

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMÁTICAS
&
UNIVERSIDAD POLITÉCNICA DE MADRID
ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACIÓN



Máster en Tratamiento Estadístico-Computacional de la Información

PRÁCTICA FINAL

TÉCNICAS DE MONTE CARLO

Madrid, noviembre de 2023

ANTONIO FERNÁNDEZ GARCÍA
ANDRÉS HERENCIA LÓPEZ-MENCHERO
ÁLVARO FERRERO MORENO

1 | ENUNCIADO

Un aeropuerto ficticio dispone, en su filtro de seguridad de pasajeros, de un total de 4 líneas de inspección de equipaje. Actualmente, están teniendo problemas para cumplir con un parámetro de calidad muy importante que viene impuesto por el DORA, que es un documento de regulación nacional para el sector aeroportuario. En el, se definen un conjunto de condiciones y compromisos que adquiere el aeropuerto durante un periodo de 5 años.

Este indicador: **TEPP-01 Tiempo de espera en control de seguridad de los pasajeros** establece lo siguiente: "El 98 % del total de pasajeros debe pasar el control de seguridad en un tiempo inferior a 10 minutos". El no cumplimiento de este parámetro conlleva importantes sanciones económicas que el aeropuerto quiere evitar.

Desde la dirección del aeropuerto, se quiere evaluar si se podría solucionar instalando una línea adicional de inspección de equipaje de mano. Como este proceso es muy costoso y delicado, se pretende simular el comportamiento del filtro y obtener resultados que ayuden a decidir al aeropuerto a decidir si acometer o no la obra necesaria, u optar por otras alternativas.

Este proyecto trata de simular el funcionamiento de un filtro de seguridad de un aeropuerto. Este aeropuerto cuenta, de inicio, con cuatro líneas de inspección. Los pasajeros pasan por una única cola y posteriormente son llevados a una de las líneas libres que estén abiertas en ese momento.

Se establecen las siguientes características:

- Las **fuentes** son tres, atendiendo al tipo de pasajeros que llegan al filtro, catalogados en función de su tiempo de paso en el proceso. Pretende representar la heterogeneidad del pasaje en cuanto a la preparación previa a pasar por el filtro.
 - Esta fuente genera pasajeros de 3 tipos (p_1, p_2, p_3), considérense perfil rápido, medio y lento. Estos tiempos de servicio están dados por las variables $t_p = \{t_{p_1}, t_{p_2}, t_{p_3}\}$.
- Las **llegadas** están dadas por el número de pasajeros que llegan en cada momento al inicio de la zona de cola previa a los servidores.
 - Esta tasa de llegadas es un proceso estocástico dado por el número de personas que llega en cada hora $P(t)$. Es decir, las llegadas son un proceso markoviano cuyo parámetro varía por horas: $\lambda(t) = \{\lambda_0, \lambda_1, \dots, \lambda_n\}$.
 - Esta $P(t)$ se obtiene basada en datos reales, pero el sistema se adapta a cualquier conjunto de datos que se le pase.
- Solo hay una **cola** que tiene un tiempo de cola t_q . Esta cola sigue el estándar FIFO, *First-In First-Out*.
- La cola da lugar a los **servidores**, que son las líneas de seguridad. El número de servidores en activo varía entre 1 y 5 en función de la demanda (tamaño de la cola).
 - El tiempo de servicio está dado por el tipo de pasajero que llega al sistema. Por tanto, la tasa de servicio depende de este parámetro: $\mu(t_p) = \frac{1}{t_p}$, $t_p = \{t_{p_1}, t_{p_2}, t_{p_3}\}$.

