ON A DUALITY PRINCIPLE IN PROCESSES OF SERVICING MACHINES WITH DOUBLE CONTROL

JEWGENI DSHALALOW
Department of Mathematics
Florida Institute of Technology
Melbourne, Florida 32901, U.S.A.

ABSTRACT

This paper examines a doubly controlled process of servicing machines. The classical system treated by Takács is equipped with \( m + 1 \) unreliable machines served by one repairman. In the present modification of this model, the failure rates and the repair time may be controlled with respect to the state of the system. The process describing the number of intact machines is considered. To derive its steady state distribution in the form of a simple explicit formula, the author introduces an auxiliary model with \( m \) unreliable machines and a single repairman who keeps working even when all machines are intact. This result is based on a duality principle applied to the process above.

Key words: Servicing Machines, Multi-Channel Queue, Control, Semi-Regenerative Process, Markov Chain, Semi-Markov Process, Duality Principle.

AMS Subject Classification: Primary 60K25, 60K10, Secondary 90B22, 90B25.

1. INTRODUCTION

In [2] the author studied a multi-channel loss (transportation) queueing system with control of input stream and service. The servicing facility of the system contained \( m \) parallel channels processing a stream of singly arriving customers. No customer was accepted when the servicing facility was occupied. The input stream and service were subject to a comprehensive control.

This model can also be interpreted as a system of \( m \) unreliable machines served by a single repairman with corresponding service and failure rates control. More specifically, each of the working machines can break down with a rate dependent on the total number of intact machines. The repair time also depends upon the number of intact machines. The repairman is not idle even when all machines become intact. At these times (called \textit{regenerations of the system}), the repairman leaves the system and comes back later with a new machine that immediately replaces an available defective machine. If no machine
breaks down during the repairman's absence, a substitution (of one of the working machines) takes place at the repairman's own choice, but the number of working machines does not change (it is still m), and this action is supposed not to affect the future status of the system. Since the repairman leaves the system to get new equipment, we can restrict his absence to a definite time. So the control can particularly be applied to this situation. Besides the clear advantages of this system as a doubly controlled transportation servicing model, it can also be viewed as a doubly controlled model of servicing machines.

Another more popular model of servicing machines with idle periods is one with m + 1 machines and with a single repairman who is either repairing defective machines or simply keeping watch when all m + 1 are intact. However, he is alert to start his service as soon as any working machine fails. Such a system is probably more practical than the one described above although obviously with a more difficult analysis. Without control it was originally set and studied by Takács [3] in his fundamental work. Takács solved many problems related to this system.

In the present paper, the author studies the process of servicing machines for the model with idle periods. The service time and failure rates are subject to control at moments of time appropriate for decision making. The process describing the number of intact machines is treated. By using some semi-regenerative techniques and by introducing a duality principle, the author obtains an explicit relation between the two models. The results are combined with those derived in [2].

2. MODEL DESCRIPTION

The models introduced in section 1 are formalized as follows.

Model 1. The system consists of m + 1 unreliable machines served by a single repairman. Denote by $Z_t^1$ the number of intact machines at time $t \geq 0$. When the repair of a last defective machine is completed and the total number of intact machines becomes m + 1, the repairman is idle until the next breakdown. Let $\tau_1, \tau_2, \ldots$ be the successive instants of the completion of machine repairs. The repair time of the nth machine is distributed in accordance with $A_{\xi_n}(x) \in \{A_k(x); k = 0, 1, \ldots, m\}$ (a tuple of arbitrary d.f.'s), where $\xi_n := Z_{\tau_n}^1, n = 1, 2, \ldots$. That is, the repair time depends upon the number of intact machines at the moment immediately before the completion of the preceding service. The working machines perform certain jobs. Within the interval $[\tau_n, \tau_{n+1})$ the continuous
On a Duality Principle in Processes of Servicing Machines with Double Control: Dshalalow

Durations of each job are conditionally independent given $\xi_n$ and exponentially distributed with parameter $\mu_k \in \{\mu_k, k=0, 1, \ldots, m + w \} \subset \mathbb{R}_+ \setminus \{0\}$.

MODEL 2. There are a maximum of $m$ working unreliable machines and one repairman. The total number of intact machines at time $t \geq 0$ is denoted by $Z_t$. Unlike Model 1, there are no idle periods. At certain epochs of time the repairman temporarily abandons the system to acquire a new machine. Let $T_n$ be an epoch when a machine is completely repaired. If at this time the total number of intact machines is $m$, the repairman leaves the system and returns to the system at time $T_{n+1}$ with a new machine to replace any machine that has failed during the repairman's absence. In other words, if the repairman returns to fewer than $m$ working machines, then the number of working machines increases by one. If no machine has failed to this time, a replacement still occurs but at the repairman's own choice (in this case, without any effect on the system). In both cases used machines are removed from the system. Consequently, $T_1, T_2, \ldots$ are the moments when the repairman completes a job (repair of a machine or acquisition of new equipment). At time $T_n$ he begins with the repair of a current machine or leaves the system if no defective machine is available. At the time $T_{n+1}$ he completes the repair or returns to the system with a new machine. The length of the interval $[T_n, T_{n+1})$ is distributed according to the $d.f.$

(2.1) $A_{X_n}(x) \in \{A_k(x); k = 0, 1, \ldots, m\}$

where

(2.2) $X_n := Z_{T_n}, \ n = 0, 1, \ldots$

For instance, control can be applied to the period of the repairman's absence distributed differently from the repair time. The assumption about the failure rates is as in Model 1. Namely within the interval $[T_n, T_{n+1})$ the continuous durations of each job are conditionally independent given $X_n$ and exponentially distributed with parameter

(2.3) $\mu_{x_n} \in \{\mu_k, k = 0, 1, \ldots, m + w \} \subset \mathbb{R}_+ \setminus \{0\}$.

