Using Detailed Signal and Detector Data to Investigate Intersection Crash Causation

Final Report

Prepared by:

Gary A. Davis
Indrajit Chatterjee

Department of Civil Engineering
University of Minnesota

CTS 13-05

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.
CTS 13-05		

4. Title and Subtitle		5. Report Date
Using Detailed Signal and Detector Data to Investigate Intersection Crash Causation		January 2013
		6.

7. Author(s)	8. Performing Organization Report No.
Gary A. Davis and Indrajit Chatterjee	

9. Performing Organization Name and Address	10. Project/Task/Work Unit No.
Department of Civil Engineering University of Minnesota 500 Pillsbury Drive SE Minneapolis, MN 55455	CTS Project #2010067
	11. Contract (C) or Grant (G) No.

12. Sponsoring Organization Name and Address	13. Type of Report and Period Covered
Intelligent Transportation Systems Institute Center for Transportation Studies University of Minnesota 200 Transportation and Safety Building 511 Washington Ave. SE Minneapolis, Minnesota 55455	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

http://www.its.umn.edu/Publications/ResearchReports/

16. Abstract (Limit: 250 words)

Traffic crashes may not always result in severe or fatal injuries, but they can still have nontrivial impacts on system performance, particularly during heavy traffic conditions. One way toward reducing the frequency of such incidents is to first identify the necessary circumstances that resulted in the collision. However, road crashes, particularly intersection related crashes, are complex phenomenon and often result from different combinations of causal factors. Recently, methods for recording high-resolution arterial traffic data have been developed, and it is important for traffic safety engineers to explore such high-resolution data to understand the causes of crashes. In this research one such integrated event based system, known as SMART SIGNAL, which collects and stores detailed loop detector and signal activity, was used to identify the events leading to a crash or a potential crash and illuminate the mechanisms by which traffic conditions and driver decisions interact to produce those events. Two specific event types, a signal violation crash and vehicle pedestrian crash, were evaluated. For the signal violation crash study, SMART SIGNAL data were used to identify the incident and the vehicles involved in the crash. It was then shown how high-resolution data could support a traditional reconstruction of this crash. For vehicle pedestrian interactions, detector and signal activity data were used to predict pedestrian crash risk in the absence of clearance interval at three signalized intersections. A simulation-based method was used to first estimate crash probabilities, and then a counterfactual approach to calculate the probability of the absence of the all-red phase as a necessary condition for the occurrence of the crash provided an alternate estimate of crash-reduction factors for the all-red phase.

17. Document Analysis/Descriptors		18. Availability Statement	
Traffic signals, Collisions, Intersections, Pedestrians		No restrictions. Document available from: National Technical Information Services, Alexandria, Virginia 22312	

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	47	

Using Detailed Signal and Detector Data to Investigate Intersection Crash Causation

Final Report

Prepared by:

Gary A. Davis
Indrajit Chatterjee

Department of Civil Engineering
University of Minnesota

January 2013

Acknowledgment

Table of Contents

List of Figures

List of Tables

Executive Summary

This research is primarily focused on determining the usefulness of high-resolution detector and signal data collected at urban intersections to identify the events leading to a crash or a potential crash, and thereby illuminate the mechanisms by which traffic conditions and driver decisions interact to produce crashes or near crashes. Road crashes are not simple phenomena and generally result from differing combinations of causal factors, whose relative importance can vary from case to case. For example, an intersection crash could have resulted from a combination of excessive speed, driver inattention, signal violation, insufficient sight distance, and poor gap choice, any one of which, if modified, could have prevented the crash. Although statistical-based safety studies can sometimes estimate aggregate causal effects, identifying how the actions of different drivers interact with each other and with roadway features to produce crashes requires a more microscopic approach to both the modeling of crash events and the collection of crash data.

One potentially useful, but neglected, source of information is available at signalized intersections, where pavement-based sensors are used to drive traffic-actuated control logics. Pavement-based sensors register the presence of vehicles, and their output can be used to measure macroscopic variables such as traffic volume, lane occupancy, and average speed. Recent work at the University of Minnesota has produced a package of hardware and software, which when attached to a signal controller, can record and store high-resolution data for both sensor activations and signal indications. Known as SMART SIGNAL, this package has been used to support modeling of queue formation on systems of signalized intersections and to drive a novel method for estimating arterial travel times. Road accidents generally disrupt normal traffic flows, hence it is expected that at least for some crashes, it will be possible to use the SMART SIGNAL data to identify when the crash occurred and then determine the signal indications and loop detector data associated with the crash event. In this study two types of events were investigated in-depth: signal violation crashes and potential vehicle-pedestrian conflicts at signalized intersections.

Currently, the SMART SIGNAL system has been installed at various intersections on Trunk Highway (TH) 55 and on Eden Prairie Centre Drive, in the Twin Cities metropolitan area in Minnesota. Three different intersections (two on TH 55 and one on Eden Prairie Centre Drive) were chosen for detailed analysis.

After giving a brief description of the SMART SIGNAL system, this report describes a detailed investigation of an intersection angle-crash resulting from signal violation. First, details of the event were extracted from the set of crash records for the intersections along TH 55 corresponding to the periods when SMART SIGNAL was active. Second, archived detector and signal data from SMART SIGNAL were extracted for a time window bracketing the occurrence of the crash (approximately half an hour before and after the event). The next task was to identify the crash based on occupancy data extracted from the SMART SIGNAL system. A well-known parameter change detection technique was adopted to detect the change in the pattern of occupancies before and after the incident. Once the crash was identified, the next effort was to reconstruct the event. It turned out that the detector data and the standard crash report were not sufficient by themselves to support the reconstruction. In order to illustrate how the behavior of

crash-involved drivers could be inferred from SMART SIGNAL data when standard post-crash scene information is available, a plausible hypothetical scenario based on qualitative information from the crash report was developed. A three-parameter vehicle trajectory model was developed, and a Monte Carlo simulation technique was used to estimate the conditional distribution of the parameters given the crash had occurred. Results indicated that even with a relatively strong braking effect, the crash could not be avoided, suggesting the excessive speed of the violating vehicle was a significant factor for the signal violation, and consequently the crash occurrence.

The report then describes how SMART SIGNAL data can be used to assess vehicle pedestrian crash risk in the absence of clearance intervals at three different signalized intersections. A recent field study in New York has suggested that retiming of clearance intervals resulted in a 37% relative reduction in pedestrian and bicycle crashes. Since these modifications generally led to increases in the clearance intervals, results from the study have implications regarding the mechanism by which at least some vehicle-pedestrian collisions occur. To validate such findings, a simulation-based method using SMART SIGNAL data was used to first estimate crash probabilities, and then a counterfactual approach was used to calculate the probability that the absence of the all-red phase was a necessary condition for the occurrence of the crash. This counterfactual probability can be interpreted as an alternative estimate of a crash-reduction effect, and our study suggested that provision of clearance intervals was able to prevent 100% of potential crashes, which would have occurred otherwise even with a driver's emergency braking effort. The difference between this result and the findings from the New York field study could be attributed to the mid-block or signal violation by pedestrians as additional sources of vehicle pedestrian conflicts.

Conclusions

1. It is possible, at least for some crashes, to use SMART SIGNAL data to identify the event and the vehicles involved in the event.
2. However, crash reconstruction was not possible using only SMART SIGNAL data. Additional information or field evidence, such as post-impact information, was required to draw meaningful conclusions. When such information is available, SMART SIGNAL data can provide information on driver behavior not typically available in a standard crash reconstruction.
3. For vehicle/pedestrian conflict analysis, it was possible to identify several events in which a vehicle exited an intersection after the signal indication turned red. Using, reasonable assumptions about pedestrian behavior based on past literature it was possible to then simulate the effect of the presence of an all-red clearance interval on the occurrence of vehicle-pedestrian conflicts.
4. Finally it is expected that in the future as far as traffic safety studies are concerned, high-resolution data collection and storage system, such as SMART SIGNAL, can make a significant contribution if used in conjunction with the other existing data collection procedures, such as video-based methods and other trajectory extraction techniques.

Chapter 1.   Background and Project Objectives

1.1   Background

Traffic crashes may not necessarily result in fatal or severe injuries, but they can still have adverse impacts of system performance, resulting in travel delays, congestion and excessive emissions. Ultimately, reducing the frequency of crashes will require identifying their causes, and then implementing appropriate countermeasures to remove or counter those causes. Road crashes are not simple phenomena, and generally result from differing combinations of causal factors, whose relative importance can vary from case to case. For example, at an intersection a crash could have resulted from a combination of excessive speed, driver inattention, signal violation, insufficient sight distance, and poor gap choice, any one of which, if modified, could have prevented the crash. Although statistical-based safety studies can sometimes estimate aggregate causal effects, identifying how the actions of different drivers interact with each other, and with roadway features, to produce crashes requires a more microscopic approach to both the modeling of crash events and the collection of crash data. This need has in fact been recognized in the research plan for the Strategic Highway Research Program 2 (SHRP2) Safety Program, which is conducting an extensive longitudinal study using specially-instrumented vehicles, together with limited development of video-based methods for collecting data at particular sites. Prominent among naturalistic driving studies are ongoing SHRP 2 study, the Virginia Tech Transportation Institute's 100-Car Study (*1*), and the Automotive Collision Avoidance System field test conducted by the University of Michigan Transportation Institute. In these studies volunteers drive instrumented vehicles which continuously collect and record measurements such as vehicle position, speed, direction, acceleration, as well as radar-based range and range-rate measurements for other vehicles. It is then possible, using this vehicle-based data, to apply trajectory-based modeling techniques to reconstruct crash related events, and in turn estimate the posterior distribution of important event parameters such as braking accelerations, reaction times, and critical headways (*2*). However, both the vehicle-based and site-based approaches require a temporary deployment of special equipment, and because crashes tend to be rare events, the number of actual crashes captured by such deployments will necessarily be limited. In the foreseeable future then, learning how and why crashes occur will still require post-hoc investigations of actual crash events, where the information is limited to what is available from inspection of the involved vehicles and investigation of the crash scene.

One potentially useful, but neglected, source of information is available at signalized intersections, where pavement-based sensors are used to drive traffic-actuated control logics. Pavement-based sensors register the presence of vehicles, and their output can be used to measure macroscopic variables such traffic volume, lane occupancy, and average speed. Many urban freeway systems which use pavement-based sensors to drive ramp-metering logics also archive the aggregated measures from the sensors, which provide useful data for a number of transportation-related research problems. For signalized intersections however, data archiving is more the exception than the rule and, as far as we know, archiving data at the level of the individual vehicle is not typically done. Recent work at the University of Minnesota has produced a package of hardware and software which, when attached to a signal controller, can record and store high-resolution data for both sensor activations and signal indications. Known as SMART SIGNAL, this package has been used to support modeling of queue formation on

systems of signalized intersections, and to drive a novel method for estimating arterial travel times (*3*). To illustrate the potential value of high-resolution detector data, consider the classic problem where an estimate of a vehicle's speed is desired, and the available information consists of a measurement of skid-mark length. Letting *d* denote this measurement and *v* denote the speed of the vehicle at the beginning of the skid-mark, *v* is then given by

$$v = \sqrt{2 \times \mu \times g \times d}, \tag{1}$$

where, μ is the coefficient of tire-pavement friction and g the gravitational acceleration.

If all quantities on the right-hand side of equation (1) are known with certainty then the estimated speed will also be certain, but such categorical certainty is almost never to be had. One effective way to allow for uncertainties is the Bayesian approach, where prior information regarding plausible values for *v* and *μ* are combined with an assessment of the measurement error for *d* to produce a posterior distribution for *v*. Suppose then that we have, in addition, an occupancy time from a detector with known finite measurement error. Assuming the length of the vehicle is known, a new posterior estimate of the speed can be computed, combining the information from the observed occupancy time and the skid mark measurement. Figure 1.1 compares the posterior distributions of the speed computed for the two data scenarios showing the potential for detector information to improve the precision of a speed estimate.
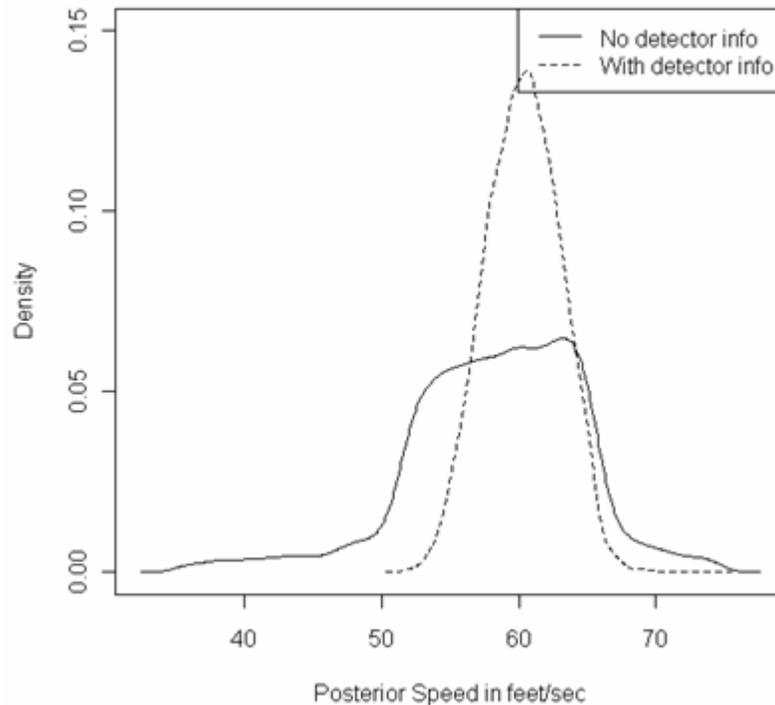
Figure 1.1: Posterior estimate of the speed from skid-mark and occupancy data.

1.2   Objectives

The immediate objective of this research is to determine the degree to which the SMART SIGNAL system for collecting and storing detailed loop detector and signal activity data can be used to identify the events leading to a crash, and illuminate the mechanisms by which traffic conditions and driver decisions interact to produce crashes. This will be done by using SMART SIGNAL data to reconstruct individual crashes, and to identify potential crash events to be simulated, with a focus on crashes resulting from signal violations or near-violations. In addition to crashes involving vehicles only, a typical pedestrian vehicle interaction at signal crossing would be explored, and the role of all-red clearance interval in preventing vehicle pedestrian conflicts will be assessed using simulation.

