RAIRO-Oper. Res. 48 (2014) 399–413
DOI: 10.1051/ro/2014014

RAIRO Operations Research
www.rairo-ro.org

THE $M^X/M/1$ QUEUE WITH WORKING BREAKDOWN

Zaiming Liu and Yang Song[1]

Abstract. In this paper, we consider a batch arrival $M^X/M/1$ queue model with working breakdown. The server may be subject to a service breakdown when it is busy, rather than completely stoping service, it will decrease its service rate. For this model, we analyze a two-dimensional Markov chain and give its quasi upper triangle transition probability matrix. Under the system stability condition, we derive the probability generating function (PGF) of the stationary queue length, and then obtain its stochastic decomposition, which shows the relationship with that of the classical $M^X/M/1$ queue without vacation or breakdown. Besides, we also give the Laplace-Stieltjes transform (LST) of the stationary waiting time distribution of an arbitrary customer in a batch. Finally, some numerical examples are given to illustrate the effect of the parameters on the system performance measures.

Keywords. $M^X/M/1$ queue, working breakdown, probability generating function (PGF), Laplace–Stieltjes transform (LST), waiting time distribution, stochastic decomposition.

Mathematics Subject Classification. 68M20, 60K20, 90B22.

1. Introduction

The queueing models with an unreliable server have been studied extensively, and in real world, there can not exist any perfect server which always operates as a new one. The machine in manufacturing system is a good example for that. Therefore, in many papers, the server is assumed to be subject to an unpredictable breakdown and be sent to repair immediately. The server is not available during the repair time.

Received May 15, 2013. Accepted February 24, 2014.

[1] School of Mathematics and Statistics, Central South University, Changsha 410075, P.R. China. yy_0605xx@126.com

Article published by EDP Sciences

© EDP Sciences, ROADEF, SMAI 2014

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

Jayaraman *et al.* [5] have considered a queueing system with general bulk service that alternates stochastically between operational state and failed state. Gray *et al.* [3] have discussed a queueing model with multiple types of server breakdowns, for which each type of breakdown requires a finite random number of stages of repair. Furthermore, working vacations queueing model with multiple types of server breakdowns has been considered in Jain [6]. And a batch arrival queue with randomized vacation and unreliable server has given by Ke and Huang [7]. Besides these, the unreliable server is related with the retrial queue as well. Falin [4] has discussed a $M/M/1$ retrial queue with retrials due to server failures. Wang and Cao [14] have given an analysis for reliability of the retrial queue with server breakdowns and repairs. A discrete-time retrial queue with unreliable server could be found in Atencia and Moreno [1], Wang and Zhang [13].

Anyway, the unreliable server is always supposed to stop working completely when it is breakdown. However, there are also exceptions. Sridharan and Jayashree [11] have studied a finite queue with normal, partial and total failures. And recently, Kalidass and Kasturi [8] proposed the concept of working breakdown, which means that the service may not be interrupted when the server is breakdown. This concept does make sense in real life. For example, when a computer is infected with a virus, it may still be able to perform at a lower rate. And the machine replacement problem provides another example. When the failed machine is sent to repair, a backup one will work immediately at a lower service rate. It is worth to note that there is some difference between the working breakdown and the working vacation. A working breakdown occurs randomly, no matter how many customers are waiting in the queue, while a working vacation occurs only when the system is empty. The working vacation policy is brought by Servi and Finn [10], and there comes out plenty of researches since then. Liu *et al.* [9] have presented some stochastic decompositions in $M/M/1$ queue with working vacations. Baba [2] and Xu *et al.* [16] have studied the batch arrival queue with working vacations. In our paper, we generalize the work of Kalidass and Kasturi [8].

The rest of this paper is organized as follows. Firstly, we give a model description in Section 2. Based on the system stability condition, we deduce the PGF of the stationary queue length and its stochastic decomposition in Section 3, then the expected queue length and some probability formulas are obtained. In Section 4, according to the properties of the conditional Erlang distribution, we get the LST of the stationary waiting time distribution of an arbitrary customer in a batch. Then in Section 5, some numerical examples are given to show the effect of parameters on the system performance. And we conclude this paper in the last section.

2. Model descriptions

We consider a batch arrival $M^X/M/1$ queue with working breakdown. Customers arrive at the system according to a Poisson process with rate λ. Batch sizes are independent and identically distributed with the probability distribution

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

function (pdf) $\{g_i\}_{i=1}^{\infty}$, the probability generating function (PGF) $G(z) = \sum_{i=1}^{\infty} g_i z^k$, the expectation g and the second order moment $g^{(2)}$. A working breakdown only occurs when the server is busy and in the normal state. In the normal working period, the service time S follows an exponential distribution with parameter μ. When the server is subject to a random breakdown, the service rate decreases to ν at once, that is, the service time $S^{'}$ follows an exponential distribution with parameter ν when the server is defective. The life time T and the repair time V of the server are assumed to be exponentially distributed with parameters α and θ, respectively. The processes above mentioned are mutually independent, and the service discipline is first come first served (FCFS).

