J. Appl. Prob. **37**, 1092–1098 (2000)
Printed in Israel

ANALYTICAL SOLUTION OF FINITE CAPACITY $M/D/1$ QUEUES

OLIVIER BRUN * AND
JEAN-MARIE GARCIA,* ** *LAAS-CNRS*

Abstract

Although the $M/D/1/N$ queueing model is well solved from a computational point of view, there is no known analytical expression of the queue length distribution. In this paper, we derive closed-form formulae for the distribution of the number of customers in the system in the finite-capacity $M/D/1$ queue. We also give an explicit solution for the mean queue length and the average waiting time.

Keywords: Finite capacity $M/D/1$ queues; embedded Markov chain; queue length distribution; mean number of customers; average waiting time

AMS 2000 Subject Classification: Primary 60K25
Secondary 68M20

1. Introduction

In the performance evaluation of modern telecommunication systems, deterministic service time queueing models are frequently applied for system modelling. For instance, such models are often used to describe the cell scale queueing problem in the ATM multiplexer. According to the system to be modelled, several deterministic service time models have been considered (see [8], [12] and [5]): $M/D/1$, $\mathrm{Geo}^N/D/1$, $N * D/D/1$ and $\sum D_i/D/1$.

The $M/D/1$ queue [6] is by far the simplest and most general model and it has a variety of applications not only in the telecommunication area, but also in operations research, computer science and many other engineering areas. This model, already studied by Erlang in 1909, is a queueing system with Poisson arrivals and deterministic (constant) service time. This model is appropriate for continuous deterministic service time queueing systems, which input can be seen as 'completely random' or as a superposition of a large number of processes. This follows from the fact that in situations with many sources, each having a small generation rate, the actual arrival process approximately follows a Poisson process.

Although this model has been studied for a long time and is well solved from a computational point of view, there are still few explicit analytical results for the finite capacity queue (denoted $M/D/1/N$). As discussed in Section 2, apart from the special case of no waiting room at all (Erlang's loss model), there is no closed-form formula for the queue length distribution. Moreover, there is no known analytical expression of the mean number of customers in the system.

In this paper, we derive closed-form formulae for the distribution of the number of customers in the system. We also give an explicit formula for the mean number of customers.

Received 27 *September* 1999; revision received 21 February 2000.
* Postal address: Laboratoire d'Analyse et d'Architecture des Systèmes du CNRS, Toulouse, France.
** Email address: jmg@laas.fr

The paper is organized as follows. Section 2 recalls the basic relations for $M/G/1/N$ queues. Section 3 is devoted to stationary analysis of $M/D/1/N$ queues. We derive closed-form formulae for the steady-state queue-length distribution, for the mean queue length and the average waiting time. In Section 4, we propose an algorithm for efficiently computing the queue length distribution.

2. The basic relations for $M/G/1/N$ queues

The $M/D/1/N$ model is a finite capacity queueing system, with $N-1$ places in the waiting room. Customers arrive according to a Poisson process at rate λ. They are serviced according to the FCFS (First Come First Served) discipline and the service time of each customer is the same constant T. Customers who upon arrival see a full system are rejected and do not further influence the system (lost customer cleared). Since the finite waiting room acts as a regulator on the queue size, no *a priori* assumption about the utilization factor $\rho = \lambda T$ is needed.

A computational scheme for the time average probabilities $\mathrm{P}_j(N)$ is known for the more general $M/G/1/N$ model [10]. Denoting by $A_{j,k}$ the expected amount of time that k customers are present during a service time that is started with j customers in the system, we get the following relations:

$$\mathrm{P}_k(N) = \lambda \mathrm{P}_0(N) A_{1,k} + \sum_{j=1}^{k} \lambda \mathrm{P}_j(N) A_{j,k}, \qquad 1 \le k \le N. \tag{1}$$

These relations allow recursive computation of $\mathrm{P}_1(N)/\mathrm{P}_0(N)$, $\mathrm{P}_2(N)/\mathrm{P}_0(N), \ldots$. The unknown $\mathrm{P}_0(N)$ is computed by normalizing probabilities.

Although the $M/D/1/N$ queue can be solved by means of the above relations, there is no analytical solution for the steady-state distribution.

3. The finite capacity $M/D/1$ queue

In this section we derive the important parameters of an $M/D/1/N$ queue. To this end, we consider the embedded Markov chain associated with the $M/D/1/N$ queue and derive the queue length distribution at departure epochs. It should be pointed out that if the probability distribution obtained at departure epochs happens also to be valid at all points in time in the infinite capacity case, this is no longer true in the finite capacity case. However, the probability distribution at departure epochs will be used to derive the stationary probability distribution at all points in time.

