Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
RES#2016-06		

4. Title and Subtitle	5. Report Date
Traffic Queue Prediction and Warning System	August 2019
	6. Performing Organization Code
7. Author(s) Yuandong Liu, Bumjoon Bae, Lee D. Han, and Asad Khattak	8. Performing Organization Report No.
9. Performing Organization Name and Address University of Tennessee, Center for Transportation Research 309 Conference Center Building Knoxville, TN 37996-4133	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. RES2016-06
12. Sponsoring Agency Name and Address Tennessee Department of Transportation 505 Deaderick Street, Suite 900 Nashville, TN 37243	13. Type of Report and Period Covered Final Report Oct 1, 2015 -June 30, 2019
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The purpose of the study was to explore and develop analytical models and field-data based algorithms for the identification or detection of the "end-of-queue" (EOQ) in freeway incident or congestion situations. TDOT's real-time ITS traffic data was the primary basis for this study, so the algorithms could be verified in the field and deployed in the future using the same database. The crowdsourced WAZE database was also included in this study to afford TDOT the flexibility of managing incident cases outside of the urban TMC/RDS data coverage areas. The study reviewed the state-of-the-practice of real-time queue length and end-of-queue identification methodologies, evaluated the suitability of real-time traffic data sources as potential input, and developed queue detection and prediction algorithms based on these data sources. The study also assessed the risks associated with the end-of-queue crash prediction model and evaluated means for warning the public of a non-recurring/unexpected queue situation, particularly outside of the major urban areas. Based on the means for warning the public, some implementation strategies were identified. This study developed a technical framework that can detect and predict the EOQ location, which is a crucial step towards protecting the queue. The system developed in this study dynamically detects the EOQ and predicts its movement in spatiotemporal domains based on real-time traffic data using traffic flow models. One of the biggest benefits of this endeavor is crash prevention. The system can proactively manage the queue, with focus on non-recurrent events, and reduce rear-end collision risks. Timely dissemination of the EOQ information could effectively slow down approaching drivers, divert drivers further upstream, and, hence, reduce the safety hazards of non-recurring events. Potential implementation strategies for warning the motoring public include infrastructure-based devices, such as changeable message signs, and vehicle-based mechanisms, such as WAZEtype navigation apps and DSRC/5G communications. In the era of connected and automated vehicles (CAV), more effective EOQ detection and warning systems could become a standard feature for cars as well as freeway operation centers.

17. Key Words	18. Distribution Statement
END-OF-QUEUE, EOQ, QUEUE PROTECTION, CROWDSOURCING DATA, INCIDENT MANAGEMENT, INCIDENT DETECTION, WAZE DATA	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	58	$113,420.34

FINAL TECHNICAL PROJECT REPORT

TRAFFIC QUEUE PREDICTION AND WARNING SYSTEM

Project #: *IG1858992*

RES#2016-06

UT#R01-1313-554

Submitted to

TDOT Research Development and Technology Program

by

Yuandong Liu, Bumjoon Bae, Lee D. Han, and Asad Khattak

of

Center for Transportation Research
The University of Tennessee

Submitted: August 2019 // Final Revision: October 2020

DISCLAIMER

This research was funded through the State Planning and Research (SPR) Program by the Tennessee Department of Transportation and the Federal Highway Administration under RES#2016-06 Traffic Queue Prediction and Warning System.

Acknowledgement

The authors are grateful for the sponosorships, the many research ideas, and the guidance in the development of this final report from Tennessee Departmnt of Transportion and Fedearl Highway Administration. The work here was presented at TRB Annual Conference and was published in Transportation Research Part B in 2019.

Executive Summary

The purpose of the study is to explore and develop analytical models and field-data based algorithms for the identification or detection of the "end-of-queue" (EOQ) in freeway incident or congestion situations. The Tennessee Department of Transportation's (TDOT's) real-time Intelligent Transportation System (ITS) traffic data is the primary basis for this study so that the algorithms could be verified in the field and deployed in the future using the same database. Crowdsourced WAZE database information was also included in this study to afford TDOT the flexibility of managing incident cases outside of the urban TMC/RDS data coverage areas.

The study reviewed the state-of-the-practice of real-time queue length and end-of-queue identification methodologies, evaluated the suitability of real-time traffic data sources as potential input, and developed queue detection and prediction algorithms based on these data sources, which was the primary objective of this endeavor. The study also assessed the risks associated with the end-of-queue crash prediction model and evaluated means for warning the public of a nonrecurring/unexpected queue situation, particularly outside of the major urban areas. Based on the means for warning the public, some implementation strategies were identified.

This study developed a technical framework that can detect and predict EOQ location, which is a crucial step towards protecting the queue. The system developed in this study dynamically detects the EOQ and predicts its movement in spatiotemporal domains based on real-time traffic data using traffic flow models. One of the biggest benefits of this endeavor is crash prevention. The system can proactively manage the queue, with focus on non-recurrent events, and reduce rear-end collision risks. Timely dissemination of the EOQ information could effectively slow down approaching drivers, divert drivers further upstream, and, hence, reduce the safety hazards of non-recurring events. Potential implementation strategies for warning the motoring public include infrastructure-based devices, such as changeable message signs, and vehicle-based mechanisms, such as WAZE type of navigation apps and DSRC/5G communications. In the era of connected and automated vehicles (CAV), more effective EOQ detection and warning systems could become a standard feature for cars as well as freeway operation centers.

Table of Contents

List of Figures

List of Tables

Introduction

Tennessee Department of Transportation (TDOT) is committed to ensuring highway safety and roadway efficiency. In addition to maintaining its roadway infrastructure, TDOT is heavily invested in the operational aspects of traffic under normal as well as incident conditions. These investments include, but are not limited to, traffic management centers (TMC) in the State's four major metropolitan areas, the incident management program and HELP truck fleets, the thousands of highway sensors and CCTV cameras, and many dozens of variable messages signs (VMS) in each metropolitan area.

In 2018, some 125,488 roadway incidents, such as crashes, disabled vehicles, and debris, were identified, reported, and managed by TDOT's Incident Management Program, just in the four major metropolitan areas alone. Many of these incidents, as well as work zones and special events, can lead to long and unexpected queues that cause not only delay and inconvenience but also a hazardous situation at the tail end-of-queue (EOQ) that can catch unsuspecting motorists by surprise as they approach at high speed upon the slow, or even stopped, traffic suddenly.

The aims of this study are, therefore, to leverage TDOT's existing real-time traffic data sources and developing an end-of-queue prediction and warning system that can dynamically predict the queue location and warn the motoring public of the potential dangers. By studying different cases with comprehensive data, incident logs, and video footage, the UT research team developed a model capable of taking in real-time data to detect and further predict the "end-of-queue" (EOQ) location and movement dynamically.

While the concept of EOQ is intuitively not difficult to understand, there is no consistent definition of this term in literature or practice. The term "queue" has been defined in various ways in the literature. Highway Capacity Manual 2010 [1]defines a queue as "a line of vehicles waiting to be served" in a system and a "queued state" as "a condition when a vehicle has slowed to less than 5 mph". Stephanopoulos et al.[2] define "queue length" for an intersection as "the length of the roadway section behind the stop line where traffic conditions range from the capacity to jammed density" in a flow-density diagram. In spite of the different and insufficient queue definitions for freeway facilities in the literature, a common condition is that a queue is formed when the system demand exceeds its capacity[2]. It is difficult to measure the traffic demand directly from traffic flow data when the flow is at or near capacity at a bottleneck. However, one can infer the presence of excessive demand if high densities and low speeds are observed upstream of the bottleneck[2]. In traffic flow theory, a breakdown is the transition from uncongested to congested flow and observed as a speed drop occurring with queue formation[1]. The speed drop, or breakdown point/location, is also hazardous to drivers. Therefore, for this project, to detect and predict the EOQ location is to identify and predict the traffic phase transition over wide spatiotemporal domains. More results on the current state of this matter will be presented in Chapter II, which will be followed by an assessment of various data sources TDOT has at its disposal for this study in Chapter III.

Chapter IV presents detailed descriptions of the methodology developed for queue identification and prediction. Two main algorithms were developed in this study. One algorithm utilizes TDOT Remote Traffic Microwave Sensor (RTMS) data system, which covers all four major urban areas in Tennessee. The other algorithm utilizes WAZE traffic jam reports (more details in the latter part of this report) and identifies the congested regions, as well as predicts its transition borders. For roads, especially interstate highways, not covered by RTMS, the second methodology can be deployed to estimate and predict the EOQ.

Chapter V proposes a methodology to assess the risks associated with EOQ and provides TDOT a better understanding of the hazard of the EOQ. This study also assesses quick information dissemination strategies for warning motorists approaching the vicinity of the end-of-queue. The state-of-the-practice review as well as TDOT current practice are comprehensively evaluated in Chapter VI. Another focus of this study looks at the location and setup of the HELP trucks upstream of the end-of-queue. Chapter VII present a conceptual framework implementing the detection, prediction, and management of EOQ scenarios.

Chapter I. State of the Practice Review

This section reviews the queue prediction strategies. Queue warning and protection strategy practice are further discussed in Chapter VI. Previous studies focused on estimating vehicular queue length at a signalized intersection for traffic signal performance measures or signal optimization since queue length is an important factor to be managed for intersections. They can be classified into two major categories: 1) cumulative traffic input-output approach and 2) traffic shock wave approach.

The cumulative traffic input-output approach describes the maximum queue length based on cumulative demand and capacity curves over time, while the shock wave approach estimates queue lengths by identifying shock wave speeds over time.

Input-Output Approach

As we expect, queue length is a function of traffic demand and capacity. The model proposed initially by Webster [3] calculated the time of the queue dissipation and effective queue size by input-output analysis. After the start-up lost time from the onset of the green signal, the vehicles in a queue are discharged at saturation flow and after the onset of a red signal, another queue grows based on an assumed arrival rate. Since the effective queue size is defined as the number of cars in the queue waiting for service at an instant in jam density [4], a constant average density throughout a cycle is assumed in the range between the congested and capacity. However, it has been pointed that density is time-varying within a cycle and the assumption of constant average density can lead to miscalculation of the effective queue size [4].

Sharma et al. [5] used two input-output models. One is a simple model in which only an advance detector is used to track vehicle arrivals and another model uses both advance and stop bar detectors to utilize the headway information too. The root mean squared error of both models was less than 15% of the length of a vehicle for average maximum queue length by evaluation with field data.

These models could not estimate queue lengths or would produce inaccurate estimation results when the EOQ extended beyond the detector because arriving vehicles would not be detected [6].

Shock Wave Approach

Lighthill and Whitham [7], and Richards [8] explained traffic flow phenomena on the basis of shock wave theory using a theoretical fundamental diagram, called LWR theory. In their model, the flow rate is assumed as a function of the vehicle density [7]. Although the shock wave theory is derived from the law of conservation of the number of vehicles on the road, which accounts for the traffic flows into and out from a roadway segment, the queue length estimation models in this catagory use the shock wave speed directly.

Kernel [9] asserted that the LWR theory cannot explain some empirical traffic flow phenomena, including: a probabilistic speed breakdown occurring spontaneously at a bottleneck due to an internal local disturbance in traffic flow (i.e. transition from free flow to synchronized flow); and a self-organizing congested pattern which consists of synchronized flow upstream at a bottleneck with wide moving jams in it.

Geroliminis and Skabardonis [10] proposed an analytical model for predicting platoon arrival profiles and queue length along signalized arterials. They employed a Markov decision process to

model traffic dispersion behaviors between successive signal intersections and then used shock wave speeds based on the LWR theory to estimate queue lengths. In comparison with simulated data, the difference of predicted queue length was less than four vehicles.

Liu et al. [6] proposed a real-time queue length estimation method for congested signalized intersections using event-based signal and vehicle detection data. They applied LWR shock wave theory to identify break points where traffic flow states change at a loop detector location. Then, the maximum queue length can be estimated at the intersecting point of a discharge and departure shock wave speed.

Location-Based Information Approach

Recent efforts for estimating queue lengths in real-time employ the location information of probe vehicles in a queue. Comert and Cetin [11] proposed a conditional probability model to estimate the expected queue length and its variance. Based on the assumption that the marginal probability distribution of queue length is known, and the vehicle arrivals follow the Poisson distribution, they found that the location information of the last probe vehicle in a queue is sufficient for queue length regardless of the market penetration of probe vehicles. However, the finding is limited since it is based on a priori knowledge of the marginal distribution and derived for undersaturated conditions. Ban et al. [12] estimated the maximum and minimum queue lengths by detecting critical pattern changes of intersection travel times or delays based on the GPS log information.

Although these studies can also be classified as the shock wave approach, using the location information of an individual vehicle can be distinguishable from the earlier studies where fixed location sensor data were used.

Conclusion

Although the previous studies have mostly focused on the estimation of queue lengths for a signalized intersection with a single link, these estimation approaches can be employed for this project, i.e. uninterrupted traffic flow. Traffic queue occurs due to traffic signals for a signalized intersection, while it occurs due to a traffic incident or natural bottleneck for interstate highways. If there are fixed traffic sensors, such as loop detectors, successively in a study area, the input-output approach can be applied. The growing queue can be detected over time by using multiple detectors in upstream. The shock wave theory can also be used in the same sense. By capturing shock wave speeds for successive detector stations, the location of queue end can be estimated collectively. If individual vehicle trajectory data are available in real-time for the highway where the detectors are deployed, the estimation result can potentially be improved and/or validated.

The expected challenges for each approach are as follow:

- Input-Output: Since multiple highway links – the roadway segments between two detector stations – should be considered for this project, calculating accurate inflow and outflow of a study area would be difficult due to on-/off-ramp flows and limitations on spatial coverage and temporal resolution of detector data, e.g. 30 seconds.
- Shock wave: This method uses a traffic flow, q, and density, k relationship. In general, a q-k relationship is estimated linearly as a concave line so that shock wave speed is calculated by selecting two single points on the line. This cannot reflect the variance or probabilistic phenomena in the q-k relationship, which is, in theory a smooth curve but in practice, a collection of stochastically distributed points with much randomness. In addition, unlike

the signalized intersection case where one of both traffic states for a shock wave speed is the jam density, a traffic incident or bottleneck may often not be a complete stop of the flow or complete blockage of all lanes. The error in shock wave speed estimation may produce significantly inaccurate queue length.

- Location-Based Information: This type of data is usually unavailable, particularly for real-time traffic analysis.

Chapter II. Real-Time Data Source Assessment

Using the proper data source is of crucial importance to detecting and predicting the end-of-queue in real-time. The data not only have to satisfy operational objectives in identifying and predicting the queue but also fit within the means, in terms of cost and technology capability parameters, of the agency. To this end, we assessed multiple traffic data sources that Tennessee Department of Transportation (TDOT) is already invested in. Not surprisingly, these data sources have different data fields, road network coverages, time intervals, limitations, and so on.

RTMS/RDS/Active ITS

Remote Traffic Microwave Sensor (RTMS) is a type of presence-detecting sensor. These sensors are mounted on poles adjacent to the roadway as in Figure II-1. When vehicles pass the side-fired microwave radar zone on the roadway, the RTMS system measures traffic volume, speed, and occupancy for each lane. For the side-firing configuration, it is known the sensor can monitor up to 8-12 lanes of traffic [13]. RTMS is insensitive to inclement weather and requires relatively low cost to install. Its accuracy is not as good as inductive loop detectors.

Figure II-1 Side-mounted configuration of RTMS

TDOT has installed the RTMS sensors approximately every 0.5 miles on the interstate highways in four metropolitan areas including Knoxville (see Figure II-2), Chattanooga, Nashville, and Memphis. From each RTMS station, the traffic data (traffic count, speed, and occupancy) are collected and transmitted to the traffic management center (TMC) of each region every 30 seconds. The data collecting time interval, or resolution, of RTMS is the highest compared to the other data sources.

Real-time traffic data are essential for advanced traffic management and traveler information systems. For example, incident management strategies such as automatic incident detection as well as end-of-queue detection and warning, require traffic speed and occupancy updated every 20 to 30 seconds [14]. Therefore, RTMS data could be the best data source for EOQ detection and prediction. This, however, does not necessarily imply or dictate that the time resolution for end-of-queue detection should be 20 or 30 seconds. Depending on the requirements of the algorithm to be developed in this study and TDOT's operational goals, the time resolution of the data may vary. Figure II-3 shows an example of traffic speed using the RTMS data for I-40 eastbound segment between exit 374 and exit 388 on August 4, 2016.

