Predicting Traffic Congestion through Mining Sensed Traffic Data

Hengfeng Li
Supervised by: Lars Kulik and Kotagiri Ramamohanarao
Department of Computing and Information Systems, University of Melbourne

Objectives
Traffic congestion leads to an increase in travel time and inaccurate time estimation for a trip. Learning from spatio-temporal traffic data can be considered as a key step to curb the effects of traffic congestion. Thus, we aim to:

- mine and analyse spatio-temporal traffic data, like density, mean speed, trip time
- model traffic condition based on road segments
- predict traffic congestion based on travel speed
- provide smarter navigating suggestions based on travel time and load balance

Introduction
Traffic congestion occurs when traffic demand of a road exceeds its capacity. Drivers slow down their speeds, and as a result traffic flow decreases dramatically. Being a critical matter, predicting this phenomena has drawn the attention of traffic authorities for years. In order to improve traffic condition via providing smarter trip recommendation, it is of importance to discover and predict traffic congestion. So far, our contributions consists of three aspects:

- Developing microscopic traffic simulator
- Surveying definitions of traffic congestion
- Modelling traffic from mining taxi trajectory

Microscopic Traffic Simulator

A traffic simulator benefits repeatable experiments and customised scenarios. Hence, a microscopic traffic simulator using real road network has been developed, which is based on the following models: Intelligent driver model (IDM) [1] and MOBIL lane-changing model [2]. Developing from a small scenario (Figure 1), we implement a decentralised traffic simulator with real road network. (Figure 2). The information of road network, including locations of intersections, traffic lights, and tram stops, is extracted from OpenStreetMap (http://www.openstreetmap.org/).

Definitions of Congestion

Congestion should be a concept to define how good traffic condition of a road segment is. The congestion of a route can be estimated by aggregating conditions of all road segments in this route. In the existing literature, several congestion definitions have been proposed:

- Traffic volume / flow
- Travel time
- Occupancy ratio (OR) / density
- Time mean speed or space mean speed
- Time headway or space headway
- Congestion index (CI)

These definitions, however, are interrelated.

Mining Taxi Trajectory

In order to model traffic condition of a city, we need to collect sensed data from vehicles by using GPS devices or loop detectors. It is hard to collect the traffic data from all vehicles, so mining from sampling data, such as buses or taxis, is helpful to understand and model the traffic of road network.

A taxi trajectory dataset was collected in Beijing (the capital of China) for a month on May 2009, see Table 1.

Table 1: Beijing taxi dataset and its road network

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of taxis</td>
<td>6,500</td>
</tr>
<tr>
<td>GPS points per day</td>
<td>≈ 4 millions</td>
</tr>
<tr>
<td>Sampling rate</td>
<td>1 – 5 mins</td>
</tr>
<tr>
<td>Area size</td>
<td>40 × 30 km²</td>
</tr>
<tr>
<td>Graph's nodes</td>
<td>131,479</td>
</tr>
<tr>
<td>Graph's edges</td>
<td>221,653</td>
</tr>
</tbody>
</table>

Future Work

In order to prove the correctness and accuracy of our microscopic traffic simulator, a calibration needs to be done by employing real traffic sensed data. In addition, the modelling of traffic is now using a simple approach to keep mean relative speed ratio for different time periods, which should be improved by using more sophisticated learning models, e.g., conditional random fields.

References


Acknowledgements

The taxi trajectories were obtained from Complex Engineered Systems Lab, Tsinghua University. Thanks for sharing this real world sensory dataset.

Contact Information

- Email: hengfeng.li@unimelb.edu.au
- Office: 6.05, Doug McDonell Building
- Office Phone: +61 3 8344 1549