



**UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA CIVIL**

**ANÁLISIS DE PATRONES ANÓMALOS EN LA SECUENCIA DE  
TRANSACCIONES DE PAGO EN EL SISTEMA DE  
TRANSPORTE PÚBLICO DE SANTIAGO**

**MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL**

**CLAUDIO ALEJANDRO NAVARRETE DELGADO**

**PROFESOR GUÍA:  
MARCELA MUNIZAGA MUÑOZ**

**MIEMBROS DE LA COMISIÓN:  
CAROLINA PALMA ALVARADO  
CRISTIAN CORTÉS CARRILLO**

**SANTIAGO DE CHILE  
SEPTIEMBRE 2012**

RESUMEN DE LA MEMORIA PARA OPTAR AL  
TÍTULO DE INGENIERO CIVIL  
POR : CLAUDIO NAVARRETE DELGADO  
FECHA: MARTES 11 DE SEPTIEMBRE DE 2012  
PROFESOR GUÍA: SRA. MARCELA MUNIZAGA M.

## **“ANÁLISIS DE PATRONES ANÓMALOS EN LA SECUENCIA DE TRANSACCIONES DE PAGO EN EL SISTEMA DE TRANSPORTE PÚBLICO DE SANTIAGO”.**

El presente trabajo de título tiene por objetivo analizar ciertos patrones anómalos en las transacciones realizadas por los usuarios de Transantiago, sistema de transporte público de Santiago, realizado mediante la tarjeta de pago integrado llamada bip!, los cuales impiden obtener la estimación de bajada en paradero de los usuarios mediante la metodología aplicada para aquello, que actualmente logra obtener un 80% de estimaciones de bajada.

El objetivo específico de este trabajo es proponer soluciones desde el punto de vista metodológico a aquellas anomalías detectadas y estudiadas que impiden una obtención completa de la estimación de bajada de los usuarios de Transantiago. Estas soluciones deben ser factibles de incorporar en el modelo, tanto del punto de vista metodológico como del punto de vista computacional.

Como parte de este trabajo se revisó tanto el modelo de asignación de bajada en paraderos, en el cual se basa este trabajo, como experiencias internacionales en el uso de información extraída a partir de sistemas de pago similares al de Santiago. Ambas revisiones permiten dar a conocer la base en la cual se trabajará en encontrar y analizar aquellas problemáticas que afecten la metodología.

Se realiza además un análisis estadístico de los problemas encontrados, con el fin de tener un estudio más acabado de las anomalías, y desprender de ellas las causas que las originan. Esto conlleva hacer un análisis a parte importante del 20% de estimaciones de bajada que no son posibles de obtener. Entre los principales problemas se destacan el multipago, los problemas de datos de los buses y la validación en zonas paga.

En estos tres casos se abordan problemas relacionados con estimación de paradero de bajada nula como estimaciones de bajada incorrectas producto de fallas en el modelo. Luego de analizar las principales anomalías y sus causas, se proponen soluciones factibles orientadas a dos ámbitos: una mejora en los datos de ingreso y cambios en las definiciones de patrones asociados a la forma de viajar de los usuarios de transporte público. Esto si bien no permitirá que el 100% de las estimaciones sean obtenidas, si logrará aumentar con respecto al porcentaje actual de estimaciones.

## Agradecimientos

En estas líneas deseo compartir mis sentimientos al finalizar este largo camino universitario que comenzó en el mes de Marzo de 2004 y que está por finalizar. Sin duda muchos quienes fueron partícipes de esta labor que realicé no los alcanzaré a nombrar, pero quiero que todo aquel que me ayudó de la más mínima forma sienta mi agradecimiento.

En primer lugar a todos aquellos compañeros y luego amigos con quienes compartí largas horas de estudio universitarias con todas sus variantes posibles: Pablo D., Felipe A., Álvaro S., Pablo H., Flavio D., Danilo P., Sebastián A., Héctor B., Felipe F., Jorge G., José R., Jaime M. y muchos más que ahora olvido. Claramente no dejaré de nombrar a mis compañeros, amigos y futuros colegas de la división de Transporte César N., Diego S., Érico G., Margarita A., Cristóbal P., Jaime O., Claudia B., Olivia A., Camila S., Sebastián A., Diego C. y todos aquellos que conocí poco pero que no dudo se su éxito a futuro.

También a todos los profesores que tuve en la Escuela a quienes les debo los conocimientos que me llevarán a ser un profesional de calidad, íntegro y con la visión de una sociedad más equitativa y compenetrada para lograr desarrollarse. Entre ellos destaco a los profesores de la división de transporte, con quienes compartí conversaciones en los almuerzos del quinto piso que me dieron a conocer sus visiones y sueños de un país mejor. Nombró acá a Sergio Jara Díaz, Francisco Martínez, Leonardo Basso, Cristián Cortés y en especial a quien me ha guiado en esta etapa final de mi carrera, Marcela Munizaga, quien ha aguantado todas mis fallas y demoras con gran paciencia y confianza en mis resultados.

Fuera de la universidad debo nombrar a todos mis amigos con quienes compartí durante estos años de universitario semana a semana las odiseas, triunfos y derrotas de mi querido Unión San Felipe. Mis agradecimientos a Rodrigo V., Luis V., Juan R., Sebastián D., Carolina J., Juan Carlos C., Felipe B., Gonzalo R., Diego Z., Pía I., Marcos M. y José C. por su constante apoyo y fuerzas para lograr el gran objetivo que se presenta acá.

Dejo para el final al pilar fundamental de todo esto: mi familia. A mis tíos y primos que siempre creyeron en mí, en especial a mi tía Rosa, mi tío Francisco y a mis primos Francisco y Raúl, quienes durante 6 años me permitieron vivir con ellos y considerarme un integrante más de su familia. Y obviamente a mis padres y a mi hermano, de los quienes siento un tremendo orgullo por el esfuerzo que realizaron para que yo llegara a ser quien soy. De ellos estaré siempre agradecidos de la formación que me dieron y espero algún día que mis hijos puedan ser criados de la forma que lo hicieron conmigo.

Si me olvidé de alguien lo siento mucho, pero estoy seguro que sabe que es parte de esto. Siento que el mayor agradecimiento que les puedo dar es seguir siendo como me conocieron y seguir adelante en mi forma de ser.

## Índice.

<b>1. Introducción.....</b>	<b>1</b>
1.1. Aspectos Generales.....	1
1.2. Estructura del trabajo de título.....	2
1.3. Antecedentes del sistema de transporte público de Santiago.....	3
<b>2. Revisión bibliográfica y del modelo de estimación de bajada.....</b>	<b>6</b>
2.1. Revisión de antecedentes bibliográficos.....	6
2.2. Explicación del modelo de estimación de bajada.....	7
2.2.1. Buses.....	11
2.2.2. Metro.....	12
2.2.3. Zonas Paga.....	13
2.2.4. Implementación computacional y extracción de datos.....	14
2.3. Análisis de datos.....	15
2.3.1. Base de datos de estimación de bajada.....	15
2.3.2. Asignación de bajada.....	19
<b>3. Problemas del modelo en la estimación de paradero de bajada.....</b>	<b>21</b>
3.1. Introducción.....	21
3.2. Punto de bajada no caminable.....	21
3.2.1. Situación Actual.....	21
3.2.2. Análisis estadístico.....	22
3.2.3. Propuesta de solución al problema.....	27
3.3. Primera transacción del día.....	28
3.3.1. Situación actual.....	28
3.3.2. Análisis estadístico.....	28
3.3.3. Propuesta de solución.....	31
3.4. Validaciones únicas del día.....	32
3.4.1. Situación Actual.....	32
3.4.2. Análisis estadístico.....	33
3.4.3. Propuesta de Solución.....	36
3.5. Estimación de paradero de subida y bajada cercano.....	37
3.5.1. Situación actual.....	37
3.5.2. Análisis estadístico.....	39
3.5.3. Propuesta de solución.....	41

<b>4. Multipago.</b> .....	<b>42</b>
4.1. Definición y antecedentes. ....	42
4.2. Viajes Grupales. ....	43
4.2.1. Validaciones Grupales. ....	43
4.2.2. Etapas y Viajes Grupales. ....	51
4.2.3. Casos particulares de viajes grupales. ....	54
4.2.4. Propuesta de solución. ....	57
4.3. Multipago en sitios distintos. ....	59
4.3.1. Antecedentes y situación actual. ....	59
4.3.2. Transacciones masivas. ....	60
4.3.3. Análisis estadístico. ....	62
4.3.4. Propuesta de solución. ....	68
<b>5. Zonas Paga.</b> .....	<b>70</b>
5.1. Introducción. ....	70
5.2. Análisis de Datos. ....	70
5.3. Análisis de errores. ....	73
5.3.1. Análisis exploratorio. ....	73
5.3.2. Análisis global de validaciones en zonas paga. ....	75
5.3.3. Servicios mal asignados. ....	76
5.3.4. Emisiones GPS perdidas. ....	84
<b>6. Conclusiones.</b> .....	<b>87</b>
6.1. Comentarios y conclusiones. ....	87
6.2. Líneas futuras de investigación. ....	90
<b>7. Referencias Bibliográficas.</b> .....	<b>92</b>

## Índice de Tablas.

Tabla 2-1: Transacciones bip! por día de la semana de estudio. ....	16
Tabla 2-2: Validaciones por periodo del día. ....	16
Tabla 2-3: Número de buses, servicios y validaciones por Unidad de Negocio. ....	18
Tabla 2-4: Bajadas estimadas y no estimadas por tipo de transporte. ....	19
Tabla 2-5: Principales problemas de estimación de bajada. ....	20
Tabla 3-1: Bajada no caminable por sitio de transacción. ....	22
Tabla 3-2: Tamaño ejes estructurantes sector centro Santiago. ....	25
Tabla 3-3: Transacciones posteriores a bajadas no caminables por zona. ....	27
Tabla 3-4: Transacciones realizadas por media hora entre 00:00 y 04:59 horas. ....	31
Tabla 3-5: Transacciones únicas del día realizadas entre Lunes y Jueves. Periodo siguiente transacción. ....	35

Tabla 3-6: Transacciones únicas del día realizadas en día Viernes. Periodo siguiente transacción.....	35
Tabla 3-7: Transacciones únicas del día realizadas en día Sábado. Período siguiente transacción.....	36
Tabla 3-8: Transacciones con misma subida y bajada por lugar de validación.....	40
Tabla 4-1: Porcentaje acumulado de ocurrencia de viaje grupal por tiempo entre transacciones. ....	46
Tabla 4-2: Transacciones y validaciones grupales.....	47
Tabla 4-3: Total viajes grupales por tarjeta ID en semana de estudio.....	52
Tabla 4-4: Viajes grupales por día.....	52
Tabla 4-5: Comparación de número de etapas entre muestra de viajes grupales y total de viajes de semana de estudio.....	53
Tabla 4-6: Cantidad de etapas grupales.....	54
Tabla 5-1: Principales Zonas Paga en cantidad de validaciones. ....	71
Tabla 5-2: Servicios asignados por paradero y validaciones.....	72
Tabla 5-3: Parámetros estadísticos de comparación entre muestra y total de tarjetas ID. ....	73
Tabla 5-4: Cantidad de validaciones por tarjeta ID, comparación entre muestra y total para semana de estudio.....	74
Tabla 5-5: Estimación de bajada para muestra y total de tarjetas ID. ....	74
Tabla 5-6: Fallas en estimación de bajada para muestra de tarjetas ID.....	75
Tabla 5-7: Zonas paga con mayor cantidad de asignaciones de bajadas nulas. ....	77
Tabla 5-8: Sitios y transacciones para servicio G04I en EIM La Cisterna. ....	78
Tabla 5-9: Sitios y transacciones para servicio 415el en Metro Pudahuel. ....	80
Tabla 5-10: Servicios mal asignados a zonas paga y transacciones sin bajada. ....	81
Tabla 5-11: Recorridos con mayor proporción de errores en posicionamiento. ....	85
Tabla 5-12: Utilización de túneles por servicios Transantiago.....	85
Tabla 6-1: Resumen de cambios en bajada producto de cambios al modelo. ....	90

## Índice de Gráficos.

Gráfico 2-1: Distribución de transacciones según lugar de validación. ....	17
Gráfico 2-2: Distribución de transacciones según categoría de operador.....	17
Gráfico 3-1: Distribución de transacciones para día laboral.....	29
Gráfico 3-2: Distribución de transacciones para día Sábado.....	29
Gráfico 3-3: Distribución de transacciones para día Domingo.....	29
Gráfico 3-4: Transacciones únicas en el día por periodo del día para día laboral.....	33
Gráfico 3-5: Transacciones únicas en el día por periodo del día para Sábado. ....	34
Gráfico 3-6: Porcentaje de estimación de paradero de subida y bajada cercano por comuna.....	40
Gráfico 3-7: Porcentaje de estimación de paradero de subida y bajada similar por media hora. ....	41
Gráfico 4-1: Cantidad de validaciones grupales por distancia temporal, caso Metro. ....	45
Gráfico 4-2: Cantidad de validaciones grupales por distancia temporal, caso Bus.....	46
Gráfico 4-3: Distribución de grupos de viaje por tamaño de grupo y sitio de transacción. ....	47
Gráfico 4-4: Validaciones grupales por operador. ....	48
Gráfico 4-5: Incidencia de validaciones grupales por comuna. ....	49

Gráfico 4-6: Incidencia de validaciones grupales por estación de Metro.....	49
Gráfico 4-7: Incidencia de validaciones grupales por tipo de día. ....	50
Gráfico 4-8: Incidencia de validaciones grupales por periodo. ....	50
Gráfico 4-9: Cantidad de transacciones masivas. ....	61
Gráfico 4-10: Tiempo entre transacciones consecutivas. Caso Zona Paga-Bus.....	63
Gráfico 4-11: Transacciones consecutivas con distancia temporal menor a 250 segundos. Caso Zona Paga-Bus.....	63
Gráfico 4-12: Distancia entre transacciones consecutivas. Caso Bus – Bus.....	64
Gráfico 4-13: Transacciones consecutivas con distancia menor o igual a 200 metros. Caso Bus – Bus.....	65
Gráfico 4-14: Distancia entre transacciones consecutivas. Caso Bus – Metro.....	66
Gráfico 4-15: Transacciones consecutivas con distancia menor a 350 metros. Caso Bus – Metro. ....	66
Gráfico 4-16: Distancia entre transacciones consecutivas. Caso Metro - Bus. ....	67
Gráfico 4-17: Transacciones consecutivas con distancia menor a 350 metros. Caso Metro – Bus.....	68
Gráfico 5-1: Validaciones en zona paga por periodo del día.....	72

## Índice de Figuras.

Figura 2-1: Modelo de estimación de bajada.....	9
Figura 2-2: Diagrama general de tratamiento de la información para la estimación del punto de bajada.....	10
Figura 3-1: Ejemplo de viaje con etapa de viaje ausente. ....	38
Figura 3-2: Ejemplo de viaje con etapa ausente. Posible traslado realizado.....	38
Figura 3-3: Ejemplo etapa ausente. Caso primera validación del día. ....	39
Figura 4-1: Esquema de identificación de transacción grupal. ....	58
Figura 4-2: Método de detección y eliminación tarjetas masivas. ....	62
Figura 5-1: Método de detección de servicio mal asignado a sitio de zona paga.....	83

## Índice de Imágenes.

Imagen 3-1: Bajadas no caminables por zona DIRTP.....	23
Imagen 3-2: Plano Regulador comuna de Santiago (Extracto). ....	24
Imagen 3-3: Plan Regulador de Providencia.....	25
Imagen 3-4: Zonas DIRTP sector Santiago Centro. ....	26
Imagen 3-5: Zonas DIRTP eje A. Providencia – 11 de Septiembre.....	26
Imagen 4-1: Caso 1 de viaje grupal.....	55
Imagen 4-2: Caso 2 de viaje grupal.....	56
Imagen 4-3: Caso 3 de viaje grupal.....	57
Imagen 5-1: Paraderos con la mayor cantidad de transacciones posteriores a hechas en sitios RM-0720 y RM-0721.....	78
Imagen 5-2: Recorrido G22R ((M) La Cisterna – Villa España).....	79
Imagen 5-3: Paraderos con la mayor cantidad de transacciones posteriores a hechas en sitios RM-0641, RM-0642, RM-0643 y RM-0758.....	80
Imagen 5-4: Servicio 414el sector Oriente ((M) Pudahuel – Los Trapenses).....	81

# 1. Introducción

## 1.1. Aspectos Generales.

A partir de febrero de 2007, con la implementación del nuevo sistema de transporte público urbano conocido como Transantiago, se implementaron una serie de reformas al modo de viajar de los usuarios. Una de estas reformas, relacionada a la forma de pago del servicio, tiene que ver con el fin del pago al conductor con efectivo y la implementación de un sistema de pago mediante tarjeta inteligente de prepago (Smart Card) que se conoce como bip!, la que funciona mediante contacto con un validador que descuenta la tarifa del saldo.

Este tipo de pago mediante Smart Card tiene antecedentes de implementación y uso en varias ciudades del mundo a partir de la década de los 90. En cada una de ellas el nivel de penetración de uso es variable dependiendo de las demás opciones existentes para pago. Para el caso de Santiago, el nivel de penetración en buses es total, pues es el único medio de pago permitido actualmente, mientras que en Metro es el principal sistema de pago por sobre el realizado con boleto, con lo que se obtiene que la tasa de pago mediante bip! es del 97% en el total del sistema (Beltrán et al., 2011).

Todas estas transacciones realizadas mediante tarjetas inteligentes en el mundo han generado grandes bases de datos con información valiosa, que permite a los investigadores y a los planificadores de transporte en los distintos países donde se ha implementado esta tecnología, conocer el comportamiento de los usuarios a la hora de utilizar el transporte público en diferentes niveles dependiendo de la calidad y cantidad de información existente.

Para el caso de Santiago, se ha generado un modelo que permite obtener, entre otros datos, estimaciones de subida y bajada de los usuarios del sistema al nivel de paradero y/o estaciones que se ha convertido en información muy valiosa para entender el comportamiento de los usuarios de transporte público que utilizan como medio de pago la tarjeta bip! (Munizaga y Palma, 2012).

Sin embargo, este modelo no es capaz de reproducir fielmente la forma de viaje de los usuarios, debido a una serie de comportamientos que realizan los pasajeros producto de la forma en que hacen su pago, y en la información disponible que si bien es de gran calidad presenta aún ciertas incertezas que deben ser disminuidas o corregidas para lograr obtener un modelo que estime de forma precisa los patrones de viaje de los pasajeros que utilizan el transporte público de Santiago.

En este trabajo de título se realiza un análisis de estos patrones anormales de viaje que si bien son permitidos, impiden estimar datos como la bajada del modo utilizado para ciertos casos. La estructura de este trabajo se da a conocer en el capítulo 1.2 de este informe que se presenta a continuación.



## **1.2. Estructura del trabajo de título.**

El presente trabajo se estructura en base a seis capítulos. Esta estructura se clasifica en tres partes: dos capítulos introductorios al trabajo en que se dan a conocer las fuentes del trabajo a realizar, tres capítulos de contenido en que se presenta el trabajo realizado con los datos, y finalmente un último capítulo con las conclusiones.

El primer capítulo es introductorio al trabajo y en él se presentan aspectos generales del trabajo realizado, información del presente informe y antecedentes generales del transporte público de Santiago, que es el sistema a estudiar.

En el capítulo 2 se da a conocer un análisis a la bibliografía existente relacionada al trabajo de estimación de bajada a partir de datos pasivos extraídos de sistemas de pago mediante tarjeta inteligente. El principal enfoque de este capítulo es al modelo de asignación de bajada en paraderos que actualmente se utiliza para el sistema de transporte público de Santiago (Munizaga y Palma, 2012), tanto del punto de vista del funcionamiento del modelo como de los datos generados. Esto tiene como objetivo tener conocimiento del modelo a trabajar y los pasos a seguir en los siguientes capítulos para generar soluciones.

Luego, el capítulo 3 presenta cuatro problemas relacionados con la forma en que los usuarios utilizan el sistema de transporte público y cómo se comportan en su manera de viajar que el modelo no es capaz de absorber y por ende se generan problemas en la estimación de bajada. Para cada uno de estos problemas se realiza un análisis estadístico y en base a él se proponen soluciones al modelo para capturar el comportamiento y corregir la estimación.

El cuarto capítulo trabaja el concepto de multipago que corresponde al pago consecutivo en un corto período de tiempo. Para esto se trabaja en dos tipos de multipago: aquel realizado en el mismo lugar y el realizado en sitios de pago distintos. Para cada uno de estos casos, mediante un análisis estadístico se definen los parámetros que identifican su ocurrencia y se realiza un conteo para obtener el grado de ocurrencia. Finalmente se proponen soluciones que van en la línea de identificar cada caso para así poder evaluar su comportamiento de viaje.

El capítulo 5 trata el tema de las zonas paga. En él se analizan las principales causas de error en la estimación de bajada en viajes generados en cada zona paga a partir de análisis exploratorios. Con esto se identifican las zonas paga con menor estimación de bajada, se busca el problema que las afecta presentando la forma en que se identifica la causa y se propone la solución para mitigar el problema existente en el modelo. Además se analizan recorridos que ocupan los túneles de la vialidad existente en Santiago y se estudia su comportamiento desde el punto de vista de las emisiones GPS.

Finalmente, en el capítulo 6 se dan a conocer las conclusiones de este trabajo, resumiendo cada uno de los patrones estudiados y dando a conocer la cantidad de validaciones que se logran recalcular. Además se presentan las futuras líneas de

investigación en la mejora del modelo de asignación de bajada, tanto desde el punto de vista de mejora del modelo como de la evaluación del mismo mediante datos exógenos.

### **1.3. Antecedentes del sistema de transporte público de Santiago.**

El sistema de transporte público de Santiago basa su funcionamiento en una estructura de servicios alimentadores y troncales. Los servicios alimentadores cumplen la función de transportar pasajeros dentro de una zona que está definida por una o más comunas geográficamente cercanas entre sí. Para ello se ha dividido Santiago en nueve zonas, cada una con un servicio alimentador manejado por una empresa que utiliza buses para el transporte de pasajeros. Los servicios troncales en cambio son los encargados de unir estas zonas alimentadoras mediante largos recorridos por las principales avenidas de la ciudad. Para aquello se ha asignado esta tarea a seis empresas que cumplen la labor de movilizar a los usuarios en estas grandes distancias. Una de estas empresas a cargo de los recorridos troncales es el Metro, sistema de tren urbano que mediante sus cinco líneas integradas entre sí es parte del sistema. Las otras cinco empresas troncales realizan estos servicios mediante buses que utilizan la vialidad de superficie disponible en Santiago para su funcionamiento.

Santiago es la capital de Chile, en que viven cerca de 6 millones de habitantes en un área de 640 km<sup>2</sup>. La distribución de sus habitantes no es homogénea, diferenciándose claramente los sectores de la ciudad por estrato social. La forma de la ciudad es circular, en la cual la gran mayoría de los viajes en la mañana ocurren desde la periferia al sector céntrico de la ciudad y en la tarde de forma contraria. Para el año 2010 en un día laboral ocurren más de 17 millones de viajes, de los cuales 10 millones son realizados en medios motorizados, distribuidos casi equitativamente entre transporte público y privado (Sectra, 2012).

Para satisfacer esta demanda de viajes en transporte público, el sistema cuenta con cerca de 6.000 buses que realizan alrededor de 300 recorridos tanto en servicios troncales como alimentadores. En lo referido a infraestructura, existen más de 10.000 paraderos de buses y casi 200 km. de vialidad de uso preferencial para buses. En lo referente a Metro, este posee en sus 103 km. de extensión 108 estaciones distribuidas en cinco líneas. Además existen 150 estaciones de parada de buses conocidas como Zona Paga, generalmente ubicadas en puntos de alto flujo tanto de usuarios (sobre 500 pax/h) como de buses (sobre 50 buses/h), donde la validación ocurre en el sitio de parada y no arriba del bus, con el fin de permitir un aumento en los tiempos de subida de parte de los usuarios y por ende una disminución en los tiempos de ciclo del servicio, y además con el fin de combatir la evasión en estos puntos de atochamiento de usuarios (Transantiago, 2008).

El sistema de pago es realizado mediante una tarjeta inteligente conocida como tarjeta bip!, de las que se han emitido cerca de 11 millones. Estas tarjetas funcionan como tarjeta de prepago y a medida que se va utilizando en los servicios integrados en Transantiago (buses troncales, alimentadores y Metro) se va descontando la tarifa, que para la semana a analizar correspondía a \$450 en buses y Metro, salvo en horas punta donde Metro cobraba \$510 y en horas baja donde cobraba \$430. Esta forma de pago se realiza haciendo contacto entre la tarjeta y un validador donde se realiza el pago. Este validador puede estar dentro del bus junto a la puerta de acceso o en una estación

dispuesta para la espera de buses (Zona Paga) o de trenes (Metro). Como se trata de un sistema integrado de pago, permite realizar combinaciones entre los distintos modos que permiten pago mediante tarjeta bip! accediendo a un beneficio de gratuidad. Este beneficio se basa en que se descuenta el valor del total del viaje en la primera transacción realizada y permite realizar hasta dos transbordos que al validar para acceder no descuentan saldo de la tarjeta en un plazo de 120 minutos desde la primera transacción<sup>1</sup>. Existen ciertas restricciones a esta gratuidad, como es el caso de que dentro del margen temporal se hagan dos o más viajes en Metro, que se valide en los mismos servicios o que se realicen dos transacciones en un periodo corto de tiempo. Para estos comportamientos el sistema no acepta la gratuidad del transbordo y cobra la tarifa completa del viaje para todas las validaciones realizadas. Los únicos transbordos que generan un descuento de saldo para la semana de estudio son los realizados al hacer transbordo de bus a Metro en horas punta (mañana y tarde) el cual genera un cobro adicional de \$60 y el pasar en hora baja de Metro a bus que implicaba un cobro adicional de \$20.

Para los estudiantes de educación media y universitaria se dispone de una tarjeta especial que además de ser de uso personal e individualizada con sus datos, permite acceder a una tarifa rebajada que corresponde a aproximadamente el 35% de la tarifa normal. Esta tarifa para la semana de estudio correspondía a \$150. A diferencia de la tarjeta bip! normal, en esta no existen descuentos en transbordos asociados al uso de Metro en ningún horario.

Existen además otras formas de movilización mediante otros modos de transporte público tales como colectivos, taxis, buses interurbanos y trenes de cercanía (Metrotrén) pero que actualmente no están integradas en el sistema de pago de Transantiago, por lo que no es posible obtener el comportamiento de estos viajes a partir del modelo.

Uno de los grandes problemas que ha presentado este sistema de pago es el de la evasión, que corresponde a omitir la validación para acceder al modo de transporte integrado a utilizar. Esto genera además de la pérdida económica para el sistema, una pérdida en la información de transacciones realizadas que implican no observar todos los viajes realizados por el usuario en los distintos modos. Para esto y en base a este comportamiento se definen dos tipos de evasión: la asociada al usuario que no paga nunca para acceder al sistema, y la asociada a la omisión de pago de alguno de los modos utilizados durante el viaje. Al primer tipo de evasor es casi imposible seguirle su rutina de viajes, ya que en muchos casos ni siquiera posee tarjeta de pago, en cambio el evasor de algunas etapas generalmente posee tarjeta pero por causas variadas no valida en alguno de los modos abordados durante su viaje. Este evasor de algunas etapas puede ser estudiado por el modelo, pero la ausencia de algunas transacciones perjudica la información procesada y genera la pérdida de traslados observados.

Todas las transacciones realizadas son almacenadas en una gran base de datos, en la que para cada transacción se tiene información acerca del número identificador *id* y tipo de tarjeta, el sitio o vehículo donde se realiza el pago, el operador

---

<sup>1</sup> Más información en <http://www.transportedesantiago.cl/TARJETABIP/TARIFAS/index.htm>

al que fue realizado el pago, el instante en que se realiza la validación y el monto de dinero descontado de la transacción. Cada semana las cerca de 3 millones de tarjetas bip! en uso realizan alrededor de 35 millones de transacciones en los diferentes sitios habilitados para aquello. La posición exacta donde se realizó la transacción es conocida para el caso de que éstas hayan sido realizadas en estaciones de Metro o en zonas paga, por tratarse de lugares fijos y conocidos, a diferencia de las realizadas en el interior del bus, donde no se conoce a priori el lugar de la transacción por tratarse de un validador en movimiento.

Además existe otra gran base de datos generada a partir de las emisiones realizadas por los GPS instalados en los buses. Estas emisiones son generadas cada 30 segundos, y en ellas se emite información referida a posición geográfica del bus (latitud y longitud), hora, fecha y velocidad instantánea. Cada emisión además está asociada a un bus que se identifica por su placa patente y por el operador al que sirve. En una semana de viajes se recogen alrededor de 80 millones de emisiones GPS. Además se posee la información georreferenciada de los recorridos de buses, estaciones de Metro, paradas de buses y Zonas Paga, datos valiosos que complementan lo anterior.

Finalmente, el proceso de asignación de buses a servicios lo realiza la Coordinación de Transantiago mediante la definición de tres puntos, localizados al inicio, en la mitad y al final de la ruta de cada servicio, cada uno con un área de influencia de 500 metros cuadrados centrados en el punto. Luego mediante las posiciones GPS de los buses se identifica si el bus pasó por las tres áreas de influencia asociadas a un servicio, y en caso que esto ocurra se le asigna a ese bus el servicio definido por los tres puntos. Este servicio se define desde la última emisión GPS realizada en el área de influencia del punto inicial del servicio hasta la primera emisión GPS realizada dentro del área de influencia del punto final del servicio. Esto si bien genera una pérdida de información de recorrido tanto al comienzo como al final de la ruta, es actualmente una buena fuente de información de servicio bien evaluada por quienes usan esta información.

## **2. Revisión bibliográfica y del modelo de estimación de bajada**

### **2.1. Revisión de antecedentes bibliográficos.**

La revisión bibliográfica de este trabajo está enfocada en la forma como diversos investigadores y académicos han ocupado la información extraída a partir de los sistemas de pago mediante Smart Card en diferentes ciudades del mundo para obtener información sobre la movilidad de los habitantes de las ciudades en transporte público. Este enfoque busca generar una base que permita conocer cómo se realizan los procesos de cálculo de estimación de puntos de bajada y comparar la forma de obtener la información.

Un resumen de la forma en que es utilizada esta información se presenta en Pelletier et al. (2010), en el que se dan a conocer las ventajas y desventajas del uso del sistema de Smart Card en el pago de los sistemas de transporte público, y luego los usos que se le da a la información obtenida a partir de las transacciones realizadas en el sistema desde el punto de vista estratégico, táctico y operacional.

En relación a las ventajas del sistema de pago mediante tarjeta inteligente, el principal beneficio corresponde a la posibilidad de recolectar información asociada a los viajes a partir de las transacciones realizadas, las cuales complementadas con información obtenida mediante otros medios como encuestas o conteos permiten mejorar la calidad de la información de viajes de una ciudad. Además desde el punto de vista operacional el pago mediante tarjeta genera una serie de beneficios como integración tarifaria y disminución de tiempos de viaje debido a la disminución en el tiempo asociado al pago del pasaje, que deja de ser manual.

Las desventajas de este método de pago pasan por la capacidad de una óptima implementación que influya en el uso correcto del sistema de pago de parte de los usuarios, además de necesitar una gran inversión económica para implementar esta tecnología. Y desde el punto de vista de la recolección de información, esta presenta una serie de vacíos que a la hora de utilizarse para generar información de viajes se deben hacer suposiciones que no siempre corresponden a la realidad.

Una de las grandes causas que origina un uso de pago mediante tarjeta distinto al que se planifica es la forma en que el usuario utiliza este medio. En McDonald (2000) se da a conocer el caso de la implementación de París, en el que se realizaron una serie de pruebas piloto desde 1993 con los empleados del Metro de París. Ya en 1997 se pudo hacer ensayos con 1.000 pasajeros, con lo que se obtuvieron resultados positivos y se pudo dar el inicio a la implementación total a partir del año 2000. En estos ensayos, se pudieron observar aquellos comportamientos que podrían inducir a un mal uso de la tarjeta ya sea por iniciativa del usuario (fraude) o por la complejidad del sistema.

Referido al punto de hacer frente a la complejidad de un sistema de pago nuevo para los usuarios, en Deakin and Kim (2001) se presenta cómo las tecnologías asociadas al transporte se han ido implementando cada vez con mayor fuerza, con el fin

de buscar beneficios que le den cada vez mejor competitividad al sistema y a su vez una menor complejidad. Sin embargo, el proceso de implementación suele ser lento debido a las múltiples inversiones que se deben realizar, las cuales deben generar formas de uso que impliquen un beneficio para el usuario, ya sea desde el punto de vista de mejora en la calidad del viaje como en lo económico.

Una de las diferencias de los sistemas de pago en otras ciudades del mundo con el implementado en Santiago es la obligatoriedad de su uso. Esta imposición de un sistema de pago único para todos los modos integrados es un caso que no se presenta en los demás sistemas, donde el pago mediante Smart Card es parcial y se complementa con otras modalidades de pago, como se da a conocer en Mora (2010).

Si bien este trabajo está enfocado en el estudio de patrones anómalos en la sucesión de pagos mediante tarjeta inteligente, la bibliografía no presenta antecedentes al respecto. El estudio realizado por Chapleau and Chu (2007) identifica datos incorrectos o sospechosos de serlo y mediante la identificación de las fuentes de error y patrones de comportamientos históricos de los usuarios se corrigen de manera desagregada con el fin de utilizar estos datos actualizados para la generación de modelos que realicen asignaciones de paradas de subida y de recorridos.

## **2.2. Explicación del modelo de estimación de bajada.**

Previo a explicar el modelo que permite obtener información sobre la estimación de bajada de los usuarios del sistema de transporte público, es necesario dar a conocer una serie de conceptos que dan forma a la teoría detrás de este modelo y a sus resultados.

El primer concepto es el de viaje, que se define como el movimiento realizado desde un origen a un destino (Ortúzar and Willumsen, 2011). El viaje tiene como resultado un propósito, que está asociado a la actividad a realizar en el destino del viaje, lo que da a entender que el movimiento realizado para acceder a cada propósito es un viaje distinto. Además se define el origen como la posición donde comienza el viaje, y destino como la posición donde termina el viaje.

El viaje está compuesto por etapas, que son las diferentes formas de movilización para acceder al destino. Para eso se utilizan medios de transporte, los cuales pueden ser motorizados o de tracción humana o animal. Para este modelo los medios de transporte a estudiar serán aquellos del sistema de transporte público integrados al sistema de pago integrado de Transantiago, que son el modo bus y el modo Metro. Serán los traslados realizados en estos modos a los que el modelo les asociará un punto de subida y de bajada, que corresponderán al comienzo y fin de la etapa respectivamente.

El hecho de que sólo se puedan observar las etapas de los viajes realizados en transporte público con pago integrado impide obtener información sobre aquellos traslados realizados en otros medios de transporte, tales como colectivo, taxi, caminata, bicicleta, etc. Esta falta de información implica ausencia de etapas en la secuencia de viajes, lo que afecta la estimación de subida y bajada de las etapas, además de impedir

en el caso de que estos modos integrados sean utilizados en la primera o última etapa el origen o el destino del viaje respectivamente.

El movimiento que debe realizar un usuario para acceder desde un medio de transporte a otro es definido como transbordo. Este movimiento se hace a pie, y no corresponde a una etapa misma del viaje, pues sólo tiene como fin utilizar otro vehículo para continuar el viaje.

Existen dos formas en las que el modelo diferencia el fin de una etapa del fin de un viaje. La primera corresponde a la ventana horaria existente entre la hora de la bajada estimada de una transacción y la de subida asociada a la siguiente transacción. Actualmente esa ventana horaria está fijada en 30 minutos, lo que quiere decir que ventanas iguales o menores a 30 minutos corresponden a transbordos, y las mayores a 30 minutos al fin del viaje y por ende a la realización de una actividad en el destino del viaje. Este corte sin embargo genera fallas en el caso de actividades muy cortas o en transbordos muy largos, lo que implica perder características de la forma de viajar de los usuarios.

Y la segunda forma de diferenciar el fin de una etapa del fin de un viaje corresponde a analizar los abordajes consecutivos. Cuando se realizan dos transacciones consecutivas en Metro o en buses con el mismo servicio, se considera que entre las transacciones hay un destino correspondiente al fin de un viaje independiente del tiempo entre bajada y subida siguiente, pues se asume que es poco probable que alguien se baje y se vuelva a subir a la red de Metro durante un viaje. Y para el caso de dos subidas consecutivas al mismo recorrido de bus también se asume poco probable que entre ellas haya un transbordo, por lo que se asume que entre las dos transacciones se realizó alguna actividad.

Para que el modelo pueda obtener la información relacionada con los viajes en transporte público, se hace necesario utilizar la información de tres grandes bases de datos generadas a partir de los datos recolectados por Transantiago:

- Transacciones realizadas por los usuarios en cada uno de los validadores con la tarjeta de pago automático (bip!).
- Posición geográfica de cada bus, ya sea en movimiento o detenido, generada a partir de la emisión del sistema de GPS integrado en el vehículo.
- Información geográfica de la red de sistema de transporte público (rutas de viaje, paraderos, estaciones de Metro, etc.)

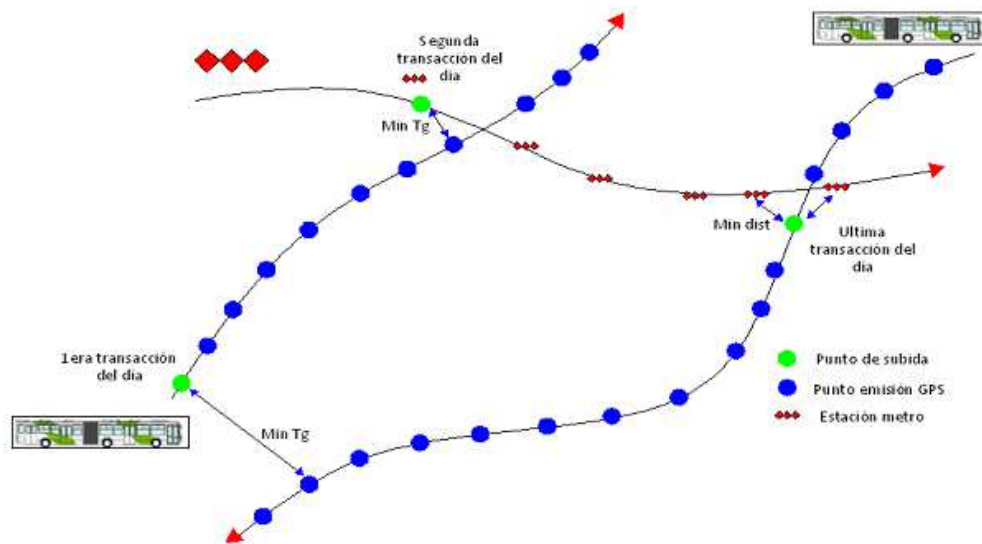
La combinación de información de estas tres tablas mediante el modelo genera la información de origen y destino del viaje, subida y bajada de cada etapa, actividad o transbordo, etc., cuya forma de obtener se dará a conocer a continuación. Sin embargo no siempre se puede obtener toda la información del viaje. Las causas son múltiples, y algunas de ellas se darán a conocer en este trabajo.

La idea principal que permite obtener la parada de bajada de la etapa del viaje se aprecia en la figura 2-1. En ella se ve que la metodología para estimar el punto de bajada de una etapa dentro del viaje de cada usuario se resume en dos supuestos básicos:

- Una etapa del viaje termina con la transacción siguiente de ésta, suponiendo que el punto de bajada de la etapa actual es cercano al punto de subida de la etapa siguiente, y además el tiempo de bajada es anterior a la hora de la próxima subida.
- La estimación del punto de bajada de la última etapa del último viaje del día considera que ocurre donde ocurrió la primera transacción del día, o sea, el usuario vuelve al origen del primer viaje del día en el último viaje del día.

