QUANTIFYING NON-RECURRING CONGESTION IMPACT ON SECONDARY INCIDENTS USING PROBE VEHICLE DATA

Hyoshin Park, University of Maryland
Ali Haghani, University of Maryland
Masoud Hamedi, University of Maryland

ABSTRACT

As a significant cost and externality to economic efficiency, congestion is partly caused by traffic incidents. For more systematic, planned, and coordinated incident management, quantifying a primary incident’s impact on secondary incidents is crucial and challenging. Many thresholds have been suggested in defining the secondary incidents, but there is no universal acceptance of a definition and corresponding set of measurement parameters. Static threshold methods cannot consider the actual representation of prevailing traffic conditions when the incidents took place. On the other hand, dynamic methods have disadvantages because necessary traffic detector data may not be available, and replication of the incidents using a simulation package can be time consuming. The novelty of this study rests in the attempt of a probe vehicle technique for capturing the dynamics of traffic evolution during the primary-crash incidents. Compared to the previous thresholds which have many errors, proposed speed contour map from Traffic Message Channel codes provides accurate feasible area for identification of secondary incidents.

INTRODUCTION

Quantification of Incident’s Impact

Congestion is a significant cost and externality to economic efficiency. Viable studies estimate that total U.S. congestion costs range from $14 billion to $200 billion annually. Fifty-five percent of total delay experienced by motorists is caused by roadway incidents in urban area population groups (TTI, 2011); and secondary crashes are estimated to cause 18% of all fatalities on freeways (TIM, 2010). Every minute that an incident remains partly cleared during peak congestion causes many more minutes of travel delay. Mitigating non-recurring congestion will return a strengthening economy.

Rubbernecking, physical impedance in the travel lanes, and other incident-related obstructions reduce capacity and impede flow. The resulting speed reduction and queue formation foment additional incidents, referred to as secondary incidents. Incident characteristics, incident duration, traffic conditions, and secondary incident occurrence are related to one another in a close way (Figure 1). Indeed, the longer an incident scene is in place caused by its characteristics, the greater the likelihood for secondary incidents. Moreover, the total time for an incident to be cleared can be increased by the occurrence of secondary incidents, and the travelers may experience ever-increasing congestion.

Figure 1. Relationship between Secondary Incident Occurrences and Contributing Factors
Incident management with systematic, planned, and coordinated resources provide tools to enhance performance of response team to reduce the duration and impact of incidents and improve the safety of motorists. To maximize the effectiveness of these operational strategies, accurate and dynamically predicted incident duration, length of queue, and chances of secondary incident occurrence can aid to real time information for travelers (e.g., variable message signs update, cell phone text messages alert). Traffic operators can assess the need to implement detour operations and any other control strategies to mitigate congestion. Furthermore, the benefits of the incident management program are identified by measuring the reduction of incident’s negative impacts. Transportation safety, however, does not seem ready to embrace this critical issue. Given the wide variety of causes and impacts of non-recurring congestion, it is especially difficult to quantify discrete, random, and complex incident nature in a system level. In addition, poor quality of incident data encumbers the accurate identification, generalization and prediction of the impact of incidents.

**Identification of Secondary Incidents Using Advanced Traffic Data**

We cannot exaggerate the importance of using a reasonable threshold in defining the secondary incidents, because it has an influence on the accuracy of the result and validity of the proposed theoretical frame. The most frequent method used to identify the secondary incidents is the static method (i.e., proximity in time and space). However, this condition is not sufficient, as it does not consider the actual representation of prevailing traffic conditions when the incident took place. In addition, this method is more likely to produce biased results as it depends on the values of chosen thresholds. On the contrary, several dynamic methods were proposed based on simulation or inductive loop detector data. However, using a deterministic queuing method in real-time is inappropriate because exact identification of traffic arrival rate and capacity reduction is difficult, and quality of raw data from inductive loop detector maybe unsatisfactory.

