Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle

S. K. Matendo

Application of Neuts' method to bulk queueing models with vacations

Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle, tome 29, n° 1 (1995), p. 93-117.

<http://www.numdam.org/item?id=RO_1995__29_1_93_0>

Recherche opérationnelle/Operations Research
(vol. 29, n° 1, 1995, p. 93 à 117)

APPLICATION OF NEUTS' METHOD TO BULK QUEUEING MODELS WITH VACATIONS (*)

by S. K. MATENDO (1)

Communicated by Bernard LEMAIRE

Abstract. – *We study a single server infinite capacity queueing system in which customers, arriving in groups according to a Poisson process, are served in batches under the usual bulk service rule. The server applies a general exhaustive service vacation policy. Computational results for the queue length at a post-departure or inactive phase termination epoch, at a post-departure epoch and at an arbitrary epoch are given. In particular, we consider a vacation model where the decision of whether to take a vacation or not, in a given inactive phase, is allowed to depend on the number of vacations already taken in this inactive phase and the number of customers waiting, compared to a specified number N. As special cases, it includes the $(T(SV); N)$-policy and the $(T(MV); N)$-policy. We discuss two applications in the analysis of production, computer and communication systems. A simple numerical example is also considered.*

Keywords: Queueing systems, vacation models, matrix analytic methods.

Résumé. – *Nous considérons un modèle d'attente à un guichet et à capacité infinie. Les arrivées successives se font par groupes d'effectif aléatoire suivant un processus de Poisson. Les clients sont servis en groupes d'effectif maximum donné. Le serveur applique une politique générale de vacances du serveur avec service exhaustif. Nous nous intéressons au calcul de la distribution stationnaire de la longueur de la file aux instants de fin de service ou de fin de période d'inactivité, aux instants de départ des clients et à un instant quelconque. Nous considérons enfin un cas particulier important où l'on permet à la décision de prendre ou non une nouvelle vacance, dans une période d'inactivité donnée, de dépendre du nombre de vacances déjà prises (dans cette période d'inactivité) et du nombre de clients qui attendent d'être servis, comparé à un nombre fixé N $(N \geqq 1)$. Deux cas particuliers importants sont la $(T(SV); N)$-politique et la $(T(MV); N)$-politique. Nous discutons de l'application du modèle à l'étude de systèmes de production, de systèmes informatiques et de systèmes de communication. Un exemple numérique est également présenté.*

Mots clés : Systèmes d'attente, files d'attente avec vacances du serveur, méthodes analytiques matricielles.

1. INTRODUCTION

A "vacation model" (vacation system) is a queueing system where the server alternates between active and inactive states. In the active state, the

(*) Received December 1991.

(1) Université de Mons-Hainaut, Place Warocqué, 17, B-7000 Mons, Belgium.
Holder of an A.G.C.D. scholarship.

server provides service to customers, so that in an "active phase", the system is never empty. An active phase is followed by an "inactive phase" during which the server is unavailable to the customers.

Vacation models have been analyzed by many authors. We refer to the surveys by Doshi [3] and Teghem Jr. [18], and to Takagi's book [17]. This type of queueing systems has been useful for computer, communication and manufacturing modelling. In these systems, the time periods during which secondary jobs or maintenances are performed can be considered as inactive phases. Another possible application can be found in the study of the performance of polling systems in which the server visits a given number of queues in a cyclic order. Here, the time periods during which the server is visiting (polling) the other queues can be regarded as inactive phases.

Several service disciplines (services policies) have been studied in the literature: exhaustive service, gated service, limited service, etc. In this paper, we consider a rather large class of exhaustive service vacation policies *i. e.* policies according to which each time the server becomes active, he works in a continuous manner until the system becomes empty. Special cases are the N-policy and the T-policy [*i. e.* the single vacation T-policy, $T\,(SV)$, or the multiple vacation T-policy, $T\,(MV)$]. For the N-policy, when the system becomes empty, the server remains inactive, inspecting the queue, until $N\,(N \geqq 1)$ customers are present. In this model, the inactive phases are "inspection phases". For the T-policy, upon becoming idle, the server leaves the system for some interval of time (of random length) called a "vacation". If he finds at least one customer waiting when he comes back to the system, the server becomes active immediately. Otherwise, in the single vacation model, he waits for the first arrival to start service, while in the multiple vacation model, the vacations are repeated until the server finds at least one customer waiting upon return from a vacation. We note that for the $T\,(SV)$-policy, an inactive phase consists of a vacation followed by a possible inspection period, while for the $T\,(MV)$-policy, an inactive phase consists of a random number of vacations.

2. THE MODEL

We consider an infinite capacity queueing system with the following assumptions:

H_1: groups of customers arrive according to a Poisson process of parameter λ. The sizes of the groups are i.i.d. random variables, with probability distribution $\{d_k\}_{k \geqq 1}$, generating function $D(z) = \sum_{k \geqq 1} d_k z^k$ $(|z| \leqq 1)$, finite expectation $E[D]$ and second order moment $E[D^2]$.

H_2: the server applies a general exhaustive service vacation policy (*see* for instance Loris-Teghem [9]) in which, as mentioned above, an active phase terminates when the system gets empty. Let V_n $(n \geqq 1)$ denote the duration of the n-th inactive phase. We assume that the random variables, V_n, $n \geqq 1$, are i.i.d., with generic variable V, finite expectation $E[V]$ and second order moment $E[V^2]$. Let Y_n $(n \geqq 1)$ be the total number of customers arriving in the n-th inactive phase; $\{Y_n\}_{n \geqq 1}$ is a sequence of i.i.d. random variables with $Y_n \geqq 1$ a.s. We denote by Y a random variable distributed as the Y_n, $n \geqq 1$, with finite expectation $E[Y]$ and second order moment $E[Y^2]$.

H_3: customers are served by a single server in batches of maximum size m $(m \geqq 1)$ under the usual bulk service rule. If the queue length at the beginning of a service exceeds m, then a group of m customers is served. If the queue is not empty, but does not exceed m, then all customers enter service. Customers are served in the order of their arrivals (customers within a batch are preordered for service or served in random order). The successive service times are conditionally independent, given the group sizes, but may depend on the number of customers in the groups. The distribution function of the service time of a group of size j, $1 \leqq j \leqq m$, is denoted by $S_j(.)$, with Laplace-Stieltjes transform (L.S.T.) $\tilde{S}_j(.)$, finite expectation $E[S_j]$ and second order moment $E[S_j^2]$. The service times are independent of the interarrival times and the durations of the inactive phases.

H_4: The traffic intensity (offered load) $\rho \equiv m^{-1} \lambda E[D] E[S_m]$ is less than one.

For the bulk arrival $M/G/1$ queue with a general exhaustive service vacation policy, a computational method has been recently applied (*see* Matendo [11]) to obtain the steady-state distributions of the queue length at a post-departure or inactive phase termination epoch, at a post-departure epoch and at an arbitrary epoch. The main purpose of the present paper is to extend to the bulk service case the computational results obtained there, using the same basic arguments.

