Probability in the Engineering and Informational Sciences, **18**, 2004, 299–313. Printed in the U.S.A.

WAITING TIME CHARACTERISTICS IN CYCLIC QUEUES

Sanne R. Smits
Department of Technology Management
Eindhoven University of Technology
5600 MB, Eindhoven, The Netherlands

Ivo Adan
Department of Mathematics and Computer Science
Eindhoven University of Technology
5600 MB, Eindhoven, The Netherlands
E-mail: iadan@win.tue.nl

Ton G. de Kok
Department of Technology Management
Eindhoven University of Technology
5600 MB, Eindhoven, The Netherlands

In this article, we study a single-server queue with FIFO service and cyclic interarrival and service times. An efficient approximative algorithm is developed for the first two moments of the waiting time. Numerical results are included to demonstrate that the algorithm yields accurate results. For the special case of exponential interarrival times, we present a simple exact analysis.

1. INTRODUCTION

The present study concerns a multiclass queuing model with cyclic interarrival and service times. This model can be used, for example, when the inflow of customers depends on the day of the week or on the hour of the day. More specifically, this model arises in the modeling of a manufacturing system producing replenishment orders for stock locations that are controlled by periodic order-up-to policies. In that situation, the interarrival times are typically deterministic, and the order sizes are location dependent.

© 2004 Cambridge University Press 0269-9648/04 $16.00

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

Cyclic queuing models have been studied by Morrice and Gajulapalli [3]. They considered a model with cyclic exponential interarrival and service times and derived bounds and exact results for the mean number in the system. Cohen [1] also studied the cyclic model and presented functional equations for the stationary waiting time distributions. These equations formulate a Hilbert boundary value problem, which can be solved if the Laplace–Stieltjes transforms (LSTs) of all the interarrival time distributions or all of the service time distributions are rational.

The central equation in this article is Lindley's equation for the waiting times. It is used to derive an iterative method to compute approximations for the first two moments of the waiting time. This method extends the one developed by de Kok [2]; it is applicable under generally distributed interarrival and service times and produces accurate results. Lindley's equation is also used to exactly determine the moments of the waiting time in the case of exponential (or Erlang) interarrival times.

The article is organized as follows. Section 2 provides a description of the model and introduces some notation. In Section 3, we present the moment-iteration method. In Section 4, we treat the special case of exponential (or Erlang) interarrival times and present an exact method for the computation of the moments of the waiting time. Numerical results are presented in Section 5, where we compare the approximations produced by the moment-iteration method with the exact results for the model with Erlang interarrival and service times and with simulation results for uniform and discrete interarrival time distributions. Finally, Section 6 is devoted to an application of the moment-iteration method.

2. MODEL DESCRIPTION

We consider a single-server queue providing FIFO service to N types of customer, numbered $1, \ldots, N$. The customers arrive in a cyclic pattern: first a type 1 customer, then one of type 2, then type 3 until type N, and then the cycle repeats. Define, for $1 \leq i \leq N$ and $k \geq 1$, the following:

$A_{i,k}$ is the time between the arrival of the kth type i customer and the previous arrival.

$B_{i,k}$ is the service time of the kth type i customer.

$W_{i,k}$ is the waiting time of the kth type i customer.

$S_{i,k}$ is the sojourn time of the kth type i customer ($= W_{i,k} + B_{i,k}$).

For each i, both $\{A_{i,k}\}_{k\geq 1}$ and $\{B_{i,k}\}_{k\geq 1}$ are sequences of independent identically distributed (i.i.d.) random variables. The two sequences are also independent of each other and independent of the sequences for different i's. For stability, we assume that the traffic intensity ρ is less than one:

$$\rho = \frac{\sum_{i=1}^{N} E[B_i]}{\sum_{i=1}^{N} E[A_i]} < 1, \tag{1}$$

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

where the generic random variables A_i and B_i have the same distribution as $A_{i,k}$ and $B_{i,k}$, respectively. It is readily verified that for $k \geq 1$,

$$W_{1,k} = (S_{N,k-1} - A_{1,k})^+,$$
$$W_{i,k} = (S_{i-1,k} - A_{i,k})^+, \qquad i = 2,\ldots,N, \tag{2}$$

where $(x)^+ = \max(x,0)$. These equations are the starting point for the iterative method presented in Section 3. This method approximates the first two moments of the stationary waiting times:

$$E[W_i] = \lim_{k\to\infty} E[W_{i,k}], \qquad E[W_i^2] = \lim_{k\to\infty} E[W_{i,k}^2], \qquad i = 1,\ldots,N;$$

the limits exist by virtue of stability condition (1).

