Analyzing & Visualizing Hot Spots Based on Mashup Technique

Handong Wang, Haixiang Zou, Yang Yue, Qingquan Li
Transportation Research Center
Wuhan University
Outline

• Background
• Problem statement & objective
• Methodology
• Prototype
• Concluding remarks
Background

- Large scale capture of vehicle trajectory
- Trajectory analysis can lead to insight on mobility, travel behavior
- Useful to LBS, transport management & planning
- Visualization tools absence when using existing analysis methods
Background -- Data Source: Taxi
480 Taxis, a Sunday
Over 40,000 pick-up & drop-off points
Problem Statement

• Dataset: taxi pick-up & drop-off points
• Objective:
  – Identifying hot spots
  – finding out time-dependent travel pattern
  – Integrating available tools under one framework
Analysis Method (1)

- Single-linkage clustering

\[ D_{KL} = \min_{i \in C_K, j \in C_L} d(x_i, x_j) \]
Analysis Method (2)

- Kernel Density Estimation

\[ f(s) = \sum_{i=1}^{n} \frac{1}{\pi r^2} k\left(\frac{d_{is}}{r}\right) \]
System Framework
## Clustering Result

<table>
<thead>
<tr>
<th>Time</th>
<th>Total points</th>
<th>No. of clusters (n&gt;50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00-11:00</td>
<td>6,997</td>
<td>16</td>
</tr>
<tr>
<td>11:00-13:00</td>
<td>5,539</td>
<td>11</td>
</tr>
<tr>
<td>13:00-16:30</td>
<td>8,812</td>
<td>25</td>
</tr>
<tr>
<td>16:30-20:00</td>
<td>8,890</td>
<td>22</td>
</tr>
<tr>
<td>20:00-24:00</td>
<td>9,913</td>
<td>23</td>
</tr>
</tbody>
</table>
Clustering Result (8:00-11:00)
Clustering result (11:00-13:00)
Clustering result (13:00-16:30)
Clustering result (16:30-20:00)
Clustering result (20:00-24:00)
Other Attempts

**Average-linkage**

Using average linkage clustering to find hot spots

**Kernel Density Estimation**

Using KDE to discover hot spots
Further Analysis
### Average Travel Distance between Clusters (20:00-24:00)

<table>
<thead>
<tr>
<th>O\D</th>
<th>C1</th>
<th>C421</th>
<th>C423</th>
<th>C621</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>--</td>
<td>9.20</td>
<td>11.7</td>
<td>6.51</td>
</tr>
<tr>
<td>C421</td>
<td>9.20</td>
<td>--</td>
<td>2.15</td>
<td>7.14</td>
</tr>
<tr>
<td>C423</td>
<td>11.7</td>
<td>2.15</td>
<td>--</td>
<td>8.65</td>
</tr>
<tr>
<td>C621</td>
<td>6.51</td>
<td>7.14</td>
<td>8.65</td>
<td>--</td>
</tr>
</tbody>
</table>
1. People tend to have larger travel scope and longer distance at the weekend night (6km²);
2. At the meantime, with smaller travel scope and more concentrated pattern at noon (<1km²);
3. In the rest of the day, people tend to travel around 3~4 km².
Flow Interactions

a) Flows origin from Cluster1

b) Flows origin from Cluster421

c) Flows origin from Cluster423

d) Flows origin from Cluster621
Flow Interaction among Clusters
(20:00-24:00)

<table>
<thead>
<tr>
<th>O/D</th>
<th>C1</th>
<th>C421</th>
<th>C423</th>
<th>C621</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>37.3%</td>
<td>0.2%</td>
<td>0.2%</td>
<td>0.3%</td>
</tr>
<tr>
<td>C421</td>
<td>3.7%</td>
<td>15.4%</td>
<td>2.9%</td>
<td>5.9%</td>
</tr>
<tr>
<td>C423</td>
<td>0.8%</td>
<td>5.9%</td>
<td>0.0%</td>
<td>29.7%</td>
</tr>
<tr>
<td>C621</td>
<td>2.9%</td>
<td>1.4%</td>
<td>0.0%</td>
<td>1.4%</td>
</tr>
</tbody>
</table>
Concluding Remarks (2)

1. The inner-cluster travel occupies the largest percentages which most possibly because people’s activities are usually occur within certain geographic scope;

2. To inter-cluster flows, closer clusters often attract more trips.

3. However, although distance is one of an important factor, it is not deterministic --- land use information is more important
Other Findings

1. People live in downtown area prefer to stay inside the area for various activity.;
2. Road traffic condition also affect people travel choices, such as not to cross the River (traffic congestion on two bridges;
Potential Applications

- Location-based services
- Location-based social network
- Travel demand analysis
- Transport & urban planning
- ... ...
Discussion

- Necessarily for refining clustering method
- Link POI with hot spot?
Thanks!

(Project is supported by MSRA)