3.1. Embedded Markov chain

Let $X_N(t)$ be the number of customers in the system at time t. Let t_n be the date of the nth customer departure. The stochastic process $\{X_N(t_n)\}_{n\ge 0}$ is a Markov chain with state space $\{0, 1, \ldots, N-1\}$. Hereafter, $q_j(t_n)$ will denote the probability that j customers are left behind by the nth departure. Also, in the sequel, α_k will denote the probability of k arrivals during a customer service. Since arrivals are Poisson distributed with rate λ and since service duration is a constant T, we have:

$$\alpha_k = \frac{\rho^k}{k!} \mathrm{e}^{-\rho}.$$

It is easy to show that the probability transition matrix of the embedded Markov chain, $\Pi = [\pi_{i,j}]$, takes the form [9]:

$$\Pi = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \cdots & \alpha_{N-2} & 1-\sum_0^{N-2}\alpha_k \\ \alpha_0 & \alpha_1 & \alpha_2 & \cdots & \alpha_{N-2} & 1-\sum_0^{N-2}\alpha_k \\ 0 & \alpha_0 & \alpha_1 & \cdots & \alpha_{N-3} & 1-\sum_0^{N-3}\alpha_k \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_0 & 1-\alpha_0 \end{bmatrix}.$$

Moreover, it is known that the stochastic matrix Π is ergodic [2]. As a consequence, the stationary distribution $Q = [q_0, \dots, q_{N-1}]$ exists and it is an eigenvector of the matrix Π, so $Q\Pi = Q$. This implies that the stationary distribution Q verifies the following linear system:

$$\begin{aligned} \alpha_0 q_0 + \alpha_0 q_1 &= q_0 \\ \alpha_1 q_0 + \alpha_1 q_1 + \alpha_0 q_2 &= q_1 \\ \alpha_2 q_0 + \alpha_2 q_1 + \alpha_1 q_2 + \alpha_0 q_3 &= q_2 \\ \vdots \qquad\qquad & \vdots \ \vdots \\ \alpha_{N-2} q_0 + \alpha_{N-2} q_1 + \cdots + \alpha_0 q_{N-1} &= q_{N-2}. \end{aligned}$$

This is a linear system of $N-1$ equations involving the N unknowns $q_0, \dots, q_{N-1}$. Hence, it allows us to express the probabilities $q_1, \dots, q_{N-1}$ in terms of q_0. Let $a_0, \dots, a_{N-1}$ be such that $q_n = a_n q_0$ for all $n \geq 0$. It is easy to see that $a_0 = 1$, $a_1 = \mathrm{e}^{\rho} - 1$ and that $a_2, \dots, a_{N-1}$ obey the following recursion:

$$a_n = \mathrm{e}^{\rho}\left(a_{n-1} - \sum_{i=1}^{n-1} \alpha_i a_{n-i} - \alpha_{n-1} a_0\right). \tag{2}$$

Now, let $(a_n)_{n\geq 0}$ be the infinite sequence whose first terms are $a_0, \dots, a_{N-1}$ and whose other terms are defined by the recursion (2). Let $A(z)$ be the z-transform of this sequence, i.e.

$$A(z) = \sum_{k=0}^{\infty} a_k z^k.$$

Lemma 1 gives an explicit formula for the generating function $A(z)$.

Lemma 1. *The generating function $A(z)$ takes the form $A(z) = (1-z)B(z)$, where $B(z) = 1/(1 - z\mathrm{e}^{\rho(1-z)})$.*

Proof. The recursion scheme (2) implies that

$$\sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} \alpha_n a_0 z^n + \sum_{i=1}^{\infty}\sum_{n=0}^{\infty} \alpha_n a_i z^{n+i-1},$$

that is,

$$A(z) = a_0 \mathrm{e}^{-\rho} \sum_{n=0}^{\infty} \frac{(\rho z)^n}{n!} + \frac{1}{z} \sum_{i=1}^{\infty} a_i z^i \sum_{n=0}^{\infty} \alpha_n z^n.$$

Hence

$$A(z) = a_0 e^{\rho(z-1)} + \frac{e^{\rho(z-1)}}{z}(A(z) - a_0),$$

which leads to

$$A(z) = \frac{1-z}{1 - ze^{\rho(1-z)}}.$$

Now, let $(b_n)_{n\geq 0}$ be the coefficients of the z-transform $B(z)$. Proposition 1 provides a closed-form expression for these coefficients.