3. MOMENT-ITERATION METHOD

Equation (2) relates the waiting time of a customer to the sojourn time of the previous customer. From this equation we get the following expression for the nth moments of $W_{i,k}$:

$$E[W_{1,k}^n] = \int_0^\infty \int_z^\infty (x-z)^n \, dF_{S_{N,k-1}}(x)\, dF_{A_{1,k}}(z),$$
$$E[W_{i,k}^n] = \int_0^\infty \int_z^\infty (x-z)^n \, dF_{S_{i-1,k}}(x)\, dF_{A_{i,k}}(z), \qquad i = 2,\ldots,N. \tag{3}$$

Here, we concentrate on the first two moments (so $n = 1,2$). If the first two moments of the sojourn time of the previous customer are known and we fit a tractable distribution to these two moments, then the above expressions with the fitted distribution can be used to compute an approximation for the first two moments of the waiting (and sojourn) time of the present customer. For the two-moment fit, we may use a mixed Erlang or hyperexponential distribution, depending on whether the squared coefficient of variation is less or greater than one (see, e.g., Tijms [4]). More specifically, let $E(S)$ and c_S^2 denote the mean and squared coefficient of variation of the sojourn time of the previous customer. If $1/k \leq c_S^2 \leq 1/(k-1)$ for some $k = 2,3,\ldots$, then the mean and squared coefficient of variation of the mixed Erlang distribution $\tilde{F}_S(\cdot)$ with density

$$\tilde{f}_S(t) = p\mu^{k-1} \frac{t^{k-2}}{(k-2)!} e^{-\mu t} + (1-p)\mu^k \frac{t^{k-1}}{(k-1)!} e^{-\mu t}, \qquad t \geq 0, \tag{4}$$

matches $E(S)$ and c_S^2, provided the parameters p and μ are chosen as

$$p = \frac{1}{1+c_S^2}\left[kc_S^2 - \{k(1+c_S^2) - k^2 c_S^2\}^{1/2}\right], \qquad \mu = \frac{k-p}{E(S)}.$$

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

If $c_S^2 > 1$, then the mean and squared coefficient of variation of the hyperexponential distribution $\tilde{F}_S(\cdot)$ with density

$$\tilde{f}_S(t) = p_1 \mu_1 e^{-\mu_1 t} + p_2 \mu_2 e^{-\mu_2 t}, \qquad t \geq 0, \tag{5}$$

matches $E(S)$ and c_S^2, provided the parameters μ_1, μ_2, p_1, and p_2 are chosen as

$$\mu_1 = \frac{2}{E(S)}\left(1 + \sqrt{\frac{c_S^2 - \frac{1}{2}}{c_S^2 + 1}}\right), \qquad \mu_2 = \frac{4}{E(S)} - \mu_1,$$

$$p_1 = \frac{\mu_1\{\mu_2 E(S) - 1\}}{\mu_2 - \mu_1}, \qquad p_2 = 1 - p_1.$$

This distribution is called the hyperexponential distribution with the gamma normalization because it has the property that its third moment as well matches that of the gamma distribution with mean $E(S)$ and squared coefficient of variation c_S^2.

The above procedure is then repeated for the next customer and so on. The resulting iteration scheme is as follows.

Iteration scheme

1. Initially, set $E[W_{i,0}] = E[W_{i,0}^2] = 0$ for $i = 1, \ldots, N$ and set $i = k = 1$.
2. Fit a tractable distribution to the first two moments of the sojourn time of the previous customer: Compute the first two moments and squared coefficient of variation of $S_{N,k-1}$ if $i = 1$ and of $S_{i-1,k}$ if $i > 1$. Then, according to (4) and (5), the fitted distribution $\tilde{F}_{S_{N,k-1}}(\cdot)$ or $\tilde{F}_{S_{i-1,k}}(\cdot)$ is a mixture of two Erlang distributions with the same scale parameter if the squared coefficient of variation is less than one; otherwise it is a hyperexponential distribution with the gamma normalization.
3. Compute $E[W_{i,k}]$ and $E[W_{i,k}^2]$ according to (3), with $F_{S_{N,k-1}}(\cdot)$ and $F_{S_{i-1,k}}(\cdot)$ replaced by the fitted distributions $\tilde{F}_{S_{N,k-1}}(\cdot)$ and $\tilde{F}_{S_{i-1,k}}(\cdot)$, respectively.
4. If $i < N$, then set $i = i + 1$ and go to step 2; if $i = N$, compute the two sums $\sum_{j=1}^{N} |E[W_{j,k-1}] - E[W_{j,k}]|$ and $\sum_{j=1}^{N} |E[W_{j,k-1}^2] - E[W_{j,k}^2]|$. If both are sufficiently small, then stop and use $E[W_{i,k}]$ and $E[W_{i,k}^2]$ as the approximation for $E[W_i]$ and $E[W_i^2]$, respectively, for $i = 1, \ldots, N$; otherwise set $k = k + 1$ and $i = 1$ and go to step 2.