Let $Q(t)$ be the number of customers in the system and $J(t)$ the state of server at time t. $J(t) = 0$ represents that the server is in a normal state, while $J(t) = 1$ represents that the server is defective. Then $X(t) = \{Q(t), J(t)\}$ is a continuous time Markov process with the state space $\Omega = \{(k, j); k \geq 0, j = 0, 1\}$. Using the lexicographical order for the states, the infinitesimal generator of the process $\{Q(t), J(t)\}$ can be written as the Block–Jacobi matrix

$$\widetilde{Q} = \begin{pmatrix} B_0 & A_1 & A_2 & A_3 \dots \\ C & A_0 & A_1 & A_2 \dots \\ & C & A_0 & A_1 \dots \\ & & C & A_0 \dots \\ & & \vdots & \vdots \quad \ddots \end{pmatrix},$$

where

$$B_0 = \begin{pmatrix} -\lambda & 0 \\ \theta & -(\lambda+\theta) \end{pmatrix}, \qquad A_i = \begin{pmatrix} \lambda g_i & 0 \\ 0 & \lambda g_i \end{pmatrix}, i \geq 1,$$

$$C = \begin{pmatrix} \mu & 0 \\ 0 & \nu \end{pmatrix}, \qquad A_0 = \begin{pmatrix} -(\lambda+\mu+\alpha) & \alpha \\ \theta & -(\lambda+\nu+\theta) \end{pmatrix}.$$

3. Stability condition and stationary queue length distribution

At first, we give the sufficient condition for ergodicity of Markov process $X(t)$ with discrete state space Ω, which is the stability condition for the system in the following Theorem 3.1. We finish its proof based on the Statement 1 in Falin [4].

Theorem 3.1. *The queueing system is stable if* $\rho = \frac{\lambda g(\alpha+\theta)}{\alpha\nu+\theta\mu} < 1$.

Proof. Let $\{t_r; r \in Z_+\}$ be a sequence of epochs at which an event occurs, the event could be a batch arrival, a departure, a working breakdown or the system becoming

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

normal after repair. $\{Q(t_r^+), J(t_r^+)\}$ is an embedded (at state transitions) Markov chain for $X(t)$. We adopt Foster's criterion here, which states that an irreducible and aperiodic Markov chain χ_k with state space Ω is ergodic if there exists a non-negative function $f(s), s \in \Omega$, called Lyapunov function, and $\varepsilon > 0$ such that the mean drift

$$x_s = E[f(\chi_{k+1}) - f(\chi_k)|\chi_k = s]$$

is finite for all $s \in \Omega$ and $x_s \leq -\varepsilon$ for all $s \in \Omega$ except perhaps a finite number. Here we take the same Lyapunov function as in [8]

$$f(s) = \begin{cases} k + B, & \text{if } s = (k, 0) \\ k + B^{'}, & \text{if } s = (k, 1) \end{cases}$$

where B and $B^{'}$ denote the constants which will be determined. For $k \geq 0$, $s \in \Omega$,

$$x_s = \begin{cases} \frac{\lambda g - \mu + \left(B^{'} - B\right)\alpha}{\lambda + \nu + \alpha}, & \text{if } s = (k, 0) \\ \frac{\lambda g - \nu + \left(B - B^{'}\right)\theta}{\lambda + \nu + \theta}, & \text{if } s = (k, 1). \end{cases}$$

The constants B and $B^{'}$ should satisfy the following inequations

$$\lambda g - \mu + \left(B^{'} - B\right)\alpha < 0$$
$$\lambda g - \nu + \left(B - B^{'}\right)\theta < 0.$$

Therefore, $\frac{\lambda g - \nu}{\theta} < B^{'} - B < \frac{\mu - \lambda g}{\alpha}$, and $\frac{\lambda g - \nu}{\theta} < \frac{\mu - \lambda g}{\alpha}$ leads to $\frac{\lambda g(\alpha+\theta)}{\alpha\nu + \theta\mu} < 1$. Denote $\rho = \frac{\lambda g(\alpha+\theta)}{\alpha\nu+\theta\mu}$. So if $\rho < 1$, the Markov process $X(t)$ is ergodic and the system is stable. □

Next we consider the stationary distribution for {Q(t),J(t)}. Assume that

$$\pi_k = (\pi_{k0}, \pi_{k1}),\ k \geq 0;$$
$$\pi_{kj} = \lim_{t\to\infty} P\{Q(t) = k, J(t) = j\},\ (k, j) \in \Omega,$$

then the equilibrium equations could be written as follows:

$$\lambda\pi_{00} = \theta\pi_{01} + \mu\pi_{10}; \tag{3.1}$$

$$(\lambda + \mu + \alpha)\pi_{k0} = \lambda\sum_{i=1}^{k} g_i\pi_{k-i,0} + \theta\pi_{k1} + \mu\pi_{k+1,0}, \quad k \geq 1; \tag{3.2}$$

$$\nu\pi_{11} = (\lambda + \theta)\pi_{01}; \tag{3.3}$$

$$(\lambda + \nu + \theta)\pi_{k1} = \lambda\sum_{i=1}^{k} g_i\pi_{k-i,1} + \alpha\pi_{k0} + \nu\pi_{k+1,1}, \quad k \geq 1. \tag{3.4}$$

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

Denote

$$Q_0(z) = \sum_{k=1}^{\infty} \pi_{k0} z^k, \quad Q_1(z) = \sum_{k=0}^{\infty} \pi_{k1} z^k, \quad |z| \leq 1,$$

then the probability generating function (PGF) of stationary queue length $L(z)$ could be written as

$$L(z) = \pi_{00} + Q_0(z) + Q_1(z).$$

Multiplying (3.1) and (3.2) by appropriate powers of z and summing over $k \geq 0$, we obtain

$$[\lambda z\,(1 - G(z)) + \alpha z + \mu(z-1)]\,Q_0(z) - \theta z Q_1(z) = \lambda z\,(G(z) - 1)\,\pi_{00}; \tag{3.5}$$

the same way to (3.3) and (3.4) leads to

$$-\alpha z Q_0(z) + [\lambda z\,(1 - G(z)) + \theta z + \nu(z-1)]\,Q_1(z) = \nu(z-1)\pi_{01}. \tag{3.6}$$

