Activity Identification from GPS Trajectories Using Spatial Temporal POIs’ Attractiveness

Lian Huang, Qingquan Li, Yang Yue
State Key Laboratory of Information Engineering in Survey, Mapping and Remote Sensing, Wuhan University
Outline

• Introduction
• Motivations
• Spatial Temporal POI’s Attractiveness (STPA)
• Activity Identification Using STPA
• Experiments and Evaluations
• Conclusions and Future Work
Introduction

- Fast development of positioning and communication technologies enables GPS devices based traffic data collection as a potential substitution of traditional travel surveying methods:
  - more accurate and reliable information.
  - the participants’ burden is reduced substantially
- As personal wearable GPS receivers become available, large scale capture of individual daily trajectories has become technically and economically feasible
  - data processing steps are required
  - Activity Identification aiming at discovering activities in trajectories since travel purposes are obviously not included in GPS traces
Introduction

- Extracting activities helps semantically organizing and interpreting GPS tracking trajectories.
- Basic components of an activity are activity location, start time, duration and purpose.
- Since GPS tracking data are basically spatial temporal movements, from time geography’s perspective of view, a home-work-park-home personal trajectory is presented as follows:
Motivations and Related Work

• Motivations
  – Available GPS tracking data obviously do not include activity purpose information
  – With previous consideration, if each activity spot’s spatial-temporal influence prism can be defined, we can identify the possible activity locations by discovering the trajectories and prisms’ relationships in time-geographical space
Spatial Temporal POI’s Attractiveness (STPA)

- **Traditional POI’s Attractiveness:**
  - Majority of the discussion on POI’s attractiveness appears in traditional transportation field
  - *gravity model*: the more distant the less attractiveness:
    - Reilly’s Law of retail gravitation:
    - Huff’s model:
  - **Posterior Analysis**: historical FCD
- **POI’s attractiveness is actual time-varying**
  - different time of day utilities different activities

---

Fig. 2. EPSCSUP: Estimated starting time preferences on major daily activities.
Spatial Temporal POI’s Attractiveness (STPA)

• Definition 1 <Spatial Temporal POIs’ Attractiveness> A POI P’s Spatial Temporal Attractiveness is a spatial temporal prism whose area of time t’s snapshot is a function of the POI’s static factors and the value of dynamic function at time t.

• Considering POIs as normal facilities, their attractiveness will be decided by some intrinsic factors. For example, larger shopping malls usually attract more people if they go for purchases. We define this kind of factors as static factors and they are described as follows:
  – The size of POI’s $S$
  – The fame of POI $\varepsilon$
  – Category of POI $C$

• POIs attractiveness will change as time flies forward;
  – This kind of time varying factor is defined as dynamic factor of STPA, and is determined by POI’s category, time of day $t$, and type of day $D$ (weekday or holiday).
Spatial Temporal POI’s Attractiveness (STPA)

Distinguished from facilities’ attractiveness in transportation research, STPA should not only be a value but also describe the influence region that will contain potential activity in trajectories.

\[
\begin{align*}
  r &= \sqrt{S^* (w_1 e_p + w_2 C)/\pi} \\
  w_1 + w_2 &= 1
\end{align*}
\]

\[
\begin{align*}
  r'(t) &= f(C, t, D) \times r \\
  STPA_{p}(x, y, t) : \begin{cases} x = r'(t) \times \cos \theta \\
                                                                    y = r'(t) \times \sin \theta \\
                                                                    t = t \end{cases}
\end{align*}
\]
In order to figure out *dynamic functions* of different type of POIs, we conducted a survey to find out people’s choices through an internet investigation in which total 247 participants vote for the result;

We divided POIs into 6 activity-related categories:

- “Dining” refers to restaurant POIs;
- “Shopping_1” refers to shopping malls/retail stores those be open from 9:00 to 21:00
- “Shopping_2” refers to functional malls/retail stores those be open from 9:00 to 18:00
- “Entertainment” includes cinema, pub and POIs related to leisure life
- “Public facilities” includes hospital, university, public park and the like
- “Others” refers to the rest of POIs
<table>
<thead>
<tr>
<th>Time/Type</th>
<th>Dining</th>
<th>Shopping_1</th>
<th>Shopping_2</th>
<th>Entertainment</th>
<th>Public facility</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:00-7:00</td>
<td>very low</td>
<td>very low</td>
<td>very low</td>
<td>very low</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>7:00-9:00</td>
<td>high</td>
<td>very low</td>
<td>very low</td>
<td>very low</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>9:00-10:00</td>
<td>medium</td>
<td>low</td>
<td>medium</td>
<td>low</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>10:00-11:00</td>
<td>medium</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
</tr>
<tr>
<td>11:00-12:00</td>
<td>very high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
</tr>
<tr>
<td>12:00-13:00</td>
<td>very high</td>
<td>low</td>
<td>low</td>
<td>medium</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>13:00-14:00</td>
<td>high</td>
<td>low</td>
<td>medium</td>
<td>medium</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>14:00-15:00</td>
<td>medium</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
</tr>
<tr>
<td>15:00-16:00</td>
<td>low</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
</tr>
<tr>
<td>16:00-17:00</td>
<td>medium</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
</tr>
<tr>
<td>17:00-18:00</td>
<td>high</td>
<td>medium</td>
<td>high</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>18:00-19:00</td>
<td>very high</td>
<td>low</td>
<td>very low</td>
<td>low</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>19:00-20:00</td>
<td>very high</td>
<td>high</td>
<td>very low</td>
<td>high</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>20:00-21:00</td>
<td>medium</td>
<td>high</td>
<td>very low</td>
<td>very high</td>
<td>medium</td>
<td>low</td>
</tr>
<tr>
<td>21:00-22:00</td>
<td>medium</td>
<td>low</td>
<td>very low</td>
<td>high</td>
<td>medium</td>
<td>low</td>
</tr>
<tr>
<td>22:00-23:00</td>
<td>low</td>
<td>very low</td>
<td>very low</td>
<td>medium</td>
<td>medium</td>
<td>very low</td>
</tr>
<tr>
<td>23:00-24:00</td>
<td>low</td>
<td>very low</td>
<td>very low</td>
<td>low</td>
<td>medium</td>
<td>very low</td>
</tr>
</tbody>
</table>

