J Syst Sci Syst Eng (Mar 2007) 16(1): 121-127 ISSN: 1004-3756 (Paper) 1861-9576 (Online)
DOI: 10.1007/s11518-006-5030-6 CN11-2983/N

Technical Note

THE M/M/1 QUEUE WITH WORKING VACATIONS AND VACATION INTERRUPTIONS*

Jihong LI [1] Naishuo TIAN [2]

[1]*College of Economics and Management, Yanshan University*
jhli_two@sina.com
[2]*College of Sciences, Yanshan University, Qinhuangdao, Hebei, 066004, China*
tiannsh@ysu.edu.cn (✉)

Abstract

In this paper, we study the M/M/1 queue with working vacations and vacation interruptions. The working vacation is introduced recently, during which the server can still provide service on the original ongoing work at a lower rate. Meanwhile, we introduce a new policy: the server can come back from the vacation to the normal working level once some indices of the system, such as the number of customers, achieve a certain value in the vacation period. The server may come back from the vacation without completing the vacation. Such policy is called vacation interruption. We connect the above mentioned two policies and assume that if there are customers in the system after a service completion during the vacation period, the server will come back to the normal working level. In terms of the quasi birth and death process and matrix-geometric solution method, we obtain the distributions and the stochastic decomposition structures for the number of customers and the waiting time and provide some indices of systems.

Keywords: Working vacation, vacation interruption, matrix-geometric solution, stochastic decomposition

1. Introduction

Working vacations is a kind of semi-vacation policy and is also a new vacation policy that was introduced by Servi and Finn (2002): a customer is served at a lower rate rather than completely stoping the service during a vacation. In the classical vacation queueing models, during the vacation period the server doesn't continue on the original work and such policy may cause the loss or dissatisfaction of the customers. For the working vacation policy, the server can still work during the vacation and may accomplish other assistant work simultaneously. So, the working vacation is more reasonable than the classical vacation in some cases.

Servi and Finn (2002) studied an M/M/1 queue with working vacation, and obtained the probability generating function of the queue

* This work was supported in part by National Natural Science Foundation of China under Grant No. 10671170.

length and the LST of the waiting time, and applied their results to performance analysis of a gateway router in fiber communication networks. Subsequently, Kim, Choi and Chae (2003), and Wu and Takagi (2003) generalized the study in Servi and Finn (2002) to an M/G/1 queue with working vacations. Baba (2005) extended the study to a GI/M/1 queue with working vacations by the matrix-analysis method.

In this paper, we will consider an M/M/1 queue with vacation interruptions and working vacations. Vacation interruption was introduced by the authors in this paper. The server can stop the vacation once some indices of the system, such as the number of customers, achieve a certain value in the vacation period. Certainly, it is possible for the server to take an interrupted vacation, so we call this policy vacation interruption. The working vacation and vacation interruption are connected in this paper and the server enters into vacation when there are no customers and he can take service at a lower rate in the vacation period. If there are customers in the system at the instant of a service completion in the vacation period, the server will come back to the normal working level. If there are no customers after a service, the server continues the vacation until after a service or a vacation ends if there are customers. In terms of the quasi birth and death process and matrix-geometric solution method, we provide expressions for the distributions of the queue length and the waiting time. Furthermore, we indicate the stochastic decomposition structures. There is the specific relationship between the results of this model and the results of the classical M/M/1 queues without vacations.

The rest of this paper is organized as follows. In section 2, we describe the quasi birth and death process of the model. Section 3 obtains the distribution of the queue length and the state probabilities of the server in steady state. Finally, in section 4, we indicate the stochastic decomposition structures of the queue length and the waiting time in such a system.

In this paper, we consider the case where the service and inter-arrival times are exponentially distributed. The cases where those times are not exponential are important in practice. We can also set a threshold for the vacation interruption. These cases are left for future study.

