Transportation Research Part C 81 (2017) 172–187

Contents lists available at ScienceDirect

Transportation Research Part C

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

A Bayesian Network model for contextual versus non-contextual driving behavior assessment

Xiaoyu Zhu [a,*], Yifei Yuan [b], Xianbiao Hu [a], Yi-Chang Chiu [c], Yu-Luen Ma [d]

[a] Metropia Inc., 1790 East River Road, Suite 140, Tucson, AZ 85718, USA
[b] Department of Systems and Industrial Engineering, University of Arizona, 1127 E. James E. Rogers Way, Tucson, AZ, USA
[c] Department of Civil Engineering and Engineering Mechanics, University of Arizona, 1209 E. Second St, Tucson, AZ, USA
[d] Department of Finance, Insurance, Real Estate and Law, University of North Texas, Denton, TX 76203, USA

ARTICLE INFO

Article history:
Received 22 August 2016
Received in revised form 9 February 2017
Accepted 26 May 2017

Keywords:
Contextual driving risk analysis
Bayesian Network model
Information-aggregation
Regression models

ABSTRACT

Driving behavior is generally considered to be one of the most important factors in crash occurrence. This paper aims to evaluate the benefits of utilizing context-relevant information in the driving behavior assessment process (i.e. contextual driving behavior assessment approach). We use a Bayesian Network (BN) model that investigates the relationships between GPS driving observations, individual driving behavior, individual driving risks, and individual crash frequency. In contrast to prior studies without context information (i.e. non-contextual approach), the data used in the BN approach is a combination of contextual features in the surrounding environment that may contribute to crash risk, such as road conditions surrounding the vehicle of interest and dynamic traffic flow information, as well as the non-contextual data such as instantaneous driving speed and the acceleration/deceleration of a vehicle. An information-aggregation mechanism is developed to aggregates massive amounts of vehicle GPS data points, kinematic events and context information into drivel-level data. With the proposed model, driving behavior risks for drivers is assessed and the relationship between contextual driving behavior and crash occurrence is established. The analysis results in the case study section show that the contextual model has significantly better performance than the non-contextual model, and that drivers who drive at a speed faster than others or much slower than the speed limit at the ramp, and with more rapid acceleration or deceleration on freeways are more likely to be involved in crash events. In addition, younger drivers, and female drivers with higher VMT are found to have higher crash risk.

1. Introduction

Driving behavior is generally considered to be one of the most important factors in crash occurrence, yet due to the stochastic nature of driving, measuring and modeling driving behavior continues to be a challenging topic today (Sagberg et al., 2015). In most prior studies, driving behavior data was collected from self-reported surveys, such as the Driving Behavior Questionnaire (DBQ) (Rowe et al., 2015), driving simulator (Chen et al., 2013) or actual in-vehicle observations (Toledo and Lotan, 2006). The relationship between driving behavior and crash involvement is generally assumed and studied (Guo

* Corresponding author.
E-mail addresses: xiaoyu.zhu@metropia.com (X. Zhu), yifeiyuan@email.arizona.edu (Y. Yuan), xb.hu@metropia.com (X. Hu), chiu@email.arizona.edu (Y.-C. Chiu), Yu-Luen.Ma@unt.edu (Y.-L. Ma).

http://dx.doi.org/10.1016/j.trc.2017.05.015
0968-090X/

et al., 2010; Paefgen et al., 2012, 2014; Sun et al., 2017), with the primary purposes being to distinguish the driving behavior of safe drivers versus unsafe drivers (Jun et al., 2011), to identify critical safety-related events (Wu et al., 2014), and to understand the variant in driving behavior. With these studies, individual driving behavior can be assessed and used as a representative of crash risk, and broader applications can be built upon this foundation, such as novice driver education (Simons-Morton et al., 2013), feedback to improve driving (Toledo et al., 2008; Ellison et al., 2015), and Pay-as-You-Drive (PAYD) insurance programs (Ferreira and Minike, 2010; Paefgen et al., 2014).

