A QUEUEING SYSTEM WITH CONSTANT REPEATED ATTEMPTS AND BERNOULLI SCHEDULE

I. Atencia[1], Department of Applied Mathematics, E.T.S.I. de Informática, University of Málaga, Málaga, Spain
G. Bouza[2], Department of Applied Mathematics, Facultad de Matemática y Computación, University of Havana, Cuba
R. Rico, University of Málaga, Málaga, Spain

ABSTRACT
We consider a single-service queueing system with a waiting room of infinite capacity. Customers arrive according to a Poison stream with rate $\lambda > 0$. A customer who finds the server occupied at the time of arrival joins with probability p a retrial group, that will be called orbit, and with complementary probability q a waiting room in order to be served. Service times are general and retrial times are inversely proportional to the number of customers in the orbit. We derive the stationary distribution of the embedded Markov chain and also the joint generating function of the number of customers of both groups in the steady state regime. The results agree with known results for special cases.

Key words: Retrial Systems, Ergodicity, Embedded Markov chain.

MSC: 90B22.

RESUMEN
Consideremos un sistema de cola con un solo servidor y con una órbita y una sala de espera, ambas de capacidad infinita, asociadas. Los tiempos de llegada entre dos clientes consecutivos siguen una distribución Poison de razón $\lambda > 0$. Un cliente que encuentre el servidor libre al llegar al sistema empieza a servirse instantáneamente, si el servidor está ocupado, se incorpora al grupo de reintentos (órbita) con probabilidad p y con probabilidad q = 1 - p, a la sala de espera, para ser servidos. Los tiempos de servicios son generales y los tiempos de reintentos, inversamente proporcionales a la cantidad de clientes en la órbita. En este trabajo, derivamos la distribución estacionaria para la cadena inducida de Markov chain y también la función generadora, conjunta, del número de clientes en ambos grupos en el estado estacionario. Se comprueba además que los resultados constituyen generalizaciones de casos especiales.

1. INTRODUCTION

Queueing systems with repeated attempts occur frequently in practice, for example in computer and communication systems. The main characteristic of a single-server queue with repeated attempts is that a customer who finds the server occupied upon arrival must leave the system and will re-initiate his request after some random time. This kind of systems has been widely studied, see Yang **et al**. (1987).

The purpose of the present work is to study a retrial queueing system with a single-server in which the input flow of primary arrivals follows a Poisson stream of rate $\lambda > 0$. An arriving customer receives immediate service if it finds the server idle; otherwise he joins, with probabilities q = 1 - p and p respectively, a buffer or a retrial group, the orbit, both of infinity capacity. The first customer in the buffer begins its service as soon as the server is free. If a customer in the orbit make an attempt for service and the server is busy, it will come back to the orbit and retry for service at a later random time. Consequently it will be admitted for service only if the priority queue is empty. The retrial time, that is, the time between two consecutive attempts by the same customer is exponentially distributed with rate $\frac{\alpha}{n}$ where n is the number of customers in the orbit. This constant policy was introduced by Fayolle (1986). Service times are general with common distribution function B(x). The Laplace- Stieltjes transform of B(x) is denoted by $\beta(s)$, for $s \geq 0$, and its corresponding n^{th}-moments are denoted by β_n, $n \in \mathbb{N}$. In Choi **et al**. (1990) is studied a queueing system similar to ours but with retrial times exponentially distributed with rate proportional to the numbers of customers in the orbit. At an arbitrary time t, the system can be described by the process:

$$X(t) = (C(t), N_1(t), N_2(t)), \xi(t))$$

E-mail: [1]iatencia@ctima.uma.es
[2]gema@matcom.uh.cu

where C(t) denotes the state of the server, 0 or 1, according to whether the server is busy or occupied, $N_1(t)$ and $N_2(t)$ are the number of customers in the priority queue and in the orbit respectively, and if C(t) = 1, $\xi(t)$ denotes the elapsed time of the customer currently being served.

The paper is organized as follows. In the next Section we study the embedded Markov chain finding out a necessary and sufficient condition for the ergodicity of the chain. The analysis of this queueing system in steady state using the methods of supplementary variables and generating functions is given in Section 3. Finally, in Section 4 we provide some numerical results.

