DISCRETE-TIME APPROXIMATION OF THE MACHINE INTERFERENCE PROBLEM WITH GENERALLY DISTRIBUTED FAILURE, REPAIR, AND WALKING TIMES

Manfred Mittler
Universität Würzburg
Lehrstuhl für Informatik III
Am Hubland, D-97074 Würzburg, Germany

Christoph Kern
University of British Columbia
Department of Computer Science
2366 Main Mall, Vancouver, B.C. Canada V6T 1Z4

Summary. In this paper we present an approximate discrete-time analysis of the machine interference problem with generally distributed working, failure, and walking times. We analyze the corresponding asymmetric polling system consisting of N separate closed queuing stations (buffer and server), each of which contains a single customer. These queuing stations are attended to by a common server in a cyclic manner. The walking times between the stations are non-zero. Upon service completion of the common server a customer requires service by its dedicated server. Then it is routed back to its corresponding buffer where it has to wait for service by the common server once again.The analysis is exact except for the assumption that the contributions of the queuing stations to the cycle time of the common server are mutually independent. We first calculate the stationary state probabilities using an iterative procedure and then derive the cycle time and waiting time distributions and their first and second moments. Several numerical examples comparing approximate results to simulations illustrate the accuracy of our method.

Keywords. Polling System, Finite Population Model, Discrete-Time Analysis

1 Introduction

To analyze the machine interference problem (MIP, see for example Stecke and Aronson 1985) where a single repairman walks cyclically from one machine to the next in non-zero time and repairs those machines that are broken down, we consider a basic cyclic single buffer polling system. We extend the basic polling model by the following modifications: In addition to the common server S, station i has its own dedicated server L_i. Each station contains exactly one customer. Upon service completion at the dedicated server this customer is routed to the station's buffer Q_i, where it waits for service by the common server. After termination of the service period at the common server the customer returns to the corresponding dedicated server. Thus, the resulting queuing system consists of N separate closed queuing stations (buffer and server), each of which contains a single customer and is attended to by a common server. We therefore refer to this polling model as a *closed* polling model or polling model with *finite population*.

In the context of the MIP, the service times of the dedicated servers correspond to machine failure times whereas operator repair times are represented by the service times of the common server. In the terminology of polling systems, the walking times of the patrolling operator are referred to as switchover times.

The contribution of this paper is as follows: based on the discrete-time method developed by Tran-Gia (1986) we present an approximation for the cycle time and waiting time distribution function of the closed polling system with a single customer per service station where all service times are assumed to be arbitrarily distributed. This model corresponds exactly to the machine interference problem with a single patrolling operator. We utilize the approach presented by Tran-Gia (1992), assuming that the contributions by the queuing stations to the cycle time are mutually independent. Our analysis is exact except for this assumption. In contrast to Tran-Gia (1992), we do not assume independence between the arrival process and the scanning process of the polling server.

2 Literature

Polling systems became popular in the performance evaluation of cyclic service systems arising out of manufacturing and telecommunication systems (see Takagi (1990, 1991) and Haverkort 1994).

Mack et al. (1957) and Mack (1957) analyzed the symmetric MIP with a patrolling operator, Poisson arrivals, constant switchover times, and either constant service times or service times following discrete distributions. Apart from these publications, all other references deal with open polling systems. Hashida and Kawashima (1981) obtained approximate results for an asymmetric single buffer polling system with exponentially distributed interarrival times and generally distributed service and switchover times. Takagi (1985) analyzed the aforementioned system in the symmetric case. Ibe and Cheng (1989) and Takine et al. (1988) presented an exact solution of Hashida and Kawashima's model. They employed the method of station times to obtain the Laplace-Stieltjes-transforms of both waiting time and cycle time distribution functions. Tran-Gia and Raith (1988) investigated an asymmetric polling system with finite buffers, limited-1 service discipline, and Poisson arrivals. Service and switchover times were generally distributed. Their approximation yields mean waiting times and the first two moments of cycle times. Tran-Gia (1992) extended this analysis to generally distributed interarrival times using his discrete-time approach (see Tran-Gia 1986). Chung et al. (1994) derived the Laplace-Stieltjes-transforms of the waiting time and cycle time distribution functions for a single buffer polling system with Markovian service discipline. Finally, we refer to the work of Takine et al. (1990) who investigated two different single buffer polling systems, the *conventional* system and a *buffer relaxation* system, which are also considered by Takine et al. (1988). In the latter a customer releases the buffer upon beginning of service whereas in the conventional system the buffer remains occupied until the service time has elapsed. If interarrival times are exponentially distributed, the open conventional system corresponds exactly to the finite population polling system as described in the previous section.

3 Analysis

The model is analyzed in the discrete-time domain, i.e., the time axis is divided into equidistant intervals of length Δt. Events can occur only at the boundaries between these unit intervals. This implies that samples of the random variables involved are restricted to values which are integer multiples of Δt. Furthermore, we assume that arrival events are scheduled before polling events.

