### Location-Based Analysis of Car-Following Behavior During Braking Using Naturalistic Driving Data

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Civil Engineering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjce-2019-0314.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>03-Jul-2019</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Tawfeek, Mostafa; University of Alberta, civil and environmental engineering; El-Basyouny, Karim; University of Alberta, Civil Engineering</td>
</tr>
<tr>
<td>Keyword:</td>
<td>transp. &amp; urban planning &lt; MANUSCRIPT CLASSIFICATION, Transportation, transportation &lt; Computer Applications</td>
</tr>
<tr>
<td>Is the invited manuscript for consideration in a Special Issue?:</td>
<td>Not applicable (regular submission)</td>
</tr>
</tbody>
</table>

https://mc06.manuscriptcentral.com/cjce-pubs
Location-Based Analysis of Car-Following Behavior During Braking Using Naturalistic Driving Data

Mostafa H. Tawfeek, Ph.D.*a

Karim El-Basyouny, Ph.D., P.Eng. b1

a Department of Civil Engineering, Ain Shams University, Cairo, Egypt

b Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada

*Corresponding author: mostafa.m.h@eng.asu.edu.eg

b1 basyouny@ualberta.ca
ABSTRACT

This study investigates the car-following behavior during braking at intersections and segments. Car-following events were extracted from a naturalistic driving dataset, mapped using ArcGIS, and analyzed to differentiate between the intersection- and segment-related events. The intersection-related events were identified according to an intersection influence area, which was estimated based on the stopping sight distance and the speed limit. Five behavioral measures were quantified based on exploring the Probability Density Functions (PDF) for intersection- and segment-related events. The results showed that there were significant differences between the PDFs of the measures for both cases. Moreover, it was indicated that drivers tend to be more aggressive at intersections when compared to segments. Thus, it is crucial to consider the driver’s location when investigating driver behavior. The quantified behavioral measures are a rich data source that can be used for car-following microscopic modeling, surrogate safety analysis, and driver assistance systems development.

Keywords: Driver Behavior, Naturalistic Driving Data, Car-following during Braking, Driving Aggressiveness, Intersection Influence Area, Probability Density Functions
INTRODUCTION

Car-following behavior is an extensive area of research due to the high frequency of this particular behavior. Previous studies have presented insights into this behavior that were beneficial for various purposes, such as car-following models’ development and calibration (Treiber et al. 2000; Rakha and Crowther 2003; Rakha 2009; Tang et al. 2012; Hammit et al. 2018), understanding driver behavior variation (Wang et al. 2010; Jensen et al. 2011; Martinez et al. 2017), rear-end near-crashes and crashes investigation (Arbabzadeh and Jafari, 2018; Carney et al., 2016; Gettman et al., 2008; Wu and Thor, 2015; Yang Zheng et al., 2014), and Driver Assistance System (DAS) development (Doi et al. 1994; Fujita et al. 1995; Hirst and Graham 1997; Barber and Clarke 1998; Kiefer et al. 1999; Bella and Russo 2011; Tawfeek and El-Basyouny 2018).

Various behavioral measures have been used to model and predict car-following behavior for the above-mentioned purposes. For instance, Wang et al. (2018) predicted driver braking behavior in car-following situations using vehicle speed, time-to-collision (TTC), relative speed, and range (i.e., the distance between the host vehicle and the leading vehicle) as input behavioral measures. Moreover, the car-following behavior was modeled by predicting the following vehicle acceleration and using relative speed, range, vehicle speed, and driver reaction delay as inputs for the model (Khodayari et al. 2012). Also, Bella and Russo (2011) estimated the range as a threshold for rear-end collisions warning using vehicle speed and relative speed. Montgomery et al. (2014), Vogel (2003), and Bella et al. (2014) investigated normal and risky car-following behavior using TTC. Moreover, aggressive and risky behavior were identified using jerk (Bagdadi and Várhelyi 2011; 2013; Feng et al. 2017).
Despite the extensive research conducted in the car-following behavior area (Saifuzzaman and Zheng 2014), a limited number of studies focused on car-following while braking; however, braking is the most frequent response to a car-following critical safety situation (Adams 1994, Lee et al. 2007). Moreover, there is a lack of previous research on quantifying and comparing car-following behavioral measures while braking at different locations of the road network (e.g., intersections versus segments). Therefore, this study aims at investigating the variation in the car-following behavior during braking at different road network locations (i.e., at intersections and on midblock segments). To reach this main objective, the following sub-objectives should be achieved:

- Differentiate between the car-following behavior at intersections and on segments by defining an intersection influence area where the driver behavior changes.
- Quantify and compare various behavioral measures (e.g., minimum following distance, acceleration, TTC, and jerk) in both cases (i.e., at intersection areas and on segments).

