

Programación Lineal Entera aplicada al Problema de Reprogramación del Cronograma Ferroviario ante grandes interrupciones

Agustín Mosteiro
Departamento de Computación
FCEyN, Universidad de Buenos Aires
Pabellón I, Ciudad Universitaria, CABA, Argentina
Teléfono: +54 11 4576 3390 ext. 701
Email: amosteiro@dc.uba.ar

Juan José Miranda Bront (Director)
Universidad Torcuato Di Tella
CONICET
Buenos Aires, Argentina
Email: jmiranda@utdt.edu

Federico Pousa (Co-director)
Departamento de Computación
FCEyN, Universidad de Buenos Aires
Pabellón I, Ciudad Universitaria, CABA, Argentina
CONICET
Email: fpousa@dc.uba.ar

Resumen

Las operaciones del transporte ferroviario son influenciadas diariamente por eventos no planeados que afectan la infraestructura y pueden producir demoras e interrupciones en la red. Como consecuencia, hay un alto impacto en la calidad del servicio percibida por los pasajeros. Bajo estos escenarios, una de las claves es construir un nuevo cronograma para recuperarse de la interrupción e informar a los pasajeros lo más rápido posible, usualmente en unos pocos minutos.

En los últimos años, ha habido una tendencia en la literatura a considerar modelos de PLEM como prototipos de sistemas automatizados de decisión para hacer frente a este tipo de situaciones. Esta investigación se basa principalmente en los trabajos de Louwerse y Huisman [5] y Veelenturf et al. [6] para el problema de reprogramación de cronograma.

En base a nuestro conocimiento, la mayor parte de la investigación relativa al problema de reprogramación de cronograma se dedica a incorporar nuevas características al modelo, mientras que pocos trabajos se enfocan en algoritmos exactos alternativos (ver Cacchiani et al. [3] para una revisión actualizada). Mejoras en los algoritmos permitirían considerar redes de mayor tamaño como así también ciertos escenarios en los cuales los tiempos de cómputo exceden los tiempos disponibles. Nuestra contribución apunta en esa dirección. Extendemos el trabajo de Veelenturf et al. [6], proponiendo una formulación PLEM alternativa basada en la llamada *Formulación por Índice de Tiempo* considerada en Dash et al. [4] para el Problema del Viajante de Comercio con Ventanas de Tiempo. Mostramos que la formulación propuesta tiene una relajación LP más ajustada. También presentamos resultados computacionales, incluyendo comparaciones respecto a los tiempos de cómputo y a la calidad de las relajaciones LP. Finalmente discutimos futuras direcciones de investigación.

Palabras claves: Gestión de interrupciones, Reprogramación de cronograma ferroviario, Programación Lineal Entera, Formulación por Índice de Tiempo.

I. INTRODUCCIÓN AL PROBLEMA DE GESTIÓN DE INTERRUPCIONES EN TRANSPORTE FERROVIARIO

Las operaciones del transporte ferroviario son influenciadas diariamente por eventos no planeados que pueden producir desde pequeñas demoras hasta grandes interrupciones. En consecuencia, se tiene un impacto directo en la calidad del servicio. Bajo este contexto, uno de los principales objetivos es recuperarse de estas situaciones lo antes posible.

La investigación en el área de *Gestión de Interrupciones* es desafiante desde un punto de vista científico, pero al mismo tiempo prometedora desde un punto de vista práctico. En efecto, incrementar la porción de mercado del transporte público, en especial del transporte ferroviario, es considerada una de las soluciones principales a los problemas de movilidad en grandes ciudades. Además, el ferrocarril es considerado un modo ecológico de transporte. Luego, incrementar el uso del servicio de la red ferroviaria es una cuestión de alta prioridad para muchos gobiernos.

El problema principal es que muchos potenciales pasajeros tienen una mala percepción del nivel de servicio ofrecido. La aplicación de modelos y algoritmos de recuperación para atacar las demoras e interrupciones, implementados mediante *Sistemas de soporte a decisiones* (DSS, por sus siglas en inglés), es clave para elevar la calidad del servicio percibida por los pasajeros.

En nuestro trabajo nos enfocamos en el problema de recuperación bajo grandes interrupciones (2 horas o más). En dichos casos, el problema no puede ser resuelto mediante pequeños ajustes al cronograma, sino que se debe realizar una reprogramación completa tanto del cronograma como así también del material rodante y la tripulación.

Como consecuencia de las interrupciones, la capacidad de la red se ve afectada (disminuída) mientras dure la interrupción. El cronograma original puede volverse imposible de cumplir, siendo necesario no solo retrasar algunos servicios, sino también cancelar otros. Según datos de la CNRT para la red ferroviaria de pasajeros del área Metropolitana de Buenos Aires¹, aproximadamente el 6 % del las operaciones del 2016 se vio afectada por cancelaciones, llegando en algunos meses a estar cerca del 10 %.

¹Estadísticas obtenidas de la CNRT (Comisión Nacional de Regulación del Transporte), <https://www.cnrt.gob.ar/estadisticas-del-transporte-ferroviario>

Las grandes interrupciones son muy difíciles de manejar por los operadores ferroviarios, debido a que causan numerosos cambios en el sistema y se deben tomar decisiones rápidamente. En particular, algunas de las decisiones principales a tomar son: qué servicios cancelar, cuáles retrasar (y cuánto), cómo recircular el material rodante y cómo reprogramar la tripulación, cumpliendo las restricciones de seguridad y convenios laborales. En base a esto, el apoyo de una herramienta automatizada para resolver el problema es altamente deseable.

Actualmente, estos problemas son resueltos en la práctica mediante la implementación de un plan de contingencia, seleccionado a partir de un conjunto de planes disponibles. Uno de los principales problemas de esta solución es que el conjunto de casos predefinidos no puede capturar todos los posibles escenarios que pueden surgir. Además, las decisiones son tomadas en base a la experiencia del operador, con poca o nula utilización de herramientas de IT que asistan a la toma de decisiones. En este contexto, en caso de que cambie la persona cumpliendo esta tarea, la solución ante un mismo escenario seguramente también se verá modificada. En los países más avanzados, la tendencia es a incorporar DSS como herramienta de asistencia.

Debido a la complejidad del problema de recuperación, en general su resolución en tiempo real en redes de mediano o gran tamaño es computacionalmente inviable. Debido a esto, el problema es decompuesto en subproblemas más pequeños, cada uno atacando etapas específicas del proceso. Típicamente en la literatura, el problema general se subdivide en los problemas de *Reprogramación de cronograma*, *Reprogramación del material rodante* y *Reprogramación de la tripulación*. En la Figura 1 podemos ver un diagrama que explica como estas distintas etapas se combinan para resolver el problema completo. El diagrama es incluido a partir del trabajo de Besinovic et al. [2], donde los autores desarrollan un *framework* en el cual comunican distintos modelos conocidos para resolver cada etapa.

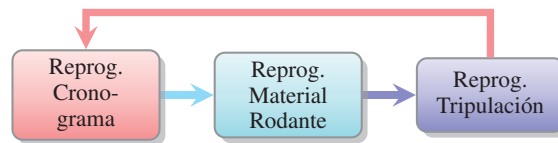


Figura 1: Diagrama del problema de Gestión de Interrupciones, subdividido en etapas. Basado en el trabajo de Besinovic et al. [2].

En cuanto a la Figura 1, más allá de los detalles particulares de la implementación, es interesante destacar que uno de los desafíos principales es la factibilidad. Cuando se obtiene una solución para una etapa, la misma es utilizada como entrada para el próximo problema, el cual podría resultar infactible. En dicho caso, la solución de la etapa anterior tendrá que ser alterada para poder obtener una solución factible en la etapa siguiente. Dicho esquema iterativo se repite hasta encontrar una solución factible para el problema completo.

En nuestro trabajo, nos enfocamos en el problema de Reprogramación de cronograma. Nos basamos principalmente en los trabajos relacionados de Louwerse y Huisman [5] y Veelenturf et al. [6]. Ambos estudios modelan el estado de la red y la capacidad de sus recursos mediante una *Red Evento-Actividad*. Es una estructura típicamente usada en problemas de este tipo.

Los dos trabajos mencionados formulan el problema como un problema de *Programación Lineal Entera*, construido a partir de la red evento-actividad mencionada. Además, dichos estudios consideran, al menos parcialmente, algunas restricciones de la asignación del material rodante, con el objetivo de facilitar la factibilidad en la próxima etapa.

Una de las diferencias principales entre el trabajo de Veelenturf et al. [6] y Louwerse y Huisman [5] es que este último no considera la transición desde el cronograma original hacia el nuevo cronograma generado por el modelo, ni tampoco el traspaso de este al original nuevamente. Se asume además que la red se encuentra vacía al momento en que ocurre el bloqueo de vías. Tener en cuenta estos detalles es clave para la aplicación

de estas herramientas en casos reales. Por dicho motivo, nuestro estudio se basa particularmente en el trabajo de Veelenturf et al. [6].

Como se mencionó anteriormente, el desafío es por un lado construir un nuevo cronograma que minimice el impacto en la calidad del servicio. Por otro lado, se poseen solo algunos pocos minutos para decidir el nuevo cronograma, con el objetivo de obtener una solución en tiempo real. Mejoras en los modelos y algoritmos permitirían considerar redes de mayor tamaño como así también escenarios en donde los tiempos de cómputo exceden los tiempos deseados.

El objetivo de nuestro trabajo apunta en esta dirección. Para lograrlo, proponemos una formulación de PLE alternativa basada en la llamada *Formulación por Índice de Tiempo* considerada en Dash et al. [4] para el *Problema del Viajante de Comercio con Ventanas de Tiempo*. Por las características de la formulación utilizada en Veelenturf et al. [6] y por las ventajas del nuevo modelo propuesto, demostramos que nuestra formulación alternativa posee una relajación lineal más ajustada, aportando un resultado teórico que relaciona ambos modelos. Asimismo, realizamos experimentos preliminares sobre un caso de estudio de una red del área Metropolitana de Buenos Aires. En dicha experimentación comparamos la calidad de la relajación lineal de ambas formulaciones como así también sus tiempos de cómputo. Observamos como los modelos se presentan factibles para su aplicación en la práctica. En particular nuestra formulación es capaz de resolver una mayor cantidad de instancias y disminuir los tiempos de cómputo sobre un subconjunto de ellas.

II. RED EVENTO-ACTIVIDAD Y FORMULACIÓN PLE DEL PROBLEMA

Nuestro problema surge en el contexto de un servicio de trenes y una red asociada en la cual, a partir de una contingencia, se produce un **bloqueo** de algún tramo de vías de la red. Dicho bloqueo vuelve infactible la **programación** original del servicio de trenes. El objetivo del problema es recalcar el cronograma en tiempo real, minimizando el impacto en la calidad del servicio para los pasajeros. Para cumplir con este objetivo, el modelo debe decidir qué trenes son **cancelados** (total y/o parcialmente) y cuáles son **retrasados**, conformando un nuevo programa que será utilizado hasta reestablecer el servicio normal.

Llamamos T al conjunto de todos los trenes y S al conjunto de todas las estaciones de la red.