For a discrete random variable (r.v.) X, we use the following notation. The distribution of X, $\mathrm{P}\{X = k \cdot \Delta t\}$, is denoted by $x(k)$, whereas $\mathrm{P}\{X \leq k \cdot \Delta t\}$ is the distribution function of X. In each case, k is a non-negative integer. To simplify the notation, we use the normalized notation $\mathrm{P}\{X = k\}$ instead of $\mathrm{P}\{X = k \cdot \Delta t\}$. Further, we denote by x, $x^{(m)}$, and c_X the mean, the mth moment, and the coefficient of variation (CoV) of X, respectively. The expression $x_1(k) \circledast x_2(k)$ is the discrete convolution of the distributions $x_1(k)$ and $x_2(k)$. Finally, we denote by $x^{[\circledast m]}(k)$ the m-fold discrete convolution of the distribution $x(k)$ with itself.

The service times of the central server and of the dedicated servers are denoted by B_i and A_i, respectively. We assume that both B_i and A_i are i.i.d. random variables with arbitrary distribution functions $B_i(t)$ and $A_i(t)$, respectively. The switchover time (walking time) U_i from station i to station $i+1$ is assumed to be i.i.d. with distribution function $U_i(t)$. Stations are numbered cyclically such that the predecessor station of station 1 is station N.

3.1 Polling Process

Figure 1 depicts a sample path of the polling process, as seen from station i. Let $q_i(t)$ be the r.v. of the occupancy of buffer Q_i at time instant t. Then $q_i(t) = 0$ denotes a free buffer and $q_i(t) = 1$ indicates an occupied buffer. Let $T_{P,i}^{(n)}$ denote the nth *polling instant* at station i. Let $q_i^{(n)} \triangleq q_i(T_{P,i}^{(n)})$ be the r.v. of the occupancy of Q_i at the polling instant. In addition, let $\pi_i^{(n)} \triangleq \mathrm{P}\{q_i^{(n)} = 1\}$ the probability that S finds Q_i occupied at $T_{P,i}^{(n)}$.

The r.v. for the service time at station i in the nth cycle is denoted by $B_i^{(n)}$. Obviously, $B_i^{(n)}$ is not defined if $q_i^{(n)} = 0$. The time instant $T_{D,i}^{(n)}$, at which S leaves station i, is called *server departure instant*. Hence, we have $T_{P,i}^{(n)} = T_{D,i}^{(n)}$ if $q_i^{(n)} = 0$.

The time interval from $T_{D,i-1}^{(n)}$ to $T_{P,i}^{(n)}$ is called *switchover time*; the corresponding r.v. is denoted by $U_i^{(n)}$. The time spent by a customer in its local server is called *interarrival time* (from S's point of view). Let $A_i^{(n)}$ denote the r.v. of the interarrival time starting at $T_{D,i}^{(n)}$. Let $W_i^{(n)}$ denote the *waiting time* ending at $T_{P,i}^{(n)}$, which comprises the time from the instant at which J_i enters Q_i to the subsequent polling instant $T_{P,i}^{(n)}$. Both $W_i^{(n)}$ and $A_i^{(n)}$ are not defined if $q_i^{(n)} = 0$.

The r.v. for the length of the time interval from $T_{D,i-1}^{(n)}$ to $T_{D,i}^{(n)}$, the *station time* of station i in the nth cycle, is denoted by $C_{S,i}^{(n)}$. The sum of the N station times comprising a full cycle is called *cycle time*. It is equal to the time interval from $T_{D,i}^{(n)}$ to $T_{D,i}^{(n+1)}$. The corresponding random variable is given by

$$C_i^{(n)} = \sum_{l=i+1}^{N} C_{S,l}^{(n)} + \sum_{l=1}^{i} C_{S,l}^{(n+1)}. \qquad (1)$$

Finally, we define the *intervisit time* as the interval from $T_{D,i}^{(n)}$ to $T_{P,i}^{(n+1)}$, whose r.v. is denoted by $V_i^{(n)}$.

If a station's buffer is found empty at the polling instant in a particular cycle, and therefore no service period occurs, this cycle is called an *idle* cycle from this station's point of view. If a cycle is an idle cycle, the cycle time equals the intervisit time. The sequence of cycles starting from the end of a service period $B_i^{(n_0)}$ (i.e., the beginning of the interarrival time $A_i^{(n_0)}$) until the end of $A_i^{(n_0)}$ is called a *cycle group*. A cycle group with the arrival occurring in its kth cycle is called a *k-cycle-group*. For a k-cycle-group starting at $T_{D,i}^{(n_0)}$, we define the *l-polling-instants*

$$T_{G,i}^{(n_0+l)} \triangleq \begin{cases} T_{D,i}^{(n_0)}, & l = 0, \\ T_{P,i}^{(n_0+l)}, & 0 < l \leq k. \end{cases} \qquad (2)$$

Note that a k-cycle-group consists of $k-1$ idle cycles and one cycle in which a customer is to be served.