Sin embargo, actualmente existe en el modelo un proceso de recalcular de paradero de bajada para el caso de la última validación del día. Si no es posible encontrar bajada para esa transacción utilizando como cierre la primera del día, se utiliza como cierre la primera del día siguiente.

Figura 2-1: Modelo de estimación de bajada.

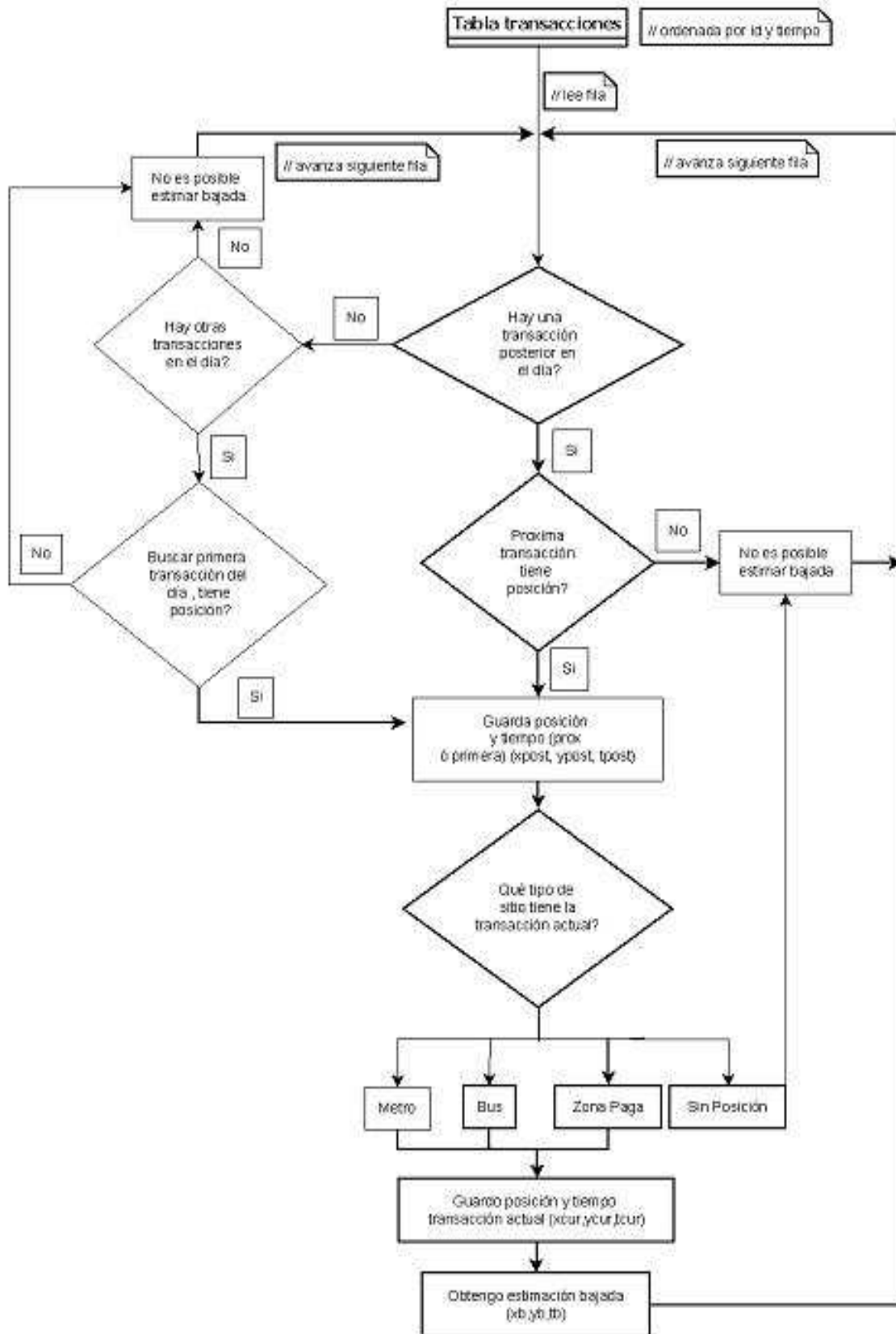


Fuente: Mora (2010).

Como existen tres tipos de lugares donde realizar la transacción, se subdividen en casos: Bus, Metro y Zona Paga. Estos si bien están integrados entre sí a la hora de realizar el pago, presentan información diferente a la hora de ser recolectada y debe ser tratada a su vez de manera diferente. Pero el tratamiento general es similar, el que se presenta en el diagrama de la figura 2-2. Este algoritmo además de reconocer los puntos de bajada, puede obtener aquellos comportamientos que imposibilitan estimar la bajada, por lo que presenta una excelente base de trabajo para esta memoria.



Figura 2-2: Diagrama general de tratamiento de la información para la estimación del punto de bajada.



Fuente: Mora (2010).

### 2.2.1. Buses.

Para obtener las bajadas de las etapas cuya transacción es realizada en el interior de un bus, el modelo debe ser capaz de obtener esta información a partir de las múltiples posiciones de cada bus y de las distintas configuraciones de ruta existentes.

Previo a la búsqueda del punto de bajada se debe tener en consideración un problema que puede afectar la estimación. Esto es suponer que sólo con un criterio de distancia física es suficiente para estimar la bajada de la etapa anterior basado en la ubicación de la siguiente transacción. Un ejemplo de este caso ocurre cuando un recorrido de bus ocupa una misma ruta (calle) tanto para su recorrido de ida como para el regreso, y un pasajero toma este recorrido para viajar algunas cuadras para luego retornar ocupando el mismo servicio pero en sentido contrario, tomándolo en la vereda del frente. Si se ocupara el criterio de distancia mínima, el modelo supondría que el usuario realizó el viaje de ida y vuelta, y que se bajó frente al sitio de subida pues es el más cercano en distancia al sitio de la siguiente transacción. Esto generaría entre otros problemas una sobreestimación del tiempo de viaje.

Es por esto que se utiliza para obtener la bajada de la etapa del viaje a partir de la trayectoria del bus conocida por la base de datos de emisiones de GPS, el tiempo generalizado como una función a ser minimizada (Munizaga y Palma, 2012).

El modelo busca un punto de bajada estimado, definido por dos coordenadas de posición y una temporal ( $x_b$ ,  $y_b$ ,  $t_b$ ) ubicado en la trayectoria del bus, definida por sus emisiones GPS que minimice el tiempo generalizado con la siguiente subida, también definida por coordenadas espaciales y temporal. Este tiempo generalizado asociado a la posición GPS  $i$  se define en la ecuación 2-1 como la suma del tiempo asociado al punto  $i$  ( $t_i$ ) y una variable que representa el tiempo estimado de caminata entre la bajada y la siguiente subida multiplicado por un factor de disgusto de caminata.

El tiempo de caminata es estimado a partir de la distancia euclidiana entre la posición  $i$  y la posición de la siguiente transacción ( $d_{i-post}$ ) dividido por una velocidad media de caminata. El factor de disgusto por caminata, llamado factor de penalización es obtenido a partir de modelos de elección discreta como la desutilidad el tiempo de caminata sobre la del tiempo de viaje en vehículo. Con esto se puede obtener el tiempo de viaje asociado con la posición  $i$  como  $t_i$  menos el tiempo de la transacción inicial.

$$Tg_i = t_i + f_c \cdot \frac{d_{i-post}}{v_c} \quad \text{Ecuación 2-1}$$

Donde:

- $i$ : Punto GPS con coordenadas ( $x_i$ ,  $y_i$ )
- $t_i$ : Tiempo asociado al punto ( $x_i$ ,  $y_i$ )
- $d_{i-post}$ : Distancia euclidiana entre ( $x_i$ ,  $y_i$ ) y ( $x_{post}$ ,  $y_{post}$ )
- $v_c$ : Velocidad promedio de caminata
- $f_c$ : Factor de penalización

Luego la búsqueda se realiza sobre todas las posiciones  $i$  de la trayectoria del bus, que se encuentran dentro de una distancia caminable  $d$  a la posición de la siguiente transacción. Así, el problema de optimización queda como:

$$\begin{aligned} & \min_i Tg_i \\ \text{s. a.} \quad & d_{i-post} < d \\ & t_i < t_{post} \end{aligned} \qquad \text{Ecuación 2-2}$$

Finalmente, aquella posición  $i$  que tenga  $\min Tg_i$  se denomina como  $Tg$  ( $x_{tg}$ ,  $y_{tg}$ ,  $t_{tg}$ ).

Se implementa además una ventana de tiempo en la búsqueda de emisiones GPS en la trayectoria del bus a partir del momento de la transacción. Este parámetro del modelo que es fijado depende de las características de cada una de las rutas troncales y alimentadoras. Esta ventana tiene como fin hacer más eficiente la búsqueda de información dada la gran base de datos existente.

Otro parámetro que es requerido en el modelo es el referido a la distancia de caminata  $d$ . Esta distancia es la que está dispuesta a caminar la persona entre la bajada y la posición de la siguiente subida. Esta distancia podría depender de variables como tipo de persona, tipo de ciudad, clima, seguridad del tramo y otros factores. Esta distancia actualmente está fijada en 1.000 metros. Esto se obtuvo basado en un análisis exploratorio a 17 tarjetas bip! hecho por Mora (2010) en donde se muestra que el 9,4% de las etapas de viaje de la muestra terminan en un punto que dista a más de 1.000 metros de la siguiente transacción. Actualmente el modelo posee un filtro de distancia que omite aquellas bajadas cuya distancia entre la bajada estimada y la transacción posterior sea superior a los 1.000 metros.

Para concluir, si la ecuación 2-2 no posee solución, se asume que existe un viaje o etapa perdida que haya sido probablemente realizada usando otra tarjeta bip! o en otro medio de transporte no integrado. Así, ese viaje es catalogado como “sin estimación de punto de bajada”.

### 2.2.2. Metro

El método para obtener los puntos de subida y bajada de las etapas de viaje realizadas en Metro es similar al caso del bus, pero con diferencias que facilitan su implementación. En primer lugar se tiene que los puntos de subida y bajada son las estaciones de Metro. La subida se conoce al observar la estación en que se realizó la transacción, y la bajada se obtiene buscando la estación más cercana a la siguiente transacción realizada dentro de un rango menor a 1.000 metros, que corresponde a la distancia caminable, donde el criterio para definir esta distancia límite caminable es similar al caso del bus. Y del mismo modo, en caso que no se encuentre una estación de Metro dentro de este rango de 1.000 metros, se asume que hay una etapa de viaje perdida y el punto de bajada no puede ser estimado.

Cuando se encuentra la estación de Metro de bajada que cumple el criterio de distancia caminable, queda por definir la hora en que ocurrió esa bajada. Para obtener

este dato, se debe conocer la ruta que utilizó el usuario para llegar al punto de bajada dentro de la red de metro. Para esto se ocupa el algoritmo de Dijkstra (1959) de rutas mínimas, que utiliza como costos los tiempos de viaje entre estaciones y tiempos de detención en cada estación. Finalmente se tiene que a partir del punto de subida conocido previamente se obtiene la ruta utilizada por el usuario hasta la bajada a partir del criterio de costo mínimo de viaje.

Una vez que se obtiene la ruta, y conocida la hora de subida (dato presente en la transacción) en base al plan operativo de Metro se obtienen los tiempos de espera en andén y de viaje en tren. Además a partir de mediciones en terreno se obtienen los tiempos de transbordo entre estaciones (en caso que ocurra transbordo entre líneas) y el tiempo de caminata entre el torniquete y el andén, tanto para el acceso como para la salida. Con esto se puede obtener el tiempo total de viaje sumando estos componentes, y por ende, la hora de bajada de la etapa de Metro.

### 2.2.3. Zonas Paga.

Para obtener la estimación de bajada en aquellas etapas de viaje cuya transacción de subida fue realizada en una Zona Paga el proceso se complejiza debido a que no se conoce a priori el recorrido utilizado por el usuario, pues éste puede utilizar cualquiera de los que utilizan como parada la zona paga.

Para obtener el servicio que utilizó el usuario para llegar a su bajada se debe hacer un tratamiento previo a los datos, en el que una vez definida la zona paga utilizada como subida y la hora a partir de la transacción realizada en la misma, se asigna un servicio del conjunto de recorridos que utilizan esta zona paga.

Este proceso de asignación de servicio en zona paga se basa en un enfoque de rutas mínimas sin restricción de capacidad propuesto por Chriqui y Robillard (1975). Este enfoque considera que el pasajero elige su alternativa de viaje dependiendo de dos factores: el tiempo de espera (esperanza) y los tiempos de viaje observados históricamente. Se usa el criterio sin restricción de capacidad debido a que no se conoce la cantidad de usuarios que van en el bus ni en el paradero, por lo que no se puede asumir congestión en alguno de los dos o en ambos.

Se tiene que:

- Esperanza del tiempo de espera:  $\left\{ \frac{k}{f_1}, \frac{k}{f_2}, \dots, \frac{k}{f_n} \right\}$

Donde  $k$ : constante de espera= $\frac{1}{2}$  (se asume llegada uniforme de pasajeros al paradero y llegada de buses a intervalos constantes)  
 $f_i$ : frecuencia del servicio  $i$ , estimada a partir de la velocidad promedio observada de la ruta, en horario establecido para zona paga.

- Tiempo en vehículo:  $\{tv_1, tv_2, \dots, tv_n\}$

Donde  $tv_i$ : tiempo de viaje en ruta  $i$ . Este tiempo considera tanto el tiempo en ruta como el tiempo de caminata (se estima a partir del mínimo tiempo generalizado, que se obtiene de  $\min_p t_p + 2,5 \frac{d_{p-b}}{v_c}$ ).

$p$ : paradero en ruta  $i$

$v_c$ : velocidad de caminata (1,5 m/s)

2,5: factor de peso, que evidencia desagrado de caminar en relación al tiempo de viaje en vehículo (Mora, 2010).

Con estos parámetros se obtiene el tiempo total esperado de viaje (TVE). Para obtener el conjunto de líneas comunes ( $L_a$ ) se minimiza el tiempo total esperado de viaje, lo cual se representa en la ecuación 2-3.

$$\text{Sea } X = \{x_1, x_2, \dots, x_n\} \quad x_i = \begin{cases} 1 & \text{si } l_i \in L_a \\ 0 & \text{en caso contrario} \end{cases} \quad \text{Ecuación 2-3}$$

$$\min_x TVE = \left\{ \frac{k + \sum_i tv_i \cdot f_i \cdot x_i}{\sum_i f_i \cdot x_i} \right\}$$

Finalmente, una vez obtenido el conjunto de líneas comunes, el modelo asigna el primer servicio que pase posterior a la transacción realizada en la zona paga que pertenezca a este conjunto, y que además cumpla con el requisito de que el servicio tenga paradero dentro de un radio de 1.000 metros con respecto a la ubicación de la siguiente transacción.

Al igual que en los casos de bus y Metro, si no se encuentra paradero de bajada dentro del rango de 1.000 metros, se asume que hay una etapa de viaje perdida y por ende no se estima paradero de bajada, dejando la transacción con "Bajada No Caminable".

#### 2.2.4. Implementación computacional y extracción de datos.

Para procesar todas estas bases de datos se ocupa en primer lugar el software PostgreSQL 8.3 como almacenador de datos, el que utiliza para mejorar la velocidad de búsqueda una serie de registros especiales. El código con el que se procesa esta información se construye en C++, utilizando para ello libpqxx como interfaz con la base de datos. Todos estos resultados son visualizados en Google Earth utilizando un API para formato KML. La forma en que se realiza este procesamiento es la explicada previamente, en la que se calculan los diferentes parámetros a utilizar.

Para el trabajo a realizar en esta memoria se tiene que los datos generados a partir de las estimaciones de bajada presentan una cantidad importante de información que debe ser extraída de la manera más eficiente y exacta posible, con el fin de que los datos generados sean fiables tanto en contenido como en lógica. Si bien estos datos serán mostrados en un capítulo posterior, en esta etapa se da a conocer el software utilizado para tal recolección y su modo de funcionamiento.

Una primera fuente de recolección es el software para lectura de base de datos llamado PostgreSQL (similar al del lectura de las bases de datos en el procesamiento

de éstas), que permite la ejecución de simples rutinas de consultas acerca de los parámetros que sean requeridos por el usuario, datos que son recolectados y visualizados a partir de la herramienta pgAdmin3, la que entrega la información en formato compatible con Excel, para poder mostrarla de una manera más amigable a quienes deseen verla.

La segunda forma de recolección de información es el lenguaje de programación Python, el cual presenta una estructura similar a los demás lenguajes más conocidos (Java, C++, etc.) pero que tiene cierta ventaja sobre la lectura de archivos planos de texto (txt, csv), formato en el que se encuentra la base de datos. Con este lenguaje es posible obtener ciertos conteos que se ven más complejos hacerlos en SQL.

Finalmente, el uso de Microsoft Excel es necesario principalmente para el proceso final de la información, que incluye conteos finales de los datos a considerar y sirve además como plataforma para la creación de gráficos que ayuden a explicar de mejor manera los comportamientos que se desean dar a conocer.

## **2.3. Análisis de datos.**

### **2.3.1. Base de datos de estimación de bajada.**

La tabla de asignación de bajadas en la que se trabaja para obtener información respecto a viajes de los usuarios posee una serie de características del punto de vista estadístico que vale la pena visualizar previamente antes de observar los patrones anómalos.

Si bien para cada validación se tienen más de 120 campos con información, entre los que se presentan fecha y hora de la misma, sitio de validación, paradero de subida y de bajada estimada y errores de modelación, sólo se revisará aquella que es interesante de evaluar y la que dará datos relevantes para este estudio.

La base de datos contiene la información de validaciones realizadas mediante el uso de tarjeta bip! entre los días Lunes 31 de Mayo y Domingo 6 de Junio de 2010. Esto tanto para validaciones en estaciones de Metro, al interior de los buses que pertenecen al sistema de transporte público Transantiago y a las Zonas Pagas habilitadas por Transantiago en ciertos puntos de Santiago de alta congestión en paraderos.

Una primera información a conocer tiene relación con la cantidad de validaciones por día de la semana y periodo del día. Esto se da a conocer en las tablas 2-1 y 2-2, donde se puede ver que el 83,3% de las validaciones ocurren el día laboral (Lunes a Viernes), un 10,5% ocurren en día Sábado y un 6,2% son hechas el día Domingo. En promedio cada día laboral ocurren aproximadamente 6.350.000 transacciones, con un alza hacia los días Lunes y Viernes. Además, al analizar los periodos del día en día laboral predominan las validaciones en los períodos punta mañana y tarde. Pero el fin de semana cambia este comportamiento, pues el periodo que predomina es el de Mediodía, debido a la diferencia de actividades de los usuarios de Transporte Público en día laboral y en fin de semana.

Tabla 2-1: Transacciones bip! por día de la semana de estudio.

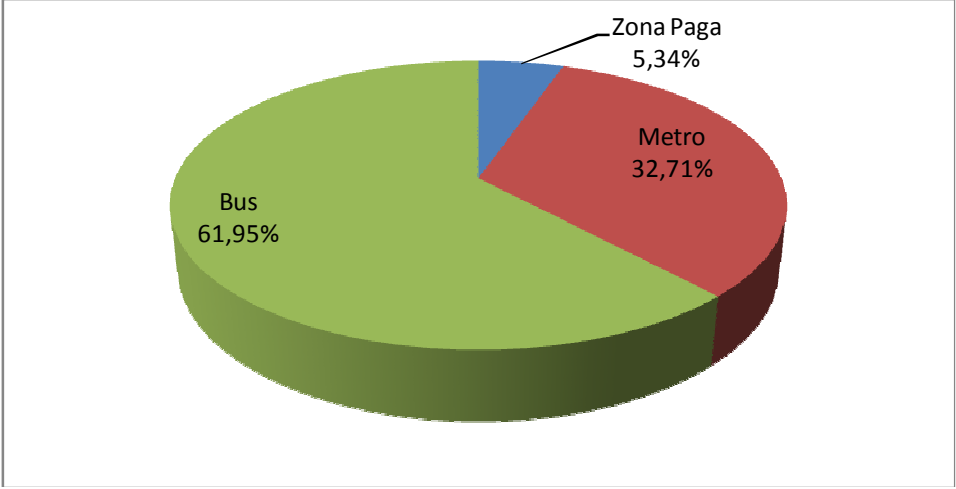
Fecha	Día	Tipo día	Transacciones
31-05-2010	Lunes	Laboral	6.484.377
01-06-2010	Martes	Laboral	6.303.518
02-06-2010	Miércoles	Laboral	6.225.939
03-06-2010	Jueves	Laboral	6.259.172
04-06-2010	Viernes	Laboral	6.478.077
05-06-2010	Sábado	Sábado	3.990.844
06-06-2010	Domingo	Domingo	2.377.022
<b>Total Semana</b>			<b>38.118.949</b>

Tabla 2-2: Validaciones por periodo del día.

Tipo Día	Periodo	Horario	Duración [horas]	Validaciones	Total Día
Laboral	Pre Nocturno	00:00 - 00:59	1:00	66.490	31.751.083
	Nocturno	01:00 - 05:29	4:30	75.279	
	Transición Nocturno	05:30 - 06:29	1:00	695.598	
	Punta Mañana	06:30 - 08:29	2:00	5.459.789	
	Transición Punta Mañana	08:30 - 09:29	1:00	2.219.867	
	Fuera de Punta Mañana	09:30 - 12:29	3:00	4.341.402	
	Punta Mediodía	12:30 - 13:59	1:30	2.441.289	
	Fuera de Punta Tarde	14:00 - 17:29	3:30	5.796.958	
	Punta Tarde	17:30 - 20:29	3:00	7.453.501	
	Transición Punta Tarde	20:30 - 21:29	1:00	1.438.461	
	Fuera de Punta Nocturno	21:30 - 22:59	1:30	1.445.791	
	Pre Nocturno	23:00 - 23:59	1:00	316.658	
Sábado	Pre Nocturno Sábado	00:00 - 00:59	1:00	30.322	3.990.844
	Nocturno Sábado	01:00 - 05:29	4:30	33.875	
	Transición Sábado Mañana	05:30 - 06:29	1:00	46.678	
	Punta Mañana Sábado	06:30 - 10:59	4:30	952.568	
	Mañana Sábado	11:00 - 13:29	2:30	622.111	
	Punta Mediodía Sábado	13:30 - 17:29	4:00	1.110.402	
	Tarde Sábado	17:30 - 20:29	3:00	708.001	
	Transición Sábado Nocturno	20:30 - 22:59	2:30	422.798	
	Pre Nocturno Sábado	23:00 - 23:59	1:00	64.089	
Domingo	Pre Nocturno Domingo	00:00 - 00:59	1:00	26.838	2.377.022
	Nocturno Domingo	01:00 - 05:29	4:30	28.476	
	Transición Domingo Mañana	05:30 - 09:29	4:00	229.115	
	Mañana Domingo	09:30 - 13:29	4:00	543.063	
	Mediodía Domingo	13:30 - 17:29	4:00	660.077	
	Tarde Domingo	17:30 - 20:59	3:30	629.291	
	Transición Domingo Nocturno	21:00 - 22:59	2:00	232.929	
	Pre Nocturno Domingo	23:00 - 23:59	1:00	27.233	

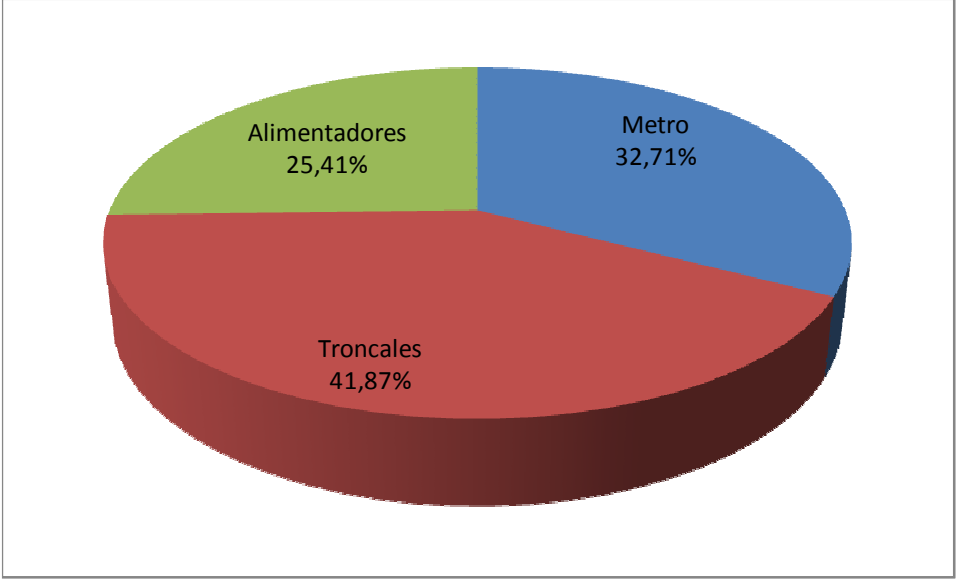
Como se mencionó en un comienzo, las validaciones son hechas en tres lugares distintos dependiendo de la necesidad de viaje del usuario: Buses, Metro y Zonas Pagas. Así, 23.614.102 validaciones son hechas en bus, 12.470.020 son hechas en la red de Metro y 2.034.917 son realizadas en Zonas Pagas habilitadas. En el gráfico 2-1 se muestra la distribución de transacciones según lugar de validación, donde se puede observar el predominio del uso del bus por parte de los usuarios de transporte público. Aún a pesar de las extensiones de líneas de Metro que se han hecho en los últimos años, el uso del modo bus sigue acaparando casi dos tercios del total de viajes (sumando Bus y Zona Paga).

Gráfico 2-1: Distribución de transacciones según lugar de validación.



Otro dato importante a conocer es la distribución de validaciones según la categoría del operador. Se tiene que 15.961.496 validaciones se hicieron en servicios troncales, 9.687.433 validaciones se hicieron en servicios alimentadores y las ya conocidas 12.470.020 validaciones hechas en Metro. En el gráfico 2-2 se aprecia la distribución porcentual de estas validaciones.

Gráfico 2-2: Distribución de transacciones según categoría de operador.





Para la fecha en estudio, Transantiago contaba con cinco operadores que cubrían los recorridos troncales, nueve que cubrían las zonas alimentadoras y el Metro que funciona como troncal. Además, algunos recorridos del troncal 3 fueron pasados a manos de operadores del Troncal 1 y 4 producto de compensaciones debido a la futura extensión del Metro a la comuna de Maipú que les iba a bajar de manera importante el número de usuarios a estos operadores.

Las validaciones hechas en cada uno de los operadores de Transantiago para la primera semana de Junio de 2010 y la cantidad de buses que recibieron esas validaciones son los mostrados en la tabla 2-3, en la que se da a conocer que cuatro grandes operadores son los que reciben la mayor cantidad de transacciones durante la semana, correlacionado con el número de buses que tienen en servicio. El troncal 3 en un comienzo fue asignado a Buses Gran Santiago S.A., sin embargo el día 28 de Noviembre de 2009 debió ceder parte de sus recorridos a la empresa Buses Vule S.A. producto de la finalización del contrato de operación de Buses Gran Santiago. Es por eso que para esa fecha los recorridos 300 presentan dos unidades de negocios distintas.<sup>2</sup>

Tabla 2-3: Número de buses, servicios y validaciones por Unidad de Negocio.

Concesionario		Nº Buses	Servicios	Validaciones
Nombre	Unidad de Negocio			
Inversiones Alsacia S.A.	T1	676	25	2.850.031
Subus Chile S.A.	T2	1061	38	3.976.004
Buses Gran Santiago S.A.	T3	341	10	945.638
Express de Santiago Uno S.A.	T4	1049	30	4.015.053
Buses Metropolitana S.A.	T5	594	18	2.820.389
Buses Vule S.A.	T6	432	16	1.354.381
Buses Gran Santiago S.A.	Zona B	420	31	1.257.949
Redbus Urbano S.A.	Zona C	317	26	1.085.747
STP Santiago S.A.	Zona D	264	20	885.675
Unión del Transporte S.A.	Zona E	238	20	997.085
STP Santiago S.A.	Zona F	402	28	1.486.685
Trans Araucarias S.A.	Zona G	251	21	818.002
Trans Araucarias S.A.	Zona H	166	18	542.775
Comercial Nuevo Milenio S.A.	Zona I	437	25	1.549.259
Comercial Nuevo Milenio S.A.	Zona J	298	28	1.064.256

Fuente: Elaboración Propia.

<sup>2</sup> "Buses Vule inicia operación del Troncal 3" 27/11/2009  
<http://www.transantiago.gob.cl/showNovedadesByIdAction.do?ID=1370>

### 2.3.2. Asignación de bajada

El fin de este trabajo es obtener una mejora en la estimación de bajada en paraderos que genera el modelo. Es por esto que se muestra a continuación un análisis de estimaciones de bajada, tanto acertadas como fallidas.

En primer lugar, se da a conocer la cantidad total de validaciones que presentan estimación de bajada y aquellas que no lo tienen, que se muestra en la tabla 2-4. En este análisis a nivel de etapa de viaje (transacción) se aprecia que aquellas transacciones realizadas en Metro son las que permiten un mayor éxito en estimar bajada de etapa de acuerdo al modelo de asignación de paradero. Luego sigue de cerca el modo Bus y algo más alejado pero no menos exitoso los viajes que tienen como origen las Zonas Pagas.

Tabla 2-4: Bajadas estimadas y no estimadas por tipo de transporte.

Modo	Bajadas				Total Transacciones
	Estimadas	%	No Estimadas	%	
Bus	19.204.477	81,33	4.409.535	18,67	23.614.012
Metro	10.758.587	86,28	1.711.433	13,72	12.470.020
Zona Paga	1.557.911	76,56	477.006	23,44	2.034.917
<b>Total</b>	<b>31.520.975</b>	<b>82,69</b>	<b>6.597.974</b>	<b>17,31</b>	<b>38.118.949</b>

Una de las causas de estos porcentajes corresponde a la confiabilidad en los datos y en la información disponible para integrar en el modelo. El hecho de que la información necesaria para realizar las estimaciones de bajada para Metro sea menor genera mayores facilidades a la hora de estimar la estación de abandono. Además, al tratarse de una estructura rígida no presenta los vaivenes de cambios en los recorridos ni problemas en las frecuencias que pueden tener los buses en superficie. En el polo opuesto, las zonas pagas presentan el problema de la confiabilidad en la asignación de servicio para cada transacción, lo que genera incertidumbre a la hora de una correcta estimación.

En la tabla 2-5 se dan a conocer los principales problemas que impiden lograr estimaciones de bajada. En esta tabla se pueden apreciar cinco grandes problemas que se deben abordar. Lo primero a considerar previo a un análisis estadístico de los problemas es que éstos al ser sumados dan un valor superior al total de transacciones sin estimación de bajada. Esto ocurre porque sucede que una transacción tiene dos o más causas de estimación de bajada nula. Sin embargo, la tabla da a conocer la magnitud de cada uno de los problemas. La explicación de cada uno de estos problemas es la siguiente:

- Bajada con destino no caminable: aquella donde la distancia entre la bajada de la transacción y la subida de la siguiente transacción es mayor a 1.000 metros.

- Transacción única en el día: se definen aquellas que son realizadas de manera única en el día, y no se les estima bajada pues no existe una siguiente transacción de subida.
- Etapa muy corta: etapa cuya distancia entre el paradero de bajada de la transacción y de subida de la siguiente transacción es menor o igual a 50 metros.
- Datos malos de buses: error en la emisión GPS del bus que implica una pérdida de información del recorrido y por ende de la etapa de viaje.
- Multipago: transacciones consecutivas realizadas en el mismo sitio de pago con una distancia temporal menor o igual a tres minutos.

Un análisis mayor a cada uno de estos problemas será realizado a medida que se vaya identificando el problema dentro de cada capítulo.

Tabla 2-5: Principales problemas de estimación de bajada.

Motivo No Estimación	Transacciones por modo				% sobre total de Transacciones
	Bus	Metro	Zona Paga	Total	
Bajada con destino no caminable	1.932.000	624.513	230.172	2.786.685	7,31%
Transacción única en el día	1.327.596	665.904	62.015	2.055.515	5,39%
Etapa muy corta	577.520	242.473	102.077	922.070	2,42%
Datos malos de buses	414.841	0	87.204	502.045	1,32%
Multipago	288.298	193.617	17.515	499.430	1,31%

Como se dijo en capítulos anteriores, este trabajo está enfocado en tres áreas a revisar: los problemas en el modelo asociados a comportamiento de usuarios, el análisis de validaciones del tipo Multipago y las estimaciones de bajada realizadas a partir del pago en Zona Paga, por lo que el hecho de conocer por qué está fallando el modelo permite discernir las problemáticas presentes en los tópicos a tratar mencionados anteriormente.

### **3. Problemas del modelo en la estimación de paradero de bajada.**

#### **3.1. Introducción.**

El modelo de asignación de paradero de bajada presenta en la actualidad problemas producto de múltiples factores que afectan la calidad de la estimación, principalmente debido a comportamientos de los usuarios que originalmente no se tenían contemplados o debido a simplificaciones que no siempre ocurren en la realidad.

Estos errores se traducen en dos problemas que se generan al estimar bajada: que se estime una bajada errónea o que no se logre estimar bajada. El primer caso es complejo de analizar, ya que determinar si una bajada está mal estimada no es algo que se pueda deducir, pues no se puede descartar ningún comportamiento de parte de los usuarios que a vista del modelador se pueda considerar como “anormal”. El segundo caso es más simple de ver en la base de datos, pero la complejidad radica en entender la causa de esa estimación nula. Si bien la base de datos entrega en sus columnas finales las causas de error, éstas son muy generales y no permiten obtener mayor información referida al verdadero problema que impide generar la estimación del paradero de bajada.

En este capítulo se abordará una serie de problemáticas en el modelo de asignación de bajada que se pueden corregir basándose en ciertos comportamientos observables de la conducta de viaje de los usuarios, ocupando la experiencia de viaje personal como la observada en el diario vivir.

El sistema de trabajo y análisis de problemáticas será similar para cada una de las encontradas. En primer lugar se da a conocer la situación actual del patrón anómalo, tomando en base lo que hace el modelo y las causas de esta definición de trabajo. A continuación se muestran análisis estadísticos y conteos del problema desde el punto de vista de las transacciones con problemas, desagregando el conteo en diferentes parámetros dependiendo de la necesidad de estudio. Finalmente se presenta la propuesta de solución a la problemática o en caso de que no exista se deja planteado el problema para un análisis más detallado a futuro.

#### **3.2. Punto de bajada no caminable.**

##### **3.2.1. Situación Actual.**

Originalmente, el filtro de distancia conocido como bajada no caminable que corresponde a aquellas transacciones cuya distancia entre la bajada obtenida mediante la función de tiempo generalizado y la siguiente transacción es mayor a 1.000 metros (ver Ecuación 2-2) tiene como objetivo eliminar la incertidumbre sobre la forma de acceso a la siguiente transacción, la que podría haberse hecho en otros modos no registrados en el sistema (bicicleta, colectivo, taxi o auto particular) o bien debido a una posible evasión en el pago del pasaje lo que deja sin registro una de las etapas del viaje.

Sin embargo, realizando un análisis un poco más detallado del comportamiento de los usuarios, el hecho de que una persona camine más de 1.000 metros entre la posible bajada y la siguiente transacción no es tan poco probable. Una causa de que esto ocurra, y que no se trate de alguna de las hipótesis de incertidumbre descritas anteriormente, es debido a que en ciertos sectores de Santiago, principalmente los grandes barrios comerciales (Centro de Santiago, eje Providencia-11 de Septiembre, etc.) la gente tiende a caminar bastante con fines de consumo o recreación (comprar, vitrinear, pasear, etc.) lo que muchas veces los puede alejar de la bajada.

Un primer análisis tiene base en definir qué sectores de Santiago concentran principalmente estas bajadas no caminables, con el fin de estudiar las actividades que se realizan en la zona y analizar si existe alguna correlación entre lo explicado recientemente y las cantidades de bajadas nulas.

Para esto se utilizan como referencia todas aquellas transacciones realizadas con posterioridad a la transacción considerada como no caminable. Si bien lo lógico hubiera sido usar las bajadas estimadas de las transacciones no caminables, este proceso necesitaba hacer correr el modelo por completo para la semana de estudio, lo que retrasaba el presente trabajo. Por lo tanto se decide utilizar la transacción posterior, teniendo precaución en generar propuestas que estén acordes a la información existente.

### 3.2.2. Análisis estadístico.

Para la base de datos utilizada en este trabajo, de las 38.118.949 transacciones realizadas en el sistema durante la semana de estudio, son 2.786.685 transacciones que no se les estima bajada debido a este problema, las que representan el 7,31% del total de validaciones, lo que lo convierte en el principal problema de estimación de bajada en relación a aquellas que no son estimadas.

El análisis de transacciones con bajada no caminable por lugar de transacción se da a conocer en la tabla 3-1. En ella se aprecia que este comportamiento proporcionalmente se presenta principalmente en transacciones hechas en zonas paga y buses.

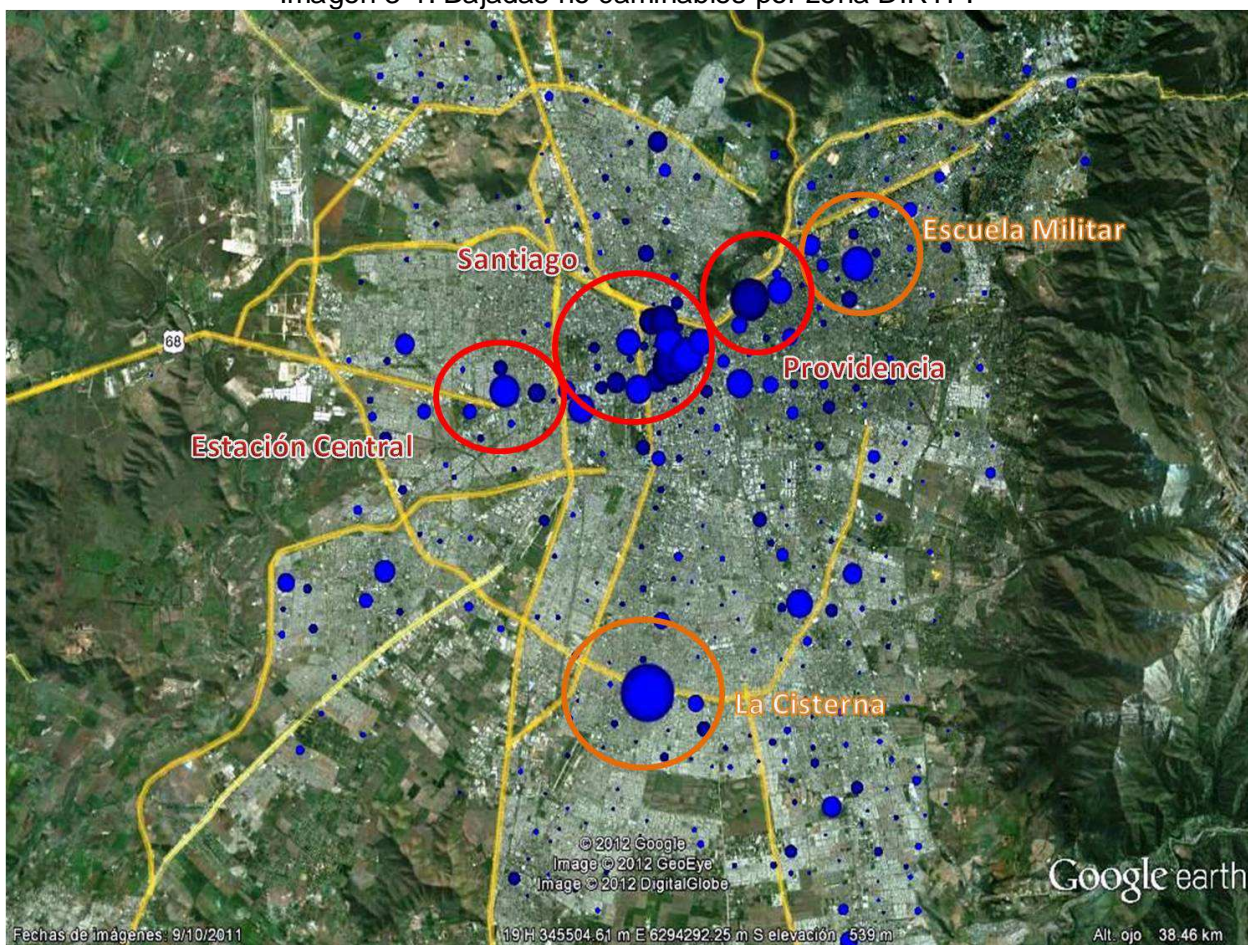
Tabla 3-1: Bajada no caminable por sitio de transacción.

