Busy Periods and Busy Cycles In Bulk-Arrival Queueing Systems

by

Sudha Jain
Department of Statistics
University of Toronto

and

Phil Reiss
Department of Statistics
University of Toronto

Technical Report No. 9813, August 20, (1998)

TECHNICAL REPORT SERIES

University of Toronto

Department of Statistics

BUSY PERIODS AND BUSY CYCLES IN BULK-ARRIVAL QUEUEING SYSTEMS

SUDHA JAIN AND PHILIP REISS

Abstract. Several results are presented concerning mean busy periods and mean busy cycles for certain general classes of queueing systems with bulk arrivals. These results lead to methods of estimating mean busy period and cycle length from Monte Carlo simulations. Using data from simulations, we compare these estimation methods with a more straightforward technique: calculating the sample mean busy period and cycle over a sequence of simulated cycles.

1. Introduction

Busy periods and busy cycles of a queueing process are important characteristics for the performance analysis of the system. Variation in busy periods and busy cycles depends on the number of the servers and the arrival and service time distributions of the system. Customers' waiting time in the queue and the queue-length of the system are related to busy periods. The time interval begins with an arrival to an empty system and ends the next time the system empties out (see, for example, Gross and Harris (1985)).

Much queueing theory literature is devoted to deriving analytical results for the busy periods and busy cycles for various types of queueing models. Takacs (1962), Chaudhry and Templeton (1983), and Gross and Harris (1985) discussed the busy periods and busy cycles for single and infinite servers. Shanbhag (1966) investigated the behviour of a busy period in infinite-server queues with batch arrivals. Dvurecenskij (1984) studied the busy period of order n in infinite-server queues. He derived the exact distribution, characteristic functions and all moments of the busy periods of the $GI^x/D/\infty$ queue. Stadje (1985) derived the distribution function of the length of a busy period and the distribution of the number of busy periods in which exactly i customers arriving in a given interval are served. Liu and Shi (1996) established a framework to deal with busy periods and busy cycles and related measures of $GI^x/G/\infty$ queueing systems. Recently, Falin and Templeton (1997) discussed the busy periods in a single-server batch arrival retrial queue.

In this paper, our objective is to develop theoretical aspects of the busy periods and busy cycles as well as implementations of these results using Monte Carlo simulations for various types of queueing processes with batch arrivals. We consider stationary systems. However, statistical fluctuations of the system are observed by the first and second moments of the quantity of the performance measures of a queueing system.

The paper is organized as follows. Theoretical results are developed in section 2, while section 3 is devoted to numerical results, comments and conclusions based on our simulation work. Section 4 suggests possible directions for further research. Simulation results were obtained using the C++ programming language. Interested researchers can obtain a copy of the program from the authors.

(e) $\sigma_c^2 = \dfrac{e^{r\lambda/\mu}}{\lambda^2 \sum_{k=0}^{r-1} \frac{(r\lambda/\mu)^k}{k!}} \left[1 + \dfrac{1}{r} + \dfrac{e^{r\lambda/\mu} - 2\sum_{k=0}^{r} \frac{(r\lambda/\mu)^k}{k!}}{\sum_{k=0}^{r-1} \frac{(r\lambda/\mu)^k}{k!}} \right].$

Proof. These formulas follow from Theorem 2 upon substituting for $\overline{F}(\mu^{-1})$, $\int_0^{\mu^{-1}} x\,dF(x)$, $\int_0^{\mu^{-1}} x^2 dF(x)$ the particular values these three expressions take on in the case of Erlang arrivals. These values are obtained by first deriving a general formula for $I_n \equiv \int_0^{\mu^{-1}} (r\lambda x)^n e^{-r\lambda x} dx$.

Letting $C_n = \frac{(r\lambda/\mu)^n e^{-r\lambda/\mu}}{r\lambda}$, we have, for $n > 0$,

$$
\begin{aligned}
I_n &= \left. \frac{(r\lambda x)^n e^{-r\lambda x}}{-r\lambda} \right|_0^{\mu^{-1}} + n \int_0^{\mu^{-1}} (r\lambda x)^{n-1} e^{-r\lambda x} dx \\
&= nI_{n-1} - C_n \qquad \text{(recursive formula)} \\
&= n\left[(n-1)I_{n-2} - C_{n-1}\right] - C_n \qquad \text{(by applying the recursive formula to } I_{n-1}\text{)} \\
&= n(n-1)I_{n-2} - nC_{n-1} - C_n \\
&= n(n-1)\left[(n-2)I_{n-3} - C_{n-2}\right] - nC_{n-1} - C_n \\
&= \frac{n!}{(n-3)!} I_{n-3} - \frac{n!}{(n-2)!} C_{n-2} - \frac{n!}{(n-1)!} C_{n-1} - \frac{n!}{n!} C_n \\
&= \cdots \\
&= n! I_0 - \sum_{k=1}^{n} \frac{n!}{k!} C_k \\
&= n! \int_0^{\mu^{-1}} e^{-r\lambda x} dx - \sum_{k=1}^{n} \frac{n!}{k!} \frac{(r\lambda/\mu)^k e^{-r\lambda/\mu}}{r\lambda} \\
&= n! \left. \frac{e^{-r\lambda x}}{-r\lambda} \right|_0^{\mu^{-1}} - \sum_{k=1}^{n} \frac{n!}{k!} \frac{(r\lambda/\mu)^k e^{-r\lambda/\mu}}{r\lambda} \\
&= \frac{n!}{r\lambda} \left[1 - e^{-r\lambda/\mu} - \sum_{k=1}^{n} \frac{(r\lambda/\mu)^k e^{-r\lambda/\mu}}{k!} \right] \\
&= \frac{n!}{r\lambda} \left[1 - e^{-r\lambda/\mu} \sum_{k=0}^{n} \frac{(r\lambda/\mu)^k}{k!} \right].
\end{aligned}
$$