4. EXPONENTIAL INTERARRIVAL TIMES

In this section, we consider the special case that the interarrival time A_i is exponentially distributed with mean $1/\lambda_i$, $i = 1, \ldots, N$. Let W_i and S_i be the waiting time and the sojourn time in steady state of a type i customer, with LSTs $W_i(s)$ and $S_i(s)$, respectively. Letting $k \to \infty$ in (2), it follows that

$$W_i = (S_{i-1} - A_i)^+, \qquad i = 1, \ldots, N,$$

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

where, by convention, a type 0 customer is the same as a type N customer. From these equations we get

$$\begin{aligned}
W_i(s) &= E(e^{-s(S_{i-1}-A_i)^+}) \\
&= P(S_{i-1}-A_i<0) + E(e^{-s(S_{i-1}-A_i)}1_{[S_{i-1}-A_i\geq 0]}) \\
&= P(S_{i-1}-A_i<0) + E(e^{-s(S_{i-1}-A_i)}) - E(e^{-s(S_{i-1}-A_i)}1_{[S_{i-1}-A_i<0]}) \\
&= P(A_i>S_{i-1}) + \frac{\lambda_i}{\lambda_i-s}S_{i-1}(s) \\
&\quad - E(e^{s(A_i-S_{i-1})}|A_i>S_{i-1})P(A_i>S_{i-1})
\end{aligned}$$

for the transforms. By the memoryless property, the overshoot $(A_i - S_{i-1}|A_i > S_{i-1})$ is again exponential with parameter λ_i, so

$$\begin{aligned}
W_i(s) &= P(A_i>S_{i-1}) + \frac{\lambda_i}{\lambda_i-s}S_{i-1}(s) - \frac{\lambda_i}{\lambda_i-s}P(A_i>S_{i-1}) \\
&= \frac{\lambda_i}{\lambda_i-s}S_{i-1}(s) - \frac{s}{\lambda_i-s}P(A_i>S_{i-1}).
\end{aligned}$$

Note that $P(A_i > S_{i-1})$ is the probability that a type i customer does not have to wait. Let us write

$$P(A_i > S_{i-1}) = 1 - \Pi_i$$

(Π_i is the probability of waiting). Further, using $S_i(s) = W_i(s)B_i(s)$ where $B_i(s)$ is the LST of B_i, the equations for the transforms $W_i(s)$ can be written in the form

$$(s-\lambda_i)W_i(s) + \lambda_i B_{i-1}(s)W_{i-1}(s) = (1-\Pi_i)s, \qquad i = 1,2,\ldots,N. \tag{6}$$

From these equations we can solve $W_N(s)$, yielding

$$W_N(s) = \frac{\sum_{i=1}^{N}(1-\Pi_i)s/\lambda_i \prod_{j=1}^{i-1}(1-s/\lambda_j)\prod_{j=i}^{N-1}B_j(s)}{\prod_{i=1}^{N}B_i(s) - \prod_{i=1}^{N}(1-s/\lambda_i)}. \tag{7}$$

Of course, the other transforms $W_i(s)$ are given by similar (symmetrical) expressions. To determine the unknown probabilities Π_i, we proceed as follows. First, we have to satisfy

$$W_N(0) = 1. \tag{8}$$

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

For the denominator in (7), it holds that

$$\prod_{i=1}^{N} B_i(s) - \prod_{i=1}^{N}\left(1 - \frac{s}{\lambda_i}\right) = \left(\sum_{i=1}^{N} \frac{1}{\lambda_i} - \sum_{i=1}^{N} E[B_i]\right) s + O(s^2), \qquad s \to 0, \tag{9}$$

and for the numerator,

$$\sum_{i=1}^{N} \frac{(1-\Pi_i)s}{\lambda_i} \prod_{j=1}^{i-1}\left(1 - \frac{s}{\lambda_j}\right) \prod_{j=i}^{N-1} B_j(s) = s \sum_{i=1}^{N} \frac{1-\Pi_i}{\lambda_i} + O(s^2), \qquad s \to 0.$$

Hence, from (8) we get

$$\sum_{i=1}^{N} \frac{1-\Pi_i}{\lambda_i} = \sum_{i=1}^{N} \frac{1}{\lambda_i} - \sum_{i=1}^{N} E[B_i]. \tag{10}$$

Further, since $W_N(s)$ is well defined for $\mathrm{Re}(s) \geq 0$, it follows that whenever the denominator in (7) vanishes for some s with $\mathrm{Re}(s) \geq 0$, the numerator should also vanish. In the Appendix, we will prove that the denominator has exactly N zeros s with $\mathrm{Re}(s) \geq 0$, say $s_0(=0), s_1, \ldots, s_{N-1}$. We assume that these zeros are all distinct. Because the numerator of (7) must also vanish at $s = s_1, \ldots, s_{N-1}$, we obtain the following equations:

$$\sum_{i=1}^{N} \frac{(1-\Pi_i)s_k}{\lambda_i} \prod_{j=1}^{i-1}\left(1 - \frac{s_k}{\lambda_j}\right) \prod_{j=i}^{N-1} B_j(s_k) = 0, \qquad k = 1, \ldots, N-1.$$

Together with (10), this forms a set of N equations for N waiting probabilities $\Pi_1, \ldots, \Pi_N$; it has a unique solution, because under the condition of stability (1), there is a unique stationary waiting time distribution and thus also a unique solution $W_i(s)$. This completes the determination of the transforms $W_i(s)$, as given by (7).

