

Obtención de Información Valiosa a partir de Datos de Transantiago

Pablo Beltrán, Coordinación Transantiago (actualmente en Cityplanning)

Cristián E. Cortés, Universidad de Chile

Antonio Gschwender, Coordinación Transantiago

Richard Ibarra, Universidad de Chile

Marcela Munizaga, Universidad de Chile

Carolina Palma, Coordinación Transantiago (actualmente en Cityplanning)

Meisy Ortega, Coordinación Transantiago (actualmente realizando estudios de doctorado en MIT)

Mauricio Zúñiga, Universidad de Chile

Corresponding autor: Marcela Munizaga
mamuniza@ing.uchile.cl
Tel: +56-2-9784649; Casilla 228-3, Santiago

RESUMEN

En este artículo se presenta un resumen de la investigación conjunta que han sostenido la Universidad de Chile con el Ministerio de Transportes y Telecomunicaciones, orientada al desarrollo de métodos sofisticados para procesar información y estimar indicadores a partir de las bases de datos de posicionamiento GPS y transacciones bip! de Transantiago. La información pura de transacciones permite realizar análisis estadísticos de los perfiles horarios de la demanda y del uso del medio de pago y modos de transporte y servicios. Al cruzar esa información con la de posicionamiento de buses, es posible también incorporar la dimensión espacial al análisis. Se describe las bases de datos con las cuales se trabaja, y se muestra para distintos cortes temporales dos aplicaciones de la metodología en desarrollo; la primera es la estimación de velocidad comercial del sistema de transporte público, donde los análisis consideran casos de desagregación tanto espacial como temporal, obteniendo un muy buen diagnóstico del rendimiento del sistema de transporte público en este sentido; la segunda es el procedimiento de estimación de paradero de bajada, del cual se logra obtener perfiles de carga de buses y servicios, además de matrices origen destino de viajes en transporte público, que pueden ser desagregados espacial y temporalmente. Estos resultados son sólo una muestra del enorme potencial que tiene este procesamiento de datos.

ABSTRACT

This paper summarizes the research conducted at Universidad de Chile in collaboration with Ministerio de Transportes y Telecomunicaciones, oriented towards the development of sophisticated methods to process GPS and smartcard data and obtain level of service and operation indicators. Statistical analysis of the daily transaction patterns and use of the smartcard per mode are made directly using smartcard data, and very interesting information can be obtained from that analysis. However, when incorporating GPS data, the spatial dimension can also be incorporated. In this paper we describe the data available and show the results of applying the proposed methods

for different time cuts. This includes speed profiles, taking into account time and space disaggregation. It also includes a procedure to estimate alighting bus-stop, from which OD matrices and load profiles can be obtained. These results are only a small sample of what can be obtained from the data, and show the enormous potential of this data processing methodology.

Keywords: smartcard, GPS, speed of buses, OD matrix, public transport.

1. INTRODUCCIÓN

La utilización de información pasiva, generada automáticamente tras la implementación de sistemas de pago con tarjeta electrónica y equipamiento de GPS en buses, es un tema que despierta gran interés en investigadores, entidades reguladoras y operadores de servicios de transporte público mayor de superficie. La oportunidad de contar con información masiva a bajo costo ha sido identificada por varios autores (ver por ejemplo Bagchi and White, 2005; Park and Kim, 2008), incluso se ha desarrollado metodologías para obtener información valiosa a partir de esos datos (Trepanier et al, 2007).

Sin embargo, en la mayoría de los casos la información está limitada por la tasa de penetración del uso de la tarjeta electrónica como medio de pago por un lado, y por la disponibilidad de información de posicionamiento de vehículos por otro. En este sentido, Transantiago es un caso privilegiado, dado que la tarjeta bip! es el único medio de pago en buses, y en Metro tiene una tasa de penetración muy alta, de modo que cerca del 97% del total de las transacciones de pago del sistema se realizan mediante tarjeta bip!.

La Universidad de Chile y el Ministerio de Transportes y Telecomunicaciones tomaron el desafío que significa procesar esta gran cantidad de datos y obtener información valiosa; para lo cual han estado colaborando de forma sistemática, utilizando las nuevas tecnologías disponibles para contribuir al mejoramiento continuo del sistema. Para esto se ha contado con apoyo de fondos de investigación (PBCT, FONDECYT, MILENIO, FONDEF).

En este artículo se sintetiza los resultados obtenidos de esta colaboración hasta la fecha. En la sección 2 se describe los datos extraídos para estos fines de las bases de posicionamiento y tarjeta bip!. Luego en la sección 3 se hace un análisis de la información de transacciones para tres cortes temporales. En el punto 4 se presenta el análisis de velocidades para distintos cortes temporales. En la sección 5 se sintetiza los avances en el procesamiento del cruce de la información, incluyendo la estimación de paradero de subida y bajada, obtención de matrices origen/destino de viajes y perfiles de carga.

2. DESCRIPCIÓN DE DATOS

El sistema de transporte público Transantiago, implementado en Santiago de Chile en febrero 2007, es un sistema basado en una estructura de servicios de buses troncal-alimentador, donde el Metro es la red estructurante de los viajes de la ciudad. El sistema está conformado por nueve áreas de operación de alimentadores, cubriendo la totalidad de la ciudad, y por cinco operadores de servicios de buses troncales, que corresponden a empresas de mayor tamaño, que sirven los ejes principales de la ciudad y pueden tener superposiciones parciales de servicios entre ellos, y con las áreas alimentadoras.

El sistema de pago considera que cada pasajero paga (o válida) una tarifa al ingresar, permitiéndole hacer hasta dos transbordos en dos horas. La estructura de pago es ligeramente diferente en los buses y el Metro; en los primeros, la única forma de pago es el de una tarjeta electrónica denominada bip!, con modalidad de prepago, que registra todas las transacciones. En el Metro, en cambio, además de permitirse el pago con tarjeta bip!, se puede comprar un boleto unitario para el viaje, con el que no se accede a la gratuidad de los transbordos. La tarifa también es diferente por modo, siendo en general levemente superior la tarifa adulta del Metro. La tarifa para estudiantes es única y rebajada en todo el sistema.

Si un pasajero adulto utiliza un bus en su primera etapa y luego transborda a Metro, al momento de acceder a Metro se le carga la diferencia entre las tarifas. Dado que todas las líneas de Metro están conectadas físicamente, los cambios de línea no requieren el uso de la tarjeta bip!. Pero, si algún pasajero re-ingresa a Metro o se repite un mismo servicio de buses, se le volverá a cobrar la tarifa, aunque esté dentro de la ventana de tiempo de dos horas.

El inicio de Transantiago estuvo marcado por serios problemas en su funcionamiento, y una de las consecuencias de esas dificultades iniciales fue que se instaló en parte de los usuarios la conducta de evadir, que pasó a ser un problema grave en los buses, no así en Metro, donde existe control físico de acceso (torniquetes) y guardias. Según mediciones encargadas por el Ministerio de Transportes y Telecomunicaciones, la evasión no se distribuye homogéneamente en la ciudad, encontrándose diferencias significativas por zona de la ciudad y por operador (Fundación Chile, 2010).

El sistema Transantiago contempla por diseño la recolección automática de información, tanto de pago de la tarifa como de posicionamiento de los buses. El medio de pago está diseñado de forma tal que la entidad recaudadora almacena la información de todas las transacciones bip!, identificando la hora de la transacción y la componente del sistema en que ésta se realizó (bus; estación de pago extra-vehicular o zona paga; o estación de Metro), además del tipo de usuario (adulto o estudiante) y el monto cobrado (tarifa o diferencial de tarifa). Por otra parte todos

los buses están equipados con sistema GPS, indicando la posición del vehículo.

La base de datos de posicionamiento contiene información geocodificada de la localización de los buses, que es registrada cada 30 segundos. Esta base de datos almacena cerca de 80 millones de registros por semana de más de 6.000 buses diferentes (patentes) y permite realizar un seguimiento de la ubicación de los buses. Adicionalmente, a partir de información complementaria, se puede identificar para la mayoría de los casos, el servicio o línea que cada bus está prestando.

El presente trabajo se ha desarrollado a partir de datos correspondientes a diversas semanas (lunes a domingo) extraídas de: septiembre 2008, marzo 2009, abril 2009, agosto 2009 y junio 2010. Los datos han sido utilizados para probar y desarrollar distintas componentes de la metodología, por lo que no todas las semanas serán mencionadas en todas las secciones, haciendo comparaciones sólo cuando hay coincidencia metodológica entre ellas.

3. ANÁLISIS GLOBAL DE DATOS DE TRANSACCIONES BIP!

En esta sección se sintetiza la información que puede ser obtenida de la base de datos de transacciones sin procesamiento adicional, ello implica que a este nivel no es posible saber con rigurosidad si una etapa corresponde a un viaje completo o sólo a parte de este, pues el monto del cobro no necesariamente está relacionado con el fin de un viaje, sino con reglas de pago. En la Tabla 1 se muestra el número de tarjetas bip! observadas en el sistema cada día de una semana para tres cortes temporales: marzo 2009, agosto 2009 y junio 2010. Para que una tarjeta sea registrada como observada en el sistema, debe haber realizado al menos una transacción de pago en ese día. Se ve que en un día laboral cualquiera, el número de tarjetas utilizadas bordea los 1,9 millones, mientras que en fines de semana se observan del orden de 1,3 millones en día sábado y 0,8 millones en día domingo. Para todos los tipos de día, los mayores valores se observan para junio 2010. Estas diferencias pueden deberse a cambios estacionales, disminución de la evasión o aumento en el número de usuarios. Con la información disponible, no es posible determinar cuál de estos efectos es predominante. Por otra parte, llama la atención la importante reducción del número de tarjetas bip! observadas en días de lluvia, como el día Martes de la semana correspondiente a Agosto 2009.