We can now calculate $\overline{F}(\mu^{-1})$, $\int_0^{\mu^{-1}} x\,dF(x)$, $\int_0^{\mu^{-1}} x^2\,dF(x)$ for the $E_r^x/D/\infty$ case, in which the interarrival density is $f(x) = \frac{(r\lambda)^r x^{r-1} e^{-r\lambda x}}{(r-1)!}$. We have

$$\begin{aligned}
\overline{F}(\mu^{-1}) = 1 - \int_0^{\mu^{-1}} dF(x) &= 1 - \int_0^{\mu^{-1}} \frac{(r\lambda)^r x^{r-1} e^{-r\lambda x}}{(r-1)!}\,dx \\
&= 1 - \frac{r\lambda}{(r-1)!}\, I_{r-1} \\
&= 1 - \frac{r\lambda}{(r-1)!}\frac{(r-1)!}{r\lambda}\left[1 - \sum_{k=0}^{r-1} \frac{(r\lambda/\mu)^k e^{-r\lambda/\mu}}{k!}\right] \\
&= e^{-r\lambda/\mu} \sum_{k=0}^{r-1} \frac{(r\lambda/\mu)^k}{k!}\,, \qquad (1)
\end{aligned}$$

$$\begin{aligned}
\int_0^{\mu^{-1}} x\,dF(x) &= \int_0^{\mu^{-1}} \frac{(r\lambda x)^r e^{-r\lambda x}}{(r-1)!}\,dx \\
&= \frac{I_r}{(r-1)!} \\
&= \frac{r!}{r\lambda(r-1)!}\left[1 - e^{r\lambda/\mu} \sum_{k=0}^{r} \frac{(r\lambda/\mu)^k}{k!}\right] \\
&= \frac{1}{\lambda}\left[1 - e^{-r\lambda/\mu} \sum_{k=0}^{r} \frac{(r\lambda/\mu)^k}{k!}\right]\,, \qquad (2)
\end{aligned}$$

and

$$\begin{aligned}
\int_0^{\mu^{-1}} x^2 dF(x) &= \int_0^{\mu^{-1}} \frac{(r\lambda)^r x^{r+1} e^{-r\lambda x}}{(r-1)!}\,dx \\
&= \frac{I_{r+1}}{r\lambda(r-1)!} \\
&= \frac{(r+1)!}{r^2\lambda^2(r-1)!}\left[1 - e^{-r\lambda/\mu} \sum_{k=0}^{r+1} \frac{(r\lambda/\mu)^k}{k!}\right] \\
&= \frac{r+1}{r\lambda^2}\left[1 - e^{-r\lambda/\mu} \sum_{k=0}^{r+1} \frac{(r\lambda/\mu)^k}{k!}\right]\,. \qquad (3)
\end{aligned}$$

Substituting (1), (2) and (3) into the conclusion of Theorem 2 leads to the Corollary.

□

We now turn from $GI^x/D/\infty$ to $M^x/G/c$ and $GI^x/M/c$ ($c \in \mathbf{Z}$ or $c = \infty$). Liu and Shi prove analogues of Theorem 2 for the latter two systems with $c = \infty$, but the variance formulas they derive are harder to apply than that for $GI^x/D/\infty$. Here we present only the Liu-Shi formulas for expected busy period and cycle. Using direct probabilistic arguments, rather than the Laplace-Stieltjes transform methods Liu and Shi employ, we extend these formulas to the $c \in \mathbf{Z}$ case.

Theorem 3 (Liu and Shi (1996), Corollary 5). *For the $M^x/G/c$ system ($c \in \mathbf{Z}$ or $c = \infty$),*

$$\bar{b} = \frac{1}{\lambda P_0} - \frac{1}{\lambda}$$

and

$$\bar{c} = \frac{1}{\lambda P_0} ,$$

where P_0 is the steady-state probability that the system is idle.

