Assessing High-resolution Connected Vehicle Data for TSMO Applications

Joyoung Lee, Ph.D

Associate Professr

Department of Civil & Environmental Engineering

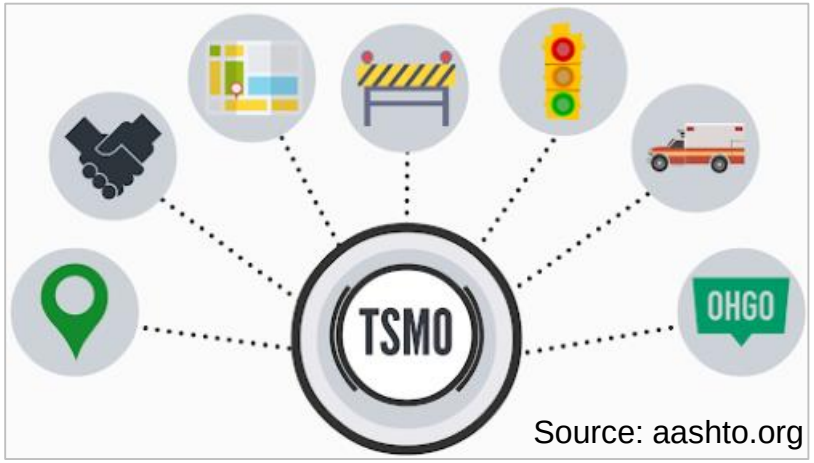
ITS Resources Center

New Jersey Institute of Technology

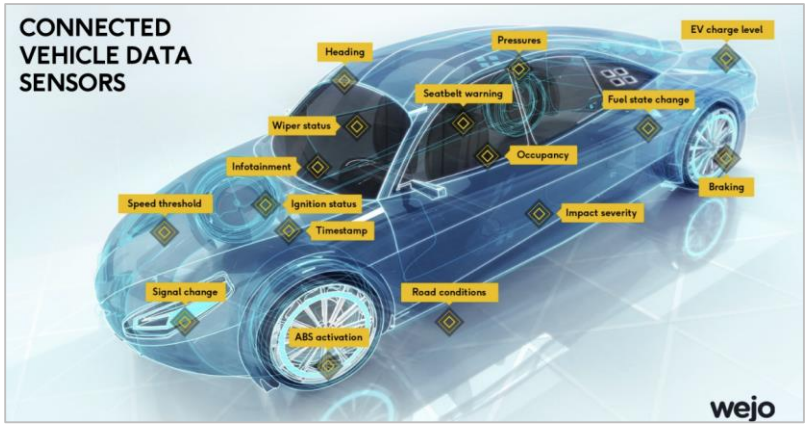
NJIT Team: Branislav Dimitrijevic, Ph.D., Liuhui Zhao, Ph.D. ,
Dejan Besenski, Ph.D., Joyoung Lee, Ph.D.

Background

- Crowd-sourced connected car data is becoming prolific, owing to advanced vehicle telematics integrated in passenger vehicles
  - ubiquitous collection of high-precision individual vehicle waypoints
  - work without ad hoc infrastructure (e.g., loop detectors, CCTV cameras)
  - much more scalable
- TSMO applications could be implemented more cost-effectively by leveraging commercially available connected car data



Source: aashto.org



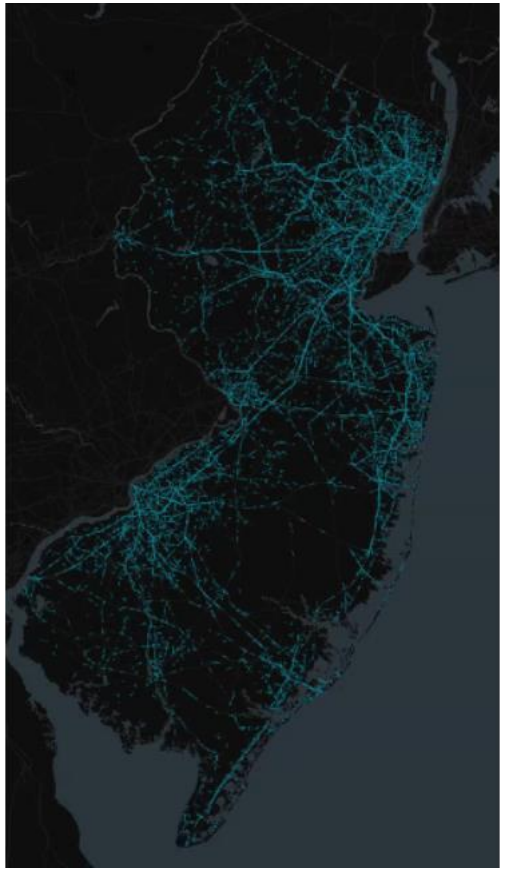
Data Overview (coverage & resolution)

- Cover most NJ highways
- Temporal resolution
  - Movement: consistent 3-second ping
  - Event: only generated when a qualified event occurs (e.g., hard braking)
- Spatial resolution (precision)
  - 6 decimal points of latitude and longitude
- Data volume: Movement data 25x more records than Event data (constant ping vs. event-based)