3.2 State Probabilities

Our approximate analysis is based on the assumption that the station time random variables $C_{S,i}^{(n)}$, $i = 1, \ldots, N$, $n \in \mathbb{N}$, are mutually independent. We first derive the steady-state probabilities $\pi_i \triangleq \lim_{n \to \infty} \pi_i^{(n)}$, $i = 1, \ldots, N$, for a buffer being occupied at a polling-instant, expressed in terms of the intervisit time distribution. Subsequently, we obtain the reverse relationship between the intervisit time distribution and the buffer occupancy probabilities. These two relationships can be combined into an iterative algorithm which allows the numerical computation of these and other performance measures.

In order to establish the first of the two relations, we identify an embedded Markov chain in the polling process. The following arguments can be carried out for each station in isolation. To simplify the notation we will henceforth drop the station index i.

Noting that we do not assume the intervisit time to be geometrically distributed, which consequently does not in general exhibit the memoryless property, it becomes clear that the state description of the embedded Markov chain must account for the elapsed (or the remaining) interarrival time. We arrive at such a state description by extending the state space by a *supplementary variable* (see Cox 1955) for the number of the polling instants within a cycle group.

Let $L^{(n)}$ denote the r.v. for the number of the current l-polling-instant within the current cycle group, if $q^{(n)} = 0$; $L^{(n)} = 0$ otherwise. We consider the stochastic process $\{(q^{(n)}, L^{(n)})\}_{n\in\mathbb{N}}$, $(q^{(n)}, L^{(n)}) \in \{(1,0),(0,1),(0,2),\dots\}$, embedded at polling instants $T_P^{(n)}$. We will show that, under the assumption of independent station times, the probability for an arrival within the next cycle depends only on the state at the current polling instant and not on the past development of the process. The process $\{(q^{(n)}, L^{(n)})\}_{n\in\mathbb{N}}$ is therefore a Markov process. Its state probabilities are defined by $\pi_{k,l}^{(n)} \triangleq \mathrm{P}\{q^{(n)} = k, L^{(n)} = l\}$.

Consider an arbitrary cycle group starting at $T_D^{(n_0)}$. Let

$$P_{A,l} \;\triangleq\; \mathrm{P}\{\text{arrival in } (T_G^{(n_0+l)}, T_G^{(n_0+l+1)}] \mid A^{(n_0)} > T_G^{(n_0+l)} - T_G^{(n_0)}\}, \quad l \geq 0, \tag{3}$$

be the probability for an arrival within cycle $l+1$ of the cycle group, under the condition that no arrival occurred during the previous l cycles. Then, the following possible state transitions can be observed during a cycle:

- from state $\left(q^{(n)} = 1, L^{(n)} = 0\right)$ into state
  - $\left(q^{(n+1)} = 1, L^{(n+1)} = 0\right)$ with prob. $P_{A,0}$,
  - $\left(q^{(n+1)} = 0, L^{(n+1)} = 1\right)$ with prob. $1 - P_{A,0}$,
- from state $\left(q^{(n)} = 0, L^{(n)} = l\right)$, $l \geq 1$, into state
  - $\left(q^{(n+1)} = 1, L^{(n+1)} = 0\right)$ with prob. $P_{A,l}$,
  - $\left(q^{(n+1)} = 0, L^{(n+1)} = l+1\right)$ with prob. $1 - P_{A,l}$.

We arrive at the following state equations for the embedded Markov chain:

$$\begin{aligned}
\pi_{1,0}^{(n+1)} &= P_{A,0}\cdot \pi_{1,0}^{(n)} + \sum_{l=1}^{\infty} P_{A,l}\cdot \pi_{0,l}^{(n)}, \\
\pi_{0,1}^{(n+1)} &= (1 - P_{A,0})\cdot \pi_{1,0}^{(n)}, \\
\pi_{0,l}^{(n+1)} &= (1 - P_{A,l-1})\cdot \pi_{0,l-1}^{(n)}, \qquad l > 1.
\end{aligned} \tag{4}$$

Assuming the existence of the limiting probabilities $\pi_{k,l} \triangleq \lim_{n\to\infty} \pi_{k,l}^{(n)}$, and taking into account the normalizing condition $\pi_{1,0} + \sum_{l=1}^{\infty} \pi_{0,l} = 1$, we obtain the steady-state probabilities:

$$\begin{aligned}
\pi_{0,l} &= \pi_{1,0}\cdot \prod_{j=0}^{l-1}(1 - P_{A,j}), \\
\pi_{1,0} &= \left[1 + \sum_{l=1}^{\infty}\prod_{j=0}^{l-1}(1 - P_{A,j})\right]^{-1}.
\end{aligned} \tag{5}$$

The only state for which $q = 1$ is state $(q = 1, L = 0)$ with $q \triangleq \lim_{n\to\infty} q^{(n)}$ and $L \triangleq \lim_{n\to\infty} L^{(n)}$. Therefore, the steady-state probability of a customer being present at the polling instant is equal to $\pi_{1,0}$.

