Investigating the Characteristics of Secondary Crashes on Freeways

Hong Yang, Ph.D. (Corresponding Author)
Post-Doctoral Associate
Rutgers Intelligent Transportation Systems (RITS) Laboratory
Department of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
623 Bowser Road, Piscataway, NJ 08854, USA
Phone: (732) 445-0579 x127
Fax: (732) 445-0577
Email: yanghong@rci.rutgers.edu

Bekir Bartin, Ph.D.
Research Associate
Rutgers Intelligent Transportation Systems (RITS) Laboratory
Department of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
623 Bowser Road, Piscataway, NJ 08854, USA
Phone: (732) 445-5496
Fax: (732) 445-0577
Email: bbartin@rci.rutgers.edu

Kaan Ozbay, Ph.D.
Professor & Director
Rutgers Intelligent Transportation Systems (RITS) Laboratory
Department of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
623 Bowser Road, Piscataway, NJ 08854, USA
Phone: (732) 445-2792
Fax: (732) 445-0577
Email: kaan@rci.rutgers.edu

Word count: 5232 Texts + 2 Tables + 7 Figures = 7482
Abstract: 176
Initial Submission Date: August 1, 2012
Resubmission Date: November 15, 2012

TRB Paper ID: 13-4866
Transportation Research Board’s 92nd Annual Meeting, Washington, D.C., 2013
ABSTRACT
Prevention of secondary crashes is one of the priorities in traffic incident management. However, limited information on secondary crashes has largely impeded the selection of appropriate countermeasures. The primary goal of this paper is to improve the understanding of secondary crashes, which is achieved by two major steps. First, an analysis framework is developed to accurately identify secondary crashes by integrating rich traffic sensor data with the statewide crash data sets. Second, the characteristics of the identified secondary crashes are investigated in detail. Secondary crashes that occurred on a 27-mile section of a major highway in New Jersey were mined using the proposed analysis framework. A thorough examination of their characteristics was then performed. Empirical findings on the frequency of secondary crashes, their spatio-temporal distributions, clearance time, crash type, severity, and major contributing factors were highlighted. These preliminary results can help transportation agencies make more informed decisions on mitigating secondary crashes and improve their incident management operations. To complement the results, further in-depth investigations based on more high-resolution sensor data and high-quality incident records are suggested.
INTRODUCTION

Traffic incidents such as disabled vehicles, vehicle fires, cargo spills and motor vehicle crashes frequently disrupt traffic flow. These nonrecurring events account for more than half of all traffic delays in urban areas and almost for all delays in rural areas (1). More importantly, they also expose other vehicles to the risk of sustaining a secondary crash (2). For instance, a study by Tedesco et al. (3) found that the crash risk is more than six times higher in the presence of an earlier crash. Similarly, Karlaftis et al. (4) found that the likelihood of secondary crashes can increase by about 2.8 percent with the increase of an additional minute to clear an initial incident. United States Department of Transportation (USDOT) estimated that secondary crashes alone were responsible for about 18 percent of all fatalities on freeways and about 20 percent of all crashes (5,6). In addition, secondary crashes impose additional delays to motorists. Secondary crash prevention thus was highlighted as one of the most priority issues in traffic incident management (5).

To prevent secondary crashes, it is crucial to examine their characteristics and provide necessary countermeasures accordingly. Previous research has made effort to investigate these crashes but have had limited success because of the lack of high-quality data and the complexity of identifying such crashes (7,8,9). Now traffic sensors such as inductive loop and microwave radar detectors are being used extensively for traffic data collection. Large sets of high-quality and high-time resolution traffic data that are available for many roadways can be used to in conjunction with crash databases to identify secondary crashes.

The objective of this paper is two-fold. First, it develops a framework for identifying secondary crashes by integrating rich traffic data with the available crash database. Second, it focuses on examining the characteristics of identified secondary crashes such as their frequency, spatio-temporal distributions, clearance time, crash type, severity, and major contributing factors. Detailed examination of their characteristics can help decision makers select the most appropriate countermeasures to mitigate secondary crashes.

