

Modelo matemático mediante líneas de espera para el desarrollo de un simulador

Guillermo Durán*, Gregorio Glas†, Gastón Freire†, Adriana Galli†, María del Carmen Pérez†, Martín Rados*, Javier Marengo*,‡, Cristián E. Cortés§, Raúl Espinoza§, Pablo A. Rey§

*Universidad de Buenos Aires

Email: gduan@dm.uba.ar, radosm@gmail.com, jmarengo@dc.uba.ar

†Universidad Nacional de Avellaneda

Email: gglas@undav.edu.ar, gastonandresfreire@yahoo.com.ar, agalli@undav.edu.ar, mdc@undav.edu.ar

‡Universidad Nacional de General Sarmiento

§Universidad de Chile

Email: ccortes@ing.uchile.cl, raul.espinoza.muirhead@gmail.com, pablo.rey.cl@gmail.com

Abstract—El presente proyecto se desarrolló en el marco de la convocatoria a proyectos cofinanciados UBA-UNDAV del año 2012. El objetivo fue proveer una herramienta para la mejora de la gestión del mantenimiento del casco céntrico de una ciudad mediana para lo que se decidió desarrollar un simulador. Con los datos del relevamiento inicial se confeccionó un modelo matemático que consistió en un conjunto de variables aleatorias de acuerdo a la Teoría de Colas para alimentar a un simulador de tránsito. Los datos relevados permitieron elaborar los estimadores (media y varianza muestral) para los parámetros de las variables aleatorias definidas. Este modelo matemático proveyó de lógica al simulador. éste fue desarrollado para SUMO (Simulation of Urban Mobility) versión 0.25, mediante un script Python que interactúa con la plataforma, más un conjunto de utilitarios entre los que se cuentan editores y graficadores.

I. INTRODUCCIÓN

Marco del proyecto

El proyecto se desarrolló bajo la convocatoria a proyectos de investigación cofinanciados UBA-UNDAV del año 2012. El desarrollo de la investigación fue llevado a cabo por docentes de ambas universidades, un estudiante (tesista) y contó con la colaboración de expertos en temas de Transporte de la Universidad de Chile.

El problema en estudio fue la acumulación de pasajeros en espera en las veredas de las calles céntricas de la localidad de Avellaneda ¹durante horas pico, en donde se producen líneas de espera indiferenciadas con ocupación total de las veredas.

Caracterización del problema y de la solución

El casco céntrico de Avellaneda se caracteriza por una actividad comercial y gastronómica intensa y una gran concentración de edificios de viviendas, oficinas, entidades bancarias y colegios. Se estima que en hora pico se produce una afluencia de aproximadamente 3000 pasajeros en 5 cuadras cuyas veredas están parcialmente ocupadas por mesas de locales gastronómicos, kioscos de diarios y revistas, puestos de

¹Partido de Avellaneda, Provincia de Buenos Aires (Argentina)

flores, cabinas telefónicas entre otros obstáculos que reducen el área de circulación y espera.

Para arribar a una solución se desarrolló un simulador que permitiera ensayar escenarios y observar los efectos de posibles cambios. Este simulador se basó en un modelo matemático desarrollado mediante un sistema de colas cuyos parámetros se calcularon en base al relevamiento de datos.

II. DATOS PARA EL MODELO MATEMÁTICO BASE DEL SIMULADOR.

Relevamiento de datos

Se realizaron dos tipos de relevamiento de datos: 1) ubicación de las paradas de colectivos y 2) afluencia de pasajeros y colectivos.

Para el primer caso se midió la ubicación de las paradas de colectivos, la distancia entre ellas, la distancia desde la esquina y las líneas de colectivo en cada parada. Estos datos se utilizaron para parametrizar el mapa de entrada del simulador de la zona en estudio.

Acerca de las mediciones sobre afluencia de pasajeros y colectivo, se tomó fecha y hora de la observación y, por cada parada en la zona relevada se registró: horario de llegada de cada línea, el tiempo de detención, la cantidad de pasajeros en ascenso y descenso, una estimación cualitativa de la ocupación del colectivo y la cantidad de pasajeros que permanecían en la parada una vez que el colectivo había partido. Se registraron 386 mediciones y con ellas se confeccionaron los estimadores para los parámetros de las variables aleatorias que conformaron el modelo matemático.

III. MODELO MATEMÁTICO

Para poder mostrar escenarios realistas durante la simulación era necesario introducir aleatoriedad a las variables en juego, interviniendo tiempos de espera y frecuencias, tanto en lo que respecta a la llegada de vehículos como en el tiempo de ascenso y partida de los vehículos observándose que estos valores diferían según la línea de colectivos.

Lo mismo para la llegada de pasajeros a las líneas de espera que varía a lo largo del día. El primer problema a resolver

era establecer entonces la distribución de probabilidad más adecuada para las variables aleatorias. SIO, Simposio Argentino de Investigación Operativa

El equipo así consideró que el problema de ocupación de las veredas podía ser analizado desde la Teoría de Colas en la que los tiempos de espera, el arribo de pasajeros y vehículos y la atención de clientes (ascenso al colectivo) podía ser encuadrado como procesos de nacimiento y muerte definiendo variables aleatorias exponenciales de parámetro lambda en la mayoría de los casos y conformando un sistema complejo formado por varias colas, factores aleatorios y resultados determinísticos.

