

## Modelado mediante redes de colas abiertas con realimentación de la sala de urgencias de un hospital público

### Modeling a public hospital's emergency department with feedback open queue networks

Gonzalo Everardo Aceves Gómez\*, Ricardo Armando González Silva\*\*✉, Héctor Alfonso Juárez López\*\*, Rodolfo Rafael Medina Ramírez\*\*\*, José Antonio Vázquez Ibarra\*\*\*\*

Aceves Gómez, G. E., González Silva, R. A., Juárez López, H. A., Medina Ramírez, R. R., & Vázquez Ibarra, J. A. Modelado mediante redes de colas abiertas con realimentación de la sala de urgencias de un hospital público. *Investigación y Ciencia de la Universidad Autónoma de Aguascalientes*, 26(74), 48-57.

#### RESUMEN

Este trabajo propone un modelo de red de colas abiertas con realimentación, de la sala de urgencias, para entender su comportamiento y tomar decisiones estratégicas. Se establece una matriz de ruta de probabilidades en la red de colas para determinar de manera general las variantes de comportamiento de este modelo, generando así un rango de escenarios con diferentes patrones de comportamiento; con los resultados numéricos, se analizan los indicadores de eficiencia de la Teoría de Colas de las tres locaciones: consultorios, laboratorios y observación-yesos-suturas, las cuales se modelan con  $M/M/s$ ,  $M/G/1$  y  $M/M/1$ , respectivamente. Los resultados del análisis numérico

**Palabras clave:** modelo de colas; sala de urgencias; matriz de ruta de probabilidades; indicadores de eficiencia.

**Keywords:** queueing model; emergency room; probability route matrix; efficiency indicators.

Recibido: 9 de junio de 2017, aceptado: 15 de enero de 2018

\* Maestría en Ciencia y Tecnología, Universidad de Guadalajara-Centro Universitario de los Lagos. Av. Enrique Díaz de León #1144, Paseos de La Montaña, C. P. 47460, Lagos de Moreno, Jalisco, México. Correo electrónico: gonzalo.aceves@academicos.udg.mx

\*\* Departamento de Ciencias Exactas y Tecnología, Universidad de Guadalajara-Centro Universitario de los Lagos. Av. Enrique Díaz de León #1144, Paseos de La Montaña, C. P. 47460, Lagos de Moreno, Jalisco, México. Correo electrónico: rgonzalez@culagos.udg.mx; hjuarez@culagos.udg.mx

\*\*\* Departamento de Posgrado e Investigación, Universidad Politécnica de Aguascalientes. Calle Paseo San Gerardo No. 207, Fracc. San Gerardo, C. P. 20342, Aguascalientes, Aguascalientes, México. Correo electrónico: rmedinamx@gmail.com

\*\*\*\* Programa Académico de Ingeniería Industrial, Universidad Politécnica de Aguascalientes. Calle Paseo San Gerardo No. 207, Fracc. San Gerardo, C. P. 20342, Aguascalientes, Aguascalientes, México. Correo electrónico: antonio.vazquez@upa.edu.mx

✉ Autor para correspondencia

muestran la sensibilidad del comportamiento de la sala de urgencias con base en los valores de la matriz de ruta de probabilidades y en cuáles valores funciona óptimamente.

#### ABSTRACT

This work proposes a model of open queue network with feedback, from the emergency department, to understand their behavior and make strategic decisions. In the queueing network, a probability route matrix is established to determine the general behavior variants of this model. It generates a range of scenarios with different behavior patterns; with the numerical results, we analyze the efficiency indicators of Queueing Theory on the three locations: clinics, laboratories, and observation-gypsum-sutures, which are modeled as  $M/M/s$ ,  $M/G/1$  and  $M/M/1$ , respectively. The numerical results show the sensibility of the emergency department behavior and the optimally performance, using the probability route matrix values.

#### INTRODUCCIÓN

El flujo de pacientes es uno de los problemas más importantes a estudiar con el objetivo de mejorar la eficiencia en los servicios de atención de la salud. De acuerdo con la investigación de operaciones (IO), consiste en la circulación de los pacientes a través de un conjunto de estaciones en un centro de atención médica (Hillier & Lieberman, 2001). Desde una perspectiva clínica, representa la progresión del estado de salud de un paciente. La comprensión del flujo de pacientes proporciona valiosa información a los médicos, a los administradores y a los mismos pacientes acerca de las necesidades de cuidado

de la salud asociados con la progresión de la enfermedad o el estado de recuperación. El estudio del flujo de pacientes desde la IO sirve para apoyar las actividades operativas de un centro de atención, con la asignación eficaz de recursos y planificación de capacidad (Côté, 2000).

Se han identificado diversos factores que impactan en el aumento de los tiempos de espera de la sala de urgencias (SU) de un hospital, como el aumento de la población o las necesidades de salud debidas a las condiciones ambientales (Derlet, Richards, & Kravitz, 2008). Un hospital debe tratar a sus pacientes de manera oportuna, especialmente a aquellos con necesidad de cuidados críticos, lo cual es muy complicado en la práctica, sobre todo en las instituciones públicas de salud, debido a la alta demanda y a los recursos limitados.

