Big Data for Transportation by StreetLight Data

Curt Thye
Director South Central Region
Why use Big Data for Transportation?
Conventional Data Collection and Models No Longer Meet the Needs of Transportation Professionals

- **Expensive**
- **Time-Consuming**
- **Conducted Rarely**
- **Small Sample Sizes**
- **Incomplete Information**
- **Cumbersome Data Integration**
Transportation Behavior Today Is Changing Rapidly

To Keep Up with Fast-Changing Travel Behavior, We Need Data That:

1. Describes Current Behavior
2. Measures Change Over Time
3. Is Diagnostic & Predictive
Introduction to StreetLight Data
We Offer The Best Big Data and Analytics Platform

Together

StreetLight InSight®:
The Only On-Demand Platform For Running
Actionable Transportation Analytics
StreetLight InSight® turns big data into actionable transportation analytics on demand.

**Metrics:**
- Origin-Destination
- Travel distance, time, circuity
- Traveler trip purpose & demographics
- 2016 AADT
- Commercial Trips & More

**Services:**
- Accessible and Confidential
- Congestion Analyses
- StreetLight InSight®
- Mass Data + Contextual Data
- Demand Management
- Travel Demand Models
- Transit Project Evaluation and Prioritization
- Air Quality/ GHG Estimation
- Internal/External Studies
- Freight Modeling
- Accessibility Studies
- Public Engagement
- Travel Demand Models
- Transit Project Evaluation and Prioritization
- Air Quality/ GHG Estimation
- Internal/External Studies
- Freight Modeling
We Deliver Additional Value Above and Beyond Our StreetLight InSight Platform

1. Data Science Expertise
   We Bring 100+ Years of Collective Experience in Data Science and Engineering to Transportation

2. Evaluations of Big Data
   We Select the Best Range of Real-World Data Sources from a Fast-Changing, Emerging Technology Landscape

3. Privacy Commitment
   We are Committed to Privacy and We Follow Privacy-by-Design Principles that Protect Consumers and You
We Offer the Best Combination of Data Resources for Understanding Travel Behavior

**Locational Big Data**
- **Navigation-GPS Data**
  - Segments commercial trucks
  - 23B+ data points/month
- **Location-Based Services Data**
  - Larger sample size
  - 60B+ data points/month

**Contextual Data**
- Road Network Maps
- Parcel Data
- American Community Surveys
Big Data Applications for Transportation and Transit
Big Data is Just A Buzzword…
Different Types of “Big” Geospatial Data Offer Different Benefits for Modelers

<table>
<thead>
<tr>
<th></th>
<th>Cellular</th>
<th>Navigation-Based GPS</th>
<th>Location-Based Services</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial Precision</strong></td>
<td>200-1000 meters</td>
<td>5 meters</td>
<td>5 meters – 50 meters</td>
</tr>
<tr>
<td><strong>Frequency of Data Pings</strong></td>
<td>Irregularly; every 15 min – hours</td>
<td>Regularly; every 1 sec – 1 min</td>
<td>Variable; sometimes triggered by location change</td>
</tr>
<tr>
<td><strong>Type of Trip</strong></td>
<td>Blends personal and commercial trips</td>
<td>Differentiates personal and commercial trips</td>
<td>Blends personal and commercial trips (for now)</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td>Varies by telco: ~10% of population for small telcos and ~25-30% for large telcos</td>
<td>Varies by region; ~1-4% personal trips; ~10-12% of commercial trips (for INRIX)</td>
<td>Medium – 62M US and Canadian devices in our database (~23% of US adult population)</td>
</tr>
</tbody>
</table>
Location Data from Mobile Devices Can Show When, Where, and How People Move

Key Benefits

• Accurate
• Precise
• Comprehensive
• Anonymous

Note: This image shows a filtered subset of data to improve visibility. The data is from September 2016 in Fremont, CA.

