

Método de estimación de la puntualidad en redes ferroviarias a través de la función densidad de probabilidad de los retrasos primarios

Alejandro Escudero Santana¹, Jesús Muñozuri Sanz¹, Luis Onieva Gimenez¹

¹ Dpto. de Ingeniería de Organización. Escuela Superior Ingenieros. Universidad de Sevilla. Avd. Descubrimientos s/n 41092. Sevilla. aescudero@esi.us.es, munuzuri@esi.us.es, onieva@esi.us.es

Palabras clave: Redes ferroviarias, fiabilidad, puntualidad.

1. Introducción

La presente herramienta surgió por la necesidad de pronosticar el retraso de un tren de mercancías. Era necesario cuantificar la influencia que tiene en el tiempo de transito de un convoy los distintos servicios que operan simultáneamente en la red ferroviaria.

Pesé a la existencia en la literatura de una gran cantidad de artículos que abordan el tema de la fiabilidad o la puntualidad en el ferrocarril (Escudero et al, 2009), no se encontró ninguno que afrontara tal problema de forma completa. La mayoría de los métodos centran sus esfuerzos en la búsqueda de horarios que mejoren la puntualidad, pero presuponen una mejora de la puntualidad como consecuencia de un aumento de la separación entre trenes y no estiman la misma. Otros definen de manera cualitativa los problemas de congestión a partir de los niveles de utilización de la capacidad de la línea. Los que presentan un enfoque algo más similar y cuantifican de alguna manera el nivel de puntualidad o el retraso, estudian sólo en el tráfico en elementos muy básicos de la línea como puede ser una vía simple con apartaderos o una estación (Carey, 1999), olvidándose de la influencia del resto de la red.

En el capítulo 2 se explicarán los mecanismos de propagación de los retrasos, explicando la diferencia entre retrasos primarios y retrasos secundarios. En el capítulo 3 se hará una extensa descripción de método, donde se explicarán los datos que deben conocer el mismo y las salidas que podrá proporcionar. En el capítulo 5 se describirán una serie de aplicaciones de la metodología seguida. Por último, en el capítulo 6 se mostrarán una serie de conclusiones.

2. Retrasos primarios vs. Retrasos secundarios

Considérese un tren que se desplaza desde una estación de origen a una estación de destino. Se llama t_{\min} al tiempo de transito bajo condiciones ideales, mientras, t , es el tiempo actual de transito bajo condiciones reales. La diferencia entre ambas cantidades, $t_1 = t - t_{\min}$, se define como tiempo perdido. Parte del tiempo perdido está ya programado en los horarios, existiendo diferentes razones para este hecho. Es extremadamente raro que se programen los trenes respecto al mínimo tiempo de transito, debido a que esto produciría, ante cualquier perturbación, demasiados conflictos entre trenes. Para reducir este riesgo de conflictos se introducen márgenes de tiempo que son añadidos a los tiempos de transito o de parada mínimos. A la diferencia entre el tiempo de transito programado acorde a los horarios y el tiempo mínimo de transito, t_{\min} , es llamado margen de tiempo o buffer temporal, t_b . Eso significa que el tiempo perdido, t_1 , es la suma de un retraso programado, t_b , y del retraso no programado, t_d . Este retraso no programado varía de un recorrido a otro de forma aleatoria.

El tiempo de retraso dependerá del tamaño del buffer temporal que se programe. Eso significa que el actual tiempo de tránsito puede ser expresado como $t = t_{\min} + t_b + t_d(t_b)$; suma de mínimo tiempo de tránsito, buffer temporal programado y el retraso aleatorio. La suma de los dos primeros componentes de la ecuación anterior será el tiempo de tránsito programado según los horarios. Un objetivo cuando se realiza el diseño de una programación horaria podría ser minimizar el tiempo de tránsito esperado, $E(t)$; o bien minimizar $t_{\min} + t_b + \alpha \cdot E(t_d(t_b))$, donde α es el incremento relativo del valor del tiempo para retrasos inesperados, ya que es un hecho que los consumidores, pasajeros o empresas que envían sus mercancías, dan un mayor valor al tiempo para retrasos inesperados que al tiempo de tránsito que ellos conocen.

