



Ciencia e Ingeniería Neogranadina  
ISSN: 0124-8170  
[revistaing@unimilitar.edu.co](mailto:revistaing@unimilitar.edu.co)  
Universidad Militar Nueva Granada  
Colombia

Jiménez Narváez, Marvin Luis; Soto Mejía, José A.  
UTILIZACIÓN DE TARJETAS INTELIGENTES PARA ESTIMAR MATRICES ORIGEN-  
DESTINO. APLICACIÓN AL SISTEMA MEGABÚS, PEREIRA  
Ciencia e Ingeniería Neogranadina, vol. 26, núm. 2, 2016, pp. 73-93  
Universidad Militar Nueva Granada  
Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=91146925005>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

# UTILIZACIÓN DE TARJETAS INTELIGENTES PARA ESTIMAR MATRICES ORIGEN-DESTINO. APlicación AL SISTEMA MEGABÚS, PEREIRA

## ESTIMATING AN ORIGIN-DESTINATION MATRIX USING SMART CARDS. AN APPLICATION TO THE MEGABÚS TRANSPORT SYSTEM IN PEREIRA

Marvin Luis Jiménez Narváez<sup>1</sup>, José A. Soto Mejía<sup>2</sup>

**Fecha de recepción:** 21 de enero de 2016

**Fecha de revisión:** 29 de abril de 2016

**Fecha de aprobación:** 2 de junio de 2016

Referencia: M. L. Jiménez Narváez, J. A. Soto Mejía (2016). Utilización de tarjetas inteligentes para estimar matrices origen-destino. Aplicación al sistema megabús, Pereira. Ciencia e Ingeniería Neogranadina, 26 (2), pp. 73-93, DOI: <http://dx.doi.org/10.18359/rcin.1661>

### RESUMEN

En la planificación operativa de los sistemas de transporte es fundamental contar con una caracterización de la demanda del servicio por medio de una matriz origen-destino (OD) de viajes. Esta matriz es estimada frecuentemente mediante técnicas de muestreo estadístico que ofrecen resultados de gran calidad pero requieren diseño e implementación compleja y de alto costo. Así, se hace pertinente aprovechar la información de los registros de las tarjetas inteligentes usadas como medios de pago en los sistemas que tienen esta tecnología. Esta fuente de información permite obtener estimaciones de matrices OD de gran calidad con un costo bajo. En este trabajo se implementa una metodología para la estimación de una matriz origen-destino entre estaciones de buses en el sistema de transporte público de Megabús en la ciudad de Pereira. La metodología

1. Ing. Industrial, M.Sc. (c) Investigación Operativa y Estadística, investigador en el Grupo Análisis Envolvente de Datos Universidad Tecnológica de Pereira, Pereira, Colombia, marvinjimenez@utp.edu.co

2. Físico, PhD. Ciencias de la Computación, Profesor titular, director del Grupo Análisis Envolvente de Datos, Universidad Tecnológica de Pereira, Pereira, Colombia, jomejia@utp.edu.co

permite depurar la base de datos teniendo en cuenta las particularidades del sistema, como las formas de ingreso, tipo de tarjeta, reventa de pasajes, y además considera la posibilidad de que un usuario pueda decidir bajarse en una estación intermedia cercana a su estación destino mediante la parametrización de un factor de penalización que está asociado con el tiempo de viaje caminando en la función objetivo. La inclusión del factor no influye significativamente en el porcentaje total de viajes estimados pero sí ajusta algunos valores de la matriz origen-destino, al considerar los descensos en estaciones intermedias. El método posibilita estimar matrices OD para diferentes ventanas de tiempo de cualquier día de la semana y distintas épocas del año.

**Palabras clave:** matriz origen-destino, tarjeta inteligente, sistema de transporte público, Megabús.

## ABSTRACT

In the operational planning of transport systems, it is essential to have a characterization of the demand for the service by means of an Origin-Destination (OD) Matrix of travels. This matrix is frequently estimated using statistical sampling techniques, which offer high-quality results but require complex design and high cost. It is therefore appropriate to take advantage of the information that is generated from the records of the smart cards used as means of payment in the systems that provide this technology. This source of information allows for obtaining estimates of high-quality Origin–Destination (OD) matrix with a low cost. In this work, it is implemented a methodology for the estimation of an Origin–Destination (OD) matrix between bus stations in the public transport system of MEGABÚS in the city of Pereira. The methodology allows for a debugging in the database taking into account the peculiarities of the system as the forms of income, type of card, and resale of tickets. In addition, it considers the possibility that a user can decide to step down at an intermediate station close to his destination station using the parameterization of a penalty factor that is associated with the walking-travel time in the objective function. The inclusion of the factor does not significantly influence the total percentage of estimated travels. Nevertheless, it recalculates some values of the Origin-Destination (OD) Matrix considering the alight in intermediate stations. The proposed method makes possible to estimate OD matrices for different time windows for any day of the week and different times of the year.

**Keywords:** origin-destination matrix; smart card; public transport system; Megabús.

## INTRODUCCIÓN

Para la planificación de un sistema de transporte público es indispensable contar con una matriz origen-destino (MOD) de alta calidad, que permita caracterizar la demanda del servicio. Sin embargo, los métodos más frecuentes y aparentemente más adecuados para la estimación de matrices OD, como el muestreo, implican un alto costo y gran complejidad en su diseño, implementación y análisis. Estas implicaciones se ven magnificadas en sistemas de transporte integrados de alta complejidad en las grandes ciudades con redes de tráfico de gran congestión, donde se hace necesario fraccionar geográficamente y dividir en ventanas de tiempo adecuadas que contemplen distintas horas del día, días de la semana y temporadas del año.

Por lo anterior, la introducción de la tecnología de la información en los sistemas de transporte público, como el uso de las tarjetas inteligentes (*smart cards*) como medio de pago, se convierte en una gran alternativa para la recopilación, almacenamiento y análisis de datos. Esto permite contar con información valiosa con la que se pueden obtener estimaciones de matrices origen-destino de gran calidad a un menor costo. Desde su aparición, el uso de las tarjetas inteligentes ha venido en aumento y han sido implementadas como medio de pago en diferentes ciudades alrededor del mundo, como San Francisco [1], Portland [2], Nueva York [3], Chicago [4], Ámsterdam (Países Bajos) [2-5], Changchun (China) [6], Gatineau-Quebec (Canadá) [7-9], Santiago (Chile)<sup>3</sup> y Pereira (Colombia)<sup>4</sup>, entre otros. A excepción

de Pereira y Santiago, en las ciudades antes nombradas la tarjeta inteligente es utilizada como un medio de pago parcial, en algunos casos con descuento en el costo del pasaje.

El objetivo principal de esta investigación es aprovechar la información recopilada a partir de los registros de pago de las tarjetas en el sistema de Megabús para estimar una MOD que permita caracterizar la demanda del servicio dentro de la zona troncal del sistema y sirva de insumo para la planificación operativa del mismo. Este problema ha sido abordado por otros investigadores desde distintas ópticas en algunos sistemas de transporte público, se han propuesto diferentes metodologías y considerado algunos supuestos que permiten utilizar la información de los registros de las tarjetas inteligentes y así obtener una MOD de calidad.