2. A Quasi Birth and Death Process Model

Consider a classical M/M/1 queue with an arrival rate λ and a service rate μ_b. Upon the completion of a service, if there is no customer in the system, the server begins a vacation and the vacation duration V follows an exponential distribution with the parameter θ. During a vacation period, arriving customers can be served at a mean rate of μ_v. Upon the completion of a service in the vacation, if there are also customers in the queue, the server ends the vacation and comes back to the normal working level. Otherwise, he continues the vacation until there are customers after a service or a vacation ends. Meanwhile, when a vacation ends, if there are no customers, another vacation is taken. Otherwise, the server also switches the rate to μ_b and a regular busy period starts. In this service discipline, the server may come back from the vacation without completing the vacation. And, the server can only go on vacations if there is no customer left in the system upon the completion of a service.

Meanwhile, the vacation service rate can be only applied to the first customer arrived during a vacation period.

We assume that the inter-arrival times, the service times, and the working vacation times are mutually independent. In addition, the service discipline is first in first out (FIFO).

Let $Q_v(t)$ be the number of customers in the system at time t and let

$$J(t)=\begin{cases}0, & \text{the system is in a working vacation period at time } t,\\ 1, & \text{the system is in a regular busy period at time } t.\end{cases}$$

Then { $Q_v(t)$, $J(t)$ } is a quasi birth and death process (QBD) with the state space

$$\Omega=\{(0,0)\}\bigcup\{(k,j):\quad k\ge 1,\quad j=0,1\}.$$

Evidently, when there is no customer, the server only stays in the vacation period.

Using the lexicographical sequence for the states, the infinitesimal generator can be written as

$$Q=\begin{bmatrix}-\lambda & C_0 & & & \\ B_1 & A & C & & \\ & B & A & C & \\ & & \vdots & \vdots & \vdots\end{bmatrix}.$$

where $C_0=(\lambda,0);\ B_1=(\mu_v,\mu_b)^T$;

$$C=\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}, B=\begin{bmatrix}0 & \mu_v\\ 0 & \mu_b\end{bmatrix},$$

$$A=\begin{bmatrix}-(\lambda+\theta+\mu_v) & \theta\\ 0 & -(\lambda+\mu_b)\end{bmatrix}.$$

To analyze this QBD process, it is necessary to solve for the minimal non-negative solution of the matrix quadratic equation

$$R^2B+RA+C=0 \tag{1}$$

This solution is called the rate matrix and denoted by R. Obviously, we have

Lemma 1 *If* $\rho=\lambda(\mu_b)^{-1}<1,$ *the matrix equation (1) has the minimal non-negative solution*

$$R=\begin{bmatrix}\dfrac{\lambda}{\lambda+\theta+\mu_v} & \rho\dfrac{\lambda+\theta}{\lambda+\theta+\mu_v}\\ 0 & \rho\end{bmatrix}. \tag{2}$$

Proof. Because the matrices A, B, C of (1) are all upper triangular, we can assume that R has the same structure as

$$R=\begin{bmatrix}r_{11} & r_{12}\\ 0 & r_{22}\end{bmatrix}.$$

Substituting R^2 and R into (1), we obtain the following set of equations:

$$\begin{cases}-(\lambda+\theta+\mu_v)r_{11}+\lambda=0,\\ \mu_b r_{22}^2-(\lambda+\mu_b)r_{22}+\lambda=0,\\ \mu_v r_{11}^2+\mu_b r_{12}(r_{11}+r_{22})+\theta r_{11}-(\lambda+\mu_b)r_{12}=0.\end{cases} \tag{3}$$

To obtain the minimal non-negative solution of (1), we obtain $r_{22}=\rho$ (the other root is $r_{22}=1$) in the second equation and r_{11} can be derived in the first equation of (3) and is noted by r. Evidently, $0<r_{11}<1$ for any $\theta>0$ and $\mu_v>0$. Substituting r and ρ into the last equation of (3), we get the expression for r_{12}.

Based on the expression for R in Lemma 1 and Theorem 3.1.1 in Neuts(1981), it is easy to prove that the QBD process is positive recurrent if and only if $\rho=\lambda\mu_b^{-1}<1$.

