Analytical study of a discrete time retrial queue with balking customers and early arrival scheme

Abdel-Karim Aboul-Hassan, Sherif I. Rabia and Ahmed Kadry

Eng. Math & Physics Dept., Faculty of Eng., Alexandria University, Alexandria, Egypt

This paper studies a discrete time Geo/G/1/1 retrial queue with balking customers in which an arriving customer that finds the service facility busy will either join the infinite buffer orbit with probability α_0 or leave the system with probability $1-\alpha_0$. Inter-retrial times are independent and follow a geometric distribution. Early arrival scheme is assumed. Analytic formula for the joint probability generating function of the remaining service time of the customer currently in the server and the number of customers in the orbit is derived. Moreover, the probability generating function of the total number of customers in the system is also obtained. Some special cases are considered.

تدرس هذه المقالة طابورا ذا محاولات متكررة وخاصية وصول مبكر يعمل فى الزمن المتقطع. زمن الخدمة يتبع توزيعا عاما والزمن بين وصول العملاء يتبع توزيعا هندسيا. العميل الذى يجد مركز الخدمة مشغولا عند وصوله قد يلتحق بمدار لانهائى باحتمال α_0 أو يغادر النظام باحتمال $1-\alpha_0$. الوقت بين المحاولات يتبع توزيعا هندسيا. تم الحصول على الدالة المولدة للاحتمالات لزمن الخدمة المتبقى وعدد العملاء فى المدار. كما تم الحصول على الدالة المولدة للاحتمالات لإجمالى عدد العملاء فى النظام. تعرض البحث كذلك لبعض الحالات الخاصة.

Keywords: Retrial queues, Discrete time, Early arrival scheme, Probability generating function

1. Introduction

Recently, there is a great interest in analyzing discrete time queueing systems. This interest is motivated by the application of discrete time queueing systems in analyzing computer and communication systems working in a slotted time environment (see [1 - 3]). For example [1], the BISDN (broadband integrated services digital network) provides a common interface for carrying different types of information including data, voice and video. Information in the BISDN is transported in discrete units of 53-octet ATM (Asynchronous Transfer Mode) cells. The time required to transmit such cells within the same system is constant and can be considered as the basic time slot. Analysis of such slotted time systems is based mainly on discrete time queueing systems.

The research in the area of discrete time queueing systems focused mainly on discrete time classical queueing systems. Few works have appeared in the area of discrete time retrial queues [4]. One of the earliest papers in this field is that of Yang and Li [5]. They considered a discrete time Geo/G/1/1 retrial queue with geometric retrial times. They obtained the joint generating function of the number of customers in the server and in the orbit in the steady state. Moreover, they developed a recursive formula for computing the steady state probabilities. In [6], Li and Yang considered a discrete time PH/Geo/1/1 retrial queue with geometric retrial times. A discrete time retrial queue with two types of customers was analyzed by Choi and Kim [7]. A similar model was studied by Li and Yang [8] where a recursive formula was presented to compute the marginal steady state distribution of the number of customers in the queue and in the orbit. A discrete time Geo/G/1/1 retrial queue with general retrial times was analyzed recently by Atencia and Moreno [4]. Nobel [9] considered a discrete time Geo/G/1/1 retrial queue with batch arrivals where the number of arrivals during any time slot follows a general distribution and the number of arrivals during different time slots are independent.

In all of the above discrete time retrial queues models, the customers are assumed to be persistent. The customer can not depart from the system before his required service is completed. A parallel study is needed for the case of impatient customers. Impatience in retrial queues is modeled by assuming that [10] an arriving customer to a busy server joins the orbit with probability α_0 and departs completely from the system without being served with probability $1-\alpha_0$. Moreover, if upon making the n^{th} retrial, a customer finds the server busy, he returns to the orbit with probability α_n and departs from the system with probability $1-\alpha_n$. Three important choices for the probabilities α_n appeared in the literature of the continuous time retrial queues [10]: (1) $a_0 < 1$, $\alpha_n = 1$; $n \geq 1$. (2) $a_0 = 1$, $\alpha_n = a < 1$; $n \geq 1$. (3) $\alpha_n = a < 1$; $n \geq 0$. The first case represents balking customer, the second represents reneging customers and the third combines both types of customers.