Proof. By the alternating renewal theorem, if $\bar{i}$ is the mean idle period,

$$\frac{\bar{b}}{\bar{i}} = \text{ ratio of busy to idle time } = \frac{1 - P_0}{P_0} .$$

Since arrivals are exponential, $\bar{i} = \frac{1}{\lambda}$ and so

$$\bar{b} = \bar{i}\,\frac{1 - P_0}{P_0} = \frac{1}{\lambda}\,\frac{1 - P_0}{P_0} = \frac{1}{\lambda P_0} - \frac{1}{\lambda} ,$$

therefore $\bar{c} = \bar{b} + \bar{i} = \frac{1}{\lambda P_0}$. □

Corollary (Chaudhry and Templeton (1983), p. 324). *For $M^x/G/1$, $\bar{b} = \frac{\bar{a}}{\mu - \lambda \bar{a}}$.*

Proof. Here $P_0 = 1 - \rho = 1 - \frac{\lambda \bar{a}}{\mu}$, and so

$$\begin{aligned} \bar{b} &= \frac{1}{\lambda\left(1 - \frac{\lambda \bar{a}}{\mu}\right)} - \frac{1}{\lambda} \\ &= \frac{1 - \left(1 - \frac{\lambda \bar{a}}{\mu}\right)}{\lambda\left(1 - \frac{\lambda \bar{a}}{\mu}\right)} \\ &= \frac{\frac{\bar{a}}{\mu}}{1 - \frac{\lambda \bar{a}}{\mu}} \\ &= \frac{\bar{a}}{\mu - \lambda \bar{a}}. \end{aligned}$$

□

Theorem 4 (Liu et al. (1996), Corollary 6). *For the $GI^x/M/c$ system ($c \in \mathbf{Z}$ or $c = \infty$),*

$$\bar{b} = \frac{1 - P_0}{\mu P_1} \qquad \text{and} \qquad \bar{c} = \frac{1}{\mu P_1} ,$$

where P_1 is the steady-state probability that there is exactly one customer in the system.

Proof. Let $P_1(s)$ be the probability that there is one customer at time s. The expected number of cycle endings in $(s, s + \Delta s]$ is

$$\begin{aligned} & Pr(\text{system contains one customer at } s, \text{ which is served by } s + \Delta s) + o(\Delta s) \\ = \;& Pr(\text{one customer at } s)\; Pr(\text{service in } (s, s + \Delta s] |\text{ one customer at } s) + o(\Delta s) \\ = \;& P_1(s) \mu \Delta s + o(\Delta s), \text{by the exponentiality of service.} \end{aligned}$$

Integrating from 0 to t and taking the limit as $t \to \infty$, we obtain

$$\begin{aligned} & \lim_{t\to\infty} \frac{1}{t} E(\text{number of cycle endings by } t) \\ = & \lim_{t\to\infty} \frac{1}{t} \int_0^t \mu P_1(s) ds \\ = & \mu P_1 \end{aligned}$$

since $\lim_{s\to\infty} P_1(s) = P_1$. Hence $\lim_{t\to\infty} E\left[\frac{\text{cycles}(t)}{t}\right] = \mu P_1$.

We also have $\frac{t}{\text{cycles}(t)} \xrightarrow[t\to\infty]{} \bar{c}$ w.p. 1 (by the SLLN, which applies since the cycles are independent), whence

$$\lim_{t\to\infty} \frac{\text{cycles}(t)}{t} = \bar{c}^{-1} \text{w.p. } 1 \ .$$

Let $g(t)$ denote the number of departures, up to time t, from the lowest-numbered currently active server. Since every cycle ending is a departure of this sort, $\frac{\text{cycles}(t)}{t}$ is dominated by $\frac{g(t)}{t}$. During busy periods, departures from the lowest-numbered active server occur as a Poisson process with rate μ. Thus

$$\begin{aligned} E\left[\frac{g(t)}{t}\right] &\leq E\left[\frac{g(t)}{\text{busy}(t)}\right] \\ &= \mu < \infty, \end{aligned}$$

and so the dominated convergence theorem may be applied to $\frac{\text{cycles}(t)}{t}$ to obtain

$$\begin{aligned} \bar{c}^{-1} = E(\bar{c}^{-1}) &= E\left[\lim_{t\to\infty} \frac{\text{cycles}(t)}{t}\right] \\ &= \lim_{t\to\infty} E\left[\frac{\text{cycles}(t)}{t}\right] \\ &= \mu P_1. \end{aligned}$$

Hence $\bar{c} = \frac{1}{\mu P_1}$.

The value of $\bar{b}$ follows since

$$\begin{aligned} \frac{\bar{c}}{\bar{b}} = \frac{\bar{b} + \bar{i}}{\bar{b}} &= 1 + \frac{\bar{i}}{\bar{b}} \\ &= 1 + \frac{P_0}{1 - P_0} \qquad \text{(alternating renewal theorem)} \\ &= \frac{1}{1 - P_0} \ ; \end{aligned}$$

therefore

$$\begin{aligned} \bar{b} &= \bar{c}(1 - P_0) \\ &= \frac{1 - P_0}{\mu P_1} \ . \end{aligned}$$

□

3. Computer simulation results

In this section we present numerical results. We carried out our simulations at the University of Toronto mathematics department computing facility with the following parameter values:

$\lambda = 3$;
$\mu = 10$;
r (number of arrival phases for Erlang-arrival queues) $= 4$;
k (number of service phases for Erlang-service queues) $= 5$;
number of cycles simulated per run $= 1000$.

These values were set in the preamble to our C++ simulation program, and each can be changed by simply revising the appropriate line of the preamble.

Adjusting the arrival batch size distribution is a bit more complicated. To do this one must rewrite the "batch" function and also revise the introductory output explaining the batch distribution. For the simulations, the batches were taken to contain 1, 2 or 3 customers, each with probability $\frac{1}{3}$.