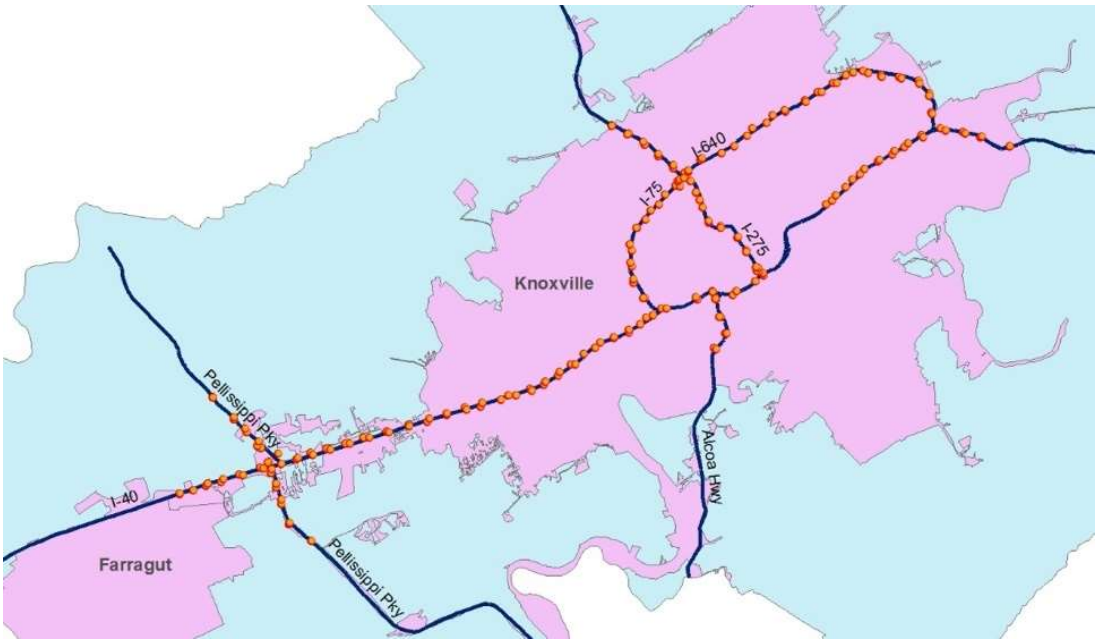

Figure II-2. RTMS stations in Knoxville

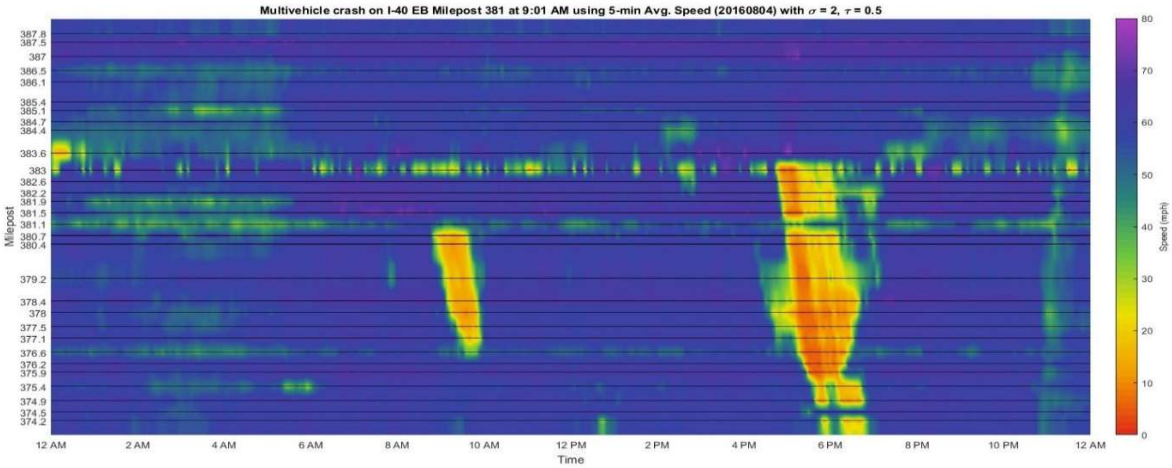

Figure II-3. Speed Heatmap of I-40 EB (374- to 388-mile segment) using RTMS

The main limitation of the RTMS is the missing data issue. In general, all traffic sources we assess will experience data loss of different natures and different magnitudes at different times. Some of the field sensors have deteriorated or even ceased to function over time. Since most of the remotely deployed sensors rely on solar panel charged batteries that lose capacity over time, data loss is likely to occur during nighttime, in winter months, and during utility construction projects. Figure II-4 shows an occurance of data missing on an I-24 segment between exit 54 and exit 62 in Nashville.

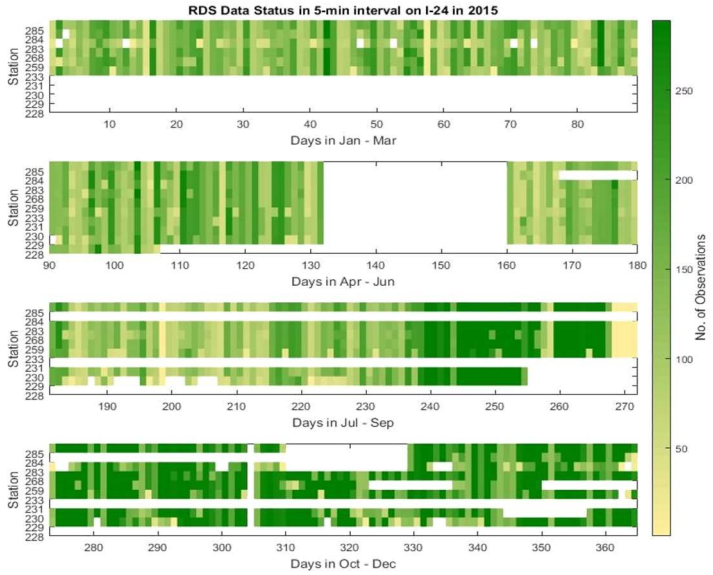

Figure II-4. Four cases of missing values of RTMS data for I-24 in 2015

HERE

HERE is a system of speed data provided by a commercial vendor. TDOT has a contract with the vendor, HERE, to download the speed data of roadway links in the state of Tennessee for a limited time period. HERE provides link speed data and it includes the average speed, standard deviation, and the 10th-90th percentile speeds of each link. The sources of HERE speed data are state installed roadway sensors and GPS information of probe vehicles. HERE covers urban streets in the state of Tennessee, as well as the Interstate highways.

TDOT has a license to download historical, i.e., not real-time, HERE data aggregated in 5 minutes or longer than 5 minutes. This means the use of HERE data for real-time operational purposes is greatly restricted. In addition, the data does not include traffic volume or occupancy variables, which are sometimes required for implementing automatic incident detection or end-of-queue detection algorithms. As such, HERE is not an appropriate data source for real-time traffic management implementations.

WAZE

WAZE is a commonly used navigation smartphone app that enables drivers to obtain and share real-time traffic and road information. Similar to other navigation apps, WAZE actively collects the user's GPS information, with permission, and provides travel time, speed, routing, and roadway information. WAZE has developed a website for TDOT to monitor the traffic conditions on the main interstate they designated in each of the four regions in Tennessee. The travel time of each road link in the monitored interstates is updated every one to two minutes. Since March 2017, WAZE allows its Connected Citizen Program (CCP) partner to add new route segments and provides travel time on these segments. This functionality affords TDOT more flexibility in managing dynamic traffic conditions.

WAZE app users can report accidents, traffic jams, work zones, etc. to WAZE platform right from their smartphone. WAZE then disseminates the reports to all users in the vicinity. TDOT

partnered with WAZE through CCP in 2016. WAZE offers live feeds of traffic logs with nine types of reports that come directly from the users (see Figure II-5) in the state of Tennessee via a real-time download link.

]

Figure II-5. WAZE traffic report types

These data and logs are currently archived by the University of Tennessee research team at minute-by-minute basis. Event logs are useful supplements for traffic incident detection and EOQ detection purposes. For locations where RTMS sensors are not deployed or not functional, incident and traffic jam reports from WAZE can provide insightful information. Figure II-6 demonstrates the spatial distribution of WAZE jam reports in Knoxville area**.** Each red dot in Figure III-6 represents an incident reported by a WZAE app user while traveling through the affected link segment or a jam report generated by Google's algorithm. The green "blobs" around the red dots, and hence the highway network, are meant to represent the density of these incident and jam reports. Darker and greater green blobs are resultant from more frequent congestions and incidents on these highway segments.

Figure II-6. The spatial distribution of WAZE jam reports, Knoxville

Traffic Cameras

TDOT provides live feeds (15 fps, 320x240) of approximately 500 traffic surveillance cameras on its Interstate highways in the metropolitan areas of the state via the *TDOT SmartWay* [15] website (see Figure II-7). The main use of the traffic cameras is to monitor traffic conditions in real-time by TMC staff. TDOT TMCs can rotate, tilt, zoom in and zoom out each camera.

The video footage from the traffic cameras can be the best resource for the validation of the EOQ detection and prediction algorithms. For that purpose, however, there are challenges in the current TDOT system. First, the live feeds of the video footage are not recorded. Because the current usage is mostly for traffic monitoring in real-time, the TMC discards the footage instead of archiving it for potential research or training purposes. Second, the current system does not timestamp the live feeds[1]. Depending on the conditions of various links in the network communication chain, time lags exist between the time when the videoed event transpires and the time it is viewed "live". The estimated lag ranges from a few seconds to several minutes. Thus, the time lag might become a significant issue on the precision of the EOQ detection and prediction models.

Figure II-7. Example of video footage

The video footage from the traffic cameras can be potentially used for detecting the back end-of-queue directly by combining with image processing techniques. Figure II-8 shows an attempt by UT research team with motion detection techniques on some video footage from *SmartWay*. Similar to how video detectors work, motion detection or other image processing techniques can measure traffic volume, speed, and occupancy. However, it is not feasible to read and analyze individual image frames for over 500 cameras in real-time. Thus, it is better for the image processing technique to cooperate with automatic incident detection algorithms, implemented using other data sources such as RTMS or WAZE. For example, whenever an incident is detected or reported on a certain roadway segment, the image processing module can then be automatically triggered and, subsequently, implemented on the cameras in close proximity.

[1] The time stamps in Figure II-7 were inserted by UT research group and they indicate the time when the live feeds were displayed through the *TDOT SmartWay* website and not necessarily the actual time when the event transpired.

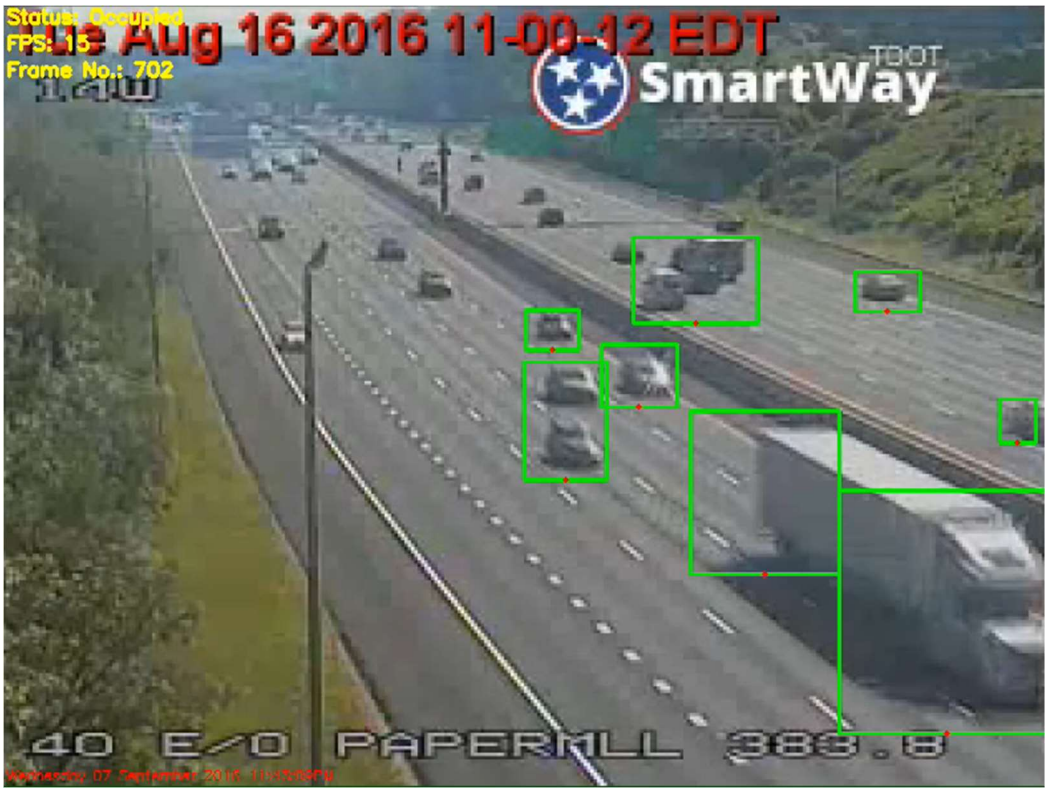

Figure II-8. Sample display of data extraction from TDOT SmartWay video footage

INRIX (NPMRDS)

INRIX supplies traffic data to the United States Federal Highway Administration (FHWA), state, and regional agencies to assess travel reliability, congestion, and emissions. National Performance Management Research Data Set (NPMRDS) is the default dataset for calculating the new US Federal's third performance management rulemaking (PM3) per 81FR23806 and freight performance measures. NPMRDS, delivered in partnership with University of Maryland at FHWA's expense, allows agencies to access massive amounts of planning-grade data from January 2015 forward on the National Highway System (NHS) road network at no cost.

INRIX provides 5-min speed data on monthly basis. In other words, the data is not available in real-time and thus is not suitable for real-time implementation in its NPMRDS offering. TDOT is currently in the process of purchasing a higher resolution(1-min) speed data from INRIX. If the data is available in real-time, it will be a potential data source for end-of-queue identification. For traffic operational needs, real-time high-resolution data sources are essential.

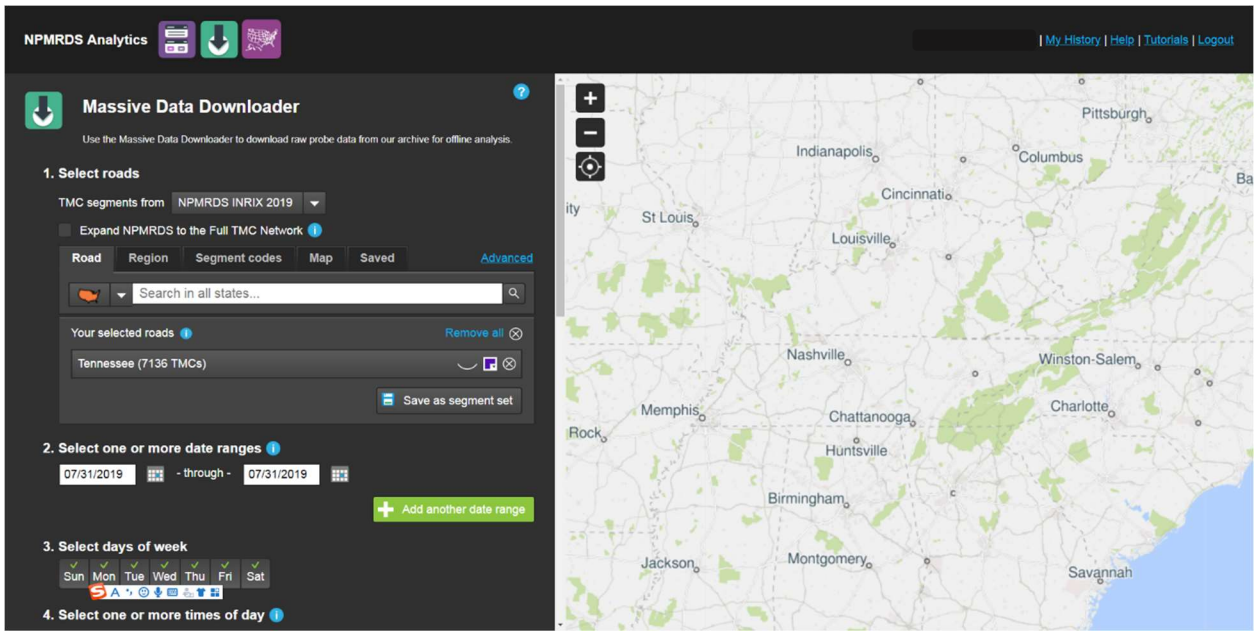

Figure II-9. NPMRDS Massive Data Downloader Interface

Summary

Table II-1 shows the summary of the data sources including (RTMS), HERE, WAZE, and traffic cameras operated by TDOT. As presented in this table, the data sources managed by TDOT have different characteristics in terms of coverage, time-resolution etc. This study aims to develop an automated EOQ detection and prediction algorithm which requires a data set and can feed into the algorithm. Among the five major data sources, RTMS, WAZE and Traffic camera supplies real-time traffic information and are suitable for real-time EOQ detection and prediction.