In section 3, we use Neuts' method to obtain the steady-state distribution of the queue length at service completion or inactive phase termination epochs. We then relate the steady-state distribution of the queue length at service completion epochs (section 4) and at an arbitrary epoch (section 5) to the former distribution. In section 6, we particularize the results to the case of a specific vacation policy which extends the policy considered in Kella [4]. In that paper, Kella considers an $M/G/1$ queue with server vacations where the decision of whether to take a new vacation or not, when the system is empty, depends on the number of vacations already taken in the current inactive phase. Using standard renewal arguments, Kella obtained the L.S.T. and the first two moments of the steady-state distribution of the waiting time. He also discussed optimization results for the vacation policy. Note that the exhaustive service discipline considered by Kella includes the $T\,(SV)$-policy and the $T\,(MV)$-policy. We generalize Kella's scheme to the case where the decision of whether to take a vacation or not is allowed to depend on the number of vacations already taken in the current inactive phase and the number of customers waiting, compared to a specified number (say N). This includes the $(T\,(SV);\,N)$-policy [combination of the N-policy and the $T\,(SV)$-policy] and the $(T\,(MV);\,N)$-policy [combination of the N-policy and the $T\,(MV)$-policy] introduced in Loris-Teghem [8]. In section 7, we present two potential applications in production processes and in computer and communications systems. We also discuss a simple numerical example. Concluding remarks follow in section 8.

Notations

For $t \geqq 0$ and $n \geqq 0$, let $P\,(n,\,t)$ denote the probability that n customers arrive during the interval $]0,\,t]$, $a_n\,(t)$ and $h_{l,\,n}\,(t)$ $(1 \leqq l \leqq m-1)$ the probability that the end of a service of a batch of size m and l, respectively, starting at time 0, occurs no later than time t, and during the service there were n arrivals. Let $c_n\,(t)$ denote the probability that the end of an inactive phase, starting at time 0, occurs no later than time t, and during the inactive phase there were n arrivals.

Then

$$P\,(n,\,t) = \sum_{k=0}^{n} e^{-\lambda\,t}\,\frac{(\lambda\,t)^k}{k!}\,d_n\,(k), \tag{1}$$

where $\{d_n\,(k)\}_{n \geqq 0}$ is the k-fold convolution of the probability distribution $\{d_n\}_{n \geqq 1}$ [with $d_n\,(0) = \delta_{0\,n}$ the Kronecker delta],

$$\left.\begin{aligned} a_n\,(t) &= \int_0^t P\,(n,\,x)\,dS_m(x), \\ h_{l,\,n}\,(t) &= \int_0^t P\,(n,x)\,dS_l\,(x), \end{aligned}\right\} \tag{2}$$

and

$$c_n(t) = \Pr[V \leqq t, Y = n]. \tag{3}$$

We note that $a_n(t)$, $h_{l,n}(t)$ and $c_n(t)\,(c_0(t) = 0)$ are probability mass-functions on $[0, +\infty[$ (*i. e.* non-decreasing functions, which lie between 0 and 1 but do not necessary tend to 1 at $+\infty$). For $\operatorname{Re}(s) \geqq 0$, let $\tilde{a}_n(s)$, $\tilde{h}_{l,n}(s)$ and $\tilde{c}_n(s)$ denote the Laplace-Stieltjes transform (L.S.T.) of $a_n(t)$, $h_{l,n}(t)$ and $c_n(t)$, respectively. From (1) and (2), it can be seen that the joint transforms $\tilde{a}(z, s) = \sum_{n \geqq 0} \tilde{a}_n(s) z^n$ and $\tilde{h}_l(z, s) = \sum_{n \geqq 0} \tilde{h}_{l,n}(s) z^n$ $(\operatorname{Re}(s) \geqq 0, |z| \leqq 1)$ of the number of customers arrived and the total service time used, during the service of a batch of size m and l, respectively, are explicitely given by

$$\left.\begin{aligned} &\tilde{a}(z, s) = \tilde{S}_m[s + \lambda - \lambda D(z)] \\ \text{and}\quad &\tilde{h}_l(z, s) = \tilde{S}_l[s + \lambda - \lambda D(z)], \qquad \text{for} \quad 1 \leqq l \leqq m - 1 \end{aligned}\right\} \tag{4}$$

The joint transform $\tilde{c}(z, s) = \sum_{n \geqq 1} \tilde{c}_n(s) z^n$ of the number of customers arrived during an inactive phase and of the duration of that inactive phase does not have an explicit closed form since, as mentioned in Matendo [11], $c_n(t)\,(n \geqq 1)$ [which has the same meaning as $C_n(t)\,(n \geqq 1)$ define there], depends on the specific type of vacation policy considered.

In the following, a_n, $h_{l,n}$ and c_n refer to

$$a_n(+\infty)\,(= \tilde{a}_n(0)), \quad h_{l,n}(+\infty)\,(= \tilde{h}_{l,n}(0)) \quad \text{and} \quad c_n(+\infty)\,(= \tilde{c}_n(0)),$$

respectively [*i. e.* the probability that n customers arrive during the service of a batch of size m, the service of a batch of size l $(1 \leqq l \leqq m - 1)$, and an inactive phase, respectively].

Their corresponding generating functions will be denoted by $a(z)$, $h_l(z)$ and $c(z)$, respectively.

Observe that for $n \geqq 0$,

$$\left.\begin{aligned} &a_n = \sum_{j=0}^{n} \Psi_j\, d_n(j), \\ \text{and}\quad & \\ &h_{l,n} = \sum_{j=0}^{n} \Psi_{l,j}\, d_n(j), \qquad \text{for} \quad 1 \leqq l \leqq m - 1, \end{aligned}\right\} \tag{5}$$

where, for $j \geqq 0$,

$$\left.\begin{aligned} \Psi_j &= \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} dS_m(t), \\ \text{and} \\ \Psi_{l,j} &= \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} dS_l(t), \qquad \text{for} \quad 1 \leqq l \leqq m-1, \end{aligned}\right\} \tag{6}$$

[Ψ_j and $\Psi_{l,j}$ $(1 \leqq l \leqq m-1)$ represent the probability of j arrival batches during the service of a batch of size m and l, respectively].

