

Marco Teórico

G2 - Modelos Estocásticos y Simulación

Primer Parcial, Quinto Punto

Juan Esteban Alarcón Bravo

David Sneider Ovalle Pineda

Gustavo Alberto Puentes Romero

Las "colas" son un aspecto de la vida moderna que nos encontramos continuamente en nuestras actividades diarias. En el contador de un supermercado, accediendo al Metro, en los Bancos, etc., el fenómeno de las colas surge cuando unos recursos compartidos necesitan ser accedidos para dar servicio a un elevado número de trabajos o clientes.

El estudio de las colas es importante porque proporciona tanto una base teórica del tipo de servicio que podemos esperar de un determinado recurso, como la forma en la cual dicho recurso puede ser diseñado para proporcionar un determinado grado de servicio a sus clientes.

Debido a lo comentado anteriormente, se plantea como algo muy útil el desarrollo de una herramienta que sea capaz de dar una respuesta sobre las características que tiene un determinado modelo de colas.

1. Definiciones Iniciales

1.1. Teoría de Colas

Es el estudio matemático del comportamiento de líneas de espera. Esta se presenta, cuando los "clientes" llegan a un "lugar" demandando un servicio a un "servidor", el cual tiene una cierta capacidad de atención. Si el servidor no está disponible inmediatamente y el cliente decide esperar, entonces se forma la línea de espera.

1.2. Cola

Una cola es una línea de espera y la teoría de colas es una colección de modelos matemáticos que describen sistemas de línea de espera particulares o sistemas de colas. Los modelos sirven para encontrar un buen compromiso entre costes del sistema y los tiempos promedio de la línea de espera para un sistema dado.

1.3. Sistema de Colas

Los sistemas de colas son modelos de sistemas que proporcionan servicio. Como modelo, pueden representar cualquier sistema en donde los trabajos o clientes llegan buscando un servicio de algún tipo y salen después de que dicho servicio haya sido atendido. Podemos modelar los sistemas de este tipo tanto como colas sencillas o como un sistema de colas interconectadas formando una red de colas. Este modelo puede usarse para representar una situación típica en la cual los clientes llegan, esperan si los servidores están ocupados, son servidos por un servidor disponible y se marchan cuando se obtiene el servicio requerido.

El problema es determinar qué capacidad o tasa de servicio proporciona el balance correcto. Esto no es sencillo, ya que un cliente no llega a un horario fijo, es decir, no se sabe con exactitud en que momento llegarán los clientes. También el tiempo de servicio no tiene un horario fijo.

Los problemas de “colas” se presentan permanentemente en la vida diaria: un estudio en EEUU concluyó que, por término medio, un ciudadano medio pasa cinco años de su vida esperando en distintas colas, y de ellos casi seis meses parado en los semáforos.

2. Teoría de Colas

En muchas ocasiones en la vida real, un fenómeno muy común es la formación de colas o líneas de espera. Esto suele ocurrir cuando la demanda real de un servicio es superior a la capacidad que existe para dar dicho servicio. Ejemplos reales de esa situación son: los cruces de dos vías de circulación, los semáforos, el peaje de una autopista, los cajeros automáticos, la atención a clientes en un establecimiento comercial, la avería de electrodomésticos u otro tipo de aparatos que deben ser reparados por un servicio técnico, etc.

Todavía más frecuentes, si cabe, son las situaciones de espera en el contexto de la informática, las telecomunicaciones y, en general, las nuevas tecnologías. Así, por ejemplo, los procesos enviados a un servidor para ejecución forman colas de espera mientras no son atendidos, la información solicitada, a través de Internet, a un servidor Web puede recibirse con demora debido a congestión en la red o en el servidor propiamente dicho, podemos recibir la señal de líneas ocupadas si la central de la que depende nuestro teléfono móvil está colapsada en ese momento, etc.

2.1. Origen

El origen de la Teoría de Colas está en el esfuerzo de Agner Kraup Erlang (Dinamarca, 1878 - 1929) en 1909 para analizar la congestión de tráfico telefónico con el objetivo de cumplir la demanda incierta de servicios en el sistema telefónico de Copenhague. Sus investigaciones acabaron en una nueva teoría denominada teoría de colas o de líneas de

espera. Esta teoría es ahora una herramienta de valor en negocios debido a que un gran número de problemas pueden caracterizarse, como problemas de congestión llegada-salida.

3. Modelo de formación de colas

En los problemas de formación de cola, a menudo se habla de clientes, tales como personas que esperan la desocupación de líneas telefónicas, la espera de máquinas para ser reparadas y los aviones que esperan aterrizar y estaciones de servicios, tales como mesas en un restaurante, operarios en un taller de reparación, pistas en un aeropuerto, etc. Los problemas de formación de colas a menudo contienen una velocidad variable de llegada de clientes que requieren cierto tipo de servicio, y una velocidad variable de prestación del servicio en la estación de servicio.

Cuando se habla de líneas de espera, se refieren a las creadas por clientes o por las estaciones de servicio. Los clientes pueden esperar en cola simplemente porque los medios existentes son inadecuados para satisfacer la demanda de servicio; en este caso, la cola tiende a ser explosiva, es decir, a ser cada vez más larga a medida que transcurre el tiempo. Las estaciones de servicio pueden estar esperando por que los medios existentes son excesivos en relación con la demanda de los clientes; en este caso, las estaciones de servicio podrían permanecer ociosas la mayor parte del tiempo. Los clientes pueden que esperen temporalmente, aunque las instalaciones de servicio sean adecuadas, porque los clientes llegados anteriormente están siendo atendidos. Las estaciones de servicio pueden encontrar temporal cuando, aunque las instalaciones sean adecuadas a largo plazo, haya una escasez ocasional de demanda debido a un hecho temporal. Estos dos últimos casos tipifican una situación equilibrada que tiende constantemente hacia el equilibrio, o una situación estable.