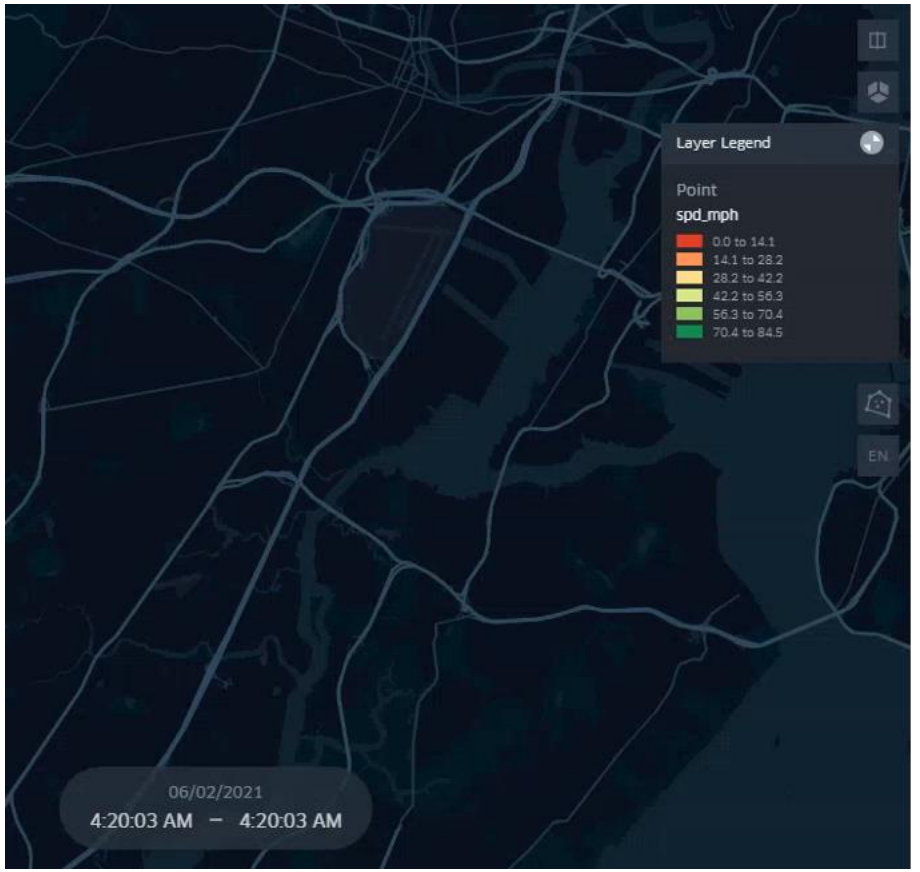
Example-Movement Data

Example-Event Data

Vehicle Trajectory Reconstruction

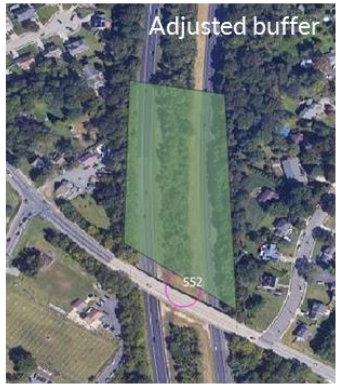


Animation: Statewide (AM peak)

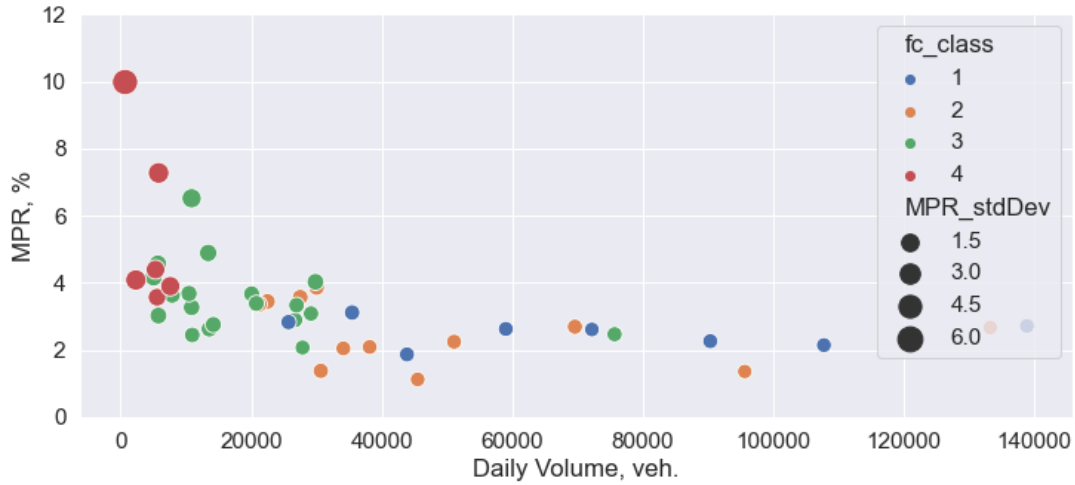


Animation: Individual journey

Data Quality (market penetration)

