A comparative study of urban mobility patterns using large-scale spatial-temporal data

DIT’18 17th November 2018
About DataSpark

- A mobility intelligence company.
- Generate insights from large-scale spatial-temporal mobile network data (up to 1TB per day).
- Applications and products in transportation, telco network planning, advertising, etc.
Motivation

Extracted mobility insights from large-scale spatial-temporal data in Singapore and Sydney:

- Travel frequency, distance and duration
- Mode of transportation
- Purpose of travel
- Mobility patterns and index
Prior Work

**Mobility Genome™ framework**

A framework to analyse spatio-temporal data at scale. Main components include:

- Preprocessing (trajectory, oscillation removal)
- Mobility units (Stay Points, Trips)
- Mobility mining (home/work, mode of transport)

T. A. Dang, J. Deepak, J. Wang, S. Luo, Y. Jin, Y. Ng, A. Lim, and Y. Li, “Mobility genome - a framework for mobility intelligence from large-scale spatio-temporal data”
Prior Work

Trip Extraction

1. Spatio-temporal data points
   \[ P_i = (\lambda_i, x_i, y_i, t_i) \]

3. Stay Point
   A location where a person has stayed for a while (based on a distance and duration threshold)

2. Trajectory
   \[ T(\lambda) = \{P_1, P_2, P_3, ..., P_n\} \]

4. Trip
   The movement between two stay points
Prior Work

Trip Attributes

Estimating trip distances

Straight line
Prior Work

Trip Attributes

Estimating trip distances

Straight line  Point-to-point
Prior Work

Trip Attributes
Estimating trip distances

- Straight line
- Point-to-point
- Line Simplification
Prior Work

Trip Attributes

Estimating trip distances

Inferring trip purpose

Type of activity associated with trip:
- Home ➔ Work
- Home ➔ Others
- Others ➔ Home
- Work ➔ Home
- Work ➔ Others
- Others ➔ Work
Prior Work

Trip Attributes

Estimating trip distances

Inferring trip purpose
Type of activity associated with trip:
- Home \( \rightarrow \) Work
- Home \( \rightarrow \) Others
- Others \( \rightarrow \) Home
- Work \( \rightarrow \) Home
- Work \( \rightarrow \) Others
- Others \( \rightarrow \) Work

Classifying Mode of transport
Focused on 2 primary modes: Road and train

A random forest model is trained using:
- labelled trips from a location sensing app and manual labelling
- Relevant features like distance to train and bus lines, heading change, acceleration, etc.
Trip and Origin-Destination Analysis

Number of trips per person

- Median number of trips per person is 3.
Trip and Origin-Destination Analysis

Number of trips per person

- Median number of trips per person is 3.
- Mode number of trips per person is 2, most likely a home-work trip.
Trip and Origin-Destination Analysis

**Number of trips per person**

- Median number of trips per person is 3
- Mode number of trips per person is 2, most likely a home-work trip.
- The long-tail distributions show that a small fraction of people behaves differently, having 10 up to 35 trips in a day. This could be due to the nature of their jobs (e.g. taxi drivers or delivery workers).
Trip and Origin-Destination Analysis

Number of trips per hour

- Temporal patterns of trips can serve as an indicator of traffic demand to help improve traffic conditions and reduce congestion (e.g. incentivising commuters that travel at off-peak hours).
Both cities has peaks during the morning and evening rush hours.
The peak hours are 7am and 6pm in Singapore and 8am and 5pm in Sydney, suggesting a longer workday in Singapore.
Trip and Origin-Destination Analysis

Number of trips per hour

| Source: Singapore Ministry of Manpower and Australian Bureau of Statistics |
|---------------------------------|------------------|
| Number of work hours per week of an average full-time worker |
| Singapore                        | 45.1             |
| Sydney                           | 40.7             |
Trip and Origin-Destination Analysis

Trip duration and distance

<table>
<thead>
<tr>
<th>Median trip duration (min)</th>
<th>Median trip distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singapore</td>
<td>24</td>
</tr>
<tr>
<td>Sydney</td>
<td>30</td>
</tr>
</tbody>
</table>

- The median trip duration and distance in both cities are similar and they are comparable with the data from public sources.
Trip and Origin-Destination Analysis

Origin-destination flows

- OD data records the population flow from one place to another.
- Trips generated from mobile network data provides insights on a much larger scale, and with greater reliability as compared to traditional survey data.
Trip and Origin-Destination Analysis

