Data Driven Approach To Characterize And Forecast The Impact Of Work Zones On Freeway Mobility Using Probe Vehicle Data

By:

Mohsen Kamyab, PhD

Intelligent Transportation Systems

September 2020
Presentation Overview

1. Introduction & Motivation
   • Why should we care about work zone impact on mobility?

2. Objectives
   • Can we improve our traffic operation and reduce the impact?

3. Methodology
   • Introducing a data drive approach to characterize and forecast the impact.

4. Results & Conclusion
Motivation: Federal Highway Administration (FHWA) Calls For Improvement In Work Zone Mobility Management

FHWA calls for transportation agencies to:

- Develop policies to manage work zone mobility
- Develop systematic approaches for mobility performance measurement
- Update and refine policies to optimize mobility performance

FHWA Policy development and implementation process

Source: Implementing the Rule on Work Zone Safety and Mobility, FHWA, 2005.
Work Zone Traffic Management Key Concerns

Mobility:
- 24% of nonrecurring freeway delays are due to work zone projects
- 888 million hours were lost in 2014
- User dissatisfaction

Environment:
- Millions of gallons of fuel used annually
- Emission detrimental to public health
- On average, 300 million gallons of fuel are lost every year as a result of work zones presence.
Work Zone Traffic Management Key Concerns: Safety

Safety:
- As more work is required to maintain the roadways, more risk is introduced:
  - Commuter’s safety
  - Worker’s Safety

National Level
- Total Work Zone Fatal Crashes by type of highway:
  - 2016: Total 688, Interstate 261, Arterial 355
  - 2017: Total 710, Interstate 283, Arterial 348

2018 Michigan
- Total: 16 Fatalities (Previous 5-year average: 19.7)
- Truck Involved: 15 Fatal Crashes (Previous 5-year average: 16.3)
- Bus Involved: 16 Fatalities
- Pedestrian Involved: 5 Fatal Crashes (Previous 5-year average: 3.7)
- Worker Involved: 5 Fatalities (Previous 5-year average: 2.7)

Source: workzonesafety.org

Source: FHWA
Specific Research Objectives

The specific objectives of this study were to:

1. Develop a systematic approach to measure and visualize the impact of work zones
2. Predict the impact future work zones will have on interstate’s mobility
3. Develop a high-level decision-making process to better plan future work zones
Research Overview

Available Data
- Work zone activities
- Traffic Speed
- Roadway Geometry
- Traffic Incidents

WORK ZONE BIG DATA IS AVAILABLE NOW

Data Fusion
- Database
- Fuse work zone data

BIG DATA ANALYTICS IS THE EMERGING SOLUTION

Performance Measurement
- IF YOU CANNOT MEASURE IT, YOU CANNOT IMPROVE IT

Impact Prediction
- MAKING SMARTER DECISIONS
Probe Vehicle Data Introduction

- GPS devices broadcast microwave signals
- GPS receivers collect this data to determine location and time
- Using location and time, probe’s speed is derived
Probe Vehicle Data Usage In Navigation Systems

Morning Rush Hour

![Map of traffic during morning rush hour]

Off-peak Hour

![Map of traffic during off-peak hour]

Source: https://www.google.com/maps
Probe Vehicle Data Overview

Probe Vehicle Data Providers

INRIX

here

TOMTOM

Probe Vehicle Data Collection Process

Probe Vehicle Data Segmentation Scheme

<table>
<thead>
<tr>
<th>Segment</th>
<th>Timestamp</th>
<th>Speed (mph)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2/1/2015 14:00</td>
<td>56</td>
</tr>
<tr>
<td>B</td>
<td>2/1/2015 14:00</td>
<td>50</td>
</tr>
<tr>
<td>C</td>
<td>2/1/2015 14:00</td>
<td>45</td>
</tr>
<tr>
<td>D</td>
<td>2/1/2015 14:00</td>
<td>53</td>
</tr>
<tr>
<td>E</td>
<td>2/1/2015 14:00</td>
<td>65</td>
</tr>
<tr>
<td>A</td>
<td>2/1/2015 14:01</td>
<td>59</td>
</tr>
<tr>
<td>B</td>
<td>2/1/2015 14:01</td>
<td>53</td>
</tr>
<tr>
<td>C</td>
<td>2/1/2015 14:01</td>
<td>48</td>
</tr>
<tr>
<td>D</td>
<td>2/1/2015 14:01</td>
<td>56</td>
</tr>
<tr>
<td>E</td>
<td>2/1/2015 14:01</td>
<td>68</td>
</tr>
</tbody>
</table>
Develop a systematic approach to measure and visualize the impact of work zones.
Highway Segments Selection

