June 29, 2016

Big Data Analytics for Transportation Planning

2016 ITS Biennial Conference, Taipei, Taiwan

Ann Skudlark, AT&T Labs Research
Jean-Francois Paiement, Colin Goodall
Alexei Pozdnukhov, Mogeng Yin, Sid Feygin, UC Berkeley
### Smart Traffic Design is Critical, Congestion is getting worse

#### 2015 Urban Mobility Scorecard

<table>
<thead>
<tr>
<th>Year</th>
<th>1982</th>
<th>2000</th>
<th>2010</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay</td>
<td>18</td>
<td>37</td>
<td>40</td>
<td>42</td>
<td><strong>42</strong></td>
</tr>
<tr>
<td>- Hours in SF Bay Area</td>
<td>-78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>$400</td>
<td>$810</td>
<td>$930</td>
<td>$950</td>
<td><strong>$960</strong></td>
</tr>
<tr>
<td>- $1,675 in SF Bay Area</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>$42B</td>
<td>$114B</td>
<td>$149B</td>
<td>$156B</td>
<td><strong>$160B</strong></td>
</tr>
</tbody>
</table>

Texas A&M Transportation Institute
SmartBay Project – an extensive simulation of Bay Area traffic patterns from transportation and mobility data

Anonymized and aggregated mobile cell data provides a reliable and efficient source for understanding transportation flows at scale.

Project objectives:
- Improve planning, design and operations of transportation systems
- Decrease travel times, reduce congestion and environmental impact
- Better manage traffic flow in case of accidents and road closures
**SmartBay region**

9 counties of SF Bay Area  
7.5M population  
3.4M commuters  
350K public transport  
75% solo occupancy  
28 minutes average commute  

**SmartBay simulation model** generated from multiple data sources

Road network  
Public transit  
Land use  
Anonymized cell data  

Focus here is on a typical working day
Road network

Number of lanes, speed limits, and vehicular capacity is obtained for each of the roads shown; additionally, detectors provide actual traffic counts.

Public transit

9 major bus agencies
light & commuter rail
BART
schedules and capacities
SmartBay: Activity-based Demand Modeling

Travel is a sequence of tours/trips between spatially distant locations, with inter-related activities, within time constraints.

Modeling approach: HSMM

Hidden semi-Markov Model, state is tours & trips

Tours and trips

Schematic shows primary tour commuting between home and work with an intermediate stop, and secondary tours based on home and work.

The primary and secondary tours contain 7 trips, taken in the order shown at distinct times.
Inference, building the demand model

- Extract aggregate patterns of mobility from anonymized cell phone data
- 1M+ plus commuters in SF Bay Area
- Approximation of home, work, and other important locations
- Input Output hidden Markov model (IOHMM):
  - Hidden activity, including its time and duration, is influenced by date and time, and the previous activity, hence we include time dependent covariates in the estimation of the output; and location/land use: e.g., residential, commercial (businesses, restaurants) area etc.

No activity can be attributed to any specific phone user
**SmartBay:- Human Activity Inference (cont’d)**

**Demand generation**
- By TAZ traffic analysis zone (1454 in SF Bay Area)
- Project aggregates from cell tower coverage areas to TAZ areas
- Rescaling, population census figures / cell phone user counts
  - Adjustments resulted in up-to-date origin-destination flows which reflected new urban developments, as well as employment re-distribution due to evolution of IT sector in Silicon Valley.

**Traffic micro-simulation for SmartBay**
- Virtual agents, virtual commuter with a sequence of activities
- Probabilistically sampled from the generative demand models

**Instance of simulation**
- MATSim software on AT&T Research computing infrastructure
- Up to 30% of the commuting population, 1.1M simulated agents
- Driving and public transport as modes in the model
- Compare link traffic volumes to CA DOT inductive loop detector counts
  - The model was found to meet the Federal Highway Authorities accuracy specifications.

*CDRs are not used in the micro-simulation. It is impossible to recover any specific AT&T customer in the output of the simulation.*
SmartBay: Virtual Agents

Comparison of simulation (blue) and loop detector (orange) volumes at two locations. Volumes are 0 – 4,250  and 0 – 6,500  vehicles/hour over 24 hours.
Conclusions - AT&T UC Berkeley SmartBay

• Provides a rich scientific experimentation platform
• SmartBay allows evaluating “what-if” scenarios for
  • Road network improvements
  • Public transport developments
  • Major future construction initiatives
• First use case provided commute estimates for the City of San Francisco Treasure Island planning initiative.
• Selected by the City of San Francisco – one of the DoT Smart City Challenge finalists as a key research partner to provide platform and analysis for improved daily commute planning.
  • Investigating combining morning commute data with historical data to better predict afternoon commute.
Acknowledgements

This work was funded in part by a research gift from AT&T to UC Berkeley and by the California Department of Transportation