If the parameters or batch distribution are changed, one must make sure that the traffic intensity $\rho = \frac{\lambda \bar{a}}{\mu}$ remains below 1. With the settings used for the simulations we have

$$\rho = \frac{\lambda \bar{a}}{\mu} = \frac{3 \cdot 2}{10} = \frac{3}{5} .$$

Table 1 provides data from simulations of single-server batch arrival systems in which both arrivals and service may be chosen to be deterministic, exponential or Erlang. In each case $B1$ and $C1$ are the sample means of the busy period and cycle, respectively, taken from a run of 1000 cycles. That is to say,

$$B1 = \frac{1}{1000}(\text{elapsed busy time over a 1000-cycle run}) = \hat{\bar{b}}(t)$$

and

$$C1 = \frac{1}{1000}(\text{elapsed time over a 1000-cycle run}) = \hat{\bar{c}}(t),$$

where $t =$ run time.

The $B1$ and $C1$ values appearing in the table are actually *sample means* (and sample standard deviations) *of sample means*. For example, the meaning of the .20010 (.002410) in the first row and first column of Table 1 is as follows. Two hundred fifty runs, each comprising 1000 cycles, were performed for the $D^x/D/1$ system. For each run the sample mean busy period $B1$ was computed. These 250 values of $B1$ were then found to have sample mean .20010 and sample standard deviation .002410.

Theorem 1 provides us with a second estimator, $B2$, of the mean busy period. Namely, since the expected busy period for $GI^x/G/1$ is $\frac{1}{\mu P_0^+}$, we take

$$B2 = \frac{1}{\mu \widehat{P_0^+}(t)}$$

TABLE 1. Single-server simulation data (250 runs)

			ARRIVALS		
			Deterministic	Exponential	Erlang-4
SERVICE	Deterministic	B1	.20010 (.002410)	.49865 (.02298)	.29460 (.008582)
		B2	.20009 (.002411)	.49868 (.02298)	.29459 (.008588)
		SCC	1.001 (!)	1.000	1.000
			True mean $\frac{1}{5}$	True mean $\frac{1}{2}$	
		C1	.33334 (0)	.83138 (.02441)	.49185 (.008919)
		C2		.83201 (.02299)	
		SCC		.905	
			True mean $\frac{1}{3}$	True mean $\frac{5}{6}$	
	Exponential	B1	.27697 (.01317)	.49772 (.02729)	.33202 (.01709)
		B2	.27616 (.009023)	.49862 (.02386)	.33199 (.01267)
		SCC	.956	.975	.960
				True mean $\frac{1}{2}$	
		C1	.46003 (.01239)	.83126 (.02897)	.55356 (.01745)
		C2		.83195 (.02386)	
		SCC		.897	
				True mean $\frac{5}{6}$	
	Erlang-5	B1	.23159 (.006311)	.50087 (.02313)	.30441 (.01106)
		B2	.23139 (.005344)	.50078 (.02194)	.30446 (.01002)
		SCC	.952	.989	.979
				True mean $\frac{1}{2}$	
		C1	.38549 (.005307)	.83608 (.02560)	.50706 (.01162)
		C2		.83411 (.02195)	
		SCC		.897	
				True mean $\frac{5}{6}$	

.20010 (.002410) denotes mean .20010, standard deviation .002410
SCC = sample correlation coefficient

where t is the total time of a run and $\widehat{P_0^+}(t)$ is an estimator of P_0^+ given by

$$\widehat{P_0^+}(t) = \frac{\text{number of departures up to t leaving behind an empty system}}{\text{deps(t)}}$$

$$= \frac{\text{cycles}(t)}{\text{deps}(t)} ,$$

as above, or

$$\widehat{P_0^+}(t) = \frac{1000}{\text{deps}(t)} .$$

In the case of exponential interarrival times, since the mean idle period is known to be $\frac{1}{\lambda}$ (by forgetfulness), a second estimator of mean busy cycle is given by

$$C2 = B2 + \frac{1}{\lambda} .$$

Where values of both mean busy period (or cycle) estimators were obtainable, their sample correlation coefficient taken over the 250 runs is shown in the table.

Remarks on Table 1.

1. Mean values of $B1$ and $B2$, and of $C1$ and $C2$, are always within .001 of each other — except for the $M^x/E_k/1$ case, in which $C1$ and $C2$ differ by about .002. Indeed, each of the twelve estimates $B2$ and $C2$ can be shown to lie within an approximate 95% confidence interval for the corresponding sample mean-based estimates $B1$ and $C1$. Consider, for instance, the busy cycle estimates in the table for $M^x/E_k/1$. The 250 values obtained for $C1$ averaged .83608 with sample standard deviation .02560, implying an approximate 95% confidence interval for $\bar{c}$ of

$$.83608 \pm 1.96\frac{.02560}{\sqrt{250}} = (.83291, .83925),$$

which contains the mean $C2$ value of .83411.