In recent years, the vehicle probe industry is emerging as a viable means to monitor network-wide traffic flow, delivering both speed and travel time information. This is a new opportunity to use real-time estimation of queue length to identify feasible area for secondary incident. An adjusted boxplot is used to separate the non-recurrent congestion from recurrent congestion that is present on the road at the time and place of an identified accident. With the speed contour output from probe vehicle data, the analysis of the secondary incidents becomes a geometrical exercise of determining of the extent of the feasible region. The secondary incidents and primary incidents are examined using the time-space evolution of disturbance boundaries while considering only the impact of isolated incidents.

**SECONDARY INCIDENT FILTERING**

**Impact Area for Secondary Incidents**

The identification of an incident as secondary to a primary incident is described in Figure 2. First, from the reported incident characteristics and traffic conditions, incident duration can be predicted. Upstream of the incident, a queue has formed from the time the primary incident “P” occurred until the service rate exceeds the queue arrival rate. This entire timeline is defined as temporal impact area, and spatial impact area changes over time according to the evolution of traffic flow. Then, incident “S” within this impact area is classified as a secondary incident. Queue dissipation covers the time the incident is cleared to the time that normal flow returns.

---

**Figure 2. Incident Timeline and Feasible Area**
Static and Dynamic Feasible Area

Several methods are available in the literature to indicate the existence of a feasible area for secondary incidents occurrence, most of which focused on temporal and spatial thresholds related to primary incidents. For example, Raud (1997) defined any incident within static thresholds of 1 mile and 15 minutes as a secondary incident. Similar static feasible areas were also discussed by Karlaftis et al., 1999; Moore et al., 2004; Hirunyanitiwattana and Mattingly, 2006; Zhan et al., 2008, Khattak et al., 2009. Most studies used Annual Average Daily Traffic as aggregated traffic related information.

On the other hand, Haghani et al. (2006) initiated the study for dynamic thresholds, which is identified from the shockwave that arises as a consequence of the incident in the simulation model. A set of preselected time intervals are employed in seeking the impact area during a specific time interval for each incident. Mathematically represented impact area requires smaller time interval for greater accuracy of the estimation method. Chou and Miller-Hooks (2010) also considered the dynamics associated with traffic state by using simulation-based secondary incident filtering method (SBSIF). Regression implementation of SBSIF has a significantly reduced misclassification rate as compared with static methods. Replication of the incidents using a simulation package can be time consuming.

Sun and Chilukuri (2010) presented a methodology in order to improve upon the existing method of static thresholds by formulating dynamic boundaries. An incident progression curve was used to indicate that dynamic and static methodology can differ by more than 30%. However, the third order polynomial equation were identically applied to all incidents, regardless of traffic volume. In addition, archived incident data might have limited queuing information.

Zhan et al. (2009) and Vlahogianni et al. (2010) defined the spatio-temporal boundary for each secondary crash based on maximum queue length and the queue duration induced by the crash. The similar approach has been studied by using deterministic queue model to estimate associated delays (Zhang and Khattak, 2010, 2011; Khattak et al., 2012). If a spillback is observed, the queue length is estimated by a deterministic D/D/1 model with estimated arrival and departure flow rates based on Highway Capacity Manual methodologies. However, it is inappropriate to use a deterministic queuing approach for real-time application in non-recurring congestion, since it assumes exact identification of arrival rate and capacity reduction (Fu and Rilett, 1997).

Vlahogianni et al. (2012) improved secondary incident detection methodology by capturing the propagation of wide moving jam generated at the upstream and the downstream to calculate the changing queue length. As criticized by Schönhof and Helbing (2009), a congestion caused by accident might not classify the pronounced stop-and-go waves as “wide moving jams.” Hence, secondary incident with the characteristic spatial structure of a “general pattern” (“synchronized flow”, “pinch region” with jam formation, and a region of “wide moving jams”) cannot be perfectly explained by their models. Table 1 summarizes existing secondary incident identification method.

