Accident Analysis and Prevention 41 (2009) 98–107

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Exploring precrash maneuvers using classification trees and random forests

Rami Harb[a,*], Xuedong Yan[b], Essam Radwan[c], Xiaogang Su[d]

[a] Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816-2450, United States
[b] Center for Transportation Research, University of Tennessee, 309 Conference Center Bldg, Knoxville, TN 37996-4133, United States
[c] Center for Advanced Transportation Systems Simulation, Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816-2450, United States
[d] Department of Statistics and Actuarial Science, University of Central Florida, Orlando, FL 32816-2450, United States

ARTICLE INFO

Article history:
Received 24 February 2008
Received in revised form 26 August 2008
Accepted 29 September 2008

Keywords:
Accident avoidance
Rear-end accidents
Head-on accidents
Angle accident
Decision trees
Random forests

ABSTRACT

Taking evasive actions vis-à-vis critical traffic situations impending to motor vehicle crashes endows drivers an opportunity to avoid the crash occurrence or at least diminish its severity. This study explores the drivers, vehicles, and environments' characteristics associated with crash avoidance maneuvers (i.e., evasive actions or no evasive actions). Rear-end collisions, head-on collisions, and angle collisions are analyzed separately using decision trees and the significance of the variables on the binary response variable (evasive actions or no evasive actions) is determined. Moreover, the random forests method is employed to rank the importance of the drivers/vehicles/environments characteristics on crash avoidance maneuvers. According to the exploratory analyses' results, drivers' visibility obstruction, drivers' physical impairment, drivers' distraction are associated with crash avoidance maneuvers in all three types of accidents. Moreover, speed limit is associated with rear-end collisions' avoidance maneuvers and vehicle type is correlated with head-on collisions and angle collisions' avoidance maneuvers. It is recommended that future research investigates further the explored trends (e.g., physically impaired drivers, visibility obstruction) using driving simulators which may help in legislative initiatives and in-vehicle technology recommendations.

Published by Elsevier Ltd.

1. Introduction

Many safety experts do not refer to traffic collisions as accidents and consider that they could and should have been avoided (What Causes Accidents, 2007). In fact, under critical traffic conditions leading to motor vehicle accidents, drivers can either take or not take evasive actions to avoid the crashes. Taking evasive actions vis-à-vis critical traffic situations grants drivers an opportunity to avoid the crash occurrence or at least diminish its severity. According to Uc et al., avoiding a crash requires continuous monitoring of neighboring vehicles, and anticipating and adjusting to changes in their speeds and positions under the pressure of time, which rely on multiple cognitive abilities (Uc et al., 2006). According to General Estimate System (GES) data for years 2002, 2003, and 2004 (NHTSA, 2004), over 30% of the drivers involved in rear-end accidents, over 25% of the drivers involved in head-on accidents, over 20% of drivers involved in angle collisions, and over 5% of the drivers involved in sideswipe collisions had no evasive actions prior to the accident occurrence (see Fig. 1). The reader may refer to Section 2.1 for detailed description of the GES data set used in this study. The noteworthy number of drivers not performing corrective actions prior to crashes underscores the need to explore the factors associated with the crash avoidance maneuvers.

The National Highway Traffic Safety Administration (NHTSA) shows great interest in accident avoiding technologies such as lane departure warning and brake assist (Safety Regulators Shifting Focus to Accident Avoidance, 2006). As a result, many automobile manufacturers studied and included collision avoidance warning systems (CAWS) in new car models that are designed to notify drivers about potential hazards from roadway departure and other vehicles (Araki et al., 1996; Chen et al., 1997; Hirst and Graham, 1997; NHTSA, 2002; Tilin, 2002; Clement and Taylor, 2006; Maltz and Shinar, 2007; Sengupta et al., 2007). Moreover, in an effort to enhance traffic accident avoidance skills, several agencies offer driver training courses for accident avoidance. These classes are designed to improve driving skills and help drivers reduce human errors when faced with hazardous traffic situations. All soldiers, civilian employees, and contractor employees who drive army-owned or leased vehicles must complete the training course (Wheeled Accident Avoidance, 2006).

Several researchers studied the reliability of different collision avoidance systems. For instance, Bliss and Acton (2003) studied the unreliability of collision avoidance systems and their effect on

* Corresponding author.
E-mail addresses: Ramyz5@yahoo.com (R. Harb), yxd22222@yahoo.com (X. Yan), Aeradwan@mail.ucf.edu (E. Radwan), Xsu@mail.ucf.edu (X. Su).

0001-4575/$ – see front matter. Published by Elsevier Ltd.
doi:10.1016/j.aap.2008.09.009

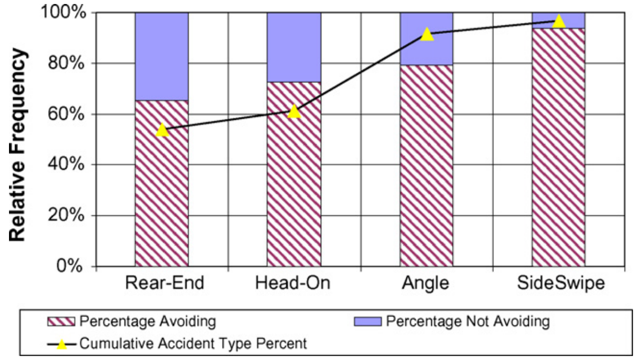

Fig. 1. Percentage of drivers trying to avoid accidents.

driving performance. It was concluded that reliable alarms generate more reaction from drivers (Bliss and Acton, 2003). Any mistrust of the alarm systems will result in reduced alarm response frequency (Bliss, 1993). Moreover, in the final report for the Automotive Collision Avoidance Systems (ACAS) Program, NHTSA states that drivers will likely ignore alarms that are not reliable (NHTSA, 2000). Therefore, it is crucial to explore the overall aspects and traits associated with the avoidance maneuvers to help develop efficient and reliable collision warning systems.