Origin-destination flows

- All top origin subzones belong to residential areas.
- The top destination subzone is the airport and from the remaining subzones, we can infer that most people are headed for work based on the type of industries or workplaces associated with the subzones.
Trip and Origin-Destination Analysis

Trip purpose

Trips to CBD - Weekday

- On a weekday in both cities, most trips to the CBD occur during the morning peak hours.
- There are much lesser work trips in Sydney than in Singapore.
Trip and Origin-Destination Analysis

Trip purpose
Trips to CBD - Weekday

<table>
<thead>
<tr>
<th>Main industries in Sydney's CBD</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial and insurance services</td>
<td>12.4%</td>
</tr>
<tr>
<td>Professional, scientific and technical services</td>
<td>15.1%</td>
</tr>
<tr>
<td>Accommodation and food services</td>
<td>21.8%</td>
</tr>
</tbody>
</table>

- The accommodation and food services industry is the most dominant.
- The large number of hotels, restaurants and places of interest in Sydney's CBD is likely to attract more tourists, providing a possible explanation for the lower proportion of work trips and validates our inferences of activity types.

Source: Australian Bureau of Statistics
Trip and Origin-Destination Analysis

Mode of transport

- The proportion of train trips in Singapore is much greater than in Sydney.
- The numbers are comparable with census data, giving us confidence in our mode detection methodology.
- Singapore is a small country with a denser and highly connected train network, while train lines in Sydney are more concentrated near the CBD and less accessible from other areas. As such, people further away from the train network would opt for alternative modes.

<table>
<thead>
<tr>
<th>Proportion of train trips</th>
<th>LBS</th>
<th>Public datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singapore</td>
<td>25.6%</td>
<td>23.9%</td>
</tr>
<tr>
<td>Sydney</td>
<td>5.74%</td>
<td>5.4%</td>
</tr>
</tbody>
</table>

Source: Singapore Household Interview Travel Survey, Singapore Public Data Portal and Sydney Household Travel Survey
Trip and Origin-Destination Analysis

Commute route

Top five roads with the most number of trips headed towards the CBD during morning peak hours

- Understanding the utilization of the road network would help transport planners to better predict traffic flow and volume.
- We applied a probabilistic map matching algorithm to project the trip data to actual road segments.
Mobility Motifs

**Definition:** Closed sequences of transitions among locations

**Methodology**

1. Select the motifs to be matched.
2. Group all OD pairs by individual sorted by time.
3. Identify the motifs; the underlying algorithm checks for loops. For example, consider the trip which falls under the classification of Motif 5 with the following pattern and OD pairs, i.e., A \(\rightarrow\) B \(\rightarrow\) C \(\rightarrow\) A and \((A_1, B_1), (B_2, C_1), (C_2, A_2)\). The algorithm will check if \(A_1\) matches \(A_2\), \(B_1\) matches \(B_2\), and \(C_1\) matches \(C_2\) based on a distance threshold \(d_{min}\).
Mobility Motifs

Distribution of detected motifs

![Graphs showing the distribution of detected motifs in Singapore and Sydney, with bars representing percentage of motifs and arrows indicating movement patterns.](image)
Mobility Motifs

Distribution of detected motifs

- Distributions in both cities are similar, with Motifs 2, 5, 8, and 12 as their top 5 pattern.
- As the motif patterns get progressively complicated, the number of individuals with such patterns decreases.
Mobility Index

Dimensions of mobility

1. Frequency of movement: the number of times a person goes anywhere
2. Spatial range: how far a person moves

Measures of mobility

<table>
<thead>
<tr>
<th>Measures of mobility</th>
<th>Dimensions of mobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of trips</td>
<td>Frequency</td>
</tr>
<tr>
<td>Number of unique grids visited</td>
<td>Spatial range</td>
</tr>
<tr>
<td>Number of unique stay points visited</td>
<td>Spatial range</td>
</tr>
<tr>
<td>Total distance travelled (km)</td>
<td>Frequency and spatial range</td>
</tr>
<tr>
<td>Total travel time (hr)</td>
<td>Frequency and spatial range</td>
</tr>
</tbody>
</table>

