Transportation Research Part C 82 (2017) 138–160

Contents lists available at ScienceDirect

Transportation Research Part C

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

Use of ubiquitous probe vehicle data for identifying secondary crashes

Hong Yang [a,*], Zhenyu Wang [a], Kun Xie [b], Dong Dai [c]

[a] Department of Modeling, Simulation & Visualization Engineering, Old Dominion University, Norfolk, VA 23529, United States
[b] Department of Civil and Urban Engineering, Tandon School of Engineering, New York University, Brooklyn, NY 11201, United States
[c] Novartis Pharmaceuticals Corporation, NJ 07936, United States

ARTICLE INFO

Article history:
Received 25 August 2016
Received in revised form 21 June 2017
Accepted 22 June 2017

Keywords:
Secondary crashes
Probe vehicle data
Traffic incident
Clustering
Impact area
Shockwave

ABSTRACT

Secondary crashes are non-recurrent incidents that frequently affect traffic operations and safety. They are an important performance measure in evaluating traffic incident management programs. Although several methods (e.g., static, contour map-based, and shockwave-based) have been introduced to identify secondary crashes, the applications of the existing methods are often limited by their shortcomings such as the needs for extra incident information, assumptions, simplified model structures, etc. As an alternative, this paper aims to develop a new data-driven analysis framework to support the identification of secondary crashes. Unlike existing methods, the proposed approach is concentrated on exploring the untapped potential of ubiquitous probe vehicle data for secondary crash analysis. It consists of three major components: detection of the impact area of a primary crash, estimation of the boundary of the impact area, and identification of secondary crashes within the boundary. The first component uses clustering methods to highlight the congested area induced by a primary crash. The second component develops meta-heuristic optimization algorithms to approximate the boundary of the congested area. With the estimated boundary, a novel identification method is introduced to automatically identify secondary crashes within the boundary. The performance of the proposed approach has been tested under a set of simulation scenarios. The test results show that the proposed approach based on the ant colony optimization can best describe the impact area and re-identify up to 95 percent of the simulated crashes. Although the performance of the proposed approach is related to the market penetration rate, the results suggest that a relatively low market penetration rate can already achieve promising performance.

1. Introduction

Traffic incidents often prevent transportation systems from operating in an efficient and safe manner. Vehicle crashes are one of the most frequent types of incidents. These non-recurrent incidents often lead to major backups along the roadways, sometimes for miles, throughout the United States every year. Based on the estimates of the National Highway Traffic Safety Administration (NHTSA), the total economic cost of the motor vehicle crashes in the U.S. was over $240 billion in 2010 (Blincoe et al., 2015). More specifically, the monetary losses in terms delay, extra fuel consumption, and emissions were about $28 billion.

* Corresponding author.
E-mail addresses: hyang@odu.edu (H. Yang), zwang002@odu.edu (Z. Wang), kun.xie@nyu.edu (K. Xie), dong.dai@novartis.com (D. Dai).

http://dx.doi.org/10.1016/j.trc.2017.06.016
0968-090X/

Other than the notable traffic congestion issue, crashes occurred on roadways also increase exposure to secondary crash hazards (FHWA, 2009). It was reported that the risk of having a crash can increase more than six times post the occurrence of a crash (Tedesco et al., 1994). In addition, the likelihood of a secondary crash (SC) increases by 2.8 percent if the primary crash presents for an additional minute (Owens et al., 2009). This often exposes road users and incident responders to higher risk. In addition, it is difficult for the rescue crews to reach and clear the crash scene. It was estimated that secondary crashes accounted for about twenty percent of all crashes and eighteen percent of all fatalities on the US freeways (O'Laughlin and Smith, 2002; Owens et al., 2010). Thus, preventing secondary crashes can result in millions of economic benefit (Peterson, 2015).

According to the Federal Highway Administration (FHWA) program, namely "TIM Performance Measures Focus States Initiative", the number of secondary crashes is considered a core performance measurement (Rensel et al., 2012). For example, many state agencies considered the determination and the reduction of secondary crashes in allocating funding for Road Rangers and the development of TIM programs (MnDOT, 2004; Owens et al., 2009; Pecheux et al., 2014). Reducing the likelihood of secondary crashes was also considered in the planning of freeway service patrol (FSP) deployment (Lou et al., 2011). To reduce the risk of secondary crashes more successfully, the mechanism and characteristics of secondary crash occurrence need to be well understood. For instance, when, where, and how do they occur? Prior to address these questions, however, an immediate question is to identify the secondary crashes. However, almost all crash report forms in practice do not have a label to mark whether a crash is a SC. As a complementary solution, a few studies have proposed several post-event analysis methods to address the issue. For example, Chou and Miller-Hooks (2010) used regression models for determining the corner points of the impact area for identifying secondary crashes. Others considered queuing models (Sun and Chilukuri, 2005; Zhang and Khattak, 2010) and speed contour maps (Chung, 2013; Yang et al., 2013c) to relate a SC to a primary crash (PC). However, these methods often require a wide range of simplified estimation procedures and assumptions (e.g., spatiotemporal windows and simplified models). In addition, the identification of secondary crashes is a challenging task due to the heterogeneous nature of day-to-day traffic conditions.

Considering the potential issues of existing methods, this paper aims to develop a methodological framework to help identify secondary crashes more precisely. The proposed methodological framework consists of three core components including the detection of the impact area of a PC, quantitatively estimating the boundary of the impact area, and the identification of secondary crashes. The first component intends to develop a data-driven approach to exploit the ubiquitous probe vehicle data for highlighting the impact area of a PC. The second component develops the procedure to estimate the boundary of the impact area. The third component automates the detection of secondary crashes based on the estimated boundary. For the sake of implementation, the proposed methodological framework is described in detail. This paper contributed to the literature by introducing a novel approach to leverage the underused probe vehicle data for supporting secondary crash analysis. The proposed approach can capture the progression of the crash impact more accurately and enables the accelerated determination of secondary crashes with quantitative evidence.

Existing literature on identifying secondary crashes will be discussed in the following section. Then the specific problem under study is stated. The proposed methodology is detailed in the fourth section followed by a case study in the fifth section. The results and discussion on testing the proposed approach are described in the sixth section. Finally, the conclusions and implications are presented.

2. Literature review

This paper has performed a state-of-the-art review of the existing studies on identifying secondary crashes. Most of existing studies were conducted in the recent two decades, with a focus on developing different methods/procedures for capturing the impact area of primary crashes. The available methods can be grouped into four categories, including static spatiotemporal threshold-based, queuing model-based, speed contour map-based, and the shockwave-based approaches. The relevant studies of each category were reviewed and discussed below.

2.1. Static spatiotemporal threshold-based method

Most of the existing studies focused on identifying secondary crashes by defining fixed spatiotemporal thresholds to depict the impact area of a crash. For example, Raub (1997a,b) assumed that secondary crashes are those occurred upstream within one mile and within a time frame of primary incident clearance plus 15 min. Many others adopted similar criteria but with some variations on the spatial and temporal thresholds. Table 1 summarized the current studies on secondary crash identification. Obviously, there is no consistent criterion to define the time-space window (thresholds). Although some studies considered the variation of each crash's clearance time, all of them remain static (or in-kind). The subjective spatiotemporal thresholds applied to all conditions (regardless of traffic conditions, roadway geometry, incident characteristics, weather, etc.) are deemed to be the weakness of these simple static approaches. Thus, the performance of these approaches is often questionable due to: (a) overestimation – the spatiotemporal thresholds are too large; and (b) underestimation – the spatiotemporal thresholds are too small.