In the remainder of this section, we show how the moments of the waiting times can be determined. As a starting point, we take (6). Substituting the Taylor series

$$W_i(s) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} E[W_i^k] s^k$$

and

$$B_{i-1}(s) W_{i-1}(s) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sum_{j=0}^{k} \binom{k}{j} E[W_{i-1}^j] E[B_{i-1}^{k-j}] s^k,$$

we obtain

$$\sum_{k=0}^{\infty} \frac{(-1)^{k-1}}{k!} \left(k E[W_i^{k-1}] + \lambda_i E[W_i^k] - \lambda_i \sum_{j=0}^{k} \binom{k}{j} E[W_{i-1}^j] E[B_{i-1}^{k-j}] \right) s^k$$
$$= (1-\Pi_i)s.$$

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

Equating the coefficients of s^k on the left- and right-hand sides yields

$$1 + \lambda_i(E[W_i] - E[W_{i-1}]) - \lambda_i E[B_{i-1}] = (1 - \Pi_i), \qquad i = 1, \dots, N, \tag{11}$$

and for $k > 1$,

$$kE[W_i^{k-1}] + \lambda_i(E[W_i^k] - E[W_{i-1}^k]) - \lambda_i \sum_{j=0}^{k-1} \binom{k}{j} E[W_{i-1}^j]E[B_{i-1}^{k-j}] = 0, \qquad i = 1, \dots, N. \tag{12}$$

Adding (12) over all $i = 1, \dots, N$ gives the equation

$$\sum_{i=1}^{N} kE[W_i^{k-1}] - \sum_{i=1}^{N} \lambda_{i+1} \sum_{j=0}^{k-1} \binom{k}{j} E[W_i^j]E[B_i^{k-j}] = 0.$$

This can be rewritten as (replace k by $k+1$)

$$\sum_{i=1}^{N} E[W_i^k](k+1)(1 - \lambda_{i+1}E[B_i]) = \sum_{i=1}^{N} \lambda_{i+1} \sum_{j=0}^{k-1} \binom{k+1}{j} E[W_i^j]E[B_i^{k+1-j}], \tag{13}$$

which is valid for $k \geq 1$. Using (11)–(13), all moments $E[W_i^k]$ can now be computed recursively. First, note that the addition of (11) over all i gives an identity; therefore, we can omit one of these equations. Then, together with (13) for $k = 1$, we have a set of N equations, from which the first moments $E[W_i]$, $i = 1, \dots, N$, can be computed. To find the second moments, we use (12) and (13) for $k = 2$, where we can again omit one equation in (12). Also, note that the lower (first) moments occurring in these equations are now known. The third moments follow from (12) and (13) for $k = 3$, and so on. It will be clear that we can repeat using (12) and (13) to successively compute all moments (if they exist).

Remark 4.1 (Erlang Interarrival Times): If the interarrival time A_i is Erlang-$r(\lambda_i)$ distributed, then we can think of the interarrival time as consisting of r subinterarrivals. By associating with each subarrival an arrival of a customer, where the first $r-1$ customers have zero service times and the last, the rth customer, has a service time B_i, we can analyze this case along the same lines as the exponential case.

Remark (Zeros): From (1) and (9) it follows that $s_0 = 0$ is a simple zero of the denominator in (7), but it might happen that some of the other zeros coincide. For example, if $s_1 = s_2$, then $s = s_1$ is also a double zero of the numerator in (7); thus, an additional equation for the probabilities Π_i can be obtained by requiring that the derivative of the numerator also vanishes at $s = s_1$.

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

5. NUMERICAL RESULTS

In this section, we validate the approximative moment-iteration method by simulation (and exact results, if available). Table 1 shows 30 different settings for the mean and squared coefficient of variation of the interarrival and service times for a model with 2 customer types. In each setting, the distributions of the interarrival and service times are matched to the mean and squared coefficient of variation according to the recipe described in Section 3. The results are listed in Table 2; the simulation results are based on 10 independent replicas of 6×10^6 arrivals.