Traffic direction:

Upstream area: 5 miles prior to work zone start mile marker

Work Zone area: Segments falling between work zone start and end mile markers

Downstream area: 3 miles after work zone end mile marker
Delay Measurement

Work zone VS typical traffic condition:

- Provides realistic delay measurement
- Account for corridors that are congested even when work zone is not present
Queue Measurement

Using probe vehicle data:

- Segments with queueing condition (speed less than 15 mph) can be identified.
- Queue propagation to the upstream segments can be tracked.
# Proposed Performance Measures

<table>
<thead>
<tr>
<th>Metric</th>
<th>What does it Measure?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User Delay</strong></td>
<td></td>
</tr>
<tr>
<td>Total Delay</td>
<td>Cumulative travel time delay experienced by users throughout the lane-closure duration</td>
</tr>
<tr>
<td>Longest User Delay</td>
<td>Longest travel time delay experienced by users</td>
</tr>
<tr>
<td><strong>Presence of Queueing Condition</strong></td>
<td></td>
</tr>
<tr>
<td>Longest Queue Length (mile)</td>
<td>Longest length of queue caused by lane-closure</td>
</tr>
<tr>
<td>Longest Queue Duration (min)</td>
<td>Longest time that at least one segment of highway was performing in queueing condition.</td>
</tr>
<tr>
<td>Total Queue Duration(hours)</td>
<td>Cumulative times that at least one segment of highway was performing in queueing condition.</td>
</tr>
<tr>
<td>Number of Queues</td>
<td>Number of times that queueing condition formed on the highway.</td>
</tr>
</tbody>
</table>
Corridor-level Mobility Assessment

CDF plots

- Represents travel time variation
- Monitors travel time reliability
- Useful for high-level monitoring

Radar plot

- Summarizes traffic condition over hours of a day
- Represents aggregated traffic metrics
- Identifies problematic hours

Cumulative Frequency (%)

Travel Time (min.)

- Typical
- Work Zone

Magnitude

Hours of Day - 10 PM

Evening peak ends

Morning Peak starts
Segment-level Mobility Assessment: Spatial Assessment

Volcano plot

- Summarizes traffic condition for each segment
- Represents cumulative traffic measures
- Identifies problematic segments
- Useful for high-level monitoring
Spatiotemporal Mobility Assessment Using Speed Heat-map

- Provides an instant overview of mobility
- Useful to extract queue metrics such as queue length and duration
Work Zone Mobility Audit

Overview

- Work Zone ID: 104811
- County: Oakland County
- Roadway: I-75
- Closure type: Single Lane Closure
- Direction: Northbound
- Start Milemarker: 73.4
- End Milemarker: 76.1
- Workzone Start: 2016-08-24 07:00
- Workzone End: 2016-09-13 10:00

Comments:

- Work Zone Typical Traffic
- Travel Time (min)
- Work Zone Speed Heatmap
- Weekday AM
- Weekday Mid
- Weekday PM
- Weekend

Legend

- Workzone
- Weekdays
- Typical
- Weekdays
- Workzone
- Weekends
- Typical
- Weekends

Spatial Characterization

- Workzone duration
- Prior year

Temporal Characterization

- % of Time Congested
- Temporal Characterization
- Congestion Hours

Delay & LOTTR Metrics

<table>
<thead>
<tr>
<th>Stats</th>
<th>AM</th>
<th>Mid</th>
<th>PM</th>
<th>Weekend</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg Delay (min)</td>
<td>0.5</td>
<td>0.3</td>
<td>2.6</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Max Delay (min)</td>
<td>72.5</td>
<td>11.0</td>
<td>19.9</td>
<td>5.2</td>
<td>72.5</td>
</tr>
<tr>
<td>Total Delay (min)</td>
<td>1338.8</td>
<td>1139.8</td>
<td>2841.9</td>
<td>243.1</td>
<td>6675.8</td>
</tr>
<tr>
<td>LOTTR</td>
<td>1.2</td>
<td>1.3</td>
<td>1.5</td>
<td>1.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Queueing Metrics