To better understand this strong agreement, let us reexamine the definitions of $B1$ and $B2$. Since $B1 = \frac{\text{busy}(t)}{1000}$ and $B2 = \frac{1}{\mu P_0^+(t)} = \frac{1}{\mu \frac{1000}{\text{deps}(t)}} = \frac{\text{deps}(t)}{1000\mu}$, we have

$$\frac{B1}{B2} = \mu\frac{\text{busy}(t)}{\text{deps}(t)} \ .$$

But the mean service time, i.e. the mean busy time per departure, is $\frac{1}{\mu}$, so that $\frac{\text{busy}(t)}{\text{deps}(t)} \xrightarrow[t\to\infty]{} \frac{1}{\mu}$ w.p. 1 and $\mu\frac{\text{busy}(t)}{\text{deps}(t)} \to 1$ w.p. 1. Hence the ratio of $B1$ to $B2$ equals the ratio of $\frac{\text{busy}(t)}{\text{deps}(t)}$ to its limiting value $\frac{1}{\mu}$, where t is the total time of a run.

2. $B2$ appears to have the same variance as $B1$ for $GI^x/D/1$, slightly less for $GI^x/E_k/1$ and noticeably less for $GI^x/M/1$. As $B1 = \frac{\text{busy}(t)}{1000}$ and $B2 = \frac{\text{deps}(t)}{1000\mu}$, this would indicate that the number of departures per cycle is less variable than busy time per cycle (i.e. busy period), when service is non-deterministic.

When $C2$ exists (i.e. for $M^x/G/1$), its variance is less than that of $C1$. This is unsurprising since the $C1$ estimate incorporates observed idle periods while $C2$ takes the idle period to be a constant $\frac{1}{\lambda}$.

3. The very high sample correlation coefficients for $B1, B2$ reflect the closeness of $\frac{B1}{B2} = \mu\frac{\text{busy}(t)}{\text{deps}(t)}$ to 1 when t is the total time of a run. $C1$ and $C2$ are not as strongly correlated (about .9) since these two estimates take into account uncorrelated idle-period values (observed values for $C1$, the constant value $\frac{1}{\lambda}$ for $C2$).

4. True mean busy periods and cycles are known for four of the nine systems in Table 1. In each of these four cases, the estimates are seen to be quite accurate: the true value can be shown to fall within an approximate 95% confidence interval of both estimates (calculated as in Remark 1) in every instance.

For the three exponential-arrival systems, Theorem 2 gives $\bar{b} = \frac{\bar{a}}{\mu-\lambda\bar{a}} = \frac{2}{10-3\cdot 2} = \frac{1}{2}$, while

$$\begin{aligned} \bar{c} &= \bar{b} + \frac{1}{\lambda} \qquad (\frac{1}{\lambda} = \text{expected idle period}) \\ &= \frac{1}{2} + \frac{1}{3} = \frac{5}{6} \ . \end{aligned}$$

For the $D^x/D/1$ system, $\bar{c} = \frac{1}{3}$ because each arrival (deterministically at intervals of $\frac{1}{\lambda} = \frac{1}{3}$) initiates a new busy cycle. This is because each arriving batch is served in time $\frac{1}{\mu} = \frac{1}{10}$, $\frac{2}{\mu} = \frac{2}{10}$ or $\frac{3}{\mu} = \frac{3}{10}$, depending on the batch size — but at any rate the batch completes service before the next arrival. We then have $\bar{b} = \frac{\bar{a}}{\mu} = \frac{2}{10} = \frac{1}{5}$.

Note that the ease with which $\bar{b}$ and $\bar{c}$ have been found for $D^x/D/1$ is an artifact of our choice of parameters. If the maximum batch service time $\frac{\text{max. batch size}}{\mu}$ exceeds the interarrival time $\frac{1}{\lambda}$ (as would occur in our case if four-customer batches were possible) then not every arrival finds an empty system and so some cycles are more than one arrival interval long. (It is only the *mean* batch service time $\frac{\bar{a}}{\mu}$, but not the maximum service time, which must be less than $\frac{1}{\lambda}$ in order to satisfy $\rho < 1$.)

TABLE 2. Infinite-server simulation data (250 runs)

		Busy periods	Busy cycles
$E_r^x/D/\infty$	SM	.102692 (.0004630)	.345592 (.005188)
	LS (exact values)	.102672	.344983
	SCC	–	–
$M^x/E_k/\infty$	SM	.14627 (.002502)	.48012 (.01138)
	LS	.14621 (.005467)	.47954 (.005469)
	SCC	.463	-.768
$E_r^x/M/\infty$	SM	.15968 (.004748)	.39843 (.006737)
	LS	.15995 (.002522)	.39932 (.01087)
	SCC	.0974	-.0200

The three systems for which data is presented in Table 2 are special cases of the systems, studied by Liu and Shi (1996), which are the subject of Theorems 2, 3 and 4. The "general" distributions of $GI^x/D/\infty$, $M^x/G/\infty$ and $GI^x/M/\infty$ are each taken to be Erlang: $E_r = E_4$ for Erlang arrivals and $E_k = E_5$ for Erlang service.

SM stands for sample mean, i.e. the sample-mean estimators $\hat{\bar{b}}(t)$ and $\hat{\bar{c}}(t)$. LS denotes Liu-Shi values. In the $E_r^x/D/\infty$ case, the Corollary to Theorem 2 has been used to find exact values of $\bar{b}$ and $\bar{c}$; regarding the exact standard deviations, see Remark 3 below.