Car-following events investigated in this study are extracted from a naturalistic driving dataset. The naturalistic driving data outperforms other driver behavior data collection methods (i.e., controlled field tests or driving simulators) by capturing drivers’ normal behavior since the data is collected over a relatively long time period. Moreover, the data is recorded in the participants’ own vehicles during the daily driving routine. Therefore, the naturalistic driving data eliminates any chance that the driver could adjust his/her behavior because of the feeling of being observed or being in a testing environment (Dingus et al. 2006; 2015; Klauer et al. 2010).

The extracted car-following events are then mapped onto the road network using ArcGIS to differentiate between intersection and segment (i.e., non-intersection) events. The probability distributions of the behavioral measures are generated for each location for three speed groups.
(i.e., low, medium, and high). To further explore the differences in driver behavior at different road network locations, the probability distributions are compared using statistical testing.

3 LITERATURE REVIEW

One of the crucial car-following behavior aspects is the braking maneuver, since braking was indicated as the first response by the following vehicle driver to leading vehicle’s deceleration, hard brake, or stopping (Adams 1994; Lee et al. 2007). The braking maneuver can be split into three stages, namely, pre-braking, during braking, and post-braking; however, more attention will be directed to the during braking stage since it is the main focus of this study. The pre-braking stage includes the driver’s intention to brake and the start of the braking maneuver (i.e., the perception, decision, and the start of the action) (Bella and Russo 2011; Chatterjee and Davis 2014; Wang et al. 2018). On the other hand, the post-braking stage occurs when a braking event ends with a situation where a near-crash or a crash occurred after an evasive maneuver, such as hard braking, or sudden steering, or moving the foot from the brake pedal, or accelerating in other safe situations (Bagdadi 2013; Bagdadi and Várhelyi 2013; Wang et al. 2014; Perez et al. 2017).

The during braking stage starts once the driver starts to brake. In this stage, the driver might prepare to stop or slow down for different reasons, such as approaching a controlled intersection or approaching an end of a queue. The during braking stage ends when the driver reaches a stopping condition or moves his or her foot from the brake pedal. Several behavioral measures were investigated to study driver behavior during braking. For example, using the Naturalistic Driving Study (NDS) data to investigate rear-end events, Montgomery et al. (2014) studied TTC of following behavior while braking. The analysis included comparisons between differences in
TTC for various ages and gender. It was found that there was a statistically significant difference between various demographics of drivers while braking. In addition, Chen et al. (2016) compared the distributions of TTC and enhanced TTC (ETTC) at braking events in normal car-following situations. ETTC is similar to TTC but considers the relative acceleration between the following and the leading vehicles. The analysis of the data showed that the ETTC has less variation between drivers with different speed bins.

The TTC and headway were compared as safety indicators in different locations around an intersection with a stop sign on the minor road (Vogel 2003). TTC and headways during braking were collected using 7-point stations on main and minor roads. Two of these stations were setup near the intersection (17.5 meters away from the center of the intersection) on the main road with no stations near the intersection on the minor road. The results showed that there was no difference between the percentages of small headway values (less than 1 second) close and far away from the junction. On the other hand, the comparison of TTC values showed that one of the two stations near the intersection had the largest percentage of small TTC values (Vogel 2003). It is worth mentioning that the results of this study were compared visually, and no statistical testing was conducted.

Bella et al. (2014) compared time headway and TTC on highways during different times of the day (i.e., day and night), following different vehicle types (i.e., passenger cars and heavy vehicles), and travelling lane type (i.e. left or right lane) to investigate the difference between the two measures. The results showed that time headway and TTC are independent, which indicates that both measures provide different insights about drivers’ car-following behavior. Based on the results of this study, TTC is more effective than time headway to identify the impending risk of a rear-end collision (Bella et al. 2014).
To evaluate safety, comfort, and trust of an FCW system, the visual input of the driver and his behavior when following a leading vehicle during braking were analyzed (Wada et al. 2007). Driver behavior was expressed in terms of the distance to the leading vehicle and “performance index for approach and alienation”, which was proposed as the driver’s visual input. This index is based on the area change of the leading vehicle on the retina, which is a function of gap distance and relative velocity. The relation between the proposed index and the gap distance were validated by a test driver a real driving trip and 4 drivers in a driving simulator. The results imply that the ratio between the proposed index and the gap could be used to assess the unsafe driving situation and develop driver assistant systems (Wada et al. 2007).