Para modelar nuestro problema de reprogramación de cronograma, nos basamos en una *red evento-actividad*. Formalmente, representamos a la red como un grafo dirigido $N = (E, A)$, donde E es el conjunto de vértices (eventos) y A el conjunto de ejes (actividades). El modelo de PLE se define en base a la red evento-actividad mencionada.

II-A. Eventos

El conjunto de eventos $E = E_{tren} \cup E_{inv}$ está conformado por un conjunto E_{tren} de **eventos de tren** y un conjunto E_{inv} de **eventos de inventario**.

Cada evento $e \in E_{tren}$ corresponde a un arribo o salida de un tren de una determinada estación. Denominamos t_e al tren asociado al evento e . El horario original en el cual el evento e está programado es notado como q_e . El horario de cada llegada o salida tiene un límite máximo de retraso en el cual puede ser reprogramado, llamamos d_e a dicho límite. Finalmente, llamamos x_e al nuevo horario en el cual el evento e es programado. Luego, nuestro modelo debe determinar los valores de x_e para cada $e \in E_{tren}$ tales que cumplan $q_e \leq x_e \leq q_e + d_e$.

Para que los trenes puedan circular necesitan utilizar un determinado conjunto de recursos cuya capacidad disponible depende de la red sobre la que se opera. Los eventos $e \in E_{inv}$ representan este inventario de recursos disponibles. Existen 3 tipos de recursos, a saber, vías de estación, vías de tramos (o secciones) y formaciones de material rodante. Cada estación tiene una capacidad determinada, que no es más que la cantidad de vías dentro de la misma. De la misma manera, la capacidad de un tramo de vías estará dada por la cantidad de vías entre las estaciones que conecta. Mientras las vías de una estación solo pueden ser ocupadas por un único tren en un tiempo determinado, las vías entre estaciones pueden ser ocupadas por

varios trenes al mismo tiempo, siempre y cuando los mismos circulen en la misma dirección. Cada tren utiliza una formación de material rodante durante todo su recorrido. Al finalizar el mismo, puede suceder que el material rodante que utilizó sea traspasado a otro tren cuyo recorrido comienza en dicha estación, o que las unidades sean guardadas en un patio de maniobras, si es que la estación posee uno. Luego, además de poseer una capacidad de vías, las estaciones poseen una capacidad de material rodante al inicio de cada día.

Llamamos i_e a la cantidad de recursos que deja disponible el evento de inventario e . Por ejemplo, si el evento e representa las vías disponibles en la estación s , i_e será la cantidad de vías de s .

II-B. Actividades

Los eventos se relacionan entre sí mediante dos tipos de actividades. Por un lado, los eventos de un mismo tren están relacionados de manera ordenada en base a su cronograma original, llamamos a estas relaciones **actividades de tren** y las notamos A_{tren} . Por otro lado, eventos de distintos trenes se relacionan en caso de que puedan traspasarse recursos entre sí, llamamos a estas relaciones **actividades de recursos** y las notamos A .

Para modelar el traspaso de recursos, decimos que una actividad $a = (e, f) \in A$, definida desde el evento e hacia el evento f ($e, f \in E$) denota el hecho de que el evento e le traspasa un recurso a f , o lo que es lo mismo, f utiliza un recurso que le deja e . Es importante notar que esta relación fuerza un orden temporal (i.e. e ocurre antes de f si le traspasa el recurso). Otra cuestión a destacar es el hecho de que los mismos eventos e y f pueden estar unidos por más de una actividad dado que hay más de un tipo de recurso, pero a lo sumo puede haber una única actividad del mismo recurso entre dos eventos. Luego, entre dos eventos e y f , puede haber a lo sumo 3 actividades.

Al momento de traspasarse recursos entre trenes, deben respetarse ciertos tiempos de seguridad. A partir de esto, cada actividad $a = (e, f)$ tiene asociado un **tiempo mínimo** L_a el cual debe respetarse para que el recurso que utilizó el evento e esté disponible para ser usado por f .

Para un tipo de recurso dado, un evento e tiene actividades definidas hacia distintos eventos, pero solo puede traspasar su recurso a uno de ellos. Similarmente, un evento f tiene posibilidades de tomar su recurso necesario de distintos eventos, pero solo puede tomarlo de un evento particular. En base a esto, las actividades determinan los posibles órdenes en que cada recurso se puede ir traspasando secuencialmente entre los eventos. Nuestro modelo deberá decidir cómo elegir las actividades buscando la mejor reprogramación posible.

Las actividades de tren son más sencillas en el sentido que el modelo no debe decidir si los eventos que vincula se relacionan o no, ya que eso ya está determinado de antemano por el cronograma original. Además, entre dos eventos de un mismo tren solo puede haber una única actividad de tren, la cual existe si los eventos son contiguos en la programación original. Sin embargo, lo que si deben respetar los eventos de un mismo tren en la nueva programación son los tiempos mínimos de viaje entre dos estaciones y de espera dentro de una estación. Luego, las actividades $a \in A_{\text{tren}}$ tienen asociado un L_a que representa esta duración mínima.

II-C. Bloqueo de vías

El bloqueo de una sección de vías vuelve infactible la programación original, siendo este el motivo que da origen a nuestro problema. El bloqueo ocurre en un **tramo** determinado, bloqueando una **cantidad** de vías durante un período de cierta **duración**.

El bloqueo puede afectar solo algunas vías (bloqueo parcial) o el total de la capacidad del tramo (bloqueo total). Es importante notar que en caso de un bloqueo total, durante el transcurso del mismo ningún tren podrá circular por la sección afectada.

Para modelar el bloqueo se considera un horizonte de tiempo discreto. Se consideran 3 instantes de tiempo, el **tiempo de inicio del bloqueo** al cual notamos τ_1 , el **tiempo de fin del bloqueo** al cual notamos τ_2 y el **tiempo límite de transición**, que modela el instante en que los trenes dejan de seguir el nuevo cronograma y vuelven a estar programados bajo el cronograma original, el cual notamos τ_3 . El modelo debe asegurar

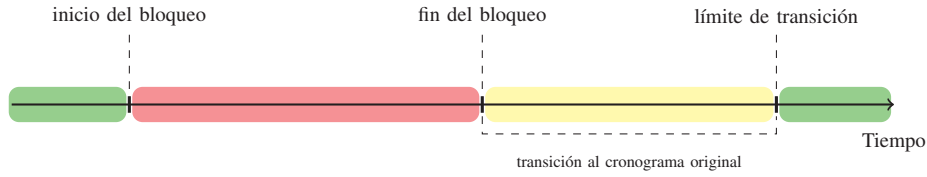


Figura 2: Bloqueo y sus instantes de tiempo

que a partir de τ_3 todos los trenes puedan circular sin ser retrasados ni cancelados. La Figura 2 muestra los instantes mencionados en una línea de tiempo.

A la hora de recalcular el programa original, debe establecerse una relación entre los retrasos y las cancelaciones. Un tren cancelado debería tener un impacto negativo mucho mayor que una llegada o salida retrasada, sin embargo, podría ser preferible tener unas pocas cancelaciones que eviten una gran cantidad de retrasos. Por dicho motivo, se incluye en el modelo una **penalización por tren cancelado**, la cual notamos λ_t para $t \in T$, y una **penalización por unidad de retraso**, la cual notamos μ_e para $e \in E_{tren}$, que modelan esta relación.

II-D. Formulación PLE

A continuación se presenta la formulación PLE del trabajo de Veelenturf et al. [6], llamamos a la misma Orig². Además de la variable x_e para determinar el nuevo horario del evento e , se agregan las variables de decisión y_t y z_a para modelar si el tren $t \in T$ es cancelado ($y_t = 1$) y si la actividad $a \in A$ es seleccionada ($z_a = 1$) respectivamente.

$$\text{Minimizar} \quad \sum_{t \in T} \lambda_t y_t + \sum_{e \in E_{tren}} \mu_e (x_e - q_e) \quad (1)$$

$$\text{sujeto a} \quad x_e - q_e \geq 0 \quad \forall e \in E_{tren} \quad (2)$$

$$x_e - q_e \leq (1 - y_{t_e}) d_e \quad \forall e \in E_{tren} \quad (3)$$

$$x_f - x_e \geq L_a \quad \forall a = (e, f) \in A_{tren} \quad (4)$$

$$\sum_{a \in C} z_a + y_{t_e} = 1 \quad \forall e \in E_{tren}, C \in A^-(e), C \neq \emptyset \quad (5)$$

$$\sum_{a \in C} z_a + y_{t_e} \leq 1 \quad \forall e \in E_{tren}, C \in A^+(e) \quad (6)$$

$$\sum_{a \in C} z_a \leq i_e \quad \forall e \in E_{inv}, C \in A^+(e) \quad (7)$$

$$x_f - x_e + M_a(1 - z_a) \geq L_a \quad \forall a = (e, f) \in A, e, f \in E_{tren} \quad (8)$$

$$z_a = z_{a'} \quad \forall (a, a') \in B_{tr} \quad (9)$$

$$z_a \leq z_{a'} \quad \forall (a, a') \in B_{mr}^{est} \quad (10)$$

$$y_{\beta_t} \geq y_{\alpha_t} \quad \forall t \in T^{blq} \quad (11)$$

$$y_{\beta_t} \geq y_{\gamma_t} \quad \forall t \in T^{blq} \quad (12)$$

$$\sum_{a \in A_{mr}^-(e_{mr,d}^k)} z_a + (i_{e_{mr}^k} - \sum_{a \in A_{mr}^+(e_{mr}^k)} z_a) \geq i_{e_{mr,d}^k} \quad \forall k \in S \quad (13)$$

$$x_e \geq \tau_2 z_a \quad \forall a = (e_{tr}^{\tau_2+}, e) \in A_{tr}^+(e_{tr}^{\tau_2+}) \quad (14)$$

$$y_t, z_a \in \{0, 1\}, x_e \in \mathbb{N} \quad \forall t \in T, \forall a \in A, \forall e \in E_{tren} \quad (15)$$

²Sobre la formulación de Veelenturf et al. [6], realizamos algunas extensiones no incluidas en su trabajo. En particular, las desigualdades (10) (13) y (14) corresponden a extensiones de nuestro trabajo.

La función (1) modela nuestro objetivo de encontrar una reprogramación buscando minimizar el impacto en la calidad de servicio. Para esto, por un lado se suma una penalización por cada tren cancelado y por otro lado se suma una penalización por cada unidad de retraso de un evento en la nueva programación. Luego, si un tren es reprogramado en el mismo horario que lo estaba originalmente, su impacto será nulo. El modelo debe buscar la mínima cantidad de cancelaciones y retrasos dentro de las posibilidades que brinda la red al estar limitada por el bloqueo.

Las restricciones (2) aseguran que los eventos no puedan ser reprogramados antes de su horario original. Asimismo, las restricciones (3) aseguran que los eventos respeten el límite máximo de retraso en caso de que sus respectivos trenes no sean cancelados. Vale notar que en caso de que un tren sea cancelado, el modelo fuerza a que el nuevo horario de sus eventos sea igual al horario original. Esto último es importante para que no se pese la penalización por retraso de sus eventos sino solamente por la penalización por la cancelación del tren. Si bien una solución óptima forzaría esto mismo sin necesidad de incluir la restricción explícitamente, al hacerlo estamos limitando el espacio de posibles soluciones.

