Transportation Research Part B 58 (2013) 44–57

Contents lists available at ScienceDirect

Transportation Research Part B

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

Modeling the time to the next primary and secondary incident: A semi-Markov stochastic process approach

ManWo Ng [a,b,*], Asad Khattak [c], Wayne K. Talley [d]

[a] Department of Modeling, Simulation and Visualization Engineering, 1318 Engineering and Computational Sciences Building, Old Dominion University, Norfolk, VA 23529, USA
[b] Department of Civil and Environmental Engineering, 1318 Engineering and Computational Sciences Building, Old Dominion University, Norfolk, VA 23529, USA
[c] Department of Civil and Environmental Engineering, 322 J.D. Tickle Bldg, University of Tennessee, Knoxville, TN 37996, USA
[d] Department of Economics, Old Dominion University, Norfolk, VA 23529, USA

ARTICLE INFO

Article history:
Received 16 April 2013
Received in revised form 19 September 2013
Accepted 23 September 2013

Keywords:
Semi-Markov process
Stochastic process
Probability
Incident
Primary incident
Secondary incident

ABSTRACT

Incidents are notorious for their delays to road users. Secondary incidents – i.e., incidents that occur within a certain temporal and spatial distance from the first/primary incident – can further complicate clearance and add to delays. While there are numerous studies on the empirical analysis of incident data, to the best of our knowledge, an analytical model that can be used for primary and secondary incident management planning that explicitly considers both the stochastic as well as the dynamic nature of traffic does not exist. In this paper, we present such a complementary model using a semi-Markov stochastic process approach. The model allows for unprecedented generality in the modeling of stochastics during incidents on freeways. Particularly, we relax the oftentimes restrictive Poisson assumption (in the modeling of vehicle arrivals, vehicle travel times, and incidence occurrence and recovery times) and explicitly model secondary incidents. Numerical case studies are provided to illustrate the proposed model.

1. Introduction

Incidents are notorious for their delays to road users, with studies estimating that they account for up to 50% of traffic congestion on urban freeways (e.g. see Khattak et al., 2009, 2011; Zhang and Khattak, 2010a,b and the references therein). Because of the importance, numerous researchers have – through the statistical analysis of incident data – provided valuable insights into the determinants of incidents (e.g. weather conditions, vehicle factors, road geometry, travel demand, and road capacity) and their associated delays, e.g. see Al-Deek et al. (1996), Garib et al. (1997), Karlaftis et al. (1999), Zhan et al. (2008) and the references therein.

One possible classification of incidents is primary incident (the first incident that takes place) versus secondary incident (an incident that occurs within a certain temporal and spatial distance from the primary incident). The duration of the secondary incident might be contained within the primary incident or it can extend the duration of the primary incident. Furthermore, there can be multiple secondary incidents. While there are numerous empirical and simulation-based studies on incident modeling (e.g. see Jones et al., 1991; Sheu et al., 2004; Sisiopiku et al., 2007; Zhang and Khattak, 2011; Zhang et al., 2011; Khattak et al., 2012; Vlahogianni et al., 2012), to the best of our knowledge, an analytical model that can be used for incident (both primary and secondary) management planning and that explicitly considers both the stochastic as well as the

* Corresponding author at: Department of Modeling, Simulation and Visualization Engineering, Department of Civil and Environmental Engineering, 1318 Engineering and Computational Sciences Building, Old Dominion University, Norfolk, VA 23529, USA. Tel.: +1 757 683 6665; fax: +1 757 683 3200.
E-mail address: mng@odu.edu (M.W. Ng).

0191-2615/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trb.2013.09.013

dynamic nature of traffic is non-existent. This paper addresses this critical gap in the literature by presenting a stochastic process that allows for unprecedented generality in the modeling of stochastics during primary and secondary incidents on freeways. Moreover, unlike previous related studies that typically focus on the prediction of incident duration and delay, our focus is on analyzing the average time to the next primary and secondary incident. As such, the proposed model directly complements current empirical and simulation studies in our fundamental understanding of incident dynamics which, in turn, has the potential to suggest new or improved incident management strategies.

