A Single Server Retrial Queue with Different Types of Server Interruptions

Tewfik Kernane

A Single Server Retrial Queue with Different Types of Server Interruptions

Tewfik Kernane
Department of Mathematics
Faculty of Sciences
King Khalid University
Abha, Saudi Arabia
tkernane@gmail.com

February 11, 2009

Abstract

We consider a single server retrial queue with the server subject to interruptions and classical retrial policy for the access from the orbit to the server. We analyze the equilibrium distribution of the system and obtain the generating functions of the limiting distribution.

1 Introduction

Queueing systems with retrials of the attempts are characterized by the fact that an arrival customer who finds the server occupied is obliged to join a group of blocked customers, called orbit, and reapply after random intervals of time to obtain the service. These systems are useful in the stochastic modeling of much of situations in practice. We can find them in aviation, where a plane which finds the runway occupied remakes its attempt of landing later and we say in this case that it is in orbit. In Telephone systems where a telephone subscriber who obtains a busy signal repeats the call until the required connection is made. In data processing, we find them in protocol of access CSMA/CD. They appear in the modeling of the systems of maintenance and the problems of repair among others. For details on these models, see the book of Falin and Templeton [4] or the recent book by Artalejo and Gomez-Corral [2].

We study in this article single server retrial queues with various types of interruptions of the server. From a practical point of view, it is more realistic to consider queues with repetitions of calls and the server exposed to random interruptions. Queueing models with interruptions of service proved to be a useful abstraction in the situations where a server is shared by multiple queues, or when the server is subject to breakdowns. Such systems were studied in the literature by many authors. Fiems et *al.* [6] considered an M/G/1 queue with various

types of interruptions of the server and our work is a generalization to the case of retrial queues. White and Christie [17] were the first to study queues with interruptions of service by considering a queueing system with exponentially distributed interruptions. Times of interruptions and services generally distributed are considered by Avi-Itzhak and Naor [3] and Thiruvengadam [14]. Other generalisations were considered in the literature (phase-type [5], approximate analysis [15], Markov-modulated environment [13] and [10], processor sharing [11]). Gaver [8] considers the case where the service is repeated or repeated and begin again after the interruption.

We consider in this paper a single server retrial queue with server interruptions and the classical retrial policy where each customer in orbit conduct its own attempts to get served independently of other customers present in the orbit. We can then assume that the probability of a retrial during the time interval $(t, t+dt)$, given that j customers were in orbit at time t, is $j\theta dt + \circ(dt)$. Kulkarni and Choi [9] studied a single server linear retrial queue with server subject to breakdowns and repairs and they obtained the generating functions of the limiting distribution and performance characteristics. Artalejo [1] obtained sufficient conditions for ergodicity of multiserver retrial queues with breakdowns and a recursive algorithm to compute the steady-state probabilities for the M/G/1 linear retrial queue with breakdowns. The detailed analysis for reliability of retrial queues with linear retrial policy was given by Wang, Cao and Li [16].

The remainder of paper is organized as follows. In the following section, we describe the model and give the necessary and sufficient conditions so that the system is stable. In section 3, we analyze the equilibrium distribution of the system in study.

2 Model Description

Consider a single server queueing system in which customers arrive in accordance with a Poisson process with arrival rate λ. If at the instant of arrival the customer finds the server free, it takes its service and leaves the system. Otherwise, if the server is busy or in interruption, the arriving customer joins an unlimited queue called orbit and makes retrials for getting served after random time intervals. We consider the classical policy where each customer in orbit conducts his own attempts to obtain service independently from the other customers present in the orbit. We can then assume that the probability of a retrial during the time interval $(t, t+dt)$, given that j customers were in orbit at time t, is $j\theta dt + \circ(dt)$. Service times constitute a series of independent and identically distributed (i.i.d.) random variables with common distribution function $B(t)$, density function $b(t)$, and corresponding Laplace–Stieltjes transform (LST) $\beta(s)$ and finite first two moments $\beta_k = (-1)^k \beta^{(k)}(0)$, $k = 1, 2$. Interruptions of the service may occur according to a Poisson process with rate ν if the server is busy and this type of interruption can be disruptive with probability p_d (or rate $\nu_d = p_d\nu$) or non-disruptive with probability $p_n = 1 - p_d$ (or rate $\nu_n = p_n\nu$). In the case of a disruptive interruption the customer being served repeats his service at the end of the interruption, in the other type the customer continues his stopped service. If