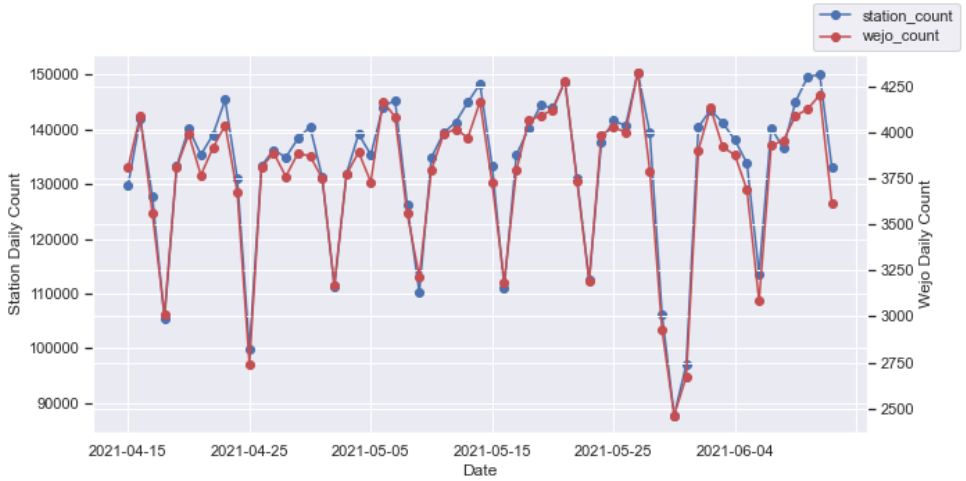


Weight-in-motion stations & geo buffers

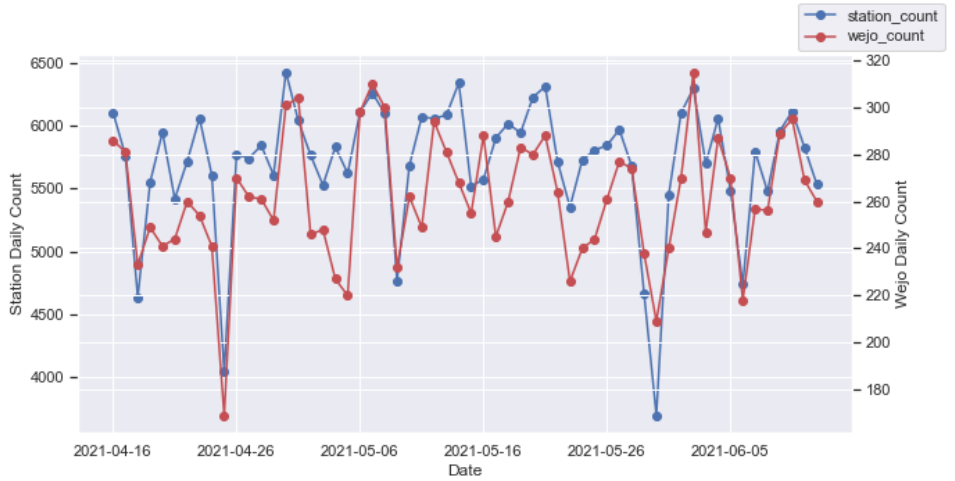


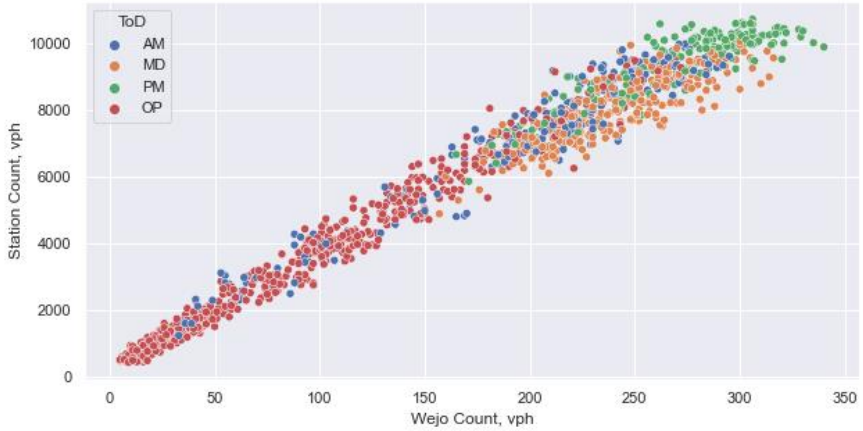
Functional Classification	MPR mean	MPR median	MPR Std. Dev	Avg. Hourly Volume	Avg. Daily Traffic
1	2.55%	2.62%	0.76%	2,913	69,906
2	2.31%	2.28%	1.07%	2,329	55,890
3	3.25%	3.15%	1.48%	853	20,465
4	4.39%	3.69%	2.65%	282	6,763

Data Quality (station variation)

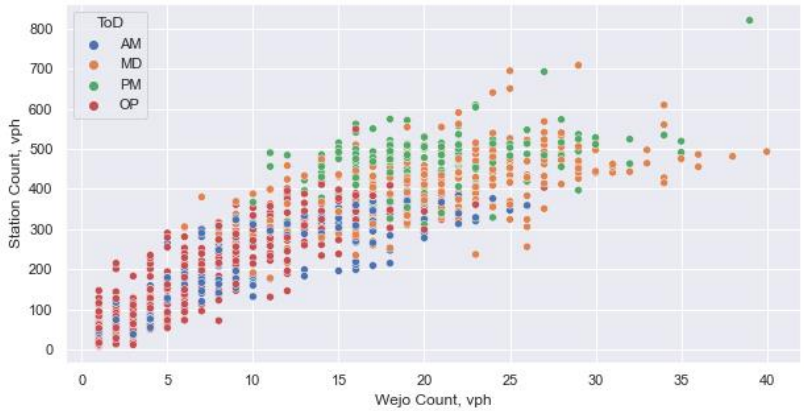


Left: WiM station on NJ-3 (r^2=0.977)