In summary, previous studies addressed investigating the variation in the driver behavior based on the age and gender, comparing between different behavioral measures such as TTC, headway, and ETTC (Vogel 2003; Bella et al. 2014; Montgomery et al. 2014; Chen et al. 2016), or evaluating FCW based on drivers’ comfort and trust (Wada et al. 2007; Montgomery et al. 2014). Although a considerable number of studies discussed and modeled car-following behavior, car-following while braking, specifically, did not get the same attention. Moreover, these situations (i.e., car-following during braking) were not quantified using behavioral measures other than the headway and TTC. Furthermore, most of these studies highlighted some insights about these situations (e.g., the ETTC had better results than the TTC) to describe these situations and recommended the consideration of such insights in the development of Driver Assistance Systems (DAS). Also, the studies found in the literature did not address the effects of the surrounding infrastructure (e.g., intersections) on driver behavior.

Therefore, this paper investigates whether the driver behavior will change at a specific location on the network. In other words, the hypothesis is that there is a difference in driver following
behavior when driving on midblock segments and around intersections. Moreover, this paper will characterize driver behavior during braking on these locations (i.e., intersections and segments) by providing the statistical distribution for several behavioral measures including following distance, acceleration, relative speed, TTC, and jerk. The results of this paper present a rich data source for different transportation applications such as enhancing context-aware DAS, microscopic modeling of driver behavior including car-following models calibration, traffic safety (i.e., surrogate safety analysis), and reliability analysis for highway design.

**METHODOLOGY**

To characterize driver behavior during braking, the naturalistic driving data was used to identify car following events as well as extract various driver behavioral measures (e.g. following distance, acceleration, relative speed, TTC, and jerk). The candidate car following events during braking were split into three groups based on the host vehicle speed and plotted on a map using ArcGIS.

**Data Description**

NDS data, which is considered as big data, is a rich source of detailed real-world driving behavioral and performance data. Big data has a wide range of definitions in the literature. For instance, big data is defined as the type of data which is hard to store, manage, process, and visualize (Cox and Ellsworth 1997, Manyika et al. 2011). In addition, big data could be defined as the methods and technologies which enable unveiling hidden values from large, complex, and diverse datasets (Hashem et al. 2015). Moreover, big data could be described by high volume (i.e. large storage), variety (i.e. data format), high velocity (i.e. change in data over time), veracity (i.e. data credibility and certainty), and value (i.e. data usefulness) (Elragal 2014). Big
data has a significant role in the automotive and transportation industry transformation. For instance, big data is used to develop intelligent vehicles applications such as Self-Driving cars, Autonomous Vehicles, and Connected Vehicles (Luckow et al. 2015). In addition, big data, in the form of NDS, was used for understanding driver behavior to develop robust intelligent vehicles applications (Wang et al. 2017). NDS was carried out to investigate driver behavior and performance by collecting detailed data on volunteer drivers, vehicles, and the surrounding environment during normal everyday driving.

The NDS, which is used in this paper, was collected as part of the Safety Pilot Model Deployment (SPMD). SPMD was a data collection effort under real-life conditions with about 3,000 vehicles in Ann Arbor, Michigan, US (Henclewood and Rajiwade 2015). SPMD was part of the “Connected Vehicle (CV) Safety Pilot Program” research initiative which was aimed at evaluating the safety benefits of CV technologies. The dataset was collected from 63 host vehicles equipped with Integrated Safety Devices (ISD) for two months (October 2012 and April 2013) (USDOT 2018). The data was stored in two .csv files with a total file size of around 40.7 GB. For more details about the data, readers are directed to (Tawfeek and El-Basyouny 2018).