the server is idle, another type of interruptions may occur according to a Poisson process with rate ν_i. We call this type *idle interruption*. The lengths of the consecutive disruptive (non-disruptive, idle time) interruptions constitute a series of i.i.d. positive random variables with distribution function $B_d(t)$ ($B_n(t)$, $B_i(t)$), density function $b_d(t)$ ($b_n(t)$, $b_i(t)$), corresponding Laplace–Stieltjes transform (LST) $\beta_d(s)$ ($\beta_n(s)$, $\beta_i(s)$) and finite first two moments β_k^d, (β_k^n, β_k^i) $k = 1, 2$.

Denote by $N(t)$ the number of customers in orbit at time t. Let $C(t)$ be the state of the server at time t : $C(t) = F$ if the server is free (and functions normally), $C(t) = S$ if the server is busy (and functions normally), $C(t) = D$ if the server is on a disruptive interruption, $C(t) = N$ if the server is on a non-disruptive interruption, $C(t) = I$ if the server is taking an idle interruption. We introduce the random variables $\xi(t)$, $\xi_D(t)$, $\xi_N(t)$ and $\xi_I(t)$ defined as follows. If $C(t) = S$ then $\xi(t)$ represents the elapsed service time at time t; if $C(t) = D$, then $\xi_D(t)$ represents the elapsed disruptive interruption time at t; if $C(t) = N$, then $\xi_N(t)$ represents the elapsed non-disruptive interruption time at t; and if $C(t) = I$ then $\xi_I(t)$ is the elapsed idle interruption time at t.

3 Stability Analysis

We first study the condition for the system to be stable. The following theorem provides the necessary and sufficient stability condition.

Theorem 1 *The system with classical retrial policy and interruptions is stable if and only if the following condition is fulfilled*

$$\frac{\lambda\left(1-\beta(\nu_d)\right)}{\nu_d\beta(\nu_d)}\left(1+\nu_d\beta_1^d+\nu_n\beta_1^n\right)<1. \tag{1}$$

Proof. Let $\{s_n;\, n\in\mathbb{N}\}$ be the sequence of epochs of service completion time. We consider the process $Y_n = (N(s_n+), C(s_n+))$ embedded immediately after time s_n. It is readily to see that $\{Y_n;\, n\in\mathbb{N}\}$ is an irreducible aperiodic Markov chain. To determine the stability of the system it remains to prove that $\{Y_n;\, n\in\mathbb{N}\}$ is ergodic under the suitable stability condition. Let us first consider the generalized service time $\widetilde{S}$ of a customer which includes, in addition to the original service time S of the customer, possible interruption times during the service period of the customer. Fiems et *al.* [6] showed that the generalized service time has the Laplace transform

$$\widetilde{\beta}(s)=\frac{\left[s+\nu-\nu_n\beta_n(s)\right]\beta\left(s+\nu-\nu_n\beta_n(s)\right)}{\left[s+\nu-\nu_n\beta_n(s)\right]-\nu_d\beta_d(s)\left[1-\beta\left(s+\nu-\nu_n\beta_n(s)\right)\right]},$$

hence its expected value is given by

$$E\widetilde{S}=-\widetilde{\beta}'(0)=\frac{\left(1-\beta(\nu_d)\right)}{\nu_d\beta(\nu_d)}\left(1+\nu_d\beta_1^d+\nu_n\beta_1^n\right).$$

