Accident Analysis and Prevention 90 (2016) 128–139

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Investigating driver injury severity patterns in rollover crashes using support vector machine models

Cong Chen[a], Guohui Zhang[a,*], Zhen Qian[b], Rafiqul A. Tarefder[a], Zong Tian[c]

[a] Department of Civil Engineering, University of New Mexico, Albuquerque, NM 87131, USA
[b] Department of Civil and Environmental Engineering and Heinz College, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[c] Department of Civil and Environmental Engineering, University of Nevada, Reno, NV 89557, USA

ARTICLE INFO

Article history:
Received 14 September 2015
Received in revised form 20 January 2016
Accepted 17 February 2016
Available online 1 March 2016

Keywords:
Driver injury severity
Rollover crash
Support vector machine model
Kernel function
Traffic safety

ABSTRACT

Rollover crash is one of the major types of traffic crashes that induce fatal injuries. It is important to investigate the factors that affect rollover crashes and their influence on driver injury severity outcomes. This study employs support vector machine (SVM) models to investigate driver injury severity patterns in rollover crashes based on two-year crash data gathered in New Mexico. The impacts of various explanatory variables are examined in terms of crash and environmental information, vehicle features, and driver demographics and behavior characteristics. A classification and regression tree (CART) model is utilized to identify significant variables and SVM models with polynomial and Gaussian radius basis function (RBF) kernels are used for model performance evaluation. It is shown that the SVM models produce reasonable prediction performance and the polynomial kernel outperforms the Gaussian RBF kernel. Variable impact analysis reveals that factors including comfortable driving environment conditions, driver alcohol or drug involvement, seatbelt use, number of travel lanes, driver demographic features, maximum vehicle damages in crashes, crash time, and crash location are significantly associated with driver incapacitating injuries and fatalities. These findings provide insights for better understanding rollover crash causes and the impacts of various explanatory factors on driver injury severity patterns.

1. Introduction

As reported by the National Highway Traffic Safety Administration (NHTSA), there were 3009 rollover crashes in the U.S. in 2012, which accounted for 10% of all fatal crashes in the country (National Highway Traffic Safety Adminstriation, 2013). These numbers were even higher for New Mexico. According to the New Mexico Department of Transportation (NMDOT) (New Mexico Department of Transportation, 2012), rollover crashes accounted for 5.2% of total statewide reported crashes, but resulted in 34.6% of total fatal crashes and 36.2% of occupant fatalities. Statistics also revealed that rollover crashes were mostly single-vehicle involved and accounted for 35% of all fatalities in single-vehicle crashes (Fréchède et al., 2011). The significant loss of life resulting from rollover crashes indicates the emergent need of comprehensive and in-depth investigation of rollover crash mechanisms. Numerous studies have been conducted to examine rollover crashes and their contributing factors, injury outcome patterns, and effective countermeasures. Due to the significant weight and size, rollover crashes are most likely to occur when heavy vehicles, such as pickup trucks, semi-trailers, and farming tractors are used. For instance, Farmer and Lund (2002) concluded that light trucks experience a higher potential of rollover crashes than passenger cars. Significant studies were also performed to investigate the injury patterns in rollover crashes. For example, Huelke and Compton (1983) discovered that ejected occupants have a 17 times higher risk of more serious and fatal injuries than restrained occupants. Head, neck, and spine injuries are the primary injuries in rollover crashes because of the roof deformation and its crushing impact on human cephalic and vertebral parts (Conroy et al., 2006; Funk et al., 2012; Mandell et al., 2010). To reduce crash risk and injury severities in rollover crashes, various preventive countermeasures have been proposed, tested, and implemented in many peer studies (Chen et al., 2012; Harris et al., 2011; Liu and Koc, 2013; Mangado et al., 2007; Reynolds and Groves, 2000; Wu et al., 2014). For example, Liu and Koc (2013) devised a mobile application based on an IOS system for monitoring tractor running stability and reporting rollover incidents with detailed spatial, temporal, and other relative information.

* Corresponding author. Fax: +1 505 277 1988.
E-mail address: guohui@unm.edu (G. Zhang).

http://dx.doi.org/10.1016/j.aap.2016.02.011

Statistical models are the primary method used in traffic crash analyses (Chen et al., 2015a,b; Liu et al., 2015b; Lord and Mannering, 2010; Wu et al., 2015). However, these models are based on certain assumptions regarding data and model structures, which inevitably pose limitations in these studies. For instance, ordered logit models have been widely used in crash severity analyses to model the ordinal nature of injury severity outcomes under the proportional odds assumption that the impacts of a contributing factor are identical across different ordinal levels, which in most circumstances will not hold. When these assumptions are violated, statistical bias or erroneous results are induced (Lord and Mannering, 2010; Savolainen et al., 2011). To overcome the limitations of statistical models in traffic crash analyses, non-parametric models, such as neural network, classification and regression tree (CART), decision tables, have been introduced and widely utilized (Abdelwahab and Abdel-Aty, 2001; Chang and Chien, 2013; Chang and Wang, 2006; Chen et al., 2016). Among these non-parametric models, support vector machine (SVM) is a modeling technique developed to address classification and regression problems. SVM models have been increasingly implemented in transportation research to address traffic flow forecasting (Cheu et al., 2006; Huang, 2015; Wei and Liu, 2013; Yu et al., 2013; Zhang et al., 2013), crash frequency assessment (Li et al., 2008; Ren and Zhou, 2011; Suárez Sánchez et al., 2011; Yu and Abdel-Aty, 2013a,b), and travel mode and travel pattern prediction (Allahviranloo and Recker, 2013; Sheng and Xiao, 2015). Recently, several peer studies have also applied SVM to examine injury severities at crash level, i.e. the most severe occupant injury outcome in a crash (Guo et al., 2012; Li et al., 2012; Yu and Abdel-Aty, 2014). For example, Guo et al. (2012) proposed a pedestrian recognition model applicable to intelligent transportation systems based on AdaBoost algorithms and SVM to reduce pedestrian suffering in traffic crashes. However, SVM models have never been used to examine injury severity patterns in rollover crashes. Besides, in previous crash severity analyses, SVM models considered each crash as a unit and examined the most severe injury outcome in each crash, which may not enough reveal the detail injury severity patterns in the crash. For instance, assuming that there are equal numbers of people involved in two different crashes where there is only one fatality in one crash but ten fatalities in the other, in traditional crash severity analysis, these two crashes are both considered as fatal crashes, but in fact the second one is much more severe than the first in terms of number of fatalities. Therefore, the authors are motivated to conduct this research to investigate the individual injury patterns taking each driver/vehicle record as a research unit.