A. Elección del modelo

En el presente trabajo se ha hecho foco en el flujo de personas que llegan a las paradas de colectivos y en cómo los colectivos que van arribando a las mismas van recogiendo pasajeros, proceso en el cual se genera una dinámica que puede modelarse matemáticamente mediante lo que se denomina "Procesos de Nacimiento y Muerte". Los nacimientos se deben a la llegada de pasajeros hacia las paradas, y las muertes aluden a aquellas personas que salen de las mismas como consecuencia de abordar el colectivo que esperaban. Un proceso de nacimiento y muerte alude a un proceso estocástico $(N_t)_{t \geq 0}$ en tiempo continuo y con espacio de estados discreto $E = \mathbb{N}_0$, en el cual el sistema se encuentra en estado E_n en el instante t cuando $N_t = n$ y la probabilidad de que N_t tome el valor n está dada por $P_n(t) = P(N_t = n)$. En su expresión más general, el proceso estocástico se dirá de nacimiento y muerte si y sólo si existen dos sucesiones de números reales no negativos $(\lambda_n)_{n \in \mathbb{N}}$ y $(\mu_n)_{n \in \mathbb{N}}$ tales que se cumplen las siguientes condiciones:

- 1) Los cambios de estado permitidos son desde E_0 hasta E_1 o desde E_n hasta E_{n-1} o E_{n+1} para todo $n \in \mathbb{N}$.
- 2) Si el sistema se encuentra en estado E_n en el instante t , la probabilidad de que pase al estado E_{n+1} en algún momento entre t y $t+h$ es $\lambda_n \cdot h + o(h)$. Y si $n \geq 1$ la probabilidad de que pase al estado E_{n-1} en algún momento entre t y $t+h$ es $\mu_n \cdot h + o(h)$, donde $o(h)$ a cualquier función de una variable verificando:

$$\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$$

- 3) La probabilidad de que el sistema tenga más de un cambio de estado entre los instantes t y $t+h$ es $o(h)$.

La distribución de probabilidades de cada N_t , es decir $P(N_t = n) = P_n(t)$, es un proceso diferencial que evoluciona a través del tiempo, pudiéndose encontrar una relación entre lo que ocurre en el instante $t+h$ y lo que ocurre en el instante t , a partir de una suerte de cociente incremental que permita luego estudiar la variación de la probabilidad en un instante t . Y una vez conocida la variación entre los instantes t y $t+h$ es posible obtener el cociente incremental para $P_n(t)$, cosa que permite establecer una relación entre $P'_n(t)$, $P_n(t)$, $P_{n-1}(t)$ y $P_{n+1}(t)$.

Como según el Teorema de la Probabilidad Total:

$$\begin{aligned} P_n(t+h) &= P(N_{t+h} = n) \\ &= P(N_{t+h} = n | N_t = n-1) \cdot P(N_t = n-1) + \\ &+ P(N_{t+h} = n | N_t = n) \cdot P(N_t = n) \\ &+ P(N_{t+h} = n | N_t = n+1) \cdot P(N_t = n+1) + \\ &+ \sum_{m \in \mathbb{N}_0 \setminus \{n-1, n, n+1\}} P(N_{t+h} = n | N_t = m) \cdot P(N_t = m) \end{aligned}$$

Como:

$$\begin{aligned} &P(N_{t+h} = n | N_t = n) \cdot P(N_t = n) \\ &= (1 - (\lambda_n h + o(h)) - (\mu_n h + o(h)) - o(h)) \cdot P(N_t = n) \\ &= P(N_t = n) - \left(\lambda_n + \mu_n + \frac{o(h)}{h} \right) h \cdot P(N_t = n) \end{aligned}$$

entonces:

$$\begin{aligned} P_n(t+h) &= \\ &(\lambda_{n-1} h + o(h)) \cdot P_{n-1}(t) + P_n(t) - \\ &\left(\lambda_n + \mu_n + \frac{o(h)}{h} \right) h \cdot P_n(t) + \\ &(\mu_{n+1} h + o(h)) \cdot P_{n+1}(t) + o(h) \end{aligned}$$

y puede reescribirse la igualdad anterior obtenemos el cociente incremental:

$$\begin{aligned} \frac{P_n(t+h) - P_n(t)}{h} &= \\ &= \left(\lambda_{n-1} + \frac{o(h)}{h} \right) \cdot P_{n-1}(t) - \\ &- \left(\lambda_n + \mu_n + \frac{o(h)}{h} \right) \cdot P_n(t) + \\ &+ \left(\mu_{n+1} + \frac{o(h)}{h} \right) \cdot P_{n+1}(t) + \frac{o(h)}{h} \end{aligned}$$

Tomando $\lim_{h \rightarrow 0}$ en la igualdad, del lado izquierdo surge $P'_n(t)$ y del lado derecho todos los términos $\frac{o(h)}{h}$ tienden a cero, obteniendo:

$$P'_n(t) = \lambda_{n-1} P_{n-1}(t) - (\lambda_n + \mu_n) P_n(t) + \mu_{n+1} P_{n+1}(t)$$

De esta forma el sistema de ecuaciones diferenciales que gobierna las distribuciones de N_t es:

$$P'_n(t) = \begin{cases} \lambda_{n-1} P_{n-1}(t) - (\lambda_n + \mu_n) P_n(t) + \mu_{n+1} P_{n+1}(t) & , \text{ si } n \geq 1 \\ -\lambda_0 P_0(t) + \mu_1 P_1(t) & , \text{ si } n = 0 \end{cases}$$

con las condiciones iniciales usuales:

$$P_0(0) = 1 \quad P_n(0) = 0 \quad (\forall n \geq 1)$$

B. Elección de variables exponenciales para las colas

Supongamos que no hubiera muertes, es decir $\mu_n = 0$ para todo $n \geq 1$ y que la tasa de nacimientos fuera independiente de la población, es decir $\lambda_n = \lambda$ para todo $n \in \mathbb{N}_0$. En ese caso el sistema se reduce a:

$$P'_n(t) = \begin{cases} \lambda P_{n-1}(t) - \lambda P_n(t) & , \text{ si } n \geq 1 \\ -\lambda P_0(t) & , \text{ si } n = 0 \end{cases}$$

