Volume & Turning Movements Project

Steering Committee Meeting #3

JANUARY 26, 2017

Conference call number:
Introductions

Kaveh Sadabadi, PhD  
Research Faculty  
Center for Advanced Transportation Technology  
University of Maryland  
kfarokhi@umd.edu  
(301) 405-1352

Denise Markow, PE  
TSMO Coordinator  
I-95 Corridor Coalition  
dmarkow@i95coalition.org  
(603) 321-0990

Dr. Stanley Young, PhD, PE  
National Renewable Energy Laboratory (NREL)  
U.S. Department of Energy  
Stanley.young@nrel.gov  
(301) 792-8180

Other Project Contacts

UMD Project Manager/Contracts: Kathy Frankle (UMD-CATT), 301-405-8271 or kfrankle@umd.edu

Steering Committee Coordinator: Joanna Reagle (KMJ Consulting, Inc.), 610-228-0760 or jreagle@kmjinc.com
### Attendees

<table>
<thead>
<tr>
<th>Agencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHWA</td>
</tr>
<tr>
<td>Maryland DOT/SHA</td>
</tr>
<tr>
<td>MWCOG</td>
</tr>
<tr>
<td>NHDOT</td>
</tr>
<tr>
<td>PennDOT</td>
</tr>
<tr>
<td>Port Authority of NY &amp; NJ</td>
</tr>
<tr>
<td>South Carolina DOT</td>
</tr>
<tr>
<td>Virginia DOT</td>
</tr>
<tr>
<td>USDOT</td>
</tr>
<tr>
<td>HERE</td>
</tr>
<tr>
<td>INRIX</td>
</tr>
<tr>
<td>TTI</td>
</tr>
<tr>
<td>NREL</td>
</tr>
<tr>
<td>I-95 Corridor Coalition</td>
</tr>
<tr>
<td>UMD CATT</td>
</tr>
<tr>
<td>UMD CATT Lab</td>
</tr>
<tr>
<td>I-95 Corridor Coalition (support)</td>
</tr>
</tbody>
</table>
Please confirm that your line is muted

*6

Thank you!
Agenda

• Very brief project overview
• Survey results
• Analysis of Maryland data
• New project one-pager
• Steering committee input
• Next steering committee meeting/webinar
Project Overview

- Initiated in 2013 as part of the I-95 Corridor Coalition MCOMP proposal

- Goal of project is to accelerate the timeframe to achieve viable volume and turning movement data through probe data

- Hypothesis: Information in existing probe data can be used to infer volume levels both for real-time operations and for historical planning applications
Project Background

• Network wide volume and turning movement data remain key missing dimensions for operational awareness and assessing system performance

• Turning movement data is only available in special studies

• There is a need for 24x7x365 volume (or density) estimates across the network
8 Projected Benefits / Uses

OPERATIONS

• Improved incident management monitoring and action
• Enhanced work zone monitoring, impact analysis, and safety
• More accurate user delay cost reporting
• Special event management
• Improved traffic signal system timing management

PLANNING

• Additional insight to anticipate and verify “jam” conditions
• Detailed after-action reviews
• Improved system performance evaluation
• Expanded project / program assessment
• Advance travel demand modeling accuracy
• Better address air quality, emissions requirements and energy analysis inquiries
Approach / Tasks / Objectives

- Define a framework for delivery of probe-based volume & turning movement data.
- Understand, document, & share data needs for various DOT applications requiring such data.
- Create a calibration and validation testbed to assist vendors’ initial development efforts.
- Provide representative data products.
- Set appropriate expectation for data fidelity, form, granularity, and usability.
- Anticipate the need for an ongoing calibration network.
- Estimate resources needed to maintain/operate a national calibration/validation testbed.
Survey Structure / Stats

• 14 completed surveys
  • 11 member agencies
  • 3 outside of coalition
• 10 additional partial surveys
• 18 survey questions that queried
  • Need for data
  • Accuracy
  • Format
• Prompted for comments
• A full report available
Overview of Survey Results

VOLUME DATA

• Great interest from a planning perspective for all aspects of this type of data.