Lugar de Transacción	Transacciones		Porcentaje [%]
	Sin estimación de bajada	Total Semana	
<b>Zona Paga</b>	230.172	2.034.917	11,31
<b>Bus</b>	1.932.000	23.614.102	8,18
<b>Metro</b>	624.513	12.470.020	5,01
<b>Total</b>	2.786.685	38.119.039	7,31

La zonificación ocupada corresponde a la realizada para el modelo DIRTP (DICTUC, 2003) que Transantiago ha ido actualizando, para completar un total de 847 zonas dentro del Gran Santiago, siendo la zonificación disponible más desagregada dentro de la base de datos existente. Esta desagregación permite obtener información más detallada de aquellos transbordos o actividades que supongan traslados mayores a

los 1.000 metros entre bajada y siguiente subida. En la imagen 3-1 se aprecia la distribución espacial de este problema en base a los datos obtenidos de las transacciones hechas en la semana de estudio. Se observa que este problema afecta principalmente a lugares como Santiago Centro, Providencia, Estación Central y puntos de transbordo masivos como La Cisterna y Escuela Militar.

Imagen 3-1: Bajadas no caminables por zona DIRTP.



El problema que se presenta al hacer el supuesto que estas zonas recién nombradas son las más afectadas por el problema de bajada no caminable es que a su vez son las zonas que presentan la mayor cantidad de transacciones en la semana de estudio, lo que podría suponer que los resultados son producto de una correlación entre el total de validaciones y los errores existentes. Sin embargo, no se debe descartar este problema por lo mencionado, ya que se presenta dentro de la lógica de comportamiento de los usuarios el hecho de realizar grandes caminatas en estos sectores para los propósitos recientemente mencionados (actividades de ocio en general).

El estudio del caso de transacciones con punto de bajada no caminable se aplicará a dos zonas comerciales ubicadas en el eje Alameda-Providencia:

- Casco histórico de Santiago (Santiago Centro).
- Eje Providencia – 11 de Septiembre.

Se definen estas zonas principalmente porque se posee mayor información de ellas referido a su ubicación y límites. Esto con el fin de obtener los tamaños y los posibles desplazamientos máximos y promedio que puedan realizar las personas a lo largo de estas zonas.

Los límites a aplicar para cada una de las zonas se basan en lo declarado por cada una de las comunas en sus respectivos planes reguladores como zona de servicio o comercial. Estas zonas, que coinciden con aquellas con mayores cantidades de bajadas no caminables, se dan a conocer en las imágenes 3-2 y 3-3 correspondientes a extractos de los planos reguladores de las comunas de Santiago y Providencia.

Imagen 3-2: Plano Regulador comuna de Santiago (Extracto).



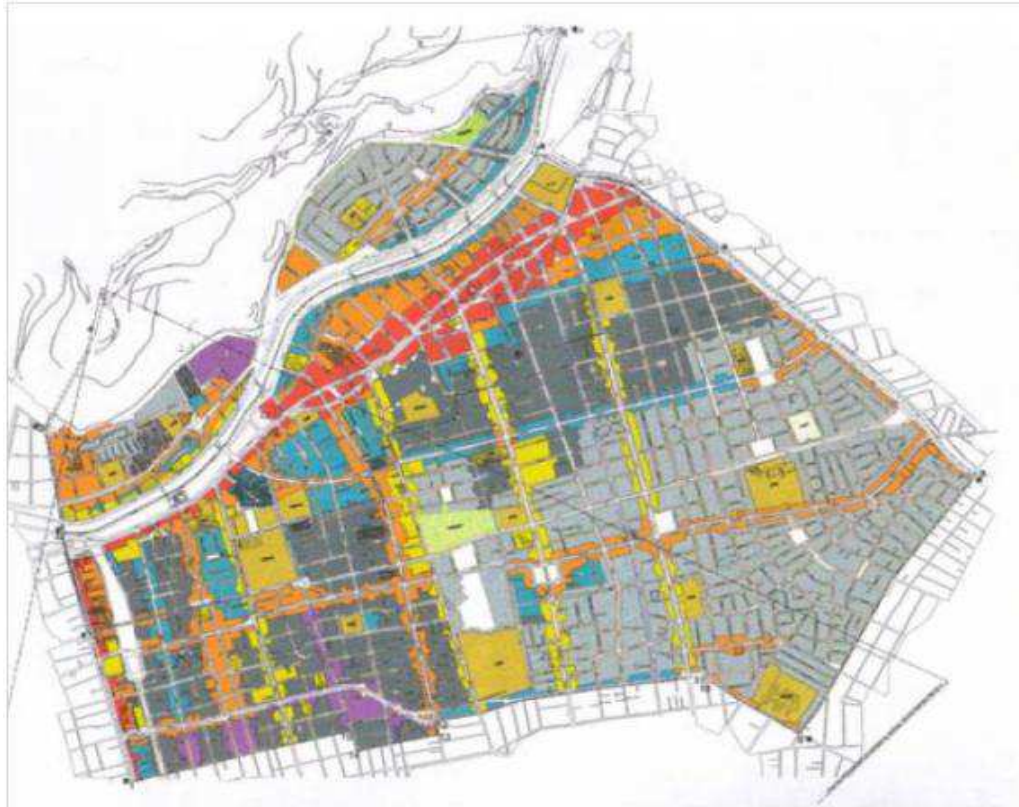
Fuente: Plano Regulador PRS – 01F de Zonificación General. Ilustre Municipalidad de Santiago, Septiembre 2011.

Para la comuna de Santiago (Imagen 3-2) se aprecia que aquella zona de color morado es la que presenta un uso principalmente comercial, servicios y cultural. Está delimitada por la Autopista Central al poniente, el Río Mapocho al Norte, Plaza Baquedano al oriente y Av. Libertador Bernardo O’Higgins al sur. Esta zona, que generalmente es conocida como el “centro de Santiago” posee variados paseos peatonales de varias cuadras que permiten el desplazamiento fluido de peatones. Esta característica lo hace un sector adecuado para la caminata.

En la imagen 3-3 se da a conocer la zonificación de la comuna de Providencia. Se observa que en el eje Av. Providencia – 11 de Septiembre se concentra principalmente el uso de suelo de tipo comercial (color rojo). Los límites de esta zona son Av. Salvador por el Poniente y Av. Tobalaba por el oriente.

La diferencia entre estas dos zonas tiene que ver con la forma geométrica de ambas. Mientras la zona comercial de Santiago presenta una forma triangular, la de Providencia presenta una forma más lineal, siguiendo el eje de Av. Providencia. Esto conlleva a tomar decisiones diferentes en la forma de abordar el rango de caminata a elegir.

Imagen 3-3: Plan Regulador de Providencia.



Fuente: Plan Regulador Comunal de Providencia. Ilustre Municipalidad de Providencia, 2007.

Para el caso de Providencia, la situación es más sencilla. El largo del eje entre los límites mencionados recientemente es de aproximadamente 2.950 metros. Tomando como referencia una rapidez promedio de caminata de  $1,5^3$  m/s todo el eje se podría caminar en un poco más de 30 minutos.

El caso de Santiago Centro es más complejo. Al existir diferentes rutas donde los peatones se pueden trasladar, no es posible definir una sola ruta y buscar el largo de ésta. Para esto se definen distancias máximas según eje. Las distancias de estos ejes se dan a conocer en la tabla 3-2. En ella se aprecia que el eje más largo es el correspondiente a la Alameda, que además corresponde a uno de los ejes comerciales y de servicios más importantes de Santiago.

Tabla 3-2: Tamaño ejes estructurantes sector centro Santiago.

<b>Eje</b>	<b>Inicio</b>	<b>Fin</b>	<b>Distancia [m]</b>
Alameda Bdo. O'Higgins	Av. Manuel Rodríguez	Av. Vicuña Mackenna	2.585
Av. Manuel Rodríguez	Av. Bdo. O'Higgins	Balmaceda	1.600
Costanera Mapocho	Av. Manuel Rodríguez	Av. Vicuña Mackenna	2.300

La cantidad de bajadas no caminables por zona, como se dijo anteriormente, no es posible definir las por zona. Sin embargo, se dan a conocer las subidas siguientes a

<sup>3</sup> Valor obtenido de REDEVU.



esas bajadas no caminables por zona de ambos sectores propuestos a estudiar. En la imagen 3-4 se muestran los centroides geográficos de las zonas correspondientes al centro de Santiago, mientras que en la imagen 3-5 se presenta la misma información pero para el eje de avenida Providencia.

Imagen 3-4: Zonas DIRTP sector Santiago Centro.

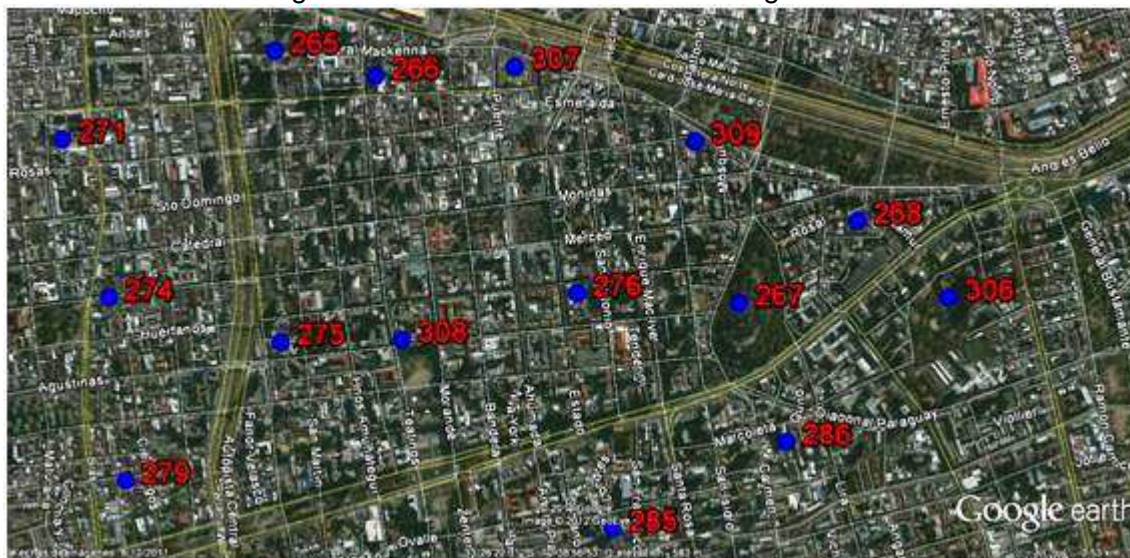


Imagen 3-5: Zonas DIRTP eje A. Providencia – 11 de Septiembre



La cantidad de transacciones posteriores a estas bajadas no caminables se dan a conocer en la tabla 3-3. Se aprecia que sumando las transacciones no caminables por zona, se tiene que en Santiago Centro son 136.553 y en el eje de Av. Providencia son 51.507, lo que da un total de 188.060 transacciones no caminables, que corresponden al 6,75% del total de bajadas no caminables (que son 2.786.685) de la semana de estudio. Si bien se trata de un porcentaje pequeño dentro del total de transacciones con el problema, se entiende que es un estudio para dos zonas, y que al realizar el mismo procedimiento para nuevas zonas se podría aumentar el porcentaje de corrección.

Tabla 3-3: Transacciones posteriores a bajadas no caminables por zona.

Sector	No. Zona	Transacciones
Santiago Centro	265	1.703
	266	13.916
	267	1.523
	268	3.840
	271	1.287
	274	13.848
	275	5.041
	276	15.592
	279	3.567
	285	22.020
	286	18.909
	306	13.898
	307	15.860
	308	1.880
309	3.669	
Eje Av. Providencia	175	22.405
	176	5.191
	178	9.437
	179	14.474

### 3.2.3. Propuesta de solución al problema.

Como propuesta de solución a este problema, se analiza eliminar el criterio de tramo no caminable para viajes entre las zonas recién mencionadas. Esto quiere decir que para viajes en el mismo sector (categorizado por las zonas mencionadas en la tabla 3-3) se libere la opción de caminar menos de 1.000 metros para transbordos o la realización de actividades, eliminando el filtro en este caso.

Si bien se definió en un comienzo que los datos conocidos correspondían a las transacciones siguientes a aquellas bajadas estimadas como no caminable, el modelo es capaz de generar una “posible bajada” que si bien se encuentra en un rango de distancia con la subida mayor a 1.000 metros se puede obtener computacionalmente. Actualmente esa bajada no se presenta en la base de datos, pero una vez generada, se utiliza para el criterio definido previamente.

Se propone además el estudio futuro de bajadas no caminables para sectores periféricos de Santiago. Este problema se presenta debido a una serie de factores aún no estudiados a cabalidad, pero en base a la experiencia vista en medios de comunicación y la conocida por el autor, los usuarios de transporte público pueden caminar muchas cuadras para lograr subirse a los buses, especialmente en las mañanas. La principal causa de este comportamiento es que los usuarios buscan ir hacia los cabezales de inicio de recorrido, esto debido a que existe una gran demanda en esos paraderos que impiden a los usuarios subirse en el más cercano a su hogar

que se supone corresponde al cercano a la bajada del día anterior. Al pasar los buses llenos, los usuarios deben realizar estos grandes traslados para poder acceder al sistema. Y otra causa de la aparición de bajadas no caminables tiene relación con el uso de modos no integrados entre etapas realizadas en Transantiago, las cuales al no quedar registradas se consideran como una caminata que posiblemente sea de gran distancia y que el modelo la considere como no caminable. Por esto se hace necesario estudiar la incidencia de los modos no integrados dentro del viaje en modos integrados a Transantiago.

### **3.3. Primera transacción del día.**

#### **3.3.1. Situación actual.**

El proceso actual que realiza el modelo de asignación de bajada para estimar el paradero de bajada de la última transacción del día se basa en utilizar el paradero más cercano a la subida estimada de la primera transacción del día como bajada de la última validación. Esto con el fin de cerrar el ciclo de viajes del pasajero, tomando como base el hecho de que la personas vuelve a su origen al final del día.

Sin embargo, para el modelo el día de viajes de un usuario comienza a las 00:00 horas, lo que implica que se define la primera transacción del día como la primera realizada después de esa hora. Esto genera que aquellas transacciones que ocurren de madrugada y que podrían corresponder a viajes de regreso de un ciclo de viajes que comenzó el día anterior queden asignadas como los primeros viajes de un nuevo ciclo del día actual.

El modelo debe tener en cuenta que el día de viaje de los usuarios generalmente no comienza a las 00:00 horas, sino que durante la mañana. A partir de esta disyuntiva se debe entender la diferencia entre un día cronológico (que va desde la 00:00 horas hasta las 23:59 horas) de un día de viajes, que puede tener otro horario de inicio y término.

Es por eso que se propone en este estudio el analizar una nueva hora de inicio de día de viaje, analizando estadísticamente la hora que se definirá como comienzo del día de viaje.

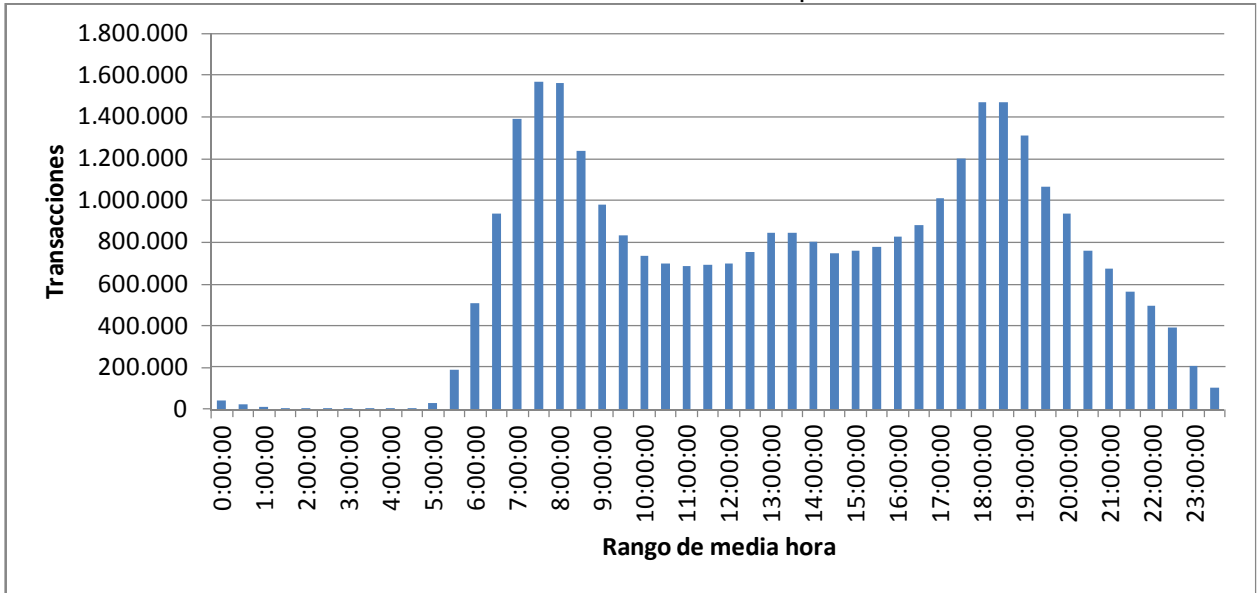
#### **3.3.2. Análisis estadístico.**

Para poder observar el comportamiento de viaje de los habitantes de Santiago en aquellos modos integrados en Transantiago, lo primero que se debe analizar es la cantidad de transacciones realizadas por media hora, clasificación que actualmente es la más desagregada a la hora de contar viajes por periodo de tiempo.

En primer lugar se hará esta clasificación por tipo de día: Laboral, Sábado y Domingo. Para el caso de día laboral (Lunes a Viernes) la distribución de transacciones por media hora se da a conocer en el gráfico 3-1. En él se puede apreciar que a partir de las 5:00 horas comienzan a aparecer de manera importante las transacciones en día laboral. Si bien distan bastante de la hora peak de transacciones (7:30 horas) se tiene

conocimiento que estas transacciones corresponden a usuarios que se deben trasladar grandes distancias para acceder a su destino.

Gráfico 3-1: Distribución de transacciones para día laboral.



El gráfico 3-2 da a conocer la distribución de transacciones por media hora para el día Sábado de la semana de estudio. Para este día el punto de inflexión en el aumento de transacciones durante la madrugada ocurre nuevamente a las 5:00 horas. Sin embargo, se observa un alza importante a partir de las 5:30 horas.

El comportamiento del día Domingo se da a conocer en el gráfico 3-3. En el histograma se observa nuevamente que a las 5:00 horas comienza el aumento en el número de transacciones, al igual que en los 2 casos anteriores.

Gráfico 3-2: Distribución de transacciones para día Sábado.

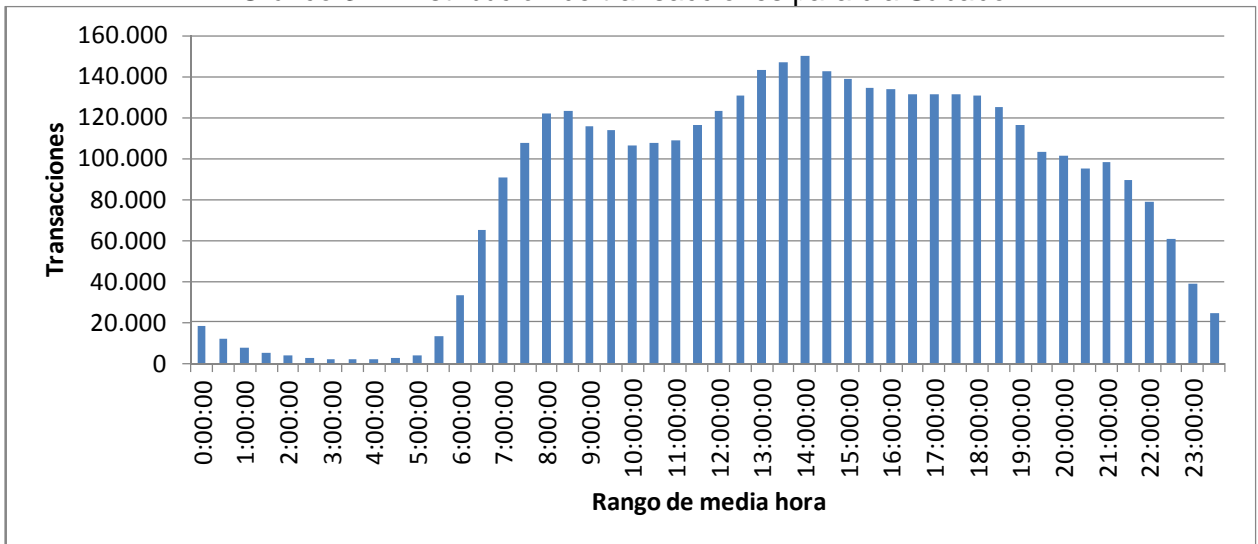
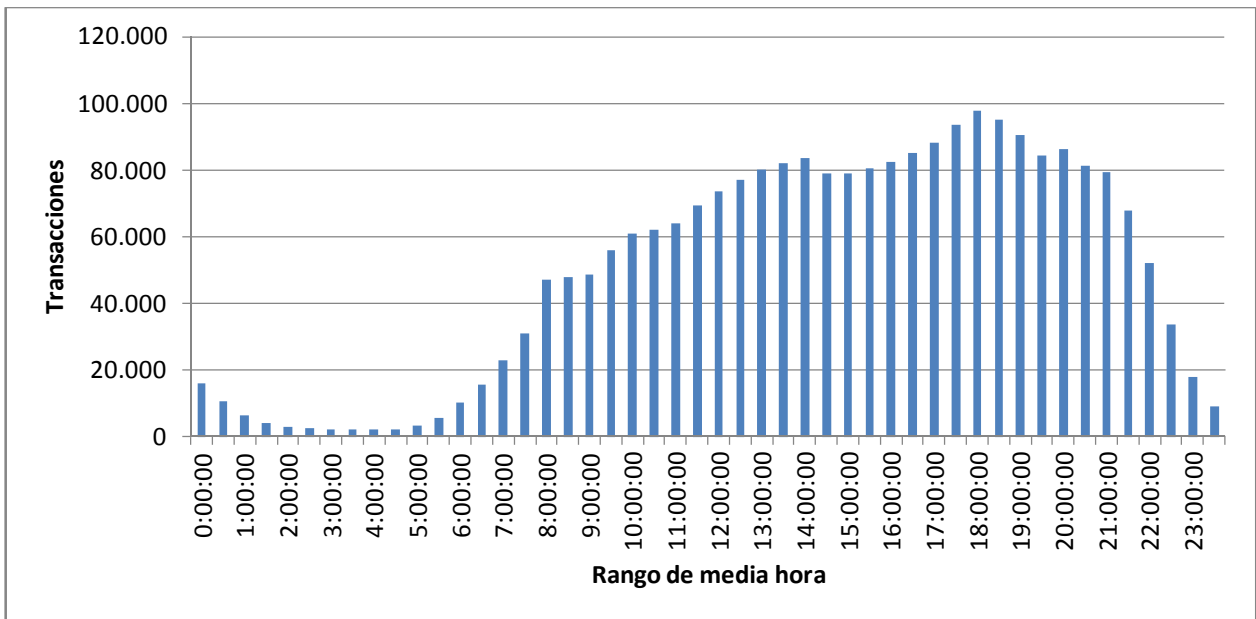


Gráfico 3-3: Distribución de transacciones para día Domingo.



En la tabla 3-4 se dan a conocer la cantidad de transacciones realizadas entre las 00:00 horas y las 04:59 horas de cada uno de los días de la semana de estudio, que se espera sea el rango horario donde se encuentre la hora crítica de transacciones mínimas. Una primera consideración al respecto es que estas transacciones son sólo realizadas en modo Bus, esto debido a que Metro no realiza servicios en ese tramo horario. De esta tabla se logra apreciar una situación interesante: los fines de semana la cantidad de transacciones realizadas en este tramo horario son considerablemente mayores que en día de semana, posiblemente producto del aumento de viajes con motivos de ocio, esparcimiento y entretención. Sin embargo, se aprecia que en la media hora comprendida entre las 04:30 y las 04:59 horas este comportamiento comienza a decaer y los viajes se asemejan más a los de un día laboral. Y en el tramo horario que comienza a las 05:00 horas los días laborales tienen predominancia nuevamente en relación a la generación de viajes.

Otro análisis que se puede hacer es mirar los días laborales. Mientras la noche del Lunes es la que presenta menor cantidad de transacciones, la del viernes es la con mayor de transacciones. Lo del Viernes tiene una explicación similar a lo que ocurre los fines de semana (viajes de ocio) aunque en menor escala respecto al fin de semana.

Al observar el total de viajes por tramo horario durante la semana, el lapso de menor cantidad de transacciones ocurre en el periodo comprendido entre las 3:30 y las 4:00 de la mañana. Si bien existen dos días en la semana de estudio que no presentan la menor cantidad de transacciones en ese tramo horario (Lunes y Sábado), se trata de diferencias bajas respecto del total de transacciones en la madrugada.

Claramente el hecho de no poder obtener información acerca de motivos de viaje de manera desagregada impide realizar un análisis más acabado sobre esto, pero con la información que se posee ya queda de manifiesto que la jornada de viajes comienza a las 4:00 horas, por lo que se propone definir esta hora como nuevo comienzo del día. Esto además simplifica el trabajo del modelo de estimación de bajada, pues el hecho de

definir distintas horas de comienzo de viaje por tipo de día implicaría generar mayores complicaciones al modelo.

Tabla 3-4: Transacciones realizadas por media hora entre 00:00 y 04:59 horas.

Media Hora	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Total
0:00:00	4.996	9.315	8.531	9.088	10.536	18.291	16.039	<b>76.796</b>
0:30:00	3.161	5.104	4.672	5.110	5.977	12.031	10.799	<b>46.854</b>
1:00:00	1.474	2.325	2.298	2.830	3.247	8.085	6.319	<b>26.578</b>
1:30:00	684	1.280	1.397	1.621	2.210	5.337	4.314	<b>16.843</b>
2:00:00	395	843	871	1.072	1.409	3.918	2.963	<b>11.471</b>
2:30:00	296	575	662	760	1.026	2.820	2.442	<b>8.581</b>
3:00:00	214	539	615	682	929	2.444	2.238	<b>7.661</b>
3:30:00	288	439	542	616	898	2.264	2.184	<b>7.231</b>
4:00:00	484	673	802	916	1.181	2.137	2.281	<b>8.474</b>
4:30:00	1.359	1.516	1.557	1.598	1.992	2.530	2.344	<b>12.896</b>
5:00:00	6.038	5.895	6.004	6.014	6.213	4.340	3.391	<b>37.895</b>
<b>Total Día</b>	<b>19.389</b>	<b>28.504</b>	<b>27.951</b>	<b>30.307</b>	<b>35.618</b>	<b>64.197</b>	<b>55.314</b>	<b>261.280</b>

Realizando un análisis estadístico, se tiene que existen 202.015 transacciones realizadas en la semana de estudio entre las 00:00 horas y las 03:59 horas que corresponde al 0,53% del total de transacciones de la semana. Se trata de una fracción muy baja, debido a que existen pocas actividades en ese periodo que generen o atraigan viajes.

En la base de datos hay 147.348 transacciones realizadas entre las 00:00 horas y las 03:59 horas categorizadas como primera transacción del día, que corresponden al 72,94% de las transacciones realizadas en el tramo horario entre 00:00 y 03:59 horas. Estas transacciones son las que dejarán de ser categorizadas como primera del día y pasarán a ser de las últimas realizadas el día anterior.

### 3.3.3. Propuesta de solución.

El proceso de solución se basa en una nueva definición de día de viaje, que sea distinto al día cronológico que se conoce. Esto es un día de viaje de 24 horas, pero que comience a las 04:00 horas y termine a las 04:00 horas del día siguiente.

Para esto se debe en primer lugar definir nuevamente el concepto de primera transacción del día. Esta, que actualmente se define como la primera transacción realizada después de las 00:00 horas del día de estudio, se debe cambiar a la primera transacción realizada después de las 04:00 horas del día de estudio.

Luego, se debe también redefinir el concepto de última transacción del día. Para esto, se debe cambiar el concepto actual en que esta última transacción se define como aquella realizada antes de las 23:59 horas del día de estudio. Con el cambio propuesto, la última transacción del día debe ser la última realizada antes de las 03:59 horas del día siguiente al de estudio.

Con estas nuevas definiciones, el modelo de asignación de bajada considerará nuevas transacciones como primera y última del día, por lo que la asignación de bajada para la última transacción del día se hará con aquella que fue la primera del día y que además ocurrió posterior a las 04:00 horas.

Finalmente, el modelo deberá utilizar el mismo criterio que utiliza actualmente para estimar la bajada de la última transacción del día, es decir:

- Se estima la bajada de la última transacción del día a partir de la ubicación de la primera transacción del día, es decir, el usuario vuelve al origen de su primer viaje del día.
- En caso que no se logre estimar bajada con el criterio anterior, se estima la bajada de la última transacción del día a partir de la primera transacción del día siguiente, es decir, siguiendo el ciclo de estimación de bajada normal (estimar bajada con la subida siguiente).

### **3.4. Validaciones únicas del día.**

#### **3.4.1. Situación Actual.**

Otro problema que presenta el modelo es el hecho de que se hagan validaciones únicas, es decir, que durante un día de viajes el usuario sólo realice una transacción con su tarjeta bip! en alguno de los validadores del sistema. Esto genera un problema en la capacidad del modelo de estimar bajada, debido a que una de las condiciones de estimar bajada cuando ocurre la última validación del día (que en este caso al ser única es la primera y la última) es que se busca como bajada el paradero de la ruta más cercano a la primera validación del día. Esto basándose en la idea de que un usuario vuelve a su punto de origen al finalizar la actividad.

Como este caso se trata de una única validación en el día, no hay una transacción previa en el día que se pueda usar para estimar bajada. Esta rigidez en la estimación del modelo que impide estimar paradero de bajada para este caso y que hasta hoy no ha sido tratada, se puede ver desde otro punto de vista.

Si suponemos que un usuario que realiza una única transacción en un día, y esta transacción es realizada al finalizar el día, se puede buscar su bajada tomando en consideración la primera validación del día siguiente, siempre y cuando ésta ocurra durante el comienzo de la siguiente jornada. Con esto se asume que el pasajero durante la noche vuelve a su lugar de origen y al día siguiente sale del mismo. Este procedimiento es similar al caso en que se vuelve a la primera subida del mismo día, sólo que se ocupa el día siguiente.

Esta hipótesis planteada tiene ciertas restricciones que actualmente son parte del modelo. En primer lugar que exista un paradero de bajada para la validación única cercano a aquel utilizado como subida al día siguiente. Lo siguiente es que la validación ocurra el día siguiente, por lo que se debe tener cuidado con utilizar la siguiente validación si ésta ocurre en el día subsiguiente o posterior. Finalmente que la diferencia temporal entre ambas validaciones sea menor a 24 horas, pues ahí ya entraríamos al

caso de que el usuario en uno de los viajes (posiblemente en el de ida) esté utilizando alguno de los medios de transporte no integrados en Transantiago. Y además se debe tener en consideración los cambios propuestos al concepto de primera y última validación del día descritos en el capítulo anterior (3.3. Primera transacción del día).

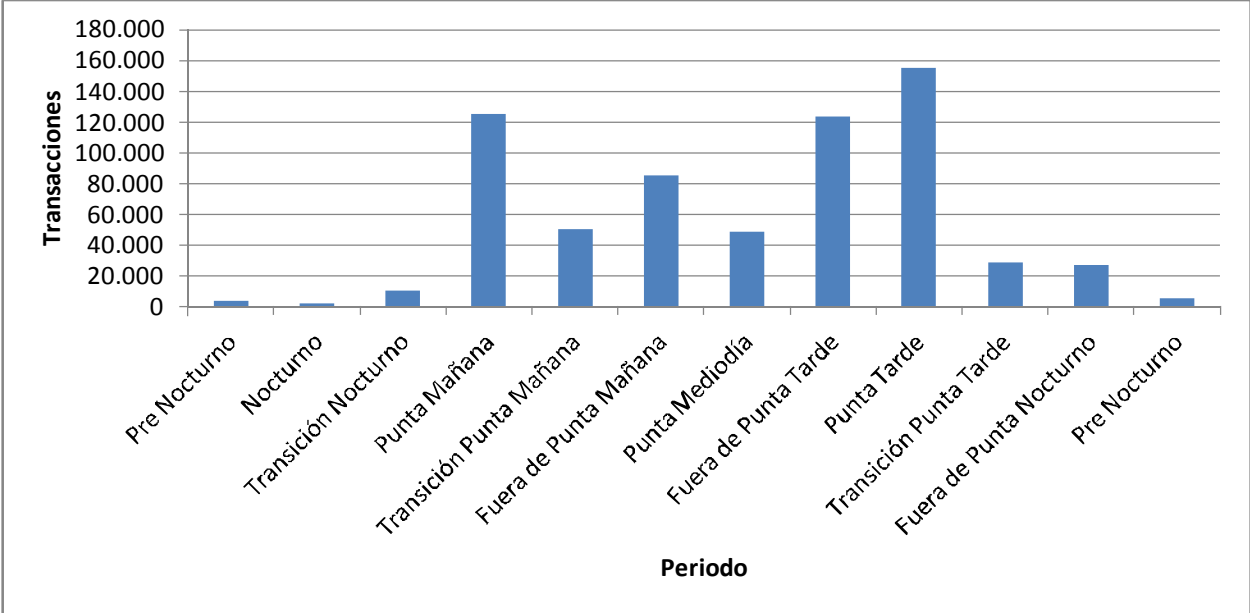
**3.4.2. Análisis estadístico.**

Para la semana en estudio, se tiene que existen 2.055.515 transacciones únicas en el día, que corresponden al 5,39% del total de transacciones realizadas. Este es el segundo mayor problema del modelo en cantidad de estimaciones de bajada nulas, tras aquellas bajadas no caminables, estudiadas en el acápite anterior.

Un primer filtro a realizar con el fin de estudiar este comportamiento y darle solución implica eliminar del estudio todas aquellas transacciones únicas del día que no presentan una transacción al día siguiente, pues a ellas no se les puede analizar la bajada. Esto en primer lugar deja fuera todas aquellas validaciones únicas del día realizadas el día Domingo, pues no existe información en la base de datos sobre el día siguiente (la semana se cuenta de Lunes a Domingo). Además quedan fuera las últimas validaciones de la semana que sean únicas en el día. Finalmente, aquellas tarjetas con una sola transacción durante la semana también quedan fuera del estudio. Este filtro deja disponibles para estudio 724.543 observaciones, que corresponden al 35,25% del total de transacciones únicas en el día.

Las transacciones hechas en día laboral (Lunes a Viernes) por período del día se aprecian en el gráfico 3-4. Se puede apreciar que el momento de mayor ocurrencia de este caso es en los periodos Fuera Punta Tarde y Punta Tarde, que es concordante con los periodos de mayor cantidad de transacciones en el sistema para los días laborales (ver tabla 2-2).

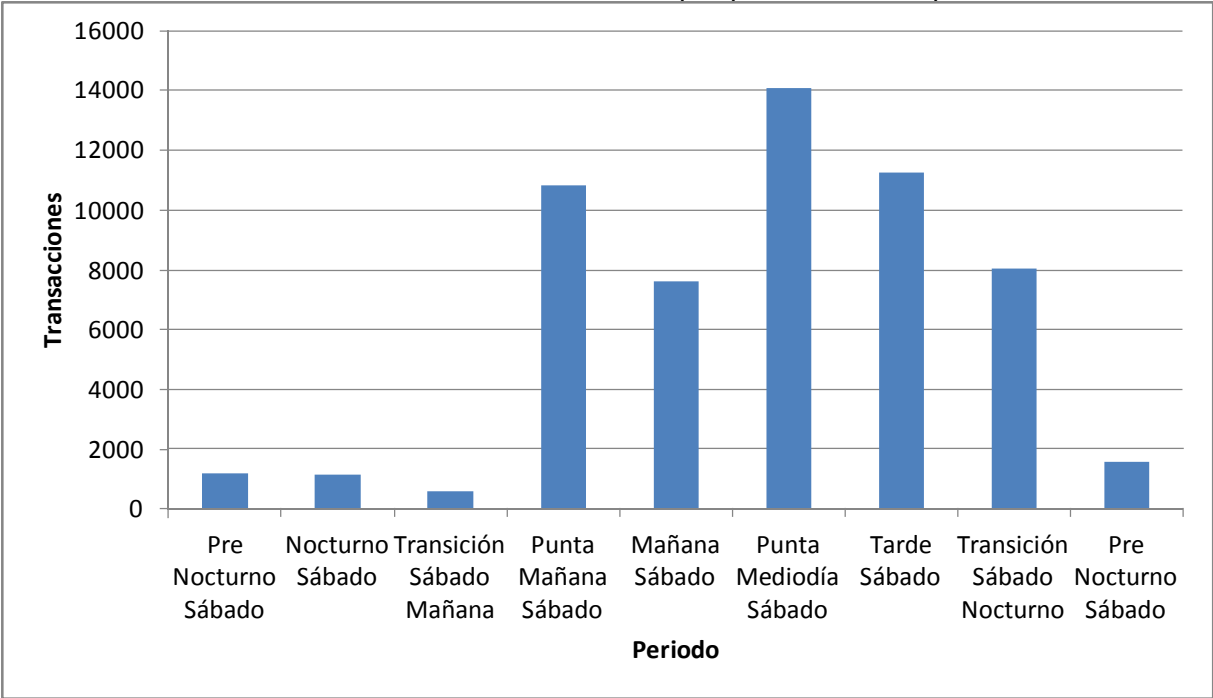
Gráfico 3-4: Transacciones únicas en el día por periodo del día para día laboral.





Las transacciones únicas realizadas el día Sábado clasificadas por período del día de estudio se dan a conocer en el gráfico 3-5. Se observa una distribución similar al de las presentes en día Sábado, donde el periodo del día con mayor cantidad de transacciones corresponde al Punta Mediodía. Sin embargo, en este caso hay más transacciones únicas del día en el periodo Tarde que en Punta Mañana, a diferencia del caso de transacciones totales del día donde ocurre lo contrario (ver tabla 2-2).

Gráfico 3-5: Transacciones únicas en el día por periodo del día para Sábado.



