Developing a Real-Time Incident Decision Support System (IDSS) for the Freight Industry

RITARS-12-H-UMN

Final Report

Prepared For
OST - Office of the Assistant Secretary for Research and Technology
U.S. Department of Transportation

Prepared By

Chen-Fu Liao & Max Donath
University of Minnesota

January 2015

Table of Contents

List of Figures

List of Tables

Glossary

AADT	Annual Average Daily Traffic
ADT	Average Daily Traffic
ASOS	Automated Surface Observing System
ATIM	Advanced Traffic Incident Management
ATRI	American Transportation Research Institute
AVL	Automatic Vehicle Location
AWOS	Automated Weather Observing System
CAD	Computer-Aided Dispatch
DOT	Department of Transportation
DOF	Degree of Freedom
DPSAC	Data Policy Standing Advisory Committee
DR	Detection Rate
DS	Data Set
DSA	Data Sharing Agreement
EB	Eastbound
EMS	Emergency Medical Services
EMSRB	Emergency Medical Services Regulatory Board
FAR	False Alarm Rate
FHWA	Federal Highway Administration
FPM	Freight Performance Measures
GIS	Geographic Information Systems
GMT	Greenwich Mean Time
GPS	Global Positioning System
HCAADT	Heavy Commercial Annual Average Daily Traffic
ICAO	International Civil Aviation Organization
ID	Identification
IDSS	Incident Decision Support System
IRC	Inter-Regional Corridor
IEEE	Institute of Electrical and Electronics Engineers
ITS	Intelligent Transportation Systems
Km	Kilometer
LAT	Latitude
LON	Longitude
MAD	Median Absolute Deviation
MM	Mile Marker
MN511	Minnesota 511 System
MnCARS	Minnesota Condition Acquisition and Reporting System
MnDOT	Minnesota Department of Transportation
MPH	Mile per Hour

MPO	Metropolitan Planning Organization
MSP	Minneapolis – St. Paul
MTTD	Mean Time to Detect
MTO	Minnesota Traffic Observatory
N-CAST	National Corridors Analysis & Speed Tool
NB	Northbound
NCHRP	National Cooperative Highway Research Program
OFCVO	Office of Freight and Commercial Vehicle Operations
OST-R	Office of the Assistant Secretary for Research and Technology
POV	Privately Owned Vehicle
RWIS	Road Weather Information System
SB	Southbound
SQL	Structured Query Language
TCMA	Twin Cities Metro Area
TTS	Text to Speech
UMN	University of Minnesota
USDOT	U.S. Department of Transportation
UTC	Coordinated Universal Time
WB	Westbound
WIM	Weigh-In-Motion
XML	Extensible Markup Language

Acknowledgement

This project is sponsored by the USDOT, Office of the Assistant Secretary for Research and Technology (OST) under contract #RITARS-12-H-UMN. We are highly indebted to the American Transportation Research Institute (ATRI) and ALK Technologies Inc. for providing in-kind match, software tools, data, and staff support for making this project possible. We also would like to thank the Minnesota Department of Transportation (MnDOT) and the Minnesota State Patrol for providing data and staff support. Many thanks and appreciations also go to Saura Verma for incident data analysis and Sandeep for developing the Android application interface.

Executive Summary

Our nation's economy is highly dependent on reliable and cost-effective truck-freight transportation. Delays to truck movement are of particular concern to the nation. Building upon our previous effort, we developed an Incident Decision Support System (IDSS) that uses GPS-equipped commercial trucks as probe vehicles on key freight corridors such as inter-regional corridors (IRC) and supplemental freight routes in Minnesota. An automatic incident detection algorithm was developed: (a) to be used by truck drivers and dispatchers who need to make necessary routing and operations decisions when incidents occur, (b) to support incident decision making by state DOTs and (c) for traveler information purposes.

The objective of this project was to develop an Incident Decision Support System (IDSS), that uses GPS-equipped commercial trucks as probe vehicles on freight corridors, to detect incidents and provide incident information in real-time to truck operators on rural corridors where infrastructure-based sensing is unavailable.

Archived truck GPS data, MN511 traffic data, Computed Aided Dispatch (CAD) data from the Minnesota State Patrol, and weather data in the Twin Cities Metro Area (TCMA) were obtained for this study. GPS-based truck probe data was obtained from the American Transportation Research Institute (ATRI) who since 2002, has partnered with the FHWA and the trucking industry to continuously collect GPS data on key national corridors, using nearly 500,000 commercial trucks in North America.

The MN511 data provides basic incident information such as time, location, incident description, presence of stalled vehicle, presence of debris, etc. The Computer Aided Dispatch (CAD) data from the Minnesota State Patrol and incident management team contains additional information such as patrol arrival/leaving time, number of vehicles involved, etc. Weather information from the Automated Weather Observing System (AWOS) and Automated Surface Observing System in Minnesota was also used for modeling incident delays.

A bivariate incident detection methodology was developed by using the travel time computed from raw GPS truck data and the travel time difference between the current interval and the previous interval for each roadway segment. The methodology was implemented using a Java program, which analyzed the probe vehicle data and compared the results to normal conditions. The incident detection algorithm was validated by comparing our incident detection results to those based on MnDOT’s MN511 incident and loop detector data on freight corridors with sufficient GPS samples in the Twin Cities metropolitan area. It is important to note that the incident detection algorithm, based on probe vehicle data, was developed for rural roadways where loop detector data is not available.

A pilot implementation was developed to demonstrate a concept of operations for the IDSS, in which incident information is provided to truck operators. The pilot implementation used available archived data for demonstration purposes. In the future, a real-time data feed from ATRI can be incorporated into the IDSS when real-time data becomes available.

For this pilot implementation and demonstration, an in-vehicle "app" using a smartphone device was developed to provide incident information to the truck operator. When an incident is detected, the app provides incident information through an auditory text-to-speech (TTS) interface as the vehicle approaches the incident. For example, if the distance to an incident is xx miles ahead, the app will announce "incident occurred xx miles ahead 'yy' minutes ago. The TTS speech engine also identifies the route name of the incident assuming that the driver is aware of his or her current route so that he or she can identify whether the incident is on the same route or on a nearby route.

The app was tested on a passenger vehicle driving on I-94 westbound near the University of Minnesota campus in Minneapolis where an incident occurred based on historical data. For testing and validation purposes, we set the time on the smartphone app to a timestamp based on the archived data when the incident actually occurred. As the test vehicle drove towards the location where the incident occurred, the app successfully detected available incident information through our web servlet and announced the corresponding information (where and when) to the driver during the experiments.

We are aware of the potential distraction of providing travel information through the smartphone. The user interface and human factors issues will need further investigation in order to determine the best way to ultimately present the incident information to the driver. Eventually, the incident information can be provided through existing in-vehicle devices already in the truck cabin.

An automatic delay estimation process, which determines when traffic conditions return to normal, was also developed and tested. Individual incident durations were calculated using the time difference between the update time of the "last" message and the start time of the "first" message, related to the incident, from MN511 and CAD data. The research team computed the incident delay for selected incident sites where loop detector data was available. In order to develop this incident delay prediction model, the research team tried several pattern classification algorithms, but with limited success.

Predicting incident delay reliably is very difficult. No one has successfully developed such a model before. We believe that developing an incident duration prediction model will require additional detailed incident information such as class of involved vehicles, number of injuries, number of lanes blocked, number of deployed emergency vehicles, etc.

Many traveler information systems provide incident alerts to drivers based on reports by travelers, or by the incident responders. Most incident detection methodologies have been developed based on loop detectors or vision sensors. Our approach, based only on probe vehicles, can detect traffic disruption automatically and alert vehicle operators of incidents ahead, especially useful for roadways without instrumentation.

1. Introduction

Rapid detection of incidents on the highway can reduce the impact of the resulting traffic congestion and the risks associated with secondary incidents. Incident information provides a key decision support element for freight dynamic route guidance which gives the truck driver or the dispatcher real-time route-specific information allowing them to make the best decision about whether to wait out the incident or take an alternate route. We developed an Incident Decision Support System (IDSS) using statistical pattern recognition and modeling techniques to detect an incident.

Current MN511 system has very good coverage on instrumented roadways in the metropolitan area. Incident information is usually broadcast through dynamic message signs on major roadways. We validated our incident detection results against incident reports from the Minnesota 511 system and developed an in-vehicle system to provide incident information to commercial vehicle operators as they are heading toward the incident location. Our approach can provide incident information to travelers on roadways with no instrumentation when there are sufficient probe GPS data samples.

As part of this research project, we had access to 24 months of American Transportation Research Institute’s (ATRI) truck GPS data in the Twin Cities Metro Area (TCMA). The archived data allowed us to verify our incident detection results with MN511 data which we have been archiving since October 2012. We feel that drivers from the MnDOT Freeway Incident Response Safety Team (FIRST) would be great candidates for future testing and evaluating the in-vehicle IDSS.

Because of our use of historical data, the validation and testing occurred in a simulated environment. In our example, a driver travels on a roadway with the IDSS timestamp set on a historical day and hour. The in-vehicle system then announces any upcoming incident information (historical) to travelers as they are on the same roadway and heading toward an incident. Currently, we are not yet receiving real-time truck GPS data feed, however, the real-time data feed can be made available for this application according to the data provider. A pilot implementation was developed to demonstrate a concept of operations.

1.1 Objectives

Our nation's economy is highly dependent on reliable and cost-effective truck-freight transportation and delays to trucks are of particular concern to the nation. The objectives of this project are to (1) develop an Incident Decision Support System (IDSS) that uses GPS-equipped commercial trucks as probe vehicles on key freight corridors in Minnesota, and (2) provide incident information to truck operators.

1.2 Literature Review

1.2.1 Freight Performance Measures

The trucking industry represents the largest portion of domestic freight movement in the United States. According to the ATA U.S. Freight Transportation Forecast for 2021, the trucking industry's share is about 68% of total tonnage; trucks move more than 80% of freight revenue. Safe and efficient trucking services are essential, not only to provide door-to-door freight transportation, but also to ensure the effective operation of other freight modes and facilities.