TABLA 1: Tarjetas bip! Observadas por Día de la Semana y Corte Temporal

Día	Marzo 2009	Agosto 2009	Junio 2010	Promedio
Lunes	1.884.367	1.882.538	2.005.399	1.924.101
Martes	1.853.728	1.572.904*	1.951.067	1.902.398
Miércoles	1.852.403	1.828.848	1.920.149	1.867.133
Jueves	1.840.945	1.878.540	1.929.650	1.883.045
Viernes	1.913.471	1.890.905	2.000.924	1.935.100
Promedio Laboral	1.868.983	1.870.208	1.961.438	1.902.352
Sábado	1.199.641	1.248.704	1.339.858	1.262.734
Domingo	807.917	800.770	875.166	827.951

*Día de lluvia, se excluye del cálculo de promedios y otros indicadores

Definiremos como “etapa” de un viaje, al acceso con el medio de pago (tarjeta bip!) a cualquier bus, zona de pago extra-vehicular (zona paga) o a la Red de Metro⁶.

⁶ Los transbordos entre líneas de Metro no se cuentan aquí como etapas adicionales, dado que no quedan registrados.

En la Figura 1 se muestra un histograma de uso de tarjetas, clasificándolas por la cantidad de transacciones (etapas) diarias realizadas para cada tipo de día y corte temporal. Se puede ver que la moda es dos en todos los cortes temporales y tipos de día. En día laboral el segundo valor más frecuente es cuatro. Por otra parte, hay un número importante de tarjetas que realiza sólo una transacción en el día observado, especialmente los días de fin de semana, en que uno es el segundo valor más frecuentemente

observado. Al analizar las diferencias entre cortes temporales, se observa nuevamente un mayor número de validaciones para la semana de junio 2010. Sin embargo, este incremento se presenta de manera proporcional para todas las tarjetas. Es decir, la distribución del porcentaje con respecto al total de tarjetas que realizan “n” etapas no varía considerablemente de un corte temporal a otro. Las diferencias se pueden atribuir al mayor número de tarjetas observadas en el sistema.

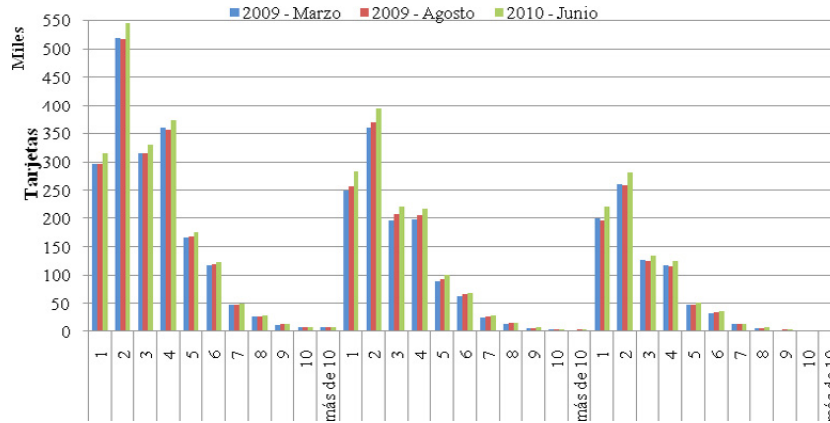


FIGURA 1: Uso de Tarjetas (número de etapas diarias) por Tipo de Día y Corte Temporal

En la Figura 2 se muestra el total de etapas diarias por tipo de día y corte temporal, incluyendo las etapas o transacciones con boleto unitario efectuadas en estaciones de Metro, que corresponden a aproximadamente un 3% del total de las etapas o transacciones. Cada una de las etapas de viaje contadas en la figura, está asociada a una transacción. El promedio diario de etapas, tanto con tarjeta bip! como con boleto unitario de Metro, para un día laboral en los tres cortes temporales bordea los 6,3 millones. Este cae a 3,9 millones para el día sábado y algo más de

2,3 millones en día domingo. Para los meses de marzo y agosto 2009 la cantidad de etapas diarias no tiene diferencias importantes, sin embargo en junio 2010 se observa un mayor valor, lo cual es consistente con lo reportado en la Tabla 1 y Figura 1. La diferencia en el total de transacciones, si se compara marzo 2009 con junio 2010, es de 4,5% en día laboral, 12,2% para el día sábado y 7,3% para el día domingo.

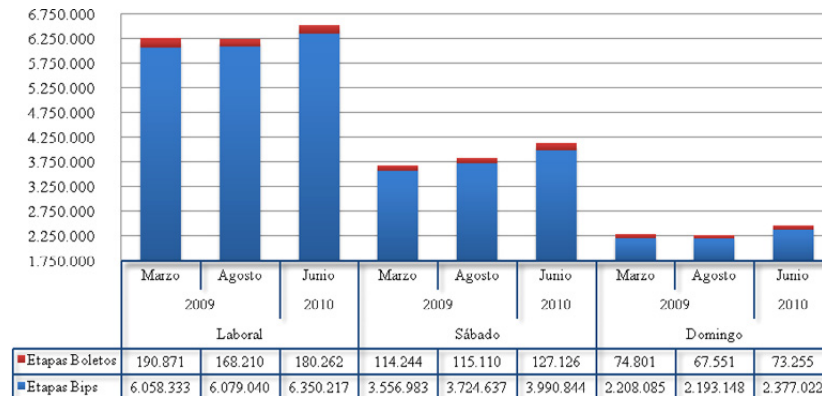


FIGURA 2: Total Etapas por Día y Corte Temporal (bip! y boletos unitarios Metro)

Analizando la cantidad de etapas por tramo tarifario, es posible realizar una primera aproximación a la cantidad de etapas por viaje que se realiza en el sistema. En la Tabla 2 se muestra la cantidad de etapas por tramo tarifario para cada uno de los cortes temporales analizados. Los valores de cada pago se agruparon por tramo tarifario considerando las alzas de tarifa que se produjeron durante los periodos de análisis. Se puede ver que, en promedio, para los tres cortes temporales, el 60% de las etapas corresponde a un primer pago (tarifa adulto bus o Metro, y tarifa escolar) y el 40% corresponde a transbordo. Estos valores no varían mayormente según tipo de día. Sin embargo, de marzo 2009 a

junio 2010 se produjo un incremento en la cantidad de primeras etapas, de un 59% a un 61% en día laboral, de 59% a 62% en día sábado y de 60% a 62% en día domingo. Esto último puede deberse a que entre un corte y otro se comenzó a cobrar por la repetición de un mismo servicio, entendiéndose que se trata de un nuevo viaje. El resultado más relevante a nivel de las etapas de viaje, teniendo a la vista la tarifa cobrada, es la reducción de las combinaciones Bus-Bus o Metro-Bus, a nivel de porcentaje respecto al total de etapas diario, seguramente explicada por las modificaciones operacionales entre Marzo 2009 y Junio 2010.

TABLA 2: Número de Etapas por Tipo de Día, Tarifa y Periodo *

Tipo Día	Tramo Tarifario	Marzo 2009	Agosto 2009	Junio 2010
Laboral	Tarifa Adulto	2.299.861 39,32%	2.247.701 38,54%	2.553.039 41,26%
	Tarifa Adulto Metro Horario Punta	351.889 6,02%	352.013 6,04%	386.715 6,25%
	Tarifa Escolar	806.390 13,79%	834.706 14,31%	819.568 13,24%
	Trasbordo Bus-Bus, Metro-Bus, Bus-Metro	2.216.697 37,90%	2.180.356 37,39%	2.192.604 35,43%
	Trasbordo Metro-Bus Horario Bajo	-	43.319 0,74%	47.323 0,76%
	Trasbordo Bus- Metro Horario Punta	174.501 2,98%	173.788 2,98%	189.092 3,06%
	Sábado	Tarifa adulto	1.610.890 47,13%	1.619.775 45,79%
	Tarifa escolar	414.128 12,12%	478.991 13,54%	475.734 12,30%
	Trasbordo Bus-Bus o Metro-Bus	1.392.611 40,75%	1.438.733 40,67%	1.482.077 38,30%
	Trasbordo Metro-Bus Horario Bajo	-	260 0,01%	439 0,01%
Domingo	Tarifa adulto	1.014.683 47,97%	970.126 46,71%	1.174.741 51,05%
	Tarifa escolar	257.749 12,19%	262.352 12,63%	252.762 10,98%
	Trasbordo Bus-Bus, Metro-Bus, Bus-Metro	842.629 39,84%	844.268 40,65%	873.485 37,96%
	Trasbordo Metro-Bus Horario Bajo	-	-	4 0,00%

* Se excluye boletos unitarios de Metro y escolares básicos.