Table 1. Existing Secondary Incident Identification Method

<table>
<thead>
<tr>
<th>Static method</th>
<th>Dynamic method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raud (1997) 1 mi, 15 min</td>
<td>Haghani et al. (2006) Simulation-based method employing shockwaves</td>
</tr>
<tr>
<td>Moore et al. (2004) 2 mi, 2 hr</td>
<td>Sun and Chilukuri (2010) Incident progression curve</td>
</tr>
<tr>
<td>Hirunyanitiwattana and Mattingly (2006) 2 mi, 60 min</td>
<td>Vlahogianni et al. (2010) Observed maximum queue length and the queue duration</td>
</tr>
<tr>
<td>Khattak et al. (2009) 1 mi, Clearance time + 15 min</td>
<td>Vlahogianni et al. (2012) ASDA model automatically tracks the propagation of moving traffic jams</td>
</tr>
</tbody>
</table>
METHODOLOGY

Impact of Non-Recurrent Congestion

As discussed, deterministic queue model approach may have over- or under- estimated the actual delay of primary-secondary incident events. This approach assumes that precise identification of capacity reductions can be available. However, accurate measurement of incident capacity reduction is not straightforward. Still, many researchers primary rely on simulation analysis (e.g., CORSIM, VISSIM) that have the ability to replicate rubbernecking by proportionally increasing the distances at which the vehicles are following one another (Haghani et al., 2006). Deterministic queuing model may be adequate for after-incident evaluation, for which information on the traffic volume and incident situation is readily available. Moreover, unsatisfactory quality of raw data from inductive loop detector can decrease accuracy of incident detection rate.

On the contrary, methodology of this study is based on real-time traffic conditions for one-minute interval from each TMC segment. Here, we need to define whether each TMC segment is under the non-recurrent congestion or not. However, separation of the non-recurrent congestion from recurrent congestion is a hard task because there is no correct answer. It is more challenging than discriminating congestion from non-congestion.

![Figure 3. A Tool for Real Time Estimation of Queue](image)

As shown in Figure 3, queue estimation tool is currently used in real time at the CATT Lab. Traffic conditions can be determined by comparing the current reported speed to the reference speed for each segment of road. Reference speed values are calculated based upon the 85th-percentile of the observed speeds on that segment for all time periods, which establishes a reliable proxy for the speed of traffic at free-flow for that segment. TMC segments represent bottleneck or congestion in which the actual travel speed drops below 60% of the reference speed longer than 5 minutes. Once the travel speed returns to a value greater than 60% of the reference speed for 10 minutes, TMC segment is no more in congestion.

Another tool was discussed by Chung and Recker (2012). Distance of a data point from the mean in standard deviation units was used to separate any particular crash speed from the distribution of crash-free speeds. They applied constant threshold empirically found by using the relationship between randomly selected accident samples and their associated traffic data. The determinant threshold value discriminates two regions: congested regions affected by a crash and uncongested regions.

However each accident data may have different thresholds and following difficulties in successfully representing the nature of accidents, compared to non-accident cases. We need a robust method to consider accident impact with the measurement of each distribution. This paper proposes a method that can be applied to all distributions, even without finite moments.

An Adjusted Boxplot Method

The boxplot is one of the most frequently used graphical tools for visualizing the distribution of continuous data (Tukey, 1977). It can be constructed by putting a line at the height of the sample median $Q_2$, drawing a box from the first quartile $Q_1$ to the third quartile $Q_3$. The length of this box equals the inter-quartile range, $IQR = Q_3 - Q_1$, as a robust measure of the scale. All points outside the interval in Equation 1 can be classified as potential crash cases.
\[ Q_1 \pm 1.5 \text{IQR}; \quad Q_3 \pm 1.5 \text{IQR} \]  

However, observations outside the fence are not necessarily real crash-cases that behave differently from the majority of the data. At thick tailed symmetric distributions, many regular observations will exceed the outlier cutoff values defined in Equation 1, whereas data from thin tailed distributions will hardly exceed the fence (Hoaglin et al. 1983). We use the medcouple (MC) to measure the skewness of a univariate sample from a continuous distribution \( F \),

\[ MC = \text{med}_{x_i \leq q_2 = x_j} \hat{h}(x_i, x_j) \]  

for all \( x_i \neq x_j \), kernel function \( h \) is defined as

\[ \hat{h}(x_i, x_j) = \frac{(x_j - q_2) - (q_2 - x_i)}{x_j - x_i} \]  