Trucks usually occupy more than twice the space of passenger vehicles on the roadway and they carry a higher value of goods. Truck delay due to traffic congestion or other environmental factors have a more significant impact on our nation's economy than automobile delay. The Federal Highway Administration (FHWA) has developed a national congestion monitoring program that uses archived traffic detector data for measuring traffic congestion and travel reliability (Turner et al., 2004; Pu, 2011;). NCHRP Synthesis Report 384 (Kuzmyak, 2008) identified the challenges that many metropolitan planning organizations (MPOs) are facing in forecasting and modeling freight transportation. Many MPOs model heavy trucks as a surrogate for modeling freight activity because trucks account for more than 80% of freight movement in most metropolitan areas. The FHWA and the American Transportation Research Institute (ATRI) recently released findings on the level of truck congestion at 250 freight significant highway locations. Five highway interchange locations in the Twin Cities Metropolitan Area (TCMA) were included in this study (ATRI, 2011).

Schofield and Harrison (2007) reported the status of freight performance measures used in DOTs nationally and suggested a set of relatively broad performance measures including mobility, reliability, economic, safety/environment, and infrastructure for emerging users. Varmar (2008) compiled, organized, and analyzed freight data by mode, performance measure and indicator categories. The report suggested that there is a need to: (1) determine what performance measures or indicators are relevant and most important for freight planning support, and (2) identify freight significant strategic corridors and nodes.

The Minnesota Department of Transportation (MnDOT) Office of Freight and Commercial Vehicle Operations (OFCVO) has identified and included travel time by mode as one of its four performance indicators (MnDOT's Statewide Freight Plan, 2005). MnDOT has also deployed Automatic Traffic Recorders (ATR) and Weigh-In-Motion (WIM) systems statewide for measuring truck weight and classifications with varying axle configurations at highway speeds. Existing ATR and WIM sensors collect truck volume and speed information at selected locations statewide, but they do not provide truck travel time information. On-board GPS systems that collect truck location at a constant polling rate, present an excellent data source for monitoring travel time and reliability. In the past, GPS-based truck trip data was not available and was difficult to collect due to the proprietary nature of the data.

1.2.2 Probe Vehicle Based Performance Measures

With the prevalence of GPS receivers on vehicles and portable navigation devices, probe vehicle based data collection has been increasingly attractive to the transportation community. The GPS based vehicle location data has been used to estimate traffic states and derive travel time information for traffic monitoring (Lund and Pack, 2010; Guo et al., 2008; Smith, 2006; Nanthawichit et al., 2003). Probe vehicle data, when fused with loop detector data and other data sources, can provide more complete and continuous coverage of traffic monitoring. Turner et al. (2011) outlined the primary data requirements for congestion-related performance measures and introduced core data elements and various metadata to ensure data consistency among data providers. They also examined legal and institutional issues related to privacy and Freedom of Information (FOIA) with regard to implementation.

Travel time reliability is one of the key measures of freight performance along interstates or interregional corridors in the nation (Lomax et al., 2003; TTI, 2006). Pu (2011) examined several reliability measures and recommended a median-based buffer index (a measure which compares the 95th percentile of travel time to the median travel time) or a percent on-time rate as more appropriate to handle heavily skewed travel time distributions.

Since 2002, FHWA has established a partnership with the American Transportation Research Institute (ATRI) to measure average truck travel speed on major freight-significant corridors (Jones et al., 2005). A spatial data processing methodology was evaluated and refined by Liao (2008) to improve the effectiveness of freight performance measures. Analyzing truck speed, volume and travel time by location can also help identify network impediments and variations of seasonal flow changes (Liao, 2009). Derived vehicle speed and travel time from the GPS and terrestrial wireless system used by the trucking industry provides potential opportunities to support freight planning and operation on the surface transportation system.

A majority of commercial vehicles are equipped with on-board Automatic Vehicle Location (AVL) systems that collect truck locations at a fixed polling rate. The continuous trajectory information presents an excellent data source for monitoring travel time and reliability. However, GPS-based truck trip data usually are not available and are more difficult to collect due to the proprietary nature of the data. Commercially available travel time information (for example, from INRIX) provides some coverage using aggregated general traffic speed data from loop detectors and other probe vehicle based data sources. However, heavy commercial vehicles are considerably underrepresented in this type of data source.

McCormack and Hallenbeck (2006) used 25 portable GPS data collection units with 1-second polling rate to gather truck positioning data for measuring freight movements along freight significant corridors in Washington State. The study concluded that GPS data can be collected cost effectively and can provide an indication of roadway performance. Based on processed truck speed data, a route model including analyses of truck travel time, delay and reliability can be developed to better understand current freight network performance, freight origin to destination flows, and to study possible solutions to future freight demand growth (Short & Jones, 2008).

In its initial phase, the FHWA FPM initiative measured average travel rates on five freight-significant corridors (Jones et al., 2005). ATRI analyzed the severity of 30 key freight bottlenecks in the U.S. interstate system (Short et al., 2009). Freight bottlenecks occurring at highway interchanges were analyzed using a freight congestion index. Possible causes for the bottlenecks may include roadway geometry (e.g., grade, curvature, and sight distance), capacity (number of lanes), toll booths, speed limit, weather, truck volume vs. general traffic volume, and available lanes of travel for trucks.

MnDOT completed a study on truck parking analysis. The goal was to develop the information necessary to support decisions regarding future approaches to the truck parking problem in Minnesota (Maze et al., 2010). Short and Murray (2008) demonstrated the capability of utilizing FPM data for truck parking analysis. Another application is to utilize the FPM data to evaluate the travel time and delay at border crossings. FHWA conducted a study to address the need to reduce the hours of delay for commercial motor vehicles passing through ports-of-entry (FHWA, 2002). However, manual truck data collection at border crossing plaza is labor intensive and expensive.

Recently, FHWA has led an effort to assess and validate the appropriateness of using GPS data from commercial vehicles to derive mobility and reliability performance measures and to support congestion monitoring on the highway system. Four key factors, including average daily traffic (ADT) per lane, percent of heavy vehicle, grade, and congestion level, were investigated. The preliminary findings indicated that (1) estimates of speed from FPM data are sufficiently accurate for performance measurement on most roadways in the United States, (2) FPM speed estimates show a consistent negative bias due to differences in operating characteristics of trucks and autos, and (3) grade and congestion have the greatest effect on FPM data accuracy among the four key factors evaluated (FHWA, 2012).

1.2.3 National Corridors Analysis & Speed Tool (N-CAST)

ATRI in coordination with the FHWA recently announced (10/22/2012) a beta release of a Freight Performance Measures (FPM) tool that expands on the scope and functionality of the original FHWA-sponsored "FPMWeb" application (www.freightperformance.org/). The National Corridors Analysis & Speed Tool (N-CAST, www.atri-online.org/n-cast) provides key roadway performance and truck mobility information for the U.S. Interstate Highway System. The N-CAST database includes the average speed and a proportion of total GPS data points for each one-mile segment during each AM peak (6-10AM), mid-day (10AM-3PM), PM peak (3PM-7PM), and off peak (7PM-6AM) periods. The N-CAST tool has the potential to be integrated with existing truck data sources to generate critical performance measures (such as delay and reliability) to provide technical guidance to stakeholders in the freight industry.

1.2.4 Incident Detection

Rapid incident detection can reduce the impact of traffic congestion and the risks associated with secondary incidents. It is also critical for improving freight mobility, just-in-time deliveries, reducing unnecessary idling, and improving safety. An incident delay estimation model was developed to estimate the potential delay when an incident is identified. Estimated delay information provides a key decision support element for freight dynamic route guidance which gives the freight driver or the dispatcher real-time route-specific information allowing them to make the best decision about whether to wait out the incident or take an alternative route.

Ozbay & Kachroo (1999) raised three basic issues (surveillance, algorithm, and verification issues) concerning incident detection. New sensors using different technologies have been adopted by DOTs. Sensor reliability, performance under different environmental condition, accuracy, real-time performance, and cost play pivotal roles in the selection of a detection system. Two types of algorithms are commonly used for incident detection and delay estimation on freeways: point-based and spatial-based algorithms.

A number of incident detection algorithms have been described in the literature. However, these algorithms were based on roadway point data, for example, loop detectors or fixed traffic detectors. The point-based approach uses comparative or pattern recognition, statistics, traffic modeling, and artificial intelligence based algorithms for incident detection. As such, these systems do not adequately generate continuous roadway/traffic conditions in the real world since they are neither ubiquitous nor free from malfunction and error.

The spatial measurement based approach uses video camera or probe vehicles which are becoming more available for traffic engineering applications. In recent years, probe vehicle based approaches have been applied in limited instances, with most probe based incident detection algorithms having only been tested in a simulation environment (Baykal-Gursoy et al., 2006; Li et al., 2006; Zeng & Songchitruksa, 2010) and/or limited to metropolitan areas (Giuliano 1989; Yu et al, 2007).

Martin et al. (2000) evaluated a range of incident detection technologies for use in the Utah

Department of Transportation’s (UDOT) Advanced Traffic Management System (ATMS). Based on the research findings, the research recommended that cellular telephone technology be used as the primary form of incident detection.

1.2.5 Incident Duration and Delay

An incident is defined as any occurrence of events that affects roadway capacity (Giuliano, 1989). Incident duration is the time taken to remove an incident and recover the road capacity. It varies significantly depending on numerous factors, including incident type, location, response time, and clearance time. It is almost impossible to predict incident duration with acceptable accuracy even when a great deal of historical data is available.

Predicting traffic incident delay is a challenging task in Advanced Traffic Incident Management (ATIM). The duration of an incident delay consists of the incident time period (detection, response, and clearance time periods) and the recovery time. In the recovery period, all obstacles are removed from the roadway and the traffic queue begins to resolve until the traffic is restored back to the normal condition. Although traffic recovery time is crucial to determining incident induced delay, relatively few studies have focused on modeling post-incident traffic recovery time (Saka et al., 2008; Zeng & Songchitruksa, 2010).

The most widely used technique to estimate incident delay is the use of deterministic queuing based delay estimation technique proposed by Morales (1989) who used a simple deterministic queuing model as an analytical procedure for estimating delay under a specific incident scenario. Traditionally, incident duration and delay have been typically modeled in the form of lognormal distributions (Golob et al., 1987 and Sullivan, 1997), the log-logistic hazard-based model (Jones et al., 1991; Nam & Manning, 2000), and the truncated regression model (Khattak et al., 1995). Giuliano (1989) included incident type and occurrence time of day in an incident time estimation model based on statistical distributions. Khattak et al. (1995) formulated a clearance time prediction model using incident type and severity as the most significant impact factors.

