Dissecting traffic flows in congestion areas using GPS data

Feng Liu, Muhammad Adnan, Ansar yasar, Davy Janssens

Institute for MOBility (IMOB), Hasselt University, Belgium
Presentation Structure

- Problem statement and the goal of this study
- GPS data
- Methodology
- Preliminary results
Track & Know Introduction

The Track & Know platform is designed to provide a variety of tools that can operate stand alone or in conjunction with each other.

- Track & Know Big data Platform
  - Track & Know Tool boxes
    - Big data Processing
    - Big data Analytics
    - Big data Visual Analytics
  - Big data Use Cases
    - Three Pilots use cases; Mobility, Insurance and Health

The Mobility use case (Vodafone Innovus (VFI)) can be summed into 3 categories:

- Error detection and correction (GPS location, speed etc.)
- Route features aggregation and prediction (trajectory statistics, hot spots, prediction of a route and its components like average speed, mileage, fuel consumption)
- Driver behavior (categorize, correlate road condition with fuel and driver behavior)
**Problem statement**

- **Urban agglomerations**
- **Economic development**
  
  **Spatial layout is being reshaped**

- **Existing transport infrastructure**

---

**Growing mismatch**

between travel demand and transport services

---

**Serious problems**  
(e.g. congestion and pollution) !
The Goal

A method to

- **Analyze** mobility demand and underlying transport systems
- **Identify** areas with serious mismatch problems (congestion)
- **Dissect** traffic flows in the detected congestion areas
- **Find** alternative routes and **Suggest** policy measures
Presentation Structure

• Problem statement and the goal of this study

• GPS data

• Methodology

• Preliminary results
GPS data:

- Data: by Vodafone Innovus
  - Location based data
  - 7500 vehicles
  - 216GB
  - Period: 01/06/2017 – 01/06/2018
  - Vehicle types (15): Passenger Car, Bus, Van,...

- Data used for the preliminary results:
  - 12,213 Files under 55 Folders
  - 533 vehicles (7.1% of the total)
  - 3.46 GB
  - Period: 08/10/2017 – 12/10/2018 (5 weekdays)
Presentation Structure

• Problem statement and the goal of this study

• GPS data

• Methodology

• Preliminary results
Methodology:

1. City-wide travel pattern modeling
2. Congestion zone identification
3. Traffic flow dissecting
4. Alternative route searching

(5 steps!)
Step 1: Trip extraction

Each GPS point: $p_i(\text{lat}_i, \text{lon}_i, t_i, v_i, \text{engineStatus}_i, \text{vehicle}_i, \text{vehicleT}_i)$

Stop points ($\text{engineStatus}_i=0$ and $v_i=0$)

Non-stop points

Trips ($\text{Distance}>\text{TH}_{\text{trip}}$)

Distance $\text{Distance}>\text{TH}_{\text{trip}}$?

$p_1 \rightarrow p_2 \rightarrow \ldots \rightarrow p_i \rightarrow \ldots \rightarrow p_j \rightarrow \ldots \rightarrow p_n$

Non-stop points

Stop points

Distance $\text{Distance}>\text{TH}_{\text{trip}}$?
Travel time: \[ T_{Trip} = t_d - t_o \]

Link \( (p_o, p_d) \):
\[ \text{Link}(p_o, p_d) = \sum_{i = 0}^{d - 1} D(p_i, p_{i+1}) \]
Step 2: City-wide travel pattern modeling

- Spatial partition
  - zones: $Grid_X \times Grid_Y$
  - $r_z$, $z=1...Grid_X \times Grid_Y$

- Temporal partition
  - $TimeP$: different time periods of a day
  - $DayT$: different types of the day
• Zone-traffic-condition-matrix: $Z(r_z, TimeP, Day, DayT)$
• OD-travel-pattern-matrix: $OD(r_o, r_d, TimeP, Day, DayT)$
Zone-traffic-condition-matrix $Z(r_z, TimeP, Day, DayT)$: average driving speed of each zone