We proceed to the derivation of the conditional arrival probability $P_{A,l}$. Again, consider an arbitrary cycle group starting at $T_D^{(n_0)}$. For sake of a concise notation, we shift the time index so that $T_D^{(n_0)} = T_D^{(0)}$. We start with the definition

$$P_{B,l} \;\triangleq\; \mathrm{P}\{T_G^{(l)} - T_G^{(0)} < A^{(0)} \leq T_G^{(l+1)} - T_G^{(0)}\}$$

of the unconditional probability for an arrival in exactly the $(l+1)$th cycle of a cycle group ($l \geq 0$). Due to the assumption of independent station times, consecutive intervisit times must be independent as well. Thus, the length of l consecutive intervisit times is distributed according to the l-fold convolution of the intervisit time distribution. Hence,

$$\mathrm{P}\{T_G^{(l)} - T_G^{(0)} = k\} \;=\; v^{[\circledast l]}(k), \tag{6}$$

because the time interval from $T_G^{(0)}$ to $T_G^{(l)}$ is composed of exactly l consecutive idle cycles, and thus l consecutive intervisit times. With an analogous argument, after some algebraic manipulations, we obtain

$$\begin{aligned}
P_{B,l} &= \mathrm{P}\{T_G^{(l)} - T_G^{(0)} < A^{(0)} \leq T_G^{(l)} - T_G^{(0)} + V^{(l)}\} \\
&= \sum_{j=0}^{\infty} A(j)\cdot\left(v^{[\circledast(l+1)]}(j) \;-\; v^{[\circledast l]}(j)\right).
\end{aligned} \tag{7}$$

The second step follows by conditioning the result of the first step on $T_G^{(l)} - T_G^{(0)} = j$ and $V^{(l)} = k$. Then, setting $n_0 = 0$ in Eqn. (3) yields

$$P_{A,l} \;=\; \frac{P_{B,l}}{1 - \mathrm{P}\{A^{(0)} \leq T_G^{(l)} - T_G^{(0)}\}}. \tag{8}$$

The denominator of this expression can be simplified using the same arguments as in the previous paragraph, which results in

$$\mathrm{P}\{A^{(0)} \leq T_G^{(l)} - T_G^{(0)}\} \;=\; \sum_{k=0}^{\infty} A(k)\cdot v^{[\circledast l]}(k). \tag{9}$$

Combining Eqns. (8) and (9) with Eqn. (7), we finally obtain

$$P_{A,l} \;=\; \frac{\displaystyle\sum_{k=0}^{\infty} A(k)\cdot\left(v^{[\circledast(l+1)]}(k) - v^{[\circledast l]}(k)\right)}{\displaystyle 1 - \sum_{k=0}^{\infty} A(k)\cdot v^{[\circledast l]}(k)}. \tag{10}$$

Note that $P_{A,l}$ is a function only of l and the distributions $v(k)$ and $a(k)$, which were assumed to be known and constant throughout this subsection. Thus, the state transition probabilities of the process $\{(q^{(n)}, L^{(n)})\}_{n\in\mathbb{N}}$ depend on the current state only and not on its past history. Therefore, the process indeed possesses the Markov property.

3.3 Intervisit Times In the previous subsection, we derived the steady-state probability of a customer being present in the buffer of a station i at polling instants, under the assumption that the intervisit time distribution is known. Recall that the calculations can be carried out for each station separately.

We now assume in turn that the buffer occupancy probabilities π_i, $i = 1,\dots,N$, are known. Following the arguments developed in Tran-Gia (1992), we obtain the station time distribution

$$c_{S,i}(k) \;=\; \pi_i\cdot[u_i(k) \circledast b_i(k)] \;+\; (1 - \pi_i)\cdot u_i(k). \tag{11}$$

Due to the assumption of independent station times, the intervisit time and cycle time distributions are given as, respectively,

$$\begin{aligned}
v_i(k) &= \mathop{\circledast}_{j=1}^{i} c_{S,j}(k) \circledast u_i(k) \circledast \mathop{\circledast}_{j=i+1}^{N} c_{S,j}(k), \\
c(k) &= c_{S,1}(k) \circledast \cdots \circledast c_{S,N}(k),
\end{aligned} \tag{12}$$

where $\mathop{\circledast}_{j=1}^{i} c_{S,j}(k) \;=\; c_{S,1}(k) \circledast \cdots \circledast c_{S,i}(k)$. Note that the assumed mutual independence of station times implies the independence of consecutive cycle and intervisit times. We used this property in the previous subsection.

