Connected Traveler: Advanced Activity, Vehicle Data and Information Services Using Smartphones

Tamer Abdulazim
PhD Candidate
University of Toronto

11/06/2013
Overview

• Context and Motivation
• Research Challenges
• Conceptual Framework
• Experiments and Results
  – Collecting Location Data
  – Inferring Traveller’s Activity
  – Detecting Transportation Mode
• Case Study: Log a Day
Who wants more data?

Decision Makers

Planners

Operators

Data

Researchers

Travellers
Travel Diary

8:00AM  left home
Walking
8:30AM  at work
Subway  (station, route ...)
5:00PM  left work
Car     (distance, fuel ...)
5:15PM  at coffee shop
5:30PM  at grocery store
Research Challenges

- Survey burden
- Missing information
- Data post-processing
- Technological challenges e.g. Battery, GPS Signal
Research Objectives

• How to collect location data in a battery-friendly way?

• How to infer traveller’s activity with minimum burden?

• How to determine transportation mode in a generic way?

11/06/2013
Framework Overview

Data

Location

Activity

Transportation Mode

Network

Foursquare

Accelerometer Gyroscope

11/06/2013
Experiment 1

Collecting Location Data

11/06/2013
Collecting Location Data

“How to track location in a battery-friendly way?”

- Use network-based location provider

- Keep record of “landmarks”, e.g. Wi-Fi network, for frequently visited locations
## Collecting Location Data

<table>
<thead>
<tr>
<th>Participant Id</th>
<th>Network-based Location Accuracy (meters)</th>
<th>Data collection period (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>768</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>990</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>620</td>
<td>58</td>
</tr>
<tr>
<td>4</td>
<td>681</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>743</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>486</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>1130</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>123</td>
<td>43</td>
</tr>
<tr>
<td>9</td>
<td>908</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>378</td>
<td>45</td>
</tr>
</tbody>
</table>
Experiment 2

Inferring Traveller’s Activity
Infer Traveller’s Activity

“How to infer traveller’s activity with minimum burden?”

• Land use information
• Location-based services: Foursquare, Google Places, Yelp, Facebook, and YellowPages

Foursquare has 20+ million users, over 2 billion check-ins

11/06/2013
Acquiring Land Use Information

- **49,262** postal codes
- **62,493** venues
- **390** location categories
- Top location has **13,816 unique** check-in

Shop & Service (21%)
Travel & Transport (6%)
Great Outdoors (5%)
Arts & Entertainment (4%)
Professional & Other Places (19%)
College & University (3%)
Food (14%)
Nightlife Spots (3%)
Residence (25%)

11/06/2013
Experiment 3

Detecting Transportation Mode
Detecting Transportation Mode

“How to determine transportation mode in a generic way?”

• Smartphone **Motion Sensors**
  – Accelerometer
  – Gyroscope
  – Orientation

• **Machine Learning** Classifier
  – Neural Network
  – Support Vector Machine
  – **Random Forest**
Experiment Setup

• Date pre-processing
  – Average gyroscope data
  – Combine $n$ records (window size)
• Modes: walk, run, bike, car, bus and subway
• 133,886 records ($n = 7$)
• 65% training and 35% testing set
• Random Forest Classifier in Weka
## Mode Classification Results

### Classified As

<table>
<thead>
<tr>
<th>Classified As</th>
<th>Bus</th>
<th>Subway</th>
<th>Car</th>
<th>Bike</th>
<th>Running</th>
<th>Walking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus</td>
<td>2767</td>
<td>0</td>
<td>3</td>
<td>21</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Subway</td>
<td>3</td>
<td>8417</td>
<td>26</td>
<td>36</td>
<td>3</td>
<td>51</td>
</tr>
<tr>
<td>Car</td>
<td>0</td>
<td>45</td>
<td>11153</td>
<td>24</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>Bike</td>
<td>4</td>
<td>8</td>
<td>0</td>
<td>12937</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Running</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>24</td>
<td>2053</td>
<td>30</td>
</tr>
<tr>
<td>Walking</td>
<td>0</td>
<td>51</td>
<td>18</td>
<td>92</td>
<td>3</td>
<td>8614</td>
</tr>
</tbody>
</table>