Solving the above equations (3.5) and (3.6), we get

$$Q_0(z) = \frac{\lambda z\,(G(z) - 1)\,[\lambda z\,(1 - G(z)) + \theta z + \nu(z-1)]\,\pi_{00} + \nu\theta z(z-1)\pi_{01}}{[\lambda z\,(1 - G(z)) + \alpha z + \mu(z-1)]\,[\lambda z\,(1 - G(z)) + \theta z + \nu(z-1)] - \alpha\theta z^2}, \tag{3.7}$$

$$Q_1(z) = \frac{\lambda\alpha z^2\,(G(z) - 1)\,\pi_{00} + \nu(z-1)\,[\lambda z\,(1 - G(z)) + \alpha z + \mu(z-1)]\,\pi_{01}}{[\lambda z\,(1 - G(z)) + \alpha z + \mu(z-1)]\,[\lambda z\,(1 - G(z)) + \theta z + \nu(z-1)] - \alpha\theta z^2}. \tag{3.8}$$

Take the denominator of $Q_0(z)$ and $Q_1(z)$ as $f(z)$, then

$$f(z) = \xi(z)\,[\mu(1-z) - \lambda z\,(1 - G(z))] - \alpha z\,[\nu(1-z) - \lambda z\,(1 - G(z))]\,,$$

where $\xi(z) = \nu(1-z) - \lambda z\,(1 - G(z)) - \theta z$ is a convex function and has the unique root $z = \tilde{z}$ in $(0, 1)$. Since

$$\begin{aligned}
f(0) &= \mu\nu > 0,\\
f(\tilde{z}) &= -\alpha\theta z^2 < 0,\\
f(1) &= 0,\\
f^{'}(0) &= -\lambda\mu - \lambda\nu - \theta\mu - \alpha\nu - 2\mu\nu < 0,\\
f^{'}(1) &= -\lambda(\alpha + \theta)g + \theta\mu + \alpha\nu > 0,
\end{aligned}$$

and a unique stationary distribution exists due to the ergodicity of Markov process $X(t)$, we suppose γ be the unique root of $f(z)$ which lies in $(0, 1)$, so the numerator of $Q_0(z)$ will vanish at $z = \gamma$. Substituting $z = \gamma$ into the numerator of the right side of (3.7), we get

$$\pi_{01} = \frac{\lambda\,(1 - G(\gamma))\,\xi(\gamma)}{\nu\theta(1-\gamma)}\pi_{00}.$$

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

Using the normalizing condition $\lim_{z\to 1} L(z) = 1$, we have

$$\pi_{00} = \frac{\theta(1-\rho)}{\theta + \frac{\rho}{g}\frac{1-G(\gamma)}{1-\gamma}\xi(\gamma)}.$$

Therefore,

$$\begin{aligned} L(z) &= \pi_{00} + Q_0(z) + Q_1(z) = \pi_{00} \\ &\quad + \frac{\lambda z\,(1-G(z))\,[\nu(1-z)-\sigma(z)] + \frac{\lambda(1-G(\gamma))\xi(\gamma)}{\theta(1-\gamma)}(1-z)\,[\mu(1-z)-\sigma(z)]}{f(z)}\pi_{00} \\ &= \frac{f(z) + \lambda z\,(1-G(z))\,[\nu(1-z)-\sigma(z)] + \frac{\lambda(1-G(\gamma))\xi(\gamma)}{\theta(1-\gamma)}(1-z)\,[\mu(1-z)-\sigma(z)]}{f(z)} \\ &\quad \times \frac{\theta g(1-\rho)}{\theta g + \rho\xi(\gamma)\frac{1-G(\gamma)}{1-\gamma}}, \end{aligned} \tag{3.9}$$

where $\sigma(z) = \alpha z + \lambda z(1-G(z)) + \theta z$ and $\nu(1-z) - \sigma(z) = \xi(z) - \alpha z$. As we already know that the PGF of the classical $M^X/M/1$ queue is

$$L_0(z) = \frac{(\lambda g - \mu)(1-z)}{(\lambda+\mu)z - \mu - \lambda z G(z)}, \tag{3.10}$$

then the PGF of stationary queue length $L(z)$ can be decomposed further as follows:

$$\begin{aligned} L(z) = L_0(z)\frac{g(1-\rho)}{\mu-\lambda g}&\frac{(\lambda+\mu)z-\mu-\lambda z G(z)}{\xi(z)\,[(\lambda+\mu)z-\mu-\lambda z G(z)] + \alpha z\,[\xi(z)+\theta z]} \\ &\times \frac{\theta(1-\gamma)\,(\mu\xi(z)-\alpha\nu z) + \lambda\xi(\gamma)\,(1-G(\gamma))\,[\mu(1-z)-\sigma(z)]}{(1-\gamma)\theta g + \rho(1-G(\gamma))\xi(\gamma)}. \end{aligned} \tag{3.11}$$

Let

$$\begin{aligned} L_d(z) = \frac{g(1-\rho)}{\mu-\lambda g}&\frac{(\lambda+\mu)z-\mu-\lambda z G(z)}{\xi(z)\,[(\lambda+\mu)z-\mu-\lambda z G(z)] + \alpha z\,[\xi(z)+\theta z]} \\ &\times \frac{\theta(1-\gamma)\,(\mu\xi(z)-\alpha\nu z) + \lambda\xi(\gamma)\,(1-G(\gamma))\,[\mu(1-z)-\sigma(z)]}{(1-\gamma)\theta g + \rho(1-G(\gamma))\xi(\gamma)}, \end{aligned} \tag{3.12}$$

and note that $\lim_{z\to 1^-} L_d(z) = 1$, which means $L_d(z)$ is also a PGF. Rewrite (3.11) as

$$L(z) = L_0(z) \times L_d(z).$$

Now we summarize these in the following stochastic decomposition theorem.