Las desigualdades (4) fuerzan a que se respeten los tiempos mínimos de viaje entre estaciones y de espera en una estación. En caso de que se utilicen como dichos valores las duraciones del cronograma original, si una actividad $a = (e, f)$ conecta los eventos e y f y el evento e es retrasado en el nuevo programa (i.e. $x_e > q_e$), el evento f también deberá retrasarse para que el tren pueda cumplir con la nueva programación.

II-E. Restricciones de recursos

Previo a describir las siguientes desigualdades, dividimos las actividades de recursos asociadas con cada evento $e \in E$ en dos: **actividades de entrada** y **actividades de salida**.

Como describe su nombre, las actividades de entrada de un evento e , notadas $A^-(e)$, corresponden a todas las actividades definidas desde otros eventos hacia el evento e . Es decir, $A^-(e) = \{a = (f, e) \in A \mid f \in E\}$. Una actividad de entrada $a = (f, e)$ de e significa que el evento e puede tomar el recurso utilizado por f al menos L_a minutos luego de que f ocurra.

Análogamente, se definen las actividades de salida de un evento e , notadas $A^+(e)$, como $A^+(e) = \{a = (e, f) \in A \mid f \in E\}$. Una actividad de salida $a = (e, f)$ de e significa que el evento e puede traspasar el recurso utilizado a f para que lo utilice al menos L_a minutos luego de que e ocurra.

Las actividades de entrada y salida de un evento representan las distintas posibilidades que tiene dicho evento para tomar y traspasar los recursos que se necesitan para circular.

Para cada evento e , se define un subconjunto de sus actividades por cada tipo de recurso. Luego, se definen *i*) los subconjuntos $A_{tr}^-(e) \subset A^-(e)$ y $A_{tr}^+(e) \subset A^+(e)$ como las actividades de e correspondientes al recurso de vías de tramos, *ii*) los subconjuntos $A_{est}^-(e) \subset A^-(e)$ y $A_{est}^+(e) \subset A^+(e)$ como las actividades de e correspondientes al recurso de vías de estaciones, y *iii*) los subconjuntos $A_{mr}^-(e) \subset A^-(e)$ y $A_{mr}^+(e) \subset A^+(e)$ como las actividades de e correspondientes al recurso de material rodante. Los conjuntos de entrada y salida quedan redefinidos entonces como la unión de estos subconjuntos: $A^-(e) = \{A_{tr}^-(e), A_{est}^-(e), A_{mr}^-(e)\}$, $A^+(e) = \{A_{tr}^+(e), A_{est}^+(e), A_{mr}^+(e)\}$.

Por limitaciones de espacio, no detallamos aquí como se definen las actividades entre los eventos para cada tipo de recurso. Se puede encontrar una explicación detallada en Veelenturf et al. [6].

Las igualdades (5) modelan cómo los eventos toman sus recursos necesarios para circular. Un evento debe tomar una unidad de cada recurso a través de una actividad o, en caso de no poder hacerlo, su tren debe ser cancelado. Es importante notar que el conjunto de igualdades fuerza que los eventos tomen un único recurso de cada tipo y no más. Otro detalle a tener en cuenta es que, en caso de que un evento no tenga actividades de entrada para un tipo de recurso (porque el evento no necesita tomar recursos de ese tipo), la igualdad no se define, dado que forzaría a cancelar el tren asociado al evento.

Las desigualdades (6) modelan cómo los eventos traspasan recursos. Para cada tipo de recurso, un evento puede traspasar el recurso utilizado a lo sumo a otro evento. Al tratarse de una desigualdad en este caso,

podría pasar que un evento no traspase sus recursos utilizados, modelando así el caso del último evento que utiliza un recurso determinado. También vale notar como en caso de que el tren del evento sea cancelado, no puede traspasar ningún recurso, esto tiene sentido dado que al ser cancelado, el evento tampoco pudo haber tomado ningún recurso por las igualdades (5).

Las desigualdades (7) modelan el primer uso de cada unidad de recurso a través de los eventos de inventario. Las actividades de salida desde estos eventos determinan cuales son los posibles eventos de tren que pueden tomar estos recursos por primera vez para luego ir traspasándolos (como marcan las restricciones (5) y (6)) hacia otros eventos. Un evento de inventario $e \in E_{inv}$ posee una capacidad i_e de recursos disponibles, las desigualdades (7) fuerzan a que los eventos de inventario no den más recursos de los que tienen disponibles. Es importante notar que los eventos de inventario no tienen actividades de entrada (i.e. $A^-(e) = \emptyset, \forall e \in E_{inv}$), su función es la de brindar los recursos disponibles de la red a los primeros eventos que los necesiten. Más aún, como los eventos de inventario corresponden a un tipo de recurso determinado, solo poseen actividades de salida para ese tipo de recurso, para los otros tipos de recurso restantes sus conjuntos de salida son vacíos.

Las desigualdades (8) aseguran que se cumplan los tiempos de seguridad de los traspasos de recursos. Las desigualdades son similares a las (4) que modelan los tiempos mínimos de viaje y de espera, la diferencia es que deben tener en cuenta el hecho de que la actividad $a = (e, f)$ puede no ser utilizada (el evento f puede tomar el recurso de otro evento distinto de e). Para estos casos, al tener $z_a = 0$, un valor de M suficientemente grande hace que la desigualdad no imponga condiciones. Dicho valor depende de la actividad a , y por ese motivo notamos a la cota M_a . Se puede ver fácilmente que tomando $M_a = L_a + (q_e + d_e) - q_f$ es suficiente para que la desigualdad sea válida.

II-F. Trenes sobre el área de bloqueo

Supongamos que se produce un bloqueo parcial del tramo de vías (k, l) por el cual el tren $t \in T$ tiene su recorrido programado, en dirección desde k hacia l . Luego, si la salida desde la estación k está programada **durante** el bloqueo (i.e. entre τ_1 y τ_2), los eventos que estaban asociados con el tren t son particionados y asociados con tres nuevos trenes parciales, llamados α_t, β_t y γ_t . Llamamos $T^{blq} \subseteq T$ al conjunto de trenes en esta situación, es decir, $t \in T^{blq}$ si su recorrido pasa por el área del bloqueo y su salida hacia el tramo del bloqueo está programada durante el transcurso del mismo.

Continuando con el ejemplo, llamemos e^{com} y e^{fin} a los eventos correspondientes al comienzo y fin de recorrido del tren t respectivamente. e_k^{arr} y e_l^{arr} denotan los arribos a las estaciones k y l respectivamente, y de manera análoga e_k^{sal} y e_l^{sal} modelan las salidas desde k y l . Podemos ver cómo estarán contruidos los nuevos trenes parciales a partir de los eventos del tren t . El tren α_t contiene los eventos desde e^{com} hasta e_k^{arr} . El tren β_t contiene únicamente a los eventos e_k^{sal} y e_l^{arr} . Mientras que el tren γ_t posee los eventos desde e_l^{sal} hasta e^{fin} .

Las desigualdades (11) y (12) de Orig aseguran que si el tren β_t es operado (la parte del tren original t sobre el área afectada) entonces tanto α_t como γ_t son también operados. Esto es equivalente a operar el tren t completo del cronograma original. Sin embargo, si β_t es cancelado, luego α_t y γ_t pueden, de manera independiente, circular o ser cancelados.

Las actividades de recursos relacionadas con los eventos de los trenes parciales son exactamente las mismas que para los eventos del tren original t , con pequeños ajustes para asegurar que si ninguno de los trenes parciales se cancela, todos usen el mismo material rodante. Por otro lado, en caso de que β_t sea cancelado, el material rodante usado por α_t y γ_t puede ser utilizado por otros trenes desde las estaciones k y l respectivamente.

II-G. Armado del conjunto de eventos E

El horario de todos los eventos que ocurren previo a τ_1 es el horario programado originalmente, ya que hasta ese punto el bloqueo aún no comenzó y sabemos que la programación original es correcta. Además, los trenes asociados a estos eventos no pueden ser cancelados, ya que se encuentran andando previo al inicio del bloqueo.

Por otro lado, debemos asegurar que a partir del instante τ_3 , los trenes vuelvan a circular de acuerdo al cronograma original. Luego, cualquier evento programado en un tiempo posterior a τ_3 no puede ser retrasado. Asimismo, cualquier tren que comience su recorrido en un tiempo posterior a τ_3 no puede ser cancelado.

Debido a estas observaciones, no es necesario que el conjunto de eventos E contenga los eventos de todos los trenes programados en el cronograma original. Es suficiente con incluir únicamente los eventos programados después de τ_1 y antes de τ_3 , dado que para el resto se utilizan los horarios originales. En base a esto, es importante notar que no solo el tamaño de la red tendrá un gran impacto en la dificultad del problema, también la duración del bloqueo será otro factor clave a considerar.

Vale aclarar que se deben tener algunas consideraciones especiales en los límites de E para por un lado modelar correctamente la disponibilidad de los recursos en τ_1 y por otro asegurar que los trenes puedan retomar el cronograma original a partir de τ_3 .

II-H. Otras consideraciones del problema

Para modelar correctamente el problema, es necesario tener diversas cuestiones en consideración. Como consecuencia, se deben realizar adaptaciones tanto a la red evento-actividad como a la formulación PLE. A continuación resumimos las principales cuestiones a tener en cuenta:

- Dada una vía de un tramo y dos trenes circulando en la misma dirección, no se debe permitir que el tren que sale después se adelante sobre el otro. Las desigualdades (9) modelan este comportamiento.
- Se debe tener en cuenta la relación entre el traspaso de material rodante y el traspaso de la vía de la estación correspondiente. Para esto, se incluyen al modelo las desigualdades (10).
- Para garantizar la operatoria normal a partir de τ_3 se deben garantizar las formaciones de material rodante necesarias en cada estación. Las desigualdades (13) apuntan a este objetivo.
- Se debe definir cómo modelar los trenes que se encuentran circulando en los límites de E .
- Se debe realizar un manejo detallado de los trenes circulando en la sección del bloqueo al inicio del mismo.
- El bloqueo limita la capacidad de vías de un tramo determinado. Luego, el inventario del recurso de sección de vías para el área afectada debe reflejar la capacidad en las distintas etapas del bloqueo. Parte de este funcionamiento se logra a partir de las desigualdades (14).

III. FORMULACIÓN TIF

A continuación presentamos la *Formulación por Índice de Tiempo* (TIF, por sus siglas en inglés) para el problema de Reprogramación de cronograma. El objetivo es analizar si es posible mejorar los resultados de la formulación Orig.

La estrategia principal para lograrlo consiste en desarrollar una formulación de PLE alternativa que posea una relajación PL más ajustada. Para obtenerla, modelamos las restricciones de tiempo entre los eventos sin utilizar desigualdades *big-M*, las cuales se encuentran presentes en Orig y son conocidas por producir relajaciones PL débiles en general. Para cumplir este propósito, extendemos el enfoque planteado en Dash et al. [4] para el TSPTW para nuestro problema de Reprogramación de cronograma.

III-A. Nuevos conjuntos de variables

Primero extendemos la notación introducida previamente, basada en su mayoría en el trabajo de Veenturf et al. [6].