• Real time volume data has a higher perceived value for incident management monitoring than for traveler information.

• The preferred volume metric was vehicle flow (vph) as opposed to percent capacity or vehicle density.

• The needed level of accuracy for flow data to support application -within 10% of roadway capacity.

• The minimum time interval/aggregation -15 minute intervals for real time, however archived flow data could be longer intervals of 30 minutes to 1 hour.

• An overwhelming additional desirable attribute was the percentage (or volume) of heavy duty trucks.

TURNING MOVEMENT DATA

• The need for archived turning movement data is more defined for planning and performance measures.

• There is a perceived need for real time turning movement data, particularly for detours and evacuation.

• The need for and use of the data for day to day operations application is less defined than for archived turning movement data.

• With respect to turning movements, there was no clear preference for a defined metric.

• Either estimates of volume in each direction or percent of turning movements in each direction were acceptable.

• If percentage of turning vehicles was reported, accuracy (and precision) to within 10% is preferred.

• Similar to volume, turning movements should be reported at 15 minute time aggregations.
Statewide Traffic Volume Estimation using GPS Traces: Machine Learning Approach

Przemysław Sekuła, Nikola Marković

Center for Advanced Transportation Technology (CATT)
Department of Civil & Environmental Engineering
University of Maryland

psekula@umd.edu, nikola@umd.edu

Presented by Stanley Young

January 17, 2016
Outline

1. Introduction
2. Data
3. Regression
4. Conclusions
Detailed information about traffic volumes is of utmost importance both for transportation planning and traffic engineering.

Maryland:
- 84 permanent automatic traffic recorder (ATR) stations
- Provide hourly counts broken down to vehicle classes
- However, we need to know volumes throughout the state

How do we estimate volumes at road links without ATR stations?
Research Questions

- Can we use GPS traces to estimate hourly volumes throughout the state of Maryland?

INRIX probes generate GPS traces

- Can we compare GPS traces and ATR data, learn the underlying relations, and then estimate volumes throughout the network?
- How accurate would those estimates be?
GPS Traces

- 4 months of INRIX data during 2015 (February, June, July, October)
  - 20 million trips, which include 1.4 billion way points
  - 112 GB of data

Way points are typically 1 sec apart

Providers and vehicle classes

- Fleet: 59%
- Consumer: 31%
- Mobile: 10%

- Vehicles ≤ 14,000 lb: 22%
- Vehicles ∈ (14,000, 26,000) lb: 33%
- Vehicles ≥ 26,000 lb: 45%
GPS Traces

- Average hourly penetration rates of GPS traces are computed at 12 ATR locations based on all 4 months of data.

- Average hourly penetration rates vary from 0.18% to 0.72%, with the median of 0.57%.
- Average hourly GPS volumes vary from 22.3 to 62.3 vehicles, with the median of 37.3 vehicles.
Additional Data

- **Speed info**
  - Speeds estimated directly from GPS traces
  - Speeds from RITIS

- **Road characteristics**
  - Type of the road (IS, US, MD), number of lanes, speed limit, directions separated (yes/no)

- **Weather info**
  - Temperature, humidity, pressure, visibility, wind speed, precipitation, conditions (e.g., clear, cloudy, fog)

- **Incident reports (work in progress)**
  - All the information available on RITIS (e.g., work zones, collisions, disabled vehicle, obstructions)
Correlation

Positive correlation between
- ATR and GPS counts
- ATR/GPS counts and number of lanes
- Speed and speed limit

Negative correlation between
- ATR/GPS counts and speed

Yellow/magenta: positive/negative correlation
Regression

- Infer relation between GPS and ATR volumes at the hourly level
- Use this relation to estimate volumes at roads without ATR stations

Models applied
- ANN*
- Random forest
- SVM

* Results shown based on ANN
Calibration and Evaluation

- Use 11 ATR stations for training and the remaining ATR for testing

- Data from 11 ATR stations are used for regression
  - 66,000 data points for training

- Data from the remaining ATR are used for evaluation
  - 6,000 data points for testing

- Repeat this 12 times and report test results for all 12 locations
Test results for all 12 sensors are reported in terms of $R^2$ and error.