- Each matrix element
  - Total number of points $M_z$ in $r_z$
  - Average speed of the points $V_z \equiv V_z(r_z, TimeP, Day, DayT) = \frac{\sum_{k=1}^{M_z} v_z^{(k)}}{M_z}$
OD-travel-pattern-matrix $OD(r_o, r_d, TimeP, Day, DayT)$: travel demand between origin and destination zones

- Each matrix element
  - Total number of trips between $r_o$ and $r_d$
  - Average travel time, speed and route directness over all the trips
Average travel time:

\[ U_{o,d} = U_{o,d}(r_o, r_d, TimeP, Day, DayT) = \frac{1}{M_{o,d}} \sum_{k=1}^{M_{o,d}} U_k \]

Average travel speed:

\[ V_{o,d} = V_{o,d}(r_o, r_d, TimeP, Day, DayT) = \frac{1}{M_{o,d}} \sum_{k=1}^{M_{o,d}} V_k \]

Average route directness:

\[ R_{o,d} = R_{o,d}(r_o, r_d, TimeP, Day, DayT) = \frac{1}{M_{o,d}} \sum_{k=1}^{M_{o,d}} R_k \]
Step 3: Congestion zone identification

- Zone-traffic-condition-matrix: $Z(r, TimeP, Day, DayT)$

- Zones with congestion on a day
  - Total number of points: $M_z > TH_{MZ}$
  - Average speed: $V_z < TH_{VZ}$

- The probability of congestion on multiple days
  - $P > TH_P$
Step 4: Traffic flow dissecting in a congestion area

- OD-travel-pattern-matrix: $OD(r_o, r_d, TimeP, Day, DayT)$
- Project each OD trips into zones

$p_o \rightarrow \cdots p_i \rightarrow p_{i+1} \rightarrow \cdots \rightarrow p_d \quad r_o \rightarrow r_g \rightarrow r_z \rightarrow r_d$
Step 5: Alternative route searching

- Search for alternative routes that avoid the congestion zones
Congestion
Presentation Structure

- Problem statement and the goal of this study
- GPS data
- Methodology
- Preliminary results
Result 1: Trips

- 23934 trips over 533 vehicles and 5 weekdays
- 8.98 trips/vehicle and day
Fig. Distribution of average speeds and number of trips per half an hour
Fig. Seven Time periods

- 7-9: morning commute (35.2 km/h)
- 10:30-15: noon rush (34.3 km/h)
- 17-19: evening commute (41.5 km/h)
- 19-21: evening rush (39.2 km/h)

Average speed (km/h)
Result 2: City-wide travel pattern modelling

- Zone-traffic-condition-matrix: $Z(r_z, TimeP, Day, DayT)$
- OD-travel-pattern-matrix: $OD(r_o, r_d, TimeP, Day, DayT)$

- $z=1,...,400 \times 400$, Each zone: 1.62 km$^2$
- $TimeP=1...7$
- $Day=1...5$
- $DayT=\text{weekdays, weekends and holidays}$

In this case study, morning commute (7-9am) on weekdays!
Result 3: Congestion zones ($M_z > 50$/day, $V_z < 20$, and $P \geq 0.8$)

Fig. The average speed and number of points each zone in the morning
Fig. Congestion zones (125)
Result 4: Traffic flow dissecting