Con esto se muestra que en general no existe un patrón especial de comportamiento a la hora de generarse este problema de validación única en el día, salvo para el día Sábado donde se observa una mayor cantidad de este tipo de transacciones en la tarde que en la mañana.

Para este análisis se pretende utilizar para los días Lunes a Jueves desde el periodo Punta Tarde (17:30 horas) hasta el Pre-Nocturno Final (23:00 horas) en el caso de las transacciones únicas del día, y compararlas con las transacciones del día siguiente (Martes a Viernes) realizadas entre los periodos Pre Nocturno (00:00 horas) y Fuera de Punta Mañana que finaliza a las 12:29 horas. Las validaciones que cumplen las condiciones de espacio temporal previamente nombradas, clasificadas por periodo se dan a conocer en la tabla 3-5, donde se logra apreciar que las transacciones únicas realizadas en periodo Punta Tarde aportan parte importante de las validaciones a estudiar. Además se da lo esperado de que las validaciones del día siguiente realizadas en periodo Punta Mañana son las que más se combinan con las realizadas en periodo Punta Tarde del día anterior (36.478). En total se tienen 105.256 transacciones únicas en el día con posibilidad de obtener estimación de bajada para este primer caso.

Tabla 3-5: Transacciones únicas del día realizadas entre Lunes y Jueves. Periodo siguiente transacción.

Periodo Transacción Día Siguiente (Martes a Viernes)	Horario	Período Transacción Única (Lunes a Jueves)			
		Punta Tarde	Transición Punta Tarde	Fuera Punta Nocturno	Pre Nocturno
		17:30-20:29	20:30-21:29	21:30-22:59	23:00-23:59
<b>Pre Nocturno</b>	00:00 - 00:59	793	196	261	447
<b>Nocturno</b>	01:00 - 05:29	792	148	217	65
<b>Transición Nocturno</b>	05:30 - 06:29	3.041	530	401	82
<b>Punta Mañana</b>	06:30 - 08:29	36.478	5.856	4.286	521
<b>Transición Punta Mañana</b>	08:30 - 09:29	15.837	2.855	2.162	296
<b>Fuera Punta Mañana</b>	09:30 - 12:29	20.408	4.592	4.187	805
<b>Total</b>		<b>77.349</b>	<b>14.177</b>	<b>11.514</b>	<b>2.216</b>

El segundo tramo a analizar corresponde a aquellas transacciones únicas realizadas el día Viernes y aquellas realizadas en la mañana del día Sábado. Para esto, se usa el mismo tramo horario para el día Viernes que para Lunes a Jueves, es decir, desde el periodo Punta Tarde (17:30 horas) hasta las 23:59 horas. Para el día siguiente, que corresponde al Sábado, se utilizan las transacciones realizadas entre las 00:00 horas (Pre-Nocturno) y las 13:29 horas (Punta Mañana Sábado). En la tabla 3-6 se dan a conocer los valores de estos pares de periodos. Al igual que para el caso de los días laborales, la existencia de transacciones únicas en el día se aprecian principalmente para los periodos Punta Tarde en Viernes y Punta Mañana en Sábado. Este caso concentra 10.391 transacciones de las 26.285 transacciones que se hacen en la tarde del Viernes y tienen como primera del día siguiente la mañana del Sábado.

Tabla 3-6: Transacciones únicas del día realizadas en día Viernes. Periodo siguiente transacción.

Periodo Transacción Día Siguiente (Sábado)	Horario	Período Transacción Única (Viernes)			
		Punta Tarde	Transición Punta Tarde	Fuera Punta Nocturno	Pre Nocturno
		17:30-20:29	20:30-21:29	21:30-22:59	23:00-23:59
<b>Pre Nocturno</b>	00:00 - 00:59	469	132	211	260
<b>Nocturno</b>	01:00 - 05:29	406	158	258	112
<b>Transición Mañana</b>	05:30 - 06:29	338	76	118	32
<b>Punta Mañana</b>	06:30 - 10:59	10.391	2.443	2.395	413
<b>Mañana</b>	11:00 - 13:29	5.626	1.440	1.261	286
<b>Total</b>		<b>17.230</b>	<b>4.249</b>	<b>4.243</b>	<b>1.103</b>

Finalmente, se analizan aquellas transacciones únicas del día realizadas en día Sábado, entre las 17:30 horas y las 23.59 horas, y aquellas cuya primera transacción es Domingo entre las 00:00 horas y las 13:29 horas. En la tabla 3-7 se presentan las transacciones por periodo, donde se da a conocer que se tienen en total 12.316 validaciones. Al igual que en los otros casos, y como era de esperar, el par de periodos

con mayor cantidad de ocurrencias es el Punta Tarde de Sábado con Punta Mañana del Domingo, con 3.813 oportunidades.

Tabla 3-7: Transacciones únicas del día realizadas en día Sábado. Período siguiente transacción.

Periodo Transacción Día Siguiente (Domingo)	Horario	Período Transacción Única (Sábado)		
		Tarde	Transición Nocturno	Pre Nocturno
		17:30-20:29	20:30-22:59	23:00-23:59
<b>Pre Nocturno</b>	00:00 - 00:59	341	397	432
<b>Nocturno</b>	01:00 - 05:29	307	531	135
<b>Transición Mañana</b>	05:30 - 09:29	1.776	1.591	171
<b>Mañana</b>	09:30 - 13:29	3.813	2.504	318
<b>Total</b>		<b>6.237</b>	<b>5.023</b>	<b>1.056</b>

### 3.4.3. Propuesta de Solución.

Este problema tiene una solución no muy compleja y fácil de implementar. Lo primero que se debe hacer es identificar la transacción como única del día, y analizar si la siguiente transacción corresponde al día posterior.

Luego, se filtran aquellos periodos a utilizar para el proceso. Esto quiere decir que se seleccionan aquellas validaciones únicas realizadas en la tarde (desde 17:30 horas en adelante) y se buscan aquellas validaciones del día siguiente que sean realizadas en los periodos de la mañana, que dependiendo del caso tienen diferente horario de término.

Finalmente, una vez obtenidos aquellos casos posibles de estimar bajada, se realiza el proceso de estimación de paradero de bajada dependiendo del tipo de transporte utilizado y con los criterios establecidos de distancia entre posible bajada y subida al día siguiente.

Se tiene que existen 144.397 transacciones con posibilidad de ser estimada parada de bajada siendo las únicas del día. Del total de transacciones únicas del día este valor corresponde al 7,02%. Pero si analizamos sólo aquellas que tienen posibilidad de ser obtenida estimación de bajada luego de la eliminación de aquellas que no poseen transacción al día siguiente o son únicas en la semana, este valor porcentual aumenta al 19,93%.

Con esto queda claro que no todas las validaciones únicas del día tendrán posibilidad de estimar paradero de bajada, pues dependerán de la hora del día en que hayan sido realizadas y si presentan transacciones al día siguiente. Sin embargo se logrará obtener un aumento en las bajadas estimadas.

### **3.5. Estimación de paradero de subida y bajada cercano.**

#### **3.5.1. Situación actual.**

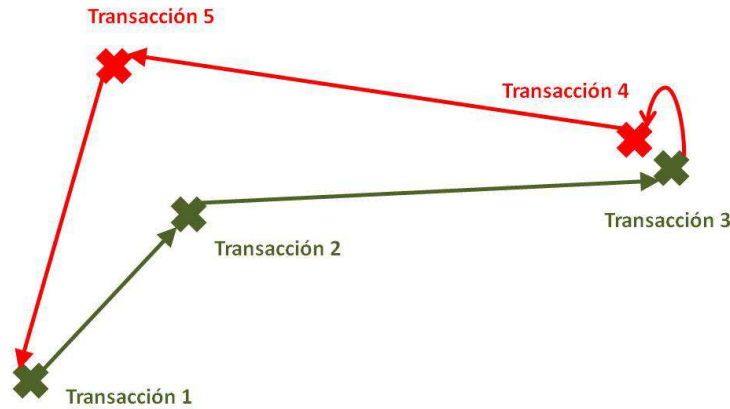
El modelo de asignación de paradero de bajada presenta problemas para un caso bastante complejo de analizar estadísticamente: las etapas del viaje no observadas. Esto, que es generado por dos grandes razones, la evasión en el pago y el uso de otro medio de transporte en una de las etapas del viaje, no es a simple vista posible de observar en el conjunto de transacciones presentadas en la base de datos.

Sin embargo, el comportamiento en sí es posible deducirlo en base a una falla en el modelo que causa este problema. Cuando el modelo asigna para una transacción paradero de subida y de bajada cercanos del punto de vista espacial lo que está diciendo es que existieron dos transacciones consecutivas realizadas en paraderos geográficamente adyacentes o en el mismo paradero. Con esto, uno supone que falta una etapa de viaje entre ellas, lo que implica que el modelo al usar la subida de la segunda transacción para estimar la bajada de la primera, genere un viaje de pocos metros. El modelo actualmente clasifica estas transacciones como “etapa muy corta”, y lo hace para aquellas bajadas estimadas que se encuentran a 50 o menos metros de la subida medidos en línea recta.

Para poder analizar este caso, se debe tener en consideración la diferencia temporal entre las dos transacciones consecutivas. En el caso de que la distancia temporal sea de horas, uno puede suponer que lo que ocurrió es la ausencia de una etapa del viaje, pues lo que debe haber ocurrido es la realización de un viaje de ida y vuelta. Sin embargo, si la distancia temporal entre las transacciones es menor y éstas se realizaron en paraderos cercanos geográficamente o en el mismo paradero, se puede suponer que lo que ocurrió fue el uso de una sola tarjeta para la realización de dos o más viajes de distintos usuarios, dando la idea de un préstamo de la tarjeta para realizar la transacción que permite el abordaje al bus o a Metro. Este caso de distancia temporal menor entre transacciones será revisado en el capítulo 4 (Multipago) en el que se analizará en mayor detalle las causas del comportamiento y las condiciones que lo diferencian respecto al estudio de paradero de bajada y subida cercano.

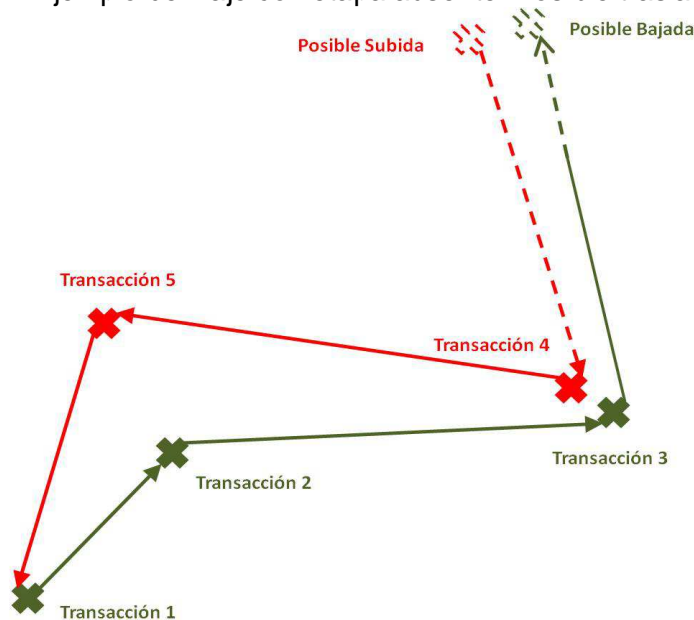
En la figura 3-1 se muestra esquemáticamente lo que sucede para un traslado de ida (ruta verde) y vuelta (ruta roja) a un punto en dos viajes de tres etapas cada uno, en el que se supone por simplicidad que en cada punto asignado con una X ocurrió la bajada y la posterior transacción de subida (cercanos geográficamente, a menos de 51 metros). En él se aprecia que la transacción 3 y 4 tienen como punto de subida paraderos muy cercanos geográficamente. Es por eso que el modelo entiende que se hizo un viaje con inicio en el punto donde se realiza la transacción 3 y que finaliza en un punto cercano donde se realiza la siguiente transacción, que en este caso es la 4. Como son muy cercanas ambas, se asume que el viaje tuvo subida y bajada cercana, y por ende, el traslado es pequeño. Además, esto genera un dato erróneo de viaje que afecta los cálculos de matrices origen destino, pues genera un viaje que no corresponde a la realidad, agregando valores a estas matrices.

Figura 3-1: Ejemplo de viaje con etapa de viaje ausente.



Lo que se espera que haya sucedido desde el punto de vista de movimientos del pasajero es lo que se denota en el esquema de la figura 3-2. En ella se presenta con línea segmentada aquellos datos no conocidos. Como la transacción 3 tiene en su información el servicio y sentido abordado, se conoce la ruta del viaje (línea verde continua). Sin embargo, no se conoce la bajada de esta transacción (cruz segmentada verde), la subida luego de haber realizado la actividad (cruz segmentada roja) ni el recorrido abordado (línea segmentada roja), sea este en modo integrado a Transantiago u otro modo. Este desplazamiento desde la cruz segmentada roja hasta la posición de la transacción 4 es de la que no se tiene información. Puede haber existido un uso de otro medio de transporte externo al sistema de pago de Transantiago (colectivo, taxi, bicicleta, caminata, etc.) o simplemente hubo una evasión en el pago del viaje en rojo segmentado que no quedó registrado en la base de datos de transacciones.

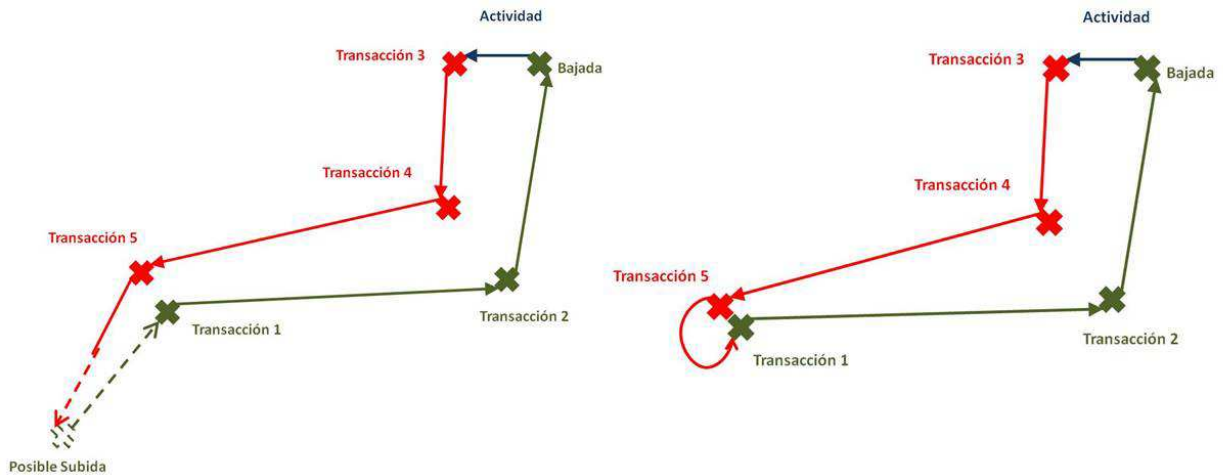
Figura 3-2: Ejemplo de viaje con etapa ausente. Posible traslado realizado.



Un caso particular corresponde a aquel cuando el primer viaje del día no posee transacción. Esto influye negativamente en los datos obtenidos, pues implica no

conocer el origen del viaje, y por ende, el posible hogar del usuario. En la figura 3-3 se presenta este caso, donde la transacción 5 busca para estimar bajada la subida de la transacción 1. Como ésta es muy cercana (mismo sector) el modelo estima bajada para la transacción 5 el mismo sector de subida, pues es el paradero más cercano a la subida de la transacción 1. Sin embargo, la transacción 5 tiene la información del servicio y sentido del mismo, por lo que se conoce la ruta de destino al hogar, pero no el lugar exacto.

Figura 3-3: Ejemplo etapa ausente. Caso primera validación del día.



Este caso particular se da a conocer porque existe una solución posible para erradicar la falla, la que se dará a conocer más adelante.

### 3.5.2. Análisis estadístico.

De las 38.118.949 transacciones realizadas, 922.070 son categorizadas como “etapa muy corta” (ver tabla 2-5) que representa el 2,42% del total de transacciones de la semana. Sin embargo, existen estimaciones de paradero de subida y bajada similar para una transacción cuya distancia euclidiana entre subida y bajada estimada es mayor a 50 metros. La causa de esta omisión es la distancia entre paraderos consecutivos, la cual es superior al criterio dado y por ende se genera una asignación de paradero de bajada que corresponde al mismo paradero de subida. Sumando estos casos se tiene que existen 1.074.347 transacciones con paradero de subida y bajada cercano o similar, que corresponden al 2,82% del total de transacciones de la semana.

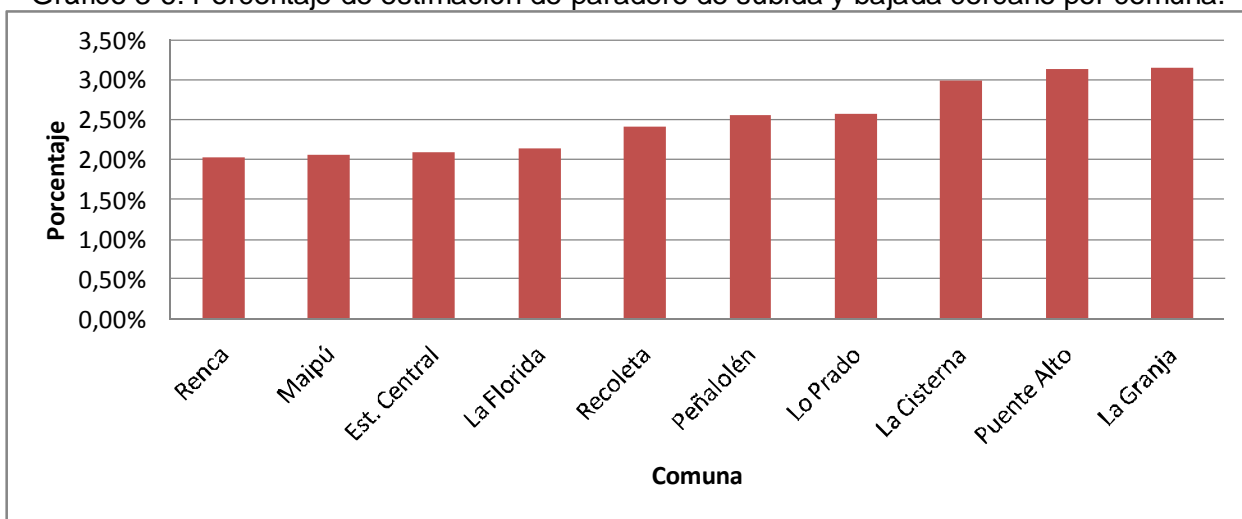
Dependiendo del lugar donde se realiza la transacción, el porcentaje de ocurrencia de este problema es el que se presenta en la tabla 3-8. De esta tabla se pueden observar que en zonas pagas el comportamiento es mayor que en el caso de las validaciones arriba de los buses. Una causa posible de esto es que el acceso a las zonas pagas ocurra en otro medio de transporte, considerando que las zonas pagas están ubicadas en lugares concurridos que tienen buena accesibilidad de medios como los colectivos.

Tabla 3-8: Transacciones con misma subida y bajada por lugar de validación

Zona Validación	Subida y bajada cercana	Total transacciones	Porcentaje [%]
<b>Zona Paga</b>	120.015	2.034.917	5,90
<b>Bus</b>	711.859	23.614.102	3,01
<b>Metro</b>	242.473	12.470.020	1,94

Otro análisis importante de conocer la comuna de origen de las transacciones. En el gráfico 3-6 se dan a conocer de manera porcentual las diez comunas con mayor incidencia de este problema, en el que se puede apreciar que el comportamiento se asocia principalmente a comunas residenciales, por lo que se asume que generalmente este problema puede afectar el estimar la bajada correspondiente al hogar del pasajero.

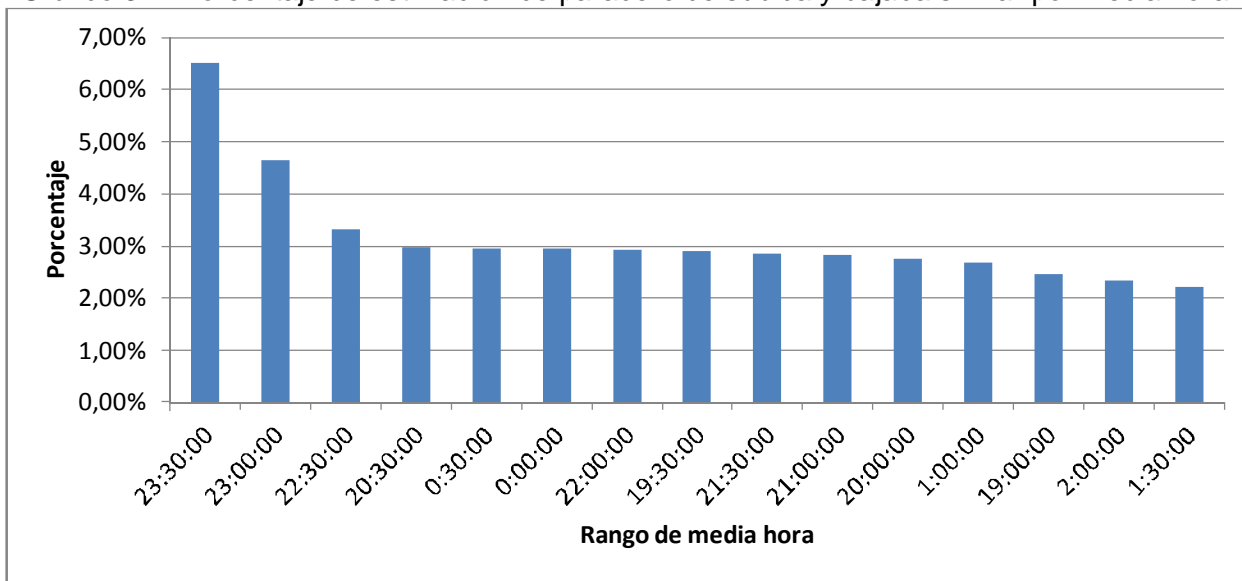
Gráfico 3-6: Porcentaje de estimación de paradero de subida y bajada cercano por comuna.



Para definir esto, se realiza el mismo análisis para la hora del día, clasificando por tramos de media hora este comportamiento de manera porcentual sobre el total de las transacciones realizadas agrupadas en tramos de media hora. Este análisis se ilustra en el gráfico 3-7, en el que se puede apreciar que de manera porcentual este comportamiento principalmente ocurre durante la noche. Esto complementa lo estudiado en el ejemplo 2 (figura 3-3) donde se aprecia que al no existir una primera validación en el día, la última que se realiza genera la estimación de subida y bajada similar. Por lo que este problema se genera principalmente con el hecho de no validar en el primer viaje de la mañana.

Las causas de esto son variadas, pero se puede suponer que ante la necesidad de acceder rápido a ciertos puntos neurálgicos de la ciudad se ocupan otros medios de transporte no integrados (colectivo, auto particular) que permiten llegar de manera más cómoda y rápida a estos lugares. Ya en la tarde, sin la necesidad de tener un horario de llegada obligado, el usuario se relaja y utiliza los modos integrados de transporte. Claramente otra gran causa de esto es que se evada el pago del pasaje en la mañana, pero no tan solo por no querer pagar el pasaje, sino que también porque los buses operan con niveles de carga muy elevados, a veces sobrepasando la capacidad, y permiten a los pasajeros acceder por las puertas traseras de las máquinas sin pagar el pasaje en el validador, por lo que el viaje no queda registrado.

Gráfico 3-7: Porcentaje de estimación de paradero de subida y bajada similar por media hora.



Además el hecho de que el comportamiento se aprecie principalmente en horas de la noche genera una conexión con lo explicado en el capítulo 3.3. referido al cambio de hora de inicio del día de viajes. Al proponer un día de viaje que comience a las 04:00 horas, todas aquellas transacciones realizadas en horario nocturno y que actualmente estén catalogadas como última transacción del día se les estimará bajada con la nueva primera transacción del día, lo que debería reducir este problema.

### 3.5.3. Propuesta de solución.

Como se ha explicado hasta ahora, cualquier hipótesis que se plantee sobre las causas de este problema son meras especulaciones, pues no se posee la información suficiente para encontrar la causa del error, a pesar de que se puede suponer a priori que existen dos grandes causas que influyen en esto: el no pago en una de las etapas del viaje en el sistema, y el uso de otros medios de transporte no integrados.

Lo primero a definir es que el paradero de bajada asignado no corresponde a la realidad, por lo que se propone que al encontrar este problema se anule la estimación del paradero de bajada y todas sus agregaciones geográficas (comuna, zona EOD, etc.), para no alterar el conteo de matrices origen destino producto de esto.

Pero existe una forma de obtener el paradero de bajada para estos casos, en el caso de que esta transacción sea la última del día. Cuando se intenta cerrar el ciclo de viajes con la primera transacción del día, y esta no corresponde al inicio original de la sucesión de viajes, se produce este problema. Pero esto se puede mitigar utilizando la primera transacción del día siguiente. Esto, suponiendo que encontrará un paradero de bajada distinto al encontrado en un primer momento. Esta solución sólo sirve si esta ausencia de etapa del viaje inicial sucede de manera poco frecuente. Si el comportamiento de no realizar alguna transacción en la primera etapa del día es algo cotidiano, sea la causa que sea, simplemente no se podrá obtener paradero de bajada asociado al origen del viaje.



## 4. Multipago.

### 4.1. Definición y antecedentes.

El concepto de multipago se basa en el hecho de que un usuario utilice su tarjeta bip! para realizar el pago de la tarifa propia y la de otro pasajero, ya sea para realizar un viaje conjunto de manera completa, parcial o de ruta completamente diferente. Esto, basándose en que el otro usuario no posee tarjeta de pago o no posee de saldo disponible para realizar la transacción, por lo que se ve obligado a valerse de la tarjeta del compañero de viaje para acceder al sistema.

A priori este comportamiento debería estimar mismo paradero de subida y bajada para la primera transacción, debido a que no hubo desplazamiento espacial de la tarjeta, y por ende del usuario. Sin embargo, lo que lo hace diferente al caso explicado en el capítulo 4.5 es la diferencia temporal entre las transacciones. Para el caso de multipago estas transacciones son realizadas de manera consecutiva con una diferencia temporal pequeña, la que dependiendo del comportamiento tiene un tiempo límite asociado. Es por esto que actualmente el modelo no asigna bajada a este comportamiento, y lo diferencia con el caso de mismo paradero de subida y bajada.

Este comportamiento, del que casi no se tienen antecedentes de existencia en otros sistemas de transporte, ni menos estudios acabados de porcentaje de ocurrencia, se permite en Transantiago debido a múltiples factores. Entre ellos tenemos:

- La tarjeta bip! no es personal. Esto quiere decir que el uso de ella se lo puede dar quien la posea y de la forma que quiera. Sin embargo, la Tarjeta Nacional Estudiantil, utilizada por escolares y universitarios para el pago con tarifa rebajada es personal.
- No existen reglas asociadas a la prohibición del uso compartido de la tarjeta de pago en el momento de realizar la transacción.
- Con el actual sistema de tarjetas de prepago, al quedarse sin saldo al momento de validar en un lugar donde la carga de dinero a la tarjeta se hace complejo, se procede a este comportamiento en caso de que se viaje con algún acompañante que se quede sin saldo.

Sin embargo, del punto de vista económico no es conveniente realizar este comportamiento, pues al realizar dos validaciones en el mismo validador el sistema de pago no lo reconoce como un transbordo y no se aplica el proceso de tarifa integrada, cobrando para cada validación tarifa completa. Si este comportamiento se repite en cada una de las etapas del viaje el descuento de la tarifa de la tarjeta se vuelve a realizar.

Actualmente, el modelo clasifica estas etapas de viaje como multipago cuando existen dos o más validaciones consecutivas en un lapso de tiempo menor que 180 segundos (3 minutos). Esto, pues este tiempo entre validaciones permite suponer que no existió una etapa de viaje entre ambas validaciones consecutivas, sino que ocurrió algún comportamiento que debe ser analizado. Se toma en consideración que se cuenta como validación multipago aquella precedente a la validación final del

comportamiento, es decir, en caso de que haya una validación y en menos de 180 segundos ocurra otra validación, se considera la primera como multipago. Si la segunda precede a una tercera que también es hecha antes de 180 segundos entonces la segunda validación también se considera como multipago, y así sucesivamente hasta que exista entre dos validaciones consecutivas una distancia temporal mayor a 180 segundos. En el modelo las transacciones consideradas como multipago no registran estimación de paradero de bajada.

Para entender esta forma de pago y su influencia en el modelo de estimación de paradero de bajada, se estudiarán dos casos. En primer lugar aquel referido a viajes grupales, es decir, aquellos viajes compartidos a lo largo de la ruta, ya sean compartidos total o parcialmente en cada una de las etapas. Y el segundo caso que hace referencia al pago múltiple en sitios cercanos, caso en el que se comparte la tarjeta pero para realizar dos viajes completamente distintos a partir de transacciones realizadas en sitios cercanos.

En cada uno de estos casos se analizará lo que actualmente estima el modelo como bajada, se realizará un análisis estadístico para definir patrones de tiempo y distancia entre validaciones consecutivas que definan este proceso, luego se procederá a un análisis de los casos existentes, en el que se estudiará la frecuencia en la ocurrencia de cada uno de ellos y explicaciones que den a conocer antecedentes más detallados, para finalmente proponer soluciones a aquellas problemáticas encontradas que puedan ser arregladas.

## **4.2. Viajes Grupales.**

### **4.2.1. Validaciones Grupales.**

La condición de validación grupal corresponde a la característica de viaje en la que dos o más usuarios se trasladan en el sistema de transporte público de Santiago de manera conjunta. Este traslado bien puede ser en una etapa dentro de un viaje, en el viaje completo, o incluso desarrollarse en viajes ida y vuelta durante el día o la semana.

El suponer que dos o más personas se mueven en grupos es una tarea compleja, pues como se dice al principio, esta tarea se basa en suponer un comportamiento grupal de viaje en base a las validaciones de la tarjeta bip! y no necesariamente es lo que sucede. Pero bajo ciertos parámetros y condiciones, es lógico suponer el viaje grupal.

La principal necesidad de registrar este comportamiento es generar, si es posible, una estimación de bajada basado en el modelo de asignación de paraderos de bajada. Si bien en lo general el modelo de estimación de bajada es el mismo, hay ciertos detalles que deben ser corregidos para asignar de buena manera este comportamiento.

Las validaciones grupales, paso previo para encontrar etapas y viajes grupales, son definidas en base a un grupo de condiciones establecidas que sugieren un comportamiento grupal a la hora de viajar. Lo primero es definir la base de este comportamiento: son validaciones hechas por la misma tarjeta bip! (Misma ID). Esto,

pues el viaje grupal que definiremos cae en el concepto de multipago, que corresponde a validaciones consecutivas en períodos cortos de tiempo. Estas validaciones, al estar tan juntas temporalmente, no dan cabida a establecerlas como un viaje, sino como algún comportamiento anómalo que se debe estudiar.

La siguiente condición que define la validación grupal es que las validaciones consecutivas sean hechas en el mismo lugar. Esto con el fin de que se asuma que el grupo de usuarios viajan juntos y deben acceder a los mismos medios de transporte. Para cada tipo de sitio de validación (arriba del bus, en Metro o en Zona Paga) las condiciones si bien son similares poseen ciertas características particulares.

- Para el caso de validación arriba del bus, se considera mismo sitio cuando las validaciones consecutivas son realizadas en el mismo bus, caracterizado en la base de datos de estimación de bajada por la patente del mismo y el servicio-sentido correspondiente.
- Para el caso del Metro, las validaciones deben ser realizadas en la misma estación de Metro.
- Para el caso de Zonas Pagas, se requiere que las validaciones sean hechas en el mismo paradero, identificado por las coordenadas geográficas. Esto, pues puede haber más de un validador en la zona paga, y las transacciones si bien son hechas en distintos validadores del mismo paradero, son parte de un movimiento conjunto de usuarios.

Finalmente, se debe definir la condición temporal para asignar validaciones grupales. Esto no es posible de suponer mediante la lógica o la inspección visual, por lo que se realiza un análisis estadístico que permita definir la cota superior de tiempo que defina una validación grupal.

Para la obtención de este valor, se toman en consideración todos los pares de validaciones consecutivas que cumplan las condiciones nombradas recientemente para cada uno de los casos de sitio de validación (Bus, Metro, Zona Paga), además de establecer un rango temporal entre validaciones menor a 30 minutos (1.800 segundos). Este valor supone un rango amplio para el análisis y además corresponde al rango de tiempo propuesto para diferenciar un transbordo entre etapas de un viaje de una actividad (Munizaga y Palma, 2012). Si bien se podrían haber tomado otros valores bajo otros criterios, lo importante es identificar dónde se concentran principalmente en el aspecto temporal estas validaciones grupales.

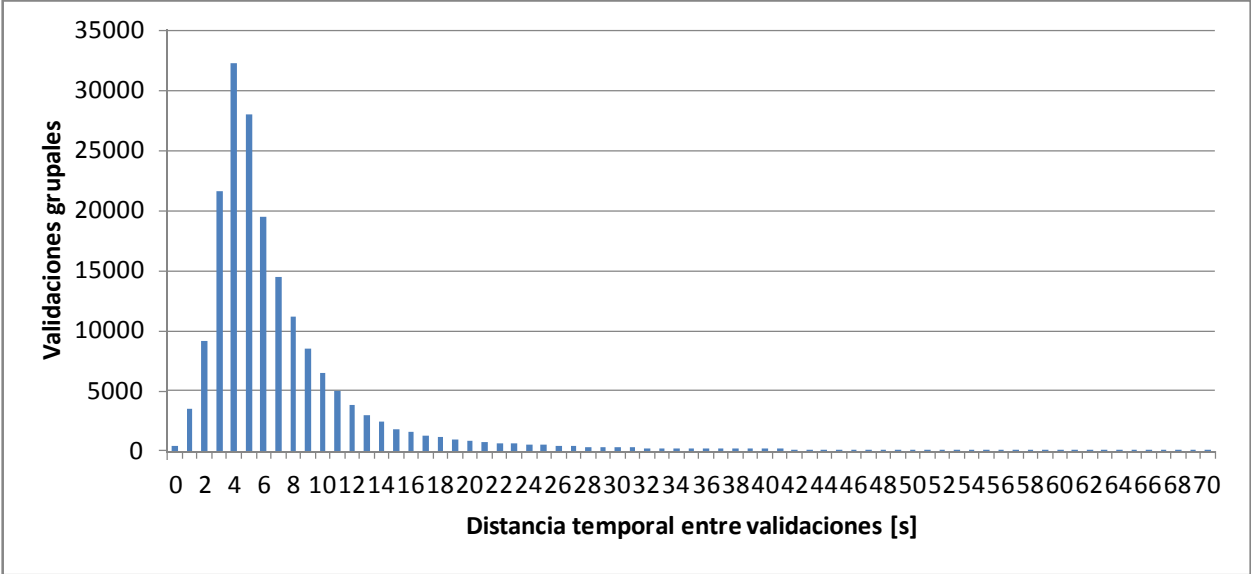
El trabajo a realizar consiste en catastrar todos aquellos pares de validaciones que cumplan los requisitos anteriores, y revisar tanto gráfica como estadísticamente los tiempos obtenidos, para definir un criterio formal.

Al filtrar todas estas condiciones dentro de la base de datos de estimación de bajada, se obtiene 394.834 validaciones grupales en las cuales se cumplen los requisitos establecidos anteriormente. Para estudiar de manera más detallada este comportamiento se analiza por separado aquellas hechas de manera consecutiva en Metro y en buses.

Para el caso de validaciones hechas en Metro el gráfico 4-1 da a conocer la cantidad de transacciones que cumplen con el requisito de validación grupal, que en total son 197.959, clasificadas por segundos entre transacciones. Con el fin de presentar un gráfico que muestra el comportamiento temporal de forma que sea entendible visualmente, se dan a conocer las cantidades de validaciones grupales para diferencias temporales de hasta 70 segundos, que corresponden al 94,8% del total de series de validaciones grupales hechas en Metro.

En el gráfico se aprecia que la mayoría de las validaciones grupales ocurre con distancias entre validaciones menores a 30 segundos, lo que parece razonable pensando en que el pago mediante Smart Card ocupa un lapso pequeño de tiempo. De hecho, los valores máximos se alcanzan entre los 3 y 7 segundos. A partir de los 40 segundos ya se denota una estabilización en la cantidad de validaciones grupales, que se mantiene para el tramo que continúa hasta los 1.800 segundos, plazo definido previamente como límite de estudio de validaciones grupales. Las validaciones grupales hechas en Metro con tiempo entre validaciones menor o igual a 40 segundos corresponden al 92,93% de las validaciones grupales con distancia temporal menor o igual a 1.800 segundos.

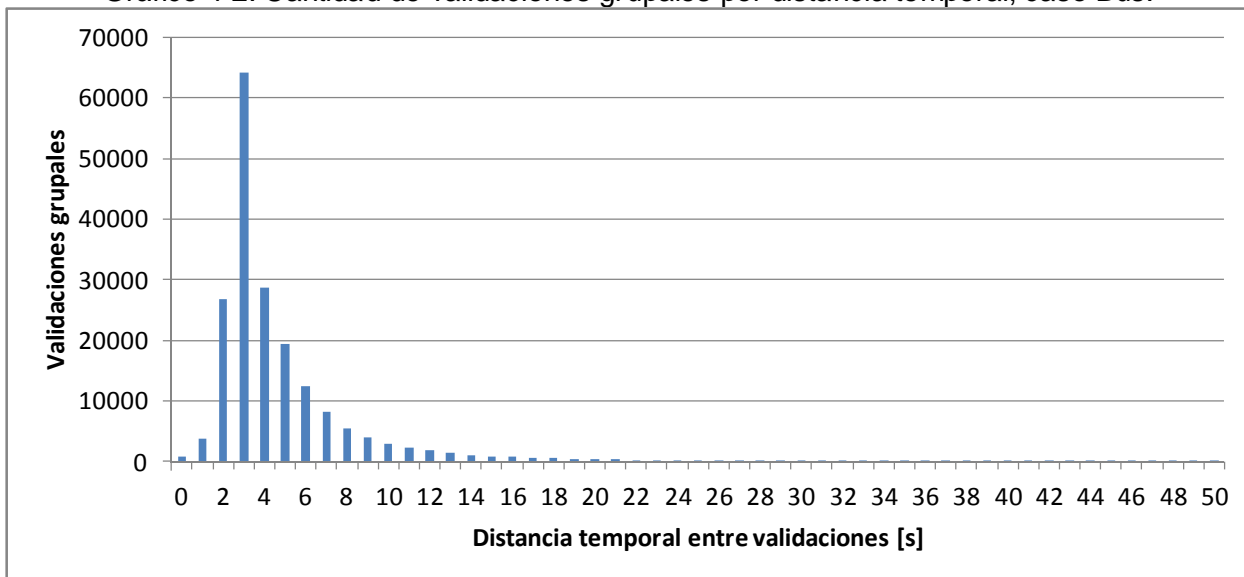
Gráfico 4-1: Cantidad de validaciones grupales por distancia temporal, caso Metro.



Para el caso de aquellas validaciones consecutivas realizadas en bus, entre las que se incluyen aquellas realizadas en los validadores al interior de los buses como en zonas pagas, el gráfico 4-2 da a conocer la distribución de diferencias temporales de las 196.875 validaciones grupales. Al igual que para el gráfico 4-1 del Metro, en este caso con motivo de una mejor visualización del comportamiento se acotará el rango temporal a mostrar entre 0 y 50 segundos.