From (4) it is easy to see that

$$\left.\begin{aligned} a(z) &= \tilde{a}(z, 0) = \tilde{S}_m[\lambda - \lambda D(z)], \\ \text{and} \\ h_l(z) &= \tilde{h}_l(z, 0) = \tilde{S}_l[\lambda - \lambda D(z)] \end{aligned}\right\} \tag{7}$$

Further, for $j \geqq 0$, let

$$\left.\begin{aligned} \Psi_j^* &= \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} [1 - S_m(t)]\, dt, \\ \text{and} \\ \Psi_{l,j}^* &= \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} [1 - S_l(t)]\, dt, \qquad 1 \leqq l \leqq m-1 \end{aligned}\right\} \tag{8}$$

From (1) and (8) we have for $n \geqq 0$,

$$\left.\begin{aligned} a_n^* &\equiv \int_0^{\infty} P(n, t)[1 - S_m(t)]\, dt = \sum_{j=0}^{n} \Psi_j^* d_n(j), \\ \text{and} \\ h_{l,n}^* &\equiv \int_0^{\infty} P(n, t)[1 - S_l(t)]\, dt = \sum_{j=0}^{n} \Psi_{l,j}^* d_n(j), \\ &\text{for} \quad 1 \leqq l \leqq m-1 \end{aligned}\right\} \tag{9}$$

Remark 1: When the service time distributons $S_k(.)$, $1 \leqq k \leqq m$, are of phase type (*see* for instance Latouche [5] and Neuts [12]), explicit expressions can be obtained for the sequences $\{\psi_j\}_{j \geqq 0}$, $\{\psi_j^*\}_{j \geqq 0}$, $\{\psi_{l,j}\}_{j \geqq 0,\, 1 \leqq l \leqq m-1}$ and $\{\psi_{l,j}^*\}_{j \geqq 0,\, 1 \leqq l \leqq m-1}$. For more details, we refer to th. 5.1.5 in Neuts [12].

3. THE QUEUE LENGTH AT SERVICE COMPLETION OR INACTIVE PHASE TERMINATION EPOCHS

In this section, we derive the stationary probability distribution of the queue length at service completion or inactive phase termination epochs. As in Matendo [11], $N(t)$ represents the number of customers in the system at time t, θ_n $(n \geqq 0)$ is the epoch of the n-th transition (*i. e.* service completion or inactive phase termination), ζ_n the number of customers arrived in the time interval $]\theta_n, \theta_{n+1}]$, $T_n = \theta_{n+1} - \theta_n$ and $N_n = N(\theta_n +)$.

It can be easily seen that for $n \geqq 0$,

$$N_{n+1} = (N_n - m)^+ + \zeta_n$$
$$(x^+ = \max(x, 0))$$

and

$$\begin{aligned} &\Pr[N_{n+1} = j, T_n \leqq x \mid N_0, T_0, N_1, T_1, \ldots, T_{n-1}, N_n = i] \\ &\quad = \Pr[N_{n+1} = j, T_n \leqq x \mid N_n = i] \qquad (\equiv Q_{ij}(x), x \geqq 0) \end{aligned}$$

Therefore, the sequence $\{(N_n, T_n)\}_{n \geq 0}$ is a Markov renewal process (M.R.P.) on the state space $\{i \geqq 0\} \times [0, +\infty[$. Its transition probability matrix $Q(x)$, $x \geqq 0$, is given by

$$Q(x) =$$

$$\begin{pmatrix}
c_0(x) & c_1(x) & \cdots & c_{m-1}(x) & c_m(x) & c_{m+1}(x) & \cdots & c_{2m-1}(x) & \cdots \\
h_{1,0}(x) & h_{1,1}(x) & \cdots & h_{1,m-1}(x) & h_{1,m}(x) & h_{1,m+1}(x) & \cdots & h_{1,2m-1}(x) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
h_{m-1,0}(x) & h_{m-1,1}(x) & \cdots & h_{m-1,m-1}(x) & h_{m-1,m}(x) & h_{m-1,m+1}(x) & \cdots & h_{m-1,2m-1}(x) & \cdots \\
a_0(x) & a_1(x) & \cdots & a_{m-1}(x) & a_m(x) & a_{m+1}(x) & \cdots & a_{2m-1}(x) & \cdots \\
0 & a_0(x) & \cdots & a_{m-2}(x) & a_{m-1}(x) & a_m(x) & \cdots & a_{2m-2}(x) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & a_0(x) & a_1(x) & a_2(x) & \cdots & a_m(x) & \cdots \\
0 & 0 & \cdots & 0 & a_0(x) & a_1(x) & \cdots & a_{m-1}(x) & \cdots \\
0 & 0 & \cdots & 0 & 0 & a_0(x) & \cdots & a_{m-2}(x) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & a_0(x) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix} \tag{10}$$

where $a_j(x)$, $h_{l,j}(x)$ $(1 \leqq l \leqq m-1)$ and $c_j(x)$, $j \geqq 0$, are the probability mass-functions defined in (2) and (3). This M.R.P. is of $M/G/1$ type (*see* Neuts [12]). To see it, we partition the matrix $Q(x)$ into $m \times m$ blocks to obtain

$$Q(x) = \begin{pmatrix} C_0(x) & C_1(x) & C_2(x) & \dots \\ A_0(x) & A_1(x) & A_2(x) & \dots \\ 0 & A_0(x) & A_1(x) & \dots \\ 0 & 0 & A_0(x) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \qquad x \succeq 0, \qquad (11)$$

where

$$C_j(x) = \begin{pmatrix} c_{mj}(x) & c_{mj+1}(x) & \dots & c_{mj+m-1}(x) \\ h_{1,mj}(x) & h_{1,mj+1}(x) & \dots & h_{1,mj+m-1}(x) \\ \vdots & \vdots & \vdots & \vdots \\ h_{m-1,mj}(x) & h_{m-1,mj+1}(x) & \dots & h_{m-1,mj+m-1}(x) \end{pmatrix},$$

$$\text{for} \quad j \succeq 0,$$

and

$$A_0(x) = \begin{pmatrix} a_0(x) & a_1(x) & \dots & a_{m-1}(x) \\ 0 & a_0(x) & \dots & a_{m-2}(x) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_0(x) \end{pmatrix},$$

$$A_j(x) = \begin{pmatrix} a_{mj}(x) & a_{mj+1}(x) & \dots & a_{mj+m-1}(x) \\ a_{mj-1}(x) & a_{mj}(x) & \dots & a_{mj+m-2}(x) \\ \vdots & \vdots & \vdots & \vdots \\ a_{mj-m+1}(x) & a_{mj-m+2}(x) & \dots & a_{mj}(x) \end{pmatrix}, \text{ for } j \succeq 1.$$

This induces a different state space description of the M.R.P.: $\{(i, j) : i \geqq 0, 1 \leqq j \leqq m\} \times [0, +\infty[$. The state (i, j) corresponds to a queue length $im + j - 1$ after a transition. For the special case where $m = 1$, $A_n(t)$ and $C_n(t)$, $(n \geqq 0)$, defined above reduce to (1) and (5) in Matendo [11].

For $n \geqq 0$ and $|z| \leqq 1$, let $A_n = A_n(+\infty)$, $C_n = C_n(+\infty)$,

$$A(z) = \sum_{k \succeq 0} A_k z^k, \qquad C(z) = \sum_{k \succeq 0} C_k z^k.$$

$$\text{Put } A = A(1) = \sum_{k \succeq 0} A_k, \qquad C = C(1) = \sum_{k \succeq 0} C_k$$

and

$$\underline{\alpha} = A'(1)\,\underline{e} = \sum_{n \succeq 1} n\, A_n\, \underline{e},$$

where $\underline{e}$ is a column vector of ones.

