

UNIVERSIDAD DE CHILE

Facultad de Ciencias Físicas y Matemáticas

Departamento de ciencias de la computación

CC6908 Introducción al trabajo de título



*Visualización de estructuras
espaciales de desplazamiento a
partir de datos de transporte público*

8 de septiembre de 2014

Profesor Guía:
Marcela munizaga

Profesor Co-guía:
Benjamín Bustos

Memorista:
Felipe A. Hernández G.
fhernand@dcc.uchile.cl
09-90977379

Semestre: 02/2014
SANTIAGO DE CHILE

Índice

1. Introducción	1
2. Motivación	3
3. Objetivos	4
3.1. Objetivo General	4
3.2. Objetivos específicos	4
4. Metodología	5
5. Bibliografía	7

1. Introducción

En los últimos años la utilización de tecnología en el transporte público ha ido en aumento debido a varios factores, mayor regulación, usuarios más exigentes, aumento en seguridad, etc. Lo anterior ha llevado al sistema público de transporte a implantar diversos dispositivos que permiten controlar los aspectos mas relevantes al momento de transportar una persona de un punto a otro.

Dentro de las tecnologías más usadas podemos nombrar AVL (Automatic Vehicle Location), que permite conocer la posición geográfica de un vehículo en todo momento con un margen de error bajo y los sistemas AFC (Automated Fare Collection) que automatizan el proceso de pago, en particular, nos interesa el basado en tarjetas de pago, que albergan un chip que permite mantener un saldo para que sea utilizado al abordar un bus, esto implica la existencia (paralelamente) de dispositivos asociados a los buses o paraderos¹ que permitan registrar el correcto descuento del valor asociado al pasaje, concepto que llamaremos validación.

Transantiago² es el sistema de transporte público de Santiago de Chile que implementa las tecnologías nombradas anteriormente, por lo que hoy en día se sabe que se realizan aproximadamente 6.000.000 de validaciones durante un día laboral³, lo que genera una cifra cercana a los 35.000.000 de transacciones a la semana (incluyendo sábado y domingo) con aproximadamente 3.000.000 de tarjetas de pago. Por otro lado, hay 80.000.000 de emisiones proveniente de la tecnología AVL del sistema. Al procesar estos datos en conjunto es posible identificar el paradero de origen, recorrido utilizado para desplazarse y el paradero de destino, este último requiere un procesamiento adicional basado en una metodología desarrollada por [2] que logra una identificación acertada en el 80 % de las validaciones.

La estructura espacial moderna de las ciudades ha sido formada, en gran medida, por avances en transporte y comunicaciones [3]. La forma en la cual se mueven los habitantes de una ciudad ha ido modificando la estructura de esta, motivados por la transferencia de recursos como materiales, dinero, personas e información. Considerando una persona como un transportador de recursos de un área urbana a otra es que se identifican las siguientes

¹Lugar físico donde un bus de transporte público se detiene para que personas ingresen y/o desciendan a el.

²fue implementado a partir del año 2007.

³Lunes, martes, miércoles, jueves o viernes.

estructuras espaciales urbanas [1]:

- Centros de flujo: Se refiere a las áreas que sirven para conectar otro par de áreas para transferencia de personas. Funcionan como puentes espaciales entre distintas áreas.
- Centros: Se refiere a áreas que concentran personas. Pueden diferir de los Centros de flujo, pero a menudo, son lo mismo.
- Bordes: Se refiere a límites socioeconómicos generados a partir de la agrupación de paraderos que divide la ciudad en pequeños barrios que llamamos comunidades.

Para detectar estos elementos en la ciudad de Santiago haremos uso de los datos obtenidos de su sistema de transporte público utilizando para ello el análisis de redes (grafos) y algoritmos de procesamiento de estructuras de grafos a gran escala como infomap, además de herramientas computacionales para diseñar una interfaz de visualización de los datos.

2. Motivación

Hoy en día el sistema de transporte público de Santiago cuenta con una gran cantidad de datos pasivos que van en aumento día a día producto de las validaciones que realizan los usuarios y los servicios que se ejecutan. Dado lo anterior, existe una gran base de datos que mantiene un potencial de información que puede mejorar la planificación y operación del sistema. Sin embargo, con las capacidades de procesamiento actuales no es posible obtener información que complemente los resultados de los indicadores usados actualmente.