We are concerned here with balking customers. More specifically, we consider a discrete time retrial queue with geometric inter-arrival and inter-retrial times, general service times, one server, no queue and balking customers. System evolution is controlled by the early arrival scheme [11]. As in [5], we derive an explicit expression for the joint generating function of the number of customers in the orbit and the remaining service time. Moreover, the generating function of the total number of customers in the system (the server and the orbit) is derived. These generating functions could be used to extract important performance measures such as average system size and average waiting time. However, it seems hard to use the obtained generating functions to derive an explicit expression for probability distribution. Hence, in our incoming paper [12] we develop (starting from these generating functions) a recursive scheme to compute the steady state probabilities. For a parallel study concerning a similar system which is controlled by the late arrival scheme see [13].

This paper is organized as follows. In Section 2, we describe the mathematical model in more details and give the notations that will be used throughout this work. In Section 3, an explicit expression for the joint generating function of the number of customers in the orbit and the remaining service time of the customer in service is derived. Moreover, the generating function of the total number of customers in the system (the server and the orbit) is obtained. Some special cases are treated in section 4. Conclusion and some open problems are presented in section 5.

2. The model

We are concerned with a retrial queueing system with a single server and without a waiting room. It is assumed that the time axis is divided into intervals of equal length called time slots. All the system events occur at the boundary of these time slots. More specifically, the evolution of the system is controlled by the early arrival scheme [11]. In this scheme, it is assumed that the arrivals (either external or coming from the orbit) during time slot m ($m \geq 0$) occur at the beginning of this time slot. On the other hand, service completion during time slot m occurs by the end of this time slot. During any time slot m, a single external arrival occurs (independently of all the other system events) with probability p. In other words, it is assumed that the time between consecutive arrivals follows a geometric distribution. If the arriving customer finds the server busy, then he joins an infinite orbit (to retry his demand later) with probability α_0 and departs completely from the system with probability $1-\alpha_0$. Service time (counted in time slots) of any customer is assumed to follow a general distribution s_j, $j = 1, 2, 3, \ldots$, where $s_j = Pr\{$ service time = j time slots $\}$. Customers in the orbit retry (independently of each other) their service during each time slot with probability $1-r$. In other words, the time between retrials for each customer follows a geometric distribution with parameter $1-r$. If both an external arrival and a retrial occur at the same time and the server is idle, then the external arrival begins his service immediately and the other customer returns to the orbit to retry obtaining his service later. If no external arrival and more than one retrial occur at the same time to an idle server, then one customer is selected at random to start service and the other customers return to the orbit to retry obtaining their service later.

Define Nm to be the number of customers in the orbit at the beginning of the time slot m ($m \geq 0$). Since a general service time distribution is assumed, we have to keep track of the remaining service time also. Define Xm to be the number of the remaining time slots for the customer in the server at the beginning of the time slot m ($m \geq 0$). The system state is given by the two dimensional stochastic process $\{(X_m,N_m),\ m \geq 0\}$. Since both inter-arrival and inter-retrial times follow a geometric distribution, then the stochastic process $\{(X_m,N_m),\ m \geq 0\}$ is a Markov chain. The next section treats the problem of obtaining the steady state distribution of this process.

3. Steady state distribution of $\{(X_m,N_m),\ m \geq 0\}$

$\{(X_m,N_m),\ m \geq 0\}$ is a two dimensional Markov chain with state space $\{0, 1, 2, \ldots\} \times \{0, 1, 2, \ldots\}$. Define $P_{i,k}$ as the joint steady state probability of this stochastic process. More specifically $P_{i,k} = Pr\{X_m = i\ ,\ N_m = k\}$. The joint probability generating function of $P_{i,k}$ is given in Theorem 1 below. However, we have to define the following generating functions first:

$$S(z) = \sum_{j=1}^{\infty} s_j z^j \ ,$$

$$R_i(z) = \sum_{k=0}^{\infty} P_{i,k} z^k \ ,$$

$$P(x,z) = \sum_{i=1}^{\infty} R_i(z) x^i \ .$$

Theorem 1 The joint probability generating function of $P_{i,k}$ is given by*:*

$$P(x,z) = \frac{x}{x-1+\alpha_0 p(1-z)} \frac{p(1-z)[S(x)-S(1-\alpha_0 p(1-z))]}{S(1-\alpha_0 p(1-z))-z} R_0(z), \qquad (1)$$

where

$$R_0(z) = \frac{1-\alpha_0 pS'(1)}{1+(1-\alpha_0)pS'(1)} \frac{\prod_{k=0}^{\infty} G(r^k z)}{\prod_{k=0}^{\infty} G(r^k)}, \qquad (2)$$

and

$$G(z) = \frac{S(1-\alpha_0 p(1-z))-z}{(\bar{p}+pz)S(1-\alpha_0 p(1-z))-z}\,\bar{p}$$