Si $n = 0$ entonces:

$$P'_0(t) = -\lambda P_0(t)$$

La solución general de esta ecuación diferencial se obtiene muy fácilmente y es:

$$P_0(t) = c \cdot e^{-\lambda t}$$

Además como $P_0(0) = 1$ entonces debe ser $c = 1$ por lo que:

$$P_0(t) = e^{-\lambda t}$$

Por inducción en n puede probarse que la solución para n en general es:

$$P_n(t) = \frac{e^{-\lambda t} (\lambda t)^n}{n!}$$

C. Procesos de nacimiento y muerte

Consideremos un proceso estocástico $(N_t)_{t \geq 0}$ en tiempo continuo y con espacio de estados discreto $E = \mathbb{N}_0$. Diremos que el sistema se encuentra en estado E_n en el instante t cuando $N_t = n$ y denotemos por $P_n(t) = P(N_t = n)$. El proceso se dirá de *nacimiento* y *muerte* si y sólo si existen dos sucesiones de números reales no negativos $(\lambda_n)_{n \in \mathbb{N}}$ y $(\mu_n)_{n \in \mathbb{N}}$ tales que se cumplen las siguientes condiciones:

- 1) Los cambios de estado permitidos son desde E_0 hasta E_1 o desde E_n hasta E_{n-1} o E_{n+1} para todo $n \in \mathbb{N}$.
- 2) Si el sistema se encuentra en estado E_n en el instante t , la probabilidad de que pase al estado E_{n+1} en algún momento entre t y $t+h$ es $\lambda_n \cdot h + o(h)$. Y si $n \geq 1$ la probabilidad de que pase al estado E_{n-1} en algún momento entre t y $t+h$ es $\mu_n \cdot h + o(h)$.
- 3) La probabilidad de que el sistema tenga más de un cambio de estado entre los instantes t y $t+h$ es $o(h)$.

1) *Distribución de probabilidades de N_t* : Para determinar la distribución de probabilidades de cada N_t , es decir $P(N_t = n) = P_n(t)$, no podemos hacerlo directamente debido a que se trata de un proceso diferencial que evoluciona a través del tiempo. Lo que sí podemos hacer es encontrar una relación entre lo que ocurre en el instante $t+h$ y lo que ocurre en el instante t , una suerte de cociente incremental que permita luego estudiar la variación de la probabilidad en un instante t .

Una vez conocida la variación entre los instantes t y $t+h$ es posible obtener el cociente incremental para $P_n(t)$, cosa que permite establecer una relación entre $P'_n(t)$, $P_n(t)$, $P_{n-1}(t)$ y $P_{n+1}(t)$.

Para empezar si utilizamos el teorema de la probabilidad total obtenemos:

$$\begin{aligned} M_n &= \mathbb{N}_0 - \{n-1, n, n+1\} \\ P_n(t+h) &= P(N_{t+h} = n) \\ &= P(N_{t+h} = n | N_t = n-1) \cdot P(N_t = n-1) \\ &\quad + P(N_{t+h} = n | N_t = n) \cdot P(N_t = n) \\ &\quad + P(N_{t+h} = n | N_t = n+1) \cdot P(N_t = n+1) \\ &\quad + \sum_{m \in M_n} P(N_{t+h} = n | N_t = m) \cdot P(N_t = m) \end{aligned}$$

El primer término contempla la probabilidad de que haya habido un nacimiento entre el instante t y $t+h$. El segundo de que el sistema no hubiera cambiado, el tercero contempla

la posibilidad de una muerte, y por último, la sumatoria contempla que hubiera habido más de un cambio.

Si tenemos presente que:

$$\begin{aligned} &P(N_{t+h} = n | N_t = n) \cdot P(N_t = n) \\ &= (1 - (\lambda_n h + o(h)) - (\mu_n h + o(h)) - o(h)) \cdot P(N_t = n) \\ &= P(N_t = n) - \left(\lambda_n + \mu_n + \frac{o(h)}{h} \right) h \cdot P(N_t = n) \end{aligned}$$

entonces:

$$\begin{aligned} P_n(t+h) &= (\lambda_{n-1} h + o(h)) \cdot P_{n-1}(t) \\ &\quad + P_n(t) - \left(\lambda_n + \mu_n + \frac{o(h)}{h} \right) h \cdot P_n(t) \\ &\quad + (\mu_{n+1} h + o(h)) \cdot P_{n+1}(t) + o(h) \end{aligned}$$

Reescribiendo la igualdad anterior obtenemos el cociente incremental:

$$\begin{aligned} \frac{P_n(t+h) - P_n(t)}{h} &= \left(\lambda_{n-1} + \frac{o(h)}{h} \right) \cdot P_{n-1}(t) \\ &\quad - \left(\lambda_n + \mu_n + \frac{o(h)}{h} \right) \cdot P_n(t) \\ &\quad + \left(\mu_{n+1} + \frac{o(h)}{h} \right) \cdot P_{n+1}(t) + \frac{o(h)}{h} \end{aligned}$$

