

Analysis of Polling Systems with General Input Process and Finite Capacity

Phuoc Tran-Gia, *Member, IEEE*

Abstract—The class of polling systems, i.e., multiqueue systems with cyclic service, plays an important role in the performance evaluation of various computer and communication systems, e.g., switching systems with distributed control and token-passing local-area networks. Although the behavior of users and the corresponding incoming traffic characteristics in such systems are increasingly complex, most analytical performance studies in the literature are based on the assumption of Poisson input processes. In this paper, an approximate discrete-time analysis of polling systems with finite capacity of waiting places and nonexhaustive service (or more precisely, limited service by one) is presented, considering general renewal input traffic. The analysis method is based on the use of efficient discrete convolution operations based on fast convolution algorithms, e.g., the fast Fourier transform (FFT). To illustrate the accuracy of the approximation and its dependency on system parameters, numerical results are given.

I. INTRODUCTION

IN modeling and performance evaluations of a broad spectrum of computer and communication systems, e.g., investigations of communication structures in switching systems with distributed control and token-passing local-area networks, etc. the class of polling models, i.e., multiqueue systems with cyclic service is often employed. Owing to the architecture of such systems and the respective new services, the traffic mix of subscribers and users on telecommunication networks and user behavior are of high complexity. There is therefore an increasing tendency for the need of more complex processes to describe traffic streams in such environments.

To include more realistic modeling elements in the class of polling systems, two main objectives are taken into account in this paper: i) the consideration of general renewal processes as inputs, and ii) the modeling component in respect of finite waiting capacity of devices or stations in polling systems. From the analytic viewpoint, this study takes advantage of discrete-time analysis methods, where efficient convolution and transformation algorithms [e.g., the fast Fourier transform (FFT)] are employed.

Polling systems have been the subject of numerous studies in the literature [1]–[16]. A survey on the analysis of polling systems can be found in Takagi [15]. Various polling mechanisms, like cyclic or priority order (cf., [11]), and several

service disciplines, e.g., exhaustive, nonexhaustive or gating, have been considered. Some of these studies take into account the switchover time, i.e., the time interval spent by the server in switching over from one queue to the succeeding one. In most of these studies, input processes are assumed to be Poisson, and the queues of the polling system to have infinite capacity.

An approximation of polling systems under symmetrical load conditions, constant switchover time and gating service is given by Leibowitz [10]. Cooper and Murray have considered a system with gating or exhaustive service and zero switchover time [4]. Their approach has been generalized by Eisenberg [5] and Hashida [6] to nonzero switchover time. The case of two queues with nonexhaustive service and nonzero switchover time was taken into account by Boxma [1] where an exact solution has been derived. An approximate analysis technique for polling systems with nonexhaustive service and general switchover time has been developed by Kuehn [8], [9]. Approximate formulas for some general classes of polling systems can be found in Boxma and Meister [2] and Bux and Truong [3]. Morris and Wang [12] and Raith [13] have provided analytical approaches to deal with polling systems with multiple servers, while Raith and Tran-Gia [14] considered the influence of the contention at the receiving part of a polling system in a more general context. Polling systems with finite capacity of waiting places and Poisson input traffic have been analyzed approximately in Tran-Gia and Raith [16].

One of the main features of this study is the use of analysis methods operating in the discrete-time domain. The discrete-time approach is justified by the fact that the parameterization of model components is often based on data measured in terms of histograms. The discrete-time nature of model components can be registered in a number of modeling processes, e.g., in performance investigations of packet-switching systems, time-slotted systems, etc.

There is a number of studies [17]–[22] which deal with the analysis of discrete-time models. In Hunter [18] and Kobayashi [19] surveys can be found. Examples for the use of discrete-time analysis methods can be seen in the case of $G/G/1$ queues where numerical solutions are given [17], [20], [21] based on the discrete-time form of the Lindley's relation [23], [24]. An interesting algorithm for calculation of the waiting-time distribution function of the $G/G/1$ queue has been presented by Ackroyd [17], where methods used in signal processing theory (in both time and frequency domains) and fast convolution algorithms are employed. In [21], calculation algorithms can be found for the idle time and interdeparture distributions of the class of $G/G/1$ queues with discrete-

Paper approved by the Editor for Computer Network Performance of the IEEE Communications Society. Manuscript received November 25, 1987; revised May 16, 1988 and July 19, 1990. This paper was presented in part at the Third International Conference on Data Communications Systems and their Performance, Rio de Janeiro, Brazil, June 1987.

The author is with Lehrstuhl für Informatik III, Universität Würzburg, Am Hubland D-8700, Würzburg.
IEEE Log Number 9106304.