<table>
<thead>
<tr>
<th>Stats</th>
<th>Queue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Duration (min)</td>
<td>185.0</td>
</tr>
<tr>
<td>Total Duration (min)</td>
<td>2300.0</td>
</tr>
<tr>
<td>Max Length (miles)</td>
<td>6.8</td>
</tr>
<tr>
<td># of Queues</td>
<td>39.0</td>
</tr>
</tbody>
</table>
WZMA Software is built to:

1. Utilizes probe vehicle data and work zone configuration information as an input
2. Automatically creates WZMA for individual or multiple work zones
3. Archives the mobility statistical summary for further mobility assessment
Work Zone Mobility Audit Tool Overview

SQL Database

- TMC Info
  - TMC Code
  - TMC Start MM
  - TMC End MM
  - TMC Length

- WZ Info
  - WZ ID
  - County
  - Roadway Name
  - Closure Type
  - TMC list
  - Start Date-Time
  - End Date-Time
  - Start Mile Marker
  - End Mile Marker

- Probe Data
  - TMC
  - Timestamp
  - Speed
  - Travel Time
  - Confidence
  - Avg Speed

- System Requirements
  - RStudio Installation
  - SQL Server Management
  - SQL ODBC Connection
  - Software Code
  - Software files

Connect

Output

WSP

WAYNE STATE UNIVERSITY

21
WZMA For A Weekend Work Zone With Severe Impact
Work Zone Mobility Audit

Overview

Work Zone ID: 103603_1
County: Ottawa County
Roadway: I-69
Closure type: Single Lane Closure
Direction: Eastbound
Start Milemarker: 130.8
End Milemarker: 141.8
Workzone Start: 2016-08-22 06:00
Workzone End: 2016-09-02 08:12

WZMA For A Weekend Work Zone With No Impact

Work Zone Travel Time vs Typical Traffic

Work Zone Congestion: Spatial Characterization

Work Zone Congestion: Temporal Characterization

Delay & LOTTR Metrics

Queueing Metrics

Wayne State University
Mobility Forecasting Using Machine Learning

Predict the impact future work zones will have on interstate’s mobility
Work Zone Mobility Forecasting Objectives

• Objective was to learn from historical work zones and predict mobility for future work zones.

• Can we predict speed for each segment throughout work zone presence?

Speed heatmap for a single-lane closure
Using Historical Speed Distribution To Predict Work Zone Mobility

- Historical speed distribution represents variation of mobility behavior
- Using this variation, corridor’s vulnerability is characterized
- This helps algorithms to better predict work zone impact
How To Train Machine Learning Algorithms?

Raw Traffic Data
- Work zone activities
- Traffic Speed
- Roadway Geometry
- Traffic Incidents

Data Preprocessing
- Training Set
- Validation Set
- Test Set

Machine Learning Model

Evaluation

Predicting
- Predict Work Zone Impact
- Work zone Traffic Management Planning

Training

Future Work Zone Project
Case Studies For Prediction Purposes

- 1,160 work zone projects occurring on Michigan Interstates from 2014 to 2017
- Including single-lane and double-lane closures
- Lane-closures were in place at least for one day to maximum 15 days
Mobility Prediction Using Classification Algorithms

Segment’s speed data discretization

<table>
<thead>
<tr>
<th>Speed (mph)</th>
<th>Class Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20</td>
<td>A</td>
</tr>
<tr>
<td>20-40</td>
<td>B</td>
</tr>
<tr>
<td>40-60</td>
<td>C</td>
</tr>
<tr>
<td>60-80</td>
<td>D</td>
</tr>
</tbody>
</table>

Classification algorithms

- XGBoost
- Random Forest
- Artificial Neural Network

Model Training & Validation

- Cross Validation
- Resampling

Evaluation

- Balanced Accuracy
How To Preprocess The Data To Improve Prediction Performance?