3.4 Iterative Algorithm We complete the analysis of the polling process by combining the results from the previous two subsections into an iterative algorithm (cf. Tran-Gia 1992).

1. $\pi_i \leftarrow$ *initial value* $\in (0,1),\ i = 1, \dots, N$
2. **repeat**
2.1 $\pi'_i \leftarrow \pi_i,\ i = 1, \dots, N$
2.2 *compute* $v_i(k)$ *from* $\pi'_i,\ i = 1, \dots, N$
2.3 *compute new values for* $\pi_i,\ i = 1, \dots, N$
2.4 **until** *termination criterion satisfied*
3. *compute* $c(k)$ *and further performance measures.*

Step 2.2 uses Eqns. (11) and (12), while Step 2.3 employs Eqns. (10) and (5). Step 3 applies Eqn. (13). Additional performance measures, in particular the waiting time distribution, will be derived in the sections below. An appropriate termination criterion would be $\sum_{i=1}^{N} |\pi_i - \pi'_i| < \epsilon$ with a given precision $\epsilon > 0$ (cf. Tran-Gia and Raith 1988).

3.5 Waiting Time We now derive relations for the steady-state waiting time distribution function of a particular station. It will turn out that this function can be expressed in terms of the distribution functions of the interarrival time and the intervisit time, taking into account the assumptions from the previous section. To achieve this goal we first introduce the notion of *residual interarrival times*. In the following, we again omit the subscript i referring to a particular station in order to simplify the notation.

3.5.1 Residual Interarrival Times Consider an arbitrary l-polling-instant $T_G^{(r)} = T_G^{(n+l)}$ in steady-state for an arbitrarily chosen $l,\ l \geq 0$. Then the *residual interarrival time* $A_{res}^{(r)}$ is defined as the time interval from this l-polling-instant until the end of the current interarrival time. If a message arrives in a particular cycle, its waiting time corresponds exactly to the difference of the intervisit time and the residual interarrival time of that particular cycle. In addition to the assumptions from Sections 3.1 – 3.3 we assume that the limiting distribution $a_{res}(k) \triangleq \lim_{r\to\infty} a_{res}^{(r)}(k)$ exists; the corresponding random variable is denoted by A_{res}. Let $L^{(r)}$ be the r.v. for the number of the current polling instant within a cycle group (see Section 3.1). Define $P_{P,l} \triangleq \mathrm{P}\{L^{(r)} = l\}$ as the probability that, in steady-state, an arbitrary observer who randomly selects a polling instant $T_G^{(r)}$ sees the system at a l-polling-instant. Then, using the law of total probability, we can write

$$a_{res}^{(r)}(k) \quad = \quad \sum_{k=0}^{\infty} \mathrm{P}\{A_{res}^{(r)} = k \mid L^{(r)} = l\} \cdot P_{P,l}. \qquad (13)$$

Next, we derive the conditional probabilities from Eqn. (13). Consider an l-polling-instant $T_G^{(n+l)}$ in steady-state for a *fixed* $l \geq 0$. At this l-polling-instant the $(l+1)$st intervisit time $V^{(n+l)}$ after the beginning of the interarrival time $A^{(n)}$ (starting at instant $T_G^{(n)}$, the 0-polling-instant of the cycle group under consideration) is initiated. Thus,

$$\left(A_{res}^{(n+l)} \,\middle|\, L^{(n+l)} = l\right) \quad = \quad A^{(n)} \; - \; \sum_{j=0}^{l-1} V^{(n+j)}. \qquad (14)$$

We now introduce the random variable $\tilde{A}'_l$ defined by

$$\tilde{A}'_l \quad \triangleq \quad A \; - \; \sum_{j=0}^{l-1} V_j. \qquad (15)$$

Here, the random variables V_j are mutually independent and follow the same distribution as the random variable V. Note that $\tilde{A}'_l$ may have negative realizations for $l > 0$. Contrarily, $(A_{res}^{(n+l)} \mid L^{(n+l)} = l)$ assumes only non-negative values because instant $T_G^{(n+l)}$ is an l-polling-instant only if the corresponding cycle group comprises at least l idle cycles. Since we assume that successive intervisit times are mutually independent (see Section 3.3), in steady-state, $(A_{res}^{(n+l)} \mid L^{(n+l)} = l)$ has the same distribution as

$$\tilde{A}_l \quad \triangleq \quad (\tilde{A}'_l \mid \tilde{A}'_l > 0). \qquad (16)$$