TABLE 1. Different Settings for a Model with Two Customer Types

$E[B_1]$	$c^2_{B_1}$	$E[A_1]$	$c^2_{A_1}$	$E[B_2]$	$c^2_{B_2}$	$E[A_2]$	$c^2_{A_2}$
0.3336	0.9399	1	0.8556	0.7808	1.2154	1	0.3256
0.3262	0.8795	1	1.2677	0.4777	1.3142	1	1.7559
0.7425	0.7869	1	1.9567	0.7672	1.4500	1	1.4146
0.5946	1.6637	1	0.6724	0.3286	0.7612	1	1.8433
0.5593	1.4135	1	1.594	0.4273	1.0227	1	0.3088
0.8825	1.1094	1	1.3081	0.8808	1.1256	1	1.0411
0.3946	0.6998	1	1.2938	0.8865	1.2776	1	0.6950
0.4902	0.9227	1	0.3806	0.7737	1.3612	1	1.7138
0.6998	0.4399	1	1.6961	0.7628	1.6648	1	1.2275
0.4858	0.8035	1	1.3663	0.8056	1.5750	1	1.4856
0.839	1.0517	1	0.4574	0.5701	0.7840	1	1.7097
0.7512	1.2887	1	1.3432	0.3304	1.2458	1	1.3743
0.6172	1.7934	1	1.3552	0.6019	1.1206	1	1.3740
0.8599	0.8801	1	0.5121	0.4384	0.6856	1	1.7459
0.6676	1.1854	1	1.5321	0.6752	0.5605	1	1.5766
0.8033	1.9531	1	1.4527	0.5419	0.8690	1	1.7902
0.5931	0.2726	1	1.3465	0.3258	1.3380	1	0.3060
0.3981	1.0487	1	0.9558	0.8055	0.6195	1	0.3989
0.4461	1.8262	1	1.7511	0.4292	1.3720	1	1.9803
0.4262	1.0165	1	0.4411	0.5804	0.5267	1	1.8240
0.6289	0.4006	1	1.5219	0.3476	0.6719	1	1.2117
0.5666	1.3757	1	1.3174	0.5500	1.6264	1	1.2899
0.8573	0.2094	1	1.8976	0.8043	1.6817	1	0.7575
0.7525	0.7981	1	1.6347	0.6474	0.7044	1	1.4072
0.4007	0.7300	1	0.4923	0.5535	0.4308	1	0.9953
0.3539	0.8127	1	0.7832	0.4452	1.5441	1	0.2640
0.6022	1.2763	1	0.4926	0.7937	1.4976	1	0.9524
0.4668	1.5906	1	1.2386	0.7049	1.6126	1	1.8389
0.7909	0.3568	1	1.3089	0.6698	1.3944	1	1.9856

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

Note that the waiting time characteristics of the two customer types can differ substantially when their interarrival time and service time characteristics are different. Table 2 indicates that the approximation method produces accurate estimates for the first two moments of the waiting times. This is confirmed by a much more extensive investigation, the results of which are now summarized.

We consider models with deterministic, uniform, exponential or (mixed) Erlang, and hyperexponential distributions for the interarrival and service times. The models are indicated with the well-known three-letter code $A/B/1$; it means that

TABLE 2. Results for Different Settings of a Model with Two Customer Types

Moment Iteration				Simulation			
$E[W_1]$	$E[W_1^2]$	$E[W_2]$	$E[W_2^2]$	$E[W_1]$	$E[W_1^2]$	$E[W_2]$	$E[W_2^2]$
0.6871	1.9184	0.5456	1.4561	0.6895	1.9216	0.5454	1.4612
0.4515	0.8193	0.3727	0.6417	0.4519	0.8182	0.3706	0.6440
3.4108	29.3421	3.4157	29.1897	3.4262	29.3963	3.4278	29.2493
0.5277	1.2622	0.5868	1.5251	0.5235	1.2677	0.5882	1.5222
0.4183	0.9046	0.5459	1.2136	0.4153	0.8989	0.5462	1.2023
7.5649	130.9852	7.5719	131.0929	7.5747	128.8335	7.5817	128.9578
1.4301	6.3987	1.2416	5.4483	1.4279	6.2962	1.2377	5.3500
1.3074	5.3483	1.1130	4.4958	1.3056	5.2616	1.1076	4.4158
2.6543	18.5757	2.6402	18.1529	2.667	18.9389	2.6434	18.5326
1.8505	10.1854	1.6991	9.2076	1.8516	9.9634	1.6959	8.9868
1.6599	7.8088	1.7235	8.2086	1.6649	7.9895	1.7264	8.3945
0.8725	3.0285	1.0504	3.6878	0.8732	2.9375	1.0499	3.5984
1.4049	6.4153	1.4128	6.5671	1.4107	6.2725	1.4230	6.4154
1.1287	3.9331	1.2513	4.4406	1.1282	4.0422	1.2467	4.5577
1.7679	8.1164	1.7641	8.2278	1.7607	8.3844	1.7632	8.4844
2.3407	16.6898	2.4518	17.7260	2.3474	16.1483	2.4634	17.1840
0.1674	0.1904	0.3123	0.3110	0.1698	0.1938	0.3099	0.3213
0.6621	1.4162	0.5440	1.1663	0.6581	1.4618	0.5477	1.2014
0.6787	1.8049	0.6799	1.8501	0.6804	1.7454	0.6836	1.7890
0.4988	0.7705	0.3648	0.5864	0.4957	0.7968	0.3660	0.6085
0.4186	0.6394	0.5387	0.8303	0.4187	0.6668	0.5368	0.8607
1.0300	3.8423	1.0383	3.8382	1.0349	3.6907	1.0404	3.6890
4.5293	47.2243	4.5878	47.3631	4.5074	49.0084	4.5562	49.1670
1.9365	9.4356	1.9944	9.7612	1.9372	9.8670	1.9953	10.1904
0.2915	0.3161	0.2089	0.2306	0.2918	0.3285	0.2107	0.2407
0.1884	0.2785	0.1858	0.2241	0.1916	0.2839	0.1844	0.2357
1.7357	9.5550	1.6233	8.9008	1.7431	9.1600	1.6293	8.5018
1.4752	7.0664	1.3458	6.4338	1.4822	6.7821	1.3538	6.1468
2.5168	15.4079	2.5474	15.3549	2.5147	16.1702	2.5333	16.1406

