IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 3 Ver. VII (May. - Jun. 2016), PP 08-14
www.iosrjournals.org

Mathematical Modeling and Analysis of Finite Queueing System with Unreliable Single Server

Anupama[1], Prof Anjana Solanki[2]
[1]*Deptt. of Applied Mathematics, Research Scholar, GBU, Greater Noida (U. P.) India,*
[2]*Deptt. of Applied Mathematics, GBU, Greater Noida (U. P.) India,*

Abstract: *In this paper we study the queueing system of unreliable server of production system which causes server breakdowns. The analysis of a finite-capacity N queueing system with a single server where service rule is FIFO has been done. A customer, on arrival to an idle server, is immediately taken for service; or else he joins the tail of the queue. During the service of a customer, one or more breakdowns may occur. It takes a random amount of time to clear the interruption. We consider a queueing system with troubled and non-troubled server breakdowns. Both troubled and non-troubled breakdowns may start when there is a customer in service. The customer repeats its service after a troubled interruption, and continues its service after a non-troubled interruption. Using a Laplace Transform approach, we obtain various performance measures such as service times with troubled and non –troubled interruption, the moments of the queue size and retard times. Special cases also considered in this paper. Numerical illustrations of the system behavior are also provided.*
Keywords: *Queuing systems of finite capacity, server vacations, server breakdowns, FIFO*

I. Introduction

Vacation queues have been investigated for over two decades as a very useful tool for modeling and analyzing computer systems, communication networks, manufacturing and production systems and many others.

Consider a production system that is subject to failures. Failures occur during production or when the system is idle and may either lead to the complete loss of the item that is in production, or to a temporary production halt. In the former case, the production of the “interrupted” item has to start all over. In the latter case, production may continue after the failure is fixed. In queueing theory, periods of temporary service unavailability are referred to as server vacations or server breakdowns. Queueing models with service breakdowns have proved to be a useful abstraction in situations where a service facility is shared by multiple queues, or where the facility is subject to failure. Some queueing problems with the service station subject to breakdowns[1] and the N -policy for an unreliable server with delaying repair and two phases of service [2] is discussed in 1963 and 2009 respectively. From the vantage point of a single queue, the facility leaves for an interruption whenever it attends another queue. When the facility is subject to failures, the breakdowns correspond to the time one needs to fix the failure, or the time one spends on preventive maintenance [7]. According to Ibe and Trivedi [11], White and Christie [5] were the first to study queues with service breakdowns. They consider an M/M/1 queueing system with exponentially distributed breakdowns. Generally distributed service times and breakdowns are considered by Avi-Itzhak and Naor [1] and by Thiruvengadam [8]. The former authors assume that the server remains with the queue for an exponentially distributed amount of time. Federgruen and Green [9] then relax this assumption to phase-type distributed ”on-periods”. Van Dijk [11] provides an approximate analysis of a system with exponentially distributed service times but with generally distributed on- and off periods, whereas Takine and Sengupta [15] and Masuyama and Takine [4] study queueing systems with service breakdowns in a Markov-modulated environment. A processor sharing queueing system with exponentially distributed on-periods and generally distributed off-periods is studied by N´u˜nez Queija [13]. All these authors assume that customers resume service after the interruption. Gaver [3] considers the cases where service is repeated, or repeated and resampled after the interruption. The latter operation mode is also studied by Ibe and Trivedi [12] for a two station polling system. In the present contribution, we consider M/G/1 queueing systems with combined troubled and non-troubled renewal-type breakdowns. We consider two variants of the system. The system with recompiling assumes that the service time after a troubled interruption is resampled. For the system without recompiling, the service time after a troubled interruption equals the original service time. In terms of the abstract production system introduced earlier, the breakdowns correspond to the time that is needed to recover from a failure. The system with recompiling can then capture the uncertainty in the production process,the production time after the failure is not necessarily the same whereas the system without recompiling allows one to consider a system that produces different types of products with unequal production time requirements. The remainder of this contribution is organized as follows. In the following two sections, the queueing model is described in detail and analysed respectively. We then illustrate our approach by means of some numerical examples and draw conclusions.