Depending on whether driving information is context relevant, driving behavior can be contextual or non-contextual. Generally speaking, the motion of a vehicle, described by vehicle kinematic parameters like speed and acceleration/deceleration, is relied on to measure driving behavior (Af Wåhlberg, 2008; Paefgen et al., 2011; Ellison et al., 2015) which could be referred to as a non-contextual driving behavior approach, as it solely applies vehicle information, but does not consider the surrounding environment of the driver that may contribute to crash risk such as road conditions surrounding the vehicle of interest and dynamic traffic flow information. Research utilizing non-contextual driving behavior approach has its limitation. For example, if a driver exhibits stop-and-go or abrupt accelerate/decelerate behavior, without supplemental information on traffic conditions, it is usually difficult to tell if this is simply due to the driver's behavior, or caused by heavy traffic conditions. Relative speed to the average traffic speed is another example. If a driver is observed to be consistently driving at a speed much higher than the other people around him/her, it's reasonable to suspect that most likely this driver is driving more aggressively than average drivers. Following this rationale, this paper proposes to define driving behavior by incorporating both habitual driving behavior (i.e. driving style) and instantaneous reactions under various contexts (i.e. contextual actions), and aims to evaluate the benefits of utilizing context-relevant information in the driving behavior assessment process (i.e. contextual driving behavior assessment approach), compared with prior studies without context information (i.e. non-contextual approach).

Prior research with the contextual driving behavior assessment approach is rather limited. Jun et al. investigated whether the exposure of crash-involved drivers and driving performance were different on the basis of disaggregated analyses by facility types and trip start times through a 14-month study (Jun et al., 2007). In a prior study, by considering contextual driving behavior, including relative speed, which is measured by the combination of vehicle moving speed, real-time traffic speed and speed limit of the road, we verified that the contextual driving behaviors have strong relationships with the driver's crash involvement (Zhu et al., submitted for publication). However, certain research questions are still yet to be answered, which leads to the research objectives of this paper.

1. Are there any benefits of utilizing contextual driving behavior assessment approach as opposed to non-contextual approach, and if so, to what extent?
2. With the challenges in information aggregation and the various data size volume for each driver, how do we extract contextual and non-contextual driving behavior from the detailed longitudinal vehicle telematics data at the GPS point level, combined with corresponding contextual and spatial-temporal features at the road segment level?
3. How could we relate individual driving behavior with crash involvement frequency when data is aggregated at the driver level, and explain the relationship?

To answer the questions above, a hierarchical Bayesian Network (BN) model is proposed as the methodology foundation with the vehicle trajectory data and corresponding spatial-temporal contextual data as input. The Bayesian method is effective in handling massive trajectory data, and an information-aggregation mechanism and regression models are built to fulfill the second and third research objectives. The advantages of the proposed information-aggregation mechanism lies in its flexibility in the data aggregation process, where massive amounts of GPS level vehicle kinematics context data and drivel-level data can have their own statistical property, and its capability in handling various volumes of data for each driver. With the regression model embedded in the BN structure, the relationship between driving behavior and crash risk can be investigated and explained, and a close estimate of the driver's crash risk can be obtained. With the proposed framework, two BN models are further presented to quantitatively evaluate the benefit of contextual driving behavior, by comparing the first model with only vehicle kinematics events but without contextual features, and the second model with both vehicle kinematics events and context-relevant information. The contextual information used in this study includes roadway type, which indicates various road facilities types such as freeway, arterial, ramp, etc., and speed related information including speed limit and average traffic speed information. The detailed data utilized and methodology will be described in Section 3.

With this modeling framework, an understanding of the contextual driving behaviors and their relationships with crash involvement can be achieved. The outcome of the research can be used to actively advise the drivers, and encourage safe driving behaviors under specific scenarios to reduce the incidences of traffic crashes. In the sections below, literature focused on driving behavior assessment from vehicle kinematics will firstly be reviewed. Next, the proposed hierarchical Bayesian Network model with three main components will be presented. The following case study section describes the data, compares the results of two BN models, and discusses the BN model for contextual driving behavior assessment. The research contributions and limitations will be summarized at the end of this paper.