On the other hand, the Liu-Shi expressions for $\bar{b}$ and $\bar{c}$ in the $M^x/E_k/\infty$ and $E_r^x/M/\infty$ cases cannot be evaluated exactly, since they contain the unknowns P_0 and P_1. Thus, for these two systems, Liu-Shi *estimators* for $\bar{b}$ and $\bar{c}$ have been created. This is done by replacing P_0 by $\hat{P}_0(t)$, the observed proportion of the run time for which the system is empty, and P_1 by $\hat{P}_1(t)$, the observed proportion of time for which there is exactly one customer in the system.

The numbers appearing in the LS rows for $M^x/E_k/\infty$ and $E_r^x/M/\infty$ are sample means and standard deviations taken from 250 values (one from each run) of the Liu-Shi estimators. By contrast, the numbers in the LS row for $E_r^x/D/\infty$ are simply a priori means and standard deviations of $\bar{b}$ and $\bar{c}$ (Liu and Shi, Corollary 2). This is why there is no sample correlation coefficient to calculate in the $E_r^x/D/\infty$ case.

Remarks on Table 2.

1. In each case, the sample-mean estimates and LS values agree to within .001. The latter value falls within an approximate 95% confidence interval for the sample mean except for the $E_r^x/M/\infty$ busy cycles.

2. The LS estimator is more variable than the sample-mean estimator for $\bar{b}$ in $M^x/E_k/\infty$ and $\bar{c}$ in $E_r^x/M/\infty$, but less variable for $\bar{b}$ in $E_r^x/M/\infty$ and $\bar{c}$ in $M^x/E_k/\infty$.

3. Consider the standard deviation figures in the first row of the table. These are estimates (from a sample of 250 runs) of the standard deviation of the mean of the 1000 busy periods or cycles in a run. But the standard deviation of the mean of 1000 independent busy periods is $\frac{\sigma_b}{\sqrt{1000}} = \frac{\sigma_b}{10\sqrt{10}}$, where σ_b is the standard deviation of individual busy periods. Hence .0004630 is an estimate of $\frac{\sigma_b}{10\sqrt{10}}$, and

$$10\sqrt{10}(.0004630) = .01464$$

is an estimate of σ_b. Using the Corollary to Theorem 3, the true value of σ_b is found to be.0149258. Similarly

$$10\sqrt{10}(.005188) = .1641$$

is an estimate of σ_c, whose true value, by the Corollary, is .163013.

4. Since $1-\hat{P}_0(t) = \frac{\text{busy time}(t)}{t} = \frac{\text{busy time}(t)/1000}{t/1000} = \frac{\hat{\bar{b}}}{\hat{\bar{c}}}$, it is clear that $\hat{P}_0(t)$ and $\frac{\hat{\bar{b}}}{\hat{\bar{c}}}$ are negatively correlated. However, it is less clear that $\hat{P}_0(t)$, or $\hat{P}_1(t)$, is correlated with $\hat{\bar{b}}$ or $\hat{\bar{c}}$ individually. Therefore, the latter two (the sample-mean estimators) stand in no obvious relationship to the LS estimators, which depend on $\hat{P}_0(t)$ and $\hat{P}_1(t)$. Hence the small sample correlation coefficients for $E_r^x/M/\infty$ are unsurprising. By contrast, the strong negative correlation of the busy cycle estimators in the $M^x/E_k/\infty$ case (a finding which was replicated in several retrials) is something of a mystery.

TABLE 3. Mean over 100 runs of sample mean busy periods and busy cycles, for different numbers of servers.

(a) Busy period.

system \ c	1	2	3	10	∞
$E_r^x/D/c$	.29473	.16157	.10475	.10349	.10273
$M^x/E_k/c$	.49939	.19592	.15610	.14606	.14650
$E_r^x/M/c$	.33524	.17274	.15985	.15984	.15998

(b) Busy cycle.

system \ c	1	2	3	10	∞
$E_r^x/D/c$	.49144	.38425	.34595	.34480	.34509
$M^x/E_k/c$	.83333	.52925	.48848	.48084	.47965
$E_r^x/M/c$	.55613	.40899	.39913	.39910	.39904

Convergence to infinite-server values as $c \to \infty$. Table 3 gives the means, over 100 runs, of the sample-mean estimators of $\bar{b}$ and $\bar{c}$, for $E_r^x/D/c$, $M^x/E_k/c$ and $E_r^x/M/c$ with $c = 1, 2, 3, 10, \infty$. The table shows that, as we would expect, $\bar{b}$ and $\bar{c}$ converge downward to their infinite-server values as the number of servers for a given system goes to infinity. (Although in practice a many-server system resembles an infinite-server system, for the simulations a common algorithm was used for all finite-server systems of a given type, and a different algorithm for the corresponding infinite-server system. Hence the observed convergence served to confirm the correctness of both algorithms.)

Note that in each case, the 3-server system figures are quite close to the infinite-server figures — the differences between $M^x/E_k/3$ and $M^x/E_k/\infty$ (about .01 for both $\bar{b}$ and $\bar{c}$) being by far the greatest. The 10-server numbers approximate the infinite-server numbers so well that in some cases they are actually slightly lower than the latter. Recall that $\rho = \frac{3}{5}$ in our simulated systems; for higher ρ values more servers would be required to approximate the infinite-server case equally well.