II. Descriptions Of The Model And Notations

We consider a finite capacity queueing system that adheres a first-in-first-out service discipline. Customers arrive in accordance with a Poisson process with arrival rate λ and their service times constitute a series of independent and identically distributed (i.i.d.) random variables with common density function s(t), $t \geq 0$ and corresponding Laplace-Stieltjes transform (LST) $S(\alpha)$, $\alpha > -R_S$. Here $-R_S$ is the boundary of the region of convergence of the LST.

The customers are served by a single server that takes an interruption from time to time. Breakdowns occur in accordance with a Poisson process with rate v while the server is serving customers and with rate v_i when the server is idle. No new breakdowns start during server breakdowns. When an interruption occurs while a customer receives service, the interruption is either troubled (with probability p_d) or non-troubled (with probability $p_n = 1 - p_d$). The server will continue the interrupted service time if the interruption is non-troubled and will repeat the entire service time if the interruption is troubled. For further use, we also introduce the interruption rates $v_d = vp_d$ and $v_n = vp_n$ of the troubled and the non-troubled breakdowns respectively. The lengths of the consecutive troubled (non-troubled, idle time) breakdowns constitute a series of i.i.d. positive random variables with density function $v_d(t)$ ($v_n(t)$, $v_i(t)$), $t \geq 0$ and corresponding LST $V_d(\alpha)$ ($V_n(\alpha)$, $V_i(\alpha)$) .

Two different queueing models are considered. The model with recompiling assumes that the service time after a troubled interruption is resampled. For the model without recompiling, the service time after a troubled interruption equals the original service time.

The queueing model with a choice of service breakdowns described in this section is based on the following notation:

Parameters:

λ: positive arrival rate of Poisson input process.
p: probability that the server takes a single vacation at a service completion epoch.
v: mean vacation time of the server.

Random variables:

v_d: Interruptions rates for troubled breakdowns
v_n: Interruptions rates for non troubled breakdowns
v_i : Interruptions rates for idle period
β : An independent exponentially distributed random variable

Probability distribution functions (PDF):

$v_d(t)$: PDF of troubled breakdowns
$v_n(t)$: PDF of non troubled breakdowns
$v_i(t)$): PDF of idle period

Laplaces- Stieltjes transforms (LST):

$V_d(\alpha)$: LST of troubled breakdowns
$V_n(\alpha)$: LST of non troubled breakdowns
$V_i(\alpha)$: LST of idle period
$T(\alpha)$: The LST of the effective service time of a random customer

Some important Notations:

T(x) : The effective service time of a customer with service time x
S : The service time of a random customer
V_d : The length of a random troubled interruption
Vn : The length of a random non-troubled interruption
$E[V_d]$: The mean length of a troubled interruptions
$E[V_n]$: The mean length of non-troubled interruption

III. System Analysis

To simplify the analysis, we first consider the effective service times for the models with the special cases. This then allows for a unified queueing analysis afterwards.

Let a customer's effective service time be defined as the amount of time between the epoch where the server starts serving this customer for the first time and the epoch where this customer leaves the system. As such, the effective service time includes all service breakdowns during the customer's service time as well as time that is lost due to service repetitions.

3.1 Special Cases:

3.1.1 The service time after a troubled interruption

We first focus on the model with recompiling. Let T(x) denote the effective service time of a customer with service time x (before being resampled, if this is the case), and let S, V_d and Vn denote the service time of a random customer, the length of a random troubled interruption and the length of a random non-troubled interruption respectively. Further, let β denote an independent exponentially distributed random variable with mean 1/v, and let Q denote an independent random variable that takes the values d and n with probability p_d and p_n respectively. The variable β corresponds to the time between the start of the customer's service and the following interruption, and the variable Q corresponds to the type of this interruption. We then find the following expression for the effective service time

$$T(S) = \begin{cases} S & \text{for } \beta \geq S, \\ \beta + V_d + T^*(S^*) & \text{for} \beta < S \text{ and } Q = d, \\ \beta + V_n + T(S - \beta) & \text{for} \beta < S \text{ and } Q = n. \end{cases}$$

