Transportation Research Part C 71 (2016) 19–31

Contents lists available at ScienceDirect

Transportation Research Part C

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

Non-recurrent congestion analysis using data-driven spatiotemporal approach for information construction

Zhuo Chen [a], Xiaoyue Cathy Liu [a,*], Guohui Zhang [b]

[a] Department of Civil & Environmental Engineering, University of Utah, 110 Central Campus Drive, Suite 2000, Salt Lake City, UT 84112, United States
[b] Department of Civil and Environmental Engineering, University of Hawaii at Manoa, 2540 Dole Street, Holmes 383, Honolulu, HI 96822, United States

ARTICLE INFO

Article history:
Received 25 August 2015
Received in revised form 1 July 2016
Accepted 4 July 2016

Keywords:
Incident
Congestion
Information extraction
Spatiotemporal
Delay

ABSTRACT

A systematic approach to quantify Incident-Induced Delay (IID) is proposed in this study. The paper complements existing literature by developing a data-driven method to dynamically determine the spatiotemporal extent of individual incidents. The information construction process can be further used to uncover a variety of features that are associated with any specific incidents for optimal freeway management. Additionally, this study contributes two particular highlights: secondary incident identification and K-Nearest Neighbor (KNN) pattern matching. Secondary incident identification, as a pre-processing for IID estimation, disentangles the convoluted influences of subsequent incidents. The proposed method uses KNN pattern matching, an essentially heuristic search process to separate the delay solely induced by incidents from the recurrent congestion. The proposed algorithm on IID quantification was implemented on Interstate 15 in the state of Utah using data obtained from 2013. Results and implications are presented. Hot spot analysis is conducted that can be potentially used for incident mitigation and to inform investment decisions. The proposed methodology is easily transferable to any traffic operation system that has access to sensor data at a corridor level.

Published by Elsevier Ltd.

1. Introduction

Congestion, caused by recurring and non-recurring sources, is a major concern of freeway performance assessment. As the most important contributor to traffic congestion, incidents account for approximately 50–60% of delays on U.S. highways (Bertini and McGill, 2003). The downside of incidents becomes even more pronounced in metropolitan areas where severe and frequent recurrent congestion occurs on a daily basis. A variety of incident management programs have been launched in recent years to monitor and respond to incidents in an effort to effectively minimize this negative impact (Allen et al., 2013; Khattak et al., 2010; Lou et al., 2011). An accurate and efficient estimation of Incident-Induced Delay (IID) can help facilitate corridor reliability assessment. It can assist with bottleneck identification (e.g. roadway geometric design deficiencies). Benefits can also accrue when corresponding strategies are implemented to enhance safety and smooth traffic, such as ramp metering, and variable speed limit (Hegyi et al., 2005; Hellinga and Mandelzys, 2011).

IID can vary significantly in different settings, depending on ambient traffic, roadway configurations, incident severity, lane blockage, etc. An implementable algorithm to dynamically identify IID at the individual incident level for performance assessment and modeling purposes would be most ideal. This paper uses a spatiotemporal method to extract information

* Corresponding author.
E-mail address: cathy.liu@utah.edu (X.C. Liu).

http://dx.doi.org/10.1016/j.trc.2016.07.002
0968-090X/Published by Elsevier Ltd.

from roadway sensors for IID estimation. We show that the proposed method not only captures the dynamic evolution of an incident, but can also disentangle the convoluted impact of non-recurrent vs. recurrent congestions. The algorithm can be trained by the data itself, leveraging the relationship between historical recurrent data and new information incurred by the dynamic evolution of an incident.

Previous studies on IID modeling will be discussed in the following section. We will argue that IID modeling has not been thoroughly conducted at the individual incident level that can provide an accurate and efficient estimation, owing to analysis methods that either had theoretically stringent assumptions or looked at only one-dimensional changes in traffic data. Then the proposed algorithm for IID estimation and the mechanism for distinguishing different congestions are unveiled via a fuller, spatiotemporal analysis with a comprehensive data set from I-15. Implications and results are discussed at the end.