Estos retrasos inesperados aparecerán por la existencia de perturbaciones o fallos imprevistos. Por ejemplo, si existe algún problema técnico con el sistema de señalización, los trenes deberán pasar por la vía a una velocidad reducida, en el mejor de los casos. Esto por supuesto retrasaría el paso de algunos trenes, causando lo que habitualmente se conoce como retraso primario o exógeno. Estos trenes que han sido afectados con el retraso, llegarían tarde a la siguiente estación. Si este error de señalización se produce en una vía simple, causaría además retrasos en los trenes con los que el tren retrasado se cruza; a estos retrasos se les conoce como retrasos secundarios o reaccionarios. La separación entre retraso exógeno y retraso secundario es una diferenciación adquirida por la mayor parte de los investigadores.

Los retrasos exógenos o primarios suelen ser aquellos causados por eventos externos. Los retrasos primarios son debidos a fallos en los sistemas de señalización, roturas o desperfectos en las vías, mayor densidad de pasajeros o cargas de la prevista, problemas con el material rodante, ineficiencias del personal, malas condiciones climatológicas... Estas incidencias provocan mayores tiempos de parada o de tránsito en el tren que los padece, sufriendo este un retraso.

Los retrasos secundarios o reaccionarios son aquellos que suceden debido a la iteración entre los retrasos exógenos y las planificaciones horarias. Si un tren llega tarde a una estación porque ha sufrido un retraso primario, puede hacer que otros trenes también se retrasen (ver figura 1). Que este retraso ocurra o no depende de la planificación de los horarios y de la separación extra existente entre trenes, buffers temporales.

Fuente: Gibson et al., 2002

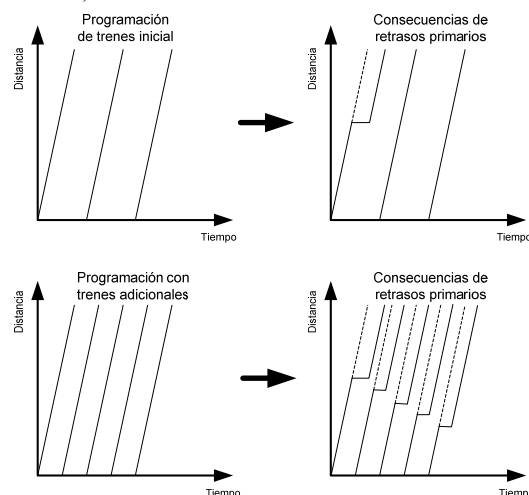


Figura 1. Efecto de la separación existente entre trenes sobre los retrasos secundarios

A través de un aumento de la separación existente entre trenes es posible conseguir una disminución de los retrasos secundarios, y por tanto lograr una mejora en la fiabilidad de la red. Un adecuado diseño de los horarios proporcionará una mejora en la puntualidad. Sin embargo, para reducir los retrasos exógenos es necesario ir al origen de estos retrasos, no pudiendo ser mejorados a través de un buen diseño de los horarios. Un buen mantenimiento de la red y una adecuada formación del personal contribuirían considerablemente a la disminución de tales retrasos.

3. Descripción del método

El método desarrollado se trata de una herramienta escalar, ya que permite alcanzar el nivel de precisión que se estime oportuno o que se considere necesario para la aplicación a realizar. La red puede ser descrita desde nivel de conexiones entre estaciones hasta el nivel mínimo de cantones.

La herramienta, a partir de los parámetros de entrada que definen la red y los servicios que en ella se prestan, va realizando una serie de iteraciones y previendo el retraso con el que se podría ver afectado un tren. Se va calculando la influencia que retrasos en los trenes previos tienen sobre un determinado tren. Es por tanto una herramienta destinada a ver como se propagan retrasos entre los trenes, determina los retrasos secundarios y considera los retrasos primarios como conocidos.

3.1. Entradas de la herramienta

Para la determinación de la puntualidad se necesita previamente conocer una serie de factores que influyen en la misma. Es necesario tener información suficiente sobre la red física, los servicios prestados en la misma, los horarios de tales servicios y los tipos de trenes que los operan. También es conveniente llevar una estadística sobre los retrasos primarios que se producen en las vías.