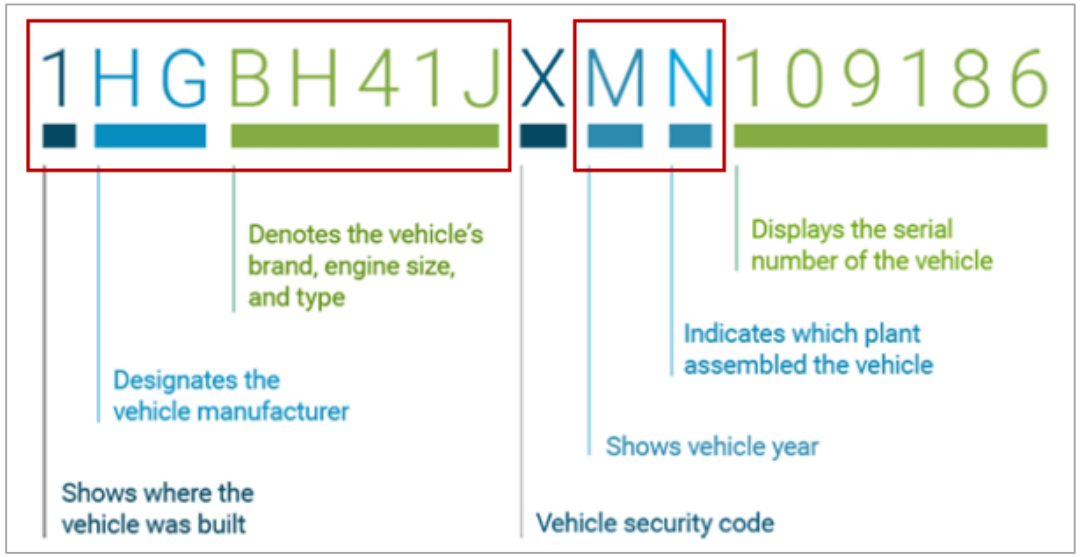




Right: WiM station on NJ-168 (r^2=0.735)

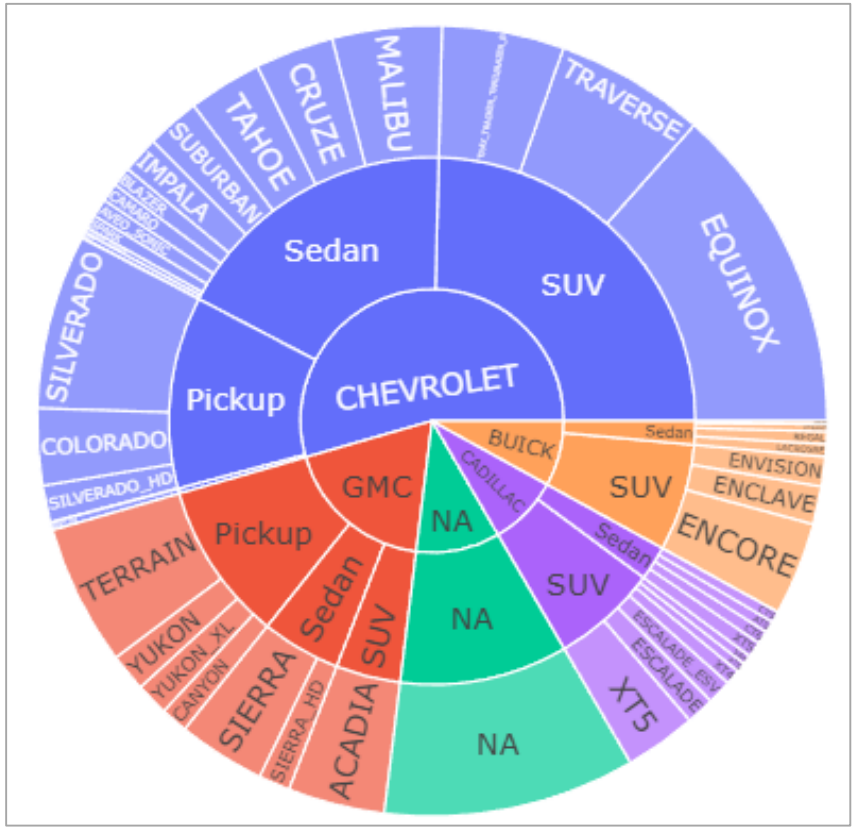


Data Overview (vehicle information)



"SquishVIN" (source: decodethevin.com)

- 8-digit "SquishVIN" number provides vehicle make, year, model, body type
- Most of the Movement data (89.7%) has "SquishVIN"
- Two EV models: Chevrolet Blot and Spark