Fu (2004) developed a fuzzy queuing model to predict possible delay of a vehicle near an incident location based on real-time information, traffic demand, queuing condition and lane closure. Fu (2004) demonstrated through simulation that incident delay prediction from a deterministic model is highly sensitive to the uncertainty of traffic conditions. Boyles & Waller (2007) took incident duration distribution from a Bayesian classification and lognormal distribution to account for uncertainty in incident duration prediction. An analytical formula for total incident delay was developed and tested in simulation using four different traffic demand profiles. They concluded that failing to properly account for uncertainty will possibly result in underestimating incident delays by up to a factor of two.

Skabardonis et al. (1996 & 1997) developed a methodology for estimating incident delay using data collected along a segment of highway I-880 in the San Francisco Bay Area. In addition, Skabardonis et al. (1999) studied the incident patterns on I-10 in Los Angeles, identified major factors affecting incident frequency, and compared the results with previous analyses on I-880. They concluded that the resulting development and analyses could help improve incident management and support the development and calibration of incident detection algorithm in simulation models.

Ji et al. (2011) developed incident recovery and delay models based on a macroscopic cell transmission model (CTM) to reproduce the traffic behavioral phenomena. They found that the recovery time increases significantly with the increase of traffic demand, which has a more significant influence over incident time than recovery delay.

1.2.6 Incident Detection Performance Measures

The following parameters have commonly been used to measure the performance of incident detection algorithms. The parameters are detection rate (DR), false alarm rate (FAR), and mean time to detect (MTTD).

1. Detection Rate (DR):

$$DR = \frac{Number\ of\ detected\ incidents}{Total\ number\ of\ recorded\ incidents} \times 100\%$$

2. False Alarm Rate (FAR):

$$FAR = \frac{Number\ of\ detected\ incidents\ that\ are\ NOT\ incidents}{Total\ number\ of\ actual\ incidents} \times 100\%$$

3. Mean Time to Detect (MTTD):

$$MTTD = \frac{1}{N}\sum_{i=1}^{N}(t_{id} - t_{io})$$

Where,

t_{io} is the time incident actually occurred,
t_{id} is the time incident was detected by the algorithm, and
N is the number of incidents.

1.3 Report Organization

The rest of this report is organized as follows. Chapter 2 describes the truck GPS data, MN511 data, MN state patrol’s CAD data and the weather data. Data analysis and methodologies are discussed in Chapter 3 and 4 respectively. Pilot implementation of the IDSS system through a smartphone app is presented in Chapter 5. Project summary and conclusion is included in Chapter 6.

2. Summary of Data

24 months of truck GPS data that covers the Greater Twin Cities metro area were acquired from ATRI for this study. In order to study the relationships among variables under both incident and non-incident conditions, detected incidents were validated against incident reports from the MN 511 system. In Minnesota, the incident information is reported and updated in a Computer Aided Dispatch (CAD) system by the state patrol or operators in the regional traffic management centers. Reported incident data automatically populate the MN 511 system, which is accessible through an Extensible Markup Language (XML) data protocol.

In addition to the CAD system, Minnesota has also developed and implemented the Minnesota Condition Acquisition and Reporting System (MnCARS) through a pooled-fund project with Iowa, Washington and Missouri. The MnCARS is an Internet-based application used by MnDOT Districts and the Minnesota State Patrol to enter data about road conditions, restrictions and incidents. The MnCARS data is integrated into a database that is then accessible to travelers though the MN 511 system.

Weather information from the Automated Weather Observing System (AWOS) and Automated Surface Observing System (ASOS) in MN was also analyzed to evaluate the weather parameters that affect the incident duration and delays.

2.1 Truck GPS Data

Since 2002, the ATRI has partnered with FHWA and the trucking industry to continuously collect GPS data on key national corridors, using nearly 500,000 commercial trucks in North America. This massive amount of truck GPS data can provide public agencies at both the federal and regional level with tools that can increase understanding of freight activity, identify impediments along the freight network, and provide for near-real-time operations decision-making.

The University of Minnesota (UMN) has established a data sharing agreement with ATRI. The data attributes to be reported for each record include a unique vehicle number, latitude, longitude, and date/time. No two trucks use the same identifier. Twenty four months (January 2012 to December 2013) of truck GPS data in the Twin Cities metro area (TCMA) were obtained. A sample of GPS point cloud data is displayed in Figure 2-1. Three different sets of truck GPS data as summarized and listed in Table 2-1. Dataset A and C contain probe vehicle spot speed and latitude-longitude location information. Dataset B does not include vehicle spot speed information. Dataset A has a positioning accuracy less than 3 meters. At 95% probability, the GPS positioning accuracy of dataset B and C is about 150 and 58 meters, respectively. Corresponding tolerance is used to merge raw GPS point to a nearest roadway. Due to data privacy concerns, the vehicle ID is masked or encrypted. In addition, the vehicle ID in dataset B rotates every 15 days and the vehicle ID in dataset C changes every 24 hours. The estimated GPS pinging rate for dataset A, B and C are about 8, 18 and 1 minute with standard deviations of 15, 26, and 5 minutes, respectively. A list of ATRI truck GPS data fields for each dataset is included in Table 2-2.

Table 2-1 Summary of ATRI GPS Data

Data Set	**DS-A**	**DS-B**	**DS-C**
Time Zone	GMT/UTC	GMT/UTC	GMT/UTC
Spot Speed	Yes	No	Yes
Static ID	Yes	Rotates every 15 days	Rotates every 24 hours
Data Accuracy	Within <3 meters	Within 124-134 meters at 90% probability and 129-150 meters at 95% probability.	Within 13-56 meters at 90% probability and 15-58 meters at 95% probability.
Snap Tolerance Used (meter)	**50**	**150**	**50**
2013 Number of Truck Trips	**74,823**	**35,179**	**76,471**
2013 Raw Data Size	50,170,591	3,142,634	38,871,190
2013 Snapped	18,792,493	957,076	13,270,602
2013 Snapped Percentage	37.5%	30.5%	34.1%
Average Sampling Time (min)	8	18	1
SD Sampling Time (min)	15	26	5

Table 2-2 ATRI Truck GPS Dataset

Data Field	**DS-A**	**DS-B**	**DS-C**
1	truckid	truckid	truckid
2	readdate	readdate	readdate
3	speed	-	speed
4	heading	-	-
5	latitude	latitude	latitude
6	longitude	longitude	longitude

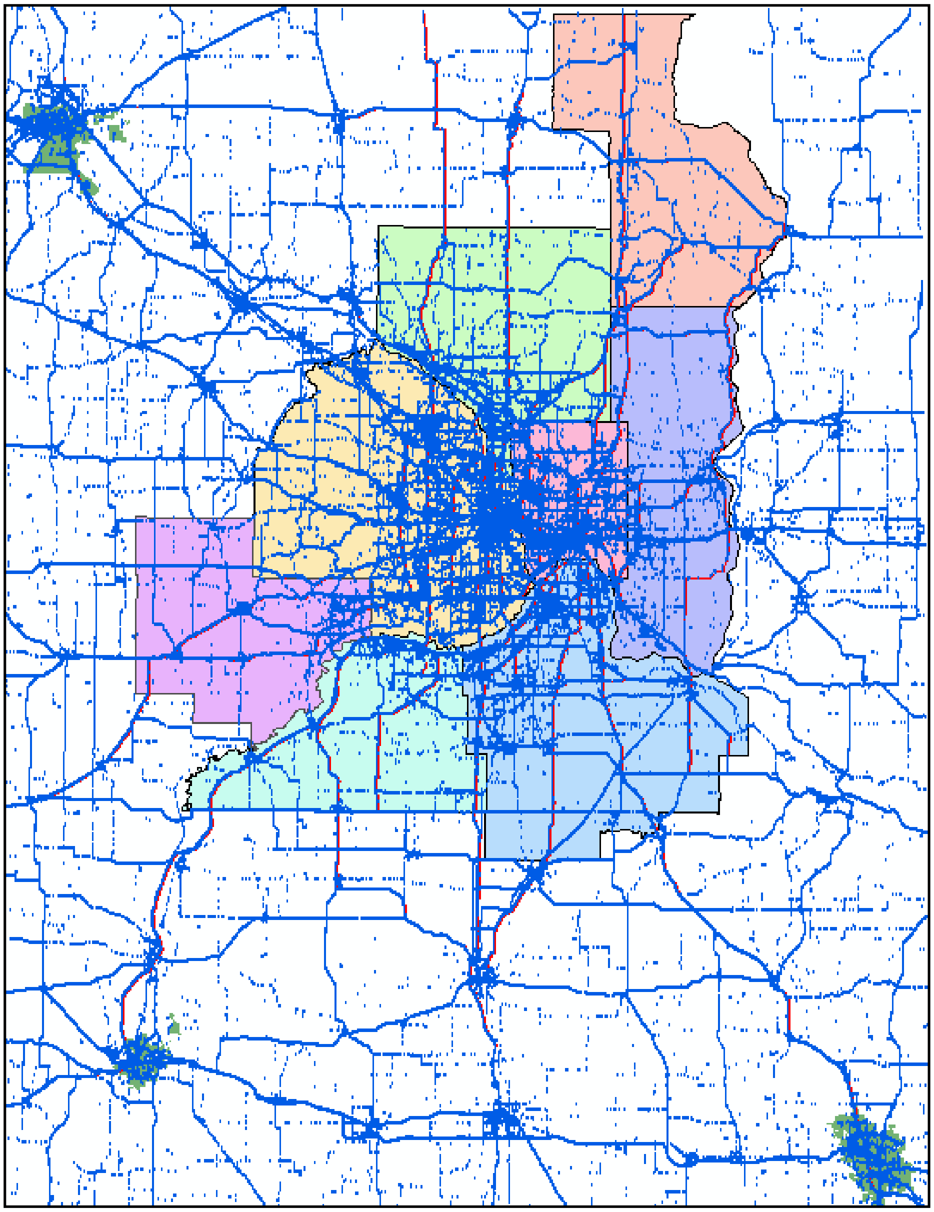


Figure 2-1 Snapshot of Truck GPS Point Cloud (Dec. 2013)

2.2 MN511 Data

A sample of MN511 Extensible Markup Language (XML) incident data is listed in Table 2-3.

