A DISCRETE-TIME RETRIAL QUEUEING SYSTEM WITH SERVICE UPGRADE

I. Atencia,
I. Fortes and
S. Sánchez
Department of Applied Mathematics
University of Málaga
29071 - Málaga, Spain
Email: iatencia@ctima.uma.es,
ifortes@uma.es, sixto@uma.es

KEYWORDS
Discrete-time model; Retrials, Markov chain; Sojourn time; Stochastic decomposition.

ABSTRACT
This paper considers a discrete-time retrial queueing system with displacements. The arriving customers can opt to go directly to the server obtaining immediately its service or to join the orbit. In the first case, the displaced customer goes directly to the orbit. The arrivals follow a geometrical law and the service times are general.

We study the Markov chain underlying the considered queueing system obtaining the generating function of the number of customers in the orbit and in the system as well as the stationary distribution of the time that a customer spends in the server. We derive the stochastic decomposition law and as an application we give bounds for the proximity between the steady-state distributions for our queueing system and its corresponding standard system.

INTRODUCTION
Queueing systems with retrial queues have been widely used to model many practical problems in telephone switching systems, telecommunication networks and computers competing to gain service from a central processing unit. For a detailed review of the main results and the literature on this topic the reader is referred to (Artalejo and Gómez-Corral 2008). In the context of continuous time systems discrete time Markov chains play an important role. Defined as embedded chains they provide us with techniques to deal with continuous time single servers or even networks when the system is observed only at specific time points, e.g., arrival or departure instances. For an introduction see Kleinrock’s monograph, Volume I, Part III (Kleinrock 1975), and for the application of these methods in performance analysis of computer and communication systems (Kleinrock 1976).

Performance prediction in communication switching queues, job processing in computers, etc., are always influenced by customers behavior, and the provision of this additional information will be useful in upgrading the service. This mechanism is called a synchronized or triggered motion (e.g., Artalejo, 2000; Gelenbe and Label, 1998). For the inverse order discipline, we refer to (Pechinkin and Svischeva 2004; Pechinkin and Shorgin 2008) as well as (Cascone et al. 2011).

Let us note that in our work is considered, with a certain probability, a LCFS discipline, that is, a customer can obtain service immediately after his arrival displacing the customer that is in the server.

Works related to discrete-time systems with server interruptions with or without expulsions and vacations can be found, including those by (Fiems et al. 2002; 2004; Vinck and Bruneel 2006; Morozov et al. 2011) as well as (Atencia and Pechinkin 2013; Atencia et al. 2013a; 2013b; Atencia 2014).

In the early times of queueing theoretical applications in telecommunication systems networks of transmission lines were modeled using brute force decomposition approximations which are still in use and are considered as unavoidable in many situations. First steps to overcome the restriction to decomposition approximations were done around 1950 by coupling classical exponential queues in lines. These were used to model job-shop like production systems, transportation lines, and complex distribution lines with inventories.

DESCRIPTION OF THE QUEUEING SYSTEM
We consider a discrete-time queueing system where the time axis is segmented into a sequence of equal time intervals (called slots). Further, let the time axis be marked by 0, 1, ..., m, ... . It is assumed that all queueing activities (arrivals and departures) take place at the slot boundaries, and therefore they may occur simultaneously.
For mathematical convenience, we assume that a potential departure occurs in the interval \((m^-, m)\) and a potential arrival occurs in the interval \((m, m^+)\); that is, the arrivals occur at the moment immediately after the slot boundaries and the departures occur at the moment immediately before the slot boundaries.

It is assumed that customers arrive according to a geometrical arrival process with rate \(a\), that is, \(a\) is the probability that an arrival occurs in a slot. If an arriving customer finds the server idle, he commences his service immediately and abandons the system after service completion. Otherwise, the arriving customer with probability \(\theta\) displaces the customer that is currently being served. 