Sea $R = \{\text{tr, est, mr}\}$ el conjunto de los tipos de recursos. Notamos como $R^-(e), R^+(e) \subseteq R$ a los tipos de recursos potencialmente requeridos o liberados por el evento $e \in E_{\text{tren}}$ respectivamente.

Dado un tipo de recurso $r \in R^-(e)$, podemos reemplazar la notación $C \in A^-(e)$ utilizada en las secciones previas como $A_r^-(e) \in A^-(e)$. Una definición análoga es considerada para los conjuntos $R^+(e)$ y $A^+(e)$. Esta nueva notación permite explicitar los recursos de los subconjuntos de actividades, característica que utilizamos para definir la formulación TIF.

Para algunos casos particulares, dada una actividad $a = (e, f)$ relativa a un recurso r , puede suceder que $a \in A_r^+(e)$ pero al mismo tiempo $a \notin A_r^-(f)$. En base a esto, definimos el conjunto $A_r^{hacia}(f)$ como

$$A_r^{hacia}(f) = A_r^-(f) \cup \{a = (e, f) \in A_r^+(e) : e \in E\}$$

A partir de este conjunto podemos determinar si un recurso r pertenece a $R^-(f)$, i.e., si sucede que $A_r^{hacia}(f) \neq \emptyset$.

Dado un evento $e \in E_{tren}$, sabemos que sus posibles horarios en el nuevo cronograma están limitados por su horario original q_e y su máximo retraso d_e . Llamaremos *ventana de tiempo* de e a los distintos instantes de tiempo en los cuales el evento e puede ser programado en el nuevo cronograma, y la notaremos W_e , i.e., $W_e = \{t : q_e \leq t \leq q_e + d_e\}$.

Dados dos eventos $e, f \in E_{tren}$ y un tipo de recurso $r \in R^-(e)$ tal que $a = (f, e) \in A_r^{hacia}(e)$, para cada $t \in W_e$ definimos $I_f^e(t, r)$ como el conjunto de los posibles instantes de tiempo en los cuales el evento f puede suceder, si al transferirle a e el recurso r , éste queda disponible exactamente en el instante t . Formalmente, definimos

$$I_f^e(t, r) = \{\tau \in W_f : a = (f, e) \in A_r^{hacia}(e), \text{ máx}\{\tau + L_a, q_e\} = t\}. \quad (16)$$

Estos conjuntos son análogos a los definidos en Dash et al. [4] para derivar la TIF para el TSPTW.

Se agregan tres nuevos conjuntos de variables para desarrollar la formulación TIF. Recordemos que estamos considerando un horizonte de tiempo discreto, por lo tanto, podemos identificar los posibles instantes de tiempo en los cuales los eventos pueden ocurrir.

Primero, definimos x_e^t como una variable binaria que toma valor 1 si el evento $e \in E_{tren}$ es programado en el instante $t \in W_e$. Formalmente, las variables enteras x_e originales que indican el tiempo para el evento $e \in E_{tren}$, pueden reescribirse de la siguiente manera:

$$x_e = \sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau \quad (17)$$

Con una idea similar, definimos las variables binarias z_a^t relativas a la actividad $a = (e, f) \in A_r^+(e)$, las cuales toman valor 1 si el evento e ocurre en el instante de tiempo $t \in W_e$ y transfiere el recurso r a f . La intuición es que la variable de actividad original, z_a , es decompuesta sobre los posibles tiempos de inicio del evento de origen e .

Finalmente, dado que cada evento puede requerir más de un tipo de recurso, necesitamos tener en cuenta el tiempo en que cada recurso se encuentra listo para ser utilizado. Luego, definimos para cada $e \in E_{tren}$, $r \in R^-(e)$ y $t \in W_e$ las variables binarias w_e^t que toman valor 1 si en el instante de tiempo t el recurso r pasa a estar disponible para ser utilizado por el evento e .

$$\text{Min } \sum_{t \in T} \lambda_t y_t + \sum_{e \in E_{tren}} \mu_e \left(\sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau - q_e \right) \quad (18)$$

s.a. (5) – (7), (9) – (13)

$$\sum_{t=q_e}^{q_e+d_e} x_e^t = 1 \quad \forall e \in E_{tren} \quad (19)$$

$$\sum_{t=t'}^{q_e+d_e} x_e^t \leq \sum_{t=\text{máx}(t'+L_a, q_f)}^{q_f+d_f} x_f^t \quad \forall a = (e, f) \in A_{tren}, t' \in W_e \quad (20)$$

$$\sum_{t=q_e}^{\tau_2-1} x_e^t + z_a \leq 1 \quad \forall a = (e_{tr}^{\tau_2^+}, e) \in A_{tr}^+(e_{tr}^{\tau_2^+}) \quad (21)$$

$$\sum_{a \in A_{tr}^+(e)} z_a^t \leq x_e^t \quad \forall e \in E_{tren}, r \in R^+(e), t \in W_e \quad (22)$$

$$z_a = \sum_{t=q_e}^{q_e+d_e} z_a^t \quad \forall a = (e, f) \in A, e \in E_{tren} \quad (23)$$

$$\sum_{t=q_e}^{q_e+d_e} r w_e^t + y_{t_e} = 1 \quad \forall e \in E_{tren}, r \in R^-(e) \quad (24)$$

$$\sum_{t' \geq t} r w_e^{t'} \leq \sum_{t' \geq t} x_e^{t'} \quad \forall e \in E_{tren}, r \in R^-(e), t \in W_e \quad (25)$$

$$\sum_{\substack{(f,e) \in A_r^{hacia}(e), \\ f \in E_{tren}}} \sum_{\tau \in I_f^r(r,t)} z_{(f,e)}^\tau = r w_e^t \quad \forall e \in E_{tren}, r \in R^-(e), t \in (q_e, q_e + d_e] \quad (26)$$

$$\sum_{\substack{(f,e) \in A_r^{hacia}(e), \\ f \in E_{tren}}} \sum_{\tau \in I_f^r(r,q_e)} z_{(f,e)}^\tau + \sum_{\substack{(f,e) \in A_r^{hacia}(e), \\ f \in E_{inv}}} z_{(f,e)} = r w_e^{q_e} \quad \forall e \in E_{tren}, r \in R^-(e) \quad (27)$$

$$y_t, z_a, x_e^t, z_a^t, r w_e^t \in \{0, 1\} \quad \forall t \in T, a \in A, e \in E_{tren}, r \in R^-(e) \quad (28)$$

Notamos primero como varias desigualdades presentes en Orig, i.e. (5) - (7) y (9) - (13), son incluídas en TIF de manera equivalente.

En la nueva formulación TIF se eliminan las variables x_e incluídas la formulación Orig mediante el uso de la identidad (17), como se puede ver en la función objetivo. En el mismo sentido, las desigualdades (20) y (21) son adaptaciones de las desigualdades (4) y (14) de Orig usando las variables x_e^t . Las desigualdades (19) aseguran que cada evento sea programado en un único instante de su ventana de tiempo. En Orig, cuando un evento se cancela todos sus eventos correspondientes quedan programados en su horario original. En este caso, este comportamiento se logra mediante la combinación de las desigualdades (19) y la función objetivo.

Las desigualdades (22) garantizan que dado un evento $e \in E_{tren}$, si el mismo traspasa el recurso r mediante alguna de sus actividad de salida en el instante $t \in W_e$, se fuerza a que sea programado a partir de dicho instante. Notar como las desigualdades se agrupan por cada recurso $r \in R^+(e)$, esto es así dado que el evento e podría realizar el traspaso de más de un tipo de recurso. Es importante observar que estas restricciones son expresadas como desigualdades porque el último evento usando el recurso no es conocido *a priori*, y para dicho evento ninguna actividad de salida de ese tipo de recurso será seleccionada.

La descomposición en instantes de tiempo de las variables de actividad es explícitamente incorporada mediante las desigualdades (23). Es importante notar que las mismas se definen para las actividades $a = (e, f)$ tales que $e \in E_{tren}$. Esto se debe a que las nuevas variables z_a^t se refieren al instante de tiempo t de la ventana W_e . Para las actividades $a' = (e, f)$ tales que $e \in E_{inv}$, utilizamos únicamente las variables z_a originales.

Las desigualdades (24) - (27) modelan la correcta transferencia de recursos en el modelo TIF.

Las desigualdades (24) modelan el hecho de que dado un evento $e \in E_{tren}$ y un recurso r potencialmente requerido, el mismo pasa a estar disponible en un único instante de tiempo de la ventana de e . También se tiene en cuenta que si el tren del evento e es cancelado (i.e. $y_{t_e} = 1$) entonces ningún recurso puede estar disponible para uso por parte de e .

Luego, las desigualdades (25) establecen que un evento e puede ser programado una vez que todos los recursos que requiere se encuentran disponibles.

Mediante los nuevos conjuntos $I_f^r(r, t)$ y las nuevas variables $r w_e^t$ podemos reescribir la desigualdad (8) de Orig, de tipo *big-M*, por las desigualdades (26). Intuitivamente, si al evento $e \in E_{tren}$ se le libera el

recurso r en el instante t , debe suceder que dicho recurso sea provisto por un evento f anterior. Dicho evento será programado en algún instante de su ventana de tiempo respetando el tiempo de seguridad de la actividad (f, e) (considerado en $I_f^e(r, t)$) dado que la variable $z_{(f,e)}^r$ fuerza el nuevo horario del evento f por las desigualdades (22). Luego, dado que el evento e debe ser programado una vez que posea sus recursos disponibles (desigualdades (25)), en caso de que f le traspase una unidad de recurso r a e , sus horarios respetarán el tiempo de seguridad $L_{(f,e)}$ requerido, que es lo que se modela originalmente mediante las desigualdades (8).

Finalmente, se debe tener en cuenta el hecho de que el evento e podría obtener su unidad de recurso r desde un evento de inventario y no desde otro evento de tren anterior. Para los casos en que un evento e obtenga su recurso de esta manera, consideraremos que el mismo se encuentra disponible desde un inicio, i.e. ${}^r w_e^{q_e} = 1$, dado que no hay restricciones de tiempo sobre los eventos de inventario. En base a esto, se incluyen las desigualdades (27) que modelan el escenario para el instante de tiempo q_e mientras que las desigualdades (26) solo considera los instantes de tiempo restantes de W_e .

En base a los nuevos conjuntos de variables introducidos, y en particular a partir las últimas desigualdades descritas, eliminamos las desigualdades *big-M* incluidas en Orig mediante las restricciones (8). El objetivo detrás de esto es obtener una formulación con una relajación lineal más ajustada. Efectivamente, logramos demostrar dicha propiedad, representada a partir de la siguiente proposición.

Proposición 1. $\hat{z}_{orig} \leq \hat{z}_{tif}$

La demostración de la Proposición 1 se puede encontrar en el Apéndice.

IV. EXPERIMENTACIÓN Y RESULTADOS

IV-A. Caso de estudio: Línea Mitre

Debido a que no se dispone de *benchmarks* de acceso público, nuestro caso de estudio corresponde a la red de trenes de la **Línea Mitre**, cuyos datos son de acceso público³.