This paper is organized as follows. The next section provides an extensive review of the past studies on secondary crashes. This is followed by the description of the proposed framework for identifying secondary crashes. The proposed methodology is then applied to identify secondary crashes along a 27-mile section of New Jersey Turnpike as a case study. The analysis results and a discussion on the characteristics of the identified secondary crashes are then presented. The major findings of our analysis and their implications for future studies are summarized in the final section.

LITERATURE REVIEW

Among several studies that attempted to identify secondary crashes, two general types of methodologies were used: (a) static approaches based on fixed spatio-temporal boundaries and (b) approaches based on queuing analysis. In the static approach, crashes that occurred within the pre-defined impact range of primary incidents were classified as secondary crashes. The first and foremost example of this type of method was by Raub (8,10). He determined secondary crashes by using clearance time plus 15 minutes and a distance of one mile upstream as the identification criteria. A substantial number of studies (e.g.,4,9,11,12) further tailored these criteria to classify secondary crashes on other roadways and found that the proportion of secondary crashes varied broadly [from 2.01 percent (12) up to 35 percent (4)]. Several other studies (e.g.,7,13,14,15) also considered additional spatio-temporal criteria to capture the secondary crashes caused by the rubbernecking effect. Despite these efforts, the static approaches used in all the aforementioned studies were not adequate to identify secondary crashes. This is mainly due to the subjective selection of fixed spatio-temporal thresholds that leads to misclassification of secondary crashes.

To address the flaw of static approaches, several studies developed queuing based approaches to identify secondary crashes. These approaches either developed queuing models (16,17,18,19) or queuing progression curves (20,21,22) to capture the impact range of prior incidents and detected secondary crashes accordingly. These approaches attempted to construe the prevailing traffic conditions using the estimated queue lengths when a prior incident occurred. However, there are still limitations to these
approaches. For instance, the simplified queuing models and progression curves cannot fully account for the impact of primary incidents (23). In addition, part of the challenge stems from the lack of accurate information to model or identify the progression of the physical queues (24,25,26). Therefore, without the knowledge of detailed traffic data and incident characteristics, these approaches cannot fully capture potential secondary crashes.

After identifying secondary crashes, several studies examined the factors that influenced their occurrence and examined some possible countermeasures. For instance, Karlaftis et al. (4) and Latoski et al. (27) used a logistic regression to model the likelihood of secondary crashes given the characteristics of the primary incident. A similar modeling approach was also used by Zhan et al. (9,16), Kopitch and Saphores (14), and Khattak et al (28). These models were used to develop traffic incidents management tool (28), and evaluate countermeasures such as improved highway service patrol program (4,27) and changeable message signs (14) to reduce secondary crashes.

Khattak et al (12) also investigated the interdependence of incident durations and secondary incidents using binary Probit regression models. Despite the differences of the analyzed variables in these studies, the primary incident characteristics such as incident duration, number of lanes blocked, number of vehicles involved, time, and vehicle type were frequently found to influence the occurrence of secondary crashes.

Zhan et al. (9,16) also used logistic regression identify the factors that affect the severity of secondary crashes. They found that lane blockage duration and visibility significantly affect secondary crash severity. Zhang and Khattak (17) used ordered logit models to explore the probability of multiple secondary crashes. They found that multiple-vehicle involvement and lane blockage had a different impact on the occurrence of primary-secondary crash pairs and primary-multiple-secondary crash pairs.

Khattak et al. (19) further predicted the frequency of secondary incidents from a macroscopic level. Poisson, zero-inflated Poisson, and negative binomial regression models were estimated. They found that factors including roadway length, traffic volume, number of on-ramps, curve level, number of lanes, congestion level, truck volumes, and roadway location affected the frequency of secondary incidents.

