho+1)\mu_1\mu_2}{\mu\rho + \pi_2\mu_1 + \pi_1\mu_2} \right] \rho + \frac{(2\rho+1)\mu_1\mu_2}{\mu\rho + \pi_2\mu_1 + \pi_1\mu_2} - \mu_1 < 0$$

$$(\mu^2 - 2\mu_1\mu_2)\rho^2 + [\mu(\pi_2\mu_1 + \pi_1\mu_2) + \mu_1\mu_2 - \mu_1\mu]\rho + \mu_1\mu_2 - \mu_1(\pi_2\mu_1 + \pi_1\mu_2) < 0.$$

Thus we have to find the roots  $\rho_C$  of the quadratic equation:

$$(\mu_1^2 + \mu_2^2)\rho_C^2 + [\pi_1(\mu_2^2 - \mu_1^2) + \mu_1\mu_2]\rho_C + \mu_1(\mu_2 - \mu_1)\pi_2 = 0.$$

Therefore let us discuss some queueing models depending on those probability values  $\pi_1$  and  $\pi_2$  ( $\pi_1 + \pi_2 = 1$ ) as follows :

**Model 1.** Let  $\pi_1 = \pi_2 = 1/2$ .

This is the classical heterogeneous two-channels M/M/2 queueing system. Thus equation (9) reduces to

$$2(\mu_1^2 + \mu_2^2)\rho_C^2 + (\mu_2^2 - 2\mu_1\mu_2 - \mu_1^2)\rho_C + \mu_1(\mu_2 - \mu_1) = 0.$$

$$(2\rho_C + 1)[(\mu_1^2 + \mu_2^2)\rho_C + \mu_1\mu_2 - \mu_1^2] = 0$$

i.e.,

$$\rho_C = \frac{\mu_1^2 - \mu_1\mu_2}{\mu_1^2 + \mu_2^2}. \quad (10)$$

Let  $\mu_2 = r\mu_1$ ,  $0 \leq r \leq 1$ , then relation (10) reduces to

$$\rho_C = \frac{1-r}{1+r^2} = 1 - \frac{r(1+r)}{1+r^2}.$$

Hence  $\rho \in (0,1]$  when  $r \in [0,1)$ , and we could discard the slower second server when  $\rho < \rho_C$  for all  $r \in [0,1)$ . This case is the same as in Rubinovitch's [1] first model. The case  $r=1$  is the homogeneous queueing system in which we could not discard any server.

**Model (2):** Let  $\pi_1 = 1$  and  $\pi_2 = 0$ .

This is in fact Singh's [4] model, and Rubinovitch [1] call it an informed customer model. In this case equation (9) reduces to:

$$(\mu_1^2 + \mu_2^2)\rho_C^2 + (\mu_2^2 - \mu_1^2 + \mu_1\mu_2)\rho_C = 0$$

i.e.

$$\rho_C = \frac{\mu_1^2 - \mu_1\mu_2 - \mu_2^2}{\mu_1^2 + \mu_2^2}. \quad (12)$$

Let  $\mu_2 = r\mu_1$ ,  $0 \leq r \leq 1$ , then relation (12) becomes

$$\rho_c = \frac{1-r-r^2}{1+r^2} = 1 - \frac{r(1+2r)}{1+r^2}. \quad (13)$$

For  $0 \leq r \leq 0.618$  we have  $0 \leq \rho_c \leq 1$  and thus we could discard the slower second server  $\rho \leq \rho_c$ . But when  $0.618 < r \leq 1$ , we have  $\rho_c < 0$  which is an impossible condition and we could not discard the slower server at all. When  $r=1$  we have the homogeneous case and we could not discard any server.

**Model (3):** Let  $\pi_i = \frac{\mu_i}{\mu}$ ,  $i=1,2$

This is Saaty's [5] model, and thus equation (9) reduces to

$$\begin{aligned} (\mu_1^2 + \mu_2^2)\rho_c^2 + \left[ \frac{\mu_1}{\mu}(\mu_2^2 - \mu_1^2)\mu_1\mu_2 \right] \rho_c + (\mu_2 - \mu_1)\frac{\mu_1\mu_2}{\mu} &= 0 \\ (\mu_1^2 + \mu_2^2)\rho_c^2 + \mu_1(2\mu_2 - \mu_1)\rho_c + \frac{\mu_1\mu_2}{\mu}(\mu_2 - \mu_1) &= 0. \end{aligned} \quad (14)$$

Let  $\mu_2 = r\mu_1$ ,  $0 \leq r \leq 1$ , then relation (14) becomes

$$(1+r^2)\rho_c^2 - (1-2r)\rho_c + \frac{r(r-1)}{1+r} = 0, \quad (15)$$

which is a quadratic equation and has the following two roots

$$\rho_c = \frac{1-2r \pm \sqrt{(1-2r)^2 + \frac{4r(1-r)(1+r^2)}{1+r}}}{2(1+r^2)}$$

i.e.

$$\rho_c = \frac{1-2r \pm \left| 1-2r \right| \sqrt{1 + \frac{4r(1-r)(1+r^2)}{(1+r)(1-2r^2)}}}{2(1+r^2)}. \quad (16)$$

**Case (3-a):** The root with positive sign.

When  $r=0$ , then  $\pi_1 = 1, \pi_2 = 0$  and  $\rho_c = 1$  which is the same result as in model (2).

For  $0 < r \leq 1$ , we have  $\rho_c \geq 0$  (where equality is satisfied at  $r=0.5$ ) and in this case we could discard the slower second server.

**Case (3-b):** The second root with negative sign.

For  $0 \leq r \leq 1$ , we have  $\rho_c \leq 0$  (where equality is satisfied at  $r=0, 0.5$ ) and in this case we could not discard the slower server.

**Conclusion:** Due to this new approach our results depends on a quadratic equation of  $\rho$  and not on a cubic equation as given in Rubinovitch [1] in the case of model (3), which is much simple to solve.

**Acknowledgment:** We are sincerely thankful to both referees for their valuable and constructive comments that had improved our original paper.

#### REFERENCES

- [1] Rubinovitch, M.(1985), " The slow server problem". J. App. Prob. 22, (205-213).
- [2] Rubinovitch, M. (1985), "The slow server problem: A queue with stalling". J.App. Prob. 22, (879-892).
- [3] Krishnamoorthi, B. (1963),"On Poisson queue with two heterogeneous servers". Op. Res. 11, (321-330).
- [4] Singh, V.P. "Queue-dependent servers". J.Eng.Math. 7, (123-126).
- [5] Saaty, T.L. (1961), " Elements of Queueing Theory". Mc Graw – Hill. N.Y..