Navigation-GPS Location
Circle enlarged for visibility

LBS Data Location
Circle radius vary; they accurately reflect the spatial precision of each unique data point
Navigation-GPS Data Is Created by Devices That Do Turn-by-Turn Navigation

<table>
<thead>
<tr>
<th>Technical Characteristics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Precision</td>
<td>~5 meters</td>
</tr>
<tr>
<td>Frequency of Data Pings</td>
<td>Regularly; every 1 sec – 1 min</td>
</tr>
<tr>
<td>Type of Trip</td>
<td>Differentiates personal and commercial trips</td>
</tr>
<tr>
<td>Sample Size</td>
<td>Penetration rate varies by region – but much smaller than LBS. ~1% - 4% for personal, 12% trucks.</td>
</tr>
</tbody>
</table>
## Location-Based Services Data Is Created By Mobile Applications (Think “Apps”)

<table>
<thead>
<tr>
<th>Technical Characteristics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial Precision</strong></td>
<td>~5 meters – 25 meters</td>
</tr>
<tr>
<td><strong>Frequency of Data Pings</strong></td>
<td>Variable; usually triggered by location change</td>
</tr>
<tr>
<td><strong>Type of Trip</strong></td>
<td>Personal</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td>~23% of US adult population (62M+ US devices in our database)</td>
</tr>
</tbody>
</table>

**LBS Data Creation**

- Mobile applications (Think “Apps”)
- Spatial precision: ~5 meters – 25 meters
- Frequency of data pings: Variable; usually triggered by location change
- Type of trip: Personal
- Sample size: ~23% of US adult population (62M+ US devices in our database)
Not all Data is created equal... Sample size, Spatial and Temporal Precision are all important

<table>
<thead>
<tr>
<th>Key Characteristics of Our Locational Big Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multiple Types of Data</strong></td>
</tr>
<tr>
<td>• Location-Based Services records</td>
</tr>
<tr>
<td>• Navigation-GPS records</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
</tr>
<tr>
<td>• Covers ~23% of adult population in US and Canada</td>
</tr>
<tr>
<td>• Unbiased sample backed up with automated normalization</td>
</tr>
<tr>
<td><strong>Spatial Precision and Coverage</strong></td>
</tr>
<tr>
<td>• As precise as 5-25 meters, average better than 18 meters</td>
</tr>
<tr>
<td>• 4-carrier coverage – no rural gaps</td>
</tr>
<tr>
<td><strong>Temporal Precision</strong></td>
</tr>
<tr>
<td>• One-hour intervals</td>
</tr>
<tr>
<td>• Weekends vs. weekdays</td>
</tr>
<tr>
<td><strong>Archival Data</strong></td>
</tr>
<tr>
<td>• Monthly data periods from 2014 through “month before last”</td>
</tr>
<tr>
<td><strong>Privacy Protection</strong></td>
</tr>
<tr>
<td>• All data is de-identified by our suppliers</td>
</tr>
<tr>
<td>• No personally identifying information</td>
</tr>
<tr>
<td>• Metrics are aggregated into groups</td>
</tr>
</tbody>
</table>

This image shows a location record’s potential location at different levels of spatial precision. At 300m to 1000m spatial precision, records cannot provide corridor- or intersection-level insights.
Challenge; Big Data is often Messy Data
StreetLight Converts into Useful Transportation Metrics

Input: Big Data
- Anonymous and accurate Big Locational Data
- Road network, land use, parcel, census and more Contextual Data

Processing: RouteScience®
- Clean
- Patternize
- Contextualize
- Aggregate

Output: StreetLight InSight Metrics
When Does it Make Sense to Use Big Data?

1. Going “Back in Time”
2. Large Study Areas
3. Need Large Sample
4. Complex Project Set-Up
5. Need Complete Trips
6. For Truck Trips & Demographics
What Questions Can Big Data Analytics Answer for Planners? A Few Examples....