- Interstate speed data set is highly imbalanced.
- Far more high-speed records were present compared to low-speed records
- This can confuse training algorithms to predict records from minority classes
Resampling Techniques To Balance The Dataset

Algorithms used to resample data:

• Random Under Sampling
• Over-Sampling
• SMOTE: Synthetic Minority Over-sampling

![Diagram showing under-sampling and over-sampling techniques.](image)
Speed Heat-maps From Predicted And Actual Observation

Actual observation

Prediction
# Artificial Neural Network Architecture

## Input Attributes

<table>
<thead>
<tr>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highway Segment’s Characteristics</td>
</tr>
<tr>
<td># Of Available Lanes, Speed Limit,</td>
</tr>
<tr>
<td>Closed Lanes/# Of Lanes, Segment Start/End Point’s</td>
</tr>
<tr>
<td>Distance To Lane Closure Start/End Locations</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Work Zone &amp; Traffic Features</td>
</tr>
<tr>
<td># Of Closed Lanes, Type Of Closure, Closure Duration,</td>
</tr>
<tr>
<td>AADT, Road Functional Class</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Probe data Extraction</td>
</tr>
<tr>
<td>Historic Speed Distribution, Month Of Year, Day Of</td>
</tr>
<tr>
<td>Week, Hour Of Day, Time Duration After Lane Closure</td>
</tr>
<tr>
<td>Started, Time Duration Remaining Till The End Of</td>
</tr>
<tr>
<td>Lane Closure, Speed Distribution</td>
</tr>
</tbody>
</table>

## Input Layer

\[
\text{Input Layer} \in \mathbb{R}^{83}
\]

## Hidden Layers

- Hidden Layer \( \in \mathbb{R}^{30} \)
- Hidden Layer \( \in \mathbb{R}^{40} \)
- Hidden Layer \( \in \mathbb{R}^{50} \)

## Output Layer

\[
\text{Output Layer} \in \mathbb{R}^{1}
\]

## Speed Range

- For Segment \( i \) At Time \( j \)
## Model Performance Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Speed range (mph)</th>
<th>Evaluation Metrics</th>
<th>Balanced Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Precision</td>
<td>Recall</td>
</tr>
<tr>
<td><strong>Random Forest</strong></td>
<td>0-20</td>
<td>0.8</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>40-60</td>
<td>0.66</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>60-80</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Macro average</td>
<td>0.76</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Micro average</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td><strong>XGBoost</strong></td>
<td>0-20</td>
<td>0.78</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>40-60</td>
<td>0.67</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>60-80</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>Macro average</td>
<td>0.77</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>Micro average</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td><strong>ANN</strong></td>
<td>0-20</td>
<td>0.85</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>0.79</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>40-60</td>
<td>0.77</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>60-80</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Macro average</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Micro average</td>
<td>0.92</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Source: [https://en.wikipedia.org/wiki/Precision_and_recall](https://en.wikipedia.org/wiki/Precision_and_recall)
State-wide Mobility Assessment & Management

Develop a high-level decision-making process to better plan future work zones.
### WMZA For Individual Work Zones

#### Single Work Zone Mobility Measurement

<table>
<thead>
<tr>
<th></th>
<th>AM</th>
<th>Mid</th>
<th>PM</th>
<th>Weekend</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avg Delay (min)</strong></td>
<td>0.0</td>
<td>5.9</td>
<td>4.0</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Max Delay (min)</strong></td>
<td>5.2</td>
<td>22.8</td>
<td>26.7</td>
<td>3.3</td>
<td>26.7</td>
</tr>
<tr>
<td><strong>Total Delay (min)</strong></td>
<td>183.7</td>
<td>3585.4</td>
<td>2111.4</td>
<td>107.3</td>
<td>6151.7</td>
</tr>
<tr>
<td><strong>LOTTR</strong></td>
<td>1.3</td>
<td>1.5</td>
<td>1.9</td>
<td>1.0</td>
<td>1.7</td>
</tr>
</tbody>
</table>

#### Delay & LOTTR Metrics

#### Queueing Metrics

<table>
<thead>
<tr>
<th></th>
<th>Queue</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Max Duration (min)</strong></td>
<td>580.0</td>
</tr>
<tr>
<td><strong>Total Duration (min)</strong></td>
<td>2395.0</td>
</tr>
<tr>
<td><strong>Max Length (miles)</strong></td>
<td>3.3</td>
</tr>
<tr>
<td><strong># of Queues</strong></td>
<td>20.0</td>
</tr>
</tbody>
</table>

