CITYMAP

Smart-phone based Spatio-temporal Sensing for Annotated Transit Map Generation

Authors: Rohit Verma, Surjya Ghosh, Niloy Ganguly, Bivas Mitra, Sandip Chakraborty

Indian Institute of Technology Kharagpur, India
ISSUES WITH PUBLIC TRANSPORT
ISSUES WITH PUBLIC TRANSPORT

Google Transit Coverage Across The Globe

- Africa
- Asia
- Oceania
- Europe
- North America
- South America

Transit Coverage
The map shows:
  • The available bus route
The map shows:

- The available bus route
- All the speed breakers, turns, bus stops, termed as Points of Concern (PoCs) on the bus route
The map shows:

- The available bus route
- All the PoCs on the bus route
- Features linked with these PoCs like
  - Steep/Gentle turns or speed breakers
The map shows:

- The available bus route
- All the PoCs on the bus route
- Features linked with these PoCs like
  - Steep/Gentle turns or speed breakers
  - Probability of getting a seat at a bus stop
The map shows:

• The available bus route
• All the PoCs on the bus route
• Features linked with these PoCs like
  • Steep/Gentle turns or speed breakers
  • Probability of getting a seat at a bus stop
• Other route features like jerky road
The map shows:

- The available bus route
- All the PoCs on the bus route
- Features linked with these PoCs like
  - Steep/Gentle turns or speed breakers
  - Probability of getting a seat at a bus stop
- Other route features like **jerky road, congested patches**
A system which generates a complete annotated transit map of a city.

Annotations on the bus route:
- PoCs like speed breaker, turns, bus stops.
- Features linked to PoCs like
  - Type of turn (sharp/gentle)
  - Type of speed breaker (steep/gentle)
  - Probability of getting a seat at a bus stop
- Jerky bus route segment
- Congestion level in a bus route segment

Smartphone based crowdsourcing application
System Architecture

CityMap Smart-phone App

- Feature Extraction
- Preprocessing

Data Collection
(Inertial Sensor Data, Periodic GPS Sound Data)

Database Generation

Database

CityMap Server

- Trajectory Generation
- Route Classification
- Map Stitching

Bus data

Bus Route With Annotation

Map Annotation
CHALLENGES: DATABASE GENERATION

How to make the application energy efficient?

How to develop a zero intervention and intelligent data logging application?
The application should **consume low energy** being a continuous data collection app.
SMART DATA LOGGING

Start Logging

Continue

Stop Logging
SMART DATA LOGGING

Identifying user in bus

Signature of the acceleration along y-axis for different modes

Variation in sound data for different motorized vehicles

Logging at bus stop

Vertical acceleration values to identify the events when the user is traveling by a bus
**Build Database**

- **PoC information** in the database.

<table>
<thead>
<tr>
<th>PoC</th>
<th>GPS Coord</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus Stop 1</td>
<td>C1</td>
</tr>
<tr>
<td>Turn 1</td>
<td>C2</td>
</tr>
<tr>
<td>Speed Breaker 1</td>
<td>C3</td>
</tr>
<tr>
<td>Bus Stop 2</td>
<td>C4</td>
</tr>
<tr>
<td>Bus Stop 3</td>
<td>C5</td>
</tr>
</tbody>
</table>

- Data from inertial sensors for every sample in between each PoC.
How to generate user trajectory from the information obtained from database?
CHALLENGES: MAP ANNOTATION

- Bootstrapping initial bus route data.
- Data collected from a commuter is never tagged by him, hence we need to decide which bus route did he actually travel on.
- We also need to shortlist one bus route when two bus routes have overlapping segments and the commuter travels on this part.
Not all commuters travel from source bus station to destination bus station. Then how to generate the complete bus route from multiple commuter data?
PoC TO TRAJECTORY

- Place detected **PoCs as anchor points** on the trajectory

- **Estimate intermediate GPS Coordinates** to generate complete trajectory of the user
  - Use **Vincenty’s Formula** to estimate a point P’
  - Bring the point closer to the desired road, at P’’, using **Coordinate geometry approach**
  - Use **Snap-to-Road API** to drag point to road
Each instance of the route data would have a set of features linked to it.

We try to cluster the data into clusters based on correlation between these features.

The clusters which have a **high confidence clustering** are used and the rest discarded.

The **cluster head** is taken as the representative for the route.

A minimum of **20 instances** is required for a route to be considered as a cluster head.
Variation of speed after DWT for different data

• Every bus route can be classified using a set of features like speed of vehicle, waiting time at bus stop, jerkiness of the bus

• We use **Discrete Wavelet Transform (DWT)** to cluster trajectories belonging to same bus route
Users in different buses on same bus route may not cover the complete route. The relevant trajectories need to be stitched together to generate complete route.

Caveat: There may be overlapping routes.

