VI Congreso Galego de Estatística e Investigación de Operacións
Vigo 5–7 de Novembro de 2003

UN SISTEMA DE COLAS CON DEMANDAS REPETIDAS IMPACIENTES Y VACACIONES DEL SERVIDOR

J. Morones, S. Sánchez

Departamento de Matemática Aplicada
Universidad de Málaga

RESUMEN

Los sistemas de colas con demandas repetidas tienen muchas aplicaciones en el modelado de sistemas de comunicación. Consideramos el sistema M/M/1 con demandas repetidas, disciplina constante de reintentos y periodos de inactividad del servidor que pueden comenzar cuando el sistema está vacío. Además, tanto las demandas externas como las repetidas pueden decidir abandonar el sistema cuando encuentran el servidor ocupado. Resolvemos el sistema de ecuaciones de equilibrio para obtener la distribución estacionaria de probabilidad y calcular los principales índices de productividad. También presentamos resultados numéricos relacionando las características fundamentales del sistema.

Palabras y frases clave: Sistemas de colas, Demandas repetidas, Pérdidas geométricas, Vacaciones generalizadas.
Clasificación AMS: 60K25.

1. INTRODUCCIÓN

La administración de los actuales sistemas de comunicación es cada vez más compleja. Hoy en día, la teoría de colas con demandas repetidas encuentra su reconocimiento en las aplicaciones que tiene en este campo. La característica principal de estos modelos es la consideración de periodos de tiempo donde el servidor está libre, aún cuando existan clientes que están esperando a ser servidos. Estos clientes, que forman un grupo llamado órbita, realizan periódicamente reintentos para obtener su servicio. Las llamadas telefónicas a una centralita son un claro ejemplo de estos modelos.

El modelo que se presenta en este trabajo es una extensión de [8]. Las dos características más importantes del sistema es la consideración de demandas impacientes y de vacaciones generalizadas.

Una demanda impaciente es aquella que tiene la posibilidad de abandonar el sistema sin haber obtenido su servicio. Existen muchos tipos de demandas impacientes en función del momento en el que deciden abandonar el sistema. En este trabajo se consideran dos tipos: demandas primarias que pueden decidir abandonar el sistema nada más llegar, si encuentran el servidor ocupado y demandas repetidas que pueden decidir abandonar el sistema si encuentran el servidor ocupado en uno de sus reintentos desde la órbita.

Existen numerosas publicaciones considerando distintos modelos de sistemas de colas con demandas repetidas y un flujo de entrada de Poisson (ver, por ejemplo [2] y [11]). En [4] y en [7] se presentan resultados de los sistemas de colas con demandas repetidas impacientes y no persistentes pero considerando una disciplina clásica de reintentos y en todos ellos no se consideran vacaciones en el servidor.

La naturaleza de algunos sistemas de colas requiere la consideración de periodos de tiempo donde los servidores dejan de funcionar. Estos periodos de inactividad se conocen como vacaciones del servidor. Existen numerosas políticas para las vacaciones del servidor, teniendo cabida en todas ellas actividades tales como la atención a demandas de otros sistemas, llamadas de índole personal o mejora de la calidad del servicio.

En [1] se estudia el sistema $M/G/1$ con demandas repetidas y vacaciones del servidor, sin embargo no se consideran las demandas impacientes. En [3] se obtienen resultados para sistemas que consideran varios flujos de entrada y demandas no persistentes o vacaciones del servidor con una disciplina clásica de reintentos. En [9] y más concretamente en [6] y en [10] se obtiene resultados para estos mismos sistemas, pero considerando una disciplina constante de reintentos. En todos los casos anteriores no se mezclan las dos características de nuestro modelo.

En este trabajo se estudia el sistema de colas $M/M/1$ con demandas repetidas que considera demandas impacientes de dos tipos distintos y vacaciones del servidor.

2. DESCRIPCIÓN DEL SISTEMA

Vamos a considerar un sistema de colas con un único servidor y un flujo de entrada de Poisson de intensidad λ. Si una demanda que llega al sistema encuentra libre el servidor, ésta comenzará inmediatamente a servirse; en caso contrario, con probabilidad $1 - H_1$ abandonará el sistema, y con probabilidad H_1 se unirá a un grupo de demandas insatisfechas, que llamaremos órbita, desde donde intentará obtener servicio en tiempos aleatorios posteriores. Los tiempos entre reintentos sucesivos de una demanda individual están regidos por una ley exponencial de parámetro α/n donde n es el número de demandas de la órbita. Cuando una demanda de la órbita reintenta acceder al servidor y lo encuentra libre, comienza inmediatamente a servirse; en caso contrario puede decidir, con probabilidad $1 - H_2$, abandonar definitivamente el sistema sin haber obtenido su servicio o, con probabilidad H_2, regresar a la órbita.