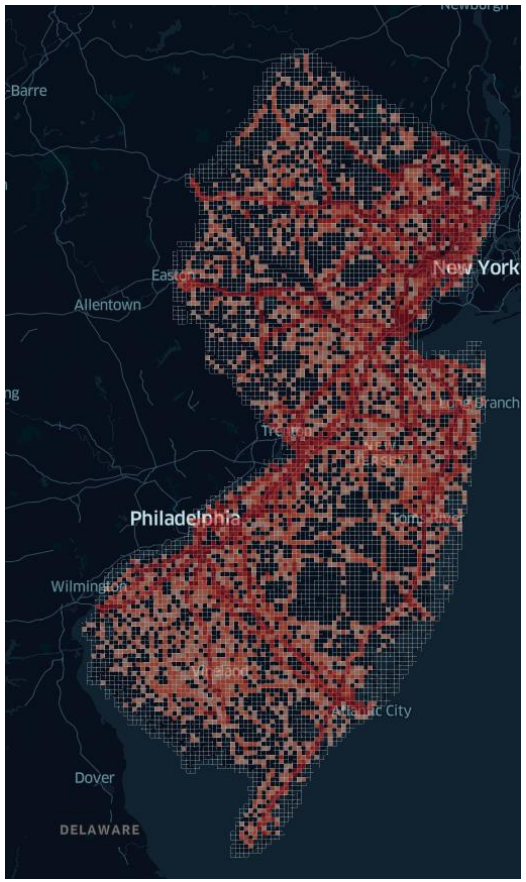


"SquishVIN" composition

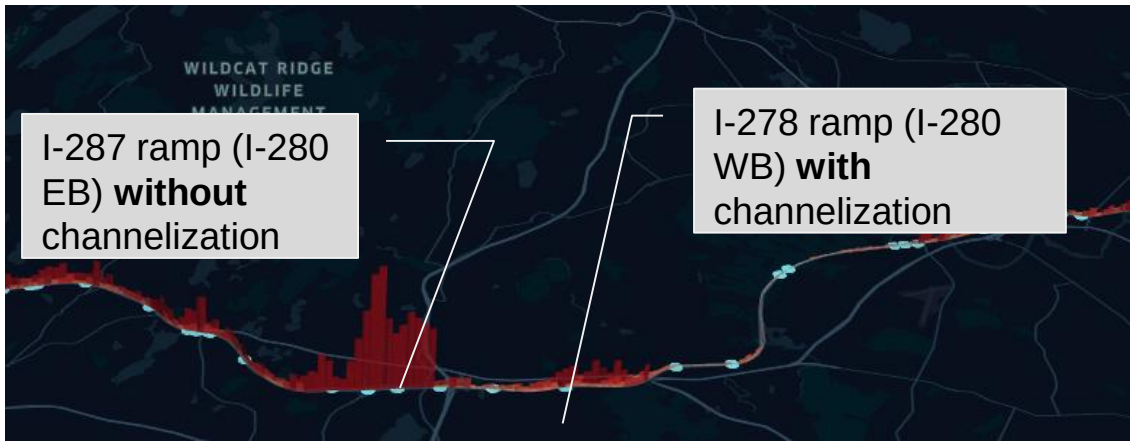
Applications

1. Roadway risk scan
2. Traffic monitoring
3. Parking facility analysis
4. Intersection management

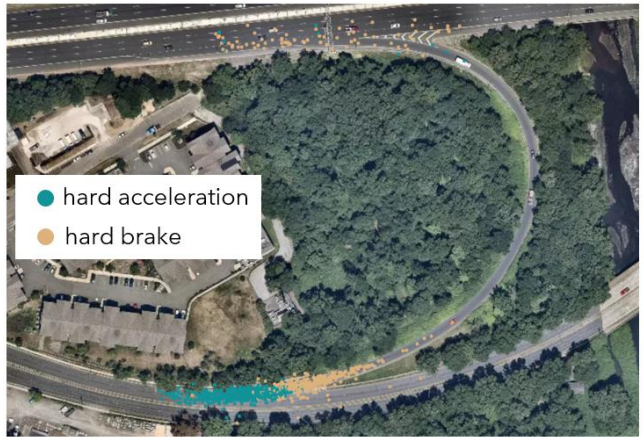
#1 Roadway Risk/Driving Behavior Scan



Statewide 0.1x0.1mi grid scan for hard-braking events

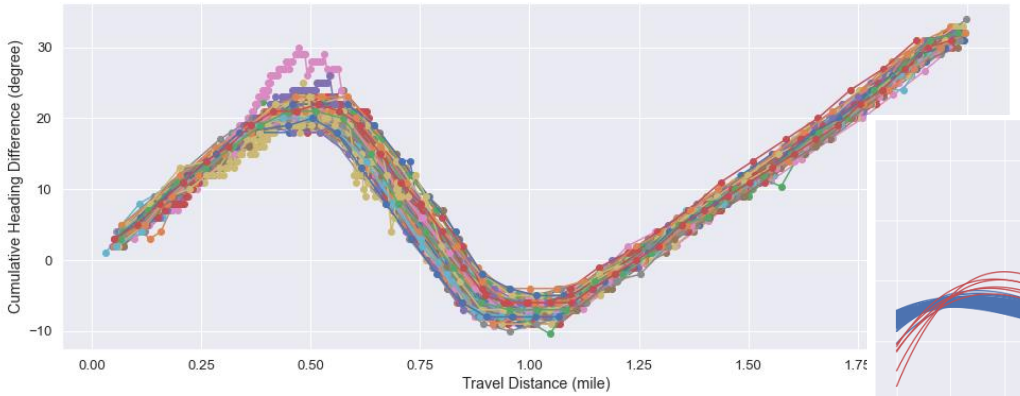


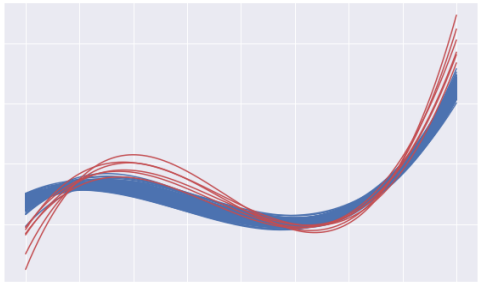
I-80 0.1-mi segment scan (I-287 & I-280 interchange shown)