En el caso de no conocerse alguno de los datos necesarios, bien porque el operador ferroviario o el administrador de la infraestructura no dispone de los datos correspondientes o bien porque se está usando la herramienta para una decisión estratégica, dichos datos deberían de ser tipificados.

La planificación horaria de los distintos servicios ferroviarios que operan en la red debe de ser detallados. En el caso de decisiones estratégicas se haría uso de horarios y servicios tipo. El método opera con la secuencia de estaciones por la que viaja un determinado servicio de la red, así como las horas de entrada y salida de las distintas estaciones, la hora a la que se prevé que pase el tren por un determinado cruce o punto crítico, y el tiempo mínimo de parada en las estaciones en el caso de existir.

Es necesario conocer también la red existente o la red que pretende ser creada con el mayor número de detalle posible o necesario. Se puede definir la red hasta el nivel de detalle que se considere necesario. En caso de no disponer de una forma explícita de los retrasos primarios de los distintos servicios, resulta de gran utilidad el conocimiento de características más específicas de la red, con el objeto de poder determinar los mismos; serían de gran ayuda datos como la velocidad máxima y media de la vía, la pendiente característica de la misma,...

Un ejemplo del grafo a grandes rasgos de una red se muestra en la figura 2, mientras que en la figura 3 se muestra una sección de vía más detallada.

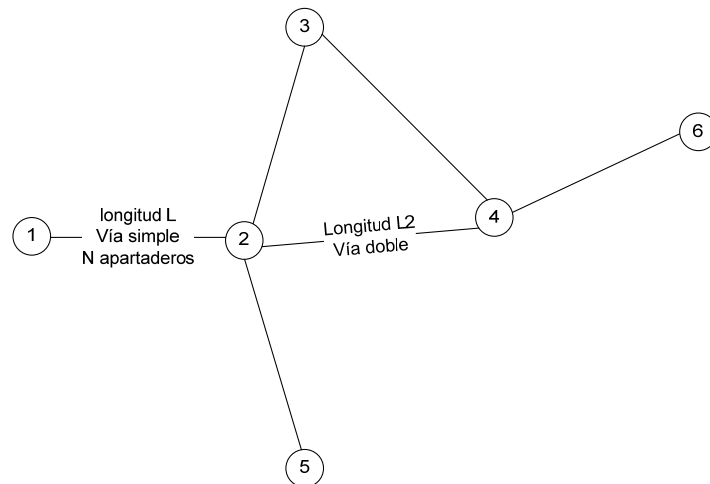


Figura 2. Grafo de red ferroviaria



Figura 3. Grafo detallado de vía simple con apeaderos

Otro de los datos necesarios para el funcionamiento de la herramienta sería la definición de los distintos tipos de locomotoras y las características de las mismas, así como la asignación que se hace de dichas locomotoras a cada uno de los servicios existentes en la red.

Como se ha comentado anteriormente, el método se basa en ver como se propagan los retrasos de unos trenes a otros, por lo que es necesario conocer los retrasos primarios de cada tren. Se hace uso la función densidad de probabilidad del retraso como variable de entrada, con el objeto de no perder información. Se tiene una función de densidad para cada tipo de tren en cada sección de la vía. Disponer de tales datos podría suponer un problema añadido, debido a que es muy probable que los operadores ferroviarios no guarden una información tan detallada; sin embargo, se podría estimar la función densidad de probabilidad a través de los datos de las características de la red (velocidad máxima, velocidad media del tramo, pendiente característica...) y de la locomotora (aceleración, poder de tracción,...). En este punto, es necesario hacer notar la importancia que tendría la inclusión de las nuevas tecnologías que van emergiendo en el ámbito de los sistemas ferroviarios; la utilización de las tecnologías de la información y la comunicación, de sistemas centralizados o distribuidos para el almacenamiento de datos y de sistemas de navegación tales como GPS o Galileo podrían facilitar enormemente el calculo de la función de probabilidad de estos retrasos.