En [3] se propone un método para estimar la estación de bajada de los usuarios del metro de Nueva York, basándose en los siguientes supuestos: 1. Después de realizar un viaje, el usuario, en otro viaje, tendrá como estación de origen la estación que fue destino del viaje anterior. 2. Al finalizar el día, el usuario volverá a la estación en donde se registró para realizar su primer viaje. En [10], en Chicago se implementó una metodología para inferir el destino de un usuario, cuando solo se registra al ingresar al sistema de transporte de buses con carril dedicado, y tiene en cuenta los siguientes supuestos: 1. El usuario se registra nuevamente para hacer su próximo viaje en el destino del viaje anterior. 2. El usuario termina su último viaje del día donde se registró en el

3. Información obtenida de: [www.transantiago.cl](http://www.transantiago.cl)

4. Información obtenida de: [www.megabus.gov.co/?page\\_id=249](http://www.megabus.gov.co/?page_id=249)

primer viaje. Si bien los supuestos son tomados de la propuesta presentada en [3], el aporte está en la implementación de la metodología en un sistema de buses con sistema de localización automática.

En [11] se aborda el mismo problema y se cruza la información de la hora y lugar de los registros de las tarjetas en cada ruta de bus con la hora de llegada del bus a cada paradero, además de tener en cuenta la ubicación geográfica de este último. De esta manera se logra estimar los puntos de bajada de los usuarios mediante el seguimiento de las validaciones realizadas durante el día. Sin embargo, en esta investigación se enfoca la atención en el problema de los trasbordos, y se establecen principalmente reglas para que dos validaciones consecutivas de un mismo usuario se consideren como dos etapas de un viaje, y no un viaje en sí. Estas reglas son: 1. Ambas validaciones fueron ejecutadas el mismo día. 2. El segundo registro se presenta en los siguientes 90 minutos después del primero. 3. La ruta en la que se presenta el segundo registro debe ser diferente a la del primero. Despues del procesamiento de los registros bajo las reglas descritas se logra generar una estimación de la MOD.

En [8] se lleva a cabo un estudio en el sistema de transporte público en Canadá; en este sistema se utiliza la tarjeta inteligente como uno de los varios medios de pago que el sistema permite. La tarjeta con la cual se accede al medio de transporte es de uso personal, además el sistema se encuentra equipado con tecnología que permite conocer la ubicación de los buses. El hecho de que existan distintos medios de pago lleva a que se presenten errores en el registro de observaciones y en el de cobros automáticos, por lo cual en [8] se centra la investigación en proponer un método para

identificar y reemplazar las observaciones erróneas o sospechosas del sistema automático de cobro de pasajes.

En [7] se muestra una metodología de estimación de matrices OD para sistemas de buses en Gatineau-Quebec (Canadá). Estos buses cuentan con sistemas que permiten conocer su ubicación, por lo cual se puede estimar la posición del registro del usuario. Los autores proponen un método para estimar el punto de descenso de los pasajeros en cada viaje, teniendo en cuenta que los usuarios solo validan cuando ingresan al sistema. Además de usar los supuestos contemplados en [3], en esta metodología se establece un criterio para estimar el lugar de descenso del usuario, que consiste en asignar como destino el paradero que se encuentra a la menor distancia de la siguiente transacción, suponiendo que algunas personas caminarían de un paradero hasta otro si es necesario, pero con un límite de recorrido denominado distancia máxima caminable. Esta última limitante se incorpora en el modelo propuesto en [7] como una restricción. Para el último registro del día, suponen que el usuario regresará al lugar de inicio, es decir, el usuario retornará a la parada en la que realizó la primera transacción del día.

En [6] se propone para la ciudad de Changchun (China) una metodología para estimar el origen y el destino de los viajes de cada usuario para un sistema que solo está constituido por buses, donde la validación de la tarjeta se da en el interior del vehículo, y se guarda automáticamente la hora de registro. Además, cada uno de los conductores debe llevar un registro de todas las horas de llegada a cada paradero. Esto permite realizar un cruce de información similar al llevado a cabo en [11]. La metodología propuesta se basa en fuertes supuestos que afirman que el lugar de destino

del viaje del primer registro del día, sería el lugar donde se presenta la última validación del día y el destino del último registro será el paradero donde se realizó la primera validación. En [4], en Chicago, se plantea un modelo para estimar el lugar que tienen como destino los usuarios del transporte público, teniendo en cuenta que este sistema está constituido por trenes y buses, cada uno de estos con sistema de cobro automático en el origen, las validaciones en los buses se realizan en el interior, mientras que en los trenes se da en las estaciones donde el usuario lo aborda. Los buses cuentan con sistema de posicionamiento, esto ayuda a la verificación de los registros de los usuarios. Se utilizan los supuestos planteados en [3] y se aplica la misma metodología de minimización de la distancia propuesta en [7]. Sin embargo, el aporte radica en la integración de los distintos modos de transporte (bus y tren), en donde las líneas de tren son trabajadas como líneas de buses, con la diferencia de que el posicionamiento de las estaciones de tren no necesita ser estimado como los puntos de paradas de los buses, puesto que la localización de las estaciones es conocida.

En [12], en São Paulo, Brasil, se aplica la metodología propuesta en [10]. En este estudio solo se consideraron los usuarios que utilizaron los vehículos con sistemas de posicionamiento geográfico. En [13], en Chile, se realiza un primer trabajo para el caso del sistema Transantiago, en donde se utiliza una metodología para estimar el destino de un usuario basada en la propuesta mostrada en [7], con ediciones importantes al enfocar el criterio de asignación de destinos en la minimización del tiempo de viaje. Esta metodología es formalizada y presentada con detalle en un siguiente trabajo realizado en el año 2012. Para una revisión detallada de la literatura, se puede ver la presentada en [14].

En [15], en Santiago de Chile y Gatineau (Canadá), se propone una metodología que contempla dos aspectos: 1. Basándose en los anteriores estudios de estimación de destino realizados en [13], se reconocen y diferencian entre actividades y viajes realizados por los usuarios. 2. A las estimaciones de destino realizadas y con información adicional se les asigna un propósito. Luego se realiza un análisis de la capacidad de la metodología y se comparan los resultados entre estas dos ciudades, para encontrar que la metodología es aplicable en ambos sistemas de transporte, y que permite, de manera general, comprender el comportamiento de los usuarios en el ámbito temporal, de propósito de viaje y naturaleza de las actividades. Esto les permitió detectar importantes diferencias entre los sistemas de ambas ciudades, principalmente en las horas del día de ejecución de las actividades, las cuales dependen mucho de las características geográficas, sociales, culturales y políticas de la ciudad. En [16], en Santiago (Chile), se implementa un método para la estimación de los destinos, basado en la minimización del tiempo de viaje generalizado entre dos puntos de registros del usuario. El método tiene en cuenta la información suministrada por el sistema de localización GPS, con el que cuentan los buses, el cual registra, cada cierto tiempo, información de la hora y posición. En la metodología se considera la posibilidad de que el punto de registro de un usuario no necesariamente es el destino del viaje iniciado en un registro anterior, puesto que los usuarios pueden decidir bajarse en un punto intermedio y caminar hacia el lugar donde realizan su próxima transacción. Por esto, en vez de la distancia, se tiene en cuenta el tiempo del viaje en bus del usuario desde el anterior registro hasta cada punto donde el sistema de localización generó una coordenada, de posición y tiempo, en una hora que se encuentra dentro de una ventana

de tiempo definida. Luego se suma el tiempo de viaje a pie desde el punto intermedio hasta el lugar del siguiente registro. El tiempo total de viaje es denominado tiempo generalizado. Adicionalmente, se incorpora la restricción de distancia máxima caminable propuesta en [4]. Lo propuesto en [16] posee gran valor al incluir la minimización del tiempo de viaje generalizado como criterio para la estimación de los destinos de los viajes, además de ser aplicado a un sistema multimodal integrado de buses y metro.