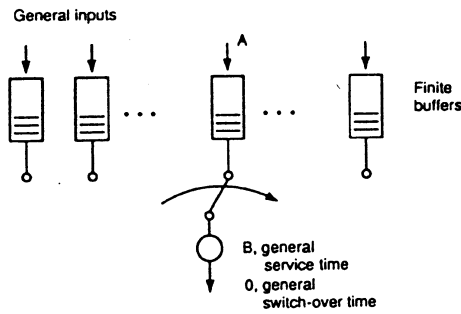


Fig. 1. The basic model.

time arrival and service processes. A solution for this class of systems with general cyclic input processes was given in Tran-Gia and Rathgeb [22], considering models of semidynamic scheduling and routing mechanisms.

Methods of discrete-time analysis will be applied in this paper to obtain an approximation algorithm for the class of polling systems with renewal inputs. In Section II, the model and its parameters will be described, while Sections III and IV give an outline of the analysis and calculation of performance measures, respectively. Numerical results are shown in Section V to illustrate the accuracy of the calculation method and its dependency on system parameters.

II. MODEL DESCRIPTION

The basic structure and related parameters of the polling model considered here are illustrated in Fig. 1. The model consists of g finite capacity queues, served nonexhaustively (or more precisely, limited service of one message per service) in a cyclic order by a single server with a generally distributed service time. After the service of a queue, the server will move to the succeeding queue. This switchover time, which models all overheads spent and procedures performed by the server to move to and scan the succeeding queue, is assumed to have a general distribution function. At the scanning epoch, i.e., at the end of the corresponding switchover time, the server will process one message in the queue, if there is at least one message waiting for service. If the queue is empty, the interscan period observed will consist of just the switchover time.

As previously mentioned, one of the main contributions of this study is the consideration of general input processes and their influence on the behavior of the polling system. Thus, the arrival processes are assumed to be general.

In principle, the analysis method presented in this paper can be applied to nonsymmetrical polling systems with queue-individual interarrival, service and switchover-time distribution functions. Also, individual sizes of the queue capacities can be chosen. The treatment of the nonsymmetrical case leads to a higher computational complexity than in the case of symmetrical systems discussed in this paper. To simplify the description of the analysis algorithm, and to focus our attention on the main objectives of the study, we shall restrict ourselves in the following treatment to the case of symmetrical systems.

III. DISCRETE-TIME ANALYSIS

A. Discrete-Time Random Variables and Notation

In the context of this analysis, we consider the random variables to be of discrete-time nature, i.e., the time axis is conceived to be divided into intervals of unit length Δt . As a consequence, samples of these random variables are integer multiples of Δt ; the time discretization is equidistant. In real systems, Δt is often given in discrete-time form, e.g., as transmission time of a bit, byte or packet. The corresponding distributions can be obtained by means of measurements, and arranged in the form of histograms.

The following notation is used for functions belonging to a discrete-time random variable (r.v.) X :

$$\begin{aligned}
 x(k) &= \Pr(X = k), & -\infty < k < +\infty & \\
 & & \text{distribution of } X & \\
 X(k) &= \sum_{i=-\infty}^k x(i), & -\infty < k < +\infty & \\
 & & \text{distribution function of } X & \\
 x_{ZT}(z) &= \sum_{k=-\infty}^{k=+\infty} x(k)z^{-k} & & \\
 & & \text{Z-transform of } x(k) & \\
 EX & & \text{mean of } X & \\
 c_X & & \text{coefficient of variation of } X &
 \end{aligned}$$

As indicated in Fig. 1, we use the following notation:

- g number of queues in the polling system,
- S queue capacity,
- A r.v. for the interarrival time of the input process at a queue, distribution $a(k)$. Since $a(0)$ can have a nonzero value, batch arrival processes with geometrically distributed batch size can also be dealt with (cf., [21]),
- B r.v. for the service time, distribution $b(k)$,
- O r.v. for the switchover time, distribution $o(k)$.

A sample path of the state process development in a queue chosen arbitrarily, say j , of the polling system is shown in Fig. 2. We observe the cycle time seen from queue j , i.e., the time interval between two consecutive scanning instants at queue j . Similar to the approach provided in [8], two types of conditional cycle times can be distinguished, denoted by the following random variables:

- C_0 r.v. for the cycle time, conditioned on an *empty* queue at the previous scanning instant (i.e., without service of queue j during the cycle considered),
- C_1 r.v. for the cycle time, conditioned on a *nonempty* queue at the previous scanning instant (i.e., with service of queue j).

During a conditional cycle time, a number of messages may arrive according to the arrival process. Depending on the type of conditional cycle, we denote the arrival group as follows:

- G_j r.v. for the number of messages arriving during a cycle of type j (C_j), $j = 0, 1$.

