Accident Analysis and Prevention 83 (2015) 90–100

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Modeling driver stop/run behavior at the onset of a yellow indication considering driver run tendency and roadway surface conditions

Mohammed Elhenawy[a,1], Arash Jahangiri[b,1], Hesham A. Rakha[c,*], Ihab El-Shawarby[b,d,1]

[a] Bradley Department of Electrical and Computer Engineering, Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Blacksburg, VA 24061, United States
[b] Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Blacksburg, VA 24061, United States
[c] Charles E. Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Blacksburg, VA 24061, United States
[d] Ain-Shams University, Cairo, Egypt

ARTICLE INFO

Article history:
Received 1 November 2014
Received in revised form 19 April 2015
Accepted 18 June 2015
Available online 28 July 2015

Keywords:
Dilemma zone
Yellow indication
Run/stop behavior
Classification

ABSTRACT

The ability to model driver stop/run behavior at signalized intersections considering the roadway surface condition is critical in the design of advanced driver assistance systems. Such systems can reduce intersection crashes and fatalities by predicting driver stop/run behavior. The research presented in this paper uses data collected from two controlled field experiments on the Smart Road at the Virginia Tech Transportation Institute (VTTI) to model driver stop/run behavior at the onset of a yellow indication for different roadway surface conditions. The paper offers two contributions. First, it introduces a new predictor related to driver aggressiveness and demonstrates that this measure enhances the modeling of driver stop/run behavior. Second, it applies well-known artificial intelligence techniques including: adaptive boosting (AdaBoost), random forest, and support vector machine (SVM) algorithms as well as traditional logistic regression techniques on the data in order to develop a model that can be used by traffic signal controllers to predict driver stop/run decisions in a connected vehicle environment. The research demonstrates that by adding the proposed driver aggressiveness predictor to the model, there is a statistically significant increase in the model accuracy. Moreover the false alarm rate is significantly reduced but this reduction is not statistically significant. The study demonstrates that, for the subject data, the SVM machine learning algorithm performs the best in terms of optimum classification accuracy and false positive rates. However, the SVM model produces the best performance in terms of the classification accuracy only.

1. Introduction

With advances in sensing, communications, and computational technologies, research in the area of vehicle safety is increasing. Most new vehicles have active safety features including anti-lock braking and adaptive cruise control systems to reduce road accidents (Jones, 2001). In the US, the Department of Transportation (DOT) reported 32,367 fatalities caused by road accidents in 2011 (Tibshirani et al., 2009). A significant percentage of these road accidents occurred at signalized intersections as a result of driver behavior in the decision/dilemma zone while approaching signalized intersections (U.S. Department of Transportation and Federal Highway Administration, 2014).

Drivers approaching a traffic signal yellow indication have to decide whether to stop or proceed through the intersection. Typically, drivers far from the intersection, when a yellow indication is initiated, tend to stop while others near the intersection tend to proceed. A dilemma zone is a spatial stretch of roadway upstream of the intersection stop line that exists when the minimum stopping distance d_s is larger than the maximum clearing distance d_r, as illustrated in Fig. 1. In this case, drivers encountering the onset of yellow interval while traveling between d_s and d_r have no valid option (i.e., they cannot stop comfortably nor can they run before the traffic signal indication turns red). The minimum stopping distance is the distance required by a vehicle to safely come to a complete stop upstream of the intersection stop bar at a reasonable deceleration level (assumed to be 3 m/s^2). The maximum clearing distance is the distance within which the

* Corresponding author. Fax: +1 540 231 1555.
E-mail addresses: mohame1@vt.edu (M. Elhenawy), arashj@vt.edu (A. Jahangiri), hrakha@vt.edu (H.A. Rakha), shawarby@vt.edu (I. El-Shawarby).
[1] Fax: +1 540 231 1555.

http://dx.doi.org/10.1016/j.aap.2015.06.016
0001-4575/

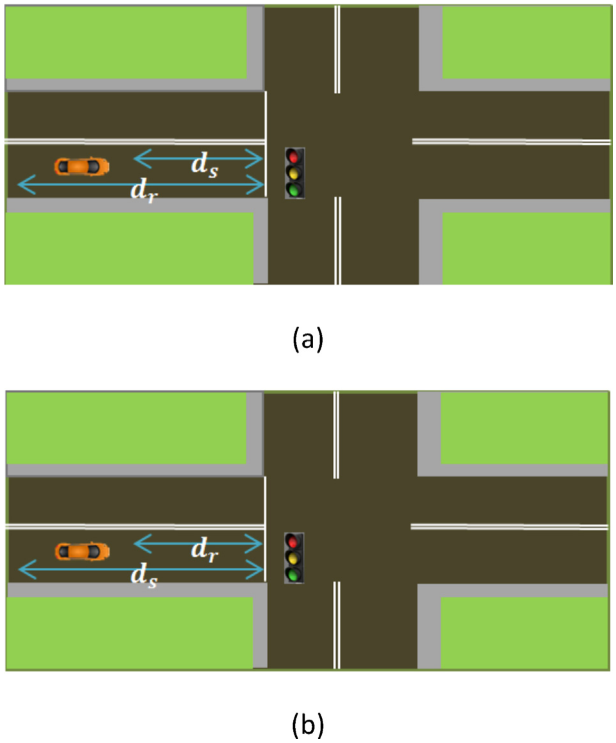

Fig. 1. Illustration of the option zone in panel (a) and DZ in panel (b).