El presente trabajo aborda el problema de la estimación de una matriz origen-destino entre estaciones (solamente zona troncal) a partir de las transacciones ejecutadas por los usuarios con la tarjeta inteligente para el sistema de transporte público, Megabús, de la ciudad de Pereira, Colombia. La estimación correcta de la MOD es un insumo necesario para la generación óptima de las tablas horarias de los buses articulados del sistema, así como en algunas estrategias de control [17]. No se cuenta con sistemas de localización en buses ni estaciones. Se utilizan varios supuestos propuestos por autores en los trabajos ya descritos, que serán detallados en la Sección 3. Para la estimación de los destinos de cada viaje se siguió la metodología propuesta en [16] por su aplicabilidad al sistema de Pereira, con adaptaciones que contemplan las especificidades propias del sistema Megabús. Se incluye una etapa que permite depurar la información y que responde a particularidades que no se presentan en los otros sistemas de transporte ya revisados. El presente artículo está organizado como sigue: en la Sección 1 se realiza una descripción del sistema de Megabús, en donde se detallan las principales características y particularidades que en este se presentan; en la Sección 2 se hace un análisis exploratorio de los datos con los

que se trabajó para la estimación de la MOD; en la Sección 3 se describe la metodología implementada; después, en la Sección 4 se muestran y discuten los resultados y, finalmente, en la Sección 5 se presentan las conclusiones y trabajos futuros.

## 1. DESCRIPCIÓN GLOBAL DEL SISTEMA

El sistema de transporte Megabús está conformado por dos "tipos de ruta": las troncales y las alimentadoras. Las troncales son de tipo BRT (*Bus Rapid Transit*) y son transitadas por los buses articulados, que se mueven principalmente en el centro de la ciudad (zona troncal). Los articulados realizan paradas solo en las estaciones de Megabús, siendo estas los únicos puntos en los que un pasajero puede subir o bajar de un bus articulado. El sistema cuenta con un total de 40 estaciones, incluyendo las de "Intercambiadores" (ver Figura 1).

Las rutas alimentadoras están definidas desde distintos puntos de partida, distribuidos en los sectores de Cuba, Dosquebradas y los alrededores de la estación "Viajero", como se muestra en la Figura 1 (el mapa no incluye las rutas alimentadoras asociadas con esta estación). Los buses alimentadores no poseen carril dedicado, y sus paradas están sujetas a la demanda del servicio en cualquier punto de la ruta.

En el sistema de Megabús el único medio de pago es la tarjeta inteligente. Esta es validada por el usuario al ingresar a una estación, o a un bus alimentador. La validación solo es realizada al ingresar al sistema. Cada tarjeta está identificada por un serial único. Este sistema de pago alimenta una base de datos donde se tiene información sobre las estaciones de las

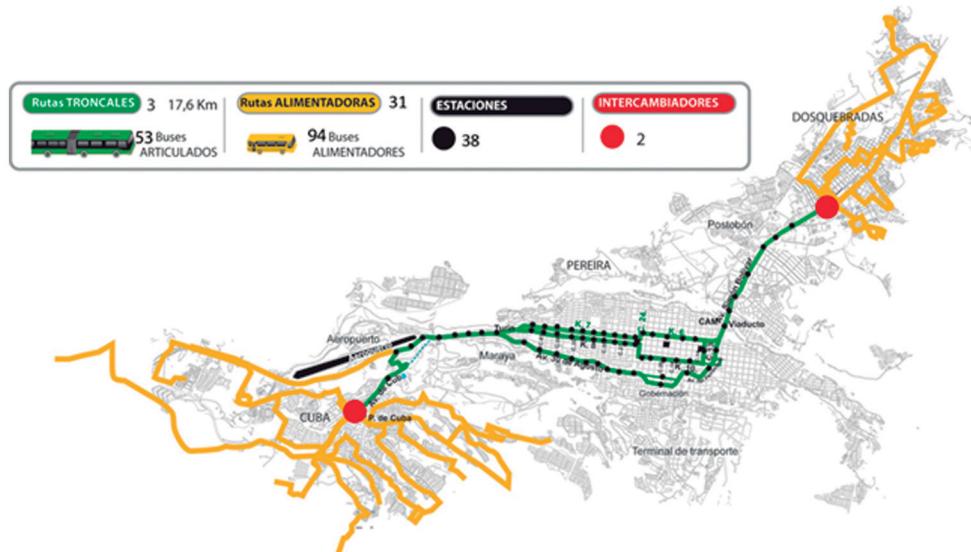


Figura 1. Mapa de cobertura del sistema.

Fuente: Megabús.

troncales o las rutas alimentadoras en las que es validada cada tarjeta, además de la hora y fecha de registro. A continuación se detallan las particularidades del sistema, que se tienen en cuenta al momento de aplicar la metodología propuesta en el presente artículo.

*Formas de ingreso al sistema.* Se puede ingresar de dos formas: entrando directamente a una estación de las rutas troncales, para luego esperar por un bus articulado, o abordando un bus alimentador en algún punto de la ruta alimentadora. Los buses alimentadores ingresan a los usuarios al intercambiador, en donde no se necesita volver a validar, y puede tomar cualquier bus articulado que lo lleve a su destino. Con la segunda forma de ingresar al sistema, es posible que el usuario utilice el bus alimentador para transportarse dentro de un determinado sector (Cuba, Viajero o Dosquebradas) sin ingresar al intercambiador correspondiente, o que al ingresar al

intercambiador aborde otro bus alimentador del mismo sector de una ruta diferente, lo cual lo convierte en demandante del servicio externo a la zona troncal.

*Tarjeta retornable.* En Megabús existe un tipo de tarjeta denominada "tarjeta blanca univiaje retornable", que es facilitada al usuario en la caja de la estación por el costo del pasaje, más un valor adicional que se le devolverá al momento de entregar la tarjeta en la estación destino. Puesto que la tarjeta retornable, después de adquirida en calidad de "préstamo", es devuelta por el usuario en la estación en la que se baja, puede ser facilitada a otro usuario bajo las mismas condiciones, por lo cual la misma tarjeta puede ser usada por diferentes usuarios el mismo día. Esto impide construir un patrón de viaje del usuario a lo largo de la semana, debido a que el registro asociado a la tarjeta no hace referencia al mismo usuario ni a sus hábitos de viaje.

*Uso múltiple de la tarjeta.* En el sistema de Pereira no es una regla que la tarjeta sea usada exclusivamente por su propietario. Tampoco existe una restricción de frecuencia de uso de la tarjeta en un intervalo de tiempo determinado, es decir, no existe prohibición alguna sobre el uso inmediato de la tarjeta una cantidad de veces en un corto tiempo, por lo que un grupo numeroso de personas puede usar la misma tarjeta para ingresar juntas a una misma estación.

*Reventa de pasajes.* Se ha observado que algunas personas recargan su tarjeta con saldo suficiente para, posteriormente, ubicarse en algunas estaciones (muy pocas, en general una o dos) o en algunos paraderos de buses alimentadores para revender los pasajes a otras personas que no poseen tarjeta de Megabús. Lo anterior lleva a que se presenten grandes cantidades de registros durante el día, con tiempos muy pequeños entre cada registro, y la mayoría en la misma estación de troncal o ruta alimentadora. Esto hace compleja la trazabilidad de los viajes realizados a partir de pasajes de *reventa* y no permite estimar una estación destino.

*Usuarios con un solo registro al día.* Otro fenómeno que se presenta está asociado con los registros unitarios, es decir, registros de tarjetas que son validadas una sola vez en el día. Esto dificulta o imposibilita la estimación de la estación de bajada de ese usuario.

*Trasbordos.* Se definen como la acción del usuario de cambiar de ruta al bajar de un bus en una estación y, sin salirse de la zona de espera de la estación, aborda otro que sigue una ruta troncal diferente. El usuario busca alcanzar una estación destino que no se encuentra directamente en la ruta inicial de su viaje pero que, mediante trasbordos, se le hace accesible.