En la presente investigación se estudia el comportamiento del flujo de pacientes de la SU de un hospital mediante una red de colas (RC). La sección 2 es el estado del arte de las RC en la SU. La sección 3 plantea los elementos de la SU, además de las relaciones fundamentales de las variables y parámetros determinados por la teoría de colas (TC). La sección 4 mide los indicadores de eficiencia del sistema de colas (SC). En la sección 5 se determinan los rangos de valores de diversos parámetros para dar paso al análisis de indicadores de eficiencia del sistema. Finalmente, se plantean las conclusiones de la investigación.

#### **Antecedentes de modelos de redes de colas de la SU**

En las investigaciones relacionadas con la modelación de la SU de hospitales mediante TC, algunas se enfocan en el flujo de pacientes, otras en el óptimo de personal o el número de camas, en los horarios, en los tipos de afección de los pacientes, etc., pero los modelos de RC son mucho menos comunes y la mayoría de ellos limitados.

Mayhew y Smith (2008) ilustran cómo el flujo de pacientes de un departamento de Emergencias (DE) puede representarse como un proceso de colas y cómo los resultados de este modelo pueden usarse para visualizar e interpretar el objetivo gubernamental de completar y descargar 98% de los pacientes dentro de 4 horas de una manera simple. Filipowicz y Kwiecień (2008) describen los SC y de RC que se utilizan con éxito para analizar el rendimiento de diferentes sistemas, que incorporan procesos de Markov con tiempos de servicio exponenciales y

un proceso de llegada de Poisson y SC con servicio individual. Presentan una aplicación de redes cerradas de BCMP en el área de salud y la evaluación del desempeño del sistema de información.

Cochran y Roche (2009) proponen un modelo de RC abiertas destinado a aumentar la capacidad de un DE para tratar a los pacientes. Se incorporan patrones de llegada no homogéneos, distribuciones de tiempo de servicio no exponenciales y múltiples tipos de pacientes. Usan los modelos  $M/M/s$  y  $M/G/s/s$ . Jiang y Giachetti (2008) amplían y mejoran el modelo de RC abiertas para analizar el tiempo del ciclo del paciente, de modo que se pueda modelar el flujo del paciente en un centro de atención de urgencias. Hacen contribuciones implementando una aproximación para colas fork/join en la red y mejorando la aproximación para múltiples servidores colas, tanto en tráfico bajo como en condiciones de mucho tráfico.

Olorunsola, Adeleke y Ogunlade (2014) analizaron el número óptimo de camas requeridas en el hospital, lo cual se puede lograr con un conocimiento adecuado del flujo del paciente para mejorar el nivel de ingreso a la atención. En esta investigación se utilizaron técnicas analíticas y simulación para estudiar una RC simple compuesta de solamente dos estaciones de servicio colocadas en tándem; se estudiaron todas las admisiones al Departamento de Emergencia y Accidentes (DAE). Usan los modelos de cola  $M/M/s$ . Véricourt y Jennings (2011) presentan un modelo de RC cerrada  $M/M/s/n$  para determinar políticas eficientes de personal de enfermería. Modelan la carga de trabajo de las enfermeras dentro de una sola unidad médica con  $n$  pacientes homogéneos.

Izady y Worthington (2012) proponen un esquema iterativo que utiliza redes con servidores infinitos, la ley de personal de raíz cuadrada y simulación para encontrar una buena solución a la planeación de horarios de trabajo y servicio eficiente, sin aumentar la carga horaria. La implementación de este algoritmo en un departamento típico de A & E sugiere que se puede obtener una mejora significativa en el objetivo, incluso sin un aumento en el total de horas de personal. Au-Yeung, Harrison y Knottenbelt (2006) desarrollan un modelo multiclase de RC markoviana de flujo de pacientes en el DAE de un importante hospital de Londres y mediante simulación de eventos discretos, proponen opciones para que la priorización por afección se optimice.

Jlassi, Mhamedi y Chabchoub (2009) desarrollaron una RC con múltiples tipos de clientes. Se usan diferentes indicadores de rendimiento e ilustran RC multiclase con el objetivo de mejorar el rendimiento de la SU del hospital Habib Bourguiba, en Sfax, Túnez. Zhu, Gong y Tang (2013) establecen un modelo general de RC de múltiples etapas con flujo de pacientes con realimentación para analizar el comportamiento del flujo de pacientes entre las etapas y el desempeño de los DE en un hospital.

## MATERIALES Y MÉTODOS

### La sala de urgencias del hospital público CHMH

Los sistemas de salud pública de México tienen como objetivo alcanzar una cobertura universal, todos y cada uno de los mexicanos debe tener acceso a los servicios de salud. En 2008 la población en México era de 106.6 millones de habitantes, de los cuales el número de personas que recibía atención médica del Seguro Popular era de 27,176 millones; es decir, aproximadamente 25.5% de la población (Gómez Dantés et al., 2011). Esto, en conjunto con la crisis económica, ha provocado que el tiempo de espera esté aumentando. Sin inversiones adicionales en instalaciones ni personal, aumentando las filas y el tiempo de espera, hay insatisfacción entre los pacientes, los familiares e incluso el personal del sistema de salud. Las SU son el principal punto de acceso a los hospitales, saturadas por aquellos que están cansados de esperar una cita (Vázquez I., González S., & Juárez L., 2014).