vehicle can cross the intersection stop bar before the conclusion of the yellow interval. If the distance of an approaching vehicle to the intersection stop bar (DTI) at the onset of a yellow indication is between d_s and d_r (i.e., $d_r < \text{DTI} < d_s$) then the vehicle is in the DZ. Alternatively, if d_s is less than d_r, then vehicles at DTIs between d_s and d_r ($d_s < \text{DTI} < d_r$) are classified as being in the option zone and have two valid choices (stop safely or proceed safely). The DZ was first introduced by Gazis et al. (1960) and has been studied in many other studies (Rakha et al., 2007; Sheffi and Mahmassani, 1981; Bonneson et al., 2002; Gates et al., 2007; Pant et al., 2005; Chang et al., 1985; Zegeer, 1978; Liu et al., 2007; Wei et al., 2011; Ghanipoor Machiani and Abbas, 2014a; Abbas et al., 2014). The design of yellow timings is made in order to avoid the creation of the DZ.

Several factors influence driver behavior at the onset of yellow that result in the potential existence of a DZ. The factors can be divided into three categories; driver-related, intersection related and vehicle-related. These factors that have been studied throughout the literature (Rakha et al., 2007; Gates et al., 2007; Liu et al., 2007; Wei et al., 2011; Ghanipoor Machiani and Abbas, 2014a; Caird et al., 2007; El-Shawarby et al., 2007; Li et al., 2012; Jahangiri et al., 2015) include the driver perception–reaction time; the driver's acceptable deceleration level; the driver's age; the driver's gender; the time-to-intersection (TTI) at the onset of yellow; the distance-to-intersection (DTI) at the onset of yellow; the vehicle's approach speed; the vehicle type; presence of side-street vehicles, pedestrians, bicycles, or opposing vehicles waiting to turn left; the arrival rate; the length of the yellow interval; the cycle length; and presence of police vehicles in the vicinity of the intersection. Moreover, El-Shawarby et al. (2015) compared driver stopping/running probabilities in clear weather and in rainy weather and found a slight shift between the two probabilities. El-Shawarby et al. (2015) correlated this shift to the decrease in the probability of stopping in case of wet pavement surface and rainy weather conditions. Consequently, the roadway surface condition was added as an input variable to the proposed classifiers. The proposed driver aggressiveness factor has not been considered in past studies.

The intersection safety needs identification report published by the Federal Highway Administration (FHWA) in July 2009 showed that in 2007, 22% of the total fatal crashes were intersection-related with an estimated cost of 27.8 US billion dollars while 44.8% of the total injury crashes were also intersection-related with an estimated cost of 51.3 US billion dollars (Coakley and Stollof, 2009). Based on the National Highway Traffic Safety Administration (NHTSA), two-thirds of all fatal crashes are caused by aggressive driving (Wei, 2008). Consequently, aggressive driving is critical in modeling driver stop/run behavior at signalized intersections; however, measuring driver aggressiveness may not be plausible. In a previous research study, five driver actions were used to measure aggressive driving behavior. These five measures include: short or long honk of the horn, cutting in front of other vehicles in a passing lane maneuver, cutting in front of other vehicles in a multi-lane passing maneuver, and passing one or more vehicles by driving on the shoulder and then cutting in (Shinar and Compton, 2004). Other studies classified drivers into three categories aggressive, conservative, and normal drivers based on their decision (stop/run) and the distance to the stop line when the traffic signal turned yellow (Liu et al., 2012). In our previous work, we proposed the use of the frequency of running a yellow indication as a measure of driver aggressiveness (Elhenawy et al., 2014). In the current study a more formal definition and formulation of the driver aggressiveness parameter is proposed. The measure proposed here is a continuous measure of aggressiveness that varies from zero (not aggressive) to one (very aggressive) as opposed to a categorical variable as was done in previous studies.

Consider a vehicle approaching a signalized intersection, our goal is to build a model that predicts driver stop or run behavior at the onset of the yellow indication. This model uses many predictors such as the TTI and driver's age to predict the driver decision. Because in real-life, different drivers behave differently, we added the proposed predictor to explain some of the variation between drivers based on their history. Such a model should be one of the main building blocks in more advanced driver assistance systems. These systems should be able to predict the driver behavior and warn them if their decision is incorrect. Moreover, it would warn the driver if there is any potential violation from other drivers/vehicles approaching the intersection. The system should ensure the algorithm produces minimum false positives in order to encourage drivers trust their output.

The past two decades have seen numerous research efforts and advances in both machine learning techniques and computer computational power. Many machine learning techniques require a large number of computations and are infeasible without computers. The available machine learning algorithms and computational power have made such techniques feasible for real-time implementation. Transportation engineers are among people who are interested in applying these algorithms to address transportation problems. This interest increases with the availability of data sets from fixed detectors, data probes and intelligent transportation systems (ITSs). Recently, some machine learning algorithms were used in the transportation field, including: classifying and counting vehicles detected by multiple inductive loop detectors (Ali et al., 2012), identifying motorway rear-end crash risks using disaggregate data (Pham et al., 2010), automatic traffic incident detection (Liu et al., 2015), real-time detection of driver distraction (Yulan et al., 2007; Tango and Botta, 2013), transportation mode recognition using smartphone sensor data (Jahangiri and Rakha, 2014, 2015), and video-based highway asset segmentation and recognition (Balali and Golparvar-Fard, 2014). Modeling driver stop/run behavior at signalized intersections is very important and is ideal for applying machine learning