2. Background

IID is defined as the extra travel delay resulting from incidents on top of the recurrent congestion (Yu et al., 2014). Previous studies on IID were based on the mechanism of delay quantification. In terms of methodology, Deterministic Queueing Theory (DQT) and Shockwave-based algorithm are most commonly adopted (Mongeot and Lesort, 2000; Wirasinghe, 1978). For DQT, delay is determined as the difference between the curve of original traffic condition and the curve of queuing process after incidents. The model is implemented with assumed capacity reduction and empirically determined incident duration functions (Li et al., 2006). The results of DQT highly depend on assumed functions, impairing the robustness of the method. Shockwave-based algorithms are developed on the basis of macroscopic traffic flow theory, treating incidents as flow perturbations. Once an incident (perturbation) occurs, shockwaves are generated and spread backward in the traffic flow. Features of perturbation (i.e., the variation of incident-induced perturbation, clearance time, and maximum queue length) can be calculated (Mongeot and Lesort, 2000). Similar to DQT, an analytical solution is not available unless simplified traffic conditions apply. Otherwise, a numerical solution is needed (Bailie et al., 2012). Rakha and Zhang (2005) found that the two aforementioned methods yield consistent results. However, at highway bottlenecks, DQT provides more accurate estimation of incident delays. Besides the mechanism-based approach, statistical method has been applied to study the general distribution of incident delay. Skabardonis et al. (2003) estimated the average and probability distribution of delay using loop detector data and showed that non-recurrent congestion accounts for only 13–30% of total delays, depending on the extent of recurrent congestion. A statistical method provides data-based estimation of IID but fails to link IID with location-specific and incident-associated characteristics. These studies offer general insights for IID estimation. Yet neither mechanism-based methods nor statistical methods are capable of quantifying IID at the disaggregate level, where important features associated with individual incidents remain unknown.

Early research on incidents' impact at disaggregate level used static method, assuming that the maximum impact area of incidents has a fixed boundary (Moore et al., 2004). For example, Moore et al. (2004) defined the impact area as 2-mile (upstream of the incident) by 2-h (post-incident). Due to the fact that such predefined boundary may not be applicable to all incidents, researchers developed dynamic threshold methods, which define the dynamic influence area of incident on the basis of analytical/empirical approach with traffic data (Imprialou et al., 2014). Efforts to uncover the dynamic impact of incidents have led to a wide-scale use of spatiotemporal analysis (Anbaroglu et al., 2014; Chung and Recker, 2012; Chung, 2011a; Snelder et al., 2013). Spatiotemporal analysis can be used to determine the additional delay within the spatial and temporal extent under the impact of incidents. It focuses on the intrinsic variations of individual incidents, and thus avoids the bias induced by the assumed relationship between delay and surrounding conditions. There are two challenges in such analysis. The first challenge is the incident's impact range in spatiotemporal domain. Under ideal situations, the spatiotemporal extent is a contiguous region originated from the moment an incident occurs. Yet in reality, disturbances exist within the region due to traffic fluctuation. The second challenge is the identification of recurrent congestion. Method is needed to disentangle the compounded impact of non-recurrent and recurrent congestions.

Common practice is to choose an empirical threshold based on historical traffic conditions and use speed or travel time as delay indicators to distinguish the two types of congestions. Spatiotemporal region with indicator value below this threshold is considered experiencing only recurrent congestion. Chung (2011a) applied spatiotemporal concept to freeway incident delay quantification. If speed falls below the threshold $\bar{s} - \alpha\sigma_s$ (where $\bar{s}$ is the average speed, σ_s is the standard deviation of speed, and α is the scaling factor), the spatiotemporal cell is considered congested. The single average speed $\bar{s}$ is used as recurrent condition indicator, which does not consider the traffic variation under incident-free scenario. Also note that α is determined empirically. Thus, any bias may result in an over- or under-estimation of the non-recurrent congestion. Chung and Recker (2012) then improved upon this method using an optimization model to minimize the probability of errors, including the speed threshold being falsely higher than the speed at uncongested cells or falsely lower than the speed at congested cells, such that the optimum value of α can be calculated and applied to the spatiotemporal extent determination. Note that both empirical and optimization methods are built on the assumption that recurrent congestion can be estimated analytically by $\bar{s} - \alpha\sigma_s$. Snelder et al. (2013) expanded the spatiotemporal method from corridor to freeway network. They assumed that the boundary of spatiotemporal extent is a parallelogram with a slope of 70 km/h (shockwave speed). To determine the recurrent congestion, they constructed a Vehicle Loss Hour (VLH) series with weighted VLH from weeks before and after an incident, and used median as the referencing case. This empirical method is simple to implement but lacks validation. Anbaroglu et al. (2014) applied the spatiotemporal clustering analysis to freeway network using links as unit. The threshold for recurrent congestion was also determined via optimization.