- What types of trips cause congestion on a particular roadway?
- What are the origins and destinations of travelers on a particular roadway?
- How do travel patterns vary during different types and times of day?
- Where do commuters live, and where do residents work?
- What are the demographic characteristics of travelers? How long are their trips?
StreetLight InSight Supports A Range of High Priority Transportation Planning and Modeling Initiatives

**Travel Demand Modeling**
Calibrate with Empirical, Comprehensive O-D Matrices

- Summer 2016
- Winter 2016

Colorado DOT used StreetLight InSight To Understand Seasonal & Weekday/Weekend Trends

**Long-Term Planning**
Study Regional Patterns & Engage With the Public

- Fehr & Peers and Napa Valley Transport. Authority Used StreetLight InSight to Study Regional Trends

**Travel Demand Management**
Scan for High-Potential Project Opportunities

- Virginia DOT, Michael Baker, and SSTI Scanned for “Displaceable Vehicle Trips” with StreetLight InSight

**Performance Measurement**
Evaluate AADT, Travel Time Reliability, & More

- Siemens Used StreetLight InSight to Study the Impact of its ITS Traffic Signals on Travel Time Reliability

**Congestion Studies**
Identify the Cause of Congestion

- The City of Lafayette, CA and Arup used StreetLight InSight to Analyze Downtown Congestion

**Project Evaluations**
Easily Conduct “Before & After” Studies

- Fehr & Peers and SANDAG used StreetLight InSight to Determine the Impact of a Toll on Behavior
Some Attendees Drive Despite Transit Options

O-D to Preset Geography (TAZs) with GPS Data – Project Results:

- Some vehicle trips originate where BART operates in San Francisco, Oakland
- Good opportunity for cities and transit agencies to focus game-day marketing efforts.
Approximately 75% of Game Attendees Live 10+ Miles from Stadium

Visitor Home-Work Analysis with LBS Data – Project Results:

• Most attendees **live or work** more than 10 miles away from the stadium.

• Thus, you can expect attendees to drive or take BART (regional transit system) as opposed to walk or bike to the stadium.
Game Attendees Are Geographically and Economically Diverse

Visitor Home-Work Analysis with LBS Data – Project Results:

• Attendees not only come from all over the Bay Area, they generally have a diverse economic background.

• Would this change if/when the stadium moves to San Francisco?
Big Data Helps Reveal the Best Opportunities for Converting Vehicle Trips to Other Modes

Scan Large Areas for Transit Opportunities

Drill Down on Refined Areas to Understand Transit Gaps
A Corridor Impact Study for a New Route

**Corridor Impact in Canada**

**Need:** Route choice between 2 highways to measure the impact of offering a new option.

**Zones:** 6 routes, 3 origin zones, 6 destinations zones.

**Metric:** O-D with Middle Filter

**Months:** May, June 2016

**Time to Run:** 2 min for each month
Prioritize Bike/Ped, Transit or Improvement projects
Big Data & Privacy
StreetLight Data Privacy Principles

Our Commitment to Privacy
• Privacy is one of the core principles at StreetLight Data.
• Our analytics only describe the movement of groups of people - not the movement of individuals.
• Privacy Officer

What We Expect from Our Data Suppliers
• Data files are not to contain personally identifiable information.

How We Protect Our Location Data Resources
• Once in house, the data sources are algorithmically processed into Metrics in a secure data repository behind a multilayered network security architecture supported by system audits and controls.
• Metrics about aggregated composite groups... “half of the group of people who shop in this shopping district live 10 miles away or more.”
• Every analysis in the StreetLight InSight® platform goes through an automated coverage and privacy check.
StreetLight Data Participates in Creation of Best Practices & Policies

Future of Privacy Forum (FPF) think tank [www.fpf.org](http://www.fpf.org)

“brings together industry, academics, consumer advocates, and other thought leaders to explore challenges posed by the technological innovation and develop privacy protections, ethical norms, and workable business practices.”

Working Groups and Issues of Interest:
• Smart Communities Working Group
• Mobile Location Analytics
• Internet of Things
• Big Data
• Connected Cars
• Government and Law Enforcement

[https://fpf.org/2017/03/30/smart-cities/](https://fpf.org/2017/03/30/smart-cities/)
We’re Here to Help!

**Contact Me**
Curt Thye
Director South Central Region
303-868-0703
Curt.Thye@streetlightdata.com

**Contact Us**
support@streetlightdata.com

**Check Our Blog**
blog.streetlightdata.com
Thank you!
Now You’re Ready to Leverage Big Data

Contact Us to Get Started:
Curt.Thye@streetlightdata.com