Table 2-3 Sample MN511 Data Queried from Database

send_org_id	Minnesota DOT	Minnesota DOT	Minnesota DOT	Minnesota DOT	Minnesota DOT
send_ctr_id	MSP/CAD System	MSP/CAD System	MSP/CAD System	MSP/CAD System	MSP/CAD System
msg_version	1	1	1	1	1
msg_number	1358543272	1358543197	1358547327	1358542508	1358546314
msg_timestamp	2/5/2013 10:01	2/5/2013 9:16	2/12/2013 1:46	2/4/2013 9:46	2/10/2013 20:16
msg_expire_time	11/25/2016 7:06	11/25/2016 6:25	12/1/2016 23:00	11/24/2016 7:00	11/30/2016 17:32
event_id	MSPCAD-P130058434	MSPCAD-P130058366	MSPCAD-P130070244	MSPCAD-P130056879	MSPCAD-P130068439
event_update	4	3	3	2	3
event_status	current	current	current	current	current
event_priority	4	4	4	4	6
headline_text	[3818]crash	[3818]crash	[3818]crash	[3818]crash	[3818]disabled vehicle
d_desc _id	1	1	1	1	1
d_desc_ text	[3818]crash	[3818]crash	[3818]crash	[3818]crash	[3818]disabled vehicle
route_designator	I-35	I-35	I-35	I-35	I-35
primary_loc (lat, Lon, Ele)	{44.0297, -93.2466, 37.9583}	{44.0725, -93.2518, 40.8984}	{44.0862, -93.2456, 41.8976}	{44.102, -93.2452, 42.9808}	{44.102, -93.2452, 42.9808}
d_loc_link_dir	negative direction	positive direction	negative direction	positive direction	positive direction
d_loc_link_lref_ver	1	1	1	1	1
update_time	2/5/2013 10:01	2/5/2013 9:16	2/12/2013 1:46	2/4/2013 9:46	2/10/2013 20:16
start_time	2/5/2013 8:53	2/5/2013 8:31	2/12/2013 0:50	2/4/2013 9:24	2/10/2013 19:39
duration_min	1:08:28	0:45:01	0:56:01	0:21:58	0:36:39

2.3 Computer Aided Dispatch (CAD) Data

Any privacy sensitive information from the CAD database was excluded from the data query. Available data fields are listed as displayed in Table 2-4. Spatial analyses were performed to match the CAD data with the MN511 incident records.

Table 2-4 CAD Data Description

Field Name	Data Descriptions
num_1	Agency "event" number
tycod	Agency specified event type used to describe the event which has occurred
typ_eng	Verbose description of the event type
sub_tycod	The sub-type of the event type.
sub_eng	The textual description of the sub_type
ecompl	Commonplace name associated with the event's location
edirpre	Direction prefix component of the event's location
edirsuf	Direction suffix component of the event's location
efeanme	Feature name component (street name) of the event's location
efeatyp	Feature type component of the event's location
estnum	Street number (house number) component of the event's location
earea	Area component of the event's location
emun	Municipality component of the event's location
loc_com	Location comments.
xstreet1	Feature name and type of one street which intersects the event's location
xstreet2	Feature name and type of one street which intersects the event's location
x_cord	X map coordinate for the event's location
y_cord	Y map coordinate for the event's location
aeven.ad_ts	Time event was actually added into system
aeven.udts	Update Date/Time Stamp.
aeven.ar_ts	Time that the first unit arrived on the scene
aeven.xdts	Closing Date/Time Stamp.

2.4 Weather Data

The Automated Weather Observing System (AWOS) and Automated Surface Observing System (ASOS) can be found at http://mesonet.agron.iastate.edu/request/download.phtml?network=MN_ASOS. A GIS map of ASOS and RWIS Weather Stations in Minnesota is displayed in Figure 2-2.

Sample ASOS data

Table 2-5 Sample Weather Data at MNPLS/CRYSTAL (MIC) Weather Station

station	MIC	MIC	MIC	MIC
valid (local timezone)	6/26/2013 6:53	6/26/2013 7:53	6/26/2013 8:53	6/26/2013 9:53
lon	-93.3539	-93.3539	-93.3539	-93.3539
lat	45.062	45.062	45.062	45.062

tmpf	73.94	77	80.06	80.06
dwpf	71.06	71.96	71.96	73.04
relh	90.73	84.46	76.36	79.2
drct	150	160	0	160
sknt	5	4	0	9
p01i	0	0	0	0
alti	29.67	29.67	29.66	29.68
mslp	1004	1004	1003.8	1004.4
vsby	4	7	9	8
gust	M	M	M	M
skyc1	CLR	CLR	FEW	CLR
skyc2	None	None	None	None
skyc3	None	None	None	None
skyc4	None	None	None	None
skyl1	M	M	1400	M
skyl2	M	M	M	M
skyl3	M	M	M	M
skyl4	M	M	M	M
metar	KMIC 261153Z 15005KT 4SM BR CLR 23/22 A2967 RMK AO2 SLP040 T02330217 10250 20222 50003	KMIC 261253Z 16004KT 7SM CLR 25/22 A2967 RMK AO2 SLP040 T02500222	KMIC 261353Z 00000KT 9SM FEW014 27/22 A2966 RMK AO2 SLP038 T02670222	KMIC 261453Z 16009KT 8SM CLR 27/23 A2968 RMK AO2 SLP044 T02670228 53004

Notes:

1. lon – longitude
2. lat – latitude
3. tmpf - temperature F
4. dwpf - dew point F
5. relh - relative humility
6. drct - wind direction
7. sknt - speed knots
8. p01i - precipitation
9. alti - altimeter (in)
10. mslp - sea level pressure (mb)
11. vsby - visibility (miles)
12. gust - wind gust
13. 'M' means that either no value was reported for that observation time or the value was reported as missing.
14. 'SCT' is scattered cloud coverage reported by the ASOS and 'BKN' is broken coverage.
15. In general, the ASOS sites do not report snowfall, only the liquid melted equivalent.
16. metar - METAR is a format for reporting weather information. Raw METAR is the most popular format in the world for the transmission of weather data. It is highly standardized through the International Civil Aviation Organization (ICAO), which allows it to be understood throughout most of the world. http://en.wikipedia.org/wiki/METAR

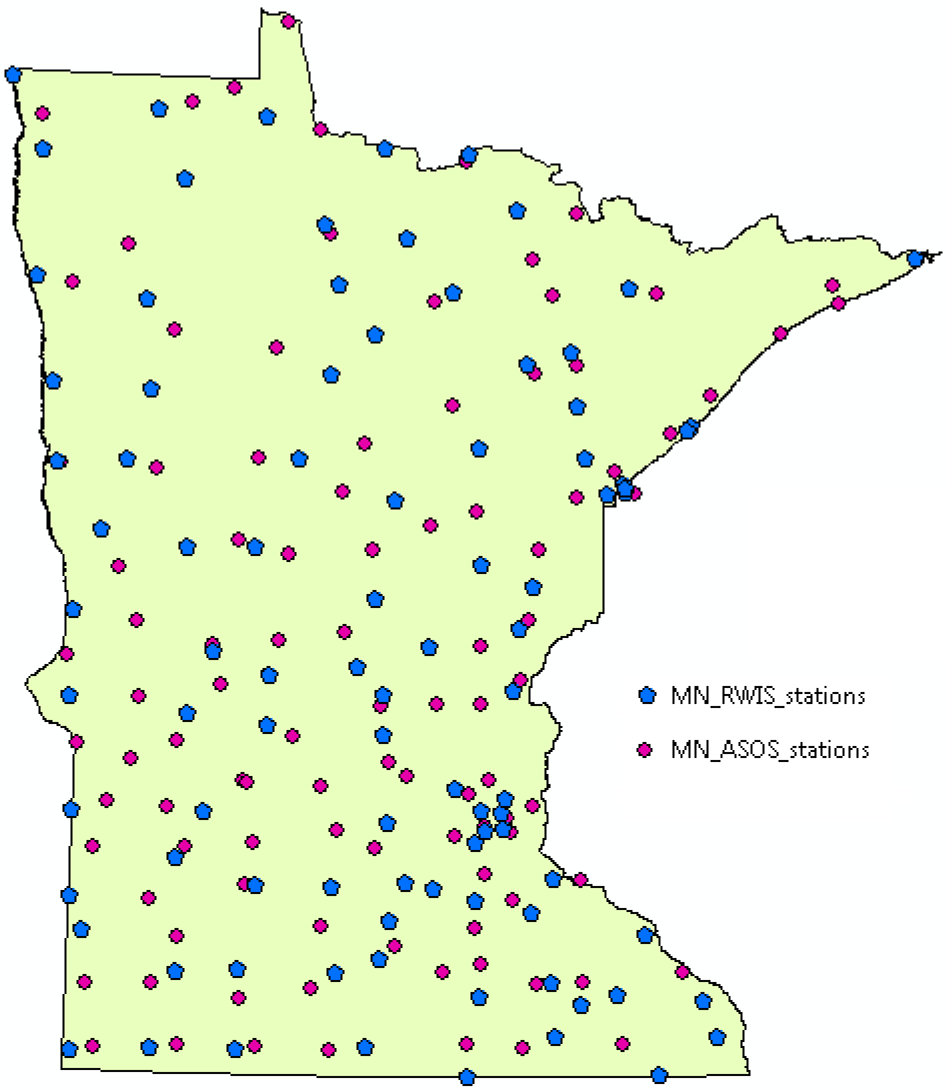


Figure 2-2 ASOS and RWIS Weather Stations in Minnesota

3. Data Analysis

Archived truck GPS data, MN511 traffic data, and Computer Aided Dispatch (CAD) data from Minnesota state patrol were analyzed.

3.1 Minnesota 511 Data Analysis

The incident duration was estimated from the incident start time to the last incident update time of each XML incident record with the same incident ID. The distributions of estimated crash incident period in February and March 2013 from the MN511 system are illustrated as follows. Figure 3-1 shows the distribution of estimated crash duration for incidents that were recorded in Feb. 2013 with mean duration of 29.5 min and median duration of 23.4 min. And Figure 3-2 displays the distribution of estimated crash duration for events that occurred in Mar. 2013 with mean duration of 28.8 min and median duration of 24.6 min.

