Transportation Research Part C 28 (2013) 28–43

Contents lists available at SciVerse ScienceDirect

Transportation Research Part C

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

Black spots identification through a Bayesian Networks quantification of accident risk index

Andreas Gregoriades [a,*], Kyriacos C. Mouskos [b]

[a] Dept. of Computer Science and Engineering, European University Cyprus, Cyprus
[b] CCNY-CUNY, Convent Ave., Marshak 910, New York, NY 10031, USA

ARTICLE INFO

Article history:
Received 18 December 2011
Received in revised form 28 September 2012
Accepted 14 December 2012

Keywords:
Road safety
Bayesian Networks
Dynamic Traffic Assignment
Crash analysis
Accident analysis

ABSTRACT

Traffic accidents constitute a major problem worldwide. One of the principal causes of traffic accidents is adverse driving behavior that is inherently influenced by traffic conditions and infrastructure among other parameters. Probabilistic models for the assessment of road accidents risk usually employs machine learning using historical data of accident records. The main drawback of these approaches is limited coverage of traffic data. This study illustrates a prototype approach that escapes from this problem, and highlights the need to enhance historical accident records with traffic information for improved road safety analysis. Traffic conditions estimation is achieved through Dynamic Traffic Assignment (DTA) simulation that utilizes temporal aspects of a transportation system. Accident risk quantification is achieved through a Bayesian Networks (BNs) model learned from the method's enriched accidents dataset. The study illustrates the integration of BN with the DTA-based simulator, Visual Interactive Systems for Transport Algorithms (VISTAs), for the assessment of accident risk index (ARI), used to identify accident black spots on road networks.

1. Introduction

Road accident statistics in Europe stress the need for more systematic mechanisms for accident analysis and prediction. According to the World Health Organization road accidents constitute one of the leading cause of death for people between the ages of 5–44 (Kapp, 2003; WHO, 2011). Given the current trends, accident fatalities are projected by 2020 to become the fifth leading cause of death worldwide resulting in an estimated 2.4 million deaths each year (WHO, 2011). At the same time, traffic accidents result in high economic losses due to traffic congestion which in turn leads to a wide variety of adverse consequences such as, traffic delays, supply chain interruptions, travel time unreliability, increased noise pollution, as well as deterioration of air quality. To combat these and the intrinsic accident risks, road safety has emerged as a priority alongside road safety management and forecasting practices. These however, suffer from major limitations and need improvement to effectively tackle this problem. One of the problems faced is data availability for the development of crash prediction and analysis models. This work contributes in this direction through the development of a prototype accident risk index quantification approach that overcomes the data availability problem by combining simulated data with historical data for the development of a BN accident prediction model.

Inherently, road networks constitute complex dynamic and uncertain systems influenced by human, technological and environmental parameters. Therefore, one of the best ways to understand the causes of road traffic accidents is to develop models capable integrating significant factors relating to human, vehicle, socio-economic, infrastructural, and environmental

* Corresponding author. Tel.: +357 22713224.
E-mail addresses: a.gregoriades@euc.ac.cy (A. Gregoriades), mouskos@utrc2.org (K.C. Mouskos).

0968-090X/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2012.12.008

properties. There are two broad categories of accident analysis methods: the qualitative and the quantitative. The former, despite its limited use, plays an important role in the process of accident analysis, modeling and forecasting. Qualitative analysis is subjective, exploratory and interpretative, while quantitative is based on the positivist philosophy and hence more widely used. Quantitative methods are classified into two principal groups: Time-series forecasting and Causality-based forecasting. The accident analysis approach proposed herein combines causality-based with a systemic technique, specifically, BN and Traffic simulation. The former is popular in the Artificial Intelligence domain and is based on the concept of Bayesian probability. BN provide a language and calculus for reasoning under uncertainty and incomplete information (Pearl, 2009). Hence, they are useful for inferring probabilities of future events, on the basis of observations or other evidence that may have a causal relationship to the event in question. Due to these characteristics BNs are becoming very popular in accident analysis. The second technique employed in the proposed approach is the use of a road traffic simulator using simulation-based DTA. The DTA produces estimates of traffic flow conditions for every 15-min time interval of a simulation. These estimates include, traffic flow and speed at link and movement level. Under well calibrated data, DTA estimates can be used as additional explanatory variables in accident prediction models. Overall, the method described herein quantifies accident risk index to predict road sections with high accident frequencies that constitute the network's black spots.