Por la flexibilidad del sistema Megabús, los usuarios pueden realizar la cantidad de trasbordos que deseen, sin límites de tiempo, sin incurrir en pagos de pasaje adicionales.

## 2. ANÁLISIS EXPLORATORIO DE LOS DATOS

Para la estimación de la MOD se utilizó la información concentrada en los registros realizados desde el lunes 19 hasta el domingo 25 de mayo de 2015. Estas fechas fueron escogidas por su generalidad al no estar afectadas por temporadas de baja demanda, como Semana Santa, vacaciones académicas o festividades de alguna naturaleza.

En la semana analizada se registraron un total de 607.582 validaciones, con un promedio de 94.731 registros diarios en días laborales; es decir, aproximadamente el 15,6% de los registros semanales corresponden a las validaciones de un día laboral. Este porcentaje disminuye a 12,73% para el sábado (77.344 registros) y a 9,31% para el domingo. En promedio, el 58,82% de las validaciones diarias son realizadas en estaciones troncales, y el 41,18% en buses alimentadores. En la Figura 2 se presenta la distribución temporal del número de registros para los días laborales. El número de validaciones por hora, en cada uno de los días de la semana evaluada, presenta un comportamiento similar.

Para los cinco días se puede apreciar claramente un pico en la mañana, entre las 6:00 y las 8:00, seguido de un valle, en el que disminuye considerablemente la demanda, hasta las 12:00 horas, donde se presenta un pico extendido de menor magnitud hasta las 17:00 horas. Sin embargo, el número de registros alcanza su máximo al finalizar la jornada laboral, entre las 17:30 y 19:30.

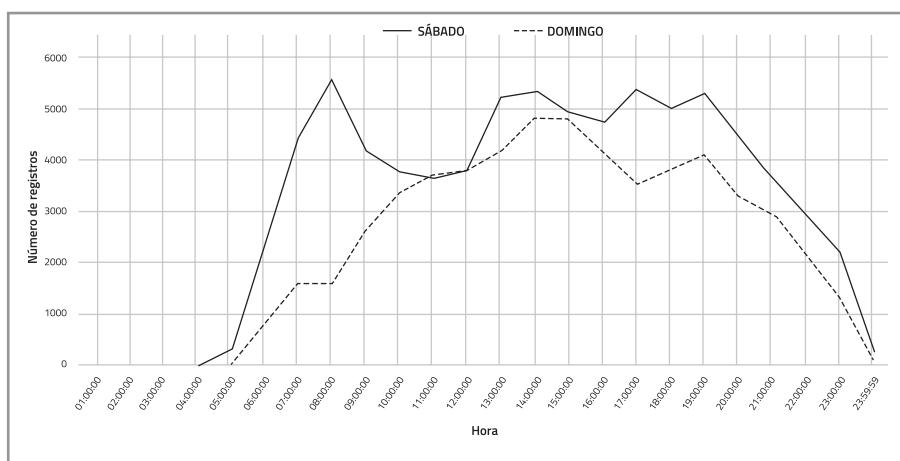


**Figura 2.** Distribución temporal de las validaciones. Días laborales.

**Fuente:** Elaboración propia.

Esta información permite identificar las ventanas de tiempo en donde se presenta la mayor demanda del servicio, orientando así los intervalos de tiempo de estimación de la MOD. Para el día sábado el comportamiento es bastante similar en el pico de la mañana y el valle hasta el mediodía (ver Figura 3). Pero desde las 13:00 horas se presenta una

demandas de magnitud similar a la del pico de la mañana, y además se mantiene cada hora con poca variabilidad hasta las 19:00 horas. La distribución de las validaciones de los domingos posee un comportamiento distinto, puesto que el pico se presenta entre las 13:30 y 15:30.



**Figura 3.** Distribución temporal de las validaciones. Sábado y domingo.

**Fuente:** Elaboración propia.

A modo de ejemplo, en la Figura 4 se presenta un diagrama de barras que muestra la cantidad de tarjetas que presentaron 1, 2, 3, 4 o más registros durante el día lunes.

Aproximadamente el 96,04% de las tarjetas poseen un número de registros, para un día lunes laboral, que no superan las 4 validaciones, mientras que 118 del total (46.560) de tarjetas se registraron más de 14 veces. Tarjetas con registros diarios atípicos sugieren la presencia del fenómeno de *reventa de pasajes*. Por otro lado, en Megabús se ha observado que las personas que revenden pasajes lo hacen, principalmente, en solo una o dos estaciones durante el día. Esta información permite establecer parámetros para la depuración de los registros que hacen referencia a la reventa de pasajes. La anterior información es utilizada en el proceso que se describe en la Sección 3.1 (Depuración de la base de datos).

### 3. METODOLOGÍA

La esencia de la metodología implementada consiste en reconstruir los viajes realizados por

los usuarios (tarjetas) dentro de la zona troncal del sistema cada día, a lo largo de un período de tiempo (días, semanas), considerando a la estación en donde se registra la tarjeta como un origen y a la estación donde se realiza la siguiente validación de la misma tarjeta, algún tiempo después, como un posible destino, siempre y cuando se presente en una estación diferente a la del origen. El proceso se llevó a cabo en varias fases: depuración de la base de datos, estimación del destino de cada viaje, pos-procesamiento de registros unitarios y obtención de la matriz origen-destino.

#### 3.1 Depuración de la base de datos

En la primera fase se abordaron cada una de las particularidades descritas en la Sección 1, iniciando con la identificación de los *usuarios con un solo registro al día*. Estos usuarios no fueron tenidos en cuenta en el proceso de reconstrucción de los viajes, por la ausencia de registros siguientes que puedan brindar información de su potencial destino. Seguidamente, se procedió a la identificación de las tarjetas (código) con las cuales se realizan

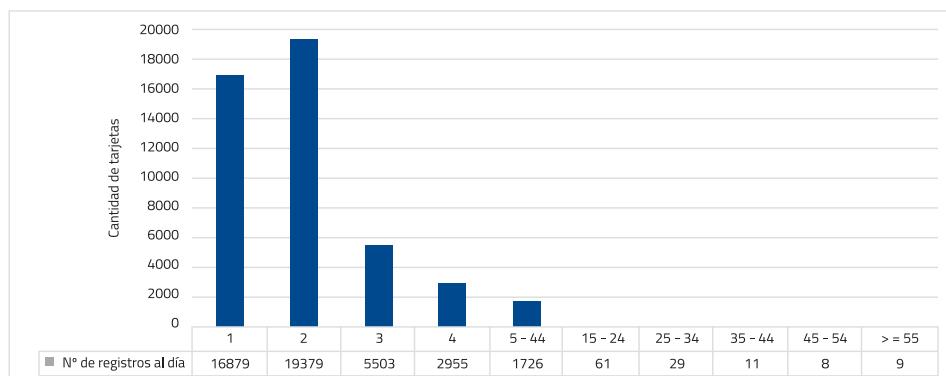


Figura 4. Cantidad de tarjetas contra el número de registros un día lunes.

Fuente: Elaboración propia.

*reventa de pasajes*. Esto se logró al incluir dos parámetros de depuración. Se clasificaron como reventa de pasajes a aquellas tarjetas que superaron el umbral de registros diarios totales (Rdt), o de registros en una misma estación (Re), considerados no atípico según lo identificado en el análisis exploratorio y lo sugerido por expertos de Megabús. Estos parámetros pueden ser modificados dentro del algoritmo de depuración. En este trabajo se definió el umbral de registros totales y de registros por estación como: Rdt = 14 y Re = 4.