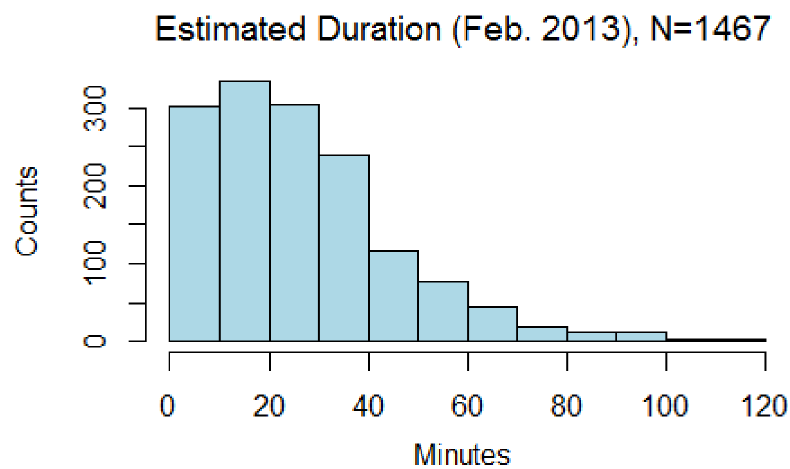


Figure 3-1 Histogram of Estimated Incident Period in Feb. 2013

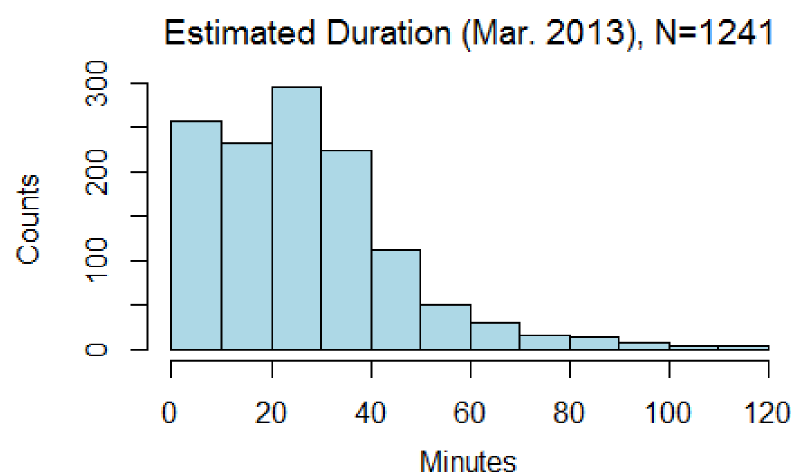


Figure 3-2 Histogram of Estimated Incident Period in Mar. 2013

Figure 3-3 and 3-4 illustrate the average daily number of crashes by hour in February and March, respectively. More crashes occur during AM and PM peak hours as compared to off-peak hours. According to the MN511 incident data, there were averaging 47 crashes per day in February and 31 crashes per day in March 2013.

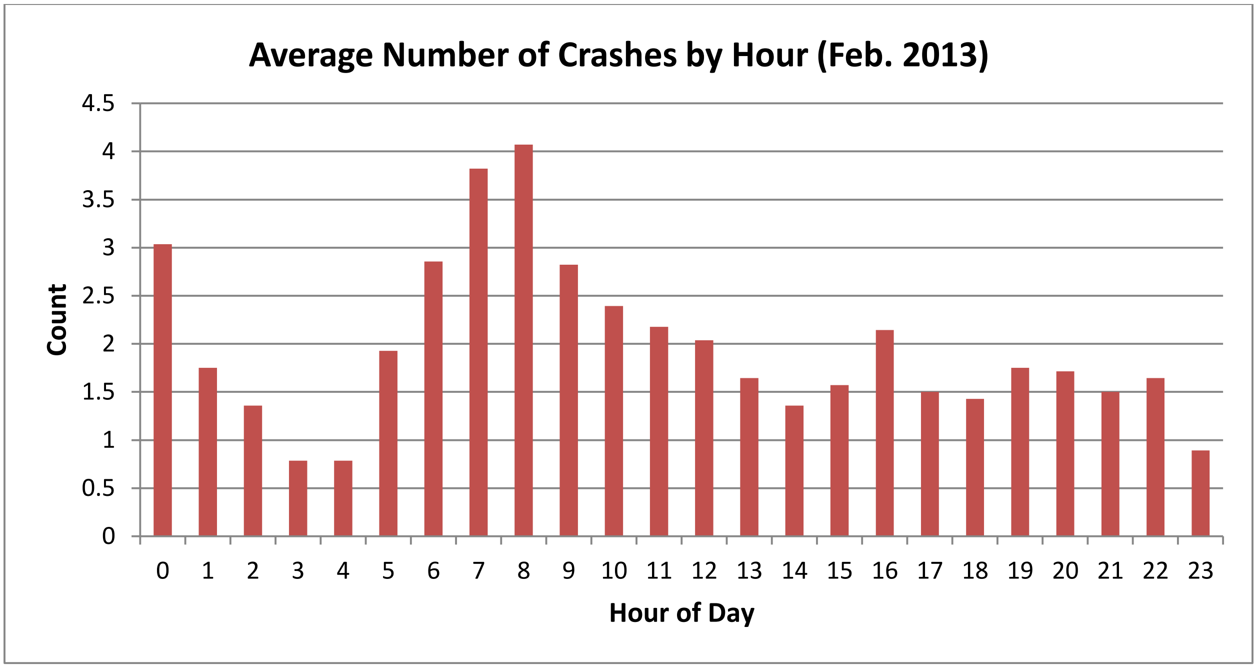


Figure 3-3 Average Daily Crash Counts by Hour (Feb. 2013)

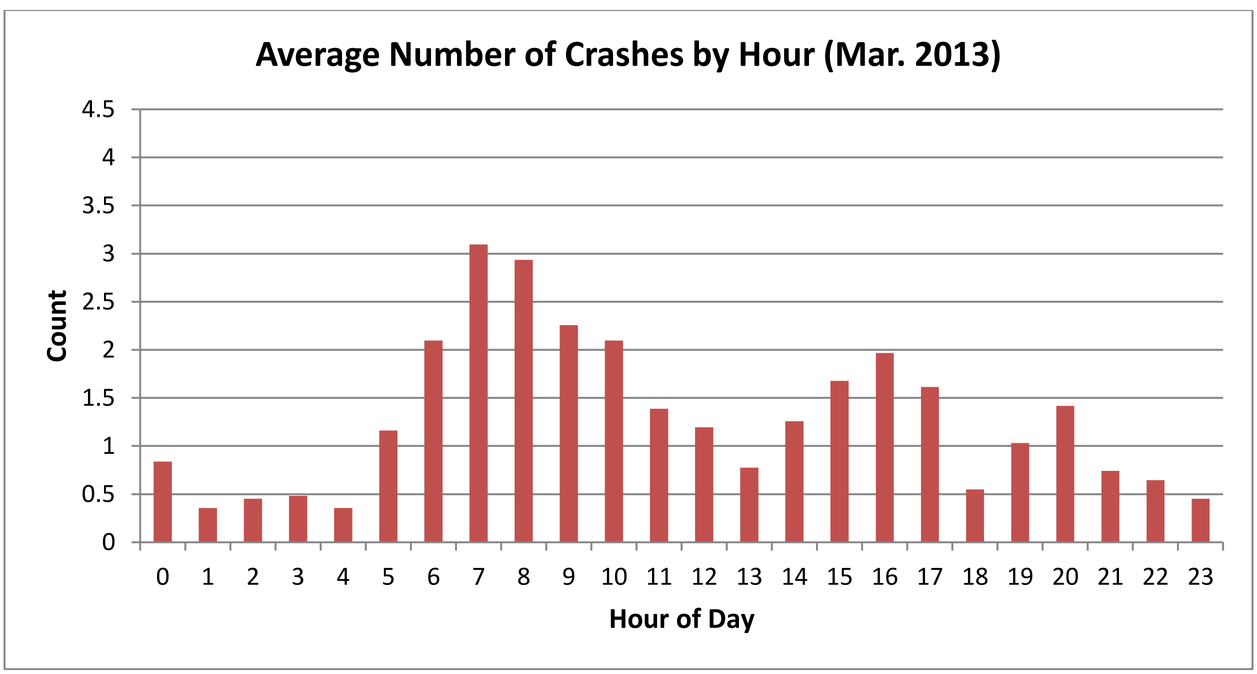


Figure 3-4 Average Daily Crash Counts by Hour (Mar. 2013)

Figure 3-5 displays the crash duration location on a Geographic Information Systems (GIS) map in March 2013. In general, incidents occurring near key state or interstate highways during peak traffic tend to generate a longer incident period which includes incident detection, response and clearance time.