Identifying Drivers’ Following Events During Braking

Despite the usefulness of the NDS data set, it is also challenging since a lot of data manipulation is required. This is a common issue with big data especially without the availability of the videos which were recorded during the NDS. Consequently, a car following identification algorithm could be used to select targeted events in the data. One of the car following identification algorithm in naturalistic data was developed by Kusano et al. (2014) which aimed at identifying host vehicle braking events while following another vehicle. This algorithm was validated and successfully identified approximately 92% of the car following events when compared to manual
inspection of video data in a 100-car NDS. This algorithm was used to investigate various aspects of driving following behavior during braking. For instance, this algorithm was used to extract car-following events from NDS to develop a FCW (Tawfeek and El-Basyouny 2018). Also, it was used to identify the situations during normal car-following situations with brake application to compare between the TTC and the Enhanced TTC as thresholds for triggering FCW (Chen et al. 2016). Moreover, the same algorithm was utilized to compare the TTC for different speed bins or different drivers’ demographic groups (i.e. age and gender) in car-following scenarios while braking (Montgomery et al. 2014; Kusano et al. 2015). Chen et al. (2015) also used this algorithm to characterize driver behavior during lane change events by studying the minimum TTC at various travel speed.

Consequently, this algorithm will be adapted to serve the purpose of this study. The following steps, which were coded in MATLAB, summarize the NDS data processing and events identification:

1. Preparing and filtering the data:
   a. Remove all object types except the object type which was recognized as a light vehicle (i.e., passenger car).
   b. Remove all instances which were marked with invalid GPS data.

2. Identifying car following events during braking using Kusano et al. algorithm (Kusano et al. 2014). For more details about the algorithm, readers are directed to (Kusano et al. 2014; Tawfeek and El-Basyouny 2018). This algorithm could be summarized in the following steps:
   a. Identify all braking events for the host vehicle only when the vehicle’s speed is above 10 mph (4.47 m/s) for at least 1 second of the braking time.
b. Identify candidate leading vehicles within 3 seconds headway in front of the host vehicle during the braking events. These candidate leading vehicles should be within 2 meters laterally from the path of the host vehicle.

c. Remove fixed objects (i.e., parked vehicle and slow vehicles in adjacent lanes). This step depends on the speed of the leading vehicle, the gap distance between the leading and the following vehicle, and the azimuth change rate. The azimuth here is defined as the angle between the host vehicle path and the candidate leading vehicle.

d. Select leading vehicle, when multiple candidate vehicles were selected from the previous step, based on the lateral distance between the host and the leading vehicle and proximity to the host vehicle.

3. Selecting incidents of interest:

a. Scan each identified following event for the minimum following distance.

b. Extract and calculate driver behavior measures occurring at the minimum following distance while braking, namely, relative speed, acceleration, jerk, and TTC.

where TTC is calculated as the ratio between the range (i.e., following distance) and the relative speed. The TTC is defined as the time left for two vehicles to collide if no evasive maneuver is performed by at least one of them (Hayward, 1972; Hydén, 1987). The TTC could be estimated based on the following equation.

\[
TTC(\text{seconds}) = \frac{r}{v_F - v_L} \quad (1)
\]

where \( r \) is the following distance in meters, \( (v_F - v_L) \) is the relative speed, \( v_F \) is the speed of the following vehicle, and \( v_L \) is the speed of the leading vehicle as shown in Figure 1.
The longitudinal jerk, which will be referred to as jerk, is the change of the acceleration over time and could be represented in the following equation:

\[
\text{Jerk (m/s}^2/\text{s}) = \frac{\Delta a_F}{\Delta t} \quad (2)
\]

where \( \Delta a_F \) is the difference between the following vehicle acceleration in two consecutive time steps and \( \Delta t \) is the time step length.

A total of 44,383 car following events during braking were identified as candidate events. The incidents, when the following distance is the minimum during the event, were prepared to be plotted on a map to differentiate between intersections and segments events as discussed in the next section.

**Mapping and Defining Intersections and Segments Following Events**

ArcGIS was used as a tool to map the identified car following events during braking and to classify each event into either an intersection or segment related event. The candidate events were imported to ArcGIS using the available position data of the host vehicle (i.e. latitude and longitude). Before starting the classification process, intersection and segment (non-intersection) events should be defined. The current literature has various definitions of the intersections influence area. For instance, both the American Association of State Highway and Transportation Officials (AASHTO) and Highway Safety Manual (HSM) defines at-grade intersections by physical and functional areas (AASHTO 2010, 2011). The physical area is the area bounded by the stop lines of both intersecting roads. On the other hand, the functional area extends upstream the intersections on the intersecting roads. This area includes driver’s decision and maneuver distances and queue storage distance. More specifically, intersection-related crashes are defined as the crashes which occur within 15 to 152 meters from the intersection.
center point (R. Stollof 2008, FHWA 2009). On the other hand, the 152 meters intersection influence area was considered to be suitable for intersections approaches according to the 40 mph speed limit only (Cottrell and Mu 2005). It is worth to mention that intersection crashes and intersection-related crashes have been defined as the crashes which occurred upstream or inside the physical area of the intersection (Wang et al. 2009). Therefore, similar to the definition of intersection crashes, the candidate events were identified as intersection events when they occurred either upstream or inside the physical area of the intersection.

