

Simulando una cola simple

Sessa, Carlos
csessa@alu.itba.edu.ar

Abramowicz, Pablo
pabramow@alu.itba.edu.ar

Gomez Vidal, Maximiliano
dgomezvi@alu.itba.edu.ar

Villa Fernández, Santiago
svillafe@alu.itba.edu.ar

RESUMEN

En este artículo se simula un sistema de cola simple y se estiman parámetros tales como la longitud media de la cola y el tiempo de atención. Se analiza también el costo operativo asociado.

Palabras clave

Modelo de cola simple, FIFO, tiempo de espera, sistema de atención al cliente, servidor simple

1. INTRODUCCIÓN

Las colas se utilizan en diversos ámbitos, tales como sistemas informáticos, transportes y operaciones de investigación, entre otros. Los objetos, personas o eventos son tomados como datos que se almacenan para su posterior procesamiento.

La cola simple, denominada $M/M/1/\infty/FIFO$ o simplemente $M/M/1$, es el sistema más sencillo de analizar mediante una simulación por eventos discretos.

El sistema de cola simple se encuentra caracterizado principalmente por el tiempo necesario para atender a un cliente y la tasa de llegada de los mismos.

En la sección 2 se describe el modelo de cola simple y los parámetros involucrados. En la sección 3 se realizan las simulaciones, se estudia el comportamiento del sistema y se analiza el costo operativo resultante. Finalmente, se exponen las conclusiones en la sección 4.

2. MODELO DE COLA SIMPLE

Este modelo asume que los clientes llegan al sistema mediante un proceso de Poisson con una tasa media de λ [clientes/hora]. También se asume que el servidor atiende a cada uno de los clientes con un tiempo de servicio distribuido en forma exponencial con media $1/\mu$.

Si cuando llega un nuevo cliente el servidor se encuentra libre, entonces el mismo es atendido en forma inmediata.

Caso contrario, el cliente debe ingresar en la cola y esperar su turno.

Resulta de interés poder establecer las probabilidades en estado estacionario, cuando el sistema se encuentra en equilibrio:

$$p_n = \lim_{t \rightarrow \infty} P\{N_t = n\} \quad (1)$$

donde N_t es la cantidad de clientes en el sistema en el instante t . Generalizando para el estado n se obtiene una relación de recurrencia cuya solución es:

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0 \quad (2)$$

con condiciones iniciales

$$p_1 = \frac{\lambda}{\mu} p_0 \quad (3)$$

El factor $\rho = \lambda/\mu$ se denomina intensidad de tráfico del sistema. De esta forma la ecuación (2) puede reescribirse como:

$$p_n = (1 - \rho)\rho^n \quad (4)$$

De esta ecuación se desprende que la intensidad de tráfico debe cumplir $\rho < 1$ para que el sistema sea estable. El caso $\rho > 1$ implica que el servidor tarda más tiempo en atender a un cliente que nuevos clientes en arribar al sistema, generando así una cola cada vez más larga.

A partir de las ecuaciones obtenidas es posible describir la longitud media de la cola L_q y el número medio de clientes en el sistema L en función de ρ :

$$L_q = \frac{\rho^2}{1 - \rho} \quad (5)$$

$$L = \frac{\rho}{1 - \rho} \quad (6)$$

El modelo no estaría completo sin el análisis del tiempo medio de espera, medida importante en lo que respecta a la performance del sistema de colas. El tiempo medio de un individuo en el sistema (W) y en cola (W_q) resultan:

$$W = \frac{1}{\mu - \lambda} \quad (7)$$

$$W_q = \frac{\rho}{\mu - \lambda} \quad (8)$$

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 200X ACM X-XXXXX-XX-X/XX/XX ...\$5.00.

3. SIMULACIÓN

En primer lugar se desea estimar el promedio temporal de clientes en el sistema y el promedio temporal de los clientes en la cola, dados por las ecuaciones:

$$L = \frac{1}{T} \int_0^T L(t) dt \quad (9)$$

$$L_q = \frac{1}{T} \int_0^T q(t) dt \quad (10)$$

La cantidad de clientes en cola en el instante t se encuentra representada por la función $q(t)$. $L(t)$ es la cantidad de clientes en el sistema.

Se comienza estudiando el comportamiento del sistema con una tasa de arribos $\lambda = 1$ y un tiempo medio de servicio $1/\mu = 1.25$, ambos indicando cantidad de clientes por hora. La cantidad total de clientes a ser atendidos por el sistema se fija en $n = 9000$.