1.3   Tasks

Since crashes generally disrupt normal traffic flows, it is expected that at least for some of the crashes it will be possible to use the SMART SIGNAL data to identify when the crash occurred, and so determine signal indications and loop detector data associated with the crash event. To accomplish the aforementioned goals, crash records for angle-type crashes were obtained for the intersections along Minnesota Trunk Highway (MNTH55) corresponding to the periods during which SMART SIGNAL data were available. Then archived SMART SIGNAL detector and signal data for time windows bracketing the occurrences of the crashes were extracted, for an exploratory analysis of traffic patterns during the periods of interest. For one of these events it was possible to identify, via disruptions in traffic flow, when the crash occurred, and we then demonstrate how high-resolution event-based data obtained from loop detectors and signal can be used to identify the vehicles involved in the crash, and how traditional crash reconstruction can be enhanced by SMART SIGNAL data.

For vehicle pedestrian interaction analysis, due to the absence of any pedestrian data, a typical pedestrian behavior at signalized intersection was assumed and information from detector and signal data was used to predict whether intervention of all-red clearance interval were able to avoid a vehicle pedestrian crash which would have occurred in the absence of it.

The remainder of the report is organized as a follows. Chapter 2 provides a brief description of the SMART SIGNAL data, the relevant intersection locations, and the crash records obtained from MnDOT. Chapter 3 gives a detailed investigation of an intersection angle-crash resulting from signal violation. Chapter 4 demonstrates how SMART SIGNAL data could be used to assess vehicle pedestrian crash risk in the absence of clearance interval at three different signalized intersections. Finally Chapter 5 presents our conclusions and recommendations.

Chapter 2. Data Collection and Reduction

2.1   SMART SIGNAL Data and Crash Records

Recently there have been several studies based on high-resolution data to evaluate arterial performance measures such as travel time, queue length estimation and travel delay (*3*). SMART SIGNAL is one such integrated event-based data collection and storage system (*4*). The data collection system simultaneously collects high-resolution event-based traffic data including every vehicle actuation over detectors located near intersections and every signal indicator change. Two types of data can be retrieved from SMART SIGNAL archives: vehicle arrival and departure times at the detectors and changes in the signal indications. The time stamp associated with each event enables identification of past traffic states based on the occupancy and gap measurements obtained directly from detector actuations. Currently, the SMART SIGNAL system has been installed at various intersections on Minnesota Trunk Highway (TH) 55 (namely, Boone Ave, Winnetka Ave, Rhode Island, Douglas Drive, TH 100 interchange), and on Eden Prairie Centre Drive, in the Twin Cities region of Minnesota. Table 2.1 shows typical data that can be obtained from the SMART SIGNAL system. The first column gives the detector number, second and the third column provides information about the vehicle arrivals at the detector and for how long it was occupied respectively. The signal status corresponding to each vehicle's arrivals was listed in the sixth column, whereas the fifth and seventh columns enumerated the beginning and the duration of signal status respectively.

Table 2.1: Sample SMART SIGNAL data.

Detector	Actuation start	occ (secs)	Phase	Signal start	status	dur (secs)
26	10:30:33.515	2.578	6	10:30:30.328	G	59.203
26	10:30:37.125	1.218	6	10:31:29.531	Y	5.5
26	10:30:39.281	0.969	6	10:32:23.828	G	55.703
26	10:30:42.921	0.813	6	10:33:19.531	Y	5.5
26	10:30:45.359	0.859	6	10:33:59.640	G	73.891
26	10:30:46.750	0.734	6	10:35:13.531	Y	5.5
26	10:31:01.156	0.422	6	10:36:25.031	G	72.5
26	10:31:02.859	0.375	6	10:37:37.531	Y	5.5
26	10:31:19.109	0.344	6	10:38:54.343	G	47.203
26	10:31:24.562	0.406	6	10:39:41.546	Y	5.485

Table 2.2 enumerates the number of crashes that occurred at the intersections where SMART SIGNAL data was available between the periods of 2008 to 2009.

Table 2.2: Crash Summaries at SMART SIGNAL enabled intersections on TH 55 from 2008 to 2009.

SYS	Intersection	Ref. Point	ACC_NUM	MONTH	DAY	YEAR	TIME	sev	diag
MNTH	Boone	184+00.558	082850120	10	10	2008	2214	C	3
MNTH	Boone	184+00.558	091410106	5	10	2009	1601	B	5
MNTH	Boone	184+00.558	091880116	6	19	2009	1402	N	1
MNTH	Boone	184+00.558	092130086	8	1	2009	1651	C	3
MNTH	Winnetka	185+00.055	082430038	8	29	2008	1607	C	5
MNTH	Winnetka	185+00.055	090490153	1	21	2009	0920	B	5
MNTH	Winnetka	185+00.055	090710109	2	27	2009	1052	N	1
MNTH	Rhode Island	185+00.227	083360131	11	12	2008	1711	N	1
MNTH	Rhode Island	185+00.227	092180022	8	5	2009	1520	N	1
MNTH	Glenwood	185+00.567	082810119	9	24	2008	1517	N	1
MNTH	Glenwood	185+00.567	090260185	8	14	2008	0841	N	1
MNTH	Glenwood	185+00.567	091270047	4	9	2009	1113	N	1
MNTH	Douglas	186+00.090	090330162	2	2	2009	1556	N	1
MNTH	Douglas	186+00.090	091690097	5	29	2009	1756	C	1
MNTH	JC TMNTH 100	186+00.622	083410015	12	5	2008	1905	N	1
MNTH	JC TMNTH 101	186+00.622	091740033	6	22	2009	1815	C	1
MNTH	JC TMNTH 102	186+00.622	091740034	6	22	2009	1749	B	1
MNTH	JC TMNTH 103	186+00.622	091950221	7	14	2009	1813	N	1
MNTH	JC TMNTH 104	186+00.622	092050153	7	24	2009	1750	C	1
MNTH	JC TMNTH 105	186+00.622	092400109	8	28	2009	0846	N	1

Table 2.2 suggests most of the crashes are rear-end (diagram code=1), and property damage only crashes (PDO). In addition to the SMART SIGNAL data, the police accident reports were also requested for each of the crashes listed in Table 2.2.

After initial review of each of the crashes, based on the nature and the severity of the crash, an angled crash at the intersection of Winnetka Ave (Accident number: 082430038) was deemed to be reasonable for our case study. The next chapter provides a detailed causal analysis of the angled crash and illustrated how SMART SIGNAL data along with police crash report could be used to identify the probable cause behind the crash.

Chapter 3.   Analysis of Signal Violation

3.1   Description of the Event

In this chapter we focus on one particular event at the intersection of TH 55 and Winnetka Ave. In addition to the SMART SIGNAL data, a crash report of the event was requested from Minnesota Department of Transportation (MnDOT). The selection of this particular event from the pool of crash records was based on our judgment that the crash was severe enough to impact the normal traffic conditions, and so make it possible to identify, from SMART SIGNAL data, exactly when the crash occurred. Factors such as type of crash (in this case angle crash), location of the crash (within the intersection), time of crash (close to evening peak hour) and collision impact were considered. Figure 3.1 shows the aerial view of the crash location, where red arrow indicating the direction of vehicle violating the signal, while the green arrow represents the vehicle with the right of way.
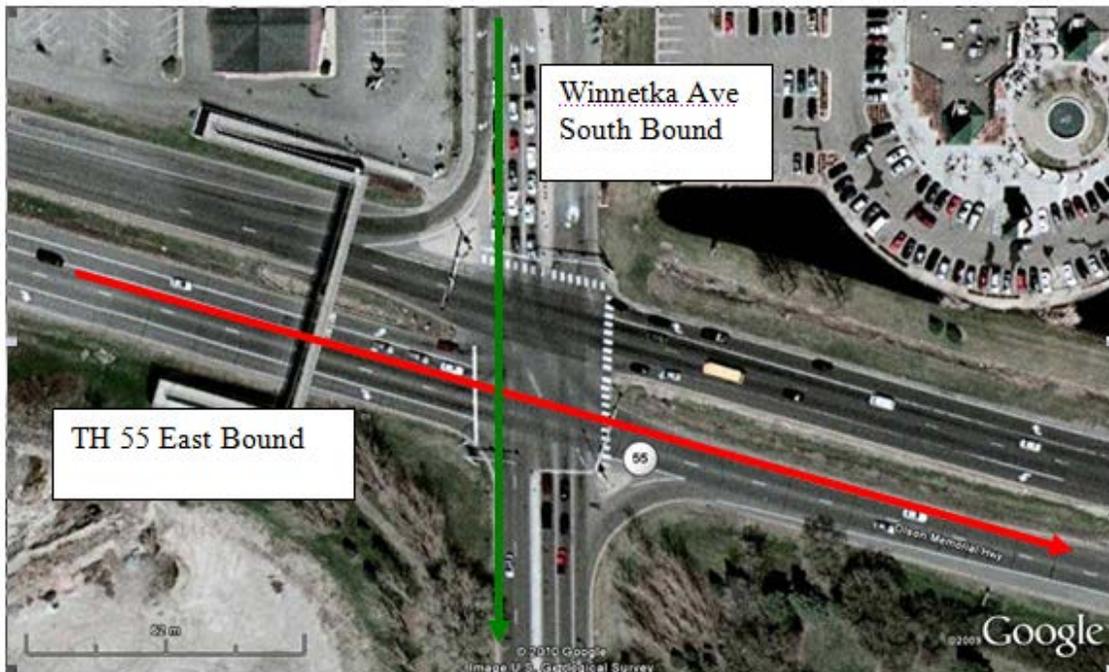
Figure 3.1: An aerial view of the crash location at TH 55/ Winnetka Ave.

TH 55 is an east-west running state highway, while Winnetka Ave is north-south county arterial. TH 55 EB (approaching the intersection) has two through and two left turn lanes with a separated right turn, while Winnetka south bound has two through lanes of which one is shared left turn lane, and a separate left-turn and right-turn lane. The following section gives the description of the event as documented in the police crash report.

According to a witness, unit 1 was driving eastbound on TH 55, while unit 2 was southbound on Winnetka and approaching the intersection with a green light. Unit 1 violated the signal and collided with unit 2. The collision then caused unit 1 to hit a NB vehicle which was waiting at a

red light. The time of the crash was recorded as 16:07 in the report. The description of the event suggests that the crash was severe enough to impact traffic conditions at the intersection. Also information about the vehicle type could help to calculate speeds of the vehicles involved in the crash based on occupancy data from the detectors.

3.2   Crash Identification Problem

The main idea behind this study is based on the observation that severe crashes tend to disrupt normal traffic flows and so it may be possible, using archived event-based data (including both vehicle-detector actuation events and signal phase change events) from SMART SIGNAL, to first determine when the crash occurred, and then identify the corresponding signal indications and the loop detector data for the involved vehicles. The crash identification problem can be treated as a type of Incident Detection Problem. Although we are not proposing a new algorithm for incident detection, it will be a helpful to review some of the relevant literature on incident detection on arterial networks.

3.3   Literature Review

In the past decade or so there has been a constant effort to improve the efficiency of existing roadways. One of the means to accomplish this goal is the introduction of incident management systems that involve initial incident detection, emergency vehicle co-ordination, clearing the incident site, informing motorists and restoring normal traffic conditions. A substantial literature is devoted to freeway incident detection, and somewhat less to signalized urban arterial incident detection, but a common feature associated with incidents is a sudden decrease in roadway capacity due to lane blockage, leading to reduction in speeds and longer lane occupancy times at detector stations. However, to be detected an incident must be severe enough for traffic data to reflect a noticeable deviation from normal conditions (*5*).

Incident detection algorithms fall broadly into two categories: (a) Pattern recognition algorithms which compare current traffic patterns with historic traffic data (the most common source of data being loop detectors), where a difference exceeding some predefined threshold value (calibrated based on the location) indicates an incident (*6, 7*). However, such algorithms may not be able to differentiate between incidents and recurring congestion. (b) Short term prediction algorithms that involve statistical procedures such as time series to predict future traffic measurements, with an incident being detected when current measurements fall outside the confidence limits of the forecast (*8, 9*). Han and May (*10*) indicated that an arterial incident detection algorithm based on artificial intelligence was more likely to be effective for incidents which are in proximity to detector locations. A similar fact was mentioned in Sethi et al. (*11*), where the authors concluded that a higher detection rate was achieved for incidents located near or at intersections, because such incidents were expected to have greater impact on traffic conditions.

Interestingly, a seemingly unrelated area receiving attention is the remote detection of changes in land cover for earth science. Generally such studies involve satellite image comparisons based on pixels, as well as traditional time series approaches to detect land changes (*12*). Change detection problems involving time series data deal with two separate problems (a) detection (whether change has occurred or not), (b) estimation (location of change point). A wide variety of literature based on both frequentist and Bayesian techniques for change detection is available (*13,*

10

14). Change detection techniques originated in manufacturing process for quality control, where time series measurements were expected to follow a certain underlying distribution and any deviation was identified as an error in the manufacturing process. Such a method is often termed as a parameter change approach, where the generating process is characterized by a parameter of the underlying distribution. A substantial literature is available for change detection based on such techniques (*14*). Often, with parameter change technique an additional test statistic is used to determine if a change point exists or not.

