Delivering Improved Alerts, Warnings, and Control Assistance Using Basic Safety Messages Transmitted between Connected Vehicles

Jun Liu, Ph.D.
Research Associate
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, TN 37996
jliu34@utk.edu

Asad J. Khattak, Ph.D.
(corresponding author)
Beaman Professor & Transportation Program Coordinator
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, TN 37996
akhattak@utk.edu

2nd Round Revision summited to Transportation Research, Part C

March, 2016

Delivering Improved Alerts, Warnings, and Control Assistance Using Basic Safety Messages
Transmitted between Connected Vehicles

Jun Liu & Asad Khattak
University of Tennessee, Knoxville

Abstract - When vehicles share their status information with other vehicles or the infrastructure, driving actions can be planned better, hazards can be identified sooner, and safer responses to hazards are possible. The Safety Pilot Model Deployment (SPMD) is underway in Ann Arbor, Michigan; the purpose is to demonstrate connected technologies in a real-world environment. The core data transmitted through Vehicle-to-Vehicle and Vehicle-to-Infrastructure (or V2V and V2I) applications are called Basic Safety Messages (BSMs), which are transmitted typically at a frequency of 10 Hz. BSMs describe a vehicle's position (latitude, longitude, and elevation) and motion (heading, speed, and acceleration). This study proposes a data analytic methodology to extract critical information from raw BSM data available from SPMD. A total of 968,522 records of basic safety messages, gathered from 155 trips made by 49 vehicles, was analyzed. The information extracted from BSM data captured extreme driving events such as hard accelerations and braking. This information can be provided to drivers, giving them instantaneous feedback about dangers in surrounding roadway environments; it can also provide control assistance. While extracting critical information from BSMs, this study offers a fundamental understanding of instantaneous driving decisions. Longitudinal and lateral accelerations included in BSMs were specifically investigated. Varying distributions of instantaneous longitudinal and lateral accelerations are quantified. Based on the distributions, the study created a framework for generating alerts/warnings, and control assistance from extreme events, transmittable through V2V and V2I applications. Models were estimated to untangle the correlates of extreme events. The implications of the findings and applications to connected vehicles are discussed in this paper.

Keywords: connected vehicles, basic safety messages, extreme events, longitudinal and lateral accelerations

INTRODUCTION

Connected vehicles are spawning new applications to address real-world safety problems, such as collisions owing to unexpected hard braking and fast lane changing. Two major applications for driving safety are vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications to inform drivers of roadway hazards and situations that they cannot immediately perceive. V2V applications enable vehicles to transmit and receive data from surrounding vehicles, and V2I allows vehicles to communicate with the infrastructure. They can receive information about traffic signal timing or incidents in real-time. These systems help drivers make informed decisions avoiding dangers posed by surrounding vehicles or by infrastructure elements (Li et al., 2011; Jin et al., 2012; Lee and Park, 2012; Goodall et al., 2013; Qin and Orosz, 2013; Hu et al., 2015). According to a US DOT report, V2V and V2I applications can substantially improve safety (Office, 2014) by reducing driver errors, which are

key contributors to crashes. The errors can include recognition errors, decision errors, and performance or nonperformance errors (NHTSA, 2008).

A connected vehicle safety pilot program, Safety Pilot Model Deployment (SPMD), underway in Ann Arbor, Michigan intended to demonstrate V2V and V2I technologies in a real-world environment (Henclewood, 2014). Approximately 3,000 vehicles are equipped with V2V communication devices, and 75 miles of roadway are instrumented with roadside equipment, mainly placed at signalized intersections. Data acquisition systems are installed in vehicles participating in the program to facilitate V2V and V2I communications. Data transmitted through V2V and V2I are called Basic Safety Message (BSM), sampled at a frequency of 10 Hz. The core contents of BSM are data elements that describe a vehicle's position (latitude, longitude, and elevation) and motion (heading, speed, and acceleration) (Henclewood, 2014). BSM also contains data pertaining to the vehicle's component status (lights, brakes, wipers) and vehicle safety information (path history, events) (Henclewood, 2014). Given the availability of advanced communication and sensor technologies such as dedicated short-range communications (DSRC), Global Positioning Systems (GPS), Radar, and Bluetooth, there is no doubt that BSMs can be successfully sent and received by vehicles and roadside equipment under various road conditions. SPMD is a demonstration of such communication technologies in a connected vehicles environment.

Using these technologies, what kind of critical information can be extracted and provided to drivers to alert them to present dangers or assist in vehicle control to help drivers make safe decisions? Driver-oriented information is supposed to be simple and informative, such as head-on collision warnings. This study proposes an original methodology, based on data analytics, to extract critical information from BSMs transmitted between connected vehicles and infrastructure.

Specifically, this study aims to 1) investigate the magnitudes and directions (longitudinal and lateral) of instantaneous driving decisions contained in BSMs; and 2) identify critical events (i.e., hard accelerations or braking and quick lane changes) by establishing reasonable thresholds. These events are the basis for actionable messages (alerts, warnings or control assists) given to drivers of connected vehicles. Instantaneous driving decisions are mainly implemented by pedal maneuvering (brake and accelerator) and steering. Pedal maneuvering changes the speed or acceleration, and steering changes the direction of vehicle motion. To uncover the full range of instantaneous driving decisions, it is important to investigate both the magnitudes and directions of instantaneous driving decisions. In addition, the current thresholds used for recognizing critical driving behaviors are not sensitive to the vehicle speeds and different driving conditions faced by drivers (De Vlieger et al., 2000; Simons-Morton et al., 2011; Kim and Choi, 2013). Furthermore, directional differences of instantaneous driving decisions, i.e., longitudinal and lateral acceleration, has not been studied widely (De Vlieger et al., 2000; Kim and Choi, 2013). Given the available high-resolution BSM data, this study contributes by establishing reasonable thresholds for identifying potentially dangerous behaviors, while considering variations in driving behaviors under different conditions. Directional differences in instantaneous driving decisions are also explored.

LITERATURE REVIEW

The literature reflects substantial increase in research activity with regards to connected vehicles, covering a wide range of topics from how connected vehicles will be adopted and used, to their

applications and implications for safety, energy, and the environment (Koulakezian and Leon-Garcia, 2011; Zeng et al., 2012; Olia et al., 2014; Zhang et al., 2014; Fagnant and Kockelman, 2015). A report by Hill et al. discusses connected vehicle infrastructure deployment approaches and strategies, including the time horizon, key issues related to drivers and vehicles, operations, benefits for state and local agencies, and early connected vehicle adopters (Hill and Garrett, 2011). Substantial work has been done to establish connected vehicle networks, such as Vehicular Ad-hoc Networks (Li et al., 2011; Zhang and Orosz, 2013). Both the public and private sectors are interested in applications and implications of connected vehicles. Applications include intersection signals (Lee and Park, 2012; Christofa et al., 2013; Guler et al., 2014; Hu et al., 2014; Wu et al., 2014; Bagheri et al., 2015; Feng et al., 2015), pavement assessments (Bridgelall, 2013), traffic queue estimation (Li et al., 2013), vehicle routing and travel time estimation (Kianfar and Edara, 2013; Tian et al., 2013; Genders and Razavi, 2015; Moylan and Skabardonis, 2015), driving behavior monitoring and warnings (Sengupta et al., 2007; Goodall et al., 2014; Chrysler et al., 2015; Doecke et al., 2015; Du and Dao, 2015; Osman et al., 2015), and fuel efficiency (Kishore Kamalanathsharma and Rakha, 2014; Liu et al., 2016).

The key to the success of connected vehicles lies in how well connectivity of vehicles and infrastructure works in real life. Recent innovations that enable connectivity include applications of V2V and V2I. supported by wireless communication technologies such as DSRC (Cheng et al., 2007; Chan, 2011), Wi-Fi (Goel et al., 2004; Chou et al., 2009), Bluetooth (Nusser and Pelz, 2000; Sugiura and Dermawan, 2005), and cellular networks (Mosyagin, 2010; Abid et al., 2012).

SPMD uses Basic Safety Messages (BSMs) to describe a vehicle's position, motion, its component status, and other relevant travel information (Henclewood, 2014). However, the BSMs are not informative to drivers when they need to make decisions based on information received through V2V or V2I applications. Most BSMs describe normal driver behaviors. However, abnormal and extreme driver behaviors determine the safety of driving the short-term. Thus, it is critical to identify abnormal or extreme behaviors from BSMs, and warn drivers through the V2V, V2I, or other connected vehicle applications.

A number of studies have focused on investigating driving behaviors. Vehicle motion (speed and acceleration) has been regarded as the core information describing driving behaviors. Fast driving is normally characterized as an aggressive, reckless or risky driving style, and speed limits are usually the threshold that determines a driver's performance (Lajunen et al., 1997a; Hoedemaeker and Brookhuis, 1998; Lajunen et al., 1998; NHTSA, 2000). However, speed choice depends mainly on the conditions of speed limits (or road conditions) and the traffic. A driver is supposed to comply speed limits, but he or she is also affected by the traffic (Åberg et al., 1997; Haglund and Åberg, 2000). Researchers give several acceleration cut-off points as thresholds to identify abnormal, extreme, and aggressive driving behaviors. Kim and Choi report thresholds for aggressive and extremely aggressive accelerations in urban driving environments (Kim and Choi, 2013), while De Vlieger et al. did similar work for calm driving, normal driving, and aggressive driving (De Vlieger et al., 2000). Simons-Morton et al. advanced the characterization of risky driving by observing the elevated gravitational force (G-force) events that are captured when longitudinal or lateral accelerations exceed certain thresholds (Simons-Morton et al., 2011; Simons-Morton et al., 2012; Simons-Morton et al., 2013). Risky driving behaviors have been found to be positively correlated with the likelihood of crashes or near-crash events (Paleti et al., 2010; Guo and Fang, 2013).

Driving occurs in various conditions, and driver behaviors may vary in different contexts.

To account for the variation of driving behaviors under different conditions, Liu et al. and Wang et al. introduced a varying acceleration threshold to identify extreme driving behaviors (Liu et al., 2014; Liu, 2015; Liu et al., 2015; Wang et al., 2015). Most previous studies investigating driver behaviors overlook the directions (longitudinal and lateral) of driving decisions. Driving decisions in two directions (longitudinal and lateral accelerations) are under-explored. This study proposes an innovative way to identify extreme driving behaviors embedded in BSMs that may provide warning messages to drivers through V2V and V2I applications.

While previous studies propose ideas for warnings or alerts to drivers using the connected vehicle applications (Goodall et al., 2014; Chrysler et al., 2015; Doecke et al., 2015; Du and Dao, 2015; Osman et al., 2015), they have not fully assimilated the value of information transmitted between connected vehicles. For example, Noble et al., (2014) used only naturalistic driving data collected through the Strategic Highway Research Program 2 (but not BSMs) for analysis, and Osman et al., (2015) used driving simulator based data. This study fully mines the geo-referenced data transmitted between vehicles and infrastructure in a real-life connected vehicle deployment. Specifically, it extracts useful information about extreme events from new data sources made possible by communication between connected vehicles.