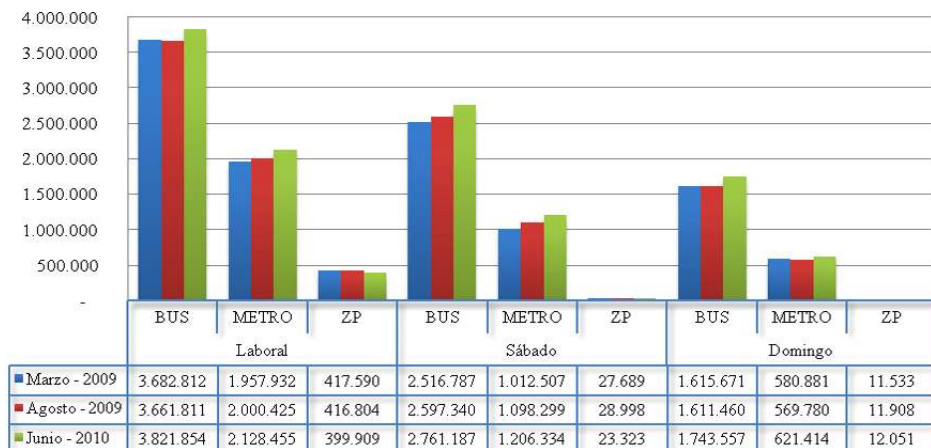


FIGURA 3: Transacciones por Tipo de Operador y Día

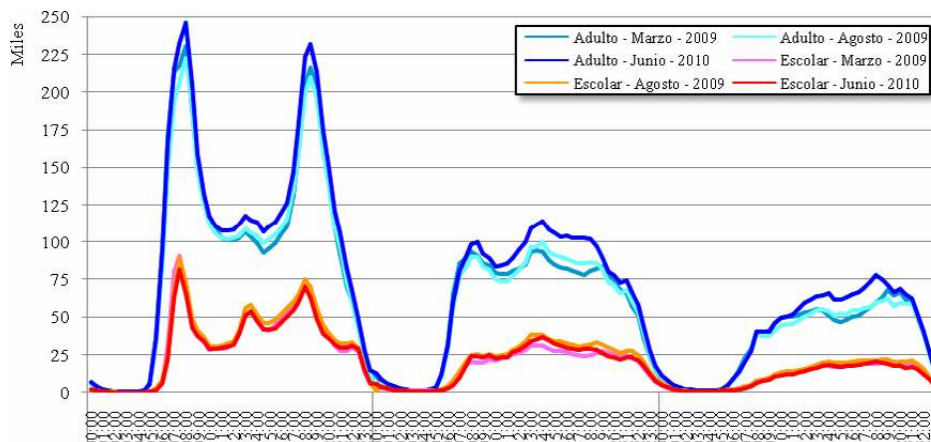


FIGURA 4: Promedio de Etapas por Usuario, Tipo de Día, Media Hora, y Corte Temporal

El viaje de los escolares de educación básica es siempre sin costo⁷, por lo cual esas transacciones no se presentan en la Tabla 2 y no fueron consideradas para el análisis de las etapas de viaje. El promedio de etapas realizadas en día laboral por escolares básicos para los tres cortes temporales corresponde a 206.009 transacciones, lo cual equivale a un 3,4% del total de etapas realizadas en día laboral. Este valor disminuye entre marzo 2009 y junio 2010 de 208.966 a 161.875 transacciones (-22,5%), probablemente debido a que esa semana el pase escolar había vencido y no todos estos usuarios habían hecho el trámite de renovación. Para día sábado el promedio temporal de transacciones realizadas por escolares básicos corresponde a 149.294 (4,0% del total de etapas día sábado), y para día domingo la cantidad de etapas efectuadas por escolares básicos es 95.152 (4,2% del total de etapas día domingo).

En la Figura 3 se muestra la distribución de etapas de viaje por modo en los distintos tipos de día. Se puede ver que en todos los cortes temporales y tipos de día, la mayoría de las transacciones se realiza en bus. En un día laboral promedio, la cantidad de etapas realizadas en bus corresponde a un 60% del total, aquellas etapas realizadas en Metro corresponden a un 33%, mientras que aquellas efectuadas en zonas paga alcanzan a un 7%. El fin de semana la participación del bus incluso aumenta con respecto al día laboral, especialmente el domingo.

En la Figura 4 se muestra los perfiles temporales de transacciones para cada tipo de día, por tipo de usuario, esto se puede entender como el perfil de hora de inicio de etapas de viaje. Se puede ver que no hay mayor variación entre cortes temporales y que la forma de la curva se mantiene estable para los distintos tipos de día (laboral, sábado y domingo). Asimismo, en los tres cortes temporales se observa que el periodo punta de los estudiantes es media hora más temprano que el de los pasajeros adultos.

En la Figura 5a se representa la participación y evolución de los tipos de transporte (Troncal, Alimentador y Metro) en los cortes temporales, observando una relativa estabilidad en la participación de ellos, cuya alteración en junio de 2010 se debe principalmente a las extensiones de las Líneas 1 y 5 de Metro, que añadieron 11,5 kilómetros y 8 estaciones al sistema de transporte público de Santiago, una de ellas compartida con otra línea ya existente.

En la Figura 5b se representa la participación por unidad de negocio, donde el anillo interior corresponde al primer corte temporal (marzo 2009) y el anillo exterior corresponde al más reciente (junio 2010). Se puede ver que las troncales 4 y 2 son las que tienen la mayor participación de mercado, en los tres cortes temporales. Sin embargo, la participación de la Troncal 2 se ha mantenido y la de la Troncal 4 ha disminuido, explicado por la misma extensión de la red de Metro.

Las unidades de negocio A2 (Zona E), A5 (Zona G) y A9 (Zona H) tienen el menor número de validaciones con respecto al total (3%, 3% y 2% del total, respectivamente). Sin embargo, no se observa una gran diferencia entre la cantidad de validaciones realizadas en las distintas unidades de negocio alimentadoras. En general, se observa estabilidad de la distribución desde un corte temporal a otro, no existiendo variaciones mayores a un 1%, con excepción de las troncales 3 y 6, que sufrieron cambios en sus contratos para el período de junio 2010 y el conjunto de servicios prestados es distinto.

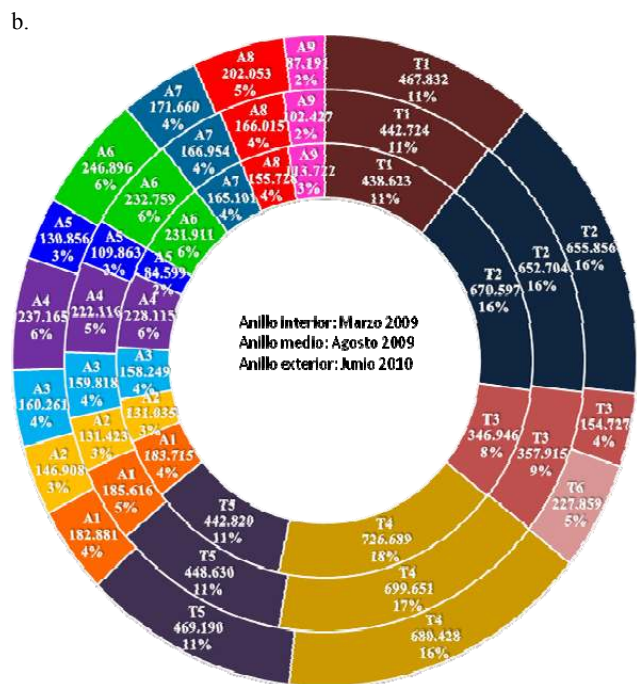
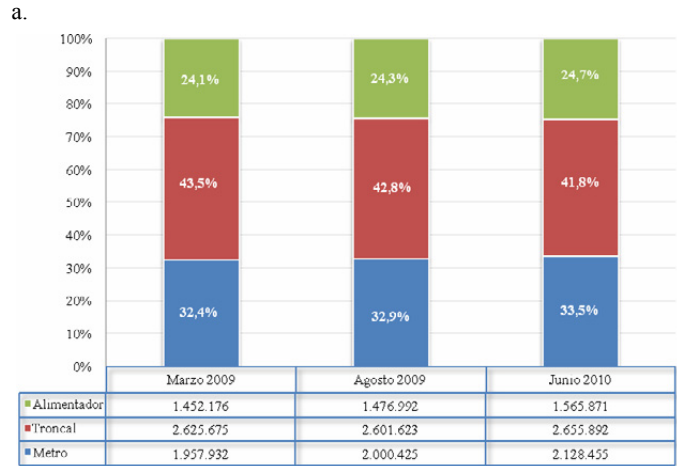


FIGURA 5: Transacciones Día Laboral por Operador, Corte Temporal y Unidad de Negocio

La Figura 6 muestra el histograma de transacciones semanales en buses en los tres cortes temporales para los servicios que presentaron más de 80.000 transacciones semanales. Como era de esperarse, los servicios con mayor número de transacciones son servicios troncales; los alimentadores aparecen hacia la mitad del histograma. Se observa una tendencia creciente del número de validaciones por servicio de un corte temporal a otro, a excepción de algunos servicios en los cuales se produce una baja de marzo 2009 a junio 2010. Dicha tendencia puede explicarse por la existencia de un margen -cada vez menor- de validaciones no asignadas a servicios (por producirse en buses para los cuales no fue posible determinar el servicio que estaban prestando); y por las diferentes cantidades de servicios de los cortes temporales.

⁷ Dado que el viaje no tiene costo para ellos, los escolares básicos no siempre validan al ingresar al sistema.