Medcouple always lies between −1 and 1. A distribution that is skewed to the right has a positive medcouple, whereas the MC becomes negative at a left skewed distribution. As shown in Hubert and Vandervieren, 2008, exponential model in the definition of our adjusted boxplot to define the boundaries of the interval.

\[ [Q_1 - \hat{h}_l(MC) \text{ IQR}; \quad Q_3 + \hat{h}_u(MC) \text{ IQR}] \]  

Additionally, we require that \( \hat{h}_l(0) = \hat{h}_u(0) = 1.5 \) in order to obtain the standard boxplot at symmetric distributions. Note that by using different functions \( \hat{h}_l \) and \( \hat{h}_u \) in Equation 5, we allow the fence to be asymmetric around the box, so that adjustment for skewness is indeed possible.

\[ \hat{h}_l(MC) = 1.5 e^{aMC}, \quad \hat{h}_u(MC) = 1.5 e^{bMC} \]  

As studied by Chung and Recker (2012), we can define speed at section \( i \) at time \( t_n \); \( S_i(t_n) \), and consider if \( S_i(t_n) \leq Q_1 - \hat{h}_l(MC) \text{ IQR} \): under crash impact area; \( S_i(t_n) > Q_1 - \hat{h}_l(MC) \text{ IQR} \): crash-free area. A continuous region affected by crashes can be described and used for identifying secondary crashes.

**CASE STUDY**

**Vehicle Probe Technology Based Data**

In previous studies, loop detectors were used to collect flow rate or speed data. It is important to point out that the results of these methods depend almost entirely on loop detectors performance. However, traffic data from loop detectors often contain invalid data (e.g., missing value, negative value, and non-zero speed with zero count) due to malfunctioning detectors, communication failures, and other reasons. Discarding values outside of the expected ranges potentially can lead to over-fitting, which diminishes the effectiveness of the model (Washington et al., 2003).

Estimating traffic state from loop detectors has another challenge for accurately representing the traffic conditions on other parts of the road link. The travel speed from loop detector can be estimated using conventional method such as g-factor method. However the g-value, related to the average effective vehicle length, may not be constant (Hall and Persaud, 1989; Pushkar et al., 1994). The percentage of long vehicles may vary significantly during different time periods and introduce huge errors in some situations (Lao et al., 2012).

Moreover, high installation and maintenance cost of fixed-point sensors drives transportation management authorities to consider both outsourcing traffic monitoring and developing new method of detection.
Recently, vehicle probe technology is increasingly becoming more attractive for real-time system state estimation, and it is a common practice for data providers to report data on Traffic Message Channel (TMC) codes. INRIX reports average speed as a normalized measure of travel time on each TMC segment, along with a score in one-minute interval. Hamedi and Haghani (2012) concluded that the INRIX travel time and speed data on the freeway segments generally satisfies the requirements of applications for real-time travel time display.

The novelty of this study rests in the attempt of explaining the dynamics of traffic evolution during the primary-crash incident using vehicle probe technology, which covers the I-695 corridor from MD-150/Eastern Blvd/Exit 38 end to the MD-151/North Point Blvd/Exit 40. INRIX provides data on all of these segments. Data from the Vehicle Probe Project comes primarily from the vehicles operating as anonymous probes. Meaningful travel time information for each TMC segment is achieved after data processing: aggregation, filtering, and smoothing. Then INRIX assigns a quality indicator to each travel time record based on real-time GPS tracks (score 30), archival data (score 10) or a combination of both (score 20); only qualities satisfying a score more than 20 are used for this study. Table 2 shows the list of TMC segments that are covered in I-695 corridor including beginning and endpoint as well as length of each TMC segment.

![Figure 4. Collision Incidents on I-695](image-url)
The incident data along this I-695 corridor are investigated. In total, 5524 incidents (e.g., disabled vehicle, weather event, road maintenance, collision incidents, vehicle on fire, debris) from May 2011 to October 2011 are collected. 614 collisions (e.g., fatality, personal injury, and property damage) and vehicles on fire incidents are regarded as candidates for primary-secondary incident pair. Based on incident location, traffic data from TMC codes are used to present traffic state of each segment. The archived incident and probe vehicle database are provided by Center for Advanced Transportation Technology Laboratory (CATT Lab) at the University of Maryland. Figure 4 illustrates the number of collision incidents for each road sections on I-695.

