Transportation Research Part A 69 (2014) 58–70

Contents lists available at ScienceDirect

Transportation Research Part A

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

Identification of freeway crash-prone traffic conditions for traffic flow at different levels of service

Chengcheng Xu [1], Pan Liu *, Wei Wang, Zhibin Li

Jiangsu Key Laboratory of Urban ITS, Southeast University, Si Pai Lou #2, Nanjing 210096, China
Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Si Pai Lou #2, Nanjing 210096, China

ARTICLE INFO

Article history:
Received 22 June 2013
Received in revised form 14 May 2014
Accepted 17 August 2014

Keywords:
Crash risk
Level of service
Bayesian inference
Markov Chain Monte Carlo (MCMC)
Random forest

ABSTRACT

The primary objective of this study was to evaluate the risks of crashes associated with the freeway traffic flow operating at various levels of service (LOS) and to identify crash-prone traffic conditions for each LOS. The results showed that the traffic flow operating at LOS E had the highest crash potential, followed by LOS F and D. The traffic flow operating at LOS B and A had the lowest crash potential. For LOS A and B, the vehicle platoon and abrupt change in vehicle speeds were major contributing factors to crash occurrences. For LOS C, crash risks were correlated with lane-change maneuvers, speed variation, and small headways in traffic. For LOS D, crash risks increased with an increase in the temporal change in traffic flow variables and the frequency of lane-change maneuvers. For LOS E, crash risks were mainly affected by high traffic volumes and oscillating traffic conditions. For LOS F, crash risks increased with an increase in the standard deviation of flow rate and the frequency of lane-change maneuvers. The findings suggested that the mechanism of crashes were quite different across various LOS. A Bayesian random-parameters logistic regression model was developed to identify crash-prone traffic conditions for various LOS. The proposed model significantly improved the prediction performance as compared to the conventional logistic regression model.

1. Introduction

Considerable effort has been devoted to understanding to what extent freeway traffic flow affects safety. Using real-time traffic flow data from inductive loop detectors, a number of studies have developed crash risk prediction models to establish a statistical relationship between the risks of crashes and traffic flow variables, such as speed, occupancy and speed variance, etc. It was found that freeway traffic flow are significantly related to the risks of crash occurrences (Oh et al., 2001; Golob and Recker, 2003, 2004; Golob et al., 2004a; Lee et al., 2002, 2003; Abdel-Aty et al., 2004; Hourdos et al., 2006; Christoforou et al., 2010; Hossain and Muromachi, 2010; Pande et al., 2011; Xu et al., 2013b, 2013c, 2013d; Golob et al., 2008; Zheng et al., 2010).

The crash risk prediction models identify the crash-prone conditions prior to crashes given real-time traffic flow data from inductive loop detectives. Such information is needed in dynamic freeway traffic management systems to apply proactive traffic control strategies that aim at reducing crash risks (Lee et al., 2006; Khoury and Hobeika, 2006; Allaby et al., 2007; Hossain and Muromachi, 2011). Most of the crash risk models that were developed by previous studies are generic

* Corresponding author at: Jiangsu Key Laboratory of Urban ITS, Southeast University, Si Pai Lou #2, Nanjing 210096, China.
E-mail addresses: iamxcc1@163.com (C. Xu), pan_liu@hotmail.com (P. Liu), wangwei@seu.edu.cn (W. Wang), lizhibin-2002@163.com (Z. Li).
[1] Tel.: +86 13801580045.

http://dx.doi.org/10.1016/j.tra.2014.08.011
0965-8564/

in nature. It was assumed that a consistent model can be applied without taking into consideration the macroscopic traffic flow states on freeways. Such assumption ignores the fact that drivers behave differently across various macroscopic traffic flow states. As a result, the mechanism of crash occurrences may be quite different in different traffic states.

Several studies developed separate crash risk prediction models for spare and heavy traffic conditions (Hauer, 2002; Lord et al., 2005; Abdel-Aty et al., 2005; Xu et al., 2012). Hauer (2002) and Lord et al. (2005) found that crash rates increase with an increase in traffic density during spare traffic conditions, while decrease as traffic density increases during heavy traffic conditions. Abdel-Aty et al. (2005) developed separate crash risk prediction models for high and low speed regimes. The results suggested that the traffic flow characteristics that contribute to crash risks are different between high and low speed regimes. Xu et al. (2012) divided freeway traffic flow into five states using occupancy measured at four nearby loop detector stations. It was found that that the traffic states significantly affect crash risks. The traffic flow variables that contribute to crash risks were quite different between spare and heavy traffic conditions.