Con respecto a las *formas de ingreso al sistema*, dado que el interés final es estimar la demanda dentro de las rutas troncales, se identificaron los usuarios que no ingresan al sistema de troncales. Lo anterior se logró al detectar aquellas validaciones que se realizaron en alguna ruta alimentadora de un sector dado (Cuba, Viajero, Dosquebradas), y que su siguiente registro, tiempo más tarde, se presentó en una ruta alimentadora perteneciente al mismo sector. Un registro posterior en un sector diferente o en una estación de troncal, supone que el usuario sí ingresó al sistema de troncales para transportarse entre sectores distintos. En el caso anterior se asigna como origen del viaje, dentro del sistema troncal, el intercambiador al cual convergen los buses alimentadores del sector en donde se registró la tarjeta. Los registros que hacen referencia a viajes que no entran a troncales son identificados en la base de datos y no son tenidos en cuenta a la hora de estimar los destinos de cada viaje. La *tarjeta retornable* fue considerada como un potencial inconveniente por no permitir reconocer patrones de viaje a lo largo de la semana. Sin embargo, esto no es un problema para el método de estimación de destinos que se implementa en este trabajo, dado que lo que se pretende es reconstruir los viajes realizados

cada día de la semana, sin tener en cuenta si son o no del mismo usuario.

En cuanto al *uso múltiple de la tarjeta*, si el caso es tal que la misma es usada por un grupo de personas de manera consecutiva en la misma estación, solo se podrá estimar el destino de una de ellas si se detecta otro registro de la tarjeta en una estación distinta tiempo después. La importancia del asunto del *trasbordo* radica en el impacto que este tiene sobre la MOD, puesto que un trasbordo no es un viaje completo, pero sí una etapa de este. Sin embargo, se consideró que en cada etapa de trasbordo se demandan rutas troncales distintas, por lo cual pueden entenderse como viajes diferentes que demandan el servicio de transporte entre distintos pares de estaciones, desde la perspectiva de la estimación de la MOD. La anterior suposición genera demandas extra en ciertos puntos del trayecto, las cuales no están respaldadas con un ingreso monetario al sistema, dado que no hay un nuevo pago en el trasbordo. En este sentido, la MOD estimada no es un insumo adecuado para la estimación de ingresos y de algunos otros indicadores.

### 3.2 Estimación del destino de cada viaje

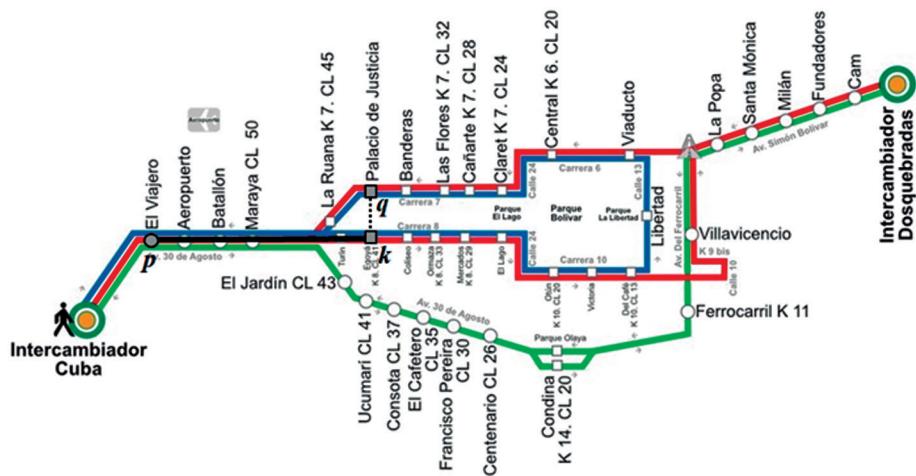
Suponiendo que un usuario ingresa al sistema y valida su tarjeta en una estación determinada, es natural creer que la estación de bajada (destino) del pasajero corresponde a la estación en la que hace su siguiente registro. Esta suposición se fundamenta en el hecho de que es razonable que el usuario se baje en la estación más cercana al sector de la ciudad al que se dirige, y de igual forma a la hora de regresar ingresará al sistema por medio de la estación más cercana al lugar en que se encuentra, la cual coincide con la estación de bajada del viaje inicial. Por lo anterior, para

estimar el destino de cada viaje solo bastaría identificar la estación en la que se realiza la siguiente validación. Sin embargo, por el diseño de las rutas de Megabús es posible que los usuarios decidan bajarse en una estación distinta a la del siguiente registro y continuar el viaje caminando, motivados por la reducción de tiempo de viaje que esto podría implicar. En la Figura 5 se presenta un mapa simplificado de las tres rutas troncales del sistema con la descripción de su recorrido.

En esta figura se ejemplifica el caso de un usuario que ingresa y valida su tarjeta en la estación *El Viajero* (nodo *p*) y su siguiente registro, tiempo más tarde, es en la estación *Palacio de Justicia* (nodo *q*). Si se supone que la estación destino del usuario es *Palacio de Justicia*, el usuario tomaría un bus de la ruta 3 (Cuba-Libertad-Cuba), y realizaría todo el recorrido hasta llegar a dicha estación. Con esta alternativa, al usuario le tomaría aproximadamente 30 minutos llegar a su

destino. Sin embargo, si el usuario decide bajarse en la estación *Egoyá* (nodo *k*) y caminar hasta la estación *Palacio de Justicia* (a 170 metros de esta) tardaría solo 9,15 minutos viajando en bus hasta la estación *Egoyá*, y 2,02 minutos caminando hasta la estación *Palacio de Justicia*, para un total de 11,17 minutos de viaje (si se considera una velocidad de caminata de 1,4 m/seg). Así, es muy plausible considerar que el usuario opte por la segunda alternativa.

Por la frecuencia con la que se pueden presentar este tipo de casos y la calidad de la estimación que otorga tener en cuenta el interés natural del usuario de tardar el menor tiempo posible en el recorrido, es recomendable escoger como destino de cada viaje la estación que implica el menor tiempo de viaje, incluido el tiempo en bus y de caminata; este tiempo es denominado tiempo generalizado [16]. No obstante, se deben considerar los casos en los que realizar todo el recorrido en bus iguale o supere por muy poco tiempo a la alternativa de



**Figura 5.** Mapa simplificado de la zona troncal del sistema.

Fuente: Megabús.

viajar en bus hasta una estación  $k$  y caminar hasta la estación destino. Esto es importante porque se debe tener en cuenta que la decisión del usuario también se ve afectada por su preferencia de transportarse en vehículo. Una solución a esta situación es penalizar el tiempo de caminata, para hacer menos atractiva la alternativa que incluye recorridos a pie. De esta manera, la estación que posea el mínimo tiempo generalizado será establecida como el destino del viaje.

Ya definido el criterio de selección del destino de un viaje, es indispensable tener en cuenta el cumplimiento de dos condiciones a la hora de considerar a una estación como posible destino:

1. La distancia entre la estación  $k$ , considerada como posible destino, y la estación del siguiente registro no debe superar la distancia máxima caminable de una persona (parámetro que se define previamente, para el caso 400 metros). Supuesto propuesto en [7].
2. La hora de registro en la estación origen, sumada al tiempo generalizado de viaje y al tiempo mínimo de actividad de un usuario en su lugar de destino, no debe sobrepasar la hora de registro de la estación donde se valida nuevamente la tarjeta. El tiempo mínimo de actividad es un parámetro de este método que debe establecerse y que tiene en cuenta que cuando un usuario se dirige a una estación es porque tiene un objetivo o actividad a realizar, la cual posee un tiempo de realización (en la presente investigación se tomó como tiempo mínimo 15 minutos). Si una de estas dos condiciones no se cumple, la estación  $k$  no se considera como posible destino. Además, si ninguna estación cumple la segunda condición, entonces el destino del viaje no es estimable. El enfoque anterior se basa en la metodología propuesta en [16] para el sistema de Transantiago, y fue adaptada al caso de Megabús. Las adaptaciones se refieren a ajustes en el valor de los parámetros y a la inclusión del

tiempo mínimo de actividad potencial tenido en cuenta en la segunda condición.