To evaluate the equation system given in (17), it remains to determine the group-size distributions of G_j , $j = 0, 1$. Keeping in mind the assumption of no correlation between the scanning and arrival processes at a queue, we observe a cycle of type j and the arrival group G_j of messages arriving in this cycle. The time $F^{(k)}$ between the previous scanning epoch and the arrival epoch of the k th message in this group is distributed according to

$$f^{(k)}(i) = a^*(i) * a(i)^{[(k-1)*]} \quad (18)$$

where $a(i)^{[m]}$ denotes the m -fold convolution of the distribution $a(i)$ with itself, and $a^*(i)$ denotes the recurrence-time distribution of the interarrival process in the discrete-time domain

$$a^*(i) = \frac{1}{EA} \left(1 - \sum_{n=0}^i a(n) \right), \quad i = 0, 1, \dots \quad (19)$$

Assuming a cycle C_j of length m , the conditional arrival-group size distribution can be given as follows:

$$\begin{aligned} g_j(k|m) &= \Pr(\text{group size is } k | \text{cycle } C_j \text{ is of length } m) \\ &= \Pr(F^{(k)} < m \leq F^{(k+1)}) \\ &= \Pr(F^{(k)} < m) - \Pr(F^{(k+1)} < m) \end{aligned} \quad (20)$$

or

$$\begin{aligned} g_j(k|0) &= \delta(k) \\ g_j(k|m) &= \sum_{i=0}^{m-1} (f^{(k)}(i) - f^{(k+1)}(i)), \quad m = 1, 2, \dots \end{aligned} \quad (21)$$

Finally, the distribution of an arrival group G_j during a cycle C_j [distribution $c_j(m)$, $m > 0$] can be obtained:

$$\begin{aligned} g_j(k) &= c_j(0) \cdot \delta(k) + \sum_{m=1}^{\infty} c_j(m) \sum_{i=0}^{m-1} \\ &\quad \cdot (f^{(k)}(i) - f^{(k+1)}(i)), \\ k &= 0, 1, \dots, j = 0, 1. \end{aligned} \quad (22)$$

C. Cycle-Time Analysis

Define C_S to be the r.v. for the cycle-time segment (cf. [16]), i.e., the time interval between the scanning instants of two consecutive queues. In a symmetrical polling system under stationary conditions with the state probabilities obtained from (17), we obtain

$$c_S(k) = x(0) \cdot a(k) + (1 - x(0)) \cdot (a(k) * b(k)) \quad (23)$$

or in Z -transform domain:

$$c_{S,ZT}(z) = a_{ZT}(z) \cdot (x(0) + (1 - x(0)) \cdot b_{ZT}(z)). \quad (24)$$

Under the assumption of independency between cycle-time segments, the Z -transforms of the conditional cycle times can be given as

$$c_{1,ZT}(z) = a_{ZT}(z) \cdot b_{ZT}(z) \cdot c_{S,ZT}^{j-1}(z). \quad (26)$$

Finally, we obtain the Z -transform of the cycle-time distribution:

$$\begin{aligned} c_{ZT}(z) &= c_{S,ZT}^j(z) \\ &= x(0) \cdot c_{0,ZT}(z) + (1 - x(0)) c_{1,ZT}(z). \end{aligned} \quad (27)$$

Equations (25)–(27) can be evaluated efficiently using algorithms belonging to the class of the fast Fourier transform (FFT) algorithms (based on the discrete Fourier transform (DFT), cf. [21], [22], [25]).

D. Discrete-Time Calculation Algorithm

The logical interrelation of the analysis steps in Sections III-B and -C above is as follows. Assuming that the conditional cycle-time distributions are known, the Markov-chain state probabilities can be calculated according to (17)–(22). On the other hand, to calculate the conditional cycle times, as formulated in (23), (25), and (26), we need to know the component $x(0)$ of the Markov-chain state-probability vector. This fact leads to a numerical calculation algorithm based on an alternating evaluation of the cycle-time distributions and the state-probability vector. The approximation algorithm is similar to the continuous-time algorithm presented in [16]. However, by taking advantage of fast convolution algorithms and the efficient evaluation of Z -transform using FFT algorithms, the analysis in the discrete-time domain allows the calculation of all the entire distributions (e.g., those of cycle times). This is in contrast, e.g., to algorithms which use a number of moments to characterize distribution functions (e.g., the two-moment matching method).

The main steps of the algorithm are as follows.

- i) Initialize the Markov-chain state-probability vector and the conditional cycle-time distributions.
- ii) Calculate the conditional cycle-time distributions according to (23), (25), and (26).
- iii) Calculate the arrival-group size distributions according to (18), (19), and (22).
- iv) Calculate the Markov-chain state-probability vector according to (17).
- v) Repeat steps ii), iii), and iv) until a convergence criterion is fulfilled. In the analysis here, the difference between two consecutive means of the state-probability vector is used as convergence criterion.

IV. SYSTEM CHARACTERISTICS

Using the equilibrium Markov-chain state probabilities obtained by the algorithm presented above, a number of performance measurements of interest can be derived, which are discussed in this section.