Remark 2: (*a*) There are many stochastic models (queues, inventories, communication systems, dams, random walks,...) which have embedded M.R.P. with block partitioned transition matrix of the form (11). These models can be efficiently solved using the matrix–analytic techniques developed by Neuts (We refer to Neuts [12] for the details). These computational methods are an alternative to classical closed-form analytical methods (based on an application of Rouché's theorem), for which many of the proposed results are generally not in a computationally tractable form.

The matrix analytic approach to the above models is based on the consideration of the minimal nonnegative (and stochastic) solution G of the nonlinear matrix equation

$$G = \sum_{l \geqq 0} A_l\, G^l, \tag{12}$$

for which the $(j,\, j')$-th entry is the conditional probability that the M.R.P. will eventually reach the set of states $\{(i,\, j) : 1 \leqq j \leqq m\}$ by entering the state $(i,\, j')$ given that it starts in the state $(i+1,\, j)$, $i \geqq 0$. Once the $m \times m$ matrix G is obtained, most important steady-state performance measures (number in the system, response time, etc.) for the related model can be written in algorithmically tractable formulas. We note that the matrix G can be computed by successive substitutions in (12) starting with the nul matrix. We also note that in the scalar case (*i. e.* $m = 1$), we have $G = 1$.

(*b*) The stochastic matrix A is a circulant. Therefore A is doubly stochastic so that its invariant probability vector, denoted by $\underline{\pi}$ (*i. e.* $\underline{\pi}\, \mathrm{A} = \underline{\pi}$, $\underline{\pi}\, \underline{e} = 1$), is given by $\underline{\pi} = m^{-1}\, \underline{e}$. Furthermore, the matrix $I - A + \underline{e}\, \underline{\pi}$ is nonsingular, where I is the $m \times m$ identity matrix.

(c) Since $\rho = \underline{\pi}\,\underline{\alpha} < 1$, $E[S_l] < +\infty\,(1 \leqq l \leqq m)$, $E[V] < +\infty$ and $E[Y] < +\infty$, the M.R.P. $\{(N_n, T_n)\}_{n \succeq 0}$ is positive recurrent and the vector

$$C'(1)\,\underline{e} = \sum_{n \succeq 1} n\,C_n\,\underline{e}$$

is finite. (*See* remark (*b*) on page 140 in Neuts [12].)

From (11), the transition probability matrix of the embedded Markov chain $\{N_n\}_{n \geqq 0}$ is given by

$$Q = Q(+\infty) = \begin{pmatrix} C_0 & C_1 & C_2 & \cdots \\ A_0 & A_1 & A_2 & \cdots \\ 0 & A_0 & A_1 & \cdots \\ 0 & 0 & A_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{13}$$

From (13), we deduce that this Markov chain is irreducible and aperiodic. The invariant probability vector $\underline{x}$, defined by $\underline{x}\,Q = \underline{x}$, $\underline{x}\,\underline{e} = 1$, then exists and is strictly positive. If we partition $\underline{x}$ as $(\underline{x}_0, \underline{x}_1, \underline{x}_2, \ldots)$, where $\underline{x}_i$ $(i \geqq 0)$ is a m-row vector, then $\underline{x}_i$ is given by

$$\underline{x}_i = \underline{x}_0\,C_i + \sum_{l=1}^{i+1} \underline{x}_l\,A_{i+1-l}, \qquad i \succeq 0 \tag{14}$$

3.1. Computation of $\underline{x}_0$

Let G as defined in (12) and let

$$L = \sum_{l \succeq 0} C_l\,G^l.$$

The (j, j')-th element of the irreducible stochastic matrix L is the probability that starting in the state $(0, j)$, the M.R.P. $\{(N_n, T_n)\}_{n \geqq 0}$ eventually returns to the set of states $\{(0, 1), \ldots, (0, m)\}$ by entering the state $(0, j')$.

We have (*see* Schellhaas [15]):

$$\underline{x}_0 = \frac{\underline{l}}{d} \tag{15}$$

where $\underline{l}$ is the invariant probability vector of L and

$$d = 1 + \frac{\underline{l}}{1-\rho}\left[C'(1)\,\underline{e} + (C - I)(I - A + \underline{e}\,\underline{\pi})^{-1}\,\underline{\alpha}\right]$$

3.2. Computation of $\underline{x}_i$, $i \geqq 1$

Once the vector $\underline{x}_0$ has been obtained, the remaining components of $\underline{x}$ are efficiently computed using the following recurrence formula established in Ramaswami [14]:

$$\underline{x}_i = \left[\underline{x}_0 \, \overline{C}_i + \sum_{j=1}^{i-1} \underline{x}_j \, \overline{A}_{i+1-j} \right] (I - \overline{A}_1)^{-1}, \qquad i \succeq 1, \qquad (16)$$

where

$$\overline{C}_l = \sum_{i \succeq l} C_i \, G^{i-l}$$

and

$$\overline{A}_l = \sum_{i \succeq l} A_i \, G^{i-l}, \qquad l \succeq 0.$$

Observe that the transition matrix of the Markov chain embedded at epochs of visists to the set of states $\{0, 1, ..., 2m - 1\}$ is given by

$$\begin{pmatrix} C_0 & \overline{C}_1 \\ A_0 & \overline{A}_1 \end{pmatrix}$$

This finite Markov chain is irreducible so that the inverse $(I - \overline{A}_1)^{-1}$ exists.

3.3. Mean queue length at service completion or inactive phase termination epochs

Let us write the invariant probability vector of Q as $(u_0, u_1, u_2 ...)$. The vector $\underline{x}_i$, $i \geqq 0$, are then given by $\underline{x}_i = (u_{mi}, u_{mi+1}, ..., u_{mi+m-1})$. $\{u_j\}_{j \succeq 0}$ is the stationary distribution of the queue length at a service completion or inactive phase termination epoch. The stationary equations $\underline{x} \, Q = \underline{x}$ can now be written as

$$u_j = u_0 \, c_j + \sum_{l=1}^{m-1} u_l \, h_{l,j} + \sum_{l=m}^{m+j} u_l \, a_{m+j-l}, \qquad j \succeq 0. \qquad (17)$$

The generating function $U_\nu(z) = \sum_{j \succeq 0} u_j z^j$, $|z| \leqq 1$, then satisfies

$$U_\nu(z)[z^m - a(z)] = u_0 [z^m c(z) - a(z)] + \sum_{l=1}^{m-1} u_l [z^m h_l(z) - z^l a(z)] \tag{18}$$

so that, differentiating twice in (18) and setting $z = 1$, the mean queue length $U'_\nu(1)$ at service completion or inactive phase termination epochs is given by

$$\begin{aligned} 2m(1-\rho) U'_\nu(1) &= u_0 [m(m-1) + 2mc'(1) + c''(1) - a''(1)] - m(m-1) \\ &\quad + a''(1) + \sum_{l=1}^{m-1} u_l [m(m-1) + 2m h'_l(1) + h''_l(1) - a''(1) \\ &\quad - 2la'(1) - l(l-1)] \end{aligned} \tag{19}$$

where

$$\begin{aligned} c'(1) &= E[Y], \\ c''(1) &= E[Y^2] - E[Y], \\ a'(1) &= \lambda E[D] E[S_m], \\ a''(1) &= \lambda^2 (E[D])^2 E[S_m^2] + \lambda E[S_m] \big(E[D^2] - E[D]\big), \\ h'_l(1) &= \lambda E[D] E[S_l], \\ h''_l(1) &= \lambda^2 (E[D])^2 E[S_l^2] + \lambda E[S_l] (E[D^2] - E[D]) \end{aligned}$$

$$(1 \preceq l \preceq m-1).$$

Note that the right-hand sides in (18) and (19) depend on the components of the vector $x_0 = (u_0, u_1, \ldots, u_{m-1})$ given in (15).