Según las hipótesis previamente comentadas, se puede considerar que el sistema sigue un modelo denotado por la notación de Kendall como $M/U/c$, con c variable entre 1 y 5. Finalmente, el sistema tiene:

- Dos o más fuentes.
- Dos o más servidores.
- Activación condicionada de los servidores.
- Además, el script tiene un mecanismo para evitar que haya más de un número predefinido de personas en la cola (Cola máxima). Cuando ocurra esto, se abrirán automáticamente el número máximo de líneas de inspección.
- Además, los tiempos de llegada varían por horas y los datos se extraen de una serie temporal. Esta dependencia temporal se ha adaptado de unos datos reales que se encuentran en el archivo excel data.xlsx

1. a) Diagrama del sistema

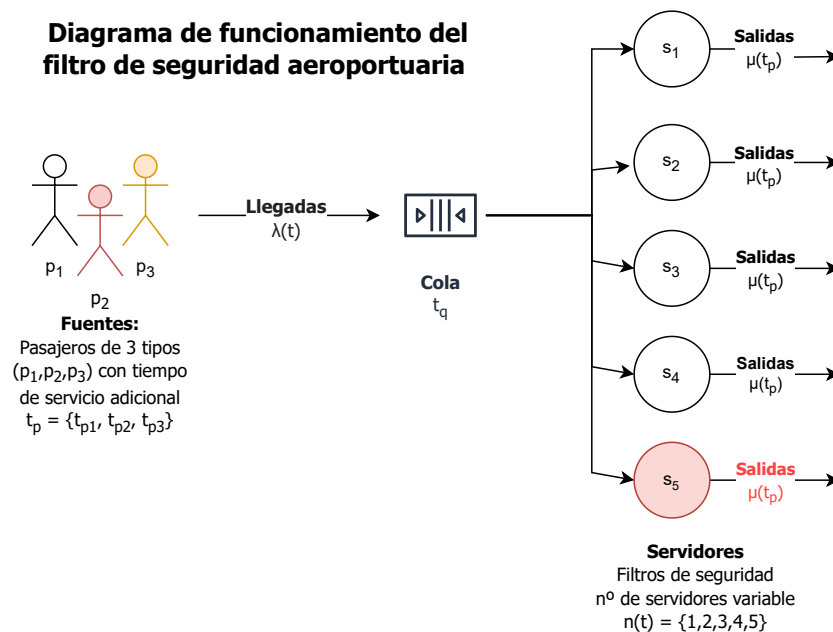


FIGURA (1.1): Diagrama general del sistema de seguridad aeroportuaria.

1. b) Objetivos

El objetivo es, por un lado, optimizar el número de servidores que tiene que estar operativo en cada momento, y por el otro, medir una serie de métricas sobre el sistema. Adicionalmente, se quiere resolver la cuestión de si el sistema mejora cuando se usan 5 líneas de seguridad en vez de 4.

El **número de servidores que deben estar abiertos en cada momento** se decide en función del número medio de personas en cola: si de media, en la última media hora, hay 5 o más personas en la cola, se activa un servidor. Por el contrario, si el número medio de personas en cola es inferior a uno, se desactiva un servidor. Además, si hay un pico de personas en el sistema (es decir, se llega a un número de personas en cola muy alto en poco tiempo), se habilitan todos los servidores en ese momento para evitar que haya congestión.

Las **métricas** que se quieren obtener son:

- **Métricas generales básicas:** Número medio de clientes en el sistema, L ; Número medio de los clientes en cola, L_q ; Tiempo medio de los clientes en el sistema, W ; Tiempo medio de los clientes en cola, W_q .
- **Métricas particulares del escenario:** Número medio de clientes por hora, N_h ; Número medio de servidores abiertos, T_s ; Tiempo total de cada servidor abierto, N_s ; Porcentaje de personas que esperan más del tiempo crítico (10 minutos) sin ser atendidos en el sistema (descrita en el enunciado), $TEPP_01$.

1. c) Recursos

- Conocimientos extraídos de la teoría de colas, los cuales se pueden consultar en Referencias.
- Software estadístico, en este caso **Python**.
- Diagramas explicativos, de flujo y del sistema, que se realizan a través de la herramienta *online* **draw.io**.
- Repositorio de código abierto para control de versiones del proyecto, **GitHub**.
- Los diagramas de flujo del sistema han sido descritos usando la herramienta *online* **code2flow**.