La función densidad de probabilidad utilizada en la herramienta es una función discreta (Ver figura 4). La elección de una variable discreta se debe principalmente a un motivo de simplificación del problema.

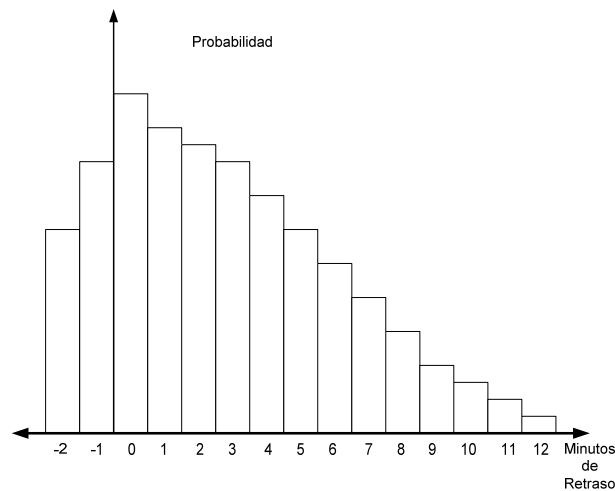


Figura 4. Función discreta de densidad de probabilidad.

Si X es una variable aleatoria discreta que describe el retraso, $x[i]$ denota la probabilidad de que el retraso sea de i minutos, y cumple las siguientes propiedades:

$$\sum_{i=-\infty}^{\infty} x[i] = 1 \quad (1)$$

$$x[i] \geq 0 \quad \forall i \quad (2)$$

$$\Pr(a \leq X \leq b) = \sum_{i=a}^b x[i] \quad (3)$$

3.2. Salidas de la herramienta

Partiendo de los datos anterior, el método dará como salida la función densidad de probabilidad del retraso de los distintos trenes en cada uno de los nodos, en estos retrasos de salida aparecen incluidos tanto los retrasos primarios (dato de entrada) como los retrasos secundarios, fruto de la interacción entre trenes.

La decisión de tener como salida tal cantidad de información es debido a que de dicho resultado pueden ser derivados una gran cantidad de resultados secundarios, tanto para un tren en particular como para la red al completo. Se puede obtener la probabilidad que tiene un tren con llegar a una estación con más retraso que un cierto valor p ; siendo esta probabilidad la suma de la función densidad de probabilidad, obtenida para dicho tren y dicha estación, desde p hasta ∞ . Se pueden tener indicadores de la puntualidad generales a partir de dichos datos, incluso indicadores ponderados dependiendo de la importancia del tren o de la estación; un indicador genérico para la valoración general de la puntualidad se muestra en la ecuación 4.

$$Indicador_global_ponderado = \frac{\sum_t \left[\sum_e \left[\sum_{i=p}^{\infty} \alpha_e^t \cdot x_e^t[i] \right] \right]}{\sum_t \left[\sum_e \alpha_e^t \right]} \quad (4)$$

donde

α_e^t es el factor de importancia que tendría el retraso del servicio t en la estación e, ponderado de 0 a 1.

$x_e^t[i]$ la probabilidad de un retraso de i minutos en la estación e del servicio ferroviario t

El hecho de ponderar respecto a distintos trenes y estaciones se debe a que un operador ferroviario no prioriza de igual manera todos sus servicios. Si un tren de alta velocidad llegase 15 minutos tarde, resultaría de cara al operador mucho más perjudicial que si esos 15 minutos de retraso afectarían a un tren de mercancías.

De igual modo podrían realizarse estimaciones mucho más complejas, se podría considerar la penalización de los servicios de distinta manera dependiendo del valor del retraso. De esta forma se tendrían en cuenta el perjuicio que tiene un retraso mayor a cierto valor para la realización de determinados transbordos entre trenes o modos.

3.3. Forma de cálculo de la herramienta

La herramienta va pronosticando los retrasos en el mismo orden que se producen los movimientos de los trenes. Los movimientos que se llevan a cabo en la red son ordenados temporalmente y uno a uno se va viendo la influencia de los movimientos anteriores en el movimiento actual.