RTMS is selected as the main data source for EOQ detection, it has highest resolution among all data sources and covers major urban roads in four main regions in Tennessee. On roads not covered by RTMS, WAZE can be used as a complementary data source. Traffic camera data is used to validate the algorithm detection and prediction results.

Table II-1. Summary of Data Source Comparison

	RTMS	HERE	WAZE[16]	NPMRDS	Traffic Camera
Data Type	· Count · Speed · Occupancy	· Speed	· Speed · Incident alert	· Speed	· Video image
Source Type	· Infrastructure-based	· Vehicle-based	· Vehicle-based	· Vehicle-based	· Infrastructure-based
Data Source	· Roadside vehicle detectors	· State installed sensors, probe vehicles, GPS	· Crowdsource (GPS and incident reports from App users)	· probe vehicles, GPS	· Roadside traffic surveillance cameras
Network Coverage	· Major Interstate routes in urban areas	· Interstate routes and urban streets throughout the state	· User-defined road links	· National Highway System	· Major Interstate routes in the urban areas
Time Resolution	· 30 seconds	· 5 minutes	· 1-2 minutes	· 5-minutes	· 15 frames per second (fps)
Limitation	· Missing values · Limited network coverage	· Missing values · Low temporal resolution · No traffic count data · No real-time data available	· No traffic count data · Inaccurate incident alerts · No historical data available · Limited network coverage	· Missing values · Low time resolution · No traffic count data · No real-time data available	· Low image resolution · Large storage space required · Low image quality during nighttime · Variable camera view angle · Limited network coverage

Chapter III. Queue Detection and Prediction Algorithm

Traffic end-of-queue (EOQ) detection and prediction were the objectives of this study. The ability to identify the EOQ and predict its location over time allows TDOT to take prompt actions to warn approaching drivers of slowing traffic ahead. As such, this study developed two dynamic queue detection models that utilize real-time traffic data to estimate the behavior of the queue and the locus of the EOQ. Also, a queue prediction algorithm is developed based on the dynamic detection results to forecast the movement of the queue.

The EOQ detection models are separately developed using two different data sources, RTMS and WAZE. The first model uses RTMS speed data to distinguish congested traffic conditions from uncongested traffic conditions and identifies the current location of queue. Unlike Waze, RTMS provides lane-specific traffic data. So this methodology can be used to identify the queue location for a specific lane. The second model dynamically clusters WAZE jam reports generated by Waze users in real-time and detects the congested region.

End-of-queue Detection and Prediction Algorithm based on RTMS Data

This section presents an EOQ detection algorithm based on the RTMS data. The algorithm uses real-time traffic detector data and automatically identifies the congested region as well the EOQ. Based on the queue detection results, the shock wave speed of the queue can be computed. Assuming the traffic status remains the same in the next time period, we can predict the movement of queue based on the shock wave speed.

Framework

This algorithm uses traffic detector data collected from each detection station to identify the traffic flow phasewhich is classified as either congested or uncongested based on the station's unique flow-density pattern in the previous days. Congestion is detected in the spatiotemporal domain by using the phase identification results from multiple stations collectively along a highway. Figure III-1 illustrates the algorithm's logical flow framework for detecting congestion in the spatiotemporal domain.

The proposed queue detection and prediction algorithm consists of these steps:

- **Traffic flow phase identification:** For each station, traffic flow phase is initially identified as one of the following classes: 'uncongested,' 'transitional,' and 'congested' flow based on speed (see Figure III-2. **Three phases in a flow-density plot.**). In the literature, the threshold value for distinguishing uncongested and congested flows ranges from 20 to 40 mph [17; 18]. As thresholds for the initial phase identification, 45 mph and 15 mph were selected in this study to capture a wider range of speed scenarios, and to ensure the minimum sample size of three data points for each flow (i.e., 0-15 mph: congested flow, 15-45 mph: transitional flow, 45+ mph: uncongested flow). The distributions of congested flow and uncongested flow are estimated in a flow-density diagram using GMMs, or Gaussian mixture models, a category of probabilistic model with all its generated data poinits derived from a finite Gaussian distribution that has no known parameters. Then, each new input data point is classified by comparing the likelihood of each distribution.

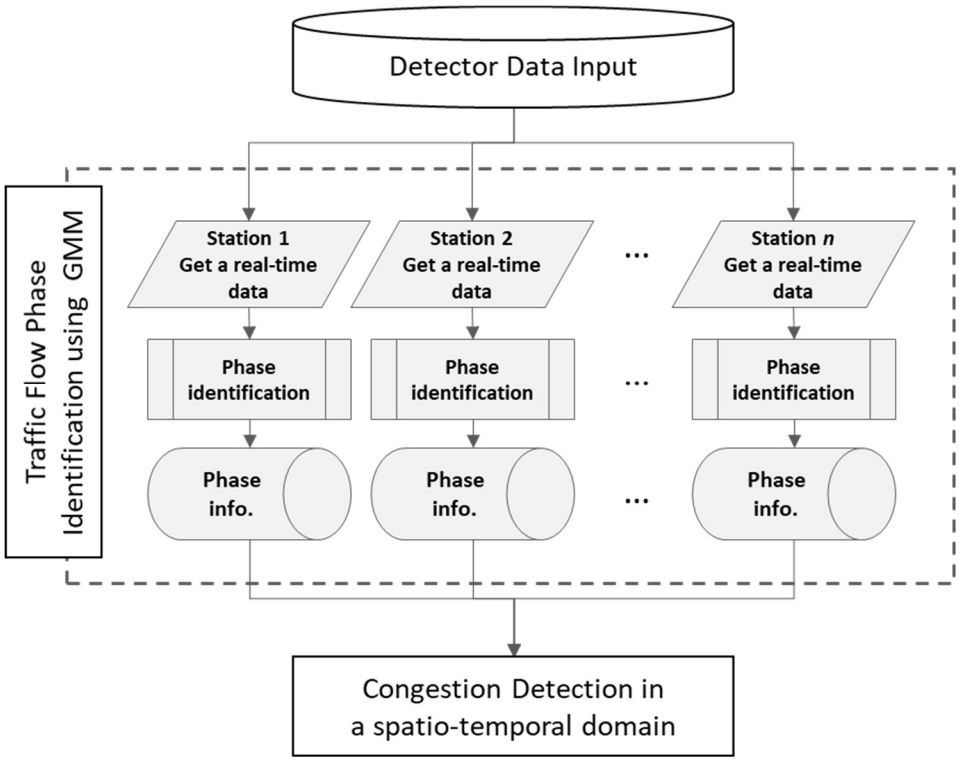

Figure III-1. Queue detection algorithm.

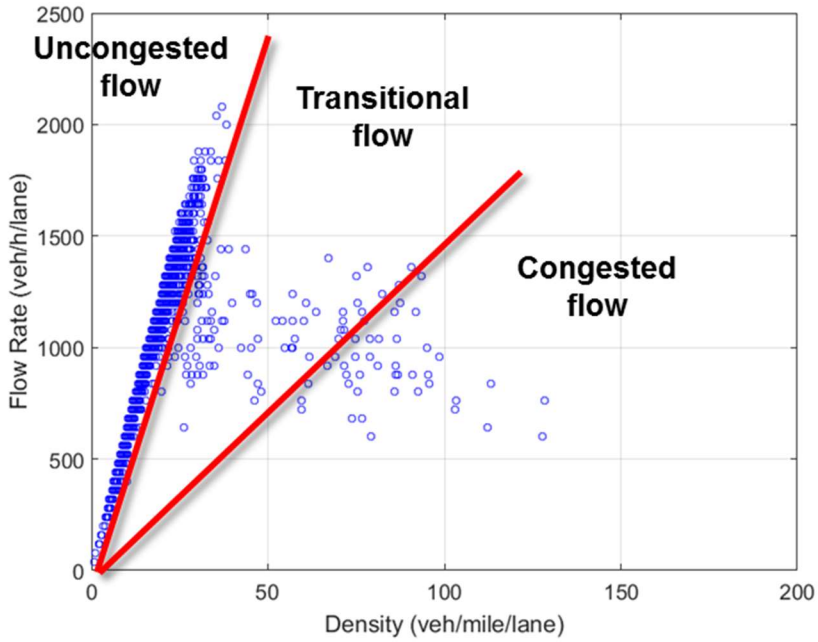

Figure III-2. Three phases in a flow-density plot.

- **Traffic congestion detection:** The phase information identified for each station in the previous step is used collectively to detect congestion occurrence at multiple locations and times. Figure III-3(a) shows a speed heat map as reference, generated by using detector data with an adaptive smoothing technique [19; 20]. Note that the horizontal lines in Figure III-3(a) represent the locations of the detector stations; Figure III-3(b) shows the examples of congested areas in the spatiotemporal domain, identified by the proposed algorithm. Each blue dot in the figure refers to the specific time and location of a detected congestion situation. To identify the boundary of the congestion and its evolution in the spatiotemporal domain, an even number of phase changes within a two-minute time window were filtered out (see Figure III-3(c)).

- **Shock wave speed computation:** The boundary of congestion detected in the previous step denotes the end-of-queue. Based on the time that the queue reaches each detector, the queue propagation speed between detectors can be computed.

Methodology

This section discusses the methodology used to identify traffic flow phase, detect traffic congestion, and calculate shock wave speed.

Gaussian Mixture Model (GMM)

Let $\boldsymbol{x_1}, \dots, \boldsymbol{x_N}$ denote a random sample of size n, where $\boldsymbol{x_j}$ is a p-dimensional random vector with the Gaussian distribution probability density function (pdf), $f(\boldsymbol{x_j})$. For a univariate random variable x, the pdf is

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\left(e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) \qquad \text{III-(1)}$$

where

μ is the mean of the random variable x,
σ is the standard diviatoin of the variable x,
e is the constant Euler's number (2.71828 …), and
$-\infty < x < \infty$, $-\infty < \mu < \infty$, and $\sigma^2 > 0$.

The normal density function can be represented as

$$f(\boldsymbol{x_j}|\mu,\Sigma) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x_j}-\mu)^T\Sigma^{-1}(\boldsymbol{x_j}-\mu)} \qquad \text{III-(2)}$$

where Σ is a covariance matrix that is positive definite, i.e., $\Sigma > 0$. The probability density function of data can be represented as a Gaussian mixture distribution, which is a linear combination of K Gaussian distributions (or components) with the set of parameters, $\boldsymbol{\Theta} = \{\alpha_{i=1\dots K}, \boldsymbol{\theta}_{i=1\dots K}\}$, for each as follows:

(a)

(b)

I

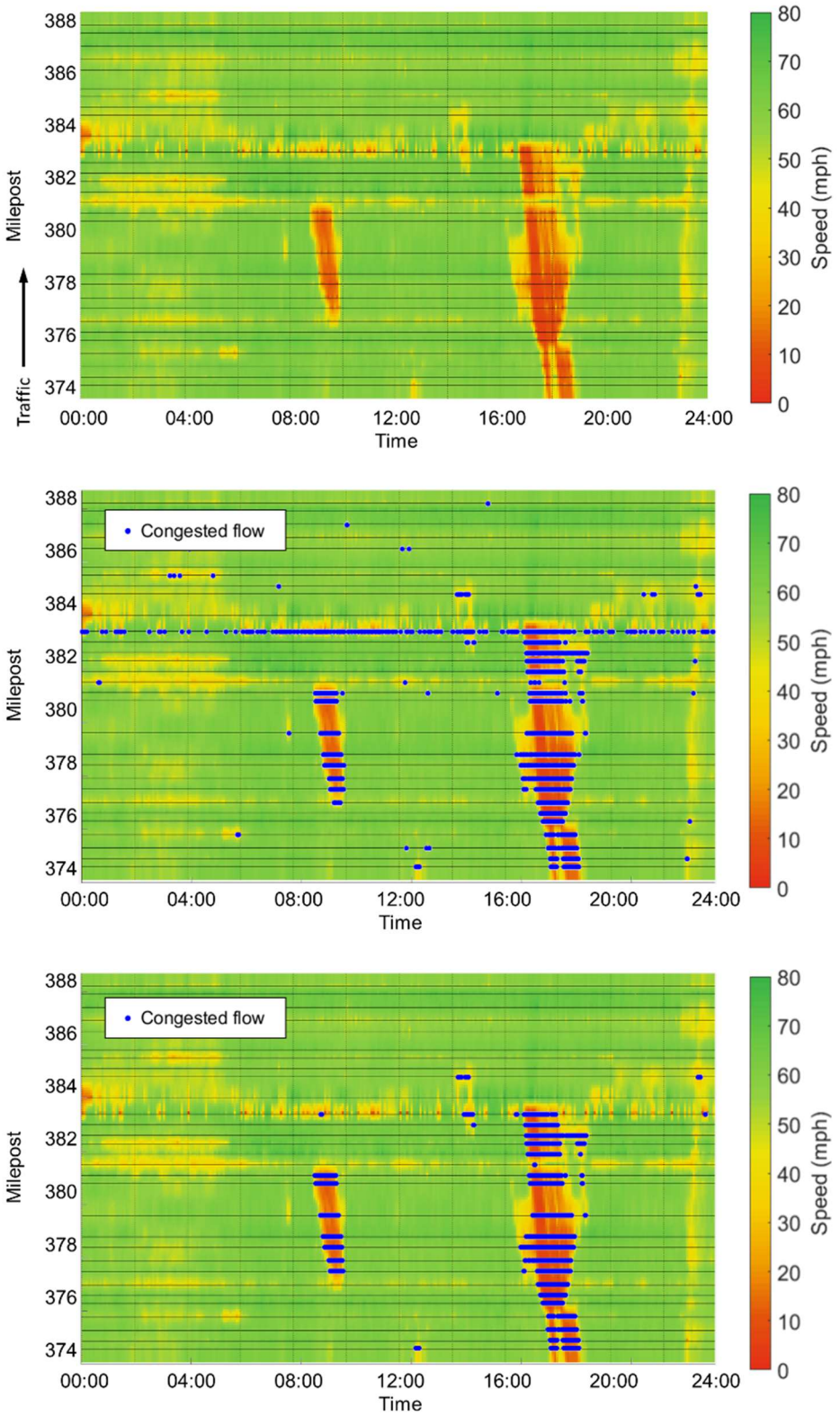

Figure III-3. An example of congestion detection (I-40 EB on August 4th, 2016): (a) speed heat map, (b) congestion detection without filtering, and (c) congestion detection with filtering.

$$f(\boldsymbol{x}_j|\boldsymbol{\Theta}) = \sum_{i=1}^{K} \alpha_i f(\boldsymbol{x}_j|\boldsymbol{\theta}_i) \qquad \text{III- (3)}$$

where, $f(\boldsymbol{x}_j|\boldsymbol{\theta}_i)$ is the Gaussian distribution with the ith parameter set $\boldsymbol{\theta_i} = \{\mu_{i=1\ldots K}, \Sigma_{i=1\ldots K}\}$, and α_i is the mixture weight of the ith component, which is nonnegative and sum to 1, that is

$$0 \leq \alpha_i \leq 1 \;\; (i = 1, \ldots, K)$$

and

$$\sum_{i=1}^{K} \alpha_i = 1.$$

The log-likelihood for $\boldsymbol{\Theta}$ is

$$\log L(\boldsymbol{\Theta}) = \sum_{j=1}^{N} \log f(\boldsymbol{x}_j|\boldsymbol{\Theta}) = \sum_{j=1}^{N} \log \left\{ \sum_{i=1}^{K} \alpha_i f(\boldsymbol{x}_j|\boldsymbol{\theta}_i) \right\} \qquad \text{III- (4)}$$

It is known that there is no closed form of the maximum likelihood estimation (MLE) for $\boldsymbol{\Theta}$ of the Gaussian mixture distribution. Therefore, the Expectation Maximization (EM) algorithm is frequently used to get the parameter estimates in GMM where the MLE is computed iteratively [21].