It can be shown that \(\bar{1} - \bar{a} = 1 - \theta\) reaches the orbit and \(\bar{a}\) retrials at the same slot when the server is idle, any of them is randomly chosen to be served and the others must return to the orbit.

**THE MARKOV CHAIN**

At time \(m^+\) the system can be described by the process \((C_m, \xi_m, N_m)\) where \(C_m\) denotes the state of the server, 0 or 1 according to whether the server is free or busy and \(N_m\) the number of repeated customers. If \(C_m = 1\), then \(\xi_m\) represents the remaining service time of the customer currently being served.

It can be shown that \(\{(C_m, \xi_m, N_m) : m \in \mathbb{N}\}\) provides a Markovian description of the queuing system under study, whose states space is

\[
\{(0, k) : k \geq 0; (1, i, k) : i \geq 1, k \geq 0\}.
\]

Our objective is to find the stationary distribution

\[
\pi_{0,k} = \lim_{m \rightarrow \infty} P[C_m = 0, N_m = k], k \geq 0
\]

\[
\pi_{1,i,k} = \lim_{m \rightarrow \infty} P[C_m = 1, \xi_m = i; N_m = k], i \geq 1, k \geq 0
\]

of the Markov chain \(\{(C_m, \xi_m, N_m) : m \in \mathbb{N}\}\).

The Kolmogorov equations for the stationary distribution are

\[
\pi_{0,k} = \bar{a} \pi_{0,k} + \bar{a} \delta_{1,1,k}, k \geq 0
\]

\[
\pi_{1,i,k} = a_s \pi_{0,k} + \bar{a} (1 - \rho^{k-1}) \delta_{1,1,k+1} + a_s \pi_{1,i,k+1} + \rho \pi_{1,i+1,k} + (1 - \delta_{1,1,k}) \rho \pi_{1,i,k+1} + \bar{a} \pi_{1,i+1,k} + (1 - \delta_{1,1,k}) \sum_{j=2}^{\infty} a_s \pi_{1,i,j-1}, i \geq 1, k \geq 0
\]

where \(\bar{\alpha} = 1 - \alpha\), \(\bar{\Theta} = 1 - \theta\) and \(\delta_{i,j}\) denotes the Kronecker’s delta.

The normalization condition is

\[
\sum_{k=0}^{\infty} \pi_{0,k} + \sum_{i=1}^{\infty} \sum_{k=0}^{\infty} \pi_{1,i,k} = 1.
\]

To solve Eqs. (1) and (2), we define the following generating functions

\[
\phi_0(z) = \sum_{k=0}^{\infty} \pi_{0,k} z^k, \quad \phi_1(x, z) = \sum_{i=1}^{\infty} \sum_{k=0}^{\infty} \pi_{1,i,k} x^i z^k
\]

and the auxiliary generating function

\[
\phi_{1,i}(z) = \sum_{k=0}^{\infty} \pi_{1,i,k} z^k, i \geq 1.
\]

Multiplying equation (1) and (2) by \(z^k\) and summing over \(k\), these equations become

\[
\phi_0(z) = \bar{a} \phi_0(rz) + \bar{a} \phi_{1,1}(rz)
\]

\[
\phi_{1,i}(z) = (\bar{a} + a \bar{\theta} z) \phi_{1,i+1}(z) + \bar{a} + az - a \theta z^2 s_i \phi_{1,1}(z) + a \theta \phi_{1,i+1}(z) + \frac{\bar{a} + az}{z}, i \geq 1.
\]

By substituting Eq. (3) into Eq. (4), we get

\[
\phi_{1,i}(z) = (\bar{a} + a \bar{\theta} z) \phi_{1,i+1}(z) + \bar{a} + az - a \theta z^2 s_i \phi_{1,1}(z) + a \theta \phi_{1,i+1}(z) - \frac{1 - \bar{a}}{z} a_s \phi_{0}(z), i \geq 1.
\]