Theorem 3.2. *If $\rho < 1$, the stationary queue length L can be decomposed into the sum of two independent random variables*

$$L = L_0 + L_d,$$

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

where L_0 is the stationary queue length of the classical $M^X/M/1$ queue model and has the PGF $L_0(z)$, L_d is an additional queue length introduced by the server working breakdown and has the PGF $L_d(z)$. The expressions of $L_0(z)$ and $L_d(z)$ are given by (3.10) (3.12), respectively.

Corollary 3.3. *The expectation of the stationary queue length is given by*

$$E(L) = E(L_0) + E(L_d),$$

where

$$E(L_0) = \frac{\lambda(g + g^{(2)})}{2(\mu - \lambda g)}, \tag{3.13}$$

and

$$\begin{aligned} E(L_d) = {} & \frac{g(1-\rho)}{\mu - \lambda g} \frac{1}{(1-\gamma)\theta g + \rho(1-G(\gamma))\xi(\gamma)} \\ & \times \left\{ \frac{\lambda\left(g + g^{(2)}\right)\left[\theta(1-\gamma)(\alpha\nu + \theta\mu) + \lambda\xi(\gamma)(1-G(\gamma))(\alpha+\theta)\right]}{2(\alpha\nu + \theta\mu)(\rho - 1)} \right. \\ & + \frac{(\mu - \lambda g)\left[\theta(1-\gamma)(\lambda\mu g - \theta\mu - \alpha\nu - \mu\nu) + \lambda\xi(\gamma)(1-G(\gamma))(\lambda g - \alpha - \theta - \mu)\right]}{(\alpha\nu + \theta\mu)(\rho - 1)} \\ & + \left[2(\lambda g - \nu - \theta)(\mu - \lambda g) + 2\alpha(\lambda g - \nu) + \lambda(\alpha + \theta)\left(g + g^{(2)}\right)\right] \\ & \left. \times \frac{(\mu - \lambda g)\left[\theta(1-\gamma)(\alpha\nu + \theta\mu) + \lambda\xi(\gamma)(1-G(\gamma))(\alpha+\theta)\right]}{2(\alpha\nu + \theta\mu)^2(\rho-1)^2} \right\}. \end{aligned} \tag{3.14}$$

Proof. We obtain the expectation of the stationary queue length by differentiating (3.10) (3.12) separately, and then substituting $z = 1$. Especially, $E(L_d) = \lim_{z \to 1^-} L^{'}(z)$, after using the L'Hospital's rule, we get (3.14). □

Corollary 3.4. *The probability that the server is normal and idle is given by π_{00}, the probability that the server is defective and idle is given by π_{01}. Let P_n be the probability of the server in normal and busy and P_d the probability of the server being defective and busy, then*

$$P_n = Q_0(1) = \frac{\lambda\theta g}{\alpha\nu + \theta\mu} \frac{\theta g(1-\gamma) + \xi(\gamma)\,(1 - G(\gamma))}{\theta g(1-\gamma) + \rho\xi(\gamma)\,(1 - G(\gamma))},$$

$$P_d = Q_1(1) - \pi_{01} = \frac{\lambda g}{\nu(\alpha\nu + \theta\mu)} \frac{\alpha\nu\theta g(1-\gamma) + \xi(\gamma)\,(1 - G(\gamma))\,[\lambda g(\alpha + \theta) - \theta\mu]}{\theta g(1-\gamma) + \rho\xi(\gamma)\,(1 - G(\gamma))}.$$

Remark 3.5. For each arrival, if the batch size is one, *i.e.*, $P(X = 1) = 1$, the above results are accordant with the known results in Kalidas and Kasturi [8].

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

4. Stationary waiting time analysis

In this section, we will obtain the Laplace-Stieltjes transform (LST) of the stationary waiting time of an arbitrary customer in a batch. Denote $W^*(s)$ as the LST of the waiting time with distribution W, $B^*(s)$ and $C^*(s)$ as the LST of the service time in the normal and defective period, respectively, where $B^*(s) = \frac{\mu}{s+\mu}$ and $C^*(s) = \frac{\nu}{s+\nu}$. Firstly, we list several properties of conditional Erlang distribution, which will be used for calculating $W^*(s)$. Let $\{S_i\}_{i=1}^{\infty}$ be a sequence of service times when the server is normal, and denote $S^{(h)} = \sum_{i=1}^{h} S_i$, $h \geq 0$, especially $S^{(0)} = 0$, then $S^{(h)}$ follows an Erlang distribution with parameters h and μ, and its LST is $(\frac{\mu}{s+\mu})^h$. Another independent random variable T follows an exponential distribution with parameter α, we give the following lemma and the more details can be found in Xu, Zhang [15] and Xu *et al.* [16].

Lemma 4.1. (i) $P\left(S^{(h)} < T < S^{(h+1)}\right) = \frac{\alpha}{\alpha+\mu}\left(\frac{\mu}{\alpha+\mu}\right)^h$, *and* $P\left(T > S^{(h)}\right) = \left(\frac{\mu}{\alpha+\mu}\right)^h$ *for* $h \geq 0$;

(ii) *if* $S^{(h)} < T < S^{(h+1)}$, $h \geq 0$, $\left(T|S^{(h)} < T < S^{(h+1)}\right)$ *follows an Erlang distribution with parameters* $h+1$ *and* $\alpha+\mu$, *and its LST is* $\left(\frac{\alpha+\mu}{s+\alpha+\mu}\right)^{h+1}$;

(iii) *if* $S^{(h)} < T$, $h \geq 1$, $\left(S^{(h)}|S^{(h)} < T\right)$ *follows an Erlang distribution with parameters* h *and* $\alpha+\mu$, *and its LST is* $\left(\frac{\alpha+\mu}{s+\alpha+\mu}\right)^h$.