2. EMBEDDED MARKOV CHAIN

Let τ_l be the time at which the l^{th} served customer leaves the system, $N_{1,l} = N_1(\tau_l^-)$ and $N_{2,l} = N_2(\tau_l^-)$, the number of customers in the priority queue and in the orbit just before the time τ_l. For $N_{1,l}$ and $N_{2,l}$, we have the following fundamental recursive equations:

$$N_{1,l} = \begin{cases} N_{1,l-1} - 1 + \nu_{1,l} & \text{if} \quad N_{1,l-1} \geq 1 \\ \nu_{1,l} & \text{if} \quad N_{1,l} - 1 = 0 \end{cases} \tag{2.1}$$

$$N_{2,l} = \begin{cases} N_{2,l-1} + \nu_{2,l} & \text{if} \quad N_{1,l-1} \geq 1 \\ N_{2,l-1} - B_l + \nu_{2,l} & \text{if} \quad N_{1,l-1} = 0 \end{cases} \tag{2.2}$$

where $\nu_{1,l}$ and $\nu_{2,l}$ are the number of customers which arrive in the priority queue and in the orbit respectively during the service time of the 1^{th} customer, and $B_l = 1$ if the l^{th} served customer proceeds from the orbit and $B_l = 0$ otherwise.

We will denote by

$$k_{m,n} = \int_0^\infty \frac{(\lambda xp)^n}{n!} \frac{(\lambda xq)^m}{m!} e^{-\lambda x} dB(x)$$

the joint distribution of the number of customers that arrives to the priority queue and the orbit during a service time.

It is easy to see that

$$G(z_1, z_2) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} k_{m,n} z_1^m z_2^n = \beta(\lambda - \lambda p z_2 - \lambda q z_1)$$

The sequence of random vectors $Y_l = (N_{1,l}, N_{2,l})$, $l \geq 0$ forms a Markov chain with states space $\mathbb{N}^2$. This Markov chain is the embedded Markov chain of our queueing system. It is not difficult to see that $\{Y_l, l \in \mathbb{N}\}$ is irreducible and aperiodic.

The first question to be investigated will be the ergodicity of the chain. Because the recursive structure of Equations 2.1 and 2.2 we shall use Foster´s criterion (Pakes, 1969) which says that an irreducible and aperiodic chain $\{Y_l, l \in \mathbb{N}\}$ with states' space S is ergodic if there exits a non-negative function (test function) f(s), $s \in S$ and $\varepsilon > 0$ such that the mean drift

$$x(s) = E[f(Y_{l+1}) - f(Y_l) \mid Y_l = s]$$

is finite for all $s \in S$ and $x_s \leq -\varepsilon$ for all $s \in S$ except perhaps a finite number.

We will consider the following test function

$$f(m, n) = am + n$$

where a is a parameter which will be determined later on.

We readily obtain from Equations 2.1 and 2.2 that the mean drifts $x_{m,n}$ are given by

$$x_{m.n} = E[f(Y_l) - f(Y_{l-1}) \mid Y_{l-1} = (m,n)] = \begin{cases} -a(1-\rho q) + \rho p, & \text{if} \quad m \geq 1 \\ a\rho q + \rho p - \dfrac{u(n)\alpha}{\lambda + u(n)\alpha}, & \text{if} \quad m = 0 \end{cases}$$

where u(x) is the Heavyside function and $\rho = \lambda\beta_l$.

Taking in account that conditions $-a(1 - \rho q) + \rho p < 0$ and $a\rho q + \rho p - \dfrac{u(n)\alpha}{\lambda + u(n)\alpha} < 0$ are equivalents to the conditions:

$$a > \frac{\rho p}{1-\rho p} \text{ and } a < \frac{\alpha - \rho p(\lambda + \alpha)}{(\lambda + \alpha)\rho q}, \; n \geq 1$$

respectively, the unknown parameter a can be found if the interval $\left(\dfrac{\rho p}{1-\rho p}, \dfrac{\alpha - \rho p(\lambda+\alpha)}{(\lambda+\alpha)\rho q}\right)$ is not empty, that is, iff

$$\frac{\rho p}{1-\rho p} < \frac{\alpha - \rho p(\lambda + \alpha)}{(\lambda + \alpha)\rho q}$$

Given that the former inequality is equivalent to the condition

$$\rho < \frac{\alpha}{\alpha - \lambda p} \tag{2.3}$$

we have that this is a sufficient condition for the ergodicity of the embedded Markov chain. We will see later on that the condition 2.3 turns out to be also necessary. Our next objective is to find the stationary distribution

$$\pi_{i,n} = \lim_{l \to \infty} P[N_{1,l} = i, N_{2,l} = n]$$

of the embedded Markov chain $\{Y_l, l \in \mathbb{N}\}$.