A. Message Blocking Probability

Again, the assumption is made that the scanning and arrival processes are uncorrelated. We first consider the following conditional blocking probability defined by

$$B(k) = \Pr(\text{message is blocked} | X = k)$$

Probability for a message in an arrival group to be rejected, conditioned on a number $X = k$ of

messages waiting in the queue observed at the previous scanning epoch.

Two cases have to be distinguished, $X = 0$ and $X > 0$. Derivation of the conditional blocking probability for case $X > 0$, will be described in detail.

For $X > 0$, i.e., the current cycle time is of type G_1 , we observe an arbitrarily chosen test message in the group arriving during the cycle. The probability of the test message being in an arrival group of size i is $i \cdot g_1(i)/EG_1$. Assuming further the group size i , blocking will occur for $k - 1 + i > S$ where a fraction of $k - 1 + i - S$ of messages will be rejected. Accordingly, the probability of the test message being in the rejected fraction is $(k - 1 + i - S)/i$. Finally, for case $X > 0$ we obtain the following *conditional blocking probability* of messages:

$$\begin{aligned} B(k) &= \sum_{i=S-k+2}^{\infty} \frac{k-1+i-S}{i} \cdot \frac{i \cdot g_1(i)}{EG_1} \\ &= \frac{1}{EG_1} \sum_{i=S-k+2}^{\infty} (k-1+i-S) \cdot g_1(i). \\ k &= 1, \dots, S. \end{aligned} \quad (28)$$

The case of an empty queue at the previous scanning instant ($X = 0$) can be analyzed analogously:

$$\begin{aligned} B(0) &= \sum_{i=S+1}^{\infty} \frac{i-S}{i} \cdot \frac{i \cdot g_0(i)}{EG_0} \\ &= \frac{1}{EG_0} \sum_{i=S+1}^{\infty} (i-S) \cdot g_0(i). \end{aligned} \quad (29)$$

By eliminating the condition, we arrive at the blocking probability for messages arriving at a queue of the polling system

$$\begin{aligned} B_A &= \sum_{k=0}^S x(k) \cdot B(k) \\ &= \frac{1}{EG_0} x(0) \sum_{i=S+1}^{\infty} (i-S) \cdot g_0(i) \\ &\quad + \frac{1}{EG_1} \sum_{k=1}^S x(k) \sum_{i=S-k+2}^{\infty} (k-1+i-S) \\ &\quad \cdot g_1(i). \end{aligned} \quad (30)$$

B. Arbitrary-Time State Probabilities

From the Markov-chain state-probability vector, which describes the state process at regeneration points chosen immediately prior to the scanning epochs, the state process seen by an arbitrary outside observer, which will be characterized by means of an arbitrary-time state-probability vector, can be derived. Using this probability vector, further system characteristics, e.g., the mean waiting time for messages, can be obtained.

The state process will now be observed at an instant t^* chosen arbitrarily. Since we are operating in the discrete-time

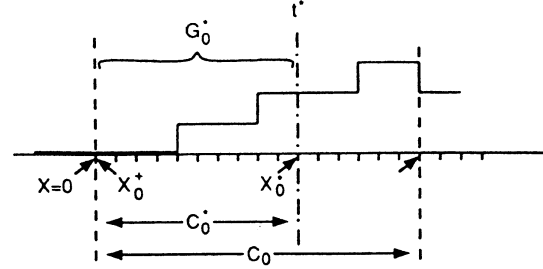


Fig. 3. Observation of state process at an arbitrary instant.

domain, t^* is conceived to be immediately prior to a discrete-time epoch (cf., Fig. 3). The probability P_j^* ($j = 0, 1$) that the observation instant t^* lies in a conditional cycle of type C_j can be given as follows, using results of semi-Markov processes:

$$P_0^* = x(0) \frac{EC_0}{EC} \quad (31)$$

$$P_1^* = 1 - P_0^* = (1 - x(0)) \frac{EC_1}{EC}. \quad (32)$$

We further consider the distribution of the recurrence cycle time with r.v. C_j^* , which represents the discrete-time interval from the previous scanning instant up to the observation point t^* :

$$c_j^*(k) = \frac{1}{EC_j} \left(1 - \sum_{i=0}^k c_j(i) \right), \quad k = 0, 1, \dots, j = 0, 1. \quad (33)$$

Denoting G_j^* ($j = 0, 1$) to be the r.v. of the group of messages arriving during a recurrence cycle time of type C_j , we obtain a similar expression for the group-size distribution as in (22) as follows:

$$\begin{aligned} g_j^*(k) &= c_j^*(0) \cdot \delta(k) + \sum_{m=1}^{\infty} c_j^*(m) \sum_{i=0}^{m-1} \left(f^{(k)}(i) - f^{(k+1)}(i) \right). \\ k &= 0, 1, \dots, j = 0, 1. \end{aligned} \quad (34)$$

As shown in Fig. 3, the following r.v. for the state process in the queue being observed are introduced.