As mentioned, this model is identical to the doubly controlled $m$-channel loss queueing system studied by the author [2], where $Z_t$ denotes the number of customers in
the system at time $t$. To explain equivalence between both systems, we use the mutual notation below.

At time $T_n$ a customer departs from a source and at time $T_{n+1}$ arrives at the system. The customer is served by one of the free parallel channels available, or is lost by the system if the servicing facility is busy. The length of the interval $[T_n, T_{n+1})$ is distributed as in (2.1) and (2.2), and the servicing policy is determined by (2.3). Within the interval $[T_n, T_{n+1})$ the service durations of customers in each of the channels are conditionally independent given $X_n$ and exponentially distributed with parameter $\mu_{X_n}$.

In addition to the above formalism, we use the following convention. Throughout the paper, we drop the superscripts in all functionals of $A(x)$ and of $\mu$ if the corresponding indices in $A(x)$ and $\mu$ are missing or if brevity dictates as in the cases of $A_m(x)$ and $\mu_m$.

3. CONNECTION BETWEEN THE MODELS

It can be shown that $\tau_1, \tau_2, \ldots$ is a sequence of stopping times relative to the canonic filtering $\sigma(Z_u^1; u \leq t)$, that $T_1, T_2, \ldots$ is a sequence of stopping times relative to $\sigma(Z_u; u \leq t)$, and that the processes $(\Omega_1, \mathcal{U}_1, (P_x)_{x \in E}, (Z^1_i; t \geq 0)) \rightarrow E = \{0,1,\ldots,m+1\}$ and $(\Omega, \mathcal{U}, (P_x)_{x \in E}, (Z^1_i; t \geq 0)) \rightarrow E = \{0,1,\ldots,m\}$ are semi-regenerative relative to these sequences (cf. Definition in Dshalalow [2, p. 54]). Consequently, $(\Omega_1, \mathcal{U}_1, (P_x)_{x \in E}, (\xi_n; n=1,2,\ldots)) \rightarrow E$ and $(\Omega, \mathcal{U}, (P_x)_{x \in E}, (X_n; n=1,2,\ldots)) \rightarrow E$ are embedded Markov chains (MC). Since the idleness of the repairman in the first model is distributed exponentially, it is easy to see that both MC's are stochastically equivalent and they are obviously ergodic. Let $(\Omega_1, \mathcal{U}_1, (P_x)_{x \in E}, (Y^1_i; t \geq 0)) \rightarrow E_1$ and $(\Omega, \mathcal{U}, (P_x)_{x \in E}, (Y_i; t \geq 0)) \rightarrow E$ be the semi-Markov processes associated with the sequences of stopping times above. Both are ergodic and their limiting probabilities are expressed through the invariant probability measure $P$ of the MC$(X_n)$ or the MC$(\xi_n)$. (Since $(X_n)$ and $(\xi_n)$ are stochastically equivalent, only one of them, say $(X_n)$, will be mentioned further.) Next we need the limiting probabilities

$$y^1_k \equiv \lim_{i \rightarrow \infty} P^x_y \{Y^1_i = k\} = \frac{P_k M_k}{PM}, k \in E$$

(cf. Çinlar [1], p. 342) where $M_k$ can be easily derived as
(3.2) \[ M_k := E[T_k] = \begin{cases} a_k & , k = 0, 1, \ldots, m - 1 \\ a + \frac{1}{\mu(m+1)} & , k = m \end{cases} \]

and \( PM \) is the scalar product of \( P \) and \( M = (M_0, M_1, \ldots, M_m)^T \) which can be expressed by the formula

(3.3) \[ PM = PA + P_m \frac{1}{\mu(m+1)} \]