The one-step transition probabilities are given by

$$p_{(0,m),(i,n)} = \frac{\lambda}{\lambda + u(m)\alpha} k_{i,n-m} + \frac{u(m)\alpha}{\lambda + u(m)\alpha} k_{i,n-m+1}, \; m = 0,\ldots,n; \; i = 0,1,2,\ldots$$

$$p_{(0,n+1),(i,n)} = \frac{\alpha}{\lambda + \alpha} k_{i,o}, \quad i = 0,1,2,\ldots$$

$$p_{(h,m),(i,n)} = k_{i-h+1,n-m}, \qquad m = 0,\ldots,n; \; h = 1,\ldots,i+1$$

The Kolmogorov equations for the distribution $\pi_{i,n}$ are

$$\pi_{i,n} = \sum_{m=0}^{n} \pi_{0,m} \frac{\lambda}{\lambda + u(m)\alpha} lk_{i,n-m} + \sum_{m=1}^{n+1} \pi_{0,m} \frac{\alpha}{\lambda + \alpha} k_{i,n-m+1} + \sum_{h=1}^{i+1} \sum_{m=0}^{n} \pi_{h,m} k_{i-h+1,n-m} \tag{2.4}$$

By considering the generating function

$$\Phi(z_1, z_2) = \sum_{1=0}^{\infty} \sum_{n=0}^{\infty} \pi_{i,n} z_1^i z_2^n$$

equation 2.4 can be transformed in the equation:

$$\Phi(z_1, z_2) = \left[\lambda\Psi(z_2) + \alpha\frac{\Psi(z_2) + \Psi(0)}{z_2} + \frac{\Phi(z_1, z_2) - \Phi(0, z_2)}{z_1}\right] . \, \beta(\lambda - \lambda p z_2 - \lambda q z_1) \tag{2.5}$$

where $\Psi(z)$ is the auxiliary partial generating function

$$\Psi(z) = \sum_{n=0}^{\infty} \frac{\pi_{0,n}}{\lambda + u(n)\alpha} z^n$$

Noting that

$$\Phi(0, z_2) = (\lambda + \alpha)\Psi(z_2) - \alpha\Psi(0)$$

the expression 2.5 takes the following form:

$$z_2\Phi(z_1, z_2)[z_1 - \beta(\lambda - \lambda p z_2 - \lambda q z_1)] = \left\{ [\lambda z_2(z_1 - 1) + \alpha(z_1 - z_2)]\Psi(z_2) + \frac{\alpha\pi_{0,0}}{\lambda}[z_2 - z_1] \right\} . \beta(\lambda - \lambda p z_2 - \lambda q z_1)$$

(2.6)

We now consider the function

$$f(z_1, z_2) = z_1 - \beta(\lambda - \lambda p z_2 - \lambda q z_1)$$

For each fixed z_2 with $|z_2| < 1$, let's take $f(z_1, z_2)$ as a function of z_1. If $|z_1| = 1$ we have that $Re(\lambda - \lambda p z_2 - \lambda q z_1) > 0$. It is known that $|\beta(s)| < 1$ if $Re(s) > 0$. Then we must have

$$|(z_1 - \beta(\lambda - \lambda p z_2 - \lambda q z_1) - z_1| = |\beta(\lambda - \lambda p z_2 - \lambda q z_1)| < 1 = |z_1|$$

By Rouche's theorem it follows that for each z_2 with $| z_2 |< 1$ there exists a unique solution $z_1 = g(z_2)$ of the equation $f(z_1, z_2) = 0$ in the unit disk, i.e.

$$f(g(z_2), z_2) = g(z_2) - \beta(\lambda - \lambda p z_2 - \lambda q g(z_2)) = 0$$

It is easy to show that:

1. $g(1) = 1$, $g'(1) = \dfrac{\rho p}{1-\rho q}$

2. $g''(1) = \dfrac{\lambda^2 p^2 \beta_2}{(1-\rho q)}$

Inserting $z_1 = g(z_2)$ in Equation 2.6 yields

$$\Psi(z_2)[\lambda z_2(g(z_2) - 1) + \alpha(g(z_2) - z_2)] = \frac{\alpha\pi_{0,0}}{\lambda}[g(z_2) - z_2] \quad (2.7)$$