Según lo anterior, el problema que se busca resolver en esta memoria es interesante de abordar debido a que ayudará a entender la estructura espacial de los viajes realizados por la población y permitirá diseñar servicios pensando en las necesidades observadas de los usuarios. Esto ayudará a:

- Mejorar las posibilidades de realizar actividades en el entorno de la zona de residencia.
- Disminución de la demanda en los Centros de flujo.
- Disminución en el tiempo requerido para trasladarse hasta el punto de interés para una comunidad determinada.
- Mejorar las condiciones de viaje de grupos vulnerables.

También es interesante académicamente debido a la masividad de los datos, ya que se deberá implementar una estrategia de procesamiento que permita manejar millones de registros y además realizar un análisis espacial para la detección de centros de flujo, centros y bordes basada principalmente en análisis de grafos masivos.

3. Objetivos

3.1. Objetivo General

“Diseñar una herramienta que permita identificar estructuras espaciales de movimiento en la ciudad de Santiago utilizando datos pasivos y masivos de transporte público.”

3.2. Objetivos específicos

1. Construir modelo de red para la ciudad de Santiago.
2. Identificar patrones de viaje, centros y puntos de alto flujo de pasada.
3. Desarrollar una herramienta de visualización.

4. Metodología

El desarrollo de esta memoria considera la siguiente metodología de trabajo:

1. Investigación bibliográfica

Se está realizando una recopilación y redacción de las ideas y estrategias más relevantes que aporten y justifiquen la base teórica de esta memoria.

2. Estudio de la data.

Se realizará un análisis de los datos existentes para comprender concretamente las bases de datos y datos requeridos.

3. Definición de estrategia de pre-procesamiento de datos.

Se investigará sobre las estrategias de pre-procesamiento y elegirá la que mejor se adapte en base al estudio realizado en el item anterior. Dentro de esta etapa de normalización y selección de la data para realizar los análisis.

4. Construcción de la red de nodos.

En esta etapa se realizará la construcción de un grafo dirigido con nodos a partir de la data estudiada.

5. Análisis de la red

- a) Definición de propiedades básicas.

Aquí asociaremos un atributo de las estructuras urbanas a cada propiedad matemática de un grado a partir de las ideas obtenidas de la investigación bibliográfica.

- b) Definición de centralidades.

- 1) Centro de flujo.

Se define el concepto de Centro de flujo en un grafo (betweenness centrality) y se propone una fórmula para medirlo.

- 2) Centro

Aquí estudiaremos y definiremos la estrategia para detectar centros de la ciudad ocupando el algoritmo PageRank.

c) Estructura de comunidad.

Para la detección de estructuras de comunidad se utilizará el software infomap.

6. Análisis espacial

a) Interpolación espacial.

Lo relevante de esta etapa es relacionar una zona geográfica a un paradero de bus de manera de poder particionar la ciudad.

b) Cálculo estadístico.

En esta etapa se realizará la asociación de las comunidades detectadas a las áreas geográficas establecidas en la interpolación espacial.

7. Análisis de los resultados.

Se estudiarán los resultados obtenidos.

8. Definir visualizaciones y nivel de interactividad de las visualizaciones.

A partir del punto anterior se definirán las visualizaciones a realizar y las posibles interacciones que puedan haber en cada una de ellas.

9. Diseño de aplicación de visualización.

Se desarrollará una aplicación que permita ver cada una de las implementaciones definidas en el punto anterior.

5. Bibliografía

- [1] *Detecting the dynamics of urban structure through spatial network analysis.* **Autores:** Chen Zhong, Stefan Müller Arisona, Xianfeng Huang, Michael Batty & Gerhard Schmitt **URL:** <http://www.tandfonline.com/doi/abs/10.1080/13658816.2014.914521>
- [2] *Estimation of a disaggregate multimodal public transport Origin–Destination matrix from passive smartcard data from Santiago, Chile* **Autores:** Marcela A. Munizaga, Carolina Palma **URL:** <http://www.sciencedirect.com/science/article/pii/S0968090X12000095>
- [3] *Urban spatial structure* **Autores:** Alex Anas, Richard Arnott, Kenneth A. Small **URL:** <http://www.sjsu.edu/people/allen.tai/courses/urbp255/s2/Urban%20Spatial%20Structure.PDF>
- [4] *Interactive Data Visualization: Foundations, Techniques, and Applications* **Autores:** Matthew O. Ward, Georges Grinstein, Daniel Keim **URL:** <http://www.amazon.com/Interactive-Data-Visualization-Foundations-Applications/dp/1568814739>