<table>
<thead>
<tr>
<th>Region_x_id_old</th>
<th>Region_y_id_old</th>
<th>Average_Speed_Zone</th>
<th>Number_Of_Points_Zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>219</td>
<td>5.7049180328</td>
<td>183</td>
</tr>
<tr>
<td>2</td>
<td>219</td>
<td>6.2380659701</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>328</td>
<td>6.6515161515</td>
<td>66</td>
</tr>
<tr>
<td>4</td>
<td>180</td>
<td>6.8055555556</td>
<td>72</td>
</tr>
<tr>
<td>5</td>
<td>208</td>
<td>7.8961038961</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>213</td>
<td>113</td>
<td>116</td>
</tr>
<tr>
<td>7</td>
<td>280</td>
<td>84</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>240</td>
<td>9.74</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>217</td>
<td>10.245283019</td>
<td>53</td>
</tr>
<tr>
<td>10</td>
<td>204</td>
<td>10.324323423</td>
<td>74</td>
</tr>
<tr>
<td>11</td>
<td>201</td>
<td>11.053571249</td>
<td>56</td>
</tr>
<tr>
<td>12</td>
<td>175</td>
<td>12.792627737</td>
<td>137</td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>12.85936825</td>
<td>63</td>
</tr>
<tr>
<td>14</td>
<td>369</td>
<td>11.205479452</td>
<td>219</td>
</tr>
<tr>
<td>15</td>
<td>181</td>
<td>12.696202532</td>
<td>79</td>
</tr>
<tr>
<td>16</td>
<td>157</td>
<td>12.722627737</td>
<td>137</td>
</tr>
<tr>
<td>17</td>
<td>175</td>
<td>12.73719487</td>
<td>156</td>
</tr>
<tr>
<td>18</td>
<td>182</td>
<td>12.9037669</td>
<td>239</td>
</tr>
<tr>
<td>19</td>
<td>348</td>
<td>12.94268596</td>
<td>173</td>
</tr>
<tr>
<td>20</td>
<td>286</td>
<td>13.4</td>
<td>55</td>
</tr>
<tr>
<td>21</td>
<td>350</td>
<td>13.10343783</td>
<td>115</td>
</tr>
<tr>
<td>22</td>
<td>348</td>
<td>13.226315789</td>
<td>190</td>
</tr>
<tr>
<td>23</td>
<td>286</td>
<td>13.8</td>
<td>55</td>
</tr>
<tr>
<td>24</td>
<td>346</td>
<td>13.43636363</td>
<td>55</td>
</tr>
<tr>
<td>25</td>
<td>286</td>
<td>13.592592593</td>
<td>21</td>
</tr>
<tr>
<td>26</td>
<td>327</td>
<td>13.755102041</td>
<td>735</td>
</tr>
<tr>
<td>27</td>
<td>177</td>
<td>13.833719298</td>
<td>228</td>
</tr>
<tr>
<td>28</td>
<td>19</td>
<td>14.207792208</td>
<td>77</td>
</tr>
<tr>
<td>29</td>
<td>345</td>
<td>14.23655914</td>
<td>186</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>14.381578947</td>
<td>76</td>
</tr>
<tr>
<td>31</td>
<td>375</td>
<td>14.407407407</td>
<td>81</td>
</tr>
<tr>
<td>32</td>
<td>325</td>
<td>14.46573425</td>
<td>73</td>
</tr>
<tr>
<td>33</td>
<td>15</td>
<td>14.5108656566</td>
<td>92</td>
</tr>
<tr>
<td>34</td>
<td>347</td>
<td>14.67935633</td>
<td>135</td>
</tr>
<tr>
<td>35</td>
<td>183</td>
<td>14.568668407</td>
<td>1915</td>
</tr>
<tr>
<td>36</td>
<td>192</td>
<td>14.57655656</td>
<td>471</td>
</tr>
</tbody>
</table>

35, (183, 228), 14.6km/h, 383/day
(183,228): 217 trips

- Start and/or end: 41.9%
  - Start: 16.1%
  - End: 4.1%
  - Inside: 21.7%

- Passing: 58.1%
  - 6->3: 17.1%
  - 1->8: 6.9%
  - 6->8: 4.1%
  - 7->3: 4.1%
  - 3->6: 3.7%
  - 7->2: 2.7%
• **Passing**: 58.1%
  - 6->3: 17.1%
  - 1->8: 6.9%
  - 6->8: 4.1%
  - 7->3: 4.1%
  - 3->6: 3.7%
  - 7->2: 2.7%
Fig. Major routes
Result 5: Alternative route searching
Future work

• *Extend* the size of the data to be used

• *Find* alternative routes and *compare* these routes with the currently observed ones

• *Visualize* the results using visualization tools

• *Suggest* policy measures

• *Change* regular grids into spatial compartment?
Thank you!

(feng.liu@uhasselt.be)