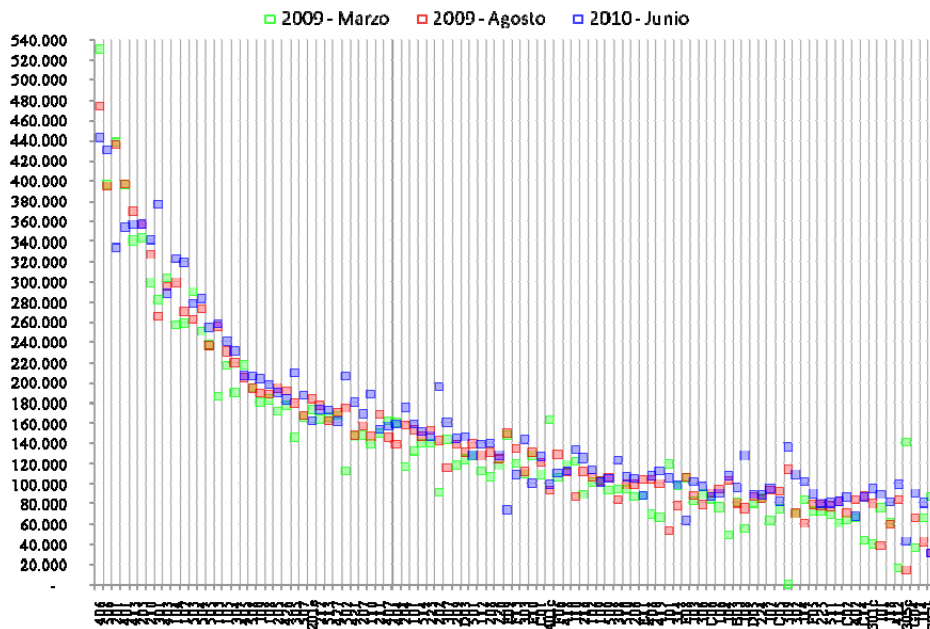


FIGURA 6: Histograma Transacciones en Buses por Servicio

4. ANÁLISIS DE VELOCIDADES DE BUSES

En términos operacionales, la velocidad comercial de los buses es una variable clave en el desempeño de los sistemas de transporte público de superficie, que afecta tanto el nivel de servicio otorgado a los usuarios como los costos de operación del sistema. Monitorear el estado de esta variable fundamental, que se ve afectada por factores diversos como la congestión en tráfico mixto, la programación de los semáforos o el grado de saturación de los paraderos, puede ser complejo en el caso de redes densas con gran cantidad de buses operando.

Dado que se dispone de gran cantidad de información, el desafío consiste en desarrollar métodos de procesamiento que permitan su correcta interpretación y manejo. Los dispositivos GPS registran la posición del bus cada 30 segundos cuando está en movimiento, con un margen de error de 15 metros. Aproximadamente el 5% de las emisiones corresponden a registros de control que son tomados cada 5 minutos, si el bus no está en movimiento.

En la Tabla 3 se muestra el total de buses diferentes que registraron emisiones GPS en cada semana, el promedio de emisiones diarias para días laborales, sábado y domingo, y el número de servicios-sentido. Se puede ver que el sistema ha ido aumentando la cantidad de servicios-sentido entre cortes temporales, con un incremento de 8% entre septiembre 2008 y junio 2010. El número de emisiones en día laboral se comporta de forma similar al número de buses -lo cual es razonable dado que el sistema opera a capacidad en esos días- con la excepción de septiembre 2008, que presenta una proporción mayor de emisiones debido a un mal uso del botón de pánico (activación frecuente por error que genera emisiones GPS cada 10 segundos) en los comienzos del sistema.

La reducción de la cantidad de buses, en contraste con el incremento en la cantidad de servicios, tiene su explicación en una fuerte renovación de buses a principios del 2010, que redujo su número, manteniendo la oferta de plazas.

TABLA 3: Buses, Emisiones GPS y Servicios en Cada Semana Analizada

	Sep 2008	Mar 2009	Ago 2009	Jun 2010
Número de buses diferentes observados en la semana	6.050	6.306	6.324	6.036
Emisiones GPS promedio día laboral [millones de pulsos]	12,4	12,0	12,1	11,8
Emisiones GPS día sábado [millones de pulsos]	9,9	9,1	9,2	9,3
Emisiones GPS día domingo [millones de pulsos]	8,4	7,9	7,9	7,9
Servicios-sentido (incluye variantes, cortos y expresos)	650	654	666	702

En la Figura 7 se muestra la trayectoria de todos los buses que realizaron una ruta en particular (506 de poniente a oriente) en un día laboral de abril 2009. En esta figura, el tiempo está en el eje X y la distancia recorrida en el eje Y, por lo que la pendiente de cada trayectoria representa la velocidad del bus. Por su parte, la separación horizontal muestra el intervalo de tiempo entre dos buses consecutivos. Se aprecian diferentes concentraciones de buses a lo largo del tiempo; por ejemplo, menos trayectorias en el período nocturno, que se ve en el lado izquierdo del gráfico. Asimismo, los intervalos son más regulares al comienzo de las trayectorias -abajo en el gráfico- que al final de éstas -arriba- (efecto pelotón o *bunching*).

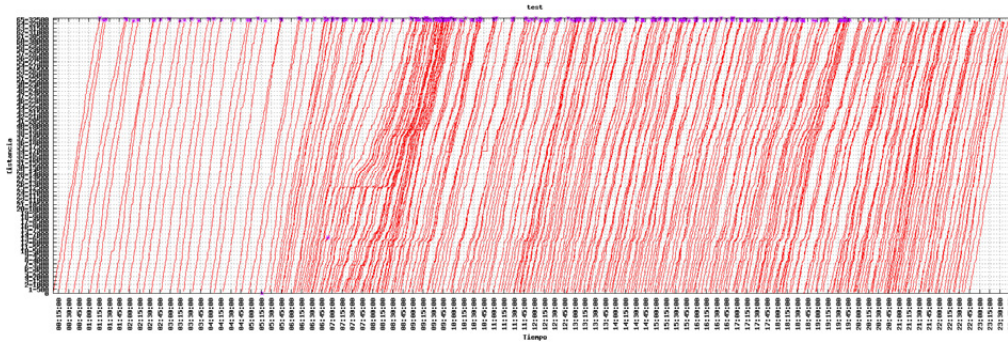
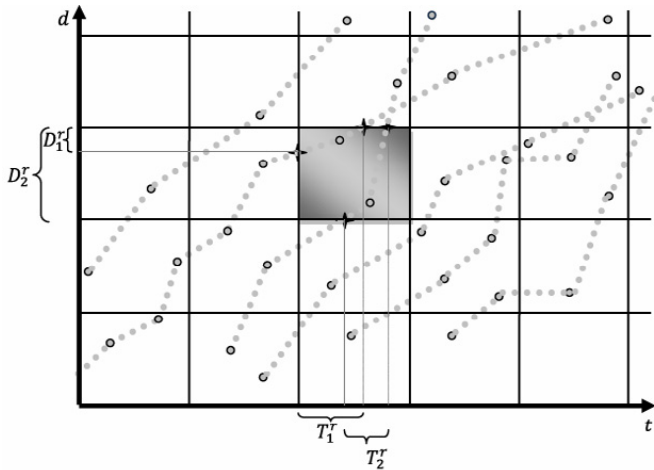


FIGURA 7: Ejemplo de Diagrama Espacio-Tiempo de una Línea de Buses

Para estimar la velocidad comercial de los buses que pertenecen a cada servicio, discriminando tanto espacial como temporalmente, se desarrolló una metodología (Cortés et al, 2010) que computa las trayectorias espacio-tiempo de los buses que están prestando el servicio en el tramo-periodo correspondiente. Para esto se utiliza la definición de una grilla como la mostrada en la Figura 8.



Fuente: Cortés et al (2010)

FIGURA 8: Grilla sobre las Trayectorias Espacio-Tiempo

La velocidad comercial promedio \bar{S} se calcula a partir de los tiempos y distancias recorridas por todos los buses observados en cada elemento de la grilla espacio-tiempo. Es decir, si D_i^r y T_i^r son la distancia e intervalo que el bus i permanece dentro de la celda r :

$$\bar{S} = \frac{\sum_i D_i^r}{\sum_i T_i^r} \quad (1)$$

Para expresar gráficamente las velocidades comerciales resultantes, se definió la escala de colores que se muestra en la Tabla 4.

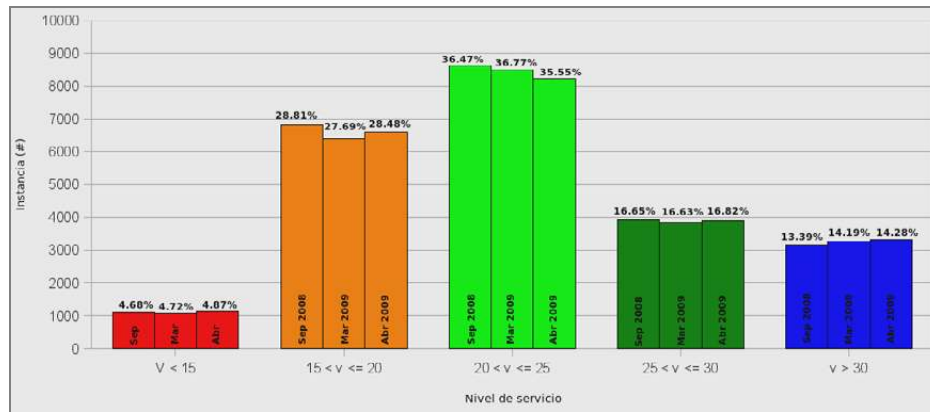
TABLA 4: Clasificación de Velocidades Comerciales y Nivel de Servicio

Nivel de servicio	Rango velocidad comercial (km/h)	Color
Excelente	> 30	Azul
Muy bueno	>25 a ≤30	Verde oscuro
Bueno	>21 a ≤25	Verde claro
Aceptable	>19 a ≤21	Amarillo
Malo	>15 a ≤19	Naranja
Muy malo	≤ 15	Rojo
	Sin información	Gris

En la Figura 9a se muestra los resultados obtenidos de la forma más agregada posible bajo una óptica espacial, comparando la velocidad global obtenida para tres semanas diferentes. Se grafica número de instancias (servicio-sentido, media hora) versus velocidad comercial promedio. Los cálculos de velocidades comerciales fueron realizados en cada caso para toda la semana laboral (de 6:00 a 23:59), considerando la distancia recorrida por todos los buses de cada servicio-sentido en cada periodo de tiempo (definido en 30 minutos).