One such popular parameter change technique was introduced by Page (*15*) based on a statistic called cumulative sum (CUSUM) to detect a changes in a mean value. A wide application of CUSUM can be found in many change detection studies (*16, 17, 18*), and here we will use the CUSUM statistic to detect changes in detector occupancy time. The basic idea behind CUSUM is based on prior knowledge of an expected measure of a process. CUSUM then calculates the cumulative deviation of the current measure from its expected value. For example let $\{t_1, t_2,...,t_n\}$ be a time series of observations with a known expected value, say μ. Then CUSUM statistic, CS_k at k time point is calculated as

$$CS_k = \sum_{i=1}^{k}(t_i - \mu) \qquad (2)$$

Let us illustrate this with a simple example. Suppose, given a time series $\{t_1, t_2, ... t_{300}\}$ which is a sequence of two Gaussian processes as follows

$$t_i \sim N(0, 0.20) \quad \forall \, i < 150$$
$$t_i \sim N(0.15, 0.20) \quad \forall \, 150 \leq i \leq 300 \qquad (3)$$

A CUSUM statistic is calculated by applying Equation (2) to data simulated according to (3) assuming $\mu=0$, is shown in Figure 3.2.
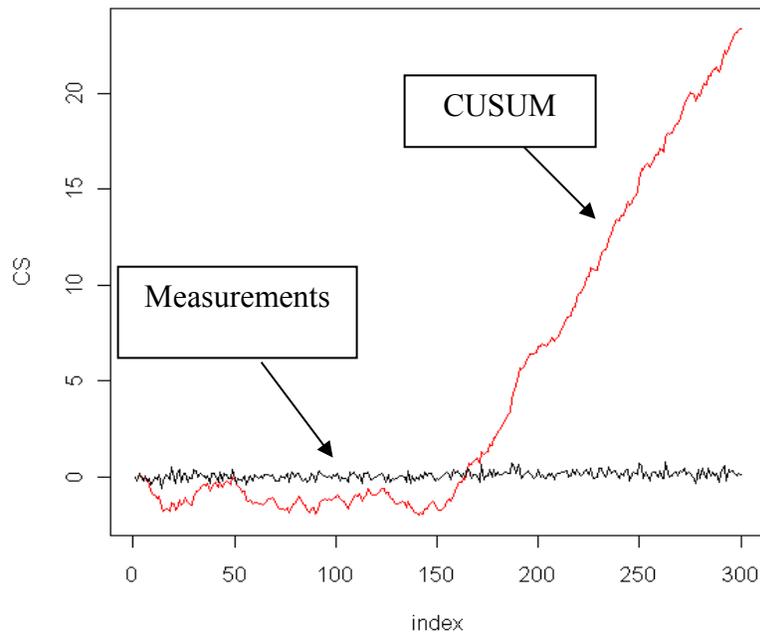
Figure 3.2: CUSUM statistic for a sequence of two Gaussian processes.

Figure 3.2 shows that the CUSUM statistic is around zero until 150th time step after which it rapidly increases monotonically.

3.4   Identification Methodology

In this study, once the time of the event was approximately located from the crash report archived detector and signal data from SMART SIGNAL system were obtained for a time window bracketing the crash occurrence (half an hour before and after the incident as documented in the crash report). Then a local query was set using Excel to segregate the detector actuation events (i.e. occupancy time) based on whether the corresponding signal phase was red or green. Figure 3.3 shows the detector layout of the intersection. Sequences of occupancy plots corresponding to separate green and red phases for detectors 4 and 3, located on east bound TH 55 about 400 feet upstream of the intersection, were prepared. Figure 3.5 shows the occupancy time measurements around 16:07 for detector 4 when the signal phase was red for the approach. It is evident from the figure that around 16:06 an event with very low occupancy time (i.e. high speed) was observed. For comparison, the general occupancy pattern during the red phase for the same approach was evaluated for an adjacent period of the incident (see Figure 3.4).
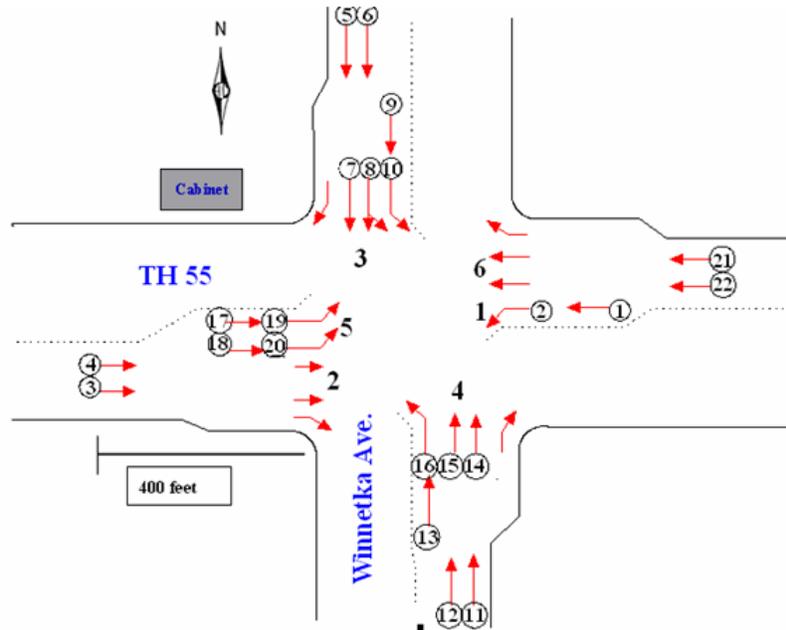
Figure 3.3: SMART SIGNAL detector layout of Winnetka Ave/ TH 55.
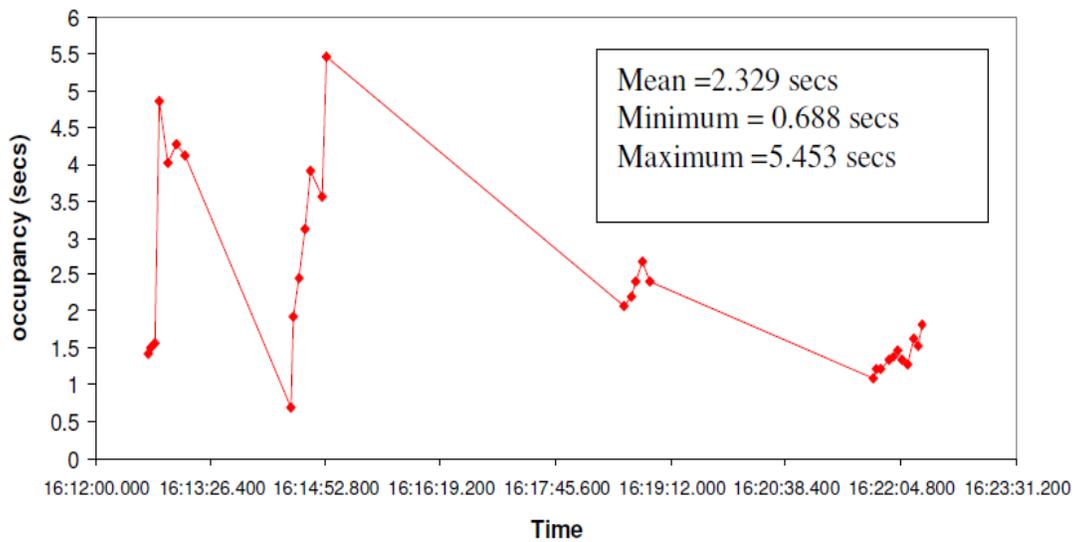
Figure 3.4: General occupancy pattern for detector 4 (red phase).

Descriptive statistics, with mean=2.329 secs, minimum=0.688 secs and maximum=5.453 secs, for general occupancy pattern suggested that the occupancy of 0.235 secs (from Figure 3.5) is below the general pattern, which indicates a potential target vehicle with much higher speed in comparison to other vehicles approaching the intersection during the red phase. However, at this point we might not be able to identify with certainty this vehicle as the vehicle ultimately

13

involved in the crash. For this we examined the possibility of the immediate recorded vehicles around the target vehicle based using the CUSUM statistic.
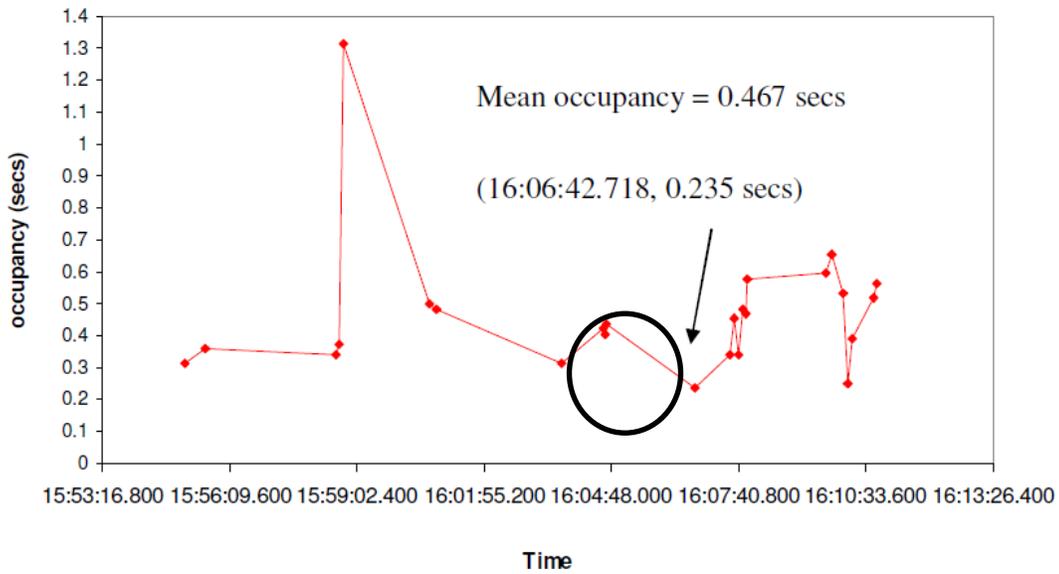
Figure 3.5: Occupancy for detector 4 during red phase and identifying vehicle 1.

To extract further information about the arrival of the proposed target vehicle at detector station 4 it was found that when the vehicle arrived at the detector, the signal for the phase had been red for 3.698 secs. We looked at similar occupancy plots for detector station 3 during the red phase and could not find any potential vehicle with a similar low occupancy time. Next, an effort was made to determine if a change in traffic conditions could be detected after the incident. For this purpose, the occupancy of the detectors 3 and 4 corresponding to the green phase approximately around the time of crash was investigated. Our hypothesis states, given that a congestion-inducing incident happened, that the traffic states (in our study, occupancy) would be significantly different from the traffic states under normal traffic conditions. The occupancy pattern for detector 4 during the green phase around the reported crash time, i.e., 16:07 was plotted, as shown in Figure 3.6. The plots clearly indicate a distinct change in the pattern of occupancies at approximately 16:06. The mean green phase occupancies before 16:06 was found to be 0.343 secs compared to the much higher average high occupancy of 0.782 secs after 16:06. Figure 3.7 shows a similar pattern with a gradual increase in occupancy at detector 3 during the green phase following the incident.

14
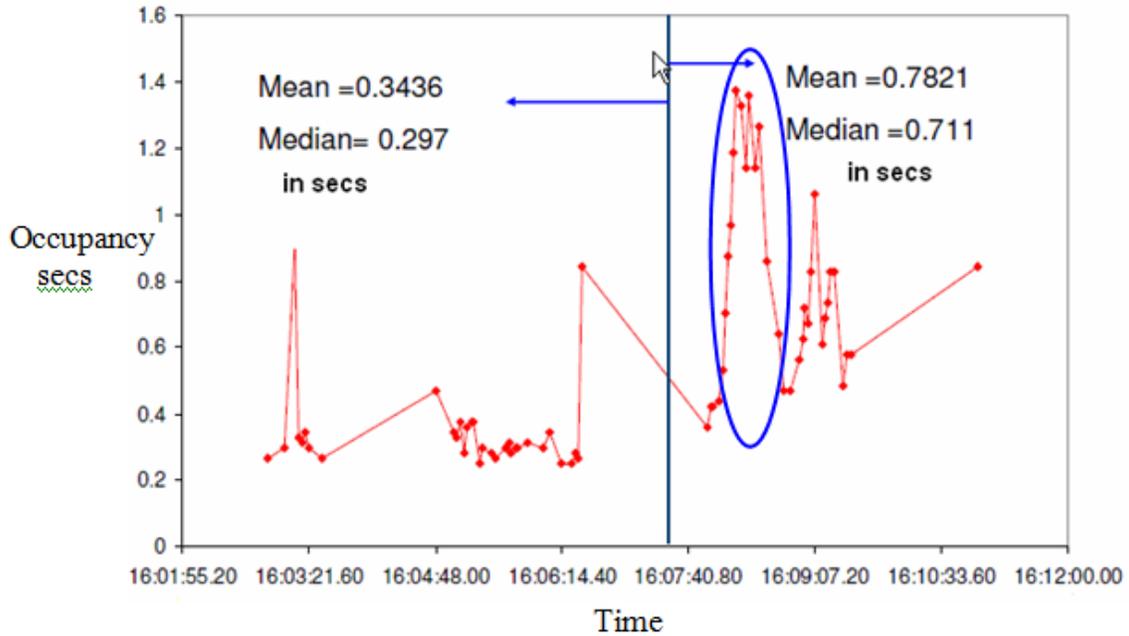
Figure 3.6: Post-incident occupancy pattern at detector 4 during green phase.
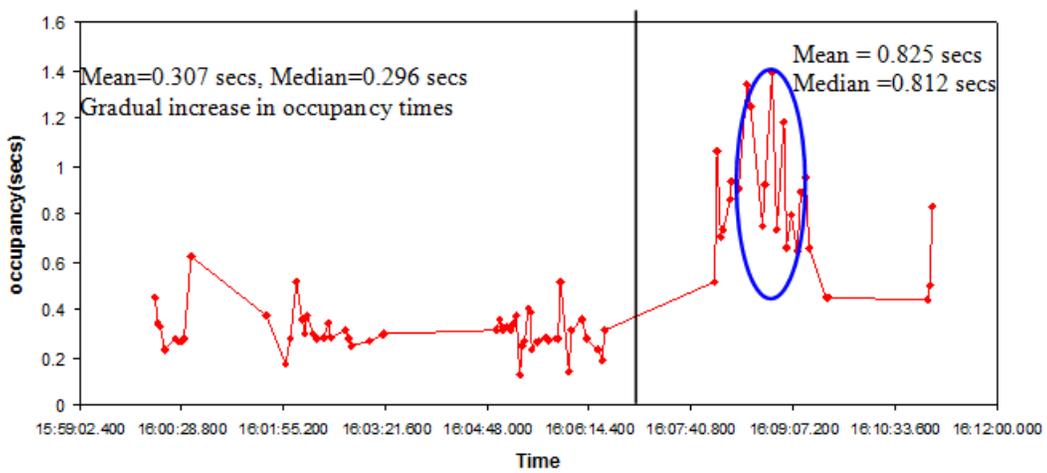
Figure 3.7: Post-incident occupancy pattern at detector 3 during green phase.