La Línea Mitre se encuentra enmarcada dentro de la red de servicios del area metropolitana de Buenos Aires. La misma cuenta con una red de 192km de extensión, conecta 56 estaciones y posee 38 formaciones operativas las cuales operan 317 servicios diarios. Es la segunda línea de mayor extensión (y cantidad de estaciones y formaciones) solo por detrás de la línea Roca. Debido a que la línea completa es de gran tamaño, decidimos utilizar únicamente el ramal **Retiro-Tigre**, que cuenta con 17 estaciones, cada una de ellas cuenta con dos vías y cada par de estaciones está unido por tramos de también dos vías. Es el ramal más utilizado de la Línea Mitre. En 2015, del total de 41.5 millones de pasajeros (con boleto pago), 26.7 millones viajaron en este ramal.

Para nuestros experimentos, tomamos el cronograma utilizado de Lunes a Viernes hábiles, el cual está compuesto de 162 servicios. Cada servicio demora 60 minutos desde una cabecera a la otra, y parten con una frecuencia de 14 minutos. En el Apéndice se incluye un extracto del cronograma para los trenes que viajan desde Retiro hacia Tigre⁴.

En cuanto al material rodante, al igual que en el trabajo de Veelenturf et al. [6], solo consideramos la cantidad mínima de unidades disponibles al inicio del día para garantizar la circulación de los trenes en operatoria normal. Es decir que no son consideradas unidades de repuesto.

IV-B. Parámetros

En cuanto a los parámetros relacionados con eventos, el horario original q_e de cada evento e queda determinado a partir del cronograma del ramal Retiro-Tigre. La capacidad i_e de los eventos de inventario queda determinada a partir de la descripción realizada de nuestra red.

³<http://www.trenmitre.com.ar/horarios.php>

⁴El cronograma de la Figura 5 difiere sutilmente de los datos mencionados (y utilizados en nuestros experimentos) dado que el mismo corresponde al período 2016 y los datos utilizados corresponden al período 2015.

Otro parámetro relativo a los eventos que debe ser definido es el mayor retraso permitido d_e para cada evento e , cuyo valor tiene una gran relevancia. Valores chicos para d_e implican menos opciones de reprogramación, obligando a cancelar más trenes. Por otro lado, valores grandes dan lugar a muchas más opciones, generando problemas en principio más difíciles, pudiendo causar un fuerte impacto en los tiempos de cómputo. Además, como vimos en la Sección III, los nuevos conjuntos de variables introducidas para la formulación TIF dependen fuertemente del máximo retraso, por lo que al aumentar su valor, aumentará el número de variables y restricciones con el que trabajará dicho modelo.

Para nuestros experimentos, utilizaremos distintos valores para d_e , a saber: 2, 3, 5 y 7 minutos, los cuales son similares a los seleccionados en Veelenturf et al. [6].

En cuanto a los parámetros relativos a las actividades, utilizamos como tiempos mínimos de viaje de cada actividad (i.e. L_a) a los tiempos de viaje del cronograma 5. Luego, para el nuevo cronograma estamos asumiendo que un tren no puede recorrer un tramo entre estaciones más rápidamente que en el programa original, ni tampoco forzamos que el tiempo de viaje tenga que ser mayor. Asimismo, en el cronograma 5 los arribos y posteriores salidas en una estación se encuentran programados en el mismo horario. Luego, todos los valores L_a relativos a actividades de espera en una estación en nuestro modelo serán 0, permitiendo el mismo comportamiento que en el programa original.

Los parámetros L_a de las actividades de recursos, es decir, los tiempos de seguridad, son valores que dependen de las características de la infraestructura de la red y de las restricciones operativas. Dichos datos no son de acceso público para nuestro caso de estudio. Luego, tomamos como valores para estos parámetros los mismos que se utilizan en Veelenturf et al. [6].

Finalmente, debemos considerar las dos penalidades presentes en nuestro modelo, pesadas en la función objetivo, la penalización λ_t por cancelación del tren t y la penalización μ_e por unidad (i.e. minuto) de retraso del evento e .

Para mantener un alto nivel de servicio, uno de los objetivos es operar el mayor número posible de trenes. Luego, cancelar un tren es penalizado con un valor mucho mayor que un retraso de un evento. Para las penalizaciones por retraso μ_e , consideramos una penalidad de valor 1 por cada unidad (minuto) de retraso con respecto al cronograma original. En cuanto a las cancelaciones, primero consideramos el tiempo total de recorrido de un servicio, el cual en nuestro caso es de 60 minutos. Tomamos tres valores para λ_t que dependen de la duración: 60, 1500 y 3000.

IV-C. Escenarios de bloqueo

Uno de los aspectos claves del bloqueo es su duración, dado que determina la cantidad de eventos (y por consiguiente de actividades) a incluir en nuestra red evento-actividad. Luego, es de esperar que el valor de este parámetro tenga un fuerte impacto en los tiempos de cómputo.

Trabajamos con tres duraciones distintas de bloqueos, alterando el valor del parámetro τ_2 , las cuales son de 50, 100 y 200 minutos.

También utilizamos distintos valores de τ_3 , que indica el tiempo límite en que debe realizarse la transición desde el nuevo cronograma al original. Los valores utilizados son: 10, 30, 50 y 100 minutos de transición.

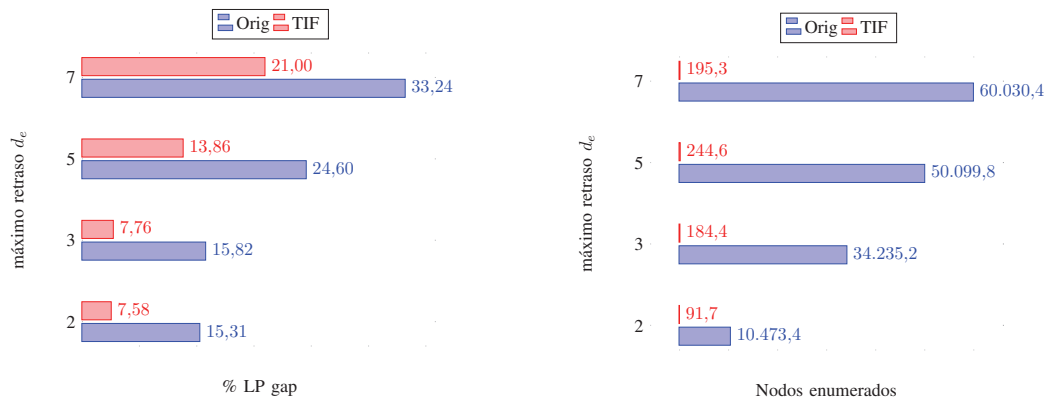
La ubicación del bloqueo (i.e. el tramo de vías que afecta) es otro dato a considerar. Las cuatro opciones consideradas son: Belgrano-Nuñez, Olivos-La Lucila, San Isidro-Beccar y San Fernando-Carupá.

Finalmente, para cada escenario de bloqueo descrito, consideramos la opción de bloqueo parcial (una vía bloqueada) y de bloqueo total (ambas vías bloqueadas).

A partir de la variación de estos parámetros del bloqueo y de la variación de los retrasos máximos y la penalización por cancelación mencionadas en la sección previa, se generan un total de 1152 instancias con las cuales realizaremos nuestros experimentos. Llamamos MITRE al conjunto de instancias.

IV-D. Comparación Orig vs TIF

Para comenzar el análisis, haremos foco en lo que fue nuestro estudio teórico de la Sección III, la mejora en la relajación de la formulación TIF por sobre Orig. Para esto, estudiaremos la métrica $LP\ gap$. Esta métrica determina la diferencia relativa entre la relajación de nuestro problema y el valor de la solución óptima.



(a) Comparación % LP gap Orig vs TIF. Los resultados se muestran agrupados por el parámetro d_e . Para cada valor, se reporta el promedio de la métrica.

(b) Comparación de nodos enumerados Orig vs TIF. Los resultados se muestran agrupados por el parámetro d_e . Para cada valor, se reporta el promedio de la métrica.

Nos interesa validar nuestro resultado teórico con resultados experimentales y estudiar también cuál es la magnitud de la diferencia en la relajación.

Todos los experimentos fueron realizados bajo un entorno de pruebas con un Intel Core i7, 16GB RAM, sistema operativo Ubuntu 16.04 LTS, utilizando C++ y el paquete de optimización CPLEX versión 12.6.3. Se resuelven ambos modelos mediante *Branch-and-Cut*, con un límite de tiempo de ejecución de 20 minutos, dado que valores mayores no son considerados admisibles en un contexto de tiempo real.

En la Figura 3a mostramos los resultados de nuestros experimentos para la métrica mencionada.

Como se puede apreciar, los resultados fueron los esperados. Podemos ver cómo nuestra formulación TIF muestra una mejora que ronda aproximadamente entre el 35 % y el 50 % del valor del LP gap. Vale mencionar que verificamos que al medir la varianza los resultados fueron similares.

Otro dato interesante a notar es el hecho de que para ambas formulaciones, a mayor valor del parámetro d_e , mayor es el valor del LP gap. Esto se corresponde con el hecho de que dicho valor tiene un alto impacto en la dificultad de las instancias. Notar que esto no ocurre únicamente para TIF, donde el impacto de d_e es directo sobre la cantidad de variables, sino que Orig también se ve afectado de manera similar.

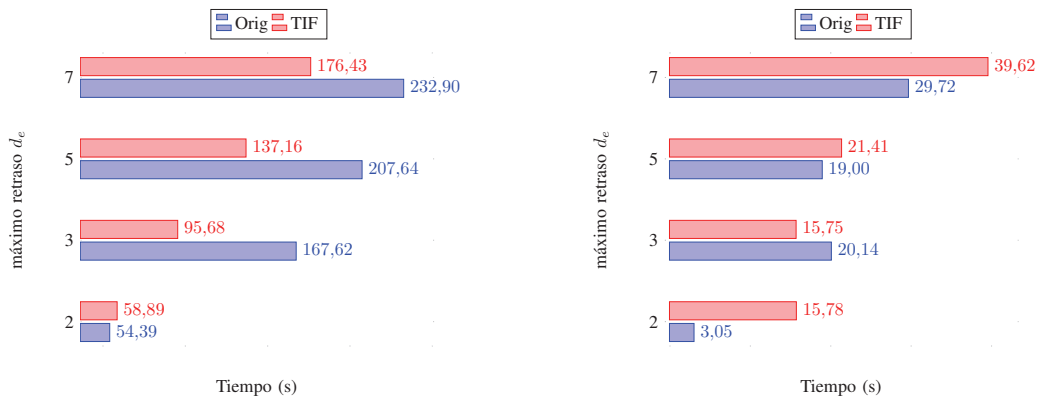
Cuando el máximo retraso permitido aumenta, las instancias comienzan con un valor de la relajación más lejano al de la solución óptima. Esto podría dificultar la terminación del algoritmo, como analizaremos a continuación.

Además de lo mencionado sobre el LP gap, es interesante analizar la métrica de cantidad de nodos enumerados por el algoritmo *Branch-and-Cut* por cada formulación. La Figura 3b muestra el valor de esta métrica para nuestro conjunto de instancias.

Como se puede apreciar, la diferencia es muy grande, de alrededor de dos órdenes de magnitud. Si bien aún no incluimos el análisis sobre los resultados a nivel del tiempo de cómputo, como veremos más adelante, el mismo se encuentra en el mismo orden de magnitud para ambas formulaciones. Es decir, aún cuando ambas formulaciones tienen tiempos de ejecución similares, Orig realiza una enumeración de nodos muy superior a TIF.