Tomando $\lim_{h \rightarrow 0}$ en la igualdad, del lado izquierdo surge $P'_n(t)$ y del lado derecho todos los términos $\frac{o(h)}{h}$ tienden a cero, obteniendo:

$$P'_n(t) = \lambda_{n-1} P_{n-1}(t) - (\lambda_n + \mu_n) P_n(t) + \mu_{n+1} P_{n+1}(t)$$

En el caso de que $n = 0$ el término λ_{n-1} no está presente y μ_0 tampoco porque no puede haber una muerte cuando aún no hay nacimientos:

$$P'_0(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t)$$

Por lo tanto, el sistema de ecuaciones diferenciales que gobierna las distribuciones de N_t es:

$$P'_0(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t)$$

$$P'_n(t) = \lambda_{n-1} P_{n-1}(t) - (\lambda_n + \mu_n) P_n(t) + \mu_{n+1} P_{n+1}(t), \text{ si } n \geq 1$$

En general, a tiempo inicial $N_0 = 0$, por lo que las condiciones iniciales para este sistema de ecuaciones diferenciales es:

$$P_0(0) = 1 \quad P_n(0) = 0 \quad (\forall n \geq 1)$$

2) *Análisis del proceso estacionario*: La fase más importante de un proceso de nacimiento y muerte suele ser la estacionaria, que es cuando el sistema se encuentra adaptado a las tasas de natalidad y mortalidad y su evolución se independiza del tiempo t .

Si asumimos que en el intervalo de tiempo $[a, b]$ $P_n(t_1) = P_n(t_2)$ para todo $t_1, t_2 \in [a, b]$, entonces $\forall t \in [a, b]$

$P_n(t) = p_n$ y además $P'_n(t) = 0$ para todo $t \in (a, b)$. Las ecuaciones se reducen a:

$$\begin{aligned} (\lambda_n + \mu_n) \cdot p_n &= \lambda_{n-1} \cdot p_{n-1} + \mu_{n+1} \cdot p_{n+1} \\ p_1 &= \frac{\lambda_0}{\mu_1} \cdot p_0 = \end{aligned}$$

Para el caso $n = 1$ vemos que utilizando la ecuación para p_1 en la de más arriba:

$$\begin{aligned} (\lambda_1 + \mu_1) \cdot p_1 &= \lambda_0 \cdot p_0 + \mu_2 \cdot p_2 \\ \Leftrightarrow (\lambda_1 + \mu_1) \cdot \frac{\lambda_0}{\mu_1} p_0 &= \lambda_0 \cdot p_0 + \mu_2 \cdot p_2 \\ \Leftrightarrow \frac{\lambda_1 \lambda_0}{\mu_1} p_0 + \cancel{\lambda_0 p_0} &= \cancel{\lambda_0 p_0} + \mu_2 \cdot p_2 \end{aligned}$$

de donde podemos despejar p_2 como:

$$p_2 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} \cdot p_0$$

Vemos que salta a la vista la forma general de p_n y conjeturamos que para todo $n \geq 1$:

$$p_n = \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1} \cdot p_0$$

cosa que demostraremos por inducción en n .

El caso $n = 1$ ya está establecido. Supongamos ahora que la propiedad es válida para $k \leq n$ y demostremos que entonces vale para $n + 1$:

$$\begin{aligned} (\lambda_n + \mu_n) \cdot p_n &= \lambda_{n-1} \cdot p_{n-1} + \mu_{n+1} \cdot p_{n+1} \\ (\lambda_n + \mu_n) \cdot \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1} \cdot p_0 &= \lambda_{n-1} \cdot \frac{\lambda_{n-2} \cdots \lambda_0}{\mu_{n-1} \cdots \mu_1} \cdot p_0 + \mu_{n+1} \cdot p_{n+1} \\ \frac{\lambda_n \cdots \lambda_0}{\mu_n \cdots \mu_1} \cdot p_0 + \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_{n-1} \cdots \mu_1} \cdot p_0 &= \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_{n-1} \cdots \mu_1} \cdot p_0 + \mu_{n+1} \cdot p_{n+1} \\ \mu_{n+1} \cdot p_{n+1} &= \frac{\lambda_n \cdots \lambda_0}{\mu_n \cdots \mu_1} \cdot p_0 \\ p_{n+1} &= \frac{\lambda_n \cdots \lambda_0}{\mu_{n+1} \cdots \mu_1} \cdot p_0 \end{aligned}$$

Vemos entonces que la propiedad vale también para $n + 1$ y eso demuestra que es válida para todo $n \in \mathbb{N}$.

Si para cada $n \in \mathbb{N}$ definimos las constantes:

$$c_n = \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1} \quad \Rightarrow \quad p_n = c_n \cdot p_0$$

Además como:

$$\sum_{n=0}^{+\infty} p_n = 1$$

ya que $p_n = P(N = n)$ y la suma de todas ellas debe ser igual a 1, entonces:

$$\begin{aligned} p_0 + p_0 \sum_{n=1}^{+\infty} c_n &= 1 \\ \Leftrightarrow p_0 \left(1 + \sum_{n=1}^{+\infty} c_n \right) &= 1 \end{aligned}$$

Vemos entonces que:

$$p_0 = \frac{1}{1 + \sum_{n=1}^{+\infty} c_n}$$

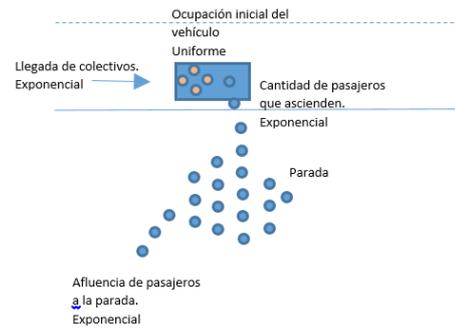


Fig. 1. Modelo matemático.