Solution: Stitch only if the correlation coefficient between route and trajectory features is high.
Experiments were carried on by **50 volunteers** on **11 routes** in **3 cities** for a period of **around 3 years**.

The volunteers travelled at different bus routes, with the application installed in their mobiles.

They tagged the ground truth data like, location of PoCs, congested patches, bus route travelling on, broken roads, etc, along with continuous GPS data for the first 3 months.

### Volunteer Details

<table>
<thead>
<tr>
<th>No. of Volunteers</th>
<th>30 (K), 10 (B), 10 (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age Group</td>
<td>18-25 yrs</td>
</tr>
<tr>
<td>Incentive</td>
<td>1500 INR/month</td>
</tr>
</tbody>
</table>
## EVALUATION: MAP GENERATION AND ROUTE DIFFERENTIATION

<table>
<thead>
<tr>
<th>Route Name</th>
<th>Route Length</th>
<th>Daily Avg Travel Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>K1</td>
<td>17 km</td>
<td>3.12 hrs</td>
</tr>
<tr>
<td>K2</td>
<td>14</td>
<td>2.76</td>
</tr>
<tr>
<td>K3</td>
<td>20</td>
<td>4.32</td>
</tr>
<tr>
<td>K4</td>
<td>10</td>
<td>0.48</td>
</tr>
<tr>
<td>B</td>
<td>19</td>
<td>1.2</td>
</tr>
<tr>
<td>D</td>
<td>22</td>
<td>3.84</td>
</tr>
</tbody>
</table>

### Map Generation Error on Different Routes

![Map Generation Error Graph](chart.png)

### Accuracy of Bus Route Differentiation

#### City | Route Number
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BBS</td>
<td>306</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P 0.93</td>
</tr>
<tr>
<td></td>
<td>R 0.88</td>
</tr>
<tr>
<td></td>
<td>A 0.83</td>
</tr>
<tr>
<td></td>
<td>P 0.92</td>
</tr>
<tr>
<td></td>
<td>R 0.92</td>
</tr>
<tr>
<td></td>
<td>A 0.85</td>
</tr>
<tr>
<td>KOL</td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>S9</td>
</tr>
<tr>
<td></td>
<td>P 1.0</td>
</tr>
<tr>
<td></td>
<td>R 0.9</td>
</tr>
<tr>
<td></td>
<td>A 0.9</td>
</tr>
<tr>
<td></td>
<td>P 0.96</td>
</tr>
<tr>
<td></td>
<td>R 0.96</td>
</tr>
<tr>
<td></td>
<td>A 0.93</td>
</tr>
<tr>
<td>DGP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1</td>
</tr>
<tr>
<td></td>
<td>P 0.92</td>
</tr>
<tr>
<td></td>
<td>R 0.92</td>
</tr>
<tr>
<td></td>
<td>A 0.86</td>
</tr>
<tr>
<td></td>
<td>P 0.93</td>
</tr>
<tr>
<td></td>
<td>R 0.93</td>
</tr>
<tr>
<td></td>
<td>A 0.88</td>
</tr>
</tbody>
</table>
EVALUATION: OVERALL SYSTEM

\[ RMF = \frac{\text{Erroneously matched route length}}{\text{total route length}} \]

- Low RMF values in K2 and K3 is because of high landmark density compared to others.
- \( p_{\text{skip}} \) impacts value of RMF the least.
- \( p_{\text{sit}} \) impacts value of RMF the most.
EVALUATION: COMPETING SYSTEMS

- Competing system only uses GPS information and hence fails in classifying overlapping segments
- With minimal GPS usage, CityMap consumes much less energy
Increasing the data collection period improves the results for both annotations and map generation.

Continuous data collection for a period of **20 days** decreases the errors many folds after which the algorithm stabilizes.
Source: M.G. Road
Destination: City Centre
Number of Speed Breakers = 2
Number of Bus Stops = 12
Number of Bad Road Patches = 0
Available Buses:
A-Zone:
Type of Bus : Mini
Avg Speed : G
Speed Before Landmarks : Y
Bus Stop skipping probability : Y
Sitting probability : Y
Jerkiness : Y
CONCLUSION

- **CityMap** is the first of its kind implementation for *city transit map generation using smartphone sensors*.

- **CityMap** effectively annotates bus routes along with other route and road features with more than *90% of accuracy*.

- The generated map can be utilized as an infrastructure for a large set of public transit based applications like
  - Navigation application
  - Route Recommendation application
  - Dynamic Rerouting application
Follow the work of Complex Network Research Group (CNeRG), IIT KGP at:
Web: http://www.cnerg.org
Facebook: https://web.facebook.com/iitkgpcnerg
Twitter: https://www.twitter.com/cnerg
Data: https://doi.org/10.17632/92yrxty5gn.1