Since speed is the main contributor to a driver’s decision (i.e. perception-reaction) distance and maneuver (i.e. braking) distance, the speed limit should be considered when deciding the influence area (Fitzpatrick et al. 2000). Consequently, the candidate events were split into three categories based on the host vehicle speed, namely; low, medium, and high speeds. The boundaries of low, medium, and high speeds were defined as less than 25 mph (40 km/h), between 25 mph and 55 mph (88 km/h), and more than 55 mph, respectively. These boundaries are chosen based on the speed limits in the City of Ann Arbor, Michigan (City of Ann Arbor 2013). Based on these categories, intersections influence area for each category was set and defined as shown in Figure 2.

The intersection influence distance is calculated based on the driver’s Stopping Sight Distance (SSD) which include driver’s decision and maneuver distances as shown in the following equation.

\[
SSD (\text{meter}) = 0.278 \times v \times PRT + \frac{v^2}{254\left(\frac{a}{g} \pm G\right)}
\]  

(3)

where \(v\) is vehicle speed in km/h, \(PRT\) is driver’s Perception-Reaction Time in seconds, \(a\) is vehicle deceleration in m²/s, and \(G\) is roadway longitudinal slope.
The SSD was estimated for low, medium, and high-speed categories as 30, 80, and 150 meters respectively based on AASHTO recommendations for PRT of 2.5 seconds and deceleration of 3.41 m²/s (AASHTO 2011). Therefore, the intersection influence distance shown in Figure 2 will be 30, 80, and 150 meters for low, medium, and high-speed categories. Based on this distance, the intersection influence area was determined and any candidate event within this area was classified as an intersection event or, otherwise, considered as a segment event as shown in Figure 2. All 44,383 candidate events were classified based on the position of the host vehicle as mentioned before.

RESULTS AND DISCUSSION

The identified events were mapped using ArcGIS to differentiate between intersection and segment events in three-speed groups. As shown from Table 1, the high-speed group does not include any intersection events because the lower limit for this speed group is 88 km/h (55 mph) and it is difficult to have a braking event at an intersection with such a high speed. Similarly, the medium-speed group has fewer intersection events as it has a lower limit of 40 km/h and an upper limit of 88 km/h. In other words, the higher the speed is, the lower the instances of observed car-following events during braking will be recorded. For the low-speed group, although the difference in the percentage between the intersection events (43.3%) and segment events (56.7%) is not significant, the number of intersection events is less because of the characteristics of the car-following behavior during braking which is highlighted in the extraction procedure. The event extraction identifies the car-following event during braking based on several conditions, as explained previously. One of these conditions is that the speed of the host vehicle should be above 10 mph (4.47 m/s) for at least 1 second of the braking time.
This condition excludes the stop and go behavior which occurs more frequently at intersections than segments and contributes to the reduction of the intersection events.

The results were exported from ArcGIS to be processed in MATLAB. To characterize the driver behavior for each event type, the probability density functions (PDF) for minimum following distance, acceleration, TTC, relative speed, and jerk were developed in each speed group for both intersection and segment events. The actual values of the measures were fitted to 17 continuous distribution types, namely; Beta, Birnbaum-Saunders, Exponential, Extreme value, Gamma, Generalized Extreme Value, Generalized Pareto, Inverse Gaussian, Logistic, Log-logistic, Log-normal, Nakagami, Normal, Rayleigh, Rician, \( t \) Location-Scale, and Weibull. Table 2 shows the distribution’s name and parameters which best fits the actual values of the above-mentioned behavioral measures for each event type and speed group. The distribution was considered to best fit the actual values of the behavioral measures since the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) had the lowest values. As noted from Table 2, different distribution types were fitted to the actual values of the behavioral measures to ensure that the reported distributions replicate the actual driver behavior, which is defined in this study by five behavioral measures (i.e., acceleration, following distance, TTC, relative speed, and jerk).