The CUSUM statistic, introduced in the previous section, was used to detect deviations from the general pattern in lane occupancies before and after the incident, during the green phase. Figure 3.8 shows the CUSUM plot along with the occupancy for detector 3. The figure detects a change in the occupancy trend after 16:06:27.078.

The initial variation in the CUSUM statistic is due to the randomness usually observed in any arterial traffic system. The prior information about the mean occupancy was calculated based on occupancy data from previous time periods under normal traffic conditions. One of the interesting facts about CUSUM, which can be observed from the above plot, is that once the

15

process deviates from its general trend due to the incident, CUSUM increases very rapidly and also monotonically. This monotonic behavior of CUSUM is quite effective from a detection algorithm point of view. For example, if we want to write an algorithm based on the CUSUM statistic to detect an incident, the only thing we have to find is the point where this monotonic behavior begins, which is relatively easy to implement.
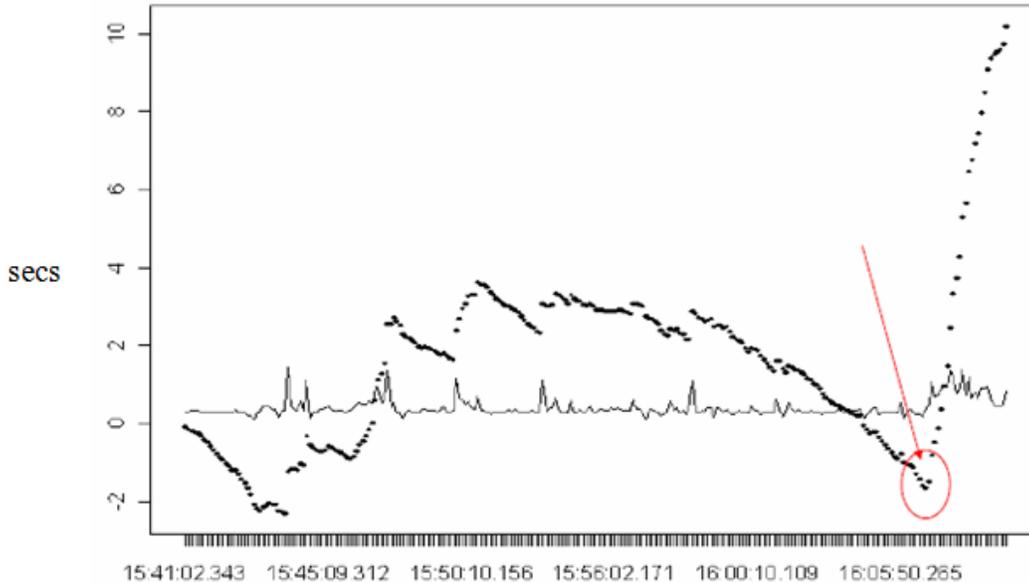
Figure 3.8: CUSUM statistic from occupancy data at detector 3.

The CUSUM statistic plot as shown in Figure 3.9 suggested the change in the pattern of the occupancy during green time had occurred after 16:06. And when we looked at Figure 3.6, the occupancy of the vehicle just before the potential target vehicle (with occupancy 0.235 secs) was 0.438 at 16:04:42. Since the effective length of the vehicle can be determined from the vehicle's make and model given on the crash report (21.68 feet), the point estimate of the speed of the vehicle could be computed as 49.49 feet/sec and hence consequently the estimated time to arrival to the potential conflict point (which is about 435 feet away from the detector location) would be 16:04:42+ (435.34/49.49 secs) =16:04:50.79, which clearly does not conform with the CUSUM result. Also, if we look at the occupancy data observed for the vehicle just recorded following the identified vehicle, we can see from Figure 3.6 that the occupancy was measured as 0.343 secs at 16:07:29.953. Similarly, the estimated speed of the vehicle would be 63.20 feet/sec and the estimated time to arrive at the potential conflict point is 16:07:36.84. However, a very high occupancy time (around 170 secs) for detector 8 with green phase on SB Winnetka Ave was recorded at 16:06:52.906, indicating an incident occurring before that time. Hence, based on the results from the CUSUM statistics, the estimated time of crash is between 16:06:27.078 and 16:06:52.906. Therefore, the most probable vehicle which was involved in the crash from Figure 3.6 was identified with occupancy of 0.235 secs at 16:06:42.718.

One key point to be noted here is that, although we were able to identify a change in the general occupancy pattern based on CUSUM statistic, it is not certain at this point whether this change in

16

the pattern is due to incident or to recurrent congestion, such as spill back from a downstream link. To distinguish between incident congestion and recurrent congestion it was decided to check the occupancy pattern of the main-line (i.e., TH 55 EB) detectors at the downstream link, located approximately 460 feet downstream of the crash location around 16:06, as suggested in the study by Gall and Hall (*19*). If a low occupancy pattern is found, indicating high speed, this would suggest no spillback phenomenon and hence the congestion at the upstream link could be attributed to the incident. The observed low mean occupancy time for the downstream detectors suggested that there was no spillback phenomenon. Hence, the congestion observed upstream at Winnetka Ave can be attributed to the incident.

The next step is to identify unit 2, the vehicle which was SB on Winnetka Ave during the green phase. The crash report indicated that unit 2 was a 2002 Buick LeSabre. The speed estimate from a single detector is given by

$$\text{Speed} = \text{Effective length} / \text{Occupancy time}, \tag{4}$$

where, Effective length = sum of the vehicle and detector length. For unit 2 effective length = (16.6+6) feet, where length of the detector is 6 feet.

The point of collision based on the approach directions of the two vehicles was found be 123.86 feet from detector stations 7 and 8. Occupancy data for detectors 7 and 8 were extracted corresponding to the green phase around 16:07, the time of collision occurrence as recorded in the crash report. Unit 2 was then identified as the most probable vehicle recorded at detector 7 or 8 with an estimated speed, based on the occupancy time, to arrive at the collision point before the upper bound identified above. That is, unit 2 was the vehicle recorded at detector 7 at 16:06:45.937 with occupancy time of 1.438 secs. Hence, an estimated arrival time of unit 2 at the collision point is 16:06:45.937 + (123.86/22.6/1.438) = 16:06:53.818. This estimated arrival time for unit 2 was calculated based on constant speed; however, the signal phase data suggest that the unit 2 was discharging during green phase and may have some acceleration and so arrived at the collision point before the estimated time, but within the bound established above.

3.5   Crash Reconstruction

Drawing from Baker's (*20*) notion of crash reconstruction of determining how a crash occurred, we next tried to address a question that, given the initial speed estimates and locations (both space and time) from detector data, what could be learned about the behavior of the drivers involved in the crash. To answer this, trajectories of the two approaching vehicles were modeled by numerically solving a system of ordinary differential equations.

Vehicles were initially assumed to travel at uniform speed along straight lines. Given initial speed and location, at each time step of 0.01 secs the vehicle's speed and locations were updated using a simple Euler's formula, and each time the separation distance (distance between the centre of two vehicles), *d* was calculated to check whether two vehicles had collided or not. Figure 3.9 shows the plot of separation distance between the two vehicles assuming constant speeds, and it is clear that had the drivers not taken any action, the crash would not have occurred. Also, it was observed from the simulation that unit 1 arrived at the collision point

17

much earlier than unit 2, suggesting that for crash to happen, unit 1 had to decelerate and/or unit 2 had to accelerate.
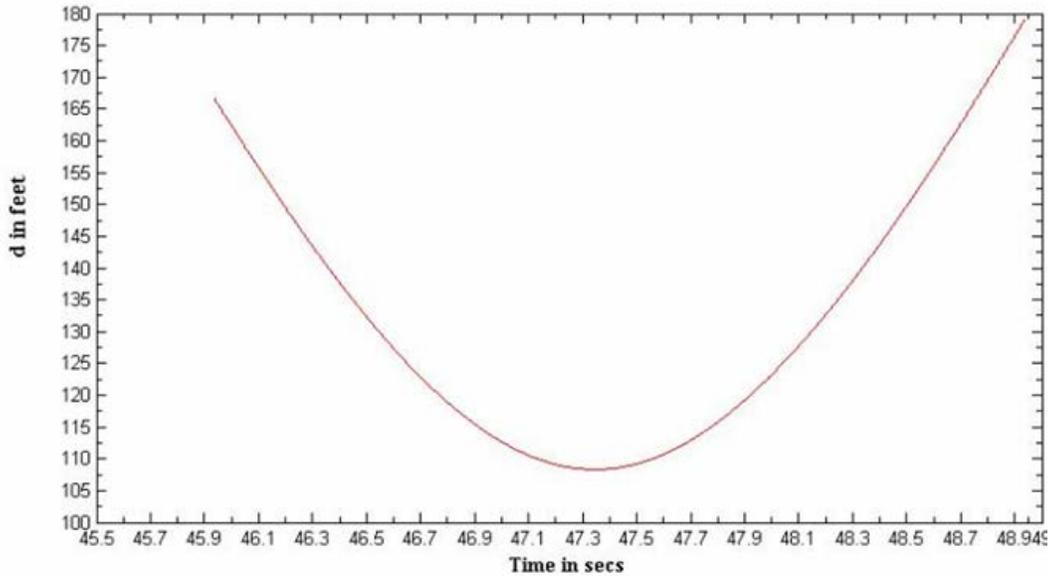
Figure 3.9: Separation distance based on initial speed estimates.

This initial simulation indicates that one or both drivers needed to decelerate in order for the collision to take place, so that a complete description of this event would require the reactions times and deceleration rates of the involved drivers. Unfortunately, long reaction times can be compensated by more extreme braking rates, so these quantities are underdetermined from the SMART SIGNAL data. Traditionally, accident reconstruction involves estimation of the speeds of the vehicles at the point of collision from post-collision information such as final resting positions for the vehicles, skid marks and damage. Such detailed information was not available for this particular case. Hence here, we would like to illustrate how, had such post-collision details been available, it would be possible to estimate driver behaviors contributing to the crash. For this purpose, a plausible hypothetical post-collision scenario was added, where the following assumptions on the final position of the vehicles were made.

1. The crash report suggested that unit 1, after colliding with unit 2, rolled down to eventually collide with unit 3 which was stopped at red light NB Winnetka (approximately 60 feet from collision point). Based on the report it was assumed that subsequent impact with unit 3 was minimal, i.e., unit 1 came to a complete stop just at the point of collision with unit 3.
2. Unit 2 was assumed to skid after it collided with unit 1 with a post-impact skid mark suggesting that unit 2 came to a complete stop somewhere between 5 to 20 feet from collision point, with an angle of departure (Φ) for unit 2 as 30°. Figure 3.10 illustrates the hypothetical collision scenario.

The post-impact speeds of the two vehicles were evaluated based on the stopping distance formula,

18

$$V_i = \sqrt{2 \times f_i \times g \times s_i}, \quad \text{where } g = 32 \text{ feet/sec}^2, \ i = 1, 2 \tag{5}$$

For unit 1, f_1 (rolling resistance) was assumed to be 0.15 and $s_1 = 60$ feet and for unit 2, f_2 (skidding resistance) was assumed to be 0.7 and s_2 between 5 and 20 feet,

$$\text{i.e., } \sqrt{2 \times 0.7 \times 32 \times 5} < V_2 < \sqrt{2 \times 0.7 \times 32 \times 20} \quad \text{(in feet/sec)} \tag{6}$$
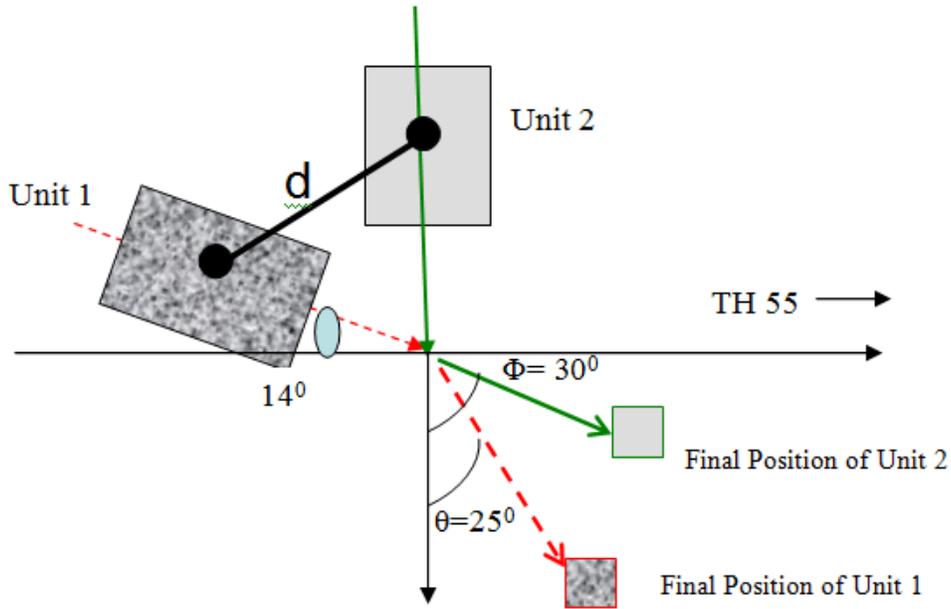
Figure 3.10: Collision diagram under hypothetical scenario.

The next step was to evaluate the pre-impact speeds of the vehicles based on point-mass collision theory (*21*). Let v_1 *and* v_2 denote the pre-impact speed, and m_1 and m_2 be the corresponding masses of unit 1 and unit 2 respectively. Then the conservation of momentum equation in the x and y directions are given by,

$$m_1 \times v_1 \times Cos(14°) = m_1 \times V_1 \times Cos(65°) + m_2 \times V_2 \times Cos(60°) \tag{7}$$

$$m_1 \times v_1 \times Sin(14°) + m_2 \times v_2 = m_1 \times V_1 \times Cos(25°) + m_2 \times V_2 \times Cos(30°) \tag{8}$$

Substituting the values for V_1 and V_2 into (7) and (8), bounds for v_1 and v_2 were established as

$$16.245 < v_1 < 24 \quad and \quad 26.7 < v_2 < 39 \tag{9}$$

Finally, the collision set was established as

$$\{(d, v_1, v_2) : d < d_{crit}, 16.245 < v_1 < 24.0 \text{ and } 26.7 < v_2 < 39.0\}, \tag{10}$$

19

where, d_{crit} is the crash closeness threshold.