Proof

The balance equations of the Markov chain $\{(X_m,N_m),\ m \geq 0\}$ are given by:

$$P_{0,0} = \bar{p}P_{0,0} + \bar{p}P_{1,0} + (1-\alpha_0)pP_{1,0}, \qquad (3)$$

$$P_{0,k} = \bar{p}r^k P_{0,k} + \bar{p}P_{1,k} + (1-\alpha_0)pP_{1,k} + \alpha_0 pP_{1,k-1}, \quad ,k \geq 1. \qquad (4)$$

$$P_{i,0} = \bar{p}P_{i+1,0} + (1-\alpha_0)pP_{i+1,0} + ps_i P_{0,0} + \bar{p}(1-r)s_i P_{0,1}, \quad ,i \geq 1. \qquad (5)$$

$$P_{i,k} = \bar{p}P_{i+1,k} + (1-\alpha_0)pP_{i+1,k} + \alpha_0 pP_{i+1,k-1} + ps_i P_{0,k} + \bar{p}(1-r^{k+1})s_i P_{0,k+1}, \quad ,i,k \geq 1, \qquad (6)$$

together with the normalizing equation:

$$\sum_{i=0}^{\infty}\sum_{k=0}^{\infty} P_{i,k} = 1. \qquad (7)$$

Multiplying both sides of (4) by z^k, making summation from $k = 1$ to ∞, and using the boundary condition in (3) yields:

$$R_0(z) = \bar{p}R_0(rz) + (1-\alpha_0)p(1-z))R_1(z). \qquad (8)$$

Applying a similar procedure to (5) and (6) we get:

$$R_i(z) = (1-\alpha_0 p(1-z))R_{i+1}(z) + (ps_i + \frac{\bar{p}s_i}{z})R_0(z) - \frac{\bar{p}s_i}{z} R_0(rz)$$

Substituting for $R_0(rz)$ in the above equation from (8) gives:

$$R_i(z) = (1-\alpha_0 p(1-z))R_{i+1}(z) + \frac{ps_i(z-1)}{z}R_0(z) + \frac{(1-\alpha_0 p(1-z))}{z}s_i R_1(z).$$

Multiplying both sides by x^i and summing over i we get:

$$\frac{x-1+\alpha_0 p(1-z)}{x}P(x,z) = \frac{p(z-1)}{z}R_0(z)S(x) + \frac{S(x)-z}{z}(1-\alpha_0 p(1-z))R_1(z). \quad (9)$$

A relation between $R_0(z)$ and $R_1(z)$ can be obtained by putting $x = 1-\alpha_0 p(1-z)$ in (9):

$$R_1(z) = \frac{p(1-z)S(1-\alpha_0 p(1-z))}{[S(1-\alpha_0 p(1-z))-z](1-\alpha_0 p(1-z))}R_0(z). \quad (10)$$

Substituting for $R_1(z)$ in (9) using the above equation we get:

$$\frac{x-1+\alpha_0 p(1-z)}{x}P(x,z) = \frac{p(1-z)[S(x)-S(1-\alpha_0 p(1-z))]}{S(1-\alpha_0 p(1-z))-z}R_0(z).$$

Rearranging terms yields eq. (1) of the theorem.

Now, in order to find $R_0(z)$, we substitute for $R_1(z)$ in (8) using (10). Hence,

$$R_0(z) = \frac{S(1-\alpha_0 p(1-z))-z}{(\bar{p}+pz)S(1-\alpha_0 p(1-z))-z}\bar{p}R_0(rz) = G(z)R_0(rz), \quad (11)$$

where $G(z)$ is as defined above. Applying (11) recursively we get:

$$R_0(z) = R_0(0)\prod_{k=0}^{\infty}G(r^k z). \quad (12)$$

Putting $z = 1$ in the above equation leads to:

$$R_0(0) = \frac{R_0(1)}{\prod_{k-0}^{\infty}G(r^k)}. \quad (13)$$

In order to get $R_0(1)$ we set $x = 1$ and take the limit as z tends to 1 in (1), then:

$$P(1,1) = \frac{pS'(1)}{1-\alpha_0 pS'(1)}R_0(1). \quad (14)$$

From the normalizing equation (7) it is clear that $R_0(1) + P(1, 1) = 1$. Hence, $R_0(1)$ has the following form:

$$R_0(1) = \frac{1-\alpha_0 pS'(1)}{1+(1-\alpha_0)pS'(1)}. \quad (15)$$

Substituting with (15) into (13) and then substituting with the result into (12) yields $R_0(z)$ as given in (2). This completes the proof.