METHODOLOGY

Recently, NHTSA and Society of Automotive Engineers (SAE) have come out with levels of automation that range from no automation to full automation (SAE, 2014). Figure 1 shows the how drivers can transition from controlling all aspects of the dynamic driving task to relinquishing control of these tasks. Based on the taxonomy, this study focuses on Level 0 in the SAE taxonomy. Within Level 0, the study transitions from driving without alerts, warnings, or intervention systems to using these for enhanced driving safety. Increasingly, vehicles are incorporating driver decision support systems, while drivers retain control of steering and braking controls, except in crash imminent situations. Alerts, warnings, and control assists can be divided into two broad categories:

1) Internal to the functioning and performance of the driver or vehicle. Examples of these include warnings about hard accelerations or braking, or frequent lane changes, sharp turns, or functioning of wipers, head- and tail-lights, turn signals, etc.

2) External to the vehicle. These are warnings that relate to proximity of other vehicles, objects, infrastructure, and the environment, and include forward collision warning or lane departure warning.

This paper considers both types of alerts, warnings, and control assists, which are still part of Level "0." Such warnings are based on BSM data, and can eventually lead to higher levels of automation.

SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	Monitoring of Driving Environment	Fallback Performance of Dynamic Driving Task	System Capability (Driving Modes)
Human driver monitors the driving environment						
0	No Automation	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	Driver Assistance	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes
2	Partial Automation	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	System	Human driver	Human driver	Some driving modes
Automated driving system ("system") monitors the driving environment						
3	Conditional Automation	the *driving mode*-specific performance by an *automated driving system* of all aspects of the dynamic driving task with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	System	Human driver	Some driving modes
4	High Automation	the *driving mode*-specific performance by an automated driving system of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	System	Some driving modes
5	Full Automation	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	All driving modes

FIGURE 1 Six levels of driving automation. Source: SAE, J3016, 2014 (SAE, 2014).

Data Description - Basic Safety Message

The data used in this study are from BSMs sent and received by vehicles and roadside equipment participating the SPMD in Ann Arbor, Michigan (Henclewood, 2014). SPMD provides a rich database for research on connected vehicles. The data are stored in a transportation data sharing system, called Research Data Exchange (RDE, https://www.its-rde.net/home), maintained by the Federal Highway Administration under US DOT. This study uses datasets collected from participating vehicles equipped with Data Acquisition Systems. Datasets contain vehicles' instantaneous driving statuses (sampled at 10 Hz) of position (altitude, latitude and longitude), motion (speed and acceleration), the status of major components (accelerator, brakes, lights, cruise control, and wipers), and instantaneous driving contexts (surrounding objects, and distance to closest objects). Table 2 presents the detailed descriptions of key data variables. More variable descriptions are available in SPMD Data Handbook (Henclewood, 2014).

One-day sample data were used in this study. Observations with errors (e.g., speeds > 200 mph and altitude > 30,000 ft) were removed from the sample. The final one-day sample contains 968,522 records of basic safety messages, from 155 trips made by 49 vehicles. The sum of trip durations is about 26.9 hours, and the average duration per trip is about 10.4 minutes. Most of the trips were made within the road networks of Ann Arbor, Michigan, and some long trips reached the neighboring towns of Dexter, Chelsea, and Livonia in Michigan. Table 3 shows the descriptive statistics of selected variables in the final datasets. Based on the error-checked descriptive statistics and the distributions, the data seemed to be of reasonably good quality.

Figure 2 presents the spatial distribution of sampled data. Distributions of variables seemed reasonable in terms of magnitude and spatial characteristics.

TABLE 2 Variable Descriptions from Safety Pilot Model Deployment, Ann Arbor, Michigan

Variable		Description
Position	Altitude	A GPS-based estimate of height above sea level (height above the reference ellipsoid that approximates mean sea level)
	Latitude	Current degree of latitude at which the vehicle is located
	Longitude	Current degree of longitude at which the vehicle is located
Motion	Speed (host vehicle)	Current vehicle speed, as determined from the vehicle's transmission
	Longitudinal Acceleration	Longitudinal acceleration measured by an Inertial Measurement Unit (IMU)
	Lateral Acceleration	Lateral acceleration measured by an IMU
Vehicle Maneuvering	Accelerator Pedal	Reflects the amount the accelerator pedal is displaced with respect to its neutral position
	Brake Pedal	Indicates whether the brake light is on or off
	Cruise Control	Indicates whether cruise control is active/engaged
	Turn Signal	Provides information regarding the state of the vehicle turn signals
Driving Context	Number of objects	Number of identified objects, as determined by the Mobileye sensor
	Distance to the closest object	Position of the closest object, relative to a reference point on the host vehicle, according to the Mobileye sensor
	Relative speed of the closest object	Longitudinal velocity of the closest object, relative to the host vehicle according to the Mobileye sensor

Source: SPMD Data Handbook (Henclewood, 2014).

TABLE 3 Descriptive Statistics of Selected Variables

Variable		Mean/ Percentage	Std. Dev.	Min	Max
Position	Altitude (ft)	724.603	81.076	496.388	1345.492
	Latitude (degree)	42.307	0.135	42.044	42.977
	Longitude (degree)	-83.745	0.291	-85.635	-83.280
Motion	Host Vehicle Speed (mph)	38.507	23.249	0.000	83.346
	Longitudinal Acceleration (ft/s^2)	-0.367	2.107	-21.818	22.420
	Lateral Acceleration (ft/s^2)	-0.090	2.246	-22.310	22.330
Vehicle Maneuvering	Accelerator Pedal (%)	13.299%	-	0.000	1.000
	Brake Pedal (engaged)	20.186%	-	0.000	1.000
	Cruise Control (engaged)	38.822%	-	0.000	1.000
	Turn Signal (None)	93.847%	-	0.000	1.000
	Turn Signal (Left)	3.927%	-	0.000	1.000
	Turn Signal (Right)	2.226%	-	0.000	1.000
Driving Context	Number of objects	1.618	1.469	0.000	9.000
	Distance to the closest object (ft)	132.919	116.053	0.410	839.690
	Relative Speed of the closest object (mph)	-3.893	21.829	-157.984	159.941

Note: Sample size = 968,522 records.

FIGURE 2 Spatial Distribution of Trajectory Data in the Final Datasets

Conceptual Framework

The main objective of this study is to use data analytics to extract critical information embedded in BSMs sent and received by vehicles and infrastructure. Figure 3 shows the proposed framework to explore the BSMs and compile the raw BSMs into messages that can be communicated to drivers. These messages can inform the host vehicle drivers (i.e., raw BSMs are from the same vehicle) to adjust their driving behaviors, and also to give warnings and control assistance to remote vehicle drivers (i.e., raw BSMs are from other vehicles) to avoid potential dangers. In real-world environments, real-time BSMs are compiled into real-time advisory or warning messages, directed to local drivers through V2V and V2I applications.

Understanding instantaneous driving volatility was one of the most challenging aspects of this study; this understanding can be accomplished by the BSM Compiler designed in Figure 3. Data sampled at a high frequency, 10 Hz, yielded deeper insights into instantaneous driving behaviors. This study used various data visualization tools to show the extent of instantaneous driving volatility, including distributions of longitudinal and lateral acceleration, speed-based distributions of instantaneous yaw rate, three-dimensional distributions of longitudinal acceleration-lateral acceleration-speed, and driving volatility on different road types. This paper provides data visualization details. Then, extreme driving events will be identified in accordance with special rules. One of the rules this study used is speed-based thresholds (Liu et al., 2014), since driving behaviors vary at different speeds, implying different driving contexts. These rules are discussed along with data visualization and analysis. Further, these extreme events were linked to instantaneous vehicle control statuses and driving contexts to understand why they occur. Finally, advisory or warning messages and vehicle control assistance to drivers can be generated.
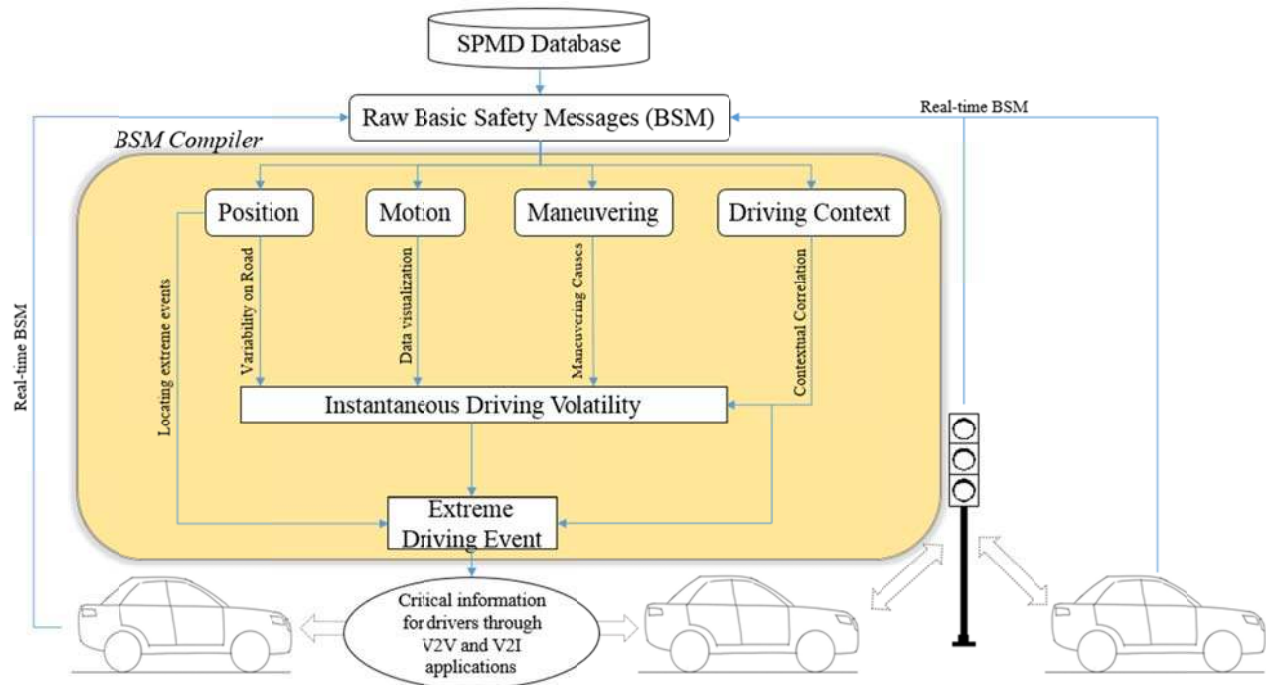
FIGURE 3 Conceptual Framework

