No. 559

Analysis of a Discrete-Time Queueing System with Time-Limited Service

Hideaki Takagi (Tsukuba University)
and
Kin K. Leung (AT&T Bell Laboratories)

November 30, 1993

Analysis of a Discrete-Time Queueing System with Time-Limited Service

Hideaki Takagi
Institute of Socio-Economic Planning, University of Tsukuba
1-1-1 Tennoudai, Tsukuba-shi, Ibaraki 305, Japan
and
Kin K. Leung
AT&T Bell Laboratories
Crawfords Corner Road, Holmdel, New Jersey 07733, U. S. A.

Abstract

We analyze a discrete-time, single-server queueing system in which the length of each service period is limited. The server takes a vacation when the limit expires or the queue empties, whichever occurs first. In the former case, the preempted service is resumed after the vacation without loss or creation of any work. This system models the transmission of message frames from a station on timed-token local-area networks (for example, FDDI and IEEE 802.4 token bus). We study a process of the unfinished work and a joint process of the queue size and the remaining service time. By using a technique of discrete Fourier transforms to determine some unknown functions in the governing equations, we numerically obtain exact mean waiting times.

Keywords: Queue, discrete-time system, server vacation model, time-limited service, token passing protocol, timed-token protocol, local area network

1. Introduction

Timed-token passing protocols are the medium access control (MAC) protocols used in many local area networks such as the Fiber Distributed Data Interface (FDDI) network [1] and the IEEE 802.4 token bus [2]. In token passing protocols, an explicit or implicit token is circulated among the stations in a network so that only the station in possession of the token is allowed to transmit. In timed-token protocols, the time at which a station can continue to transmit is limited according to a certain rule that depends on the congestion of the network as well as the priority of messages under transmission. Performance aspects of timed-token protocols have been studied by a number of researchers, for example, [14, 15, 20]. The performance of token passing protocols has often been studied by using a queue with server vacations and a polling system, which is a set of queues attended by a single server in cyclic order [4, 19]. If we focus on one station of the network and treat the time period in which other stations are transmitting as a "vacation" of the server for that station, we get a queue with server vacations, while the whole network can be modeled by a polling system. With these and other applications in mind, as well as from the theoretical interest on the underlying stochastic processes, many contributions have been made to the study of vacations systems [6, 18] and polling systems [16, 17]. It should be noted, however, that systems with time-limited service are dealt with only by a few of them [5, 7, 9–12]. For example, Leung [9] considers a continuous-time system with generally distributed service times and exponentially distributed time limits. Leung and Lucantoni [12] also study a continuous-time system with generally distributed service times and a constant time limit which is approximated by exponential time stages. To the best of the authors' knowledge, however, no exact results for the mean message waiting time have been published for discrete-time systems with generally distributed service times and time-limited service.

In this paper, we consider a discrete-time queueing system with time-limited service, and present a procedure for calculating the exact mean waiting time for the first time. By the benefit of discrete-time setting, we can reduce the state-transition equations to a rather simple set of equations for a finite number of unknown functions, which are determined by a technique of discrete Fourier transforms.

The rest of this paper is organized as follows. In Section 2, we describe the queueing system analyzed in the paper by introducing notation and explaining the service mechanism. In Section 3, we show two methods for calculating the mean unfinished work in the system at an arbitrary time. If the service time is geometrically distributed, we can also obtain the mean response time (waiting time plus service time) from the mean unfinished work. The joint process of the queue size and the remaining service time during a service period is considered in Section 4, and the joint process of the queue size and the elapsed vacation time during a vacation is analyzed in Section 5. In Section 6, we give an expression for the mean response time in terms of the queue size distribution. In Section 7, we present an iterative procedure for solving unknown functions that appear in Sections 4 and 5. Combining these results, we can compute the exact mean response time. In Section 8, we demonstrate the feasibility of our approach by showing two sets of numerical examples. When the service time is geometrically distributed, it is shown that the mean response times obtained from the analysis of the unfinished work agree with those from the analysis of the joint process of the queue size and the remaining service time. We conclude the paper in Section 9 by summarizing the results and suggesting future research subjects.

2. Model

The parameters and the operations of a queueing system we analyze in this paper are described as follows. The system has a single server, an infinite queueing capacity, and an infinite population of customers. The time axis of the system is segmented into a sequence of equal intervals of unit duration called *slots*. Services and vacations of the server can be started only at slot boundaries, and their durations are always integral multiples of a slot duration.

The number Λ of customers that arrive in each slot is an independent and identically distributed random variable. The probability distribution of Λ is specified by

$$\lambda(k) := \text{Prob}[\Lambda = k] \qquad k = 0,1,2,\cdots \qquad ; \qquad A(z) := \sum_{k=0}^{\infty} \lambda(k) z^k \tag{2.1}$$

We call $A(z)$ the probability generating function (PGF) for Λ. We denote by λ and $\lambda^{(2)}$ the mean and the second factorial moment, respectively, of Λ. That is,

$$\lambda := E[\Lambda] = A^{(1)}(1) \qquad ; \qquad \lambda^{(2)} := E[\Lambda(\Lambda-1)] = A^{(2)}(1) \tag{2.2}$$

where $A^{(i)}(1) := d^i A(z)/dz^i|_{z=1}$ for $i = 1,2,\cdots$.

Let us denote by X the service time (measured in slots) of each customer. Each service is started and completed at exact slot boundaries. The probability distribution of X is specified by

$$b(l) := \text{Prob}[X = l] \qquad l = 1,2,\cdots \qquad ; \qquad B(u) := \sum_{l=1}^{\infty} b(l) u^l \tag{2.3}$$

The mean and the second moment of the service time are denoted by b and $b^{(2)}$, respectively:

$$b := E[X] = B^{(1)}(1) \qquad ; \qquad b^{(2)} := E[X^2] = B^{(2)}(1) + B^{(1)}(1) \tag{2.4}$$

where $B^{(i)}(1) := d^i B(u)/du^i|_{u=1}$ for $i = 1, 2, \cdots$. The traffic intensity ρ is assumed to be less than unity for stability:

$$\rho = \lambda b < 1 \tag{2.5}$$

Since the queueing process depends on the exact arrival points during a slot and the slot in which the service to a customer that arrives when the system is empty is started, we consider a *late arrival model.* In this model, customers arrive late during a slot, just prior to the end of the slot. Therefore, arriving customers see a departing customer, if any, about to leave, and the departing customer leaves behind the customers that have just arrived.

We study a time-limited service system with multiple vacations, in which the length (measured in slots) of each service period is limited by M slots. That is, the server takes a vacation when the limit expires or the queue empties, whichever occurs first. The parameter M is called the *maximum server attendance time* by Leung and Eisenberg [10, 11], who first analyzed continuous-time time-limited service systems. There are a *gated time-limited service* system in which only those customers present in the system at the beginning of a service period can be served during the service period, and an *exhaustive time-limited service* system in which those customers that arrive during a service period can also be served in the same service period. In this paper, we restrict ourselves to the study of a discrete-time system with exhaustive time-limited service. We assume that the service preempted when the maximum server attendance time expires is resumed in the next service period without loss or creation of work.