3) *Condición necesaria y suficiente para la estacionariedad:* Para que p_0 esté bien definido, vemos que entonces es necesario que:

$$\sum_{n=1}^{+\infty} \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1}$$

sea convergente. Por lo tanto una condición necesaria para la estacionariedad es que dicha serie converja. Lo interesante además, es que también vale la recíproca: si la serie converge entonces el proceso de nacimiento y muerte es estacionario.

Por lo tanto la condición necesaria y suficiente para la estacionariedad de un proceso de nacimiento y muerte es la convergencia de dicha serie.

D. Variables aleatorias y estimadores

Se consideró un sistema en el que confluyen distintas variables. La Figura 1 muestra un esquema de las variables consideradas.

Se definieron así las siguientes variables aleatorias:

- Afluencia de pasajeros a la parada (por parada): la línea de espera es un atributo de la parada. El uso de una parada está relacionado con la proximidad de bancos o colegios, con lo que se define una variable por parada. Se mide en personas que llegan por segundo. Esta variable es ajustada por un trapezoido para simular la creciente congestión en hora pico y su extinción paulatina. La población es fija, pues los clientes del sistema de colas (pasajeros) es una cantidad finita que son las personas que trabajan en bancos, comercios y colegios y los estudiantes que finalizan el horario lectivo.
- Llegada de colectivos a la parada (por línea de colectivo): un mismo vehículo transita las distintas paradas con lo que se ha definido que esta variable aleatoria se defina por línea y está relacionada con la frecuencia de la línea. Mide los segundos que demora en llegar el próximo colectivo con lo que su unidad de medida es segundos.
- Cantidad de pasajeros que ascienden (parada, línea de colectivo): Determina la cantidad de pasajeros que suben a un colectivo. Su unidad de medida es personas por

TABLA I
OCUPACIÓN INICIAL DE LOS VEHÍCULOS

| Categoría cualitativa | Rango cuantitativo (número de pasajeros) |
|-----------------------------|--|
| Vacío | 0 a 5 |
| Con asientos libres | 6 a 20 |
| Con lugar pero sin asientos | 26 a 40 |
| Lleno | 41 a 64 |
| Tan lleno que no se detiene | Más de 65 |

TABLA II
VARIABLES ALEATORIAS DEL MODELO.

| Variable/Factor | Relacionada con | Distribución |
|-------------------------------------|----------------------------|--------------|
| Afluencia de pasajeros | Parada | Exponencial |
| Llegada de colectivos a la parada | Línea de colectivos | Exponencial |
| Cantidad de pasajeros que ascienden | Línea de colectivo, Parada | Exponencial |
| Tiempo de ascenso | General | Normal |
| Ocupación inicial del vehículo | General | Uniforme |

segundo y está relacionada con el tiempo que transcurrió desde el último colectivo arribado.

- Tiempo de ascenso (único): representa el tiempo promedio que demora una persona en ascender y que afecta al tiempo total de detención del vehículo. Está relacionado con la cantidad de pasajeros que suben.
- Ocupación inicial del vehículo (general): indica el grado de ocupación del vehículo al inicio y condiciona la cantidad de pasajeros que pueden ascender efectivamente al vehículo. Si bien distintos vehículos pueden albergar a distintas cantidades de pasajeros sentados y de pie, se ha estimado una capacidad única de 75 pasajeros. En cuanto a la ocupación, su probabilidad se estableció en base a mediciones cualitativas observacionales según la correspondencia en Tabla I.

La Tabla II resume las variables aleatorias definidas.

Posteriormente el script Python implementó la orquestación e interacción de las variables citadas.

E. Validación del modelo

La validación del modelo se realizó refinando el modelo para obtener resultados compatibles con las observaciones. Esta etapa fue extensa y condujo a diversos ajustes y redefiniciones, entre ellos refinamiento en el cálculo de los parámetros de las distribuciones. En esta búsqueda es que se propuso el planteamiento de un polígono para simular el congestionamiento y descongestionamiento de la hora pico. Básicamente se hicieron tres tipos de controles:

- a) de los resultados arrojados por el simulador contra los valores observados (por planilla),
- b) comparativos dentro de la simulación. c) de evolución de

TABLA III
COMPARACIÓN ENTRE SIMULACIONES Y OBSERVACIONES.

| Métrica | Promedio de ejecuciones | Observado |
|----------------------------------|-------------------------|-----------|
| Vehículos en sistema | 228,75 | 219 |
| Pasajeros atendidos | 2580,85 | 2844 |
| Pasajeros en espera al finalizar | 73 | 113 |

los datos a lo largo de la simulación (gráficos)

Dentro del primer tipo, las métricas que se definieron fueron:

- 1) Cantidad total de pasajeros: observado versus promedio de simulaciones
- 2) Total de colectivos que finalizaron el recorrido (indicador de congestionamiento): sólo simulador como control cualitativo
- 3) Pasajeros que quedaron en las paradas al finalizar la observación/simulación.

La Tabla III muestra los valores obtenidos por simulación (promedios) y relevamiento.

La Figura 2 muestra la comparación entre los valores obtenidos y los relevados.

Dentro del grupo de valores comparativos, las métricas para hacer los controles de coherencia de los resultados de la ejecución fueron: 1) Total de colectivos en el sistema: observado versus promedio de simulaciones

2) Cantidad de pasajeros en los vehículos: otro valor controlado es la ocupación final del colectivo en pasajeros.