El valor medio de clientes en cola teórico para los parámetros especificados resulta aproximadamente 3.2 clientes. La estimación luego de 50 simulaciones resulta 3.241, presentando un error relativo del 0.295%.

Resulta sumamente importante estimar también el tiempo medio que tarda cada cliente desde que arriba al sistema hasta que termina de consumir el servicio.

En la figura 1 se observa como varía el tiempo medio en el sistema W a medida que la cantidad de clientes arribados aumenta.

En la figura 2 se observa que el tiempo medio de un individuo, tanto en la cola como en el sistema, se incrementa a medida que aumenta el tiempo medio que toma al servicio atender al cliente. Para este gráfico se toman 18 valores distintos del parámetro μ y se estudia su impacto en el tiempo medio de un individuo tanto en la cola como en el sistema. Se realiza un análisis similar en la figura 3, esta vez con distintos valores del parámetro λ .

Es factible suponer que se incurre en determinados costos al mantener al sistema en estado operativo. Los mismos pueden separarse en el costo asociado a un cliente que espera, sea $c_{\text{hora cliente}} \frac{\$}{\text{hora}}$, y el costo propio de utilización del servidor, $s_{\text{hora}} \frac{\$}{\text{hora}}$.

A partir de estas definiciones es posible computar el costo total medio, dado por:

$$C = \frac{1}{T} \int_0^T c(t) dt \quad (11)$$

donde $c(t)$ es el costo total y T es el tiempo de operación.

En las figuras 4 y 5 se presenta la evolución del costo total medio al variar los parámetros μ y λ respectivamente.

4. CONCLUSIONES

En un sistema de cola simple la estabilidad depende exclusivamente de la intensidad de tráfico. Se observa que pequeñas variaciones en la tasa de arribos o el tiempo medio de servicio pueden implicar grandes costos asociados.

Contar con información precisa de la cantidad de clientes a atender y la frecuencia de llegada permite poder refinar la forma en que se presta el servicio, buscando minimizar el tiempo de espera de los clientes y también el tiempo en que cada servidor se encuentra inactivo.

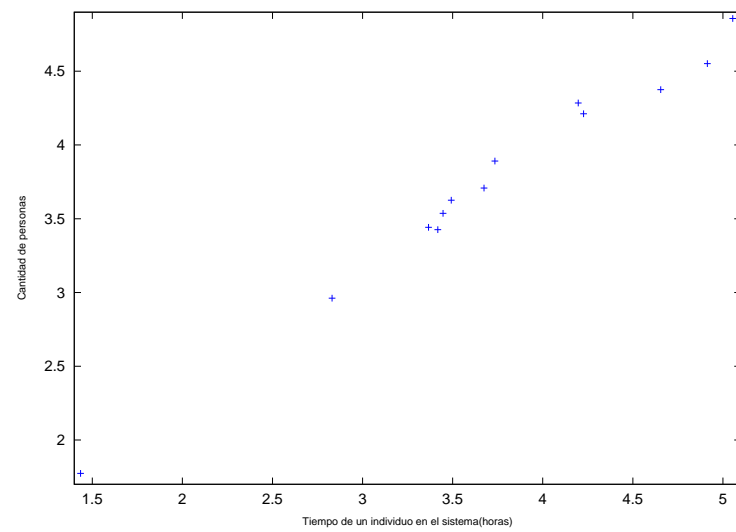


Figura 1: Valor medio de individuos en el sistema vs. Tiempo de un individuo en el sistema