2 | MODELO Y ELEMENTOS DEL SISTEMA

2. a) Diagrama de flujo

En la Figura 5.1 se muestra el diagrama de flujo del sistema de simulación. En él se describe el funcionamiento del sistema. El bloque de OPTIMIZACIÓN SERVIDORES se expone en la Figura 5.2.

- El estado inicial y final (obtención de métricas y visualización de datos) aparece con una elipse en color naranja.
- Los bloques condicionales y las condiciones de parada aparecen en rombos naranjas. MIENTRAS QUE equivale a un WHILE. SI es un condicional IF.
- La actualización de las variables aparece en un recuadro azul.

2. b) Elementos del sistema

El script de simulación está escrito en la función *main.py* y tiene dos métodos: *aeropuerto()* y *metricas()*. El primer método obtiene los resultados de la simulación en forma de *dataframe*, el segundo las métricas. Posteriormente, representa los resultados y escribe por consola las métricas.

Parámetros

- `seed`: semilla para la generación de números aleatorios (por defecto 12345).
- `time`: tiempo donde se para la simulación. Fijado en 720 minutos (14 horas).
- `freq`: parámetro que dicta cada cuánto se revisa el número de servidores abierto. Fijado en 30 minutos.
- `Cola_max`: parámetro que define el número crítico de personas en cola que hace que se activen todos los servidores simultáneamente. Por defecto es 35.

- `cmax`: número máximo de filtros que pueden estar activos simultáneamente. Puede ser 4 ó 5.
- `service_time`: tiempos que tarda cada usuario en completar el servicio. Por defecto [0.5, 1, 1.5].
- `probabilities_nextS`: probabilidad de generar pasajeros de tipo 1, 2 ó 3. Por defecto [0.25, 0.6, 0.15].

Variables

Variables asociadas al tiempo:

- `t`: tiempo de simulación en minutos. Actúa como reloj del sistema.
- `nextL`: tiempo estimado de nueva llegada.
- `nextS`: lista que guarda los tiempos de fin de servicio estimados para cada filtro.
- `stay`: espacio de tiempo que el sistema no cambia de estado.
- `n`: variable que cuenta el número de ciclos (veces que el reloj del sistema ha alcanzado *freq* minutos).

Variables asociadas a las llegadas:

- `Llegadas`: contador que indica cuando ha llegado un nuevo pasajero.
- `lambdaL`: parámetro λ asociado al proceso markoviano de llegadas (varía por horas).

Variables asociadas a la cola:

- `Cola`: variable que guarda el número de personas en cola.

Variables asociadas al servicio:

- `c`: número de filtros activados.
- `Servicio`: lista de valores binarios que dice cuando un filtro está activo (1).

Rutinas

- **Inicialización**: se inicializan las variables anteriores.
- **Fin**: se genera el dataframe y se obtienen las métricas (ejecución del *main.py*).
- **Almacenamiento de variables**: en cada ciclo del bucle principal se almacenan los valores de las variables monitorizadas.

El resto de rutinas están detalladas en el código.

3 | CÓDIGO

El código puede ser consultado en el archivo comprimido adjunto (son varios scripts de Python y un excel de entrada de datos) y/o en el siguiente repositorio de GitHub: [enlace](#). Para ejecutarlo, simplemente es necesario ejecutar el archivo *main.py*. Se mostrarán en la terminal diferentes métricas y se generará un gráfico resumen de la simulación. En el archivo *README.md* del repositorio GitHub también se proporcionan instrucciones para ejecutar el script.

4 RESULTADOS

A continuación se presentan y analizan los resultados obtenidos tras ejecutar la simulación y se plantean unas conclusiones para el problema específico planteado.

Primer escenario: cuatro líneas de inspección

La gráfica de la izquierda de la Figura 4.1 muestra la evolución en porcentaje de las llegadas acumuladas producidas a lo largo de la simulación, así como la evolución del número de personas en cola. Se observan 3 picos principales en la cola que alcanzan valores próximos a 50 personas, un límite razonable para el problema expuesto.