Dos funciones han sido creadas para calcular la propagación de los retrasos; en una de ellas se le añade a la función de probabilidad del retraso a la salida de una estación la función de probabilidad del retraso primario que se incurriría en el trayecto hasta el siguiente nodo (Ver figura 5), en la otra se modifica la función densidad de probabilidad del retraso de salida de una estación con la probabilidad de que la vía que necesita ser usada esté todavía ocupada (Ver figura 6).

Estas funciones pueden ser descritas matemáticamente. En el caso de la primera de ellas (figura 5), si X_{outA} es la PDF del retraso del tren a la salida del nodo A y X_{A-B} es la PDF del retraso primario que el tren puede llegar a tener en el tramo entre A y B, entonces la PDF del retraso a la entrada del nodo B, X_{inB} , es:

$$x_{inB}[i] = \sum_{\forall(j+k=i)} (x_{outA}[j] \cdot x_{A-B}[k]) \quad \forall i, j, k \in \mathbb{N} \quad (5)$$

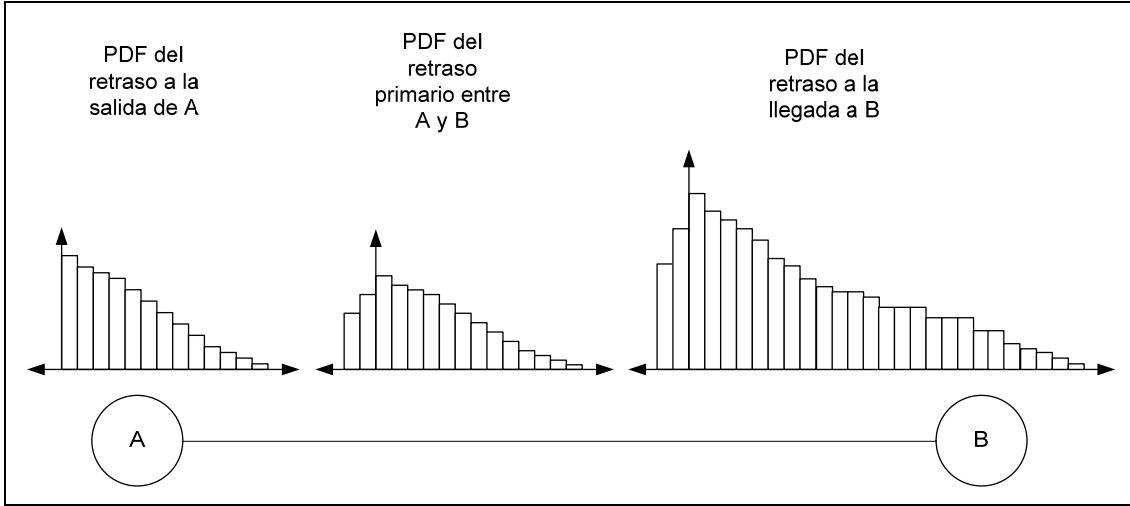


Figura 5. Modificación de la PDF del retraso por el paso de una estación a otra

En el caso de la segunda de las funciones (figura 6), si X_{inV} es la PDF del retraso a la entrada del nodo A de un tren V, y X_{inW} es la PDF del retraso de un tren W que está ocupando la vía que necesita utilizar el tren V, entonces es la PDF del retraso a la salida del nodo A de un tren V, X_{outV} , es:

$$x_{outV}[0] = \left(\sum_{j=-\infty}^{t_{out}^V - t_{in}^V - t_h^V} x_{inV}[j] \right) \cdot \left(\sum_{k=-\infty}^{t_{out}^V - t_{in}^W} x_{inW}[k] \right) \quad \forall j, k \in \mathbb{N} \quad (6)$$

$$\begin{aligned} x_{outV}[i] = & x_{inV}[t_{out}^V - t_{in}^V - t_h^V + i] \cdot x_{inW}[t_{out}^V - t_{in}^W + i] + \\ & + x_{inV}[t_{out}^V - t_{in}^V - t_h^V + i] \cdot \left(\sum_{j=-\infty}^{t_{out}^V - t_{in}^W + i - 1} x_{inW}[j] \right) + \quad \forall i, j, k \in \mathbb{N} \quad (7) \\ & + x_{inW}[t_{out}^V - t_{in}^W + i] \cdot \left(\sum_{i=-\infty}^{t_{out}^V - t_{in}^V - t_h^V + i - 1} x_{inV}[i] \right) \end{aligned}$$

donde:

t_{in}^V es el tiempo pronosticado de llegada del tren V a la estación.