Tanto para las demandas primarias como para las repetidas, los tiempos de servicio de cada una de ellas son variables aleatorias independientes distribuidas exponencialmente con parámetro ν_1.

Si la órbita se encuentra vacía en el instante de finalización del servicio de una demanda, el servidor puede comenzar un periodo de vacaciones con probabilidad θ_1, o permanecer desocupado a la espera de servir una nueva demanda con probabilidad $\bar{\theta}_1 = 1 - \theta_1$. Durante el periodo de vacaciones, las demandas que acceden al sistema se comportan como si el servidor estuviese ocupado. Si en el momento de finalización del periodo de vacaciones hay, al menos, una demanda en la órbita, el servidor vuelve a estar activo permaneciendo desocupado a la espera de servir una nueva demanda primaria o repetida. Pero si la órbita está vacía al finalizar el periodo de vacaciones, el servidor puede volver a comenzar un nuevo periodo de vacaciones con probabilidad θ_2, o permanecer desocupado a la espera de servir una nueva demanda con probabilidad $\bar{\theta}_2 = 1 - \theta_2$. Los tiempos que duran los periodos de vacaciones son variables aleatorias independientes distribuidas exponencialmente con parámetro ν_2.

3. ECUACIONES DE EQUILIBRIO DEL SISTEMA

El comportamiento estocástico del sistema considerado en este trabajo puede describirse mediante el proceso $\chi(t) = (C(t), N(t))$ con $t \geq 0$, donde

- $C(t)$ indica el estado del servidor en el instante t: 0 si está libre, 1 si está ocupado o 2 si está de vacaciones.

- $N(t)$ es el número de demandas que hay en la órbita en el instante t.

Designaremos por:

$$p_i(n) = \lim_{t \to \infty} P[C(t) = i; N(t) = n] \quad \text{para todo} \quad i = 1; 2; 3 \quad \text{con} \quad n \in \mathbb{N}$$

a las probabilidades estacionarias de los estados del sistema.

Como la distribución de los tiempos de servicio es exponencial, el proceso $\chi(t)$ es un proceso de Markov cuyo espacio de estados es $S = \{0; 1; 2\} \times \mathbb{N}$. Del estudio de la transición entre estados obtenemos el siguiente sistema de ecuaciones de equilibrio:

$$[\lambda + (1 - \delta_{0;n})\alpha]\, p_0(n) = \nu_1 p_1(n) + \nu_2(1 - \delta_{0;n}\theta_2) p_2(n) \tag{1}$$

$$\begin{aligned}[\nu_1 + \lambda H_1 + (1 - \delta_{0;n})\alpha(1 - H_2)]\, p_1(n) &= (1 - \delta_{0;n})\lambda H_1 p_1(n-1) + \\ &+ \alpha(1 - H_2) p_1(n+1) + \lambda p_0(n) + \alpha p_0(n+1)\end{aligned} \tag{2}$$

$$\begin{aligned}[\nu_2 + \lambda H_1 + (1 - \delta_{0;n})\alpha(1 - H_2)]\, p_2(n) &= (1 - \delta_{0;n})\lambda H_1 p_2(n-1) + \\ &+ \alpha(1 - H_2) p_2(n+1) + \delta_{0;n}\nu_1\theta_1 p_1(n)\end{aligned} \tag{3}$$

para todo $n \in \mathbb{N}$, siendo $\delta_{i;j}$ el símbolo de Kronecker.

4. RESOLUCIÓN DEL SISTEMA DE ECUACIONES DE EQUILIBRIO

Para la resolución del sistema de ecuaciones (1)-(3) introducimos las siguientes funciones generatrices:

$$P_i(z) = \sum_{n=0}^{\infty} p_i(n) \cdot z^n \quad \text{para todo} \quad i = 1; 2; 3 \quad \text{con} \quad z \in [0; 1].$$