En la misma figura (Figura 4.1), en la gráfica de la derecha, se observa el comportamiento de las líneas de inspección durante la jornada: La línea negra continua indica el número de servidores abiertos a lo largo del día. Las cruces de color representan si el servidor está siendo utilizado en ese momento.

Este gráfico permite comprobar el correcto funcionamiento del sistema. Verifica que los servidores solo son utilizados cuando están activos, y que la cola y el número de llegadas es coherente con lo esperado.

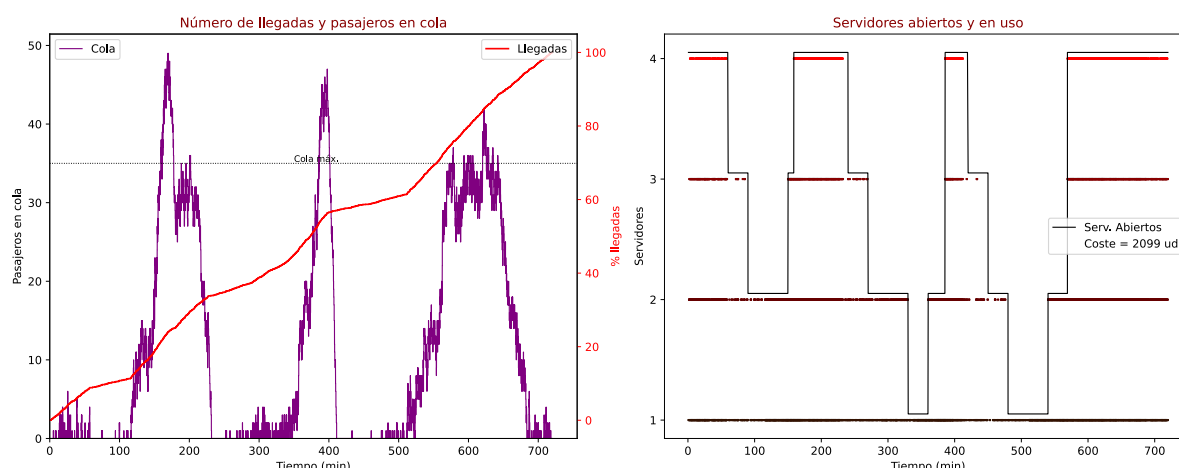


FIGURA (4.1): Simulación del primer escenario. Evolución de la tasa de llegadas y de la cola. Utilización de los servidores.

En el gráfico se ha incluido una variable adicional llamada **coste**. Se trata de una variable adimensional que representa el desembolso económico que conlleva al aeropuerto tener una línea abierta (por cuestiones operativas y de personal). Este coste se calcula por cada minuto que está abierta una línea. La suma de estas cantidades representa una medida de coste total de la operación en la jornada.

Segundo escenario: cinco líneas de inspección

Según se observa en la Figura 4.2, la cola máxima producida disminuye notablemente su valor (algo que cabría esperar). También se muestra cómo cambia el tiempo y el número de servidores que permanece abierto. Particularmente, se observa cómo el hecho de disponer de una quinta máquina a primera hora de la mañana reduce el número de servidores necesarios posteriormente, cubriéndose la mayor parte del escenario con 1, 2 y 3 servidores. Consecuentemente, se comprueba como el coste ha sido notablemente menor.