According to Tran-Gia (1988) the distribution of the difference of two discrete random variables is given by the discrete cross-correlation of their distributions. Thus, the distributions of $\tilde{A}'_l$ and $\tilde{A}_l$ are as follows:

$$\tilde{a}'_l(k) \quad = \quad a(k) \circledast v^{[\circledast l]}(-k), \qquad (17)$$

$$\tilde{a}_l(k) \quad = \quad \tilde{a}'_l(k) \cdot \left[1 - \tilde{A}'_l(0)\right]^{-1}, \quad k > 0. \qquad (18)$$

Taking the limit for $r \to \infty$ the limit of the left hand side of Eqn. (13) results to

$$a_{res}(k) \quad = \quad \sum_{l=0}^{\infty} \tilde{a}_l(k) \cdot P_{P,l}. \qquad (19)$$

It remains to derive the probability $P_{P,l}$. For this purpose, consider an arbitrary cycle group. Since, by Eqn. (6), a cycle group is a j-cycle-group with probability $P_{B,j}$, the mean number of polling instants within an arbitrary cycle group is $\sum_{j=0}^{\infty} (j+1) \cdot P_{B,j}$. Note that a j-cycle-group comprises exactly $(j+1)$ polling instants (including the 0-polling-instant). The mean number of l-polling-instants within an arbitrary cycle group is $\sum_{j=l}^{\infty} P_{B,j}$, because every cycle group comprising at least l cycles contains a single l-polling-instant. Hence, in steady-state, instant $T_G^{(r)}$ randomly selected by an arbitrary observer is an l-polling-instant with probability

$$P_{P,l} \quad = \quad \sum_{j=l}^{\infty} P_{B,j} \cdot \Big[\sum_{j=0}^{\infty} (j+1) \cdot P_{B,j}\Big]^{-1}. \qquad (20)$$

3.5.2 Waiting Time Distribution Function Using the distribution of the residual interarrival time we are now able to calculate the waiting time distribution. If an arrival occurs during the intervisit time $V^{(m-1)}$ the random variable of the waiting time $W^{(m)}$ obeys the following equation:

$$W^{(m)} \quad = \quad V^{(m-1)} \; - \; A_{res}^{(m-1)}.$$

Taking the limit for $m \to \infty$ and assuming that the limiting distributions exist, we obtain the steady-state probability distribution function of the waiting time

$$W(k) \quad = \quad \mathrm{P}\{V - A_{res} \leq k \mid A_{res} \leq V\}, \qquad (21)$$

where condition $A_{res} \leq V$ is equivalent to the condition that an arrival occurs in an idle cycle. The waiting time is zero if the residual interarrival time equals the intervisit time because, by convention, arrival events are processed before polling events if both occur at the same discrete time instant. Using the law of total probability and conditioning Eqn. (21) by $V = j$, we obtain after some algebraic manipulations

$$W(k) = \frac{\sum_{j=0}^{\infty} \sum_{m=j-k}^{j} a_{res}(m) \cdot v(j)}{\sum_{j=0}^{\infty} A_{res}(j) \cdot v(j)}, \qquad k \geq 0.$$

Clearly, $W(k) = 0$ for $k < 0$. Note that the waiting time distribution $W(k)$ depends only on the distribution of the interarrival time (cf. Eqn. (17)) and the intervisit time (cf. Eqns. (17) – (18)). Furthermore, to calculate probability $P_{P,l}$ from Eqn. (20) we only need probability $P_{B,j}$. By Eqn. (7), the latter is a function of distributions $v(k)$ and $a(k)$. Thus, we have shown that, indeed, the waiting time distribution is solely a function of the intervisit and interarrival time distribution.

4 Numerical Results

To provide some insight into the approximation accuracy, we apply the proposed method to several test cases. We consider a symmetric system of $N = 8$ stations. Because of the symmetry, we omit the station index from the system parameters and performance measures. The switchover time is assumed to be deterministic and normalized to one unit of time, that is $u = 1$ and $c_U = 0$. We focus on the impact of variations in the second moments (given through the coefficients of variation c_B and c_A) of the service and interarrival times on the system behavior. The mean service time is kept fixed at $b = 10$, while we show performance measures for the mean interarrival time $a \in \{40, 80, 120\}$.

The input distributions were derived from their given first and second moments according to the moment matching method developed by Kühn (1979) and subsequent discretization according to Tran-Gia (1988). In the tables below we also provide simulation results, along with 99% student-t confidence intervals for the means, and relative errors (r.e.).

Table 1 depicts the mean waiting time w for service time CoVs of $c_B = 0.2$, 1, 2 and interarrival time CoVs of $c_A = 0.2$, 2. The results exhibit that, given a fixed mean interarrival time, the larger c_A and c_B the longer the mean waiting times. A comparison of the results for $c_A = 0.2$ and $c_A = 2$ shows that w is mainly determined by c_A whereas c_B has less impact. The relative errors of the mean waiting time are small in most cases. They significantly exceed the 10 % margin only for large interarrival time CoV.