In the multiple vacation model, if the server finds the system not empty upon returning from a vacation, it starts to work immediately and continues to work until the limit expires or the system becomes empty again. Otherwise, the server begins another vacation immediately, and repeats vacations in this manner until it finds at least one waiting customer upon returning from a vacation. The length V of each vacation (measured in slots) is assumed to be an integral multiple of a slot duration, and an independent and identically distributed random variable. The PGF for V is denoted by $V(u)$. Thus we have

$$v(l) := \text{Prob}[V = l] \qquad l = 1, 2, \cdots \qquad ; \qquad V(u) := \sum_{l=1}^{\infty} v(l) u^l \tag{2.6}$$

3. Unfinished Work

Let $U^{(0)}(u)$ denote the PGF for the unfinished work $U^{(0)}$ at the end of a vacation (i.e., at the beginning of a service period). From the decomposition property for the unfinished work shown by Boxma and Groenendijk [3], the PGF $U(u)$ for the unfinished work immediately after an arbitrary slot boundary is given by

$$U(u) = U_{\text{Geo}^X/\text{G}/1}(u) \bullet \frac{U^{(0)}(u)}{V\{A[B(u)]\}} \bullet \frac{1 - V\{A[B(u)]\}}{E[V]\{1 - A[B(u)]\}} \tag{3.1}$$

where $U_{\text{Geo}^X/\text{G}/1}(u)$ is the PGF for the unfinished work immediately after an arbitrary slot boundary in the corresponding $\text{Geo}^X/\text{G}/1$ system without vacations, given by

$$U_{\text{Geo}^X/\text{G}/1}(u) = \frac{(1-\rho)(1-u)A[B(u)]}{A[B(u)] - u} \tag{3.2}$$

The second factor on the r.h.s. of (3.1) is the PGF for the unfinished work at the end of a service period. The third factor is the PGF for the work brought to the system by the customers that arrive during the backward recurrence time of a vacation. From (3.1), we have

$$E[U] = \frac{\lambda b^{(2)} + \lambda^{(2)} b^2 + \rho(1-2\rho)}{2(1-\rho)} + E[U^{(0)}] - \rho E[V] + \frac{\rho E[V(V-1)]}{2E[V]} \tag{3.3}$$

Hence, if $E[U^{(0)}]$ is known, we can obtain the mean unfinished work $E[U]$ immediately after an arbitrary slot boundary.

In order to find the distribution of $U^{(0)}$, we consider a modified process of the unfinished work such that, starting with $U^{(0)}$ at the beginning of the first slot, once the system becomes empty it remains to be empty afterwards. Thus, the modified unfinished work immediately after the Mth slot is the unfinished work at the end of a service period in the original system. Let Θ_c be the length (measured in slots) of the busy period started with the unfinished work whose PGF is given by $U^{(0)}(u)$. The PGF $\Theta_c(u)$ for Θ_c is given by

$$\Theta_c(u) = U^{(0)}\{uA[\Theta(u)]\} \tag{3.4}$$

where $\Theta(u)$ is the PGF for the length (measured in slots) of a busy period started with a single customer in a standard Geo^X/G/1 system without vacations. It is given as a solution to the equation

$$\Theta(u) = B\{uA[\Theta(u)]\} \tag{3.5}$$

We define $U_l^{(n)}$ as the probability that the modified unfinished work is l slots immediately after the nth slot, where $l \geq 0$ and $n = 1, 2, \cdots$. Let $U_l^{(0)}$ be the probability that the unfinished work $U^{(0)}$ at the beginning of a service period is l slots. We introduce the generating function $U^{(n)}(u)$ for $\{U_l^{(n)}; l \geq 0\}$ by

$$U^{(n)}(u) := \sum_{l=0}^{\infty} U_l^{(n)} u^l \qquad n = 0, 1, 2, \cdots \tag{3.6}$$

and the w-transform $U(u; w)$ for $\{U^{(n)}(u); n = 0, 1, 2, \cdots\}$ by

$$U(u; w) := \sum_{n=0}^{\infty} U^{(n)}(u) w^n \tag{3.7}$$

Our treatment of the unfinished work follows the approach taken by Leung [8] for the queue size in a continuous-time system. We start with the relation

$$U^{(n+1)}(u) = \frac{A[B(u)]}{u}\left[U^{(n)}(u) - U^{(n)}(0)\right] + U^{(n)}(0) \qquad n = 0, 1, 2, \cdots \tag{3.8}$$

which is converted into

$$\frac{U(u; w) - U^{(0)}(u)}{w} = \frac{A[B(u)]}{u}[U(u; w) - U(0; w)] + U(0; w) \tag{3.9}$$

Solving (3.9) for $U(u; w)$, we get

$$U(u; w) = \frac{uU^{(0)}(u) + \{u - A[B(u)]\} w U(0; w)}{u - wA[B(u)]} \tag{3.10}$$

From the condition that this function is analytic at $u = wA[\Theta(w)]$ where the denominator on the r.h.s. is null, we determine that

$$U(0; w) = \frac{U^{(0)}\{wA[\Theta(w)]\}}{1 - w} \tag{3.11}$$

Substituting (3.11) into (3.10), we get

$$\sum_{n=0}^{\infty} U^{(n)}(u)w^n = \frac{uU^{(0)}(u)}{u - wA[B(u)]} + \frac{\{u - A[B(u)]\}wU^{(0)}\{wA[\Theta(w)]\}}{(1-w)\{u - wA[B(u)]\}} \quad (3.12)$$

In order to invert the w-transform in (3.12), we expand its r.h.s. in powers of w as

$$\begin{aligned}\sum_{n=0}^{\infty} U^{(n)}(u) \quad w^n &= U^{(0)}(u)\sum_{n=0}^{\infty}\left[\frac{A[B(u)]}{u}\right]^n w^n \\ &+ w\left[1 - \frac{A[B(u)]}{u}\right]\left[\sum_{n=0}^{\infty}\left[\frac{A[B(u)]}{u}\right]^n w^n\right]\left[\sum_{n=0}^{\infty}\overline{U}^{(0)}(n)w^n\right]\left[\sum_{n=0}^{\infty} w^n\right]\end{aligned} \quad (3.13)$$

where

$$\overline{U}^{(0)}(n) := \frac{1}{n}\frac{d^n}{dw^n}U^{(0)}\{wA[\Theta(w)]\}|_{w=0} \qquad n = 0, 1, 2, \cdots \quad (3.14)$$

Therefore, the transform inversion of (3.12) is given by

$$U^{(n)}(u) = U^{(0)}(u)\left[\frac{A[B(u)]}{u}\right]^n + \left[1 - \frac{A[B(u)]}{u}\right]\sum_{k=0}^{n-1}\overline{U}^{(0)}(k)\sum_{m=0}^{n-k-1}\left[\frac{A[B(u)]}{u}\right]^m$$

$$n = 0, 1, 2, \cdots \quad (3.15)$$

Note that $U^{(M)}(u)$ for the modified unfinished work corresponds to the PGF for the unfinished work at the end of a service period in the original system, which is given by

$$U^{(M)}(u) = U^{(0)}(u)\left[\frac{A[B(u)]}{u}\right]^M + \left[1 - \frac{A[B(u)]}{u}\right]\sum_{k=0}^{M-1}\overline{U}^{(0)}(k)\sum_{m=0}^{M-k-1}\left[\frac{A[B(u)]}{u}\right]^m \quad (3.16)$$

We then have the following equation for $U^{(0)}(u)$:

$$U^{(0)}(u) = U^{(M)}(u)V\{A[B(u)]\} \quad (3.17)$$

It is clear that $U^{(M)}(u)$ in (3.17) can be obtained from (3.16), which is in turn a function of $U^{(0)}(u)$. Alternatively, $U^{(M)}(u)$ is given by (3.8) as a recursive function of $U^{(0)}(u)$. As a result, $U^{(0)}(u)$ can be solved numerically. From our numerical experience, the latter approach of using (3.8) to find $U^{(0)}(u)$ is more efficient, because $\overline{U}^{(0)}(n)$ in (3.14) need not be computed. Once $U^{(0)}(u)$ is found, $E[U^{(0)}]$ as well as $E[U]$ in (3.3) becomes known.