Estos valores son calculados por el script de simulación y registrados en el log de corrida.

Un ejemplo de lote de ejecuciones con sus datos de control es la que puede observarse en la Tabla IV.

Dentro del último grupo, las métricas definidas fueron:

- 1) Curva de la evolución de la cantidad de pasajeros en las paradas: permite observar el crecimiento o decrecimiento de la cantidad de pasajeros en espera. Se confecciona un gráfico por parada y los valores corresponden a los lambda de la parada.
- 2) Curva de la duración del recorrido: permite detectar posibles congestionamientos que afecten a la duración del recorrido en forma creciente. Se confecciona un gráfico por línea de colectivo.

Un ejemplo de estas curvas son las de la Figura 3 en donde se muestran los distintos valores que toman las variables correspondientes a lo largo de la ejecución.

IV. DESARROLLO DEL SIMULADOR

Para el desarrollo del simulador se seleccionaron herramientas de software libre. En primer lugar, se eligió SUMO², una plataforma para simulaciones de tráfico que ofrece una serie de APIs; en particular se utilizó la versión 0.25 que introduce una mejora a la api de TraCI³ necesaria para interactuar con

²SUMO: es una suite de simulación de tráfico. http://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931_read-41000/

³TraCI: "Traffic Control Interface" ofrece acceso a la simulación de tráfico recuperando sus variables y permitiendo manipular su comportamiento. <http://sumo.dlr.de/wiki/TraCI>

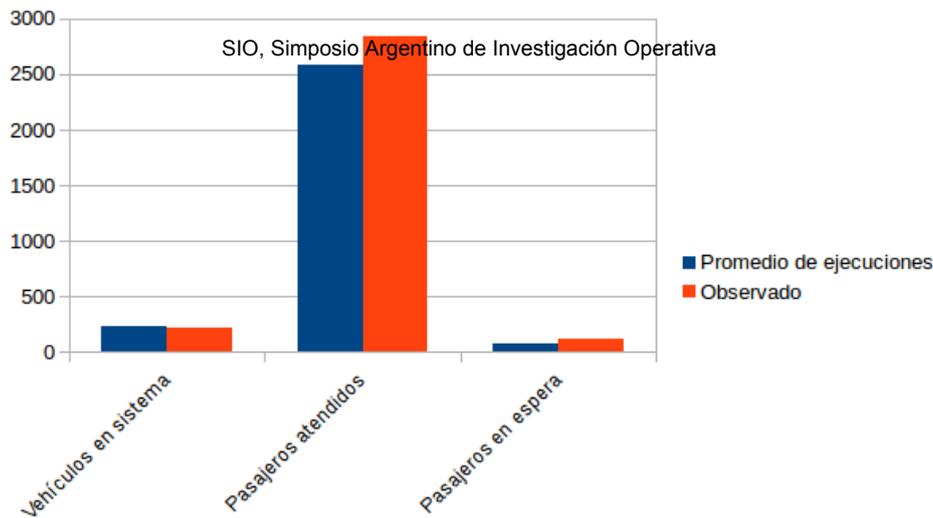


Fig. 2. Comparación entre simulación y relevamiento.

TABLA IV
LOTE DE EJECUCIONES DE SIMULACIÓN

| Corrida -6 a las paradas | Personas que arriban atendidos | Total pasajeros al finalizar | Personas en espera | Total de vehículos su recorrido | Vehículos que concluyeron al finalizar | Vehículos en el sistema (seg) | Tiempo promedio de recorrido |
|-----------------------------|--------------------------------------|------------------------------------|-----------------------|---------------------------------------|--|-------------------------------------|---------------------------------|
| 1 | 2670 | 2644 | 26 | 243 | 218 | 25 | 324 |
| 2 | 2673 | 2631 | 42 | 228 | 194 | 34 | 419 |
| 3 | 2670 | 2452 | 218 | 216 | 170 | 46 | 339 |
| 4 | 2671 | 2634 | 37 | 240 | 219 | 21 | 326 |
| 5 | 2649 | 2637 | 12 | 237 | 205 | 32 | 339 |
| 6 | 2663 | 2643 | 20 | 220 | 189 | 31 | 490 |
| 7 | 2637 | 2634 | 3 | 230 | 198 | 32 | 482 |
| 8 | 2650 | 2480 | 170 | 219 | 212 | 7 | 244 |
| 9 | 2636 | 2351 | 285 | 239 | 219 | 20 | 322 |
| 10 | 2668 | 2616 | 52 | 224 | 213 | 11 | 250 |
| 11 | 2670 | 2568 | 102 | 233 | 218 | 15 | 323 |
| 12 | 2668 | 2653 | 15 | 233 | 211 | 22 | 340 |
| 13 | 2623 | 2534 | 89 | 217 | 208 | 9 | 291 |
| 14 | 2622 | 2610 | 12 | 242 | 209 | 33 | 306 |
| 15 | 2660 | 2513 | 147 | 215 | 186 | 29 | 376 |
| 16 | 2661 | 2543 | 118 | 237 | 201 | 36 | 467 |
| 17 | 2626 | 2604 | 22 | 226 | 215 | 11 | 273 |
| 18 | 2658 | 2621 | 37 | 227 | 216 | 11 | 360 |
| 19 | 2646 | 2635 | 11 | 228 | 211 | 17 | 223 |
| 20 | 2656 | 2614 | 42 | 221 | 199 | 22 | 243 |
| Promedio | 2653,85 | 2580,85 | 73 | 228,75 | 205,55 | 23,2 | 336,85 |

las funcionalidades que ofrece SUMO desde el script que implementa el modelo matemático.