TABLE 4. Two estimators of $\bar{b}$ for $M^x/G/1$ with unknown parameters (1000 runs).

	SERVICE		
	Deterministic	Exponential	Erlang
$\hat{\bar{b}}$	.499102 (.021236)	.499544 (.027990)	.500663 (.023070)
$\frac{1-\hat{P}_0}{\hat{\lambda}\hat{P}_0}$	.499954 (.017566)	.499972 (.025601)	.501237 (.020086)
SCC	.82269	.90127	.84743

Unknown parameters. Now suppose all we know about a queueing system is the following: arrivals come in batches at exponentially distributed intervals, and a single server handles customers one at a time. The parameters λ, μ, $\bar{a}$ are unknown. Given the relevant measurements taken over a sequence of 1000 busy cycles, how can we estimate the mean busy period?

The first estimator that suggests itself is the sample mean $\hat{\bar{b}}(t) = \frac{\text{busy}(t)}{\text{cycles}(t)} = \frac{\text{busy}(t)}{1000}$, where t is the elapsed time for 1000 cycles.

A second idea is to utilize Theorem 1 and estimate $\bar{b} = \frac{1}{\mu P_0^+}$ by $\frac{1}{\hat{\mu}(t)\widehat{P_0^+}(t)}$, where $\hat{\mu}(t) = \frac{\text{deps}(t)}{\text{busy}(t)}$ and $\widehat{P_0^+}(t) = \frac{\text{cycles}(t)}{\text{deps}(t)}$ as in the proof of that theorem. But the resulting estimator is just

$$\frac{1}{\frac{\text{deps}(t)}{\text{busy}(t)}\,\frac{\text{cycles}(t)}{\text{deps}(t)}} = \frac{\text{busy}(t)}{\text{cycles}(t)} = \hat{\bar{b}}(t)\ ,$$

i.e., exactly the same as our first estimator.

A different estimator emerges from Theorem 3, which asserts that $\bar{b} = \frac{1-P_0}{\lambda P_0}$. Using the estimators

$$\begin{aligned}
\hat{\lambda}(t) &= \frac{\text{arrs}(t)}{t}\ , \\
\hat{P}_0(t) &= \frac{\text{idle}(t)}{t\ ,} \\
1-\hat{P}_0(t) &= \frac{\text{busy}(t)}{t}
\end{aligned}$$

(where idle(t) = idle time in $[0,t)$), we set

$$\begin{aligned}
\bar{\bar{b}}(t) = \frac{1-\hat{P}_0}{\hat{\lambda}\hat{P}_0} &= \frac{\frac{\text{busy}(t)}{t}}{\frac{\text{arrs}(t)}{t}\frac{\text{idle}(t)}{t}} \\
&= \frac{t\,\text{busy}(t)}{\text{arrs}(t)\,\text{idle}(t)}.
\end{aligned}$$

Both estimators were calculated for each of 1000 1000-cycle runs for $M^x/G/1$, with $G = D, M, E_k$. The results appear in Table 4. Note that the parameter values are the same as above; they are "unknown" only in the sense that they are not used for the estimation of $\bar{b}$.

As noted in connection with Table 1, the true $\bar{b}$ is $\frac{1}{2}$. The second estimator $\bar{\bar{b}}(t)$ is slightly closer to this value for two of the three systems ($M^x/D/1$ and $M^x/M/1$) and has slightly lower variance for all three. The two estimators are very strongly positively correlated for all three systems.

4. Some possibilities for further study

Our results in the previous section suggest that theoretically-derived estimators for mean busy period and cycle may be more efficient than estimators based on sample means, especially in the single-server case (see Remark 2 on Table 1, and the closing paragraph of the previous section). Further simulations would be required, however, in order to draw any broad conclusions. In this section we suggest some additional questions which might be investigated.

1. For a given system one can vary ρ by varying just one of the parameters, say λ. Then, for each ρ value, one can observe how many servers are required to bring the mean busy period and cycle to within a given range of the corresponding infinite-server values. For instance, with the parameters used in the above simulations, $c = 3$ is sufficient to bring $E_r^x/M/c$ estimates of $\bar{b}$ and $\bar{c}$ to within .001 of the $E_r^x/M/\infty$ values.

It is then possible to graph ρ against c_ρ, the smallest number of servers for which $\bar{b}$ and $\bar{c}$ appear "close" to the infinite-server values. (Actually the notation c_ρ is misleading, since the desired number of servers presumably depends not only on ρ but on λ, μ and the interarrival, service and batch distributions, which determine ρ.)