Since the distributions of all the measures were not normally distributed, a nonparametric statistical test (i.e. Kolmogorov-Smirnov (KS) test) was used to check the significance of the difference between the intersection/segment event distributions in each speed group. The null hypothesis is that the values for each measure of the intersection and segment events are from the same continues distribution. As shown in Table 2, the results of the KS test showed that all the measures of intersection events were different from segment events at 95% confidence level.
except the jerk in the medium speed group. When comparing the jerk parameter for intersection and segment events in the medium speed group, the $p$-value was estimated at 0.0514, which is marginally insignificant at the 95% confidence level but significant at 90% confidence level. These results indicate that driver behavior at intersections is substantially different from driver behavior on segments. Such results are crucial, for example, when developing any driving assistance system because drivers will receive the warning differently in varied contexts. In other words, drivers could accept a warning relayed to them on midblock segments while considering the same warning at intersections as an early warning which will have an impact on the entire system acceptance. Also, these results can be integrated into the development or the calibration of car-following models to enhance the modeling of such behavior.

Table 2 also shows the mean values of the selected behavioral measures in the third column between brackets. For the jerk, the mean was calculated for the negative and positive jerk values separated. However, the distributions’ parameters of the jerk (in the fifth, sixth, and seventh columns) were estimated based on all jerk values (i.e. negative and positive values). As shown in Table 2, measures’ mean values for intersection events were less than segment events. For instance, the mean value of the acceleration of low-speed intersection events (i.e. -1.68 m/s$^2$) was less than the acceleration of low-speed segment events (i.e. -1.43 m/s$^2$). This indicated that the drivers tend to have smaller deceleration rates and larger following distance during braking when driving on segments than when driving on intersections. This observation was confirmed by the values of TTC, where TTC of car-following during braking events at intersections were less than segment events. However, the differences between the jerk values were not big but they were statistically significant as discussed earlier. On the other hand, the relative speed at intersections
was significantly higher than on segments. Moreover, the mean values of the behavioral measures in the medium-speed group were larger than the values of the low-speed group.

In summary, by investigating the mean values of the behavioral measures, it was found that drivers approaching intersections had higher deceleration rates, smaller minimum following distance, smaller TTC, higher relative speeds and larger jerk than when driving on segments. This insight implies that drivers are more likely to exhibit aggressive behavior when approaching intersections. This aggressive behavior is a result of drivers desire to cross intersections without having to stop. For example, when a driver is approaching a signalized intersection that is close to the end of its green cycle, he or she will intentionally accelerate to catch the green time and avoid stopping at the intersection even if this will cause them to closely follow the leading vehicle. Moreover, this insight confirms the hypothesis stated earlier in this paper that the driver behavior at intersections is different from segments.

SUMMARY AND CONCLUSIONS

This paper quantified five driver behavioral measures (i.e., acceleration, minimum following distance, relative speed, TTC, and jerk) for car-following during braking situations. Using NDS, this quantification is carried out for events occurred at intersections and midblock segments. In addition, the selected behavioral measures for intersection and segment events were compared. To achieve the objectives of this paper, the car-following events during braking were extracted from the NDS. These events were mapped onto the road network using ArcGIS to identify intersection and segment events. For the identification of intersection events, a clear definition of the intersection influence area based on measurable design parameters (i.e., SSD and the speed limit) was presented. This study defined the intersection influence area as the area where the driver
behavior is affected substantially by an intersection ahead which will impact the driver’s maneuvering decisions including car-following, braking, and lane changing decisions. This influence area includes the physical area of the intersection which is bounded by the stop lines and the area upstream the stop line of the intersection limited by the SSD which is estimated based on the speed of the driver and the speed limit. The intersection influence area was defined for three speed groups (i.e., low, medium, and high) based on driver’s SSD and speed. Then, the probability distributions for the behavioral measures for each speed group for intersections and segments were developed and fitted to continuous probability distributions. In addition, intersection/segment driver following behavior during braking was compared by checking the significance of the difference between the behavioral measures distributions in each speed group.

The results of this study revealed that there was a considerable difference between driver behavior during braking at intersections and on midblock segments. This difference was defined as the differences between the behavioral measures distributions of intersection and segment events. Using the KS test, these differences were statistically significant at 95% confidence level. In addition, this study documented the distributions’ names and parameters which best fitted the behavioral measures. By investigating the mean values of the behavioral measures, it was found that drivers approaching intersections had higher deceleration rates, smaller minimum following distance, smaller TTC, higher relative speed, and larger jerk than when driving on segments. This finding implies that drivers are more likely to exhibit aggressive behavior when approaching intersections. This finding highlights the importance of considering the driver location when dealing with driver behavior. For example, according to the current literature, drivers classification procedures are usually based only on some selected behavioral measures (e.g., acceleration, TTC, jerk, etc.) regardless the location (Wang et al. 2010; Rodriguez

Considering the findings of this study, the proposed procedures in the literature could misclassify a driver as an aggressive driver when most of the data was collected in high intersection density areas (e.g., urban areas or downtown areas). Therefore, as a main contribution of this study, the driving data should be split based on the location before classifying drivers’ behavior. In addition, the driving assistance systems should consider the surrounding environment (i.e. proximity to intersections) by modifying these systems to context-aware driving assistance systems. Such context-aware systems are expected to be more robust and reliable since it will accommodate the changing drivers’ surrounding environment.