Theorem 4.2. *If the queue system is stable, the LST of the stationary waiting time of an arbitrary customer in a batch could be written as*

$$\begin{aligned} W^*(s) = {} & \frac{1-G\left(B^*(s+\alpha)\right)}{g\left[1-B^*(s+\alpha)\right]}\left[\pi_{00}+Q_0\left(B^*(s+\alpha)\right)\right] \\ & + \frac{1-G\left(C^*(s+\theta)\right)}{g\left[1-C^*(s+\theta)\right]}Q_1\left(C^*(s+\theta)\right) \\ & + \frac{\alpha\nu}{g\left[\alpha\nu+(\nu-\mu)s\right]}\frac{\left[\pi_{00}+Q_0\left(C^*(s)\right)\right]\left[1-G\left(C^*(s)\right)\right]}{1-C^*(s)} \\ & - \frac{\alpha\nu}{g\left[\alpha\nu+(\nu-\mu)s\right]}\frac{\left[\pi_{00}+Q_0\left(B^*(s+\alpha)\right)\right]\left[1-G\left(B^*(s+\alpha)\right]\right.}{1-B^*(s+\alpha)} \\ & + \frac{\theta\mu}{g\left[\theta\mu+(\mu-\nu)s\right]}\frac{Q_1\left(B^*(s)\right)\left[1-G\left(B^*(s)\right)\right]}{1-B^*(s)} \\ & - \frac{\theta\mu}{g\left[\theta\mu+(\mu-\nu)s\right]}\frac{Q_1\left(C^*(s+\theta)\right)\left[1-G\left(C^*(s+\theta)\right)\right]}{1-C^*(s+\theta)}. \end{aligned}$$

Proof. We divide $W^*(s)$ into two parts and calculate them separately.

• Case 1: If a batch of customers arrives in the state $(k,0)$, and the jth customer in the batch is the tagged customer, then the waiting time of the tagged customer equals to the total customers' service times ahead of it. Obviously, there are k

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

customers out of its batch and $j-1$ customers in its batch. Let Y be the order of the position of a randomly selected tagged customer in a batch and denote $r_j = P(Y=j)$, then from Sections 2–5 in Wollf [12], we know that $r_j = \frac{\sum_{i\geq j} g_i}{E(X)} = \frac{1}{g}\sum_{i=j}^{\infty} g_i$. Especially, we denote $\tilde{S}^{(h)}$ as the sum of service times of h customers. Then

$$\begin{aligned} W_{k0}^*(s) &= \sum_{j=1}^{\infty} r_j \left[\int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(\tilde{S}^{(k+j-1)} < x\right)\right] \\ &= \sum_{j=1}^{\infty} r_j \int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(\tilde{S}^{(k+j-1)} < x, \tilde{S}^{(k+j-1)} < T\right) \\ &\quad + \sum_{j=1}^{\infty} r_j \int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(\tilde{S}^{(k+j-1)} < x, \tilde{S}^{(k+j-1)} > T\right). \end{aligned} \tag{4.1}$$

The first part of the formula (4.1) equals

$$\begin{aligned} &\sum_{j=1}^{\infty} r_j \left[P\left(S^{(k+j-1)} < T\right)\int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(S^{(k+j-1)} < x | S^{(k+j-1)} < T\right)\right] \\ &\qquad = \sum_{j=1}^{\infty} r_j \left(\frac{\mu}{\alpha+\mu}\right)^{k+j-1}\left(\frac{\alpha+\mu}{s+\alpha+\mu}\right)^{k+j-1}. \end{aligned}$$

In the second part of the formula (4.1),

$$\begin{aligned} &\int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(\tilde{S}^{(k+j-1)} < x, \tilde{S}^{(k+j-1)} > T\right) \\ &= \sum_{h=0}^{k+j-2} \int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(\tilde{S}^{(k+j-1)} < x, S^{(h)} < T < S^{(h+1)}\right) \\ &= \sum_{h=0}^{k+j-2} P\left(S^{(h)} < T < S^{(h+1)}\right)\int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(\tilde{S}^{(k+j-1)} < x | S^{(h)} < T < S^{(h+1)}\right) \\ &= \sum_{h=0}^{k+j-2} P\left(S^{(h)} < T < S^{(h+1)}\right) \\ &\quad \times \int_0^{\infty} \mathrm{e}^{-sx}\mathrm{d}P\left(V + S'^{(k+j-h-1)} < x | S^{(h)} < T < S^{(h+1)}\right) \\ &= \sum_{h=0}^{k+j-2} \frac{\alpha}{\alpha+\mu}\left(\frac{\mu}{\alpha+\mu}\right)^{h}\left(\frac{\alpha+\mu}{s+\alpha+\mu}\right)^{h+1}\left(\frac{\nu}{s+\nu}\right)^{k+j-h-1}, \end{aligned}$$

where $S'^{(k+j-h-1)}$ means the sum of the $k+j-h-1$ service times when the server is in defective. Since $B^*(s) = \frac{\mu}{s+\mu}$ and $C^*(s) = \frac{\nu}{s+\nu}$, by some calculations

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

we have

$$W^*_{k0}(s)=\sum_{j=1}^{\infty} r_j\left[B^*(s+\alpha)^{k+j-1}+\frac{\alpha\nu}{(\nu-\mu)s+\alpha\nu}\left(C^*(s)^{k+j-1}-B^*(s+\alpha)^{k+j-1}\right)\right].$$