Creemos que la mejora en la relajación de TIF y el hecho de que trabaje con una formulación más ajustada, permite obtener mejores cotas duales y por consiguiente encontrar el óptimo enumerando una cantidad de nodos mucho menor.

Por otro lado, podemos ver cómo los valores de ambas métricas crecen según se incrementa la complejidad de las instancias. Al trabajar con instancias de mayor tamaño, el algoritmo de *Branch-and-Cut* debe enumerar una mayor cantidad de nodos hasta su terminación. Esto se condice también con los resultados observados sobre el LP gap.



(a) Comparación de tiempo Orig vs TIF. Los resultados se muestran agrupados por el parámetro d_e . Para cada valor, se reporta el promedio de la métrica.

(b) Comparación de tiempo Orig vs TIF para las instancias de MITRE en que ambas formulaciones terminan por optimalidad. Los resultados se muestran agrupados por el parámetro d_e . Para cada valor, se reporta el promedio de la métrica.

Nos concentramos ahora en el análisis a nivel del tiempo de cómputo. Como pudimos corroborar, nuestra formulación posee una mejor relajación, con lo cual el algoritmo *Branch-and-Cut* inicia con una mejor cota dual. Esto podría mejorar los tiempos de cómputo en el caso de encontrar una solución integral óptima, al estar más cerca de probar optimalidad y así poder terminar. Además, al manejarnos con una formulación más ajustada, esto generará mejores relajaciones en cada nodo del árbol, con lo cual se incrementan también las posibilidades de podar el árbol en base a la cota primal. Por otro lado, nuestra formulación TIF trabaja con una cantidad de variables mucho mayor que Orig, esto puede tener un gran impacto en los tiempos de cómputo, por el costo de resolver cada subproblema asociado. Bajo este escenario, es posible que la desventaja inicial que posee Orig respecto del LP gap sea compensada al enumerar más rápidamente.

Previo a analizar los resultados, es importante tener en cuenta la cantidad de instancias resueltas por optimalidad (i.e. sin llegar al límite de 20 minutos) por cada formulación sobre el total de instancias MITRE. Recordemos que nuestro conjunto MITRE está compuesto por 1152 instancias.

Mientras la formulación Orig resuelve 1008 instancias (87.5 % del total), nuestra formulación TIF logra resolver 1083 (94 % del total). Es decir, nuestra formulación es capaz de resolver más del 6 % del total de instancias que la formulación Orig. Consideramos que este resultado es de suma importancia. Aún en el caso en que nuestra formulación obtenga peores resultados que Orig sobre el resto de las instancias, puede ser preferible utilizar TIF igualmente, dado que resuelve por optimalidad una mayor cantidad.

Analicemos ahora los resultados para nuestro conjunto MITRE, la Figura 4a muestra el promedio del tiempo de cómputo para las 1152 instancias.

Podemos observar que si tomamos el promedio sobre todas las instancias, nuestra formulación tiene un tiempo de cómputo menor que Orig. Esta ventaja es particularmente importante en las instancias de mayor complejidad ($d_e = 7$), dado que en dichos casos los tiempos de ejecución se acercan a valores restrictivos para el uso en tiempo real. Además, para estas instancias de mayor dificultad, TIF también logra resolver una mayor cantidad de instancias que Orig, como podemos ver en el Cuadro I.

Como se puede observar, la ventaja en la cantidad de instancias resueltas de TIF por sobre Orig no se da solo en las instancias de menor tamaño, sino que se mantiene para todas las instancias, y en particular para las de mayor dificultad.

Si bien los resultados de la Figura 4a son interesantes, en la misma estamos calculando el promedio del tiempo de cómputo sobre todas las instancias. En particular, como hemos analizado, hay un subconjunto que TIF logra resolver pero Orig no, llegando al tiempo límite de 1200 segundos e impactando fuertemente sobre

d_e	2	3	5	7
Instancias resueltas Orig	276	252	242	238
Instancias resueltas TIF	278	275	268	262

Cuadro I: Cantidad de instancias resueltas por cada formulación para cada opción del parámetro d_e . Para cada valor de d_e , hay un total de 288 instancias.

el promedio. En base a esto, incluimos en la Figura 4b una comparación de tiempos teniendo en cuenta solo las instancias en las cuales ambas formulaciones terminan por optimalidad. Vale aclarar que estas instancias son en total 1002, representando un 87 % del total.

Como podemos ver, si eliminamos de nuestro análisis a las instancias en donde alguna de las formulaciones no logra encontrar el óptimo, los resultados cambian notoriamente. En general, sobre estas instancias que podemos considerar más fáciles de resolver, Orig obtiene mejores resultados. Igualmente, podemos ver cómo para ciertos casos ($d_e = 3$) TIF obtiene mejores resultados incluso cuando nos limitamos sobre este conjunto.

Para cerrar los resultados, remarcamos que todos los experimentos aquí realizados nos marcan tendencias interesantes, pero conforman una primera etapa dentro de un estudio empírico donde hay mucho más terreno por abarcar. El objetivo final de este tipo de modelos es el de poder brindar una solución a problemas en instancias variadas de la vida real.

V. CONCLUSIONES Y TRABAJO FUTURO

El problema de Gestión de Interrupciones en redes ferroviarias es un problema complejo, pero con un alto potencial de uso en la práctica. Al trabajar a nivel operacional, las limitaciones en los tiempos de cómputo agregan un desafío extra comparado a otros tipos de problemas (e.g. Planificación Táctica). Nuestra contribución consistió en extender el trabajo de Veelenturf et al. [6], proponiendo una nueva formulación de PLE para el problema. Como característica principal, podemos destacar que nuestra formulación posee una relajación lineal más ajustada que la formulación original y además posee una estructura mucho más rica la cual brinda mucho espacio para aplicar técnicas de PLE.

Se realizó una completa experimentación comparando ambos modelos sobre un caso de estudio de una red ferroviaria del area Metropolitana de Buenos Aires. Sobre estos experimentos, podemos destacar

- Nuestra formulación es capaz de mejorar considerablemente el % LP gap, mostrando una mejora aproximada entre 35 % y 50 % dependiendo el caso.
- Nuestra formulación logra resolver por optimalidad aproximadamente un 6.5 % más de instancias que la formulación original.
- Comparando los tiempos de cómputo sobre el total de instancias, TIF obtuvo mejores resultados en promedio. Limitando el análisis a instancias donde ambos modelos resuelven por optimalidad, el modelo original presenta mejores resultados, pero incluso sobre dicho subconjunto nuestra formulación presenta mejores resultados en algunos casos.
- En general, pudimos observar que ambos modelos son comparables y factibles para su aplicación en la práctica.

El trabajo realizado plantea numerosos ejes de estudio a futuro. Algunos de ellos consisten en extender el presente trabajo buscando obtener algoritmos que reduzcan los tiempos de cómputo, por ejemplo mediante el desarrollo de heurísticas y la incorporación de un algoritmo de planos de corte al *Branch-and-Cut*. Por otro lado, resta realizar un estudio cualitativo de los cronogramas generados, buscando evaluar su potencial aplicación en la práctica bajo las condiciones operativas locales. Para ello es indispensable contar con información de infraestructura específica, las reprogramaciones ejecutadas ante interrupciones y herramientas adecuadas de visualización.

APÉNDICE

Relajación lineal TIF vs Orig

En una primera instancia, incluimos en TIF un conjunto de desigualdades que no son necesarias para que la formulación sea válida, dado que toda solución entera óptima las cumple. Sin embargo, su inclusión simplifica el análisis al comparar las relajaciones lineales de ambas formulaciones.

Dado un evento $e \in E_{tren}$, agregamos a TIF las siguientes desigualdades

$$y_{t_e} \leq x_e^{q_e} \quad (29)$$

La formulación Orig cumple esta restricción a partir de las desigualdades (3). En la formulación TIF se eliminan las variables x_e utilizando en su lugar las variables x_e^t . Se puede ver fácilmente como cualquier solución factible de TIF que no cumpla (29) se puede modificar por una solución con mejor valor de la función objetivo y que cumpla las desigualdades (29).

En segundo lugar, asumimos que un preprocesamiento básico es aplicado tanto sobre la formulación Orig como sobre la formulación TIF con el objetivo de eliminar variables que no pueden ser factibles en el modelo.

Dada una actividad $a = (e, f)$ tal que $q_e + L_a > q_f + d_f$, asumimos que dicha actividad no es incluida en la red evento-actividad. Para la formulación TIF, también asumimos que si $\tau + L_a > q_f + d_f$, para $\tau \in W_e$ entonces $z_a^\tau = 0$. Intuitivamente, estamos eliminando las actividades $a = (e, f)$ que, en caso de estar activas, violarían el retraso máximo permitido para el evento f .

Sean \hat{z}_{orig} y \hat{z}_{tif} los valores de las relajaciones PL de las formulaciones Orig y TIF respectivamente. Sean también P_{orig} y P_{TIF} los conjuntos de soluciones asociados con las formulaciones Orig y TIF, y P_{orig}^{LIP} y P_{TIF}^{LIP} los polítopos asociados con su correspondiente relajación PL. Recordemos que estamos considerando que TIF incluye las desigualdades (29) como parte de la formulación.

Demostremos primero las siguientes observaciones, las cuales son utilizados luego durante la demostración principal.

Observación 1. *Dados $f \in E_{tren}$ y $r \in R^-(f)$, entonces*

$$\sum_{\tau=q_f}^{q_f+d_f} (\tau - q_f) x_f^\tau \geq \sum_{\tau=q_f}^{q_f+d_f} (\tau - q_f)^\tau w_f^r$$

Proof: La desigualdad se obtiene sumando las desigualdades (25) para $t = q_f + 1, \dots, q_f + d_f$. ■

Observación 2. *Dado un evento $e \in E_{tren}$, un recurso $r \in R^-(e)$, un instante $t \in W_e$ y una actividad $a = (f, e) \in A_r^{hacia}(e)$ tal que $f \in E_{tren}$, la siguiente desigualdad es válida*

$$r w_e^t = \sum_{\substack{(g,e) \in A_r^-(e) \\ g \in E_{tren}}} \sum_{\tau \in I_g^e(t,r)} z_{(g,e)}^\tau \geq \sum_{\tau \in I_f^e(t,r)} z_{(f,e)}^\tau$$

Proof: Para $\tau \in (q_e, q_e + d_e]$ la desigualdad se obtiene considerando la restricción (26) para el evento e y el recurso r y restringiendo la sumatoria solamente sobre el evento f . Para $\tau = q_e$ la desigualdad se obtiene considerando la restricción (27) de una manera similar y además descartando la sumatoria sobre los eventos de inventario. ■

Observación 3. *Sean $e, f \in E_{tren}$ y $r \in R^-(f)$ tal que $a = (e, f) \in A_r^{hacia}(f)$. Luego, la siguiente igualdad se mantiene*

$$\sum_{t=q_f}^{q_f+d_f} \sum_{\tau \in I_e^f(t,r)} z_{(e,f)}^\tau = \sum_{\tau=q_e}^{q_e+d_e} z_{(e,f)}^\tau = z_a$$

Proof: Recordemos que por definición, $\forall t \in W_f, I_e^f(t, r) \subseteq W_e$. Para cada instante de tiempo $t \in W_f$, se consideran los posibles tiempo de inicio $\tau \in I_e^f(t, r)$ para la actividad $a = (e, f)$, incluyendo aquellos

en los que la transferencia del recurso termina antes del horario original del evento f , i.e. $\tau + L_a \leq q_f$. Además, dado que la expresión del lado izquierdo considera todos los instantes de tiempo de la ventana de f , si existe un instante de tiempo $\tau \in W_e$ el cual no pertenece a $I_e^f(t, r)$ para ningún $t \in W_f$, entonces se debe a que desde τ no es posible respetar el retraso máximo permitido del evento f , i.e. $\tau + L_a > q_f + d_f$. Pero recordemos que para estos casos estamos asumiendo el preprocesamiento $z_a^\tau = 0$. Por consiguiente, podemos incluir dichos términos en la sumatoria derecha y la igualdad se mantiene. Finalmente, podemos reemplazar esta sumatoria por la variable z_a en base a las igualdades (23). ■

A continuación demostramos la Proposición 1, la cual conforma el resultado principal respecto a las formulaciones Orig y TIF.