The paper is organized as follows: An overview of the literature and related work is presented first. Next, an outline of the method is shown. Following this, the theoretical underpinnings of BN are illustrated. Subsequently, the approach followed to implement VISTA and BN models is presented. Next, a description of the integration of the two techniques is provided and finally, results from this integration are presented before conclusions are drawn.

2. Literature review

Historically, DTA was pioneered by the United States Federal Highway Administration (FHWA) through sponsoring the development of two mesoscopic DTA modes the DYNASMART-P (Mahmassani et al., 2000) and the DYNAMIT (Ben-Akiva et al., 2002), at the University of Texas and Massachusetts Institute of Technology respectively. Parallel to these efforts, Ziliaskopoulos and Lee (1996) developed the RouteSim mesoscopic simulator and the Visual Interactive System for Transport Algorithms DTA (VISTA-DTA) (Ziliaskopoulos et al., 2004). Currently, many more simulation-based models have been developed around the world as the transport agencies are embracing them as a tool to evaluate various infrastructure and operational network improvements. These include DynusT (Chiu et al., 2008), Dynameq (Florian and Mahut, 2005; Florian et al., 2008), AIMSUN (Barcelo and Casas, 2002), TRANSMODELER (Caliper Corporation, 2009), INTEGRATION (Aerde et al., 1996) and METROPOLIS (de Palma and Marchal, 2002). The DTA model used in this study is realized in VISTA (Ziliaskopoulos et al., 2004; VISTA, 2002; Ziliaskopoulos and Barrett, 2006), which at convergence it reaches a local DUE. The main output of a DTA model is the OD DUE of vehicles' trajectories based on six seconds or less simulation time step. Hence, the analyst can aggregate the traffic flow characteristics at the link, movement, path, sub-network, network level as desired.

The principal characteristics of simulation-based DTA models are: (1) A dynamic Origin–Destination (OD) matrix, estimated using a combination of techniques such as, OD surveys, traffic counts, path trajectories via location estimation devices i.e. GPS and wireless roadside vehicle readers. (2) The DTA model propagates the OD demand using a mesoscopic traffic simulator such as Daganzo's (1994) cell transmission model at every six seconds or less. Vehicles move in packets from one cell to the next subject to the traffic flow theory laws of density, flow and speed. (3) Each vehicle moves along a time-dependent shortest path that is determined at each iteration. (4) The model converges to a Dynamic User Equilibrium (DUE) that states that no user can unilaterally improves his/her travel time (cost) by changing his/her departure or desired arrival time and path within the assignment time interval. Therefore, for each path with vehicles and for a specific time interval, each path will eventually have the same travel time as the other paths for each OD pair. (5) DTA models converge to a local DUE. Global DUE is computationally intractable while until now no model reported has ever claimed global convergence.

A thorough review of the main characteristics of DTA models is presented by Peeta and Ziliaskopoulos (2001). An evaluation of DTA models conducted by Parvathy et al. (2012)) used the following discriminators: simulation unit (link and/or cell), simulation time step (less/equal to 6 s), modeling of signals (pre-timed and/or actuated), stop signs, use of zone connectors, lane connectivity modeling (implicit or explicit), equilibration method (gradient, MSA), iterations needed to reach a DUE (30–60), modeling of a generalized cost function, and the computing platform. VISTA is a DUE convergent model that uses the simplicial decomposition algorithm and is the only Internet-based DTA model running on Linux whereas the remaining are Windows/PC based. This study also demonstrates that the interest in implementing DTA models is increasing, while it points to the challenges of calibration that requires substantially more data than Static Traffic Assignment (STA) models and the slow convergence (Mahut et al., 2004) that requires hours versus minutes for STA models. The main characteristics of DTA models, their differentiation from STA models, the issues of stability and convergence can be found at the primer (Chiu et al., 2011). The primer provides support to the implementation of DTA models as a tool to estimate the traffic flow conditions at 15-min time intervals. This necessitates modeling with sufficient accuracy the aggregated demand (in 15-min time intervals), roadway geometry, traffic control and traveler information devices/services.

Microscopic traffic simulators may produce additional traffic flow characteristics that may provide more explanatory variables on the causation of accidents such as the distribution of the vehicle speeds, the headway distribution among vehicles, acceleration/deceleration and gap acceptance in lane changing maneuvers. Various microscopic simulators could be employed for accident analysis such as CORSIM, VISSIM, PARAMICS, WATSIM (FHWA, 2003) and others. Their main