Multiplying the above equation by \(x^i\) and summing over \(i\) yields

\[
x - (\bar{a} + a \bar{\theta} z) \phi_1(x, z) =
\]

\[
= \left[\frac{\bar{a} + az - a \theta z^2}{z} S(x) - (\bar{a} + a \bar{\theta} z)\right] \phi_{1,1}(z) + a \theta z S(x) \phi_1(1, z) - \frac{1 - \bar{a}}{z} a_S(x) \phi_0(z)
\]
Putting $x = 1$ in Eq. (5), we have

$$a \varphi_1(1, z) = \frac{a + az + a\theta z}{z} \varphi_1,1(z) - \frac{a}{z} \varphi_0(z)$$

(6)

and substituting (6) into (5), we get

$$x - (\bar{a} + a\theta z) \varphi_1(x, z) = \frac{(\bar{a} + az + a\theta z)S(x) - z(\bar{a} + a\theta z) \varphi_1,1(z) - 1 - \bar{\theta z}}{z} - aS(x)\varphi_0(z).$$

(7)

Inserting $x = \bar{a} + a\theta z$ in the above equation we get

$$\varphi_{1,1}(z) = \frac{a(1 - \bar{\theta}z)(\bar{a} + a\theta z)}{(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)} \varphi_0(z)$$

(8)

and substituting (8) into (7) we finally have

$$\varphi_1(x, z) = \frac{S(x) - S(\bar{a} + a\theta z)}{x - (\bar{a} + a\theta z)} \times \frac{ax(1 - \bar{\theta}z)(\bar{a} + a\theta z)}{(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)} \varphi_0(z)$$

(9)

In order to find $\varphi_0(z)$ we substitute $\varphi_{1,1}(rz)$ in (3), obtaining

$$\varphi_0(z) = G(rz)\varphi_0(rz),$$

(10)

where

$$G(z) = \frac{(1 + \theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)}{\bar{a}(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)} \bar{a}$$

It follows, by using Eq. (10) recursively, that

$$\varphi_0(z) = \varphi_0(1) \prod_{i=1}^{\infty} G(r^i z).$$

(11)

If $z = 1$, we get

$$\varphi_0(0) = \varphi_0(1) \prod_{i=1}^{\infty} G(r^i z).$$

(12)

From the normalization condition, that can be written as

$$\varphi_0(1) + \varphi_1(1, 1) = 1$$

we find the value of $\varphi_0(1)$:

$$\varphi_0(1) = \frac{(1 + \bar{\theta} z)S(\bar{a} + a\theta) - (\bar{a} + a\theta)}{\bar{\theta} S(\bar{a} + a\theta)}.$$  

(13)

The stability condition is given by

$$S(\bar{a} + a\theta) > \frac{\bar{a} + a\theta}{1 + a\theta}.$$  

(14)

The convergence of the infinite product in (11) is established in the following way:

Let us express $G(z)$ as

$$G(z) = 1 + F(z)$$

where

$$F(z) = \frac{[(\bar{a} + a\theta z) - S(\bar{a} + a\theta z)]}{(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)} \bar{a}$$

It can be easily shown that if the stability condition (14) is fulfilled the following inequality holds for $0 \leq z \leq 1$:

$$(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) > z(\bar{a} + a\theta z),$$

Therefore the function $F(z)$ is positive in the interval [0, 1].

The infinite product in (11) can be rewritten as

$$\prod_{i=1}^{\infty} G(r^i z) = \prod_{i=1}^{\infty} [1 + F(r^i z)].$$

(15)

It is well known that the infinite product in (15) converges if and only if the series $\sum_{i=1}^{\infty} F(r^i z)$ is convergent, which is obvious since

$$\lim_{r \to 1} \frac{F(r^i z)}{F(r^i z)} = r < 1.$$  

We can summarize the above results in the following theorem:

Theorem 1 If the stability condition (14) is satisfied, the stationary distribution of the Markov chain \{$(C_m, \xi_m, N_m)$ : $m \in \mathbb{N}$\} has the generating functions:

$$\varphi_0(z) = \varphi_0(1) \prod_{i=1}^{\infty} G(r^i z)$$

$$\varphi_1(x, z) = \frac{S(x) - S(\bar{a} + a\theta z)}{x - (\bar{a} + a\theta z)} \times \frac{ax(1 - \bar{\theta}z)(\bar{a} + a\theta z)}{(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)} \varphi_0(z)$$

where

$$G(z) = \frac{(1 + \theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)}{\bar{a}(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)} \bar{a}$$

Corollary 1

1. The probability generating function of the orbit size (i.e. of the variable $N$) is given by

$$\Psi(z) = \varphi_0(z) + \varphi_1(1, z) = \frac{\theta z S(\bar{a} + a\theta z) + (\bar{a} + a\theta z)(1 - z)}{(\bar{a} + az + a\theta z)S(\bar{a} + a\theta z) - z(\bar{a} + a\theta z)} \varphi_0(z)$$
2. The probability generating function of the system size (i.e. of the variable $L$) is given by

$$\Phi(z) = \varphi_0^z + z \varphi_1(z) = \frac{(\bar{a} + az)(1 - \bar{\theta}z)\bar{S}(\bar{a} + a\bar{\theta}z) - z(\bar{a} + a\bar{\theta}z)}{\theta S(\bar{a} + a\bar{\theta})[(1 + \bar{a}\bar{\theta})\bar{S}(\bar{a} + a\bar{\theta}) - (\bar{a} + a\bar{\theta})]} \varphi_0(z)$$

Corollary 2

1. The mean orbit size is given by

$$E[N] = \Phi'(1) = \frac{1}{\theta S(\bar{a} + a\bar{\theta})[(1 + \bar{a}\bar{\theta})\bar{S}(\bar{a} + a\bar{\theta}) - (\bar{a} + a\bar{\theta})]} \times$$

$$\left[ (\bar{a} + a\bar{\theta})^2 + |a\bar{\theta} - (\bar{a} + a\bar{\theta})(1 + a\bar{\theta})| \bar{S}(\bar{a} + a\bar{\theta}) + \right.$$ 

$$\left. + \bar{a}\theta \bar{S}(\bar{a} + a\bar{\theta}) - a\theta(\bar{a} + a\bar{\theta})\bar{S}(\bar{a} + a\bar{\theta}) + \sum_{i=1}^{\infty} \frac{G(r^i)}{G(r^i)} r^i \right]$$

2. The mean system size is given by

$$E[L] = \Phi'(1) = E[N] + \rho$$

where $\rho = 1 - \varphi_0(1)$.

3. The mean time a customer spends in the system is given by

$$W = \frac{E[L]}{a}$$

• Special case: When $r = 0$, $G(r^i) = 1$, $\forall i \geq 1$ and $\varphi_0(z) = \varphi_0(1)$ we have

$$\Phi(z) = \frac{(\bar{a} + az)(1 - \bar{\theta}z)\bar{S}(\bar{a} + a\bar{\theta}z) - z(\bar{a} + a\bar{\theta}z)}{(\bar{a} + az + a\bar{\theta}z)\bar{S}(\bar{a} + a\bar{\theta}z) - z(\bar{a} + a\bar{\theta}z)} \varphi_0(z)$$

which is the probability generating function of the system size in the model $Geo/G/1/\infty$ with optional LCFS discipline in (Atencia and Pechinkin 2013).

**SOJOURN TIME OF A CUSTOMER IN THE SERVER**

In this section we are going to find the distribution of the time that a customer spends in the server. Since the service can be interrupted if a new customer arrives to the system, the sojourn time in the server consists of separate time intervals.