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

Table 3. Parameter Settings

Model	Parameter	Value
$D/G/1$	$E[B_i]$	$U(0.2,0.99)$
	$c^2_{B_i}$	$U(0.2,2)$
	A_i	1
$U/G/1$	$E[B_i]$	$U(0.3,0.99)$
	$c^2_{B_i}$	$U(0.2,2)$
	$\min A_i$	0.7
	$\max A_i$	1.3
$M/M/1$	$E[B_i]$	$U(0.3,0.99)$
	$E[A_i]$	1
$E_k/M/1$	$E[B_i]$	$U(0.3,0.99)$
	$E[A_i]$	1
	k	$U(\{1,2\})$
$E_k/E_l/1$	$E[B_i]$	$U(0.3,0.99)$
	l	$U\{(1,2,3,4\})$
	$E[A_i]$	1
	k	$U(\{1,2\})$
$H_2/H_2/1$	$E[B_i]$	$U(0.3,0.99)$
	$c^2_{B_i}$	$U(1,2)$
	$E[A_i]$	1
	$c^2_{A_i}$	$U(1,2)$

all customer types have the same type of interarrival time distribution, indicated by A, and the same type of service time distribution, indicated by B, but, of course, the parameters of the distributions depend on the customer type. For each case of interarrival and service time distributions, we randomly generate 10^3 settings of the parameters, and each setting will be evaluated for 2, 5, and 25 customer types. If only mean and squared coefficients of variation are specified, we fit a distribution according to the recipe described in Section 3. Table 3 gives an overview of the parameter settings, where $U(a,b)$ denotes the uniform distribution on (a,b), and $U(S)$ denotes the (discrete) uniform distribution on the points in the set S.

Note that, according to Table 3, we will generate settings with different traffic intensities. We divide the settings into three categories: low load $(0.4 < \rho < 0.6)$, medium load $(0.6 \leq \rho < 0.8)$, and high load $(\rho \geq 0.8)$. We compare the estimates produced by the moment-iteration method with the exact results of Section 4 in case of exponential or Erlang interarrival times and with simulation results otherwise. Again, the simulation results are based on 10 independent replicas of 6×10^6 arrivals. In Tables 4 and 5, we display percentage errors; $\bar{\delta}^{E[W]}$ and $\max \delta^{E[W]}$ denote, respectively, the *mean* and *maximum* percentage error of the average of the

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

TABLE 4. Average Percentage Errors of the Moment-Iteration Method

Load	Ave. Error	N	Model $D/G/1$	$U/G/1$	$M/M/1$	$E_k/M/1$	$E_k/E_l/1$	$H_2/H_2/1$
$0.4 \leq \rho < 0.6$	$\bar{\delta}^{E[W]}$	2	2.99	1.99	2.19	3.59	5.78	0.35
		5	4.14	3.97	0.93	1.54	3.31	0.37
		25	5.89	5.43	0.91	0.73	0.82	0.81
	$\bar{\delta}^{\sigma(W)}$	2	6.30	6.45	2.09	3.95	7.02	3.38
		5	7.25	6.83	1.29	3.46	8.27	3.31
		25	9.39	8.85	0.34	1.33	7.59	3.86
$0.6 \leq \rho < 0.8$	$\bar{\delta}^{E[W]}$	2	1.93	2.38	1.30	1.93	3.60	0.36
		5	3.04	3.36	0.78	0.86	3.04	0.35
		25	3.93	3.56	1.27	0.70	0.75	0.69
	$\bar{\delta}^{\sigma(W)}$	2	4.14	3.11	1.04	1.81	5.00	4.67
		5	5.87	5.97	0.69	1.10	7.08	4.57
		25	5.60	4.95	0.73	0.77	7.56	4.48
$0.8 \leq \rho$	$\bar{\delta}^{E[W]}$	2	1.25	1.48	1.80	2.31	1.69	0.36
		5	1.13	0.97	0.81	0.77	1.13	0.38
		25	2.13	1.21	0.95	0.56	0.60	0.40
	$\bar{\delta}^{\sigma(W)}$	2	3.73	4.77	2.84	0.52	3.04	4.86
		5	3.61	5.89	0.76	0.76	3.09	5.80
		25	3.41	3.40	0.64	2.70	4.53	4.20

waiting time over all customer types and all settings generated in a category, and $\bar{\delta}^{\sigma(W)}$ and max $\delta^{\sigma(W)}$ denote, respectively, the *average* and *maximum* percentage error of the standard deviation of the waiting time over all customer types and all settings generated in a category.

We can conclude that the moment-iteration algorithm produces accurate results, especially for high loads; typically, the mean of the waiting times is more accurately estimated than the standard deviation. For low loads, the percentage error of the moments of the waiting time might be high; the absolute error, however, will be modest (in comparison with the service times). Further, it seems that the type of distribution of the interarrival times does not influence the quality of the estimates produced by the moment-iteration method; the traffic intensity and the squared coefficients of variation of the service times are more crucial.