Proposition 1. *The coefficients a_n are given by $a_0 = 1$ and, for $n \geq 1$, $a_n = b_n - b_{n-1}$, where $b_0 = 1$ and, for $n \geq 1$,*

$$b_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!}(n-k)^k e^{(n-k)\rho} \rho^k.$$

Proof. We have

$$\sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} b_n z^n - z \sum_{n=0}^{\infty} b_n z^n,$$

hence

$$\sum_{n=0}^{\infty} a_n z^n = b_0 + \sum_{n=1}^{\infty} (b_n - b_{n-1}) z^n,$$

which proves that $a_0 = 1$ and that $a_n = b_n - b_{n-1}$ for all $n \geq 1$. We now turn to the second part of the proof. Let us define the z-transform $F(z)$ as follows:

$$F(z) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{(-1)^k}{k!}(n-k)^k e^{(n-k)\rho} \rho^k z^n.$$

Thus

$$F(z) = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^k}{k!} m^k e^{m\rho} \rho^k z^k z^m.$$

Since

$$\sum_{m=0}^{\infty} \frac{(-1)^k}{k!} m^k \rho^k z^k = e^{-m\rho z}$$

we have that

$$F(z) = \sum_{m=0}^{\infty} e^{m\rho(1-z)} z^m.$$

The series $F(z)$ converges if $|e^{\rho(1-z)} z| < 1$. In particular, if $|z| < 1$ we have the following explicit expression for the z-transform $F(z)$,

$$F(z) = \frac{1}{1 - ze^{\rho(1-z)}}$$

and thus we find that $F(z) = B(z)$ which proves that $b_0 = 1$ and

$$b_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!}(n-k)^k e^{(n-k)\rho} \rho^k, \qquad \forall n \geq 1.$$

Remark 1. Letting $\mathrm{P}_k(t) = ((\lambda t)^k/k!)\mathrm{e}^{-\lambda t}$, we have the following simple expression for the coefficients $(b_n)_{n\geq 0}$:

$$b_n = \sum_{k=0}^{n} \mathrm{P}_k((k-n)T), \qquad \forall n \geq 0.$$

We can now derive an explicit expression for the steady-state probability distribution at departure epochs.

Lemma 2. *At equilibrium, the probability distribution $Q = [q_0, \ldots, q_n]$ of the number of customers left behind by a departure is given by*

$$q_0 = \frac{1}{b_{N-1}}, \qquad q_n = \frac{b_n - b_{n-1}}{b_{N-1}}, \qquad n = 1, \ldots, N-1.$$

Proof. The probability normalization implies that $\sum_{k=0}^{N-1} q_k = 1$. Thus,

$$q_0 = \frac{1}{\sum_{k=0}^{N-1} a_k} = \frac{1}{b_{N-1}}$$

and, moreover,

$$q_n = a_n q_0 = \frac{b_n - b_{n-1}}{b_{N-1}}, \qquad n = 1, \ldots, N-1.$$

3.2. Steady-state probability distribution

It should be pointed out that in the finite capacity case, the steady-state probability distribution Q of the number of customers left behind by a departure differs from the queue length distribution $\mathrm{P} = [\mathrm{P}_0(N), \ldots, \mathrm{P}_N(N)]$ in the $M/D/1/N$ queue. However, in the case of $M/G/1/N$ queues, it is known that the following relation holds [2]:

$$q_j = \frac{\mathrm{P}_j(N)}{1 - \mathrm{P}_N(N)}.$$

Using this relation the following theorem states our main result.

Theorem 1. *The probability distribution of the number of customers in the system is given by*

$$\begin{aligned}
\mathrm{P}_0(N) &= \frac{1}{1 + \rho b_{N-1}}, \\
\mathrm{P}_N(N) &= 1 - \frac{b_{N-1}}{1 + \rho b_{N-1}}, \\
\mathrm{P}_j(N) &= \frac{b_j - b_{j-1}}{1 + \rho b_{N-1}}, \qquad j = 1, \ldots, N-1.
\end{aligned} \tag{3}$$

Proof. We have

$$q_j = \frac{\mathrm{P}_j(N)}{1 - \mathrm{P}_N(N)}, \quad j = 0, \ldots, N-1. \tag{4}$$

A simple conservation law implies that

$$\lambda(1 - \mathrm{P}_N(N)) = \frac{1}{T}(1 - \mathrm{P}_0(N)).$$

Thus

$$1 - \mathrm{P}_N(N) = \frac{1}{\rho}(1 - (1 - \mathrm{P}_N(N))q_0).$$

Together with $q_0 = 1/b_{N-1}$, this leads to

$$1 - \mathrm{P}_N(N) = \frac{b_{N-1}}{1 + \rho b_{N-1}}. \tag{5}$$

Other probabilities are derived using (4).

Using Theorem 1, it is straightforward to find the mean number of customers X_N in the $M/D/1/N$ queue.