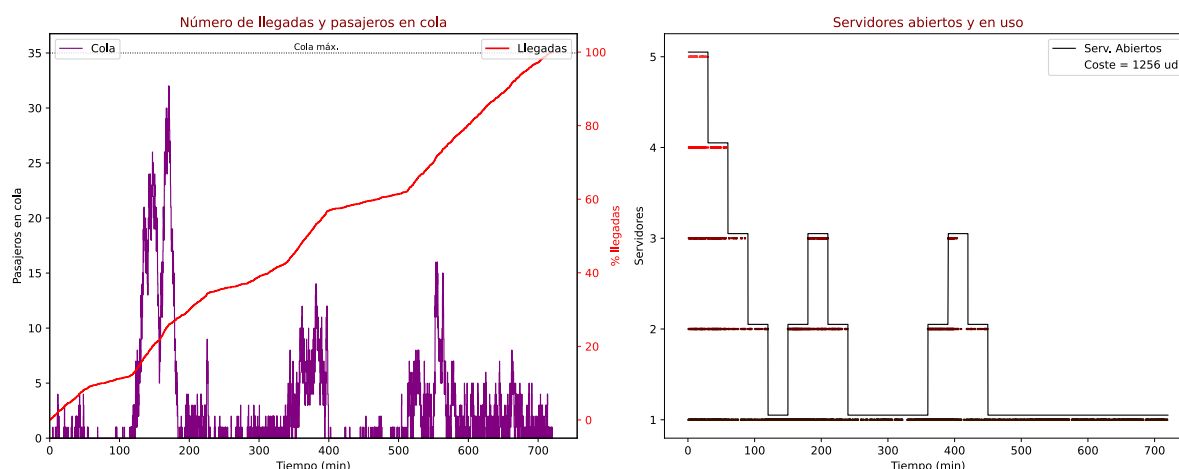


FIGURA (4.2): Simulación del día completo usando 5 servidores

En la Tabla 4.1 se resumen las métricas obtenidas de ambas simulaciones. Como análisis general, se observa que los tiempos medios se reducen considerablemente, así como el número de personas en cola de todas las métricas que se analizan.

TABLA (4.1): Métricas obtenidas tras simular 4 y 5 líneas de inspección.

Métrica	4 líneas	5 líneas	Unidades
Tiempo medio de los clientes en cola (Wq)	4.03	1.44	min
Número medio de los clientes en cola (Lq)	13.37	3.37	pax
Número medio de clientes en el sistema (L)	16.31	5.08	pax
Tiempo medio de los clientes en el sistema (W)	4.98	2.39	min
Número medio de servidores abiertos (Ts)	2.94	1.71	servidores
Coste total de la operativa:	2099.00	1256.00	Ud.
Porcentaje de pasajeros en menos de 10min (TEPP_01):	96.10	99.59	%

Para concluir el ejercicio, se presentan unas conclusiones valorando los resultados obtenidos en las simulaciones:

Desde el aeropuerto, aunque se valora positivamente el mejor desempeño de las métricas de tiempos y pasajeros en cola mostradas en la Tabla 4.1, los dos últimos valores de esta son los que tendrán mayor peso a la hora de decidir qué acción tomar.

El indicador clave que ha propiciado este estudio, **TEPP-01 Tiempo de espera en control de seguridad de los pasajeros**, después de simular una quinta línea de inspección de equipaje mano, ha entrado en el rango adecuado, obteniendo un volumen de pasajeros, cuyo tiempo de paso por el filtro es menor a 10 minutos, superior al 98 %.

Además, después de realizar la simulación, se ha observado una ventaja no prevista: que el **coste** diario de operar con esta máquina inicial es notablemente inferior a continuar con el escenario actual de 4 servidores. A la larga, esto supondrá un ahorro en los costes operativos del aeropuerto, así como un beneficio para la empresa.

Esta combinación de ambas métricas será una razón de peso para impulsar a la dirección de realizar la inversión necesaria para llevar a cabo la implantación de una quinta línea de inspección de equipaje de mano.