Here T*(·) and S*are independent random variables that are distributed as T(·) and S respectively. Let T(α|x) denote the LST of the effective service time, given that the service time equals x(before being resampled, if this is the case). Further, let T(α) denote the LST of the effective service time of a random customer. In view of the former expression and by conditioning on β, we find that T(α|x) obeys following integral equation:

$$\begin{aligned} T(\alpha\backslash x) &= e^{-(\alpha+v)x} + v_d V_d(\alpha) T(\alpha) \frac{1 - e^{-(\alpha+v)x}}{\alpha + v} + vnVn(\alpha) \int_0^x e^{-(\alpha+v)y} T(\alpha\backslash x - y)\, dy \\ &= e^{-(\alpha+v)x} + v_d V_d(\alpha) T(\alpha) \frac{1 - e^{-(\alpha+v)x}}{\alpha + v} + vnVn(\alpha) e^{-(\alpha+v)x} \int_0^x e^{-(\alpha+v)y} T(\alpha\backslash y)\, dy, \end{aligned} \quad (1)$$

for x ≥ 0. One then easily shows that a solution of the equation (1) has the form,

$$T(\alpha\backslash x) = \frac{v_d V_d(\alpha) T(\alpha)}{\alpha + v - vnVn(\alpha)} + C(\alpha) e^{-(\alpha + v - vnVn(\alpha))x}, \quad (2)$$

where C(α) is an unknown function. Since the effective service time of a zero length service time equals zero, we have T(α|0) = 1. Therefore, plugging x = 0 in equation (2) and solving for C(α) yields,

$$C(\alpha) = 1 - \frac{v_d V_d(\alpha) T(\alpha)}{\alpha + v - vnVn(\alpha)} \quad (3)$$

Combining the equations (2) & (3) and integrating over all possible service times with respect to the service time distribution then yields the following expression for the LST of the effective service times:

$$T(\alpha) = \frac{v_d V_d(\alpha) T(\alpha)}{U(\alpha)} + \left(1 - \frac{v_d V_d(\alpha) T(\alpha)}{U(\alpha)}\right) S(U(\alpha)) \quad (4)$$

Here we introduced $U(\alpha) = \alpha + v - v_n V_n(\alpha)$ to simplify notation. Solving for T(α) finally leads to

$$T(\alpha) = \frac{S(U(\alpha)) U(\alpha)}{U(\alpha) - v_d V_d(\alpha)(1 - S(U(\alpha))} \quad (5)$$

By means of the moment generating property of LSTs, we obtain expressions for the various moments of the effective service times. In particular, the first moment is given by

$$E[T] = \frac{1-S(v_d)}{v_d S(v_d)}\left(1 + v_d E(V_d)\right) + v_n E[V_n]$$
(6)

Where $E[V_d]$ and $E[V_n]$ denote the mean length of a troubled and non-troubled interruption respectively. Similar expressions can be obtained for higher moments. In particular, the k^{th} moment can be expressed in terms of the LST $S(\alpha)$ and its derivatives evaluated in jv_d ($j = 1 \ldots k$), and the moments up to order k of the breakdowns. Notice that for $v_d > 0$ existence of the moments of the effective service time does not require the existence of the moments of the service time.

3.1.2 The service time after a troubled interruption equals the original service time.

For the model without recompiling, we may proceed analogously. It is however more convenient to rely on the results of the model with recompiling. As before, let $T(\alpha/x)$ denote the effective service time of a customer given that the customer's service time equals x. It is then easy to see that under the assumption that the service times are fixed and equal to x, $T(\alpha/x)$ equals the LST for the model with recompiling. This notion immediately leads to the following expression for the conditional LST of the effective service times:

$$T(\alpha/x) = \frac{e^{-U(\alpha)x}U(\alpha)}{U(\alpha) - v_d V_d(\alpha)(1 - e^{-U(\alpha)x})}$$

Integrating over all possible service times with respect to the service time distribution then yields the following expression for the LST of the effective service times for the model without recompiling:

$$T(\alpha) = \int_0^N \frac{s(x)e^{-U(\alpha)x}U(\alpha)}{U(\alpha) - \mathbf{v_d V_d(\alpha)(1 - e^{-U(\alpha)x})}}dx \qquad (7)$$

.