We will denote by $b_k$ the probability that the sojourn time of a customer in the server lasts exactly $k$ slots. The distribution $\{b_k, k \geq 0\}$ is given by

$$b_0 = 0$$

$$b_k = (\bar{a} + a\bar{\theta})^{k-1} + \sum_{i=1}^{k} a\theta(\bar{a} + a\bar{\theta})^{-1} S_{i+1} b_{k-i}, \quad k \geq 1.$$ 

The corresponding GF is given by

$$b(x) = \frac{[1 - (\bar{a} + a\bar{\theta})x]\bar{S}(\bar{a} + a\bar{\theta}x)}{(1 - x)(\bar{a} + a\bar{\theta}) + a\theta S'(\bar{a} + a\bar{\theta})x}$$

The mean time of a customer in the server is given by

$$\bar{b} = b'(1) = \frac{(\bar{a} + a\bar{\theta})[1 - S'(\bar{a} + a\bar{\theta})]}{a\theta S'(\bar{a} + a\bar{\theta})}$$

Let us note that the stability condition (14) can be written as

$$\rho = ab < 1$$

**STOCHASTIC DECOMPOSITION LAW**

The stochastic decomposition law has been analyzed extensively for the queueing systems with server vacations. This property allows to study the system by considering separately the distribution of the system size without vacations and the additional system size due to vacations. Specifically, the stochastic decomposition law establishes that the number of customers in the system can be decomposed as sum of two independent random variables: one being the number of customers in the corresponding standard system and the other random variable can have different probabilistic meanings depending on how the vacation discipline has been defined.

In the context of our system, we present the following stochastic decomposition of the system size distribution

$$\Phi(z) = \frac{(1 - \rho)(\bar{a} + az)(1 - \bar{\theta}z)\bar{S}(\bar{a} + a\bar{\theta}z) - z(\bar{a} + a\bar{\theta}z)}{(\bar{a} + az + a\bar{\theta}z)\bar{S}(\bar{a} + a\bar{\theta}z) - z(\bar{a} + a\bar{\theta}z)} \varphi_0(z) \Phi_0(z)$$

where the first fraction corresponds to the probability generating function of the number of customers in the $Geo/G/1/\infty$ queueing system with LCFS discipline and the second fraction is the probability generating function of the number of blocked customers given that the server is idle. This result can be summed up in the next theorem.

**Theorem 2** The random variable ”system size” ($L$) can be decomposed as the sum of two independent random variables, one of which is the number of customers in the $Geo/G/1/\infty$ queueing system with optional LCFS discipline ($L_0$) and the other is the number of repeated customers given that the server is idle ($M$). That is $L = L_0 + M$.

As a consequence of the stochastic decomposition law, we provide a measure of the proximity between the steady-state distribution of the $Geo/G/1/\infty$ with optional LCFS discipline and the ours.
Theorem 3. The following inequalities hold

\[
2(1 - \rho - \pi_{0,0}) \leq \sum_{j=0}^{\infty} |P[L = j] - P[L_0 = j]| \leq 2 \frac{1 - \rho - \pi_{0,0}}{1 - \rho}
\]

Proof 1. The proof is based on the decomposition’s law property of the total number of customers in the system. In terms of distributions, this property means that \(P[L = j]\) is a convolution of the distributions \(P[L_0 = j]\) and \(P[M = j]\), that is,

\[
P[L = j] = \sum_{k=0}^{j} P[L_0 = k]P[M = j - k].
\]

From the above equation we obtain

\[
|P[L = j] - P[L_0 = j]| \leq (1 - \delta_{0,j}) \sum_{k=0}^{j-1} P[L_0 = k]P[M = j - k] + P[L_0 = j](1 - P[M = 0]) = P[L = j] + P[L_0 = j](1 - 2P[M = 0]).
\]

Therefore, summing all over the states we obtain

\[
\sum_{j=0}^{\infty} |P[L = j] - P[L_0 = j]| \leq \sum_{j=0}^{\infty} P[L = j] + (1 - 2P[M = 0]) \sum_{j=0}^{\infty} P[L_0 = j] = 2(1 - P[M = 0]) = 2 \left(1 - \frac{\pi_{0,0}}{1 - \rho}\right) = 2 \frac{1 - \rho - \pi_{0,0}}{1 - \rho}
\]