2. The formula $\bar{b} = \frac{\bar{a}}{\mu - \lambda\bar{a}}$ for $M^x/G/1$ gives rise to two formulas for $\bar{c}$ in the $M^x/G/1$ case:

$$\bar{c} = \bar{b} + \frac{1}{\lambda} = \frac{\bar{a}}{\mu - \lambda\bar{a}} + \frac{1}{\lambda} ,$$

$$\bar{c} = \frac{\bar{b}}{\rho} = \frac{\bar{a}}{\mu - \lambda\bar{a}} \Big/ \frac{\lambda\bar{a}}{\mu} = \frac{\mu}{\lambda(\mu - \lambda\bar{a})} .$$

Clearly these are equal. The first of these formulas leads to the estimator

$$\begin{aligned}\bar{\bar{c}}(t) &= \bar{\bar{b}}(t) + \frac{1}{\hat{\lambda}(t)} \quad (\text{ where } \bar{\bar{b}}(t) \text{ is the second estimator used in Table 4}) \\ &= \frac{t\,\mathrm{busy}(t)}{\mathrm{arrs}(t)\,\mathrm{idle}(t)} + \frac{t}{\mathrm{arrs}(t)} \\ &= \frac{t}{\mathrm{arrs}(t)}\left[\frac{\mathrm{busy}(t)}{\mathrm{idle}(t)} + 1\right] \\ &= \frac{t^2}{\mathrm{arrs}(t)\,\mathrm{idle}(t)} .\end{aligned}$$

We thus have a second estimator for $\bar{c}$ in the $M^x/G/1$ system, which could be compared with the sample mean $\hat{c}(t)$ in a table analogous to Table 4.

3. The simulation program allows for nine different systems with a given finite number of servers, since the interarrival and service distributions may each be deterministic, exponential or Erlang. Liu and Shi's results (for $GI^x/D/\infty$, $M^x/G/\infty$, $GI^x/M/\infty$) cover only seven of the nine corresponding infinite-server systems. Missing are $D^x/E_k/\infty$ and $E_r^x/E_k/\infty$. Can the results of Liu and Shi be extended to these two cases?

4. The standard deviation estimates, e.g. in Table 1, can be used to adjust the run length so as to reduce the standard deviation of the sample mean estimates to a desired level.

For example, $B1$ for the $M^x/E_k/1$ system was found to have mean .50087 and standard deviation .02313, with 1000-cycle runs. Suppose, however, that we are not satisfied with a standard deviation above .01 for $B1$. What number n of cycles per run will reduce the standard deviation of $B1$ to .01?

Since the standard deviation of $B1$ is inversely proportional to the square root of the run length, we have (approximately)

$$\sqrt{\frac{n}{1000}} = \frac{.02313}{.01} ;$$

therefore

$$n = 1000\left(\frac{.02313}{.01}\right)^2 \approx 5350 ,$$

i.e., 5350-cycle runs will reduce the standard deviation of $B1$ to approximately .01.

A related problem is to find a priori results relating the number of runs to the accuracy of various estimators' mean values.

5. A waiting-time approach to estimating $\bar{b}$ and $\bar{c}$ can be derived from p. 33 of Law (1975). Let D be the mean total time in the system for all customers in a cycle. Using this notation

and that of Gross and Harris (1985), (1.4) of Law (1975) says that

$$L_q = \frac{D}{\bar{c}};$$

therefore

$$\bar{c} = \frac{D}{L_q} = \frac{D}{\lambda W_q} \quad \text{(by Little's result).} \tag{4}$$

In the one-server case we also have

$$\bar{b} = \rho\bar{c} = \frac{\lambda\bar{a}}{\mu}\frac{D}{\lambda W_q} = \frac{\bar{a}D}{\mu W_q}. \tag{5}$$

If waiting-time data are available we may take

$$\widehat{W}_q = \text{ observed mean time in queue}$$

and

$$\hat{D} = \text{ observed mean total time in system for all customers in a cycle,}$$

and then use (11) and (12) to form the estimators

$$\stackrel{\cong}{b} = \frac{\bar{a}\hat{D}}{\mu\widehat{W}_q},$$

$$\stackrel{\cong}{c} = \frac{\hat{D}}{\lambda\widehat{W}_q}.$$

5. Acknowledgements

The first author acknowledges with thanks the financial support provided by NSERC Canada. The authors are grateful to Profs. J. G. C. Templeton and Vladimir Vinogradov for their advice and comments.

6. References

Chaudhry, M. L. and Templeton, J. G. C. (1983). *A First Course in Bulk Queues.* Wiley, New York.

Dvurecenskij, A. (1984). The busy period of order n in the GI/D/∞ queue. *J. Appl. Prob.* **21**, 207-212.

Falin, G. I. and Templeton, J. G. C. (1997). *Retrial Queues.* Chapman & Hall, London.

Gross, D. and Harris, C. M. (1985). *Fundamentals of Queueing Theory*, 2nd edition. Wiley, New York.

Law, A. M. (1975). Efficient estimators for simulated queueing systems. *Management Sci.* **22**, 30-41.

Liu, L. and Shi, D. H. (1996). Busy period in $GI^x/G/\infty$. *J. Appl. Prob.* **33**, 815-829.

Shanbhag, D. N. (1966). On infinite server queues with batch arrivals. *J. Appl. Prob.* **3**, 274-279.

Shedler, G. S. (1987). *Regeneration and Networks of Queues.* Springer-Verlag, New York.

Stadje, W. (1985). The busy period of the queueing system $M/G/\infty$. *J. Appl. Prob.* **22**, 697-704.

Takacs, L. (1962). *Introduction to the Theory of Queues.* Oxford University Press, New York.

DEPARTMENT OF STATISTICS, UNIVERSITY OF TORONTO, TORONTO, ONTARIO, CANADA M5S 3G3