The EM algorithm consists of two steps: E for expectation and M for maximization.

E-step: Let $\boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{z})$ denote the complete data vector which consists of the observed data $\boldsymbol{x}$ and its posterior probability membership variable of the K components $\boldsymbol{z} = \{\boldsymbol{z_1}, \ldots, \boldsymbol{z_K}\}$, where each $\boldsymbol{z_i}$ is an N-length vector $[z_{i1}, \ldots, z_{iN}]^T$. The complete data log-likelihood for $\boldsymbol{\Theta}$ is

$$\log L(\boldsymbol{\Theta}|\boldsymbol{y}) = \sum_{i=1}^{K} \sum_{j=1}^{N} z_{ij} \log\{\alpha_i f(\boldsymbol{x}_j|\boldsymbol{\theta}_i)\} \qquad \text{III- (5)}$$

where

$$z_{ij} = P(i|\boldsymbol{x}_j, \boldsymbol{\Theta}) = \frac{\alpha_i f(\boldsymbol{x}_j|\boldsymbol{\theta}_i)}{\sum_{l=1}^{K} \alpha_l f(\boldsymbol{x}_j|\boldsymbol{\theta}_l)} \;\; , \quad \text{for } i \in 1, \ldots, K, j \in 1, \ldots, N. \qquad \text{III- (6)}$$

Then, the conditional expectation of the log-likelihood of the complete data $\boldsymbol{y}$ given the parameter estimate on (t)th iteration can be written as

$$\mathrm{Q}(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(t)}) = E[\log L(\boldsymbol{\Theta}^{(t)}|\boldsymbol{y})]. \qquad \text{III- (7)}$$

where $\mathrm{Q}(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(t)})$ denotes the conditional expectation of $\boldsymbol{\Theta}$, if $\boldsymbol{\Theta}^{(t)}$ is true.

M-step: The parameter set of the (*t+1*)th iteration is determined based on the estimated z_{ij}. The mixture weights would be given simply as

$$\alpha_i^{(t+1)} = \frac{1}{N}\sum_{j=1}^{N} z_{ij}\,, \quad \text{for } i \in 1, \dots, K. \qquad \text{III- (8)}$$

$\boldsymbol{\theta}_i^{(t+1)}$ that maximizes $\mathrm{Q}\left(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(t)}\right)$ can be found from $\frac{\partial \mathrm{Q}\left(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(t)}\right)}{\partial \boldsymbol{\theta}_i} = \mathbf{0}$ and the new mean and covariance matrix are

$$\boldsymbol{\mu}_i^{(t+1)} = \frac{\sum_{j=1}^{N} z_{ij}\boldsymbol{x}_j}{\sum_{j=1}^{N} z_{ij}} \qquad \text{III- (9)}$$

and

$$\Sigma_i^{(t+1)} = \frac{\sum_{j=1}^{N} z_{ij}\left(\boldsymbol{x}_j - \boldsymbol{\mu}_i^{(t+1)}\right)\left(\boldsymbol{x}_j - \boldsymbol{\mu}_i^{(t+1)}\right)^T}{\sum_{j=1}^{N} z_{ij}}. \qquad \text{III- (10)}$$

The E- and M-steps are repeated until either the difference $\log L\left(\boldsymbol{\Theta}^{(t+1)}\right) - \log L\left(\boldsymbol{\Theta}^{(t)}\right)$ becomes smaller than a convergence value or the number of iterations reaches the preselected maximum value. The convergence value of 0.000001 and the maximum iteration of 1000 were used in this study. The initial parameter values were selected by using the *k*-means clustering algorithm, where the mixture probability and covariance matrix across *k* clusters were assumed to be identical, then the centroid of each cluster is computed based on the Mahalanobis distance, which is a measure of the distance between a point and a distribution as introduced by P.C. Mahalanobis in 1936.

Figure III-4 shows an example of estimated distributions using GMM. A nonlinear discrimination line has been drawn by comparing the likelihoods of both distributions. Note that the shape of line varies with the estimated GMMs by station.

Traffic Flow Identification

Once the GMMs of the congested and uncongested traffic phases are estimated for each station, new data points fed into the algorithm are classified into either phase by comparison of likelihoods. Based on the Equation (4),

$$\begin{cases} \text{phase} = congested, & \text{if } \log L\left(\boldsymbol{\Theta}_{congested}|\boldsymbol{x}^{new}\right) > \log L\left(\boldsymbol{\Theta}_{uncongested}|\boldsymbol{x}^{new}\right) \\ \text{phase} = uncongested, & \text{otherwise.} \end{cases} \qquad \text{III- (15)}$$

where, $\boldsymbol{x}^{new}$ is the new data vector of flow and density, $\boldsymbol{\Theta}_{congested}$ and $\boldsymbol{\Theta}_{uncongested}$ are the sets of parameters for the congested flow and the uncongested flow mixture models, respectively.

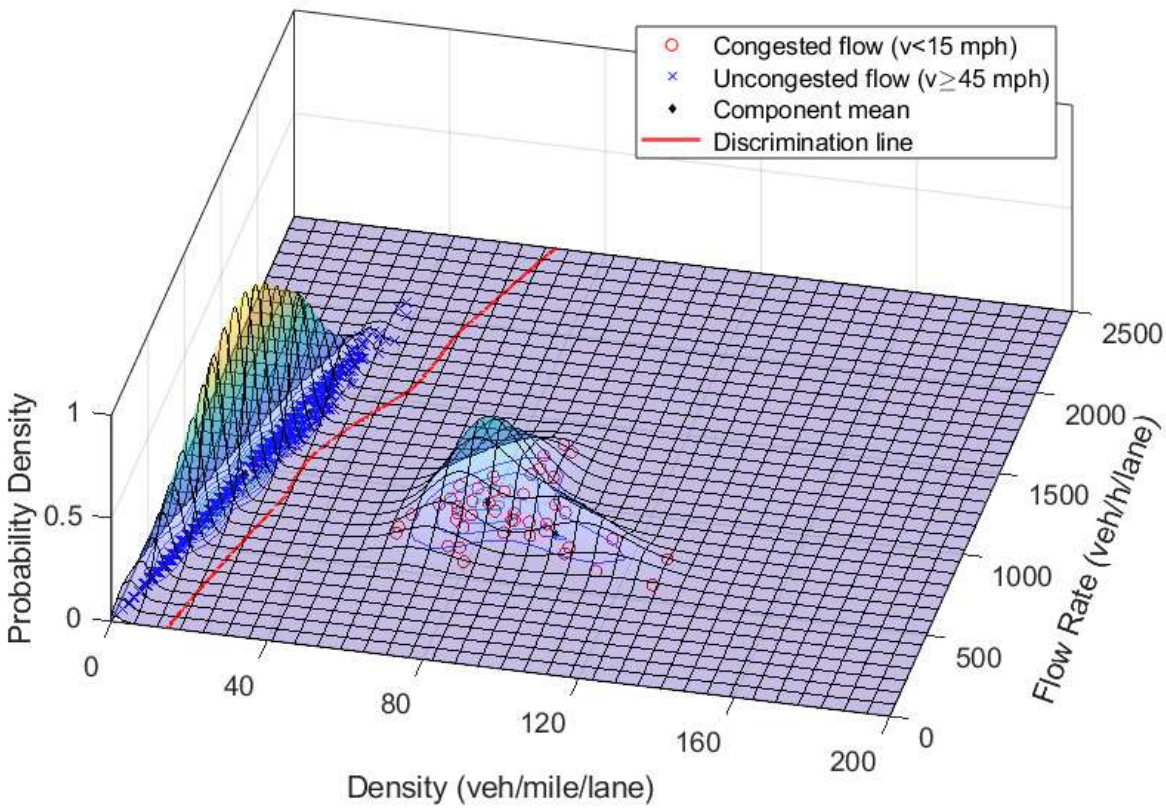

Figure III-4. An example of estimated probability density distributions using GMM

Shock Wave Speed Calculation

In traffic flow theory, a shock wave refers to boundary conditions in a time-space domain that represents a discontinuity in flow-density states[22]. Based on the well-known traffic flow theory of *flow=speed×density*, the shock wave speed between two states is defined as the change in flow divided by the change in density as follows

$$\omega_{AB} \frac{q_A - q_B}{k_A - k_B} \qquad \text{III- (16)}$$

where A and B denote different traffic flow-density states, ω_{AB} is shock wave speed when a state changes from A to B, q is flow, and k is density.

However, applying Equation (16) is not suitable for tracing a queue in the time-space domain in real-time. Unlike the theoretical concave curve or triangular shape in a flow-density diagram, real traffic flow-density data plots often show a reversed lambda shape and very chaotic movements on the right (congested) side (see Figure III-5) [23; 24]. Therefore, shock wave speeds calculated from real data are too sensitive for the purpose of this study. In addition, the speeds from Equation (16) represent shock waves at a given station as depicted in Figure III-6(a), not a link between stations. This is an important point as we can only calculate shock wave speed at a given detector station because that is the only place we have traffic flow and vehicle density information. Since we do not have flow and density state information between detector stations, we cannot directly calculate the shock wave speed along the link.

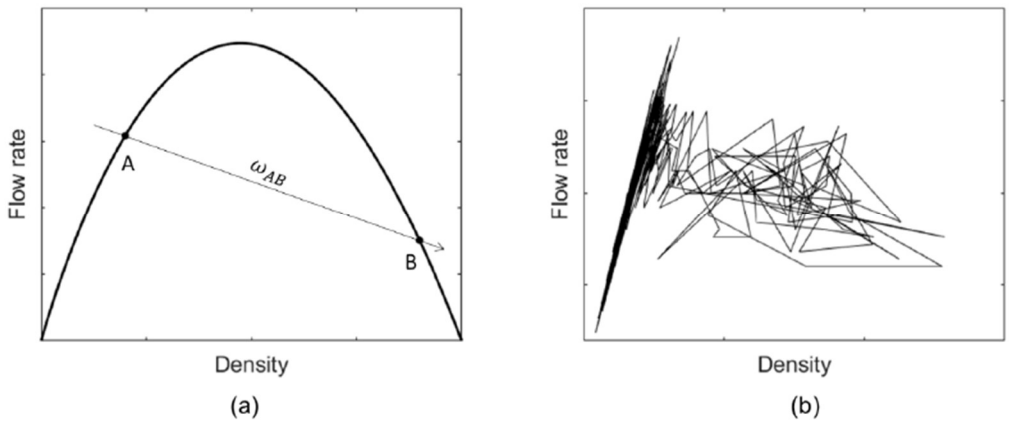

Figure III-5. Flow-density relationship: (a) theoretical flow-density curve and shock wave speed and (b) real traffic data (station at 374.2-mile EB on August 4, 2016, 4-9PM).

Just like Figure III-3 in the previous chapter, Figure IV-6 consists of time-space diagrams where the y-axis is distance (or location) with traffic going towards the top of the page while the x-axis is time progressing towards the right side of the page. When a queue occurs in a traffic stream, the shock wave travels upstream (backwards) or in the opposite direction of the traffic flow. Because one can only move forward in time, the direction (arrow) of this type of shock wave is always to the bottom-right of the page.

Shock wave speeds are calculated empirically between a pair of stations along a highway. Two different approaches were tested as shown in Figure III-6(b) and Figure IV-6(c). The first approach is to use the arrival time differences between two neighboring stations. The second approach is to use the arrival time difference between the first downstream station where a queue starts to form and each upstream station that the queue reaches. The shock wave speeds from the first approach can have a greater variation, whereas the second reduces the variation while a queue is propagating upstream. These shock wave speeds are used to predict the queue arrival time at the next upstream station.

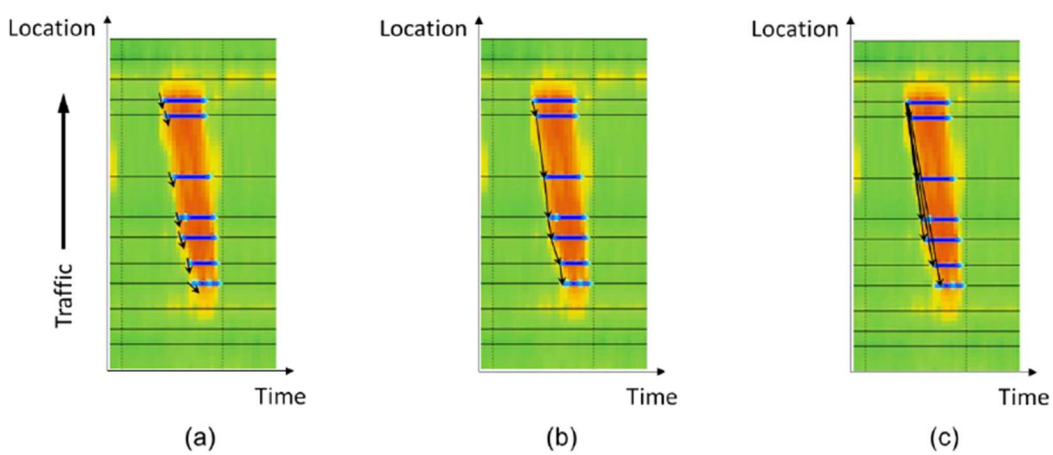

Figure III-6. Waze speed calculation(a) at each station, (b) between two neighboring stations, and (c) between the first downstream station and each upstream station.

By assuming that the shock wave detected at the current station and time required to travel to the next upstream stations at the same speed, the arrival time at the next station can be predicted.

Queue Detection and Prediction Algorithm based on WAZE Jam Reports

The previous section identifies the queue using RTMS data. In this section, a different algorithm was devised to cluster WAZE jam and accident reports to determine the spatiotemporal extent of congestion. An end-of-queue identification procedure is further proposed. The data used for this method are the WAZE jam reports, which have been discussed in Chapter III.

Overall, the proposed queue detection and prediction algorithm consists of three steps:

- **WAZE jam reports clustering:** WAZE users report incidents once they observe a crash or experience traffic congestion; the clustering algorithm logically groups those reports in spatiotemporal domain, thus identifying the congestion region.
- **End-of-queue identification:** Multiple Waze user reports associated with the same congestion or queueing event, when charted on a time-space diagram, may seem like random points at first. But a closer examination would show that these points, each representing the time and location when the report was submitted, roughly form a congestion region, on the time-space diagram, see Figure IV-6. The EOQ, which moves mostly upstream, but may occasionally linger or even shrink back downstream, is the lower boundary of this dotted region on the diagram.
- **Shock wave speed computation:** By connecting the two report points defining the lower boundary of the congestion area, which denotes the EOQ, on a time-space diagram, the shock wave speed from one report point to another can be obtained. Assuming the shock wave is continuous at this speed, at least for a short time period, the queue movement can be predicted.

Methodology

This section discusses the spatiotemporal density-based spatial clustering of application with noice (DBSCAN) algorithm, defined below, used to cluster WAZE reports and identify congestion regions. Then, an automatic boundary point identification procedure is developed. Based on the results of the identification procedure, the shock wave speed computation method is presented.

Spatiotemporal DBSCAN (ST-DBSCAN) Algorithm

In order to support two-dimensional spatial data clustering, Derya proposed a Spatiotemporal DBSCAN (ST-DBSCAN) algorithm that extends the conventional DBSCAN algorithm by adding a temporal dimension to take into account the time correlations among objects. [25; 26].

The difference between ST-DBSCAN and DBSCAN is that the neighborhood radius ε in DBSCAN is separated into two radii: the spatial neighborhood radius ε_s and temporal neighborhood radius ε_t in ST-DBSCAN. Therefore, three parameters will be used in ST-DBSCAN algorithms. ε_s, ε_t , and minPts, the minimum number of points required to form a cluster. A point, p, is said to be in the epsilon-neighborhood (eps-neighborhood) of another point, q, if and only if p is within the ε_s-neighborhood and ε_t-neighborhood of q. If q has more than minPts eps-neighborhood, q is called a core point. Similarly, the other concepts in ST-DBSCAN should be also extended accordingly based on DBSCAN. (More details on DBSCAN can be found in Clustering Analsyis textbooks.)

ST-DBSCAN first obtains the eps (ε_t, ε_s) neighbors of each data point and identifies core points within the set radius containing more than minPts neighbors. It then locates the connected components of core points initially disregarding all non-core points. Then, non-core points are attributed to an adjacent cluster if possible. All other non-core points are deemed as noise. A more detailed description of the algorithm can be found in Birant's paper[25].