- The $R^2$ varies from 0.61 to 0.94, with the median of 0.82.
- The error varies from 14% to 48%, with the median of 27%.
Example Results

Typical

Station: 21, direction: eastbound, R^2 = 0.792

Station: 21, direction: westbound, R^2 = 0.822

Worst

Station: 27, direction: southbound, R^2 = 0.529

Best

Station: 71, direction: northbound, R^2 = 0.945
Example Results

- Typical
  - Station: 21, direction: eastbound, R² = 0.792
  - Station: 21, direction: westbound, R² = 0.822

- Worst
  - Station: 27, direction: southbound, R² = 0.529

- Best
  - Station: 71, direction: northbound, R² = 0.945
Contribution of GPS Traces

- Re-train and re-test the model without GPS traces

\[ R^2 \text{ for test data} \]
\[ \text{Error for test data} \]

- The \( R^2 \) varies from 0.49 to 0.90, with the median of 0.73
- The error varies from 16% to 54%, with the median of 37% (27%)
Conclusions

- We can use ML to estimate hourly volumes with average accuracy of:
  - 27% if GPS traces are available ($R^2 = 0.82$)
  - 37% otherwise ($R^2 = 0.73$)

- Estimated volumes can be used for computation of performance measures in RITIS
Acknowledgments

Contact Information
Przemysław Sekuła, Nikola Marković
Center for Advanced Transportation Technology
University of Maryland
psekula@umd.edu, nikola@umd.edu
From statistics to meaning….

- Results of initial MD analysis – from a roadway VOLUME perspective

<table>
<thead>
<tr>
<th>Volume (from)</th>
<th>Volume (to)</th>
<th>Number of samples</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>500</td>
<td>8479</td>
<td>68.89%</td>
</tr>
<tr>
<td>500</td>
<td>1000</td>
<td>7701</td>
<td>35.34%</td>
</tr>
<tr>
<td>1000</td>
<td>1500</td>
<td>5946</td>
<td>32.26%</td>
</tr>
<tr>
<td>1500</td>
<td>2000</td>
<td>6352</td>
<td>22.64%</td>
</tr>
<tr>
<td>2000</td>
<td>2500</td>
<td>5879</td>
<td>20.42%</td>
</tr>
<tr>
<td>2500</td>
<td>3000</td>
<td>4711</td>
<td>18.62%</td>
</tr>
<tr>
<td>3000</td>
<td>3500</td>
<td>3221</td>
<td>18.59%</td>
</tr>
<tr>
<td>3500</td>
<td>4000</td>
<td>2650</td>
<td>20.63%</td>
</tr>
<tr>
<td>4000</td>
<td>4500</td>
<td>3235</td>
<td>19.56%</td>
</tr>
<tr>
<td>4500</td>
<td>5000</td>
<td>3621</td>
<td>14.89%</td>
</tr>
<tr>
<td>5000</td>
<td>5500</td>
<td>3313</td>
<td>10.80%</td>
</tr>
<tr>
<td>5500</td>
<td>6000</td>
<td>2554</td>
<td>10.61%</td>
</tr>
<tr>
<td>6000</td>
<td>6500</td>
<td>2304</td>
<td>10.16%</td>
</tr>
<tr>
<td>6500</td>
<td>7000</td>
<td>1876</td>
<td>12.29%</td>
</tr>
<tr>
<td>7000</td>
<td>9266</td>
<td>1592</td>
<td>16.81%</td>
</tr>
</tbody>
</table>
From statistics to meaning....

