On Movement Activities Estimation for Mobile Opportunistic Networking

Austrian-Italian Workshop on Future Internet Challenges
Università di Pavia - May 4, 2011

Andrea Hess
University of Vienna,
Research Group Entertainment Computing

Overview

- Movement feature extraction from every-day trips
- Method for estimating type of movement activity based on features
- Final objectives
  - Investigate use of activity information in opportunistic networks
  - Estimate preferable forwarding nodes in delay-tolerant routing
- First results described in:
Outline of the Talk

• Motivation
  - Mobility-assisted forwarding in opportunistic networks
• Deriving features from mobility traces
• Use mobility features to estimate activities
• Experimental results
  - Empirical CDF of mobility features
  - Movement activity classification based on Naïve Bayes
• Conclusion and future work

Mobility-Assisted Forwarding in Opportunistic Networks

• Concept of opportunistic networking
  - Disseminate data in store-and-forward manner by mobile devices connecting ad hoc
  - Exploit networking opportunities of moving devices
• Select best forwarding node
  - predict forwarding capabilities of each device
  - e.g., expected traveling distance and likelihood of revisiting locations
• How are forwarding metrics affected by movement patterns?
Effects of Activities on Networking

- **Movement Activity**
  - composition of movements, name corresponds to trip purpose

- Bayes classification approach
  - detect activity given a specific mobility feature vector

\[
P(A_j|V_i) = \frac{P(V_i|A_j)P(A_j)}{P(V_i)}
\]

- \( A_j \) ... movement activity
- \( V_i \) ... feature vector observed

- GPS-based positioning method
  - offers finer granular movement information than cellular or WLAN data

- Define feature set based on literature survey
Typical Features Extracted - Summary of Literature Survey

- **Spatial characteristics**
  - **Prevalence** - fraction of time user spends at an AP
  - spatial distribution -> 'location visiting preferences', 'hotspot regions'
  - **Activity range** - area covering all locations that have been visited

- **Temporal characteristics**
  - **Pause time**
  - **Persistence** - time a user stays continuously connected to one AP

- **Spatio-temporal characteristics**
  - **Revisit metrics**, e.g., 'periodical re-appearances', 'return time'
  - **Meeting metrics**, e.g., 'inter-contact time', 'contact time', 'inter-meeting time', 'time distance'

---

Feature Extraction I

- **Direct metrics derived from traces**
  - 1. **Velocity**
    - speed between two positions measured consecutively (position sampling interval is 1 s)
  - 2. **Direction changes**
    - difference between current direction and direction measured 20 m earlier

- **Spatial metrics**
  - 3. **Flight length**
    - length of path (in meters) traveled between two consecutive pauses
  - 4. **Mobility range**
    - distance of GPS position to the center of rectangle covering trip
Feature Extraction II

- Temporal metrics
  - 5. Pause time
    duration between two consecutive movement phases ($v<0.5 \text{ m/s for } t\geq 5 \text{ s}$)
  - 6. Start time
    hour of day trip started

- Combined metrics
  - 7. Number of revisits of a position
    position is assumed to be the same if within position radius of 20 m
  - 8. Time between revisits

Study of Movement Features for Activity Estimation

- Can activities be recognized based on mobility feature set?

- 4 types of movement activities considered
  - Way to work
  - Evening activity
  - Shopping activity
  - Tourist activity

- Data set of 252 trips
  - GPS traces of daily trips of 13 test persons
  - semantic information about trips noted by test persons

- Empirical CDF for 8 features
  - determine suitability for activity estimation

- Naïve Bayes classifier
  - categorize trips into 4 activities
Features Observed - Ex 1: Flight Length

- **Tourist**: 80% below 230 m
  corresponds to the behavior of tourists walking between sights
- **Way to Work** and **Evening**: longer flight lengths due to public means of transport and cars

Features Observed - Ex 2: Time between Revisits

- **Way to Work**: 99% of return times < 160 s
  (e.g., walking back a street after getting off a bus)
- **Evening**: highest values due to longer pauses and similar return paths
- **Shopping**: 90% of positions revisited after 20 min,
  small number up to 75 min
Features Observed - Ex 3: Start Time

- **Way to Work**: 55% from 6:00 to 10:00, wider range for second bunch of trips
- **Evening**: 17:00-23:00 as one would expect
- **Shopping and Tourist**: Tourist curve shows steeper ascent in the afternoon, Shopping trips start earlier in the morning

![Graph showing ECDF of start times for different activities](image)

**Movement Activity Recognition - Classification Results**

- Training of Naïve Bayes classifier
  - 50% of trips in each activity category
- Overall success rate of 80.65%
  - Classification matrix:

<table>
<thead>
<tr>
<th>Assigned label</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evening (A)</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Shopping (B)</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Tourist (C)</td>
<td>0</td>
<td>6</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Way to Work (D)</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>68</td>
</tr>
</tbody>
</table>

- Trips classified wrongly expose similarities
  - e.g., some Tourist trips show similarities with Shopping trips
  - Way to Work trips classified as Evening trip if taking place with pauses at a later hour
**Fitting Set of Movement Features**

- Test all feature combinations
  - classify test set by using reduced feature set (1 to 7 features out of 8)
  - Mean success rate [%] and standard deviation:

<table>
<thead>
<tr>
<th>Number of features (n)</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>79.74</td>
<td>77.71</td>
<td>76.40</td>
<td>73.76</td>
<td>71.39</td>
<td>67.68</td>
<td>62.30</td>
</tr>
<tr>
<td>Stdv</td>
<td>1.32</td>
<td>1.95</td>
<td>2.63</td>
<td>2.93</td>
<td>4.07</td>
<td>4.71</td>
<td>3.65</td>
</tr>
</tbody>
</table>

- Combination of six features
  - best combination of six features achieves 80.65% as well
  - omitting important features - num. of revisits, start time - yields lowest rate

- Combination of two features
  - num. of revisits and start time achieved best result

**Conclusion and Future Work**

- Activity estimation based on mobility features
  - 4 typical movement activities, 8 features
  - 80.65% success rate for Naïve Bayes approach

- Potential use of activity estimation
  - forwarding in mobility-assisted networks
  - estimate user-caused network traffic
  - applications: situation-aware services

- Current and Future Work
  - extend data set of daily trips
  - correlation betw. activities and forwarding behaviour
  - propose best forwarding node estimation for routing mechanism