The mean unfinished work $E[U]$ immediately after an arbitrary slot boundary can also be expressed as

$$E[U] = E[L_q]b + \rho\left[\frac{b^{(2)}}{2b} + \frac{1}{2}\right] + P_e E[X_e] \quad (3.18)$$

where $E[L_q]$ is the mean number of customers waiting in the queue, P_e is the probability that the service of a customer is being interrupted by a vacation, and $E[X_e]$ is the mean remaining service time of the interrupted customer, all being observed immediately after an arbitrary slot boundary. If $E[L]$ denotes the mean number of customers in the system immediately after an arbitrary slot boundary, we have

$$E[L] = E[L_q] + \rho + P_e \quad (3.19)$$

The probability P_e that the service of a customer is being interrupted by a vacation immediately after an arbitrary slot boundary is given by

$$P_e = (1-\rho)[1-U^{(M)}(0)] \tag{3.20}$$

If we know $E[L]$, the mean response time $E[T]$ of a customer is given from Little's theorem

$$E[L] = \lambda E[T] \tag{3.21}$$

For generally distributed service times, it is difficult to obtain $E[T]$ from the unfinished work because of the unknown P_e and $E[X_e]$. However, in a special case in which the service time is geometrically distributed, $E[X_e] = b$ due to the memoryless property of geometric distributions. As a result, $E[T]$ can be obtained from $E[U]$ as follows. In this case, we have

$$B(u) = \frac{(1-\beta)u}{1-\beta u} \qquad 0 < \beta < 1 \tag{3.22}$$

which gives

$$\frac{b^{(2)}}{2b} + \frac{1}{2} = b \tag{3.23}$$

Substituting (3.23) and $E[X_e] = b$ into (3.18), we have

$$E[U] = E[L_q]b + \rho b + P_e b = E[L]b \tag{3.24}$$

From (3.21) and (3.24), we get

$$E[T] = \frac{E[U]}{\rho} \tag{3.25}$$

From (3.3) and (3.25), we can express $E[T]$ explicitly in terms of $E[U^{(0)}]$ which has been computed earlier.

4. Joint Distribution of the Queue Size and the Remaining Service Time

We next consider the joint process of the queue size and the remaining service time during a service period for the exhaustive time-limited service system. Let $L^{(n)}$ and $X_+^{(n)}$ be the queue size and the remaining service time, respectively, immediately after the nth slot in a service period, where $n = 0, 1, 2, \cdots, M$. Note that $L^{(0)}$ and $X_+^{(0)}$ are the queue size and the remaining service time, respectively, at the beginning of the service period, and M is the maximum length (in slots) of the service period. Denoting by ξ the slot number since the beginning of a service period, we define

$$\Pi_k^{(n)}(l) := \text{Prob}[\xi = n, L^{(n)} = k, X_+^{(n)} = l] \qquad k \geq 1, l \geq 1, n = 0, 1, 2, \cdots, M \tag{4.1}$$

The probability that the system is empty at the beginning of a service period (thus another vacation is started immediately) is defined by

$$P_0^{(0)} := \text{Prob}[L^{(0)} = 0] \tag{4.2}$$

We then get the following set of equations for $\{\Pi_k^{(n)}(l); k \geq 1, l \geq 1, 0 \leq n \leq M\}$ and $P_0^{(0)}$:

$$\Pi_k^{(n+1)}(l) = \sum_{j=1}^{k} \Pi_j^{(n)}(l+1)\lambda(k-j) + \sum_{j=1}^{k+1} \Pi_j^{(n)}(1)\lambda(k-j+1)b(l)$$

$$k \geq 1, \quad l \geq 1, \quad n = 0, 1, 2, \cdots, M-1 \tag{4.3}$$

$$\Pi_k^{(0)}(l) = \left[\sum_{n=0}^{M-1} \Pi_1^{(n)}(1)\lambda(0) + P_0^{(0)}\right] f_k b(l) + \sum_{j=1}^{k} \Pi_j^{(M)}(l) f_{k-j} \qquad k \geq 1, l \geq 1 \tag{4.4}$$

$$P_0^{(0)} = \left[\sum_{n=0}^{M-1} \Pi_1^{(n)}(1)\lambda(0) + P_0^{(0)}\right] f_0 \tag{4.5}$$

where $b(l)$ is the probability that the service time of a customer is l slots, where $l = 1, 2, \cdots$, and f_k is the probability that k customers arrive during a vacation, where $k = 0, 1, 2, \cdots$. Note that the first term on the r.h.s. of (4.4) represents a case in which a service period ends in the $n+1$st slot leaving the system empty, where $n = 0, 1, \cdots, M-1$, and a case in which the system is empty at the beginning of a service period and the next vacation is therefore started immediately. The second term represents a case in which a service period is terminated in the Mth slot leaving j customers and l slots of the remaining service time in the system.

Two relationships can be derived from (4.3)-(4.5). From (4.3), we get

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \Pi_k^{(n+1)}(l) = \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \Pi_k^{(n)}(l) - \Pi_1^{(n)}(1)\lambda(0) \qquad n = 0, 1, 2, \cdots, M-1 \tag{4.6}$$

which is the probability that the system is immediately after the $n+1$st slot in a service period for $n = 0, 1, 2, \cdots, M-1$. From (4.4) and (4.5), we get

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \Pi_k^{(0)}(l) = \sum_{n=0}^{M-1} \Pi_1^{(n)}(1)\lambda(0) + \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \Pi_k^{(M)}(l) \tag{4.7}$$

which equates the probability of the system being at the beginning of a service period to the probability that the system is at the end of a service period immediately after an arbitrary slot boundary.

We introduce the generating function $\overline{\Pi}_l^{(n)}(z)$ for $\{\Pi_k^{(n)}(l); k \geq 1\}$ by

$$\overline{\Pi}_l^{(n)}(z) := \sum_{k=1}^{\infty} \Pi_k^{(n)}(l) z^k \qquad l \geq 1, \quad n = 0, 1, 2, \cdots, M \tag{4.8}$$

and the generating function $\Pi^{(n)}(z, u)$ for $\{\overline{\Pi}_l^{(n)}(z); l \geq 1\}$ by

$$\Pi^{(n)}(z, u) := \sum_{l=1}^{\infty} \overline{\Pi}_l^{(n)}(z) u^l \qquad n = 0, 1, 2, \cdots, M \tag{4.9}$$

