Open and Big -
Data Frontiers in Regional Planning

Christopher M. Puchalsky, Ph.D.
Deputy Dir., Transportation Planning
Outline

- Stage setting – the need for a network
- Open data feeding the network
- Exporting back to GTFS
- Some applications – open collaboration with non traditional partners on open data

For more details:
Delaware Valley Regional Planning Commission

- Metropolitan Planning Organization (MPO)
- 2 States
- 9 Counties
- 351 Municipalities
- 5.6 Million Population
- 3,800 sq. miles
- ~115 employees

Activities –
- Long Range Plan (LRP)
- Transportation Improvement Program (TIP)
- Wide range of planning and technical support for regional partners
Need for a Network

Travel forecasting models:

- Air Quality Conformity Analysis
- Project engineering, design, and prioritization
Travel Demand Model Networks

• The need:
  – Accurate representations of regional highway and transit networks

• The past:
  – “hand” code from paper maps, schedules, etc.
  – or, combine a multitude of different data sources

• The innovation:
  – Fuse OpenStreetMap (OSM) and GTF (i.e. “Google-transit”), then and add extra data for modeling
The OpenStreetMap

- www.osm.org
  - Started in 2004
  - Organization: OSM foundation
    - non-profit, based in the U.K.

- Volunteers
  - They generate the map
    - Upload data from their private GPS devices
    - Edit directly on www.osm.org

- Data distribution
  - Free of charge
  - Can be used for any commercial or non-commercial purpose

- Data content
  - Routable street network plus other geography
  - U.S. data derived from an import of the 2005 TIGER file
GTFS Overview

• TriMet/Google developed specification
• Widely adopted standard for public transit
• Series of text files with comma-delimited values
• (GTFS = General Transit Feed Specification)
Open Data Mash-up for Modeling

- Data integration
  - Data objects of different origin are merged
  - New relationships are created

<table>
<thead>
<tr>
<th>Node</th>
<th>Stop Area</th>
<th>Link</th>
<th>Stop Point</th>
<th>Service Pattern</th>
<th>Scheduled Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>Number</td>
<td>From Node To Node</td>
<td>Number</td>
<td>Line Name</td>
<td>Route Name</td>
</tr>
</tbody>
</table>

Legend:

- 1 or more
- 0 or more
- Exactly 1

from OSM

from GTFS

Travel Demand Data
Integrated Street & Transit Network

© in part by OSM and CC-by-SA
TIM 2.0 Highway Network

New, accurate topology (& routable)  Legacy DVRPC network model
Giving Back – Exporting to an Open Format

(an example with transit networks; new open standards on matrices)
Reasons for Exporting

- Combining multiple transit feeds
- Validation/Debugging transit network
- Use in FTA New Starts STOPS
- Portability of Transit Network
VISUM Imported Network
VISUM Exported Network (WKTPoly shape)
Now for something completely different…

- A few applications of the open data derived network
Sketch-level Transit Supply/Demand Analysis

Uses two regional datasets:

- **DVRPC Transit Score**
  - Screening tool that uses population density, job density, and the density of 0-car households to assess likely transit demand and transit modal appropriateness

- **TAZ-level Transit Supply**
  - 24-hour transit vehicle departures by TAZ (From GTFS data)

Grouping both datasets into 5 score bins and subtracting one from the other permits a sketch-level regional assessment of transit demanded vs. transit available.
TRANSPORT SERVICE PROVIDED
vs.
TRANSPORT SERVICE DEMANDED
Demand > Supply
Supply > Demand

NOTE:
This is a sketch planning calculation based on two sketch planning tools:
DVRPC Transit Scores (2010), which estimate transit demand based on
population, job, and 0-car HH densities,
and DVRPC TAZ-level 24-hour transit
service volumes (rail + bus) from GTFS
data and the TIM 2.0 travel demand
model.

DVRPC
DELWARE VALLEY
REGIONAL
PLANNING COMMISSION

NOV. 2013
CyclePhilly – User Generated GPS Data

Start Trip!

Weather Alert:
Partly cloudy throughout the day.

21 saved trips:
- Commute 8/15/14 6:07 PM 4.3 miles, 19 minutes.
- Commute 8/15/14 8:06 AM 4.3 miles, 21 minutes.
- Commute 7/28/14 5:53 PM 4.3 miles, 28 minutes.
- Commute 7/23/14 5:35 PM

Hourly Count by Purpose:
- Commute
- Errand
- Exercise
- Other
- School
- Shopping
- Social
- Work-related

Purpose Legend:
Open network to “snap” open data

- OSM + Bicycle GPS Data
  - Base: OSM Street Network (2010?)
  - Input: GPS Traces from CyclePhilly application
  - Output: ‘Snapped’ GPS traces to street network

- Increases utility of GPS data
- Facilitates analysis of bicycle data
- ++Efficiency by computerized processing
- 5+GB of GPS data
Uses of “Snapped” Data

- **Macro-level**
  - Demographic data of bicyclists
  - Origin-Destination
  - Correction Factors (seasonal, weather)
- **Micro-level**
  - Bicycle speeds per road segment and facility type
  - Bicycle model calibration (Impedance formula, route choice)
Thanks to:

- Will Tsay
- Greg Krykewycz