Individual off-ramp scan

Vehicle set with cumulative heading difference



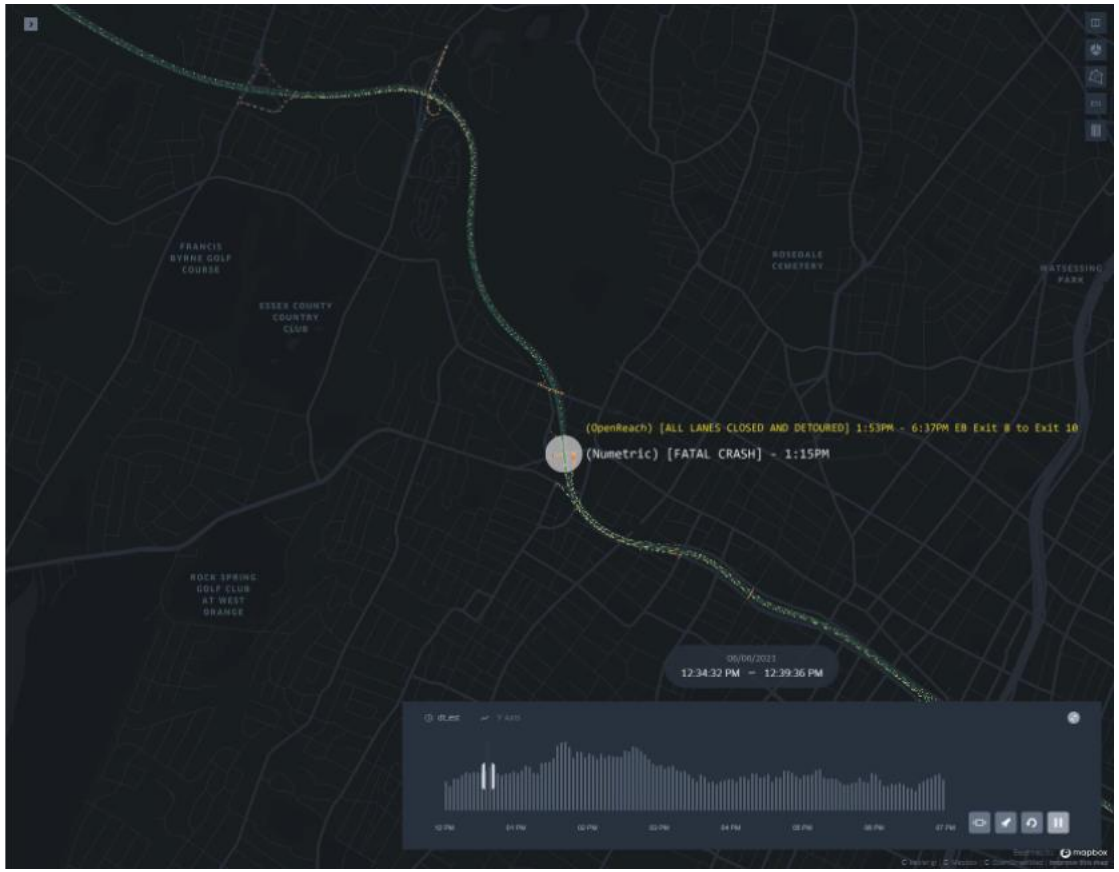


Two groups clustered

- *Red lines represent a different pattern from blue lines*
- *Back to vehicle trajectories, these red lines come from vehicles making lane changes due to crash lane closure*

- *Different road geometries (e.g., with on-ramp/off-ramp, etc.) needs to be checked for the algorithm robustness*

#2 Traffic Monitoring



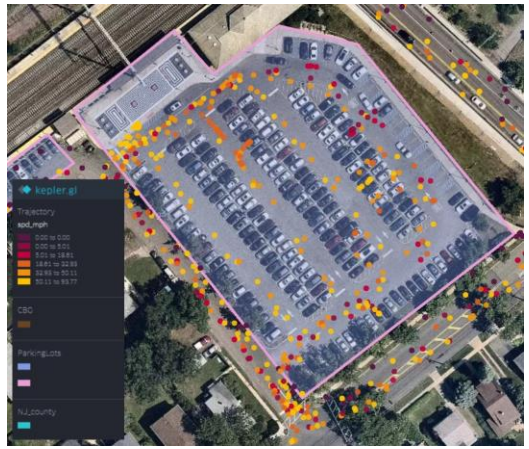
Animation: crash impact

- I-280 Traffic (Fatal Crash Zoom-In)
  - Numetric Start: 2021-06-06 1:15PM
  - OpenReach Start: 2021-06-06 1:53PM
    - EB All lanes closed and detoured
  - OpenReach Close: 2021-06-06 6:37PM
- Traffic
  1. Crash occurred
  2. Delays in both direction
  3. All-lane closure: no EB traffic
  4. Detour: congested upstream interchange @ Prospect Ave