4. THE QUEUE LENGTH AT POST-DEPARTURE EPOCHS

The queue lengths following departures and the times between departures define another M.R.P. on the state space $\{i \geqq 0\} \times [0, +\infty[$, with block-partitioned transition probability matrix $Q_1(x)$ given by

$$Q_1(x) = \begin{pmatrix} B_0(x) & B_1(x) & B_2(x) & \dots \\ A_0(x) & A_1(x) & A_2(x) & \dots \\ 0 & A_0(x) & A_1(x) & \dots \\ 0 & 0 & A_0(x) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \qquad x \succeq 0, \qquad (20)$$

where

$$B_j(x) = \begin{pmatrix} h_{0,mj}(x) & h_{0,mj+1}(x) & \dots & h_{0,mj+m-1}(x) \\ h_{1,mj}(x) & h_{1,mj+1}(x) & \dots & h_{1,mj+m-1}(x) \\ \vdots & \vdots & \vdots & \vdots \\ h_{m-1,mj}(x) & h_{m-1,mj+1}(x) & \dots & h_{m-1,mj+m-1}(x) \end{pmatrix},$$

$$j \geqq 0,$$

with

$$h_{0,j}(x) = \sum_{l=1}^{m-1} c_l(x) \star h_{l,j}(x) + \sum_{l=m}^{m+j} c_l(x) \star a_{m+j-l}(x), \qquad \text{for} \quad j \succeq 0.$$

We note that $Q_1(x)$ differs from $Q(x)$–given by (11)–only in the entries on the first line: $B_i(x)$ instead of $C_i(x)$. The invariant probability vector $\underline{y} = (\underline{y}_i)_{i \geqq 0}$ of $Q_1 = Q_1(+\infty)$ (*i. e.* $\underline{y}\, Q = \underline{y}$, $\underline{y}\underline{e} = 1$) can be computed in a way similar to that described in section 3 for the computation of the stationary probability vector $\underline{x} = (\underline{x}_i)_{i \geqq 0}$.

Remark 3: Let $\underline{y} = (v_0, v_1, v_2, \dots)$ so that

$$\underline{y}_i = (v_{mi}, v_{mi+1}, \dots, v_{mi+m-1}), \qquad i \geqq 0,$$

where v_j $(j \geqq 0)$ is the stationary probability that j customers remain in the system after a service completion. The following relation, noticed in Matendo [11], holds:

$$u_i = \frac{u_0}{\nu_0}\, \nu_i + u_0\, c_i, \qquad i \succeq 1, \qquad (21)$$

so that

$$\nu_0 = \frac{u_0}{1 - u_0} \qquad (22)$$
$$\nu_i = \left(\frac{1}{1 - u_0}\right)(u_i - u_0\, c_i), \qquad i \succeq 1.$$

To obtain (21), we note that u_0/ν_0 is the probability that a transition is a service completion, u_0 the probability that a transition is an inactive phase termination. Since c_i is the probability that i customers arrive in an inactive phase, the result follows from applying the law of total probability.

The result for v_0 in (22) follows from (21) upon summation over i. The corresponding result for v_i $(i \geqq 1)$ is now obvious.

Using (18), the generating function $V_\nu(z) = \sum_{i \succeq 0} \nu_i\, z^i$ $(|z| \preceq 1)$ is seen to be given by

$$V_\nu(z)\,[z^m - a(z)] = \nu_0 \left[a(z)\,(c(z) - 1) + \sum_{l=1}^{m-1} c_l\,(z^m\, h_l(z) - z^l\, a(z))\right] + \sum_{l=1}^{m-1} \nu_l\,[z^m\, h_l(z) - z^l\, a(z)] \qquad (23)$$

and from (19), we deduce that the mean queue length $V'_\nu(1)$ after service completions is given by

$$\begin{aligned} &2m\,(1 - \rho)\,V'_\nu(1) \\ &\quad = \nu_0\,[c''(1) + 2m\,\rho\,c'(1)] + a''(1) - m\,(m - 1) \\ &\qquad + \sum_{l=1}^{m-1} (\nu_l + \nu_0\, c_l)\,[m\,(m - 1) + 2m\,h'_l(1) + h''_l(1) \\ &\qquad - a''(1) - 2\,la'(1) - l\,(l - 1)] \end{aligned} \qquad (24)$$

5. THE STATIONARY QUEUE LENGTH AT AN ARBITRARY EPOCH

In this section, we relate the stationary queue length distribution at an arbitrary point of time, denoted by $\{p_i\}_{i \geqq 0}$, to the sequence $\{u_i\}_{i \geqq 0}$. As observed in Matendo [11], the distribution $\{p_i\}_{i \geqq 0}$ can also be related to the stationary distribution $\{v_i\}_{i \geqq 0}$ of the queue length at a service completion epoch using the matrix $Q_1(\,.\,)$. Our approach [based on using the matrix $Q(\,.\,)$. instead of the matrix $Q_1(\,.\,)$] is motivated by the fact that the expression for the entries on the first line is simpler in $Q(\,.\,)$ than in $Q_1(\,.\,)$. Once the relation between the sequences $\{p_i\}_{i \geqq 0}$ and $\{u_i\}_{i \geqq 0}$ is obtained, the probabilities p_i can be related to the probabilities v_i, $i \geqq 0$, by using (22).

We now outline the method of obtaining the relation between the sequences $\{p_i\}_{i \geqq 0}$ and $\{u_i\}_{i \geqq 0}$ (We refer to Matendo [11] for the details).

We assume that time $t = 0$ corresponds to a transition in the M.R.P. $\{(N_n, T_n)\}_{n \geqq 0}$ and that $N_0 = i_0 \geqq 0$.

First, we introduce the fundamental mean E^* of $Q(.)$ which is the inner product of the vector $\underline{x}$ and the column vector of the row sum means $\int_0^\infty x\, dQ(x)\, \underline{e}$ of the matrix $Q(\,.\,)$. From (10), it can be easily seen that the first m rows of this column vector are given by $E[V]$, $E[S_1]$, ..., $E[S_{m-1}]$, and all other rows by $E[S_m]$, so that

$$E^* = u_0\, E[V] + \sum_{i=1}^{m-1} u_i\, E[S_i] + E[S_m] \left(1 - \sum_{j=0}^{m-1} u_j\right) \tag{25}$$

We note that in the stationary version of the queue, E^* may be interpreted as the average time between two consecutive transitions (*i. e.* service completion or inactive phase termination). Secondly, we express the time dependent probabilities,

$$P_{i_0 i}(t) = \Pr[N(t) = i | N_0 = i_0], \qquad i \geqq 0,$$

in terms of the Markov renewal matrix corresponding to $Q(\,.)$ and of the functions

$$K_{i_0 i}(t) = \Pr[N(t) = i,\, \theta_1 > t | N_0 = i_0], \qquad \text{for} \quad i \geqq i_0,$$

by using the law of total probability.

