Analysis of the urban travel structure using smartcard and GPS data from Santiago, Chile

Marcela Munizaga, Felipe Hernández, Universidad de Chile
Ricardo Hurtubia, Pontificia Universidad Católica de Chile
Introduction

• Cities are complex and they will become even more complex in the future
• If we want to improve life quality in cities, we need to understand the activities/travel structure
• OD surveys: detailed information, based on respondents' declarations, usually small samples
• Passive data: provided (and limited) by the technological devices available, limited human intervention, usually huge samples
Objective

- To analyze the urban spatial structure of Santiago de Chile using passive data from the public transport system (AFV, AVL)
Background

Methods developed for:

- **Alighting estimation**
  - Zhao et al 2007; Trepanier et al 2007; **Munizaga & Palma 2012**
- **Trip linking, transfer behavior**
  - Seaborn et al 2009; Charikov & Erath 2011; **Devillaine et al 2013**
- **Trip purpose estimation**
  - Charikov & Erath 2011; **Devillaine et al 2013**; Lee & Hickman 2014
- **Zone of residence estimation**
- **Analysis of the dynamics of urban structure in Singapore using passive data**
  - Zhong et al. 2014
Case of study: Santiago

Santiago. Capital City of Chile:
- Population: 6.6 million
- 2 million households
- Area: 1,400 km²
- 34 Municipalities
- 18M trips per day
  - 28% public transport
  - 35% walk
  - 26% car
### Description of Santiago in 6 large zones

<table>
<thead>
<tr>
<th>Region</th>
<th>Persons/HH</th>
<th>Monthly Income</th>
<th>Daily Trips/person</th>
<th>Public Transport Rate</th>
<th>Vehicles/HH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>North</strong></td>
<td>3.39</td>
<td>899</td>
<td>2.51</td>
<td>0.48</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>West</strong></td>
<td>3.36</td>
<td>775</td>
<td>2.65</td>
<td>0.50</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>South</strong></td>
<td>3.43</td>
<td>772</td>
<td>2.89</td>
<td>0.44</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>East</strong></td>
<td>2.95</td>
<td>2100</td>
<td>3.67</td>
<td>0.26</td>
<td>1.19</td>
</tr>
<tr>
<td><strong>Centre</strong></td>
<td>2.35</td>
<td>1076</td>
<td>2.83</td>
<td>0.58</td>
<td>0.40</td>
</tr>
<tr>
<td><strong>South-East</strong></td>
<td>3.37</td>
<td>967</td>
<td>2.44</td>
<td>0.43</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Source: 2012 ODS
Transantiago:

- Multimodal integrated public transport system (bus, metro) introduced in 2007
- Equipped with Smartcard and GPS
- Transactions sequence → travel structure

- 6,500 buses all equipped with GPS
- 70 km of segregated busways
- 10,000 bus stops
- 125 bus stations (off-bus fare collection)
- Metro: 5 lines, 100 km, 54 trains
- Only smartcard payment in buses (global penetration rate 97%)
- 4.6M trips per day
The Data

- Buses GPS: 1 record every 30s, 80–100 M records per week
- bip! transactions: 35-40 M records per week
- Other information:
  - Routes paths
  - Route assignments
  - Position of bus stops
  - Position of Metro stations
  - Position of bus stations
• Estimation of alighting stop

\[
\min_i T_g = t_i + \frac{d_i > x_{post} y_{post}}{s_{walk}} \times (\text{walk} / \text{travel}) \\
\text{s.t. } d_{post} \leq d
\]
### Processing: Stages and Trips

**Trip purpose estimation**
- **Work**: Card type is adult, activity is longer than 2h, trip before is not the last of the day
- **Study**: Card type is student, activity is longer than 2h, trip before is not the last of the day
- **Home**: Trip before is the last of the day
- **Other**: Activity is between 1 min and 2h long. Trip before is not the last of the day

**Criteria to distinguish destination from transfer**
- Time elapsed
- Transaction sequence
- Frequency of PT services
- Ratio: distance on the route / Euclidean distance
Zone of residence estimation for frequent users

Day 1. 07:18 am
Day 2. 07:38 am
Day 3. 10:53 am
Day 4. 09:02 am

Enrichment
Zone of residence estimation for frequent users

Enrichment
Data

• Period: 15/04/2013 - 21/04/2013

• Trips observed: 25,244,791
  – 1 stage 63.93%
  – 2 stages 29.91%
  – 3 stages 5.89%
  – 4 or more stages 0.27%

• Smart cards: 3,419,038

• Bus stops: 11,183

• Metro stations: 100

Trips used: 18,535,153 (73.42%)
1,294,049 (39%) satisfied the minimum number of trips observed in the time window.

888,970 (68%) have radii less than 1,000 m circumscribing all first-in-the-morning transactions.
We want to recover information

Network

Bus stops

Administrative limits

URBAN SPATIAL STRUCTURE
Methodology

Network
Weighted directed graph $G = (N, L, W)$

- $N$: Bus stops
- $L$: Link between two bus stops (trip)
- $W$: Trip volume

Center $\leftrightarrow$ Page rank

Hubs $\leftrightarrow$ Betweenness centrality (network, trips)

Communities $\leftrightarrow$ Infomap algorithm
Results -> Page Rank
Results -> Network Betweenness Centrality
Results - Trip stages Betweenness Centrality
Results -> Communities (10 most significant)
Results -> Communities (10 most significant)
Results -> Communities (10 most significant)
Results -> Communities (10 most significant)
Results -> Communities (10 most significant)

North
West
South
Southeast
East
Results -> Communities (10 most significant)
Results -> Communities (10 most significant)
Results-> Communities (10 most significant)
Results -> Communities (10 most significant)
Results -> Communities (10 most significant)
Results -> Communities (10 most significant)
Synthesis and Conclusions

• Successful application of network analysis methodology (Zhong et al) to Santiago
• Santiago different than Singapore
• Identification of CBD center-east (Page Rank)
• Transfer points (hubs) do not vary during the day
  – Can be identified with original and modified methodology
• Communities -> identifiable, but all have a connection to the CBD
Future Work

• Incorporate other data sources (mobile phone traces, GPS data from other transport modes)
• Analyze different periods of time (per hour, per day, weekday, weekend, etc)
• Analyze network structure
• Look at the evolution
Thanks!
References