Para el caso de bus las diferencias temporales entre transacciones son menores, y se concentran principalmente entre los 0 y 20 segundos, con una moda de 3 segundos. El porcentaje de validaciones grupales que caen dentro del rango menor o igual a 20 segundos corresponden al 95,62% del total de la muestra estudiada (0 a 1.800 segundos).

Gráfico 4-2: Cantidad de validaciones grupales por distancia temporal, caso Bus.



Como una forma de unificar criterios para ambos modos de viaje, y facilitar la labor de identificación de validaciones grupales, se propone como límite temporal para definir una validación grupal 40 segundos. Este límite implica que del total de validaciones consecutivas estudiadas que presentan un rango temporal entre 0 y 1.800 segundos, se considerarán como validaciones grupales 375.674, que corresponde al 95,15% del total estudiado.

En la tabla 4-1 se resume el total de validaciones grupales tanto para modo Metro como para modo Bus. En él se aprecia que se logra obtener una mayor participación de las validaciones consecutivas realizadas en buses que en Metro. La causa de esta diferencia con el Metro es la presencia de torniquetes en este último que retrasa la segunda validación hasta que pase el primer usuario, lo que obliga a ambos a validar por separado, mientras que en los buses una sola persona hace la validación para él y su o sus acompañantes.

Este análisis de distancia temporal posiblemente deje algunos casos especiales fuera de estudio, en el que entre validaciones ocurra alguna situación no observada por el modelo. Sin embargo, concentra parte importante del comportamiento y la forma de pago grupal esperada, en que las validaciones ocurran en un lapso corto de tiempo.

Tabla 4-1: Porcentaje acumulado de ocurrencia de viaje grupal por tiempo entre transacciones.

MODO	TIEMPO = 40 seg.	
	Cantidad	Porcentaje [%]
<b>Metro</b>	183.969	92,93
<b>Bus</b>	191.705	97,37
<b>Total</b>	375.674	95,15

A continuación, una vez obtenidas todas las validaciones grupales dentro de la semana de muestra, se dan a conocer datos acerca de las validaciones que conlleva a entender la causa de este tipo de comportamiento entre los usuarios.

De un total de 38.118.949 validaciones registradas en la semana de estudio, se tiene que 736.356 validaciones (1,93% del total de validaciones) caen en este rango en el cual participan 243.044 tarjetas ID diferentes que representan el 7,01% del universo total de tarjetas utilizadas (3.466.848 tarjetas bip!).

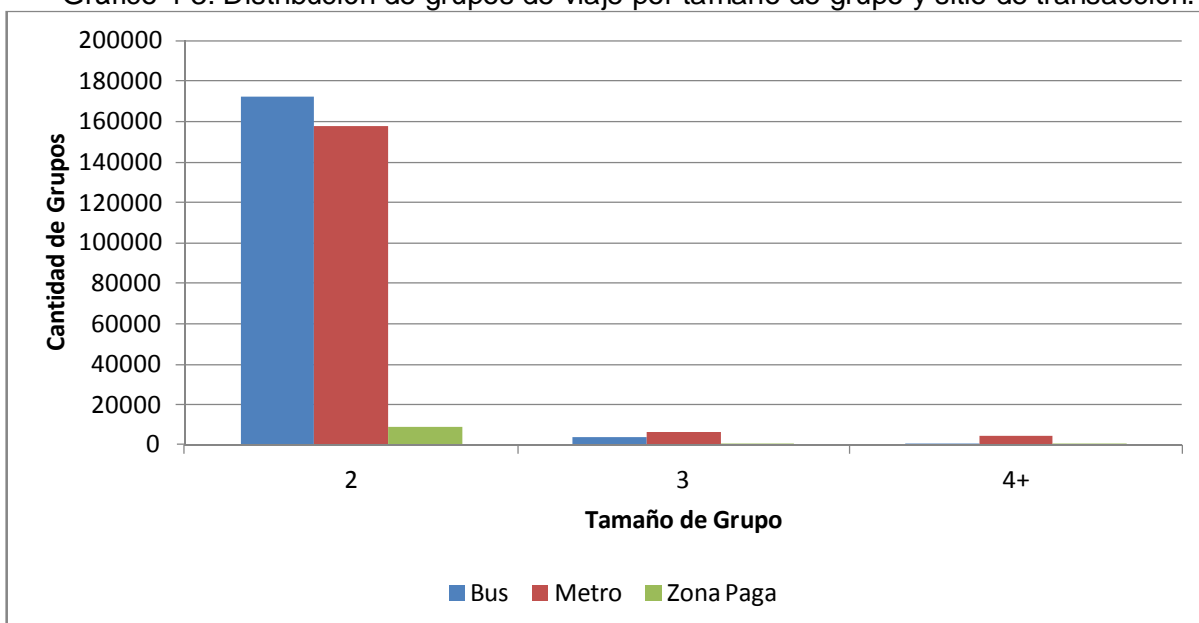
El siguiente análisis tiene por objetivo conocer los tamaños de los grupos por lugar de validación. Para esto se observa en la tabla 5-2 la distribución de tamaños de estos grupos para los casos de transacciones realizadas en bus, zonas paga y Metro. El promedio obtenido a partir de la tabla 4-2 da a entender que en Metro los grupos de validaciones tienden a ser un poco mayores que en buses y zonas paga. Este análisis será desarrollado posteriormente.

Tabla 4-2: Transacciones y validaciones grupales.

Modo	Transacciones Grupales	Etapas Grupales	Tamaño Promedio Grupo
Metro	357.970	169.588	2,111
Bus	359.460	176.892	2,032
Zona Paga	18.926	9.263	2,043
<b>Total</b>	<b>736.356</b>	<b>355.743</b>	<b>2,070</b>

La distribución de tamaños de transacción grupal, basado en la cantidad de validaciones grupales por sitio se puede apreciar en el gráfico 4-3. En éste se aprecia que mayoritariamente los grupos de dos validaciones son los que predominan en este comportamiento (95,36% del total de transacciones grupales). Sin embargo los grupos de tres (3,04%) y mayores de tres validaciones (1,60%) no son despreciables y deben ser estudiados para su análisis. Otro análisis que se realiza y se complementa con lo presentado en la tabla 4-2 corresponde a la participación que tienen las validaciones grupales hechas en Metro para grupos mayores de 2 validaciones, donde es mayoritaria su aparición en el tren urbano por sobre las validaciones hechas en buses y zonas paga.

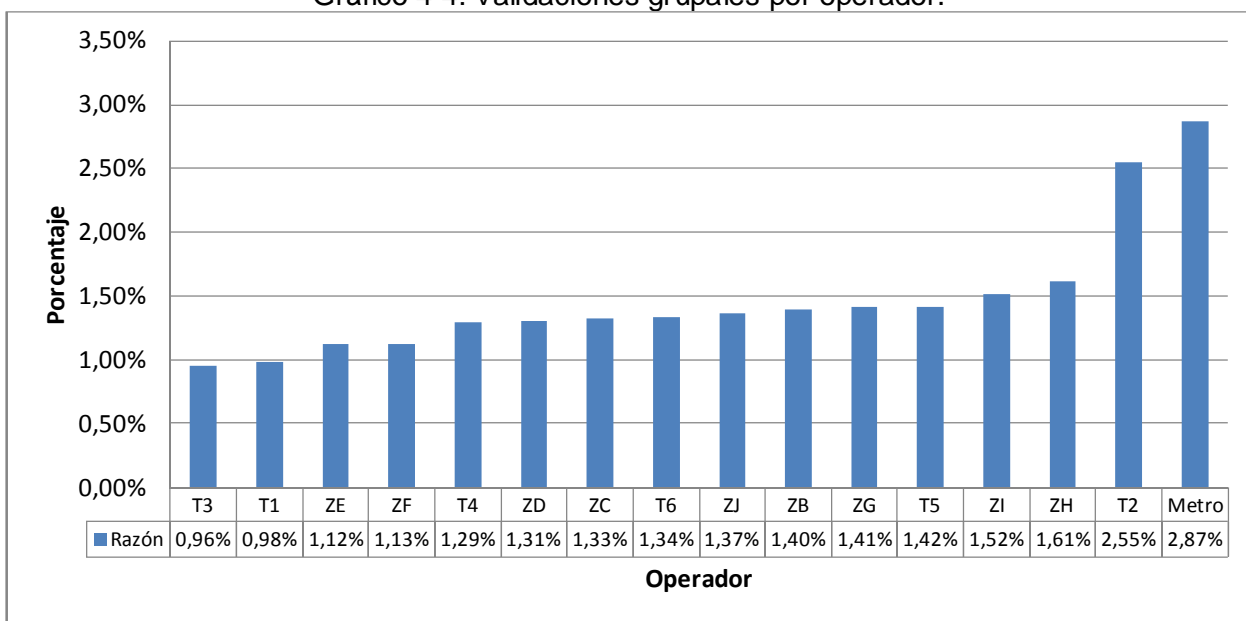
Gráfico 4-3: Distribución de grupos de viaje por tamaño de grupo y sitio de transacción.



Desde el punto de vista del sitio donde se realizan las transacciones, en Metro el 2,87% de las validaciones efectuadas (357.970) se comportan como grupales. Para el caso de las validaciones grupales efectuadas en Bus (359.460) éstas corresponden al 1,52% del total de validaciones en buses, y en Zonas Pagas (18.926) son el 0,93% de ellas.

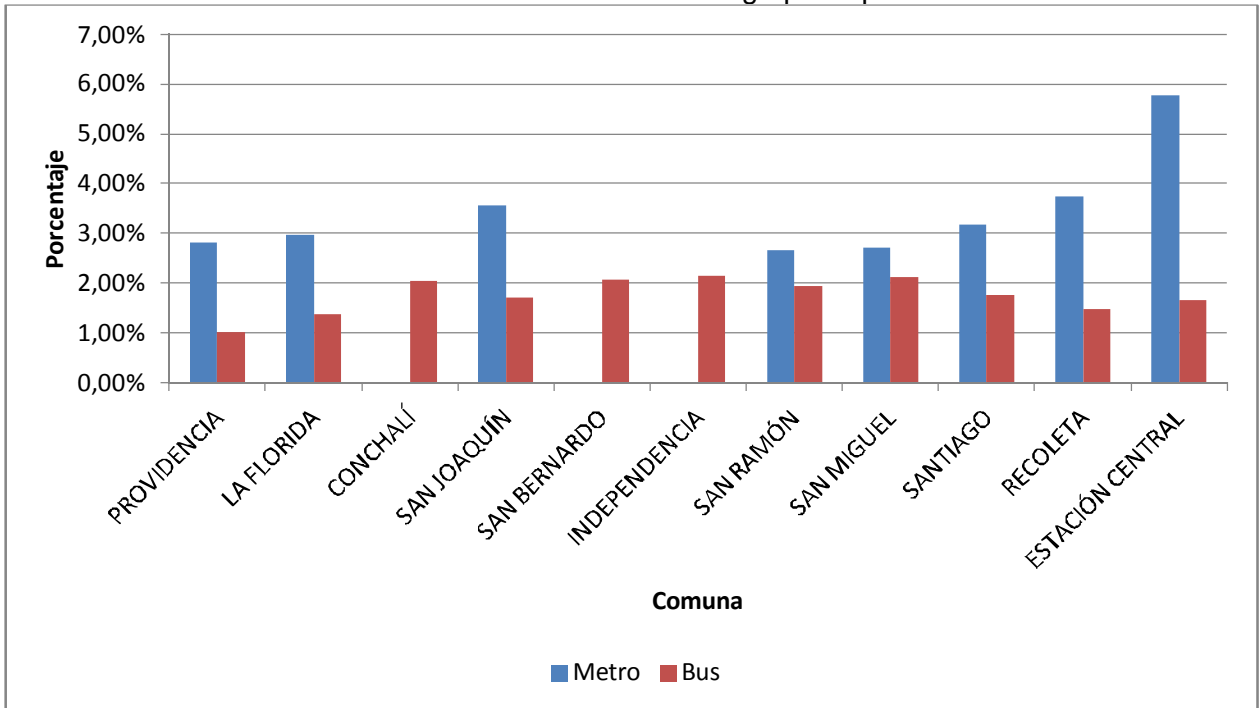
Un análisis más acabado tiene relación con el observado en la cantidad de transacciones grupales realizadas por operador, el cual se da a conocer en el gráfico 4-4. De este gráfico se denota una participación mayor de viajes grupales tanto en el Metro como en el Troncal 2 por sobre los demás operadores. Esto es posible que suceda debido a la existencia de torniquetes, que impiden que alguien que no tenga tarjeta bip! o no tenga saldo en ella se vea obligado a que su acompañante valide por él. Y el hecho de que Metro tenga un valor mayor que el troncal 2 se podría deber a que la evasión de torniquetes en los buses es mayor que en Metro, debido a la poca fiscalización existente, por lo que el acompañante prefiere omitir el pago y de alguna forma evadir el torniquete.

Gráfico 4-4: Validaciones grupales por operador.



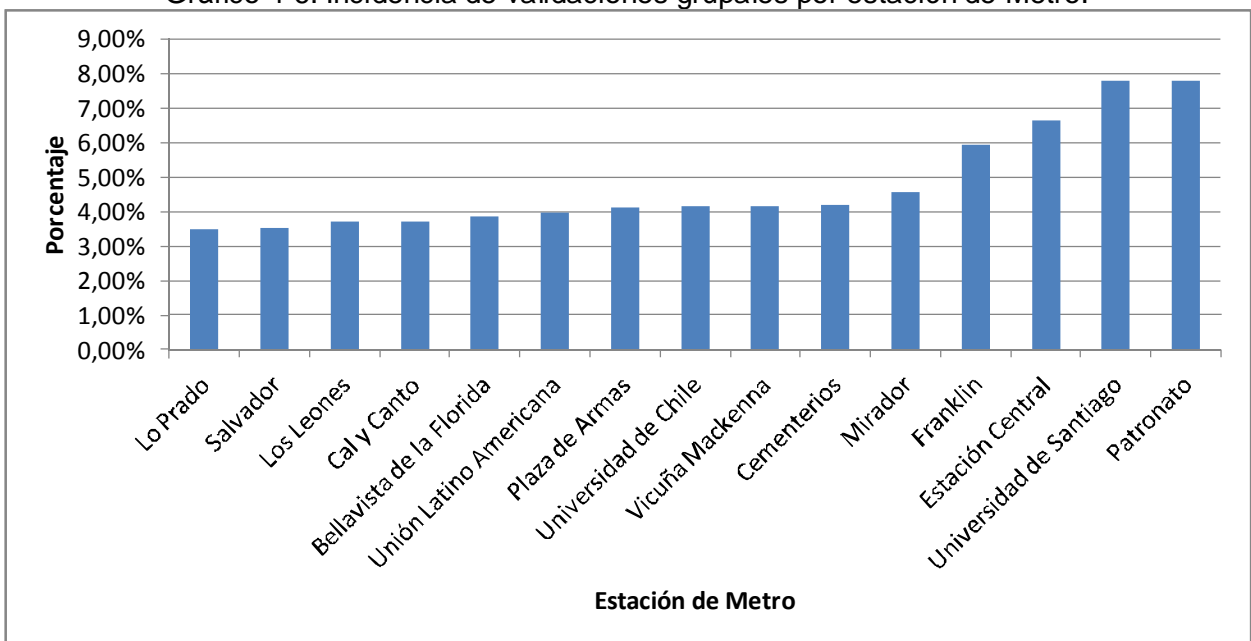
Otro análisis a realizar corresponde a observar este comportamiento por Comuna-Subida. En el gráfico 4-5 se aprecia el comportamiento de validaciones grupales respecto al total de validaciones en las 12 comunas con mayor incidencia de estos viajes. Se denota que este comportamiento ocurre principalmente en comunas céntricas de la capital.

Gráfico 4-5: Incidencia de validaciones grupales por comuna.



Este dato en sí no genera mucha información, pero al complementarlo con el gráfico 4-6 relacionado con las estaciones de Metro que presentan mayor incidencia de validaciones grupales se tienen más antecedentes al respecto de este comportamiento. En él se puede apreciar que las 15 estaciones de Metro (y servicios) con mayor incidencia de este comportamiento corresponden a zonas comerciales y de terminales de buses. Son lugares de atracción tanto comercial (Patronato, Franklin, Mirador, etc.) como turístico (Cementerios, Plaza de Armas, etc.).

Gráfico 4-6: Incidencia de validaciones grupales por estación de Metro.





La distribución de este comportamiento por tipo de día y períodos es la que se muestra en los gráficos 4-7 y 4-8, en los que se denota que principalmente los períodos donde ocurre este comportamiento se presentan en el fin de semana, y principalmente en horas de la mañana y la tarde.

Gráfico 4-7: Incidencia de validaciones grupales por tipo de día.

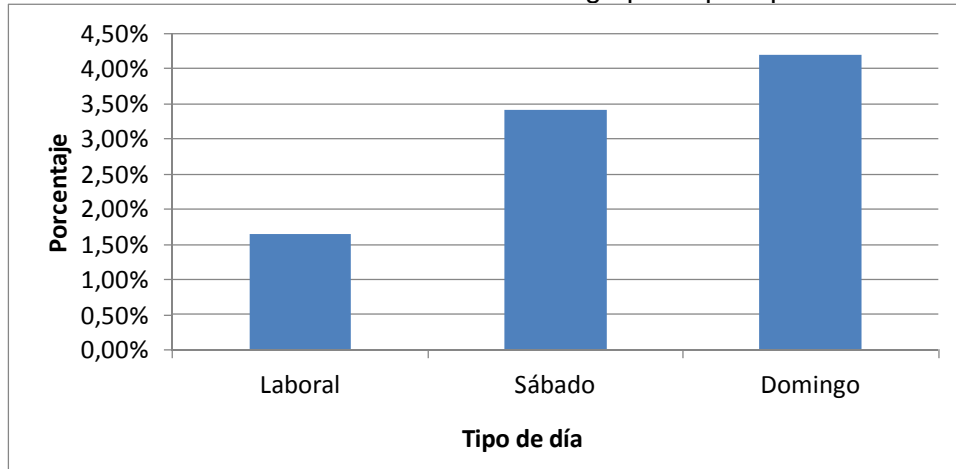
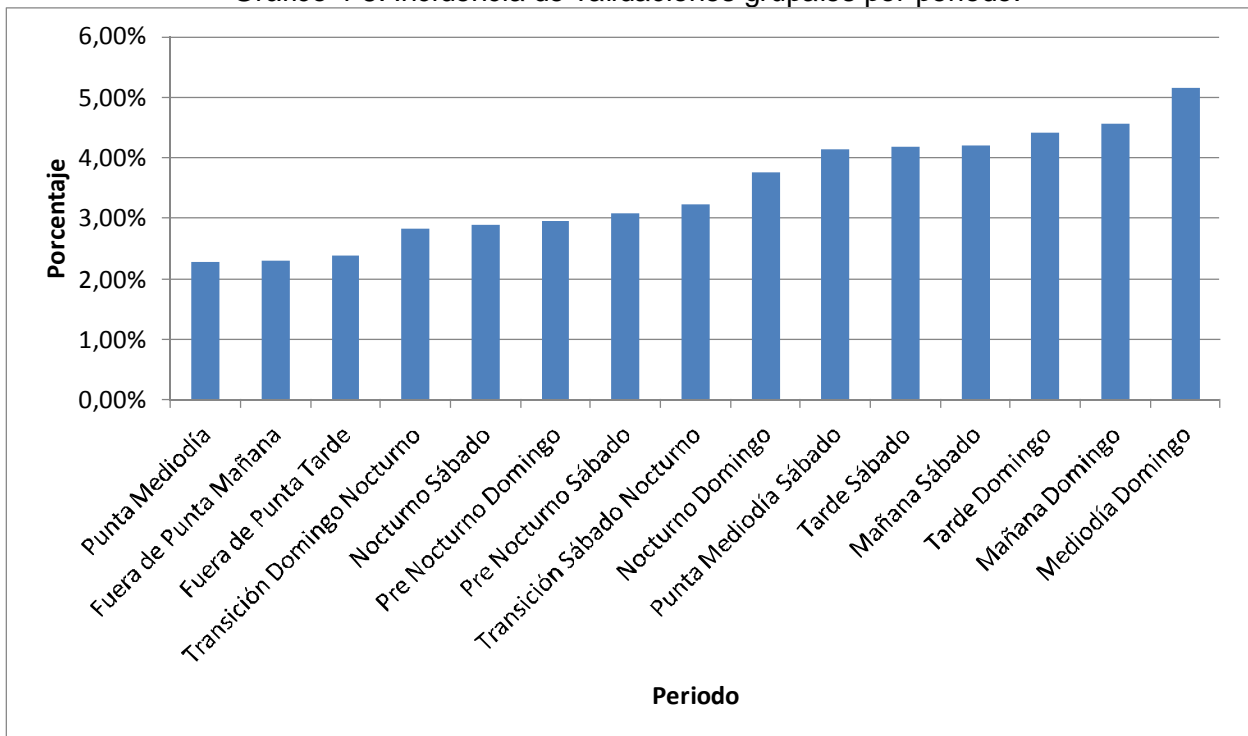


Gráfico 4-8: Incidencia de validaciones grupales por periodo.



Con esta información, se infiere que las validaciones grupales, asociadas a viajes en grupos corresponden a actividades relacionadas con el ocio, pues no ocurren en horas punta en la semana, sino que en periodos donde la gente principalmente se dedica a realizar actividades como compras, paseos familiares, salidas a lugares de esparcimiento tanto diurno como nocturno, etc., y complementando esta información con los lugares dónde se realizan las validaciones, y por ende las zonas donde se desarrollan tales actividades, se complementa lo propuesto. Además, otra inferencia

que se realiza es que gente que viene de fuera de Santiago realiza este comportamiento de manera importante, pues zonas de terminales de buses y trenes presentan una alta incidencia, asociado a la necesidad de viajar dentro de Santiago en familia o en grupo sin comprar tarjetas bip! para cada integrante.

#### **4.2.2. Etapas y Viajes Grupales.**

Una vez analizados los comportamientos que originan estas transacciones grupales, se busca estudiar cómo se mueven estos grupos de usuarios en el sistema de transporte público, estudiando además si estos grupos mantienen su cantidad de validaciones y las causas que pueden llevar a cambios en las cantidades de validaciones.

En primer lugar, se dan a conocer las definiciones de los conceptos de etapa grupal y viaje grupal:

- Etapa grupal: corresponde a una transacción grupal que es antecedida o sucedida por otra transacción, sea grupal o no, con una diferencia temporal entre la bajada y la siguiente transacción menor a 30 minutos.
- Viaje grupal: conjunto de etapas grupales que tienen como fin el momento en que los usuarios realizan una actividad al bajarse del último modo de viaje. Esto queda definido por la existencia de una brecha temporal mayor a 30 minutos entre la bajada estimada que da fin a la etapa y la siguiente transacción realizada, sea esta grupal o no.

La forma en que se mueven las personas en grupo es algo muy complejo de analizar. Esto debido a que no hay un patrón de comportamiento común en lo que se refiere a cantidad de validaciones por etapa. Además, el factor de la evasión impide tener una estructura completa de todos los viajes grupales existentes. Es por esto que para analizar este comportamiento se toma una muestra de tarjetas ID que cumplen la condición de tener al menos una transacción grupal dentro de sus transacciones semanales.

Debido a la complejidad de automatizar la búsqueda de viajes grupales dentro del total de transacciones grupales existentes, se toma una muestra con el fin de realizar este proceso inspectivo de manera manual.

Estos viajes grupales se identifican siguiendo los parámetros establecidos anteriormente, y para un mejor entendimiento de éstos se consideran como parte del viaje grupal aquellas etapas con una sola validación pero que cumplan los criterios temporales y espaciales descritos, pues se necesita entender el comportamiento completo del viaje y no sólo las etapas en las que se viaja en grupo.

Para obtener información acerca de la forma de viaje de estas transacciones grupales, se tomaron en total datos de viajes grupales de 110 tarjetas ID, las cuales presentan en total 223 viajes grupales, lo que da un promedio de 2,03 viajes grupales por tarjeta ID.

Una primera información tiene que ver con la cantidad de viajes grupales por tarjeta ID en la semana, valores que aparecen en la tabla 4-3. Se infiere de esta tabla que la mayoría de las tarjetas (77,27% del total de la muestra) realiza a lo más dos viajes grupales por semana, generando la idea que no se trata de un comportamiento común en el día a día de los usuarios sino que se trata generalmente de casos aislados.

Tabla 4-3: Total viajes grupales por tarjeta ID en semana de estudio.

<b>Viajes por ID</b>	<b>Total</b>	<b>Porcentaje [%]</b>
1	59	53,64
2	26	23,64
3	12	10,91
4	5	4,55
5	2	1,82
6	2	1,82
7	1	0,91
8	1	0,91
9	1	0,91
10	1	0,91

En la tabla 4-4 se presenta el análisis de viajes grupales por día del total de la muestra de 223 viajes grupales. Nuevamente se aprecia que el realizar un viaje grupal por día es el comportamiento más habitual, pues el 47,98% de los viajes grupales son únicos en el día. Esto conlleva a suponer que los viajes grupales de ida y vuelta no son el común de comportamiento de viaje grupal. El caso de 2 viajes grupales en el día (ida-vuelta) aparece 33 veces en esta muestra, lo que representa el 29,6% del total de viajes de la muestra (66 viajes).

Tabla 4-4: Viajes grupales por día.

<b>Viajes Grupales por día</b>	<b>Frecuencia</b>	<b>Total Viajes</b>	<b>Porcentaje [%]</b>
1	107	107	47,98
2	33	66	29,60
3	10	30	13,45
4	5	20	8,97

Finalmente, un último análisis referido a esta muestra de tarjetas tiene relación con el número de etapas que tienen los viajes grupales estudiados. Para esto en la tabla 4-5 se da a conocer una tabla comparativa que describe la cantidad de viajes desagregados por cantidad de etapas para la muestra de tarjetas y para el total de validaciones de la semana de estudio.

Se puede apreciar que para el comportamiento grupal predominan los viajes de una etapa, generando la idea de que este tipo de viajes generalmente se realiza para viajes simples sin mucha combinación de modos, debido a que el sistema de pago se complejiza cuando existen dos o más validaciones consecutivas y los usuarios ven perjudicado el beneficio de optar a pago cero para las combinaciones a realizar.

Tabla 4-5: Comparación de número de etapas entre muestra de viajes grupales y total de viajes de semana de estudio.

Etapas	Muestra tarjetas		Total semana estudio	
	Cantidad Viajes	Porcentaje [%]	Cantidad Viajes	Porcentaje [%]
1	137	61,43	11.853.539	50,74
2	68	30,49	8.677.454	37,14
3	16	7,17	2.421.224	10,36
4	2	0,90	407.977	1,75
+4	0	0,00	2790	0,01

Posterior a este análisis de la muestra, un nuevo estudio tiene por finalidad observar el comportamiento de los grupos de validaciones a lo largo de un viaje durante el día. Previamente es importante diferenciar tres tipos de viaje, dependiendo del número de validaciones: viaje grupal completo, donde se mantiene el tamaño del grupo durante todo el viaje, viaje grupal parcial, donde el tamaño del grupo cambia durante el viaje (aumenta o disminuye) y viaje grupal con etapa siguiente individual, donde aparece posterior a las etapas grupales una individual que termina con el viaje grupal. La tabla 4-6 da a conocer este comportamiento, dividiendo el comportamiento en los tres tipos recientemente nombrados.

Este análisis, a diferencia del anterior realizado con una muestra de tarjetas, no presenta un viaje completo, sino que la parte del viaje en que el comportamiento es grupal.

La tabla no da a conocer los tamaños de los grupos, sino que la duración en etapas de los viajes grupales encontrados. Se aprecia en esta tabla que los viajes grupales de una etapa son los que predominan en el conjunto de viajes grupales (94,17% del total de viajes grupales). Además se da a conocer que 50.803 viajes presentan una siguiente etapa individual, es decir, que luego de la o las etapas grupales existe una etapa individual dentro del viaje.

Para el caso de viajes de dos o más etapas se incluye el dato acerca si los grupos se mantienen en cantidad de integrantes (completo) o no (parcial). Para el caso de los viajes grupales de dos etapas, de 18.696 identificados se tiene que 16.610 mantienen la cantidad de usuarios en la siguiente etapa, dejando a 831 viajes de dos etapas grupales con un número diferente de usuarios. Además, de los 18.696 viajes de dos etapas grupales encontrados en la semana de estudio, se tiene que 1.255 poseen una tercera etapa pero con validación individual.

Estas validaciones individuales posteriores afectan de manera importante el proceso de viaje grupal, pues cortan estos viajes y los convierten en individuales, salvo que se pueda suponer que en esa transacción individual el o los otros integrantes del grupo hayan evadido el pasaje pero continúen en viaje juntos. Si se considera el caso extremo, es decir que estos grupos desaparecen y se convierten en viajes individuales, los viajes completos grupales son 286.760.

Tabla 4-6: Cantidad de etapas grupales.

Etapas	Viaje grupal			Total Viajes
	Completo	Parcial	Etapas individuales posteriores	
5	2	0	-	2
4	32	2	5	39
3	730	35	47	812
2	16.610	831	1.255	18.696
1	266.254	-	49.496	315.750
<b>Total</b>	<b>283.628</b>	<b>868</b>	<b>50.803</b>	<b>335.299</b>

#### 4.2.3. Casos particulares de viajes grupales.

La forma en que los viajes grupales se mueven por Santiago es tan variada que hace muy difícil establecer una estructura de comportamiento que pueda analizar y estimar de manera correcta la bajada de cada uno de los integrantes del viaje.

Sin embargo, existen una serie de comportamientos de viajes que se presentan con mayor frecuencia y que es importante darlos a conocer para ver si existe o no posibilidad de estimar parcial o totalmente las bajadas en cada uno de los paraderos tanto para las etapas de cada viaje como para el destino del mismo.

A continuación se dan a conocer algunos casos y las posibilidades de estimación de bajada.

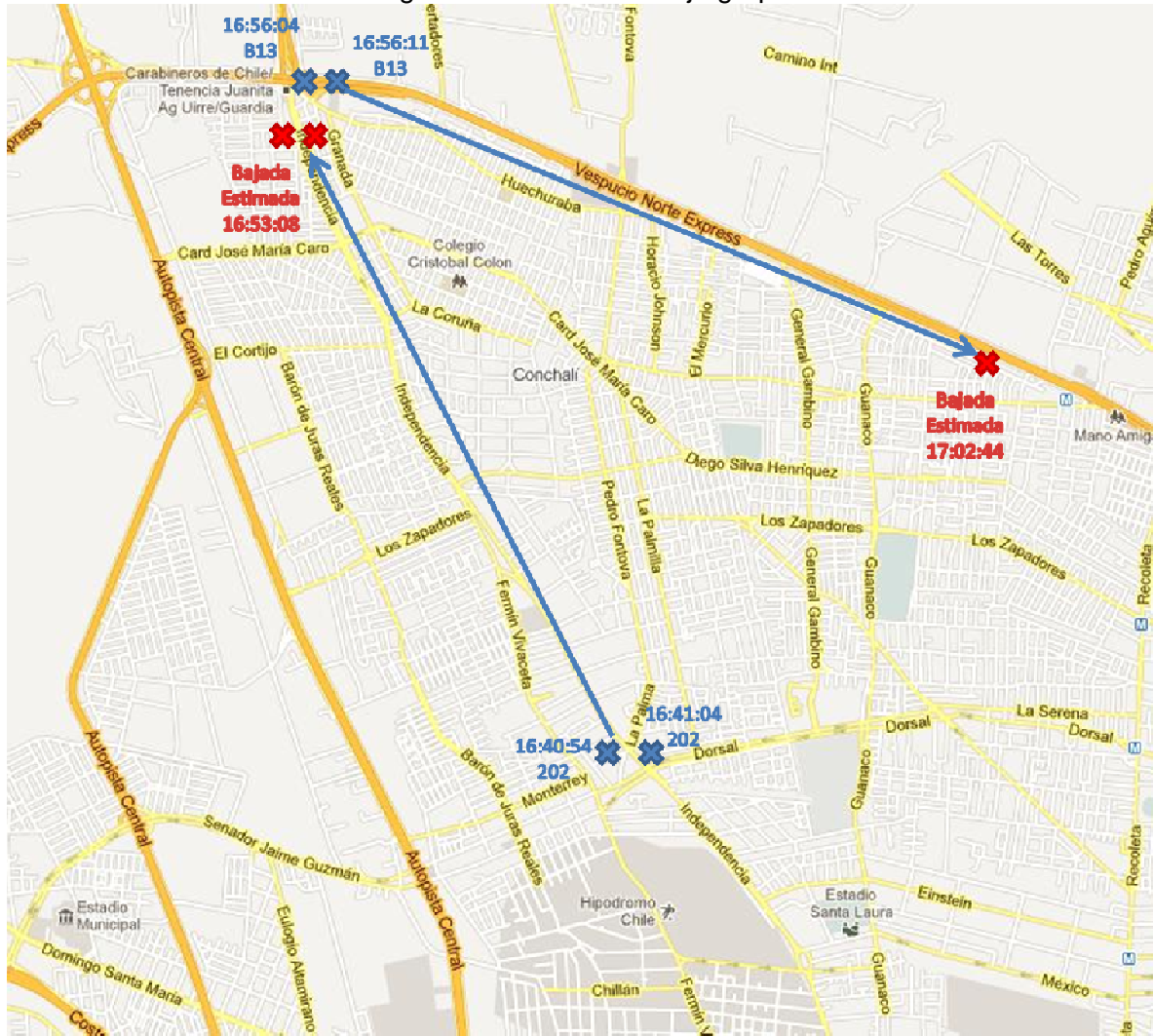
- Caso 1

Este viaje, que se aprecia en la imagen 4-1, corresponde a un viaje de dos etapas cuyo comportamiento corresponde a dos transacciones grupales de dos validaciones cada una caracterizadas por cruces azules, la primera en la zona de Dorsal con Independencia y la segunda en la zona de Vespucio Norte con Independencia. El modelo asume para la bajada de la primera etapa del viaje una sola bajada en un paradero cercano al de subida de la segunda etapa del viaje, y para la bajada de la segunda etapa del viaje asume sólo una bajada en la zona del Terminal Vespucio Norte (que corresponde a la subida de la primera validación de la tarjeta ID del día).

Para este viaje grupal es posible asumir que la bajada de la primera etapa del viaje, realizado en el recorrido 202, es la misma para ambos que en la imagen se presentan con dos cruces rojas, pues la siguiente transacción también es grupal. Además se puede asumir como paradero de bajada el calculado por el modelo, que si bien lo realiza para sólo una de las dos validaciones es posible replicarlo para la otra validación.

En lo que se refiere a la segunda etapa del viaje (recorrido B13), no es posible asumir la bajada de ambos en el mismo lugar, pues no hay forma de suponer que ambos bajaron juntos. Este caso en el que el destino final del viaje no sea posible suponerlo como bajada común es la principal dificultad a la hora de estimar bajadas en viajes grupales. Esta bajada no grupal se aprecia sola en el mapa con una cruz roja.

Imagen 4-1: Caso 1 de viaje grupal.



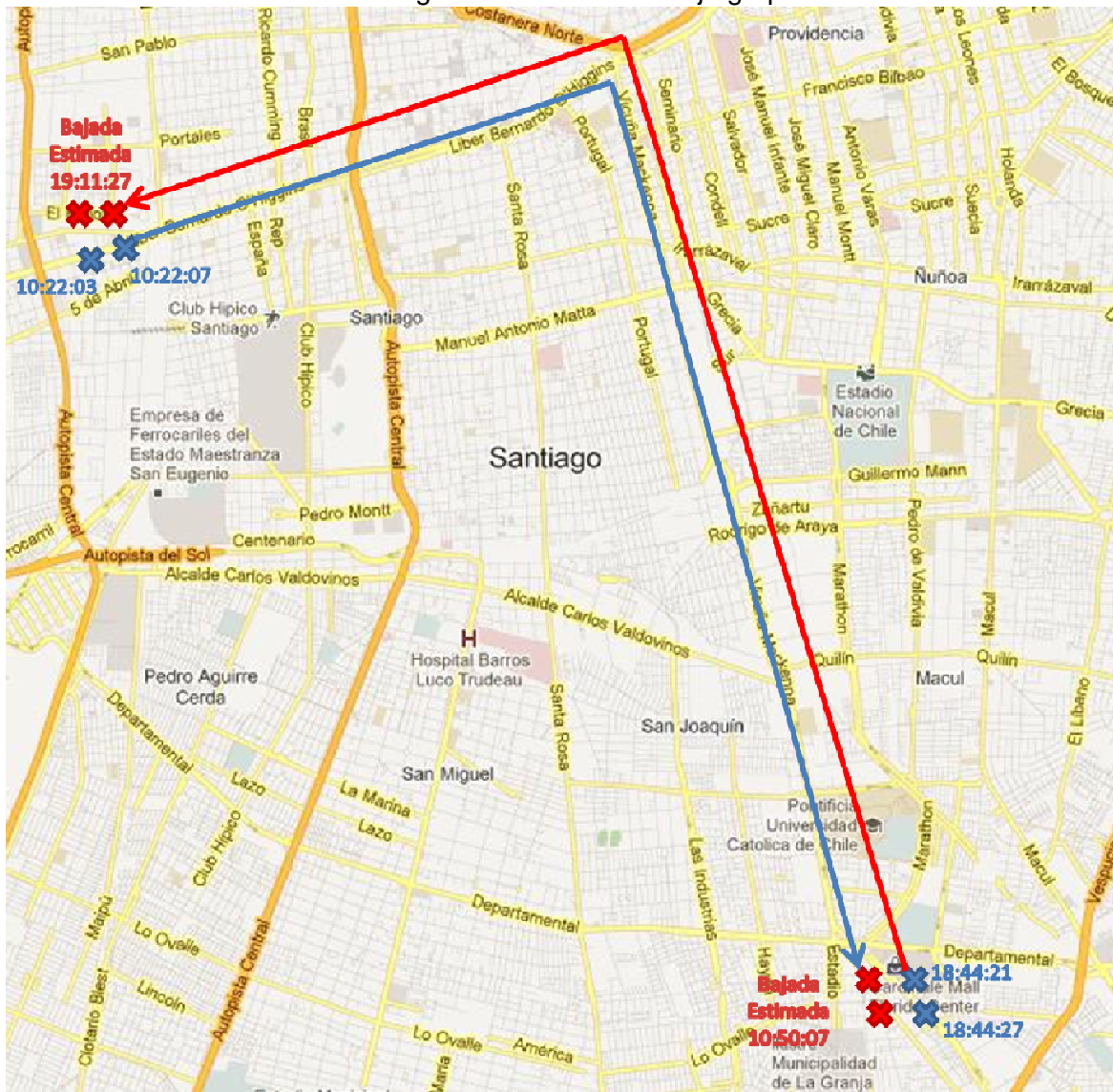
- Caso 2

El segundo ejemplo es un caso particular pero interesante, que se aprecia en la imagen 4-2. Corresponde a dos transacciones grupales de dos validaciones cada una, que además son las únicas validaciones que registra la tarjeta ID en el día y en la semana de estudio. La primera transacción grupal es realizada en el metro Universidad de Santiago y corresponde al primer viaje, que se muestra con una flecha azul. La segunda transacción grupal es realizada en el metro Mirador y corresponde al segundo viaje caracterizado con una flecha roja. Ambos pares de transacciones se aprecian con cruces azules en la figura.

Para el primer viaje el modelo estima como bajada la estación de metro Mirador, solo para una de las dos validaciones. Siendo que no hay otra transacción entremedio es posible suponer que el viaje tiene como bajada para ambas validaciones el metro Mirador. Para el segundo viaje del día el modelo estima como bajada para una de las validaciones el metro Universidad de Santiago (por ser la última validación del día asocia como bajada la primera del mismo día). Para este caso asumir que ambas validaciones tienen la misma bajada es más razonable que el caso anterior, pues no

existen otros viajes asociados a la tarjeta. Las bajadas grupales son caracterizadas por cruces rojos.