Perhaps one input that needs more explanation is the collision closeness threshold (d_{crit}) which was set to be 14.7 feet in the simulation model based on the dimension of the vehicles involved in the crash and their angle of approach (76°). The simulation model used this threshold value to decide whether a crash has occurred or not and hence it is desirable to make sure that the value was reasonable.

A three-parameter model was assumed for this study, (a) unit 2 acceleration (*acc2* in feet/sec^2), (b) unit 1 deceleration (*dcc1* in feet/sec^2) and (c) for driver 1, perception-reaction time (*rt* in secs) which was defined as the time elapsed between unit 1 arriving at the detector 4 and the initiation of braking. Our objective was to find the joint distribution of the three parameters (*acc2, dcc1* and *rt)* conditioning on the crash occurrence. A Monte Carlo simulation technique based on rejection sampling was adopted to sample the desired distribution (*22*). The initial distribution for *acc2* was chosen as Uniform between 0 to 10 feet/sec^2, where 10 feet/sec^2 was deemed to be a reasonable upper bound for acceleration. *dcc1* was sampled from Uniform between -5 feet/sec^2 and -21 feet/sec^2. The upper bound for *dcc1* was chosen from the fact that for the given initial speed estimates and location of the vehicles any deceleration weaker than -5 feet/sec^2 would not result in a collision. The lower bound was chosen based on the study by Fambro et al. (*23*), where the mean emergency braking rate was taken as -.65g. For the perception-reaction time of unit 1, initial samples were drawn from Uniform [0, (45.937-42.718)]. The upper bound was chosen on the basis of the fact that if unit 1 had a perception-reaction time longer than 3.219 secs, the crash would not have occurred even for the strongest braking rate. Once the three parameters were sampled, for each time step, the location and speed of the two vehicles were updated using the simple Euler's method as mentioned before, and the separation distance (*d* in feet) was computed. If the condition for crash in equation 10 was satisfied then a collision was recorded. Monte Carlo Simulation was computed in *R* statistical software (*24*). Table 3.1 shows the descriptive statistics of the parameters obtained from the simulation. The point estimate of the collision time from the simulation was found to be 16:06:50.481, which is within the bound suggested by CUSUM statistics.

Table 3.1: Sample Statistics from Monte Carlo Simulation.

Parameters	Mean	Standard deviation
acc2(feet/sec^2)	4.65	0.529
dcc1(feet/sec^2)	-10.157	0.632
rt (secs)	0.337	0.251

3.6   Mode Estimation

Traffic accident reconstruction can be treated as an effort to determine how an accident has occurred, and often involves uncertainties regarding the behavior of the drivers involved in the crash. One way to identify a probable cause is to estimate the mode of the posterior distribution

of the parameters obtained from Monte Carlo simulation. The mode can be defined as the most probable combination of factors that resulted in the particular collision.

The study of mode estimation from a sample can be found in Parzen (*25*), where the author defined a family of kernel estimators for a distribution *f(x)* and the mode estimates were shown to be asymptotically normal and consistent. Later on, Silverman *(29),* Wand and Jones (*26*) explored more in this area of kernel density estimation. Initially the focus was on univariate distributions but later the theory was extended to the multivariate case. The next section gives a brief introduction to Kernel density estimation.

3.6.1 *Kernel Density Estimation*

Suppose a s-variate random sample of size *n*, X_1, X_2, X_3,......., X_n is drawn from an unknown distribution *f*. Then the Kernel density estimate of *f* is given by,

$$\hat{f}(x:H) = n^{-1}\sum_{i=1}^{n}K_H(x-X_i), \tag{11}$$

where, $x = (x_1,x_2,.........,x_s)^T$. If *K(x)* is the kernel, which is a probability density and *H* is bandwidth matrix, which is a symmetric and positive definite matrix, then

$$K_H(x) =| H |^{-1/2} K(H^{-1/2}x) \tag{12}$$

The choice of K is not as important as the choice of bandwidth matrix, as the amount of kernel smoothing is controlled by the bandwidth matrix (*27*). A substantial literature is available on optimal bandwidth selection (*28, 29*). Here, without going into the details we highlight the key notion underlying the bandwidth selection procedure. The form of *K(x)* is taken as standard normal in this study, i.e.

$$K(x) = (2\Pi)^{-s/2} \exp(-\frac{1}{2}x^T x) \tag{13}$$

The common measure of performance of $\hat{f}$ used in the literature is Mean Integrated Squared Error (MISE),

$$MISE(H) = E\int_{R^s}\{\hat{f}(x,H) - f(x)\}^2 dx \tag{14}$$

The optimal bandwidth is given by

$$H_{opt} = \arg\min_{H} MISE(H) \tag{15}$$

over all possible symmetric positive definite *s* by *s* matrices.

Two separate optimal bandwidth selectors were used in this study: A Plug-in bandwidth selector (with pre-sphering of data) and a Least squares cross validation based on leave-one-out estimator. Details are given in the literature mentioned above. Bandwidth results from the two selectors when compared were found similar, so the *H* matrix evaluated by Plug-in bandwidth selector (based on sample covariance matrix) was used for further analysis. The "Ks" package (*30*) for multivariate kernel smoothing was used in the R statistical language.

Evaluating the mode of a multivariate distribution is itself a challenging task. However, some progress has been made in the past (*31, 32*). One of the simple techniques, adopted in this study, was recommended by Abraham et al. (*33*). The mode is estimated by maximizing the kernel estimate over the set of sample values. For large enough samples and under certain mild conditions this estimate is found to be consistent and converges almost surely to the true value. More formally, the estimated mode is defined as follows.

If S_n denotes the set of sample points $\{X_1, X_2,..., X_n\}$ then the estimated mode α_n is defined as

$$\alpha_n \in \{x \in S_n : f_n(x) = \max_{1 \le i \le n} f_n(X_i)\}, \tag{16}$$

where, $f_n(X_i)$ is the Kernel density evaluated at sample point X_i.

To compare with Abraham's estimate for mode, an optimization routine in *R* is called, which returns a vector $\{acc2, dcc1 \text{ and } rt\}$ that maximizes the Gaussian kernel density estimate (based on the evaluated bandwidth matrix) according to equation 11. Abraham's estimates are found to be $\{acc2=4.935 \text{ feet/sec}^2, dcc1= -9.978 \text{ feet/sec}^2, rt= 0.262 \text{ secs}\}$, where the mode estimates using optimization routine were $\{acc2=4.76 \text{ feet/sec}^2, dcc1= -9.94 \text{ feet/sec}^2, rt= 0.245 \text{ secs}\}$. Both the estimates suggested most likely the driver violating the traffic signal has to brake at a very strong rate around -10 feet/sec^{2-}, but still could not avoid the violation and consequently the crash with the other vehicle which was accelerating at a rate of 5 feet/sec^2.

Chapter 4.   Vehicle Pedestrian Crash Risk Assessment

4.1   Vehicle Pedestrian Crash Reduction

Design of intersection signal timing, in particular traffic signal-change interval, i.e. yellow indication followed by all-red phase plays a critical role in preventing vehicle pedestrian collision. A recent study by Retting et al.(*34*) conducted at standard four legged signalized intersections located in New York state, reported a significant 37% relative reduction in the pedestrian and bicycle crashes by modifying the traffic signal change interval as per as Institute of Transportation Engineers(ITE) recommendations. Since these modifications generally led to increases in the clearance intervals, the result found by Retting et al. has implications regarding the mechanism by which at least some vehicle-pedestrian collisions occur. A plausible explanation is that a non-negligible fraction of drivers approached the study intersections in such a way that they would not clear the intersection when pedestrians received a walk signal.. The effect of longer all-red intervals would then be to delay pedestrian entry into the intersection until these problem vehicles had cleared the intersections.

Verifying whether or not this explanation is in fact accurate requires more detailed information regarding vehicle-pedestrian crashes than is usually available in traditional crash records. Questions regarding such individual crash causation have been addressed several times in the past from accident reconstruction point of view. For example, researchers at the University of Adelaide's Center for Automotive Safety Research (formerly the Road Accident Research Unit), used standard deterministic reconstruction method to investigate the causal impact of speeding in fatal vehicle-pedestrian crashes (*35*). Another related effort could be seen in the study of speed limit reduction on pedestrian accidents by Pasanen and Salmivaara (*36*). Video recording of pedestrian collisions were made at an intersection near the downtown of Helsinki. Sequence of events such as pedestrian entering in the path of the vehicle, but the driver of the vehicle unable to stop before hitting the pedestrian, were extracted from the video data and combining with simple equation of motions, causal model for collision event was developed to answer the relevant research question of whether collision would have been avoided if the vehicle had not been exceeding the recommended speed limit. Although Pasanen and Salmivaara's study was interesting in its exhibiting the mechanism behind many pedestrian-vehicle conflicts, the luxury of video recording will not be available in most practical scenarios. In the absence of such elaborate data collection technology traditional detector data along with the signal status information could be used to analyze any probable vehicle pedestrian conflicts. More particularly, given initial estimates of a vehicle's position and speed from detector data, together with the beginning and ending times of signal intervals, the effectiveness of all-red phase in preventing vehicle-pedestrian conflicts could be evaluated.

4.2   Data Source for Vehicle Pedestrian Conflict Analysis

As mentioned in the previous chapter, currently the SMART SIGNAL system has been installed at various intersections on Trunk Highway (TH) 55 and on Eden Prairie Centre Drive, in the Twin Cities region of Minnesota. However, not all intersections being suitable for our study (for example, intersections with large setback distance to target lane, i.e. distance from the point where pedestrian waiting for traffic to the target lane), we chose three intersections with specific

approaches. Typically, data recorded between 10:30 am to 5 pm were analyzed for each intersection. Most approaches had 3 second yellow intervals followed by 1.5 to 2 second all-red intervals. The speed limit of all local arterials was 30 mph, except for one approach on a trunk highway, where the speed limit was 55 mph. Table 4.1 below lists some relevant characteristics of the sites chosen for our study.

Table 4.1: Characteristics of the sites chosen for our study.

Site	Approach	length of intersection (feet)	speed limit (mph)	no. of cycles	set back distance (feet)	all-red phase (secs)	# of* events
1(a)	NB	153.5	30	160	7	2.5	7
1(b)	WB	169	55	160	7	1.5	28
2(a)	SB	146	30	162	8	3.0	10
2(b)	NB	135	30	162	8	3.0	16
3	EB	140	30	305	7	3.0	32

* Number of events indicates those events (potential vehicle/pedestrian conflicts) that were investigated, based on the vehicle's entry into the intersection after the beginning of yellow phase.

Figures 4.1, 4.2 and 4.3 provide the layout of the three intersections used for the vehicle pedestrian analysis. The highlighted detectors were added later to the intersections.
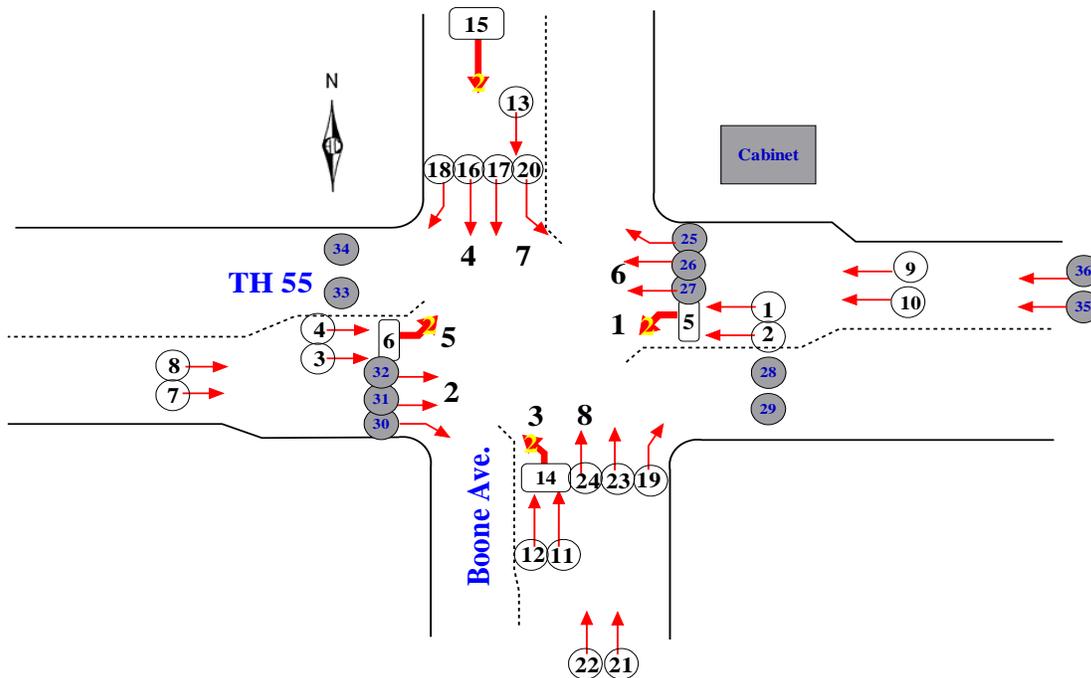
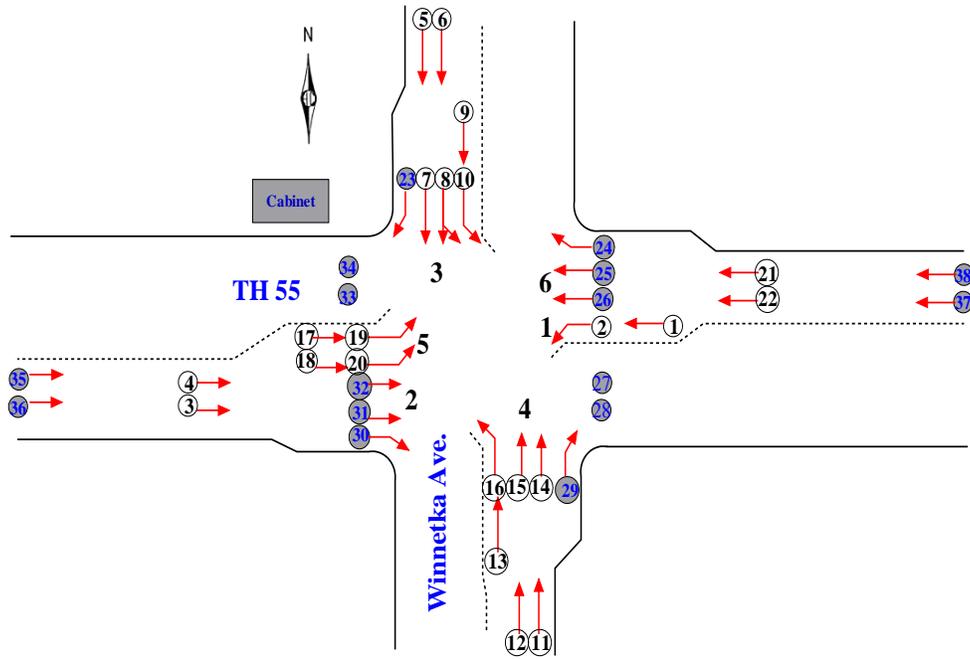**Figure 4.1: Detector layout of the Boone Ave /TH 55 intersection.**