6. APPLICATION

In this section, we apply the moment-iteration method to the manufacturing problem mentioned in Section 1. We consider a manufacturing system with N stockpoints,

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

Table 5. Maximum Percentage Errors of the Moment-Iteration Method

			Model					
Load	Max. Error	N	$D/G/1$	$U/G/1$	$M/M/1$	$E_k/M/1$	$E_k/E_l/1$	$H_2/H_2/1$
$0.4 \le \rho < 0.6$	$\max \delta^{E[W]}$	2	12.61	14.52	4.62	4.35	9.04	1.11
		5	16.02	24.87	2.23	7.23	10.43	1.62
		25	19.42	26.91	2.86	3.64	4.42	5.30
	$\max \delta^{\sigma(W)}$	2	18.08	18.73	5.34	7.05	18.82	9.97
		5	22.27	23.24	5.92	6.38	19.82	8.36
		25	23.26	25.55	3.45	5.90	11.60	6.12
$0.6 \le \rho < 0.8$	$\max \delta^{E[W]}$	2	8.87	13.29	3.90	4.42	8.42	1.20
		5	9.77	13.57	4.03	3.81	15.56	1.58
		25	12.24	12.66	2.76	4.29	6.14	4.40
	$\max \delta^{\sigma(W)}$	2	14.71	18.77	5.27	4.08	11.56	11.57
		5	16.49	19.88	4.24	4.01	15.50	9.44
		25	19.35	22.10	4.30	4.16	11.61	7.96
$0.8 \le \rho$	$\max \delta^{E[W]}$	2	4.04	4.06	2.10	3.59	7.20	1.63
		5	7.54	6.88	2.37	1.85	8.40	0.88
		25	7.84	8.33	1.96	1.11	2.42	1.54
	$\max \delta^{\sigma(W)}$	2	9.62	16.21	2.41	2.21	9.84	10.81
		5	11.46	17.24	2.70	1.59	8.33	6.67
		25	12.53	17.60	2.24	2.08	10.86	6.62

one for each item, and one production facility, which produces all the items in a FIFO order. The objective is to determine the sojourn time of the replenishment orders placed by the stockpoints. Figure 1 is a schematic representation of the model for $N = 4$.

The stockpoints are controlled by periodic order-up-to policies. The periodic order-up-to policy operates as follows. Every R_i units, the inventory position of stockpoint i is inspected and a replenishment order is placed at the production facility to raise the inventory position up to the order-up-to level. The inventory position is defined as the physical inventory level plus the stock on order minus the backorders. We assume that the review period of each item is identical ($R_i = R$). The processing times of replenishment orders from stockpoint i are assumed to be i.i.d. random variables with mean $E[B_i]$ and standard deviation $\sigma(B_i)$. The first time the inventory position is inspected is denoted by t_i^0 $(i = 1, \ldots, N)$. The values of R_i, $E[B_i]$, $\sigma(B_i)$, and t_i^0 $(i = 1, \ldots, N)$ are listed in Table 6.

Now, we are interested in the sojourn time or production lead time S_i of replenishment orders of stockpoint i; obviously, the production lead time S_i is required for setting an appropriate stock level.

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

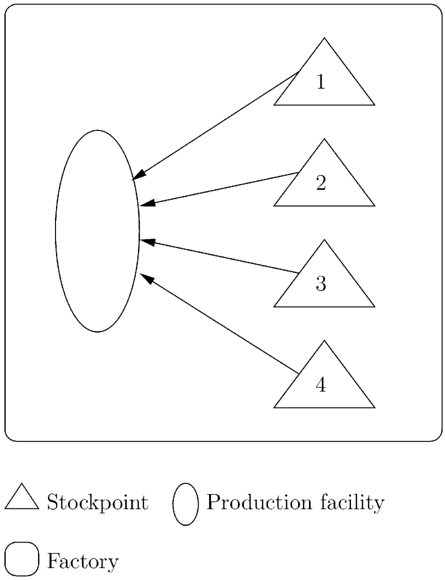

FIGURE 1. Schematic representation of the model.

The time A_i between the arrival of an item i replenishment order and item $i-1$ replenishment order follows from t_i^0 and R_i $(i = 1, \ldots, N)$. Because A_i is deterministic, we are dealing with a cyclic $D/G/1$ queue. Given A_i and the first two moments of B_i, we can determine the first two moments of W_i and S_i $(i = 1, \ldots, N)$ by using the moment-iteration method presented in Section 3. The results are presented in Table 7.