Proof: La idea detrás de la demostración es mostrar que $P_{\text{TIF}}^{LP} \subseteq P_{\text{Orig}}^{LP}$. Si esto sucede, $\hat{z}_{\text{tif}} \in P_{\text{Orig}}^{LP}$ y por consiguiente $\hat{z}_{\text{orig}} \leq \hat{z}_{\text{tif}}$.

Sea $\hat{x} \in P_{\text{TIF}}^{LP}$ una solución (fraccionaria) factible de la relajación PL de TIF, necesitamos mostrar que \hat{x} es también factible en P_{Orig}^{LP} . Debemos notar que TIF tiene muchas desigualdades en común con Orig, i.e., aquellas incluidas en Orig en las cuales las variables x_e no son consideradas explícitamente. Luego, el problema se reduce a mostrar que \hat{x} satisface las desigualdades (2), (3), (8), (4) y (14) de Orig.

Como mencionamos previamente, la formulación TIF no posee las variables x_e incluidas en las desigualdades enumeradas. Luego, para que \hat{x} sea factible en P_{Orig}^{LP} necesitamos transformar la solución para que este compuesta por las variables utilizadas en Orig. Dicha transformación se realiza mediante la igualdad (17) que vincula las variables x_e^t con las variables x_e . Por otro lado, las variables y_t y z_a se mantienen equivalentes, mientras que las variables z_a^t y $r w_e^t$ simplemente son descartadas.

A continuación analizamos cada una de las desigualdades mencionadas.

1. *Desigualdades (2).* Comenzamos mostrando que \hat{x} satisface las desigualdades (2). Por (17) y (19) tenemos que

$$x_e = \sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau \geq \sum_{\tau=q_e}^{q_e+d_e} q_e x_e^\tau = q_e.$$

2. *Desigualdades (3).* Primero reescribimos las desigualdades (3) como

$$x_e \leq q_e + (1 - y_{t_e})d_e.$$

Sabemos que \hat{x} satisface las desigualdades (29) agregadas a TIF, que establecen que $x_e^{q_e} = 1$ si el tren t_e es cancelado, para soluciones enteras. Reescribiendo las desigualdades (19) tenemos que

$$y_{t_e} \leq x_e^{q_e} = 1 - \sum_{\tau=q_e+1}^{q_e+d_e} x_e^\tau,$$

y por lo tanto

$$\sum_{\tau=q_e+1}^{q_e+d_e} x_e^\tau \leq 1 - y_{t_e}. \quad (30)$$

Usando esta desigualdad y analizando la definición de las variables x_e en (17), tenemos que

$$\begin{aligned}
 x_e &= \sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau = q_e x_e^{q_e} + \sum_{\tau=q_e+1}^{q_e+d_e} \tau x_e^\tau \leq q_e x_e^{q_e} + \sum_{\tau=q_e+1}^{q_e+d_e} (q_e + d_e) x_e^\tau \\
 &= q_e x_e^{q_e} + q_e \sum_{\tau=q_e+1}^{q_e+d_e} x_e^\tau + d_e \sum_{\tau=q_e+1}^{q_e+d_e} x_e^\tau \\
 &\leq q_e \sum_{\tau=q_e}^{q_e+d_e} x_e^\tau + d_e \sum_{\tau=q_e+1}^{q_e+d_e} x_e^\tau = q_e + d_e \sum_{\tau=q_e+1}^{q_e+d_e} x_e^\tau \\
 \text{por (30)} &\leq q_e + d_e(1 - y_{t_e}).
 \end{aligned}$$

3. *Desigualdades* (8). A continuación analizamos las desigualdades (8) siguiendo una idea similar a la utilizada en Dash et al. [4]. Sean $e, f \in E_{tren}$, $r \in R^-(f)$ y $a = (e, f) \in A_r^{hacia}(f)$. Reescribimos la desigualdad como

$$x_e - M_a(1 - z_a) + L_a \leq x_f,$$

y recordemos que $M_a = L_a + (q_e + d_e - q_f)$. Definimos $\bar{x}_e^t = x_e^t - z_a^t$, y notamos que

$$\sum_{\tau=q_e}^{q_e+d_e} \bar{x}_e^\tau = \sum_{\tau=q_e}^{q_e+d_e} x_e^\tau - \sum_{\tau=q_e}^{q_e+d_e} z_a^\tau = 1 - z_a.$$

Utilizando esta igualdad, podemos reescribir las variables x_e como

$$\begin{aligned}
 x_e &= \sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau = \sum_{\tau=q_e}^{q_e+d_e} \tau \bar{x}_e^\tau + \sum_{\tau=q_e}^{q_e+d_e} \tau z_a^\tau \leq (q_e + d_e) \sum_{\tau=q_e}^{q_e+d_e} \bar{x}_e^\tau + \sum_{\tau=q_e}^{q_e+d_e} \tau z_a^\tau \\
 &\leq (q_e + d_e)(1 - z_a) + \sum_{\tau=q_e}^{q_e+d_e} \tau z_a^\tau
 \end{aligned} \tag{31}$$

Reemplazando la variable x_e por esta expresión, obtenemos la siguiente cota superior para el lado izquierdo de la desigualdad (8):

$$\begin{aligned}
 x_e - M_a(1 - z_a) + L_a &= x_e - (L_a + q_e + d_e - q_f)(1 - z_a) + L_a \\
 &= x_e + L_a z_a - (q_e + d_e)(1 - z_a) + q_f(1 - z_a) \\
 \text{por (31)} &\leq L_a z_a + q_f(1 - z_a) + \sum_{\tau=q_e}^{q_e+d_e} \tau z_a^\tau \\
 \text{por (23)} &= q_f(1 - z_a) + \sum_{\tau=q_e}^{q_e+d_e} (\tau + L_a) z_a^\tau.
 \end{aligned} \tag{32}$$

Ahora obtendremos una expresión que es una cota inferior de x_f , es decir, del lado derecho de la desigualdad (8). Para mostrar los resultados intermedios, nos basamos en las observaciones 1-3 analizadas anteriormente.

$$\begin{aligned}
 x_f &= \sum_{\tau=q_f}^{q_f+d_f} \tau x_f^\tau = \sum_{\tau=q_f}^{q_f+d_f} (q_f + \tau - q_f) x_f^\tau = q_f + \sum_{\tau=q_f}^{q_f+d_f} (\tau - q_f) x_f^\tau \\
 \text{(por Observación 1)} &\geq q_f + \sum_{\tau=q_f}^{q_f+d_f} (\tau - q_f)^r w_f^\tau \\
 \text{(por Observación 2)} &\geq q_f + \sum_{\tau=q_f}^{q_f+d_f} (\tau - q_f) \sum_{\tau_0 \in I_e^f(\tau, r)} z_{(e,f)}^{\tau_0} \\
 \text{(por Observación 3)} &\geq q_f - q_f z_a + \sum_{\tau=q_f}^{q_f+d_f} \tau \sum_{\tau_0 \in I_e^f(\tau, r)} z_{(e,f)}^{\tau_0} \tag{33}
 \end{aligned}$$

La Observación 3 relaciona la sumatoria sobre la ventana de tiempo de f con la sumatoria sobre la ventana de e . Además, dado $\tau \in W_f$, sea $\tau_0 \in I_e^f(\tau, r)$, sabemos que se cumple que $\tau \geq \tau_0 + L_a$. Luego, podemos reemplazar la expresión (33) por

$$x_f \geq q_f(1 - z_a) + \sum_{\tau_0=q_e}^{q_e+d_e} (\tau_0 + L_a) z_{(e,f)}^{\tau_0}. \tag{34}$$

Como podemos ver, la expresión (34) es equivalente a (32). Luego, dicha expresión actúa por un lado como una cota superior del lado izquierdo de la desigualdad (8) y por otro como una cota inferior de su lado derecho. Por dicho motivo, \hat{x} satisface la desigualdad (8).

4. *Desigualdades* (4). Inicialmente reescribimos la desigualdad de la siguiente manera

$$x_e + L_a \leq x_f$$

Tengamos en cuenta que estas desigualdades se definen para actividades $a = (e, f) \in A_{\text{tren}}$. Los valores de q_e y q_f son los horarios del cronograma original, los cuales dependen del tiempo de viaje o espera según corresponda. Inicialmente, asumimos que en el nuevo cronograma los tiempos mínimos de viaje o espera (i.e. L_a) son iguales o mayores al cronograma original, por ejemplo, debido a la menor disponibilidad de recursos de la red. Luego, sucede que $q_e + L_a \geq q_f$.

Sumamos la desigualdad (20) para cada $t = q_e \dots q_e + d_e$ y obtenemos

$$\sum_{t=q_e}^{q_e+d_e} \sum_{\tau=t}^{q_e+d_e} x_e^\tau \leq \sum_{t=q_e}^{q_e+d_e} \sum_{\tau=t+L_a}^{q_f+d_f} x_f^\tau. \tag{35}$$

Analizando el lado izquierdo de la desigualdad, tenemos que

$$\sum_{t=q_e}^{q_e+d_e} \sum_{\tau=t}^{q_e+d_e} x_e^\tau = \sum_{\tau=q_e}^{q_e+d_e} (\tau - q_e + 1) x_e^\tau = \sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau + (-q_e + 1) \sum_{\tau=q_e}^{q_e+d_e} x_e^\tau = x_e - q_e + 1. \tag{36}$$

En cuanto al lado derecho de (35), suponemos que $q_e + d_e + L_a < q_f + d_f$ y separamos la sumatoria de la siguiente manera

$$\begin{aligned}
 \sum_{t=q_e}^{q_e+d_e} \sum_{\tau=t+L_a}^{q_f+d_f} x_f^\tau &= \sum_{\tau=q_e+L_a}^{q_e+d_e+L_a} (\tau - (q_e + L_a) + 1)x_f^\tau + \sum_{\tau=q_e+d_e+L_a+1}^{q_f+d_f} (d_e + 1)x_f^\tau, \\
 &\leq \sum_{\tau=q_e+L_a}^{q_e+d_e+L_a} (\tau - (q_e + L_a) + 1)x_f^\tau + \sum_{\tau=q_e+d_e+L_a+1}^{q_f+d_f} (\tau - (q_e + L_a) + 1)x_f^\tau, \\
 &= \sum_{\tau=q_e+L_a}^{q_f+d_f} (\tau - (q_e + L_a) + 1)x_f^\tau. \tag{37}
 \end{aligned}$$

La expresión obtenida en (37) es equivalente a la que se obtiene en caso de que $q_e + d_e + L_a \geq q_f + d_f$, quedando ambos escenarios descriptos. Luego, reemplazando (36) y (37) en la desigualdad (35) obtenemos

$$\begin{aligned}
 x_e - q_e + 1 &\leq \sum_{\tau=q_e+L_a}^{q_f+d_f} (\tau - (q_e + L_a) + 1)x_f^\tau, \\
 (\text{por } q_e + L_a \geq q_f) &\leq \sum_{\tau=q_f}^{q_f+d_f} (\tau - (q_e + L_a) + 1)x_f^\tau, \\
 &= \sum_{\tau=q_f}^{q_f+d_f} \tau x_f^\tau + (-q_e - L_a + 1) \sum_{\tau=q_f}^{q_f+d_f} x_f^\tau, \\
 &= x_f - q_e - L_a + 1.
 \end{aligned}$$

Llegando a la expresión buscada.