One of the main goals of understanding driver behavior is to develop innovative traffic safety solutions. Hence, the results of this study could be beneficial for the development of collision avoidance/warning systems such as FCW. Forasmuch as driver behavior at intersections is different from the behavior on segments, it is recommended to develop a FCW algorithm which accounts for the driver location in the network. To develop such FCW algorithm(s), further analysis of the results should be conducted to select events of interest based on the distribution of acceleration, jerk, and TTC.

In addition, the results could enhance the calibration process of microsimulation models by updating the distributions of drivers’ behavioral measures in simulation packages as suggested by (Tawfeek et al. 2018). In addition, the PDF distributions of the behavioral measures could be used in modeling car following behavior mathematically as suggested by Gonzalez et al. (2014). The PDF distributions generated from this study is expected to lead to more robust simulation results because of the consideration of the driver location and the fact that these distributions
were generated based on NDS. The NDS grants the recording of normal driving behavior without exposing the drivers to a test setting as in test track which will affect their behavior.

Moreover, this study shed light on the importance of integrating the NDS with the road network map to deeply investigate the driver behavior. The pairing of the NDS and the road network, using ArcGIS, LiDAR data, or other means, is a very promising area of research which will allow insightful findings in driver behavior understanding. Therefore, as an extension of this study, an automated approach to differentiate between intersection and midblock segment events could be explored. This automated approach could enhance the definition of the intersection influence area which has a wide range of definitions in the literature (e.g. the suggested definition in this study which was based on the SSD and vehicle speed).

REFERENCES


Wang, J., Kodaka, K., and Li, K. 2014. Driving risk assessment using cluster analysis based on naturalistic driving data. 17th International IEEE Conference on Intelligent Transportation...


1 LIST OF FIGURES

2 Figure 1. The main parameters of a car-following situation.

3 Figure 2 Sketch of intersection influence area on the left and an example of the candidate
intersection events (red dots) and segment events (blue dots) on the right.
### Table 1 Events breakdown by event type and speed group

<table>
<thead>
<tr>
<th>Speed Group</th>
<th>Intersection Events</th>
<th>Segment Events</th>
<th>Total Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>11403</td>
<td>14962</td>
<td>26365</td>
</tr>
<tr>
<td>Medium</td>
<td>1995</td>
<td>12614</td>
<td>14609</td>
</tr>
<tr>
<td>High</td>
<td>0</td>
<td>3409</td>
<td>3409</td>
</tr>
<tr>
<td></td>
<td>13398</td>
<td>30985</td>
<td>44383</td>
</tr>
</tbody>
</table>
Table 2 The parameters of the fitted distributions to the actual PDFs