Note that the traffic intensity of the cyclic $D/G/1$ queue describing the manufacturing system is 0.91. Hence, Table 4 indicates that in this case (with $N = 4$), we can expect errors in the mean waiting time close to 1% and in its standard deviation close to 4%; this seems sufficiently accurate for practical purposes. From Table 6, it

TABLE 6. Input Parameters

	i			
	1	2	3	4
R_i	105.33	105.33	105.33	105.33
$E[B_i]$	19.24	25.20	27.15	24.52
$\sigma(B_i)$	7.05	8.02	5.34	4.81
t_i^0	0	21.06	48.69	78.45

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press

TABLE 7. Numerical Results

	i			
	1	2	3	4
A_i	26.88	21.06	27.63	29.79
$E[W_i]$	5.42	5.76	6.11	5.81
$\sigma(W_i)$	6.92	7.72	8.42	7.59
$E[S_i]$	24.66	30.96	33.26	30.33
$\sigma(S_i)$	9.88	11.13	9.97	8.98

follows that $c_{B_1} = 0.37$, $c_{B_2} = 0.32$, $c_{B_3} = 0.20$ and $c_{B_4} = 0.20$. In Table 7, we see that replenishment orders from stockpoint 1 clearly benefit from the smaller variation in processing times of replenishment orders from stockpoints 3 and 4.

References

1. Cohen, J.W. (1996). Periodic Pollaczek waiting time processes. In C.C. Heyde, Y.V. Prohorov, R. Pyke, & S.T. Rachev (eds.), *Athens Conference on Applied Probability and Time Series Analysis: Applied Probability in Honor of J.M. Gani*, pp. 361–377. New York: Springer-Verlag.
2. de Kok, A.G. (1989). A Moment-iteration method for approximating the waiting-time characteristics of the $GI/G/1$ queue. *Probability in the Engineering and Informational Sciences* 3: 273–287.
3. Morrice, D.J. & Gajulapalli, R.S. (1994). A single server queue with cyclically indexed arrival and services rates. *Queueing Systems* 15: 165–198.
4. Tijms, H.C. (1994). *Stochastic models: An algorithmic approach.* Chichester: Wiley.

APPENDIX

In this appendix, we will prove that the denominator of (7) has N zeros s with $\text{Re}(s) \geq 0$. To do so, we will use Rouche's theorem, which reads as follows. Let $f(s)$ and $g(s)$ be analytic functions inside and on a smooth contour C and suppose that $|g(s)| < |f(s)|$ on C. Then, $f(s)$ and $f(s) + g(s)$ have the same number of zeros inside C. Of course, we may also replace $f(s) + g(s)$ by $g(s) - f(s)$ in this formulation; actually, that is the form we will use.

We first assume that there is some $\epsilon > 0$ such that the transforms $B_i(s), i = 1, \ldots, N$, are analytic for all s with $\text{Re}(s) > -\epsilon$. This assumption holds, for example, for service time distributions with a finite support or an exponential tail. Now, take

$$f(s) = \prod_{i=1}^{N} \left(1 - \frac{s}{\lambda_i}\right)$$

and

$$g(s) = \prod_{i=1}^{N} B_i(s).$$

So $g(s) - f(s)$ is the denominator of (7). As a contour, we take the circle C_δ with center $\max_i \lambda_i$ and radius $\delta + \max_i \lambda_i$ with $0 < \delta < \epsilon$. It is easily verified that for all s on C_δ,

$$|g(s)| \leq \prod_{i=1}^{N} B_i(\mathrm{Re}(s)) \leq \prod_{i=1}^{N} B_i(-\delta) = g(-\delta), \qquad f(-\delta) \leq |f(s)|. \tag{A.1}$$

Note that

$$g'(0) = -\sum_{i=1}^{N} E[B_i], \qquad f'(0) = -\sum_{i=1}^{N} \frac{1}{\lambda_i}.$$

Hence, the stability condition (1) states that $g'(0) > f'(0)$. Thus, for sufficiently small $\delta > 0$, it holds that $g(-\delta) < f(-\delta)$, which implies together with (A.1) that $|g(s)| < |f(s)|$ for all s on C_δ. Rouche's theorem now guarantees that $f(s)$ and $g(s) - f(s)$ have the same number of zeros inside C_δ. Since $f(s)$ has N zeros inside C_δ, the same holds for $g(s) - f(s)$. Letting δ tend to zero, it follows that $g(s) - f(s)$ has exactly N zeros inside or on the circle C_0. There are no other zeros in the right half plane since for all s with $\mathrm{Re}(s) \geq 0$ outside C_0 we have

$$g(s) \leq \prod_{i=1}^{N} B_i(\mathrm{Re}(s)) \leq \prod_{i=1}^{N} B_i(0) = 1 = f(0) < |f(s)|.$$

To complete the proof, we must remove the initial assumption that for some $\epsilon > 0$, the transforms $B_i(s)$ are analytic for all s with $\mathrm{Re}(s) > -\epsilon$. To this end, first consider, instead of B_i, the truncated service times $\min(B_i, K)$, where $K > 0$ is some constant. For these truncated service times, the claim for the zeros of the denominator (7) holds; by letting K tend to infinity, the claim also follows for the original service time distributions.

Remark: In fact, we not only showed that the existence of N zeros s with $\mathrm{Re}(s) \geq 0$ but also that they are located inside or on the circle with center $\max_i \lambda_i$ and radius $\max_i \lambda_i$; this may be useful for the numerical calculation of the zeros.

https://doi.org/10.1017/S0269964804183022 Published online by Cambridge University Press