Furthermore, the moment generating property of LSTs allows us to obtain expressions for the various moments of the effective service times. In particular, the first moment is given by

$$E[T] = \frac{s(-v_d) - 1}{v_d}(1 + v_d E[V_d] + v_n E[V_n]), \qquad (8)$$

For $v_d < R_S$. Higher moments can also be expressed in terms of the LST of the service times. In particular, the kth moment can be expressed in terms of the LST $S(\alpha)$ and its derivatives, evaluated in v_d ($j = 1 \ldots k$), and the moments up to order k of the lengths of the troubled and non-troubled breakdowns. Therefore, the kth moment exists whenever the k^{th} moments of the breakdowns exist and whenever

$$\chi\, v_d < Rs \qquad (9)$$

The existence of the k^{th} moments for $\chi v_d = R_S$ depends on the behaviour of the LST $S(\alpha)$ and its derivatives for $\alpha = -R_S$.

The expressions (1) and (7) for the LSTs of the effective service times for the model without and with recompiling respectively now enable us to present a unified queueing analysis of both models.

IV. Queue Size

Let U_k denote the number of packets in the queue upon departure of the kth customer. We then find,

$$U_{k+1} = \begin{cases} U_k + N_{k+1} - 1 & \text{for} \quad U_k > 0, \\ \Gamma_{k+1} + N_{k+1} & \text{for} \quad U_k = 0. \end{cases}$$

Here N_k denotes the number of customer arrivals during the kth customer's effective service time and Γ_k denotes the number of arrivals during the "remaining interruption time" upon arrival of the kth customer. The remaining interruption time of the kth customer is defined as the time between the arrival instant and the end of the interruption if service is interrupted when the customer arrives or equals zero if the server is available. Let $U_k(z)$ denote the probability generating function of U_k. In view of the former expression, we find

$$U_{k+1}(z) = \frac{1}{z}\big(U_k(z) - U_k(0)\big)T\big(\lambda(1-z)\big) + U_k(0)\eta\big(\lambda(1-z)\big)T\big(\lambda(1-z)\big),$$

where $\eta(\alpha)$ is the LST of the remaining interruption time upon arrival of a random customer that arrives in an empty system, and where $T(\alpha)$ is given by (1) or (7) depending on the model under consideration.
One can show that the queueing system under consideration reaches steady state whenever λ E[T] < 1 (where E[T] is given by either (1) or (7). Let $U(z) = \lim_{k\to\infty} U_k(z)$ denote the probability generating function of the queue content at departure epochs in steady state. From the former expression and the normalization condition U(1) = 1, we obtain

$$U(z) = \frac{1-\lambda E[T]}{1+\lambda E[\eta]}\frac{\eta\big(\lambda(1-z)\big)-1}{z-T\big(\lambda(1-z)\big)}T\big(\lambda(1-z)\big), \tag{10}$$

With $E[\eta]$ the average remaining interruption time.

We now determine the LST $\eta(\alpha)$. Let Γ_A and $\Gamma_V$denote exponentially distributed random variables with mean $1/v_i$ and $1/\lambda$ respectively. The former variable corresponds to the amount of time till the next server interruption after the server becomes idle. The latter random variable corresponds to the arrival time of the first arrival after the epoch where the server becomes idle. We may now decompose the remaining interruption time as follows,

$$\eta = \begin{cases} 0 & \text{for } \Gamma_A < \Gamma_V, \\ \Gamma_V + V - \Gamma_A & \text{for } \Gamma_V \le \Gamma_A < \Gamma_V + V, \\ \eta^* & \text{for } \Gamma_A > \Gamma_V + V \end{cases} \tag{11}$$

Here η^* is an independent copy of η . In view of the former expressions, the LST of η satisfies

$$\eta(\alpha) = \frac{\lambda}{\lambda+v_i} + v_i\lambda\frac{V_i(\alpha)-V_i(\lambda)}{(\lambda+v_i)(\lambda-\alpha)} + \eta(\alpha)\frac{v_iV_i(\lambda)}{\lambda+v_i} \tag{12}$$

Solving for $\eta(\alpha)$ leads to

$$\eta(\alpha) = \frac{\lambda(\lambda-\alpha)+v_i\lambda(V_i(\alpha)-V_i(\lambda))}{\big(\lambda+v_i-v_iV_i(\lambda)\big)(\lambda-\alpha)} \tag{13}$$

The moment-generating property of LSTs further yields

$$E[\eta] = \frac{v_i}{\lambda}\frac{\lambda E[V_i]-\big(1-V_i(\lambda)\big)}{\lambda+v_i\big(1-V_i(\lambda)\big)} \tag{14}$$

where $E[V_i]$ denotes the mean length of an interruption that starts when the server is idle.