In this study, the proximity of two reports is defined at both spatial and temporal levels. In addition to adding a temporal dimension, we have improved the DBSCAN algorithm in two important aspects. First, the distance functions such as Euclidian distances, which are widely adopted in existing spatial clustering algorithms, are adequate for most spatial clustering applications, but are not satisfactory for road network clusters. Instead, we use realistic road network distances to capture the spatial correlation among reports. Second, the current clustering algorithm is static and does not meet real-time detection requirements, we proposed a real-time implementation of ST-DBSCAN. The construction of the distance function and real-time implementation of ST-DBSCAN is demonstrated and discussed.

Distance Function

Temporal distances *(Δt)*

The temporal distance is simply computed as the report time differences between every two reports i and j in seconds, which may or may not be for the same incident.

$$\Delta t_{ij} = t_i - t_j \qquad \text{III- (17)}$$

If $\Delta t_{ij}$is greater than zero, then report j occurred after event i, and vice versa.

Spatial distances *(Δd)*

The spatial distance is mostly measured by Manhattan distance, Euclidean distance, or Minkowski distance given coordinates in spatial clustering studies[25]. These distance functions are suitable in most spatial clustering situations. However, directly measuring the spatial distances between two points may result in clusters that have small Euclidean distances but don't have road connections among elements, as illustrated in Figure III-7. This figure demonstrates a part of a typical interstate highway road network. The five red points inside the red dotted circle are spatially close to each other and give small Euclidean distances. Three of them are located on the westbound side of the road network while the remaining two points are on the eastbound highway and on the interchange. Although close in Euclidean distance, these points are isolated to each other and cannot be connected via any path. This has been pointed out in several previous studies as well[27].

Because of this, we use the actual road network distance (the distance of the route that connects any two reports) to measure the spatial distance in this study. Because there might exist multiple routes or paths associating one report with another, the conventional Dijkstra shortest path algorithm is implemented to obtain the shortest path between the two reports and the corresponding spatial distances.

$$\Delta d_{ij} = Dijkstra(i, j) \qquad \text{III- (18)}$$

If Δd_{ij} is greater than zero, then report j is located downstream of report i, and vice versa.

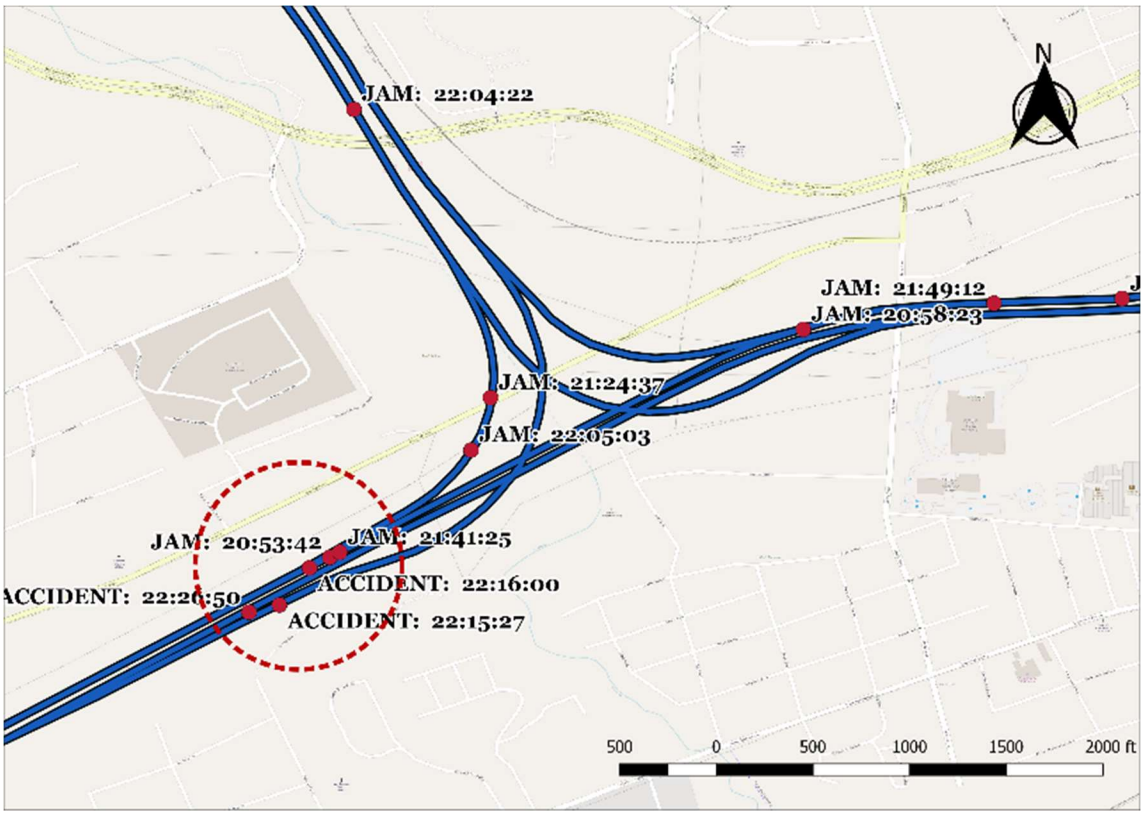

Figure III-7. Jam reports and road network connectivity

Dynamic ST-DBSCAN Algorithm

While ST-DBSCAN expands the abilities of DBSCAN by adding a temporal dimension, it is still a static clustering procedure. Our problem demands an algorithm that can be implemented in real-time when the queue in question is forming and propagating. In addition, the ST-DBSCAN algorithm has difficulties distinguishing two different clusters that start at different times and locations but propagate and merge together over time. An example is illustrated in figure 4. The title of each subplot describes the threshold used. for instance, with respect to Figure IV-8(a), it indicates the distance threshold is 3 miles and time threshold is 40 minutes, denoted as d = 3 mi and t = 40 m in the figure. All four subplots show the ST-DBSCAN clustering results using different thresholds. The points with the same color belong to the same cluster. Color shades represent different congestion levels. It is clear that different thresholds produce almost the same results (the cluster that colored with purple) for this specific case with static ST-DBSCAN algorithm. Although this big cluster actually consists of at least two clusters that feature different points, adjusting either the distance threshold or the time threshold is insufficient to produce meaningful results. In order to differentiate clusters from one another in this situation, we propose a real-time implementation of ST-DBSCAN that forms clusters dynamically and can discriminate clusters that start at different locations and times.

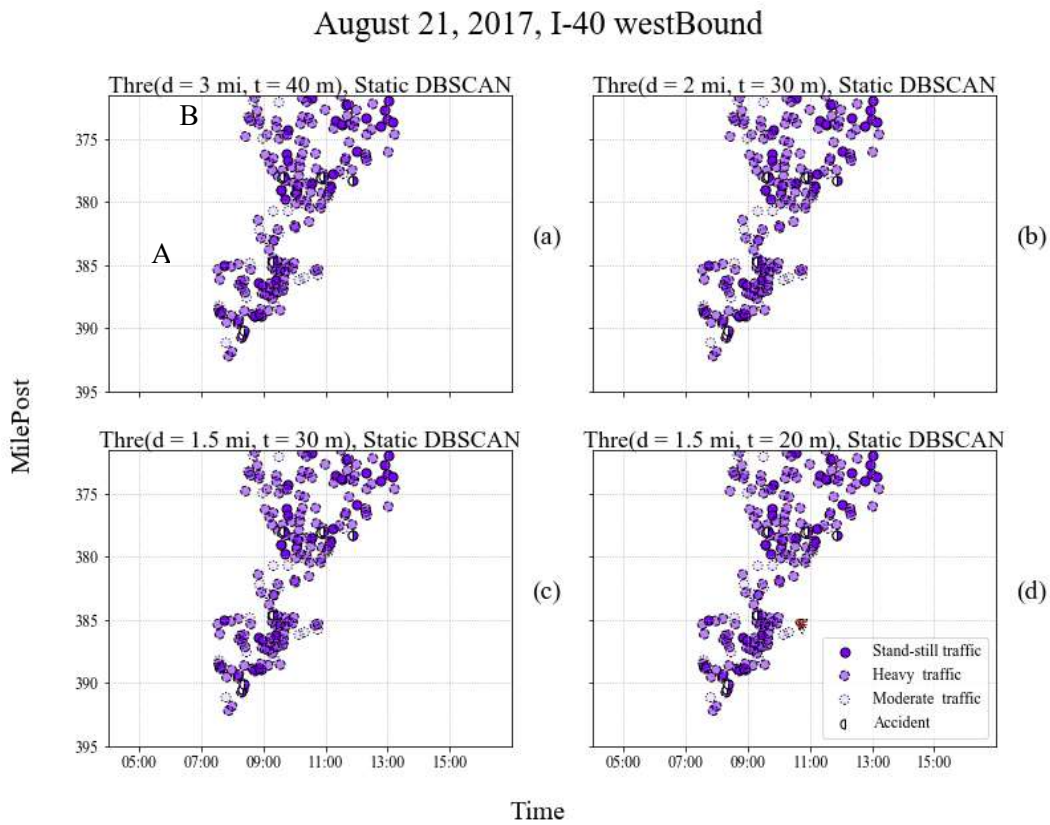

Figure III-8. ST-DBSCAN clustering results

In real-time ST-DBSCAN, there are still two distance parameters, spatial neighborhood radius(ε_s) and temporal neighborhood radius(ε_t). A point (p) is the ε-neighborhood of point (q) if and only if the point (p) is within the ε_s-neighborhood and ε_t-neighborhood of point q.

The algorithm starts with retrieving the eps (ε_t, ε_s) neighbors of each newcoming point. If the point has more than minPts neighbors, each neighbor is assigned to either the labeled dataset (already labeled with a cluster ID) or unlabeled dataset based on its current label status. If all neighbors are not labeled previously, then a new cluster starts and the neighbors are assigned to the new cluster. If all the labeled neighbors belong to the same cluster A, all the unlabeled neighbors are assigned to cluster A as well. If the neighbors belong to different clusters, each point in the unlabeled dataset is assigned to a specific cluster based on whether that point is within a spatiotemporal zone bounded by the prevailing shock wave speeds from the points in an existing cluster.

The pseudocode of the algorithm is described in detail in Figure IV-9. D is a streaming dataset composed of jam reports; it is continuously updated by Waze. *Assign_cluster* is a function used if the neighbors of the subjected point associate with two or more clusters. This function defines the boundary of two adjacent clusters. In this paper, considering the propagation feature of traffic queue, any point at the boundary of two or more clusters is assigned to the cluster to which its propagation direction is the most consistent for its location.

```
Real Time ST-DBSCAN (D, ε_t, ε_s, minPts)
    Cluster_id = 0
    for each point P in dataset D
        mark P as visited
        ε_s-neighborhood = spatialNeighbors(D, P, ε_s)
        ε_t-neighborhood = timeNeighbors(D, P, ε_t)
        N = ε_s-neighborhood ∩ ε_t-neighborhood
        if sizeof (N) < minPts:
            mark P as Noise
        else
            N = N ∪ P
            Split N into N_unlabel and N_label
            nCluster = unique(N_label) /* nCluster store the number of unique
                                          clusters of all points in N_label */
            If  nCluster <> 0:
                Cluster_id = Cluster_id + 1
                Label each point in N with Cluster_id
            Else if nCluster <> 1:
                Each point Q_u in N_unlabel is labeled with the only
                cluster
            Else if nCluster > 1:
                For each point Q_u in N_unlabel:
                    Assign_cluster (Q_u, N_label)


Assign_cluster (Q_u, N_label)
    For each point Q in N_label:
        v_{Q,Q_u} = Δd_{Q,Q_u} / Δt_{Q,Q_u}.   /* v_{Q,Q_u} represent propagation speed from Q to Q_u, */
        v_diff = v_s - v_{Q,Q_u}   /* v_s is shockwave speed of the cluster Q belongs to*/
    Assign Q_u to the cluster that Q belongs to, Q has smallest v_diff.
```

Figure III-9. Example of real-time DBSCAN implementation pseudocode

This is an effective real-time implementation of ST-DBSCAN algorithm. The biggest difference between a real-time algorithm and a static algorithm is that in a real-time algorithm, for each new coming report, *P*, if eps-neighborhood reports occur before report *P* are less than minPts, a new cluster starts. However, in a static algorithm, this report might be assigned to an existing cluster if it could be connected through some reports that occur subsequently. It is worth noticing that the real-time algorithm only applies to the case the minPts equals 2. As will be illustrated in the following section, 2 is an appropriate choice to deal with Waze data. If minPts is greater than 2, this algorithm may produce unsatisfactory clustering results.

End-of-queue Identification and Shock Wave Speed Computation

While clustering the jam reports, a small procedure is implemented to identify the EOQ dynamically. The 'term' queue has been defined in various ways in the literature. The most commonly adopted is vehicle speed less than some predefined threshold. Highway Capacity Manual 2010[(1)] defines a

queued state as 'a condition when a vehicle has slowed to less than 5 mph'. Some literature adopts a higher speed threshold with respect to highways. For instance, vehicle speed lower than 60km/h(38mph) is regarded as the EOQ in this paper[28]. With respect to Waze data, because no accurate speed information is associated with each report, each jam report is regarded as within queue status, identifying the EOQ is, therefore, identifying the jam reports that comprising the boundary of each cluster.

In traffic flow theory, a shock wave refers to boundary conditions established that demarks the time-space domain of one flow state from another. In some situations, the shock wave can be rather smooth. In other situations, the shock wave can be a very significant change in flow states, for instance when high-speed vehicles approach a queue of stopped or nearly stopped vehicles. In general, there are six different types of shock waves: frontal stationary shock wave, forward forming shock wave, backward recovery shock wave, rear stationary shock wave, backward forming shock wave, and forward recovery shock wave[29].

Among different types of shock waves, the most commonly encountered that has significant safety impacts on drivers is the backward forming shock wave. The term 'backward' implies the shock wave is moving backward or upstream in the opposite direction of traffic. This type of shock wave must always be present if congestion occurs, a typical example is congestions caused by lane blockage due to the likes of an unexpected accident or a work zone area.

In this section, we will separately define backward forming shock wave front and forward recovery shock wave front. As shown in figure 6, we first define d_{0i}, and t_{0i}, separately represent the distance and time interval from the first report to the ith report, where the red point represents the first report in this cluster and the green point represents the ith report. Note that d_{0i} is less than zero, t_{0i} is greater than zero in the example.

Backward forming shock wave front: a set of reports that for any report i, if there doesn't exist another report j in the same cluster that $t_{0j} < t_{0i}$ and $d_{0j} < d_{0i}$, then report i is defined as backward forming front or in other words, back of queue at the time report i is made.

Forward recovery shock wave front: a set of reports that for any report i, if there does not exist another report j in the same cluster that $t_{0j} > t_{0i}$ and $d_{0j} < d_{0i}$, then report i is defined forward recovery front or in other word, back of queue at the time report i is made.

From the definition, it is obvious that backward forming front and forward forming front all indicate the EOQ location at the time the report is generated. With the identified EOQ reports, the shock wave speed can be computed in two ways,

$$v_s = \frac{d_{i-1,i}}{t_{i-1,i}} \qquad \text{III-(19)}$$

or

$$\overline{v_s} = \frac{d_{0i}}{t_{0i}} \qquad \text{III- (20)}$$

where $t_{i-1,i}$ represent the time difference of report i and its most recent previous report $i-1$, $d_{i-1,i}$ represent the distance difference of report i and its most recent previous report $i-1$. Both i and $i-1$ are identified as shock-wave fronts. Equation IV-(19) computes the shock wave speed from the previous report to the current report, while equation IV-(20) computes the average shock wave speed from the very beginning report in the cluster to the current report.