3. Queue Length Distribution

If $\rho<1$, let (Q_v,J) be the stationary limit of the QBD process $\{Q_v(t),J(t)\}$. Let $\pi_0=\pi_{00}$; $\pi_k=(\pi_{k0},\pi_{k1}),\quad k\ge 1$

$$\pi_{kj}=P\{Q_v=k,J=j\}=\lim_{t\to\infty}P\{Q_v(t)=k,J(t)=j\},$$

$(k,j)\in\Omega.$

With regard to the queue length, we have

Theorem 1 *If* $\rho < 1$, *the stationary probability distribution of* (Q_v, J) *is*

$$\begin{cases} \pi_{k0} = K(\dfrac{\lambda}{\lambda+\theta+\mu_v})^k, \quad k \ge 0, \\ \pi_{k1} = K\rho\dfrac{\lambda+\theta}{\lambda+\theta+\mu_v}\sum_{j=0}^{k-1}\rho^j(\dfrac{\lambda}{\lambda+\theta+\mu_v})^{k-1-j}, \\ \qquad k \ge 1. \end{cases} \tag{4}$$

where $K = \dfrac{\theta+\mu_v}{\lambda+\theta+\mu_v}(1+\dfrac{\theta+\lambda}{\lambda+\theta+\mu_v}\dfrac{\rho}{1-\rho})^{-1}$.

Proof. In terms of the matrix-geometric solution method (Neuts, 1981), we have

$$\pi_k = (\pi_{k0}, \pi_{k1}) = (\pi_{10}, \pi_{11})R^{k-1}, \quad k \ge 1. \tag{5}$$

and $(\pi_{00}, \pi_{10}, \pi_{11})$ satisfies the set of equations

$$(\pi_{00}, \pi_{10}, \pi_{11})B[R] = 0 \tag{6}$$

where

$$B[R] = \begin{bmatrix} A_{00} & C_0 \\ B_{10} & RB+A \end{bmatrix} = \begin{bmatrix} -\lambda & \lambda & 0 \\ \mu_v & -(\lambda+\theta+\mu_v) & \lambda+\theta \\ \mu_b & 0 & -\mu_b \end{bmatrix}.$$

Substituting $B[R]$ into the above relation, we obtain the set of equations

$$\begin{cases} -\lambda\pi_{00} + \mu_v\pi_{10} + \mu_b\pi_{11} = 0, \\ \lambda\pi_{00} - (\lambda+\theta+\mu_v)\pi_{10} = 0, \\ (\lambda+\theta)\pi_{10} - \mu_b\pi_{11} = 0. \end{cases} \tag{7}$$

Taking $\pi_{00} = K$, from the second and third equations in (7), we get

$$\pi_{10} = K(\frac{\lambda}{\lambda+\theta+\mu_v}), \quad \pi_{11} = K\rho\frac{\lambda+\theta}{\lambda+\theta+\mu_v}$$

For $k \ge 1$, note that

$$R^k = \begin{bmatrix} r^k & \rho\dfrac{\lambda+\theta}{\lambda+\theta+\mu_v}\sum_{j=0}^{k-1}\rho^j r^{k-1-j} \\ 0 & \rho^k \end{bmatrix}, \quad k \ge 1.$$

And $r = \lambda(\lambda+\theta+\mu_v)^{-1}$. Substituting (π_{10}, π_{11}) and R^{k-1} into (5), we obtain (4). Finally, the constant factor K can be determined by the normalization condition.

Thus, we can obtain the distribution of the number of customers Q_v

$$p\{Q_v = k\} = \pi_{k0} + \pi_{k1} = K(\frac{\lambda}{\lambda+\theta+\mu_v})^k + K\rho\frac{\lambda+\theta}{\lambda+\theta+\mu_v}\sum_{j=0}^{k-1}\rho^j(\frac{\lambda}{\lambda+\theta+\mu_v})^{k-1-j}, k \ge 1$$

Meanwhile, we can easily obtain the state probabilities of a server in the steady-state.

$$p\{J=0\} = \sum_{k=0}^{+\infty}\pi_{k0} = K\frac{\lambda+\theta+\mu_v}{\theta+\mu_v} = (1+\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho})^{-1}$$

$$p\{J=1\} = \sum_{k=1}^{+\infty}\pi_{k1} = K\left\{\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho} + \frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho}\frac{\lambda}{\theta+\mu_v}\right\} = \frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho}(1+\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho})^{-1}. \tag{8}$$

Remark 1: If $\mu_v = 0$, i.e., the server can not take the service in the vacation period, we obtain the classical results of the steady-state distribution for the M/M/1 queue with multiple vacations. So we can set up the relationship between the classical vacation model and our present model.