La zona de simulación se obtuvo mediante el mapa del casco céntrico de Avellaneda de Open Street Map⁴. SUMO requiere por su parte que los datos del mapa se especifiquen en formato .xml. Dicha especificación se obtuvo importando el mapa

mencionado mediante NetConvert⁵. A fin de omitir detalles innecesarios para el problema, se aplicó una simplificación a

⁴Open Street Map: <https://www.openstreetmap.org/>

⁵NetConvert: es un utilitario para la importación de redes viales digitales de distintas fuentes para SUMO. <http://sumo.dlr.de/wiki/NETCONVERT>

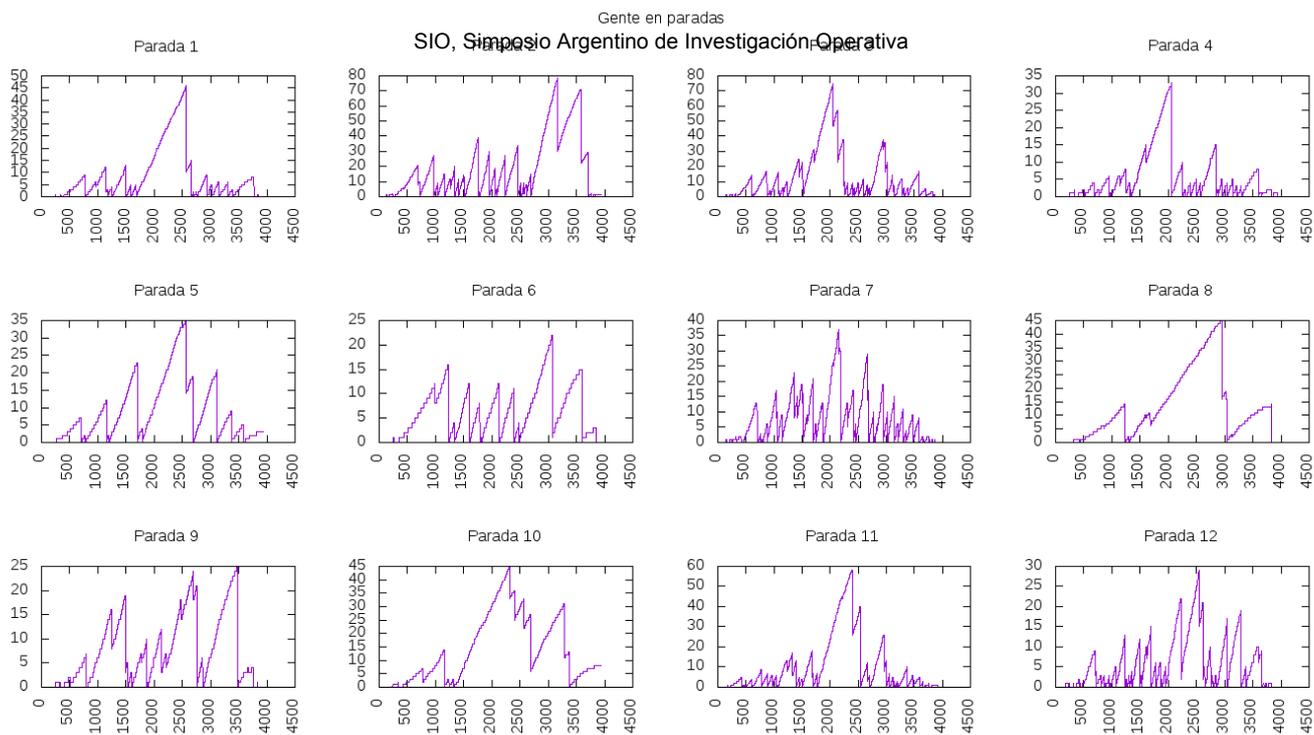


Fig. 3. Evolución de la acumulación de pasajeros en las paradas (GNUPlot)

dicho mapa con NetEdit⁶ y JOSM⁷ (herramientas de SUMO) principalmente en lo que refiere a conexiones y carriles.

La lógica de la simulación se programó con un script Python que se comunica con el simulador mediante TraCI, una api que permite analizar el estado del simulador y cada paso de simulación e introducir cambios al mismo.

Además se utilizó la biblioteca estadística SciPy⁸ para la implementación de las variables aleatorias del modelo matemático. Finalmente se utilizó GNUPlot⁹ y Shell script para la confección de gráficos a fin de visualizar de la evolución de los datos producidos durante la simulación y verificar el ajuste del modelo.

Implementación en SUMO

SUMO trabaja con un archivo de configuración que inicializará el simulador y en el que se indican las referencias a otros archivos que describen el mapa, las rutas y las paradas, algunas propiedades del GUI y otros archivos adicionales que definen el tipo "colectivo" y la ubicación de las paradas. Esta inicialización también define el puerto a usar por TraCI.

⁶NetEdit: es un editor gráfico de red para SUMO. <http://sumo.dlr.de/wiki/NETEDIT>

⁷JOSM: es un utilitario de Open Street Map para editar mapas. <https://josm.openstreetmap.de/>

⁸SciPy: es una biblioteca para Python de <https://scipy.org/>, la documentación específica puede obtenerse de <https://docs.scipy.org/doc/scipy-0.17.1/reference/>

⁹GNUPlot: es un utilitario de graficación de línea de comandos multi-plataforma. <http://www.gnuplot.info/>

Se dividió el periodo simulado en segundos ya que la ejecución se efecta segundo a segundo. Inicialmente el tiempo de detención en las paradas se inicializa en un valor alto (i.e. infinito) ya que luego eso se ajustará cuando se calcule el tiempo de detención de cada ómnibus.