También podría suceder que en el nuevo cronograma, los tiempos de viaje o espera sean menores a los originales. Para estos casos, sucede que $q_e + L_a < q_f$.

En particular, en caso de que $q_f > q_e + d_e + L_a$, la desigualdad (4) se cumple trivialmente dado que

$$x_e + L_a \leq q_e + d_e + L_a < q_f \leq x_f$$

Luego, nos enfocamos en el caso donde $q_e + L_a < q_f \leq q_e + d_e + L_a$.

Inicialmente, multiplicamos la desigualdad (20) para $t = q_e$ por $q_e + L_a$, obteniendo como resultado

$$\sum_{\tau=q_e}^{q_e+d_e} (q_e + L_a)x_e^\tau \leq \sum_{\tau=q_f}^{q_f+d_f} (q_e + L_a)x_f^\tau. \tag{38}$$

También agregamos las desigualdades (20) para $t = q_e + 1 \dots q_e + d_e$, las cuales sumadas a la desigualdad (38) dan como resultado

$$\sum_{\tau=q_e}^{q_e+d_e} (q_e + L_a)x_e^\tau + \sum_{t=q_e+1}^{q_e+d_e} \sum_{\tau=t}^{q_e+d_e} x_e^\tau \leq \sum_{\tau=q_f}^{q_f+d_f} (q_e + L_a)x_f^\tau + \sum_{t=q_e+1}^{q_e+d_e} \sum_{\tau=\max\{t+L_a, q_f\}}^{q_f+d_f} x_f^\tau. \tag{39}$$

Analizamos cada lado de la desigualdad por separado. Primero, para el lado izquierdo podemos ver

que

$$\begin{aligned} \sum_{\tau=q_e}^{q_e+d_e} (q_e + L_a)x_e^\tau + \sum_{t=q_e+1}^{q_e+d_e} \sum_{\tau=t}^{q_e+d_e} x_e^\tau &= \sum_{\tau=q_e}^{q_e+d_e} q_e x_e^\tau + \sum_{t=q_e+1}^{q_e+d_e} \sum_{\tau=t}^{q_e+d_e} x_e^\tau + L_a \sum_{\tau=q_e}^{q_e+d_e} x_e^\tau \\ \text{por (17) y (19)} &= \sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau + L_a \sum_{\tau=q_e}^{q_e+d_e} x_e^\tau = x_e + L_a. \end{aligned}$$

Ahora analizamos el lado derecho de la desigualdad (39). En particular, nos enfocamos en el segundo término de la suma.

Recordemos que por suposición $q_e + 1 \leq q_f - L_a \leq q_e + d_e$. Luego, podemos reescribir el término de la siguiente manera:

$$\sum_{t=q_e+1}^{q_e+d_e} \sum_{\tau=\max\{t+L_a, q_f\}}^{q_f+d_f} x_f^\tau = \sum_{t=q_e+1}^{q_f-L_a} \sum_{\tau=q_f}^{q_f+d_f} x_f^\tau + \sum_{t=q_f-L_a+1}^{q_e+d_e} \sum_{\tau=t+L_a}^{q_f+d_f} x_f^\tau. \quad (40)$$

Analizamos cada parte por separado:

a) El primer término se puede reescribir como

$$\sum_{t=q_e+1}^{q_f-L_a} \sum_{\tau=q_f}^{q_f+d_f} x_f^\tau = (q_f - q_e - L_a) \sum_{\tau=q_f}^{q_f+d_f} x_f^\tau.$$

b) Para el segundo término, el límite superior de la primera sumatoria puede ser reemplazado por $q_f + d_f - L_a$. En caso de que $q_e + d_e > q_f + d_f - L_a$ la nueva expresión es equivalente, dado que los términos excluidos se anulan sobre la sumatoria interna. En caso de que $q_e + d_e \leq q_f + d_f - L_a$, se consideran más términos en la sumatoria interna, y de esa manera se satisface la desigualdad. Además, al considerar la sustitución $t' = t + L_a$, obtenemos la siguiente expresión

$$\begin{aligned} \sum_{t=q_f-L_a+1}^{q_e+d_e} \sum_{\tau=t+L_a}^{q_f+d_f} x_f^\tau &\leq \sum_{t=q_f-L_a+1}^{q_f+d_f-L_a} \sum_{\tau=t+L_a}^{q_f+d_f} x_f^\tau \\ &= \sum_{t'=q_f+1}^{q_f+d_f} \sum_{\tau=t'}^{q_f+d_f} x_f^\tau. \end{aligned}$$

Combinando estas dos expresiones con (40) y reemplazando en el lado derecho de (39) obtenemos finalmente

$$\begin{aligned} x_e + L_a &\leq \sum_{\tau=q_f}^{q_f+d_f} (q_e + L_a)x_f^\tau + (q_f - q_e - L_a) \sum_{\tau=q_f}^{q_f+d_f} x_f^\tau + \sum_{t'=q_f+1}^{q_f+d_f} \sum_{\tau=t'}^{q_f+d_f} x_f^\tau \\ &= \sum_{\tau=q_f}^{q_f+d_f} q_f x_f^\tau + \sum_{t'=q_f+1}^{q_f+d_f} \sum_{\tau=t'}^{q_f+d_f} x_f^\tau \\ &= \sum_{\tau=q_f}^{q_f+d_f} \tau x_f^\tau = x_f. \end{aligned}$$

lo que completa la demostración para el último caso.

5. *Desigualdades* (14). La formulación TIF modela estas desigualdades mediante las restricciones (21). Combinando estas últimas con las desigualdades (19), tenemos que

$$\sum_{\tau=q_e}^{\tau_2-1} x_e^\tau + z_a \leq 1 = \sum_{\tau=q_e}^{q_e+d_e} x_e^\tau$$

obteniendo como resultado la siguiente expresión

$$z_a \leq \sum_{\tau=\tau_2}^{q_e+d_e} x_e^\tau.$$

A partir de esta desigualdad, podemos llegar a la expresión buscada

$$\tau_2 z_a \leq \tau_2 \sum_{\tau=\tau_2}^{q_e+d_e} x_e^\tau \leq \sum_{\tau=q_e}^{\tau_2-1} \tau x_e^\tau + \sum_{\tau=\tau_2}^{q_e+d_e} \tau_2 x_e^\tau \leq \sum_{\tau=q_e}^{q_e+d_e} \tau x_e^\tau = x_e.$$

A partir de estos cinco puntos, podemos ver que la solución (fraccionaria) $\hat{x} \in P_{TIF}^{LP}$ es también factible para P_{Orig}^{LP} , y por lo tanto la proposición $\hat{z}_{orig} \leq \hat{z}_{tif}$ queda demostrada. ■

Extracto cronograma tren Mitre, ramal Retiro-Tigre

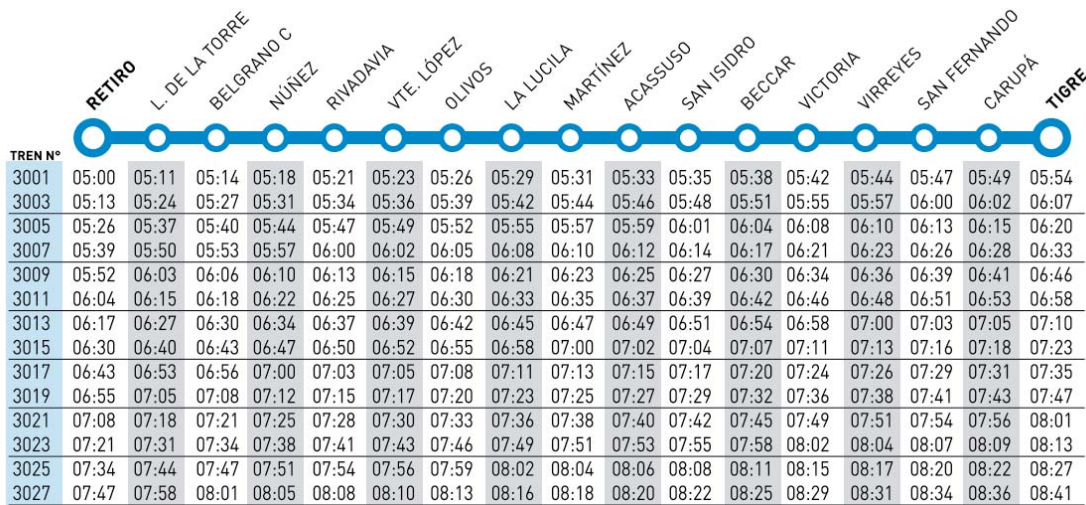


Figura 5: Ramal Retiro-Tigre. Extracto del Cronograma 2016 para trenes viajando desde Retiro hacia Tigre, utilizado en días Lunes a Viernes hábiles.

REFERENCIAS

[1] N. Ascheuer, M. Fischetti, and M. Grötschel. Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut. *Mathematical Programming*, 2001.

[2] N. Besinovic, V. Cacchiani, T. Dollevoet, R. M. Goverde, D. Huisman, M. P. Kidd, L. G. Kroon, E. Quaglietta, J. Rodriguez, P. Toth, et al. Integrated decision support tools for disruption management, 2015.

[3] V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf, and J. Wagenaar. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63(0):15–37, 2014.

[4] S. Dash, O. Günlük, A. Lodi, and A. Tramontani. A time bucket formulation for the traveling salesman problem with time windows. *INFORMS Journal on Computing*, 24(1):132–147, 2012.

- [5] I. Louwerse and D. Huisman. Adjusting a railway timetable in case of partial or complete blockades. *European Journal of Operational Research*, 235(3):583–593, 2014.
- [6] L. P. Veelenturf, M. P. Kidd, V. Cacchiani, L. G. Kroon, and P. Toth. A Railway Timetable Rescheduling Approach for Handling Large-Scale Disruptions. *Transportation Science*, 2015.