#### Spatial Characterization

#### Temporal Characterization

#### Delay & LOTTR Metrics

<table>
<thead>
<tr>
<th></th>
<th>AM</th>
<th>Mid</th>
<th>PM</th>
<th>Weekend</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avg Delay (min)</strong></td>
<td>0.5</td>
<td>0.3</td>
<td>2.6</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td><strong>Max Delay (min)</strong></td>
<td>32.5</td>
<td>11.6</td>
<td>19.9</td>
<td>5.2</td>
<td>72.5</td>
</tr>
<tr>
<td><strong>Total Delay (min)</strong></td>
<td>1330.0</td>
<td>1130.8</td>
<td>2941.0</td>
<td>243.1</td>
<td>6075.0</td>
</tr>
<tr>
<td><strong>LOTTR</strong></td>
<td>1.2</td>
<td>1.3</td>
<td>1.5</td>
<td>1.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

#### Queueing Metrics

<table>
<thead>
<tr>
<th></th>
<th>Queue</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Max Duration (min)</strong></td>
<td>185.0</td>
</tr>
<tr>
<td><strong>Total Duration (min)</strong></td>
<td>2395.0</td>
</tr>
<tr>
<td><strong>Lot Length (miles)</strong></td>
<td>6.8</td>
</tr>
<tr>
<td><strong># of Queues</strong></td>
<td>39.0</td>
</tr>
</tbody>
</table>
Selected Work Zone Projects:

- 2014 – 638 Work Zones
- 2015 – 601 Work Zones
- 2016 – 535 Work Zones
- 2017 – 344 Work Zones

More than 1,700 case studies
Work Zone Case Studies

- 1,705 work zone case studies from 2014 to 2017
- Shoulder to multiple lane closures
- One to 15 days
Ranking Interstates Based On Mobility Impact

- Identifying interstates with the highest impact on mobility.
- More information on how different work zone categories impacted mobility.
- Useful for budget allocation and high-level planning.
**Significant Projects?**

**Ranking Work Zones Using Pareto Principle**

- Determines 20% of work zones which accounted for 80% of the overall impact.
- These projects can be considered “significant” projects.
- Agencies could prioritize these work zones to improve their mobility management.
Work Zone Mobility Dashboard

Overall Impact

Median Impact

Pareto sort: significant projects

Impact Distribution

Relative Impact
Work Zone Mobility Dashboard

Overall Impact

Pareto sort: significant projects

Median Impact

Impact Distribution

Relative Impact
Making Decisions Based On Data!

Research questions:

• What is the relationship between work zone characteristics and its impact on mobility?
• Which work zone strategies work more efficiently?
• What are significant factors effecting mobility performance?
• Can we develop decision rules based on data?
Mobility Metrics
(dependent variables)

- Total Work zone delay (normalized: hour per day)
- Total Queue duration (normalized: percent of time performing in queue condition per day)
- Number of queue (normalized: per day)

Work Zone Characteristics
(independent variables)

- Work zone category (shoulder to multiple lane closure)
- Roadway ()
- AADT
- CAADT
- Closure side (Left-closure or right-closure)
- Duration (intermediate or long-term)
- Day of week (work zone starts)
- Day of week (work zone ends)
- Month of year
Decision-making Based On The Delay Metric

- Determines how many hours of delay is expected for each work zone category.

- Work zone category were the statistical significant factor splitting work zones along with closure-side and interstates.
Decision-making Based On The Queue Frequency Metric

- Determines how many queues would form as a result of work zone.
- Interstates were the statistical significant factor splitting work zones along with work zone category and AADT.
Provided Methodologies To Address The FHWA Call

Developed a machine learning method to assist in planning future lane-closures.

Developed a Business Intelligence dashboard to assess mobility in a state-wide level.

Developed a scalable framework to audit mobility performance measures for individual lane-closures.
Work Zone Mobility Audit Framework

Link to the “Auditing Work Zone Mobility Using Probe Vehicle Data” document
Link: https://www.workzonesafety.org/publication/auditing-work-zone-mobility-using-probe-vehicle-data/

To access to the source code of the WZMA tool on GitHub:
Link: https://github.com/WSUTRG/WorkZone_Mobility_Audit

2. FHWA Work Zone Data Exchange: [https://www.transportation.gov/av/data/wzdx](https://www.transportation.gov/av/data/wzdx)


Work Zone Mobility Audit

Comments:

Work Zone Typical Traffic
Travel Time (min)

Work Zone Speed Heatmap

Weekday AM
Weekday Mid
Weekday PM
Weekend

Work Zone Travel Time Reliability

Mile Marker

Work Zone Travel Time vs Typical Traffic

Overview

Information

Knowledge

Wisdom

Data Pool

WSP