X^* r.v. for the state of the queue being observed at t^*
 X, X^+ r.v. for the state of the queue observed immediately prior to and after the previous scanning epoch, respectively.

and, similar to the Markov-chain analysis [(cf., (1) and (2))] the random variables conditioned on the state prior to the previous scanning instant:

$$X_0^+ = X^+ | X = 0, \quad X_1^+ = X^+ | X > 0. \quad (35)$$

$$X_0^* = X^* | X = 0, \quad X_1^* = X^* | X > 0. \quad (36)$$

To calculate the probability vector for X^* , the following set of equations describing the interrelations between state random variables and corresponding distributions is required.

$$i) X_n = 0$$

$$X_0^+ = 0 \quad (37)$$

$$x_0^+(k) = \delta(k) \quad (38)$$

$$X_0^* = \min(X_0^+ + G_0^*, S) = \min(G_0^*, S) \quad (39)$$

$$x_0^*(k) = \pi^S(x_0^+(k) * g_0^*(k)). \quad (40)$$

ii) $X_n > 0$

$$X_1^+ = X - 1 \quad (41)$$

$$x_1^+(k) = \begin{cases} \frac{x(k+1)}{\Pr(X>0)} = \frac{x(k+1)}{1-x(0)} & k = 0, 1, \dots, S-1 \\ 0 & \text{otherwise} \end{cases} \quad (42)$$

$$X_1^* = \min(X_1^+ + G_1^*, S) \quad (43)$$

$$x_1^*(k) = \pi^S(x_1^+(k) * g_1^*(k)). \quad (44)$$

Hence, the arbitrary-time state probabilities can be given as

$$x^*(k) = P_0^* \cdot x_0^*(k) + P_1^* \cdot x_1^*(k), \quad k = 0, 1, \dots, S. \quad (45)$$

Using (31)–(33) and (37)–(44), we obtain the final set of equations to determine the arbitrary-time state probabilities:

$$\begin{aligned} x^*(k) &= \frac{EC_0}{EC} x(0) \cdot g_0^*(k) + \frac{EC_1}{EC} \sum_{i=1}^{k+1} x(i) \\ &\quad \cdot g_1^*(k-i+1), \quad k = 0, \dots, S-1 \\ x^*(S) &= \frac{EC_0}{EC} x(0) \sum_{i=S}^{\infty} g_0^*(i) + \frac{EC_1}{EC} \sum_{i=1}^S x(i) \\ &\quad \cdot \sum_{m=S-i+1}^{\infty} g_1^*(m), \quad k = S. \end{aligned} \quad (46)$$

C. Waiting Time of Accepted Messages

Having the state probabilities of a queue seen by an arbitrary outside observer, we can use Little's theorem to determine the mean waiting time of *accepted* messages

$$EW_A = \frac{EX^* \cdot EA}{1 - B_A}. \quad (47)$$

V. RESULTS AND APPROXIMATION ACCURACY

Taking as example a symmetrical polling system with $g = 5$ queues having finite capacities $S = 5$, numerical results are presented in this section to discuss the application of the calculation algorithm and the approximation accuracy. The switchover time is assumed to be deterministic equal to 0.5 EB, and the discrete-time axis is scaled to $\Delta t = 1$. Time variables are normalized to the mean service time $EB = 10 \cdot \Delta t$. The offered traffic intensity is denoted by $\rho = EB/EA$.

To validate the approximate analysis, computer simulations are provided. The simulation results will be depicted with their 95% confidence intervals, calculated using the Student- t test technique.

It should be noted here that the results presented below focus on the influence of the characteristics of input processes, which are the main new modeling components of this study.

To obtain a parametric representation of various random-process types, we consider here the interarrival and service times having distributions given by their two parameters, e.g.,

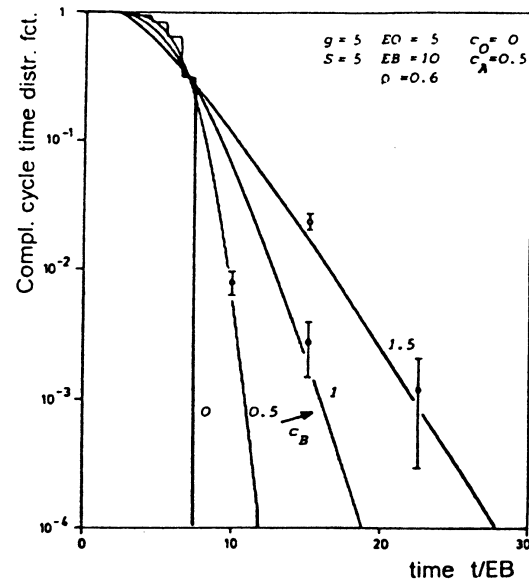


Fig. 4. Influence of arrival process on waiting time.