Then, by using the key renewal theorem it follows that the limits $p_i = \lim_{t \to \infty} P_{i_0 i}(t)$ exist (and are independent of i_0) and are given by

$$p_0 = \frac{u_0}{\lambda E^*} \tag{26}$$

$$p_i = \frac{1}{E^*}\left\{u_0 \int_0^\infty K_{0\,i}(t)\,dt + \sum_{j=1}^{i} u_j\,h^*_{j,\,i-j}\right\}, \qquad 1 \preceq i \preceq m-1$$

$$p_i = \frac{1}{E^*}\left\{u_0 \int_0^\infty K_{0\,i}(t)\,dt + \sum_{j=1}^{m-1} u_j\,h^*_{j,\,i-j} + \sum_{j=m}^{i} u_j\,a^*_{i-j}\right\}, \qquad i \succeq m, \tag{27}$$

provided that the functions $t \to K_{0\,i}(t)$ be directly Riemann integrable.

The generating function $P_\nu(z) = \sum_{i \succeq 0} p_i\,z^i$ $(|z| \preceq 1)$ satisfies

$$\begin{aligned} P_\nu(z) = p_0 + \frac{1}{E^*}\Bigg\{u_0 \int_0^\infty K_0(t,\,z)\,dt + [\lambda - \lambda\,D(z)]^{-1} \\ \times \Bigg[\sum_{j=1}^{m-1} u_j\,z^j\,(1 - h_j(z)) \\ + \left(U_\nu(z) - \sum_{j=0}^{m-1} u_j\,z^j\right)(1 - a(z))\Bigg]\Bigg\} \end{aligned} \tag{28}$$

where $K_0(t,\,z) = \sum_{i \succeq 1} K_{0\,i}(t)\,z^i$. Then, by differentiating both members in (28), we can obtain the mean queue length $\sum_{i \succeq 1} i p_i = P'_\nu(1)$ once $\int_0^\infty K_0(t,\,z)\,dt$ is known.

6. SPECIAL CASE

In this section, we will concern ourselves with the distribution $\{c_k\}_{k \geq 1}$ of Y (the total number of customers arriving in a typical inactive phase) and the functions $K_{0i}(t)$, $i \geqq 1$, for the following exhaustive service vacation policy. Upon returning from the $(i-1)$ st consecutive vacation $(i \geqq 1)$ in a given inactive phase, the server becomes active immediately if he finds at least N customers waiting. Otherwise, he decides to take another vacation–of random length–with probability σ_i and remains in the system instead with probability $q_i = 1 - \sigma_i$. In the latter case, the server remains inactive, inspecting the queue, until N customers are present. Vacation lengths are assumed to be i.i.d. random variables with distribution function $U(.)$ and L.S.T. $\tilde{U}(.)$. This extends Kella's scheme (*see* Kella [4]) to the case where the decision of whether to take a vacation or not is allowed to depend on the number of vacations already taken and the number of customers waiting, compared to a specified number N.

The $(T(SV); N)$-policy and the $(T(MV); N)$-policy introduced in Loris-Teghem [8] are special cases of the above description. For the former model $\sigma_1 = 1$ and $\sigma_i = 0\,(i \geqq 2)$, while for the latter model $\sigma_i = 1\,(i \geqq 1)$. Note also that the special case with $q_1 = 1$ gives the N-policy model. q_1 is, indeed, the probability that once the system becomes empty, the server stays inactive in the system, until N customers are present.

Let

$$b_i^{(r)} = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^i}{i!} \, dU^{(r)}(t), \qquad i \succeq 0, \quad r \succeq 0,$$

(with $b_0^{(r)} = \delta_{0r}$ the Kronecker delta), be the probability that i batches accumulate during r successive vacations, where $U^{(r)}(t)$ denotes the r-th power of convolution of $U(t)$.

We have:

$$1) \quad c_k = \sum_{r \succeq 0} \left(\prod_{l=1}^{r+1} \sigma_l \right) \sum_{n=1}^{k} \sum_{i=0}^{\min(n,N)-1} b_i^{(r)} b_{n-i} \sum_{j=i}^{N-1} d_j(i)\, d_{k-j}(n-i)$$
$$+ \sum_{r \succeq 0} \left(\prod_{l=1}^{r} \sigma_l \right) q_{r+1} \sum_{n=1}^{N} \left(\sum_{i=0}^{n-1} b_i^{(r)} \right) \sum_{j=n-1}^{N-1} d_j(n-1)\, d_{k-j},$$
$$\text{for} \quad k \succeq N, \tag{29}$$

where $\prod_{l=1}^{0} \sigma_l = 1$.

2) $K_{0i}(t) = P(i, t), \quad \text{for} \quad 1 \preceq i \preceq N-1$

$$= \sum_{r \geqq 0} \left(\prod_{l=1}^{r+1} \sigma_l \right) \sum_{k=0}^{N-1} \int_0^t P(k, y) P(i-k, t-y) [1 - U(t-y)] . dU^{(r)}(y), \quad \text{for} \quad i \geqq N, \tag{30}$$

To obtain (29), assume that the number of batches arrived in an inactive phase is $n\,(n \geqq 1)$. Then, the first term in (29) corresponds to the case where the inactive phase consists of $(r+1)$ vacations. During the first r vacations, there are i arrival batches $(i < n,\ i < N)$ and the total number of customers in these i batches is $j\,(j < N)$. During the last vacation, there are $(n-i)$ batches with $(k-j)$ customers so that the queue length at the end of the inactive phase is k. The second term corresponds to the case where the inactive phase consists of r successive vacations followed by an inspection period. During the r vacations, there are i arrival batches $(i < n)$. The total number of customers in the first $(n-1)$ batches is $j\,(j < N)$. In order that $Y = k$, the n-th arriving group must have $(k-j)$ customers. The result follows by application of the law of total probability.

To obtain (30), we note that since $N_0 = 0$, θ_1 is the duration of the first inactive phase $(i.e.\, V_1)$. For $1 \leqq i \leqq N-1$,

$$K_{0i}(t) = \Pr[N(t) = i,\ V_1 > t | N_0 = 0]$$

is merely the probability that there are i arrivals in $]0, t]$ $[i.e.\, P(i, t)]$.

For $i \geqq N$, in order that $N(t) = i$ and the duration $\theta_1 = V_1$ of the inactive phase [consisting of $(r+1)$ successive vacations for some $r \geqq 0$] is greater than t, there must be $k\,(k < N)$ arrivals during the first r vacations (the r-th vacation ends at time y). The last vacation (which starts at y), is still in course at time t and during the interval $[y, t]$, $(i-k)$ arrivals occur.