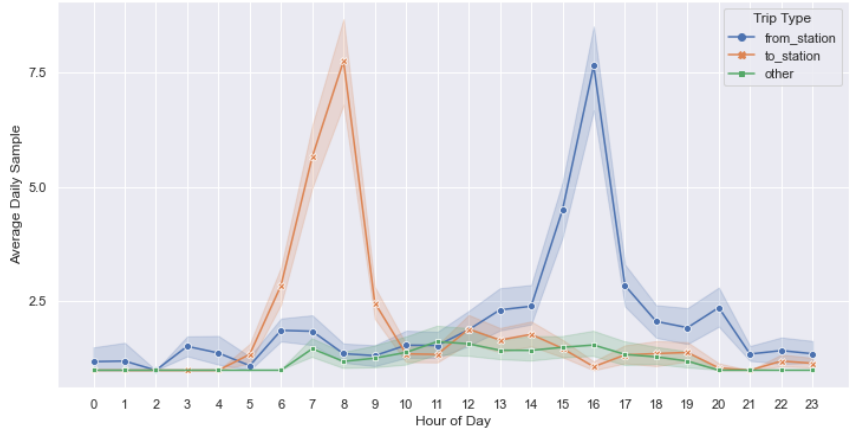
#3 Parking Analysis

Study of park-and-ride facilities

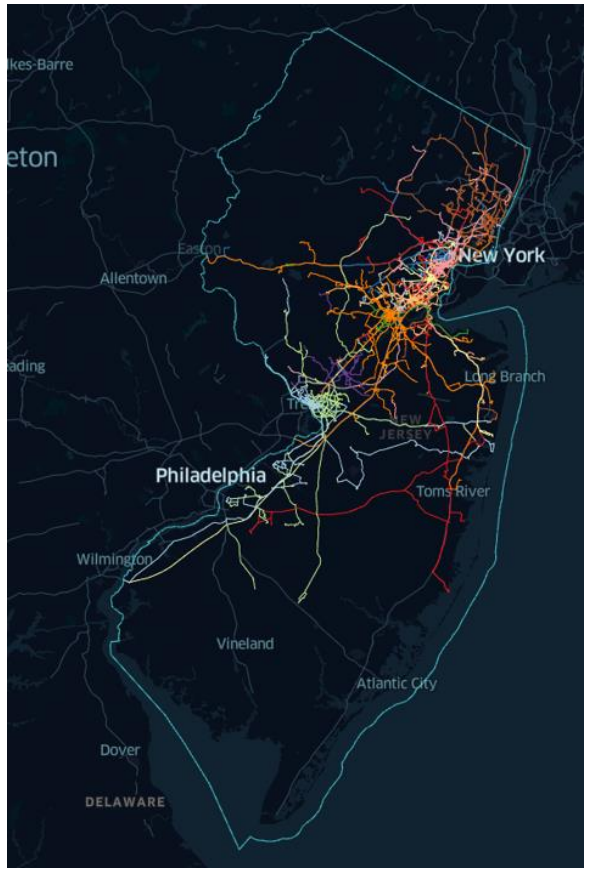
- Number of sample
- Duration at parking lot (searching for parking)
- Facility users' origin-destination (catchment area)
- Trip characteristics (travel distance, travel time to/from a station)



Vehicle movement within parking lot (Edition)



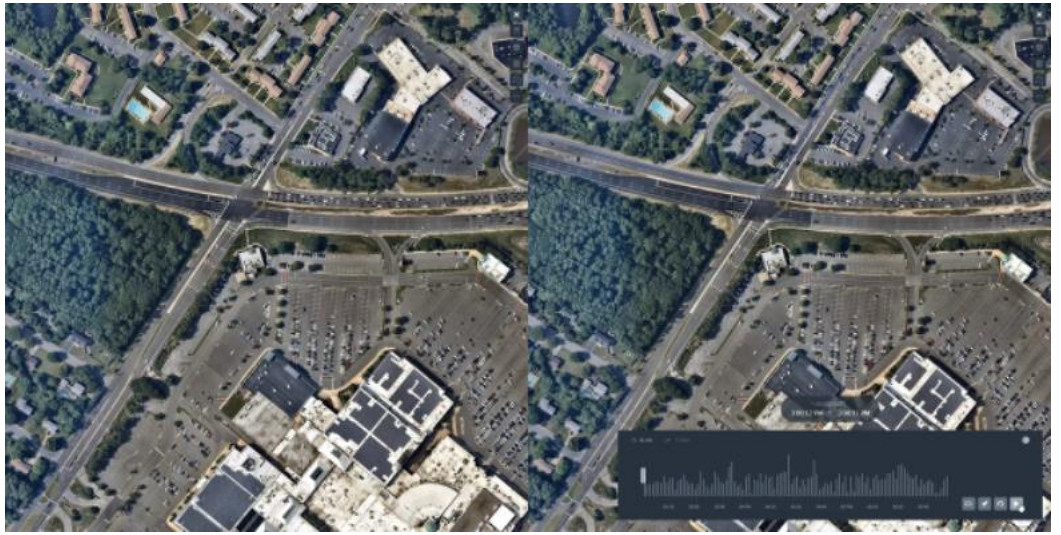
Inflow and outflow of a station (New Brunswick)



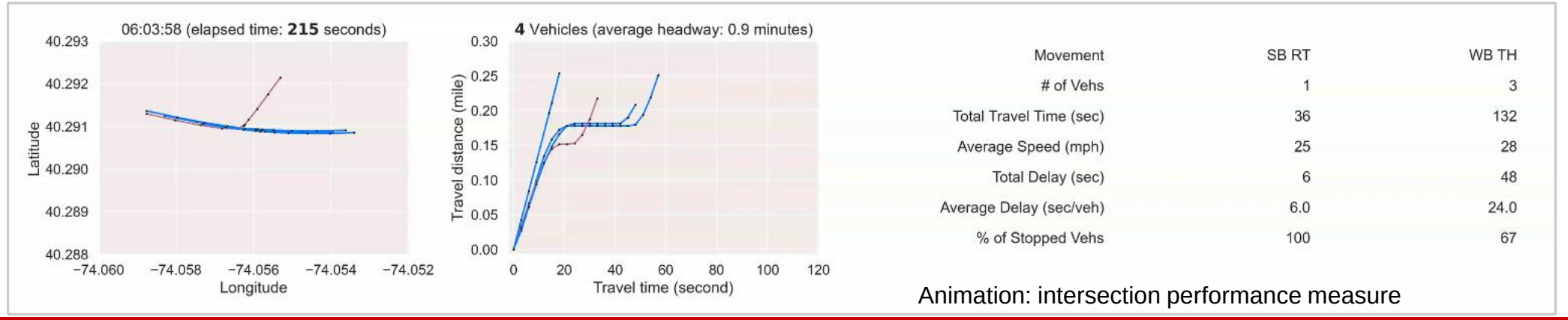
Catchment areas for stations of Northeast Corridor line