La agregación espacial es evidente, pues en la Figura 9ase considera una velocidad representativa de todo el recorrido en cada período, donde la mayoría de los servicios circulan a un nivel aceptable, pero con casi 5% de las instancias mostrando condiciones muy malas (velocidades menores a 15 km/h). También se observa que a través del tiempo se ha mantenido ese porcentaje de instancias muy malas (de 4,68 a 4,87%), pero también han aumentado las instancias excelentes en casi un punto porcentual (13,39 a 14,28%).

El nivel de agregación de la Figura 9a entrega una visión general, pero sin duda es necesario observar los fenómenos de forma más desagregada para realizar un diagnóstico adecuado. Un primer análisis corresponde a una desagregación espacial, manteniendo la periodización de 30 minutos, pero separando por tramo de recorrido. Analizando ahora sólo una semana de información (abril 2009), se hace un cálculo de velocidades más desagregadas que considera las instancias asociadas a la grilla que se muestra en la Figura 8, con celdas de 500 metros e intervalos de 30 minutos. En este caso se computa la distancia recorrida por todos los buses en cada celda de la grilla y el tiempo que requieren para esto, lo que entrega una estimación de la velocidad comercial experimentada en cada celda de acuerdo a la expresión (1).



* En esta figura se eliminó el rango amarillo, moviéndose el límite de los rangos naranja y verde claro a 20 km/h.

FIGURA 9a: Evolución de velocidades desagregadas por servicio-sentido y período*

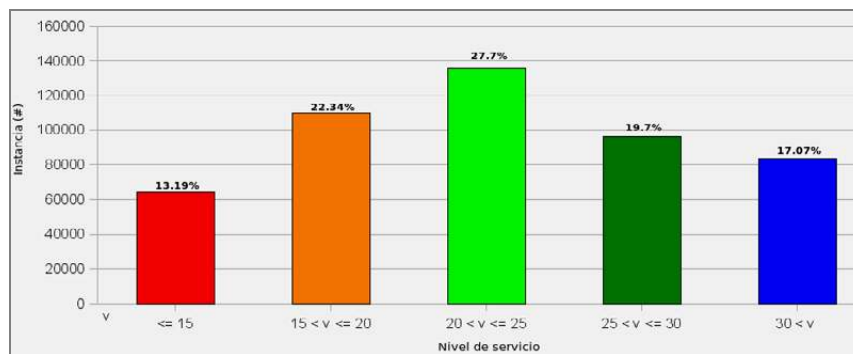


FIGURA 9b: Velocidades Desagregadas por Servicio-Sentido (tramificado cada 500 metros) y Período (30 minutos)

Comparando ambas figuras se aprecia una gran diferencia en el número de instancias con velocidades muy malas (de 4,87% a 13,19%), y una diferencia de casi diez puntos porcentuales en las instancias en condiciones aceptables (de 35,55 a 27,7%). Esto muestra los sesgos que genera considerar información agregada, especialmente en los extremos de las distribuciones, y permite destacar la importancia que debe darse a los niveles de agregación, tanto temporal como espacial, cuando se quiera tomar decisiones y definir políticas a partir de estos procesamientos.

Cabe hacer notar que las velocidades mostradas en los gráficos no están ponderadas por el número de buses afectados ni por las cargas de esos buses, por lo que no representan las velocidades percibidas por los usuarios. En efecto, es natural que las mayores concentraciones de usuarios se observen en servicios-sentidos-tramos que están sujetos a mayores niveles de congestión y, por lo tanto, presentarán velocidades bajo el promedio.

Aumentando el nivel de detalle, en la Figura 10 se muestra a modo de ejemplo un perfil de velocidades para un servicio-sentido en particular (406 sentido "ida") de la primera semana de abril de 2009 desagregado a lo largo de la ruta en tramos de 500 metros y por tiempo cada 30 minutos, según la ecuación (1). Utilizando el código de colores descrito en la Tabla 4, destacan los tramos 32 a 39 con velocidades malas y muy malas durante todo el día, mientras en otros tramos, por ejemplo de 14 a 16, se observa esto sólo durante las puntas mañana y tarde, con velocidades aceptables o buenas el resto del día.

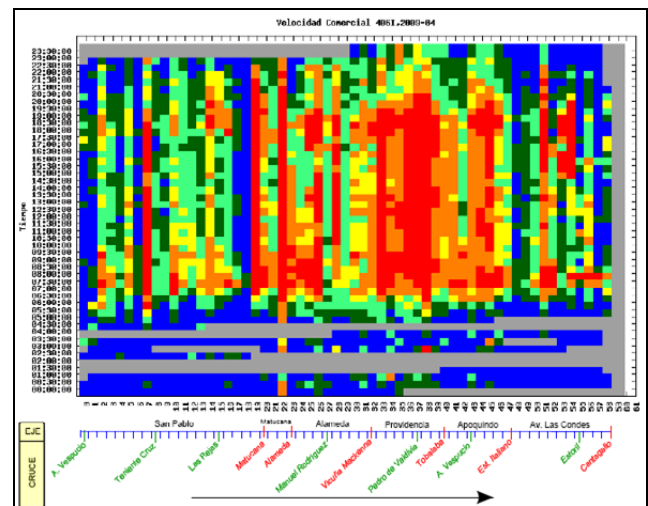


FIGURA 10: Ejemplo de Velocidad Comercial de un Servicio

Para un período en particular, la información se puede desplegar sobre un mapa de la ciudad para ver la ubicación de los sectores con problemas y orientar las medidas de mejoramiento que correspondan. En la Figura 11 se ilustra el recorrido del mismo servicio anterior con el código de colores aplicado a los segmentos de ruta, para el período 7:30 a 8:00 de la mañana. También es posible visualizar sobre el mapa todas las emisiones de los buses en algún tramo, para verificar si se encuentran concentradas por ejemplo cerca de una intersección o en un paradero, y de esta forma guiar el diagnóstico.



FIGURA 11: Ruta de un Servicio y Velocidad Comercial en Tramos Abril 2009

En la Figura 12 se muestra un histograma de velocidades agregado en intervalos de 1 km/h, diferenciando entre servicios troncales normales, troncales expresos y alimentadores. Los datos de junio 2010 arrojan que la moda de la velocidad se encuentra en 21 km/h. Se aprecia también una concentración de velocidades en torno a los 60 km/h, explicada mayoritariamente por servicios expresos.

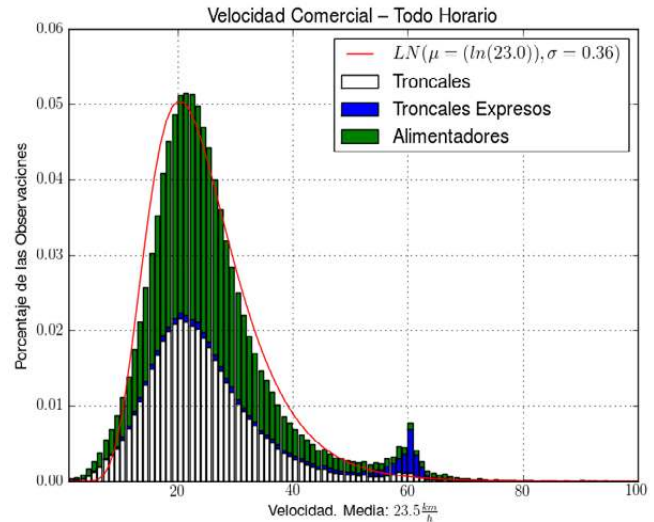


FIGURA 12: Histograma de Velocidades

A partir de los datos desagregados de velocidad de cada servicio, se creó una agregación espacial que permite ver la velocidad de cada servicio completo cada media hora. La Figura 13 muestra estas velocidades para todos los servicios de Transantiago en diferentes semanas. Los servicios están en el eje X y el tiempo en el eje Y. En la parte superior de cada gráfico están las horas correspondientes a la noche, cuando muchos servicios no operan (color gris) y los que sí lo hacen presentan velocidades muy altas (en azul). Las horas punta se pueden identificar en los horarios donde predominan los colores rojo y naranja.