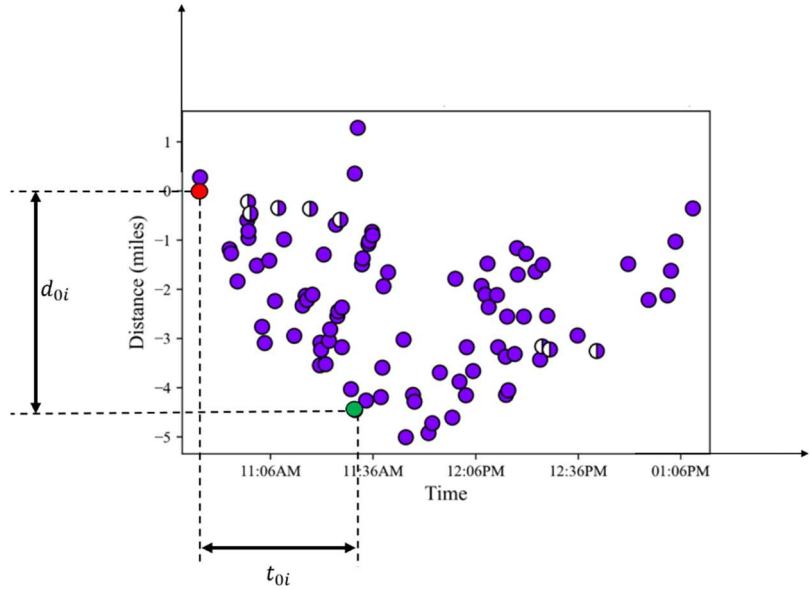

Figure III-10. Backward forming and forward recovery shock waves due to traffic incident

End-of-queue Identification and Shock Wave Speed Computation

The methodology section provides an identification method for the EOQ and shock wave speed computation. In this section, the method was applied to illustrate the procedure. According to the definition, shock wave front can be easily located. Figure III-11 demonstrates the identification results for a specific case. Points colored with green in each cluster represent the identified backward forming shock wave front.

The associated information for each shock wave front report is listed in table 2. The timestamp column denotes the time that report was made, location denotes the location of the object (milepost is used) on the Interstate Highway. The time difference column represents the time spent from the previous report to current report. The distance column represents the distance differences from the previous report to current report (negative means the present report is located upstream of the previous report). Then, the shock wave speed from one report to the next is computed as Distance/Time and stored in $\boldsymbol{v_s}$ column. The average speed from the start of the queue to the EOQ is computed as Total Distance/Total Time and stored in $\overline{\boldsymbol{v_s}}$ column. Where total time collects the overall time spent from the first report in a specific cluster to the current report, and total distance accumulates the length traveled from the first report to current reports. In other words, it represents the queue length at the current timestamp.

As shown in the table, on average, around every 8 minutes, or every 1 mile, a report was produced at the EOQ. The speed from previous to proximate reports varies while mean speed is comparatively stable. This is true in most cases; therefore, we suggest using mean speed for shock wave speed computation.

As demonstrated in the methodology section, the mean speed for moderate traffic report is around 40 mph. If we only consider standstill and heavy traffic jam reports and ignore moderate traffic jam reports, a new shock wave front can be established as well. This depends on the needs of the queue warning system.

Table III-1.EOQ and shock wave speed

Report ID	*Timestamp*	*Location (milepost)*	*Time Difference (minute)*	*Distance (mile)*	v_s *(mph)*	$\overline{v_s}$ *(mph)*
5573	15:03:52	376.45	/	/	/	/
5575	15:08:33	376.91	4.68	0.46	-5.91	-5.91
5576	15:11:20	378.04	2.78	1.13	-24.27	-12.76
5577	15:11:53	378.91	0.55	0.87	-94.37	-18.35
5586	15:22:17	379.39	10.40	0.48	-2.78	-9.56
5589	15:32:47	382.01	10.50	2.63	-15.01	-11.54
5594	15:39:59	384.30	7.20	2.29	-19.06	-13.04
5610	15:58:24	384.52	18.42	0.22	-0.71	-8.87
5614	16:00:53	385.57	2.48	1.05	-25.40	-9.59
5624	16:07:10	385.72	6.28	0.15	-1.40	-8.78
5631	16:20:32	387.50	13.37	1.78	-7.99	-8.64
5634	16:23:58	388.30	3.43	0.80	-14.03	-8.87
5639	16:28:23	389.21	4.42	0.91	-12.41	-9.06
5660	16:47:15	390.88	18.87	1.67	-5.31	-8.37
5683	17:04:17	392.38	9.98	1.68	-10.07	-8.52
5686	17:05:36	393.04	17.03	1.72	-6.06	-8.20
5701	17:13:49	393.15	1.32	0.66	-30.23	-8.42

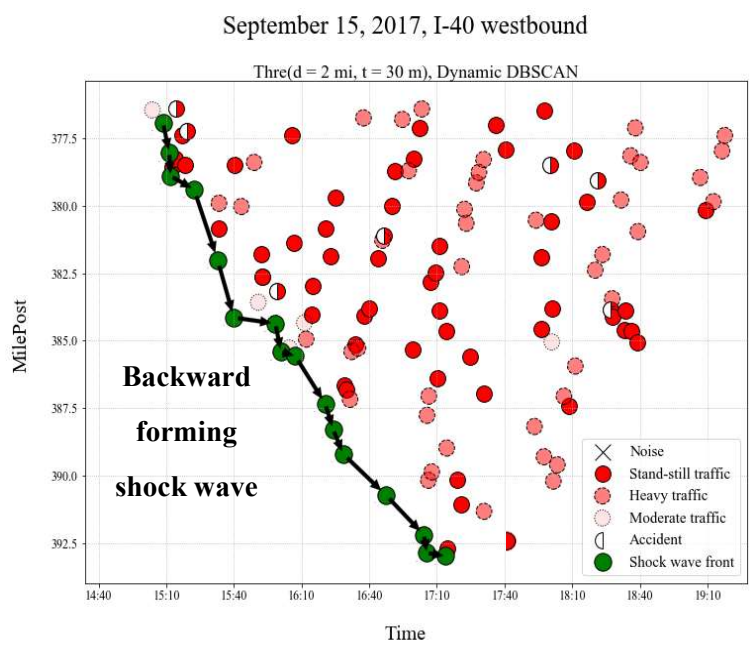

Figure III-11. Example of backward forming shock wave detection for the same accident

By assuming that the shock wave detected at the current station and time continues to travel to the next upstream station at the same speed, the arrival time at the next station can be predicted.

Chapter IV. Queue Risk Assessment

In addition to the location of the EOQ over time, the speed, speed differential and the volume of traffic in the vicinity of the EOQ can also determine the safety risks and the need for queue protection. The main risk considered in this study for the EOQ protection is rear-end collision risks. A risk assessment measurement is developed in this section.

Rear-end Risk Assessment Overview

Many safety performance measures were proposed to assess the risk of rear-end crashes including maximum deceleration rates to avoid a crash (DRAC), time to collision (TTC), proportion of stopping distance (PSD), time integrated time to collision(TIT), and crash potential index(CPI)[30]. DRAC was defined by Almquist et al. (1991) as the differential speed between the following vehicle and its corresponding lead vehicle divided by their closing time. A recent study[31] has recognized the relevance of DRAC as a measure of safety performance. DRAC explicitly considers the role of differential speeds and decelerations in traffic flow. Archer [31] suggests that a given vehicle is in "traffic conflict," or has a heightened risk for vehicular crash, if its DRAC exceeds a threshold braking value of 3.35 m/s^2. Another study [32] presents a composite measure that utilizes minimum TTC, collisions, deceleration and velocity to provide a more robust measure of driver performance in scenarios that result in collision. One study [33] assessed multiple measures and concludes that the best safety surrogate measures are time to collision, post-encroachment time, deceleration rate, maximum speed, and speed differential. Among these, time to collision, post-encroachment time, and deceleration rate can be used to measure the severity of the conflict. Maximum speed and the speed differential can be used to measure the severity of the potential collision. The potential for real-time traffic crash prediction[34] uses real-time data from loop detectors and estimate the precise time and location that a crash happened based on the upstream detector, downstream detector, and shock wave speed.

Most of the performance measures proposed in previous studies are micro-level measurements and require individual vehicle data. The application of these performance measures is restricted because at many locations, only aggregated traffic information is available. The EOQ risks are generally affected by the traffic volume, vehicle approaching speed, and speed differentials. In this study, a new safety measurement is proposed to take all the factors into consideration. The new measurement incorporates the effect of the queuing procedure into risk assessment.

Methodology

The logic behind the proposed safety measurement can be explained with the help of Figure IV-2. Assuming the traffic status in each state is homogeneous, Figure IV-2 shows simulated vehicle trajectories during the traffic phase transition period. In "Traffic State 1" vehicles travel at free-flowing or uncongested condition; in "State 2" traffic transitions from uncongested towards congested condition; and in "State 3", traffic is queuing. Between two neighboring states a shock wave is formed, e.g., shock wave A and shock wave B. A vehicle i, traveling at speed v_u starts to decelerate at time t_0 and reaches the queuing state at time $t_0 + T$ when the vehicle's speed is reduced to v_d. The following vehicle $i + 1$, traveling at speed v_u, starts to decelerate at time $t_0 + t'$ and reaches the queuing state at time $t_0 + t' + T$. Based on the homogeneous assumption, the time it takes each vehicle to travel from State 1 to State 3 is the same. For a short distance, such as 0.3 miles between two detectors, this assumption can be regarded as valid.

To avoid crash and uncomfortably hard braking, drivers tend to maintain distance to the vehicle ahead by using a minimum average deceleration rate (MADR). The traffic speed at EOQ is assumed to stay the same as it propagates backward within a short distance (two neighboring detectors, typically 0.2~0.5 miles). Therefore, based on the homogeneity assumptions, the following equations can be obtained:

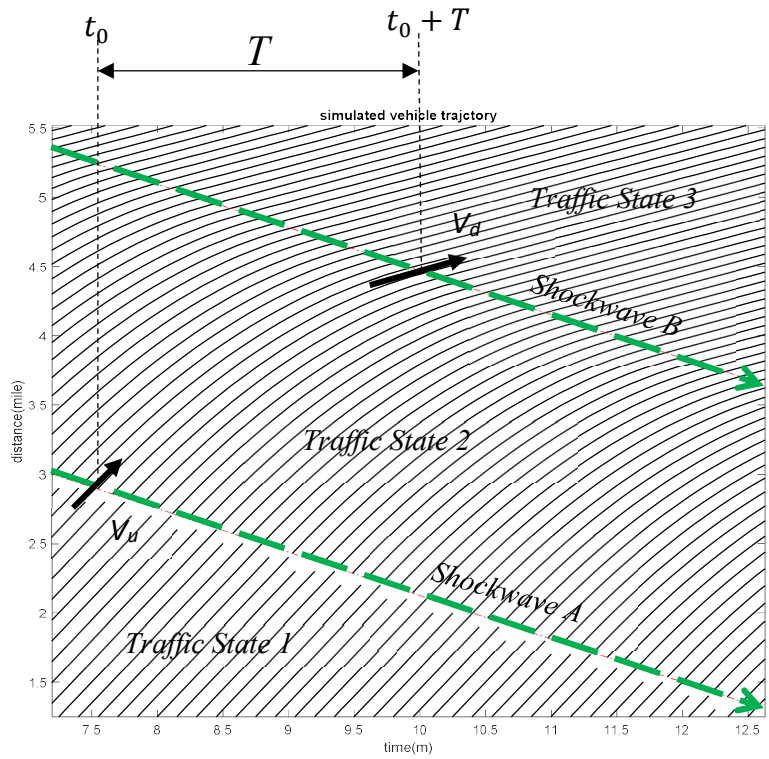

Figure IV-1. Simulated vehicle trajectories and vehicle deceleration rate

$$\omega * t + V_u * t = d \qquad \textbf{IV-(1)}$$

$$\boldsymbol{V_d - V_u = MADR * t} \qquad \textbf{IV-(2)}$$

where

ω represents shock wave speed,
V_u represent upstream vehicle speed, usually detected by upstream detectors,
V_d represent downstream vehicle speed, detected by downstream detectors, and
d represents the distance between two adjacent detectors.

The first equation computes the time t it takes for the following vehicle to reach the EOQ, the second equation indicates the minimum average deceleration rates that a following vehicle must employ so that it can reduce to the same speed of the vehicles within the queue and can avoid any irrational behaviors.

Combining the two equations, t can be eliminated, therefore we obtain:

$$MADR = \frac{(\omega + V_u) * (V_d - V_u)}{d} \quad \textbf{IV-(3)}$$

The minimum average deceleration rate (MADR) represents secondary rear-end collision risks induced by kinematic waves. Higher absolute values of MADR indicates higher risk of secondary rear-end crashes since vehicle should adopt higher deceleration rates to decelerate to a preferred speed. Lower values of MADR indicates lower likelihood of secondary rear-end collision risk. The speeds of kinematic waves were predicted by detecting the traffic state transition time, which was illustrated in Figure V-1.

Case Study

A case study is presented in this section to demonstrate how MADR is calculated. Figure V-2 is a speed heat-map showing massive queue backups in the aftermath of a multi-vehicle crash. The crash occurred at milepost 383 in the eastbound direction of I-40 in the vicinity of Knoxville, TN at 9:01 AM, August 4th, 2016. The blockage and subsequent queue buildup propagated backwards to milepost 374 and lasted about 2 hours. I-40 is a major east-west artery of the state, carrying near 200,000 vehicles per day and excessive traffic volume during peak hours. Once the incident occurred, the EOQ propagated backwards (westwards) quickly. Unsuspecting drivers approaching the back EOQ of stopped vehicles are subject to the risk of secondary crashes.

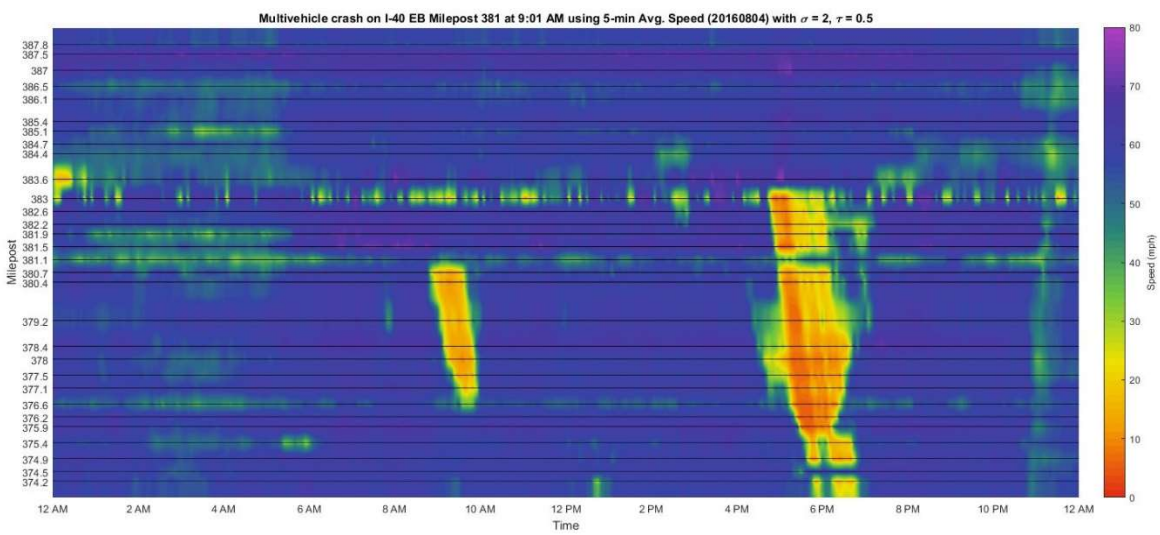

Figure IV-2. Speed heat-map

This section is equipped with TDOT's RDS detectors that report flow, speed, and occupancy every 30s for individual lanes. The spacing between two detectors ranges from 0.2 to 0.7 mile. The data from neighboring detectors were extracted to analyze the EOQ risks.

The procedure of obtaining MADR and assessing the EOQ risks is shown below,

- Detect the time that the EOQ reaches downstream detector.
- Predict the current shock wave speed.
- Assess the potential risks of vehicles that pass the upstream detectors and approach the EOQ.

The detailed procedure of risks assessment is described as follows:

Step 1. Once the EOQ reaches detector n, the shock wave speed ω will be predicted.
Step 2. Obtain the average speed reported at detector $n, n+1, n+2, \ldots, n+k$,as well as the distance between detector n and detector $n+1, n+2, \ldots, n+k$
Step 3. Use equation IV-(3) to calculate MADR values between the EOQ and upstream detectors. Table IV-1 shows the results of MADR between two neighboring detectors, the MADR value indicates the vehicle shall decelerate(-) or accelerate(+). The shock wave speed is predicted using the EOQ prediction model in the previous chapter. MADR values are updated every 30 seconds. Based on the calculation results, at time 4:47:00PM, the vehicles from upstream detector experience the highest risks.