### Accurate Impact Area for Secondary Incidents

When the speeds of vehicles return to normal after an incident, the queue has dissipated. The length of queue could be tracked based on the vehicles or interpreted from traffic speed contour plots. Based on configuration of crash impact for each segment, one can develop a daily speed contour map for the entire section of I-695 corridor (September 1, 2011, Thursday), as shown in Figure 5. For better representation, each cell was aggregated to one hour averaged traffic speed for TMC.

<table>
<thead>
<tr>
<th>TMC</th>
<th>Start Lat</th>
<th>Start Long</th>
<th>End Lat</th>
<th>End Long</th>
<th>Length(mi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110P04555</td>
<td>39.2063</td>
<td>-76.5913</td>
<td>39.2066</td>
<td>-76.6119</td>
<td>1.11</td>
</tr>
<tr>
<td>110P04520</td>
<td>39.2968</td>
<td>-76.7426</td>
<td>39.3130</td>
<td>-76.7445</td>
<td>1.12</td>
</tr>
<tr>
<td>110-04523</td>
<td>39.3820</td>
<td>-76.7376</td>
<td>39.3780</td>
<td>-76.7440</td>
<td>0.45</td>
</tr>
<tr>
<td>110-04520</td>
<td>39.3122</td>
<td>-76.7447</td>
<td>39.3118</td>
<td>-76.7447</td>
<td>0.03</td>
</tr>
<tr>
<td>110-04519</td>
<td>39.2994</td>
<td>-76.7432</td>
<td>39.2894</td>
<td>-76.7414</td>
<td>0.70</td>
</tr>
<tr>
<td>110+04527</td>
<td>39.3922</td>
<td>-76.7071</td>
<td>39.3959</td>
<td>-76.6877</td>
<td>1.08</td>
</tr>
<tr>
<td>110N04535</td>
<td>39.4017</td>
<td>-76.5629</td>
<td>39.4019</td>
<td>-76.5683</td>
<td>0.30</td>
</tr>
<tr>
<td>110P04512</td>
<td>39.2363</td>
<td>-76.6677</td>
<td>39.2391</td>
<td>-76.6685</td>
<td>0.20</td>
</tr>
<tr>
<td>110+04542</td>
<td>39.3346</td>
<td>-76.4904</td>
<td>39.3344</td>
<td>-76.4902</td>
<td>0.02</td>
</tr>
<tr>
<td>110P04514</td>
<td>39.2460</td>
<td>-76.6749</td>
<td>39.2561</td>
<td>-76.6914</td>
<td>1.15</td>
</tr>
<tr>
<td>110P04532</td>
<td>39.4131</td>
<td>-76.6040</td>
<td>39.4136</td>
<td>-76.5958</td>
<td>0.44</td>
</tr>
<tr>
<td>110P04549</td>
<td>39.2336</td>
<td>-76.5043</td>
<td>39.2320</td>
<td>-76.5071</td>
<td>0.18</td>
</tr>
<tr>
<td>110-04558</td>
<td>39.2044</td>
<td>-76.6392</td>
<td>39.2025</td>
<td>-76.6347</td>
<td>0.28</td>
</tr>
</tbody>
</table>

The incident data along this I-695 corridor are investigated. In total, 5524 incidents (e.g., disabled vehicle, weather event, road maintenance, collision incidents, vehicle on fire, debris) from May 2011 to October 2011 are collected. 614 collisions (e.g., fatality, personal injury, and property damage) and vehicles on fire incidents are regarded as candidates for primary-secondary incident pair. Based on incident location, traffic data from TMC codes are used to present traffic state of each segment. The archived incident and probe vehicle database are provided by Center for Advanced Transportation Technology Laboratory (CATT Lab) at the University of Maryland. Figure 4 illustrates the number of collision incidents for each road sections on I-695.
segments. Shaded cell represent temporal-spatial area under non-recurrent congestion. Some segments of freeway corridor experience traffic congestion during morning and afternoon peak hours, and also due to incidents.