For the sufficiency, we shall use Foster's criterion, which states that a Markov chain $\{Y_n;\, n\in\mathbb{N}\}$ is ergodic if there exists a nonnegative function $f(k)$, $k\in\mathbb{N}$,

and $\delta > 0$ such that for all $k \neq 0$ the mean drift

$$\chi_k = E\left[f(Y_{n+1}) - f(Y_n) \mid Y_n = k\right], \tag{2}$$

satisfies $\chi_k \leq -\delta$ and $E\left[f(Y_{n+1}) \mid Y_n = 0\right] < \infty$. If we choose $f(k) = k$ we obtain

$$E\left[f(Y_{n+1}) \mid Y_n = 0\right] = \lambda E\widetilde{S} = \frac{\lambda\left(1 - \beta(\nu_d)\right)}{\nu_d \beta(\nu_d)}\left(1 + \nu_d \beta_1^d + \nu_n \beta_1^n\right) < \infty,$$

and we can easily check that

$$\chi_k = \lambda E\widetilde{S} - 1 = \left[\lambda\left(1 - \beta(\nu_d)\right)/\nu_d \beta(\nu_d)\right]\left(1 + \nu_d \beta_1^d + \nu_n \beta_1^n\right) - 1.$$

If we set

$$\delta = 1 - \frac{\lambda\left(1 - \beta(\nu_d)\right)}{\nu_d \beta(\nu_d)}\left(1 + \nu_d \beta_1^d + \nu_n \beta_1^n\right)$$

then the condition (1) is sufficient for ergodicity.
To prove that the condition (1) is necessary, we use theorem 1 of Sennot et *al.* [12] which states that if the Markov chain $\{Y_n;\, n \in \mathbb{N}\}$ satisfies Kaplan's condition, namely $\chi_k < \infty$ for all $k \geq 0$ and there is an k_0 such that $\chi_k \geq 0$ for $k \geq k_0$, then $\{Y_n;\, n \in \mathbb{N}\}$ is not ergodic. Indeed, if

$$\frac{\lambda\left(1 - \beta(\nu_d)\right)}{\nu_d \beta(\nu_d)}\left(1 + \nu_d \beta_1^d + \nu_n \beta_1^n\right) \geq 1$$

then for $f(k) = k$, there is a k_0 such that $p_{ij} = 0$ for $j < i - k_0$ and $i > 0$, where $P = (p_{ij})$ is the one-step transition matrix associated to $\{Y_n;\, n \in \mathbb{N}\}$.
The stability of the system follows from Burke's theorem (see Cooper [7] p187) since the input flow is a Poisson process. ■

4 Steady-state analysis

We investigate in this section the steady-state distribution of the system. Define the functions $\mu(x)$, $\mu_D(x)$, $\mu_N(y)$ and $\mu_I(x)$ as the conditional completion rates for service, disruptive interruption, non-disruptive interruption and idle interruption, respectively, i.e., $\mu(x) = b(x)/\left(1 - B(x)\right)$, $\mu_D(x) = b_d(x)/\left(1 - B_d(x)\right)$, $\mu_N(x) = b_n(x)/\left(1 - B_n(x)\right)$ and $\mu_I(x) = b_i(x)/\left(1 - B_i(x)\right)$.

We now introduce the following set of probabilities for $j \geq 0$:

$$\begin{aligned}
p_{F,j}(t) &= P\left\{N(t) = j,\ C(t) = F\right\}, \\
p_{B,j}(t,x)dx &= P\left\{N(t) = j,\ C(t) = S,\ x \leq \xi(t) < x + dx\right\}, \\
p_{D,j}(t,x)dx &= P\left\{N(t) = j,\ C(t) = D,\ x \leq \xi_D(t) < x + dx\right\}, \\
p_{N,j}(t,x,y)dy &= P\left\{N(t) = j,\ C(t) = N,\ \xi(t) = x,\ y \leq \xi_N(t) < y + dy\right\}, \\
p_{I,j}(t,x)dx &= P\left\{N(t) = j,\ C(t) = I,\ x \leq \xi_I(t) < x + dx\right\}.
\end{aligned}$$

where $t \geq 0$ and $x, y \geq 0$.