Observe that the coefficient of $\Psi(z_2)$ in Equation 2.7 never vanishes for $z_2 \in [0, 1)$. Furthermore, using L'Hôpital rule we have

$$\lim_{z_2 \to 1} \Psi(z_2) = \frac{1-\rho}{\alpha - (\lambda p + \alpha)\rho} \frac{\alpha\pi_{0,0}}{\lambda} < \infty$$

Thus, from Equation 2.7 we obtain

$$\Psi(z_2) = \frac{g(z_2) - z_2}{\lambda z_2(g(z_2) - 1) + \alpha(g(z_2) - z_2)} \frac{\alpha\pi_{0,0}}{\lambda}, |z_2| \le 1 \quad (2.8)$$

From Equation 2.6 we obtain

$$\Phi(z_1, z_2) = \frac{\lambda z_2(z_1 - 1) + \alpha(z_1 - z_2)\Psi(z_2) + \frac{\alpha\pi_{0,0}}{\lambda}[z_2 - z_1]}{z_2[z_1 - \beta(\lambda - \lambda p z_2 - \lambda q z_1)]} \beta(\lambda - \lambda p z_2 - \lambda q z_1) \quad (2.9)$$

The value of $\pi_{0,0}$ can be derived from the normalizing condition $\Phi(1, 1) = 1$. Thus we obtain

$$\pi_{0,0} = \frac{\alpha - (\lambda p + \alpha)\rho}{\alpha}$$

A necessary condition for the ergodicity of the chain is $\pi_{0,0} > 0$ and this implies the relation 2.3 . In summary we have proved the following theorem:

Theorem 3.1. The Markov chain $\{Y_l, l \in \mathbb{N}\}$ is ergodic if and only if

$$\rho < \frac{\alpha}{\alpha + \lambda p}$$

If the ergodicity condition is satisfied, the generating function $\Phi(z_1, z_2)$ of the stationary distribution $\{\pi_{i,n}\}$ is given by

$$\Phi(z_1, z_2) = \sum_{i=0}^{\infty}\sum_{n=0}^{\infty} \pi_{i,n} z_1^i z_2^n$$

on the stationary distribution $\{\pi_{i,n}\}$ is given by

$$\Phi(z_1, z_2) = \left[\frac{[\lambda z_2(z_1 - 1) + \alpha(z_1 - z_2)](g(z_2) - z_2)}{\lambda z_2[g(z_2 - 1] + \alpha[g(z_2) - z_2]} + (z_2 - z_1)\right] \frac{\alpha\pi_{0,0}}{\lambda} \frac{\beta(\lambda - \lambda p z_2 - \lambda q z_1)}{z_2[z_1 - \beta(\lambda - \lambda p z_2 - \lambda q z_1)]}$$

where $\pi_{0,0} = \dfrac{\alpha - (\lambda p + \alpha)\rho}{\alpha}$.

3. ANALYSIS OF THE STEADY STATE PROBABILITIES

We shall firstly point out that condition 2.3, that was a necessary and sufficient for the embedded Markov chain, is equivalent to the existence of the limiting probabilities

$$p(0,n) = \lim_{t\to\infty} P[C(t) = 0,\ N_1(t) = 0,\ N_2(t) = n]$$

$$p(x,i,n) = \lim_{t\to\infty} P[C(t) = 1,\ N_1(t) = i,\ N_2(t) = n,\ x < \xi(t) \le x + dx]$$

As it is made in the classical examples displayed in Falin **et al**. (1997), we readily obtain by the supplementary variable method the system of equilibrium equations:

$$(\lambda + u(n)\alpha)p(0,n) = \int_0^{\infty} p(x,0,n)b(x)dx,\ n \ge 0 \qquad (3.1)$$

$$\frac{d}{dx}p(x,i,n) = -(\lambda + b(x))p(x,i,n) + u(n)\lambda p p(x,i,n-1) + u(i)\lambda q p(x,i-1,n), i \ge 0,\ n \ge 0 \qquad (3.2)$$

$$p(0,i,n) = \int_0^{\infty} p(x,i+1,n)b(x)dx + [1 - u(i)](\lambda p(0,n) + \alpha p(0,n+1)),\ i \ge 0,\ n \ge 0 \qquad (3.3)$$

where b(x) is the conditional completion rate at time x. In order to solve Equations 3.1-3.3 we introduce the following partial generating functions:

$$P_0(z_2) = \sum_{n=0}^{\infty} p(0,n) z_2^n$$

$$P(x,z_1,z_2)=\sum_{i=0}^{\infty}\sum_{n=0}^{\infty}p(x,i,n)z_1^i z_2^n$$

The joint generating function of the number of customers in the priority queue and in the orbit given that the server is busy is:

$$P(z_1,z_2)=\sum_{i=0}^{\infty}\sum_{n=0}^{\infty}p(i,n)z_1^i z_2^n$$

where p(i, n) is obtained neglecting, in the case C(t) = 1, the elapsed service time. It is clear that

