Using Innovative Data in Transportation Planning and Modeling

presented at
2014 Ground Transportation Technology Symposium: Big Data and Innovative Solutions for Safe, Efficient, and Sustainable Mobility

presented by
Cambridge Systematics, Inc.
Nikhil Puri

November 19, 2014
Outline

- Context – need for good data
- Applications in transportation planning and modeling
- Benefits and limitations of using (big) data
- Potential applications
Context

**Need for Good Data**

- Annual cost of traffic congestion in the U.S. – 100s of billions ($124 billion recent study by INRIX and the Centre for Economics and Business Research).
- Key infrastructure decisions based on travel demand models
  - Rely on (good) data
- Traditional data collection – intercept origin-destination, tube counters, household travel surveys
  - Still needed
- Innovative data collection sources – anonymized cell phones, GPS probes, aircraft, Bluetooth devices and toll plaza, General Transit Feed Specification (GTFS)
  - Becoming more common
  - Not without limitations
Applications in Transportation Planning and Modeling

- **Planning**
  - Understanding O-D patterns, trip lengths, imputed trip purposes, external trip patterns
  - Performance measurement – speeds
  - Travel behavior (infancy)

- **Model Applications**
  - Macroscopic and microscopic models
  - Calibration/validation of models – trip lengths, trip distribution patterns, trip purposes, speeds, counts
Central Florida Region – Spatial Distribution of Daily Trips

Note: Partial counties shown in red.

Central Florida Region Spatial Distribution (Destination)

Percentage Trips

- At Home: 26.7%
- Long Term Visitor: 18.5%
- Leaving the CFRPM Region: 1.3%
- Within CFRPM Region: 50.9%
- Through Trips: 2.6%

Trip Type

Note: Derived from AirSage Data
Over 40 percent of trips are Home Based Other Trips.

There is a large variation in Home Based Work trips between AM and PM periods.

Approximately 22 percent of trips occur during the AM peak, whereas 26 percent occur during the PM peak.

Note: Derived from AirSage Data.
Interesting findings…

- Green boundary represents Disney World area
- Summation of daily non-home-based work trips 189K trips. In comparison, observed 2013 counts was in the order of 207K annual visitors
Select Link Analysis Using Cell Phone Data

- Used anonymized cell phone data (2012)
- Select link analysis on bridges (O-Ds on bridges)
- Findings
  - Due to proximity of bridges, hard to distinguish which bridge traffic was on
  - O-D patterns compared reasonably well at county level

Source: Google Maps

Comparison of County Level Distribution Patterns - Household Travel Survey vs Cell Phone Data
Probe Speed Data

Data tells us…
- Bottleneck location and severity
- Duration of congestion
- Extents of queues
- Free-flow conditions
- Issues: Overlapping TMC links

Some Applications:
- Performance Measures
- Macro and micro transportation model validation
- Toll diversion modeling

Heat map indicating congestion locations and time

Example of overlapping TMC links
MTA (New York) BusTime® Data

- Approximately 6,000 buses throughout NYC
- 8 Million daily data points
- GPS bus data every 30 seconds
- Currently comparing data against GTFS for reliability statistics

Note: Not based on real data
Benefits and Limitations of Using (Big) Data

Benefits

» Often larger sample sizes
» Easily collected; archived data available
» Often cheaper with wider coverage

Limitations

» Spatial accuracy (cell phone data)
» Currently unable to gain additional insights on trip behavior characteristics such as mode
» Accuracy may be susceptible to weather
» Privacy concerns
Potential Applications

- Using trajectory data – route choice behavior
- Transit ridership
- Performance measurement for transit operations
- Travel behavior during events
- Emergency response
- Airport usage
- Tourism
Contact Information

NIKHIL PURI
Cambridge Systematics
npuri@camsys.com
(646) 364-5491