Several other studies examined the characteristics of secondary incidents specifically. Zhan et al. (9) analyzed secondary crashes in Florida by freeway corridors, different time periods, lane blockage and incident types. Kopitch and Saphores (14) investigated the similar characteristics of secondary crashes on a section of Interstate 5 in California. Hirunyanitiwattana and Mattingly (11) also compared the characteristics of secondary crashes with primary crashes in the California highway system. They found that urban districts have a higher proportion of secondary crashes. They also found that typical secondary crashes are rear-end, property damage only crashes that usually occur in peak periods on urban freeways with four or more lanes. Zhang and Khattak (18,23) used ordinary least squares regression models to investigate time and distance gap of secondary incidents in relation to primary incidents. They found that the time and distance of secondary incidents varied systematically with the characteristics of primary incidents.

Accurate identification of secondary crashes is a challenge. Existing approaches, whether static or queue-based, fall short of accomplishing this task. Consequently, any analysis of crash characteristics or modeling contributing factors based on the misclassified secondary crashes will lead to biased findings. These unreliable findings can lead to unreliable decisions on selecting appropriate incident management countermeasures. Therefore, a thorough examination of secondary crashes identified by reliable identification approaches is needed.

ANALYSIS FRAMEWORK
Two major tasks formed the analysis framework: (a) Identifying secondary crashes and (b) Examining the characteristics of the identified secondary crashes. Instead of using the existing static or queuing based approaches, the first task involves the development of an enhanced identification approach using the available traffic sensor data. The second task then statistically analyzes the specific features of the
Identifying Secondary Crashes

Secondary crashes are incidents that occur within the impact range of a primary incident. The impact range defines the temporal and spatial influence triggered by a primary incident. Existing static methods simply used fixed spatio-temporal thresholds to define the impact range. Others used simplified queuing models to capture the impact range. Regardless, these methods cannot adequately capture the impact range of the primary incident as discussed in the literature review section. Instead, we developed a new methodology to accurately account for the dynamic characteristics of the spatio-temporal impact of primary incidents under the prevailing traffic conditions. The proposed method focused on identifying the impact range of a primary incident using archived sensor data and detecting secondary crashes within the impact range. The detailed description of the methodology can be found in (29). The general procedure is summarized as follows.

**Step 1: Developing speed contour plots (SCP) over time and space:** Let us use FIGURE 1 (a) as an example. Each cell in the figure represents a speed measurement $V(t,s)$ from the loop detector $s$ along the freeway at the $i^*$ time interval, $\forall s = 1,2,...,S$ and $\forall t=1,2,...,T$. FIGURE 1 (a) represents a slice of a typical SCP on a freeway. A clear queue formation is observed soon after crash A.

**Step 2: Developing a representative speed contour plot (RSCP):** The RSCP generally represents daily normal traffic conditions on a freeway when no incident occurred. Such RSCP can be built based upon representative speed measurements. The $p^{th}$ percentile speed $V_{p}(t,s)$ of the historical speed measurements on $i^*$ day of week is used as the representative speed measured at detector $s$ and time period $t$, where $i=1,2,...,7$ analogous to the day of the week from Monday to Sunday. FIGURE 1 (b) shows an example of RSCP that represents a normal Wednesday traffic condition on a freeway based on data collected every 5 minutes during all other Wednesdays in 2011.