The main result of this section is provided by the following theorem:

$$P(z_1,z_2)=\int_0^{\infty}P(x,z_1,z_2)b(x)dx \tag{3.4}$$

Theorem 4.1. The stationary distribution of the process X(t) is given by the following generating functions:

$$P_0(z_2)=\frac{g(z_2)-z_2}{\lambda z_2[g(z_2)-1]+\alpha[g(z_2)-z_2]}\alpha p(0,0) \tag{3.5}$$

$$P(x,z_1,z_2)=[(\lambda z_2(z_1-1)+\alpha(z_1-z_2))P_0(z_2)+[z_2-z_1]\alpha p(0,0)]\frac{(1-B(x))e^{-(\lambda-\lambda pz_2-\lambda qz_1)x}}{z_2[z_1-\beta(\lambda-\lambda pz_2-\lambda qz_1)]} \tag{3.6}$$

$$P(z_1,z_2)=\frac{z_1-g(z_2)}{z_1-\beta(\lambda-\lambda pz_2-\lambda qz_1)}\frac{1-\beta(\lambda-\lambda pz_2-\lambda qz_1)}{1-pz_2-qz_1}\frac{1-z_2}{g(z_2)-z_2}P_0(z_2) \tag{3.7}$$

where

$$p(0,0)=\frac{\alpha-(\lambda p+\alpha)\rho}{\alpha}$$

Proof: For the generating functions $P_0(z_2)$ and $P(x, z_1, z_2)$ we can write 3.1-3.3 as

$$(\lambda+\alpha)P_0(z_2)-\alpha p(0,0)=\int_0^{\infty}P(x,0,z_2)b(x)dx \tag{3.8}$$

$$\frac{\partial}{\partial x}P(x,z_1,z_2)=-(\lambda-\lambda pz_2-\lambda qz_1+b(x))P(x,z_1,z_2) \tag{3.9}$$

$$z_1z_2P(0,z_1,z_2)=(\lambda z_1z_2+\alpha z_1)P_0(z_2)-\alpha z_1p(0,0)+z_2\int_0^{\infty}P(x,z_1,z_2)-P(x,0,z_2)b(x)dx \tag{3.10}$$

The solution of the differential equation 3.9 is given by

$$P(x, z_1, z_2) = [1 - B(x)]P(0, z_1, z_2)e^{-(\lambda-\lambda pz_2-\lambda qz_1)x} \tag{3.11}$$

substituting Equation 3.11 into Equations. 3.8, 3.10 and eliminating $P(0, 0, z_2)$ we obtain

$$z_2[z_1 - \beta(\lambda-\lambda pz_2 - \lambda qz_1)]P(0, z_1, z_2) = \lambda z_2(z_1 - 1) + \alpha(z_1 - z_2))P_0(z_2) + \alpha p(0, 0)(z_2 - z_1) \tag{3.12}$$

Considering again the function $f(z_1, z_2) = z_1 - \beta(\lambda - \lambda p z_2 - \lambda q z_1)$ and using similar arguments to those used in the proof of Theorem 3.1, we obtain from Equation 3.12 the formula 3.5. From 3.12 we have

$$P(0,z_1,z_2) = \frac{\lambda z_2(z_1-1) + \alpha(z_1 - z_2)P_0(z_2) + \alpha p(0,0)(z_2 - z_1)}{z_2(z_1 - \beta(\lambda - \lambda p z_2 - \lambda q z_1))} \tag{3.13}$$

By combining Equations. 3.4 and 3.11 we have

$$P(z_1,z_2) = P(0,z_1,z_2)\frac{1-\beta(\lambda - \lambda p z_2 - \lambda q z_1)}{\lambda - \lambda p z_2 - \lambda q z_1} \tag{3.14}$$

Using L'Hôpital rule in Equations. 3.5, 3.13 and 3.14 we get:

$$P_0(1) = \frac{1-\rho}{\alpha - (\lambda p + \alpha)\rho}\alpha p(0,0)$$

$$P(0,1,1) = \frac{\lambda}{\alpha - (\lambda p + \alpha)\rho}\alpha p(0,0)$$

$$P(1,1) = P(0,1,1)\beta_1$$

The constant p(0, 0) can be determined from the normalizing condition $P_0(1) + P(1, 1) = 1$.

