ARTICLE IN PRESS

International Journal of Transportation Science and Technology xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

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

Factors influencing the patterns of wrong-way driving crashes on freeway exit ramps and median crossovers: Exploration using 'Eclat' association rules to promote safety

Subasish Das [a,*], Anandi Dutta [b], Mohammad Jalayer [c], Apoorba Bibeka [a], Lingtao Wu [a]

[a] Texas A&M Transportation Institute, 3135 Texas A&M University, College Station, TX 77843, United States
[b] Computer Science and Engineering Department, Texas A&M University, College Station, TX 77843, United States
[c] Center for Advanced Infrastructure and Transportation, Rutgers, The State University of New Jersey, 100 Brett Road, Piscataway, NJ 08854-8058, United States

ARTICLE INFO

Article history:
Received 16 December 2017
Received in revised form 27 February 2018
Accepted 27 February 2018
Available online xxxx

Keywords:
Wrong-way driving crashes
Traffic safety
Association rules mining
Eclat algorithm

ABSTRACT

Wrong-way driving (WWD) has been a constant traffic safety problem in certain types of roads. These crashes are mostly associated with fatal or severe injuries. This study aims to determine associations between various factors in the WWD crashes. Past studies on WWD crashes used either descriptive statistics or logistic regression to identify the impact of key contributing factors on frequency and/or severity of crashes. Machine learning and data mining approaches are resourceful in determining interesting and non-trivial patterns from complex datasets. This study employed association rules 'Eclat' algorithm to determine the interactions between different factors that result in WWD crashes. This study analyzed five years (2010–2014) of Louisiana WWD crash data to perform the analysis. A broad definition of WWD crashes (both freeway exit ramp WWD crashes and median crossover WWD crashes on low speed roadways) was used in this study. The results of this study confirmed that WWD fatalities are more likely to be associated with head-on collisions. Additionally, fatal WWD crashes tend to be involved with male drivers and off-peak hours. Driver impairment was found as a critical factor among the top twenty rules. Despite several other studies identifying only the WWD contributing factors, this study determined several influencing patterns in WWD crashes. The findings can provide an excellent opportunity for state departments of transportation (DOTs) and local agencies to develop safety strategies and engineering solutions to tackle the issues associated with WWD crashes.

Introduction

Wrong-way driving (WWD) crashes occur when a driver, intentionally or unintentionally, drives in the opposite direction of traffic flow. These crashes have a higher probability of fatal consequences (1.34 fatalities per fatal crash) compared to other types of crashes (1.10 fatalities per fatal crash) since, being likely head-on or opposite-direction sideswipe collisions.

Peer review under responsibility of Tongji University and Tongji University Press.
* Corresponding author.
E-mail addresses: s-das@tti.tamu.edu (S. Das), anandixd@tamu.edu (A. Dutta), mohammad.jalayer@rutgers.edu (M. Jalayer), a-bibeka@tti.tamu.edu (A. Bibeka), l-wu@tti.tamu.edu (L. Wu).

https://doi.org/10.1016/j.ijtst.2018.02.001
2046-0430/

According to the Federal Highway Administration (FHWA), approximately 300–400 people die each year due to WWD crashes in the U.S. (Federal Highway Administration, 2017). These values highlight the need for novel preventive approaches to mitigate the number of WWD crashes, which we address in this paper.

The National Transportation Safety Board (NTSB) defines that 'WWD is vehicular movement along a travel lane in a direction opposing the legal flow of traffic on high-speed divided highways or access ramps' (NTSB, 2012). This definition restricted WWD crashes only on controlled-access highways. This report has not included WWD crashes that result from median crossover encroachments. While the majority of the studies address WWD crashes on freeways, the crash analysis on median crossover WWD crashes has not been conducted in depth. This study aims to overcome this gap.