Now, using the inequality \(|a - b| \geq a - b\), we obtain a lower bound

\[
\sum_{j=0}^{\infty} |P[L = j] - P[L_0 = j]| \geq \sum_{j=1}^{\infty} (P[L = j] - P[L_0 = j]) = 2(1 - \rho) \left(1 - \frac{\pi_{0,0}}{1 - \rho}\right) = 2 \frac{1 - \rho - \pi_{0,0}}{1 - \rho}
\]

Finally, let us observe that the distance

\[
\sum_{j=0}^{\infty} |P[L = j] - P[L_0 = j]|
\]

between the distributions of the variables \(L\) and \(L_0\) diminishes when \(r\) goes toward 0.

Let us note that the interest of the former theorem is to provide lower and upper estimates between these distributions.

**NUMERICAL RESULTS**

In this section, we present some numerical examples that describe several performance characteristics of the queueing model under study. Of course in all the below examples, the parametric values are chosen so as to satisfy the stability condition.

We will concentrate our attention on two important performance descriptors: the probability that the system is empty and the mean number of customers in the system where all of them will be plotted versus \(\theta\) which is the most characteristic parameter of the system.

We consider the case that the service time distribution is given by \(S(x) = \frac{1}{2}x + \frac{1}{6}x^2\).

In the Figure 1 the probability that the system is empty is plotted against the parameter \(\theta\). As we expected, \(\phi_0(1)\) is decreasing as function of \(\theta\). The curves which in decreasing order correspond to higher arrival rates show, as intuition tells us, that for increasing values of the parameter \(a\) the probability of an empty system decreases.

![Figure 1: The effect of \(\theta\) on \(\phi_0(1)\)](image)

The effect of the parameter \(\theta\) on \(E[L]\) is showed in Figure 2. We have presented three curves which correspond to \(a = 0.3, 0.35, 0.4\). The curves show, as it to be expected, that \(E[L]\) increases with increasing values of \(\theta\) and that they increase more rapidly as the arrival rate \(a\) increases.

![Figure 2: The mean system size versus \(\theta\)](image)
CONCLUSIONS AND FUTURE RESEARCHES

In this paper we analyze a discrete-time retrial queueing system in which an arriving customer can decide, with a certain probability, to go directly to the server displacing the customer that is currently being served or to join the orbit. We analyze the principal characteristics of the model, such as, the size of the system, the sojourn time of a customer in the server and the stochastic decomposition law. As an application of the stochastic decomposition we study the proximity between the steady-state distributions for the Geo/G/1/∞ with optional service and our queueing system.

ACKNOWLEDGEMENTS

The research was supported by the national Spanish project TIN2012-39353-C04-01.

REFERENCES


AUTHORS BIOGRAPHIES

IVAN ATENCIA graduated in 1997 and obtained his Ph.D. in Mathematics in 2000 from Málaga University, Spain. During this period he collaborated with the Department of Probability Theory and Mathematical Statistics of The Peoples’ Friendship University of Russia, Moscow. At present, he is Associate Professor in the Department of Applied Mathematics at Málaga University and actively collaborates with the Institute of Informatics Problems of RAS, Russia (Moscow). His research interests include queueing theory and stochastic modeling of communication systems. His e-mail address is: iatencia@ctima.uma.es

INMACULADA FORTES graduated in 1996 and obtained her Ph.D. in Mathematics in 2002 from Málaga University, Spain. She is Associate Professor in the Department of Applied Mathematics at Málaga University. Her research interests are queueing theory, stochastic modelling of communication systems as well as data mining and mathematical logic. Her e-mail address is: ifortes@uma.es

SIXTO SÁNCHEZ graduated in 1992 and obtained his Ph.D. in Mathematics in 2002 from Málaga University, Spain. He is Associate Professor in the Department of Applied Mathematics at Málaga University. His research interests are queueing theory, stochastic modelling of communication systems as well as mathematical logic. His e-mail address is: sixto@uma.es