Identifying Passengers who are at Risk of Reducing Public Transport Use: A Survival Time Analysis using Smart Card Data

Hiroaki Nishiuchi
(Kochi University of Technology)
Kochi city (Study area)

Study Area: Kochi City (Kochi Pref.)
Population: 340,000  (750,000)

Public Transport in Kochi City:
- Inter-City Train (JR line)
- Tram (private company)
- Bus (private company)

Smart Card Service
70 Tram Stops
1000 Bus Stops

Location of Kochi

Location of Tram and Bus Stops
Public Transport in Kochi
Smart Card “DESUCA”

Service start from 2009
- 50% of passengers
- Unique card ID
- Bus and Tram stops
- Time of Boarding and Alighting
- Card type
  - General card
  - Commuter pass
  - Elderly
  - Children
  - Handicap

My stocks of data
- 1\textsuperscript{st} - 30\textsuperscript{th} June 2010
- 1\textsuperscript{st} Sep – 31\textsuperscript{st} Dec 2012
- 1\textsuperscript{st} Sep 2014 – 31\textsuperscript{st} Oct 2015 (Number of card ID = 55,784)
- 1\textsuperscript{st} Jan 2009 – 30\textsuperscript{st} June 2017 (All data from PT authority)
Research framework

Step 1
Understand trip pattern change behavior
- Who stop to use public transport?

Step 2
Development of trip frequency change model
- Applying survival time analysis

Step 3
Identifying passengers who stop PT use
- factors affecting PT use stop behavior
- Suggest to PT authority

Change to positive spiral

- Increase number of passengers
- Increment of trip frequency

Decrease number of passengers

PT authority does not know

Change of trip pattern

Raise fare price

Cut of operation cost and service level

Reduction of income

Negative spiral of public transport in rural city

Smart Card Data

Increment of trip frequency

Suggest promotion
Trip frequency change model

Cox proportional hazard model (survival time analysis)

\[ h_k(t) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 \ldots + \beta_n x_n) \]

Previous study (Nisiuchi and Chikaraishi, 201?)

- Passengers who tend to reduce PT use are....
  - traveling same route everyday

- Passengers who tend to increase PT use are....
  - traveling long distance on holidays
  - traveling frequently use on weekdays

- Definition of PT use stop was not clear
  - it focused “his/her averaged number of trip in a week” by 4 month data
  - need to define by clear tendency of trip characteristics with longer data

Improve definition of public transport use stop
Pattern of interval of PT use days

Frequency of PT use **decrease** if interval becomes wider

<table>
<thead>
<tr>
<th>July</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th></th>
<th></th>
<th>August</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Relationship between the pattern and trip characteristics
- Considering special events, discount, point....
# Understanding daily PT use interval

Data: 1\textsuperscript{st} Sep 2014 – 31\textsuperscript{st} Oct 2015 (Number of card ID = 55,784)

<table>
<thead>
<tr>
<th>ID</th>
<th>2014/10/1</th>
<th>2014/10/2</th>
<th>2014/10/3</th>
<th>2014/10/4</th>
<th>2014/10/5</th>
<th>2014/10/6</th>
<th>2014/10/7</th>
<th>2014/10/8</th>
<th>2014/10/9</th>
</tr>
</thead>
<tbody>
<tr>
<td>6051</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6052</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6053</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6054</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6055</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6056</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6057</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6058</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6059</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6060</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6061</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6062</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6063</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

1 (red): Use, 0 (white): not use

Example; 600 passengers ⇒
Interval of daily PT use

Example of 600 passengers

Clustering daily PT use interval to

Identifying passengers who are ..... 

- **Low frequent passengers**(random use) 
- High frequent passengers 
- Sudden Stop or Start to use 
- Interval becomes wider 
- **Interval becomes Narrower**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of passengers</td>
<td>873</td>
<td>2,505</td>
<td>976</td>
<td>502</td>
<td>2,476</td>
<td>8,465</td>
<td>35,509</td>
<td>1,748</td>
<td>896</td>
<td>1,834</td>
</tr>
</tbody>
</table>
Clustering interval of daily PT use

Cluster 7 (35,509 ID)

365 days

Passengers

400 ID
Clustering interval of daily PT use

Cluster10 (1,834 ID)

Passengers

400 ID

365 days
Clustering interval of daily PT use

Cluster 3 (976 ID)

365 days

400 ID

Passengers
Clustering interval of daily PT use

Cluster 1 (873 ID)

365 days

Passengers

400 ID
Cluster 8 (1,748 ID)
Conclusions and Current works

Understanding interval of daily PT use

- Trip frequency change model by survival time analysis
- Data set of interval of daily PT use
- Understand the tendency of daily PT use interval

Current tasks

- Considering Spatial-Temporal trip characteristics
- Define and detect PT use stop behavior using 8 years data
- Understand passengers’ characteristics on PT use stop behavior
- Develop PT use frequency change model