EXTENT OF INSTANTANEOUS DRIVING VOLATILITY

Driving Volatility on the Road

To observe how a vehicle moves on a road, 10-Hz motion data (speed, longitudinal, and lateral acceleration) from BSMs were visualized on maps according to position data (longitude and latitude). Figure 4 shows two sample trips made on different types of roads. Figure 4(i) presents one trip going through downtown Ann Arbor and Figure 4(ii) shows a trip that includes freeway driving. To illustrate the volatility of instantaneous driving decisions (i.e., variability of

instantaneous accelerations), Figure 4 also shows the variance of longitudinal and lateral accelerations within one second, calculated by Equations (1) and (2) below.

$$VAR_i^{Longitudinal} = \frac{1}{n}\sum_{j=i}^{i+9}\left(a_j^{Longitudinal} - \frac{1}{n}\sum_{j=i}^{i+9}a_j^{Longitudinal}\right)^2 \qquad \text{Equation (1)}$$

$$VAR_i^{Lateral} = \frac{1}{n}\sum_{j=i}^{i+9}\left(a_j^{Lateral} - \frac{1}{n}\sum_{j=i}^{i+9}a_j^{Lateral}\right)^2 \qquad \text{Equation (2)}$$

Where,
VAR = variance of accelerations within one second;
i = time stamp in 0.1 second, i = 1.0, 1.1, 1.2, …;
j = the index of summation, $j = i,\ i+1,\ i+2,\ ..,\ i+9$;
$a_j^{Longitudinal}$ = the jth record of instantaneous longitudinal acceleration in 10 Hz;
$a_j^{Lateral}$ = the jth record of instantaneous lateral acceleration in 10 Hz;

Note that this study uses longitudinal and lateral acceleration as key measures to capture instantaneous driving decisions, i.e., how a vehicle is maneuvered instantaneously. Driving behaviors can also be captured by other measures, such as steering angles and the position of the accelerator or brake in a vehicle. Given that longitudinal and lateral accelerations are the outcomes of maneuvering by drivers, the authors prefer to use them for analysis.

As expected, driving on local roads is more volatile, in terms of variance of accelerations, than driving on freeways because conditions on local roads are more complex owing to greater variation in roadways and surrounding environment, e.g., presence of different traffic signs and signals at intersections and distractions from pedestrians, bicyclists as well as changes in land use (commercial, residential or industrial). For this reason, critical information extraction from BSMs should consider the difference of driving performance or behavior under different driving contexts. An easy way to distinguish the driving context would be travel speed, by assuming no congestions on the road. The travel speed on freeways would often be higher than on local roads, as indicated by speed limits. Figure 4 presents the instantaneous driving decisions visualized on road, showing that higher speeds are associated with smaller variations of acceleration in two directions. Ahn et al. (Ahn et al., 2002) report that higher accelerations are associated with lower speeds. Consequently, driving volatility might also be associated with speeds.
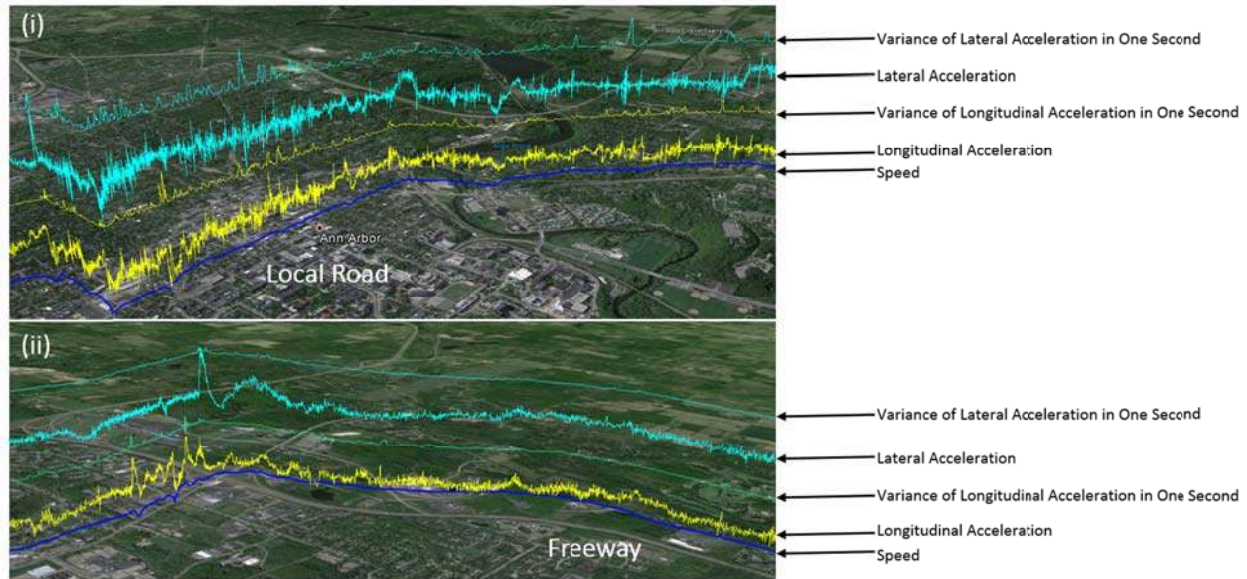
FIGURE 4 Instantaneous Driving Decisions Visualized on Road

Distributions of Acceleration

To clarify the relationship between speeds and acceleration, distributions of acceleration were visualized in two directions: longitudinal and lateral. Figure 5 presents the magnitude distribution of acceleration along speeds and the density of possible acceleration values along speeds. It shows that generally higher speeds (>50 mph) were associated with smaller acceleration magnitudes, which is partially consistent with Ahn et al. Vehicle engines have to do more work in order to maintain the same acceleration at higher speeds to overcome increasing air resistance. Therefore, the ability to accelerate or decelerate a vehicle decreases naturally at higher speeds (Wang et al., 2015). Ahn et al. (Ahn et al., 2002) points out a linear relationship between acceleration and speeds, and this study reveals a nonlinear relationship between acceleration and speed in real-life driving situations. The varying distributions of instantaneous accelerations along speeds confirm the above findings that driving behavior varies in different driving situations, as reflected by driving speeds. In addition, Figure 5 also presents the distribution of longitudinal vs. lateral accelerations. The lozenge shaped distribution implies that longitudinal and lateral accelerations do not have large magnitudes simultaneously. In terms of their magnitude, longitudinal and lateral accelerations seem to be inversely correlated (correlation = -0.8343, *p*-value < 0.01).
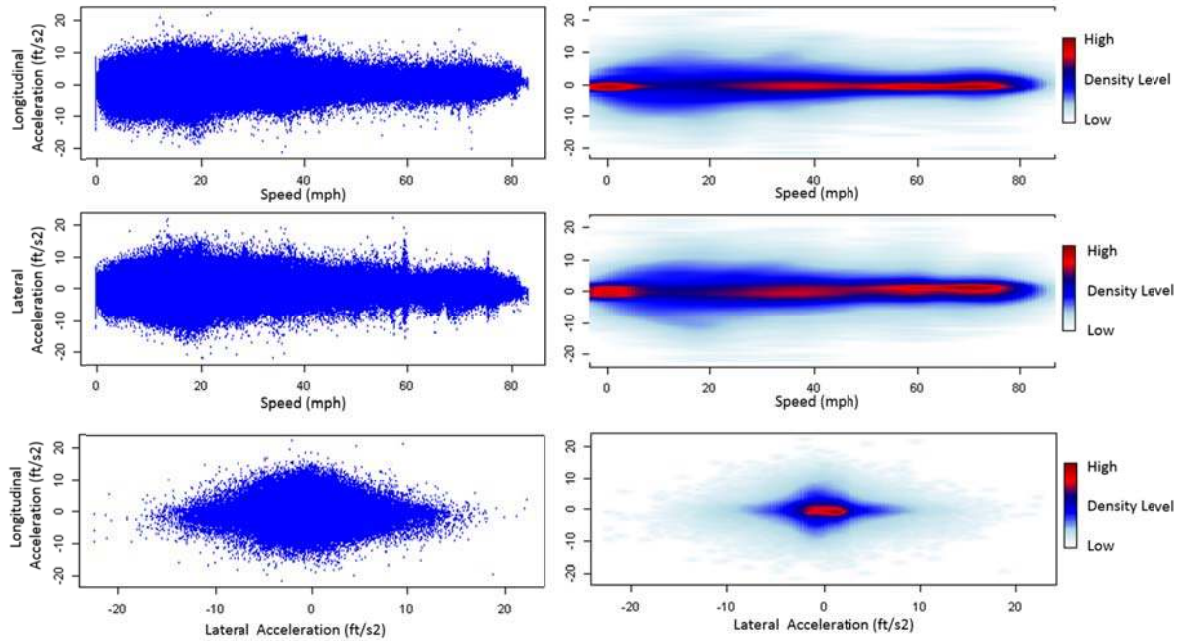
FIGURE 5 Distributions of Longitudinal and Lateral Acceleration

The distributions of longitudinal and lateral accelerations were visualized in three-dimensional space, according to their relative magnitude and direction. The resultant instantaneous acceleration of a vehicle is the sum of motion vectors of longitudinal and lateral acceleration, as shown in Equation (3).

$$\vec{A} = \vec{a}_{Longitudinal} + \vec{a}_{Lateral} \qquad \text{Equation (3)}$$

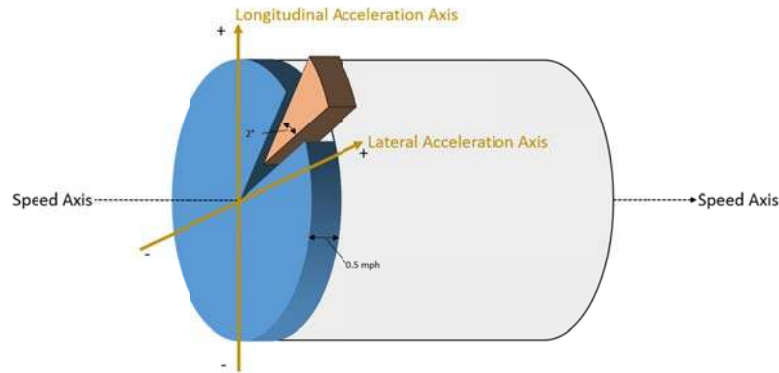The magnitude of resultant instantaneous acceleration $\vec{A}$ is

$$|\vec{A}| = \sqrt[2]{(\vec{a}_{Longitudinal}^{2} + \vec{a}_{Lateral}^{2})} \qquad \text{Equation (4)}$$

The direction of resultant instantaneous acceleration $\vec{A}$ is

$$\Delta = 180 \frac{\tan^{-1}\left(\dfrac{\vec{a}_{Longitudinal}}{\vec{a}_{Lateral}}\right)}{\pi} \qquad \text{Equation (5)}$$

Δ is the counter-clockwise angle between vehicle heading direction and the direction of resultant instantaneous acceleration.

As shown in Figure 5, varying distributions implied variations in driving behaviors at given speeds and motion in different directions. To show the magnitude of distribution at various speeds and directions, this study used different colors to indicate the relative magnitudes of accelerations by speed × sector bin in 0.5 mph × 2°. An example bin is shown in Figure 6. All magnitudes of accelerations were compared within a bin.