### Performance Measures

<table>
<thead>
<tr>
<th>Weigted Average</th>
<th>TP-rate&lt;sup&gt;1&lt;/sup&gt;</th>
<th>FP-rate&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Precision (P)&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Recall (R)&lt;sup&gt;2&lt;/sup&gt;</th>
<th>F-Measure&lt;sup&gt;3&lt;/sup&gt;</th>
<th>Accuracy&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>98.9%</td>
<td>3%</td>
<td>98.9%</td>
<td>98.9%</td>
<td>98.9%</td>
<td>98.85%</td>
</tr>
</tbody>
</table>

11/06/2013
Case Study

Log a Day and Shuttle Program

11/06/2013
Shuttle Program

• Designed by Summerhill Impact
• Encourage drivers to reduce their emission footprint
• The program has three steps:
  1. Measure **baseline** driving pattern
  2. Eco-Driving **training**
  3. **Challenge** to driving impact by 10%

11/06/2013
On-board diagnostics (OBD)

OttoView Logger

Low-cost Bluetooth OBD

11/06/2013
## Trip History

### Total trips since 2013-02-24 is 46

- **Distance:** 30.38KM  
  **Duration:** 48 mins
- **Fuel:** 5.3869L  
  2013-04-15 18:11:06

#### Don Valley Pkwy N and ON-404 N
- **Distance:** 33.50KM  
  **Duration:** 73 mins
- **Fuel:** 7.2080L  
  2013-04-15 16:47:56

#### King's College Rd and College St
- **Distance:** 0.00KM  
  **Duration:** 0 mins
- **Fuel:** 0.0056L  
  2013-04-15 16:46:18

#### Highland Park Blvd and Nipigon Ave
- **Distance:** 0.34KM  
  **Duration:** 2 mins
- **Fuel:** 0.1101L  
  2013-04-06 23:39:56

#### Yonge St
- **Distance:** 26.66KM  
  **Duration:** 29 mins
- **Fuel:** 2.8295L  
  2013-04-06 23:05:15

#### Charles St W and Bloor St W
- **Distance:** 0.53KM  
  **Duration:** 4 mins
- **Fuel:** 0.1101L  
  2013-04-06 23:39:56
Fuel consumption as a function of driving distance
<table>
<thead>
<tr>
<th>Trip Duration</th>
<th>Distance from Network/GPS</th>
<th>Distance from OBD</th>
<th>Fuel Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes</td>
<td>KM</td>
<td>KM</td>
<td>Liters</td>
</tr>
<tr>
<td>9</td>
<td>1.66</td>
<td>1.4</td>
<td>0.53</td>
</tr>
<tr>
<td>13</td>
<td>2.7</td>
<td>2.4</td>
<td>0.68</td>
</tr>
<tr>
<td>22</td>
<td>8.39</td>
<td>8.8</td>
<td>0.93</td>
</tr>
<tr>
<td>0</td>
<td>0.33</td>
<td>0</td>
<td>0.03</td>
</tr>
<tr>
<td>40</td>
<td>9.57</td>
<td>7.3</td>
<td>1.29</td>
</tr>
<tr>
<td>10</td>
<td>3.87</td>
<td>3.3</td>
<td>0.62</td>
</tr>
<tr>
<td>159</td>
<td>232.82</td>
<td>227.9</td>
<td>27</td>
</tr>
<tr>
<td>37</td>
<td>47.38</td>
<td>45.6</td>
<td>4.3</td>
</tr>
<tr>
<td>34</td>
<td>45.44</td>
<td>44.2</td>
<td>4.3</td>
</tr>
<tr>
<td>33</td>
<td>45.26</td>
<td>44</td>
<td>4.16</td>
</tr>
<tr>
<td>120</td>
<td>186.91</td>
<td>184</td>
<td>18.05</td>
</tr>
</tbody>
</table>

11/06/2013
Summary

• **Network-based** location can be used for long-term data collection

• **Foursquare** and similar services contain rich land use information

• **Smartphone motion sensors** are capable of differentiating transportation mode

• Having a smartphone app could reduce the survey cost, allow real-time processing and support two-way communication
Future Work

• Develop a real-time and dynamic survey tool

• Assess the influence of information on travellers’ behaviour
Thank you!

tamer.abdulazim@utoronto.ca