• Case 2: If a batch of customers arrives in the state $(k,1)$, it is straightforward to obtain a similar formula just as the above case:

$$W^*_{k1}(s)=\sum_{j=1}^{\infty} r_j\left[C^*(s+\theta)^{k+j-1}+\frac{\theta\mu}{(\mu-\nu)s+\theta\mu}\left(B^*(s)^{k+j-1}-C^*(s+\theta)^{k+j-1}\right)\right].$$

Finally, $W^*(s)$ is given by

$$\begin{aligned}
W^*(s) &= \sum_{k=0}^{\infty}\pi_{k0}W^*_{k0}(s)+\sum_{k=0}^{\infty}\pi_{k1}W^*_{k1}(s)\\
&= \frac{1-G\left(B^*(s+\alpha)\right)}{g\left[1-B^*(s+\alpha)\right]}\left[\pi_{00}+Q_0\left(B^*(s+\alpha)\right)\right]\\
&\quad+\frac{1-G\left(C^*(s+\theta)\right)}{g\left[1-C^*(s+\theta)\right]}Q_1\left(C^*(s+\theta)\right)\\
&\quad+\frac{\alpha\nu}{g\left[\alpha\nu+(\nu-\mu)s\right]}\frac{\left[\pi_{00}+Q_0\left(C^*(s)\right)\right]\left[1-G\left(C^*(s)\right)\right]}{1-C^*(s)}\\
&\quad-\frac{\alpha\nu}{g\left[\alpha\nu+(\nu-\mu)s\right]}\frac{\left[\pi_{00}+Q_0\left(B^*(s+\alpha)\right)\right]\left[1-G\left(B^*(s+\alpha)\right)\right]}{1-B^*(s+\alpha)}\\
&\quad+\frac{\theta\mu}{g\left[\theta\mu+(\mu-\nu)s\right]}\frac{Q_1\left(B^*(s)\right)\left[1-G\left(B^*(s)\right)\right]}{1-B^*(s)}\\
&\quad-\frac{\theta\mu}{g\left[\theta\mu+(\mu-\nu)s\right]}\frac{Q_1\left(C^*(s+\theta)\right)\left[1-G\left(C^*(s+\theta)\right)\right]}{1-C^*(s+\theta)}.
\end{aligned}$$

□

5. Numerical examples

In this section, we present a concrete example to illustrate the effect of the different parameters on the system performance measures. Suppose that the batch size of the arrival follows a geometric distribution with parameter p $(0<p<1)$, denote $\bar{p}=1-p$, thus the expectation, the second order moment and the PGF are easy to get, *i.e.*,

$$g=E(X)=\frac{1}{p},\qquad g^{(2)}=\frac{2-p}{p^2},\qquad G(z)=\frac{pz}{1-\bar{p}z},\quad |z|\le 1.$$

Figure 1 shows that the probability of the server being busy in normal or defective state decreases with increasing the value of p, which is consistent with the fact

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

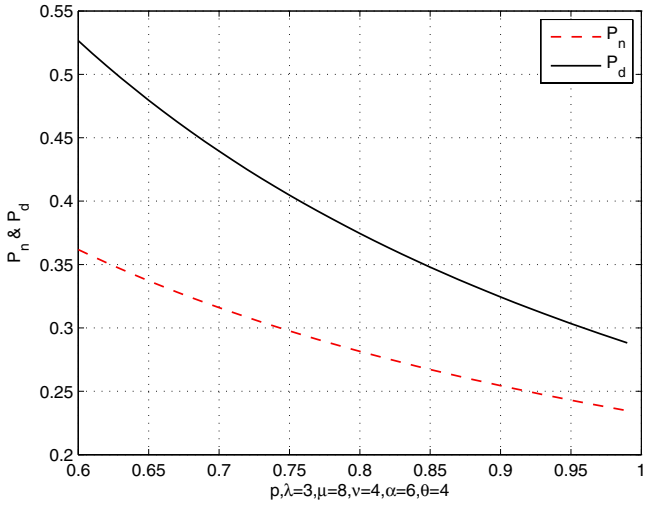


FIGURE 1. The probability of server being busy versus p.

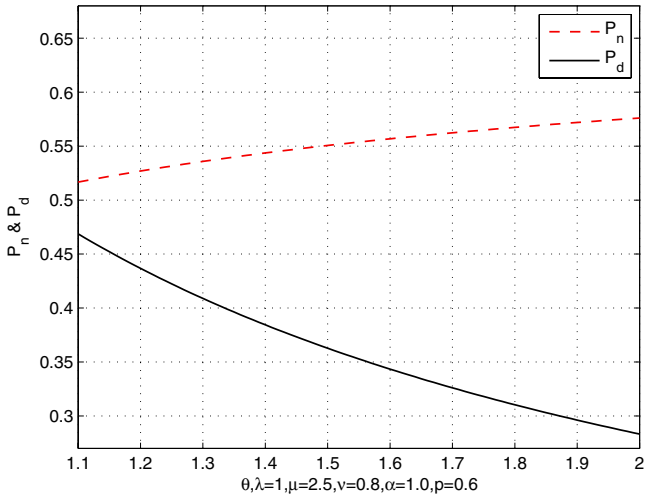


FIGURE 2. The probability of server being busy versus θ.

that, the less average batch size causes the less possibility of the server being busy. Moreover, the probability of the server being busy in defective state is bigger than that of the server is in normal state because α is bigger than θ. While in Figure 2, the curve of P_n is above the curve of P_d due to α smaller than θ. And with the average repair time decreasing, the probability of the server being busy in normal