FIGURE 6 Speed × Sector Bin

Figure 7 shows the three-dimensional distribution of accelerations at different speeds and directions. The view of sectional drawings reveals the magnitude of instantaneous accelerations. Blue implies that the magnitudes of accelerations are close to zero compared to other accelerations within the same bin, and red indicates greater magnitudes. Same color indicates that magnitudes of acceleration within different bins were at the same percentiles, forming percentile bands. Magnitudes of longitudinal and lateral accelerations varied with different speeds. Magnitudes at lower speeds were relatively larger than at higher speeds, illustrated by the percentile bands, as shown in Figure 7(iii) and (iv). The cross orthogonal shape, shown by the blue and yellow/green area in Figure 7 (ii), implies that magnitudes of accelerations that were parallel or perpendicular to the heading directions were relatively greater than those diagonal to heading directions. This confirmed that instantaneous driving decisions varied in different directions, which is useful for identifying extreme driving events such as sudden lane change behaviors.
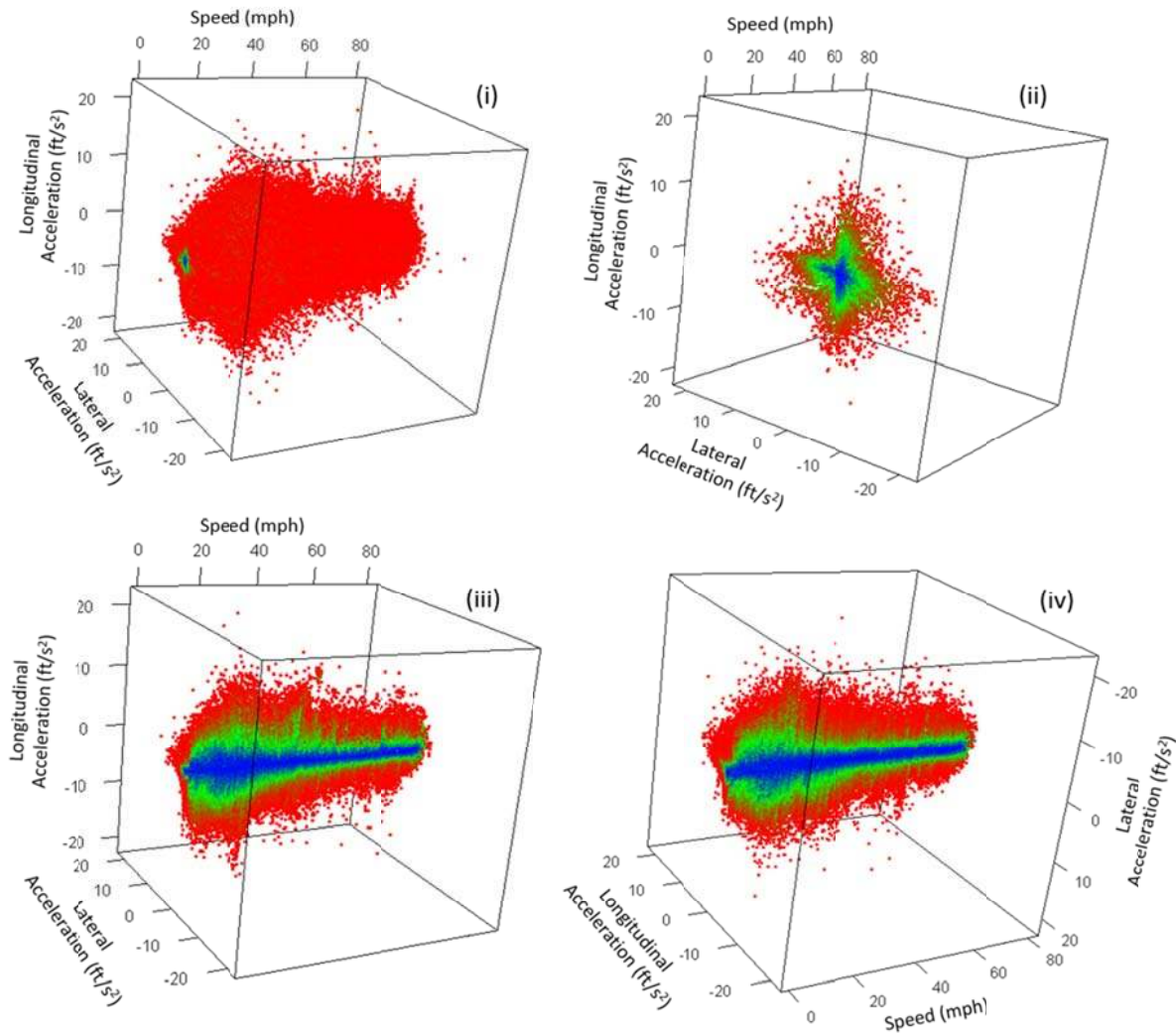
FIGURE 7 Distributions of Longitudinal and Lateral Acceleration in 3-D Space

IDENTIFICATION OF EXTREME EVENTS

Previous studies indicate that extreme driving events (e.g., hard braking or acceleration) are associated with aggressive driving behaviors. However, aggressive driving might be one reason for extreme events. Road situations and vehicle conditions such as obstacles on roads, poor pavements, slippery road surface, sharp curves, and sensitive acceleration or braking systems can also be reasons for extreme driving events. Researchers have given cut-off acceleration values as a threshold for defining extreme (aggressive) driving and calm (normal) driving (Lajunen et al., 1997b; De Vlieger et al., 2000; Langari and Won, 2005; Kim and Choi, 2013). In light of the varying distributions of instantaneous accelerations with different vehicle speeds and in various dimensions (x, y, and z), this study proposes an innovative way to identify extreme events, which are the core information generated for drivers from BSMs and transmitted through V2V and V2I applications. Our previous studies explored extreme acceleration events at various driving speeds (Liu et al., 2014; Wang et al., 2015) in one dimension of vehicle motion, i.e., longitudinal. This study uses data that contains both longitudinal and lateral accelerations.

Extreme events of acceleration in different directions potentially correspond to different warning and control assist messages. For example, if an extreme event is longitudinal acceleration, then a warning might be given to the vehicle ahead. If extreme lateral acceleration to the right is observed, then vehicles on the right hand side, i.e., in the right lane might be warned through V2V applications.

Unlike previous studies that give cut-off values as thresholds regardless of driving situations, this study proposes thresholds that change with speeds. Such changes can account for different driving situations encountered by drivers. In addition, the directional variation of acceleration distributions was considered when defining the thresholds. This study used 95th percentile values in each bin as the thresholds. The thresholds can be customized according to the acceptable levels of driving volatility. Since the thresholds were generated based on possible values in a specific bin, they vary with different speeds and in different directions. Figure 8 presents: (i) an aggregated threshold surface that is enclosed like a cylinder with varying radii at different speeds and directions; (ii) identified extreme acceleration events (i.e., gray dots out of the surface); and (iii) an enlarged image of figure (ii). Note that the threshold surface was fitted using the 95th percentile magnitudes of longitudinal and lateral accelerations within the one speed×sector bin. The radii of the enclosed surface were greater for lower speeds and narrower for high speeds, as shown in Figure 7 (iii) and (iv). Accelerations parallel or perpendicular to the heading directions had greater radii than did those diagonal to heading directions, as shown in Figure 7 (ii).
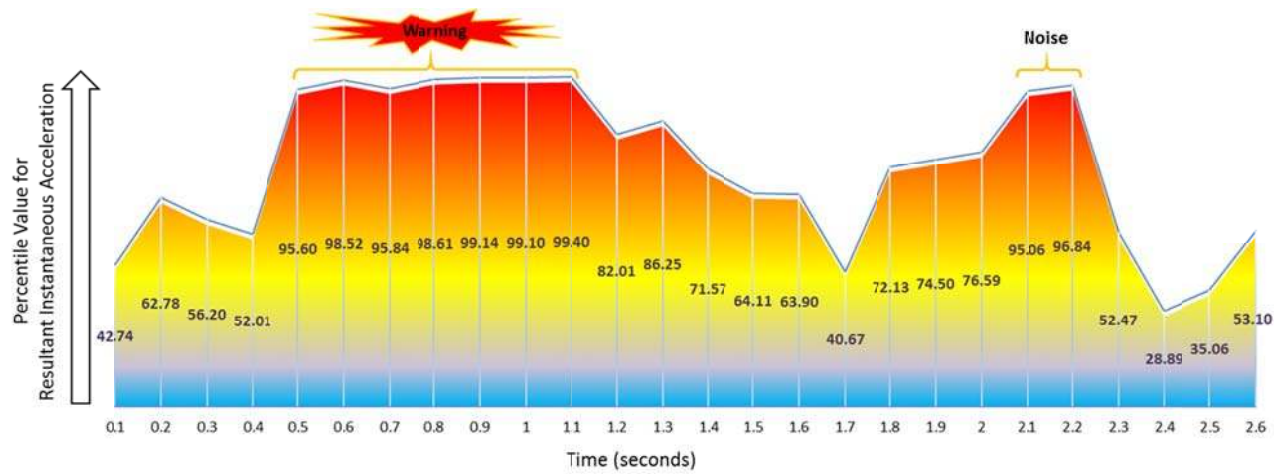
The warnings and control assists were generated based on the 95th percentile thresholds introduced above, and if more than five successive BSMs (> 0.5 seconds) beyond the 95th percentile thresholds were identified together one warning or control assist can be generated as shown in Figure 9. Figure 10 presents a sample trip with identified extreme events (see "Warnings") and their locations on a map.