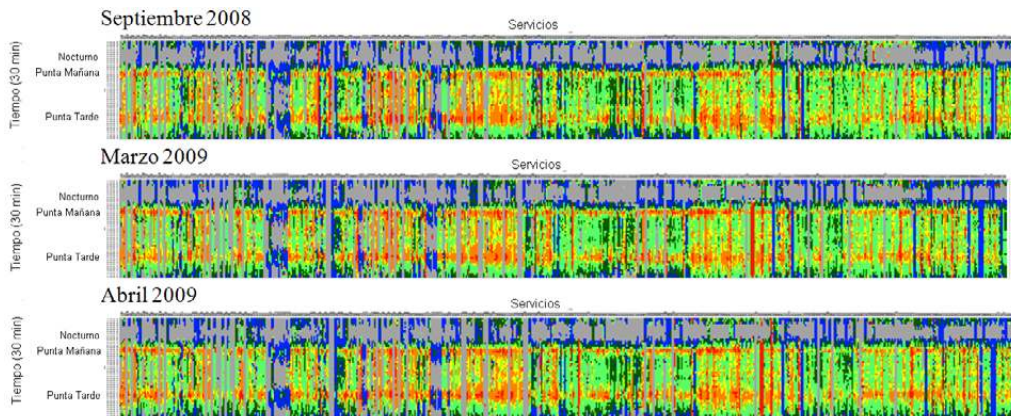


FIGURA 13: Velocidades Comerciales de Todos los Servicios cada Media Hora

Dada la gran cantidad de información desagregada de velocidades generadas, se creó una herramienta que permite identificar los tramos con peores velocidades de toda la red, para diferentes períodos del día. En efecto, se generó una lista ordenada de los tramos-servicio de menor a mayor velocidad comercial, la cual se visualiza en una herramienta web.

Al seleccionar un tramo-servicio en particular, se muestra éste en el mapa adyacente con un zoom del entorno del tramo, indicando además el sentido de circulación del tramo-servicio. La Figura 14 muestra un ejemplo de esta herramienta con los tramos-servicio de menor velocidad en punta mañana para la semana de junio 2010, considerando tramos de 500 m.

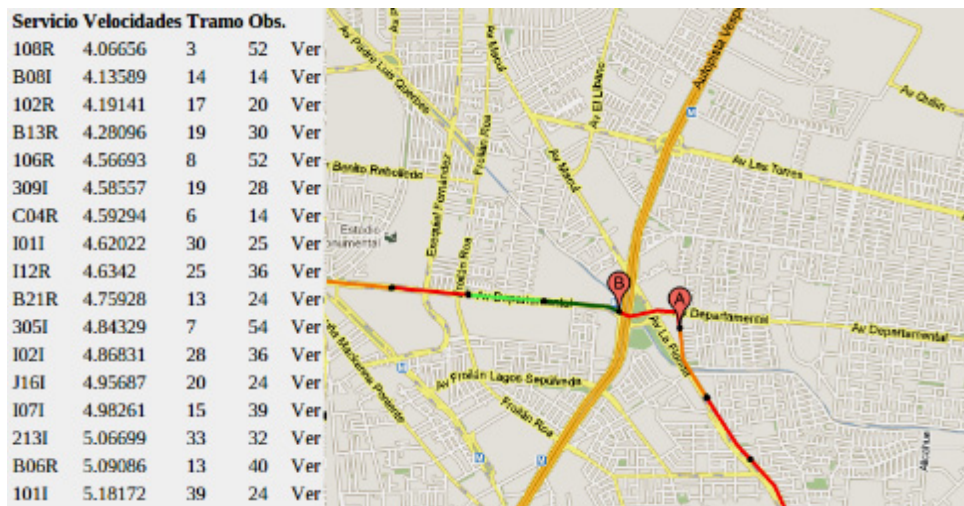


FIGURA 14: Herramienta de Visualización de Tramos con Peores Velocidades

Para observar la evolución de las velocidades a lo largo del día, se generaron gráficos de distribución de velocidades. Utilizando la misma escala de colores definida anteriormente, se grafica, para cada media hora, el porcentaje de tramos-servicio-sentido que se encuentra en cada rango de velocidades. La Figura 15a muestra la distribución de velocidades para todos los servicios en junio 2010, mientras que la Figura 15b diferencia entre los servicios troncales (izquierda) y alimentadores (derecha). En todos los casos, se aprecia el notorio empeoramiento de las velocidades en los periodos punta mañana y punta tarde, así como una leve reducción de velocidad en punta mediodía. Comparando troncales con alimentadores, se observa una mayor proporción de velocidades malas en los primeros, especialmente en punta tarde. Estos gráficos pueden generarse también de forma desagregada, por ejemplo a nivel de cada unidad de negocio.

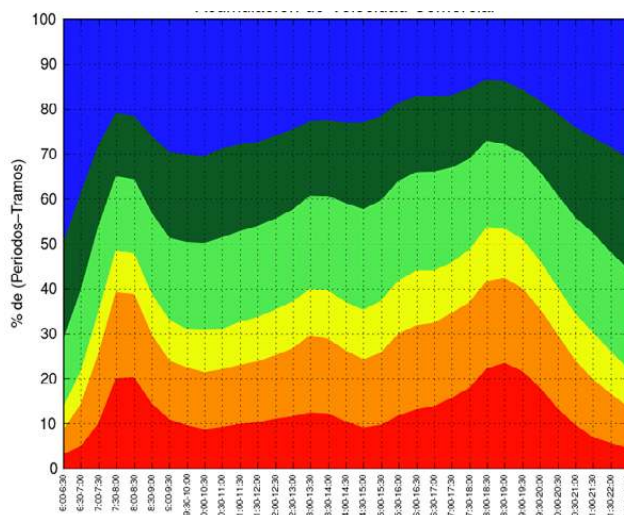
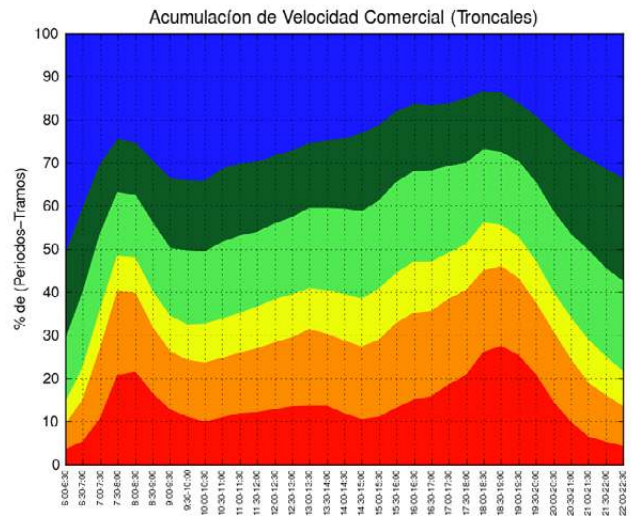


FIGURA 15a: Distribución de Velocidades por Media Hora para Todos los Servicios

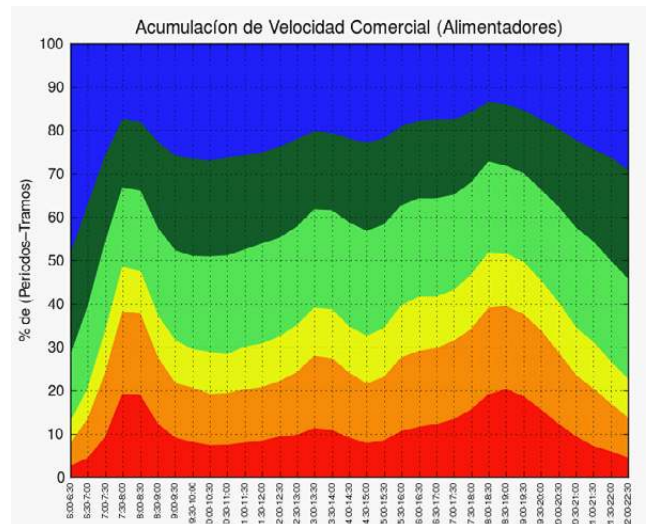


FIGURA 15b: Distribución de Velocidades por Media Hora para Servicios Troncales y Alimentadores

5. ANÁLISIS DE LA ESTRUCTURA ESPACIAL DE LA DEMANDA

El cruce de la información de transacciones con la de posicionamiento permite asociar posición geográfica a las subidas registradas en los distintos servicios, vinculando la hora de la transacción y el bus en que ésta se realizó, con la base de datos de posicionamiento. Las estaciones de Metro y las zonas paga tienen posición fija, por lo que en esos casos la posición es directamente asignada a las transacciones correspondientes. No obstante, en estos casos la hora de validación no corresponde a la hora de subida.

Las transacciones realizadas en los buses son localizadas geográficamente a través de la placa patente del bus al que corresponde el validador, pues en la base de datos de posicionamientos GPS los buses son registrados cada 30 segundos, y a partir de dichos registros se hace una interpolación lineal entre dos puntos de GPS para estimar la posición. En caso de que la interpolación se realice entre posiciones registradas con un intervalo de tiempo mayor, o con distancias excesivamente grandes, debido a falta o error en la información, se aplica filtros para que esas transacciones no sean consideradas en los procesos posteriores.

La tasa de éxito de la identificación de la posición geográfica de la validación alcanza el 98,2% en los datos de marzo 2009 y 99,9% en los de junio 2010. Además de estas fuentes de error, se debe considerar que es posible que transcurran algunos segundos entre el instante en que un pasajero aborda el bus y el instante en que valida, especialmente en el caso de buses con alto flujo de pasajeros. En consecuencia, es probable que la posición estimada de las transacciones esté sesgada hacia aguas abajo del paradero. Sin embargo, al agregar a nivel de paradero, este error es imperceptible.