Roll data

Transportation Research Center (TRC), Wuhan University
Dynamic Function

Temporal Attractiveness Factor

Time Attractive Factor

- Dining
- Shopping₁
- Shopping₂
- entertainment
- Public facilities
- Others
Given *dynamic functions* discussed in previous section, STPA of the six typical POIs are shown:

- **Dining**
- **Entertainment**
- **Shopping_1**
- **Shopping_2**
- **Public Facility**
- **Others**
Activity Identification Using STPA Prisms

Candidate POIs selection

Activity Spots Recognition

- Intersection Points of Trajectory and STPA Prism
- GPS Tracking Points
- Sub Trajectory
- Trajectory

Time

Geographical Space
Duration Extraction

Intersection Points of Trajectory and STPA Prism
GPS Tracking Points
Interpolation Points

Intersection Points of Trajectory and STPA Prism

GPS Tracking Points

Interpolation Points

0
Geographical Space

Time

$P_{t_{i-1}}(x_{i-1}, y_{i-1}, t_{i-1})$

$P_{t_i}(x_i, y_i, t_i)$

$IP_i^1\left(\frac{x_{i-1} + x_i}{2}, \frac{y_{i-1} + y_i}{2}, \frac{t_{i-1} + t_i}{2}\right)$

$IP_i^2\left(\frac{x_{i-1} + x_i}{2}, \frac{y_{i-1} + y_i}{2}, \frac{t_{i-1} + t_i}{2}\right)$

$IP_i^3\left(\frac{x_{i-1} + x_i}{2}, \frac{y_{i-1} + y_i}{2}, \frac{t_{i-1} + t_i}{2}\right)$

$t'$
Experiments and Evaluations

• Data Set
  – one month’s GPS trajectories collected by personal wearable GPS loggers from 10 volunteers;
  – road network and KIWI navigation format coded POIs of Wuhan city;
    • The POI dataset including major restaurants, shopping malls, retail stores, public facilities and entertainment places in Wuhan City;
    • POIs’ area data are collected through internet
  – We use facilities ranking website to classify POIs’ fame $\varepsilon_p$ into 5 levels: top 5, 10th to 5th, 15th to 10th, 20th to 15th, and “others”. Each level corresponds to a numerical value
  – Raw Data is imported using Google Earth
  – 3D presentation and algorithms are implemented using Matlab 7.0
Examples of Activity Spots Identification

- Red: Kara OK Club; Blue: Cyber City for Computer (Computer Retail Shop)
- Time: around 20:00-22:00
Experiments and Evaluations

Examples of Activity Spots Identification 2

• Results when Multiple POI competing
  – two “shopping_1”, one “entertainment”, one “Dining”
  – Around 15:00-17:00

WanDa Cinema is selected as Activity Spot
Experiments and Evaluations

<table>
<thead>
<tr>
<th>Scenario \ Count</th>
<th>Raw Data</th>
<th>Correct POI Identification</th>
<th>Correct Duration Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single POI</td>
<td>24</td>
<td>24 (100%)</td>
<td>23 (95.8%)</td>
</tr>
<tr>
<td>Multi-POIs</td>
<td>21</td>
<td>19 (90.5%)</td>
<td>19 (90.5%)</td>
</tr>
<tr>
<td>POI Cluster</td>
<td>13</td>
<td>9 (69%)</td>
<td>8 (61.5%)</td>
</tr>
<tr>
<td>Single Activity</td>
<td>38</td>
<td>35 (92.1%)</td>
<td>34 (89.5%)</td>
</tr>
<tr>
<td>Multi-Activities</td>
<td>10</td>
<td>7 (70%)</td>
<td>6 (60%)</td>
</tr>
</tbody>
</table>

We compared the detected results to the recall information of volunteers in extensive scenarios:

- Single POI as candidate;
- 2-4 POIs competing to be identified;
- POI detection in CBD areas where lots of POIs locate at;
- only one activity was conducted
- multiple activities happened such as “Dining” after “Entertainment”

An uncertainty of 10 minutes is used to measure duration identification accuracy.
Conclusions and Future Work

• Activity identification from raw GPS trajectories is a recently hot but still ongoing issue.
• In this paper we proposed an automatic activity detection method using POIs’ spatial temporal attractiveness. This proposed method complete judgment depends on information from POIs and trajectories and take arriving time, duration, spatial factor as well as background factor into account.
• Experiment results show that the method provides high accuracy for activity identification.
• Future Work: Using Time Geography Framework for the in-depth Activity Analysis based on personal GPS Tracking Data
Contact the author:

huangliansinc@hotmail.com