Extreme events seemed to be located at critical driving conditions, such as sharp turns and complex intersections. A zoomed-in view of the suggested warning and control assist locations shows that this is a six-way intersection consisting of three two-way roads with pedestrian sidewalks, as shown in Figure 10(ii). Three locations generated three separate warnings and control assists. Warning and control assist 1 indicated the potential for poor sight distance caused by roadside plantings on East Madison Street. Warning and control assist 2 pointed out driver behaviors influenced by the intersecting traffic from Packard Street. Warning and control assist 3 was possibly associated with the pedestrians crossing South Division Street. The methodology proposed in this study identifies extreme events and locates them accurately in space.
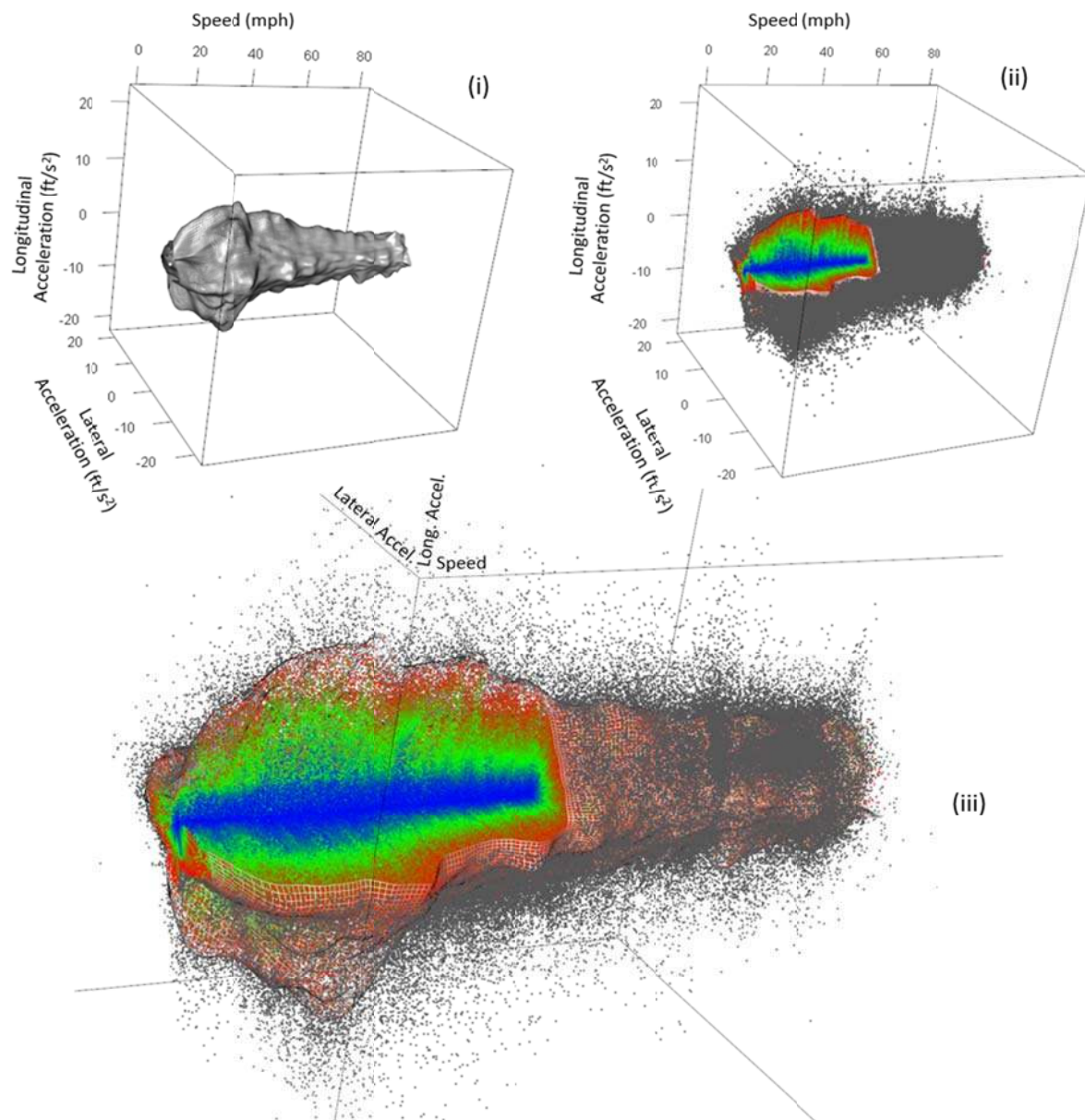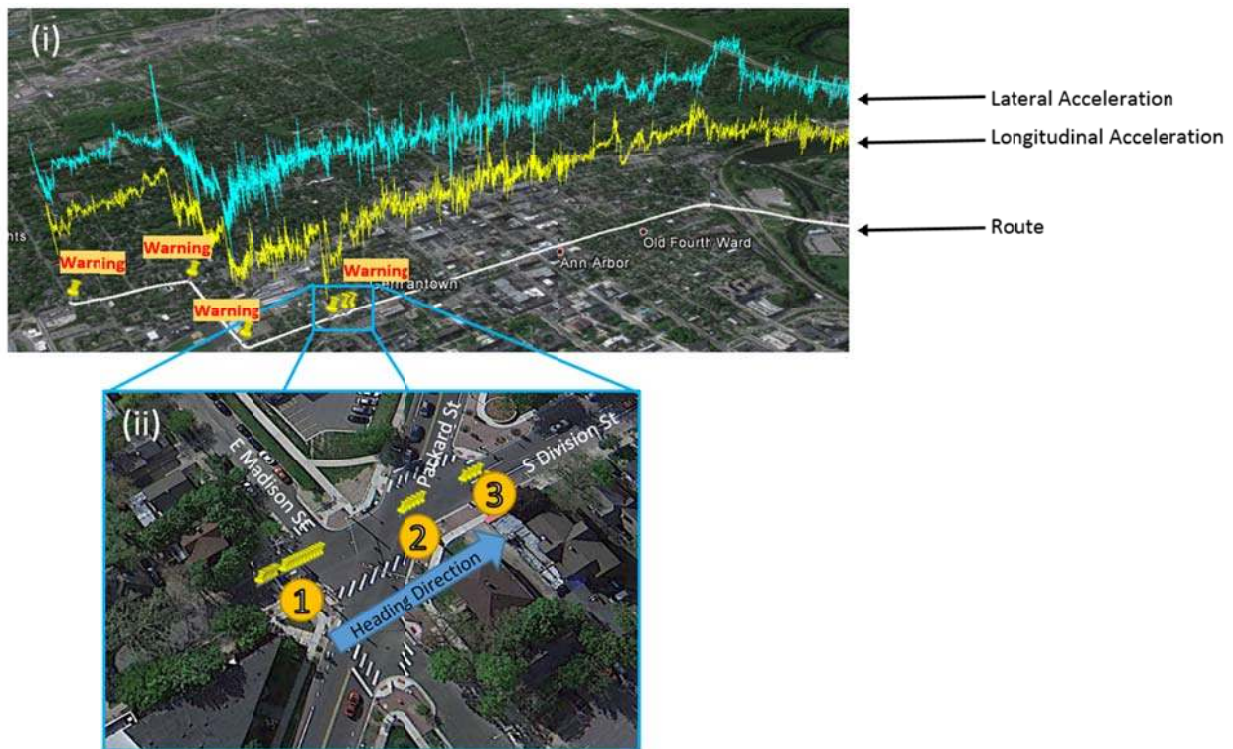
FIGURE 8 Plots of Extreme Acceleration Events (hollowed in side)

Figure 9 Generating Warnings and Control Assists

FIGURE 10 Extreme Events Identified in Space

The warnings and control assists can be further specified based on the directions of accelerations. Figure 11 presents examples of sending warnings/control assists to *Remote Vehicles* "B" and "C" that are proximate to the *Host Vehicle* "A". For example, if an event of extreme longitudinal deceleration (e.g., hard braking) occurs for the host vehicle, a rear-end collision warning or control assist can be sent to the Vehicle "B" that is travelling behind the host vehicle, in the same lane and within the warning range. If an event of extreme lateral acceleration (e.g., fast turn to the right) occurs for the host vehicle, then a side-collision warning or control assist may be given to the Vehicle "C" that is travelling on the right lane and might be affected.

Note that, such warnings and control assists are suggested for the drivers of surrounding connected vehicles. For the host vehicle driver, these warnings and control assists may not be that helpful as the extreme events have already occurred in most cases. To assist the host vehicle driver, two strategies can be considered: 1) providing the host vehicle driver post-event or post-trip feedback about their driving performance. For example, if the number of extreme events on a specific trip is excessive, before the trip ends, then the driver may be alerted to the fact and advised to drive more calmly. 2) The host vehicle driver's behaviors can be predicted in the short-term (a few seconds) by using algorithms, such as Markov decision chains (Khattak et al., 2015). If this can be done successfully, then the host vehicle driver can be given alerts and warnings in advance.

In addition to assisting driver behaviors, the idea of identifying extreme events is also potentially helpful for proactive traffic management. Hazards can be identified on road segments or intersections (such as poor sight distance and a large number of sharp turns) if many extreme events are frequently observed at the same location.
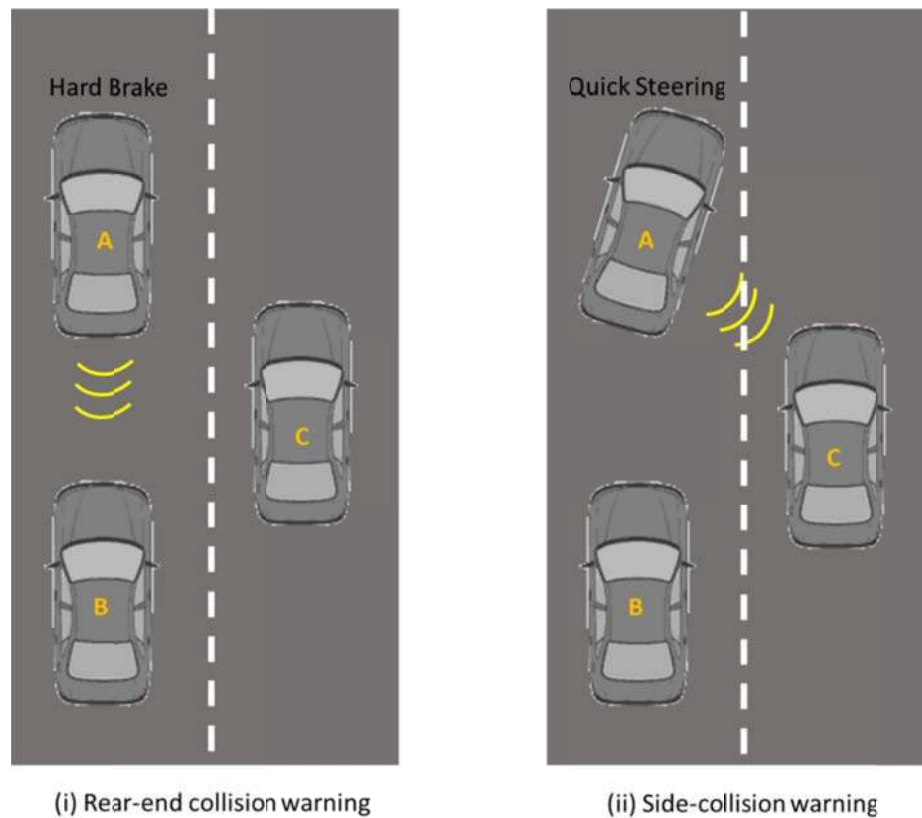

Figure 11: Examples of generating warnings that are dedicated to target vehicles

UNDERSTANDING FREQUENCY OF EXTREME EVENTS

One question to be answered after the identification of extreme acceleration events was whether these extreme events made sense in terms of vehicle maneuvering and driving conditions. To understand the correlates of extreme events, a simple regression model was estimated. The correlates included vehicle maneuvering decisions (e.g., position of accelerator, brakes, and cruise control during trip), and the instantaneous driving contexts, e.g., surrounding objects and

distance to closest objects. Modeling provides clarity on what factors are associated with higher frequency of extreme events, and helps anticipate the number of extreme events given a set of driving-related factors. Further, driving safety strategies may be developed to avoid certain driving conditions that are associated with more extreme events.