En la teoría de la formación de colas, generalmente se llama sistema a un grupo de unidades físicas, integradas de tal modo que pueden operar al unísono con una serie de operaciones organizadas. La teoría de la formación de colas busca una solución al problema de la espera prediciendo primero el comportamiento del sistema. Pero una solución al problema de la espera consiste en no solo en minimizar el tiempo que los clientes pasan en el sistema, sino también en minimizar los costos totales de aquellos que solicitan el servicio y de quienes lo prestan.

4. Objetivos de la teoría de colas

Los objetivos de la teoría de colas consisten en:

- Identificar el nivel óptimo de capacidad del sistema que minimiza el coste global del mismo.
- Evaluar el impacto que las posibles alternativas de modificación de la capacidad del sistema tendrían en el coste total del mismo.

- Establecer un balance equilibrado (“óptimo”) entre las consideraciones cuantitativas de costes y las cualitativas de servicio.
- Hay que prestar atención al tiempo de permanencia en el sistema o en la cola: la “paciencia” de los clientes depende del tipo de servicio específico considerado y eso puede hacer que un cliente “abandone” el sistema.

5. Elementos existentes en un modelo de colas

5.1. Fuente de entrada o población potencial

Es un conjunto de individuos (no necesariamente seres vivos) que pueden llegar a solicitar el servicio en cuestión. Podemos considerarla finita o infinita. Aunque el caso de infinitud no es realista, sí permite (por extraño que parezca) resolver de forma más sencilla muchas situaciones en las que, en realidad, la población es finita pero muy grande. Dicha suposición de infinitud no resulta restrictiva cuando, aun siendo finita la población potencial, su número de elementos es tan grande que el número de individuos que ya están solicitando el citado servicio prácticamente no afecta a la frecuencia con la que la población potencial genera nuevas peticiones de servicio.

5.2. Cliente

Es todo individuo de la población potencial que solicita servicio. Suponiendo que los tiempos de llegada de clientes consecutivos son $0 < t_1 < t_2 < \dots$, será importante conocer el patrón de probabilidad según el cual la fuente de entrada genera clientes. Lo más habitual es tomar como referencia los tiempos entre las llegadas de dos clientes consecutivos: consecutivos: clientes consecutivos: $T_k = t_k - t_{k-1}$, fijando su distribución de probabilidad. Normalmente, cuando la población potencial es infinita se supone que la distribución de probabilidad de los T_k (que será la llamada distribución de los tiempos entre llegadas) no depende del número de clientes que estén en espera de completar su servicio, mientras que en el caso de que la fuente de entrada sea finita, la distribución de los T_k variará según el número de clientes en proceso de ser atendidos.

5.3. Capacidad de la cola

Es el máximo número de clientes que pueden estar haciendo cola (antes de comenzar a ser servidos). De nuevo, puede suponerse finita o infinita. Lo más sencillo, a efectos de simplicidad en los cálculos, es suponerla infinita. Aunque es obvio que en la mayor parte de los casos reales la capacidad de la cola es finita, no es una gran restricción el suponerla infinita si es extremadamente improbable que no puedan entrar clientes a la cola por haberse llegado a ese número límite en la misma.

5.4. Disciplina de la cola

Es el modo en el que los clientes son seleccionados para ser servidos. Las disciplinas más habituales son:

- La disciplina FIFO (first in first out), también llamada FCFS (first come first served): según la cual se atiende primero al cliente que antes haya llegado.
- La disciplina LIFO (last in first out), también conocida como LCFS (last come first served) o pila: que consiste en atender primero al cliente que ha llegado el último.
- La RSS (random selection of service), o SIRO (service in random order), que selecciona a los clientes de forma aleatoria.

5.5. Mecanismo de servicio

Es el procedimiento por el cual se da servicio a los clientes que lo solicitan. Para determinar totalmente el mecanismo de servicio debemos conocer el número de servidores de dicho mecanismo (si dicho número fuese aleatorio, la distribución de probabilidad del mismo) y la distribución de probabilidad del tiempo que le lleva a cada servidor dar un servicio. En caso de que los servidores tengan distinta destreza para dar el servicio, se debe especificar la distribución del tiempo de servicio para cada uno.

6. Bibliografía

- Teoría de Colas, [en línea en].
https://www.google.com/search?q=definicion+de+cola+mmn&sxsrf=AOaemvKr9GVRAhZUbsrPPAuLqeT6U2ewOA%3A1637973059492&ei=Q3yhYd_PHeyuwbkPh7iPgA4&ved=0ahUKEwif_fKSpbf0AhVsVzABHQfcA-AQ4dUDCA4&uact=5&oq=definicion+de+cola+mmn&gs_lcp=Cgdnd3Mtd2l6EAMyBwghEAoQoAE6BwgAEEcQsAM6CAghEBYQHRAeSgUIPBIBMUoECEYYAEoECEYYAFC4A1jcBmDsEGgBcAJ4AIABqgGIAawCkgEDMC4ymAEAoAEBYAEIwAEB&sclient=gws-wiz#:~:text=de%20la%20Web-.TEORIA%20DE%20COLAS,-https%3A//www.ingenieria