The primary objectives of this study are to investigate the applicability of SVM models in driver injury severity analyses and use it to examine driver injury severity patterns in rollover crashes. A disadvantage of the SVM models is that it lacks the capability of automatically selecting significant factors contributing to the target variable. Therefore, a CART analysis is conducted to rank variable relative importance and identify significant variables for driver injury prediction. The rest of this paper is organized as follows: Section 2 provides a comprehensive and in-depth literature review regarding rollover crashes and SVM applications. The data description and processing procedure are introduced in Section 3, followed by the methodology design and model specifications in Section 4. Research results are explicitly discussed in Section 5, and the research limitations and overall research effort are summarized in Section 6.

2. State of the art

Rollover crashes have been widely investigated from multiple perspectives with different methods. Abundant studies have been performed to examine the crash mechanisms and injury patterns in rollover crashes. Hu and Donnell (2011) applied a multinomial logit model to examine the significant factors in predicting rollover crash severity in rural divided highways. Through logistic regression analyses, Bambach et al. (2013a,b) examined risk factors among human, vehicle, and environmental features on occupant thoracic injuries in rollover crashes. Chang et al. (2006) investigated passenger injury outcomes and the associated risk factors in motor coach rollover crashes and discovered that upside passengers thrown from seats and their downside neighbors were most likely to suffer major injuries. Rollover crashes primarily occurr among heavy vehicles, and considerable research has been conducted to explore the mechanisms in terms of contributing factors and kinetic features. Whitfield and Jones (1995) revealed that overhead weight on top of SUVs and pickups increases vehicle rollover risk. Franceschetti et al. (2014) proposed a model to simulate the kinetic features of tractor lateral rollovers considering the geometric features, tractor inertia nature, and environmental conditions at crash occurrence. Albertsson et al. (2006) investigated the injury outcome patterns, injury risk and mechanism, and the protective effect of seatbelts in rollover coach crashes. Rollover crashes are intimately associated with vehicle roof and side-structure deformation, resulting in severe head or neck injuries of vehicle occupants. Yoganandan et al. (1990) developed a kinetic 3-dimensional model using articulated body structure to examine the impacts of vehicle roof deformation and head clearance on head and neck injury risk in rollover crashes and identified the optimal roof and head clearance configuration to minimize injury risk. Freeman et al. (2012) discovered that a significant correlation exists between vehicle roof deformation and occupant head and neck injury and developed a head and neck injury score (HNIS) to predict roof crash potential in rollover crashes. Dobbertin et al. (2013) also verified the cause-effect association between vehicle roof crash deformation and occupant head, neck, and spine injury severity in rollover crashes. Bambach et al. (2013b) explored the mechanisms of spine injuries of passengers with seatbelt restraints in rollover crashes, concluding that a roof intrusion of 20 centimeters is preferable in order to avoid spine injuries.

Because of these kinetic and injury characteristics of rollover crashes, multiple protective countermeasures have been proposed to reduce driver injury risk and severity in rollover crashes. van der Westhuizen and Els (2013) verified that slow lateral maneuverers could significantly reduce the potential of vehicle rollovers. Roll-over protective structures (ROPS) are popular devices installed on vehicles to minimize rollover crash casualties. Mangado et al. (2007) developed an affordable ROPS for agricultural tractors and established a performance model to examine the effectiveness of the proposed ROPS based on their beam physical features. Chen et al. (2012) investigated the association between the lateral stiffness coefficients (LSC) of a ROPS and human injury levels in rollover crashes, concluding that appropriate LSC varies across different ROPSs. Reynolds and Groves (2000) concluded that law enforcement and driver educational programs are recommended to improve the effectiveness of ROPS in roll over crash prevention. Harris et al. (2011) assessed the performance of existing cost-effective ROPS at reducing rollover vehicle fatalities according to national standards, concluding that a re-design of these devices is necessary before implementation.

The SVM model is a relatively new method in classification problems, and has been increasingly utilized in traffic safety research, including traffic incident detection, crash frequency prediction, and crash severity investigation. Li et al. (2008) evaluated the efficiency of SVM models in predicting motor vehicle crash occurrences, concluding that SVM models outperform negative binomial models and back-propagation neural networks in crash data prediction. Suárez Sánchez et al. (2011) applied a SVM model to forecast