Now we define $Q(z)$ to be the probability generating function of the number of customers in the system including those in orbit and the one in the server. In other words,

$$Q(z) = \sum_{k=0}^{\infty} q_k z^k,$$

where q_k is the probability that there are k customers in the system. Hence,

$$Q(z) = \sum_{k=0}^{\infty}(P_{0,k} + \sum_{i=1}^{\infty}P_{i,k-i})z^k = R_0(z) + zP(1,z),$$

Using (1), we obtain directly the following theorem:

Theorem 2 The probability generating function of the steady state distribution of the total number of customers in the system is given by:

$$Q(z) = \frac{z(1-\alpha_0)+(\alpha_0-z)S(1-\alpha_0 p(1-z))}{\alpha_0[S(1-\alpha_0 p(1-z))-z]}R_0(z).$$

4. Special cases

4.1. The Geo/G/1/1 retrial queue with persistent customers

When $\alpha_0 = 1$, we obtain the Geo/G/1/1 retrial queue with persistent customers in which an arriving customer who finds the service facility busy will always join the orbit. In this case, the generating functions reduce to:

$$R(x,z) = \frac{x}{x - \bar{p} - pz} \frac{p(1-z)[S(x) - S(\bar{p} + pz)]}{S(\bar{p} + pz) - z} R_0(z), \qquad (17)$$

$$Q(z) = \frac{(1-z)S(\bar{p} + pz)}{S(\bar{p} + pz) - z} R_0(z). \qquad (18)$$

where

$$R_0(z) = (1 - pS'(1)) \frac{\prod_{k=0}^{\infty} G(r^k z)}{\prod_{k=0}^{\infty} G(r^k)}, \qquad (19)$$

and

$$G(z) = \frac{S(\bar{p} + pz) - z}{(\bar{p} + pz)S(\bar{p} + pz) - z} \bar{p}. \qquad (20)$$

4.2. The regular Geo/G/1/1 queue.

Setting $\alpha_0 = 0$ implies that an arriving customer who finds the service facility busy will depart immediately from the system. Therefore, the orbit will be always empty. This is represented by setting $r = 1$. This system can be viewed as a regular Geo/G/1/1 queue which has no waiting room. When $\alpha_0 = 0$ and $r = 1$ we get $G(r^k z) = 1$, $R_0(z) = 1/(1 + pS'(1))$ and $P(1, z) = pS'(1)/(1 + pS'(1))$. Therefore, $Q(z)$ reduces to:

$$Q(z) = \frac{1}{1 + pS'(1)} + \frac{pS'(1)}{1 + pS'(1)} z. \qquad (21)$$

This simple model can be analyzed using a different approach. The state transition diagram of this queueing system is shown in fig. 1. Thus, we can write the balance equations as follows:

$$P_0 = \bar{p}P_0 + P_1, \qquad (22)$$

$$P_k = ps_k P_0 + P_{k+1} \qquad , k \ge 1, \qquad (23)$$

where P_k is the steady state probability that the remaining service time of the customer in the server is k time slots. Define the generating function $H(z) = \sum_{k=0}^{\infty} P_k z^k$. Multiplying both sides of (23) by z^k and taking summation from $k = 1$ to ∞ we get:

$$H(z) - P_0 = pP_0 S(z) + \frac{1}{z}[H(z) - zP_1 - P_0]. \qquad (24)$$

Rearranging terms yields:

$$H(z) = \frac{P_0(z + zpS(z) - 1) - P_1 z}{z - 1}. \qquad (25)$$

Using (22) then:

$$H(z) = \frac{P_0(z + zpS(z) - 1) - pP_0 z}{z - 1}.$$

Since H(1) = 1 then it follows that:

$$P_0 = \frac{1}{1 + pS'(1)}. \qquad (26)$$

which represents the probability of an idle system. Consequently:

Pr {one customer in the system}

$$= \sum_{k=1}^{\infty} P_k = 1 - P_0$$

$$= \frac{pS'(1)}{1 + pS'(1)}. \qquad (27)$$

The results in (26, 27) coincide with those that can be extracted from (21).

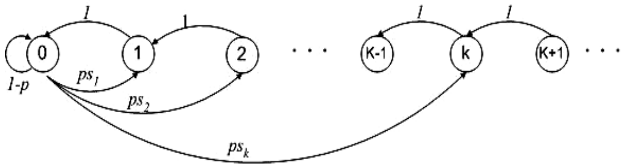

Fig. 1. State transition diagram of the regular Geo/G/1/1 queue.