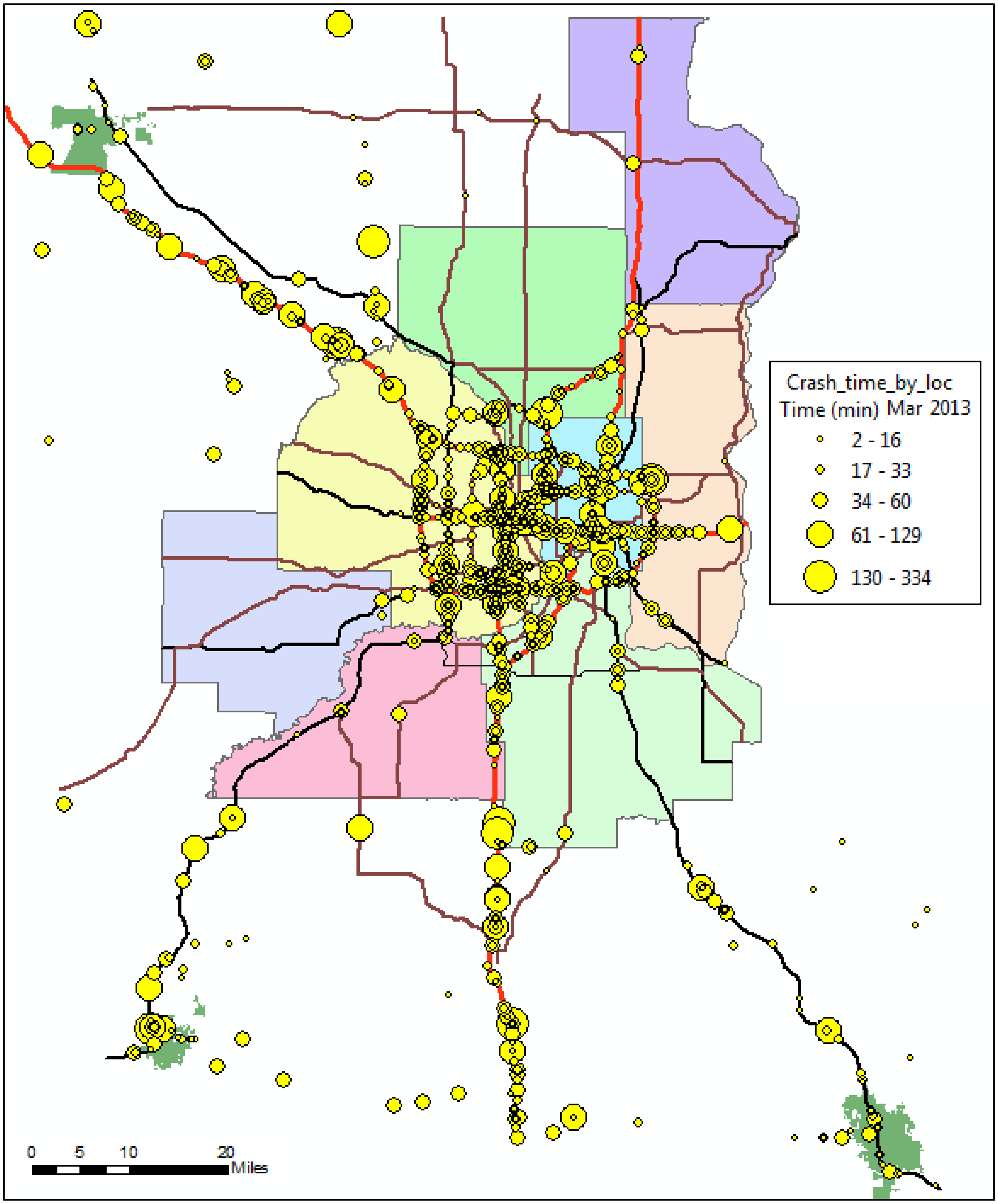


Figure 3-5 Spatial Analysis of Estimated Crash Period in TCMA in Mar. 2013

3.2 Truck GPS Data Analysis

Derived performance measures were compared with WIM data for data verification.

3.2.1 Key Freight Corridors in Twin Cities Metro Area (TCMA)

38 key freight corridors as displayed in Figure 3-6 were studied and analyzed. It consists of interstate, state highway, US highway and inter-regional corridors.

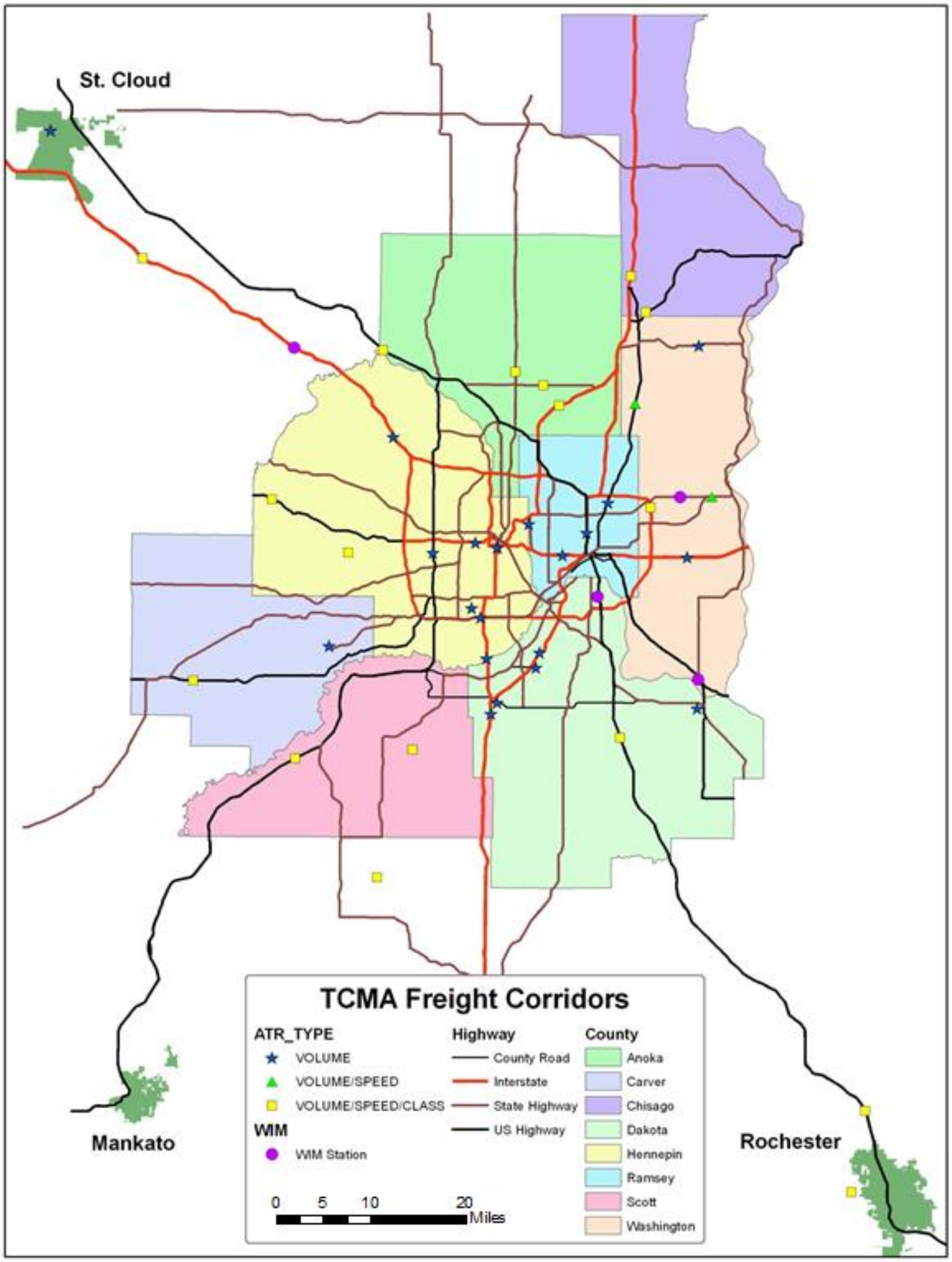


Figure 3-6 Key Freight Corridors in Twin Cities Metro Area

3.2.2 Data Processing Procedures

The flowchart that describes how raw GPS data is processed is displayed in Figure 3-7. Raw truck GPS data received from ATRI were loaded to a geospatial database (open source packages, PostgreSQL & PostGIS). SQL scripts were developed to snap each GPS point to the nearest route and to determine space mean speed (defined as the total travel distance divided by the total travel time between two GPS samples) when spot speed (instantaneous vehicle speed measured by GPS) is not available. Performance measures were aggregated by 1-mile roadway segment.

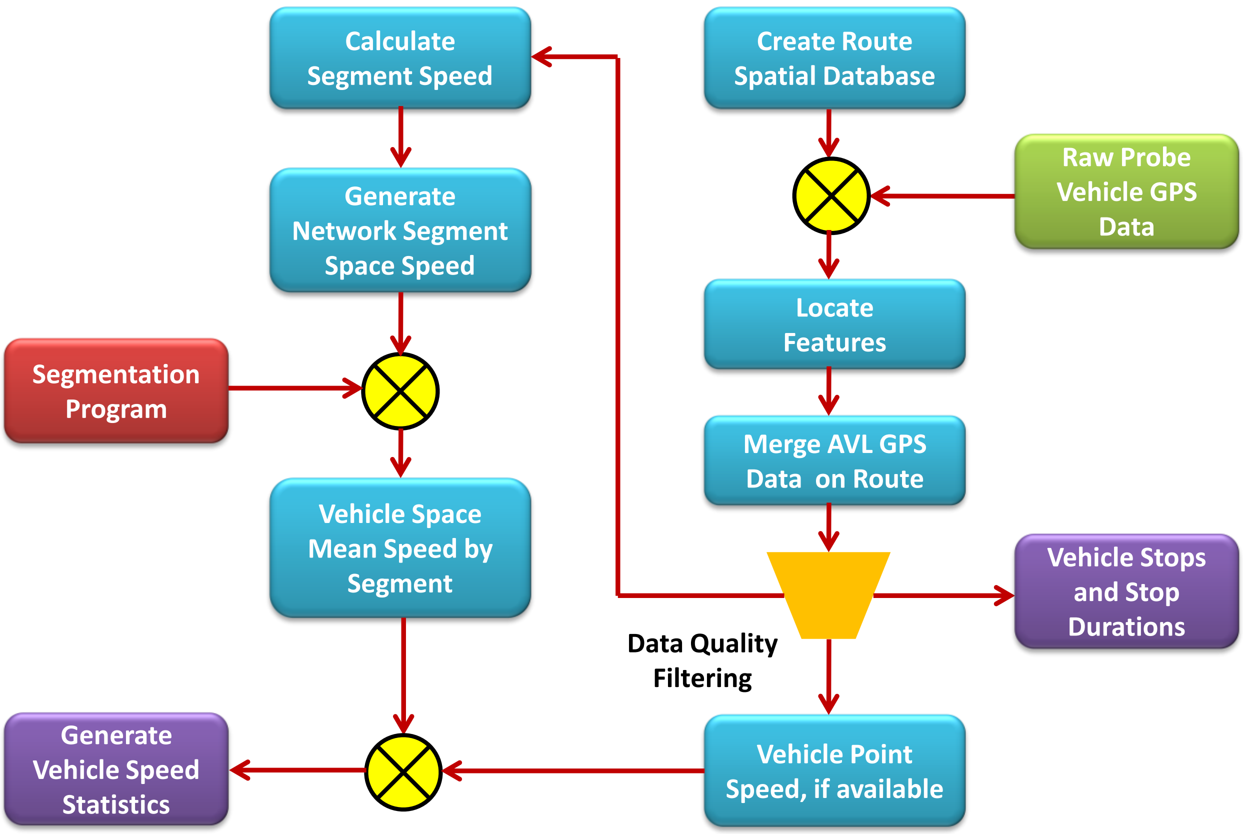


Figure 3-7 Probe Vehicle Data Analysis Flowchart

Steps:

1. Create highway route network
2. Prepare raw GPS data
3. Snap raw GPS points to a nearest route and locate snapped point distance and linear referencing value
4. Assign vehicle point GPS speed (if available) to the nearest highway segment
5. Compute vehicle speed between two consecutive GPS points and assign the average speed to roadway segments between the two consecutive GPS points
6. Analyze probe vehicle speed and travel time by roadway segment and time

3.2.3 Probe Vehicle Point Cloud

Figure 3-8 is a point cloud snapshot of probe vehicle raw GPS points in December 2012 in the Twin Cities eight-county metropolitan area. It contains over 3.5 million GPS pings from one of the three GPS datasets received from ATRI.