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

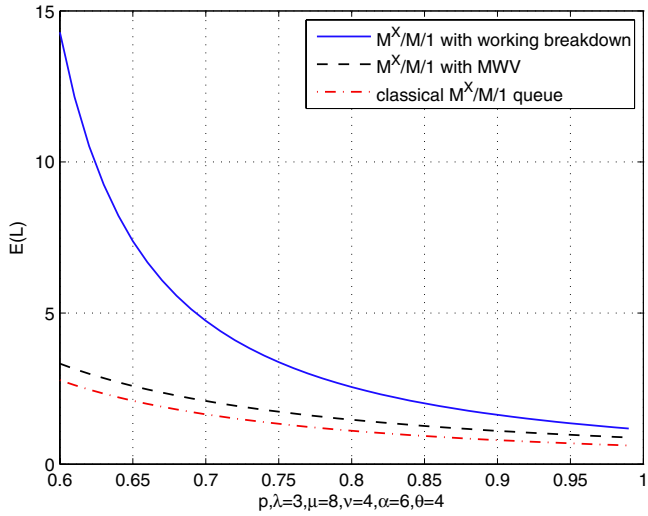


FIGURE 3. $E(L)$ versus p for different batch arrival model.

state increases, but that of the server being defective decreases. From Figures 1 and 2, we note that P_d is more sensitive to the parameter p and θ than P_n.

Next we compare the average queue length in three different queueing models, that is, the classical $M^X/M/1$ queue, the $M^X/M/1$ queue with multiple working vacation (MWV) and the $M^X/M/1$ queue with working breakdown. The concrete expressions of the mean length are presented in Appendix.

Figure 3 illustrates the change trends of average queue length in three different batch arrival queue models, it shows that the smaller the average batch size is, the shorter the average queue length is. Obviously, these change trends conform with our tuition. It is also worth to note that, for a fixed p, our model has the longest average queue length and the classical $M^X/M/1$ queue has the shortest one. This is because the queue system with working breakdown is more vulnerable than the queue with multiple working vacation (MWV), and the classical $M^X/M/1$ queue is the most stable. In Figure 4, it presents the average queue length of the $M^X/M/1$ with working breakdown versus ν for different θ. If $\nu < \mu$, for a fixed ν, the longer the average repair time leads to a longer average queue length. Especially, when ν goes to $\mu = 0.8$, θ has less and less influence on the average queue length. And if $\nu = \mu = 0.8$, the three models will be the same as the classical $M^X/M/1$ queue model.

6. Conclusions

In this paper, we consider a $M^X/M/1$ queue with working breakdown. We generalize the results in [8]. In fact, the concept of working breakdown is totally

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

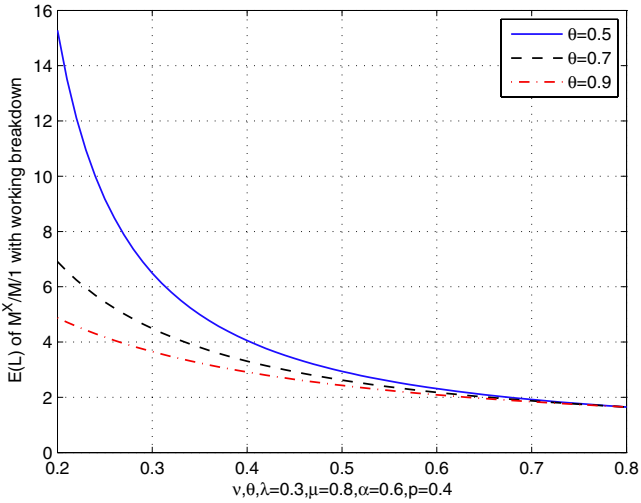


FIGURE 4. $E(L)$ versus ν for different θ in $M^X/M/1$ with working breakdown model.

different from the working vacation policy. That means if the server is subject to an unpredictable breakdown, the interrupted customer will be served immediately at a lower service rate by a backup one. Such examples could be found in the communication system, transportation, manufacturing system and so on. It could be interesting to consider this policy in queue models just as those with working vacation.

Acknowledgements. We thank the editor and the anonymous reviewers for their suggestions. This work was done during the visit to the School of Mathematics and Statistics of Carleton University (Ottawa, CA.). We thank the school for good working condition and would also like to thank China Scholarship Council for supporting our visit to Carleton University. This work was supported by National Natural Science Foundation of China (Nos. 11271373, 11361007) and Guangxi Natural Science Foundation (No. 2012GXNSFBA053010).

APPENDIX

We suppose that the batch size of the arrivals follow the same geometric distribution with parameter p $(0 < p < 1)$ in the classical $M^X/M/1$ queue, the $M^X/M/1$ queue with MWV and the $M^X/M/1$ queue with working breakdown. All other random variables follow the exponential distributions, that is, the batch arrival rate is λ, the normal service rate is μ, the service rate during the working vacation or the repair period is ν, the average working vacation time and the repair time are both $1/\theta$. If the server would be subject to the working breakdown,

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

the average life time of the server is $1/\alpha$. Next we show the expectations of the stationary queue length in the above three models.