Imagen 4-2: Caso 2 de viaje grupal.

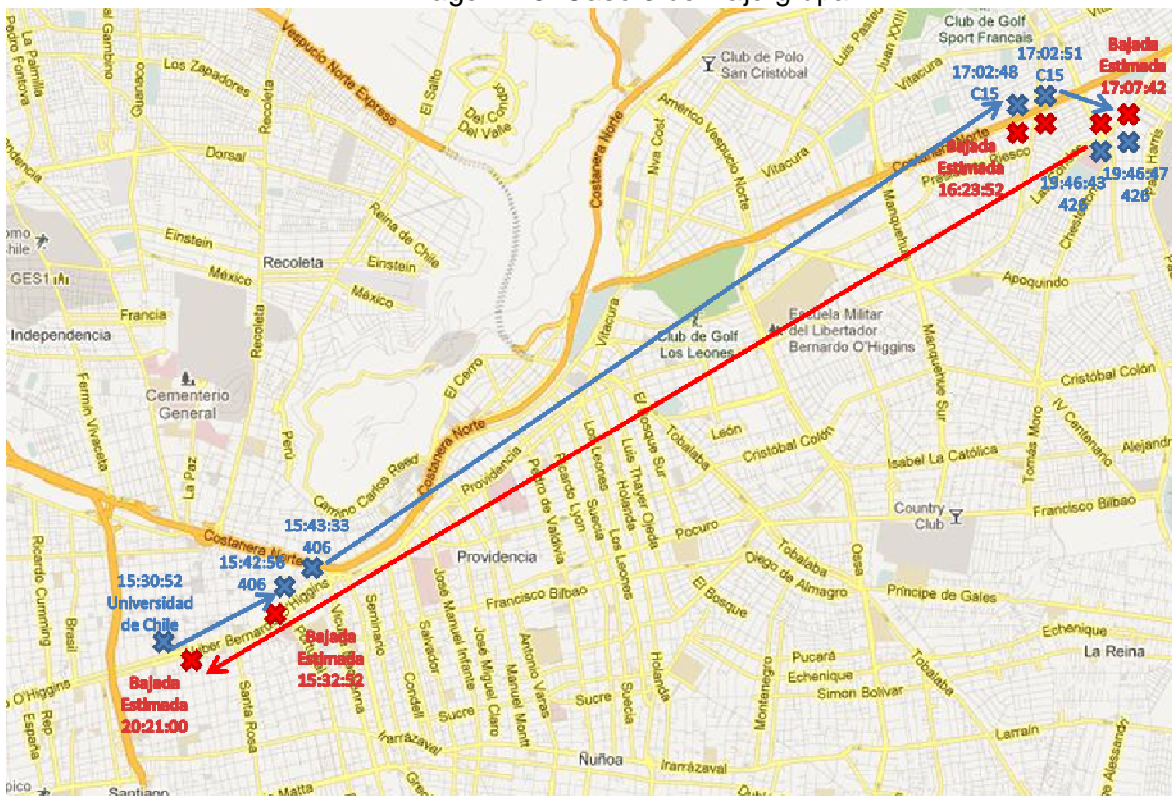


- Caso 3

En el caso 3, presentado en la imagen 4-3, se muestran todas las validaciones registradas por la tarjeta bip! durante el día. Salvo la primera validación en la estación de Metro Universidad de Chile, el resto son todas validaciones grupales. Todas ellas están representadas en la figura con cruces azules. El primer grupo de validaciones (Alameda) corresponde a un viaje grupal de dos etapas (realizados en metro y en bus de servicio 406). Las siguientes (en servicio C15 y 426), en la zona oriente, corresponden a dos viajes grupales por separado, siendo el último viaje de regreso al sector de Alameda. Salvo la última validación grupal, se puede obtener la bajada (cruces rojas) de todas las subidas. Pero, si suponemos que en la primera validación del día ambos integrantes empezaron el viaje, se puede cerrar el bucle de viaje. Pero como no hay información adicional que haga suponer que empezaron juntos el viaje, el

último viaje grupal del día no posee bajada estimada para ambos. Sí es posible estimar la bajada de uno de los integrantes en base a la primera validación del día.

Imagen 4-3: Caso 3 de viaje grupal



Un suceso importante de comentar en el caso 3, que si bien es una suposición pero tiene una lógica que se presenta en la realidad, es la diferencia temporal en la transacción grupal realizada en el servicio 406, que es cercana a los 40 segundos, muy superior al resto de diferencias temporales entre transacciones. Una posible causa de este comportamiento es que quien se unió al viaje no tenía su tarjeta bip! con saldo, por lo que al darse cuenta tuvo que recurrir a la tarjeta del acompañante para poder viajar durante el día. Como se dijo es una suposición, pero es parte del comportamiento de pago que se aprecia diariamente en Transantiago.

Estos casos dejan de manifiesto las múltiples posibilidades existentes de viaje grupal de los usuarios de Transantiago. El problema de cómo obtener la bajada entonces radica en entender el comportamiento de la tarjeta y a partir de aquello estimar la bajada del grupo o de parte de los integrantes del grupo según sea el caso.

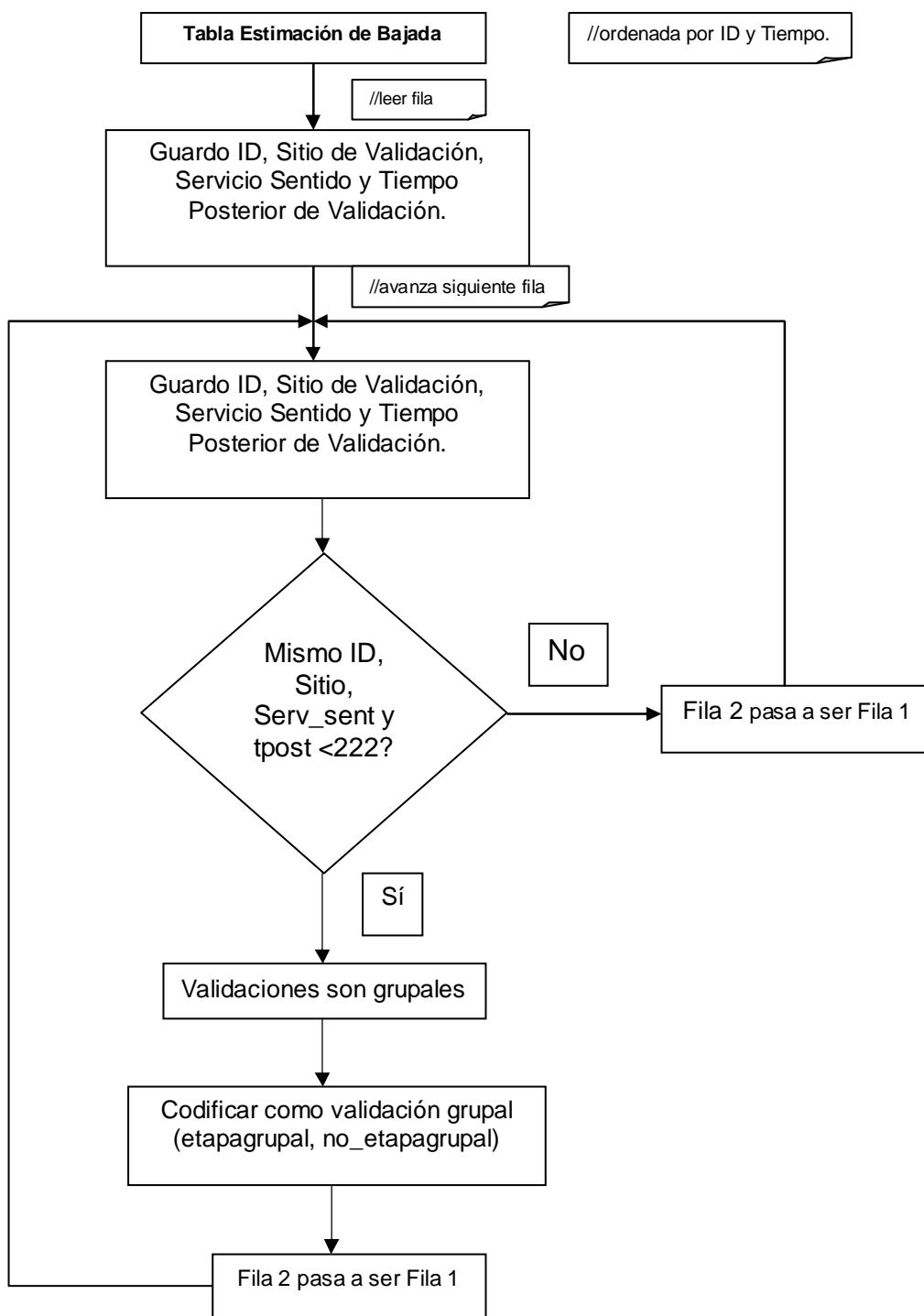
#### 4.2.4. Propuesta de solución.

La solución a este problema requiere de dos fases. En primer lugar clasificar aquellas transacciones que tienen la característica de grupal, y en segundo lugar estimar el comportamiento basándose en los fundamentos del modelo para estimar paradero de bajada, que dependiendo del caso sería para todos los integrantes del grupo de viaje, para parte de él o para ninguno.



La idea de agrupar las transacciones tiene por objetivo estudiar el traslado de la tarjeta, entendiendo por esto que las transacciones son las que definen el origen y el destino del viaje. Es por eso que se busca con este proceso entender el viaje de manera agregada, para luego en análisis posteriores que definan el uso de matrices de origen y destino para el conteo de viajes desagregar estos viajes dependiendo del tamaño del grupo. El proceso de agregación de transacciones grupales se da a conocer en la figura 4-1.

Figura 4-1: Esquema de identificación de transacción grupal.



Una vez identificada la transacción como grupal, se procede a definir si existe un traslado grupal que pueda ser definido como etapa grupal. Para esto simplemente se compara con la transacción siguiente:

- Si la transacción siguiente es grupal, y el tamaño del grupo es igual o mayor que la transacción inicial, se infiere que el grupo completo realizó el traslado por lo que se define como etapa o viaje grupal dependiendo del caso y se le estima paradero de bajada a todos los integrantes de la transacción inicial.
- Si la transacción siguiente es grupal, y el tamaño del grupo es menor que el de la transacción grupal inicial, implica que parte del grupo llegó a destino, y por ende sólo se estima paradero de bajada a la cantidad de pasajeros que hayan validado en la transacción siguiente.
- Si la transacción siguiente es simple (una sola), es un caso particular del punto anterior y quiere decir que sólo se le puede estimar bajada a uno de los integrantes del grupo.

Finalmente, se aplica el proceso conocido de estimación de paradero de bajada según sea el caso (bus, metro o zona paga), tomando en consideración que el viaje es uno, pero los usuarios que lo realizan son varios, por lo que es recomendable dejar especificado el tamaño del grupo de viaje en una columna de la base de datos para conteos posteriores de matrices origen-destino de viajes.

### **4.3. Multipago en sitios distintos.**

#### **4.3.1. Antecedentes y situación actual.**

Este caso corresponde a aquel en que se presentan dos o más validaciones consecutivas cuya distancia geográfica y temporal es pequeña. A diferencia del caso de viajes grupales, en este caso la validación se realiza en validadores distintos, por lo que no se puede suponer que el viaje realizado sea el mismo para cada una de las validaciones.

Las causas detrás de este comportamiento son complejas de discernir, pero se pueden suponer algunas causas basadas en el comportamiento de viaje de los usuarios:

- Compartir la tarjeta, es decir, prestar la tarjeta para que alguien realice la transacción en uno de los modos integrados de Transantiago y luego ocuparla en el viaje asociado al dueño de la tarjeta de pago.
- Equivocación en el pasajero al abordar un medio de transporte, debido a desconocimiento de la ruta del mismo. Esto suponiendo que valida en un sitio que tiene un recorrido que no le sirve para el propósito del viaje.
- Utilización del bus para recorridos cortos, por motivos de seguridad (trayectos complejos y oscuros) o de comodidad.

Actualmente, el modelo de estimación de paradero de bajada propone un sitio de bajada a estas transacciones, lo que puede generar un error al asumir como bajada un sitio cercano a la siguiente transacción, que se sabe que es cercano a la anterior, por lo

que propone un viaje que puede no existir en realidad (salvo el caso de que sí haga un viaje de muy corto trayecto). Esto genera que se estimen destinos de viajes erróneos que afectan los cálculos posteriores de destinos de viajes.

Con el fin de identificar este comportamiento, el análisis va enfocado a estudiar tres casos de multipago en sitios distintos, los que se presentan a continuación:

- Validación Bus-Bus: La tarjeta presenta dos validaciones consecutivas en sitios cercanos, ambas hechas en buses distintos.
- Validación Metro-Bus: La primera validación se realiza en una estación de Metro y la siguiente en un bus cercano a la estación.
- Validación Bus-Metro: Se valida inicialmente en un bus y consecutivamente en una estación de Metro próxima.

Un primer análisis tiene como objetivo definir qué distancia será definida como cercana y que tiempo será definido como acotado. Esto claramente dependerá del tipo de comportamiento a estudiar, pues existen parámetros distintos para cada uno como lo son tiempo de espera en paradero, ingreso o salida de una estación de Metro, etc. Para ello se busca un tiempo y distancia media de transbordo, es decir, lo que los usuarios del sistema demoran y caminan en promedio para cambiarse de un servicio a otro.

#### **4.3.2. Transacciones masivas.**

Un caso particular de este comportamiento, y que es observado en la base de datos, corresponde al caso de algunas tarjetas bip! que presentan una alta cantidad de transacciones durante la semana de estudio.

La causa de esto no es deducible, pero existen ciertos patrones apreciables que caracterizan este comportamiento. Uno de ellos es que las transacciones son realizadas en buses de la misma unidad de negocios, sin ser el mismo servicio, y por su masividad además son realizadas en lapsos de tiempo cortos entre ellas.

Como se trata de un tipo de viaje extraño y parece más a otro tipo de comportamiento, como por ejemplo transacciones realizadas por personal de fiscalización, de medición u otro comportamiento no explicable, se eliminan estas tarjetas de la base de datos de viajes pues no corresponden a un patrón de viaje.

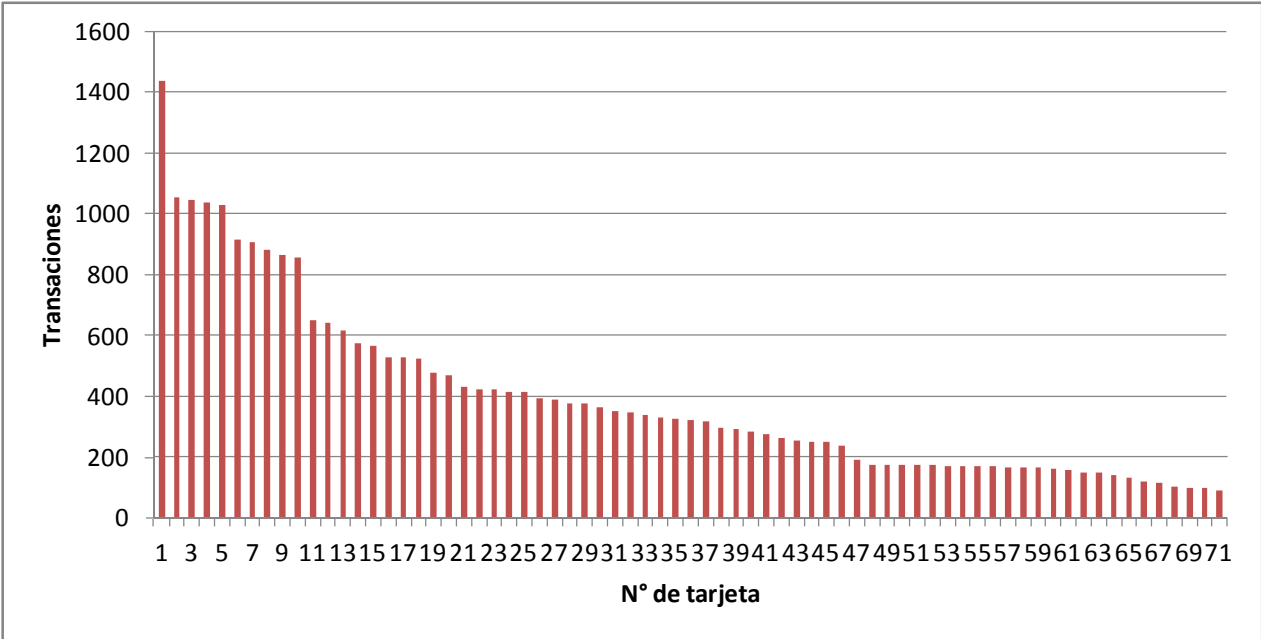
La manera de visualizar este caso se basa en observar aquellas tarjetas ID que tengan cantidades altas de transacciones durante la semana, y todas realizadas en buses del mismo operador.

En primer lugar se aplica un filtro basado en que todas aquellas tarjetas con más de 100 transacciones en la semana, y que sean del mismo operador, son candidatas a tarjetas con transacciones masivas. Luego, se investigan visualmente si hay casos de tarjetas con menos transacciones pero que presenten el mismo caso.

Luego de estos filtros, se logran obtener 71 tarjetas ID con transacción masiva. En el gráfico 4-9 se da a conocer un histograma con las cantidades de transacciones masivas para las tarjetas involucradas. Como se aprecia el comportamiento de estas

tarjetas no deja de ser despreciable, entendiendo que como promedio cada tarjeta ID registra 11 transacciones a la semana y hay tarjetas con más de 1.000 transacciones en la semana en el mismo operador.

Gráfico 4-9: Cantidad de transacciones masivas.

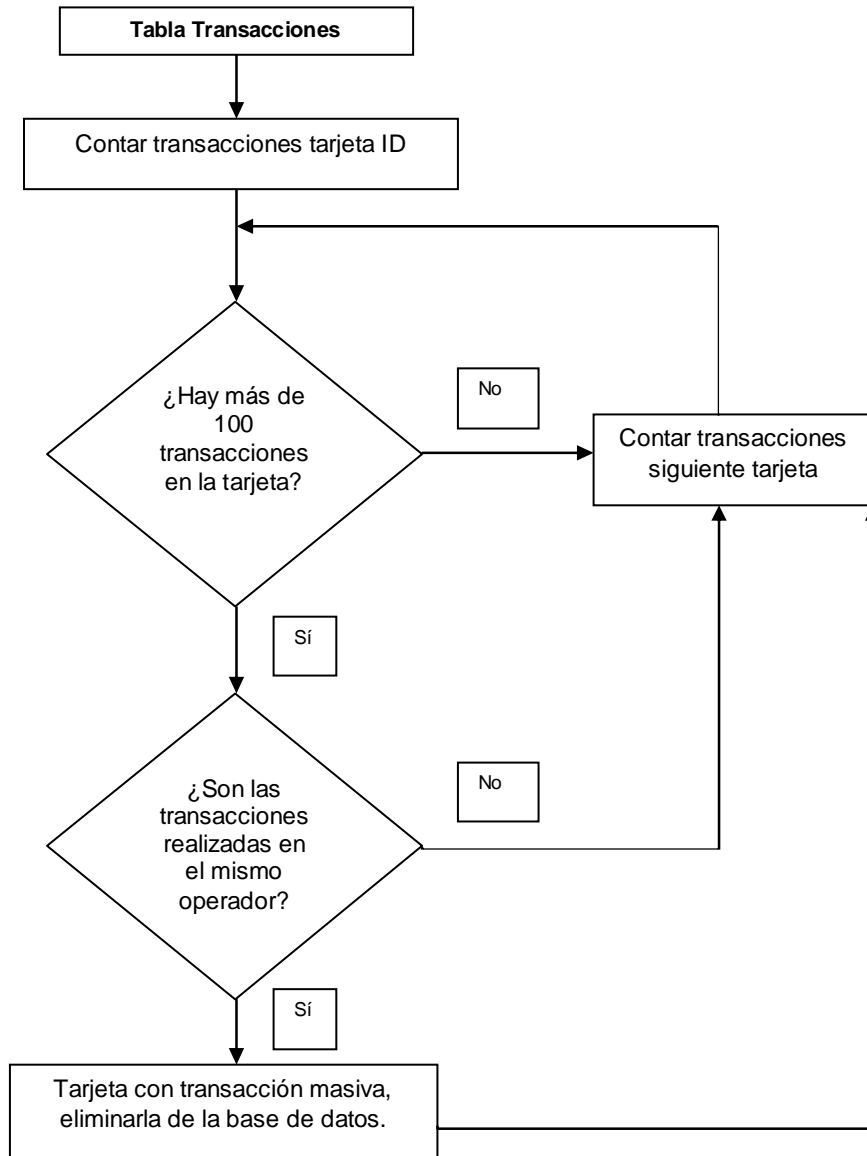


Por motivos de confidencialidad de datos, no se dan a conocer los operadores involucrados en este comportamiento ni el código ID de las tarjetas, pues el motivo de este estudio es meramente estadístico.

La inspección visual acerca de otras tarjetas con menos de 100 transacciones en la semana que presentan este comportamiento transacción masiva da como resultado sólo dos tarjetas más (con 99 y 93 transacciones) que presentan este problema, por lo que parece lógico tomar este filtro de 100 transacciones a la semana como transacción masiva.

Finalmente, se tiene que existen 28.117 transacciones de este tipo asociadas a 71 tarjetas ID. La forma de deshacerse de estas tarjetas en esta y las siguientes bases de datos es la presentada en la figura 4-2.

Figura 4-2: Método de detección y eliminación tarjetas masivas.

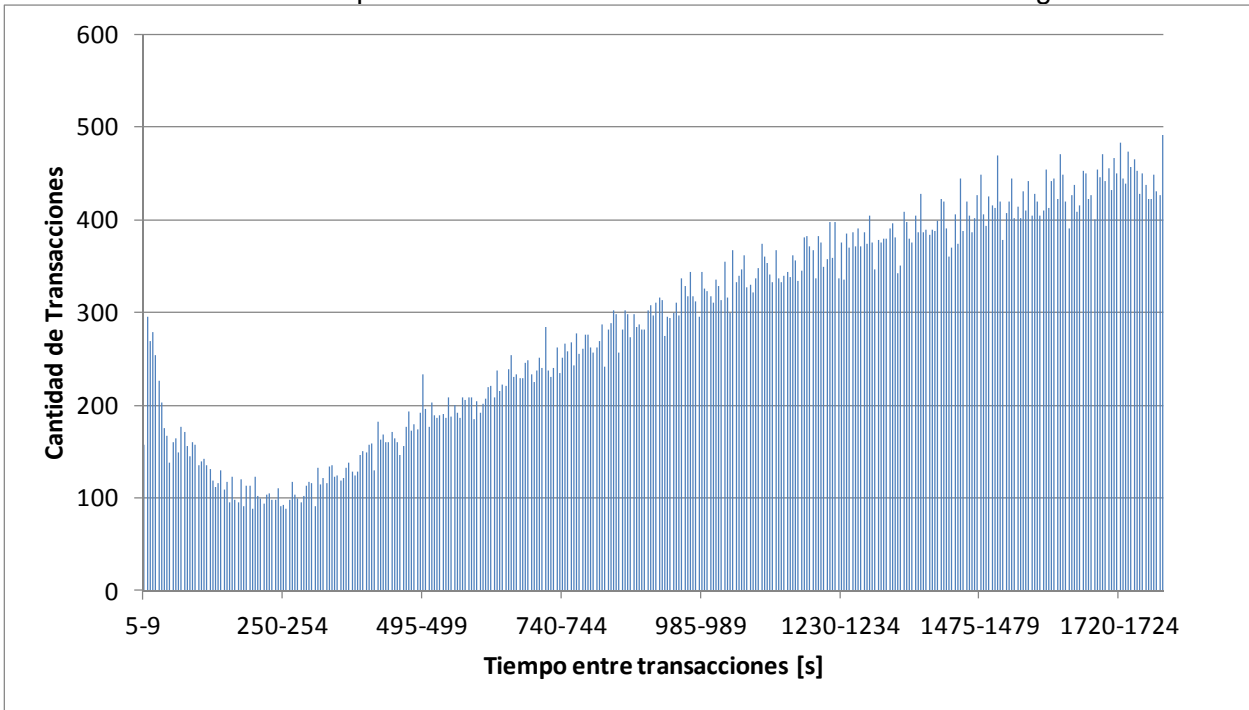


#### 4.3.3. Análisis estadístico.

i. Caso Zona Paga-Bus.

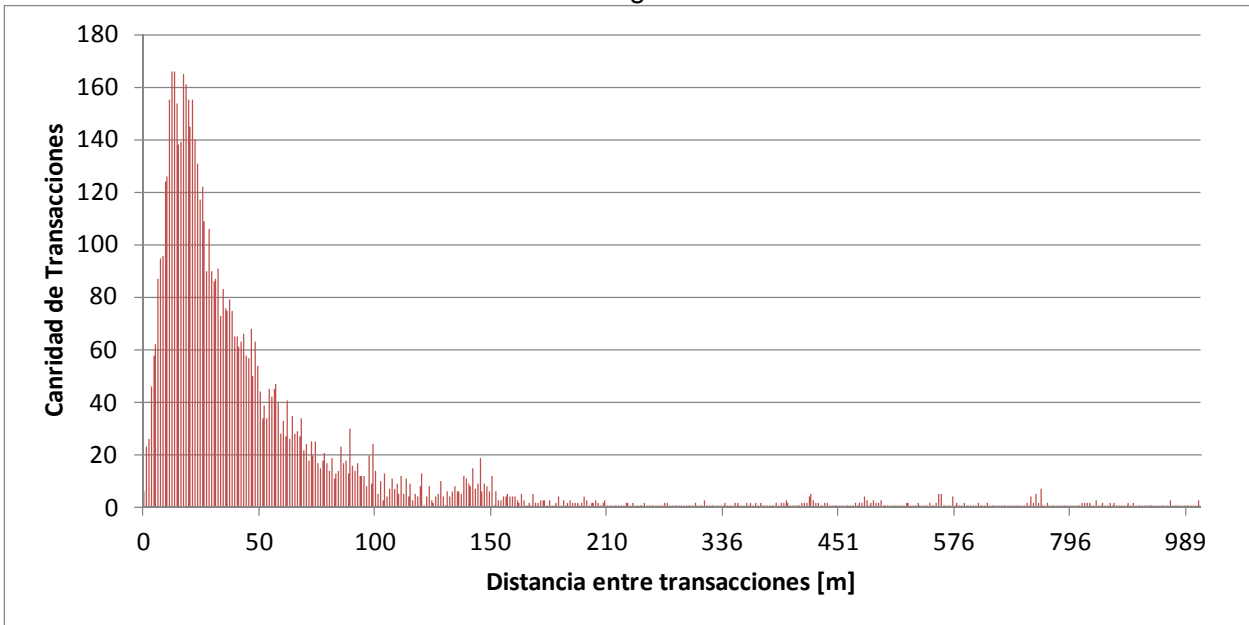
El primer caso a analizar corresponde a las transacciones consecutivas que se realizan en primer lugar en una zona paga y la siguiente se realiza en el validador de un bus. En el gráfico 4-10 se presenta la distribución de cantidad de transacciones por tiempo entre ellas para este caso. En él se aprecia que la curva de transacciones decrece hasta un tiempo entre transacciones de 250 segundos para luego volver a aumentar a medida que aumenta el tiempo entre transacciones. A partir de este valor se analizan las distancias entre transacciones para aquellas con distancia temporal menor a 250 segundos.

Gráfico 4-10: Tiempo entre transacciones consecutivas. Caso Zona Paga-Bus.



En el gráfico 4-11 se aprecian las transacciones, categorizadas por distancia entre ellas, realizadas de manera consecutiva en zona paga y bus respectivamente y cuyo tiempo entre ambas transacciones es menor o igual a 250 segundos. En él se ve que la curva decrece a medida que aumenta la distancia entre transacciones para estabilizarse en un valor cercano a los 150 metros.

Gráfico 4-11: Transacciones consecutivas con distancia temporal menor a 250 segundos. Caso Zona Paga-Bus.

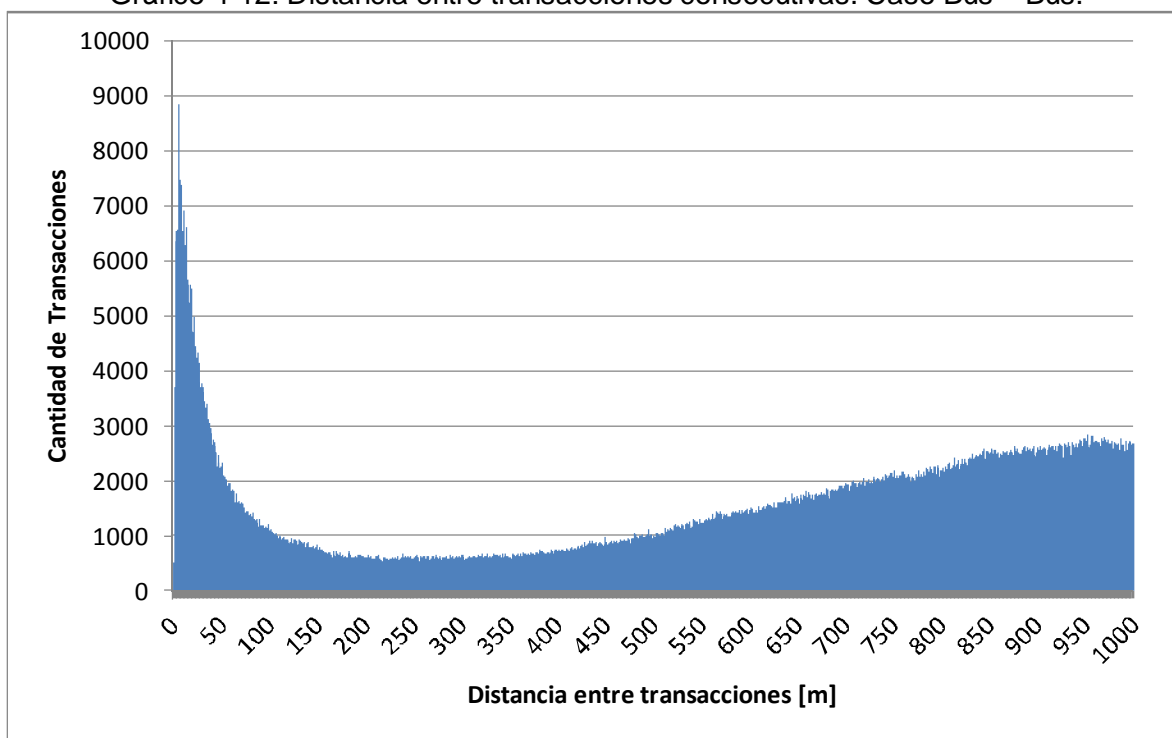


Se tiene que existen 6.340 transacciones realizadas en zona paga cuya siguiente transacción es realizada en un bus, con una distancia temporal menor a 250 segundos y una distancia espacial menor a 150 metros.

ii. Caso Bus-Bus.

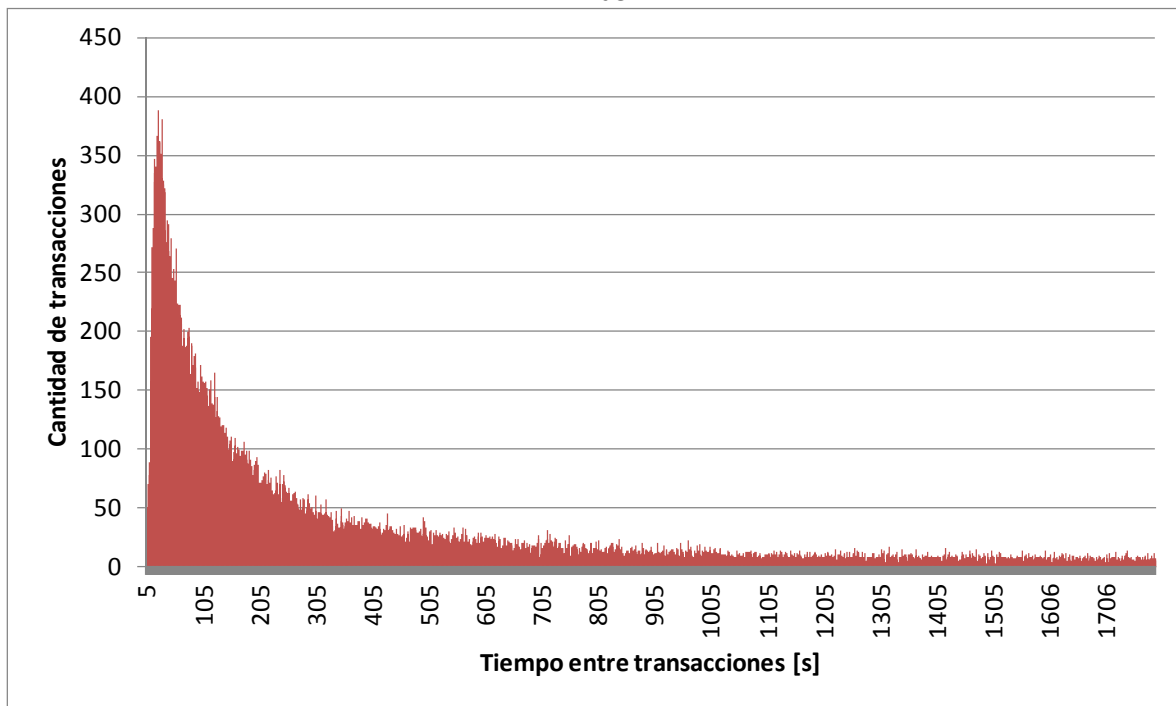
En este caso se analizan las distancias entre dos transacciones consecutivas realizadas con la misma tarjeta en bus, en las que se incluyen las realizadas en de la forma bus-zona paga. Para ello se presenta en el gráfico 4-12 las transacciones cuya distancia con la transacción siguiente es menor a 1.000 metros. Esta distancia se toma como referencia del concepto de bajada no caminable, en la que se propone que esa distancia es considerada como el máximo caminable por el usuario para acceder de un paradero de bajada a uno de subida. En el gráfico se aprecia que la distribución de transacciones tiene una moda de 5 metros de distancia, y luego decae hasta un mínimo que se ubica en la zona cercana a los 200 metros, para luego volver a crecer la cantidad de transacciones a medida que aumenta la distancia entre transacciones.

Gráfico 4-12: Distancia entre transacciones consecutivas. Caso Bus – Bus.



A partir del análisis del gráfico anterior, se define como rango de paradero cercano aquellos que están ubicados a 200 metros de distancia o menos, pues en ellos existe un comportamiento a estudiar que claramente no se acerca al concepto de traslado en transporte público, sin descartar que sea un viaje. En el gráfico 4-13 se da a conocer la distribución de transacciones por tiempo entre dos transacciones consecutivas que cumplan el criterio de distancia menor o igual a 200 metros. En él se aprecia que se logra un máximo de transacciones en el rango entre 5 y 60 segundos, para luego empezar a disminuir la cantidad de transacciones y estabilizarse a partir de los 1.000 segundos.

Gráfico 4-13: Transacciones consecutivas con distancia menor o igual a 200 metros. Caso Bus – Bus.



Se tiene finalmente que hay 53.894 transacciones que cumplen el requisito que su transacción posterior haya sido realizada a 200 o menos metros de distancia y con una diferencia de tiempo de 1.000 segundos o menos. Estas transacciones son consideradas como pago en el mismo sitio para el caso Bus-Bus.

iii. Caso Bus – Metro.

El segundo caso a estudiar corresponde a aquel en que la transacción a estudiar se realiza en un validador en bus o en zona paga y la siguiente en una estación de Metro, ambas realizadas con la misma tarjeta. En el gráfico 4-14 se da a conocer la cantidad de transacciones que cumplen la condición anterior, clasificadas por distancia geográfica entre las dos transacciones, mostrando sólo aquellas cuya distancia temporal es menor o igual a 1.000 metros. En el gráfico se aprecia que la tendencia no es tan clara debido a ciertas distancias que presentan distancias con grandes cantidades de transacciones. Esto ocurre debido a la inclusión de las zonas paga dentro del análisis, al tratarse de sitios de alta demanda y cercanos a estaciones de Metro. Sin embargo, se puede ver que el mínimo ocurre en la zona cercana a los 350 metros.

Tomando en consideración el análisis realizado para el gráfico 4-12, se observarán aquellas transacciones que cumplan el criterio mencionado anteriormente de transacción consecutiva de bus y metro, y se analizará como es el comportamiento de distancia temporal para aquellas que tienen una distancia espacial menor o igual a 350 metros. Esto se aprecia en el gráfico 4-15, el que da a conocer que a partir de una distancia temporal mayor a 600 segundos se nota una estabilización de la curva.



Gráfico 4-14: Distancia entre transacciones consecutivas. Caso Bus – Metro.