La SU está compuesta por las siguientes locaciones: registro (R), área de consultorios (C), área de choque, laboratorios (L), observación-yesos-suturas (OYS), y nueve camas. Un paciente entra a la SU de dos formas: por área de choque (llegadas por ambulancia) o por sus medios. En esta investigación solamente se considera este segundo caso. Cuando un paciente entra a la SU, debe pasar a R, para luego esperar en la fila hasta que llegue su turno de pasar a C, posteriormente es enviado a L y de vuelta a C, donde según el diagnóstico final puede pasar a OYS o salir del sistema.

La SU está dividida en tres fases y cuatro estaciones, de las cuales la segunda es de suma importancia por la complejidad de su comportamiento (nótese que la segunda fase es un sistema, con la

particularidad de que los pacientes que son atendidos después de que van a L pasan con prioridad).

### Modelo de red de colas de la sala de urgencias

Descripción y supuestos del modelo. El diagrama de flujo de pacientes de la SU representado como una RC abierta y la simbología gráfica de esta disciplina se muestra en la figura 1, en donde se observa que se agrupan varias de las locaciones: en el caso de yesos y suturas en una locación con observación. La alta está representada por una S y todo lo que representa las áreas periféricas (unidad de cuidados, tercer nivel, piso y quirófano) son la letra E.

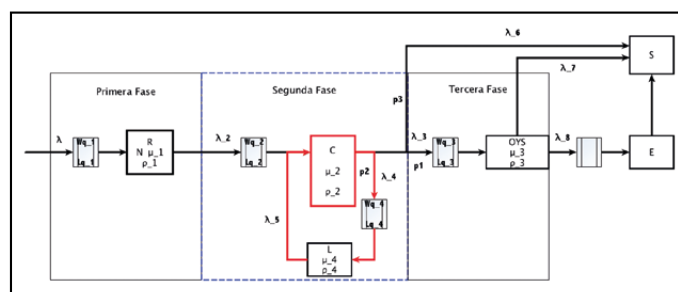


Figura 1. Diagrama de red de colas de la SU. Elaboración propia.

Un paciente puede pasar a la tercer fase con probabilidad  $p_1$ , a L con probabilidad  $p_2$  o directamente a S con probabilidad  $p_3$ . Los pacientes que entran a L regresan a C, pueden pasar a la tercer fase con probabilidad  $p_4$  o a S con probabilidad  $p_5$ . Se tiene que

$$p_1 + p_2 + p_3 = 1 \text{ y } p_4 + p_5 = 1. \quad (1)$$

Los valores de la ruta de probabilidades de los pacientes  $p_1$ ,  $p_2$  y  $p_3$  se describen en la tabla 2 (de resultados). La tasa de arribo de pacientes que entra al sistema por unidad de tiempo,  $\lambda_i$  significa el número promedio de pacientes que entran a la locación  $i$  por unidad de tiempo; así, la tasa de pacientes de entrada a los laboratorios es  $\lambda_4$  y su tasa de salida es  $\lambda_5$  (mismos que retornan a la estación 2). Un porcentaje  $p_2$  de los pacientes que pasan por la estación C pasa a la estación L, así que  $\lambda_4 = p_2 \lambda_2$ , y otro porcentaje  $p_3$  pasa a la estación S, así que  $\lambda_6 = p_3 \lambda_2$ .

La siguiente lista de notación y terminología de los indicadores de redes de colas es de Krieger (2008).

---

$\lambda$	Tasa media de llegadas.
$1/\lambda$	Tiempo medio entre llegadas
$\mu$	Tasa media de servicio.
$1/\mu$	Tiempo medio de servicio.
$L$	Número esperado de clientes en el sistema.
$L_q$	Número esperado de clientes en la cola <i>q-esima</i> .
$W$	Tiempo total de espera en el sistema
$W_q$	Tiempo promedio de espera en la cola <i>q-esima</i> .
$\pi_n$	Probabilidad de que el sistema esté ocupado (en estado estacionario)
$\pi_0$	Probabilidad de que el sistema esté vacío (en estado estacionario)
$N_i$	Número de servidores en la estación $i, i=1,2,\dots,s$
$\rho_i = \lambda_i / \mu_i$	Intensidad media de servicio en el servidor $i, i=1,2,\dots,s$
$\omega_i$	Intensidad media de servicio en la estación $i, i=1,2,\dots,s$
$\sigma^2$ o $var(S)$	Varianza de servicio solo para <i>M/G/1</i>

---

Por la naturaleza de las ecuaciones que se usarán de la TC en la RC del sistema que se propone se supondrá que hay una capacidad infinita en la cola de cada una de las estaciones, el sistema se encuentra en estado estacionario (patrón de llegadas al sistema que no cambian con el tiempo), además se considerará que el sistema está en estado estable (esto es  $\rho_i < 1$  para toda  $i$ ) (Gross & Harris, 1998, pp. 9, 33), así como que la red es un SC multiestado, en donde las fases 1 y 3 son SC tipo *M/M/1*. La fase dos es una RC con realimentación, en la cual el proceso de servicio en C es un modelo tipo *M/M/s* y el proceso de servicio en L es un modelo tipo *M/G/1*. En L se considera que tiene un tiempo de servicio con función de distribución general y media  $1/\mu_4$ , ya que gran parte del trabajo lo hace una máquina o equipo de laboratorio.