4. Stochastic Decomposition Results

We know that the classical vacation queues have the stochastic decomposition property for the queue length and the waiting time. The model in this paper also has this property. Now, we give stochastic decomposition structures of the stationary queue length and the waiting time, denoted by Q and W, respectively.

Firstly, we discuss the distribution of the number of customers in the system.

Theorem 2 *If* $\rho<1$ *and* $\mu_b>\mu_v$*, the stationary queue length* Q_v *can be decomposed into the sum of two independent random variables:* $Q_v=Q_0+Q_d$ *. Where* Q_0 *is the stationary queue length of a classical M/M/1 queue without vacations, and follows a geometric distribution with the parameter* $1-\rho$ *; the additional queue length* Q_d *has a modified geometric distribution*

$$\begin{cases} P\{Q_d=0\}=K^*\delta_1, \\ P\{Q_d=k\}=K^*\delta_2\left(1-\frac{\lambda}{\lambda+\theta+\mu_v}\right)\left(\frac{\lambda}{\lambda+\theta+\mu_v}\right)^k, \\ \qquad k\geq 1. \end{cases} \tag{9}$$

where

$$\delta_1=\frac{1}{1-\rho}\frac{\theta+\mu_v}{\lambda+\theta+\mu_v},\quad \delta_2=\frac{1}{1-\rho}\frac{\lambda-\rho\mu_v}{\lambda+\theta+\mu_v}.$$

And $K^*=(1+\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho})^{-1}$.

Proof. From Theorem 1, the probability generating function of Q_v is as follows

$$\begin{aligned} Q(z)&=\sum_{k=0}^{\infty}z^k p\{Q_v=k\} \\ &=K\left[\frac{1}{1-rz}+\rho\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{1}{1-\rho z}\frac{1}{1-rz}\right] \\ &=\frac{1-\rho}{1-\rho z}K^*\left[\frac{1-r}{1-rz}\frac{1-\rho z}{1-\rho}+\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho}\frac{(1-r)z}{1-rz}\right] \\ &=\frac{1-\rho}{1-\rho z}K^*\left[\frac{\theta+\mu_v}{\lambda+\theta+\mu_v}\frac{1}{1-\rho}+\frac{\lambda-\rho\mu_v}{\lambda+\theta+\mu_v}\frac{1}{1-\rho}\frac{(1-r)z}{1-rz}\right] \\ &=\frac{1-\rho}{1-\rho z}K^*\left[\delta_1+\delta_2\frac{(1-r)z}{1-rz}\right] \\ &=Q_0(z)Q_d(z) \end{aligned} \tag{10}$$

where $r=\lambda(\lambda+\theta+\mu_v)^{-1}$, and δ_1,δ_2,K^* are defined as in the theorem. We can verify that $\delta_1+\delta_2=(K^*)^{-1}$. Thus, we can get the (9).

Equation (9) indicates that the additional delay Q_d can be written as the mixture of two random variables $Q_d=q_0X_0+q_1X_1$, where $q_0=K^*\delta_1, q_1=K^*\delta_2$, and $X_0\equiv 0, X_1$ follows a geometric distribution with the parameter $(1-r)$ on the set $\{1,2,\cdots\}$.

With Theorem 2, we can easily get means

$$E(Q_d)=K^*\delta_2\frac{1}{1-\frac{\lambda}{\lambda+\theta+\mu_v}}=\frac{\frac{1}{1-\rho}\frac{\lambda-\rho\mu_v}{\theta+\mu_v}}{1+\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho}},$$

$$E(Q_v)=\frac{\rho}{1-\rho}+E(Q_d).$$

Now, we analyze the waiting time of each customer. Denote the LST of W by $W^*(s)$.