We present below the particular results for the three special cases mentioned above:

***The *(T (SV); N)*-model**

Let $\sigma_1 = 1$ and $\sigma_i = 0$, $i \geqq 2$. Then from (29) and (30) we get (note that $b_i^{(0)} = \delta_{0i}$, $d_j(0) = \delta_{0j}$ and $b_i^{(1)} = b_i$)

$$c_k = \sum_{n=1}^{k} b_n\, d_k(n) + \sum_{n=1}^{N} \left(\sum_{i=0}^{n-1} b_i\right) \sum_{j=n-1}^{N-1} d_j(n-1)\, d_{k-j}, \qquad k \succeq N \tag{31}$$

and

$$K_{0\,i}(t) = P(i, t), \qquad \text{for} \quad 1 \preceq i \preceq N-1$$
$$= P(i, t)\,[1 - U(t)], \qquad \text{for} \quad i \succeq N \tag{32}$$

***The (*T (MV); N*)-model**

Let $\sigma_i = 1$, $i \geqq 1$. Then from (29) and (30) it follows that (note that the second term in (29) reduces to 0)

$$c_k = \sum_{n=1}^{k} \sum_{i=0}^{\min(n, N)-1} \chi_i\, b_{n-i} \sum_{j=i}^{N-1} d_j(i)\, d_{k-j}(n-i), \qquad k \succeq N \tag{33}$$

where

$$\chi_i = \sum_{r \succeq 0} b_i^{(r)}, \qquad i \succeq 0$$

and

$$K_{0\,i}(t) = P(i, t), \qquad \text{for} \quad 1 \preceq i \preceq N-1$$
$$= \sum_{r \succeq 0} \sum_{k=0}^{N-1} \int_0^t P(k, y)\, P(i-k, t-y) \times [1 - U(t-y)]\, dU^{(r)}(y), \qquad \text{for} \quad i \geqq N \tag{34}$$

Remark 4: (*a*) Results (31) and (33) were obtained by Loris-Teghem [10], while (32) and (34) agree with the results given in Matendo [11].

(*b*) When the service time distributions do not depend on the group sizes ($i.e.\, h_l(z) \equiv a(z)$, $1 \leqq l \leqq m-1$), it follows from (23) that

$$[z^m - a(z)]\, V_\nu(z) = a(z) \left\{ \nu_0 \left[c(z) - 1 + \sum_{l=1}^{m-1} c_l \left(z^m - z^l \right) \right] + \sum_{l=1}^{m-1} \nu_l \left[z^m - z^l \right] \right\} \tag{35}$$

Let $N = 1$. From (31) we get

$$c_k = \sum_{n=1}^{k} b_n\, d_k(n) + b_0\, d_k, \qquad \text{for} \quad k \geqq 1,$$

so that

$$c(z) = \tilde{U}[\lambda - \lambda D(z)] - b_0 (1 - D(z))$$

for the $T(SV)$-policy, and from (33) we get for the $T(MV)$-policy

$$c_k = (1 - b_0)^{-1} \sum_{n=1}^{k} b_n\, d_k(n), \qquad \text{for} \quad k \geqq 1,$$

so that

$$c(z) = (1 - b_0)^{-1} (\tilde{U}[\lambda - \lambda D(z)] - b_0)$$

Substitutions in (35) lead, for both models, to the results obtained by Chatterjee and Mukkerjee [1].

***The *N*-policy**

Let $q_1 = 1$ (so that $\sigma_i = 0$, $i \geqq 1$). Then from (29) and (30) it is clear that

$$c_k = \sum_{n=1}^{N} \sum_{j=n-1}^{N-1} d_j(n-1)\, d_{k-j}, \qquad \text{for} \quad k \succeq N \tag{36}$$

and

$$\left. \begin{aligned} K_{0\,i}(t) &= P(i, t), && \text{for} \quad 1 \preceq i \preceq N-1 \\ &= 0, && \text{for} \quad i \succeq N \end{aligned} \right\} \tag{37}$$

Remark 5: The results for the N-policy model can be derived from those for the $(T(SV);\ N)$-policy by putting $U(\,.\,) \equiv 1$, $b_0 = 1$ and $b_i = 0$, $i \geqq 1$.

7. APPLICATIONS

Bulk arrival, bulk service queues arise in transportation problems (involving buses, airplanes, trains, ships, elevators, shuttles, car ferries...), manufacturing, production, inventory, computer and communication systems (processing of computer programs, polling, local area networks...), etc. We refer to Deb [2], Powel and Humblet [13], Stidham Jr. [16] and Lee and Srinivasan [7] for further details and other useful applications. It is well known (*see* Teghem Jr. [18]) that these systems can be efficiently analyzed when viewed as bulk queueing systems with vacations.

We mention that Lee and Srinivasan [6] considered the problem of determining optimal control policies for the batch arrival *M/G/*1 queue in which the server applies the N-policy and the $(T(MV);\ N)$-policy. Costs are incurred each time the server is turned on ($i.e.$, becomes available to the customers), for waiting customers and when the server is present at the queue. For both models, they derived the mean waiting time of an arbitrary customer for a given value of N and, under the above cost structure, obtained the stationary optimal policy.

In this section, we discuss two applications of our results to the modelling of some production, computer and communication systems considered in Doshi [3]. We also present a simple numerical example to illustrate the tractability of the given results.

7.1. Maintenance in production systems

Consider a machine used to produce a variety of items. When the machine becomes idle, it undergoes preventive maintenance (referred to as a vacation). If, on completion of this maintenance, some items are present, then the machine immediately starts to process the items exhaustively. Otherwise, it waits for the first item to start processing. The resulting model is a single vacation model.

7.2. Maintenance in computer and communication systems

Consider the following application used in computer and communication systems to schedule primary jobs (processing telephone calls, processing interactive and batch jobs, receiving and transmitting data,...) and maintenance work. Here, the maintenance work is divided into short segments. Jobs arrive to a central processor (server) for being processed. Whenever the system gets empty, the processor does a segment of the maintenance work (*i. e.*, takes a vacation). If, on completion of this segment, some primary jobs are present, then the processor serves the primary jobs until the system becomes empty again. Otherwise, it begins to work on a second maintenance segment and keeps on doing these segments until, on return from a maintenance segment, if finds at least one primary job waiting. The resulting model is a multiple vacation model.

We mention that for both applications (7.1 and 7.2), if the vacation distribution is explicitely given, our results can be applied to solve them if we assume that requests for service (items, jobs) arrive in batches to the machine (processor) according to a Poisson process with rate λ and are processed in groups of maximum size m (as arises, as mentioned above, in many practical applications). In both cases, we have a bulk arrival, bulk service *M/G/*1 queue, with single vacation and multiple vacations, respectively.