Así, la determinación de los destinos de cada viaje se realizó como sigue: para cada tarjeta de registro no unitario (los que presentan varios registros al día), no reventa y que hace referencia a un viaje que sí entra en la zona troncal, se identificó la estación  $p$  donde realizó la validación, luego la estación  $q$  en la que realizó su siguiente registro y se calculó, mediante la Ecuación (1) propuesta en [16], para todas las estaciones  $k$  (excluyendo  $p$ ) el tiempo generalizado de viaje de la estación  $p$  a la estación  $q$ , haciendo escala en  $k$ .

$$Tg_{pq,k} = T\upsilon_{pk} + f \cdot \frac{d_{kq}}{V_c} \quad (1)$$

$Tg_{pq,k}$ : =tiempo generalizado de viaje

$\upsilon_{pk}$ : =tiempo de viaje en bus desde la estación  $p$  a la estación  $k$

$d_{kq}$ : =distancia de caminata entre las estaciones  $k$  y  $q$

$V_c$ : =velocidad de caminata, considerada como, 1,4 m/s

$f$ : =factor de desagrado de caminar vs. transportarse en vehículo

Así, se asigna como destino la estación  $k$ , que implica el mínimo tiempo generalizado de viaje entre las estaciones  $p$  y  $q$ , y que no viole las condiciones 1 y 2, antes descritas. El modelo a resolver, para cada par de registros en las estaciones  $p, q$  está dado por el Modelo 1.

$$\begin{aligned} \text{Min}_k Tg_{pq,k} &= T\upsilon_{pk} + f \cdot \frac{d_{kq}}{V_c} \quad (\text{Modelo 1}) \\ &\text{s. a} \end{aligned}$$

$$\begin{aligned} d_{kq} &\leq d_{max} \\ t_p + Tg_{pq,k} + T_a &\leq t_q \end{aligned}$$

$d_{max}$  := distancia máxima que se considera caminable entre estaciones, 400 m  
 $t_p$  := hora de registro en la estación  $p$   
 $t_q$  := hora de registro en la estación  $q$   
 $T_a$  := tiempo mínimo de actividad por un usuario- 15 minutos

La matriz de distancias fue obtenida por medio de Google Maps, al ubicar cada par de estaciones de interés, medir la distancia de caminata, y considerar la geografía de la ciudad, el diseño de las vías, edificaciones, parques, etc. La matriz de tiempos de viaje en bus se obtuvo a partir de las tablas de los tiempos teóricos entre cada par de estaciones consecutivas de una misma ruta, las cuales fueron suministradas por Megabús. Con esta información se generaron dos matrices de tiempos de viaje entre estaciones, una con tiempos mínimos de viaje y otra con tiempos promedios. Esto permite considerar, respectivamente, dos situaciones: la primera, en la que se hace el supuesto de que los usuarios son, de alguna forma, expertos en el servicio de transporte de Megabús, por lo cual saben escoger la ruta o combinación de rutas (en el caso de trasbordo) que minimiza su tiempo de viaje y, la segunda, en la que no son expertos, por lo que un valor a considerar, en la función objetivo, como el tiempo de viaje entre dos estaciones, en las que el usuario se puede movilizar con más de una ruta, es el promedio de los tiempos teóricos de las rutas posibles. Se considera el promedio porque no son conocidos los tiempos reales de cada viaje, dado que se desconoce la ruta del bus abordado por el usuario. Antes de implementarse el modelo se define la matriz con la que se va a trabajar. Los

resultados presentados más adelante fueron obtenidos con la matriz de tiempos de viaje promedio.

Para estimar el destino del último viaje del día de cada usuario se usó el supuesto planteado en [3] y descrito en la sección Introducción. La última transacción del día es el inicio del viaje de regreso de un usuario a su primer origen, así a cada último registro del día de una tarjeta se le asignó como destino la estación en donde se presentó su primer registro, siempre y cuando las estaciones de ambas validaciones sean diferentes. En caso de ser iguales, el destino no es estimable.

### 3.3 Pos-procesamiento de registros unitarios

Después de realizar la estimación del destino de los viajes, en cada día, se hizo un procesamiento especial para los registros unitarios (ver Sección 3.1) en días laborales. Así, para un día laboral  $D$  de la semana, se identifica cada tarjeta  $i$ , que posee un único registro en el día, luego se realiza una búsqueda de registros de viajes de esta tarjeta, en otros días laborales, en los que sí se haya estimado un destino. Se tiene en cuenta la hora del registro unitario en el día  $D$ , para asociarlo a las validaciones de otro día en donde sí se pudo estimar un destino, realizadas en un intervalo de tiempo de 30 minutos alrededor de la hora de registro en el día  $D$ . Este intervalo posee como punto medio la hora del registro unitario al que se le está estimando un destino. Como es posible que más de un día cumpla con esta condición, se asignó el destino estimado en el día donde el registro, dentro de la ventana de tiempo considerada, posee la menor diferencia temporal con el registro unitario en el día  $D$ . Este pos-procesamiento busca disminuir la pérdida de información dada por los registros unitarios

y se descartan los registros provenientes de tarjetas retornables por su incapacidad de ofrecer patrones de viajes confiables.

### 3.4 Obtención de la matriz origen-destino

Para la estimación de la MOD, en un día determinado, se define una ventana de tiempo con hora inicial,  $t_{ini}$  y hora final  $t_{fin}$ . Esta ventana de tiempo se define dependiendo de la hora del día en la cual se desea conocer la demanda del servicio, como por ejemplo horas pico o valle del día.

Puesto que los destinos de algunos viajes no son estimables, la MOD conformada por los viajes con destino estimado será una subestimación de la matriz real de viajes, por lo cual se hace necesario realizar un proceso de expansión. Para el cálculo de los factores de expansión adecuados se parte del supuesto de que el comportamiento de los viajes con destino no estimado sigue el patrón de los viajes con destino estimado. Se define entonces a  $R_i^t$  como el número total de registros detectados en la estación  $i$ , dentro de la ventana de tiempo  $t=[t_{ini}, t_{fin}]$ , y  $VC_{ij}^t$  como el número de viajes reconstruidos (viajes con destino estimado) que se originan en  $i$  dentro de la ventana de tiempo  $t$  y que poseen como destino la estación  $j$ . Se calcula el factor de expansión  $F_i^t$  para los viajes originados en  $i$  dentro de la ventana  $t$  con la Ecuación (2).

$$F_i^t = \frac{R_i^t}{\sum_{j=1}^n VC_{ij}^t} \quad (2)$$

Donde  $n$  es el número de estaciones del sistema. Luego se recalcula (se ajusta) el número subestimado de viajes  $VC_{ij}^t$ , expandiendo por el factor  $F_i^t$ , para obtener el número de viajes

estimados,  $VE_{ij}^t$  con origen en dentro del intervalo  $t$  y destino  $j$ , mediante la Ecuación (3).

$$VE_{ij}^t = F_i^t VC_{ij}^t \quad (3)$$

La MOD obtenida utilizando la Ecuación (3) entrega el número de viajes totales iniciados dentro de la ventana total de tiempo  $t$ , sin embargo, para la planificación operativa del sistema es más útil contar con los flujos de pasajeros entre estaciones cada  $r$  unidades de tiempo. Suponiendo una distribución uniforme de la demanda en la ventana de tiempo  $t$ , el número promedio de viajes estimados  $\bar{VE}_{ij}^{tr}$  que se realizan en intervalos de tiempo de tamaño  $r$ , dentro de la ventana de tiempo  $t$ , con  $r \leq t$ , es calculado dividiendo el número total de viajes estimados en la ventana  $t$  por el número total de intervalos de tamaño  $r$ , que caben en la ventana de tiempo especificada (ver Ecuación 4).