Teorema. La distribución estacionaria del sistema viene dada por:

$$\begin{aligned}
P_0(z) &= \frac{\alpha\nu_1 p_0(0) + (1-z)\nu_1\alpha(1-H_2)p_1(0)}{\lambda\nu_1 z + \alpha\nu_1 - (\lambda+\alpha)[\nu_1 z + z(1-z)\lambda H_1 - \alpha(1-H_2)(1-z)]} + \\
&+ \frac{[\nu_2\theta_2 p_2(0) - \nu_2 P_2(z) - \alpha p_0(0)][z(1-z)\lambda H_1 + z\nu_1 - (1-z)\alpha(1-H_2)]}{\lambda\nu_1 z + \alpha\nu_1 - (\lambda+\alpha)[\nu_1 z + z(1-z)\lambda H_1 - \alpha(1-H_2)(1-z)]} \qquad (4)\\
P_1(z) &= \frac{1}{\nu_1}[(\lambda+\alpha)P_0(z) - \nu_2 P_2(z) - \alpha p_0(0) + \nu_2\theta_2 p_2(0)] \qquad (5)\\
P_2(z) &= \frac{\alpha(1-H_2)(z-1)p_2(0) + z\nu_1\theta_1 p_1(0)}{\nu_2 z + \lambda H_1(1-z)z + \alpha(1-H_2)(z-1)} \qquad (6)
\end{aligned}$$

con

$$p_1(0) = \frac{\lambda\theta_2}{\nu_1[\theta_2(1-\theta_1) + \theta_1]} p_0(0) \tag{7}$$

$$p_2(0) = \frac{\lambda\theta_1}{\nu_2[\theta_2(1-\theta_1) + \theta_1]} p_0(0) \tag{8}$$

siendo

$$p_0(0) = \frac{(\theta_1\bar{\theta}_2 + \theta_2)\nu_2 A}{\theta_1 B + \nu_2\theta_2 C} \tag{9}$$

la probabilidad de que el sistema está vacío y A, B y C las siguientes tres constantes auxiliares:

$$\begin{aligned}
A &= (1-H_1)\lambda(\lambda+\alpha) + (1-H_2)\alpha(\lambda+\alpha) + \alpha(\nu_1-\lambda) - \lambda^2 \\
B &= (1-H_1)\lambda(\alpha\nu_2\bar{\theta}_2 - \lambda\nu_1\theta_2) + (1-H_2)\alpha\left[\alpha(\lambda+\nu_2\bar{\theta}_2) + \lambda(\lambda+\nu_1\bar{\theta}_2)\right] + \\
&\quad + \nu_1\left[\alpha(\lambda\theta_2+\nu_2\bar{\theta}_2) + \lambda^2\theta_2\right] \\
C &= \alpha\left[(1-H_1)\lambda + (1-H_2)(\lambda+\alpha)(\rho+1) + \nu_1\right]
\end{aligned}$$

siendo $\rho = \lambda/\nu_1$ la carga del sistema.

Demostración. Mediante las funciones generatrices antes definidas el sistema de ecuaciones de equilibrio (1)-(3) se transforma en

$$(\lambda+\alpha)P_0(z) - \alpha p_0(0) = \nu_1 P_1(z) + \nu_2 P_2(z) - \theta_2\nu_2 p_2(0) \tag{10}$$

$$\begin{aligned}
&[\nu_1 + \lambda H_1 + \alpha(1-H_2)]\,P_1(z) - \alpha(1-H_2)p_1(0) = z\lambda H_1 P_1(z) + \\
&\quad + \frac{1}{z}\alpha(1-H_2)\,[P_1(z) - p_1(0)] + \lambda P_0(z) + \frac{\alpha}{z}[P_0(z) - p_0(0)]
\end{aligned} \tag{11}$$

$$\begin{aligned}
&[\nu_2 + \lambda H_1 + \alpha(1-H_2)]\,P_2(z) - \alpha(1-H_2)p_2(0) = z\lambda H_1 P_2(z) + \\
&\quad + \frac{1}{z}\alpha(1-H_2)\,[P_2(z) - p_2(0)] + \nu_1\theta_1 p_1(0)
\end{aligned} \tag{12}$$

Resolvemos este sistema de ecuaciones de la siguiente forma:

1. De las ecuaciones (10)-(12) para $z = 1$ obtenemos:

$$\nu_1\theta_1 p_1(0) = \nu_2\theta_2 p_2(0) \tag{13}$$

   y como la ecuación (1) para $n = 0$ resulta ser:

$$\lambda p_0(0) = \nu_1 p_1(0) + \nu_2(1-\theta_2)p_2(0) \tag{14}$$

   resolvemos el sistema de ecuaciones (13)-(14) y se obtienen las expresiones (7) y (8) de $p_1(0)$ y de $p_2(0)$ en función de $p_0(0)$.

2. De la ecuación (12) obtenemos la expresión (6) de $P_2(z)$ en función de $p_1(0)$ y de $p_2(0)$. Teniendo en cuenta (7) y (8) podemos obtener otra expresión de $P_2(z)$ en función de $p_0(0)$.