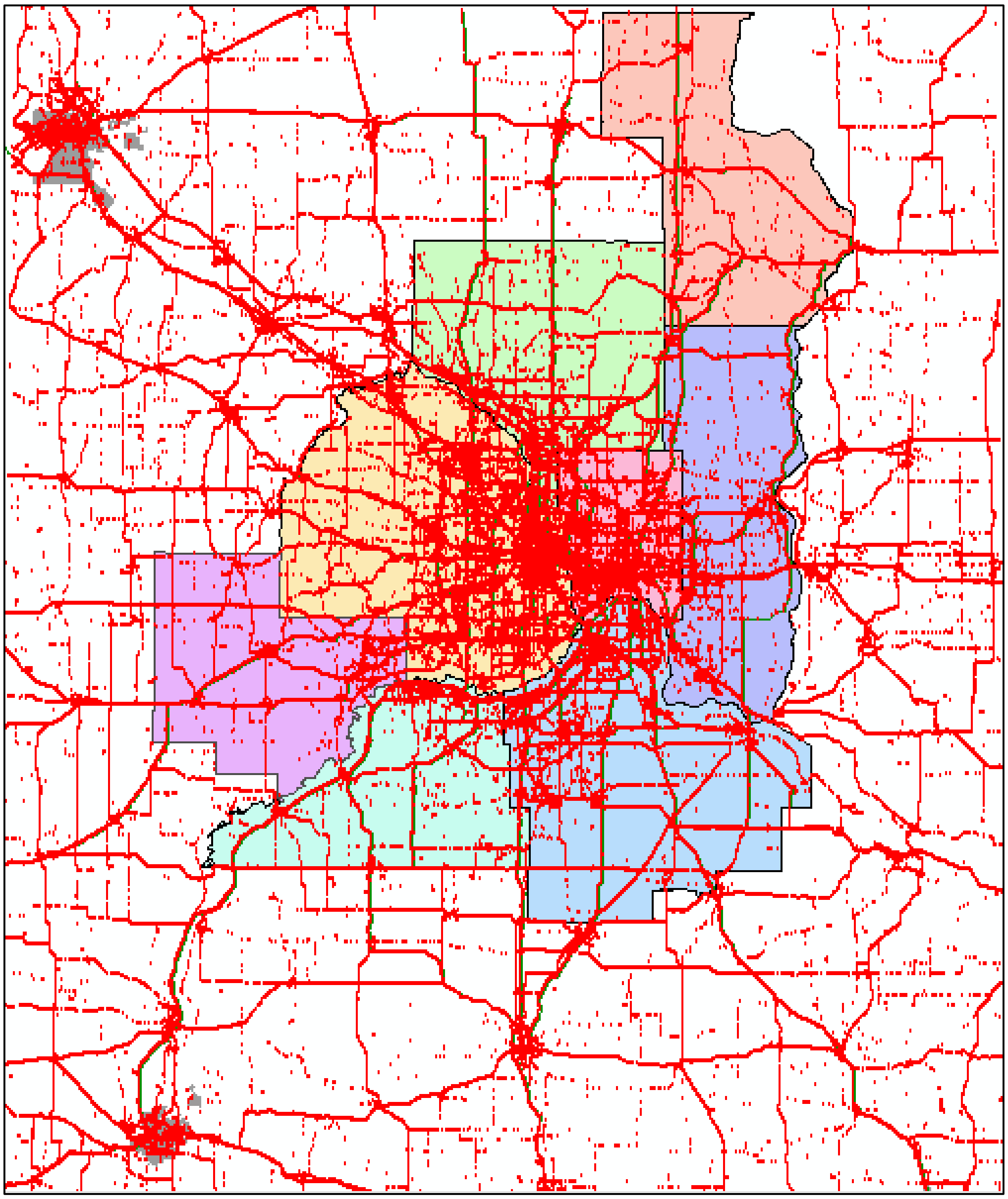


Figure 3-8 Probe Vehicle GPS Point Cloud Snapshot (Dec. 2012)

3.2.4 Processed Vehicle Speed Statistics

Figure 3-9 illustrates the computed probe vehicle speed statistics in a 1-mile segment nearby UMN campus in both directions.

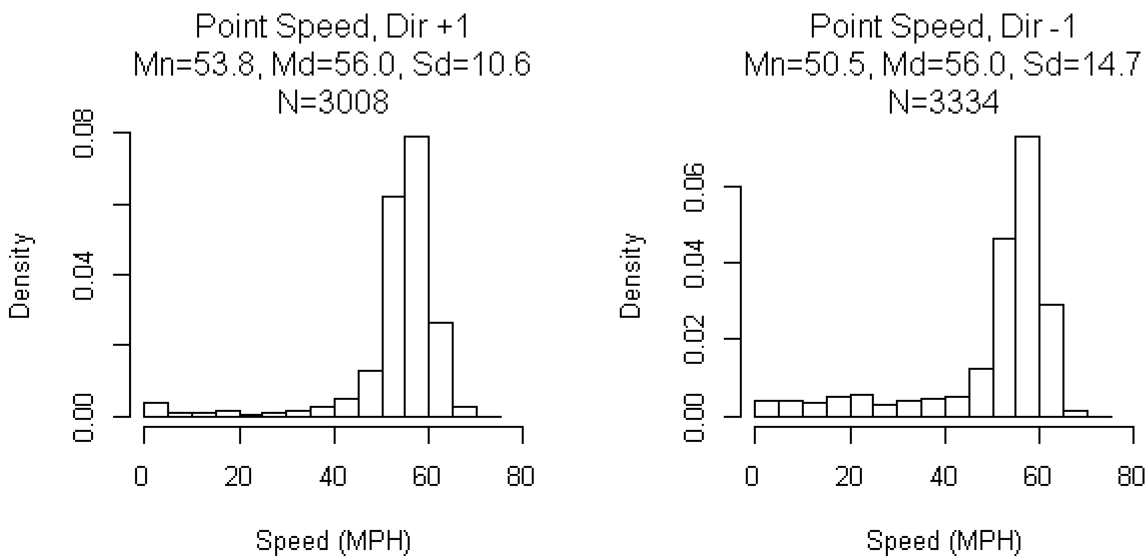


Figure 3-9 Probe Vehicle Point Speed on Route I-35W at 59th mile marker (Nearby Broadway St. NE in Minneapolis)

3.2.5 Compare Processed Results with WIM Data

We also compared the processed results with data collected from the Weight-in-Motion (WIM) station. There are four WIM stations inside the study area. A summary of the WIM stations in the Twin Cities Metro Area (TCMA) is listed in Table 3-1 as follows.

Table 3-1 Description of WIM stations

WIM ID	**36**	**37**	**40**	**42**
Route Name	MN 36	I-94	US 52	US 61
County Name	Washington	Wright	Dakota	Washington
City Name	Lake Elmo	Otsego	West St Paul	Cottage Grove
WIM Location Description	.7 MI W OF CSAH17 (LAKE ELMO AVE N) IN LAKE ELMO	1.2 MI NW OF CSAH19 (LA BEAUX AVE) IN OTSEGO	.5 MI N OF CSAH14 IN WEST ST PAUL	.4 MI S OF TH95 (MANNING AVE S), S OF COTTAGE GROVE
WIM Type	VOLUME/SPEED/CLASS/WEIGHT			

Figure 3-10 illustrates the speed comparisons and variation by hour of a day at WIM station 37. The traffic volume percentage by hour of a day is displayed in Figure 3-11. The hourly volume percentage (%) is defined as the hourly volume counts divided by the total vehicle volume. A snapshot of probe vehicle count along the highway network in TCMA between 7 and 8 AM is illustrated in Figure 3-12.