t_{out}^V es el tiempo pronosticado de salida del tren V de la estación

t_h^V es el tiempo mínimo de parada del tren V en la estación

t_{in}^W es el tiempo pronosticado en el que el tren W entra en una nueva sección de la red y abandona la sección necesitada por el tren V.

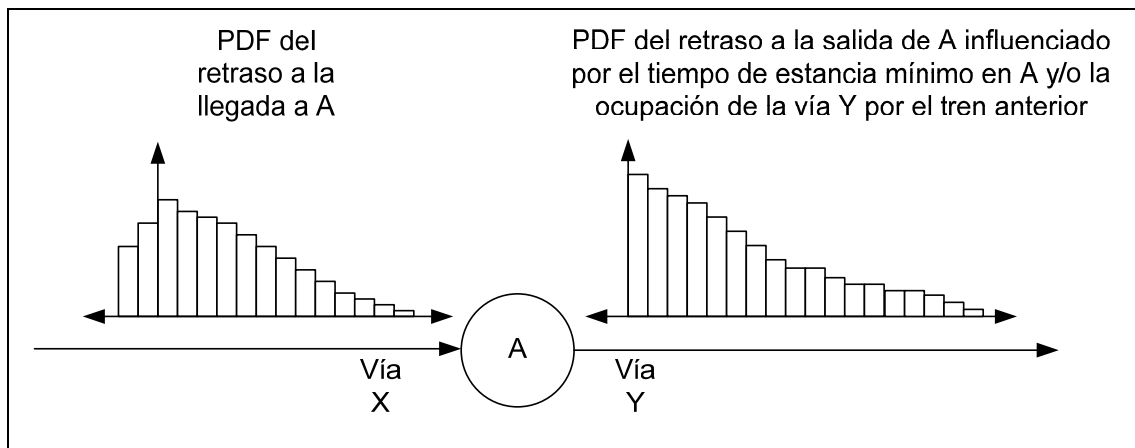


Figura 6. Modificación de la PDF del retraso por la espera antes de poder salir de una estación

Es necesario hacer notar las simplificaciones realizadas en el método. El tiempo mínimo de estancia en la estación se ha considerado un tiempo fijo t_h^V y no un tiempo estadístico, si bien es cierto que la eliminación de dicha simplificación no supondría gran modificación. Solo con considerar la estación un arco de nuestro grafo en lugar de un simple nodo estaría solucionado tal inconveniente. Otra de las simplificaciones llevadas a cabo en el método es la imposibilidad de adelantamientos no pronosticados, estos adelantamientos o cambios en la programación horaria podrían ser realizados en el caso retrasos considerables y por el bien de todos los servicios, sin embargo el método no los contempla.

4. Conclusiones

El método desarrollado permite hacer una primera estimación sobre los retrasos que pueden acumular los trenes, dará información probabilística sobre los mismos. Esta información puede tener un gran valor. Por un lado, esta información puede ser utilizada como parámetro significativo a la hora de elegir un modo de transporte. Por otro lado, puede ser utilizada para una correcta definición de los servicios. Se crearían programaciones horarias de forma que la probabilidad de retrasos en los servicios fuese lo menor posible.

Esta metodología tiene una dificultad, y es la adquisición de los datos que necesita como entrada. El hecho que necesite el conocimiento de los retrasos primarios hace que la puesta en práctica del mismo necesite de la buena voluntad de los administradores de las redes ferroviarias. Ya que esta información sólo puede ser conocida mediante estudios estadísticos.

Referencias

- Alejandro Escudero, Jesús Muñuzuri, M^a Carmen Delgado (2009). Estimación de la puntualidad en redes ferroviarias: estado del arte. CIO 2009.
- Carey, M. (1999) Ex ante heuristic measures of schedule reliability. *Transportation Research B*, 473-494