The usual arguments lead to the differential difference equations by letting $t \to +\infty$

$$(\lambda + j\theta + \nu_i)\, p_{F,j} = \int_0^{+\infty} \mu(x) p_{B,j}(x) dx + \int_0^{+\infty} \mu_I(x) p_{I,j}(x) dx, \quad (3)$$

$$\left(\frac{\partial}{\partial x} + \lambda + \nu + \mu(x)\right) p_{B,j}(x) = (1 - \delta_{0j})\, \lambda p_{B,j-1}(x) + \int_0^{+\infty} \mu_N(y) p_{N,j}(x,y) dy, \quad (4)$$

$$\left(\frac{\partial}{\partial x} + \lambda + \mu_D(x)\right) p_{D,j}(x) = (1 - \delta_{0j})\, \lambda p_{D,j-1}(x), \quad (5)$$

$$\left(\frac{\partial}{\partial y} + \lambda + \mu_N(y)\right) p_{N,j}(x,y) = (1 - \delta_{0j})\, \lambda p_{N,j-1}(x,y), \quad (6)$$

$$\left(\frac{\partial}{\partial x} + \lambda + \mu_I(x)\right) p_{I,j}(x) = (1 - \delta_{0j})\, \lambda p_{I,j-1}(x). \quad (7)$$

With boundary conditions

$$p_{B,j}(0) = (j+1)\, \theta p_{F,j+1} + \lambda p_{F,j} + \int_0^{+\infty} \mu_D(x) p_{D,j}(x) dx, \quad (8)$$

$$p_{D,j}(0) = \nu_d \int_0^{+\infty} p_{B,j}(x) dx, \quad (9)$$

$$p_{N,j}(x,0) = \nu_n p_{B,j}(x), \quad (10)$$

$$p_{I,j}(0) = \nu_i p_{F,j}. \quad (11)$$

The normalising equation is

$$\sum_{j=0}^{+\infty} p_{F,j} + \sum_{j=0}^{+\infty} \int_0^{+\infty} p_{B,j}(x) dx + \sum_{j=0}^{+\infty} \int_0^{+\infty} p_{D,j}(x) dx + \sum_{j=0}^{+\infty} \int_0^{+\infty} \int_0^{+\infty} p_{N,j}(x,y) dx dy$$

$$+ \sum_{j=0}^{+\infty} \int_0^{+\infty} p_{I,j}(x) dx = 1$$

Define the generating functions

$$\begin{aligned}
P_F(z) &= \sum_{j=0}^{\infty} p_{F,j}\ z^j, \\
P_B(x,z) &= \sum_{j=0}^{\infty} p_{B,j}(x)\ z^j, \\
P_D(x,z) &= \sum_{j=0}^{\infty} p_{D,j}(x)\ z^j, \\
P_N(x,y,z) &= \sum_{j=0}^{\infty} p_{N,j}(x,y)\ z^j, \\
P_I(x,z) &= \sum_{j=0}^{\infty} p_{I,j}(x)\ z^j,
\end{aligned}$$

for $|z| \leq 1$.
We introduce $h(z) = [\nu - \nu_n \beta_n (\lambda - \lambda z) + \lambda - \lambda z]$ and
$\chi(z) = h(z) - \nu_d \beta_d (\lambda - \lambda z)(1 - \beta(h(z)))$ to simplify notation.