Table 2 shows the waiting time CoV for configurations with $c_A = 0.2$, 1, 2 and $c_B = 0.2$, 2. Again, the approximation is fairly accurate. Observe that, for a given interarrival mean, the waiting time CoV is much more sensitive to changes in c_A in the case of $c_B = 0.2$ than for $c_B = 2$. This type of multiple dependency between parameters and performance measures complicates the prediction of their relationship. This, in turn, underpins the need for an efficient approximation procedure.

Note that our method significantly underestimates the CoV of the cycle time. This is due to the assumption of independent station times, which does not take into account any positive correlation between station times. The reason for this correlation is that longer station times at a station i imply a higher probability for an arrival at a station j in the current cycle, and thus a longer station time at j.

5 Conclusion

The comparison of numerical examples and simulation results showed the validity of our approximation. The method is well suited to its application on large symmetrical systems since, in this case, the major equation that has to be solved reduces to an N-fold convolution. The only disadvantage in applying it to entirely asymmetric systems is the increase in computational complexity which mainly depends on the length of the distributions involved. However, it is also influenced by their detailed characteristics, which are reflected in the number of iterations necessary to satisfy a given termination criterion. For example, an increase in the CoVs of the service or the interarrival time tends to increase the required numerical effort. For our case studies, the required CPU times on a Sparc20 workstation generally ranged from several seconds to around 15 minutes. For a few extreme parameter settings however, longer running times of up to four hours were observed, in which case simulation becomes the more economical approach (simulation times ranged between 30 and 50 minutes in comparison). Hence, a direction for future research is to obtain performance measure bounds that are simple to evaluate.

Acknowledgement The authors would like to thank Prof. Dr. P. Tran-Gia for the support of this work.

References

Chung, H., C. K. Un, and W. Y. Jung (1994). Performance analysis of Markovian polling systems with single buffers. *Performance Evaluation 19*, 303–315.

Cox, D. R. (1955). The analysis of non-markovian stochastic processes by the inclusion of supplementary variables. *Proceedings of the Cambridge Philosophical Society (Mathematics and Physical Sciences) 51*, 433–441.

Hashida, O. and K. Kawashima (1981). Analysis of a polling system with single user at each terminal. *Review of the Electrical Communication Laboratories 29*(3-4), 245–253.

Haverkort, B. R. (1994). Polling models: Theory and applications. In R. Marie, G. Haring, and G. Kotsis (Eds.), *Performance and Reliability Evaluation*, pp. 237–266. Wien: R. Oldenbourg.

Ibe, O. C. and X. Cheng (1989). Performance analysis of asymmetric single-buffer polling systems. *Performance Evaluation 10*, 1–14.

Kühn, P. J. (1979). Approximate analysis of general queueing networks by decomposition. *IEEE Transactions on Communications COM-27*, 113–126.

Mack, C. (1957). The efficiency of n machines unidirectionally patrolled by one operative when walking time is constant and repair times are variable. *Journal of the Royal Statistical Society,* Series B *19*(1), 173–178.

Mack, C., T. Murphy, and N. L. Webb (1957). The efficiency of n machines uni-directionally patrolled by one operative when walking time and repair times are constant. *Journal of the Royal Statistical Society,* Series B *19*(1), 166–172.

Stecke, K. E. and J. E. Aronson (1985). Review of operator/machine interference models. *International Journal of Production Research 23*, 129–151.

Takagi, H. (1985). On the analysis of a symmetric polling system with single-message buffers. *Performance Evaluation 5*, 149–157.

Takagi, H. (1990). Queueing analysis of polling models: an update. In H. Takagi (Ed.), *Stochastic Analysis of Computer and Communication Systems*, pp. 267–318. North-Holland, Amsterdam, The Netherlands.

Takagi, H. (1991). Application of polling models to computer networks. *Computer Networks and ISDN Systems 22*, 193–211.

Takine, T., H. Takagi, Y. Takahashi, and T. Hasegawa (1990). Analysis of asymmetric single-buffer polling and priority systems without switchover times. *Performance Evaluation 11*, 253–264.

Takine, T., Y. Takahashi, and T. Hasegawa (1988). Exact analysis of asymmetric polling systems with single buffers. *IEEE Transactions on Communications 36*(10), 1119–1127.

Tran-Gia, P. (1986). Discrete-time analysis for the interdeparture distribution of GI/G/1 queues. In O. J. Boxma, J. W. Cohen, and H. C. Tijms (Eds.), *Teletraffic Analysis and Computer Performance Evaluation*, pp. 341–357. Elsevier (North-Holland).

Tran-Gia, P. (1988). Zeitdiskrete Analyse in verkehrstheoretischer Modelle in Rechner- und Kommunikationssystemen. Bericht 46, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart. In German.