<table>
<thead>
<tr>
<th>Speed Group</th>
<th>Location (mean value)</th>
<th>Distribution Name</th>
<th>Parameters</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Location</td>
<td>Scale</td>
</tr>
<tr>
<td><strong>Acceleration (m/s²)</strong></td>
<td></td>
<td></td>
<td>Location-Scale</td>
<td>0.745</td>
</tr>
<tr>
<td>Low</td>
<td>Int.(-1.68)</td>
<td>t Location-Scale</td>
<td>-1.667</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>Seg.(-1.43)</td>
<td>Extreme Value</td>
<td>-1.060</td>
<td>0.663</td>
</tr>
<tr>
<td>Med.</td>
<td>Int.(-1.19)</td>
<td>Extreme Value</td>
<td>-0.847</td>
<td>0.608</td>
</tr>
<tr>
<td>Med.</td>
<td>Seg.(-1.00)</td>
<td>Extreme Value</td>
<td>-0.715</td>
<td>0.507</td>
</tr>
<tr>
<td>High</td>
<td>Seg.(-0.85)</td>
<td>Extreme Value</td>
<td>-0.624</td>
<td>0.405</td>
</tr>
<tr>
<td><strong>Following Distance (m)</strong></td>
<td></td>
<td></td>
<td>Generalized Extreme Value</td>
<td>7.127</td>
</tr>
<tr>
<td>Low</td>
<td>Int.(9.01)</td>
<td>Generalized Extreme Value</td>
<td>7.127</td>
<td>3.153</td>
</tr>
<tr>
<td>Low</td>
<td>Seg.(11.05)</td>
<td>Generalized Extreme Value</td>
<td>8.496</td>
<td>4.071</td>
</tr>
<tr>
<td>Med.</td>
<td>Int.(19.5)</td>
<td>Gamma</td>
<td>-</td>
<td>4.430</td>
</tr>
<tr>
<td>High</td>
<td>Seg.(29.07)</td>
<td>Birnbaum-Saunders</td>
<td>-</td>
<td>25.100</td>
</tr>
<tr>
<td><strong>TTC (sec)</strong></td>
<td></td>
<td></td>
<td>Birnbaum-Saunders</td>
<td>8.496</td>
</tr>
<tr>
<td>Low</td>
<td>Int.(10.48)</td>
<td>Generalized Extreme Value</td>
<td>3.556</td>
<td>3.081</td>
</tr>
<tr>
<td>Low</td>
<td>Seg.(17.23)</td>
<td>Birnbaum-Saunders</td>
<td>-</td>
<td>9.383</td>
</tr>
<tr>
<td>Med.</td>
<td>Int.(21.42)</td>
<td>Log-normal</td>
<td>2.549</td>
<td>1.087</td>
</tr>
<tr>
<td>Med.</td>
<td>Seg.(30.00)</td>
<td>Gamma</td>
<td>-</td>
<td>20.834</td>
</tr>
<tr>
<td>High</td>
<td>Seg.(31.81)</td>
<td>Birnbaum-Saunders</td>
<td>-</td>
<td>24.162</td>
</tr>
<tr>
<td><strong>Relative Speed (m/s)</strong></td>
<td></td>
<td></td>
<td>Generalized Extreme Value</td>
<td>-1.368</td>
</tr>
<tr>
<td>Low</td>
<td>Int.(-2.11)</td>
<td>Extreme Value</td>
<td>-1.368</td>
<td>1.212</td>
</tr>
<tr>
<td>Low</td>
<td>Seg.(-1.54)</td>
<td>Extreme Value</td>
<td>-0.919</td>
<td>0.964</td>
</tr>
<tr>
<td>Med.</td>
<td>Int.(-2.13)</td>
<td>Generalized Extreme Value</td>
<td>-2.208</td>
<td>2.340</td>
</tr>
<tr>
<td>Med.</td>
<td>Seg.(-1.28)</td>
<td>t Location-Scale</td>
<td>-0.487</td>
<td>0.435</td>
</tr>
<tr>
<td>High</td>
<td>Seg.(-1.23)</td>
<td>Generalized Extreme Value</td>
<td>-1.312</td>
<td>1.315</td>
</tr>
<tr>
<td><strong>Jerk (m/s²/s)</strong></td>
<td></td>
<td></td>
<td>t Location-Scale</td>
<td>0.059</td>
</tr>
<tr>
<td>Low</td>
<td>Int.(-1.47/1.88)*</td>
<td>t Location-Scale</td>
<td>0.059</td>
<td>1.082</td>
</tr>
<tr>
<td>Low</td>
<td>Seg.(-1.43/1.98)*</td>
<td>t Location-Scale</td>
<td>0.136</td>
<td>1.049</td>
</tr>
<tr>
<td>Med.</td>
<td>Int.(-1.84/2.22)*</td>
<td>t Location-Scale</td>
<td>0.147</td>
<td>1.453</td>
</tr>
<tr>
<td>Med.</td>
<td>Seg.(-1.79/2.14)*</td>
<td>t Location-Scale</td>
<td>0.169</td>
<td>1.299</td>
</tr>
<tr>
<td>High</td>
<td>Seg.(-1.78/2.19)*</td>
<td>t Location-Scale</td>
<td>0.266</td>
<td>1.529</td>
</tr>
</tbody>
</table>

*The average is calculated for (negative/positive) jerk separated

https://mc06.manuscriptcentral.com/cjce-pubs
Figure 1. The main parameters of a car-following situation.
Figure 2 Sketch of intersection influence area on the left and an example of the candidate intersection events (red dots) and segment events (blue dots) on the right.