V. Retard

Let customer retard be defined as the time between a customer's arrival and departure instant and let $D(\alpha)$ denote the LST of a random customer's delay. Since the queue content upon departure of a customer equals the number of arrivals that arrived during this customer's delay, we find

$$U(z) = D(\lambda(1-z)) \tag{15}$$

or equivalently,

$$D(\alpha) = U\left(1-\frac{\alpha}{\lambda}\right) = \frac{1-\lambda E[t]}{1+v_iE[V_i]} \tag{16}$$

The mean customer delay is given by

$$E[D] = E[T] + \frac{\lambda E[T^2]}{2(1 - \lambda E[T])} \tag{17}$$

VI. Numerical Examples

In this section, we present numerical example for the parameters derived from above analysis. In particular case customer service times has gamma distributions and length of the different types of the breakdowns are exponentially distributed. The gamma distribution is completely specified by the mean μ and standard deviation σ, and has the following LST:

$$\gamma(\alpha) = \left(1 + \alpha\frac{\sigma^2}{\mu}\right)^{-\frac{\mu^2}{\sigma^2}}$$

When $\mu = \sigma$ Gamma distribution becomes exponential distribution

Now the fraction φ indicated the server is available while there are customers in service and the mean length of the sum of an available and interruption period χ,

$$\varphi = \frac{\frac{1}{v}}{\frac{1}{v} + p_n E[V_n] + p_d E[V_d]}, \quad \chi = \frac{1}{v} + p_n E[V_n] + p_d E[V_d]$$

This also referred as the time scale parameter of the interruption process (while customers receive Service).

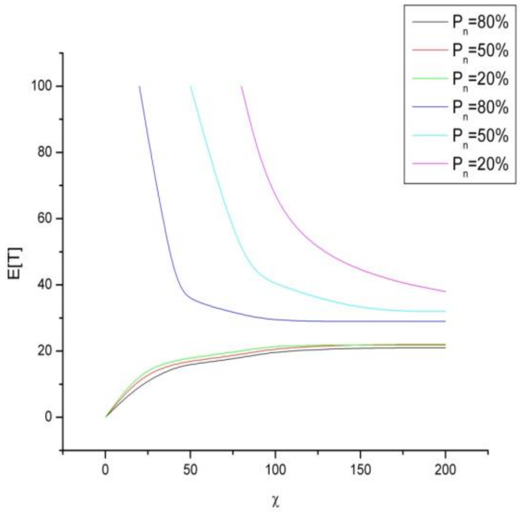

Figure (a) [Mean]

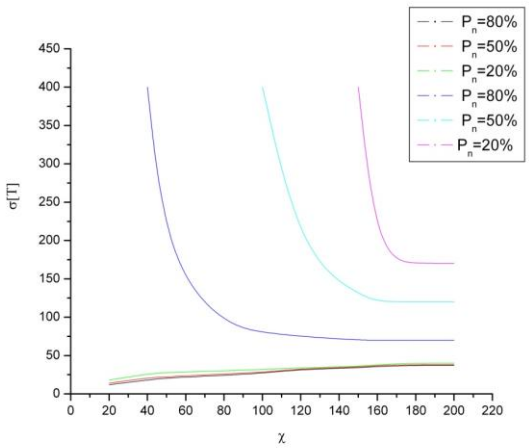

Figure (b) [Standard deviation]

In Figures (a) and (b), the mean μ and standard deviation σ of the effective service times are depicted versus the time scale parameter χ of the interruption process. Non-troubled and troubled breakdowns share the same (exponential) distribution and the server is available during φ = 80% of the time. The mean and standard deviation of the customer service times are equal to E[S] = 20 and σ[S] = 40 respectively. Different values of the probability p_n that an interruption is non-troubled are assumed, and we consider both the systems with (lower curves) and without recompiling (upper curves).