At first, the average queue length of the classical $M^X/M/1$ queue model is given by (3.13), that is,

$$E(L_0) = \frac{\lambda}{\mu p^2 - \lambda p}.$$

Then according the Corollary 1 in [2], we have the average queue length of the $M^X/M/1$ with multiple working vacation, *i.e.*,

$$E(L_{MWV}) = \frac{\lambda(\mu-\nu)\left[\theta(1-\bar{p}\beta) + (\lambda-\nu p)(1-\beta)\right]}{p\left[\mu\theta^2(1-\bar{p}\beta) + \lambda\theta(\mu-\nu)(1-\beta)\right]} + E(L_0),$$

where β is the unique root in $(0,1)$ of

$$f(z) = (\lambda + \nu\bar{p} + \theta\bar{p})z^2 - (\lambda + \nu + \theta + \nu\bar{p})z + \nu.$$

Finally, we can compute the expectation of the stationary queue length in our model directly from Corollary 3.1. Let

$$\begin{aligned}\phi(z) = &-\left[(\lambda+\mu\bar{p})(\nu\bar{p}+\lambda+\theta\bar{p}) + \alpha\bar{p}(\lambda+\nu\bar{p})\right]z^3\\ &+\left[(\lambda+\mu\bar{p})(\nu\bar{p}+\nu+\lambda+\theta) + \mu(\nu\bar{p}+\lambda+\theta\bar{p}) + \alpha(\lambda+2\nu\bar{p})\right]z^2\\ &-\left[\mu(\nu\bar{p}+\nu+\lambda+\theta) + \nu(\lambda+\mu\bar{p}+\alpha)\right]z\\ &+\mu\nu.\end{aligned}$$

Suppose γ is the unique root of $\phi(z)$ in $(0,1)$, and

$$\begin{aligned}\rho &= \frac{\lambda(\alpha+\theta)}{p(\alpha\nu+\theta\mu)} < 1,\\ \xi(z) &= \nu(1-z) - \lambda z\frac{1-z}{1-\bar{p}z} - \theta z,\\ \pi_{00} &= \frac{\theta(1-\rho)(1-\bar{p}\gamma)}{\theta(1-\bar{p}\gamma) + p\rho\xi(\gamma)},\end{aligned}$$

we have

$$\begin{aligned}E(L_d) = \pi_{00}\Bigg\{&\frac{\lambda[\alpha\nu+\theta\mu+\frac{\lambda\xi(\gamma)}{\theta(1-\bar{p}\gamma)}(\alpha+\theta)]}{p(\mu p-\lambda)(\alpha\nu+\theta\mu)(\rho-1)}\\ &+\frac{[(\lambda\mu-\theta\mu p-\alpha\nu p-\mu\nu p)+\frac{\lambda\xi(\gamma)}{\theta(1-\bar{p}\gamma)}(\lambda-\alpha p-\theta p-\mu p)]}{p(\alpha\nu+\theta\mu)(\rho-1)}\\ &+\frac{[\alpha\nu+\theta\mu+\frac{\lambda\xi(\gamma)}{\theta(1-\bar{p}\gamma)}(\alpha+\theta)][(\lambda-\nu p-\theta p)(\mu p-\lambda)+\alpha p(\lambda-\nu p)+\lambda(\alpha+\theta)]}{p^2(\alpha\nu+\theta\mu)^2(\rho-1)^2}\Bigg\},\end{aligned}$$

and

$$E(L) = E(L_d) + E(L_0).$$

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press

References

[1] I. Atencia and P. Moreno, A discrete-time $Geo/G/1$ retrial queue with server breakdowns. *Asia-Pac. J. Oper. Res.* **23** (2006) 247–271.

[2] Y. Baba, The $M^X/M/1$ queue with multiple working vacation. *Am. J. Oper. Res.* **2** (2012) 217–224.

[3] W.J. Gray, P.P. Wang and M. Scott, A queueing model with multiple types of server breakdowns. *Quality Technology and Quantitative Management* **1** (2004) 245–255.

[4] G.I. Falin, The $M/M/1$ retrial queue with retrials due to server failures. *Queueing Syst.* **58** (2008) 155–160.

[5] D. Jayaraman, R. Nadarajan and R.M. Sitrarasu, A general bulk service queue with arrival rate dependent on server breakdowns. *Appl. Math. Model.* **18** (1994) 156–160.

[6] M. Jain and A. Jain, Working vacations queueing model with multiple types of server breakdowns. *Appl. Math. Model.* **34** (2010) 1–13.

[7] J.C. Ke and K.B. Huang, Analysis of batch arrival queue with randomized vacation policy and an unreliable server. *J. Syst. Sci. Complexity* **25** (2012) 759–777.

[8] K. Kalidass and R. Kasturi, A queue with working breakdowns. *Comput. Indus. Eng.* **63** (2012) 779–783.

[9] W.Y. Liu, X.L. Xu and N.S. Tian, Stochastic decompositions in the $M/M/1$ queue with working vacations. *Oper. Res. Lett.* **35** (2007) 596–600.

[10] L.D. Servi and S.G. Finn, $M/M/1$ queues with working vacations$(M/M/1/WV)$. *Perform. Eval.* **50** (2002) 41–52.

[11] V. Sridharan and P.J. Jayashree, Some characteristics on a finite queue with normal, partial and total failures. *Microelectron. Reliab.* **36** (1996) 265–267.

[12] R.W. Wolff, Stochastic modeling and the theory of queues. Prentice Hall, Inc., New Jersey (1989).

[13] J.T. Wang and P. Zhang, A discrete-time retrial queue with negative customers and unreliable server. *Comput. Indus. Eng.* **56** (2009) 1216–1222.

[14] J.T. Wang and J.H. Cao, Reliability analysis of the retrial queue with server breakdowns and repairs. *Queueing Syst.* **38** (2001) 363–380.

[15] X.L. Xu and Z.J. Zhang, Analysis of multi-server queue with a single vacation (e,d)-Policy. *Perform. Eval.* **63** (2006) 625–638.

[16] X.L. Xu, Z.J. Zhang and N.S. Tian, Analysis for the $M^X/M/1$ working vacation queue. *Int. J. Inf. Manage. Sci.* **20** (2009) 379–394.

https://doi.org/10.1051/ro/2014014 Published online by Cambridge University Press