El patrón de llegada de los pacientes al sistema satisface una distribución de Poisson con parámetro  $\lambda$  (Vázquez Í. et al., 2014). Ya que el tiempo de recepción o registro de cada paciente es muy corto, se supondrá que  $\lambda_1 = \lambda$ . La tasa de llegada de pacientes es  $\lambda_4$  y la tasa de salida es  $\lambda_5$ . Como el sistema se encuentra en estado estable,  $\lambda_4 = \lambda_5$ . Los pacientes son atendidos secuencialmente de acuerdo con la regla de Primeras Entradas, Primeras Salidas (*PEPS*) antes de salir del sistema.

**Ecuaciones del modelo.** Si se tiene en cuenta la RC de la figura 1, la ecuación  $p_1 + p_2 + p_3 = 1$  y las consideraciones anteriores, se tienen las siguientes relaciones entre las tasas de llegadas de pacientes

de las estaciones del sistema:

$$\lambda = \lambda_1, \tag{2}$$

$$\lambda_3 = \lambda(1 - p_2 - p_3) = p_1 \lambda, \tag{3}$$

$$\lambda_2 = \lambda_1 + \lambda_5 = \lambda + \lambda_4, \tag{4}$$

$$\lambda_4 = p_2 \lambda_2. \tag{5}$$

Con apego a la notación de Gross y Harris (1998), las ecuaciones que estructuran el modelo *M/M/1* son:

$$\pi_0 = (1 - \rho) \quad (\rho = \lambda/\mu < 1), \tag{6}$$

$$\pi_n = (1 - \rho)\rho^n \quad (\rho = \lambda/\mu < 1), \tag{7}$$

$$L = \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda}, \tag{8}$$

$$W = \frac{L}{\lambda} = \frac{\rho}{\lambda(1 - \rho)} = \frac{1}{\mu - \lambda}, \tag{9}$$

$$W_q = \frac{L_q}{\lambda} = \frac{\rho}{\mu(1 - \rho)} = \frac{\rho}{(\mu - \lambda)}, \tag{10}$$

$$L_q = \frac{\rho^2}{(1 - \rho)} = \frac{\lambda^2}{\mu(\mu - \lambda)}. \tag{11}$$

Para el caso del sistema de los consultorios, modelo  $M/M/1$ , se tiene que  $s$  es el número de consultorios, sus ecuaciones son:

$$\pi_0 = \left[ \sum_{n=0}^{s-1} \frac{\omega^n}{n!} + \frac{\omega^s}{s!(1-\rho)} \right]^{-1}, \quad (12)$$

donde  $\omega = \frac{\lambda}{\mu}$  y  $\rho = \frac{\omega}{s} = \frac{\lambda}{s\mu}$ .

Además, se tiene

$$L = \frac{\lambda}{\mu} + \frac{\rho\omega^s}{s!(1-\rho)^2} \pi_0, \quad (13)$$

$$L_q = \frac{\omega^s \rho}{s!(1-\rho)^2} \pi_0 = L - \frac{\lambda}{\mu}, \quad (14)$$

$$W = \frac{1}{\mu} + \frac{\omega^s \rho}{s!(1-\rho)^2 \lambda} \pi_0, \quad (15)$$

$$W_q = \frac{L_q}{\lambda} = \frac{\rho\omega^s}{s!(1-\rho)^2 \lambda} \pi_0 = W - \frac{1}{\mu}. \quad (16)$$

El comportamiento del sistema  $M/G/1$  se determina por las ecuaciones de Pollaczek y Khinchine. Si se usa la notación de Gross y Harris (1998):

$$L = \frac{\rho^2 + \lambda^2 \text{var}(S)}{2(1-\rho)} + \rho, \quad (17)$$

$$W = \frac{\frac{\rho^2}{\lambda} + \lambda \text{var}(S)}{2(1-\rho)} + \frac{1}{\mu}, \quad (18)$$

$$W_q = \frac{\frac{\rho^2}{\lambda} + \lambda \text{var}(S)}{2(1-\rho)}, \quad (19)$$

$$L_q = \frac{\rho^2 + \lambda^2 \text{var}(S)}{2(1-\rho)}. \quad (20)$$

Un conjunto de estaciones interconectadas de SC en el que cualquier usuario al salir de una estación puede unirse a otra, o debe dejar el sistema total, se denomina RC. Las probabilidades de enrutamiento  $p_{ij}$  denotan la probabilidad que tiene un usuario de viajar desde la estación  $i$  a la estación  $j$ . Los tiempos de viaje, en general, se supone que son cero. La matriz generada por estas probabilidades,  $P=(p_{ij})$ , es llamada matriz de ruta de probabilidad de la red.

Las probabilidades de enrutamiento pueden ser dependientes del estado, en donde un estado de la red se define por el vector  $n=(n_1, \dots, n_M)$ , donde  $n_i$  es el número de clientes en las estaciones  $i=1, \dots, M$ . En nuestro caso, la matriz de ruta de probabilidades es:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{bmatrix}$$

En la figura 1 se observa que varias de las entradas de la matriz  $P$  son cero. Como solo importa analizar lo que pasa de las fases 1 a la 3 no se considerarán los valores de la ruta de la estación 3 a S o E, ni las salidas a la estación 5 (que es E), por lo cual solo se considerarán los valores de:  $p_{23} = p_1$ ,  $p_{24} = p_2$  y  $p_{25} = p_3$ .