Negative binomial (NB) models were estimated (also called extra-Poisson or Poisson-gamma model) to explore correlates of extreme event frequency. The model assumes that the observed frequency of extreme events identified from BSMs during one trip follows a NB distribution (Lord and Mannering, 2010; Guo and Fang, 2013):

$Y_i \sim NB\ (E_i \lambda_i, \alpha)$

Where, λ_i = expected rate of extreme events (i.e., frequency of extreme events over the duration or length of one trip); E_i = trip length (in miles) or duration (in minutes); and α = NB over-dispersion parameter. A larger value of α implies greater over-dispersion in data. If $\alpha = 0$, then the data follows a Poisson distribution (where mean = variance). In such a situation, the Poisson and Negative Binomial model provide identical estimates. If α is significantly greater than 0, a NB model is preferred. Formally, λ_i can be viewed as a log link function of a set of independent variables:

$Log\ (\lambda_i) = X_i \beta$ Equation (6)

Where, X_i = the matrix of independent variables; β = the vector of estimated associations between independent variables and the dependent variable (i.e., frequency of extreme events during one trip).

Table 4 shows the descriptive statistics of variables for a trip. On average, there are 7.23 extreme acceleration events per trip. The trip with the greatest number had 39 extreme events while 35 (out of 155) trips had no extreme events. Note that a zero-inflated NB model may be estimated for future studies (Anastasopoulos et al., 2010), if the sample size is expanded and excessive number of trips have zero extreme events. The longest extreme event lasted 19.1 seconds, as shown in Table 4. Noticeably, there were some extremely short trips (less than 1 minute) because drivers did not always initiate data collection at the beginning of their trips. For those trips, the data relates to one segment of a trip. Distributions of other variables seemed reasonable, based on error checking.

TABLE 4 Descriptive Statistics of Variables at Trip-Level (N=155)

	Variable	Mean	Std. Dev.	Min	Max
Warning	Number of extreme acceleration events	7.23	8.22	0.00	39.00
	Total extreme acceleration event duration (second)	7.96	10.14	0.00	58.00
	Average duration per extreme events (second)	0.80	0.62	0.00	3.35
	Longest extreme acceleration events duration (second)	2.11	2.67	0.00	19.10
Trip Attribute	Average speed (mph)	33.55	15.01	2.74	71.22
	Speed variance	228.96	158.60	0.06	728.58
	Trip length (mile) *	6.68	9.46	0.03	72.14
	Trip duration (minute) *	10.41	10.92	0.07	60.78
Vehicle Maneuvering	Average accelerator pedal displace (%)	13.97%	11.79%	0.00%	64.30%
	Brake pedal (engaged) (%)	23.08%	16.39%	0.00%	86.02%
	Cruise control (engaged) (%)	30.16%	37.78%	0.00%	100.00%
	Turn signal (left) (%)	4.71%	10.82%	0.00%	97.54%
	Turn signal (right) (%)	2.38%	5.70%	0.00%	60.03%
	Number of total turns	3.06	3.99	0.00	21.00
Driving Context	Average number of objects	1.61	0.95	0.00	3.91
	Average distance to the closest object (ft)	163.76	149.98	21.33	839.69
	Average relative speed of the closest object (mph)	-4.38	10.80	-80.86	12.03

*: There are some extreme short trips because drivers did not always initiate the data collection at the beginning of the trips. For those trips, the data refer to one segment of a trip.

Multicollinearity may exist between independent variables (e.g., trip length and trip duration), which may lead the regression model estimates to become unstable (Farrar and Glauber, 1967). In this study, the variance inflation factors (VIFs) were calculated to examine the multicollinearity of independent variables (Chatterjee and Hadi, 1986). The VIF for independent variable x_j can be calculated by

$$VIF\left(x_j\right) = \frac{1}{1 - \hat{R}_j^2} \qquad \text{Equation (7)}$$

Where $\hat{R}_j^2$ is the square of the multiple correlation coefficients that are from the regression when x_j is regressed with intercept against all the other independent variables. A variable whose VIF value is greater than 10 is possibly redundant in a regression model, since this variable may be a function of other independent variables. Based on VIF calculations, the variable of trip length had the greatest VIF value (= 9.80). A reasonable explanation is that the trip length and trip duration and average speed are highly correlated, noting that speed = trip distance/trip duration. It was decided to remove trip length from the model, and as a result all variables had a VIF value smaller than 5. This is an indication that multicollinearity between the remaining variables is not problematic.

Although some variables may not be correlated with other independent variables (i.e., VIF is significantly less than 10), they can be statistically insignificant. That is they may not be correlated significantly with the dependent variable (i.e., the frequency of extreme events) and explaining little variation in data. To deal with such situations, a backward stepwise variable elimination procedure was conducted to eliminate the variables whose estimates that were not significant at 90% confidence level (StataCorp, 2013). Variables that were kept in the final model specification have an estimate that was at least marginally significant at 90% confidence

level. The variables that were included in the final NB model include speed variance, trip duration, percent of driving time when the brake pedal was engaged, number of turns, average distance to the closest object, and average relative speed of the closest object.

Sometimes the NB model has an exposure or offset variable such as trip length or trip duration. When using an exposure variable, modelers typically assume that the likelihood of the events is changing proportionately over the duration of exposure while controlling for other variables. However, trip length or trip duration are not simply measures of exposure that are linearly related to extreme events. If a trip is very short, then the trip is likely to be undertaken on local roads and if the trip is very long, then the trip is likely to include freeway travel. Local driving may be associated with proportionately more extreme events compared with freeway driving. Therefore, to explore non-linear relationships, this study treated trip length and duration as independent variables instead of using them as exposure variables.

Variables such as the speed variance and the brake engagement may be endogenous to the dependent variable, i.e., frequency of extreme events. A Durbin–Wu–Hausman test was conducted to test for endogeneity (Davidson and MacKinnon, 1993). Results showed that these variables were not endogenous at 95% confidence level.

Table 5 shows the model specification process, i.e., variable selection and removal procedures, and the final NB modeling results. Overall, the modeling results were reasonable, providing insights about the correlates of extreme events embedded in BSMs. The overall goodness-of-fit was rather weak, since the phenomena are complex, the model is limited to a relatively small sample size (N=155), and a small number of variables. The estimated coefficients and goodness of fit will likely change as more data become available.

A likelihood ratio test showed that the over-dispersion parameter α is significantly greater than zero ($\chi^2 = 528.44$, *p*-value = 0.000), which validates the appropriateness of NB regression instead of Poisson regression. As expected, the greater variation of speeds was correlated with the more extreme events, since the identification of extreme events was based on accelerations which are the change rates of speeds. For vehicle maneuvering, more time spent on braking was associated with more extreme acceleration events. Variables that are positively correlated with more warnings include: 1) if a trip had many turns (10% level), pointing to the potential for making extreme maneuvers such as sharp turns at such locations; 2) the distance to the closest object because proximity to objects may induce greater volatility; 3) higher relative speed of the closest object; 4) longer trip duration; and 5) more brake pedal engagement. Among all factors examined, the brake pedal engagement had the greatest association with the frequency of extreme events. One percent increase in time spent on braking during a trip is associated with a 272.5% increased chances of extreme events.

TABLE 5 Negative Binomial Models for Frequency of Extreme Events
(Y=number of extreme events or suggested warnings/control assists during one trip)

Variable	VIF Calculation		Stepwise Elimination		Final NB model		
	1st Round	2nd Round	P-value	Elimination order	β	P-value	Incidence Rate Ratio
Average speed (mph)	4.87	3.70	0.363	4			
Speed variance	1.60	1.57			0.001**	0.044	1.001
Trip length (miles)	9.80						
Trip duration (minutes)	9.47	2.06			0.046***	0.000	1.047
Average accelerator pedal displacement (%)	1.70	1.70	0.539	3			
Brake pedal (engaged) (%)	4.14	4.08			1.315**	0.044	3.725
Cruise control (engaged) (%)	1.72	1.72	0.346	5			
Turn signal (left) (%)	1.37	1.36	0.847	1			
Turn signal (right) (%)	1.11	1.11	0.742	2			
Number of turns	1.94	1.91			0.049*	0.074	1.051
Average number of objects	1.36	1.36	0.156	6			
Average distance to the closest object (ft)	1.50	1.50			0.002**	0.019	1.002
Average relative speed of the closest object (mph)	1.49	1.48			0.026*	0.055	1.026
Constant					0.319	0.230	
SUMMARY STATISTICS α					1.005[#]		
Number of Observations					155		
Log Likelihood at zero, *LL(Constant only)*					-465.491		
Log Likelihood at convergence, *LL(Model)*					-437.339		
AIC					890.678		
Log Likelihood Chi-square					56.3		
Prob > χ^2					0.000		
Pseudo McFadden's R^2					0.061		

Note: ***= significant at a 99% confidence level; ** = significant at a 95% confidence level; * = marginally significant at a 90% confidence level; [#]the probability of the dispersion parameter $\alpha = 0$ is virtually zero, indicating NB model performs better than Poisson model. Pseudo McFadden's R^2 = 1 – *LL(Model)/LL(Constant only)*. Incidence Rate Ratio = e^β.

Given the potential concerns: 1) the excessive zeros in observations (35 trips, out of 155, had no extreme events), and 2) significant unobserved heterogeneity in correlates across observations, zero inflated models and random parameter models were estimated. Table 6 presents the zero-inflated model and the random-parameter models. Previous studies have provided good explanations and justifications for estimating zero-inflated models (Anastasopoulos et al., 2010) and random-parameter models (Milton et al., 2008; Anastasopoulos and Mannering, 2009, 2011; Russo et al., 2014). Compared with the NB model, zero-inflated model seems to be better, given the statistically significant (at 95% confidence level) Vuong statistic (Anastasopoulos et al., 2010). However, most estimates are not statistically significant in the model (5% level). Also, in terms of the estimated Incident Rate Ratios, zero-inflated model parameters are similar to the simple NB model. In addition, the random-parameter model showed that three variables (brake pedal engagement, number of turns and average relative speed) should be estimated using random-parameters. However, the random-parameter estimates for these variables are not statistically significant (5% level). Further, the fixed-parameters in this model are very similar to the estimates from simple NB model, implying that there may not be substantial heterogeneity in the correlates of extreme events. The correlates estimated in the simple NB model are constant across observations. This study did not

demonstrate substantial value added due to zero inflated or random-parameter models. However, this is not to say that such models should not be explored in other studies. Additional data, when available, may be able to better demonstrate the value of such models.