Finally, from Equations. 3.11, 3.13 and 3.14 we immediately obtain formulae 3.6 and 3.7 of the theorem.

Denoting by $C = 0$ or $C = 1$ the state of the server in the steady state regime we obtain by routine differentiation in 3.5 and 3.7 the following expressions of the mean queue lengths:

$$E(N_2;C=0) = \frac{(1-\rho)(1-\rho q)\lambda\rho p + \frac{\lambda^3 p^2 \beta_2}{2}}{(1-\rho q)(\alpha - (\lambda p + \alpha)\rho)} \tag{3.15}$$

$$E(N_1;C=1) = \frac{q\lambda^2\beta_2}{2(1-\rho q)} \tag{3.16}$$

$$E(N_2;C=1) = \frac{\lambda p \rho^2(1-\rho q) + \alpha\frac{\lambda^2 p\beta_2}{2}}{(1-\rho q)(\alpha - (\lambda p + \alpha)\rho)} \tag{3.17}$$

Corollary: The mean queue lengths in the two groups are given by

$$E(N_1) = \frac{q\lambda^2\beta_2}{2(1-\rho q)} \tag{3.18}$$

$$E(N_2) = \frac{\lambda p\rho(1-\rho q) + (\alpha + \lambda p)\frac{\lambda^2 p\beta_2}{2}}{(1-\rho q)(\alpha - (\lambda p + \alpha)\rho)} \tag{3.19}$$

Note: In the case p = 1 our system becomes the M/G/1 retrial with constant repeated attempts. In this case $g(z_2) = \beta(\lambda - \lambda z_2)$, $N_1 = 0$ and equations 3.5, 3.7, 3.14, 3.16, 3.18 reduce to:

$$P_0(z) = \frac{\beta(\lambda - \lambda z) - z}{\lambda z(\beta(\lambda - \lambda z) - 1) + \alpha(\beta(\lambda - \lambda z) - z)}\alpha p_{0,0}$$

$$P_1(z) = \frac{1-\beta(\lambda-\lambda z)}{\lambda z(\beta(\lambda-\lambda z)-1)+\alpha(\beta(\lambda-\lambda z)-z)}\alpha p_{0,0}$$

$$E(N; C = 0) = \frac{\lambda\rho(1-\rho)+\frac{\lambda^3\beta_2}{2}}{\alpha-(\lambda+\alpha)\rho}$$

$$E(N; C = 1) = \frac{\lambda\rho^2+\alpha\frac{\lambda^2\beta_2}{2}}{\alpha-(\lambda+\alpha)\rho}$$

$$E(N) = \frac{\lambda\rho+(\lambda+\alpha)\frac{\lambda^2\beta_2}{2}}{\alpha-(\lambda+\alpha)\rho}$$

Which are equations (2.11), (2,12), (2.16) and (2.17) respectively in Choi, B.D. **et al**. (1993).

In the case p = 0 we have the classical M/G/1 queue and eqs. (4.5), (4.7) and (4.16) become

$$P_0(1) = 1 - \rho$$

$$P(z, 1) = E[z^{N_1}; C = 1] = (1-\rho)\frac{1-\beta(\lambda-\lambda z)}{\beta(\lambda-\lambda z)-z}$$

$$E(N_1) = \frac{\lambda^2\beta^2}{2(1-\rho)}$$

respectively.

4. OPTIMAL VALUES

In some cases, every unit in the priority group and the orbit costs r and s units respectively.

In this section we will solve the problem of minimizing in p the expected cost function at the stationary state. As the case r = s = 0 is non-interesting, we will assume that at least one of them is non-zero. An interesting particular case arises when r = s = 1, where the expected total of consumers would be minimized. Let's denote the cost function as T(p). As in the previous section $E(N_1)$ $E(N_2)$ are the expected length of the queue and the orbit. The problem is the following:

$$\min T(p) = rE(N_1) + sE(N_2)$$

$$= r\frac{q\lambda^2\beta_2}{1-\rho q}+s\left(\lambda\rho\frac{p}{\alpha-\lambda p\rho-\alpha\rho}-\lambda^2\beta_2\frac{p}{2q(1-\rho q)}+\alpha\lambda^2\beta\frac{p}{2q(1-\rho q)(\alpha-\lambda p\rho-\alpha\rho)}\right)$$

s.t. $0 \le p \le 1$ where $q = 1 - p$.