## RESULTADOS

### Estimación de parámetros de la SU

El análisis del comportamiento de la SU requiere determinar los valores de los parámetros del modelo, la tasa de llegada  $\lambda$ , las probabilidades de enrutamiento  $p_{ij}$ , las tasas de servicio  $\mu_i$  y el número de servidores  $N_i$ . Los primeros fueron estimados a partir de estadísticos del hospital y los valores de  $N_i$  son parte de la información básica del sistema administrativo de la SU.

Los datos para el análisis de la demanda de pacientes en la SU fueron recabados a través de un muestreo en los libros de registro desde junio de 2009 hasta abril de 2012 (1,021 días) con un nivel de confianza de 99%. Para determinar la tasa de llegada de los pacientes se realizaron las pruebas estadísticas de bondad de ajuste no paramétricas (*Chi-cuadrada CC*); por la inconsistencia en la información y los registros incompletos, a la prueba *CC* se le aplicó la corrección de Yates para clases con frecuencia esperada menor a 5, agrupando por día (Vázquez I. et al., 2014). Con estos datos se obtuvo que la demanda se ajusta a una distribución Poisson con  $\lambda=106.713$ , con lo cual el tiempo entre llegadas tiene una función de distribución exponencial con  $\frac{1}{\lambda} = 13.494 \text{ min} = 4.446$  usuarios por hora. La tabla 1 muestra los parámetros del sistema.

Tabla 1  
Parámetros del sistema. Muestra los valores de las capacidades y características propias del sistema, así como el valor de la tasa de arribo general.

Parámetro	Valor
$N_1$	1
$N_2$	3
$N_3$	1
$N_4$	10
$\lambda$	4.446
$\mu_1$	12
$\mu_2$	2.4
$\mu_3$	2
$\mu_4$	2.2

Nota: Elaboración propia.

Los valores de  $var(S)$  fueron calculados tomando los promedios de pacientes atendidos por hora en los laboratorios de química sanguínea, examen general de orina y radiología; los que fueron 2.4, 3.0 y 4.9, respectivamente.

**Análisis numérico**

De la ecuación  $p_1+p_2+p_3=1$  se desprende que los valores de  $p_1$ ,  $p_2$  y  $p_3$  pueden variar entre 0 y 1. Sin embargo, para que el sistema se encuentre estable es necesario que  $\rho_3 = \lambda_3/\mu_3 < \alpha$  con  $\alpha$  un número menor que 1. Por la naturaleza del sistema se supondrá que su valor máximo es de 0.8. Si se considera este valor para  $\alpha$ , se toma en cuenta que  $\lambda_3=p_1 \lambda_2=p_1 (4.446)$  y que  $\mu_3=2$ , se obtiene que  $p_1=0.359$ . Además, se agrega la condición de que el valor de  $p_1$  no excederá 0.3, si se supone que máximo 30% de los pacientes pasan la locación 3. Por otro lado, si bien las condiciones de estabilidad indican que el valor de  $p_2$  es menor a 0.4, en las simulaciones se consideró que  $p_2$  puede variar entre 0.5 y 0.25, derivado de suponer que al menos 25% de los casos de urgencias requiere de estudios de laboratorio, pero no más de 50%.

Los resultados de las simulaciones realizadas variando los valores de  $p_1$ ,  $p_2$  y  $p_3$  en pasos de 0.05 se muestran en la tabla 2. Los tonos en gris corresponden a los valores de  $p_2$ .

Tabla 2  
Medidas de rendimiento del sistema

$p_1$	$p_2$	$p_3$	$I_2$	$L_2$	$W_2$	$L_{q2}$	$W_{q2}$	$I_3$	$L_3$	$W_3$	$L_{q3}$	$W_{q3}$	$I_4$	$L_4$	$W_4$	$L_{q4}$	$W_{q4}$
0.30	0.25	0.45	5.56	8.18	1.47	5.86	1.05	1.67	5.01	3.01	4.18	2.51	1.11	0.99	0.89	0.49	0.44
0.25	0.25	0.50	5.56	8.18	1.47	5.86	1.05	1.39	2.28	1.64	1.58	1.14	1.11	0.99	0.89	0.49	0.44
0.20	0.25	0.55	5.56	8.18	1.47	5.86	1.05	1.11	1.25	1.13	0.70	0.63	1.11	0.99	0.89	0.49	0.44
0.15	0.25	0.60	5.56	8.18	1.47	5.86	1.05	0.83	0.71	0.86	0.30	0.36	1.11	0.99	0.89	0.49	0.44
0.10	0.25	0.65	5.56	8.18	1.47	5.86	1.05	0.56	0.38	0.69	0.11	0.19	1.11	0.99	0.89	0.49	0.44
0.25	0.30	0.45	5.78	10.85	1.88	8.44	1.46	1.45	2.60	1.80	1.88	1.30	1.33	1.49	1.12	0.88	0.66
0.20	0.30	0.50	5.78	10.85	1.88	8.44	1.46	1.16	1.37	1.18	0.79	0.68	1.33	1.49	1.12	0.88	0.66
0.15	0.30	0.55	5.78	10.85	1.88	8.44	1.46	0.87	0.77	0.88	0.33	0.38	1.33	1.49	1.12	0.88	0.66
0.10	0.30	0.60	5.78	10.85	1.88	8.44	1.46	0.58	0.41	0.70	0.12	0.20	1.33	1.49	1.12	0.88	0.66
0.05	0.30	0.65	5.78	10.85	1.88	8.44	1.46	0.29	0.17	0.58	0.02	0.08	1.33	1.49	1.12	0.88	0.66
0.20	0.35	0.45	6.00	15.19	2.53	12.69	2.11	1.20	1.50	1.25	0.90	0.75	1.56	2.32	1.49	1.61	1.04
0.15	0.35	0.50	6.00	15.19	2.53	12.69	2.11	0.90	0.82	0.91	0.37	0.41	1.56	2.32	1.49	1.61	1.04
0.10	0.35	0.55	6.00	15.19	2.53	12.69	2.11	0.60	0.43	0.71	0.13	0.21	1.56	2.32	1.49	1.61	1.04
0.05	0.35	0.60	6.00	15.19	2.53	12.69	2.11	0.30	0.18	0.59	0.03	0.09	1.56	2.32	1.49	1.61	1.04
0.15	0.40	0.45	6.22	23.02	3.70	20.43	3.28	0.93	0.88	0.94	0.41	0.44	1.78	4.03	2.26	3.22	1.81
0.10	0.40	0.50	6.22	23.02	3.70	20.43	3.28	0.62	0.45	0.73	0.14	0.23	1.78	4.03	2.26	3.22	1.81
0.05	0.40	0.55	6.22	23.02	3.70	20.43	3.28	0.31	0.18	0.59	0.03	0.09	1.78	4.03	2.26	3.22	1.81
0.10	0.45	0.45	6.45	39.36	6.11	36.68	5.69	0.64	0.48	0.74	0.15	0.24	2.00	9.53	4.76	8.62	4.31
0.05	0.45	0.50	6.45	39.36	6.11	36.68	5.69	0.32	0.19	0.60	0.03	0.10	2.00	9.53	4.76	8.62	4.31

