Review of transportation mode detection approaches based on smartphone data

Marija Nikolić, Michel Bierlaire

May 18, 2017
Transportation mode detection (TMD)

- Environmental studies
- Urban planning
- Customized advertisements
- Context aware smartphones
- Travel demand estimation
- Real-time traffic state estimation
- Intelligent travel assistants
- PT companies
Travel surveys

Drawbacks:
Biased response
No response
Erroneous reporting
Smartphones: Mobile personal computers
Smartphone penetration

5.6 BILLION smartphone subscriptions by the end of 2019

Smartphones, mobile PCs, tablets and mobile routers with cellular connection

- 1.9 billion smartphone subscriptions
- 300 million mobile PCs, tablets and mobile router subscriptions
- 800 million mobile PCs, tablets and mobile router subscriptions

NOVEMBER 2013 ERICSSON MOBILITY REPORT
TMD: Procedure

Data → Pre-processing → Segmentation & Feature extraction

Classification → Mode detection
TMD: Sensor data

Sensors in Smartphone

Motion sensors
Position sensors
Environmental sensors
TMD: Sensor data

**Accelerometer**
- The acceleration force on all three physical axes
- Independence of any external signal sources
- Low energy consumption

**Global Positioning System (GPS)**
- The position and velocity information
- Outdoor context
- Reduced precision in dense urban environments
- Modest accuracy (50-80 meters)
- High power consumption
TMD: Sensor data

Cellular network signals: GSM

The fluctuation pattern of cell identifiers and signal strength

- Information on the position, outdoor and indoor contexts
- Precision: 50 - 200 meters, ping-pong effect

Data from mobile phone operators

- Anonymous location measurements, coarse-grained
TMD: Sensor data

WiFi
- Provides wireless connectivity to devices inside a WLAN
- Low positioning accuracy
- The most power-demanding sensor after GPS

Bluetooth
- Wireless connectivity and short range communication
- Sense devices in their vicinity
- Range: 10 - 100 meters
- Penetration rate: 7 - 11%

Barometers, thermometers, humidity sensors, cameras...
TMD: External data sources
TMD: Categories

Non-motorized

Motorized

- Person walking
- Person skating
- Bicycle
- Person sitting
- Train
- Car
- Motorbike
- Bus
- High-speed train
## TMD approaches: Comparison

<table>
<thead>
<tr>
<th>Source</th>
<th>Modes</th>
<th>Smartphone data</th>
<th>External data</th>
<th>Algorithm</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patterson et al. (2003)</td>
<td>Walking, Bus, Car</td>
<td>GPS</td>
<td>GIS</td>
<td>Bayes Model</td>
<td>84%</td>
</tr>
<tr>
<td>Muller (2006)</td>
<td>Walking, Stationary, Car</td>
<td>GSM</td>
<td>/</td>
<td>Artificial Neural Network, Hidden Markov Model</td>
<td>Average: 80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Walking: 87% Stationary: 98% Car: 75%</td>
<td></td>
</tr>
<tr>
<td>Sohn et al. (2006)</td>
<td>Walking, Stationary, Driving</td>
<td>GSM</td>
<td>/</td>
<td>Naïve Bayes, Support Vector Machines, heuristic-based methods, 2-stage boosted Logistic Regression</td>
<td>Average: 85%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Walking: 70.2% Stationary: 95.4% Driving: 84.3%</td>
<td></td>
</tr>
<tr>
<td>Mun et al. (2008)</td>
<td>Walking, Stationary, Driving</td>
<td>GSM, WiFi</td>
<td>/</td>
<td>Decision Trees</td>
<td>Average: 88%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Walking: 90.17% Stationary: 90.26% Driving: 87.83%</td>
<td></td>
</tr>
<tr>
<td>Zheng et al. (2008)</td>
<td>Walking, Biking, Driving</td>
<td>GPS</td>
<td>/</td>
<td>Graph-based</td>
<td>Average: 76.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Walking: 89.1% Biking: 66.6% Driving: 86.1%</td>
<td></td>
</tr>
<tr>
<td>Miluzzo et al. (2008)</td>
<td>Sitting, Stationary, Walking, Running</td>
<td>Accelerometer</td>
<td>/</td>
<td>JRIP rule learning</td>
<td>Average: 78.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sitting: 68.2% Stationary: 78.4% height Walking: 94.4% Running: 74.5%</td>
<td></td>
</tr>
</tbody>
</table>
## TMD approaches: Comparison