In terms of $\{\overline{\Pi}_l^{(n)}(z); l \geq 1\}$, (4.3) can be written as

$$\overline{\Pi}_l^{(n+1)}(z) = A(z)\overline{\Pi}_{l+1}^{(n)}(z) + \frac{A(z)}{z}\overline{\Pi}_1^{(n)}(z) b(l) - \Pi_1^{(n)}(1)\lambda(0) b(l)$$

$$l \geq 1, \quad n = 0, 1, 2, \cdots, M-1 \tag{4.10}$$

Taking the u-transform of (4.10) with respect to l yields

$$\Pi^{(n+1)}(z, u) = \frac{A(z)}{u}\Pi^{(n)}(z, u) + A(z)\left[\frac{B(u)}{z} - 1\right]\overline{\Pi}_1^{(n)}(z) - \lambda(0) B(u) \Pi_1^{(n)}(1)$$

$$n = 0, 1, 2, \cdots, M-1 \tag{4.11}$$

From (4.4) and (4.5), we get

$$\overline{\Pi}_l^{(0)}(z) = P_0^{(0)} \frac{V[A(z)] - V[\lambda(0)]}{V[\lambda(0)]} b(l) + \overline{\Pi}_l^{(M)}(z) V[A(z)] \tag{4.12}$$

where $f_0 = V[\lambda(0)]$ has been used. We then obtain

$$\Pi^{(M)}(z,u) = \frac{1}{V[A(z)]} \left[\Pi^{(0)}(z,u) - P_0^{(0)} \frac{V[A(z)] - V[\lambda(0)]}{V[\lambda(0)]} B(u) \right] \tag{4.13}$$

Further introducing

$$\Pi(z,u;w) := \sum_{n=0}^{M} \Pi^{(n)}(z,u) w^n \tag{4.14}$$

we can write (4.11) as

$$\left[1 - \frac{A(z)w}{u}\right] \Pi(z,u;w) = \Pi^{(0)}(z,u) - \frac{A(z)}{u} \Pi^{(M)}(z,u) w^{M+1}$$

$$+A(z) \left[\frac{B(u)}{z} - 1\right] w \sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(z) w^n - \lambda(0) B(u) w \sum_{n=0}^{M-1} \Pi_1^{(n)}(1) w^n \tag{4.15}$$

Substituting (4.13) for $\Pi^{(M)}(z,u)$ into (4.15) yields

$$\left[1 - \frac{A(z)w}{u}\right] \Pi(z,u;w) = \left[1 - \frac{A(z)w^{M+1}}{V[A(z)]u}\right] \Pi^{(0)}(z,u)$$

$$+P_0^{(0)} \frac{A(z)\{V[A(z)] - V[\lambda(0)]\} B(u) w^{M+1}}{V[\lambda(0)] V[A(z)] u}$$

$$+A(z) \left[\frac{B(u)}{z} - 1\right] w \sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(z) w^n - \lambda(0) B(u) w \sum_{n=0}^{M-1} \Pi_1^{(n)}(1) w^n \tag{4.16}$$

By setting $u = wA(z)$ in (4.16), we determine $\Sigma_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(z) w^n$ as

$$\left[1 - \frac{B[wA(z)]}{z}\right] wA(z) \sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(z) w^n$$

$$= \left[1 - \frac{w^M}{V[A(z)]}\right] \Pi^{(0)}(z, wA(z)) + P_0^{(0)} \frac{\{V[A(z)] - V[\lambda(0)]\} B[wA(z)] w^M}{V[\lambda(0)] V[A(z)]}$$

$$-\lambda(0) B[wA(z)] w \sum_{n=0}^{M-1} \Pi_1^{(n)}(1) w^n \tag{4.17}$$

The l.h.s. of (4.17) is null when $z = \Theta(w)$, because $\Theta(w)$ satisfies the equation in (3.5). It follows that

$$\lambda(0)\Theta(w) w \sum_{n=0}^{M-1} \Pi_1^{(n)}(1) w^n = \left[1 - \frac{w^M}{V\{A[\Theta(w)]\}}\right] \Pi^{(0)}(\Theta(w), wA[\Theta(w)])$$

$$+P_0^{(0)} \frac{(V\{A[\Theta(w)]\} - V[\lambda(0)])\,\Theta(w) w^M}{V[\lambda(0)]V\{A[\Theta(w)]\}} \tag{4.18}$$

where (3.5) has been used. Note that we get (4.5) if $w = 1$ in (4.18). Substituting $P_0^{(0)}$ from (4.5) into (4.18) yields the following equation with respect to w :

$$\lambda(0) w \sum_{n=0}^{M-1} \Pi_1^{(n)}(1) w^n = \left[1 - \frac{w^M}{V\{A[\Theta(w)]\}}\right] \frac{\Pi^{(0)}(\Theta(w), wA[\Theta(w)])}{\Theta(w)}$$

$$+ \frac{\lambda(0)\,(V\{A[\Theta(w)]\} - V[\lambda(0)])\, w^M \sum_{n=0}^{M-1} \Pi_1^{(n)}(1)}{\{1 - V[\lambda(0)]\} V\{A[\Theta(w)]\}} \tag{4.19}$$

which contains a set of unknown constants $\{\Pi_1^{(n)}(1); 0 \leq n \leq M-1\}$ and an unknown function $\Pi^{(0)}(z,u)$, the PGF for the joint queue size and the remaining service time at the beginning of a service period. Once they are known, we can obtain $P_0^{(0)}$ from (4.5), and $\Pi^{(n)}(z,u); 1 \leq n \leq M$ from (4.11) recursively. We will show how they can be obtained numerically in Section 7.

5. Joint Distribution of the Queue Size and the Elapsed Vacation Time

In order to obtain the queue size distribution immediately after an arbitrary slot boundary, we also need to know the system state during a vacation by keeping track of a joint process for the queue size, the remaining service time, and the elapsed vacation time at each slot boundary.

Let $\xi = (n, m)$ be the state of the system during a vacation such that the system is in the nth slot since the beginning of a vacation which is m slots long, where $n = 0, 1, 2, \cdots, m$ and $m = 1, 2, \cdots$. We define

$$\Omega_0^{(n,m)} := \text{Prob}[\xi = (n,m), L^{(n)} = 0] \qquad 0 \leq n \leq m, \quad m \geq 1 \tag{5.1}$$

$$\Omega_k^{(n,m)}(l) := \text{Prob}[\xi = (n,m), L^{(n)} = k, X_+^{(n)} = l]$$
$$k \geq 1, \quad l \geq 1, \quad 0 \leq n \leq m, \quad m \geq 1 \tag{5.2}$$

Note that $\Omega_0^{(0,m)}$ is the probability that the system is empty at the beginning of a vacation m slots long. Therefore, we have the relation

$$\Omega_0^{(0,m)} = \left[\sum_{n=0}^{M-1} \Pi_1^{(n)}(1)\lambda(0) + P_0^{(0)}\right] v(m) \qquad m \geq 1 \tag{5.3}$$

where $v(m)$ is the probability that the length of a vacation is m slots for $m = 1, 2, \cdots$. Since $\Omega_k^{(0,m)}(l)$ is the probability that there are k customers and l slots of the remaining service time in the system at the beginning of a vacation, we have the relation

$$\Omega_k^{(0,m)}(l) = \Pi_k^{(M)}(l) v(m) \qquad k \geq 1, \quad l \geq 1, \quad m \geq 1 \tag{5.4}$$

During a vacation, we have

$$\Omega_0^{(n+1,m)} = \Omega_0^{(n,m)} \lambda(0) \qquad 0 \leq n \leq m-1, \quad m \geq 1 \tag{5.5}$$

$$\Omega_k^{(n+1,m)}(l) = \sum_{j=1}^{k} \Omega_j^{(n,m)}(l)\lambda(k-j) + \Omega_0^{(n,m)}\lambda(k)b(l)$$

$$k \geq 1, \quad l \geq 1, \quad 0 \leq n \leq m-1, \quad m \geq 1 \tag{5.6}$$

For more concise expressions, let us introduce

$$\Omega^{(n,m)}(z,u) := \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \Omega_k^{(n,m)}(l) z^k u^l \qquad 0 \leq n \leq m, \quad m \geq 1 \tag{5.7}$$