Table IV-1. MADR results between detector 46 (I40 Milepost 380.4 eastbound) and detector 41 (I40 Milepost 379.2 eastbound) at different time periods

Predicted shock wave speed: 9.6 mph					
Distances between detectors: 0.8 mile					
downstream detector (46)		**upstream detector (41)**		**t (s)**	**MADR(ft/s)**
Time	**Speed (mph)**	**Time**	**Speed (mph)**		
4:47:00 PM	7	4:47:04 PM	58	42.60	-1.756
4:47:30 PM	13	4:47:34 PM	55	44.58	-1.382
4:48:30 PM	7	4:48:34 PM	47	50.88	-1.153
4:49:00 PM	13	4:49:04 PM	41	56.91	-0.722
4:49:30 PM	11	4:49:34 PM	35	64.57	-0.545
4:50:30 PM	7	4:50:34 PM	22	91.14	-0.241
4:51:00 PM	6	4:51:04 PM	18	104.35	-0.169
4:51:30 PM	0	4:51:34 PM	14	122.03	-0.168
4:52:00 PM	0	4:52:04 PM	14	122.03	-0.168

Then, we implement the procedure to assess the EOQ risks for multiple consecutive detectors. The assessment results are shown in Figure IV-3. Risk assessment heat-map. A layer of risk assessment results is imposed on the speed heat-map. The figure on the upper right corner is a zoomed-in view of the risk assessment layer. In this example, when the EOQ reaches the downstream detection, only the risk of vehicles approaching the nearest upstream detector is calculated. Color indicates the magnitude of risks, Yellow indicates a higher MADR value, indicating higher risk. Blue indicates a lower MADR value, indicating lower risk. The figure shows the change of rear-end collision risks over time.

Figure IV-3**. Risk assessment heat-map**

Chapter V. Quick Warning Mechanisms Assessment

Highway incidents have the potential to lead to secondary crashes. If high-speed traffic is not properly managed, rear-end collision is likely at the back end (upstream) of the resultant queue. It is important to have a rapid response system in place to prevent the occurrence of such events. Statistics show that rear-crashes occur mainly because of driver inattention and excessive speeding, also, because of poor visibility conditions (e.g. caused by fog, rain, and snow), where drivers may not be able to correctly assess the speed of downstream vehicles. Informing the motorist of the potential threat ahead and enforcing change in driver behavior when a queue is forming ahead can effectively avoid rear collisions and enhance safety and mobility. The queue warning mechanism serves as the first step in maintaining a safer transitioning of high-speed vehicles into controlled traffic queues.

The queue warning mechanism satisfies two basic requirements:

- warns approaching drivers of potential threats (the incident and traffic queues) ahead.
- diverts the drivers that plan to use the segment and suggests alternatives. This could both reduce the travel time of the drivers and decrease the flow entering the queues, thus reducing the impact of the event.

A typical warning mechanism is composed of two parts: 1) queue length detection and prediction system and 2) information dissemination system. The previous section establishes the methodology to detect and predict the EOQ. This section focuses on the comprehensive evaluation of the state-of-the-practice of traffic information broadcasting strategies.

The current dissemination mainly falls into two categories: 1) Infrastructure-based traveler information and 2) Vehicle-based traveler information. An assessment of the effectiveness, cost, market penetration and reliability of various mechanisms is discussed in the following section.

Infrastructure-based Queue Warning Mechanism

Infrastructure-based queue warning mechanisms are generally composed of roadside devices or overhead devices that disseminate the current traffic information to drivers. Typical traffic information comprises travel time information and abnormal traffic conditions (incident, work zones, or slowed traffic ahead). The devices/strategies used to broadcast the traffic information to road users include variable message signs (VMS), variable speed limit (VSL), and deployment of law enforcement personnel.

Variable Message Signs (VMS) and Signals

VMS are traffic control devices used to provide drivers dynamic/update-to-date traffic information route guidance and can be controlled either from a remote centralized location or locally at the site. They can alert drivers of slowed down traffic ahead and provide drivers with alternative route information. The effectiveness of VMS at guiding drivers has been extensively studied [35-40], VMS are effective in reducing traffic speed as shown by field speed measurements[41]. The researcher also shows that displaying route guidance information on the relevant VMS has a positive effect on improving traffic performance. More specifically, VMS are effective in rerouting traffic [35]. Rerouting traffic reduces the number of vehicles entering the congested area and thus the propagation of the queue. VMS is one of the most commonly used driver information dissemination devices in the world. Three types of VMS can be distinguished, namely permanent variable message sign, portable variable message sign (PVMS) and vehicle-mounted variable message sign.

Locating queue warning signs at positions sufficiently upstream of the queue condition is critical to allow drivers enough time to react safely. Permanent variable message signs locate at fixed locations and can alert drivers in advance of the traffic condition and provide alternatives. However, because permanent VMS are fixed infrastructure and because queue conditions can shift along a road facility, the effectiveness of permanent VMS is constrained by how well queue conditions align with the positioning and distribution of signs.

In conditions that require immediate attention from drivers, PVMS are more flexible and are often deployed at work zone areas to inform motorists of construction and road closures. They are suitable because the work zone is scheduled, and the queuing pattern is predictable. Portable VMS are placed upstream of the work-zone area and display queuing information according to dynamic queue detection and prediction. The general information displayed on PVMS is distance to the EOQ, such as slowdown 3-mi, slowdown 2-mi, etc.

Compared to the queue caused by the work zone, the queue caused by an incident has more complex temporal and spatial characteristics and is harder to predict. It is desirable to have a warning mechanism that can easily adjust its location with the queuing condition. Truck-mounted VMS is the most flexible among the three types of VMS and is suitable for queues resultant from highway incidents. A VMS-equipped truck can be stationed on the highway shoulder in advance of the slower traffic. As the traffic backs up further, the truck can be moved backwards along shoulders with the queue and warn the upstream traffic continuously.

Variable Message Signals

Variable message signals are used to convey the message to motorists for improving network performance, reducing congestion and incident management. MnDOT deployed intelligent lane control signals (ILCS) spaced every half-mile over every lane to warn motorists of incidents or hazards on the roadway ahead[42]. It was combined with a system that identifies lane-specific shock wave or queuing conditions on the freeway and uses existing ILCS to warn motorists upstream for rear-end collision prevention; a 22% decrease in crashes and a 54% decrease in near-crashes was observed after the deployment of the system. Since lane-by-lane variable signals are expensive, it is recommended this system be installed at critical locations where high rear-end crash risks are expected.

In practical applications, the three types of variable message signs can be used simultaneously. The combination of different types of VMS can convey traffic information to motorists more effectively.

Variable Speed Limits (VSL)

VSL utilizes traffic speed and volume detection, weather information, and road surface condition technology to determine appropriate speeds at which drivers should be traveling, given current roadway and traffic conditions. The goal of VSL systems is to adjust the speed limit of a roadway based on prevailing road, traffic, and environmental conditions to improve safety and efficiency. Studies have demonstrated positive safety impacts of such systems[40]. Variable speed limits are particularly suitable for managing queues at work zone areas. University of Michigan developed an adaptive queue warning system called Smart Drum. Smart drum automatically calculates the speed of traffic flow using speed sensors and is communicated to the control system for warning and variable speed limits[43].

Use of Law Enforcement Personnel

Law Enforcement Personnel (LEP) can provide an extra level of notification for traffic queues on higher speed roadways and enhance drivers' attention. Some states use LEPs to monitor the back of the queue. LEPs are often combined with warning message signs to alert drivers. They may be parked off the roadway between the second and third advance warning sign and may be facing traffic. They also may move upstream as needed to always provide presence and motorist warning (through flashing lights) in advance of the EOQ.

Some other strategies include deploying prepare to STOP signs. Different from general VMS, 'prepare to STOP' signs are critical when high-speed traffic is approaching nearly stopped queues. This is helpful in an accident scenario, where the queue is backed up quickly and the stopped traffic is unexpected to upstream drivers. When depoying 'Prepare to STOP' signs it is crucial to ensure adequate stopping sight distance is provided upstream of the EOQ. This is the minimal distance required for the driver to react and stop their vehicle before the EOQ.

The infrastructure-based queue warning mechanism has several advantages. The system achieves high coverage, the devices or personnel deployed on-road are visible to all drivers by their nature. However, infrastructure-based queue warning applications are fundamentally limited in their potential range and scope due to many factors such as large spacing between fixed VMS, limited geographical coverage of VMS, imprecise traffic queue information, etc. Also, the information is only available to on-road users, travelers that plan to travel cannot be informed by on-road device.

TDOT Current Practice and Suggestions

TDOT has four fully integrated Intelligent Transportation System (ITS), known as TDOT SmartWay, located in Memphis, Nashville, Chattanooga, and Knoxville. The system compromises a total of 551 cameras and 1107 roadway detection systems that support real-time traffic conditions detection and management.

TDOT is equipped with 198 overhead VMS signs in four regions (as shown in Figure VI-1). Those signs are located in urban interstate highway and provide safety information to travelers based on the traffic information provided by TDOT's comprehensive detection and surveillance system. The general message shown on these VMS is the real-time travel time information. But if congestions or incidents are detected and verified, the corresponding information such as the location of the queue, alternate routes, and expected delay is shown. As aforementioned, the effectiveness of overhead VMS is limited since they can only provide information at fixed locations. In areas where no VMS are available, TDOT often deploys PVMS located 1 to 2 miles beyond the EOQ to warn travelers of the unexpected traffic conditions ahead.

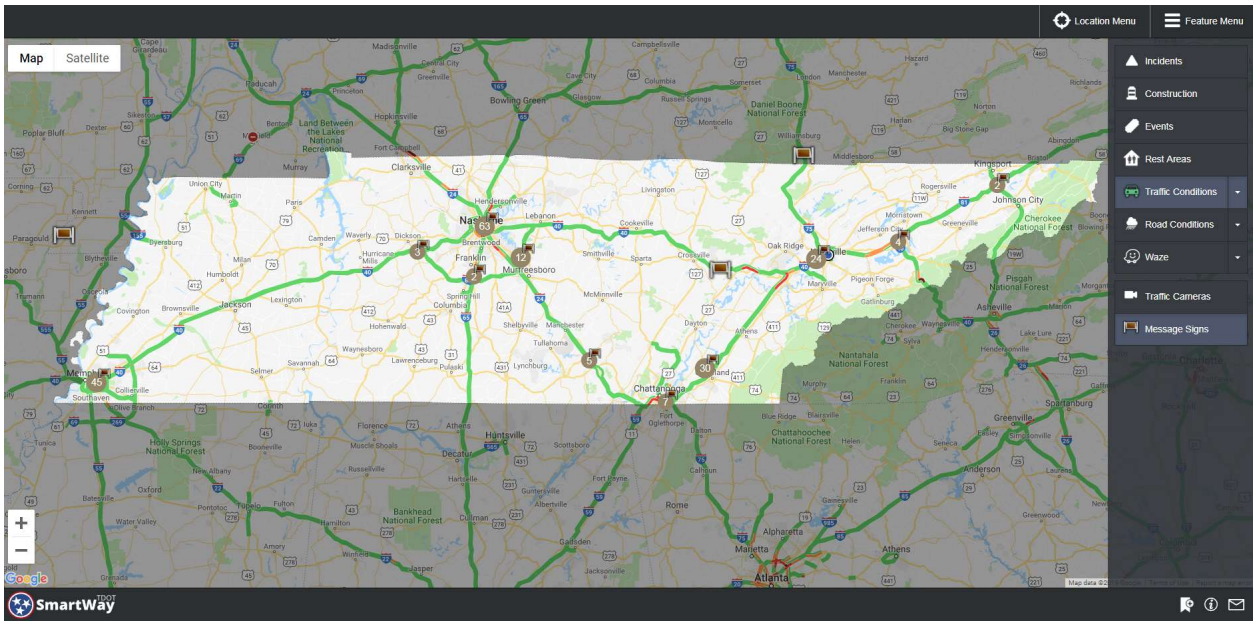

Figure V-1. The overhead variable message signs in Tennessee

TDOT currently utilizes variable speed limits on I-75 in region 2 in an area prone to fog issues but does not currently operate any congestion-based variable speed limits (VSL). TDOT is currently reviewing and developing standard operating procedures for deploying VSL as a congestion mitigation and safety measure. The application of VSL for incident management has gained importance over time as it stands to show effectiveness in queue and congestion control. Within the context of protecting the queue (PTQ), VSL utilizes PTQ algorithms to detect the occurrence of incidents and dispatch rapid responses to manage the speed of approaching traffic to protect the ends of queues. These algorithms are important in the sense that they reduce the potential for the occurrence of secondary crash incidents. The challenge with VSL in the present day is figuring out how to optimize its utilization in different situational contexts. Most traffic jams have the characteristics of moving upstream at a speed of approximately 15 km/h and remain stationary for a long time[(14)]. In optimizing the coordination of VSL to suppress shock waves, these shock waves need to be in a metastable state to allow the shock wave to spread out into a disturbance that is small enough to dissipate on its own[(15)].

TDOT operates the HELP truck program on designated mainline highways in the state's major urban areas to respond to traffic incidents in real-time. In addition, supervisor trucks in rural counties are dispatched for PTQ responsibilities when major incidents/queues are reported on nearby Interstate highways. These trucks are equipped with vehicle-mounted VMS signs and the supervisors are trained to perform PTQ activities.

Vehicle-based Queue Warning Application

Highway Advisory Radio (HAR)

Highway advisory radio systems are vehicle-based queue warning applications. Timely traffic information mostly comes from the infrastructure-based detection systems. One of the advantages of HAR is the ability to divert travelers away from congested areas and incident events.

Connected Vehicles/Dedicated Short-Range Communication (DSRC)

Dedicated short-range communication (DSRC) is a technology that allows vehicles to communicate with each other. It lets cars broadcast their position, speed, road condition, and other data to other vehicles up to a few hundred meters. The other cars can use such information to build a detailed picture of what's unfolding around them, revealing trouble that even the most careful and alert driver, or the best sensor system, would miss or fail to anticipate. There have been efforts studying the use of DSRC in queue detection and warning systems as well as rear-end collision avoidance systems. Results show the application of DSRC in these systems is promising [44-47]. But problems still exist in DSRC-based traffic queue protection systems such as relatively high rates of false alarms and missing alarms in emergency warnings. Additionally, the reliability of the system is easily affected by some factors like DSRC penetration rate, DSRC communication range, positioning accuracy, and transmission delay. Note that DSRC is not the only option for wireless data dissemination. With the much-anticipated 5G technology on the horizon, the technology adoption and deployment decision is a challenging balance between cost, functionality, and future-proofing.

Probe-Based/third-party Based Queue Warning Applications

With the advent of technology and proliferation of third-party traffic data sources such as INRIX, HERE (formerly NAVTEQ Maps), WAZE, TomTom, and others, queue warning systems can be developed using only these data or in combination with other data sources such as infrastructure data. Many DOTs have subscriptions to use traffic mobility data from third-party sources for varying purposes. Such data can fill the void for road facilities where traditional infrastructure/sensors have not been deployed to monitor traffic flow conditions and eventually used to develop queue warning applications.

The travel time information provided by the third party are of different quality, frequency, and road segmentation consistency, which increases the difficulty in incorporating the data into current traffic information systems. The quality of probe vehicle data is affected by several factors including as market penetration rates, traffic conditions, etc. Further analysis needs to be conducted by transportation agencies to verify the accuracy of probe vehicle data before incorporating the data in detecting queues for dissemination.

In addition to traffic speed information provided to traffic agencies, navigation applications such as Google Maps and WAZE also provide travel time to travelers and guide users to alternative routes. Travelers are informed of traffic information such as queue locations (colored lines on the map) and estimated travel time. This serves as a complementary method for EOQ warning. Transportation agencies can potentially collaborate with these third-party vendors and disseminate traffic information via their platform. Information such as the location of crashes, back forming queues, emergency vehicles, HELP trucks, and developing situations can be quite useful to the motoring public.

Websites and Social Media

Many DOTs also disseminate traffic information on their traffic information website and social media. This could potentially reduce traffic demand and divert drivers from congested/incident routes to alternative routes. All these serve as complementary methods to the on-site EOQ warning system.

Conclusion

In previous sections, a series of queue warning strategies and the benefits/challenges of each strategy were presented. Table V-1 summarizes current EOQ warning strategies and TDOT's deployment practice.