Vehicles upstream of an incident event, near the I-695 outer loop at I-795 / Exit 19 (TMC 110N4523) 1:25 PM, should be in a slow-moving queue. It is because of following vehicles suffer from congestion with traffic condition rapidly decelerating from normal driving speed or free-flow speed to stop-and-go traffic. When they pass the incident, they should speed up to normal driving speed or even free-flow speed. There are two additional incidents in southbound direction, and one incident in the opposite direction (northbound) within this temporal-spatial congestion area, and they have possibilities to be regarded as secondary incidents.

As shown above, speed information from probe vehicle data can capture the impact of an incident based on incident’s characteristics and prevailing traffic conditions. Figure 6 shows micro-detail feasible area expressed by queue length from TMS segment information and dynamic changes of congestion pattern. To facilitate illustration of secondary incident phenomena, 1-minute interval of speed contour plots from onset of incident to queue dissipation are investigated. To clear the primary incident, in which a truck is on its side carrying diesel fuel resulting in a two-vehicle collision with injury, two lanes are blocked for whole clearance and 13 operation units are dispatched (including police, fire, emergency medical personnel, Coordinated Highways Action Response Team units, tow services, public affairs, and flatbed). Incident’s characteristics indicate different rubbernecking phenomenon which perpetuates in the impact area with different intensities depending on the cross and longitudinal location with respect to the incident. Especially, in this larger scale event, multiple secondary incidents have higher likelihood to occur and their clearance takes longer time.

Speed reduction from the primary incident (incident “1” ID: 25001beda3bf015f) may have an impact on the possibility of secondary incident occurrence (incident “2” ID: d3ffbed1aaf015f) at I-695 Perring parkway, 3:41 PM. This may make the period of congestion even longer and cause an additional secondary incident event (incident “3” ID: 510018b2b8130160) at I-695 Loch Raven blvd., 7:22 PM. Moreover, incident influences not only on the incident direction but also on the opposite direction. Incident 4 (event ID: 2000daabb250160) can be defined as a secondary incident in the opposite direction. Then, when drivers pass the incident, they should speed up to normal driving speed. Traffic flow conditions will return to normal after 7:37PM. Therefore, incident “5” (event ID: 6f00f848a5cf015f) is not identified as a secondary incident.
Figure 5. Traffic Speed Contour for September 1, 2011 (Southbound)
Compared to static threshold methods in previous studies, the probe-based filtering method has superiority. Including the methods proposed by Raud (1997), Hirunyanitiwattana and Mattingly (2006) and Zhan et al. (2008), only the proposed method can capture incidents “2” and “3” as secondary incidents. Moreover, incident “4” in the opposite direction only can be identified method proposed by this study and Moore et al. (2004). In this way, among 614 collision incidents on I-695, 112 (18%) are classified as primary-secondary pair based on archived incident and probe vehicle data.

CONCLUSIONS AND RECOMMENDATIONS

Modern data collection technologies enabled us to look into critical factors for incident duration and establish a deliberate incident management plan. In this study, probe vehicle technique is used to capture the dynamics of traffic evolution during the primary-crash incident. Compared to static threshold method in previous studies, probe-based filtering method has better result in identifying secondary incidents. Once there is a universal acceptance of a definition and corresponding set of parameters of secondary incidents, jurisdictions can be applied to compare incident reports. Data used in previous studies are not reasonable because necessary traffic detector data may not be available and replication of the incidents using a simulation package can be time consuming. The novelty of this study rests in the attempt to use INRIX speed data on the identification of secondary incidents of freeway segments, which generally satisfies the requirements of applications for real-time travel time display. However, proposed methodology cannot be applied to freeway segments which probe vehicle data are not available. Recently, vehicle probe technology is increasingly becoming more attractive for real-time system state estimation, and it is a common practice for data-providers to report data on TMC codes. Especially in State of Maryland, freeway corridors (e.g., I-495, I-95, I-295, I-83, and I-270) are providing probe vehicle data. Including more incident data on improved road segment as well as more time period will help developing robust model. Accurate and understandable information provided by the tool may help emergency operator to make a better decision, and maximize the effectiveness of incident management.

REFERENCES