To study the accident avoidance maneuvers, many researchers employ driving simulators. Dangerous environments and hazardous traffic situations may be replicated in a safe driving simulator environment. For instance, Hancock and Ridder (2003) used a driving simulator to study the behavioral responses of drivers in the final seconds and milliseconds of induced hazardous situations that could lead to traffic accidents. Ikeda et al. (2001) studied the accident avoidance performance with respect to the age of the drivers using a driving simulator. Charlton (2007) used the University of Waikato-New Zealand driving simulator to examine the role of attentional, perceptual, and lane placement factors in drivers' behavior at horizontal curves. Ho et al. (2006) used the University of Oxford driving simulator to assess the effectiveness of vibrotactile warning signals in preventing front-to-rear-end collisions. Maltz and Shinar (2007) compared imperfect in-vehicle collision avoidance warning systems (IVCAWS) to a higher reliability collision avoidance warning system. Other studies using driving simulators entailed older drivers' avoidance capability in traffic accidents such as reduced vision field under complex work (Uno and Hiramatsu, 1995), effect of variable message signs on driver speed behavior on a section of expressway under adverse fog conditions (Kolisetty et al., 2006), deterioration in response for evasive maneuvers (Uno and Hiramatsu, 1998), deterioration of accident avoidance capability (Nishida, 1998), and rear-end accident avoidance resulting from horizontal visibility blockage (Harb et al., 2007a,b). Although the simulator studies provide a safe driving environment to test drivers' crash avoidance behaviors, these studies are generally limited by the sample size and the particular traffic scenario design. Therefore, an overall exploration analysis of the environments/drivers/vehicles factors associated with drivers' evasive maneuvers is needed, which could be a starting point for more in depth analyses using driving simulator or instrumented vehicles.

This paper studies the environments/drivers/vehicles factors associated with crash avoidance maneuvers based on the General Estimate System (GES) for years 2002, 2003, and 2004. The GES database records whether drivers took evasive actions or not prior to crashes in addition to the type of evasive actions (braking, steering, etc.). Rear-end collisions, head-on collisions, and angle collisions are analyzed individually to explore their environments, vehicles, and drivers' characteristics associated with the crash avoidance maneuver. It should be noted that sideswipe accidents were excluded from this analysis due to the small sample size. Previous studies successfully applied logistic regression to identify the statistical significance of independent variables on binary dependent variables (Hing et al., 2003; Stamatiadis and Deacon, 1995). Although logistic regression is a proper method to estimate the significance of independent variable on a dichotomous dependent variable, it makes it difficult to detect and interpret complex or high-order interactions among independent variables (Morgan and Sonquist, 1963; Su et al., 2008). Therefore, classification trees, nonparametric models, are utilized to analyze the accident avoidance maneuvers (evasive actions versus no evasive actions) for each accident type. It should be mentioned that tree methods handle interactions implicitly. In other words, one should not pick up a terminal tree node and trace up its ancestors in order to look for higher-order interactions. In addition, the random forests technique is employed to determine the independent variables' importance ranking for each accident type.

The remainder of this paper is organized as follows. Section 2 presents the crash database used in this study including the years of interest, the target variable, the data preparation, and the limitations of the data. In the same section, the classification trees and random forests techniques are briefly described. Section 3 illustrates the classification trees models and the variable importance rankings for rear-end collisions, head-on collisions, and angle collisions. Section 4 summarizes the findings of the analyses including discussions recommendations for future research.

2. Methodology

2.1. Crash database

The General Estimates System (GES) database for years 2002, 2003, and 2004 is used in this study. The GES database obtains its data from a nationally representative probability sample selected from the estimated 6.3 million police-reported crashes which occur annually. This database includes crashes that result in fatalities, injuries, and major property damage. The crash reports are chosen from areas that reflect the geography, roadway mileage, population, and traffic density of the United States. For more detailed information, refer to the GES Analytical User's Manual (NHTSA, 2004).

The GES database is a relational database consisting of three main files: accident, vehicle/driver, and person. Each file deals with a specific aspect of traffic crashes. These files may be linked as needed by the crash report case number and vehicle number. The accident file contains information on crash characteristics and environmental conditions at the time of the crash. The vehicle/driver file contains general information describing all vehicles and drivers involved in the crash. The person file contains general information describing all persons involved in the crash: drivers, passengers, pedestrians, pedal cyclists, and non-motorists.

Since 1992, five precrash variables have been added to the vehicle/driver file to identify: (1) what was the vehicle doing just prior to the critical pre-crash event (P_CRASH1), (2) what made the vehicle's situation critical (P_CRASH2), (3) what was the corrective action made, if any, to this critical situation (P_CRASH3), (4) what was the stability of the vehicle just prior to impact (P_CRASH4), (5) and what were the results of the vehicle's pre-crash stability coded in variable P_CRASH4 (P_CRASH5). To investigate the significant factors associated with drivers' crash avoidance actions; this study mainly focuses on the variables P_CRASH2 and P_CRASH3. P_CRASH2 was used to identify the vehicles/drivers that were under a critical traffic situation and might have chances