Theorem 3 *If* $\rho<1$ *and* $\mu_b>\mu_v$*, the stationary waiting time W can be decomposed into the sum of two independent random variables:* $W=W_0+W_d$ *, where* W_0 *is the waiting time of a customer in a corresponding classical M/M/1 queue without vacations, which has an exponential distribution with the parameter* $\mu_b(1-\rho)$ *; the additional delay* W_d *has the LST*

$$W_d^*(s)=K^*\left[\sigma_1+\sigma_2\frac{\theta+\mu_v}{s+\theta+\mu_v}\right]. \tag{11}$$

where $\sigma_1=\delta_1-\frac{\theta+\mu_v}{\lambda}\delta_2;\ \sigma_2=\delta_2\frac{\lambda+\theta+\mu_v}{\lambda}.$

And δ_1,δ_2 *are noted as Theorem 2.*

Proof. In the model, the queue length and the waiting time of a customer have the relationship below as in the other vacation queues:

$$Q(z)=W^{*}(\lambda(1-z)).$$

So, let $s=\lambda(1-z)$, $z=1-s\lambda^{-1}$ is replaced into the equation above, and we easily obtain the theorem above. Evidently, $\sigma_1+\sigma_2=(K^{*})^{-1}$.

Similarly, the equation (11) indicates that the stationary additional delay is the mixture of two variables $W_d=q_0Y_0+q_1Y_1$. $Y_0\equiv 0, Y_1$ follows an exponential distribution with the parameter $(\theta+\mu_v)$.

With the stochastic decomposition structure in Theorem 3, we can easily get means

$$E(W_d)=K^{*}\sigma_2\frac{1}{\theta+\mu_v}=\frac{\frac{1}{\lambda}\frac{1}{1-\rho}\frac{\lambda-\rho\mu_v}{\theta+\mu_v}}{1+\frac{\theta+\lambda}{\lambda+\theta+\mu_v}\frac{\rho}{1-\rho}},$$

$$E(W)=\frac{1}{\mu_b(1-\rho)}+E(W_d)=\frac{1}{\lambda}E(Q_v).$$

Remark 2: From the equation above, we can find that the Little's formula, $E(Q_v)=\lambda E(W)$ also exists in the M/M/1 queue with multiple working vacations and vacation interruptions.

5. Acknowledgment

The authors thank to the anonymous referees for their insightful comments and suggestions, which are very helpful to improve the presentations of the paper.

References

[1] Baba, Y. (2005). Analysis of a GI/M/1 queue with multiple working vacations. Operation Research Letters, 33: 201-209

[2] Doshi, B.T. (1986). Queueing systems with vacations - A survey. Queueing Systems, 1(1): 29-66

[3] Doshi, B.T. (1990). Single server queues with vacations. In: H. Takagi (ed.), Stochastic Analysis of the Computer and Communication Systems. North-Holland Elsevier, Amsterdam

[4] Kim, J., Choi, D. & Chae, K. (2003). Analysis of queue-length distribution of the M/G/1 queue with working vacations. In: International Conference on Statistics and Related Fields, Hawaii

[5] Neuts, M. (1981). Matrix-Geometric Solutions in Stochastic Models. Johns Hopkins University Press, Baltimore

[6] Servi, L.D. & Finn, S.G. (2002). M/M/1 queue with working vacations (M/M/1WV). Performance Evaluation, 50: 41-52

[7] Tian, N.S. & Zhang, Z.G. (2006). Vacation Queueing Models: Theory and Applications. Springer Publisher, New York

[8] Wu, D. & Takagi, H. (2003). M/G/1 queue with multiple working vacations. In: Proceedings of the Queueing Symposium, Stochastic Models and the Applications, Kakegawa, 51-6

Jihong Li is studying for her Ph.D degree in the College of Economics and Management, Yanshan University in China. Her research interests are queueing models with vacations, production management and performance analysis of Communication Networks. She has published several papers in international journals, such as Applied Mathematics and Computation (AMC) and the other journals in China.

Naishuo Tian is a professor in the College of Sciences, Yanshan University in China. His

research interests are stochastic process, queueing systems, and performance evaluation models in computer networks and supply chain systems. He has published many papers in several international journals including Queueing Systems, Stochastic Models, Performance Evaluation, Operations Research Letters, EJOR, INFOR and several other journals in China.