5 | ANEXOS

5. a) Diagrama de flujo de la simulación

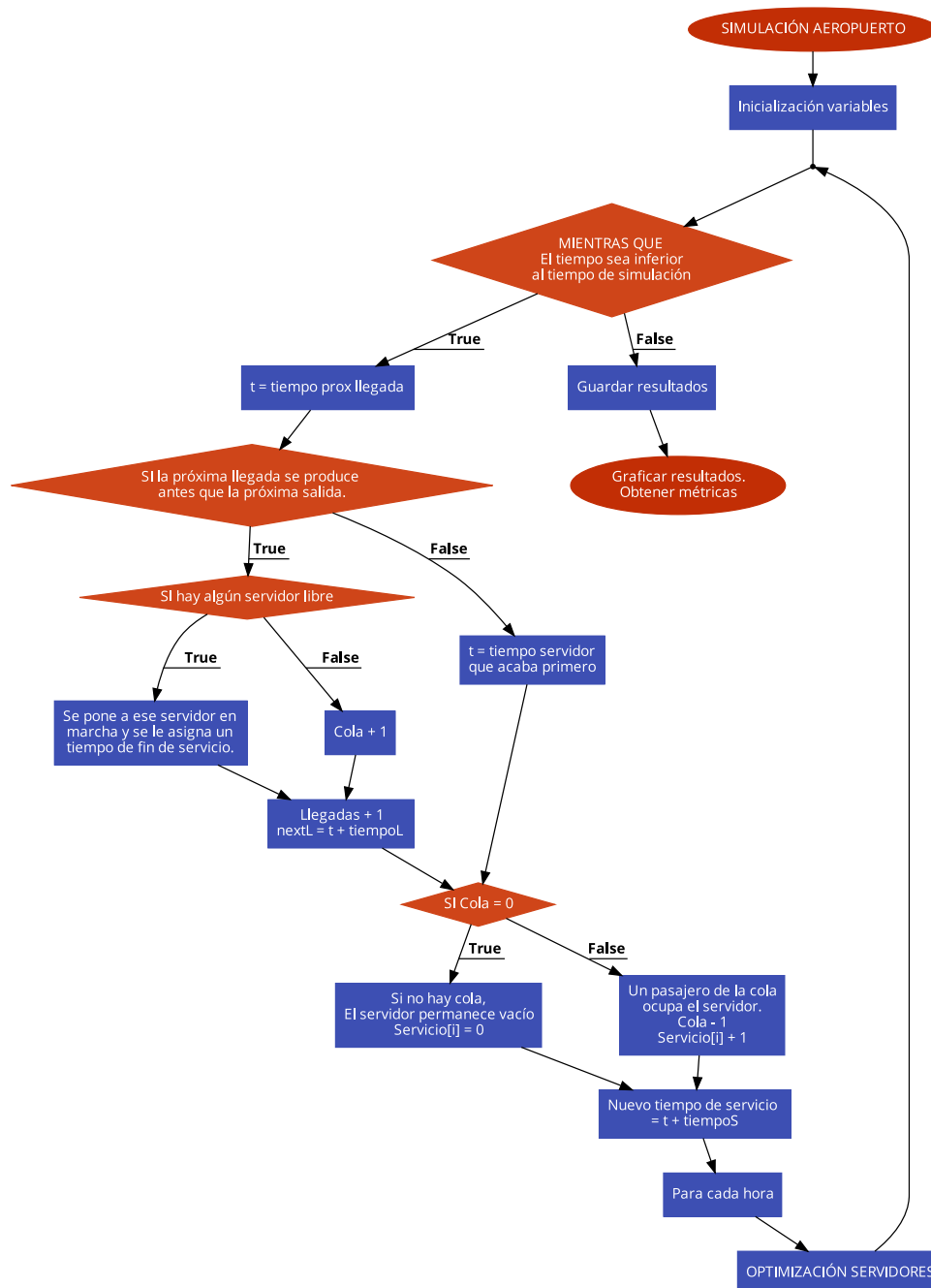


FIGURA (5.1): Diagrama de flujo del funcionamiento del sistema.