Theorem 2. *The mean number of customers of the $M/D/1/N$ queue is given by*

$$X_N = N - \frac{\sum_{k=0}^{N-1} b_k}{1 + \rho b_{N-1}}. \tag{6}$$

Proof. Since

$$X_N = \sum_{k=0}^{N} k\mathrm{P}_k(N)$$

we have

$$X_N = N + \frac{\sum_{k=1}^{N-1} k(b_k - b_{k-1}) - N b_{N-1}}{1 + \rho b_{N-1}}.$$

The expected result is then obtained after some algebra.

Theorem 3. *The mean waiting time W_N in the $M/D/1/N$ queueing system is*

$$W_N = \left(N - 1 - \frac{\sum_{k=0}^{N-1} b_k - N}{\rho b_{N-1}}\right)T.$$

Proof. Let T_N be the average system time. Application of Little's law yields

$$X_N = \lambda(1 - \mathrm{P}_N(N))T_N.$$

Using (5) and (6),

$$T_N = \frac{1}{\lambda}\frac{1 + \rho b_{N-1}}{b_{N-1}}\frac{N + N\rho b_{N-1} - \sum_{k=0}^{N-1} b_k}{1 + \rho b_{N-1}},$$

and thus

$$W_N = T_N - T = \left(N - 1 - \frac{\sum_{k=0}^{N-1} b_k - N}{\rho b_{N-1}}\right)T.$$

TABLE 1: Comparison between the numbers of operations involved in each algorithm.

Operations	Our algorithm	Algorithm (1)
$+$	$N(N-1)$	$\frac{1}{2}N(N+3)$
$-$	$2N$	N^2
$*$	$\frac{1}{2}N(N+3)$	$\frac{1}{2}N(N+7)$
$/$	N	$\frac{1}{2}N(N+3)$

4. Algorithmic issues

In this section, we propose an algorithm for efficiently computing the queue length distribution. This algorithm is compared to the one using the relation (1).

The algorithm is based upon relation (2) which allows an efficient computation of the coefficients a_n once the constants α_k have been computed. This relation can also be used to compute the coefficients b_n, since $b_0 = a_0$ and, for $k \geq 1$, $b_k = \sum_{i=0}^{k} a_i$. The queue length distribution then follows using (3).

Table 1 proposes a comparison between the number of operations involved in this algorithm and the number of operations involved in the algorithm (1). It is worthwhile noticing that the number of divisions is of magnitude N^2 in the algorithm (1) while it is only of magnitude N in our algorithm.

5. Conclusion

In this paper, we have derived a closed-form formula for the distribution of the number of customers in the finite-capacity $M/D/1$ queue. We also gave an explicit solution for the mean queue length and the average waiting time. In [1] it is shown that all derived results for the finite buffer queue are in agreement with those already known for the infinite buffer queue when the queue size grows to infinity.

Acknowledgements

This research was founded by CNRS and DELTA Partners Company, Labège, France.

References

[1] BRUN, O. AND GARCIA, J. M. (1999). Analytical solution of finite capacity $M/D/1$ queues. Res. Rept 99372, LAAS, Toulouse.
[2] COHEN, J. W. (1969). *The Single Server Queue*. North Holland, Amsterdam.
[3] CROMMELIN, C. D. (1933). Delay probability formulae. *Post Office Elec. Eng. J.* **26**.
[4] DOYON, G. (1989). *Systèmes et Réseaux de Télécommunications en Régime Stochastique* (Collection Technique et Scientifique des Télécommunications). Masson, Paris.
[5] GRAVEY, A., LOUVION, J. R. AND BOYER, P. (1990). On the Geo/D/1 and Geo/$D/1/N$ Queues. *Perf. Eval.* **11,** 117–125.
[6] KLEINROCK, L. (1975). *Queueing Systems. Volume I: Theory*. John Wiley, New York.
[7] KLEINROCK, L. (1976). *Queueing Systems. Volume II: Applications*. John Wiley, New York.
[8] ROBERTS, J., MOCCI, U. AND VIRTAMO, J. (eds) (1996). *Broadband Network Teletraffic. Final Report of Action Cost 242*. Springer, Berlin.
[9] TAKACS, L. (1962). *Introduction to the Theory of Queues*. Oxford University Press.
[10] TIJMS, H. C. (1986). *Stochastic Modelling and Analysis: A Computational Approach*. John Wiley, New York.
[11] TIJMS, H. C. (1994). *Stochastic Models: An Algorithmic Approach*. John Wiley, New York.
[12] VICARI, N. AND TRAN-GIA, P. (1996). A numerical analysis of the *Geo/D/N* queueing system. Tech. Rept 151, Institute of Computer Science, University of Würzburg.