Este procesamiento básico ya genera información interesante, permitiendo visualizar la distribución espacio-tiempo de la demanda. La Figura 16 muestra la cantidad de transacciones en bus y zonas paga agregadas a nivel de paradero, para dos periodos: de 6 a 6:15 AM y de 7:15 a 7:30 AM en un día laboral. Se puede ver como los sectores más alejados de la ciudad presentan gran número de transacciones en el horario más temprano, mientras que en el segundo periodo mostrado hay gran cantidad de transacciones en toda la ciudad. Este fenómeno se marca con mucha claridad en el eje Pajaritos-Alameda-Providencia-Apoquindo.

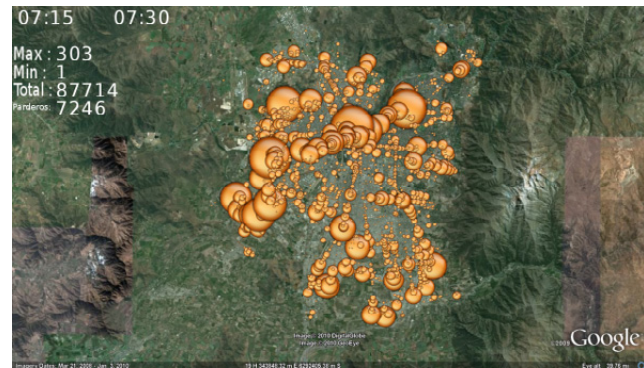
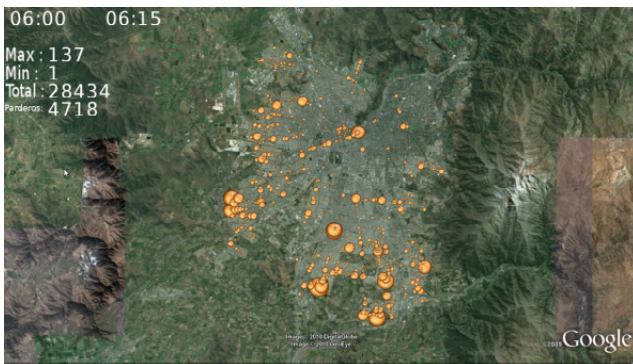


FIGURA 16: Transacciones en Paraderos de Transantiago

Para realizar otros análisis de la estructura de la demanda es necesario aplicar procesamiento adicional a estos datos. Munizaga y Palma (2012) proponen una metodología para estimar paradero de bajada, basada en la observación de la siguiente transacción de la misma tarjeta. Dependiendo del tipo de transacción (en bus, metro o zona paga), se realiza una búsqueda sobre las posibles estaciones o paraderos de bajada ubicadas en el radio de un kilómetro de la posición de la siguiente transacción, y se elige aquella que minimiza el tiempo generalizado (tiempo de viaje más tiempo de caminata penalizado).

En el caso de las transacciones en zona paga, existe la dificultad adicional de identificar el servicio que el usuario utiliza para ese segmento de viaje, lo cual se enfrenta identificando las líneas comunes (Chriqui y Robillard, 1975) y asumiendo que el usuario toma el primer bus que pasa de los servicios que constituyen el conjunto de líneas comunes. De esta manera, se logra estimar el paradero de bajada en un porcentaje superior al 80% en ambos cortes respecto al total de subidas de la semana, permitiendo analizar subidas y bajadas de pasajeros en los más de diez mil paraderos de la ciudad. En la Tabla 5 se muestra el porcentaje de éxito del proceso, así como el porcentaje asociado a las distintas fuentes de error en los casos en que no fue posible estimar paradero de bajada. En la Tabla 6 se reporta la tasa de éxito por tipo de transacción.

TABLA 5: Porcentaje de Éxito en Estimación de Paradero de Bajada a Nivel de Etapas

	Marzo 2009	Junio 2010
Número de transacciones [millones]	36	38
% éxito en la estimación de paradero de bajada	80,77	83,01
Razones de no estimación %		
Distancia no caminable	7,3	7,6
Sólo una transacción por día	5,2	5,4
Estima subida y bajada en el mismo lugar	2,43	2,39
Errores de datos	4,3	1,6

TABLA 6: Porcentaje de Éxito en Estimación de Paradero de Bajada por Modo a Nivel de Etapas

	Marzo 2009		Junio 2010	
	Transacciones	% éxito	Transacciones	% éxito
Bus	22.546.517	81,7	23.614.012	84,5
Metro	11.383.047	89,0	12.470.020	89,4
Zona paga	2.127.170	67,1	2.034.917	72,2

La estimación de paradero de bajada permite obtener subidas y bajadas de las etapas de viaje registradas, y perfiles de carga de los buses (referidos exclusivamente a quienes no evaden o eluden su pasaje) y, a través de herramientas de visualización desarrolladas permiten desplegar sobre un mapa de Google Earth las subidas y bajadas por horario y por tipo de usuario. A modo de ejemplo, en la Figura 17 se muestra las subidas de pasajeros estudiantes y adultos en los distintos paraderos del sistema. Esta herramienta también está implementada para las transacciones en Metro y para las bajadas.

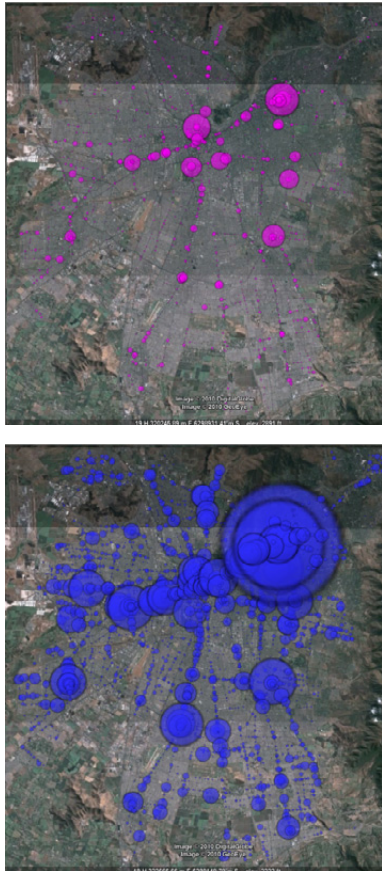


FIGURA 17: Visualización Subidas de Pasajeros Estudiante y Adulto

Adicionalmente, es posible construir una pseudo matriz origen/destino considerando el origen y destino de cada etapa. Tanto para perfiles de carga como para la pseudo matriz origen/destino es necesario aplicar factores de expansión que corrijan las estimaciones por el % de etapas a los que no se logró estimar paradero de bajada. Siguiendo la metodología propuesta por Munizaga et al (2011) aquí se considera dos tipos de factores de expansión: uno que corrige por aquellas transacciones que no tienen ni origen ni destino identificado (0,08%), aplicando un factor uniforme para amplificar todos los pares origen/destino, y un factor asociado al origen, para aquellas transacciones que teniendo posición en el origen, no se le pudo estimar paradero de bajada. Este último factor se construye a nivel de zona DIRTP y reparte las etapas sin destino asignada de manera proporcional a los destinos de las otras etapas que comparten el origen y que sí tienen destino identificado.

Otro proceso importante para obtener información valiosa a partir de estos datos es poder distinguir entre el destino de un viaje y los lugares en los cuales se realiza un transbordo. Para enfrentar este problema, en primera instancia se aplicó un criterio de tiempo, considerándose que si el tiempo transcurrido entre la bajada estimada y la siguiente validación es de más de 30 minutos, se trata del destino de una actividad y finaliza un viaje. Si es menor a 30 minutos, se trata de un transbordo. Devillaine et al (2012) proponen una metodología que permitirá mejorar ese aspecto, considerando elementos adicionales como la secuencia de servicios utilizados, la frecuencia de estos, la relación entre distancia recorrida y distancia euclidiana, y el uso de suelo en el punto de destino o transbordo.

En las Tablas 7a y 7b se presenta una agregación de las matrices origen/destino de viajes registrados a través del medio de pago electrónico, que son obtenidas a nivel de paradero. Hay que recordar que los factores de expansión no consideran la evasión y que para hacerlo es necesario realizar mediciones que permitan distinguir al evasor duro, que evade todas las etapas del viaje, del evasor casual, que evade sólo algunas, pues en un caso corresponderá aplicar un factor de expansión y en el otro un factor de corrección a la distorsión provocada. Con este sesgo, se estimaron 3.7 millones de viajes promedio (día laboral) en marzo 2009 y 4 millones en junio 2010, aproximadamente.

TABLA 7a: Matriz Origen-Destino agregada Marzo 2009

	Norte	Occidente	Oriente	Centro	Sur	Sur Oriente	Oi
Norte	157.950	36.389	51.489	78.906	22.988	18.087	365.810
Occidente	34.164	294.670	116.217	162.561	37.041	30.525	675.177
Oriente	49.382	112.436	317.606	173.157	70.812	150.056	873.450
Centro	74.593	167.516	160.132	171.932	103.127	96.399	773.700
Sur	22.222	34.877	73.977	104.116	189.216	54.216	478.624
Sur Oriente	18.379	30.450	158.839	97.234	55.614	250.057	610.572
Dj	356.690	676.338	878.261	787.906	478.798	599.339	3.777.333

TABLA 7b: Matriz Origen-Destino agregada Junio 2010

	Norte	Occidente	Oriente	Centro	Sur	Sur Oriente	Oi
Norte	176.291	38.976	57.360	86.022	24.671	20.565	403.885
Occidente	37.568	313.822	131.083	176.402	38.491	33.571	730.937
Oriente	52.693	120.714	349.172	187.507	71.893	160.822	942.799
Centro	84.344	178.518	168.284	176.849	106.518	100.024	814.538
Sur	24.658	37.350	83.124	113.291	194.636	56.133	509.192
Sur Oriente	20.115	33.326	170.041	104.565	56.563	255.868	640.478
Dj	395.668	722.707	959.064	844.636	492.772	626.982	4.041.830

No es posible hacer una comparación rigurosa entre estas matrices y, por ejemplo, la resultante de la EOD2006 (Sectra, 2006) obtenida a través de encuestas. En efecto, si bien coinciden en la importancia relativa de la diagonal, que representa los viajes intra-macrozonas, las primeras están severamente sesgadas por la evasión, que como se ha dicho no afecta a todas las zonas y modos de transporte de igual forma. Un claro ejemplo de esto último es que, al ser la zona oriente una de las que exhibe los menores niveles de evasión (Fundación Chile, 2010), la mayor cantidad de viajes en las Tablas 7a y 7b pertenecen al par Oriente-Oriente.