the mean and the coefficient of variation. For this purpose, we employ the negative binomial distribution. Thus, for an r.v. X with mean EX and coefficient of variation c_X , the distribution is

$$x(k) = \binom{y+k-1}{k} p^y (1-p)^k, \quad 0 \leq p < 1, y \text{ real} \quad (48)$$

where

$$p = \frac{1}{EX \cdot c_X^2}, \quad y = \frac{EX}{EX \cdot c_X^2 - 1}, \quad EX \cdot c_X^2 > 1.$$

As mentioned above, all the entire distributions of the random variables of interest can be obtained, using discrete-time analysis methods. An example is given in Fig. 4 where the complementary cycle-time distribution function is depicted for different values of the service-time coefficient of variation c_B . Two main effects can be seen clearly, here: i) the step-wise functions, and ii) the geometric caudal characteristics of the discrete-time distribution functions obtained.

In Figs. 5 and 6, the mean and the coefficient of variation of the cycle time are depicted as functions of the offered traffic intensity ρ for different types of interarrival and service processes. The choice of parameters in these figures is based on detailed studies showing that the mean cycle time EC is affected mainly by the type of arrival processes, while the cycle-time coefficient of variation c_C is primarily influenced by the type of service processes. It should be noted here that the coefficients of variation of these discrete-time processes are chosen to be equivalent to the deterministic ($c_A, c_B = 0$), the Erlangian of fourth order ($c_A, c_B = 0.5$), the Markovian ($c_A, c_B = 1$), and the hyperexponential ($c_A, c_B = 1.5$, cf. [16]) distributions. With respect to the traffic level, the two lower and upper bounds of the mean cycle time can be observed (cf., Fig. 5): i) the *empty cycle* at disappearing traffic intensities (sum of just switchover times), and ii) the

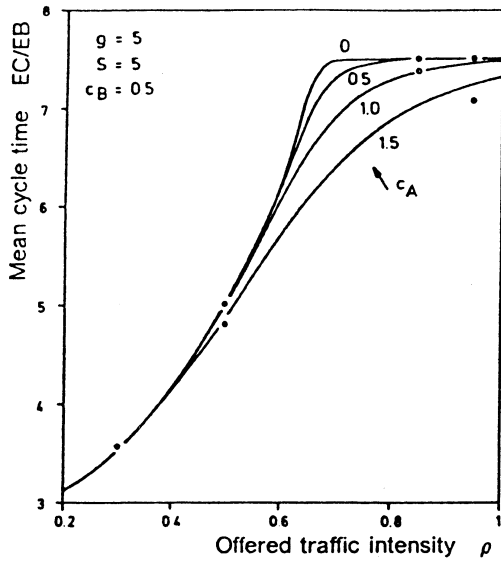


Fig. 5. Complementary cycle-time distribution function.

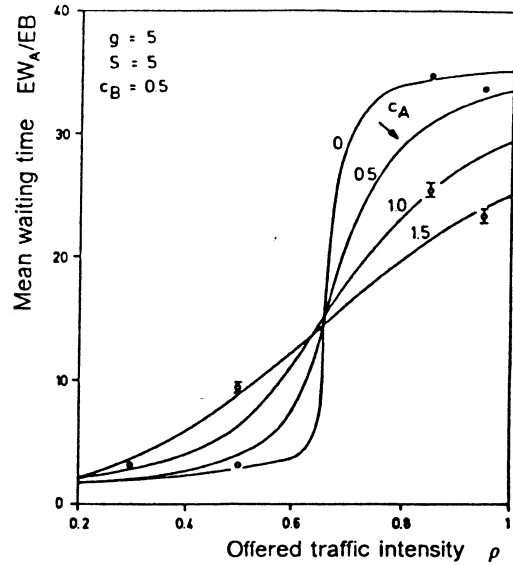


Fig. 7. Influence of service process on cycle-time coefficient of variation.

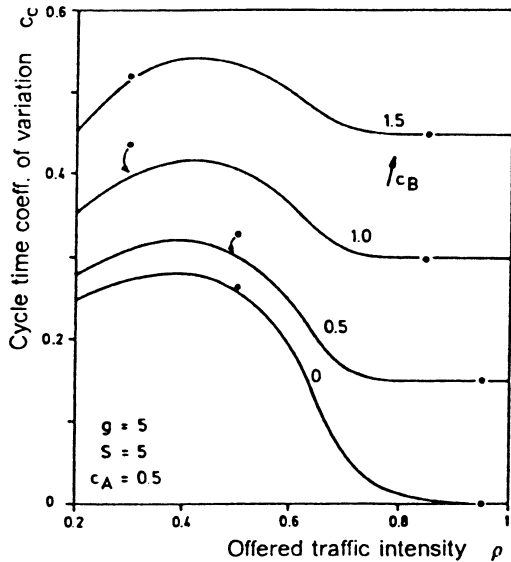


Fig. 6. Influence of arrival-process characteristics on mean cycle time.