Nota: Elaboración propia.

La tabla muestra que las variables  $W$ 's y  $L$ 's de las estaciones 2 y 4 dependen directamente del valor de  $p_2$ . Para  $p_2 > 0.35$ , y aun cuando  $\lambda = \lambda_1 = 4.446$  (lo cual es considerado un valor no muy alto), los laboratorios L tienen dificultades para atender el flujo

de pacientes que requieren sus servicios (valores de  $W_4 > 1.49$  h). Cuando  $p_2 > 0.40$  el servicio en L se vuelve totalmente ineficiente (tiempo promedio total de más de 2 h y tiempo promedio de atención  $1/\mu_4 = 0.45$  h = 27 min).

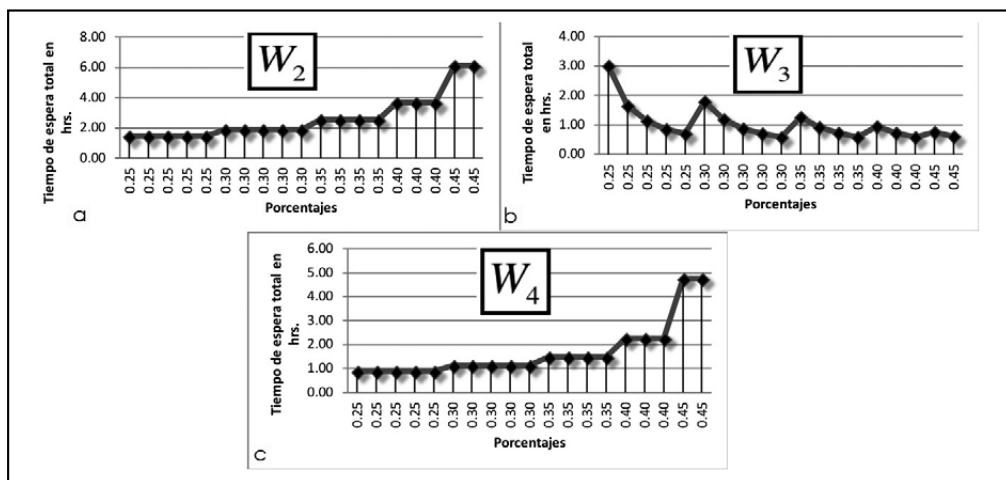


Figura 2. Las gráficas (a), (b) y (c) representan los tiempos de espera (TE) totales de (C), (OYS) y (L), respectivamente. El gráfico (a) muestra que al aumentar  $p_2$  se incrementa el TE en C. En (b), OYS el TE disminuye considerablemente cuando aumenta  $p_2$ . Sin embargo, (c) muestra que al aumentar  $p_2$ , provoca sobresaturación del TE en L. Elaboración propia.

La estación 2 presenta problemas de capacidad aun cuando los valores de  $p_2$  son bajos, ya que  $W_2$  siempre es elevado (tiempo promedio de al menos 88 min), si bien puede decirse que en el rango  $p_1 \leq 0.20$  y  $p_2 \leq 0.30$  el sistema está en situación aceptable de funcionalidad, ya que en otros rangos los tiempos de espera promedios son excesivos.