3. De la ecuación (10) obtenemos la expresión (5) de $P_1(z)$. Si sustituimos esta expresión de $P_1(z)$ en (11) obtenemos la expresión (4) de $P_0(z)$ en función de $P_2(z)$, $p_0(0)$, $p_1(0)$ y $p_2(0)$. Aplicando las relaciones anteriores podríamos obtener la expresión de $P_0(z)$ y de $P_1(z)$ en función, únicamente, de $p_0(0)$ (que no se incluyen en este trabajo por la extensión de las mismas).

4. Por último, para obtener la expresión explícita (9) de $p_0(0)$ aplicamos la condición de normalización:
$$P_0(1) + P_1(1) + P_2(1) = 1$$

$\square$

Por otro lado, de la expresión (9) de $p_0(0)$ para las probabilidades estacionarias de los estados del sistema se obtiene la siguiente condición necesaria para la ergodicidad del sistema:

$$p_0(0) > 0$$

Esta condición es una generalización de

$$\frac{\lambda(\lambda+\alpha)H_1}{\alpha[\nu + (1-H_2)(\lambda+\alpha)]} < 1$$

obtenida en [8], considerando $\theta_1 = 0$ y de

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

obtenida en [5] para el caso básico, si además consideramos $H_1 = H_2 = 1$.

5. CONCLUSIONES

En este trabajo se ha estudiado un sistema de colas con demandas repetidas que considera demandas impacientes de dos tipos y vacaciones del servidor. Como resultado se han obtenido las expresiones explícitas de las funciones generatrices de la distribución estacionaria de probabilidad del sistema. Estas expresiones nos permiten, por simple derivación, obtener la probabilidad de cualquier suceso y, en particular, los principales índices de productividad del sistema. Además, estas expresiones explícitas de la probabilidad permiten obtener resultados numéricos que relacionan las características fundamentales del sistema.

Este trabajo es un extensión del modelo presentado en [8] que no consideraba vacaciones en el servidor. Podemos comprobar que haciendo $\theta_1 = 0$ en nuestro modelo obtenemos los mismos resultados. Además, podemos comprobar que obtenemos casos particulares de [10] haciendo $H_1 = H_2 = 1$, de [6] haciendo $\theta_1 = 0$ y $H_2 = 1$ y de [5] haciendo $\theta_1 = 0$ y $H_1 = H_2 = 1$ (todos ellos, recogidos en [9]).

El trabajo futuro pasa por seguir generalizando estos modelos y considerar otras características importantes como feedback, disciplina lineal de reintentos, varios servidores, etc.

6. REFERENCIAS

[1] Artalejo, J. R. (1997): Analysis of an M/G/1 queue with constant repeated attempts and server vacations. Computer and Operations Research. 24, 493–504.

[2] Artalejo, J. R. (1999): A classified bibliography of research on retrial queues: Progress in 1990-1999. Top, 7 (2), pp. 187–211.

[3] Atencia, I., Bocharov, P.P., D'Apice, C. y Phong, N.H. (2000): A Single-Server Retrial Queue System with Multidimensional Poisson Flow. Automation and Remote Control, 61, pp. 1871–1884.

[4] Atencia, I., Bocharov, P.P. y Phong, N.H. (2001): Retrial queueing system with several input flows. Investigación Operacional, 2 (22), pp. 135–144.

[5] Atencia, I., Fortes, I., Moreno, P. y Sánchez, S. (2001): Un sistema de colas con demandas repetidas, varios flujos de entrada y reintentos constantes. XXVI Congreso Nacional de Estadística e Investigación Operativa. Úbeda (España).

[6] Atencia, I., Fortes, I., Moreno, P. y Sánchez, S. (2002): A queueing system with non-persistent customers, several input flows and constant repeated attempts. International Conference on Computational and Mathematical Methods in Science and Engineering.

[7] Falin, G.I. y Templenton, J.R. (1997): Retrial Queues. Ed. Chapman & Hall

[8] Morones, J. y Sánchez, S. (2003): Un sistema de colas con demandas repetidas impacientes y no persistentes. XXVII Congreso Nacional de Estadística e Investigación Operativa. Lérida (España).

[9] Sánchez, S. (2002). Sistemas de colas con demandas repetidas y con varios flujos de entrada. Ph. D. University of Málaga.

[10] Sánchez, S. y Fortes, I. (2003): Un sistema de colas con varios flujos de entrada, reintentos constantes y vacaciones del servidor. XXVII Congreso Nacional de Estadística e Investigación Operativa. Lérida (España).

[11] Yang, T. y Templeton, J.G.C. (1987): A survey on retrial queues. Queueing Systems, 2, pp. 201–233.