Using $\Omega^{(n,m)}(z,u)$, one can convert (5.4) and (5.6) into

$$\Omega^{(0,m)}(z,u) = \Pi^{(M)}(z,u)v(m) \qquad m \geq 1 \tag{5.8}$$

$$\Omega^{(n+1,m)}(z,u) = A(z)\Omega^{(n,m)}(z,u) + [A(z) - \lambda(0)]B(u)\Omega_0^{(n,m)}$$

$$0 \leq n \leq m-1, \quad m \geq 1 \tag{5.9}$$

The state of the system at the beginning of a vacation can be expressed by

$$\Omega_0^{(0)}(w) := \sum_{m=1}^{\infty} \Omega_0^{(0,m)} w^m \tag{5.10}$$

$$\Omega^{(0)}(z,u;w) := \sum_{m=1}^{\infty} \Omega^{(0,m)}(z,u) w^m \tag{5.11}$$

Substituting (5.3) into (5.10), we get

$$\Omega_0^{(0)}(w) = \left[\sum_{n=0}^{M-1} \Pi_1^{(n)}(1)\lambda(0) + P_0^{(0)}\right] V(w) \tag{5.12}$$

Putting (5.8) into (5.11) yields

$$\Omega^{(0)}(z,u;w) = \Pi^{(M)}(z,u)V(w) \tag{5.13}$$

Hence, $\Omega_0^{(0)}(w)$ and $\Omega^{(0)}(z,u;w)$ are known as $\{\Pi_1^{(n)}(1); 0 \leq n \leq M-1\}$, $P_0^{(0)}$, and $\Pi^{(M)}(z,u)$ are from the analysis of a service period in Section 4.

We now express the state of the system during a vacation in terms of $\Omega^{(0)}(w)$ and $\Omega^{(0)}(z,u;w)$. In order to do so, we define the generating functions for the probability that the system is empty as

$$\Omega_0^{(m)}(v) := \sum_{n=0}^{m} \Omega_0^{(n,m)} v^n \qquad m \geq 1 \tag{5.14}$$

$$\Omega_0(v,w) := \sum_{m=1}^{\infty} \Omega_0^{(m)}(v) w^m \tag{5.15}$$

From (5.5), we have

$$\Omega_0^{(n,m)} = \Omega_0^{(0,m)}[\lambda(0)]^n \qquad 0 \leq n \leq m \tag{5.16}$$

Substituting (5.16) into (5.14), we get

$$\Omega_0^{(m)}(v) = \Omega_0^{(0,m)} \frac{1 - [\lambda(0)v]^{m+1}}{1 - \lambda(0)v} \qquad m \geq 1 \tag{5.17}$$

Substituting (5.17) into (5.15) gives

$$\Omega_0(v,w) = \frac{\Omega^{(0)}(w) - \lambda(0)v\Omega^{(0)}[\lambda(0)vw]}{1-\lambda(0)v} \tag{5.18}$$

We also define

$$\Omega^{(m)}(z,u;v) := \sum_{n=0}^{m} \Omega^{(n,m)}(z,u)v^n \qquad m \geq 1 \tag{5.19}$$

$$\Omega(z,u;v,w) := \sum_{m=1}^{\infty} \Omega^{(m)}(z,u;v)w^m \tag{5.20}$$

Multiplying both sides of (5.9) by v^n and summing over $n = 0,1,2,\cdots,m-1$, we get

$$\frac{\Omega^{(m)}(z,u;v) - \Omega^{(0,m)}(z,u)}{v} = A(z)\left[\Omega^{(m)}(z,u;v) - \Omega^{(m,m)}(z,u)v^m\right]$$

$$+[A(z)-\lambda(0)]B(u)\left[\Omega_0^{(m)}(v) - \Omega_0^{(m,m)}v^m\right] \qquad m \geq 1 \tag{5.21}$$

Calculating $\Omega_0^{(m)}(v) - \Omega_0^{(m,m)}v^m$ from (5.16) and (5.17), and substituting into (5.21), we get

$$[1-vA(z)]\Omega^{(m)}(z,u;v) = \Omega^{(0,m)}(z,u) - A(z)\Omega^{(m,m)}(z,u)v^{m+1}$$

$$+[A(z)-\lambda(0)]B(u)v\Omega_0^{(0,m)}\frac{1-[\lambda(0)v]^m}{1-\lambda(0)v} \qquad m \geq 1 \tag{5.22}$$

Since $\Omega^{(m)}(z,u;v)$ is a polynomial of the mth order in v, the r.h.s. of (5.22) must be zero when $v = 1/A(z)$. Hence, we determine that

$$\Omega^{(m,m)}(z,u) = \Omega^{(0,m)}(z,u)[A(z)]^m + B(u)\Omega_0^{(0,m)}\left\{[A(z)]^m - [\lambda(0)]^m\right\} \qquad m \geq 1 \tag{5.23}$$

which gives

$$\sum_{m=1}^{\infty} \Omega^{(m,m)}(z,u)w^m = \Omega^{(0)}(z,u;wA(z)) + B(u)\left\{\Omega_0^{(0)}[wA(z)] - \Omega_0^{(0)}[\lambda(0)w]\right\} \tag{5.24}$$

Multiplying both sides of (5.22) by w^m, and summing over $m = 1,2,\cdots$, we get

$$[1-vA(z)]\Omega(z,u;v,w) = \Omega^{(0)}(z,u;w) - vA(z)\sum_{m=1}^{\infty}\Omega^{(m,m)}(z,u)(vw)^m$$

$$+[A(z)-\lambda(0)]B(u)v\frac{\Omega^{(0)}(w) - \Omega^{(0)}[\lambda(0)vw]}{1-\lambda(0)v} \tag{5.25}$$

Substituting (5.24) into (5.25), we obtain

$$\Omega(z,u;v,w) = \frac{\Omega^{(0)}(z,u;w) - vA(z)\Omega^{(0)}(z,u;vwA(z)) + B(u)\Omega_0^{(0)}[vwA(z)]}{1-vA(z)}$$

$$+\frac{[A(z)-\lambda(0)]B(u)v\Omega^{(0)}(w)}{[1-\lambda(0)v][1-vA(z)]} + \frac{\lambda(0)B(u)v\Omega^{(0)}[\lambda(0)vw]}{1-\lambda(0)v} \tag{5.26}$$

Thus we have obtained the system state during a vacation from (5.18) and (5.26) in terms of the initial state $\Omega_0^{(0)}(w)$ and $\Omega^{(0)}(z,u;w)$ given in (5.14) and (5.15).