Figure 4.2: Detector layout of the Winnetka Ave /TH 55 intersection.
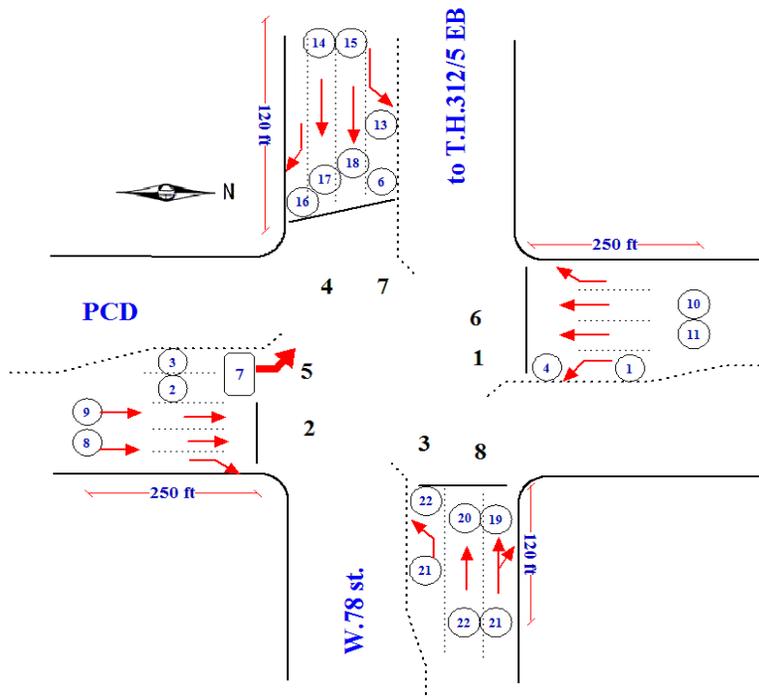

Figure 4.3: Detector layout of the Prairie Centre Drive.

4.3   Vehicle Pedestrian Conflict Analysis

As indicated earlier, our study has two inter-related objectives: to demonstrate how SMART SIGNAL data could support reconstruction of vehicle-pedestrian crashes, and to assess the degree to which all-red intervals can prevent potential crashes To accomplish the aforementioned objectives we first assume that the traffic signals have no all-red phase or clearance interval, i.e. as soon as the yellow phase is over a conflicting pedestrian walk sign begins. Suppose then, a vehicle enters the intersection when the signal turns yellow, but could not cross the intersection before the yellow phase was over. A potential conflict then arises if a typical pedestrian initiates a movement across the vehicle's path. For a particular intersection, given the vehicle's initial speed, time when it enters the intersection, and the typical pedestrian's behavior (including his/her reaction time and speed), trajectories of the two conflicting movements could be modeled to see whether or not a potential conflict turns to a crash. Such treatment of vehicle-pedestrian conflicts first appeared in Mayne's study (*37*), where information about vehicle's width and its location within the lane was also included in the collision model. However, in practical safety studies such detailed information is almost never to be had. In order to account for such lack of details, we would look at an alternative approach motivated by recent use of crash surrogate measure in traffic safety analysis (*38*). Actual observable vehicle/pedestrian crashes are rare and as a consequence it is quite difficult to assess the safety effect of an implemented countermeasure. In this study we define an event as a "Critical Event" (CE) if the vehicle reaches the conflicting point when the pedestrian is still present within the lane. For convenience, in this study the term "critical event" and crash/collision will be used interchangeably. The entire lane width is a critical zone which accounted for the uncertainty of location and physical attributes (width) of the vehicle. Figure 4.4 below shows a schematic description of a critical event.
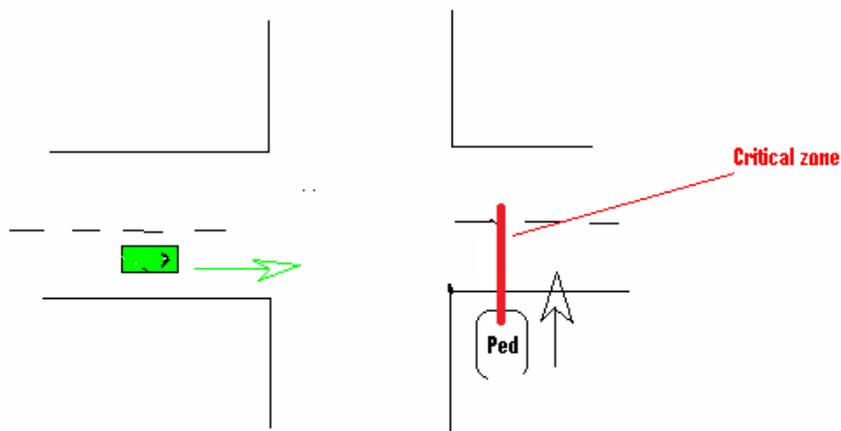
Figure 4.4: Schematic description of a Critical Event.

More formally, suppose the pedestrian takes time t_1 to enter the target lane (i.e. the lane on which vehicle is approaching), time t_2 to cross the lane including the initial reaction phase (r_p), and the vehicle reaches the critical zone at time t_3. Then the critical event indicator, Y is given by,

$$\begin{cases} Y = 1, \textit{if } t_1 \leq t_3 \leq t_2 \\ \quad 0, \quad \textit{otherwise.} \end{cases} \qquad (17)$$

To compute t_1 and t_2 we need to have knowledge of pedestrian speed (v_p) and his/her reaction time (r_p) at the beginning of the walking phase. For t_3, information about vehicles speed and its arrival at the upstream point of the intersection is required. Assume the speed and the location of the vehicle within the intersection when the pedestrian began to walk is available from detector data. If the remaining collision variables (r_p, v_p) are assumed to be random outcomes from certain underlying population with a known probability distribution, then computing the probability assigned to a subset of collision variables would give the probability that a randomly selected vehicle-pedestrian conflict leads to a critical event. Applications of such an approach based on structural knowledge of the vehicle-pedestrian collision have been described in detail in Davis (*39*).

To verify whether the potential conflict would result in a critical event, we need to identify a plausible underlying distribution of pedestrian's reaction phase (r_p) and speed (v_p). If we know the probability distribution, $F(r_p, v_p)$ then probability of a critical event could in principle be computed as

$$\Pr[\textit{Critical Event}] = \Pr[t_1 \leq t_3 \leq t_2] = \int I(t_1 \leq t_3 \leq t_2) dF(r_p, v_p) \qquad (18)$$

where, $I(.)$ is the indicator function. Rather than attempting a closed-form solution we will use Monte Carlo simulation method to evaluate the above interval, where random samples can be extracted from the underlying distribution of the collision variables, and simply evaluating the proportion of samples satisfying equation (17) would give an estimate of a probability of a critical event for a particular vehicle.

4.3.1 Braking Model

In the next step we would like to allow for the fact that the driver could take evasive action to avoid a potential conflict. For simplicity in our study we define braking as the driver's solo evasive action.

The braking model implemented in this study is an adaptation from the previous pedestrian collision model proposed by Davis (*40*), except the collision set here is more restrictive. Here is a brief description of the braking model used in this study.

Assume the car is traveling at an initial speed v_c when the pedestrian begins moving toward the street at a speed v_p. At the time the driver notices the pedestrian, the car is at a distance d_{red} from the critical zone, whereas the pedestrian is a distance s from the critical zone. In the absence of any more detailed information about when the driver notices the pedestrian, we assume the driver tends to notice the pedestrian as soon as pedestrian initiates a movement. Then

$$d_{red} = d_{init} - v_c \times t_{red} \qquad (19)$$

where, d_{init} is the intersection length (from near cross walk to far cross walk) and t_{red} is the time elapsed between signal turning red and the vehicle entering into the intersection. If it is possible for the car to pass the critical zone before the pedestrian without changing speed, this is what the driver does, otherwise after a perception reaction interval of rt, the driver brakes to a stop at a constant deceleration a. Conflict occurs if the car reaches the critical zone when the pedestrian is within the target lane.

An analytical formulation similar to Mayne (*37*) defines the critical set for d_{red} by the following equations.

$\{d_{red} : d_v \le d_{red} \le d_{11}\}$ which is equivalent to $\{t_1 \le t_3 \le t_2\}$,

$where\ d_v = v_c \times t_1,$

$If, t_{11} = \max(0, rt + \dfrac{v}{a} - t_2),$ then $d_{11} = d_s - \dfrac{1}{2}at_{11}^2$ (20)

Here stopping distance, $d_s = v_c \times rt + \dfrac{v^2}{2a}$

$\therefore$ Probabiliy of critical event for a particular vehicle,

$\Pr(CE) = \int I(d_v \le d_{red} \le d_{11}) dF(a, rt, v_p, r_p)$

Under the braking model, a Critical Event could be avoided under 3 conditions:

- Condition 1: The vehicle could able to stop before reaching the critical zone.
- Condition 2: The vehicle passes the critical zone without braking before the pedestrian arrives
- Condition 3: The vehicle is able to slow down enough to reach the critical zone after the pedestrian's exit.

Figure 4.5 provides a schematic diagram for the critical set of d_{red} indicating the three aforementioned conditions.
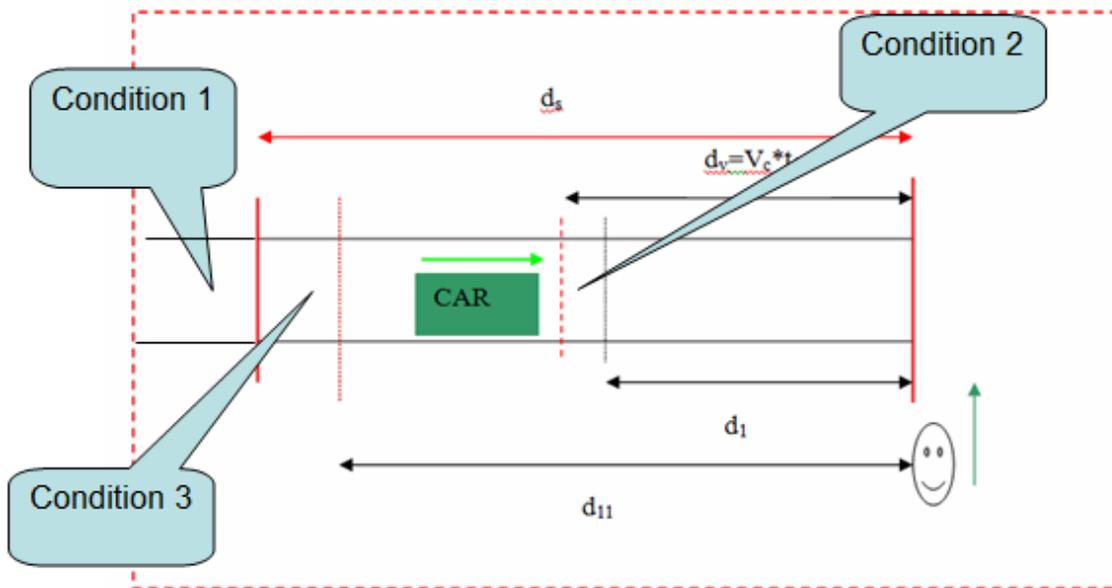
Figure 4.5: Schematic diagram for the critical set of d_{red} under braking model.

To compute the probability of a critical event we have to determine a plausible probability distribution for a, rt, $v_p$ and r_p. Without any empirical evidence supporting the nature of the aforementioned variables from the chosen sites in this study, distributions of the variables were borrowed from the past literature.

Evidence regarding the probability distributions for a driver's reaction times and braking rate during emergency stops have been reported by Koppa et al. (*41*). There, the drivers reaction time appeared to follow lognormal distribution with mean=1.07 secs and standard deviation of 0.248 secs. The emergency braking rate was also modeled as a lognormal variable with mean=0.63g and standard deviation of 0.08g, where g=32 feet/sec^2. Although previous studies, such as by Fugger et al.(*42*) suggested that pedestrian speed and reaction time can vary depending upon age and gender and nature of the pedestrian, in our study for simplicity, pedestrian speed was assumed to be uniform random variable between 2.6 feet/sec and 5.9 feet/sec (*43*). Reaction time for pedestrian was assumed to be uniformly distributed with lower and upper limit as 0.2 and 1.5 seconds.