<table>
<thead>
<tr>
<th>Source</th>
<th>Modes</th>
<th>Smartphone data</th>
<th>External data</th>
<th>Algorithm</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reddy et al. (2010)</td>
<td>Walking, Stationary, Biking, Running, Motorized</td>
<td>GPS, Accelerometer</td>
<td>/</td>
<td>Naïve Bayes, Decision Trees, k-Nearest Neighbors, Support Vector Machines, k-Means Clustering, Continuous Hidden Markov Model, 2 stage Decision Tree and Discrete Hidden Markov Model</td>
<td>Average: 93.6% Walking: 96.8% Stationary: 95.6% Biking: 92.8% Running: 91% Motorized: 93.9%</td>
</tr>
<tr>
<td>Stenneth et al. (2011)</td>
<td>Walking, Bus, Car, Train, Stationary, Biking</td>
<td>GPS, GIS</td>
<td></td>
<td>Naïve Bayes, Decision Trees, Bayesian Network, Multilayer Perception, Random Forest</td>
<td>Average: 93.7% Walking: 96.8% Bus: 88.3% Car: 87.5% Train: 98.4% Stationary: 100% Biking: 88.9%</td>
</tr>
<tr>
<td>Montoya et al. (2015)</td>
<td>Walking, Biking, Bus, Train, Tram, Motorized</td>
<td>GPS, WIFI, Accelerometer, GSM, Bluetooth</td>
<td>Road maps, Rail maps, Public transport schedules, Public transport routes</td>
<td>Dynamic Bayesian Network</td>
<td>Average: 75.8% Walking: 91% Biking: 36% Bus: 80% Train and Motorized: 81% Tram: 91%</td>
</tr>
<tr>
<td>Chen and Bierlaire (2015)</td>
<td>Walking, Biking, Car, Bus, Metro</td>
<td>GPS, Bluetooth, Accelerometer</td>
<td>Open Street Map</td>
<td>Probabilistic method</td>
<td>SI &gt; 90%</td>
</tr>
<tr>
<td>Sonderen (2016)</td>
<td>Walking, Running, Biking, Car</td>
<td>Accelerometer, Gyroscope, Magnetometer</td>
<td>/</td>
<td>Decision Tree, Random Forest, k-Nearest Neighbors</td>
<td>98%</td>
</tr>
</tbody>
</table>
Comparison: Data sources

- Typically one or two sensors used: accelerometer and GPS
- External data: rarely used (transportation network data)
- Accuracy: higher if more data sources are utilized
Comparison: Classification algorithms

- Generative models: better suited when mobile phones are used only as a sensing system
- Discriminative models: better suited when detection is intended to run on mobile devices directly
  
  Decision Trees: satisfactory accuracy while using the least resources
Comparison: Categories & Accuracy

- Predominant: stationary, walking, biking and a unique motorized transport modes
- The best accuracy: walking and stationary modes
- Key challenge: differentiation between motorized classes (bus, car, train, metro)
- External data
  - Added value in detecting various motorized modes
  - Public transportation detection
Comparison: Performance

**Generative models: Chen and Bierlaire (2015)**

- Probabilistic method: the inference of transport modes and physical paths
- Structural travel model: captures the dynamics of smartphone users
- Sensor measurement models: capture the operation of sensors
- Categories: walking, biking, car, bus and metro
- Smartphone sensors: GPS, Bluetooth, and accelerometer
- External data: transportation network
Comparison: Performance

**Discriminative models: Stenneth et al. (2011)**

- Random Forests to infer a mode of transportation
- Findings supported by other studies: Abdulazim et al. (2013); Ellis et al. (2014); Shafique and Hato (2015)
- Categories: car, bus, train, walking, biking and stationary
- Smartphone sensors: GPS
- External data: transportation network
Conclusion

- Transportation mode detection based on smartphone data

- The approaches differ in terms of:
  - The type and the number of used input data
  - The considered transportation mode categories
  - The algorithm used for the classification task

- Accuracy: higher if more data sources are utilized

- External data: essential for the detection of various motorized modes
Future directions

- Studies with larger samples and over a longer time periods
- Water transportation modes
- Utilization of GSM logs provided by the operators
- Additional data sources
  - Barometers, temperature, humidity sensors
  - Real time traffic information
  - Socio-economic and demographic data
  - Mobility and transport census data
  - Seasonal data, weather conditions
- Transportation network data: OpenStreetMap
- Public transportation data: opendata.swiss
Thank you

STRC 2017 - 17th Swiss Transport Research Conference, Ascona

Review of transportation mode detection approaches based on smartphone data

Marija Nikolić, Michel Bierlaire

- marija.nikolic@epfl.ch
References I


References II


References IV


References V