4.3. The Geo/Geo/1/1 retrial queue with balking customers.

We assume that the service time is geometrically distributed with parameter q, *i.e.,* $s_j = q(\bar{q})^{j-1}$ for $j \geq 1$, where $0 < q < 1$ and $\bar{q} = 1 - q$. Hence, we get:

$$S(z) = \frac{qz}{1-qz},$$

$$G(z) = \frac{q(1-\alpha_0 p(1-z))(1+z)-z}{q(1-\alpha_0 p(1-z))(\bar{p}+pz+z)-z}\bar{p}.$$

This formula together with (1) and (2) represent the joint probability generating function of $P_{i,k}$. This system will be treated numerically in [12].

5. Conclusions

Discrete time queueing systems have received a great interest in recent years because they are used in the modeling and analysis of modern communication systems. In this paper we analyzed a discrete time Geo/G/1/1 retrial queue with balking customers in which an arriving customer that finds the service facility busy will either join the infinite orbit with probability α_0 or leave the system with probability $1-\alpha_0$. Inter-retrial times are assumed to be independent and geometrically distributed. Early arrival scheme is assumed.

We derived analytic formulas for both the joint probability generating function of the remaining service time of the customer currently in the server and the number of customers in the orbit, and the probability generating function of the total number of customers in the system. These generating functions could be used to obtain important performance measures such as average system size and average waiting time. However, it seems hard to invert these generating functions to derive an explicit expression for the system size distribution. To resolve this problem, we present in our next paper [12], a set of recursive formulas for computing the required probabilities.

The present work can be extended in many directions. Only balking customers case was considered. Other types of impatience have to be considered. The present queueing system has no waiting room. The extension of the obtained results to the finite buffer case is under investigation. The present work was directed to obtain the distribution of the system size. A parallel study is needed for analyzing other system characteristics such as waiting time and idle and busy periods.

References

[1] H. Bruneel and B.G. Kim, Discrete-Time Models for Communication Systems Including ATM. Kluwer Academic Publishers, Boston (1993).

[2] H. Takagi, Discrete-Time Systems, Vol. 3 of Queueing Analysis - A Foundation of Performance Evaluation. North-Holland: NY (1993).

[3] M.E. Woodward, Modelling with Discrete-Time Queues. Communication and Computer Networks, Los Alamitos, California: California IEEE Computer Society Press (1994).

[4] I. Atencia and P. Moreno, “Geo/G/1 Retrial Queue with General Retrial Times,” Queueing Systems, Vol. 48, pp. 5 - 21 (2004).

[5] T. Yang and H. Li, “On the Steady-State Queue Size Distribution of the Discrete-Time Geo/G/1 Queue with Repeated Customers,” Queueing Systems, Vol. 21, pp. 199 – 215 (1995).

[6] T. Yang and H. Li, “Steady-State Queue Size Distribution of Discrete-Time PH/Geo/1 Retrial Queues,” Mathematical and Computer Modelling, Vol. 30, pp. 51– 63 (1999).

[7] B.D. Choi and J.W. Kim, “Discrete-Time Geo1, Geo2/G/1 Retrial Queueing

Systems with Two Types of Calls," Computers and Mathematics with Applications, Vol. 33, pp. 79 – 88 (1997).

[8] T. Yang and H. Li, "Geo/G/1 Discrete Time Retrial Queue with Bernoulli Schedule," Eur. J. Oper. Res., Vol. 111 (3), pp. 629 – 649, (1998).

[9] R. Nobel, "A Discrete-Time Retrial Queueing Model with One Server," (China), INFORMS/APS Conference, June (2004).

[10] V. Kulkarni and H. Liang, Frontiers in Queueing (Ed. J.H. Dshalalow), ch. Retrial queues revisited, pp. 19–34. CRC Press, Boca Raton (1997).

[11] E. Gelenbe and G. Pujolle, Introduction to Queueing Networks. John Wiley, New York, second ed. (1997).

[12] A. Aboul-Hassan, S.I. Rabia and A. Kadry, "A Recursive Algorithm for Analyzing a Discrete Time Retrial Queue with Balking Customers and Early Arrival Scheme.", Accepted for Publication in Alexandria Engineering Journal (2005).

[13] A. Aboul-Hassan, S.I. Rabia and A. Kadry, "A Discrete Time Geo/G/1/1 Retrial Queueing System with Balking Customers.", Unpublished work, (2005).

Received August 13, 2005
Accepted October 20, 2005