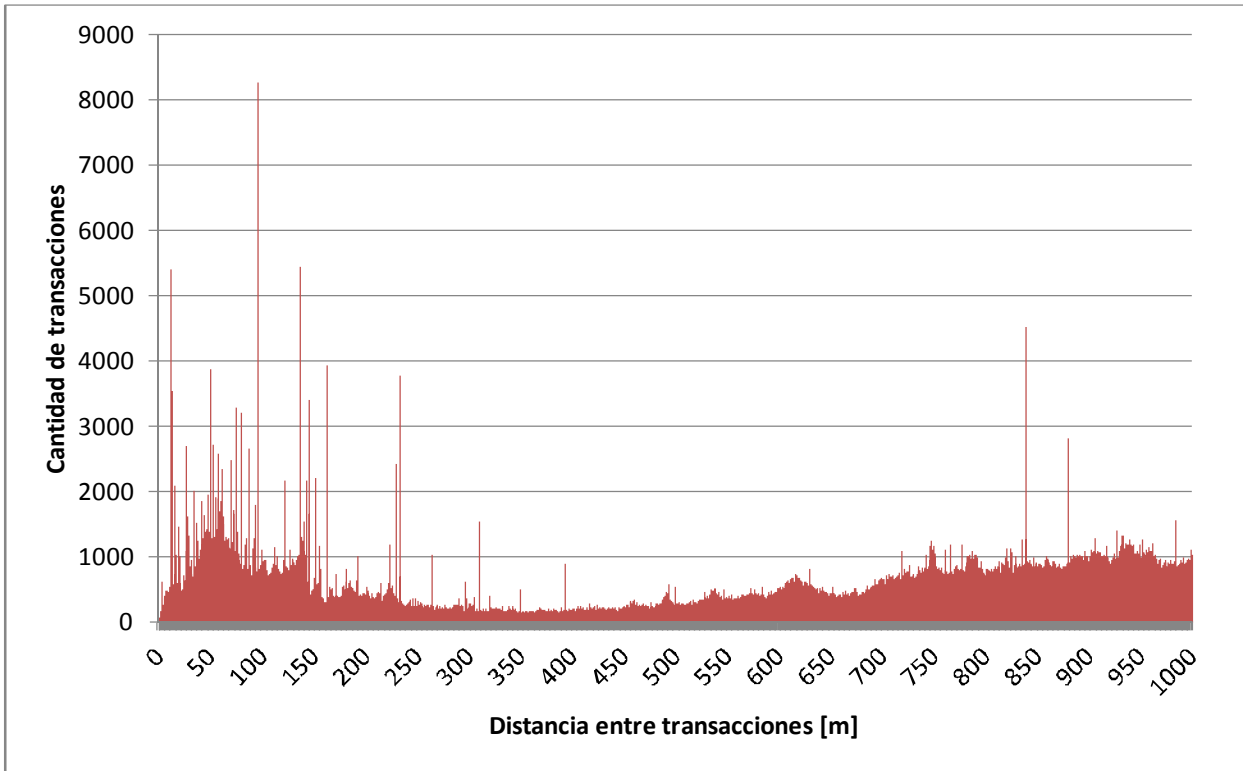
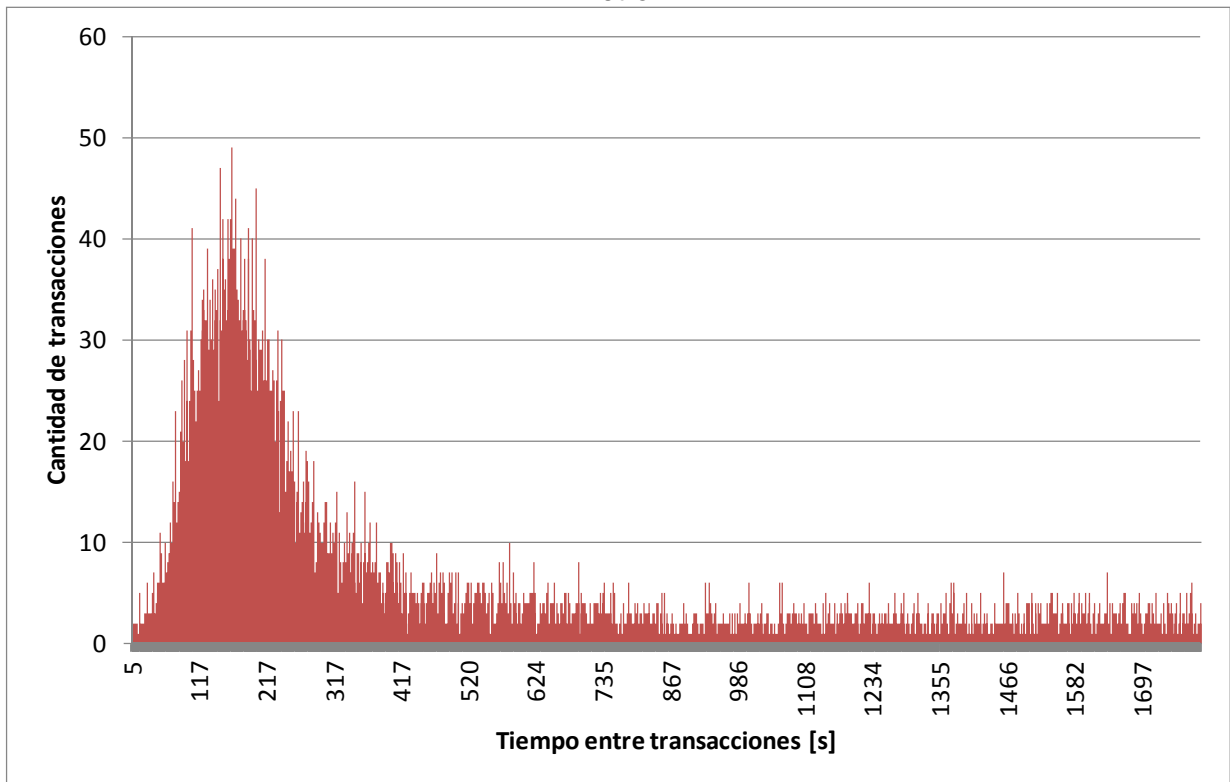


Gráfico 4-15: Transacciones consecutivas con distancia menor a 350 metros. Caso Bus – Metro.

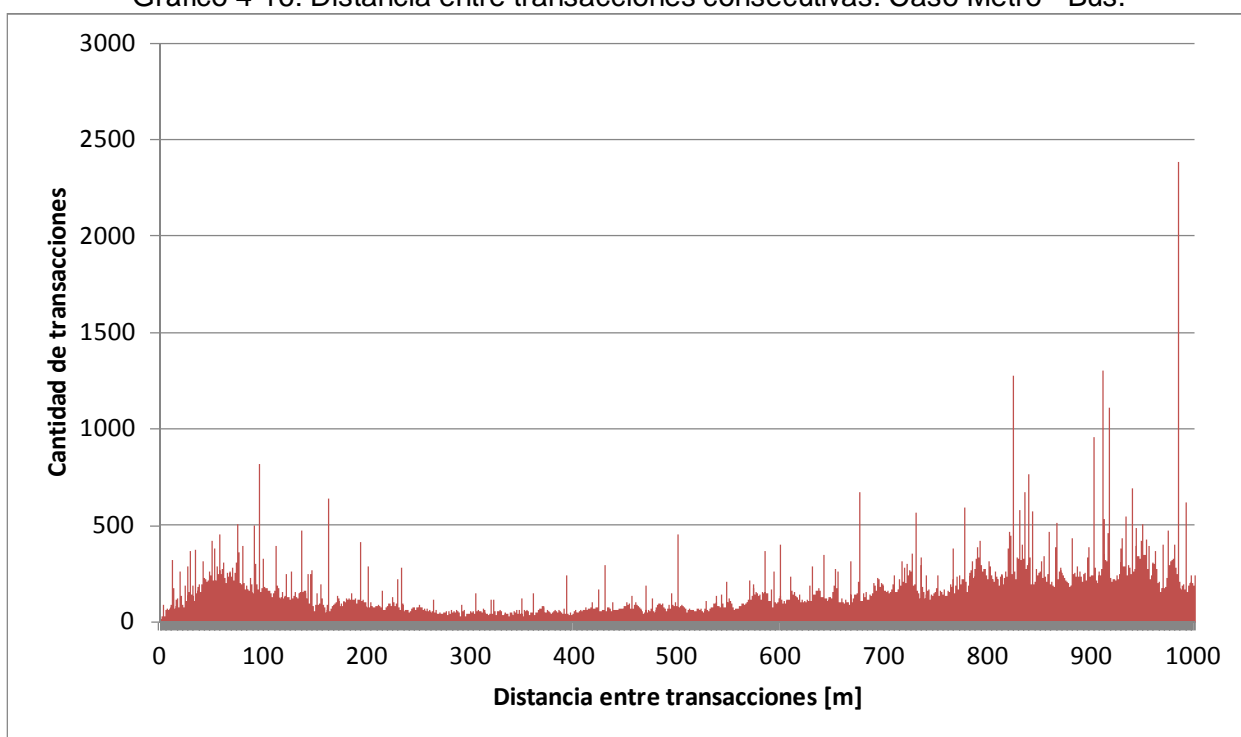


Se tiene que las transacciones consecutivas que cumplen las condiciones de sitio de validación, distancia y tiempo descritas para el caso Bus–Metro son 9.812 transacciones.

iv. Caso Metro – Bus.

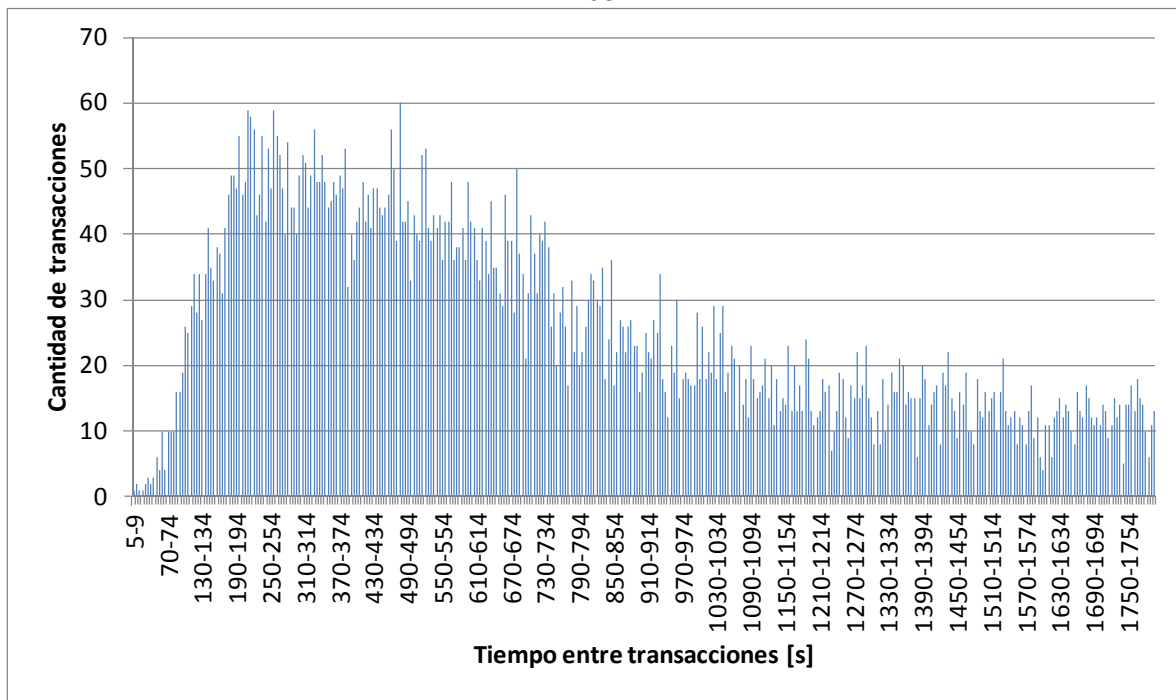
Este caso corresponde a aquel donde se realizan dos transacciones consecutivas con la misma tarjeta, la primera en una estación de Metro y la siguiente en un bus o zona paga. El primer análisis corresponde a estudiar las distancias espaciales entre las dos transacciones, el que se aprecia en el gráfico 4-16. En él se puede ver que el mínimo de transacciones ocurre para la distancia de 350 metros, al igual que en el caso anterior. Sin embargo, no se aprecia un máximo considerable para distancias menores.

Gráfico 4-16: Distancia entre transacciones consecutivas. Caso Metro - Bus.



En este caso se vuelven a observar las transacciones que cumplen los criterios de multipago en sitios cercanos cuya distancia es menor o igual a 350 metros. Desde el punto de vista de la distancia temporal entre ellas, el gráfico 4-17 da a conocer la distribución temporal entre transacciones en rangos de cinco segundos. En él se aprecia que se logra una estabilización de la curva a partir de los 1.000 segundos.

Gráfico 4-17: Transacciones consecutivas con distancia menor a 350 metros. Caso Metro – Bus.



Se tiene que existen 9.155 transacciones que cumplen los requisitos mencionados para este caso, es decir transacciones consecutivas con la misma tarjeta que sean realizadas a menos de 350 metros de distancia y con menos de 1.000 segundos de diferencia.

#### 4.3.4. Propuesta de solución.

El haber realizado análisis estadísticos para los cuatro casos anteriormente descritos tiene como objetivo estudiarlos por separado, pues se supone que los comportamientos asociados a multipago en sitios cercanos tienen distintos fines. Sin embargo, con sólo datos estadísticos no es posible suponer con certeza la ocurrencia de cada comportamiento a presentar, pero se sugieren a partir de la experiencia de viajes y de la lógica del usuario.

El primer caso a estudiar es el de transacciones consecutivas realizadas en primer lugar en zona paga y luego en el validador de un bus. Acá para los parámetros obtenidos de distancia (150 metros) y tiempo (250 segundos) se propone eliminar del registro de etapas aquella realizada en la zona paga, pues se denota la ausencia de traslado. Las causas de este comportamiento son variadas, y pueden ir desde la equivocación del usuario al validar en una zona paga equivocada, o que se olvide de la validación hecha en la zona paga y vuelva a validar en el bus. Sin embargo, el comportamiento indica que no hay traslado a medir para la transacción en zona paga para las distancias y tiempos mencionados.

El segundo caso a estudiar es el de validaciones consecutivas en bus. De los tres casos es el de mayor ocurrencia para los parámetros de distancia (200 metros) y tiempo (1.000 segundos) establecidos. Acá la distancia entre transacciones es la menor

de los tres casos, y esto sugiere que las transacciones consecutivas se realizan en paraderos adyacentes, los cuales generalmente están asociados a un grupo de paradas de buses que permiten la subida y bajada de pasajeros para sectores de alta oferta de recorridos. Para este caso se supone el préstamo de la tarjeta bip! entre dos usuarios como principal causa de este comportamiento de multipago. Esto debido a que la distancia obtenida no sugiere la ocurrencia de un viaje corto en bus, sino más bien a un traslado a pie entre paraderos. Para esto, el modelo deberá asumir que la primera transacción no genera una etapa de viaje, por lo que no se deberá estimar bajada en este caso y se catalogará la bajada como multipago en sitio cercano. Sin embargo, la segunda transacción sí se califica como etapa de un viaje, y se le estima paradero de bajada siguiendo el procedimiento habitual.

El tercer caso correspondiente a transacción consecutiva entre bus y metro se puede mirar desde otro punto de vista a partir de los datos obtenidos. El hecho que el rango de tiempo entre validaciones obtenidas del estudio (600 segundos) sea menor a los otros dos casos y la distancia sea mayor al caso anterior (350 metros) hace suponer que existió un viaje corto realizado en bus para acceder a la estación de Metro. Este comportamiento tiene dos posibles causas: la primera tiene relación con motivos de seguridad en el traslado hacia la estación de Metro, y la segunda un aprovechamiento del sistema tarifario en que el traslado corto no genera un costo adicional para el usuario. Sin embargo, no es posible apreciar en qué caso se puede hacer el corte entre un viaje corto y un préstamo de la tarjeta. Se propone para este caso mantener el concepto de viaje corto para aquellos traslados de la tarjeta mayores a 50 metros, y aquellas transacciones con distancias menores se elimine la estimación de bajada pues no supone un traslado en algún modo motorizado integrado sino un posible préstamo de la tarjeta.

Finalmente, el caso de transacciones consecutivas de metro y bus se explica de manera similar al caso de transacciones consecutivas en buses. La causa más razonable de este comportamiento es el préstamo de la tarjeta de parte del dueño de ésta a otra persona para la realización de su viaje. Sin embargo, también cabe pensar la posibilidad de que el usuario una vez hecha la validación en metro se arrepienta de su uso debido a factores tales como congestión en el andén y utilice bus para hacer el recorrido. A pesar de esto, la baja cantidad de transacciones que cumple este comportamiento no permite un mejor análisis. Se propone como solución a este problema, al igual que para los casos anteriores, eliminar la bajada estimada (si es que existe) de la primera transacción y catalogarla como multipago en sitio cercano para las condiciones de distancia (350 metros) y tiempo (1.000 segundos) entre transacciones asociadas a este comportamiento.

## **5. Zonas Paga.**

### **5.1. Introducción.**

Las zonas paga son paraderos de buses con infraestructura especial para realizar el pago en validadores instalados en el andén, lo que genera una mejora en los tiempos de subida al bus debido a la omisión del pago arriba del bus. Estas zonas paga están generalmente instaladas en lugares de alto flujo peatonal y de buses, por lo que presentan una infraestructura más grande que un paradero tipo.

Desde el punto de vista metodológico, existen dos problemas que afectan la estimación de bajada de las zonas paga. En primer lugar se tiene que al validar el usuario en el sitio, no es posible saber con exactitud qué servicio abordó pues no queda registrado en el validador, y por ende se complica estimar el paradero de bajada. Esto a diferencia de los buses que tienen asignado el servicio y el sentido que recorren. Actualmente se utiliza un criterio de selección de líneas comunes, esto en base a elegir aquel servicio que pase por la zona paga en la que validó el usuario y presente el menor tiempo generalizado de viaje.

El segundo problema que afecta la estimación de bajada, que sin duda es el más importante, corresponde a la calidad de los datos disponibles para trabajar. El hecho de que no haya concordancia entre validadores y servicios afecta el modelo pues estima de manera errónea (o simplemente no estima) las bajadas de los usuarios de transporte público. También se aprecian problemas que son de datos pero que no están directamente relacionadas con las zonas paga, como son problemas con la emisión de GPS de los buses en los túneles existentes en la rutas, pero que serán tratados en este capítulo debido a su alta incidencia en recorridos que usan zonas paga.

### **5.2. Análisis de Datos.**

Se tiene que durante la semana de estudio se realizaron 2.034.917 validaciones en Zonas Paga, que corresponden al 5,34% del total de validaciones hechas en el sistema durante esa semana (Ver Gráfico 2-1). Además, 875.365 tarjetas bip! utilizaron esta forma de viaje.

En la tabla 5-1 se dan a conocer las 15 paradas de tipo Zona Paga con mayor cantidad de validaciones en la semana de estudio. Estas zonas paga se encuentran en zonas próximas a estaciones de Metro, pues se trata de zonas de alto flujo de usuarios, quienes cambian de modo Metro a modo Bus.

Estos sistemas de pago no son todos permanentes durante el día. Debido a múltiples factores (monetarios, logísticos, etc.) son ubicados validadores en estos paraderos principalmente en horas de alta afluencia de pasajeros (Horas Punta). La cantidad de validaciones por periodo del día para los tres tipos de día (Laboral promedio, Sábado y Domingo) en zonas paga se muestra en el gráfico 5-1. Se puede apreciar en el gráfico que el uso principal de zonas paga ocurre en la tarde en día laboral, debido a que se habilitan para quienes se trasladan a sus hogares, especialmente cerca de estaciones de Metro, donde se genera congestión peatonal a diferencia de las mañanas, donde la gente aborda de inmediato el Metro una vez que

baja del bus que tomó previamente. Esto muestra la alta concentración en los orígenes de viaje en el periodo punta tarde. En cambio, aquellas zonas paga del tipo permanente son las existentes en Estaciones de Intercambio Modal, lugares de alta afluencia de pasajeros durante todo el día y que poseen una gran cantidad de servicios de parada. Estos son edificios, generalmente subterráneos de varios andenes, donde previo al ingreso está un sector habilitado con torniquetes tipo Metro y validadores para el pago de los usuarios.

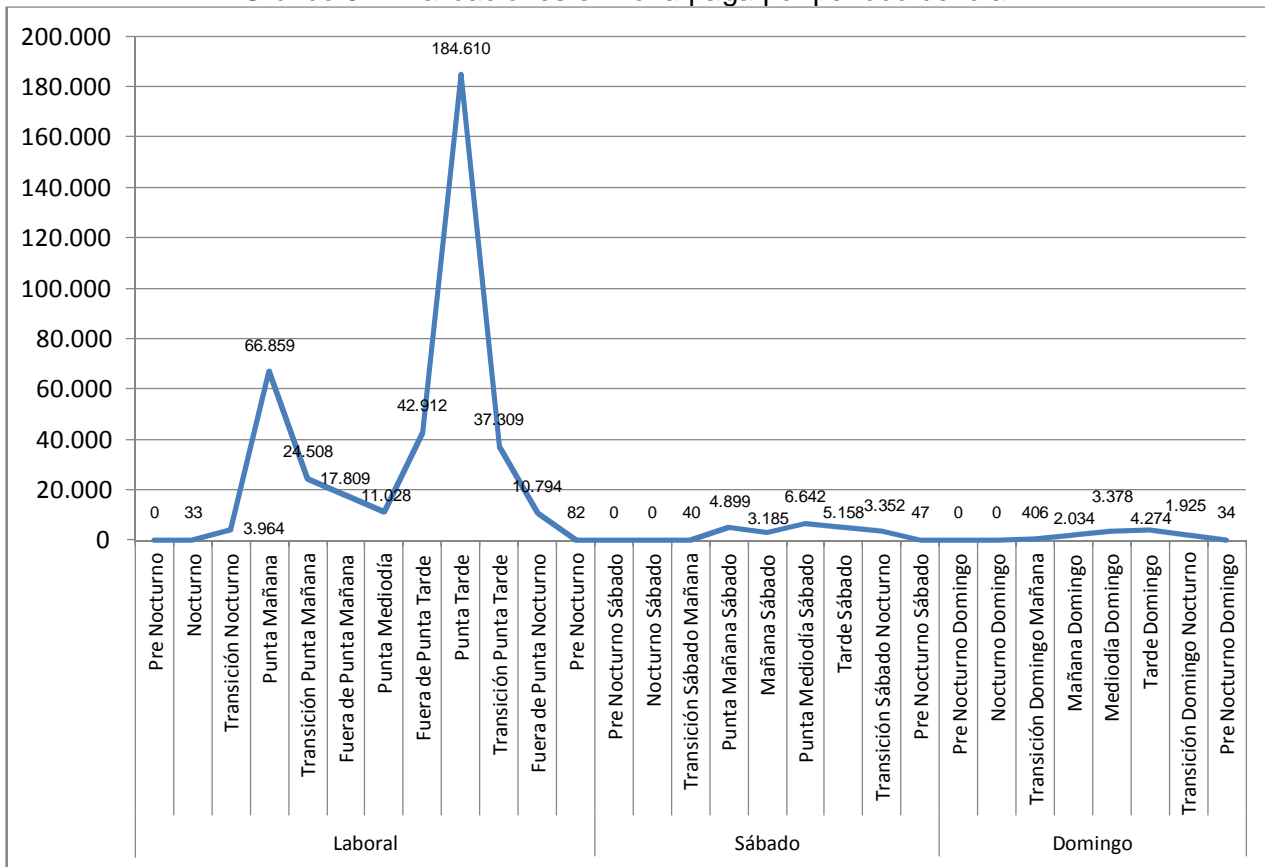
Tabla 5-1: Principales Zonas Paga en cantidad de validaciones.

<b>Sector</b>	<b>Comuna</b>	<b>Validaciones</b>
EIM La Cisterna	La Cisterna	112.086
EIM Bellavista de La Florida	La Florida	91.035
EIM Lo Ovalle	La Cisterna	84.270
Metro Escuela Militar	Las Condes	74.555
Metro Santa Lucía	Santiago	63.550
Metro Universidad de Chile	Santiago	60.928
Metro Las Rejas	Estación Central	56.825
Plaza de Puente Alto Poniente	Puente Alto	47.818
Metro Pedro de Valdivia	Providencia	46.193
Estación Mapocho	Santiago	43.514
Plaza de Maipú	Maipú	41.938
José María Caro esq. Av. Recoleta	Santiago	35.343
Hospital Sótero del Río	Puente Alto	34.507
Plaza Baquedano	Providencia	33.279
Plaza de Puente Alto Oriente	Puente Alto	32.513

El sistema presenta una cantidad de 141 paradas del tipo Zona Paga, las cuales tienen en su conjunto 487 validadores disponibles que son los que el modelo reconoce como sitios de pago. En estas paradas pueden detenerse desde uno y hasta diez servicios diferentes. Cada zona paga existente presenta una cantidad definida de servicios asignados a ella, los cuales utilizan la parada como parte de su ruta, pudiendo esta ser la parada inicial del recorrido o parada dentro del conjunto de las mismas.

De las 141 paradas tipo zona paga existentes, aquellas que tienen un solo servicio asignado son 33 paradas, que corresponden al 23,4% de las paradas existentes como zona paga. Estas, a su vez, tienen asignadas 115 validadores que corresponden al 23,61% del total de validadores presentados en este trabajo como sitios de pago. Y en estas zonas paga se realizaron 417.907 validaciones, que son el 20,54% del total de validaciones realizadas en zonas paga en la semana de estudio. Desde el punto de vista de asignación de servicio se puede decir con certeza que los usuarios de estas zonas paga usaron el único servicio disponible como parte de su viaje.

Gráfico 5-1: Validaciones en zona paga por periodo del día.



Para el resto de zonas paga con dos ó más servicios disponibles, la cantidad de paradas, sitios de pago y validaciones se presenta en la tabla 5-2. Se puede ver que existe una correlación entre la cantidad de sitios de pago y las transacciones realizadas. En la mayoría de las zonas paga arriban entre dos y seis servicios, y esto queda ratificado al observar las transacciones realizadas.

Tabla 5-2: Servicios asignados por paradero y validaciones.

Servicios	Paraderos	Sitios de Pago	Validaciones
2	36	111	328.554
3	24	81	296.956
4	14	52	221.873
5	12	46	233.016
6	14	46	273.093
7	2	9	37.307
8	5	22	178.393
10	1	5	47.818

### 5.3. Análisis de errores.

#### 5.3.1. Análisis exploratorio.

Una primera etapa en el trabajo de estudiar la capacidad del modelo de asignación de bajada es revisar en una pequeña muestra de tarjetas ID los problemas que se podrían presentar tanto en la asignación de servicio como en la estimación de paradero de bajada. Para la elección de las tarjetas ID a estudiar se realiza un proceso de elección aleatorio donde se seleccionan 102 tarjetas ID de un total de 875.365 utilizadas en la semana de estudio que validaron por lo menos una vez en una de las zonas paga dentro del sistema. El tamaño muestral es elegido en base a un criterio de capacidad temporal de revisar una por una cada tarjeta.

Un primer análisis de la muestra tiene que ver con comparar el comportamiento estadístico entre la muestra y el total de tarjetas que validaron en zona paga en la semana de estudio. Los principales parámetros se presentan en la tabla 5-3, en la que se aprecia cercanía en el promedio de viajes por tarjeta para la muestra y el total.

Tabla 5-3: Parámetros estadísticos de comparación entre muestra y total de tarjetas ID.

Parámetro	Muestra	Total (Zona Paga)
Tarjetas ID	102	875.365
Validaciones	245	2.034.917
Promedio de validaciones por tarjeta ID	2,40	2,32
Varianza de validaciones por tarjeta ID	4,3022	3,3455
Desviación Estándar de validaciones por tarjeta ID	2,0742	1,8291

Si bien existe diferencia en la varianza (y por ende en la desviación estándar) esto ocurre por la diferencia de magnitudes en la cantidad de tarjetas entre ambos, pero que a pesar de aquello son parecidos. La causa de esto se observa en la tabla 5-4, correspondiente a un conteo de validaciones por tarjeta tanto para la muestra como para el total en la semana de estudio. El hecho de que la varianza sea mayor en la muestra que en el total de tarjetas se debe a la sobrerrepresentación que tienen las tarjetas con siete ó más validaciones en la muestra. Esto genera que la muestra para esos casos no esté bien distribuida, y por ende un aumento en la varianza. Pero como se trata de casos aislados (sobrerrepresentación afecta a menos de 2% de la muestra) no genera mayores incidencias en el análisis ni en los resultados.

Finalmente, se hace una comparación entre ambas tablas en base a las transacciones con estimación de bajada. La comparación se observa en la tabla 5-5, donde se puede ver que hay una similitud en la estimación de bajada entre ambas tablas, en la que se espera un orden cercano a tres cuartos del total de transacciones con estimación de bajada.



Tabla 5-4: Cantidad de validaciones por tarjeta ID, comparación entre muestra y total para semana de estudio.

Validaciones por Tarjeta ID	Muestra		Total	
	Frecuencia [Tarjetas]	%	Frecuencia [Tarjetas]	%
1	49	20,00	420953	20,69
2	21	8,57	167117	8,21
3	11	4,49	99531	4,89
4	7	2,86	78415	3,85
5	7	2,86	63134	3,10
6	0	0,00	17663	0,87
7	4	1,63	9150	0,45
8	1	0,41	7017	0,34
9	0	0,00	5189	0,25
10	1	0,41	4597	0,23
11	0	0,00	936	0,05
12	1	0,41	530	0,03
13	0	0,00	385	0,02
14	0	0,00	314	0,02
15	0	0,00	280	0,01
16	0	0,00	54	0,00
17	0	0,00	28	0,00
18	0	0,00	25	0,00
19	0	0,00	24	0,00
20	0	0,00	20	0,00
21	0	0,00	1	0,00
23	0	0,00	1	0,00
25	0	0,00	1	0,00

Tabla 5-5: Estimación de bajada para muestra y total de tarjetas ID.

Estimación	Muestra		Total	
	Frecuencia	%	Frecuencia	%
Hay estimación de bajada	180	73,47%	1.557.911	76,56%
No hay estimación de bajada	65	26,53%	477.006	23,44%

Una característica de los viajes con zonas paga es que la transacción que se realiza en la zona paga corresponde generalmente a la última etapa del viaje. De las 245 transacciones realizadas en esta muestra en zona paga, 216 corresponden a la última etapa del viaje. Esto es una desventaja para la estimación de bajada, pues el modelo debe buscar la bajada en base a una validación posterior en la que posiblemente ocurrió alguna actividad que podría implicar traslados no contenidos dentro de la base de datos de transacciones.

El análisis de errores se debe hacer tanto para aquellas validaciones que no estiman bajada como para aquellas que estiman bajada pero presentan comportamientos anómalos.

De las 180 validaciones en las que el modelo estima bajada, en 167 no se presentan errores visibles. Las 13 validaciones con estimación de bajada pero que presentan errores corresponden a estimaciones de bajada donde se estima como paradero de bajada el paradero de subida de la transacción, lo que hace suponer que la persona subió y bajó en el mismo lugar. Además existen cinco transacciones en las que la parada de bajada estimada es cercana a la de subida. De este problema ya se habló en el capítulo 3.

Las 66 transacciones realizadas en zonas paga que no tienen estimación de paradero de bajada tienen diferentes errores se clasifican en la tabla 5-6. Se aprecian dos problemas principales: los errores de estimación de bajada debido a la ausencia de un paradero de bajada a menos de 1.000 metros de los paraderos del servicio que se utiliza y los errores de emisión de coordenadas mediante GPS del bus. Además existen errores asociados a problemas de falta de información de recorridos. Las fallas encontradas en este análisis dan una idea de lo que se debe buscar dentro de la base de datos completa de validaciones.

Tabla 5-6: Fallas en estimación de bajada para muestra de tarjetas ID.

<b>Falla</b>	<b>Frecuencia</b>
Etapa No Caminable	28
Datos malos de bus	20
Única validación del día	7
No se encontró bus para el recorrido	6
Faltan etapas	2
Validación posterior sin posición	1
Multipago en sitios distintos	1
<b>Total</b>	<b>65</b>

Ajeno al problema de zona paga es la inexistencia de otra validación en el día con el que poder estimar bajada, la falta de etapas de viaje con uso de tarjeta bip! y el multipago en sitios distintos. Estos temas, sin embargo, son tratados en otros capítulos de este trabajo.

### 5.3.2. Análisis global de validaciones en zonas paga.

Luego de haber construido una base de trabajo en la búsqueda de errores de estimación de bajada, se procede a cuantificar cada uno de los principales errores que afectan al modelo, para definir la magnitud de cada uno dentro del universo de validaciones.

Si bien no todas estas anomalías presentan soluciones factibles de realizar, es importante conocerlas para tener una idea clara de la clasificación de estos errores y

tenerlos en cuenta ante posibles soluciones futuras de aquellos que no se traten en este trabajo.

En la tabla 5-7 se presentan las zonas paga con mayor cantidad de estimación de paradero de bajada nula y el porcentaje sin bajada sobre el total de validaciones en la zona paga. Esta tabla permite tener una idea de aquellas zonas paga con menor cantidad de estimaciones de bajada. En el caso de la cantidad de validaciones sin estimación el análisis va enfocado en buscar y analizar los distintos patrones que afecten la estimación. Desde el punto de vista porcentual se pueden apreciar sitios con casi la totalidad de las transacciones sin estimación de bajada, lo que implica estudiar si se trata de un comportamiento anómalo en el modelo o en los datos entregados para el proceso de los mismos.

El proceso de trabajo se basa en analizar cada una de las zonas paga con mayores problemas, tanto en lo porcentual como en las cantidades, y buscar las fallas principales que afectan a cada una de estas zonas paga. La búsqueda de fallas tiene como base el análisis previo realizado en la muestra de 102 tarjetas ID donde se encontraron un grupo de problemas principales que afectan al proceso de estimación de paradero de bajada a partir de transacciones en zonas paga. Es en base a aquellos problemas a los que se les dará énfasis en la corrección.

### **5.3.3. Servicios mal asignados.**

El primer problema encontrado, que afecta principalmente a aquellas zonas paga con un porcentaje cercano al 100% de transacciones sin estimación de bajada, corresponde a la asignación de servicios a los distintos sitios de zonas paga existentes. Para este caso se encontraron una serie de servicios mal asignados a validadores de zonas paga, cuya búsqueda se realizó principalmente de manera manual, observando las subidas de las siguientes transacciones y analizando la posible ruta que debería seguir el servicio que corresponde a la realidad, para luego buscar dentro del conjunto de servicios el que mejor se asocie a la ruta verdadera.

Cada validador habilitado de las distintas zonas paga existentes tiene asociado un grupo de servicios que son los que se detienen en el paradero, generalmente asociado a un operador, con el fin de hacer el proceso de recaudación más expedito. La importancia para el modelo de asignación de paradero de bajada que los servicios estén correctamente asignados radica en que el modelo ocupa la información de los servicios asignados al paradero para hacer el proceso de asignación de servicio a la transacción en base al enfoque de líneas comunes y tiempo mínimo generalizado de viaje. Si los servicios asignados a los paraderos no corresponden a la realidad, el modelo no es capaz de estimar bajada debido a que la siguiente transacción se encuentra generalmente alejada de la ruta del servicio asignado a la zona paga, y no estima bajada dejando la falla como bajada no caminable.

Tabla 5-7: Zonas paga con mayor cantidad de asignaciones de bajadas nulas.

Sector	Validaciones		Sitios	Porcentaje sin bajada
	Total	Sin Bajada		
EIM La Cisterna	112.086	69.919	20	62,38%
EIM Lo Ovalle	84.270	16.961	6	20,13%
Plaza Baquedano	33.279	16.231	4	48,77%
EIM Bellavista de La Florida	91.035	15.492	4	17,02%
Metro Manquehue	28.200	14.771	7	52,38%
Metro Santa Lucía	63.550	12.406	11	19,52%
Metro Escuela Militar / Apoquindo	74.555	12.377	12	16,60%
José María Caro con Av. Recoleta	35.343	11.728	7	33,18%
Metro Santa Rosa	22.092	10.324	5	46,73%
Metro Universidad de Chile	60.928	9.614	4	15,78%
Metro Las Rejas Norte	56.825	9.233	10	16,25%
Plaza de Maipú 1	9.452	9.087	2	96,14%
Metro San Pablo	8.681	8.601	6	99,08%
Metro Pedro de Valdivia / 11 de Sept.	46.193	8.584	4	18,58%
Metro Pedro de Valdivia / Av. Providencia	9.229	8.544	2	92,58%
Metro Plaza de Puente Alto / J. L. Coó	32.513	7.890	4	24,27%
Plaza de Maipú 3	41.938	6.961	5	16,60%
Metro Plaza de Puente Alto / Sta. Josefina	47.818	6.746	5	14,11%
Metro Escuela Militar / Vespucio	31.936	6.593	6	20,64%
Estación Mapocho / Bandera	43.514	6.486	3	14,91%
Plaza de Maipú 2	6.053	5.513	2	91,08%
Metro Francisco Bilbao / Bilbao	26.667	5.492	5	20,59%
San Antonio con Moneda	28.819	4.928	4	17,10%
José Joaquín Pérez con Teniente Cruz	5.100	4.795	2	94,02%
Rotonda Irene Frei / Líder Buenaventura	4.688	4.546	4	96,97%

Para entender mejor este procedimiento se dan a conocer dos ejemplos del trabajo realizado. En primer lugar se estudia el servicio G04 Ida, cuyo origen es en la Estación Intermodal de La Cisterna y tiene como destino la esquina de Miguel Mujica con María Elena en la comuna de La Florida.<sup>4</sup>

Los validadores asignados a este servicio dentro de la EIM La Cisterna y sus respectivas validaciones totales y sin estimación de bajada son las que se muestran en la tabla 5-8. Se puede apreciar que existen dos sitios (RM-0720 y RM-0721) que presentan un alto porcentaje de estimaciones de bajada nulas. Al analizar los errores presentes en esos dos sitios se observa que los principales problemas se deben a que el modelo no encontró paradero de bajada dentro del recorrido G04I debido a que las transacciones estaban a más de 1.000 metros de distancia de la ruta.

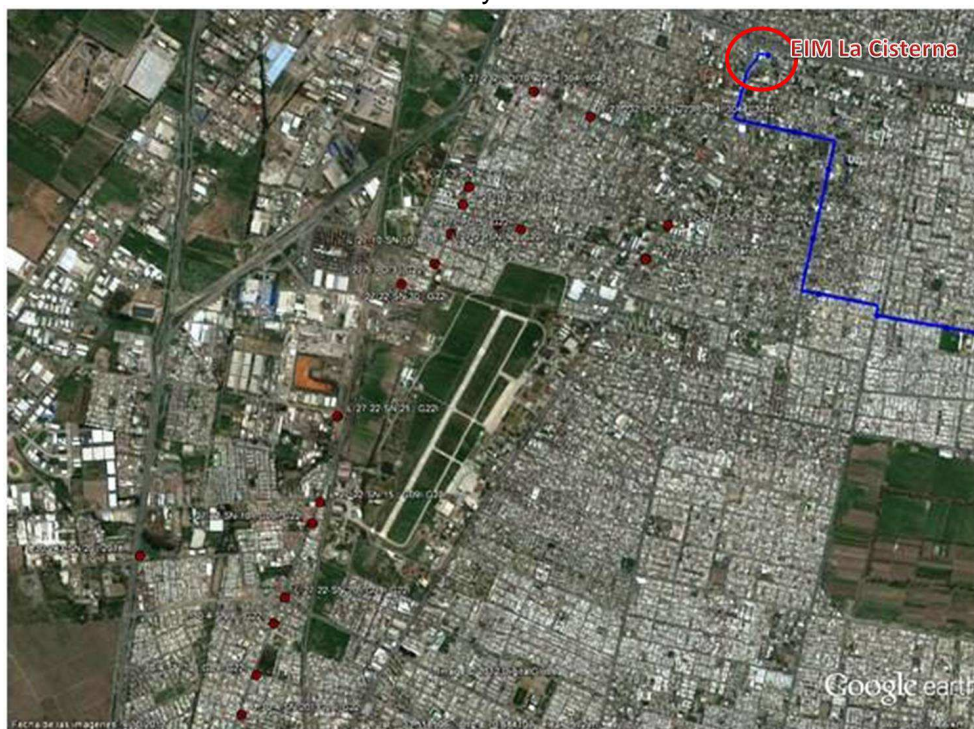
<sup>4</sup> 14º P.O. Modificado Zona G. Sitio web Transantiago.

En base a esto se observaron las ubicaciones geográficas de las transacciones siguientes a cada validación del recorrido G04I hechas en alguno de estos dos sitios con alto porcentaje de error. Con esa información se realizó un mapa en que se muestran los paraderos donde se realizaron la mayor cantidad de transacciones posteriores, mapa que se observa en la imagen 5-1. Los puntos en rojo representan los paraderos donde se realizaron la mayoría de las transacciones posteriores. Claramente se puede apreciar que hay una diferencia total entre el recorrido del servicio G04I (línea azul) y las siguientes transacciones (puntos rojos). Bajo el supuesto de que el usuario baja en un paradero cercano a la ubicación geográfica de la siguiente transacción, no existe relación entre el recorrido asociado al validador y el que realmente debe ser. Esto hace suponer que el recorrido al que pertenecen estas transacciones no es el asignado en la zona paga y por ende se debe buscar el recorrido correspondiente.

Tabla 5-8: Sitios y transacciones para servicio G04I en EIM La Cisterna.