As a byproduct of our study we demonstrate here when the underlying mechanisms governing the vehicle/pedestrian collision process is known, how using the initial speed estimate and vehicle location within the intersection from the detector data, reconstruction of a critical event can be done. We would take one event where the vehicle entered the intersection after the yellow phase at a speed of 42.63 feet/sec. Now, suppose a city engineer was asked a question about how severe a crash would have happened if the all-red phase for this approach was removed. One way to answer this question is to adopt simulation-based accident reconstruction. From the detector and signal data a location for the vehicle within the intersection when the driver observes the

pedestrian can be easily extracted (in this case around 93 feet from the conflict zone). Then using Baye's theorem, a posterior distribution of the collision input variables, given that a collision occurs can be computed. Once the updated distribution is obtained, simply substituting mean values back into the collision model would provide an estimate of the vehicle/pedestrian conflict. Figure 4.6 shows the reconstruction of the vehicle pedestrian conflict with an estimate of final impact speed of the vehicle as measure of the severity of the collision (*44*).
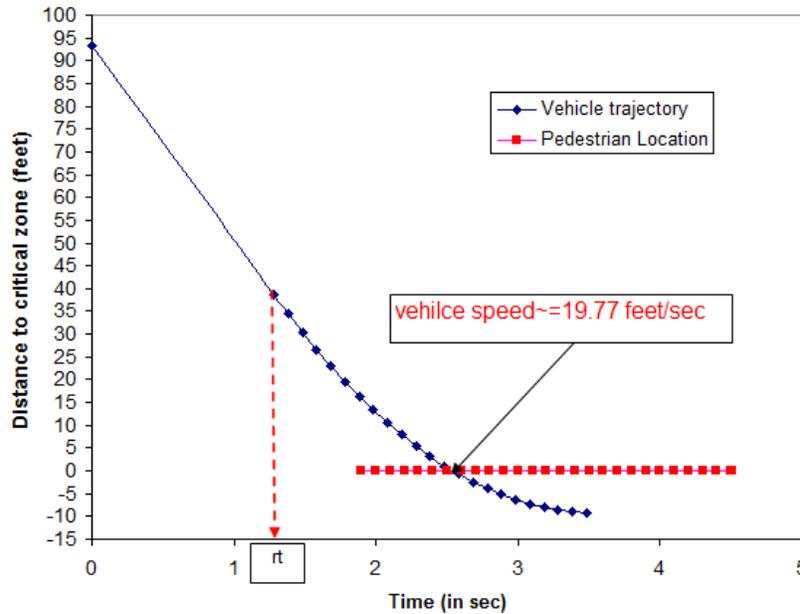
Figure 4.6: Reconstruction of vehicle/pedestrian Critical Event.

4.4   Alternate Estimation of a Crash-Reduction Factor

In the introductory section we indicated that an alternate approach to assessing crash-reduction effects could possibly be based on accident reconstruction and simulation. Adopting such an approach we would likely to answer the following question:

1. What is the probability of a critical event being avoided by driver's evasive action which would have occurred without all-red phase?
2. What is the probability of a critical event being avoided by introducing all-red phase, which would have occurred without all-red phase?
3. What proportion of critical events would be avoided by all-red phase which would not have been circumvented by driver's evasive action?

To answer the above questions a general concept, called "Probability of Necessity" (PN) proposed by Pearl (*45*) will be used. First to apply Pearl's concept we need to specify

- A set of exogenous variables (in our study such as r_p, a, v_p, r_t).
- A set of endogenous variables (such as t_1, t_2, t_3).
- A set of structural equations describing the dependency of each endogenous variable on other variables within the model.

- Finally, a probability distribution over the exogenous variables (i.e. prior distributions as mentioned in earlier sections)

As it can be seen all the four requirements are satisfied in our study. Before we proceed with the concept of Probability of Necessity (PN) let us introduce several notations relevant to the concept.

Define Z as a control variable, which identifies three situations of interest:

Z=00, if no braking and no all-red phase (AR)

=10, if braking and no AR
=01 if no braking and AR

$$\text{Now, let } Y = Y_{00} \text{ if } Z=00 \tag{21}$$

$= Y_{10}$ if Z=10
$= Y_{01}$ if Z=01

From earlier section we can write the structural relation between Y (critical event indicator) and the other variables as below:

$$
\begin{aligned}
Y_{00} &= 1, if\ v_c \times t_1 \le d_{red} \le v_c \times t_2 \\
&= 0, otherwise \\
Y_{10} &= 1, if\ d_1 \le d_{red} \le d_{11} \\
&= 0, otherwise \\
Y_{01} &= 1, if\ v_c \times t_1 \le d_{red}^1 \le v_c \times t_2 \\
&= 0, otherwise \\
where, & d_{red}^1 = d_{init} - v_c * (ar + t_{red}),
\end{aligned}
\tag{22}
$$

Then Probability of Necessity of braking (PN_b) and all-red phase (PN_{ar}) can be given by the following expressions,

$$
\begin{aligned}
PN_b &= Pr\ (Y_{10}=0|Y=1,\ Z=00) \\
PN_{ar} &= Pr\ (Y_{01}=0|Y=1,\ Z=00)
\end{aligned}
\tag{23}
$$

In other words, for example, probability of necessity for braking (PN_b) is the probability of the crash being avoided by driver's braking effort which would have occurred otherwise had the driver not taken the evasive action. The causal dependency between Y, Z, v_c, and U (background variables, such as r_p, a, v_p, rt) is given by Figure 4.7. Here the initial speed of the vehicle, v_c is assumed to be known from the detector data.
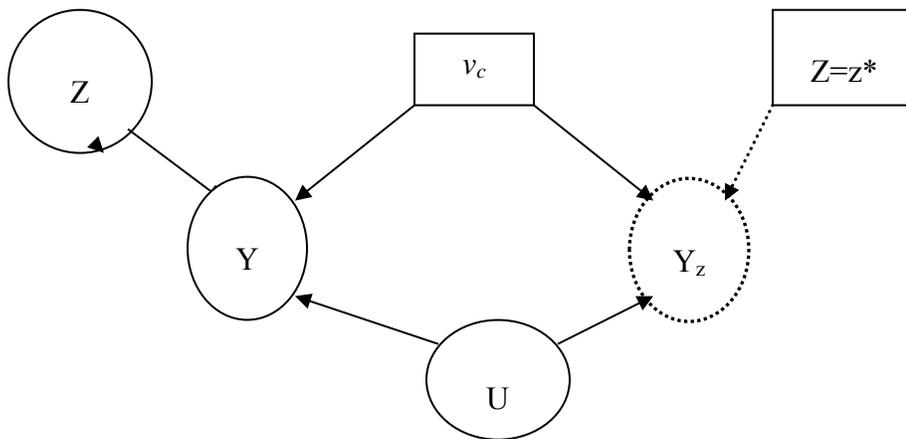
Figure 4.7: Causal dependency between *Y*, *U*, *Z* and *Vc*.

To compute respective PNs following steps are required:

- Using Baye's Theorem compute posterior distribution of exogenous variables, given the critical event occurs under initial condition, i.e. when Z=00.
- Set the value of Z=01 or 10 corresponding to braking or all-red phase effect.
- Compute the respective PNs using the posterior distribution of exogenous variables computed from step 1.

All computations were done via Markov Chain Monte Carlo (MCMC) methods implemented in a software package, named WinBUGS (*46*).

4.5   Results for Vehicle-Pedestrian Crash Risk

PN for braking and all-red for each event (i.e. vehicles entering intersection after yellow phase) from the three chosen intersections with specific approaches were computed, and are plotted in the following graphs (see Figures 4.8-4.10). The x-axis represents PN for braking and y-axis indicates corresponding PN for AR. The 45° dotted line indicates the events where PN for braking is equal to PN for AR. An event lying below the dotted line suggests that introducing all-red phase would not have any superior effect over that of the driver taking evasive action to avoid the critical event. Some of the interesting findings from the plots are discussed below.

- Plots suggest that the all-red phase, in most of the cases, leads to a 100% crash reduction.
- Except one or two cases, the all-red phase has a superior impact in preventing critical events compared to the driver's braking effort.
- PN plot from site 1(b) WB approach (see Figure 4.8) indicates that a majority of the events have a very low crash-reduction effect for driver's braking effort. The reason for this is that the WB approach for site 1 has a high speed limit (55 mph), hence vehicles which entered the intersection during the yellow interval usually have high speed. Thus

32

as a consequence, even with strong braking a conflict with pedestrian could not be prevented.

- One of the events in Figure 4.9 for site 2(a) indicates very low PN (0.08) for AR compared to a PN value of 1.0 for braking. Further investigation of this particular event indicated that this approach has low speed limit (30 mph) and the intersection has a length of 146 feet. The vehicle entered the intersection just before the start of all-red phase with low speed, and hence was not able to cross the intersection before the all red phase was over. In contrast, nominal braking effort would have been sufficient for the vehicle to come to a complete stop before arriving to the conflict zone.

(a)   (b)
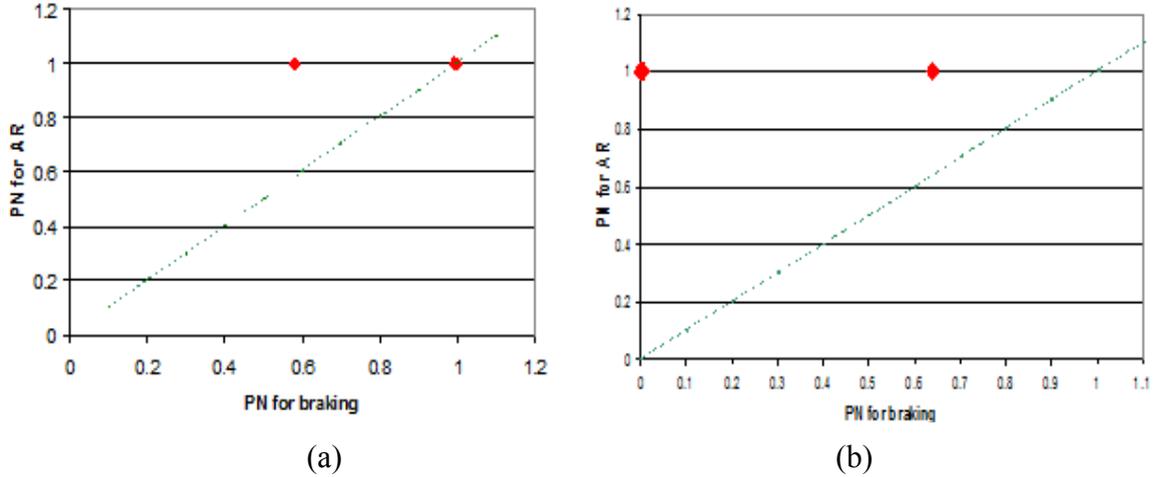
Figure 4.8: Comparing PN for braking and AR for site 1:(a) NB, (b) WB approach.

(a)   (b)
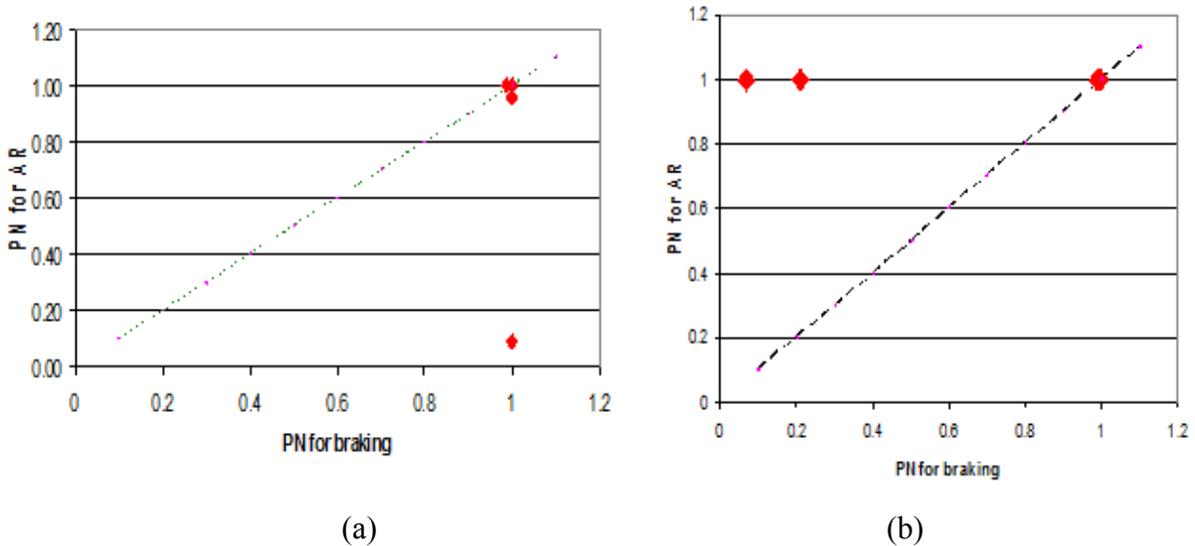
Figure 4.9: Comparing PN for braking and AR for site 2: (a) SB (b) NB approach.

33
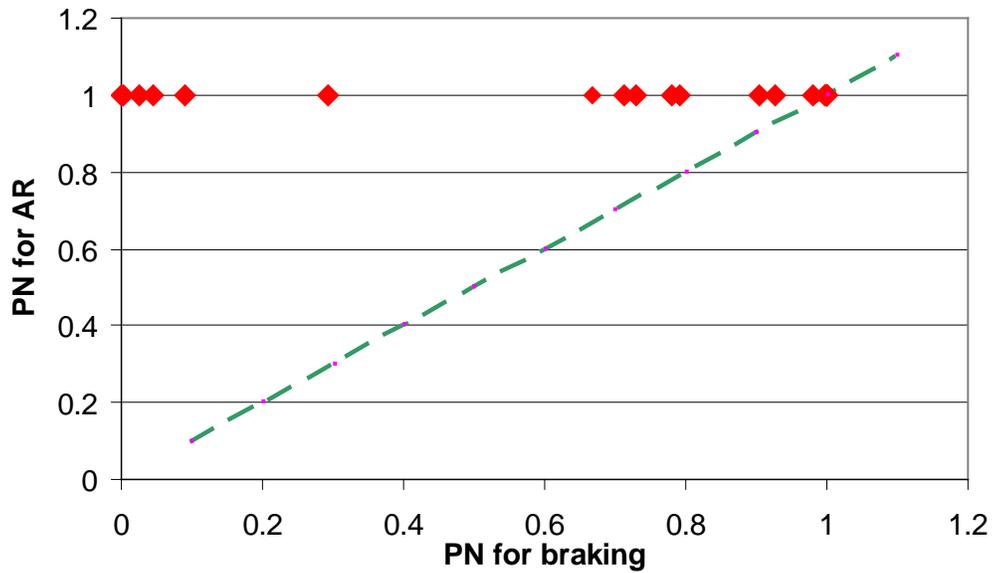
Figure 4.10: Comparing PN for braking and AR for site 3: EB approach.