$$\bar{VE}_{ij}^{tr} = \frac{r \cdot VE_{ij}^t}{t_{fin} - t_{ini}} \quad (4)$$

### 3.5 Proceso computacional

El algoritmo para el procesamiento de la información, fue elaborado en el *software* R-Project versión 3.1.2. La base de datos utilizada posee ocho campos en los que se detalla, sobre cada registro de una tarjeta, información como la fecha, hora, el nombre de la ruta alimentadora que utilizó el usuario (si es el caso), el tipo de bus que el usuario demandó (alimentador o articulado), el código ID de la tarjeta, el nombre de la estación de la

zona troncal del registro o intercambiador del sector al que pertenece el bus alimentador de la transacción, entre otros. Para la implementación del algoritmo que realizó el proceso descrito en la metodología, se ordenó previamente la base de datos por fecha, serial y hora, para tener en cada día de la semana los registros ordenados de forma temporal para cada tarjeta (serial). Luego de definir los parámetros de depuración y de estimación de los destinos se ejecutó el proceso para cada día de la semana, como sigue: se identificaron los registros unitarios, luego en los registros no unitarios se detectaron los que provienen de reventa de pasajes, seguidamente (obviando los unitarios y de reventa) con los viajes que no entran a la zona troncal, es decir, viajes que iniciaron en una ruta alimentadora y terminaron en el mismo sector y, por último, con la identificación de los registros que implicaron el uso múltiple de tarjetas. Luego, se realizó la estimación de los destinos de cada viaje utilizando el Modelo 1. Seguidamente se llevó a cabo el pos-procesamiento de los registros unitarios de cada día y se ultimó el proceso con la estimación de la MOD en la ventana de tiempo deseada en un día determinado.

#### 4. RESULTADOS Y DISCUSIÓN

El enfoque descrito en la Sección 3 permitió describir el comportamiento de la demanda del servicio de Megabús. En la Tabla 1 se presenta la distribución de frecuencia de los registros según la estación identificada como origen, para un día laboral de una semana.

Se observa un total de 97.841 validaciones en todo el sistema, siendo el Intercambiador de Cuba el punto donde se concentra la mayor demanda del servicio (26,53%), seguido por el Intercambiador de Dosquebradas y la estación

Viajero que representan, respectivamente, el 14,54% y 7,83% de los registros totales. Las 37 estaciones del sistema que no tienen rutas alimentadoras asociadas absorben un poco más del 50% de los registros. Por otro lado, solo el 5,61% de los registros totales del día hacen referencia a viajes que fueron iniciados en un bus alimentador y que encontraron su destino en el mismo sector de su origen y que, por tanto, no entraron a la zona troncal del sistema (ver Tabla 2).

**Tabla 1.** Distribución de frecuencia del origen de los viajes en un día laboral.

Estación origen	N° de registros	% de registros
Inter. Cuba	25.953	26,53
Inter. Dqs	14.228	14,54
Viajero	76.61	7,83
37 restantes	49.999	51,10
Total	97.841	100

**Fuente:** Elaboración propia.

**Tabla 2.** Distribución de frecuencia del tipo de demanda en un día laboral.

Tipo de demanda	N° de registros	% de registros
Mismo sector	5.493	5,61
Demandan troncal	92.348	94,39
Total	97.841	100

**Fuente:** Elaboración propia.

En la Tabla 3 se presenta la distribución de frecuencias según el tipo de registro. Del total de 92.348 registros (demandantes del servicio

de la zona troncal), el 6,37% son registros que provienen de la reventa de pasajes, de los cuales no se puede obtener información para la estimación de su destino. El 18,28% son registros unitarios, tarjetas que presentaron una sola validación al día y que no se les pudo estimar destino en el pos-procesamiento de registros unitarios.

**Tabla 3.** Distribución de frecuencia del tipo de registro en la zona troncal.

Tipo de registro	N° de registros	% de registros
Tratables	57.963	62,77
Reventa	5.881	6,37
Unitario	16.879	18,28
Uso múltiple	11.625	12,59
Total	92.348	100

**Fuente:** Elaboración propia.

El 12,59% de los registros provienen de uso múltiple de la tarjeta, es decir, registros de una tarjeta ejecutados en la misma estación de manera consecutiva pero que no provienen de reventa de pasajes. El 62,77% de las validaciones con demanda en zona troncal son tratables por el Modelo 1 (ver Sección 3.2).

En la Tabla 4 se presenta la distribución de frecuencias de los diferentes tipos de registro, considerando distintos valores de  $f$  en el Modelo 1, discriminado para los registros tratables, nombrados en la Tabla 3, los viajes a los que se les logró estimar su destino, los no estimados por el incumplimiento de las restricciones del modelo y los no estimados por la falta de información (cuando el último registro del día de la tarjeta fue ejecutado en la misma estación que el primer registro de la misma tarjeta).

Cuando  $f = 1$ , se está considerando con igual peso el tiempo de viaje en bus y el

**Tabla 4.** Distribución de frecuencias del tipo registro () – Lunes.

	$f = 1.0$		$f = 1.5$		$f = 2.0$	
	N° de registros	% de registros	N° de registros	% de registros	N° de registros	% de registros
Estimado	50.948	55,170	50.943	55,164	50.933	55,153
No estimable restricción	493	0,534	498	0,539	508	0,550
No estimable falta info.	6.522	7,062	6.522	7,062	6.522	7,062
Unitario, reventa o uso múltiple	34.385	37,234	34.385	37,234	34.385	37,234
Total	92.348	100	92.348	100	92.348	100

**Fuente:** Elaboración propia.

tiempo de viaje caminando; es decir, que se está suponiendo que al usuario le da igual trasladarse a pie que en un bus articulado, cuando hace la escala. Con  $f = 1.5$ , se considera que para un usuario trasladarse caminando es 50% menos atractivo que trasladarse en bus. Con  $f = 2$ , se contempla el escenario extremo en el que los usuarios están poco dispuestos a caminar y solo lo hacen cuando la ganancia en tiempo es alta. Como se observa en la Tabla 4, en los tres escenarios descritos el porcentaje de viajes estimados se mantiene alrededor del 55,1%, el de los viajes no estimados por incumplimiento de restricciones cerca al 0,55% y el de los no estimados por falta de información se mantiene constante en 7,06%. El número de viajes estimados cuando  $f = 1.5$ , disminuye en solo cinco registros con respecto al escenario en el que caminar y viajar en bus pesan lo mismo ( $f = 1$ ).

En la Tabla 5 se muestra la matriz origen-destino estimada para una hora pico de la mañana en un día laboral, agregada por macrosectores

(en la primera columna se identifica el origen, y en la primera fila el destino). En esta se observa que los sectores en los que más se originan viajes, entre las 6:00 y las 8:00 a.m., son el Oeste y Este de la zona troncal del sistema, los cuales corresponden a los sectores Cuba y alrededores, y Dosquebradas y, respectivamente, representan un 52,3% y 33,8% del total de viajes efectuados.

Se observa también en la Tabla 5 que la cantidad de viajes a cada uno de los destinos es más homogénea comparada con lo que ocurre con el número de viajes que se originan en los diferentes macrosectores. Los sectores Oeste y Centro son las zonas hacia donde más convergen los viajes del sistema troncal, con 3.053 y 3.034, respectivamente. Adicionalmente, se nota que el mayor nivel de demanda se presenta en los viajes con origen en el Oeste y destino en el Centro, con un total de 2.501, los cuales representan el 18,7% del total de los viajes en la franja horaria y día laboral ilustrado.