Del análisis anterior se desprende que tanto C como L (correspondientes a la segunda fase) son los que determinan la funcionalidad del sistema con base en  $p_2$ . La tabla 2 muestra los indicadores de eficiencia del sistema para cada variante de los valores posibles ( $p_1, p_2, p_3$ ), los cuales pueden ser interpretados de acuerdo con la figura 1. Por ejemplo, para el caso  $(p_1, p_2, p_3) = (0.10, 0.25, 0.65)$ , que se encuentra dentro del rango de valores de funcionalidad aceptable, se tiene que, en promedio, hay 2.32 pacientes recibiendo atención o en servicio de consulta ( $L_2 - Lq_2$ ), cada uno tarda 25.2 min en atención (de acuerdo con  $\mu_2$ ), en promedio hay seis pacientes aguardando ser atendidos un tiempo promedio de 63 min ( $Wq_2$ ). Además, en promedio hay un paciente cada 4 h recibiendo atención en OYS, y cada uno tarda 30 min (esto porque se considera la capacidad del sistema  $\mu_3 = 2$  y no el tiempo real de

la afección). Análogamente, en promedio, hay un paciente cada 2 h recibiendo atención, con tiempo promedio de 27 min, y un paciente esperando cada 2 h para ser atendido en L.

La figura 2 muestra la relación entre  $p_2$  y el tiempo de espera en las estaciones 2, 3 y 4. El gráfico (a) muestra que al aumentar  $p_2$  (es decir, al incrementar la demanda de atención en C por la realimentación) se incrementa el tiempo de espera en 2. Por otro lado, los incrementos en  $p_2$  generan efectos de decrecimiento en la estación 3 (b). Finalmente, en la estación 4 se provoca sobresaturación del sistema, lo que genera tiempo excesivo para las operaciones (c).

## DISCUSIÓN

El presente trabajo propone un modelo de RC abierta con realimentación simplificado, el cual difiere de los usados en las referencias citadas — en las cuales no toman en cuenta los escenarios posibles, las  $p$ 's están fijas— en la variación de los valores de la matriz de ruta de probabilidades. Los resultados de las simulaciones de la tabla fueron calculados con ecuaciones analíticas en Excel.



Figura 3. La imagen superior izquierda corresponde a la entrada a la SU, la que sigue a la derecha es la recepción. El área de espera se encuentra abajo a la izquierda y, finalmente, la imagen abajo a la derecha son los consultorios.

Fotografía del equipo de investigación.

Los resultados del análisis numérico muestran la utilización de la capacidad del sistema en un rango de escenarios.

El número esperado de pacientes en el sistema de colas de consultorios varía de ocho a 43 pacientes, en el área de laboratorios de uno a 10 pacientes; el tiempo medio de espera en la cola de consultorios va de 1 a 6 h y en laboratorios de 0.5 a 4.5 h. Esto demuestra que la longitud de las colas y los tiempos de espera son demasiado grandes, lo cual redundará en insatisfacción de los pacientes y mayor carga de trabajo para los empleados.

Con base en los resultados del modelo de la SU, durante ciertas horas se presenta congestión, si bien la tabla 2 con los parámetros de rendimiento de TC para diferentes ternas no refleja en qué momentos ocurre; debido a que el modelo no considera temporalidad, cada una de las ternas representa un escenario en el cual se encuentra el sistema en un periodo del día, semana, mes o año. Esta investigación complementa en cierto sentido la investigación de Cochran y Roche (2009), ya que en el presente estudio se determinaron rangos de valores (ternas) de la matriz de ruta de probabilidades, ya que intuitivamente la proporción de flujo de pacientes en las diversas locaciones no es algo estático.

Salvo la investigación de Zhu et al. (2013), en ninguna de las investigaciones presentadas en las referencias se considera la realimentación; aunque se presentan modelos de red de colas con tasas de llegada no homogéneas, diversos tipos de pacientes, capacidad de recursos (camas o enfermeras), sistemas con bloqueo, tiempos de ciclo de paciente o modelos jerárquicos. Parte de la investigación de Au-Yeung et al. (2006) es similar a la del presente trabajo, pues muestra una RC del departamento A&E, pero se enfoca en analizar y simular (mediante eventos discretos) combinaciones de priorización de varios tipos de pacientes para que no se afecte el tiempo de respuesta de arribos por ambulancia.

Los resultados de Jlassi et al. (2009) tienen similitud con el análisis de tiempos de espera en las estaciones; aunque como ellos tienen múltiples clases de pacientes, sus resultados muestran de manera más detallada el tipo de paciente y la locación. La investigación de Izady y Worthington (2012) tiene como objetivo generar una programación de horarios de personal para mejorar la eficiencia del servicio y reducir las estadías de los pacientes; en este sentido, la presente investigación se relaciona, pues la tabla 2 muestra en qué momentos es necesario tener más o menos personal para lograr dicho objetivo.



## CONCLUSIONES

Se presentó un modelo de RC diferente a los que se encontraron en las referencias, muestra diversos escenarios y la sensibilidad de la matriz de ruta de probabilidades. El modelo de RC que se desarrolló en la presente investigación proporciona un análisis rápido de un sistema, lo cual es bueno cuando los analistas únicamente necesitan información crítica del sistema para la toma de decisiones y estrategias.

A partir de los datos disponibles y en el contexto del tema más los supuestos se podría utilizar el modelo en la planificación de las actividades de la SU, mediante la tabla 2 de indicadores de eficiencia, para minimizar o eliminar la sobresaturación. El modelo y las simulaciones muestran que el nivel óptimo de funcionalidad de la SU ocurre para valores de  $p_2$  menores a 25%, o que si  $\mu_3$  es menor o  $p_1$  toma valores cercanos a 0.36 (es decir 36%) se congestiona el área de OYS, lo cual hace patente la necesidad de implementar estrategias de administración o médicas.