The condition at the end of a vacation, that is, at the beginning of a service period is given by

$$P_0^{(0)} = \sum_{m=1}^{\infty} \Omega_0^{(m,m)} \tag{5.27}$$

$$\Pi_k^{(0)}(l) = \sum_{m=1}^{\infty} \Omega_k^{(m,m)}(l) \qquad k \geq 1, \quad l \geq 1 \tag{5.28}$$

Using (5.16) for $\Omega_0^{(m,m)}$ in (5.27), we get

$$P_0^{(0)} = \sum_{m=1}^{\infty} \Omega_0^{(0,m)}[\lambda(0)]^m = \Omega_0^{(0)}[\lambda(0)] \tag{5.29}$$

which is the probability that no customers arrive during a vacation at the beginning of which the system was empty. Note that we can recover (4.5) from (5.12) with $w = \lambda(0)$ and (5.29). Using (5.7) and (5.24) with $w = 1$, we can convert (5.28) into

$$\Pi^{(0)}(z,u) = \sum_{m=1}^{\infty} \Omega^{(m,m)}(z,u) = \Omega^{(0)}(z,u;A(z)) + B(u)\left\{\Omega^{(0)}[A(z)] - \Omega^{(0)}[\lambda(0)]\right\} \tag{5.30}$$

The PGF $U(u)$ for the unfinished work in the system immediately after an arbitrary slot boundary is given by

$$\begin{aligned} U(u) &= \Omega_0(1,1) + \sum_{m=1}^{\infty}\sum_{n=0}^{m}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\Omega_k^{(n,m)}(l)[B(u)]^{k-1}u^l \\ &\quad + \sum_{n=1}^{M-1}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\Pi_k^{(n)}(l)[B(u)]^{k-1}u^l \qquad (5.31) \\ &= \Omega_0(1,1) + \frac{\Omega(B(u),u;1,1)}{B(u)} + \sum_{n=1}^{M-1}\frac{\Pi^{(n)}(B(u),u)}{B(u)} \end{aligned}$$

6. Queue Size and Mean Response Time

From the system state during both a service period and a vacation, we can obtain the PGFs for the queue size and the unfinished work immediately after an arbitrary slot boundary. From the mean queue size immediately after an arbitrary slot boundary and Little's theorem, we can obtain the mean response time. The mean response time can also be obtained from the calculation of the mean queue size immediately after a service completion.

The distribution of the queue size L immediately after an arbitrary slot is given by

$$\text{Prob}[L=0] = \sum_{m=1}^{\infty}\sum_{n=0}^{m}\Omega_0^{(n,m)} \tag{6.1}$$

$$\text{Prob}[L=k] = \sum_{m=1}^{\infty}\sum_{n=0}^{m}\sum_{l=1}^{\infty}\Omega_k^{(n,m)}(l) + \sum_{n=1}^{M-1}\sum_{l=1}^{\infty}\Pi_k^{(n)}(l) \qquad k \geq 1 \tag{6.2}$$

Using (5.12), (5.15), (5.18), and (5.29), we can express Prob$[L=0]$ in (6.1) as

$$\begin{aligned}\text{Prob}[L=0] &= \Omega_0(1,1) = \frac{\Omega^{(0)}(1) - \lambda(0)\Omega^{(0)}[\lambda(0)]}{1-\lambda(0)} \\ &= P_0^{(0)} + \frac{\lambda(0)}{1-\lambda(0)} \sum_{n=0}^{M-1} \Pi_1^{(n)}(1)\end{aligned} \tag{6.3}$$

From (6.2), we get

$$\sum_{k=1}^{\infty} \text{Prob}[L=k]z^k = \Omega(z,1;1,1) + \sum_{n=1}^{M-1} \Pi^{(n)}(z,1) \tag{6.4}$$

However, from (5.26) and (5.30), we have

$$\begin{aligned}\Omega(z,1;1,1) &= \frac{\Omega^{(0)}(z,1;1) - A(z)\{\Omega^{(0)}(z,1;A(z)) + \Omega^{(0)}[A(z)]\}}{1-A(z)} \\ &\quad + \frac{[A(z)-\lambda(0)]\Omega^{(0)}(1)}{[1-\lambda(0)][1-A(z)]} + \frac{\lambda(0)\Omega^{(0)}[\lambda(0)]}{1-\lambda(0)} \\ &= \frac{\Omega^{(0)}(z,1;1) - A(z)\Pi^{(0)}(z,1)}{1-A(z)} + \frac{\{\Omega^{(0)}(1) - \Omega^{(0)}[\lambda(0)]\}[A(z)-\lambda(0)]}{[1-\lambda(0)][1-A(z)]}\end{aligned} \tag{6.5}$$

Further using (5.12) and (5.13) with $u = w = 1$ and (5.29), we get

$$\Omega(z,1;1,1) = \frac{\Pi^{(M)}(z,1) - A(z)\Pi^{(0)}(z,1)}{1-A(z)} + \frac{\lambda(0)\left[\sum_{n=0}^{M-1} \Pi_1^{(n)}(1)\right][A(z)-\lambda(0)]}{[1-\lambda(0)][1-A(z)]} \tag{6.6}$$

Hence, we have expressed the queue size distribution immediately after an arbitrary slot boundary in (6.3), (6.4), and (6.6) in terms of $P_0^{(0)}$, $\Sigma_{n=0}^{M-1}\Pi_1^{(n)}(1)$, and $\{\Pi^{(n)}(z,1); 0 \le n \le M\}$. From the mean queue size $E[L]$ immediately after an arbitrary slot boundary, we can obtain the mean response time $E[T]$ by using Little's theorem given in (3.21).

We can also obtain the mean response time $E[T]$ from the calculation of the mean queue size immediately after a service completion. By collecting terms relevant to service completions in (4.3)-(4.5), the probability π_k that there are k customers in the system immediately after a service completion is given by

$$\pi_k = c \sum_{n=0}^{M-1} \sum_{j=1}^{k+1} \Pi_j^{(n)}(1)\lambda(k-j+1) \qquad k \ge 0 \tag{6.7}$$

where c is a constant. The generating function $\Pi(z)$ for $\{\pi_k; k \ge 0\}$ is then given by

$$\begin{aligned}\Pi(z) &:= \sum_{k=0}^{\infty} \pi_k z^k = c \sum_{k=0}^{\infty} z^k \sum_{n=0}^{M-1} \sum_{j=1}^{k+1} \Pi_j^{(n)}(1)\lambda(k-j+1) \\ &= c \sum_{n=0}^{M-1} \sum_{j=1}^{\infty} \Pi_j^{(n)}(1) z^{j-1} \sum_{k=j-1}^{\infty} \lambda(k-j+1) z^{k-j+1} \\ &= c\frac{A(z)}{z} \sum_{n=0}^{M-1} \sum_{j=1}^{\infty} \Pi_j^{(n)}(1) z^{j-1} = c\frac{A(z)}{z} \sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(z)\end{aligned} \tag{6.8}$$

We can determine c from the normalization condition $\Pi(1) = 1$ as

$$\frac{1}{c} = \sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(1) \tag{6.9}$$

Hence we get

$$\Pi(z) = \frac{A(z)}{z} \bullet \frac{\sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(z)}{\sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(1)} \tag{6.10}$$

The mean queue size $E[L_D]$ immediately after a service completion is given by

$$E[L_D] = \lambda - 1 + \frac{\sum_{n=0}^{M-1} \left[\frac{d\overline{\Pi}_1^{(n)}(z)}{dz} \right]_{z=1}}{\sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(1)} \tag{6.11}$$

In order to obtain the mean response time $E[T]$, let us temporarily assume that customers are served on the first-come first-served (FCFS) basis. Then, the set of customers present in the system immediately after the completion of service to an arbitrary customer consists of those arrivals during the time in which the arbitrary customer was in the system and those customers that arrived in the same slot as the arbitrary customer did and were placed behind it. Therefore, we have

$$E[L_D] = \lambda E[T] + \frac{\lambda^{(2)}}{2\lambda} \tag{6.12}$$

From (6.11) and (6.12), we obtain

$$E[T] = \frac{\sum_{n=0}^{M-1} \left[\frac{d\overline{\Pi}_1^{(n)}(z)}{dz} \right]_{z=1}}{\lambda \sum_{n=0}^{M-1} \overline{\Pi}_1^{(n)}(1)} - \frac{\lambda^{(2)} + 2\lambda(1-\lambda)}{2\lambda^2} \tag{6.13}$$

In fact, the mean response time $E[T]$ in (6.13) holds for systems with any service discipline that does not distinguish customers on the basis of their service times (including FCFS, last-come first-served, and random order of service).