The duality principle between Models 1 and 2 is based on the following consideration. Let \( \mathcal{S}_n \) and \( \mathcal{F}_n \) denote the \( n \)-th busy period and the idle period following the \( n \)-th busy period, respectively, in Model 1. Let \( \{ \mathcal{S}_t, t \geq 0 \} \) be the counting process associated with the point process \( \{ \mathcal{S}_n, n=1,2,\ldots \} \). It is readily seen that the processes \( Z_t \) and \( Z^1_t \) during their busy periods are stochastically equivalent or formally \( P^* \{ Z_t = k \} = P^* \{ Z^1_t = k \mid \mathcal{F}_t > 0 \} \), \( k \in E \), where the probability \( P^* \{ Z^1_t = k \mid \mathcal{F}_t > 0 \} \) can be expressed as

\[
P^* \{ Z^1_t = k \mid \mathcal{F}_t > 0 \} = \frac{P^* \{ Z^1_t = k, Z^1_t \in \{0, 1, \ldots, m\} \}}{P^* \{ Z^1_t \in \{0, 1, \ldots, m\} \}} \]

\[
= \begin{cases} 0 & , k = m + 1 \\ \frac{P^* \{ Z^1_t = k \}}{1-P^* \{ Z^1_t = m+1 \}} & , k \leq m \end{cases}
\]
Therefore, 

\[(3.4) \quad P^x \{ Z_i^1 = k \} = [1 - P^x \{ Z_i^1 = m + 1 \}] P^x \{ Z_i^1 = k \}, \quad k = 0, 1, \ldots, m. \]

We now find \( P^x \{ Z_i^1 = m + 1 \} \).

\[(3.5) \quad P^x \{ Z_i^1 = m + 1 \} = P^x \{ Y_i^1 = m \} P^x \{ t \leq t' | Y_i^1 = m \} \]

where

\[(3.6) \quad \lim_{t \to \infty} P^x \{ t \leq t' | Y_i^1 = m \} = \frac{1}{1 + a\mu(m+1)}. \]

Let \( \pi_k^1 = \lim_{t \to \infty} P^x \{ Z_i^1 = k \} \). Then from (3.1) - (3.3), (3.5) and (3.6) it follows that

\[(3.7) \quad \pi_{m+1}^1 = \frac{p_m}{PA\mu(m+1) + p_m}. \]

Finally, (3.4) and (3.7) yield

\[(3.8) \quad \pi_k^1 = (1 - \pi_{m+1}^1) \pi_k, \quad k = 0, 1, \ldots, m \]

where \( \pi_k = \lim_{t \to \infty} P^x \{ Z_t = k \} \) was obtained in [2, p.50-51].

4. EXAMPLES

(i) Recall that in case of Model 2 the repairman leaves the system when all machines are intact. However, the repairman may plan to leave for only a short duration. Assume his expected absence is \( a = a_m \leq \frac{1}{\mu} \). Here both \( a \) and \( \mu \) can be adjusted if necessary. For example,

\[ \mu_j = \begin{cases} \mu_0, & j = 0, 1, \ldots, m - 1 \\ \mu, & j = m \end{cases} \]
while $A_j(x)$ is subject to no restriction.

(ii) An undesirable situation occurs if during the repairman's absence the number of working machines falls below the level $m - r$, $r = 0, 1, \ldots, m - 1$. We calculate the probability of this event as

\[ \gamma_r := \lim_{t \to \infty} P^X \{ Y_t = m, Z_t < m - r \} = \sum_{n=0}^{m} \pi_{m,n} - \sum_{n=m-r}^{m} \pi_{m,n} . \]

To derive $\gamma_r$ we observe that the first sum above is

\[ \gamma_r^2 := \lim_{t \to \infty} P^X \{ Y_t = m \} = \frac{a P_m}{P A} \]

(cf. the similar formula (3.1))

Also from [2, (4.9)-(4.11) and (4.19)] one can derive by the summation of the equations

\[ P A n \pi_{jn} = \frac{1}{\mu_j} \sum_{k=0}^{n-1} P_j p_{jk} 1 \leq n \leq \min \{ j + 1, m \}, j \in E. \]

Therefore, from (4.1) - (4.3) and [2, (4.19)] we have

\[ \gamma_r = \frac{P_m}{P A} \left[ a - \frac{1}{m \mu} \sum_{n=m-r}^{m} \sum_{k=0}^{n-1} p_{m,k} \right] . \]

(iii) Now we return to the relation between the models discussed in section 3. Formulas (3.7), (3.8) and [2, (4.11), (4.12) and (4.19)] can be combined to obtain the corresponding stationary probabilities $\pi^1_k$. An elegant expression follows when $\mu_j = \mu$, $j \in E$:

\[ \pi^1_k = \frac{(m+1) P_{k,1}}{k [PA \mu (m+1) + P_m]}, \quad k = 1, \ldots, m + 1 . \]

And finally, when $PA = a$ (i.e. $a_j = a$), (4.5) implies Takács' well-known formula [3, p.198]:
\[(4.6) \quad \pi_k^1 = \frac{(m+1) P_{k-1}}{k[a\mu(m+1)+ P_m]} , \quad k = 1, \ldots, m + 1\]

where the \(d.f.'s\) \(A_j(x)\) are still arbitrary and possibly distinct and thus (4.6) holds under more general conditions.

REFERENCES

[1] Çinlar, E., *Introduction to Stochastic Processes*, Prentice Hall, Englewood Cliffs, N.J. 1975.

[2] Dshalalow, J. On a Multi-channel Transportation Loss System with Controlled Input and Controlled Service, *Journ. Appl. Math. Simulation*, 1, 1, 41 - 55, 1987.

[3] Takács, L., *Introduction to the Theory of Queues*, Oxford University Press, 1962.