TABLE 6 Zero-Inflated and Random-Parameter Models for Frequency of Extreme Events
(Y=number of extreme events or suggested warnings/control assists during one trip)

Variables	Zero-Inflated NB Model			Random-Parameter Model		
	β	P-value	Incidence Rate Ratio	β	P-value	Incidence Rate Ratio
	Count State			Fixed-Parameter		
Speed variance	0.001	0.180	1.001	0.001*	0.086	1.001
Trip duration (minutes)	0.039***	0.000	1.039	0.046***	0.001	1.047
Brake pedal (engaged) (%)	0.778	0.219	2.176	1.315*	0.088	3.725
Number of turns	0.036	0.151	1.036	0.049	0.127	1.050
Average distance to the closest object (ft)	0.001	0.113	1.001	0.002**	0.050	1.002
Average relative speed of the closest object (mph)	0.010	0.533	1.010	0.026	0.106	1.026
Constant	0.844***	0.002		0.319	0.257	
	Zero State			Random-Parameter[a]		
Speed variance	-0.013	0.307		NRP[b]		
Trip duration (minutes)	-2.261**	0.044		NRP		
Brake pedal (engaged) (%)	-2.866	0.318		2.02E-02	0.953	
Number of turns	1.003	0.144		3.79E-04	0.986	
Average distance to the closest object (ft)	-0.014*	0.086		NRP		
Average relative speed of the closest object (mph)	-0.244*	0.073		1.91E-05	0.966	
Constant	5.592**	0.025				
SUMMARY STATISTICS						
α	0.791[#]			0.995[#]		
Vuong statistic	2.790***	0.003				
Number of observations	155			155		
Log Likelihood at zero, *LL(Constant only)*	-441.457			-465.491		
Log Likelihood at convergence, *LL(Model)*	-422.782			-437.338		
Prob > $\chi2$	0.000			0.000		
Pseudo McFadden's R^2	0.042			0.060		

Note: ***= significant at a 99% confidence level; ** = significant at a 95% confidence level; * = marginally significant at a 90% confidence level; [#]the probability of the dispersion parameter $\alpha = 0$ is virtually zero, indicating NB model performs better than Passion model; [a] = Standard deviations of random-parameter; [b]= No estimable random parameter. Pseudo McFadden's $R^2 = 1 – LL(Model)/LL(Constant only)$. Incidence Rate Ratio = e^{β}.

LIMITATIONS

Data used in this study were BSMs sent and received by connected vehicles through V2V and V2I applications. Participating vehicles equipped with Data Acquisition Systems collected the BSMs. Thus, the extent of measurement errors in the data was unknown, although the results from data visualization and modeling were reasonable. In addition, the Data Acquisition Systems captured information explored in this study does not represent a complete set of driving contexts or operating environments. Additional information about roadway features and traffic signals/signs may need to be integrated with information from Data Acquisition Systems when

warnings or control assistances are about to be given. The warnings or control assists are suggested merely based on vehicle motion information, i.e., extreme acceleration or braking events. Such extreme events can benefit from validation and verification, e.g., by analyzing other information, such as videos for the driver's gestures or hand positions on steering wheels, eye tracking for distraction and fatigue, and sensor data of how a driver moves the accelerator and brake pedals.

A threat to the validity of this study is the limited sample size. Only one-day sample data from the SPMD is publicly available in Research Data Exchange (RDE, https://www.its-rde.net/home). Consequently, this study can be regarded as an exploration of BSMs transmitted by real-world connected vehicle technologies. As more data become available, the methodology of this study can be applied directly for a broader exploration, and the results of this study can be validated by expanded sample data.

CONCLUSION

The connected vehicle technologies discussed in this paper are capable of transmitting high-frequency data between vehicles and infrastructure, which has the potential to improve mobility, safety, energy consumption, and the environment. In this context, it is important to maximize the value embedded in data generated by the entire ecosystem of vehicles and infrastructure to support driver decision-making. SPMD provides data on basic safety messages, which are the core data sent and received by connected vehicles and infrastructures through V2V and V2I applications. The content of BSMs describe a vehicle's position (latitude, longitude, and elevation) and motion (heading, speed, and acceleration) (Henclewood, 2014). BSMs also contains data pertaining to the vehicle's component status (lights, brakes, wipers) and vehicle safety information (path history, events) (Henclewood, 2014). The raw BSMs are complex and not informative to drivers. This study proposes a data analytic methodology to extract critical information from raw BSMs. The information can be provided to drivers and inform them about their driving behaviors or about dangers in surrounding roadway environments. The information is simple and informative, and helps drivers make informed decisions. The research is timely and has long-term value, as connected and automated vehicles are likely to have substantial impacts throughout the world, and have seen substantial research activity.

Our previous studies have explored the critical information embedded in vehicle trajectory data, which were collected through GPS devices (Liu et al., 2014; Wang et al., 2015). This study extended the methodology of extracting critical information from BSMs transmitted between connected vehicles. This study established a fundamental understanding of instantaneous driving decisions by investigating two-dimensional instantaneous accelerations, i.e., longitudinal and lateral accelerations. Instantaneous driving volatility was visualized, and it clearly showed that driving behavior is strongly associated with driving contexts, whether driving on local roads or freeways. This study untangled the relationship between speeds and acceleration through the distribution of instantaneous accelerations at different speeds. Higher speeds (>50 mph) are associated with smaller acceleration magnitudes, which is consistent with (Ahn et al., 2002). This study further revealed a nonlinear relationship between acceleration and speed in real-life driving situations. The lozenge shaped joint distribution implies that longitudinal and lateral acceleration hardly reached a large magnitude simultaneously. In terms of magnitude, longitudinal and lateral accelerations seemed inversely correlated.

This research presents an original idea, which is to establish context-relevant alert,

warning, and control assist thresholds based on extreme event information embedded in BSMs. Most previous studies give fixed cut-off values for thresholds regardless of driving situations (Lajunen et al., 1997b; De Vlieger et al., 2000; Langari and Won, 2005; Kim and Choi, 2013). However, some of the thresholds for warnings and control assists may be flexible and can change with speeds to account for different driving situations and contexts. In addition, the directional variation of acceleration distributions is considered in establishing the thresholds. Information about extreme driving decisions can be used for control assists and provided as feedback to drivers in real time to help them shift to calmer driving, and transmitted to other drivers surrounding through V2V applications to warn them about potential dangers.

Results from rigorous statistical modeling revealed that the extreme events identified from BSMs are strongly associated with trip attributes, driver maneuvering, and driving contexts. The results provide correlates of extreme events, indicating situations where warnings and control assists may be provided to drivers. Such conditions include situations in which a driver has to step on brakes frequently (possibly due to unstable speed conditions and congested flow), and a trip which involves many turns at intersections. Note that the modeling results were limited to a relatively small sample of 155 observations and a small number of variables. The estimated results may change as more data and variables become available.

This study contributes by making sense of high-frequency geo-referenced connected vehicle data, and extracts critical information about extreme events from new data sources created by communications between connected vehicles. Connected vehicles are a relatively new and emerging area of research activity in intelligent transportation systems, with strong interest from a wide audience that includes government agencies, auto makers, practitioners and researchers who are interested in implementing connected vehicles.

The findings of this study are relevant to incorporation of alerts, warnings, and control assists in V2V and V2I applications of connected vehicles. Such applications can help drivers identify situations where surrounding drivers are volatile and they may avoid dangers by taking defensive actions. For example, rear-end or side-impact collision warnings can be given to drivers who are in close proximity of other drivers undergoing extreme longitudinal or lateral accelerations, respectively. Applications can also be developed for departments of transportation to identify locations with high density of extreme events (Abdel-Aty et al., 2007; Cheng et al., 2015). Hazards in road networks can be identified proactively on certain road segments or at intersections (such as poor sight distance and higher incidence of sharp turns) that have higher densities of extreme events.

ACKNOWLEDGEMENT

for modeling. The views expressed in this paper are those of the authors, who are responsible for the facts and accuracy of information presented herein.

REFERENCES
Abdel-Aty, M., A. Pande, C. Lee, V. Gayah, and C. D. Santos. 2007. Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. Journal of Intelligent Transportation Systems 11: 107-120.

Åberg, L., L. Larsen, A. Glad, and L. Beilinsson. 1997. Observed vehicle speed and drivers' perceived speed of others. Applied Psychology 46: 287-302.

Abid, H., T. C. Chung, S. Lee, and S. Qaisar. 2012. Performance analysis of lte smartphones-based vehicle-to-infrastrcuture communication. In: Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference on. p 72-78.

Ahn, K., H. Rakha, A. Trani, and M. Van Aerde. 2002. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. Journal of Transportation Engineering 128: 182-190.

Anastasopoulos, P. C., S. Labi, A. Bhargava, C. Bordat, and F. L. Mannering. 2010. Frequency of change orders in highway construction using alternate count-data modeling methods. Journal of Construction Engineering and Management 136: 886-893.

Anastasopoulos, P. C., and F. L. Mannering. 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accident Analysis & Prevention 41: 153-159.

Anastasopoulos, P. C., and F. L. Mannering. 2011. An empirical assessment of fixed and random parameter logit models using crash-and non-crash-specific injury data. Accident Analysis & Prevention 43: 1140-1147.

Bagheri, E., B. Mehran, and B. Hellinga. 2015. Real-time Estimation Of Saturation Flow Rates For Dynamic Traffic Signal Control Using Connected Vehicle Data. In: Transportation Research Board 94th Annual Meeting

Bridgelall, R. 2013. Connected vehicle approach for pavement roughness evaluation. Journal of Infrastructure Systems 20: 04013001.

Chan, C.-Y. 2011. Connected vehicles in a connected world. In: VLSI Design, Automation and Test (VLSI-DAT), 2011 International Symposium on. p 1-4.

Chatterjee, S., and A. S. Hadi. 1986. Influential observations, high leverage points, and outliers in linear regression. Statistical Science: 379-393.

Cheng, L., B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige. 2007. Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 GHz dedicated short range communication (DSRC) frequency band. Selected Areas in Communications, IEEE Journal on 25: 1501-1516.

Cheng, Y., M. Zhang, and D. Yang. 2015. Automatic Incident Detection for Urban Expressways Based on Segment Traffic Flow Density. Journal of Intelligent Transportation Systems 19: 205-213.

Chou, C.-M., C.-Y. Li, W.-M. Chien, and K.-c. Lan. 2009. A feasibility study on vehicle-to-infrastructure communication: WiFi vs. WiMAX. In: Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09. Tenth International Conference on. p 397-398.

Christofa, E., J. Argote, and A. Skabardonis. 2013. Arterial queue spillback detection and signal control based on connected vehicle technology. Transportation Research Record: Journal of the Transportation Research Board: 61-70.

Chrysler, S. T., J. M. Cooper, and D. Marshall. 2015. The Cost of Warning of Unseen Threats: Unintended Consequences of Connected Vehicle Alerts. In: Transportation Research Board 94th Annual Meeting

Davidson, R., and J. G. MacKinnon. 1993. Estimation and inference in econometrics. OUP Catalogue.

De Vlieger, I., D. De Keukeleere, and J. Kretzschmar. 2000. Environmental effects of driving behaviour and congestion related to passenger cars. Atmospheric Environment 34: 4649-4655.

Doecke, S., A. Grant, and R. W. Anderson. 2015. The real-world safety potential of connected vehicle technology. Traffic injury prevention 16: S31-S35.

Du, L., and H. Dao. 2015. Information Dissemination Delay in Vehicle-to-Vehicle Communication Networks in a Traffic Stream. Intelligent Transportation Systems, IEEE Transactions on 16: 66-80.

Fagnant, D. J., and K. Kockelman. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transportation Research Part A: Policy and Practice 77: 167-181.