El modelo propuesto es susceptible de implementar variantes de investigación futura, desde considerar más atributos de los patrones de llegada de los pacientes, como a estos mismos o con supuestos más naturales. Los autores del presente creen que futuras extensiones de este trabajo pueden conducir al desarrollo de configuraciones de servicios que proporcionen una mejor correspondencia entre el nivel de atención y las necesidades de los clientes que los basados en las prácticas de toma de decisiones.

Los autores desean agradecer a los directivos del Centenario Hospital Miguel Hidalgo, en especial al personal relacionado con el servicio de urgencias médicas, por su disposición, apertura y tiempo en cada una de las reuniones, así como por el acceso a la información. Se agradece el apoyo de PROMEP 103.5/10/4547 y 103.5/16/8952 por parte de los autores Rodolfo Rafael Medina Ramírez y José Antonio Vázquez Ibarra. Asimismo, el equipo de investigación agradece a los evaluadores anónimos por sus constructivos comentarios y sugerencias.

**REFERENCIAS**

- Au-Yeung, S. W. M., Harrison, P. G., & Knottenbelt, W. J. (2006). A queueing network model of patient flow in an accident and emergency department. *Proceedings of the 20th Annual European and Simulation Modelling Conference*, 60-67.
- Cochran, J. K., & Roche, K. T. (2009). A multi-class queueing network analysis methodology for improving hospital emergency department performance. *Computers and Operations Research*, 36(5), 1497-1512. doi: 10.1016/j.cor.2008.02.004
- Côté, M. J. (2000). Understanding patient flow. *Decision Line*, 31(2), 8-10.
- Derlet, R. W., Richards, J. R., & Kravitz, R. L. (2008). Frequent overcrowding in US emergency departments. *Academic Emergency Medicine*, 8(2), 151-155. doi: 10.1111/j.1553-2712.2001.tb01280.x
- Filipowicz, B., & Kwiecień, J. (2008). Queueing systems and networks. Models and applications. *Bulletin of the Polish Academy of Sciences, Technical Sciences*, 56(4), 379-390.
- Gómez Dantés, O., Sesma, S., Becerril, V. M., Knaul, F. M., Arreola, H., & Frenk, J. (2011). Sistema de salud de México. *Salud Pública de México*, 53(Supl. 2), S220-S232. Recuperado de <http://www.scielo.org.mx/pdf/spm/v53s2/17.pdf>
- Gross, D., & Harris, C. M. (1998). *Fundamentals of queueing theory* (3<sup>rd</sup> ed.). John Wiley & Sons. doi: 10.1002/9781118625651
- Hillier, F. S., & Lieberman, G. J. (2001). *Introduction to Operation Research* (7th ed.). New York, NY: McGraw Hill.
- Hospitalesmexico.com © Datos públicos. (s. f.). Directorio hospitales, clínicas y consultorios en México [Portal electrónico]. Recuperado de <https://hospitalesmexico.com/>
- Izady, N., & Worthington, D. (2012). Setting staffing requirements for time dependent queueing networks: The case of accident and emergency departments. *European Journal of Operational Research*, 219(3), 531-540. doi: 10.1016/j.ejor.2011.10.040
- Jiang, L., & Giachetti, R. E. (2008). A queueing network model to analyze the impact of parallelization of care on patient cycle time. *Health Care Management Science*, 11(3), 248-261. doi: 10.1007/s10729-007-9040-9
- Jlassi, J., Mhamedi, A. E., & Chabchoub, H. (2009). Networks of queues with multiple customer types: Application in emergency departments. *International Journal of Behavioural and Healthcare Research*, 1(4), 400. doi: 10.1504/IJBHR.2009.032157
- Krieger, U. R. (2008). *Queueing networks and Markov chains*, 2nd edition by G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi. John Wiley & Sons, Hoboken, NJ, 2006, 878 pages, ISBN 0-471-56525-3. Book review. *IIE Transactions*, 40(5), 567-568. doi: 10.1080/07408170701623187
- Mayhew, L., & Smith, D. (2008). Using queueing theory to analyse the government's 4-H completion time target in accident and emergency departments. *Health Care Management Science*, 11(1), 11-21. doi: 10.1007/s10729-007-9033-8
- Olorunsola, S. A., Adeleke, R. A., & Ogunlade, T. O. (2014). Queueing analysis of patient flow in hospital. *IOSR Journal of Mathematics*, 10(4) Ver. VI, 47-53.
- Vázquez I., J. A., González S., R. A., & Juárez L., H. A. (2014). *Modelado computacional de un servicio de urgencias considerando las variables relacionadas con el factor humano y con los tiempos de atención en consulta de acuerdo con el diagnóstico* (Tesis doctoral). Centro Universitario de los Lagos-UdeG, México.
- Véricourt, F. de, & Jennings, O. B. (2011). Nurse staffing in medical units: A queueing perspective. *Operations Research*, 59(6), 1320-1331. doi: 10.1287/opre.1110.0968
- Zhu, H., Gong, J., & Tang, J. (2013). A queueing network analysis model in emergency departments. *IEEE Xplore*, 1829-1834.