We have the following theorem

Theorem 2 *In steady state, the joint distribution of the server state and queue length is given by*

$$\begin{aligned}
P_F(z) &= P_F(1) \exp\left\{ \int_1^z \Psi(u) du \right\}, \\
P_B(x,z) &= P_B(0,z)(1 - B(x)) \exp\{-h(z)x\}, \\
P_D(x,z) &= P_B(0,z) \frac{\nu_d (1 - \beta(h(z)))}{h(z)} (1 - B_d(x)) \exp\{-(\lambda - \lambda z)x\}, \\
P_N(x,y,z) &= P_B(0,z)\nu_n (1 - B_n(y))(1 - B(x)) \exp\{-h(z)x\} \exp[-(\lambda - \lambda z)y], \\
P_I(x,z) &= \nu_i P_F(z)(1 - B_i(x)) \exp[-(\lambda - \lambda z)x],
\end{aligned}$$

where

$$\begin{aligned}
P_F(1) &= \frac{\nu_d^2 \beta(\nu_d)(1 - \rho)}{\left(1 + \nu_i \beta_1^i\right)\left[\nu_d^2 \beta(\nu_d) + \lambda \nu_n \beta_1^n (1 - \beta(\nu_d))^2 - \lambda \nu_n \beta_1^n \nu_d (1 - \beta(\nu_d))\right]}, \\
\Psi(z) &= \frac{\lambda h(z) \beta(h(z)) - [\lambda + \nu_i (1 - \beta_i(\lambda - \lambda z))] \chi(z)}{\theta (z\chi(z) - h(z)\beta(h(z)))}, \\
P_B(0,z) &= \frac{h(z)(\lambda + \theta \Psi(z))}{\chi(z)} P_F(z).
\end{aligned}$$

References

[1] J. R. Artalejo, New results in retrial queueing systems with breakdown of the servers, *Statistica Neerlandica* **48**, 23-36, (1994).

[2] J. R. Artalejo and A. Gomez-Corral, *Retrial Queueing Systems: A Computational Approach.* Springer-Verlag, 2008.

[3] B. Avi-Itzhak, P. Naor, Some queuing problems with the service station subject to breakdown, *Operations Research* **11** (3) (1963) 303–319.

[4] G.I. Falin and J.G.C. Templeton, *Retrial Queues*, Chapman & Hall, London, (1997).

[5] A. Federgruen, L. Green, Queueing systems with service interruptions, *Operations Research* **34** (5) (1986) 752–768.

[6] Fiems, D., Maertens, T., Bruneel, H., Queueing systems with different types of server interruptions, *European Journal of Operational Research* 188 (2008) 838–845.

[7] R. Cooper, *Introduction to Queueing Theory*, Elsevier North-Holland Inc., New York, 1981.

[8] D. Gaver Jr., A waiting line with interrupted service, including priorities, *Journal of the Royal Statistical Society* **B** 24 (1962) 73–90.

[9] V. G. Kulkarni and B. D. Choi, Retrial queues subject to breakdowns and repairs, *Queueing Systems* **7**, 191-208, (1990).

[10] H. Masuyama, T. Takine, Stationary queue length in a FIFO single server queue with service interruptions and multiple batch Markovian arrival streams, *Journal of the Operations Research Society of Japan* **46** (3) (2003) 319–341.

[11] R. Nunez Queija, Sojourn times in a processor sharing queue with service interruptions, *Queueing Systems* **34** (1–4) (2000) 351–386.

[12] L. I. Sennot, P. A. Humblet and R. L. Tweedie, Mean drifts and the non-ergodicity of Markov chains, *Operations Research* **31**, 783-789, (1983).

[13] T. Takine, B. Sengupta, A single server queue with service interruptions, *Queueing Systems* **26** (1997) 285–300.

[14] K. Thiruvengadam, Queuing with breakdowns, *Operations Research* **11** (1) (1963) 62–71.

[15] N. Van Dijk, Simple bounds for queueing systems with breakdowns, *Performance Evaluation* **8** (2) (1988) 117–128.

[16] J. Wang, J. Cao and Q. Li, Reliability analysis of the retrial queue with server breakdowns and repairs, *Queueing Systems* **38**, 363-380, (2001).

[17] H. White, L. Christie, Queuing with preemptive priorities or with breakdown, *Operations Research* **6** (1) (1958) 79–95.