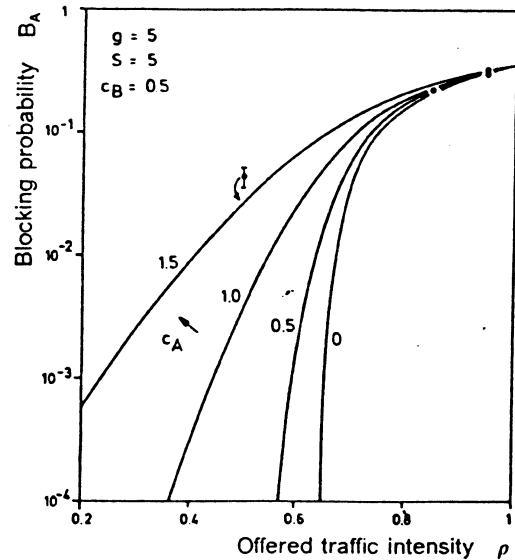


Fig. 8. Influence of arrival process on blocking probability.

maximal cycle at very high traffic levels (service at each queue during the polling cycle). For arrival processes with higher coefficients of variation, the blocking effect will become noticeable, which leads to smaller cycle lengths. As depicted in Fig. 6, the cycle-time coefficient of variation has a maximum value which increases with increasing service-time coefficient of variation.

The blocking probability of arriving messages and the mean waiting time of accepted messages are depicted in Figs. 8 and 7, respectively, where attention is again devoted to the influence of the arrival process. The mean waiting time is drawn in Fig. 7 for accepted messages. Hence, a crossover phenomenon can be recognized, which can be verified on considering the

higher blocking probability, i.e., smaller number of accepted messages, for larger values of c_A . The cross point of the curves in Fig. 7 corresponds to the value of the traffic intensity, where the blocking probability becomes significant (cf., Fig. 8). For the case of disappearing traffic intensities, the waiting time of messages consists of just the recurrence time of empty cycles.

As seen in the diagrams, with the exception of blocking probabilities at higher c_A values, the overall approximation accuracy for the system parameters given ($g = 5$, $S = 5$, $EO = 0.5 EB$) is sufficient for system-engineering purposes. However, two main restriction factors for use of the approximation should be mentioned here: i) the accuracy of the approximate analysis is of decreasing tendency for smaller

values of switchover time and higher values of c_A, c_B , and ii) the computing efforts increase overproportionally with the values of g and S , according to the lengths of probability vectors involved in convolution operations.

Although the computational scheme is complex, for the parameter range discussed in this section, the convergence behavior of the algorithm is good. For the parameters shown above in the figures, convergence has been typically reached after less than 50 iteration cycles.

VI. CONCLUSION

In this paper, an approximate algorithm for polling systems with finite capacity of waiting places and nonexhaustive service (or more precisely, limited service of one message per service) is presented. The analysis is made in the discrete-time domain, based on the evaluation of discrete convolution operations taking advantage of fast convolution algorithms, e.g., the fast Fourier transform. Attention is devoted to two essential modeling aspects: i) consideration of general renewal input traffic, and ii) assumption of finite capacity of waiting places in the system. Numerical examples are shown to illustrate the approximation accuracy of the analysis. The approximation is validated by means of computer simulations. The class of polling models considered here can be used in the performance investigation of a broad spectrum of models in computer and communication systems.

ACKNOWLEDGMENT

The author would like to thank M. Böpplé for helpful discussions and valuable programming support.