To find the relative effectiveness of all-red intervals compared to driver's braking efforts, we define a term, $\delta = \dfrac{PN_{ar} - PN_b}{1 - PN_b}$ for each event where $PN_b \neq 1$. It was found in all events where there is a non-zero probability of a critical event with the driver's braking effort, given it would have occurred under the initial conditions, δ is very close to 1. This suggests that intervention of all-red phase was able to prevent almost all the critical events which would not have been avoided with driver's braking effort.

Chapter 5. Conclusions and Recommendations

5.1 Summary

Recent traffic studies have witnessed increased use of high-resolution arterial traffic data to evaluate various traffic performance measures. It is also important for traffic safety engineers to explore such high-resolution data to improve traffic safety. This research attempts to demonstrate how one such integrated event-based data collection and storage system, developed at University of Minnesota, popularly known as the SMART SIGNAL system can be used to identify the events leading to a crash and illuminate the mechanism by which traffic conditions and driver decisions interact to produce crashes. Primarily two types of events were investigated in-depth: signal violation crashes and potential vehicle pedestrian conflicts at signalized intersections.

In Chapter 3 a case study of a signal violation crash was analyzed using the detailed loop detector and signal activity data from the SMART SIGNAL system along with a preliminary crash report. It was fairly straightforward to identify the time of the crash based on occupancy data extracted from the SMART SIGNAL system. The CUSUM statistic was used to detect the change in the pattern of occupancies before and after the incident. It was found quite effective and easy to implement, and the estimated time of collision was between 16:06:27.078 and 16:06:52.906. However, one important point to note is that this methodology may not yield good results for all cases, particularly for less severe incidents (in terms of traffic impact) or events during very light traffic conditions. Once the crash was identified, the next effort was to reconstruct the event. The detector data and crash report were not sufficient for this event, so a plausible hypothetical scenario based on qualitative information from the crash report was developed in which the final position of the vehicles was assumed as a way to illustrate how the behaviors of the drivers involved in the crash could be estimated when the post-collision information is available. The initial speed estimates for the two vehicles were obtained based on the knowledge of vehicle type recorded in the crash report and occupancy data from the detector. A three-parameter model was developed, and a Monte Carlo simulation technique was used to estimate the conditional distribution of the parameters given the crash had occurred. The estimated mode of the multivariate distribution based on the two methods was quite close, and results indicated that even with a relatively strong braking effect, the crash could not be avoided, suggesting that the excessive speed of unit 1 was a significant factor for the signal violation and consequently the crash occurrence. The reconstruction effort indicated that it would be interesting to investigate scenarios, such as fatal crashes, in which more specific details regarding post-collision status of the vehicles are available from detailed crash investigations.

In Chapter 4 data on the initial conditions of vehicles at the entry of the intersection, obtained from SMART SIGNAL, were used to explain an observed 37% reduction in pedestrian and bicycle crashes following changes in clearance intervals. Typical pedestrian behavior at signalized intersections was assumed, and a counterfactual methodology was proposed to provide an alternate estimate of the vehicle pedestrian crash-reduction factor for clearance interval could be established. Our findings suggested that providing clearance intervals prevented 100% of potential crashes, where a driver's braking effort would not have been sufficient to avoid those. However, like any other test mechanism, the proposed model does not capture all possible scenarios in which vehicle/pedestrian conflict could happen at an urban arterial. For

example, Mclean et al. (*47*) in their documentation of 176 fatal pedestrian collisions in the Adelaide area between 1983 and 1991 indicated that many of the crashes happened when pedestrians attempted to cross the street somewhere midway between two intersections. In those scenarios, it is quite doubtful that intervention of all-red phase would have any impact. Hence, the point needed to be emphasized that the discrepancy between our findings and the study by Retting et al. (*34*), where the authors reported that re-timing of the clearance intervals resulted in only a 37% reduction, indicates that further investigation into the location and nature of the crashes is needed to identify the relative frequency of that subset, which has the potential of being affected by clearance interval.

5.2   Recommendations

As mentioned before, the main purpose of this research was to evaluate the feasibility of SMART SIGNAL data for crash identification, reconstruction, and effectiveness of traffic measures at signalized intersections. Two different studies were considered to achieve our goals. Both studies, signal violation crash and vehicle pedestrian conflict, indicated that an integrated event-based data-collection system, such as SMART SIGNAL, can provide useful information in a crash or potential crash events, which are not usually available in traditional traffic safety studies. However, it was not possible to draw meaningful conclusions solely based on SMART SIGNAL data. For example, in the reconstruction of the signal violated crash, additional post-impact information was essential to estimate the trajectory parameters. Similarly for vehicle/pedestrian interaction, adequate knowledge about the pedestrian's behavior is essential to assess the crash risk in the absence of the clearance interval. Hence, it is expected that in the future as far as traffic safety studies are concerned, the SMART SIGNAL system can make a significant contribution if used in conjunction with the other existing data collection procedures, such as video-based methods, as well as other trajectory extraction techniques.

References

1. T. Dingus, S. Klauer, V. Neale, A. Petersen., S. Lee, J. Sudweeks, M. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. Doerzaph, J. Jermeland, and R. Knipling. *The 100-car Naturalistic Driving Study Phase II-Results of the 100-car Field Experiment*. Report DOT HS 810 593, National Highway and Transportation Safety Administration, Washington, DC (2006).

2. G. Davis and T. Swenson. "Collective responsibility for freeway rear-ending accidents? An application of probabilistic causal models." *Accident Analysis and Prevention*, 38 (2006) pp. 728-736.

3. H. Liu and W. Ma. "A virtual probe approach for time-dependent arterial travel time estimation." *Transportation Research, Part C* 17 (1), 11–26, 2009.

4. H. Liu and W. Ma. "A Real-time Performance Measurement System for Arterial Traffic Signals." Presented at 86th Annual Transportation Research Board meeting, Washington, D. C., 2008.

5. M.J. Culip and F. Hall (1997). "Incident Detection on an Arterial Roadway." In *Transportation Research Record 1603*, TRB, National research Council, Washington, D.C., pp. 112-118.

6. L.D. Han and A. D. May (1989). *Automatic Detection of Traffic Operational Problems on Urban Arterials*. Research Report UCB-ITS-RR- 89-15. Institute of Transportation Studies, University of California at Berkeley.

7. M. Solomon. *A Review of Automatic Incident Detection Techniques*. ADVANCE Program Technical Report NU-1d.1-1, Transportation Center, Northwestern University, Evanston, IL, 1991.

8. S.A. Ahmed and A.R. Cook. "Application of Time-Series Analysis techniques to Freeway Incident Detection." *Transportation Research Record 841*, pp. 19-21, Issue 841, 1982.

9. M.G.H. Bell and B. Thancanamootoo. "Automatic Incident Detection Within Urban Traffic Control Systems." *Proc., Roads and Traffic 2000 Conference*, Vol. 4(2), Berlin, Germany, September, 1988.

10. L.D. Han and A. D. May. *Artificial Intelligence approaches for signalized network control*. Research Report UCB-ITS-RR-89-15. Institute of Transportation Studies, University of California at Berkeley, 1988.

11. V. Sethi, N. Bhandari, F.S. Koppelman, and J.L. Schofer. "Arterial Incident Detection Using Fixed Detector and Probe Vehicle Data." *Transportation Research C*, Vol. 3C, No. 2, pp. 99–112, April, 1995.

12. R.S. Lunetta, J.F. Knight, J. Ediriwickrema, J.G. Lyon, and L.D. Worthy. "Land-cover change detection using multi-temporal MODIS NDVI data." *Remote Sensing of Environment*, 105(2):142-154, 2006.

13. P.R. Krishnaiah and B.Q. Miao. "Review about estimation of change points." *Quality Control and Reliability*, Volume 7 of Handbook of Statistics, pp. 375-402. North-Holland, Amsterdam, Netherlands 1988.

14. J. Chen and A. Gupta. "On change point detection and estimation." *Communications in Statistics: Simulation & Computation,* 30(3), 665-697, 2001.

15. E.S. Page. "On problems in which a change in a parameter occurs at an unknown point." *Biometrika*, 44(1/2), 248-252, 1957.

16. H. Chernoff and S. Zacks. "Estimating the current mean of a normal distribution which is subjected to changes in time." *The Annals of Mathematical Statistics*, 35 (3), 999-1018, 1964.

17. V.K. Jandhyala, S.B. Fotopoulos, and D.M. Hawkins. "Detection and estimation of abrupt changes in the variability of a process." *Computational Statistics & Data Analysis*, 40(1), 1-19, 2002.

18. S. Boriah. "Time Series Change Detection: Algorithms for Land Cover Change." PhD dissertation, Department of Computer Science and Engineering, University of Minnesota, 2010.

19. A.I. Gall, F. Hall. "Distinguishing between incident congestion and recurrent congestion: A proposed Logic." *Transportation Research Record 1232*, pp. 1-8, 1989.

20. J.S. Baker. "Causes and contributing factors in traffic accidents." *Traffic Accident Reconstruction*, Northwestern University Traffic Institute, Evanston, IL, 1990.

21. R.M. Brach. and R. M. Brach. *Vehicle Accident Analysis and Reconstruction Methods*. SAE International, Warrendale, PA, 2004.

22. C.P. Robert and G. Casella. *Monte Carlo Statistical Methods* (second edition), New York: Springer-Verlag, 2004.

23. D. Fambro, K. Fitzpatrick, and R. Koppa. *NCHRP Report 400: Determination of Stopping Sight Distance*. TRB, National Research Council, Washington D.C., 1997.

24. R Development Core Team (2008): A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

25. E. Parzen. "On estimation of a probability density function and mode." *The Annals of Mathematica1Statistics,* 33, 1065-1076, 1962.

26. M.P. Wand and M.C. Jones. "Comparison of Smoothing Parameterizations in Bivariate Kernel Density Estimation." *Journal of American Statistical Association*, 88, 520-528, 1993.

27. M.P. Wand and M.C. Jones. "Multivariate Plug-in Bandwidth Selection." *Computational Statistics*, 9, 97-116, 1994.

28. T. Duong and M.L. Hazelton. "Plugin bandwidth matrices for bivariate kernel density estimation." *Journal of Nonparametric Statistics*, 15, pp. 17-30, 2003.

29. B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.

30. T. Duong. ks: "Kernel Density Estimation and Kernel Discriminant Analysis for Multivariate Data." R. *Journal of Statistical Software*, Vol. 21, Issue 7, 2007.

31. L. Devroye. "Recursive estimation of the mode of a multivariate density." *The Canadian Journal of Statistics*, 7, 159-167, 1979.

32.   A. Tsybakov. "Recursive estimation of the mode of a multivariate distribution." *Problems of Information Transmission*, 26, 31-37, 1990.

33.   C. Abraham, G. Biau, and B. Cadre. "Simple estimation of the mode of a multivariate density." *The Canadian Journal of Statistics*, Vol. 31, No. 1, pp 23-34, 2003.

34.   R.A. Retting, J.F. Chapline, and A.F. Williams. "Changes in crash risk following re-timing of traffic signal change intervals." *Accident Analysis and Prevention*, 34, 215-220, 2002.

35.   R.W.G. Anderson, A.J. McLean, M.J.B. Farmer, B.H. Lee, and C.G. Brooks. "Vehicle Travel Speeds and the Incidence of Fatal Pedestrian Collisions." *Accident Analysis and Prevention*, Vol. 29, No. 5, pp. 667-674, 1997.

36.   E. Pasanen and H. Salmivaara. "Driving Speed and Pedestrian Safety in the City of Helsinki." *Traffic Engineering and Control*, 34, No. 6, 308-310, 1993.

37.   A. Mayne. "The Problem of the Careless pedestrian." *Proceedings of the Second International Symposium on the Theory of Road Traffic,* OECD, Paris, 279-285, 1963.

38.   A. Tarko, G. Davis, N. Saunier, T. Sayed, and S. Washington. "Surrogate Measures of Safety." White paper, ANB20 (3) Subcommittee on Surrogate Measures of Safety, McLean, VA, 2009.

39.   G. Davis. "Method of Estimating the Effect of Traffic Volume and Speed on Pedestrian Safety for Residential Streets." *Transportation Research Record*, 1636, TRB, 110-115, 1998.

40.   G. Davis. *Vehicle /Pedestrian Collision Model for Neighborhood Traffic Control*. Report to Minnesota Department of Transportation, St. Paul, 2001.

41.   R. Koppa, D. Fambro, and R. Zimmer. "Measuring Driver Performance in Braking maneuvers." *Transportation Research Record*, 1550, 8-15, 1996.

42.   T.F. Fugger, B. Randles, A.C. Stein, W.C. Whiting, and B. Gallagher. "Pedestrian behavior at signal-controlled crosswalks." *SAE Technical Paper*, 01-0896, 2001.

43.   L.B. Fricke. *Traffic Accident Reconstruction*. Northwestern University Traffic Institute. Evanston, IL, 1990.

44.   G. Davis. "Relating Severity of Pedestrian Injury to Impact Speed in Vehicle-Pedestrian Crashes: Simple Threshold Model." *Transportation Research Record* 1773, 108-113, 2001.

45.   J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, UK, 2000.

46.   D. Lunn, A. Thomas, N. Best, and D. Spiegelhalter. "WinBUGS- A Bayesian modeling framework: concepts, structure, and extensibility." *Statistics and Computing*, 10, (2000), 325-337.

47.   J. McLean, R.W.G. Anderson, M.J.B. Farmer, B.H. Lee, and C.G. Brooks. *Vehicle travel speeds and the incidence of fatal pedestrian collisions*, Vol. II. Federal Office of Road Safety, Transport and Communications, Adelaide, Australia, 1994.