Moreover, a threshold (N) can be introduced so that the machine (processor) starts service at the end of a vacation only if it finds at least $N\ (N \geqq 1)$ items (jobs) waiting upon returning from a vacation. Otherwise, for the single vacation model with a threshold [*i. e.*, the $(T\,(SV);\ N)$-policy], the machine remains inactive, continuously monitoring the system, until at least N items have accumulated. For the multiple vacation model with a threshold [*i. e.*, the $(T\,(MV);\ N)$-policy], the maintenance segments are scheduled repeatedly until the processor finds at least N jobs waiting.

7.3. A numerical example

We assume that customers arrive in groups according to a Poisson process with rate $\lambda = 3.0$. The probability that the batch size equals j is $d_j = 0.2$, for $j = 1, 2, ..., 5$. The customers are served in batches of maximum size $m = 2$. The service time distribution does not depend on the group size. The service times are i.i.d. random variables with an Erlang distribution with two stages and a mean of 0.2. Thus, the traffic intensity is equal to 0.9. The server applies the 2-policy (*i. e.* $N = 2$). For this problem, the queue length distributions are presented in the appendix.

8. CONCLUDING REMARKS

This paper has been devoted to a single-server bulk arrival, bulk service queue in which the server applies a general exhaustive service vacation policy. Our analysis is based on the following approach:

a) use Neuts' method (matrix-analytic methodology) to derive the stationary queue length at service completion or inactive phase termination epochs;

b) relate the steady-state distributions of the queue length at post-departure epochs and at an arbitrary epoch to the former distribution.

These computational results extend previous results for vacation models with bulk arrivals (Matendo [11]).

In particular, we considered a specific vacation policy which is an extension of Kella's vacation (Kella [4]). We showed that our results can be applied to study maintenance in production, computer and communication systems. We provided a simple numerical example to illustrate the practical use of the tractable results derived in this paper.

Our planned extension of this work is to allow the input process to be a more general (bulk) arrival process including bursty arrival processes which commonly arise in packet voice technology and communication engineering. We also are currently analyzing the waiting time distributions of the model under consideration. We expect to present these results in forthcoming papers, where also other applications and numerical results will be discussed.

REFERENCES

1. U. Chatterjee and S. P. Mukkerjee, Two bulk queueing models with vacation periods, *Cahiers du C.E.R.O., 29*, 1987, p. 85-93.
2. R. K. Deb, Optimal Control of bulk queues with compound Poisson arrivals and batch service, *Opsearch, 21*, (4), 1984, p. 227-245.
3. B. T. Doshi, Queueing systems with vacations. A survey, *Queueing Syst., 1,* 1986, p. 29-66.
4. O. Kella, Optimal control of the vacation scheme in an *M/G/*1 queue, *Opns. Res., 38,* 1990, p. 724-728.
5. G. Latouche, Distributions de type phase: tutorial, *Cahiers du C.E.R.O., 31,* 1989, p. 3-11.
6. H. S. Lee and M. M. Srinivasan, Control policies for the $M^x/G/1$ queueing system, *Management Science, 35*, (6), 1989, p. 708-721.
7. H. S. Lee and M. M. Srinivasan, The shuttle dispatch problem with compound Poisson arrivals: controls at two terminals, *Queueing Systems,* (6), 1990, p. 207-222.
8. J. Loris-Teghem, Analysis of single server queueing systems with vacation periods, *Belg. Journ. of Oper. Res., Stat. and Comput. Sc., 25* 1985, p. 47-54.
9. J. Loris-Teghem, On vacation models with bulk arrivals, *Belg. journ. of Oper. Res., Stat. and Comput. Sc., 30,* (1), 1990 *a*, p. 53-66.
10. J. Loris-Teghem, Remark on: On vacation models with bulk arrivals, *Belg. Journ. of Oper. Res., Stat. and Comput. Sc., 30,* (4), 1990 *b*, p. 53-56.
11. S. K. Matendo, Application of Neut's method to vacation models with bulk arrivals, *Belg. Journ. of Oper. Res., Stat. and Comput. Sc., 31*, (1-2), 1991, p. 34-48.

12. M. F. NEUTS, *Structured stochastic matrices of M/G/1 type and their applications,* Marcel Dekker, New York, 1989.
13. W. B. POWEL and P. HUMBLET, The bulk service queue with a general control strategy: Theoretical analysis and a new computational procedure, *Opns. Res., 34,* (2), 1986, p. 267-275.
14. V. RAMASWAMI, A stable recursion for the steady state vector in Markov chains of *M/G/*1 type, *Stoch. Mod., 4,* 1988, p. 183-188.
15. H. SCHELLHAAS, On Ramaswami's algorithm for the computation of the steady-state vector in Markov chains of *M/G/*1-type, *Stoch. Mod., 6,* 1990, p. 541-550.
16. S. STIDHAM Jr., Clearing systems and (s, S) inventory systems with nonlinear costs and positive lead times, *Opns. Res., 34,* (2), 1986, p. 276-280.
17. H. TAKAGI, *Queueing analysis, vol. 1: Vacation and Priority Systems,* Part. 1, North Holland, 1991.
18. J. TEGHEM Jr., Control of the service process in a queueing system, *Eur. J. Opns. Res., 23,* 1986, p. 141-158.

APPENDIX

The queue length distributions

i	$u_i = \Pr$ [i in the system at service compl. or inact. phase terminat. epochs]	$v_i = \Pr$ [i in the system at post-departure epochs]	$p_i = \Pr$ [i in the system at an arbitrary epoch]
0	.044293	.046346	.070691
1	.031676	.033144	.034779
2	.043118	.033993	.029463
3	.046624	.037662	.033690
4	.043315	.034200	.033745
5	.044333	.035265	.036608
6	.033517	.033217	.031936
7	.030658	.032079	.030756
8	.029640	.031014	.030115
9	.028234	.029543	.028909
10	.027137	.028395	.027835
11	.025952	.027155	.026458
12	.024856	.026008	.025377
13	.023832	.024937	.024357
14	.022768	.023823	.023291
15	.021826	.022838	.022316
16	.020853	.021819	.021329
17	.019977	.020903	.020430
18	.019092	.019977	.019538
19	.018293	.019141	.018711
20	.017481	.018291	.017891
21	.016750	.017526	.017133
22	.016007	.016749	.016382

i	$u_i = \Pr$ [i in the system at service compl. or inact. phase terminat. epochs]	$v_i = \Pr$ [i in the system at post-departure epochs]	$p_i = \Pr$ [i in the system at an arbitrary epoch]
23	.015337	.016048	.015688
24	.014657	.015336	.015001
25	.014044	.014695	.014365
26	.013421	.014043	.013736
27	.012860	.013456	.013154
28	.012290	.012860	.012578
29	.011775	.012321	.012045
30	.011253	.011775	.011517
31	.010782	.011282	.011029
32	.010303	.010781	.010545
33	.009872	.010330	.010098
34	.009434	.009871	.009655
35	.009039	.009458	.009246
36	.008638	.009038	.008840
. . .	. . .	. . .	. . .
The proba. that there are more than 36 customers	.186063	.194684	.190765
The mean queue lengths	21.535700	22.367000	21.871950