**Tabla 5.** Matriz OD agregada por macrosectores, de 6:00 a 8:00 a.m. en un día laboral.

	Oeste	Norte	Centro	Sur	Este	Total
<b>Oeste</b>	966	1.239	2.501	1.014	1.256	6.976
<b>Norte</b>	542	77	13	20	22	674
<b>Centro</b>	85	51	66	13	279	495
<b>Sur</b>	344	12	31	66	222	675
<b>Este</b>	1.116	1.322	423	1.152	503	4.516
<b>Total</b>	3.053	2.701	3.034	2.264	2.283	13.336

**Fuente:** Elaboración propia.

## 5. CONCLUSIONES

El porcentaje de viajes estimados es aproximadamente igual para los tres escenarios donde se varía el valor de  $f$ , es decir, la variación del factor de penalización para el tiempo de viaje a pie no tuvo mayor efecto sobre el porcentaje de viajes estimados totales (ver Tabla 4). No obstante lo anterior, no se puede afirmar que los viajes (origen-destino) que se estiman bajo diferentes valores de  $f$  sean exactamente los mismos, ya que se pueden presentar casos en los que para un par de registros consecutivos de una misma tarjeta, en distintas estaciones, para los diferentes valores de  $f$ , se cumplan las restricciones del modelo, y el destino estimado sea distinto, dado que para cada valor de  $f$  el tiempo de viaje generalizado puede ser diferente.

La efectividad del modelo en la estimación de los destinos de viajes realizados con registros que sí poseen información válida para su tratamiento es medida por el porcentaje de viajes estimados con respecto al total de registros tratables. Así, en el escenario con un valor de  $f = 1$ , se logró estimar el destino del viaje en un 87,89% de los casos tratables (50948/57963, ver Tablas 3 y 4).

El aporte de esta investigación está en el ajuste e implementación de una metodología de estimación de matrices OD, propuesta por otros autores, para un sistema de buses, con corredor dedicado, estaciones de parada fijas para el abordaje y descenso de los pasajeros y sin localización automática. A diferencia de lo presentado en otros trabajos, las distancias consideradas en el presente enfoque no son euclidianas y se tiene en cuenta la ubicación de las estaciones y la geografía y diseño urbano de la ciudad. La restricción de tiempo que incluye

el mínimo de la actividad de un usuario en su destino es un ajuste debido a la naturaleza particular del sistema de Megabús.

Además, el enfoque propuesto tiene en cuenta la posibilidad de que los usuarios puedan ser considerados como expertos o no del sistema, por lo cual se pueden analizar ambos escenarios al escoger la matriz de tiempos mínimos de viajes entre estaciones o la matriz de tiempos promedios. La obtención de matrices OD para el sistema de transporte público de Pereira mediante el aprovechamiento de la información en las bases de datos con los registros de las tarjetas inteligentes *versus* la utilización del muestreo en campo aporta grandes ventajas.

Se tiene la posibilidad de estimar matrices OD para distintas franjas horarias del día, en cualquier día de la semana y distintas temporadas del año, a un menor costo. Además, la información fuente puede ser actualizada constantemente por la naturaleza del sistema automático de cobro de pasajes, sin costo adicional a diferencia de la información recolectada por muestreo en campo, la cual solo es útil en situaciones temporales similares a la de la muestra.

## AGRADECIMIENTOS

Los autores agradecen a la Universidad Tecnológica de Pereira y a Colciencias, proyecto con código 1110-622 38514, por el apoyo académico y económico brindado en el desarrollo de esta investigación. Un agradecimiento especial a Megabús, por suministrar la información para la estimación de la matriz OD, mapas de rutas e información sobre el funcionamiento del sistema de transporte.

## REFERENCIAS

- [1] Buneman, K. (1984). Automated and passenger-based transit performance measures. *Transportation Research Record*, 992, pp. 23-28.
- [2] Furth, P. G., Hemily, B. J., Muller, T. H. J. & Strathman, J.G. (2006). Using Archived AVL-APC Data to Improve Transit Performance and Management. *Transportation research board*. doi: 10.17226/13907
- [3] Barry, J. J., Newhouser, R., Rahbee, A. & Sayeda, S. (2002). Origin and destination estimation in New York City with automated fare system data. *Transportation Research Record*, 1817, pp. 183-187. doi: 10.3141/1817-24
- [4] Zhao, J., Rahbee, A. & Wilson, N. (2007). Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. *Computer-Aided Civil and Infrastructure Engineering*, 22(5), pp. 376-387. doi: 10.1111/j.1467-8667.2007.00494.x
- [5] Muller, T. H. J. & Furth, P. G. (2001). Trip time analyzes: key to transit service quality. *Transportation Research Record*, 1760, pp. 10-19. doi: 10.3141/1760-02
- [6] Lianfu, Z., Shuzhi, Z., Yonggang, Z. & Ziyin, Z. (2007). Study on the method of constructing bus stops OD matrix based on IC card data. En *The International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 3147-3150. doi: 10.1109/WICOM.2007.780
- [7] Trépanier, M., Tranchant, N. & Chapleau, R. (2007). Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems*, 11(1), pp.1-14. doi: 10.1080/15472450601122256
- [8] Chapleau, R. & Chu, K. K. (2007). Modeling transit travel patterns from location-stamped smart card data using a disaggregate approach. En *The 11th World Conference on Transportation Research*, pp. 1-29.
- [9] Chapleau, R., Trépanier, M. & Chu, K. K. (2008). The ultimate survey for transit planning: Complete information with smart card data and GIS. En *The 8th International Conference on International Steering Committee for Travel Survey Conferences*, Montreal, Canadá.
- [10] Zhao J. (1977). *The Planning and Analysis Implications of Automated Data Collection System: Rail Transit OD Matrix Inference and Path Choice Modeling Examples*. (Tesis doctoral). Department of Urban Studies and Planning and the Department of Civil and Environmental Engineering. Massachusetts Institute of Technology, Boston.
- [11] Hofmann, M. & O'Mahony, M. (2005). Transfer Journey Identification and Analyses form Electronic Fare Collection Data. En *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems*. Viena, Austria, pp. 34-39. doi: 10.1109/ITSC.2005.1520156
- [12] Farzin, J. (2008). Constructing an Automated Bus Origin-Destination Matrix Using Farecard and Global Positioning System Data in São Paulo,

- Brazil. *Transportation Research Record*, 2072, pp. 30-37. doi: 10.3141/2072-04
- [13] Munizaga, M., Palma, C. & Mora, P. (2010). Public Transport OD Matrix Estimation from Smart Card Payment System Data. En: *Proceedings from 12th World Conference on Transport Research*, Lisboa, Portugal, pp. 1-16.
- [14] Pelletier, M. P., Trépanier, M. & Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C*, 19, pp. 557-568. doi: 10.1016/j.trc.2010.12.003
- [15] Devillaine, F., Munizaga, M. & Trépanier, M. (2012). Detection activities of public transport users by analyzing smart card data. *Transportation Research Record*, 2276, pp. 48-55. doi: 10.3141/2276-06
- [16] Munizaga, M. A. & Palma, C. (2012). Estimation of a disaggregate multimodal public transport origin-destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C*, 24, pp. 9-18. doi: 10.1016/j.trc.2012.01.007
- [17] Giahi, R. & Moghaddam, R. (2014). A Holding Strategy to Optimize the Bus Transit Service. En: *International Journal of Industrial Engineering & Production Research*, 25(1), pp. 33-40.