Computation of index

1. Compute all 5 mobility measures across each individual.
2. Check each measure for outliers greater (less) than $2\sigma$ from $\mu$ and reassign their values to be $\mu + 2\sigma (\mu - 2\sigma)$.
3. Normalize each data point using:
   \[
   \frac{\text{datapoint} - \min(\text{measure})}{\max(\text{measure}) - \min(\text{measure})}
   \]
4. Weigh all 5 measures equally to calculate the final index.
Mobility Index

Mobility index by city

- The median mobility index for Sydney is higher than Singapore.
Mobility Index

Median mobility index of residents by geographical region

- We divided Singapore and Sydney according to planning areas and SA3 respectively and assigned each region with a mobility index based on the median mobility index of its residents.
- Regions where we do not have statistically significant number of residents in our data were left out.
Mobility Index

Median mobility index of residents by geographical region

For both cities, areas with the most mobile residents tend to be in the suburbs while areas with the least mobile residents tend to be situated closer to the CBD.
Mobility Index
Median mobility index of residents by geographical region

For both cities, areas with the most mobile residents tend to be in the suburbs while areas with the least mobile residents tend to be situated closer to the CBD.
We divided the population into 4 different age groups, the young, young adults, adults and the old and sorted them according to their median mobility index.
Mobility Index

Mobility index by age

- In both cities, the young adults and adults, which makes up the bulk of the workforce are the most mobile.
Mobility Index

Mobility index by age

- Interestingly, the ordering of the young and the old age groups by mobility index differs between the two cities.
- In Sydney, the old are the least mobile followed by the young, which is what we would expect.
- In Singapore however, we observe the reverse.
Mobility Index

Mobility index by age

<table>
<thead>
<tr>
<th>Age</th>
<th>Singapore</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>14.2%</td>
<td>52.2%</td>
</tr>
<tr>
<td>55 - 59</td>
<td>74.9%</td>
<td>75.3%</td>
</tr>
<tr>
<td>60 - 64</td>
<td>63.6%</td>
<td>58.0%</td>
</tr>
<tr>
<td>&gt; 65</td>
<td>26.8%</td>
<td>13.4%</td>
</tr>
</tbody>
</table>

Labour force participation rate by age (2017)

- We hypothesize that mobility level correlates with one's employment status where people who are employed tend to be more mobile and it could explain the deviation seen.
Mobility Index

Mobility index by age

<table>
<thead>
<tr>
<th>Age</th>
<th>Singapore</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>14.2%</td>
<td>52.2%</td>
</tr>
<tr>
<td>55 - 59</td>
<td>74.9%</td>
<td>75.3%</td>
</tr>
<tr>
<td>60 - 64</td>
<td>63.6%</td>
<td>58.0%</td>
</tr>
<tr>
<td>&gt; 65</td>
<td>26.8%</td>
<td>13.4%</td>
</tr>
</tbody>
</table>

Labour force participation rate by age (2017)

- From the LBFR of both cities, the LFPR of the young in Singapore is the least among all age groups and significantly lesser than that of Australia, causing them to be the least mobile.
Mobility Index

Mobility index by age

<table>
<thead>
<tr>
<th>Age</th>
<th>Singapore</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>14.2%</td>
<td>52.2%</td>
</tr>
<tr>
<td>55 - 59</td>
<td>74.9%</td>
<td>75.3%</td>
</tr>
<tr>
<td>60 - 64</td>
<td>63.6%</td>
<td>58.0%</td>
</tr>
<tr>
<td>&gt; 65</td>
<td>26.8%</td>
<td>13.4%</td>
</tr>
</tbody>
</table>

Labour force participation rate by age (2017)

- The LFPR of the old (> 55) in Singapore is larger than in Australia. This is attributed to multiple factors such as favorable government policies, an increase in life expectancy and a tight job market. As such, the old in Singapore are more mobile than the young as many of them remain in the workforce past the retirement age.

Source: Singapore Ministry of Manpower and Australian Bureau of Statistics
Conclusion

Summary
- Analysed large-scale spatio-temporal data from mobile networks
- Presented various mobility insights about two cities, Singapore and Sydney
- Demonstrated the ability to support decision and policy making and complement methods to accurately capture various facets of human mobility

Future work
- Mobility motifs: add contextual location information to trips and to uncover habitual travel patterns by identifying recurrent motifs
- Others:
  - enhance spatial resolution through data fusion
  - conduct study over other regions (e.g. developing cities)
Thank you.