5. b) Diagrama de flujo de la activación y desactivación de servidores

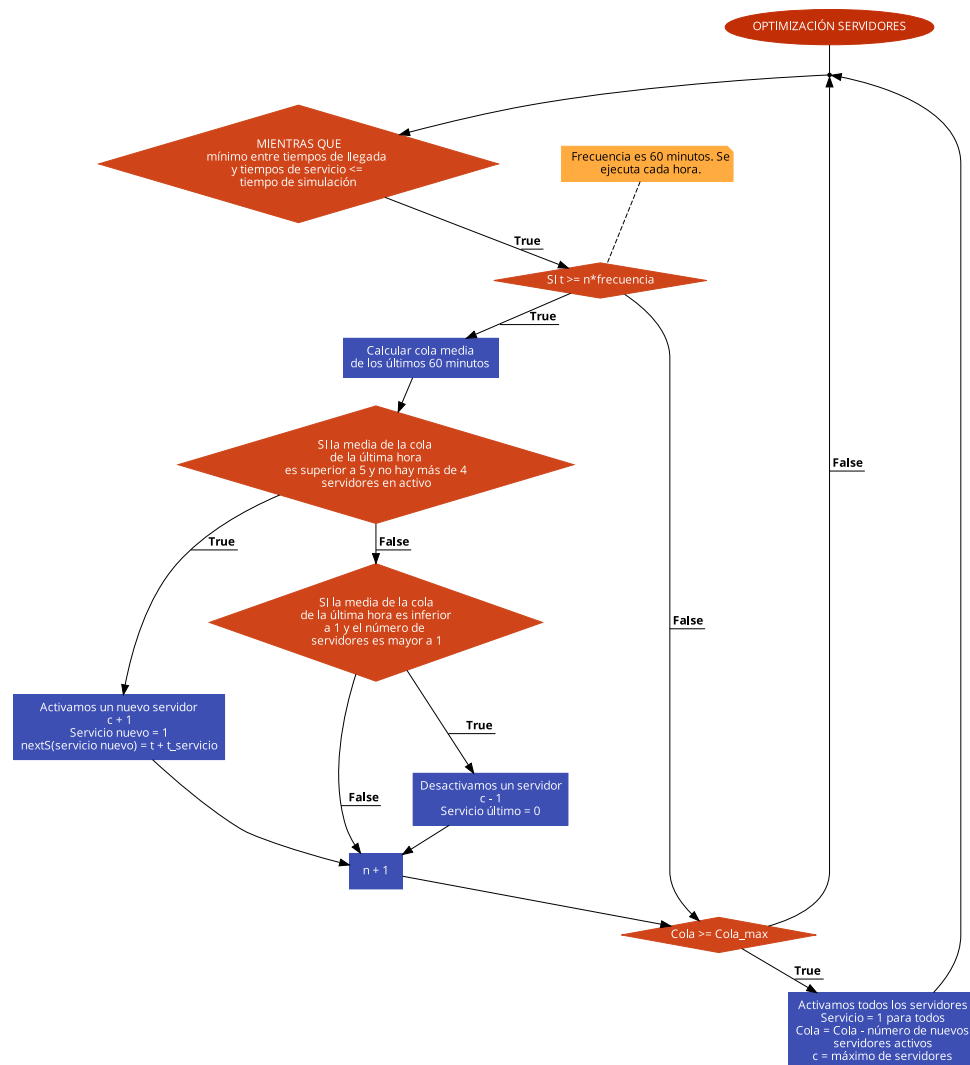


FIGURA (5.2): Diagrama de flujo que explica cómo funciona la activación y desactivación de servidores.

6 REFERENCIAS

- Apuntes de la asignatura de “Técnicas de Montecarlo” impartida en el máster.
- Santiago de la Fuente Fernández - Teoría de colas: Modelo determinista. Sistema de un aeropuerto - Universidad Autónoma de Madrid - Facultad de Ciencias Económicas y Empresariales - Departamento de Economía Aplicada.
- Gabriel Esteban Velázquez y Antonio Rufián Lizana - Modelos de Teoría de Colas - Universidad de Sevilla - Facultad de Matemáticas - Departamento de Estadística e Investigación Operativa.
- Universidad Pontificia de Comillas - Apuntes sobre Teoría de Colas.
- Se han utilizado las herramientas previamente mencionadas: GitHub, Python, draw.io y code2flow.