Transportation Mode Detection using Random Forest

Jasna Urbančič, Veljko Pejović, Dunja Mladenić

SiKDD 2018, 11 October 2018
Motivation

- APIs already support fine-grained classification of non-motorized forms of transportation (still, walking, running, cycling, and in vehicle)

- Not enough for tracking and routing purposes, specially in use-cases for urban environments
Goals

Focus on:

1. Feature extraction
2. Feature analysis

Main contribution:

Feature analysis, which revealed the impact of each feature to the classification scores
Our approach

1. Data acquisition
   1a. Mobile applications
2. Pre-processing
   2a. Resampling
   2b. Filtering
   2b. Gravity estimation
3. Feature extraction
4. Feature analysis
   4a. Correlation analysis
   4b. Statistical analysis
5. Classification
   5a. Defining feature sets
   5b. Choosing classifiers
Feature domains

- **Statistical** – mean, standard deviation, skewness, percentiles, etc.
- **Time** – integral and double integral of signal over time, zero crossings, etc.
- **Frequency** – spectral energy, spectral entropy, spectrum peak position, etc.
- **Peak** – volume, height, width of a peak, etc.
- **Segment** – peak frequency, stationary duration, etc.
Extraction of peak-based features

1. Convolute with a box window
2. Split the signal on acceleration and deceleration
3. Find peaks
4. Count or compute:
   - Number of peaks
   - Mean
   - Standard deviation
   - Skewness
   - Peak height
   - Peak width
   - Peak width height
   - Peak area

Signal → Peak-based features
Evaluation

Change model parameters

(2) Validate

Evaluate

Join datasets

Join datasets

(3) Train + Validate

Use best parameters

(4) Test and evaluate

Distribution of samples across modes

Composition of the set
## Feature sets

<table>
<thead>
<tr>
<th>Set</th>
<th>Acceleration</th>
<th>Features</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-S</td>
<td>Dynamic</td>
<td>Statistical</td>
<td>54</td>
</tr>
<tr>
<td>D-SF</td>
<td>Dynamic</td>
<td>Statistical, Frequency</td>
<td>94</td>
</tr>
<tr>
<td>D-SFP</td>
<td>Dynamic</td>
<td>Statistical, Frequency, Peak</td>
<td>172</td>
</tr>
<tr>
<td>H-S</td>
<td>Horizontal</td>
<td>Statistical</td>
<td>54</td>
</tr>
<tr>
<td>H-SF</td>
<td>Horizontal</td>
<td>Statistical, Frequency</td>
<td>94</td>
</tr>
<tr>
<td>H-SFP</td>
<td>Horizontal</td>
<td>Statistical, Frequency, Peak</td>
<td>172</td>
</tr>
<tr>
<td>ALL</td>
<td>ALL</td>
<td>ALL</td>
<td>376</td>
</tr>
</tbody>
</table>
Results

<table>
<thead>
<tr>
<th>Set</th>
<th>CA</th>
<th>RE</th>
<th>PR</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-S</td>
<td>0.48</td>
<td>0.41</td>
<td>0.39</td>
<td>0.37</td>
</tr>
<tr>
<td>D-SF</td>
<td>0.60</td>
<td>0.41</td>
<td>0.41</td>
<td>0.39</td>
</tr>
<tr>
<td>D-SFP</td>
<td>0.46</td>
<td>0.39</td>
<td>0.40</td>
<td>0.35</td>
</tr>
<tr>
<td>H-S</td>
<td>0.64</td>
<td>0.40</td>
<td>0.43</td>
<td><strong>0.41</strong></td>
</tr>
<tr>
<td>H-SF</td>
<td>0.53</td>
<td>0.39</td>
<td>0.43</td>
<td>0.36</td>
</tr>
<tr>
<td>H-SFP</td>
<td>0.50</td>
<td>0.37</td>
<td>0.40</td>
<td>0.34</td>
</tr>
<tr>
<td>ALL</td>
<td>0.47</td>
<td>0.35</td>
<td>0.40</td>
<td>0.33</td>
</tr>
</tbody>
</table>

- Using peak features in combination with the other features sets results in the decrease in F1 score.
- F1 and CA increase when we add frequency-based features in case of dynamic acceleration and decrease in case of a similar action for horizontal acceleration.
- Smaller features generally perform better than larger
Feature selection

• **Backward feature selection (elimination)**
  - We remove features from the feature set
  - 28 features – more peak-based features than statistical, dynamic and horizontal acceleration appear in similar proportions

• **Forward feature selection**
  - We add features to the feature set
  - 10 features – mostly statistical features, followed by peak-based features, extracted from dynamic acceleration

• Only one frequency-based feature in each of these feature sets
## Results

### Feature set

<table>
<thead>
<tr>
<th>Feature set</th>
<th>CA</th>
<th>RE</th>
<th>PR</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward selection (10)</td>
<td>0.70</td>
<td>0.50</td>
<td>0.47</td>
<td>0.48</td>
</tr>
<tr>
<td>Backward elimination (28)</td>
<td>0.73</td>
<td>0.50</td>
<td>0.48</td>
<td>0.49</td>
</tr>
</tbody>
</table>

### Forward selection

<table>
<thead>
<tr>
<th>T/P</th>
<th>Car</th>
<th>Bus</th>
<th>Train</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>0.78</td>
<td>0.27</td>
<td>0.05</td>
</tr>
<tr>
<td>Bus</td>
<td>0.51</td>
<td>0.40</td>
<td>0.09</td>
</tr>
</tbody>
</table>
| Train   | 0.47 | 0.21 | **0.32** |}

### Backward elimination

<table>
<thead>
<tr>
<th>T/P</th>
<th>Car</th>
<th>Bus</th>
<th>Train</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>0.83</td>
<td>0.12</td>
<td>0.05</td>
</tr>
<tr>
<td>Bus</td>
<td>0.55</td>
<td><strong>0.35</strong></td>
<td>0.10</td>
</tr>
<tr>
<td>Train</td>
<td>0.45</td>
<td>0.23</td>
<td><strong>0.32</strong></td>
</tr>
</tbody>
</table>
Conclusion

- Using feature selection we are able to improve classification scores for at least 0.04, in some cases even over 0.10.
- Most of the non-car samples are still misclassified as cars.
- Peak-based features did perform well in predefined feature sets, they consistently appear among selected features in feature selection.