Tran-Gia, P. (1992). Analysis of polling systems with general input process and finite capacity. *IEEE Transactions on Communications 40*(2), 337–344.

Tran-Gia, P. and T. Raith (1988). Performance analysis of finite capacity polling systems with nonexhaustive service. *Performance Evaluation 9*, 1–16.

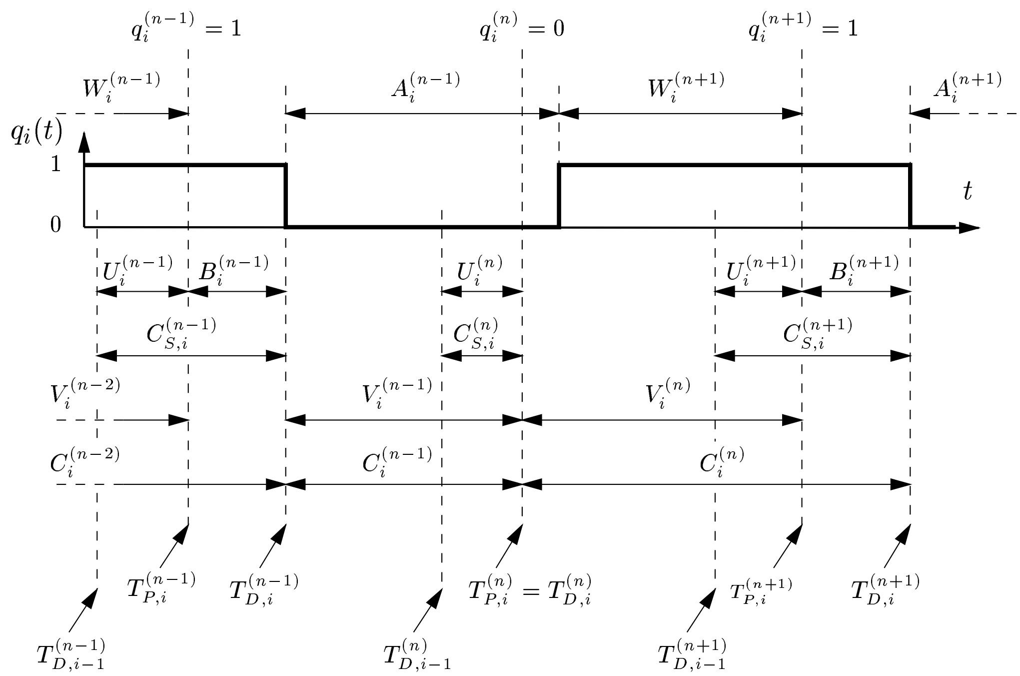

Figure 1: *Sample polling process and corresponding random variables*

Table 1: *Mean waiting time*

		$c_B = 0.2$			$c_B = 1$			$c_B = 2$		
	a	w	w_{sim}	r.e.	w	w_{sim}	r.e.	w	w_{sim}	r.e.
$c_A = 0.2$	40	37.8	38.0±0.04	-0.6%	38.8	40.0±0.09	-2.8%	42.3	43.6±0.20	-2.9%
	80	17.1	15.1±0.03	12.8%	19.5	21.7±0.08	-10.1%	26.1	30.4±0.21	-13.9%
	120	10.9	10.5±0.04	3.9%	13.2	14.9±0.08	-11.5%	19.4	23.6±0.15	-17.5%
$c_A = 2$	40	41.0	41.4±0.08	-1.0%	41.7	42.5±0.16	-1.7%	43.0	44.3±0.23	-2.8%
	80	20.8	20.6±0.10	0.9%	22.8	23.5±0.09	-3.0%	26.6	27.7±0.24	-3.8%
	120	14.1	13.2±0.07	7.5%	16.4	15.9±0.08	3.4%	20.9	20.4±0.18	2.5%

Table 2: *Waiting time coefficient of variation*

		$c_A = 0.2$			$c_A = 1$			$c_A = 2$		
	a	c	c_{sim}	r.e.	c	c_{sim}	r.e.	c	c_{sim}	r.e.
$c_B = 0.2$	40	0.26	0.25	1.2%	0.48	0.50	-3.5%	0.45	0.47	-4.9%
	80	0.72	0.74	-2.3%	0.70	0.77	-9.3%	0.65	0.76	-13.4%
	120	0.80	0.86	-6.9%	0.79	0.88	-10.4%	0.74	0.88	-15.1%
$c_B = 2$	40	1.11	1.18	-6.3%	1.07	1.14	-6.7%	1.05	1.11	-5.7%
	80	1.40	1.51	-7.4%	1.41	1.50	-6.0%	1.35	1.48	-8.6%
	120	1.62	1.74	-6.7%	1.65	1.74	-5.1%	1.53	1.71	-10.9%