For increasing values of χ, one observes that all curves converge. Large χ implies few but long breakdowns. Whenever there is such an interruption, neither recompiling nor the troubleness of the interruption comes in to play compared to the length of the interruption. For decreasing values of χ, performance of the systems with and without recompiling is completely different. For the system with recompiling, all curves converge to 0. This comes from the fact that low χ implies many short breakdowns such that service times get resampled until they are very small. For the system without recompiling, the service times remain the same after an interruption. Since decreasing χ implies that sufficiently long available periods become less probable, the mean and standard deviation of the effective service times converge to ∞. Finally notice that the vertical asymptotes for the system without recompiling come from the condition (9) for the existence of the moments

VII. Conclusions

We considered a queueing system with troubled and non-troubled breakdown. After a troubled breakdowns, service of the interrupted customer is either repeated or repeated and resampled. By means of a Laplace – stieltjes approach, we obtained performance measures such as the moments of queue content and customer Retard. We finally illustrated our approach by means of some numerical examples.

References

[1]. Avi-ltzhak, B. and Naor, P. (1963). "Some queueing problems with the service station subject to breakdowns". *Operations Research*, Vol 11, pp: 303-320.

[2]. Choudhury, G., Ke, J. C. and Tadj, L. (2009). "The N -policy for an unreliable server with delaying repair and two phases of service". *Journal of Computational and Applied Mathematics*, Vol 231, pp: 349-364.

[3]. D. Gaver Jr. (1962). "A waiting line with interrupted service, including priorities". *Journal of the Royal Statistical Society*, Vol. B24, pp: 73–90.

[4]. H. Masuyama and T. Takine. (2003). "Stationary queue length in a FIFO single server queue with service interruptions and multiple batch Markovian arrival streams". *Journal of the Operations Research Society of Japan*, Vol. 46(3), pp: 319–341.

[5]. H. White and L. Christie. (1958). "Queuing with preemptive priorities or with breakdown". *Operation Research,* Vol. 6(1), pp: 79–95.

[6]. Ke, J. C. (2005). "Modified *T* vacation policy for an *M/G/*1 queueing system with an un-reliable server and startup". *Mathematical and Computer Modelling*, Vol. 41, pp: 1267-1277.

[7]. Ke, J. C. (2006). "On *M/G/*1 system under *NT* policies with breakdowns, startup and closedown". *Applied Mathematical Modelling,* Vol. 30, PP: 49-66.

[8]. K. Thiruvengadam. (1963). "Queuing with breakdowns". *Operations Research*, Vol. 11(1), pp: 62–71.

[9]. A. Federgruen and L. Green. (1986). "Queueing systems with service interruptions". *Operations Research*, Vol. 34(5), pp: 752–768.

[10]. Li, W., Shi, D. and Chao, X. (1997). "Reliability analysis of *M/G/*1 queueing system with server breakdowns and vacations". *Journal of Applied Probability*, Vol. 34, pp: 546-555.

[11]. N. Van Dijk. (1988). "Simple bounds for queueing systems with breakdowns". *Performance Evaluation,* Vol. 8(2), pp: 117–128.

[12]. O. Ibe and K. Trivedi. (1990). "Two queues with alternating service and server breakdown". *Queueing Systems,* Vol. 7(3-4), pp: 253–268.

[13]. R. N´u˜nez Queija. (2000). "Sojourn times in a processor sharing queue with service interruptions". *Queueing Systems*, Vol. 34(1–4), pp: 351–386.

[14]. R. Cooper. (1981). "Introduction to queueing theory". *Elsevier North-Holland, Inc*., New York.

[15]. T. Takine and B. Sengupta. (1997). "A single server queue with service interruptions". *Queueing Systems*, Vol. 26, pp: 285–300.

[16]. T. Takine, (2001) "Queue length distribution in a FIFO single-server queue with multiple arrival streams having different service time distributions". *Queueing Systems*, Vol. 39, pp: 349–375.

[17]. Wang, K. H., Wang, T. Y. and Pearn, W. L. (2007). "Optimal control of *N*-policy *M/G/*1 queueing system with server breakdown and general startup times". *Applied Mathematical Modelling*, Vol. 31, pp: 2199-2212.

[18]. Wang, T. Y., Wang, K. H. and Pearn, W. L. (2009). "Optimization of the T-policy *M/G/*1 queue with server breakdowns and general startup times". *Journal of Computational and Applied Mathematics*, Vol. 228, pp: 270-278.