**Step 3: Constructing a binary speed contour plot (BSCP):** Compare $V(t,s)$ of SCP in step 1 with the corresponding $V_{p}(t,s)$ in RSCP generated in step 2. If $V(t,s) < \omega \times V_{p}(t,s)$, the speed measurement $V(t,s)$ in the original SCP is converted into $\hat{V}(t,s)=1$. Otherwise, it is denoted as $\hat{V}(t,s)=0$. This constraint defines an abnormal traffic condition if the measured speed is below the threshold $\omega \times V_{p}(t,s)$, where $\omega$ is a user defined weighting factor between 0 and 1. It assumed that a $(1-\omega) \times 100$ percent reduction in the representative (normal) speed indicates the occurrence of congestion. A small weight factor suggests an aggressive threshold to define congestion whereas a large one implies a conservative threshold. A reasonable weight factor should be determined based on highway operators’ classification of congested and non-congested conditions. To reflect the consistency of speed measurements in short time periods, the speed measurement at the $i^*$ time period is changed to $\hat{V}(t,s)=1$ if $|\hat{V}(t-1,s)-\hat{V}(t,s)|=1$ and $\hat{V}(t+1,s)=1$.

After conversion, the original SCP will be represented by a BSCP. The following FIGURE 1 (c) shows an example of converting the original SCP into a BSCP based on the RSCP of 50th percentile historical speed. In BSCP, a red cell indicates that $\hat{V}(t,s)=1$ and a green cell means $\hat{V}(t,s)=0$. Each cluster of red cells visualizes the congested area.

**Step 4: Detecting secondary crashes in the impact range in BSCP:** Let us use FIGURE 1 (d) as an example. We need to detect whether crash B and crash C are secondary crashes in relation to the first crash A. The line (e.g., line AB and AC) between each pair of potential primary-secondary crashes is estimated based on their coordinates. If the line is enclosed by the impact range, then it is suggested that this pair of crashes are primary and secondary ones. Otherwise, they are independent crashes. Readers may refer to (29) for the detailed description of the algorithm for verifying whether the line is enclosed by the impact range. The essential idea of the algorithm is to check the coordinates of some critical points on the line in relation to the impact range. If the binary speed measurement for cells that contain all the sampled points are 1, then it suggests the line is enclosed by the impact range and the corresponding
crashes are primary-secondary crashes (e.g., Crash A and B). Otherwise, they are independent crashes (e.g., crash A and C).

FIGURE 1 Framework of identifying secondary crashes using sensor data.

Step 1: Develop SCP of a given day

Step 2: Develop RSCP based on historical sensor data

Step 3: Convert SCP to BSCP (BSCP vs. RSCP)

Step 4: Detect secondary crashes in impact range
The proposed method can be used to identify the secondary crashes when a prior incident cause observable traffic impact. Some minor crashes may not cause a obvious queue but still lead to secondary crashes as a result of distracted drivers in the approaching traffic. In this case, we assume that secondary crashes only occur within half an hour and half a mile upstream of prior incidents. Crashes that occur in the opposite direction within one hour and one mile upstream of prior incidents are also assumed to be potential secondary ones. In addition to these criteria, a vehicle crash in the opposite direction is identified as secondary only if it occurred in a slow speed traffic flow.

Data Description
A 27-mile section of New Jersey Turnpike (NJTPK) between interchanges 5 and 9 was used as a case study. FIGURE 2 illustrates the site map. The section has 25 traffic sensors placed approximately at every one mile on the mainline. There are 3 lanes between interchanges 5 and 8A, and 5 lanes between interchanges 8A and 9, as shown in FIGURE 2. The posted speed limit of the section is 65 mph. Traffic data collected by each traffic sensor in 2011 are used for the analyses. The data include volume, speed, and occupancy measured at different time intervals at each sensor station. The data aggregated in 5-minute intervals were extracted to develop the speed contour plots using the proposed method.

Other than the traffic data, crash data were also collected. The corresponding crash records in 2011 were obtained from the crash database of the New Jersey Department of Transportation (NJDOT) available online (30). The records include the detailed crash characteristics such as date, time, location, crash type, number vehicles involved, vehicle characteristics, and crash severity. After the exclusion of two crashes occurred on ramps and other with unknown directions, the remaining data consist of a total of 1,188 crash records for the studied section (Southbound: 575 crashes; Northbound: 613 crashes). The secondary crashes among the 1,188 crash records was identified using the aforementioned identification approach.