Golob and Recker (2004) classified freeway traffic flow into various regimes with distinct crash characteristics. It was found that each of the traffic flow regimes has a unique profile in terms of the type of crashes that are most likely to occur. The multi-vehicle and property damage only crashes are more likely to occur in the traffic sates with high densities, while the single-vehicle and injury crashes are more likely to occur in the traffic states with low densities. The crashes occurring on left-lanes are associated with traffic states with high density and high flow conditions, while crashes occurring at other locations, especially interior lane, are associated with traffic states with low density and low flow conditions. In a subsequent study, Golob et al. (2004a, b) further classified freeway traffic flow into eight states, and explored the prevailing collisions types for each state. It was found that the prevailing collision types differ significantly across various traffic flow states. The studies conducted by Golob et al. (2004a, b) and Golob and Recker (2003, 2004) provided deep insights into the relationship between traffic flow state and the characteristics of crashes given the fact that a crash has occurred. However, research is still needed to quantitatively evaluate to what extent traffic flow states affect the risks of crash occurrences (Golob et al., 2004a; Golob and Recker, 2004).

Previous studies have demonstrated that freeway traffic can be divided into various homogenous states (Hall et al., 1992; Kerner and Rehborn, 1996; Wu 2002; Xia and Chen, 2007). So far, the most widely accepted procedure for the classification of freeway traffic state is proposed by the Highway Capacity Manual (HCM), in which freeway traffic flow is divided into six scenarios on the basis of traffic density (TRB 2010). A service level is assigned to each scenario. According to the HCM, the level of service (LOS) is a quantitative stratification of a performance measure or measures that represent the quality of service provided to travelers. The LOS associated with each freeway traffic state varies from the LOS A for the undersaturated traffic flow to the LOS F for the oversaturated traffic flow.

The HCM analytical procedure has been widely used in practical engineering applications for evaluating the effects of traffic management policies on the operations of freeway traffic. However, until recently, it is still not sure how LOS can be related to safety. The research objective of this study was to evaluate the crash risks associated with the freeway traffic flow operating at various LOS and to identify crash-prone traffic conditions for each LOS. The research results have potential to help incorporate safety into the current HCM analytical procedure that focuses primarily on delay and travel time. In addition, it was expected that by considering the differences in the crash mechanism across various traffic flow states one can better understand the contributing factors to crash risks.

The present study involves three steps. First, a pilot study was conducted to examine whether traffic flow operating at various LOS differs distinctly in terms of crash risks. More specifically, a Bayesian conditional logistic regression model was developed to identify how LOS affected the risks of crashes. The random forest (RF) analyses were then conducted to identify the major traffic flow variables that contribute to the crash risks under each LOS. With the traffic flow variables selected by the RF analyses, a Bayesian random-parameters logistic regression model was developed to evaluate real-time crash risks considering varying contributing factors across various traffic states.

2. Data

Crash and traffic data were collected from a 22-mile freeway segment from the milepost 13.5–35 on the I-880 N freeway in the United States. A total of 46 loop detector stations and three weather stations are located along the selected freeway segment (see Fig. 1). The average spacing between detector stations is less than 0.5 miles, and the standard deviation of the spacing is around 0.3 miles. All the three weather stations are located within five miles from the I-880 N freeway. The average spacing between weather stations is about seven miles. Crash and the paired real-time traffic data were collected from January 1, 2010 to December 31, 2010. A total of 509 crashes were identified and used for further data analysis.

The data were extracted from the database of the Highway Performance Measurement System (PeMS) maintained by the California Department of Transportation. The database provided 30-s raw loop detector data, including traffic volume, vehicle speed and traffic occupancy. The traffic data were collected from the nearest loop detector station to crashes (see Fig. 1). The research team extracted 15-min traffic data five minutes prior to crash occurrences. The purpose was to identify the crash-prone traffic conditions prior to the crash occurrence time to make preemptive measures possible (Abdel-Aty et al., 2004; Pande et al., 2011; Ahmed et al., 2012; Xu et al., 2013a). For example, if a crash occurred at 14:00 pm, the traffic data were extracted from 13:40–13:55 pm. The time lag of 5 min was also adopted in previous studies to develop real-time crash risk models (Abdel-Aty et al., 2004; Pande et al., 2011; Hossain and Muromachi, 2011; Ahmed et al., 2012; Xu et al., 2013a).