- Results of initial MD analysis – from a roadway VOLUME perspective

<table>
<thead>
<tr>
<th>Volume (from)</th>
<th>Volume (to)</th>
<th>Number of samples</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>500</td>
<td>8479</td>
<td>68.89%</td>
</tr>
<tr>
<td>500</td>
<td>1000</td>
<td>7701</td>
<td>35.34%</td>
</tr>
<tr>
<td>1000</td>
<td>1500</td>
<td>5946</td>
<td>32.26%</td>
</tr>
<tr>
<td>1500</td>
<td>2000</td>
<td>6352</td>
<td>22.64%</td>
</tr>
<tr>
<td>2000</td>
<td>2500</td>
<td>5879</td>
<td>20.42%</td>
</tr>
<tr>
<td>2500</td>
<td>3000</td>
<td>4711</td>
<td>18.62%</td>
</tr>
<tr>
<td>3000</td>
<td>3500</td>
<td>3221</td>
<td>18.59%</td>
</tr>
<tr>
<td>3500</td>
<td>4000</td>
<td>2650</td>
<td>20.63%</td>
</tr>
<tr>
<td>4000</td>
<td>4500</td>
<td>3235</td>
<td>19.56%</td>
</tr>
<tr>
<td>4500</td>
<td>5000</td>
<td>3621</td>
<td>14.89%</td>
</tr>
<tr>
<td>5000</td>
<td>5500</td>
<td>3313</td>
<td>10.80%</td>
</tr>
<tr>
<td>5500</td>
<td>6000</td>
<td>2554</td>
<td>10.61%</td>
</tr>
<tr>
<td>6000</td>
<td>6500</td>
<td>2304</td>
<td>10.16%</td>
</tr>
<tr>
<td>6500</td>
<td>7000</td>
<td>1876</td>
<td>12.29%</td>
</tr>
<tr>
<td>7000</td>
<td>9266</td>
<td>1592</td>
<td>16.81%</td>
</tr>
</tbody>
</table>

- The greater the volume the less the percent error
From statistics to meaning....

- Results of initial MD analysis – from a roadway CAPACITY perspective

<table>
<thead>
<tr>
<th>% Capacity (from)</th>
<th>% Capacity (to)</th>
<th>Number of samples</th>
<th>% Volume</th>
<th>% Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>10%</td>
<td>8348</td>
<td>64.23%</td>
<td>3.46%</td>
</tr>
<tr>
<td>10%</td>
<td>20%</td>
<td>7430</td>
<td>39.31%</td>
<td>5.57%</td>
</tr>
<tr>
<td>20%</td>
<td>30%</td>
<td>4985</td>
<td>32.31%</td>
<td>8.08%</td>
</tr>
<tr>
<td>30%</td>
<td>40%</td>
<td>4899</td>
<td>28.47%</td>
<td>9.92%</td>
</tr>
<tr>
<td>40%</td>
<td>50%</td>
<td>5497</td>
<td>23.15%</td>
<td>10.43%</td>
</tr>
<tr>
<td>50%</td>
<td>60%</td>
<td>5125</td>
<td>19.64%</td>
<td>10.76%</td>
</tr>
<tr>
<td>60%</td>
<td>70%</td>
<td>5818</td>
<td>16.04%</td>
<td>10.41%</td>
</tr>
<tr>
<td>70%</td>
<td>80%</td>
<td>6372</td>
<td>14.28%</td>
<td>10.71%</td>
</tr>
<tr>
<td>80%</td>
<td>90%</td>
<td>5912</td>
<td>14.17%</td>
<td>12.01%</td>
</tr>
<tr>
<td>90%</td>
<td>100%</td>
<td>5394</td>
<td>14.49%</td>
<td>13.74%</td>
</tr>
</tbody>
</table>
From statistics to meaning....