Antes de empezar a ejecutar la simulación se calcula cuántos pasajeros llegarán en cada segundo. Como la V.A. exponencial no da un valor entero se trunca tomando sólo la parte entera acumulando los decimales al segundo siguiente. También se calcula en qué segundo llegará el primer ómnibus de cada línea.

La implementación de la lógica sigue en general el esquema presentado en el siguiente pseudocódigo:

```

Variables calculadas antes
pasajeros_nuevos: pasajeros que arriban en cada segundo
pasajeros_parada: pasajeros en espera en una parada
seg: segundo corriente
seg_llegada: proximo segundo en que arribara un vehiculo

Ciclo de simulacion // 1 por segundo)
  pasajeros_parada = pasajeros_parada + pasajeros_nuevos
  Para cada linea
  Si seg == seg_llegada
    Calcular ocupacion_inicial
    Enviar la orden a SUMO // TRACi
    seg_llegada_ant = seg_llegada
    Calcular seg_llegada // v.a.
  fin fi
fin para
Para cada vehiculo
  Si llego a una parada
    Calcular pasajeros_que_ascienden
    segun (seg-seg_llegada_ant)
    Si este pasajeros_que_ascienden > pasajeros_parada
      pasajeros_que_ascienden <- pasajeros_parada
    fin si

```

```

Calcular tiempo_de_detencion
segun pasajeros_que_ascienden
  y v.a. tiempo_detencion // v.a. normal
pasajeros_parada = pasajeros_parada -
- pasajeros_que_ascienden
fin si
fin para
Fin ciclo

```

SIO, Simposio Argentino de Investigación Operativa

El modelo propuesto puede ser adaptado a cualquier ciudad definiendo los parámetros específicos correspondientes como: mapa, ubicación de las paradas, líneas de colectivo, estimaciones de los parámetros. Finalmente el uso de tecnologías libres (Python, xml, Open Street Map, Shell scripting language, SUMO, NetConvert, NetEdit, JOSM, GNUPlot) garantizan la visibilidad del desarrollo, su accesibilidad y la adaptación a otros problemas similares libremente.

Facilidades y dificultades de SUMO

Dentro de las facilidades que provee SUMO es la que importar mapas, la existencia y el funcionamiento de los semáforos, considerar los cruces y la cantidad de carriles de calles y avenidas. En particular se utilizó un mapa abierto (de Open Street Map) que se exportó a .xml para su importación a SUMO luego de ser editado y simplificado mediante dos utilitarios. También es posible modelizar las paradas, y de acuerdo a los tipos de vehículos, el simulador aplica la aceleración y velocidad correspondiente. También se puede limitar el uso de carriles a los vehículos de tipo **bus**.

Sin embargo, algunos aspectos presentaron dificultades. La modelización de paradas afecta a los vehículos de tipo bus obligándolos a detenerse cuando llegan a su parada. Si se produce una acumulación de vehículos, estos esperarán a que el anterior salga de la parada para detenerse a su vez. Este comportamiento no se observa en la realidad: en general cuando más de un colectivo de una línea se aproxima a la parada, uno de ellos se detiene y los restantes continúan. Este comportamiento tuvo que programarse para evitar el comportamiento definido por defecto. Por otro lado dentro de los atributos de la parada no se incluye la cantidad de personas en espera por lo que se debió definir una variable.

V. CONCLUSIONES

El proyecto desarrollado consistió en el desarrollo de un simulador que implementado mediante un modelo matemático para el estudio de la acumulación de pasajeros en las veredas de una zona específica relevada y para guiar la lógica de la simulación.

Para el desarrollo del modelo matemático se eligió diseñar un sistema de colas con un conjunto de variables aleatorias y mostrar la evolución de la simulación mediante una plataforma de código abierto.

El simulador se implementó con SUMO, una plataforma de Software Libre que permite su parametrización y la programación de extensiones mediante las APIs que ofrece. Una serie de gráficos permitieron estudiar la evolución de las variables de interés y el log de corrida permitió además obtener datos estadísticos de cada corrida. Ambas salidas se utilizaron para verificar el ajuste de la simulación a la realidad observada.

La introducción de variables aleatorias ha producido simulaciones de comportamiento comparable con los datos observados aunque consideramos que es posible lograr un mejor ajuste aumentando la cantidad de mediciones y refinando la construcción de los parámetros, ya que tanto la frecuencia de colectivos como la afluencia de pasajeros, entre otros elementos del modelo, no se repiten en forma idéntica día a día en la realidad.

REFERENCES

- [1] M. R. Azarang and E. García Dunna, *Simulación y análisis de modelos estocásticos*. Mc Graw Hill, 1996.
- [2] B. D. Bunday, *An introduction to queueing theory*. Arnold, 1996.
- [3] D. Gross and C. M. Harris, *Fundamentals of queueing theory*, 2nd ed. Wiley, 1985.
- [4] Z. A. Karian and E. J. Dudewicz, *Modern statistical, systems, and GPSS simulation*. CRC press, 1998.
- [5] A. M. Law and W. D. Kelton, *Simulation modelling and analysis*, 2nd ed. Mc Graw Hill, 1991.
- [6] J. Medhi, *Stochastic models in queueing theory*. Academic Press, 2002.
- [7] S. Ross, *Simulación*. Prentice Hall, 1999.
- [8] J. Barceló, Ed., *Fundamentals of Traffic Simulation*, ser. International Series in Operations Research & Management Science. Springer, 2010, vol. 145.