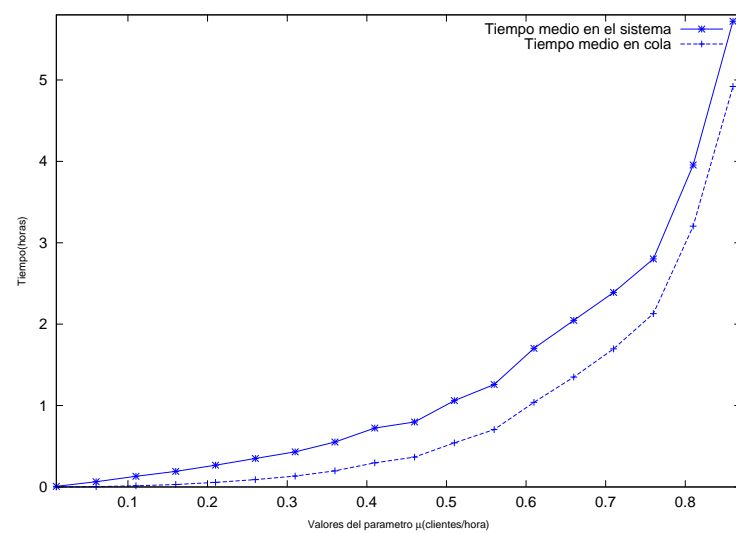


Figura 2: Tiempo medio vs. μ

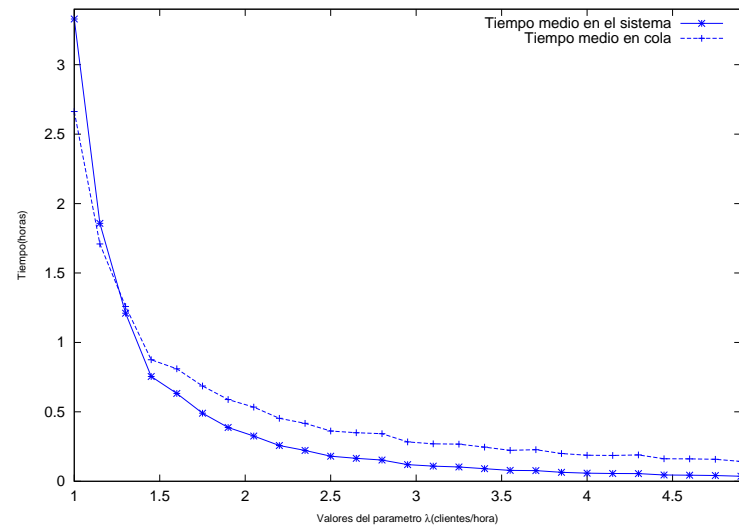


Figura 3: Tiempo medio vs. λ

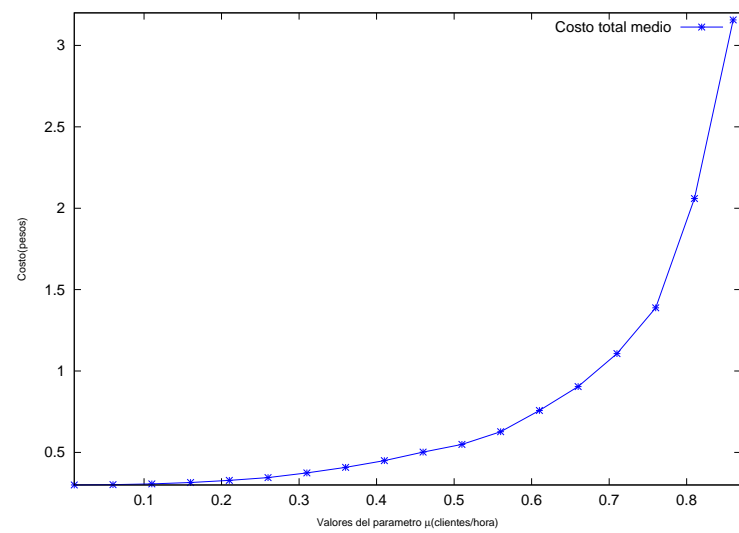


Figura 4: Costo total medio vs. μ

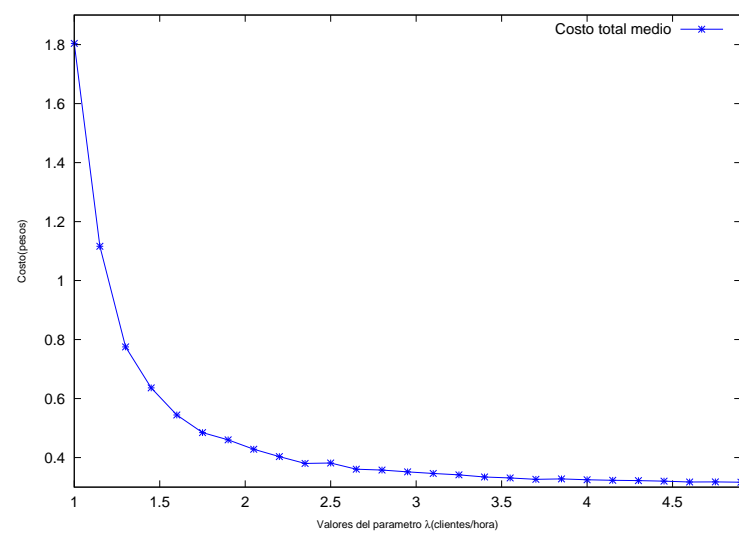


Figura 5: Costo total medio vs. λ