7. Solution by a Discrete Fourier Transform Technique

As pointed out in Section 4, the unknowns to be computed include $\{\overline{\Pi}_1^{(n)}(1); 0 \leq n \leq M-1\}$, $\{\Pi^{(n)}(z,u); 0 \leq n \leq M\}$, and $P_0^{(0)}$. We show in this section how to solve for these unknowns numerically by a technique of discrete Fourier transforms (DFTs) [8].

To begin with, let us explain the applicability of discrete Fourier transforms to the problem under study. As an example, consider the double z-transform $\Pi^{(n)}(z,u)$, where z and u are associated with the queue size and the remaining service time, respectively. For a given parameter setting under which the system is stable in the steady state, one can determine the "maximum" queue size N_q and the "maximum" service time N_s such that the probability of reaching a queue size beyond N_q or having a service time longer than N_s is negligibly small (for example, less than 10^{-8}). Since the number of feasible system states now becomes finite, $\Pi^{(n)}(z,u)$ can be represented by a set of DFTs $\{\Pi^{(n)*}(i,j); 0 \le i \le N_q, 0 \le j \le N_s\}$. Specifically, let us define $\omega_q := e^{-2\pi j/(N_q+1)}$ and $\omega_s := e^{-2\pi j/(N_s+1)}$, where $j := \sqrt{-1}$. By the definition of DFTs (see, for example, [13]), we have

$$\Pi^{(n)*}(i,j) = \sum_{k=0}^{N_q}\sum_{l=0}^{N_s}\omega_q^{ki}\omega_s^{lj}\Pi_k^{(n)}(l) \qquad 0 \le i \le N_q, \quad 0 \le j \le N_s \tag{7.1}$$

From (4.9), inverting the DFTs $\{\Pi^{(n)*}(i,j)\}$ with respect to j (which corresponds to u in (4.9)) gives us the DFTs $\{\overline{\Pi}_1^{(n)*}(i); 0 \le i \le N_q\}$ corresponding to $\overline{\Pi}_1^{(n)}(z)$. Therefore, we have

$$\overline{\Pi}_1^{(n)*}(i) = \frac{1}{N_s+1}\sum_{j=0}^{N_s}\omega_s^{-j}\Pi^{(n)*}(i,j) \qquad 0 \le i \le N_q \tag{7.2}$$

Similarly, inverting the above DFTs once more yields

$$\Pi_1^{(n)}(1) = \frac{1}{N_q+1}\sum_{i=0}^{N_q}\omega_s^{-i}\overline{\Pi}_1^{(n)*}(i) \tag{7.3}$$

Further, let $\{A^*(i); 0 \le i \le N_q\}$, $\{V^*[A^*(i)]; 0 \le i \le N_q\}$, and $\{B^*(j); 0 \le j \le N_s\}$ be the DFTs corresponding to the PGFs $A(z)$, $V[A(z)]$, and $B(u)$, respectively. By the same approach as outlined above, the DFTs $\{A^*(i); 0 \le i \le N_q\}$, $\{V^*[A^*(i)]; 0 \le i \le N_q\}$, and $\{B^*(j); 0 \le j \le N_s\}$ can be obtained from their known generating functions.

To present an iterative procedure for solving unknowns $\{\Pi_1^{(n)}(1); 0 \le n \le M-1\}$, $\{\Pi^{(n)}(z,u); 0 \le n \le M\}$, and $P_0^{(0)}$, we begin by rewriting several equations in Section 4 in terms of DFTs. From (4.11), we have

$$\Pi^{(n+1)*}(i,j) = \frac{A^*(i)}{\omega_s^j}\Pi^{(n)*}(i,j) + \Lambda^*(i)\left[\frac{B^*(j)}{\omega_q^i} - 1\right]\overline{\Pi}_1^{(n)*}(i) - \lambda(0)B^*(j)\Pi_1^{(n)}(1)$$

$$n = 0,1,2,\cdots,M-1 \tag{7.4}$$

Solving (4.5) for $P_0^{(0)}$, we get

$$P_0^{(0)} = \frac{V[\lambda(0)]}{1-V[\lambda(0)]}\sum_{n=0}^{M-1}\Pi_1^{(n)}(1)\lambda(0) \tag{7.5}$$

where $f_0 = V[\lambda(0)]$ has been applied. Similarly, using (4.13), we can obtain

$$\Pi^{(0)*}(i,j) = P_0^{(0)}\left[\frac{V^*[A^*(i)]}{V[\lambda(0)]} - 1\right]B^*(j) + \Pi^{(M)*}(i,j)V^*[A^*(i)] \tag{7.6}$$

Note that (7.4) through (7.6) allow us to find the unknowns iteratively as follows. Substituting an initial solution of $\{\Pi^{(M)*}(i,j); 0 \le i \le N_q, 0 \le j \le N_s\}$ into (7.6), $\Pi^{(0)*}(i,j)$ now becomes known for all $i = 0,1,\cdots,N_q$ and $j = 0,1,\cdots,N_s$. Based on these DFTs, $\{\Pi^{(n)*}(i,j); 1 \le n \le M\}$ can be obtained recursively from (7.4), where $\{\overline{\Pi}_1^{(n)*}(i)\}$ is computed according to (7.2). In addition, with the help of (7.3), $\Pi_1^{(n)}(1)$ is known, so is $P_0^{(0)}$ from (7.5). Finally, a new set of DFTs $\{\Pi^{(0)*}(i,j)\}$ can be found from (7.6). Therefore, the computation forms an iterative procedure and the iteration stops when the old and new sets of $\{\Pi^{(0)*}(i,j)\}$ and $P_0^{(0)}$ differ less than a convergence criterion.

A few remarks may be in order. First, the initial solution of $\{\Pi^{(M)*}(i,j)\}$ can be chosen as $\Pi^{(M)*}(i,j) = 0$ for all $i = 0, 1, \cdots, N_q$ and $j = 0, 1, \cdots, N_s$, which represents the system being empty. Second, the iteration procedure actually determines all unknowns in terms of DFTs from the set of equations for the time-limited service system. Once these unknowns have been obtained, other performance measures discussed in Sections 4 through 6 also become known. Finally, although we have only shown how to apply the DFT technique to analyze the joint queue size and the remaining service time, the same approach is also applicable to the process of unfinished work studied in Section 3. Specifically, we can convert the generating functions given by (3.8) (or (3.16)) and (3.17) into their DFT version and form an iteration procedure similar to (7.4)–(7.6). For brevity, we omit the details here.

Before presenting some numerical examples, let us comment on the selection of the "maximum" queue size and service time, N_q and N_s. Given a service time distribution and a target accuracy ϵ of the computation results, we can choose N_s such that the probability of having a service time exceeding N_s is less than $\epsilon/10$. One way to estimate the initial value of N_q is to solve the system as if the time limit M were set to infinity. Then, the needed value for N_q is identified so that the probability of reaching the queue size beyond N_q is less than $\epsilon/10$. In general, up to the limit of computer precision, the larger values of N_q and N_s, the less aliasing errors result from the DFTs. Therefore, if the computation results are not satisfactory, we can always repeat the computation by choosing larger values of N_q and N_s.

