

Hazard Identification using Bluetooth Sensors in Austin, TX





The Data:

- “Hack the Traffic” retrieved from data.austintexas.gov
- Data is collected by 193 Bluetooth detectors placed around the city
- Logging device’s MAC addresses allows an approximation of travel speeds on stretches of road.
- City also releases a dataset of reported traffic incidents.
- Utilized datasets:
 - Individual Address Files
 - Individual Traffic Match Files
 - Traffic Match Summary Records
 - Real-Time Traffic Incident Reports
 - Travel Sensors



The Problem:

- Traffic studies in Germany in the mid-90s identified 3 “phases” of traffic flow: free, synchronized, and jammed.
- The periods of greatest hazard in traffic are at the transitions between the phases, and most particularly the synchronized to jammed transition.
- Alert generation by the Bluetooth Sensor cloud could allow the city to better distribute its traffic incident response resources.
- Can we identify roads which are in dangerous transition periods by observing the data being generated by the sensors?
- **The Problem with the Problem:** we can’t assume that a constant density of vehicles are carrying bluetooth devices and so we can’t measure traffic density directly.



Proposed Study:

- Compare the Individual Traffic Match Files to the Incident Reports to determine which measurements were taken shortly before a traffic incident.
- Create an aggregate data set where points represent (record ID, location of sensor, current speed, deviation from average traffic speed, standard deviation of traffic speeds on that road, current count density, classification as “close” or “not close” to traffic incident).
- Train a classifier to recognize “close” points.
- Test that classifier.



Proposed Framework:

- Hardware: the LEAP CPU cluster.
- Software: Apache Spark