Farrar, D. E., and R. R. Glauber. 1967. Multicollinearity in regression analysis: the problem revisited. The Review of Economic and Statistics: 92-107.

Feng, Y., S. Khoshmagham, M. Zamanipour, and K. L. Head. 2015. A Real-time Adaptive Signal Phase Allocation Algorithm in a Connected Vehicle Environment. In: Transportation Research Board 94th Annual Meeting

Genders, W., and S. N. Razavi. 2015. Impact of Connected Vehicle on Work Zone Network Safety through Dynamic Route Guidance. Journal of Computing in Civil Engineering: 04015020.

Goel, S., T. Imielinski, and K. Ozbay. 2004. Ascertaining viability of WiFi based vehicle-to-vehicle network for traffic information dissemination. In: Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on. p 1086-1091.

Goodall, N. J., B. L. Smith, and B. Park. 2013. Traffic Signal Control with Connected Vehicles. Transportation Research Record: Journal of the Transportation Research Board 2381: 65-72.

Goodall, N. J., B. L. Smith, and B. B. Park. 2014. Microscopic estimation of freeway vehicle positions from the behavior of connected vehicles. Journal of Intelligent Transportation Systems: 1-10.

Guler, S. I., M. Menendez, and L. Meier. 2014. Using connected vehicle technology to improve the efficiency of intersections. Transportation Research Part C: Emerging Technologies 46: 121-131.

Guo, F., and Y. Fang. 2013. Individual driver risk assessment using naturalistic driving data. Accident Analysis & Prevention 61: 3-9.

Haglund, M., and L. Åberg. 2000. Speed choice in relation to speed limit and influences from other drivers. Transportation Research Part F: Traffic Psychology and Behaviour 3: 39-51.

Henclewood, D. 2014. Safety Pilot Model Deployment – One Day Sample Data Environment Data Handbook. Research and Technology Innovation Administration, US Department of Transportation, McLean, VA.

Hill, C. J., and J. K. Garrett. 2011. AASHTO connected vehicle infrastructure deployment analysis.

Hoedemaeker, M., and K. A. Brookhuis. 1998. Behavioural adaptation to driving with an adaptive cruise control (ACC). Transportation Research Part F: Traffic Psychology and Behaviour 1: 95-106.

Hu, J., B. Park, and A. Parkany. 2014. Transit Signal Priority with Connected Vehicle Technology. Transportation Research Record: Journal of the Transportation Research Board: 20-29.

Hu, J., B. B. Park, and Y.-J. Lee. 2015. Coordinated transit signal priority supporting transit progression under Connected Vehicle Technology. Transportation Research Part C: Emerging Technologies 55: 393-408.

Jin, Q., G. Wu, K. Boriboonsomsin, and M. Barth. 2012. Advanced intersection management for connected vehicles using a multi-agent systems approach. In: Intelligent Vehicles Symposium (IV), 2012 IEEE. p 932-937.

Khattak, A., S. Nambisan, and S. Chakraborty. 2015. Study of Driving Volatility in Connected and Cooperative Vehicle Systems. National Science Foundation.

Kianfar, J., and P. Edara. 2013. Placement of Roadside Equipment in Connected Vehicle Environment for Travel Time Estimation. Transportation Research Record: Journal of the Transportation Research Board: 20-27.

Kim, E., and E. Choi. 2013. Estimates of Critical Values of Aggressive Acceleration from a Viewpoint of Fuel Consumption and Emissions. In: 2013 Transportation Research Board Annual Meeting, Washington DC

Kishore Kamalanathsharma, R., and H. A. Rakha. 2014. Leveraging connected vehicle technology and telematics to enhance vehicle fuel efficiency in the vicinity of signalized intersections. Journal of Intelligent Transportation Systems: 1-12.

Koulakezian, A., and A. Leon-Garcia. 2011. CVI: Connected vehicle infrastructure for ITS. In: Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on. p 750-755.

Lajunen, T., A. Corry, H. Summala, and L. Hartley. 1997a. Impression management and self-deception in traffic behaviour inventories. Personality and individual differences 22: 341-353.

Lajunen, T., A. Corry, H. Summala, and L. Hartley. 1998. Cross-cultural differences in drivers' self-assessments of their perceptual-motor and safety skills: Australians and Finns. Personality and Individual Differences 24: 539-550.

Lajunen, T., J. Karola, and H. Summala. 1997b. Speed and acceleration as measures of driving style in young male drivers. Perceptual and motor skills 85: 3-16.

Langari, R., and J.-S. Won. 2005. Intelligent energy management agent for a parallel hybrid vehicle-part I: system architecture and design of the driving situation identification process. Vehicular Technology, IEEE Transactions on 54: 925-934.

Lee, J., and B. Park. 2012. Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment. Intelligent Transportation Systems, IEEE Transactions on 13: 81-90.

Li, J.-Q., K. Zhou, S. Shladover, and A. Skabardonis. 2013. Estimating Queue Length Under Connected Vehicle Technology: Using Probe Vehicle, Loop Detector, and Fused Data. Transportation Research Record: Journal of the Transportation Research Board: 17-22.

Li, Y., M. Zhao, and W. Wang. 2011. Intermittently connected vehicle-to-vehicle networks: detection and analysis. In: Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE. p 1-6.

Liu, J. 2015. Driving Volatility in Instantaneous Driving Behaviors: Studies Using Large-Scale Trajectory Data.

Liu, J., A. Khattak, and X. Wang. 2015. The role of alternative fuel vehicles: Using behavioral and sensor data to model hierarchies in travel. Transportation Research Part C: Emerging Technologies 55: 379-392.

Liu, J., X. Wang, and A. Khattak. 2014. Generating Real-Time Driving Volatility Information 2014 World Congress on Intelligent Transport Systems, Detroit, MI.

Liu, J., X. Wang, and A. Khattak. 2016. Customizing driving cycles to support vehicle purchase and use decisions: Fuel economy estimation for alternative fuel vehicle users. Transportation Research Part C: Emerging Technologies 67: 280-298.

Lord, D., and F. Mannering. 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice 44: 291-305.

Milton, J. C., V. N. Shankar, and F. L. Mannering. 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis & Prevention 40: 260-266.

Mosyagin, J. 2010. Using 4G wireless technology in the car. In: Transparent Optical Networks (ICTON), 2010 12th International Conference on. p 1-4.

Moylan, E., and A. Skabardonis. 2015. Reliability-and Median-Based Identification of Toll Locations in a Connected Vehicle Context. In: Transportation Research Board 94th Annual Meeting

NHTSA. 2000. Resource Guide Describes Best Practices For Aggressive Driving Enforcement http://www.nhtsa.gov/About+NHTSA/Traffic+Techs/current/Resource+Guide+Describes+Best+Practices+For+Aggressive+Driving+Enforcement Accessed May 15th 2014.

NHTSA. 2008. National Motor Vehicle Crash Causation Survey: Report to Congress. National Highway Traffic Safety Administration Technical Report DOT HS 811: 059.

Noble, A. M., S. B. Mclaughlin, Z. R. Doerzaph, and T. A. Dingus. 2014. Crowd-sourced Connected-vehicle Warning Algorithm using Naturalistic Driving Data.

Nusser, R., and R. M. Pelz. 2000. Bluetooth-based wireless connectivity in an automotive environment. In: Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd. p 1935-1942.

Office, T. I. T. S. J. P. 2014. Connected Vehicle Research in the United States. http://www.its.dot.gov/connected_vehicle/connected_vehicle_research.htm.

Olia, A., H. Abdelgawad, B. Abdulhai, and S. N. Razavi. 2014. Assessing the Potential Impacts of Connected Vehicles: Mobility, Environmental, and Safety Perspectives. In: Transportation Research Board 93rd Annual Meeting

Osman, O. A., J. Codjoe, and S. Ishak. 2015. Impact of Time-to-Collision Information on Driving Behavior in Connected Vehicle Environments Using A Driving Simulator Test Bed. Journal of Traffic and Logistics Engineering Vol 3.

Paleti, R., N. Eluru, and C. Bhat. 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Accident Analysis & Prevention 42: 1839-1854.

Qin, W. B., and G. Orosz. 2013. Digital effects and delays in connected vehicles: linear stability and simulations. In: ASME 2013 Dynamic Systems and Control Conference. p V002T030A001-V002T030A001.

Russo, B. J., P. T. Savolainen, W. H. Schneider, and P. C. Anastasopoulos. 2014. Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model. Analytic methods in accident research 2: 21-29.

SAE. 2014. Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems, SAE International.

Sengupta, R., S. Rezaei, S. E. Shladover, D. Cody, S. Dickey, and H. Krishnan. 2007. Cooperative collision warning systems: Concept definition and experimental implementation. Journal of Intelligent Transportation Systems 11: 143-155.

Simons-Morton, B. G., K. Cheon, F. Guo, and P. Albert. 2013. Trajectories of kinematic risky driving among novice teenagers. Accident Analysis & Prevention 51: 27-32.

Simons-Morton, B. G., M. C. Ouimet, Z. Zhang, S. E. Klauer, S. E. Lee, J. Wang, P. S. Albert, and T. A. Dingus. 2011. Crash and risky driving involvement among novice adolescent drivers and their parents. American journal of public health 101: 2362.

Simons-Morton, B. G., Z. Zhang, J. C. Jackson, and P. S. Albert. 2012. Do elevated gravitational-force events while driving predict crashes and near crashes? American journal of epidemiology 175: 1075-1079.

StataCorp. 2013. Stepwise Estimation.

Sugiura, A., and C. Dermawan. 2005. In traffic jam IVC-RVC system for ITS using Bluetooth. Intelligent Transportation Systems, IEEE Transactions on 6: 302-313.

Tian, D., Y. Yuan, J. Zhou, Y. Wang, G. Lu, and H. Xia. 2013. Real-time vehicle route guidance based on connected vehicles. In: Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing. p 1512-1517.

Wang, X., A. Khattak, J. Liu, G. Masghati-Amoli, and S. Son. 2015. What is the Level of Volatility in Instantaneous Driving Decisions? Transportation Research Part C: Emerging Technologies.

Wu, G., K. Boriboonsomsin, H. Xia, and M. Barth. 2014. Supplementary Benefits from Partial Vehicle Automation in an Ecoapproach and Departure Application at Signalized Intersections. Transportation Research Record: Journal of the Transportation Research Board: 66-75.

Zeng, X., K. N. Balke, and P. Songchitruksa. 2012. Potential Connected Vehicle Applications to Enhance Mobility, Safety, and Environmental Security, Southwest Region University Transportation Center, Texas Transportation Institute, Texas A&M University System.

Zhang, L., and G. Orosz. 2013. Designing network motifs in connected vehicle systems: delay effects and stability. In: ASME 2013 Dynamic Systems and Control Conference. p V003T042A006-V003T042A006.

Zhang, W.-B., J. A. Misener, C.-Y. Chan, K. Zhou, and J.-Q. Li. 2014. Feasibility Assessment of A Truck Automation Deployment Framework. In: Transportation Research Board 93rd Annual Meeting