8. Numerical Examples

Let us show two sets of numerical examples. In the first set, we assume a Bernoulli arrival process and a geometric distribution in (3.22) for the service time. In this case, we can obtain the mean response time from the analysis of the unfinished work as well as from the analysis of the joint process of the queue size and the remaining service time. It is therefore interesting to compare the values of the mean response time obtained from the two methods. Table 1 shows such a comparison when we assume $\beta = 0.5 (b = 2)$, $M = 10$ slots, and the Bernoulli arrival process

$$A(z) = 1 - \lambda + \lambda z \tag{8.1}$$

In addition, we assume a uniform distribution for the length of a vacation as follows:

$$V(u) = \frac{1}{5}\sum_{l=1}^{5} u^l \tag{8.2}$$

Excellent agreement of the results from both methods in Table 1 demonstrates the correctness of our analysis. The CPU time consumption for each case on a SUN SPARC workstation was 0.4 sec. to 1 min. for the method based on the analysis of the unfinished work, while it was 5 sec. to 33 min. for that of the queue size.

In the second set of examples, we assume a geometric distribution for the arrival process:

$$A(z) = \frac{1-\alpha}{1-\alpha z} \quad ; \quad \lambda = \frac{\alpha}{1-\alpha} \tag{8.3}$$

a bimodal distribution for the service time:

$$B(u) = 0.5u + 0.5u^3 \quad ; \quad b = 2 \tag{8.4}$$

and a uniform distribution for the vacation time given in (8.2). Table 2 shows the values of the mean response time calculated by the method based on the analysis of the queue size. As $M \to \infty$, our time-limited system reduces to an exhaustive service system for which we have explicitly

$$E[T] = \frac{\lambda^2 b^{(2)} - \lambda\rho + \lambda^{(2)}}{2\lambda(1-\rho)} + b + \frac{E[V(V-1)]}{E[V]} \tag{8.5}$$

Our numerical values when $M \to \infty$ agrees with (8.5). The CPU time consumption was 0.3 to 16.5 min. for each case.

9. Concluding Remarks

In this paper, we have analyzed a discrete-time vacation model with exhaustive time-limited service, which can be applied to the evaluation of message delays for a station in timed-token networks. We have studied two processes of the unfinished work in the system, and a joint process of the queue size, the remaining service time, and the elapsed vacation time. A set of unknown functions that appear in the formulation of these processes have been solved by a technique based on discrete Fourier transforms. Numerical examples have been presented and the results have been verified by two different approaches.

The present technique can be applied to similar discrete-time queueing systems. For example, we can consider a time-limited service system with *overrunning* service; that is, the service is always continued to completion even if the time limit expires in the middle of the service time. This feature appears in a standard of a token passing [2]. The approach is also applicable to a system of queues attended by a single server in cyclic order (a polling system) which has been often used as a model of token passing protocols. These are the subjects of future research.

References

[1] ANSI Standard, *FDDI Token Ring - Media Access Control,* ANSI X3.139-1987, Nov.5, 1986.

[2] ANSI/IEEE Standard 802.4, *Token-Passing Bus Access Method,* IEEE Press, New York, 1985.

[3] Boxma, O. J., and Groenendijk, W. P. 1988. Waiting times in discrete-time cyclic-service systems. *IEEE Transactions on Communications,* Vol.COM-36, No.2 (February), pp.164-170.

[4] Bux, W. 1989. Token-ring local-area networks and their performance. *Proceedings of the IEEE,* Vol.77, No.2 (February), pp.238-256.

[5] Chiarawongse, J., Srinivasan, M. M., and Teorey, M. J. The M/G/1 queueing system with vacations and timer-controlled service. Preprint.

[6] Doshi, B. T. 1986. Queueing systems with vacations – A survey. *Queueing Systems,* Vol.1, No.1 (June), pp.29-66.

[7] LaMaire, R. O. 1991. An M/G/1 vacation model of an FDDI station. *IEEE Journal on Selected Areas in Communications,* Vol.9, No.2 (February), pp.257-264.

[8] Leung, K. K. 1991. Cyclic-service systems with probabilistically-limited service. *IEEE Journal on Selected Areas in Communications,* Vol.9, No.2 (February), pp.185-193.

[9] Leung, K. K. 1993. An execution/sleep scheduling policy for serving an additional job in priority queueing systems. *Journal of the Association for Computing Machinery,* Vol.40, No.2 (April), pp.394-417.

[10] Leung, K. K., and Eisenberg, M. 1990. A single-server queue with vacations and gated time-limited service. *IEEE Transactions on Communications,* Vol.38, No.9 (September), pp.1454-1462.

[11] Leung, K. K., and Eisenberg, M. 1991. A single-server queue with vacations and non-gated time-limited service. *Performance Evaluation,* Vol.12, No.2 (April), pp.115-125.

[12] Leung, K. K., and Lucantoni, D. M. 1993. Two vacation models for token-ring networks where service is controlled by timers. To be presented at Performance '93.

[13] Oppenheim, A. V., and Schafer, R. 1975. *Signal Processing.* Prentice-Hall, Englewood Cliff, New Jersey.

[14] Pang, J. W., and Tobagi, F. A. 1989. Throughput analysis of a timer controlled token passing protocol under heavy load. *IEEE Transactions on Communications,* Vol.37, No.7 (July), pp.694-702.

[15] Sevcik, K. C., and Johnson, M. J. 1987. Cycle time properties of the FDDI token ring protocol. *IEEE Transactions on Software Engineering,* Vol.SE-13, No.3 (March), pp.376-385.

[16] Takagi, H. 1986. *Analysis of Polling Systems.* The MIT Press, Cambridge, Massachusetts.

[17] Takagi, H. 1988. Queuing analysis of polling models. *ACM Computing Surveys,* Vol.20, No.1 (March), pp.5-28.

[18] Takagi, H. 1991. *Queueing Analysis : A Foundation of Performance Evaluation, Volume 1: Vacation and Priority Systems, Part 1.* Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

[19] Takagi, H. 1991. Application of polling models to computer networks. *Computer Networks and ISDN Systems,* Vol.22, No.3 (October), pp.193-211.

[20] Yue, O.-C., and Brooks, C. A. 1990. Performance of the timed token scheme in MAP. *IEEE Transactions on Communications,* Vol.38, No.7 (July), pp.1006-1012.

ρ	$E[T]$ from Unfinished work	$E[T]$ from Queue Size
0.1	3.47301	3.47301
0.3	4.02091	4.02091
0.5	5.66906	5.66907
0.7	18.05504	18.05521

Table 1: Mean response time for the cases with a Bernoulli arrival process, a geometric service time distribution, and $M = 10$, obtained by two different approaches.

ρ	$M = 50$	$M = 100$	$M = 500$	$M = 1000$	$M = \infty$
0.1	3.52778	3.52778	3.52778	3.52778	3.52778
0.3	4.08340	4.08333	4.08333	4.08333	4.08333
0.5	5.10485	5.08381	5.08333	5.08333	5.08333
0.7	7.96217	7.51826	7.41668	7.41667	7.41667
0.9	40.31619	25.08919	19.51791	19.17385	19.08333

Table 2: Mean response time for cases with a geometric batch arrival process, a bimodal service time distribution, and selected values of M (obtained from the analysis of queue size).