REFERENCES

- [1] J. Boxma, "Two symmetric queues with alternating service and switching times," in *Proc. Performance '84*, Paris, France, 1984, pp. 409-431.
- [2] O. J. Boxma and B. Meister, "Waiting time approximations for cyclic-service systems with switch-over times," in *Proc. Joint Conf. Performance '86 ACM Sigmetrics*, Raleigh, NC, 1986, pp. 254-262.
- [3] W. Bux and H. L. Truong, "Mean-delay approximation for cyclic service queueing systems," *Perform. Eval.*, vol. 3, pp. 187-196, Aug. 1983.
- [4] R. B. Cooper and G. Murray, "Queues served in cyclic order," *Bell Syst. Tech. J.*, vol. 48, pp. 675-689, Mar. 1969.
- [5] M. Eisenberg, "Queues with periodic service and changeover times," *Oper. Res.*, vol. 20, pp. 440-451, Mar.-Apr. 1972.
- [6] O. Hashida, "Analysis of multiqueue," *Rev. El. Commun. Lab.*, vol. 20, pp. 189-199, Mar.-Apr. 1972.
- [7] O. Hashida and K. Ohara, "Line accommodation capacity of a communication control unit," *Rev. El. Commun. Lab.*, vol. 20, pp. 231-239, 1972.
- [8] J. Kuehn, "Multiqueue systems with nonexhaustive cyclic service," *Bell Syst. Tech. J.*, vol. 58, pp. 671-699, Mar. 1979.
- [9] P. J. Kuehn, "Performance of ARQ-protocols for HDX-transmission in hierarchical polling systems," *Perform. Eval.*, vol. 1, pp. 19-30, Jan. 1981.
- [10] M. A. Leibowitz, "An approximate method for treating a class of multiqueue problems," *IBM J. Res. Develop.*, vol. 5, pp. 204-209, July 1961.
- [11] D. R. Manfield, "Analysis of a polling system with priorities," in *Proc. GLOBECOM '83*, San Diego, CA, 1983, paper 43.4.
- [12] R. J. T. Morris and Y. T. Wang, "Some results for multiqueue systems with multiple cyclic servers," in *Proc. 2nd Int. Symp. Perform. Comput. Commun. Syst.*, Zurich, Switzerland, 1984, pp. 245-258.
- [13] T. Raith, "Performance analysis of multibus interconnection networks in distributed systems," in *Proc. 11th Int. Teletraffic Congr.*, Kyoto, Japan, 1985, paper 4.2A-5.
- [14] T. Raith and P. Tran-Gia, "Performance analysis of polling mechanisms with receiver blocking," in *Proc. 8th Int. Conf. Comput. Commun. (ICCC)*, Munich, W. Germany, 1986, p. C5-1.
- [15] H. Takagi, *Analysis of Polling Systems*. Cambridge, MA: M.I.T. Press, 1986.
- [16] P. Tran-Gia and T. Raith, "Multiqueue systems with finite capacity and nonexhaustive cyclic service," in *Proc. Int. Sem. Comput. Networking Perform. Eval.*, Tokyo, Japan, 1985.
- [17] M. H. Ackroyd, "Computing the waiting time distribution for the $G/G/1$ queue by signal processing methods," *IEEE Trans. Commun.*, vol. COM-28, pp. 52-58, Jan. 1980.
- [18] J. J. Hunter, *Mathematical Techniques of Applied Probability, Vol. 2: Discrete Time Models, Techniques and Applications*. London: Academic, 1983.
- [19] H. Kobayashi, "Stochastic modelling: Queueing models; discrete-time queueing systems," in *Probability Theory and Computer Science Part II*, G. Louchard and G. Latouche, Eds. London: Academic, 1983.
- [20] A. G. Konheim, "An elementary solution of the queueing system $GI/G/1$," *SIAM J. Comput.*, vol. 4, pp. 540-545, 1975.
- [21] P. Tran-Gia, "Discrete-time analysis for the interdeparture distribution of $GI/G/1$ queues," in *Proc. Sem. Teletraffic Anal. Comput. Perform. Eval.*, Amsterdam, The Netherlands, 1986.
- [22] P. Tran-Gia and E. Rathgeb, "Performance analysis of semidynamic scheduling strategies in discrete-time domain," in *Proc. INFOCOM '87*, San Francisco, CA, 1987.
- [23] D. V. Lindley, "The theory of queues with a single server," *Proc. Cambridge Philos. Soc.*, vol. 48, pp. 277-289, 1952.
- [24] W. L. Smith, "On the distribution of queueing times," *Proc. Cambridge Philos. Soc.*, vol. 49, pp. 449-461, 1953.
- [25] P. Henrici, "Fast Fourier methods in computational complex analysis," *Siam Rev.*, vol. 21, pp. 481-527, Oct. 1979.



Phuoc Tran-Gia (M'80) received the M.S. degree (Dipl.-Ing.) from the Stuttgart University in 1977 and the Ph.D. degree (Dr.-Ing.) from the University of Siegen in 1982, both in electrical engineering.

In 1977, he joined Standard Elektrik Lorenz (ITT), Stuttgart, where he worked in software development of digital switching systems. From 1979 until 1982 he worked as an assistant professor in the Department of Communications of the University of Siegen. From 1983 to 1986, he was head of a research group at the Institute of Communications Switching and Data Techniques, Stuttgart University, Germany. His research activities were in the field of environment simulations for communication systems and applied queueing theory, especially with focus on discrete-time methods. In 1986, he joined IBM Zürich Research Laboratory where he worked on the architecture and performance evaluation of computer communication systems. Since July 1988 he has been Professor in the Faculty of Mathematics and Computer Sciences (Distributed Systems) at the University of Würzburg, Germany. His current research areas are performance modelling of high-capacity LAN's, MAN's, communication and manufacturing systems as well as applications of neural networks in telecommunication environments.