In addition to the traffic sensor data and crash records, dispatcher data were also used to examine additional information related to crashes. Vehicles involved in a crash reported to the traffic operation center were recorded. They were classified based on their required assistance (e.g., mechanical service, crash, commercial tire service). Corresponding dispatcher data of the vehicles that were involved in crashes were extracted. Key information such as date, dispatch time, end time, vehicle description, roadway direction and milepost was recorded for each vehicle that required towing. In each crash, one or more involved vehicles may call for assistance. Each vehicle that called for assistance had a unique aid number. Dispatch records were linked with NJDOT crash records that are available on-line by matching information associated with direction, milepost, time, and vehicles in both datasets. FIGURE 3 shows an example of examining key information in both datasets. The example shows that a crash occurred at 14:21 (04/22/2011) at the milepost of 56.9 in the northbound direction. There were 4 vehicles involved in the crash. The vehicle table listed the detailed information about each involved vehicle. Similarly, three records in the dispatch data matched the information of the three towed vehicles based on direction,
milepost, vehicle characteristics, and time proximity. There was no dispatch record for Yukon car make (shown in vehicle table) as this vehicle was drivable and did not call for assistance. Thus, for crashes that have no matched dispatch records, we assumed they were quickly removed without calling for assistance. Once the dispatch data were linked with crash data, the crash clearance time was assumed to be the difference between the crash time in crash data and the time to complete the assistance for last vehicle in the dispatch data. For instance, the clearance time of the crash in FIGURE 3 was 101 minutes (from 14:21 to 16:02).

FIGURE 3 Illustration of the process of linking dispatch data with crash records.

DESCRIPTIVE ANALYSIS OF SECONDARY CRASHES

The following sub-sections analyze the critical characteristics of secondary crashes identified by the proposed data mining framework.

Frequency of Secondary Crashes

The method proposed to identify secondary crashes was implemented using the integrated database. FIGURE 4 shows an example of identified secondary crashes. The two plots on the left show the original SCPs with observed crashes for each direction of the studied highway section. The ones on the right show the corresponding BSCPs with the identified secondary crashes. FIGURE 4 (b) illustrates that a primary crash occurred at 13:52 in the southbound induced a secondary crash at 14:14. FIGURE 4 (d) shows that a primary crash occurred at 9:43 caused two secondary crashes at 9:53 and 10:03, respectively.

Overall, 100 secondary crashes were identified as a result of 71 primary crashes. The identified secondary crashes account for 8.42 percent of the 1,188 reported crashes occurred on the studied section. The 71 primary crashes that induced secondary crashes represented 6.98 percent of all non-secondary crashes (1,188-100=1,088). These results indicate that approximately every eleven non-secondary crashes were associated with one secondary crash. Each primary crash on average caused about 1.4 secondary crashes.

The proportion of secondary crashes identified in this case study is inconsistent with findings of previous studies [e.g., 5.22 percent in (9); 5.53 percent in (14); 3.23 percent in (16); 3 percent in (17,18,19); and 0.10 to 0.15 percent in (7)]. The difference is mainly attributable to different approaches used to identify secondary crashes. Previous static methods or queuing-model based methods used to identify secondary crashes did not fully account for the impact of primary crashes. Misclassifications by these methods are expected due to the shortcomings of either the fixed spatio-temporal boundaries or the simplified queuing models. In contrast, the identification approach in this study individually examined the impact of each primary crash and identified the secondary crashes correspondingly. The dynamic
characteristics of the actual impact of primary crashes were captured using the sensor data; thus, providing a more accurate identification of secondary crashes.