To model traffic flow, stochastic queuing theory (e.g. see Van Woensel and Vandaele, 2007) – henceforth simply referred to as queuing theory – is employed. (Validation studies of these types of traffic flow models can, for example, be found in Van Woensel and Vandaele, 2006.) One key distinctive feature of the proposed model is the relaxation/elimination of the well-known, but oftentimes restrictive, Poisson assumption. More specifically, unlike in previous work (cf. Baykal-Gursoy et al., 2009), time headways, travel times, as well as incidence occurrence and recovery times are not restricted to follow a Poisson process any longer. In addition, whereas previous work only allows for the modeling of primary incidents (e.g. see Baykal-Gursoy et al., 2009), the proposed model is able to handle secondary incidents as well, leading to a general and comprehensive incident management decision support model. It is noteworthy that secondary incidents are not only interesting from a theoretical perspective, from a practical perspective, they are perhaps even more critical to account for as they tend to divert limited resources, such as emergency medical services, towing providers and first responders from the oftentimes more severe primary incidents. This can lead to disproportionate delays and further exacerbate the unreliability of travel times (Chen et al., 2002; Ng and Waller, 2010a; Ng et al., 2011).

The remainder of this paper is organized as follows. In Section 2, a semi-Markov process is proposed to model primary incidents. Analytical results are formally derived for the expected time to the first primary incident. Section 3 presents analytical results for the case of secondary incidents, followed by a numerical case study in Section 4. Section 5 concludes the paper with a summary and conclusions.

2. A first model: time to primary incident

In this section, an initial semi-Markov process (Janssen and Manca, 2006) is presented to explicitly model the stochastic and dynamic nature of primary incidents. In Section 3, we shall extend this model to include secondary incidents. Semi-Markov processes are natural generalizations of the classical continuous-time Markov chains, in which state transitions necessarily occur after exponentially distributed sojourn times (Ross, 1995). Because of the well-known memoryless property of the exponential distribution, their use in modeling might oftentimes be inappropriate. For example, in our case, if the incident recovery time is exponentially distributed, this would imply that if an incident has not been cleared yet by time t, it would be as if the incident *just* happened at time t. Clearly, this is not the case in practice. Semi-Markov processes relax this assumption of exponentially distributed sojourn times. For a more technical discussion on semi-Markov processes, the reader is referred to Janssen and Manca (2006).

Suppose that we have a freeway segment of length L. Under normal (i.e. incident-free) conditions, following Jain and Smith (1997), this freeway segment is modeled – using Kendall's notation (Kendall, 1953) – as a state-dependent $GI/G_n/1/c$ queuing system, i.e., vehicle arrivals follow a general renewal process (indicated by GI), are processed (i.e. travel through the segment) by "one server" (i.e. the freeway) – the travel speed may vary with the congestion level (the general distribution for the travel time is indicated by G_n) – and the freeway segment can hold at most c vehicles, which is defined as the jam density multiplied by the number of lanes multiplied by L. (The interested reader is referred to Van Woensel and Vandaele (2007) for more discussion on the use of stochastic queuing models in traffic flow modeling.) When an incident occurs, the freeway reaches a state of primary incident.

To further illustrate the proposed model, consider Fig. 1. The top layer in Fig. 1 depicts the states the freeway segment can be in (i.e. the number of vehicles that occupies the freeway segment, ranging from zero vehicles to c vehicles) during normal,

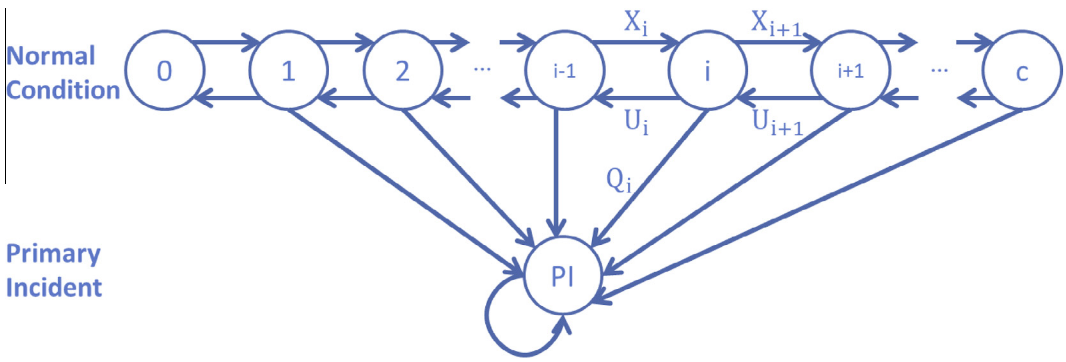

Fig. 1. State transition diagram used to investigate mean time to first primary incident.