#4 Intersection/Interchange Management

NJ-36 & Wyckoff Rd, Eatontown



Animation: intersection movements

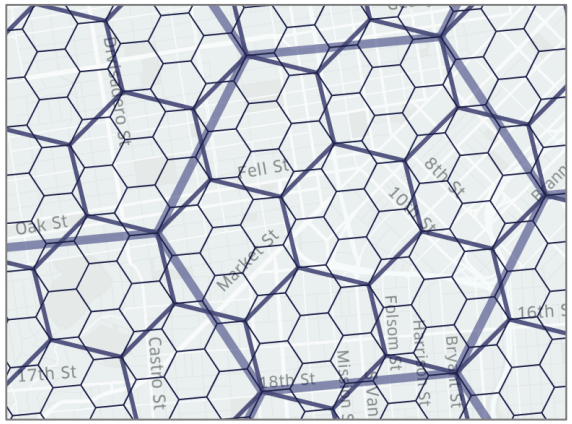


Movement	SB RT	WB TH
# of Vehs	1	3
Total Travel Time (sec)	36	132
Average Speed (mph)	25	28
Total Delay (sec)	6	48
Average Delay (sec/veh)	6.0	24.0
% of Stopped Vehs	100	67

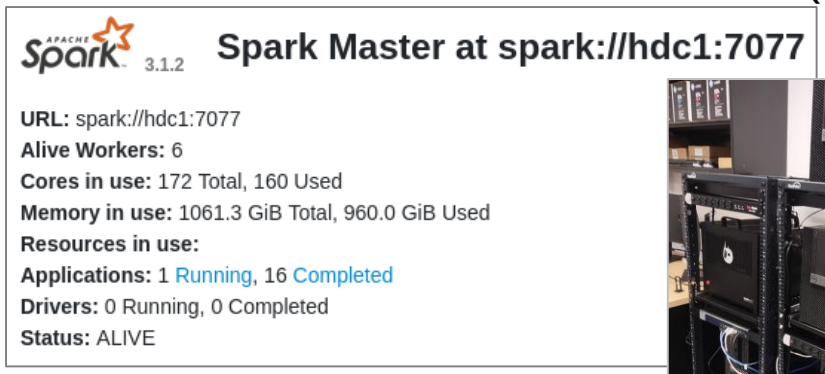
Animation: intersection performance measure

Data Processing Challenge

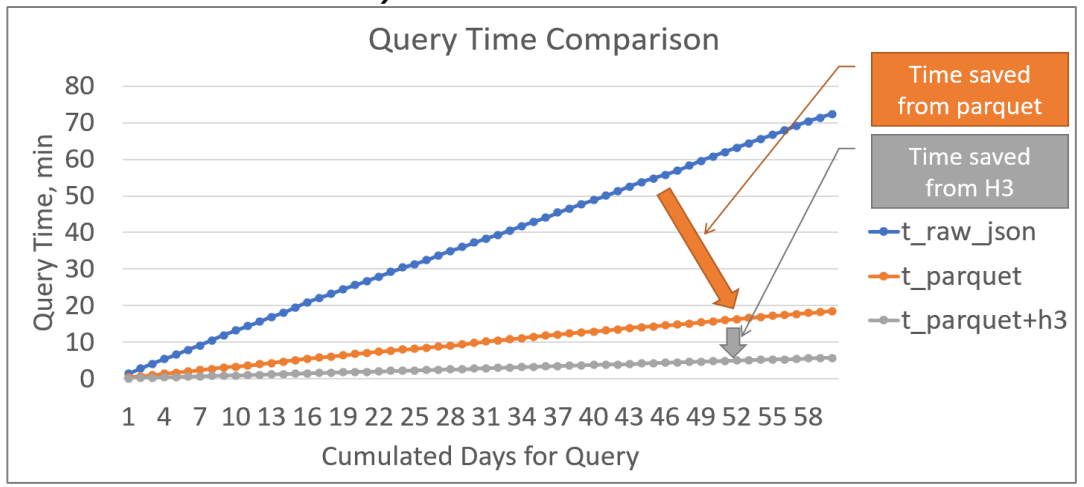
- Scale of the data: 2-mo NJ statewide data: 9TB (uncompressed) JSON, 17 billion records
- ITSRC Spark-Hadoop Cluster: 6 computation nodes (future expandable)
- Data engineering optimization
  - Data format (JSON → Parquet)
  - Spatial indexing (geocoordinate → H3 spatial index)
  - Time reduction: 3 hrs. → 2 min (extract waypoints for one intersection)



H3 hierarchical spatial index



Spark-Hadoop cluster @ ITSRC



Query time improvement

Conclusions

- The market penetration of the instrumented vehicles resides within the range of 2.4% - 4.5%, depending on roadway functional classifications.
- The trajectory data has near ubiquitous coverage in NJ roadways and found to be representative to traffic stream.
- CV data is cost-effective and provides much greater observability of the transportation system without ad hoc sensors
- Highly granular vehicle trajectory (as well as OD pair) reconstruction is feasible
- Scale (and granularity) of data may pose processing challenge

Thank You

Dr. Zijia Zhong
Sr. Transportation Engineer
zijia.zhong@njit.edu

Dr. Branislav Dimitrijevic
Assistant Professor
dimitrijevic@njit.edu

Dr. Dejan Besenski
Deputy Director
besenski@njit.edu

Intelligent Transportation Systems Resources Center
John A. Reif, Jr. Department of Civil & Environmental Engineering
New Jersey Institute of Technology