![FIGURE 4 Identified secondary crashes using the proposed method.](image)

**FIGURE 4** Identified secondary crashes using the proposed method.

**Spatio-temporal Distributions**

The temporal and spatial characteristics of secondary crashes in relation to primary crashes are shown in **FIGURE 5**. Temporally, it was found that approximately 75 percent of the secondary crashes occurred within two hours after the occurrence of prior crashes. Spatially, 50 percent of the secondary crashes occurred within two miles upstream of the prior ones. Together, 42 percent of secondary crashes occurred within two hours of the onset of primary crashes and within a distance of two miles upstream. 58 percent of secondary crashes occurred beyond these frequently used spatio-temporal thresholds in static methods. These statistics confirm that the proposed identification approach capture more secondary crashes that beyond the capability of the static method. In addition, more than half of the secondary crashes occurred as a consequence of the primary-crash induced queues that last more than two hours. These findings are helpful for developing special incident management and response programs to reduce or prevent secondary crashes. For instance, when and where to deploy advance warning system to disseminate the
primary crash information can be better determined based on the identified spatio-temporal characteristics of secondary crashes.

FIGURE 5 Secondary crashes in relation to primary crashes.

FIGURE 6 shows the distributions of the 100 secondary crashes and 1,088 non-secondary crashes by different time periods. FIGURE 6 (a) indicates that only 11 percent of the secondary crashes occurred between 0:00 and 12:00 in the morning whereas more than half of them occurred between 12:00 and 16:59. The corresponding proportions of non-secondary crashes occurred during these hours were also the highest, which in total accounted for 38 percent of all non-secondary crashes. 24 percent of the secondary crashes occurred between 17:00 and 20:00. The remaining 11 percent of secondary crashes occurred between 20:00 and 24:00. The Pearson’s correlation coefficient between the hourly distributions of non-secondary crashes and secondary crashes is 0.777.

FIGURE 6 Crash distributions by different time resolutions.

FIGURE 6 (b) presents the crash distributions by day of the week. It shows that both non-secondary crashes and secondary crashes are more likely to occur on Fridays. This finding is consistent with the previous findings of Kopitch and Saphores (14). Specifically, Fridays accounted for 26 percent of secondary crashes, which were followed by Sundays with 19 percent. The occurrence rate of secondary crashes on Monday were the lowest, with only 7 percent. The non-secondary crashes also showed similar
characteristics of distributions by day of week. The Pearson's correlation coefficient between the
distributions by day of week is 0.929.

FIGURE 6 (c) captures monthly distributions of different categories of crashes. More than 10
percent of secondary crashes occurred in each month between April and August. November and
December each also shared more than 10 percent of secondary crashes. Particularly, 15 percent of all
secondary crashes were observed in December. Secondary crashes were less frequent in other months.
Particularly, no secondary crash was observed in February and the proportion of non-secondary crashes in
the same month was also the smallest. The Pearson's correlation coefficient between the monthly
distributions of non-secondary crashes and secondary crashes is 0.641.

Despite the variations, all Pearson's correlation coefficients measured in different time periods
suggest that there is a positive association between the occurrences of non-secondary crashes and
secondary crashes. In other words, the more non-secondary crashes occur in a given time period, the more
secondary crashes are likely to occur. Thus, it is necessary to pay more attention to the time periods that
have high crash frequency.

Incident Clearance Time
FIGURE 7 (a) shows the results of the identified crashes linked with dispatcher records. It shows that 33
out of the 71 primary crashes did not have any entry in the dispatcher data. Similarly, 67 out of the 100
secondary crashes did not have any entry in the dispatcher data. It is assumed that the crashes that did not
require towing were quickly removed (clearance time \( \leq 30 \) minutes). FIGURE 7 (b) presents the
distribution of the clearance time for the identified primary and secondary crashes. Overall it shows that
the secondary crashes were quickly cleared than the primary crashes. For instance, approximately 90
percent of the secondary crashes were cleared within 90 minutes. In contrast, only about 70 percent of the
primary crashes can be cleared within 90 minutes.