• Results of initial MD analysis – from a CAPACITY perspective
From statistics to meaning....

• Results of initial MD analysis – from a **CAPACITY** perspective

<table>
<thead>
<tr>
<th>% Capacity (from)</th>
<th>% Capacity (to)</th>
<th>Number of samples</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>10%</td>
<td>8348</td>
<td>64.23%</td>
</tr>
<tr>
<td>10%</td>
<td>20%</td>
<td>7430</td>
<td>39.31%</td>
</tr>
<tr>
<td>20%</td>
<td>30%</td>
<td>4865</td>
<td>32.31%</td>
</tr>
<tr>
<td>30%</td>
<td>40%</td>
<td>5899</td>
<td>28.47%</td>
</tr>
<tr>
<td>40%</td>
<td>50%</td>
<td>5497</td>
<td>23.15%</td>
</tr>
<tr>
<td>50%</td>
<td>60%</td>
<td>5125</td>
<td>19.64%</td>
</tr>
<tr>
<td>60%</td>
<td>70%</td>
<td>5818</td>
<td>16.04%</td>
</tr>
<tr>
<td>70%</td>
<td>80%</td>
<td>6372</td>
<td>14.28%</td>
</tr>
<tr>
<td>80%</td>
<td>90%</td>
<td>5912</td>
<td>14.17%</td>
</tr>
<tr>
<td>90%</td>
<td>100%</td>
<td>5394</td>
<td>14.49%</td>
</tr>
</tbody>
</table>

• The average error with respect to capacity is **9.5% !!!**
Overview of Survey Results

• There is great interest from a planning perspective and for planning applications for all aspects of this type of data.

• Real time volume data seems to have a higher perceived value for incident management monitoring than for traveler information.

• The preferred volume metric was vehicle flow (vehicles per hour) as opposed to percent capacity or vehicle density.

• **The needed level of accuracy for flow data to support anticipated application was to within 10% of roadway capacity.**

• The minimum time interval/aggregation that was recommended was 15 minute intervals for real time, however archived flow data could be longer intervals of 30 minutes to 1 hour.

• An overwhelming additional desirable attribute was the percentage (or volume) of heavy duty trucks.
Additional/Caveats on existing analysis ...

- Current 9.5% error with respect to capacity ....
  - More data should decrease error
  - Current calculations are with respect to observed maximum capacity ...
    Error with respect to theoretical capacity even lower.

- Caveats
  - Current error metric is Average Absolute Error
  - Survey accuracy response may be more aligned with 95% (2- sigma) limit

- Bottom line – we are in the Ball Park ... Stay Tuned
On Coalition Website (VPP page)

Reached out to steering committee for quotes – Thank You

Need help enunciating anticipated benefits
36 Steering Committee Participation and Feedback

- Proactive in populating test bed data
  - Count data in 15 minute intervals (minimum) with vehicle class
- Comments on initial MD Analysis
  - Insights, suggested direction, balance of light-duty / heavy-duty
- Enunciating the benefits of this project
Questions
Wrap Up

Closing remarks

Next meeting/webinar

- Thursday, April 13, 2017
- 10:30a.m. - 12:00p.m. (EDT)

Agenda
- Specifications & validation/calibration methodology
- First look at other data sets
Thank You!

• For Questions, please contact:
  • **PI** – Kaveh Sadabadi (UMD-CATT) 301-405-1352 or kfarokhi@umd.edu
  • **Co-PI** – Denise Markow (I-95 Corridor Coalition) 301-789-9088 or dmarkow@i95coalition.org
  • **Co-PI** – Stanley Young (NREL) 301-792-8180 or Stanley.Young@nrel.gov
  • **UMD PM/Contracts** – Kathy Frankle (UMD-CATT) 301-405-8271 or kfrankle@umd.edu
  • **Logistics** – Joanna Reagle (KMJ Consulting, Inc.) 610.228.0760 or jreagle@kmjinc.com