Sitio	Servicio	Total Validaciones	Sin Asignar	% Sin asignar
RM-0718	G04I	8275	2186	26,42
RM-0719	G04I	9101	2401	26,38
RM-0720	G04I	8719	8118	93,11
RM-0721	G04I	7920	7379	93,17

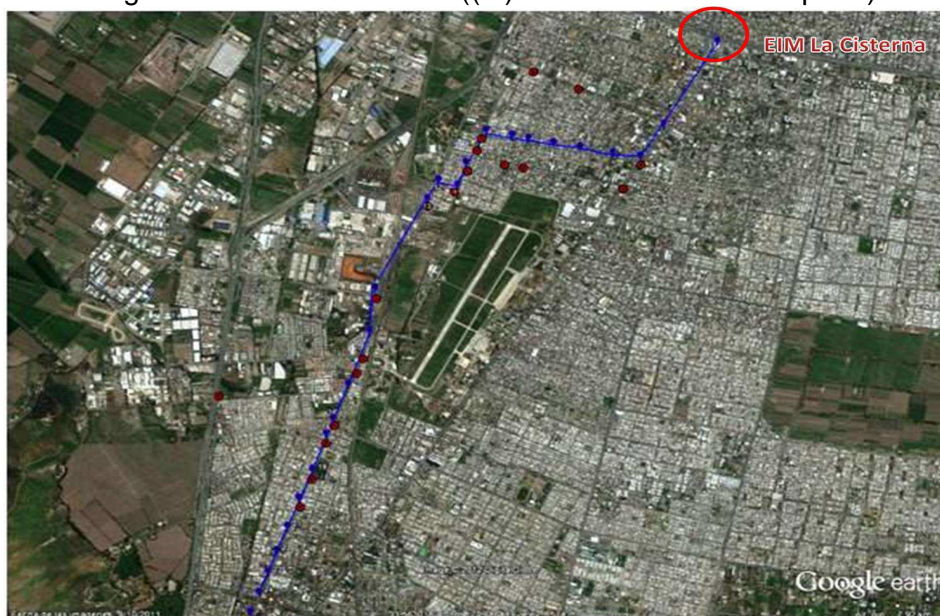
Imagen 5-1: Paraderos con la mayor cantidad de transacciones posteriores a hechas en sitios RM-0720 y RM-0721.



La etapa siguiente, una vez reconocido el error en la asignación de servicio al validador, es encontrar el servicio que corresponde a los validadores descritos. Para eso se observa previamente los recorridos que están asignados a los paraderos de subida de la siguiente transacción con el fin de obtener una idea del recorrido a buscar. Para este caso los dos recorridos que más aparecieron en los paraderos de transacción

siguiente corresponden a los servicios G09I y G22I. Bajo esto, suponemos que uno de los dos pero con sentido contrario es el que corresponde asignar a la zona paga. Al analizar por separado cada uno de ellos, sólo el servicio G22R tiene como origen la EIM La Cisterna, por lo que se convierte en el principal candidato a ser el servicio correcto. Finalmente, para corroborar esto, se ubican en un mapa los paraderos de bajada de este servicio y se comparan con los paraderos de subida de las siguientes transacciones. Tal comparación se da a conocer en la imagen 5-2, donde se aprecia claramente que la ruta del servicio G22R (en azul, de norte a sur el sentido) corresponde a los sitios RM-0720 y RM-0721. Existen algunos paraderos de subida de la siguiente transacción que están alejados de la ruta del servicio G22R, sin embargo se trata de distancias menores a 1.500 metros que perfectamente podrían ser caminables por los usuarios.

Imagen 5-2: Recorrido G22R ((M) La Cisterna – Villa España)



El siguiente análisis se realiza con el servicio 415 expreso sentido ida, cuya parada de inicio de servicio corresponde a la zona paga ubicada en Teniente Cruz con San Pablo, a las afueras de la estación de Metro Pudahuel. Este servicio, para la fecha de estudio (Junio de 2010) presenta como destino la estación de Metro Manquehue, en la comuna de Las Condes.<sup>5</sup> El análisis de las validaciones de los sitios asignados a este servicio se presenta en la tabla 5-9. Se puede ver que si bien en orden de magnitud las validaciones de este servicio son menores que para el caso anterior, el porcentaje de transacciones sin paradero de bajada es prácticamente del 100%. Al analizar el tipo de error presente en estos sitios se aprecia que la falla corresponde al error de bajada no caminable, es decir, hay más de 1.000 metros de distancia entre la posible parada de bajada y la ubicación del paradero de la siguiente transacción. Al repetirse este mismo error en todas las transacciones se presume que hay problemas con la ruta asignada.

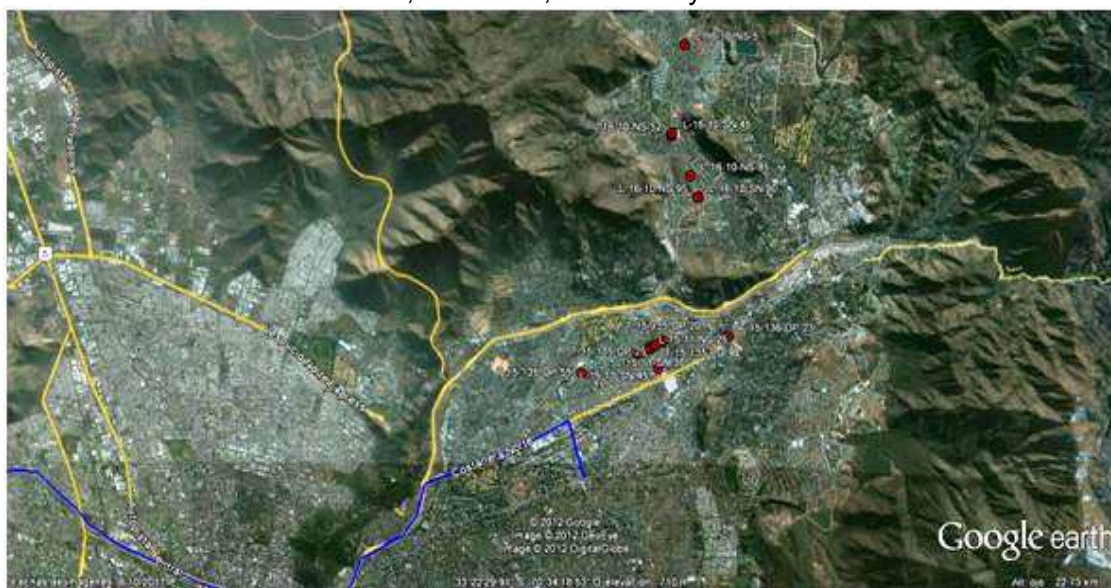
<sup>5</sup> 15° P.O. Modificado Troncal 4 Anexos 1 y 2. Sitio web Transantiago.

Tabla 5-9: Sitios y transacciones para servicio 415el en Metro Pudahuel.

Sitio	Servicio	Sin Asignar	Total Validaciones	% Sin asignar
RM-0641	415el	2556	2571	99,42
RM-0642	415el	1572	1584	99,24
RM-0643	415el	650	652	99,69
RM-0758	415el	212	212	100,00

De la misma forma que para el caso anterior, la imagen 5-3 da a conocer el esquema de análisis de transacciones, donde se observa que en rojo se muestran los paraderos de subida de la siguiente transacción que presentan mayor frecuencia y la línea azul que corresponde a la ruta final del recorrido 415el. Con esto se busca saber qué paraderos son los que aborda la gente y con eso tener una idea de el o los servicios correspondientes a esa zona paga.

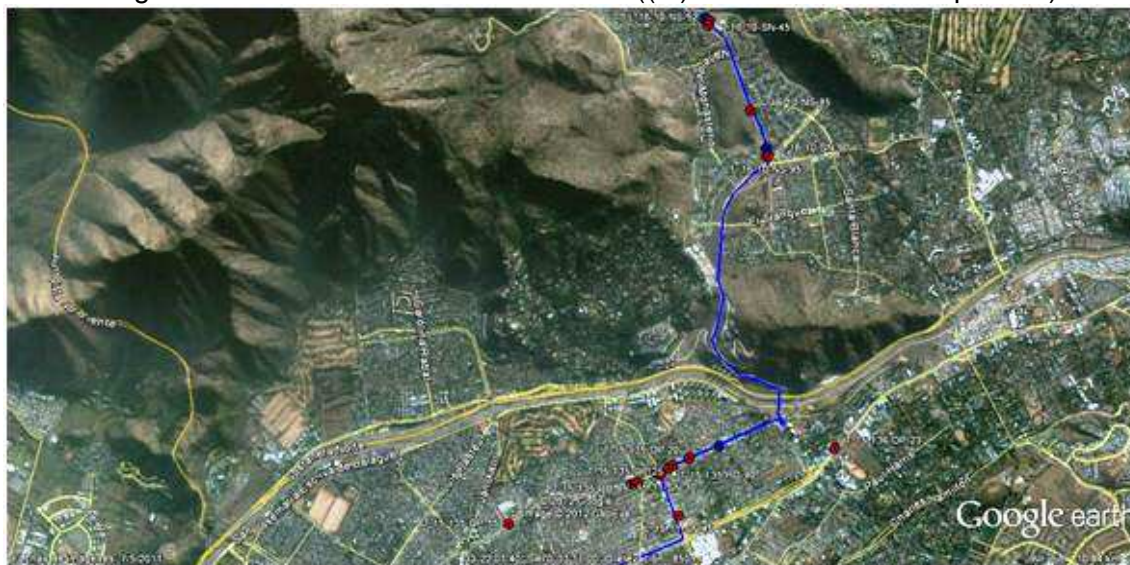
Imagen 5-3: Paraderos con la mayor cantidad de transacciones posteriores a hechas en sitios RM-0641, RM-0642, RM-0643 y RM-0758.



Finalmente, buscando todos los servicios que tienen como origen la zona paga ubicada en Teniente Cruz con San Pablo, se encuentra que aquel servicio que tiene paraderos cercanos a los grupos de puntos rojos (paraderos) descritos anteriormente en la imagen 5-3 es el recorrido 414el, cuya ruta se presenta en la imagen 5-4.

Los servicios mal asignados y los que son los correctos, además de la cantidad de validaciones que presentan bajada nula producto de esta falla se presentan en la tabla 5-10. Se puede ver que también hay casos donde hay menos servicios asignados de los correspondientes. Haciendo el mismo análisis explicado en los ejemplos se llega a la conclusión que la zona paga ubicada en Santa Rosa con Américo Vesputio (sitios RM-0408, RM-0409, RM-0410, RM-0756 y RM-0757) debe ser complementada con los servicios 2071, 2091 y sus variantes cortos y expreso. Esto debido a que muchas posibles bajadas de esta zona paga se ubican en la ruta de ambos servicios, logrando incluso obtener que el paradero de destino de mayor concurrencia es el fin del recorrido de este par de servicios. El mismo criterio corre para los sitios RM-0147 y RM-0148.

Imagen 5-4: Servicio 414el sector Oriente ((M) Pudahuel – Los Trapenses)



El total de transacciones sin estimación de paradero de bajada producto de este error en la entrega de datos corresponde a 40.205 validaciones, que corresponden al 8,43% del total de estimaciones de bajada nula producidas por transacciones en zonas paga. Considerando el total de transacciones hechas en zona paga en la semana de estudio, este problema corresponde al 1,98% de las transacciones de zona paga.

Tabla 5-10: Servicios mal asignados a zonas paga y transacciones sin bajada.

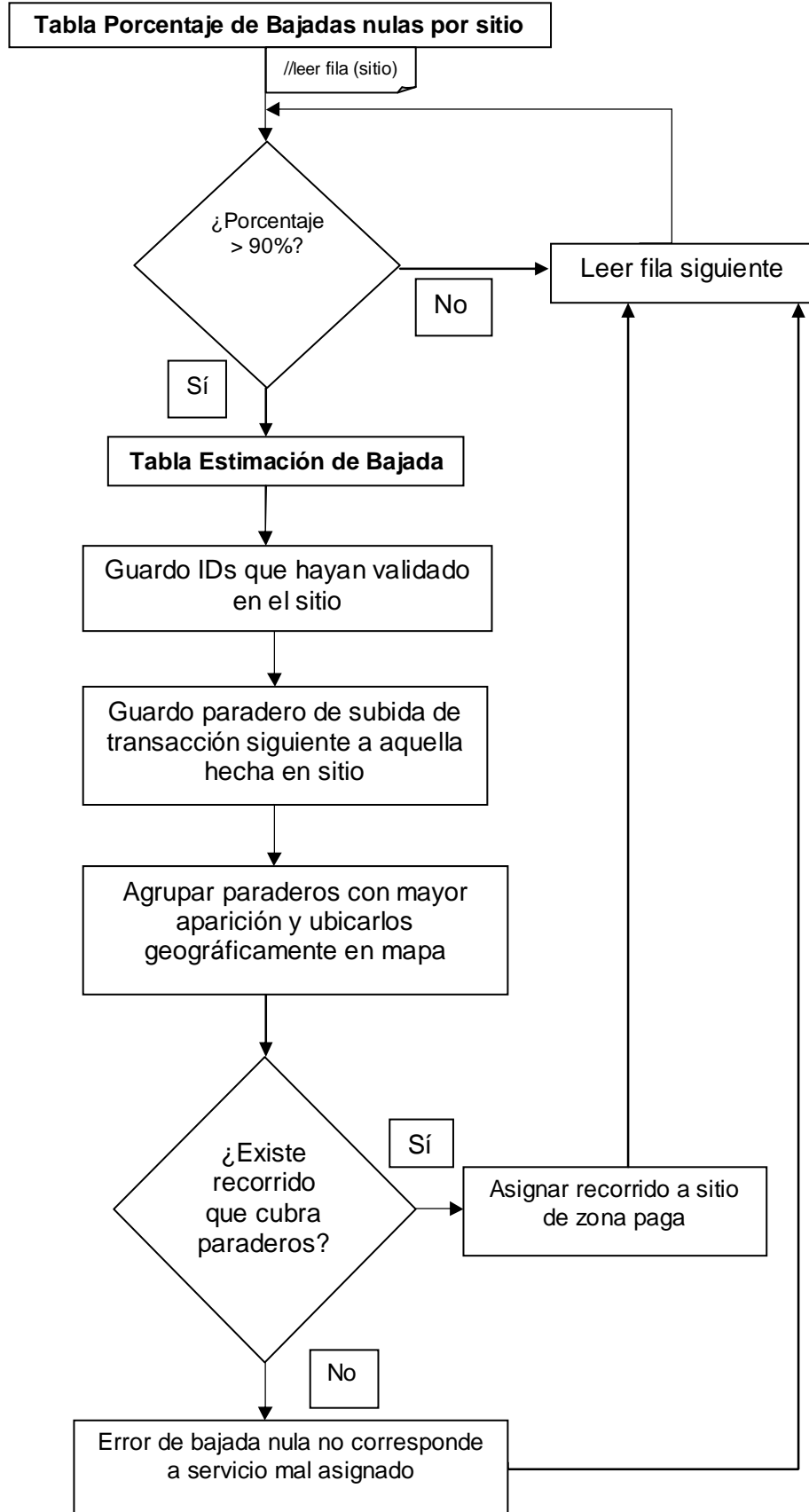
<b>Sitio</b>	<b>Servicio Asignado</b>	<b>Servicio Correcto</b>	<b>Bajadas sin estimar</b>
RM-0639	414el-420el	415el	2007
RM-0640	414el-420el	415el	1604
RM-0641	415el	414el	212
RM-0642	415el	414el	2556
RM-0643	415el	414el	1572
RM-0758	415el	414el	650
RM-0187	416eR	414eR	1149
RM-0188	416eR	414eR	1295
RM-0175	420eR	416eR	845
RM-0176	420eR	416eR	1257
RM-0720	G04I	G22R	8118
RM-0721	G04I	G22R	7379
RM-0757	206I-212I	206I-212I-207I-207cl-207el-209I-209el	3151
RM-0756	206I-212I	206I-212I-207I-207cl-207el-209I-209el	2970
RM-0410	206I-212I	206I-212I-207I-207cl-207el-209I-209el	2075
RM-0408	206I-212I	206I-212I-207I-207cl-207el-209I-209el	1076
RM-0409	206I-212I	206I-212I-207I-207cl-207el-209I-209el	1052
RM-0148	217el	217el-218el	767
RM-0147	217el	217el-218el	470



La solución para este problema tiene dos enfoques: uno metodológico para resolver los datos erróneos y uno basado en la captura de datos. El primer enfoque tiene como objetivo reparar aquellos datos que vienen erróneos. Este proceso se basa en analizar estadísticamente aquellas zonas paga que presenten mayor porcentaje de estimaciones de bajada nulas. El porcentaje de datos nulos puede variar entre zonas paga, pero se espera que para porcentajes de asignación de bajada nula superior al 90% el problema corresponda a servicio mal asignado. La solución a este problema se da a conocer en la figura 5-1, en la cual se presentan los pasos a seguir para encontrar aquellas zonas paga con servicios mal asignados. En la figura se aprecia que el procedimiento se basa en analizar aquellas zonas paga con alta tasa de estimaciones nulas de bajada, revisar el comportamiento de esos viajes sin estimación de bajada (transacción posterior), buscar los paraderos de aquellas transacciones y asignar un recorrido a la ruta que forman los paraderos encontrados. El hecho de que este procedimiento sea inspectivo no logra asegurar una eficacia total en la asignación de recorridos correctos, pues el que la asignación de bajada no sea estimada se puede deber a otros factores, pero es una herramienta que cumple el propósito de corregir datos que vienen con errores del cual no se tiene conocimiento previo al momento de correr la rutina computacional de asignación de bajada.

El segundo enfoque, en el cual se pretende que se eviten estos errores en el futuro corresponde a la capacidad de quien recolecta los datos de asignar correctamente cada sitio a su o sus recorridos. El procedimiento pasa principalmente por realizar una revisión previa de los recorridos asignados a cada zona paga. El problema radica en que puede ser que en terreno no se estén instalando los validadores correspondientes en los paraderos asignados. Al producirse un intercambio de validadores entre paraderos en una estación de transbordo los recorridos asignados a cada una de las zonas paga se cambian. Esto para el operador no es un problema, pues al tratarse todos los validadores de recorridos pertenecientes a su troncal no se ven afectados monetariamente, pero el modelo sí se ve afectado. Para eliminar este problema, una solución más factible y de menor costo temporal es generar para cada grupo de zonas paga que se encuentren cercanas y que además pertenezcan sus validadores a la misma unidad de negocios, un sitio o paradero unificado donde se asignen todos los servicios del operador, y que luego el modelo de asignación de paradero basado en líneas comunes se encargue de asignar el servicio correspondiente y la bajada en base a la siguiente transacción. Esto si bien deja a criterio del modelo el asignar un servicio a cada transacción, generaría una mayor eficacia en este proceso en vez de revisar una por una todas las zonas paga y observar si están los sitios correspondientes a cada servicio.

Figura 5-1: Método de detección de servicio mal asignado a sitio de zona paga.



#### **5.3.4. Emisiones GPS perdidas.**

Para que el modelo pueda estimar una bajada, tanto del punto de vista geográfico como del punto de vista temporal, es necesario saber la ruta que realiza el bus y su ubicación en ciertos instantes de tiempo. Para eso se utilizan las emisiones GPS que realizan los buses de Transantiago. Estas emisiones GPS son generadas por un dispositivo al interior del bus cada 30 segundos, y dan a conocer las coordenadas geográficas del bus en ese instante (latitud y longitud), además de la hora, minuto y segundo de la emisión. Además cada emisión GPS entrega información de la patente del bus, el operador al cual pertenece el bus y un valor binario de ignición del bus (motor del bus prendido si/no) (Mora, 2010).

Con esta información se puede conocer la posición del bus, lo que permite observar la ruta del mismo a través del conjunto de emisiones y es posible asociar la pasada del bus por un paradero con la transacción realizada al interior del bus de manera temporal, asociando los tiempos de emisión GPS y validación y buscando una cercanía temporal entre ambos para obtener la subida y la bajada.

Sin embargo, cuando el bus no emite señales de GPS, el modelo no es capaz de obtener la posición del bus, por lo que se pierde el recorrido. Además, el modelo actualmente es bastante rígido a la hora de filtrar aquellos buses con problemas de emisión de GPS: cuando un bus tiene problemas con alguna emisión de GPS se elimina todo el recorrido del mismo, por lo que no se estima bajada en ninguna de las transacciones asociadas a ese bus.

Un primer análisis hace referencia a aquellos recorridos con mayor tasa de problemas de GPS. Estos recorridos y la cantidad de transacciones afectadas se dan a conocer en la tabla 5-11. Aquellos recorridos que aparecen con dos servicios corresponden a zonas paga donde no se realizó la estimación de bajada, por lo que no se pudo asignar recorrido y se mantuvieron los correspondientes a esa zona paga.

Una característica importante de la mayoría de estos recorridos es que utilizan los túneles ubicados en las principales autopistas concesionadas de Santiago. En la tabla 5-12 se describe la utilización de estos túneles por recorrido y el largo de cada uno de los tramos en que se utiliza túnel por servicio. En ella se puede ver que la correlación entre los servicios con problemas de datos de recorrido y aquellos que utilizan túneles es evidente. Sin embargo, la causa de por qué esto ocurre no está completamente clara, pero se supone que corresponde a una falla en la recepción de la emisión GPS debido al impedimento físico existente.

Si bien existen recorridos que no utilizan zonas paga y se ven afectados por esta causa, la mayoría de ellos corresponde a aquellos que las utilizan. Esto, basado en que una de las funciones de las zonas paga es el mejoramiento en el tiempo de acceso de los usuarios cuando las demandas de paraderos son altas, situación que ocurre en estos casos de recorridos largos y que utilizan las autopistas concesionadas para disminuir los tiempos de viaje, donde las demandas son altas debido a esos factores y a que llegan generalmente a puntos neurálgicos de la ciudad.

Tabla 5-11: Recorridos con mayor proporción de errores en posicionamiento.

Recorrido	Falla GPS	Total	Porcentaje
410I	26814	30780	87,12%
414eR	3280	3863	84,91%
410R	39463	46862	84,21%
414el-420el-415el	4086	5100	80,12%
417el	7497	9460	79,25%
414el	45	60	75,00%
416el	4589	6188	74,16%
305cl	6901	9347	73,83%
305el	9938	13470	73,78%
402R-409I	5721	8103	70,60%
117I	25253	36196	69,77%
415eR	4731	6954	68,03%
417eR	4412	6514	67,73%
420el	43	66	65,15%
416eR	2917	5362	54,40%
218eR	1	2	50,00%
309R	29515	63118	46,76%
414el-420el	1651	3662	45,08%
409I	11413	26571	42,95%
309I	29885	71589	41,75%
420eR	928	2224	41,73%
315eR	2303	6730	34,22%

Tabla 5-12: Utilización de túneles por servicios Transantiago.

Túnel	Servicio	Tramo		Largo [m]
<b>San Cristóbal</b>	305c – 305e – 309 – 117	Completo		1825
<b>Costanera Norte</b>	312el – 408I	Enlace Vivaceta	Salida Mercado Central	770
	315eR – 410R	Ingreso Purísima	Enlace Vivaceta	1690
	409I – 410I	Ingreso Vivaceta	Salida La Concepción	3810
	414el – 415el – 416el – 417el	Enlace Vivaceta	Enlace Lo Saldes	5080
	414eR – 415eR – 416eR – 417eR	Enlace Lo Saldes	Enlace Vivaceta	5670
<b>Trinchera G. Velázquez</b>	313el	Paso Carrascal	Salida Mapocho	650
	313eR	Paso San Pablo	Paso Carrascal	1690
	218e – 221eR	Paso Antofagasta	Paso Ecuador	1588
<b>Tupper - Copiapó</b>	201el – 217eR – 222eR	Completo		490
	201eR – 217el	Completo		260

Fuente: Elaboración propia.

En lo relacionado a este problema se tienen que existen 233.917 transacciones sin estimación de bajada cuyo problema es generado por no existir datos producto de la presencia de túneles en el recorrido. Esto corresponde al 3,55% de las estimaciones de bajada no obtenidas en el total de la semana. Desde el punto de vista de las zonas paga son 55.078 transacciones afectadas por este problema, que corresponden al 11,55% de las estimaciones nulas de zonas paga.

La solución a plantear en este problema pasa por definir los tramos de ruta de cada uno de estos recorridos que pasan por túneles, para predefinir la ruta en esos casos sin emisiones GPS. El principal problema de esta solución radica en que al eliminar las emisiones de posicionamiento del bus se pierde la posibilidad del modelo de estimar el paradero de bajada del usuario al hacer el cruce entre las informaciones existentes (transacciones y posiciones). Sin embargo, dentro de los túneles no existen paraderos ni de subida ni de bajada, por lo que aquella información no se verá perjudicada a la hora de plantear la solución.

## 6. Conclusiones.

### 6.1. Comentarios y conclusiones.

El modelo de asignación de paradero de bajada, que utiliza como información tanto las transacciones realizadas por los usuarios del sistema Transantiago como las emisiones del sistema de posicionamiento global (GPS) presente en cada uno de los buses, se presenta como una poderosa herramienta para generar una gran cantidad de datos relativos al uso del transporte público integrado al pago con tarjeta bip! en el Gran Santiago. Entre los principales datos que se obtienen están los orígenes y destino de cada una de las transacciones, llegando hasta el nivel desagregado de paradero y estación de metro, lo que lo convierte en una de las mejores formas de obtener datos acerca de los viajes de los habitantes de Santiago.

Sin embargo, como todo modelo computacional, está afecto a supuestos que no siempre se cumplen a cabalidad, pues los comportamientos de los usuarios al momento de viajar y de pagar no siempre son los esperados. Además, la ciudad presenta distintas complicaciones a la hora de obtener las rutas de viaje de los buses, y éstas pueden ser topográficas, estructurales o viales. Es por esto que en este trabajo se logran analizar algunos de estos problemas con el fin de darles una solución que se acerque de mejor manera a lo que sucede en el día a día.

Se tiene que el modelo es capaz de estimar un 82,7% de los paraderos de bajada asociado a cada una de las transacciones existentes en la semana de estudio. Este trabajo se preocupó de analizar dos caminos de error en la estimación de bajadas: en primer lugar aquellas transacciones que no se les estima bajada y en segundo lugar transacciones que si bien poseen estimación de bajada, esta parece no corresponder a lo esperado.

Este trabajo está enfocado en el análisis de tres situaciones en que el modelo de estimación de bajada presenta problemas y por lo tanto se requiere atención particular para lograr mejorar la estimación en el modelo. Esto generando cambios en la estimación de bajada o suprimiendo bajadas estimadas no fiables.

La primera situación que corresponde a fallas particulares del modelo debido a que no se interiorizan ciertos comportamientos de los usuarios en éste. Dentro de esto se observan aquellos sitios de alto flujo peatonal que, por la naturaleza de los mismos, están hechos para realizar caminatas largas. Estas caminatas podrían estar fuera del rango de transbordo caminable para el modelo, pues éste supone que la gente se mueve en un radio cercano al paradero de bajada para realizar sus actividades o un transbordo a otro medio de transporte. Con la ubicación de estos sectores y la liberación de este filtro de caminata se espera que se logre obtener un aumento en las estimaciones de bajada. Se espera lograr con este proceso una estimación del 6,75% de aquellas estimaciones no caminables.

El segundo problema que busca obtener mejoras se basa en cambiar el horario del día de viaje de un usuario. Actualmente la primera transacción realizada después de las 00:00 horas se comporta como primera transacción del día, y es utilizada por el

modelo para estimar la bajada correspondiente a la última transacción del día. Sin embargo, ocurre que generalmente estas transacciones posteriores a medianoche preceden de viajes comenzados el día anterior, por lo que afectan la estimación de bajada de la última transacción del día. Para esto se propone cambiar el día de viaje y hacerlo comenzar a las 04:00 horas, y terminarlo a las 03:59 horas del día siguiente, con el fin de interiorizar el comportamiento social de las personas a la hora de realizar viajes. Si bien estas transacciones no representan un valor muy importante sobre el total semanal (0,53%) es importante entender que es un tramo horario importante que no está siendo bien utilizado en el modelo.

La tercera causa de error en el modelo se refiere a aquellas transacciones únicas en el día. Actualmente el modelo es rígido en este aspecto, anulando la bajada de aquellas transacciones que son únicas en el día por considerar que no hay un viaje de ida y vuelta. Sin embargo, existe la posibilidad de analizar las bajadas de estas validaciones únicas del día cuando son realizadas en periodo punta tarde o más adelante, y estimar bajada basándose en la validación del día siguiente que haya sido realizada durante la mañana. Esto, suponiendo que el usuario realizó con esa transacción única su último viaje. Con este proceso se espera que como máximo el 7,02% de las transacciones únicas del día se les pueda estimar bajada.

El cuarto comportamiento analizado corresponde al análisis de aquellas transacciones que presentan mismo paradero de subida y bajada. Acá se concluye que la causa es la ausencia de validaciones en una de las etapas del viaje, ya sea por la omisión en el pago de una de las etapas del viaje, o por el uso de otro medio de transporte externo al sistema integrado de pago de Transantiago que acerca a un medio integrado en el sistema. Esta corrección lo que busca es identificar aquellos casos donde ocurre esta situación y eliminar el paradero de bajada asociado a la transacción pues se trata de una bajada no válida que ensucia los datos y perjudica trabajos futuros con ellos. El resolver este problema generará la pérdida de 1.074.347 bajadas estimadas, pero que de alguna manera dejarán de causar problemas en los datos.

En resumen, se espera que estas modificaciones al modelo permitan ganar como máximo 332.457 nuevas estimaciones de bajada con una capacidad de predicción mejorada. Además habría un posible cambio de estimación de bajada en 223.385 transacciones debido al nuevo día de viajes. Finalmente se tendrá que 1.074.347 estimaciones de bajada erróneas desaparecerán de la base de datos.

La segunda gran investigación realizada corresponde a un proceso denominado Multipago. Este comportamiento, que corresponde al pago con una tarjeta bip! el pasaje de dos o más pasajeros, se da de dos maneras diferentes dependiendo de las necesidades de viaje de los usuarios.

Un primer comportamiento tiene relación con el viaje en grupo utilizando una sola tarjeta. Se llega a conclusiones que esto sucede principalmente en parejas (95,2% de los casos) y que se trata de viajes con fines de ocio y concentrados en comunas con importantes destinos comerciales y turísticos, lo que hace razonar que la gente viaja en grupo para estos comportamientos. Como el modelo no es capaz de estimar la bajada del grupo de transacciones, una vez identificado el grupo y bajo ciertos casos se

pueden obtener las bajadas de todos los integrantes del viaje, tomando como base que el grupo continúe su viaje en la transacción grupal siguiente. Se logran identificar 355.743 transacciones grupales, a las cuales se le es posible obtener paradero de bajada, dependiendo si los grupos mantienen su tamaño durante el viaje o no.

Otro comportamiento que incluye el uso de una tarjeta para que dos personas realicen sus respectivos viajes corresponde a multipago en sitios distintos. Este comportamiento se basa en que una tarjeta emita dos pagos consecutivos en validadores cercanos tanto en distancia como en tiempo. Esto hace suponer que se utilizó la tarjeta para viajes distintos, lo que implica que sólo a una de las transacciones puede estimársele bajada, pues el otro viaje no registrará nueva transacción conocida. En base a cada uno de los comportamientos presentados se tiene que existen 97.506 transacciones que cumplen este criterio y que perderán su estimación de bajada por no corresponder a una etapa de viaje.

Finalmente, el trabajo realizado en zonas pagas presenta dos aristas, sin embargo ninguna de ellas tiene que ver con el proceso previo que corresponde a la asignación del servicio que abordó el pasajero, sino más bien con errores de datos y del modelo desde el punto de vista geométrico. A pesar que se realizó un análisis de un grupo de tarjetas ID para observar posibles problemas en la asignación de servicio, que era uno de los temores iniciales al comenzar este trabajo, las fallas observadas correspondían a cuestiones metodológicas.

Luego de un análisis previo de una muestra de tarjetas se logran identificar los principales problemas asociados a las zonas pagas. Los dos problemas abordados tienen relación con los datos que utiliza el modelo para asignar bajada. En primer lugar, las fallas existentes por la mala asignación de recorridos a cada una de las zonas pagas existentes. Al estar cambiadas, o faltar servicios, el modelo utiliza servicios no correspondientes para asignar servicios a las transacciones, lo que implica que las bajadas no son las correctas. Este problema afecta a 19 sitios de pago que suman 40.205 transacciones en la semana de estudio.

Y el segundo problema que si bien no es producto de la zona paga si lo afecta principalmente es el hecho de que al pasar por los túneles existentes en Santiago los servicios generen problemas con las emisiones de GPS, lo que afecta el trazado estimado del bus y por ende la estimación de bajada desde el punto de vista geográfico y temporal. Este problema afecta a 27 servicios-sentido que ocupan los cuatro principales túneles de la ciudad, generando estimación de bajada nula para 233.917 transacciones. La solución a este problema pasa por identificar estos túneles y sectorizar los recorridos con el fin de ignorar fallas existentes en esos sectores.

En resumen, la tabla 6-1 da a conocer las transacciones en las que se ganarán bajadas, en las que se perderán bajadas y en las que se recalcularán. Estos valores posiblemente no sean los que se aprecien una vez que se implementen los cambios propuestos en este trabajo, pero dan a conocer un orden de magnitud de cada uno de los problemas. En ella se aprecia que aproximadamente se podrían ganar 2 millones de transacciones con nueva estimación de bajada. Sin embargo, producto de las características del modelo podría existir un número importante de transacciones



recalculadas que actualmente tengan estimación de bajada y producto de las correcciones las pierdan, pero esos valores se lograrán obtener sólo al momento de correr el modelo con los cambios propuestos.

Tabla 6-1: Resumen de cambios en bajada producto de cambios al modelo.

Corrección	Estimaciones de bajada		
	Ganadas	Perdidas	Recalculadas
Bajada no caminable	188.060		
Primera transacción del día			147.348
Validación única del día	144.397		
Paradero subida y bajada cercano			1.074.347
Viaje grupal	355.743		
Multipago sitio cercano		97.506	
Zona paga equivocada	40.205		
Túnel	233.917		
<b>Total</b>	<b>962.322</b>	<b>97.506</b>	<b>1.221.695</b>

Como comentario final, si bien el modelo es una herramienta muy poderosa para analizar el movimiento de los usuarios de transporte público en Santiago, nunca es algo perfecto, pues siempre se van a encontrar comportamientos ajenos a lo esperado para un viaje normal en estos modos. Es por eso que en este trabajo se intentó “humanizar” el modelo para adecuarlo de cierta forma a los distintos movimientos que realizan los usuarios de Transantiago en sus respectivos viajes en la ciudad.

## 6.2. Líneas futuras de investigación.

Siempre en estos modelos que utilizan y generan grandes cantidades de información quedan tareas pendientes que vale la pena dejarlas en conocimiento, y que en este trabajo no se analizaron porque estaba fuera del alcance del trabajo de título presente.

En primer lugar las propuestas aquí presentadas se deberán implementar en el modelo de asignación de bajada para generar las nuevas bajadas y poder analizar los resultados obtenidos, comparándolos con los estimados inicialmente.

En el aspecto de las bajadas no caminables es esperable que existan otras zonas donde se pueda mejorar el rango de distancia caminable, en particular son dos tipos de zonas a las que puede ser necesario aplicarles un estudio más acabado: las comunas periféricas de Santiago y las fajas ubicadas entre dos avenidas importantes. Se tiene conocimiento de que en esas zonas se dan importantes caminatas por factores asociados a baja accesibilidad en el caso de la periferia y a variada oferta en el caso de las fajas entre avenidas importantes.

El análisis de etapas de viaje no capturadas por las transacciones realizadas se convierte en un tema que debe ser profundizado. La falta de información referente en

primer lugar a la causa de esta falta de etapa de viaje y del modo utilizado para aquel viaje son dos líneas investigativas que deben ser miradas desde un punto de vista más desarrollado, utilizando para eso información más acabada referente a tasas de movilización en otros modos no integrados en Transantiago (colectivos, taxis, buses no integrados a TS, etc.) y a la evasión, ya sea ésta en alguna de las etapas del viaje o en todo el viaje. Si bien en ambos frentes ya hay trabajos realizándose (Encuesta Origen Destino 2012-2013 y trabajo investigativo referente a la evasión) éstos deben ser integrados en el modelo para poder mejorar su trabajo estimativo.

El mejoramiento de la información sobre la red de Metro, referida principalmente a los tiempos de traslado permitirá obtener información más precisa sobre la forma en que se mueven los usuarios dentro de la red, obteniendo información sobre la manera en que combinan entre líneas y la elección de rutas que hacen más eficientes su viaje.

Parte importante en el trabajo futuro de este modelo es validar exógenamente sus resultados, basándose en datos obtenidos a partir de diarios de viajes que se pueden conseguir a partir de encuestas de viajes, como la Encuesta Origen Destino o la Encuesta de Metro, o a partir de experiencias de usuarios que den a conocer sus viajes. Todo esto con el fin de validar los resultados obtenidos y estudiar aquellos comportamientos que no esté recogiendo el modelo.

Finalmente, se deben revisar problemas en los datos entregados para el trabajo del modelo. El hecho de que cada vez mejoren los inputs permitirá ahorrar tiempo importante en la revisión y en el post proceso de la información generada. Actualmente existen problemas relacionados con paraderos de buses de los que se tiene poca o nula información. Sin duda cada vez que se generen mejores datos de entrada el modelo hará mejor su trabajo.

## 7. Referencias Bibliográficas.

Beltrán, P., Cortés, C., Gschwender, A., Ibarra, R., Munizaga, M., Ortega, M., Palma, C., Zúñiga, M., 2011. Obtención de información valiosa a partir de datos de Transantiago, XV Congreso Chileno de Ingeniería de Transporte.

Chapleau, R., Chu, K.K., 2007. Modeling transit travel patterns from location-stamped smart card data using a disaggregate approach. 11<sup>th</sup> World Conference on Transportation Research, Berkeley, California.

Deakin, E., Kim, S., 2001. Transportation Technologies: Implications for Planning. University of California Transportation Center, Paper #536, 27 p.

DICTUC, 2003. Actualización de encuestas Origen Destino de viajes, V Etapa. Informe final a Sectra, Santiago.

DICTUC, 2008. Actualización y recolección de información del sistema de transporte urbano, III etapa. Informe final a Sectra, Santiago.

McDonald, N., 2000. Multipurpose Smart Cards in Transportation: Benefits and Barriers to Use. University of California Transportation Center Research Paper #630, 27 p.

Ministerio de Vivienda y Urbanismo, 2009. Recomendaciones para el diseño de elementos de infraestructura vial urbana. Santiago, Chile.

Mora, P., 2010. Generación de datos de patrones de viaje a partir de transacciones BIP. Memoria para optar al título de Ingeniero Civil. Santiago, Chile. Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas.

Munizaga, M., Palma C., 2012. Estimation of a disaggregate multimodal public transport Origin-Destination matrix for passive smartcard data from Santiago, Chile. Transportation Research C 24, 9-18.

Ortúzar, J. de D., Willumsen, L.G., 2011. Modelling Transport, fourth ed. Wiley, Chichester.

Pelletier, M.P., Trépanier, M., Morency, C., 2011. Smart card data use in public transit: a literature review. Transportation Research C 19, 557-568.

Plano regulador PRS-01F de zonificación general, 2012. Asesoría urbana, Ilustre Municipalidad de Santiago. Esc. 1:7800.

Plan regulador comunal de Providencia. Espacio privado: zonas de uso de suelo, 2007. Ilustre Municipalidad de Providencia. Esc. 1:5000.

SECTRA, 2012. Información de Transporte Urbano [en línea] Santiago, Chile. <<http://www.mtt.gob.cl/transporteurbano/>>. Visitada el 5 de julio de 2012.

TRANSANTIAGO, 2008. Zonas Pagas versión 4.01 [en línea] Santiago, Chile. <<http://www.coordinaciontransantiago.cl/corporativo/index.php>>. Visitada el 26 de abril de 2012.

TRANSANTIAGO, 2010. Décimo cuarto programa de operación [en línea] <<http://www.coordinaciontransantiago.cl/corporativo/index.php>>. Visitada el 7 de Diciembre de 2011.

TRANSANTIAGO, 2012. ¿Qué es el transporte público de Santiago? [en línea] <<http://www.transantiago.cl/QUIENESSOMOS/HISTORIA/index.htm>>. Visitada el 14 de Diciembre de 2011.