Finalmente, en la Figura 18 se presentan los perfiles de carga de los servicios sentido como el promedio de carga de todos los buses que salieron del punto de inicio en una hora (cohorte). Esta etapa de la investigación está aún en desarrollo, por lo que estos resultados son preliminares y las posibilidades de agregación aún están en evolución.

En cuanto a la demanda, se considera todos los pasajeros que suben a esos buses, y su estimación de paradero de bajada, debidamente ponderados por los factores de expansión correspondientes. La capacidad indicada en las figuras es el promedio de las capacidades de los buses de la cohorte correspondiente.

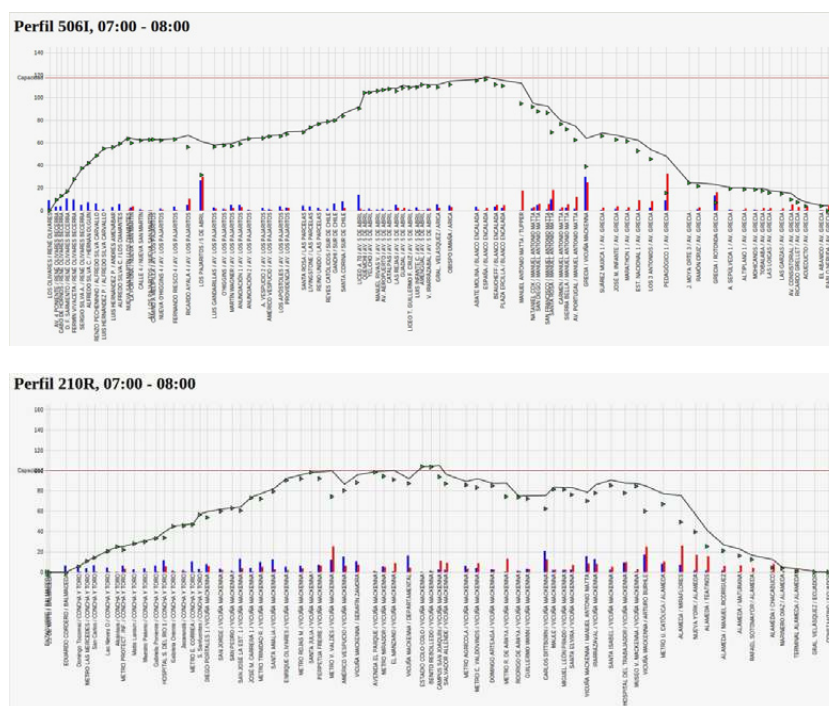


FIGURA 18: Ejemplos de Perfiles de Carga de por Servicio (pasajeros que validan)

6. CONCLUSIONES

Como parte de esta línea de investigación se ha desarrollado métodos para procesar los datos y obtener información valiosa a partir de las bases GPS y de tarjeta bip! provenientes de Transantiago. La información pura de transacciones permite realizar análisis estadísticos de los perfiles horarios de la demanda y del uso del medio de pago y modos de transporte y servicios. Al cruzar esa información con la de posicionamiento de buses, es posible también incorporar la dimensión espacial al análisis. Al aplicar el procedimiento de estimación de paradero de bajada, se logra además obtener información detallada de los viajes y transbordos en términos de tiempos de viaje, tiempos de transbordo y variabilidad de estos, así como matrices origen destino de viajes registrados en transporte público, desagregando espacial y temporalmente al nivel deseado, y perfiles de carga de buses y servicios.

Lo anterior implica que gran parte de la información que tradicionalmente analistas y planificadores requieren para proponer mejoras o reformas sobre los sistemas de transporte, es posible obtenerla a bajo costo y con alta periodicidad. Esto es de una gran importancia y potencial, pues de validarse y completarse las metodologías presentadas se podría dejar de depender de encuestas o mediciones que, generalmente por su costo, magnitud y esfuerzo se realizan con muy baja periodicidad, y entregan resultados estáticos que generalmente son usados como verdades absolutas sin realmente serlas.

En efecto, la información presentada en este documento es solo una síntesis de los resultados obtenidos hasta la fecha en el contexto de la colaboración entre la Universidad de Chile y el Ministerio de Transportes y Telecomunicaciones, pues adicionalmente a ella se ha obtenido y puede obtenerse resultados con el nivel de desagregación temporal y espacial que se requiera.

Por ejemplo, es posible obtener matrices de viaje (o entre paraderos) o perfiles de carga por tipo de día, periodo, horario, etc., lo que permitirá gestionar la demanda, y planificar o diseñar la oferta de una forma más precisa y eficiente.

Por cierto, la posibilidad de contar con matrices dinámicas derriba uno de los impedimentos naturales a la hora de comenzar a aplicar modelos dinámicos de asignación de pasajeros a redes de transporte público.

Asimismo, la observación desagregada de los usos de las tarjetas bip!, permite conocer en detalle el comportamiento de los viajes en lo referente a los servicios y combinaciones que realiza, y cómo la demanda se comporta o ha reaccionado a los diversos estímulos (tarifas), cambios operacionales (trazados, frecuencias, extensión de Metro) y reglas de uso del sistema (ventana horaria y transbordos máximos).

En particular, el detalle del comportamiento de los usuarios, respecto a los servicios y combinaciones usadas en el corte temporal Junio 2010 fue utilizado por el Ministerio de Transportes y Telecomunicaciones, y puesto a disposición de los Concesionarios, como parte del proceso de renegociación de los contratos. Igual cosa ocurrió con la estimación de subidas y bajadas del corte temporal Marzo 2009, que fue usada para construir una versión desagregada de la matriz de viajes en punta mañana, que además incluyó una expansión por evasión.

Si bien los resultados mostrados en este documento son importantes y concluyentes en muchos ámbitos, por ejemplo en materia de uso de tarjetas bip; al incremento en la validaciones totales del sistema; a la reducción relativa de las combinaciones Bus-Bus y Metro-Bus; y a los sesgos que se producen al considerar información agregada en materia de velocidades, la investigación y colaboración está aún lejos de terminar.

En efecto, las siguientes etapas en esta investigación corresponden a los mejoramientos metodológicos señalados en los capítulos precedentes y a la validación de la información generada, la cual en una primera instancia se puede hacer mediante análisis de consistencia de los datos, pero la validación definitiva se tendrá cuando se pueda comparar los patrones de viaje generados con esta metodología con patrones de viaje generados mediante encuestas. Además, se está trabajando en la generación de indicadores globales del sistema, a nivel de tiempos de espera, tiempos de viaje y velocidades globales, que permitan caracterizarlo y observar su evolución en el tiempo.

AGRADECIMIENTOS

Agradecemos el apoyo de Transantiago, de la Subsecretaría de Transportes del Ministerio de Transportes y Telecomunicaciones, el Instituto Sistemas Complejos (ICM P-05-004-F, CONICYT FBO16); FONDEF (D10I-1002); FONDECYT (1120288). Agradecemos también la contribución de Eduardo Graells, quien trabajó breve pero entusiastamente en la visualización de velocidades.

REFERENCIAS

- Bagchi, M., White, P.R.(2005) The potential of public transport smartcard data. **Transport Policy** 12, 464-474.
- Chriqui, C., Robillard, P. (1975) Common bus line. **Transportation Science** 9, 115-121.
- Cortés, C., J. Gibson, A. Gschwender, M. Munizaga, M. Zúñiga (2011). Bus commercial speed diagnosis based on GPS monitored data. **Transportation Research C** 19, 695-707.
- Devillaine, F., Munizaga, M.A., Trepanier, M. (2012) Detection of public transport user activities through the analysis of smartcard data. **Transportation Research Record**. Aceptado para publicación.
- Fundación Chile (2010) Verificación funcional del sistema de pago de Transantiago. Informe técnico de orden de trabajo 3, Transantiago.
- Munizaga, M.A., Palma, C. (2012). Estimation of a disaggregate multimodal public transport OD matrix from passive smartcard data from Santiago. **Transportation Research C** 24, 9-18.
- Park, J.Y., Kim, D.J., 2008. The potential of using the smartcard data to define the use of public transit in Seoul. **Transportation Research Record: Journal of the Transportation Research Board**, No. 2063, Transportation Research Board of the National Academies, Washington, DC, pp. 3-9.
- Sectra, 2006. Encuesta Origen Destino, Informe Ejecutivo.
- Trépanier, M., Chapleau, R., Tranchant, N., 2007. Individual trip destination estimation in transit smartcard automated fare collection system. **Journal of Intelligent Transportation Systems: Technology, Planning, and Operations** 11, 1-15.