![Graph showing incident clearance time distribution](image)

The shorter clearance time of secondary crashes should be attributable to minor damage of
vehicles involved in these crashes. It can be claimed that the more vehicles requiring towing service the
longer time it takes to clear the crash. The crash data described information on how the vehicles were
removed: (a) Driven, (b) Left at Scene, and (c) Towed (including impound and disabled). For the 155
vehicles involved in the primary crashes, about 50 percent of them were towed and 50 percent were
driven. In contrast, only 20 percent of the 220 vehicles involved in the secondary crashes were towed and
80 percent were driven.
Crash Severity and Type
Of the 100 identified secondary crashes, 20 were injury crashes and 80 were property damage only (PDO) crashes. Similarly, there were 216 injury crashes (including 2 fatal crashes) and 872 PDO crashes among the 1,088 non-secondary crashes. The proportion of the injury crashes among the secondary and non-secondary crashes was 20.0 percent and 19.9 percent, respectively. The proportion of injury crashes in secondary crashes and non-secondary crashes was not statistically different (proportional test \( p \)-value >0.1). These results suggest that the crash severity of secondary crashes do not necessarily have to be less severe than non-secondary independent crashes. This result is inconsistent with the previous finding in (11) that indicated the crash severity for secondary crashes is lower.

TABLE 1 shows that rear-end crash was the dominant secondary crash type along the studied highway section. Specifically, 75 percent of secondary crashes were rear-end, which is approximately 31 percent higher than that of non-secondary crashes. Side-swipe was the second highest crash type of the secondary crashes. 19 percent of the identified secondary crashes were side-swipe crashes, which were only 3 percent less than that of the non-secondary crashes. The rear-end crashes and side-swipe crashes together accounted for 94 percent of the secondary crashes. They were significantly higher than that of non-secondary crashes (\( p \)-value < 0.001).

Fixed object crashes and other crashes together accounted only for 6 percent of the secondary crashes. These results suggest that secondary crashes are more likely to be collisions between vehicles. Indeed, 96 percent of the secondary crashes were found to involve two or more vehicles whereas only 74 percent of the non-secondary crashes involved multiple vehicles.

### TABLE 1 Crash Distribution by Collision Type

<table>
<thead>
<tr>
<th>Collision Type</th>
<th>Independent Crash (n=1,088)</th>
<th>Secondary Crash (n=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rear end</td>
<td>43.8%</td>
<td>75.0%</td>
</tr>
<tr>
<td>Side swipe</td>
<td>22.0%</td>
<td>19.0%</td>
</tr>
<tr>
<td>Fixed object</td>
<td>19.3%</td>
<td>4.0%</td>
</tr>
<tr>
<td>Non-fixed object</td>
<td>8.5%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Overturned</td>
<td>1.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Others</td>
<td>5.4%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Contributing Circumstances
TABLE 2 shows the major contributing circumstances for secondary and non-secondary crashes. The major contributing factor to the secondary crashes was "following too closely", which was 54.0 percent. This proportion is 21.8 percent more than that of non-secondary crashes. The difference in percentages can be attributed to the possibility of vehicles being more likely to keep shorter distance gaps in the queue caused by primary crashes. Compared to the other crashes, the secondary crashes were also more likely to be associated with improper lane change and inattentive driving. These three contributing circumstances together accounted for 84.0 percent of the contributing factors for the secondary crashes whereas they only account for 63.6 percent of the contributing factors for the non-secondary crashes. These factors in turn resulted in more rear-end and side-swipe crashes as shown in TABLE 1. To mitigate secondary crashes, alerting motorists to keep a safe distance and stay in lane in the impact area of primary crashes can thus be beneficial.