WWD crashes have been a subject of intense scrutiny over the past decades. Researchers have studied WWD crashes in different states in the U.S. including, Texas, Illinois, North Carolina, Michigan, and New Mexico (Braam, 2006; Finley et al., 2014; Lathrop et al., 2010; Morena and Leix, 2012; Zhou et al., 2015). Most of the previous studies used the descriptive statistics to explore the role of different factors associated with WWD crashes on high speed roadways and freeways. Pour-Rouholamin, Kemel, and Ponnaluri have employed other techniques such as Firth's penalized likelihood logistic regression, logistic regression and generalized order logit model to analyze the WWD crash data (Kemel, 2015; Ponnaluri, 2016; Pour-Rouholamin et al., 2016). Based on several previous studies (Braam, 2006; Cooner et al., 2004; Copelan, 1989; Lathrop et al., 2010; Morena and Leix, 2012; Scaramuzza and Cavegn, 2007; Zhou et al., 2015), deadly WWD crashes appear to involve intoxicated drivers. Moreover, these studies identified other significant confounding factors associated with WWD crashes such as driver age, driver gender, and time of day. Several previous studies (Morena and Leix, 2012, Zhou et al., 2015) also found the significant role of dark roadway conditions on the likelihood of WWD crashes. Studies using parametric techniques such as logistic regression and Firth's Penalized Likelihood Logistic Regression also explored that intoxicated drivers and darkness cause WWD crashes as significant factors (Kemel, 2015; Ponnaluri, 2016; Pour-Rouholamin et al., 2016). These studies are pivotal in coming up with countermeasures to reduce WWD crashes. However, these studies either used descriptive statistics alone or employed parametric methods to model the underlying relationships in the data. It should be noted that descriptive statistics might be insufficient to untangling complex relationships between different variables, whereas parametric models make assumptions about the distribution of the independent and dependent variables that might not always be true. In the recent years, machine learning and data mining methods have been widely used in transportation safety research to overcome the assumption issues associated with statistical modeling (Chen and Xie, 2015; Sun et al., 2014; Dong et al., 2015; Das et al., 2018b,c; Khan et al., 2015; Iranitalab and Khattak, 2017; Das and Sun, 2016, 2015). Machine learning models can detect interactions between different features and mask them if necessary. Data mining methods do not depend on any assumptions because the generated rules and patterns would show either interestingness or redundancy.

Two recent studies used multiple correspondence analysis (MCA) method (Jalayer et al., 2017; Das et al., 2018a) to explore WWD crashes. Compared to the parametric methods, MCA has a lower bias as this method does not make any prior assumptions about the variables. One of the limitations of this method is that it does not enable researchers to conduct significance test on the clusters. Moreover, this technique only informs about the correlation between variables but not about causation. To overcome the limitation of MCA at the same time retaining the low bias advantage of MCA, we analyzed five years (2010–2014) of Louisiana WWD crashes (for the remainder of this paper, Louisiana wrong way crashes will be referred as WWD crashes for consistency) using association rules. Frequent pattern mining (FPM) and association rules are powerful machine learning tools to identify the relationship between different factors. FPM recognizes the set of items that tend to occur together and association rules help understand the causal relationship between a set of antecedent items with the following item. The researchers are interested in finding long patterns as a multitude of factors can cause WWD crashes. Vertical data mining technique called "Eclat" performs better when long patterns are required thus the researchers used "Eclat" to analyze the data. This study aims to identify interdependence between various factors leading to WWD crashes.

Theory

According to Aggarwal and Han, frequent pattern mining (FPM) involves finding relationships among the items in a database (Aggarwal and Han, 2014). Several recent studies have employed association rules mining to determine significant associations for different safety problems (Das et al., 2018d; Das et al., 2017a,b; Das and Dutta, 2014; Das and Sun, 2014; Geurts et al., 2005; Mirabadi and Sharifian, 2010; Montella, 2011; Pande and Abdel-Aty, 2009).

FPM can be utilized to obtain association rules. A frequent pattern is obtained by comparing the support for a pattern with a certain minimum threshold. A pattern consists of various items in the dataset. Association rules mining uses three parameters to perform the analysis. These parameters are: support, confidence, and lift. The support of a pattern is determined by how frequently that pattern occurs in the dataset. The confidence of an association rule can be obtained by identifying the ratio of the number of times antecedent and consequent occur together by the number of time the antecedent occurs in the dataset. A third parameter called "lift" helps us determine the type of association between the antecedent and the consequent. Lift value higher than 1 indicates positive interdependence between the antecedent and the consequent, whereas a value less than 1 indicates a negative interdependence. These parameters can be expressed as:

$$S(X) = \frac{N_X}{N} \tag{1}$$