First we substitute q by 1 - p and calculate the derivative respect p. After developing some algebraic work we obtain:

$$T'(p) = \lambda\frac{p^2C+2pD+E}{[2q(1-\rho q)(\alpha-\lambda p\rho-\alpha\rho)]^2}$$

$$C = \rho^2[s(2\rho^2 - 2\rho^3) + \lambda^2\beta_2(\alpha s(1 + \lambda\rho - \lambda) - 2r\lambda\rho)]$$

$$D = \rho(1 - \rho)[(2\rho_2\alpha^2\beta_2\lambda s)(1 - \rho) + 4\lambda^2 r\beta_2]$$

$$E = (1 - \rho)^2[2s(\rho^2 + \alpha^2\beta_2(1 - \rho)) + 2r\lambda\beta_2\rho\alpha]$$

Though the coefficients are complicated, the roots of T(p), $r_1 \le r_2$ can be obtained analytically because its numerator is a second degree's polynomial and its denominator is always positive.

Now we will analyze which of the four minimum's candidates, are global. There are the following different cases:

(a) r, s ≥ 0, D and E will be positive, due to $\rho < \frac{\alpha}{\alpha + \lambda p} < 1$.

- If C ≥ 0, that is $A \le \frac{B(2\rho^2 - 2\rho^3) + \lambda^2 \beta_2 (\alpha B(1 + \lambda\rho - \lambda))}{2\lambda\rho}$, then the function is strictly monotonically increasing, so the minimum will be attained at p = 0 and the maximum at p = 1.

- If C < 0, the function will have a positive root r_2 and a negative one r_1. If $r_2 \in [0, 1]$, r_2 is the maximum of T(p) and the minimum is attained at argminT(0), T(1). If $r_2 \notin [0, 1]$, the function is monotonically increasing, so the minimum will be at p = 0 and the maximum at p = 1.

(b) A,B ≤ 0 then we multiply the function T(p) by -1 and find its maximum, for the minimization.

(c) If AB < 0 we suppose that A < 0, if not we make as in (b). Then C > 0, and we have five cases

- $0 \le r_1 < r_2 \le 1$, the minimum is attained at argmin {$T(r_2)$, T(0)}, and the maximum, at argmax {T(1), $T(r_1)$}

- $0 \le r_1 < 1 < r_2$: the maximizer is r_1 and as A < 0 < B the minimizer is p = 0

- $r_1 < 0 \le r_2 < 1$, argmax(T(p))=0, and argmin(T(p)) = r_2

- $r_1 < 0 < 1 < r_2$ this case is impossible, because T(1) > T(0), so the function is non-decreasing.

- $r_1 < r_2 < 0$, $1 < r_1 < r_2$, the function is monotonically increasing on [0,1], then the minimizer is p = 0 and the maximizer p = 1.

Now let's consider A = B = 1, the above mentioned particular case: for it the sign of C is again free, depending on the specifics values that describe the retrial queue model: α, ρ, β_2, λ. If C ≥ 0 as has been discussed all clients will go to the queue, if not as T(0)¡T(1), and again the same result will be obtained.

REFERENCES

CHOI, B.D. and K.K. PARK (1990): "The M/G/1 Retrial Queue with Bernoulli Schedule", **Queueing Systems**, 7, 219-228.

CHOI, B.D.; K.H. RHEE and K.K. PARK (1993): "The M/G/1 Retrial Queue with Retrial rate Control Policy", **Prob. in the Engineering and Inf. Sciences**, 7, 29-46.

FALIN, G.I. and J.G.C. TEMPLETON (1997): **Retrial Queues**, Chapman and Hall, London.

FAYOLLE, G. (1986): "A single telephone exchange with delayed feedbacks", **Traffic Analysis and Computer Performance evaluation**, 245- 253.

PAKES, A.G. (1969): "Some conditions for ergodicity and recurrence of Markov chains", **Operations Research**, 17, 1058-1061.

YANG, T. and J.G.C. TEMPLETON (1987): "A survey on retrial queues", **Queueing Systems**, 2, 201-233.