### TABLE 2 Major Contributing Circumstances of Different Crashes

<table>
<thead>
<tr>
<th>Factor</th>
<th>Independent Crash (n=1,017)</th>
<th>Secondary Crash (n=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Following too Closely</td>
<td>32.2%</td>
<td>54.0%</td>
</tr>
<tr>
<td>Improper Lane Change</td>
<td>16.8%</td>
<td>12.0%</td>
</tr>
<tr>
<td>Driver Inattention</td>
<td>13.6%</td>
<td>18.0%</td>
</tr>
<tr>
<td>Unsafe Speed</td>
<td>15.3%</td>
<td>8.0%</td>
</tr>
<tr>
<td>Others</td>
<td>22.1%</td>
<td>8.0%</td>
</tr>
</tbody>
</table>
In addition to the contributing factors, the vehicles that contributed to the crashes were also examined. It was found that approximately 21 percent of the vehicles that caused independent crashes and primary crashes were heavy vehicles (trucks and/or bus). In contrast, the heavy vehicles accounted only for 12.0 percent of the secondary crashes, whereas passenger cars were responsible for the remaining 88.0 percent. Heavy vehicles were approximately twice as less likely to cause the secondary crashes. This may be attributed to the fact that heavy vehicles do not have the high maneuvering flexibility as passenger cars, and less likely to change lanes aggressively and follow too closely in impact area (queue) of primary crashes.

CONCLUSIONS
The main goal of this study was to identify secondary crashes on highways and analyze their characteristics in detail. Data from existing crash records, roadway traffic sensors and the available incident response data were integrated and mined for this purpose. A analysis framework developed for more accurate identification of secondary crashes was presented. Then, the descriptive analysis of the identified secondary crashes was performed. The major characteristics of these crashes were carefully investigated through a comparative analysis of secondary and non-secondary crashes. The empirical findings based on available data sources are summarized as follows:

1. The results of the descriptive analyses indicate that secondary crashes accounted for more than 8 percent of the crashes occurred on the studied freeway. The frequency of secondary crashes was positively correlated with the number of non-secondary crashes occurred in a given time period. Approximately every eleven crashes were associated with one secondary crash.

2. Almost half of secondary crashes were found to occur within two miles upstream of primary crashes. The other half occurred more than two miles away. Thus, it is crucial to prevent the formation of long queues induced by primary crashes.

3. 75 percent of secondary occurred within up to two hours of primary crashes. The remaining 25 percent occurred due to primary crashes that have impact lasting more than two hours. In order to reduce the number of secondary crashes, the primary crashes have to be quickly removed.

4. Secondary crashes were more likely to involve two or more vehicles. They were more likely to be rear-end type crashes (when compared to non-secondary crashes). These induced secondary crashes were as severe as the non-secondary crashes because they have similar proportions of the injury crashes.

5. In addition, “Following too closely”, “Driver Inattention”, and "Improper lane change" were the major reported contributing circumstances for the secondary crashes. Heavy vehicles were less likely to cause secondary crashes. Therefore, it is very important to provide advance warnings to alert approaching vehicles, particular passenger vehicles.

6. Though more vehicles were likely to be involved in secondary crashes, the average clearance time of the secondary crashes was relatively shorter than that of the primary crashes. This is mainly the less damage of vehicles in secondary crashes. It was found only 20 percent of the vehicles involved in secondary crashes required towing service, whereas half of the crashes involved in primary crashes were towed. The more towing services need, the longer time it will take to clear a crash.

This study provided some preliminary analysis on the characteristics of secondary crashes on highways. These results can help transportation agencies have a better understanding of secondary crashes and implement countermeasures to reduce or prevent secondary crashes accordingly. In addition, the analysis presented in this study also illustrates how secondary crash information can be more accurately mined from multiple existing data sources. As this study only examined limited data sources, further in-depth examinations of secondary crashes as well as primary crashes based on more high-resolution traffic sensor data and high-quality incident records are suggested. Secondary crashes caused by other non-crash incidents also deserve more investigations.
ACKNOWLEDGEMENTS

The contents of this paper reflect views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents of the paper do not necessarily reflect the official views or policies of the agencies.

REFERENCES