Table V-1. Summary of EOQ warning strategies

Strategies		TDOT Current Practice
Infrastructure Based	VMS	Equipped
	VSL	Equipped at foggy location
	LEP	THP (highway patrol)
Vehicle-Based	HAR	Equipped
	DSRC	In the process of development and deployment
	Navigation APP	Established a contract with WAZE to share information
	Website and Social Media	Equipped

The selection of different traffic queue protection strategies varies with traffic conditions. It is recommended that TDOT fully considers the advantage and disadvantages of varies strategies discussed in this section. Within the applicable means of ITS, queue warning can be achieved effectively with a number or combination of different strategies. For example, when DSRC and navigation apps are coupled with law enforcing and VSL displays, swift responses can be expected.

Chapter VI. Implementation Strategies and Conclusion

The previous chapters investigated different aspects of the queue prediction and warning system. Chapter IV presents the real time automatic EOQ detection and prediction algorithm, Chapter V discussed the dynamic risk assessment procedure, and Chapter VI developed a framework of the implementation of protecting the EOQ. To realize the safety benefits of TDOT's efforts towards protecting the queue, the automated queue prediction algorithm and the subsequent motorist warning system have to be deployed successfully in real-time. This chapter considers all the techniques developed in previous chapters and establishes an implementable framework based on TDOT's resources and the availability and maturity of various technologies. The framework identifies all pieces of the eventual system, how they would function together and the fashion they would be integrated.

Basically, the EOQ prediction and warning system developed in this study was composed of four components: data management system, EOQ prediction and risk assessment mechanism, EOQ warning mechanism, and queue protection implementation strategy. Some practical issues with the implementation of each system will be discussed in the following sections. Then, a framework that integrates each of those components is established.

Data Management System

The method proposed in this study relies on RTMS and WAZE data to detect and predict traffic status. Therefore, a dependable real-time data management system is needed to archive, maintain, process, analyze, and implement the mission-critical multimodal traffic data. TDOT now archives RTMS data in its Active ITS system. The data comes in 30-second intervals and is stored in ASCII format. This data needs to be split and matched to each detector pair for EOQ detection and prediction algorithm implementation. WAZE provides jam/accident feeds for the entire state of Tennessee in XML format. The logs are updated once a new report from the user comes in. Therefore, a mechanism needs to be developed to dynamically extract the new reports and match it to appropriate roadway segments based on its GPS coordinates.

EOQ Prediction and Risk Assessment System

The EOQ prediction algorithms developed in this study can be implemented in real-time as intended. After processing the RTMS and WAZE data, the algorithms will process the data and automatic detection and predict the EOQ location. Once the queue location in the next time period is determined by the algorithm, the rear-end collision risks at each detector location can be calculated using the surrogate safety measurement proposed in Chapter V. Some sort of Monte Carlo simulation or sensitivity or factor analysis could be performed here to take into consideration of stochastic effects to bracket the location of the end-of-queue and the magnitude of the risks.

EOQ Warning Mechanism

Infrastructure-based and vehicle-based queue warning mechanisms are comprehensively assessed in Chapter VI of this report. Infrastructure-based warning mechanisms such as VMS are suitable for broadcasting the current EOQ location to drivers. However, this information could be misleading, especially for a crash-related queue, which grows rapidly over a short period of time. Under these circumstances, based on the current traffic condition and predicted queue propagation speed, PVMS and truck-mounted VMS could be used to provide real-time estimation of the time and distance to reach the EOQ.

WAZE and Google Maps now provides colored lines on their navigation maps to indicate the congested links, which implies the approximate location of the EOQ. But even with this information, the approaching drivers are likely unaware of the risk related to the EOQ. TDOT can explore the potential of working with these applications and provide additional information such as the EOQ risks and queue propagation speeds to offer the motoring public more information.

Benefits of EOQ Detection/Prediction Algorithm

Independent to phone calls, which is the prevailing mechanism for TDOT to be informed about incidents in non-urban areas, and without the advantage of TDOT's RDS stations and CCTV cameras, the EOQ detection algorithm could take in real-time probe data, from WAZE, INRIX, or the like, as well as crowdsourced event reports, from WAZE (or Google Maps) to identify the existence of an unexpected queue, the location of the EOQ, and the propagation direction/speed of it. The biggest benefit of this endeavor is crash prevention. The algorithm can help proactively manage the queue, especially queues caused by non-recurrent events, and lessen rear-end collision risks. By disseminating the information to approaching drivers and diverting drivers further upstream to alternative routes, the system can improve the safety of non-recurring events. At the local level, where the EOQ is, the reduction in speed differential can reduce the crash probability and, if unpreventable, impact force. At the system level, this would reduce congestion delay through diversion as well as prevention of secondary crashes. Prompt and proper management in the form of queue protection could help maintain a travel time reliability in spite of non-recurring events.

Final Thoughts

The EOQ is hazardous to the motorists sitting at the end-of-queue as well as the unsuspecting drivers approaching it at cruising speed. This study developed a system that can detect and predict the EOQ location, which is a crucial step towards protecting the queue. The system developed in this study dynamically detects the EOQ and predicts its movement in spatiotemporal domains based on real-time traffic data and the traffic flow models. The EOQ location information can then be disseminated promptly to approaching and at-risk drivers in multiple ways. The deployment of HELP trucks or other emergency vehicles to "protect" the queue can raise driver awareness of the potential hazards on road ahead.

The queue detection and prediction algorithms developed in this study are based on TDOT's existing RDS traffic data supplemented with WAZE traffic logs. This allows TDOT to expand the EOQ monitoring effort beyond the four TMC areas. The queue prediction and protection system are composed of four subsystems:

1. Real-Time Traffic Data Management System;
2. EOQ Prediction and Risk Assessment System;
3. EOQ Warning System; and
4. EOQ Protection System.

These are presented in depth in earlier chapters of this report. The subsystems need to be integrated to work in unison for the entire EOQ protection effort to be effective. In the meanwhile, further work on testing the queue detection/prediction models and the implementation of new data sources, such as INRIX, should be looked into and pilot-implemented at TDOT to move forward with this vision.

To realize the beneifts of the EOQ algorithm developed in this project, the transportation agency needs to keep the three indepent, real-time streams: phone calls, crowdsourced event

reports, and traffic data based EOQ detection. Preferrably, the three sources of info should be somewhat automated and presented to the appropriate dispatching operators, at a local TMC perhaps, in a timely, concise, and geographically informative manner. Subsequent steps of dispatching and coordinating with local, typically county, supervisor vehicles to respond to perform PTQ activities would also be essesntial for realizing the crash avoidance benefits.

List of Abbreviations

5G	Fifth Generation standard for cellular technology
CAV	Connected and Auomated Vehicles
CCP	WAZE's Connected Citizen Program, renamed as WAZE for Cities Program in 2019
CCTV	Closed Circuit Television
CPI	Crash Potential Index
CTR	Center for Transportation Research at the University of Tennessee
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DRAC	Maximum Deceleration Rate to Avoid a Crash
DSRC	Dedicated Short Range Communication
EB	Eastbound
EM	Expectation Maximization
EOQ	End-of-queue
FHWA	Federal Highway Administration
fps	Frames per second
GMM	Gaussian Mixture Model
GPS	Global Positioning System
HAR	Highway Advisory Radio
HCM	Highway Capacity Manual
HERE	A traffic data system
ILCS	Intelligent Lane Control Signal
INRIX	A traffic analytics comanpy
IO	Input and Output
ITS	Intelligent Transportation Systems
LEP	Law Enforcement Personnel
LWR	The classic traffic flow theory per Lighthill, Whitham, and Richards
MADR	Minimum Average Deceleration Rate
MLE	Maximum Likelihood Estimation
mph	Miles per hour
NB	Northbound
NHS	National Highway System
NPMRDS	National Performance Management Research Data Set
pdf	Probability Density Function
PM3	Performance Measure Rule 3
PTQ	Protect the Queue
PSD	Proportion of Stopping Distance
PVMS	Portable Variable Message Sign
RDS	Radar Detection System
RTMS	Remote Traffic Microwave Sensor
SB	Southbound
SPR	State Planning and Research Program
TDOT	Tennessee Department of Transportation
THP	Tennessee Highway Patrol
TIT	Time Integrated Time to Collision
TMC	Traffic Management Center
TN	Tennessee
TTC	Time to Collision
UT	The University of Tennessee

VMS	Variable Message Signs
VSL	Variable Speed Limits
WAZE	A GPS navigation softward app owned by Google
WB	Westbound
XMP	Extensible Markup Language

REFERENCES

[1] Highway Capacity Manual 2010. Transportation research board. *National Research*, 2010.

[2] Palikareva, H., and C. Cadar. Multi-solver support in symbolic execution.In *International Conference on Computer Aided Verification*, Springer, 2013. pp. 53-68.

[3] Webster, F. V. *Traffic Signal Settings*. H.M. Stationery Office, 1958.

[4] Michalopoulos, P. G., G. Stephanopoulos, and G. Stephanopoulos. An application of shock wave theory to traffic signal control. *Transportation Research Part B: Methodological,* Vol. 15, No. 1, 1981, pp. 35-51.

[5] Sharma, A., D. Bullock, and J. Bonneson. Input-Output and Hybrid Techniques for Real-Time Prediction of Delay and Maximum Queue Length at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2035, 2007, pp. 69-80.

[6] Liu, H. X., X. Wu, W. Ma, and H. Hu. Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C: Emerging Technologies,* Vol. 17, No. 4, 2009, pp. 412-427.

[7] Lighthill, M. J., and G. B. Whitham. On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* Vol. 229, No. 1178, 1955, p. 317.

[8] Richards, P. I. Shock Waves on the Highway. *Operations Research,* Vol. 4, No. 1, 1956, pp. 42-51.

[9] Kerner, B. S. *The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory*. Springer, Berlin, Heidelberg, New York, 2004.

[10] Geroliminis, N., and A. Skabardonis. Prediction of Arrival Profiles and Queue Lengths Along Signalized Arterials by Using a Markov Decision Process. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 1934, 2005, pp. 116-124.

[11] Comert, G., and M. Cetin. Queue length estimation from probe vehicle location and the impacts of sample size. *European journal of operational research,* Vol. 197, No. 1, 2009, pp. 196-202.

[12] Ban, X., P. Hao, and Z. Sun. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transportation Research Part C: Emerging Technologies,* Vol. 19, No. 6, 2011, pp. 1133-1156.

[13] FHWA. Summary of Vehicle Detection and Surveillance Technologies Used in Intelligent Transportation Systems.In, 2007.

[14] ---. Traffic Detector Handbook: Third Edition—Volume I.In, 2006.

[15] *TDOT SmartWay*. https://smartway.tn.gov/traffic.

[16] *WAZE*. https://www.waze.com/about.

[17] Chen, C., A. Skabardonis, and P. Varaiya. Systematic Identification of Freeway Bottlenecks. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 1867, 2004, pp. 46-52.

[18] Li, H., and R. Bertini. Comparison of Algorithms for Systematic Tracking of Patterns of Traffic Congestion on Freeways in Portland, Oregon. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2178, 2010, pp. 101-110.

[19] Treiber, M., and D. Helbing. Reconstructing the spatiotemporal traffic dynamics from stationary detector data. *Cooperative Transport@tion Dynamics,* Vol. 1, 2002, pp. 3.1–3.24.

[20] Treiber, M., A. Kesting, and R. E. Wilson. Reconstructing the Traffic State by Fusion of Heterogeneous Data. *Computer-Aided Civil and Infrastructure Engineering,* Vol. 26, No. 6, 2011, pp. 408-419.

[21] McLachlan, G. J. Finite mixture models.In, New York : Wiley, New York, 2000.

[22] May, A. D. *Traffic flow fundamentals*. 1990.

[23] Edie, L. C. Car-following and steady-state theory for noncongested traffic. *Operations research,* Vol. 9, No. 1, 1961, pp. 66-76.
[24] Daganzo, C. F. A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections. *Transportation Research Part B: Methodological,* Vol. 36, No. 2, 2002, pp. 131-158.
[25] Birant, D., and A. Kut. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering,* Vol. 60, No. 1, 2007, pp. 208-221.
[26] Wang, M., A. Wang, and A. Li. Mining spatial-temporal clusters from geo-databases.In *International Conference on Advanced Data Mining and Applications*, Springer, 2006. pp. 263-270.
[27] Zhang, Y., and L. D. Han. Dijkstra-DBSCAN: Fast, Accurate and Routable Density Based Clustering of Traffic Incidents on Large Road Network.In, 2018.
[28] Dinh, T.-U., R. Billot, E. Pillet, and N.-E. El Faouzi. Real-Time Queue-End Detection on Freeways with Floating Car Data: Practice-Ready Algorithm. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2470, 2014, pp. 46-56.
[29] Ferrara, A., S. Sacone, and S. Siri. Fundamentals of Traffic Dynamics.In *Freeway Traffic Modelling and Control*, Springer, 2018. pp. 25-44.
[30] Guido, G., F. Saccomanno, A. Vitale, V. Astarita, and D. Festa. Comparing safety performance measures obtained from video capture data. *Journal of Transportation Engineering,* Vol. 137, No. 7, 2010, pp. 481-491.
[31] Archer, J. Methods for the assessment and prediction of traffic safety at urban intersections and their application in micro-simulation modelling. *Royal institute of technology*, 2004.
[32] Brown, T. L. Adjusted minimum time-to-collision (TTC): A robust approach to evaluating crash scenarios.In *Proceedings of the Driving Simulation Conference North America, No. 40*, 2005.
[33] Gettman, D., and L. Head. Surrogate safety measures from traffic simulation models. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1840, 2003, pp. 104-115.
[34] Abdel-Aty, M., A. Pande, L.Y. Hsia, and F. Abdalla. The Potential for Real-Time Traffic Crash Prediction. *ITE JOURNAL,* Vol. 69, 2005.
[35] Chatterjee, K., and M. Mcdonald. Effectiveness of using variable message signs to disseminate dynamic traffic information: Evidence from field trails in European cities. *Transport Reviews,* Vol. 24, No. 5, 2004, pp. 559-585.
[36] Peeta, S., and J. L. Ramos. Driver response to variable message signs-based traffic information.In *IEE Proceedings-Intelligent Transport Systems, No. 153*, IET, 2006. pp. 2-10.
[37] Kolisetty, V. G. B., T. Iryo, Y. Asakura, and K. Kuroda. Effect of variable message signs on driver speed behavior on a section of expressway under adverse fog conditions—a driving simulator approach. *Journal of advanced transportation,* Vol. 40, No. 1, 2006, pp. 47-74.
[38] Edara, P., C. Sun, C. Keller, and Y. Hou. Evaluating the benefits of dynamic message signs on Missouri's rural corridors.In, Missouri. Dept. of Transportation, 2011.
[39] Erke, A., F. Sagberg, and R. Hagman. Effects of route guidance variable message signs (VMS) on driver behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour,* Vol. 10, No. 6, 2007, pp. 447-457.
[40] Borrough, P. Variable speed limits reduce crashes significantly in the UK. *The Urban Transportation Monitor*, 1997.
[41] Ulfarsson, G. F., V. N. Shankar, and P. Vu. The effect of variable message and speed limit signs on mean speeds and speed deviations. *International Journal of Vehicle Information and Communication Systems,* Vol. 1, No. 1-2, 2005.
[42] Hourdos, J., Z. Liu, P. Dirks, H. X. Liu, S. Huang, W. Sun, and L. Xiao. Development of a queue warning system utilizing ATM infrastructure system development and field-testing. 2017.

[43] Sullivan, J., C. Winkler, M. Hagan, B. Kantowitz, and P. Huang. The Concept of a Smart Drum Speed Warning System. *Ann Arbor, Michigan: University of Michigan Transportation Research Institute*, 2007.

[44] Liu, Y., W.-B. Zhang, Z.-L. Wang, and C.-Y. Chan. DSRC-based end of queue warning system.In *Intelligent Vehicles Symposium (IV), 2017 IEEE*, IEEE, 2017. pp. 993-998.

[45] Ibrahim, U., M. I Hayee, E. Kwon, and M. Donath. Development of a Freeway Queue Detection and Warning System using Ad-hoc Control and DSRC based V2V Communication. *Recent Advances in Communications and Networking Technology (Formerly Recent Patents on Telecommunication),* Vol. 4, No. 2, 2015, pp. 103-116.

[46] Xu, Q., T. Mak, J. Ko, and R. Sengupta. Vehicle-to-vehicle safety messaging in DSRC.In *Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, ACM, 2004. pp. 19-28.

[47] Xiang, X., W. Qin, and B. Xiang. Research on a DSRC-based rear-end collision warning model. *IEEE Transactions on Intelligent Transportation Systems,* Vol. 15, No. 3, 2014, pp. 1054-1065.

APPENDIX