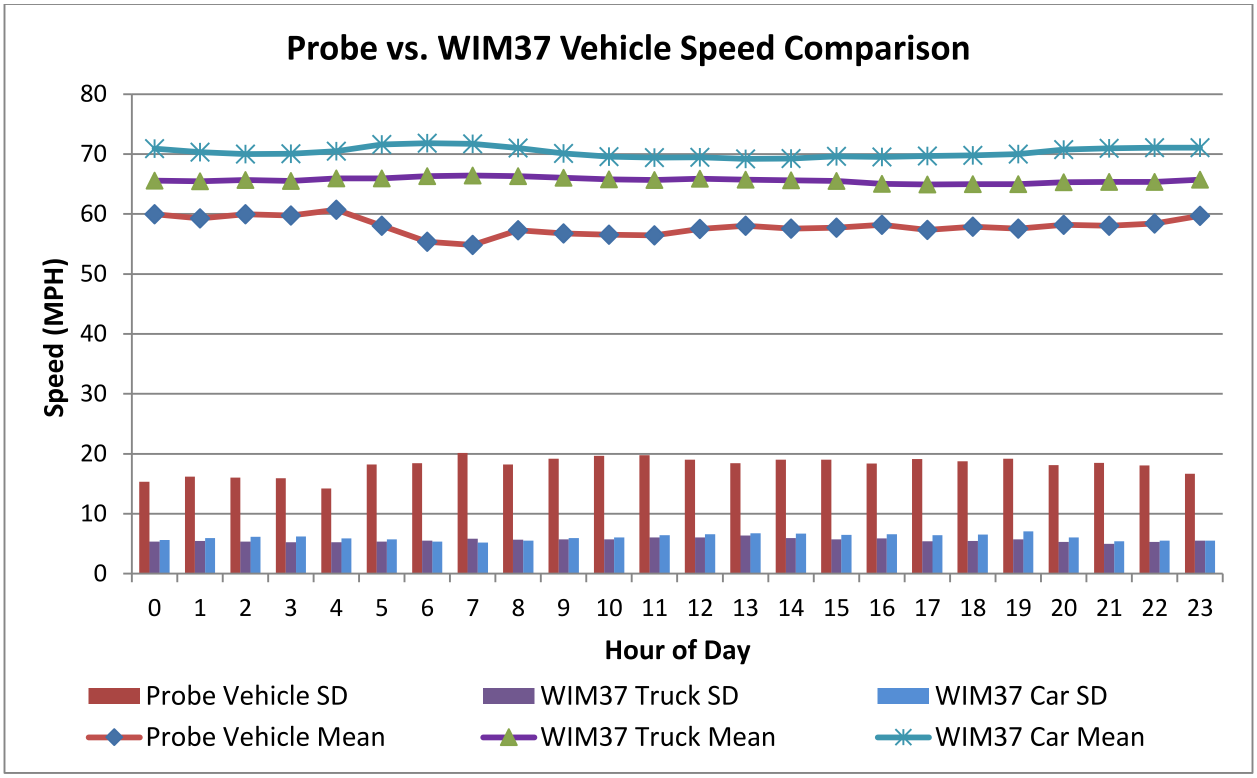


Figure 3-10 Probe Vehicle Speed vs. WIM Speed by Hour at WIM#37

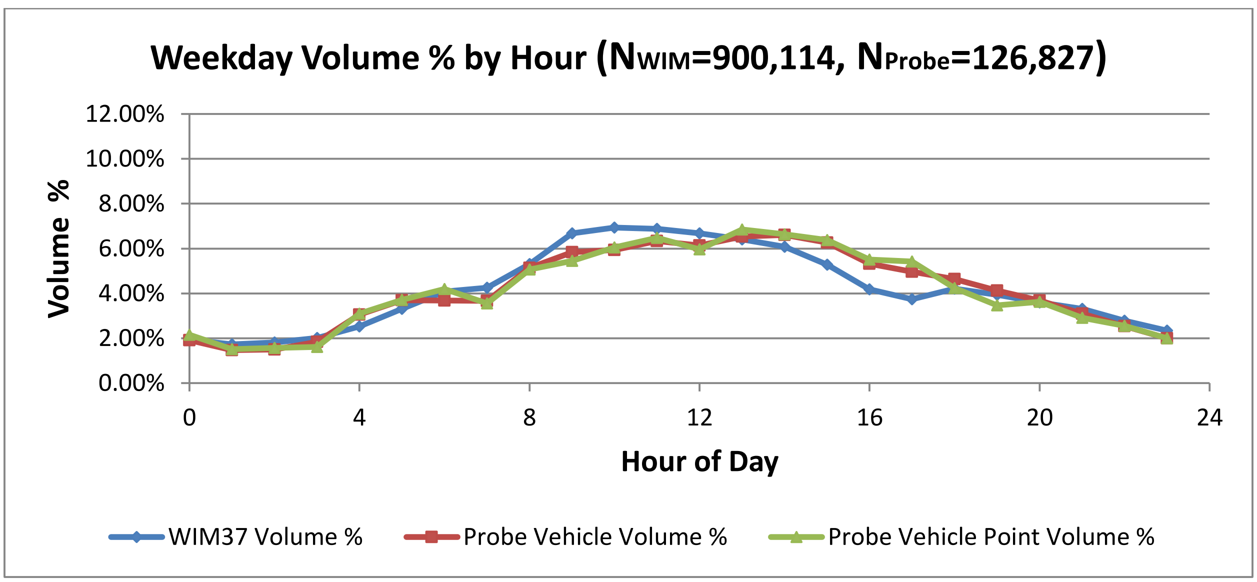


Figure 3-11 Probe Vehicle vs. WIM Volume % by Hour at WIM Station #37

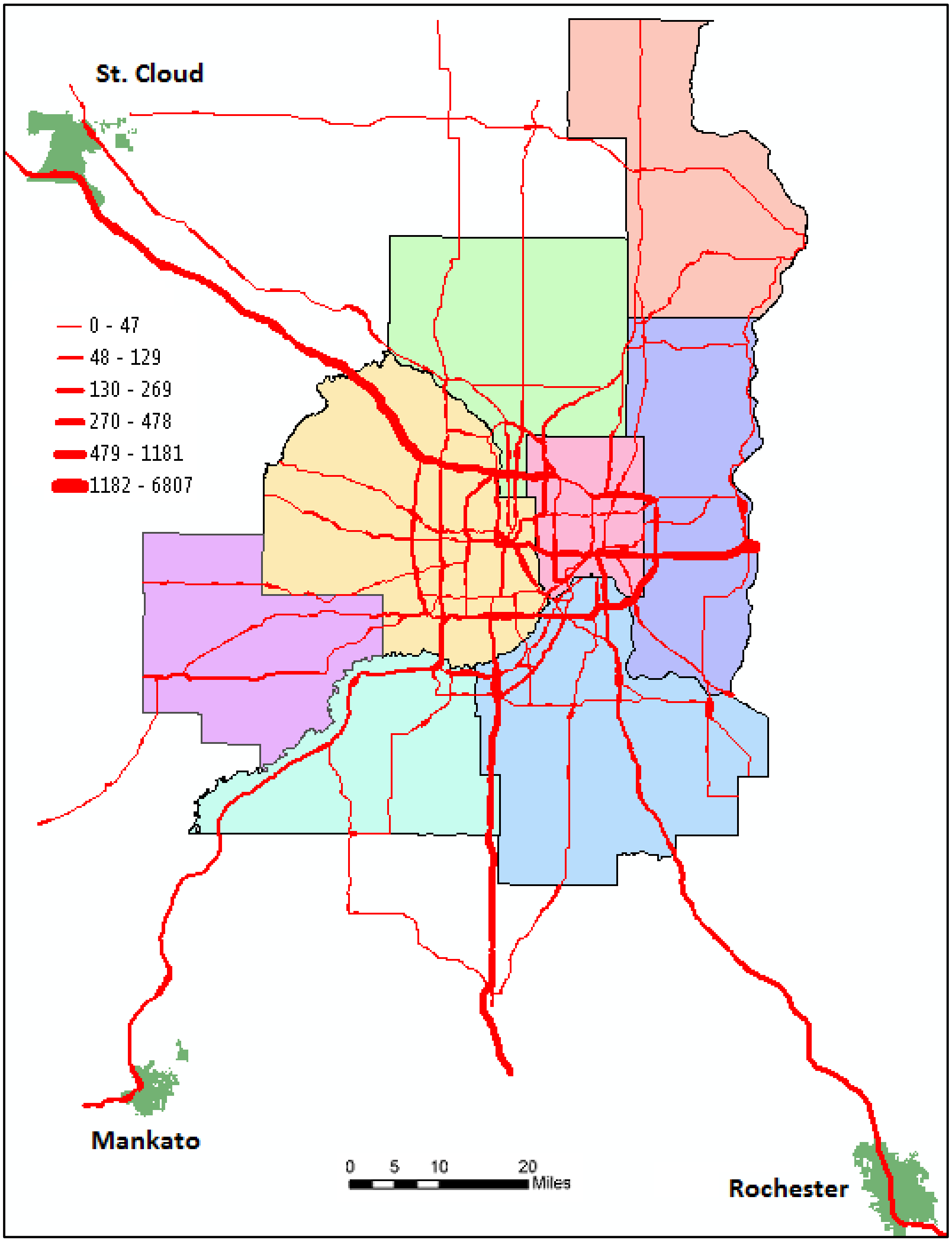


Figure 3-12 Snapshot of Probe Vehicle Count in TCMA (7~8AM in 2012)

3.3 Incident Duration Analysis and Classification

3.3.1 Estimation of Incident Duration from MN511 Data

The real-time incident data is frequently updated by the incident response team when an incident occurs. Table 3-2 listed an abbreviated sample of an incident that occurred on 9/12/2013 on I-35W. There were 4 message updates regarding this particular incident. Incident duration can be calculated based on the time difference between the update time (9:31 AM) of the "last" message and the start time (8:41AM) of the "first" message. In this example, the incident (MSPCAD-P130465520) has an estimated duration of 50 minutes (9:31 – 8:41 = 50 min). A sample of the processed incident duration estimation by incident ID is listed in Table 3-3 as an example.

Table 3-2 Abbreviated Sample of MN511 Data

msg_number	1377890597	1377890604	1377890614	1377890621
msg_timestamp	9/12/2013 8:46	9/12/2013 9:01	9/12/2013 9:16	9/12/2013 9:31
event_id	MSPCAD-P130465520	MSPCAD-P130465520	MSPCAD-P130465520	MSPCAD-P130465520
event_update	1	2	3	4
event_priority	4	4	4	4
headline_text	[3818]crash	[3818]crash	[3818]crash	[3818]crash
d_desc_phrase_text	[3818]crash	[3818]crash	[3818]crash	[3818]crash
d_loc_route_designator	I-35W	I-35W	I-35W	I-35W
d_loc_primary_loc	{44.7822, -93.2884, 3.16854}	{44.7822, -93.2884, 3.16854}	{44.7822, -93.2884, 3.16854}	{44.7822, -93.2884, 3.16854}
d_loc_link_dir	positive direction	positive direction	positive direction	positive direction
d_times_update_time	9/12/2013 8:46	9/12/2013 9:01	9/12/2013 9:16	9/12/2013 9:31
d_times_start_time	9/12/2013 8:41	9/12/2013 8:47	9/12/2013 9:04	9/12/2013 9:04

Table 3-3 Sample of Incident Duration Estimates

Message update time	**Event id**	**Estimated Duration (min)**	**Description**	**Latitude**	**Longitude**	**Estimated Start time**	**Direction**
9/12/2013 8:46	MSPCAD-P130465520	50	[3818]crash	44.7822	-93.2884	9/12/2013 8:41	positive direction
9/12/2013 9:01	MSPCAD-P130465541	36	[3818]crash	44.9705	-93.3409	9/12/2013 8:55	positive direction
9/12/2013 9:16	MSPCAD-P130465582	33	[3818]incident	45.1266	-93.4852	9/12/2013 9:13	negative direction
9/12/2013 9:31	MSPCAD-P130465603	20	[3818]stalled vehicle [3819]blocked	45.0686	-93.2633	9/12/2013 9:26	positive direction
9/12/2013 9:46	MSPCAD-P130465624	23	[3818]crash	44.8956	-93.247	9/12/2013 9:38	positive direction

3.3.2 Analysis of Incident Duration

Figure 3-13 illustrates the distribution of incidents related to crashes on major freight corridors in the Twin Cities Metro Area from Oct. 2012 to Sep. 2013. The incident duration statistics are summarized in Table 3-4. Out of the 606,376 incident records, about 12% of them are incident and the other 88% are non-incident events (i.e., driving condition is good or fair). The incident data consists of "crash" [3818], "road closed" [3819], and "debris/animals on roadway" [3822]. The crash related data is further identified into 6 groups, such as "stalled vehicle" (3,205 unique samples), "vehicle spin-out" (5,001), "incident" (975), "disabled vehicle" (3,731), "jack-knifed semi-trailer" (206), "crash" (13,460).

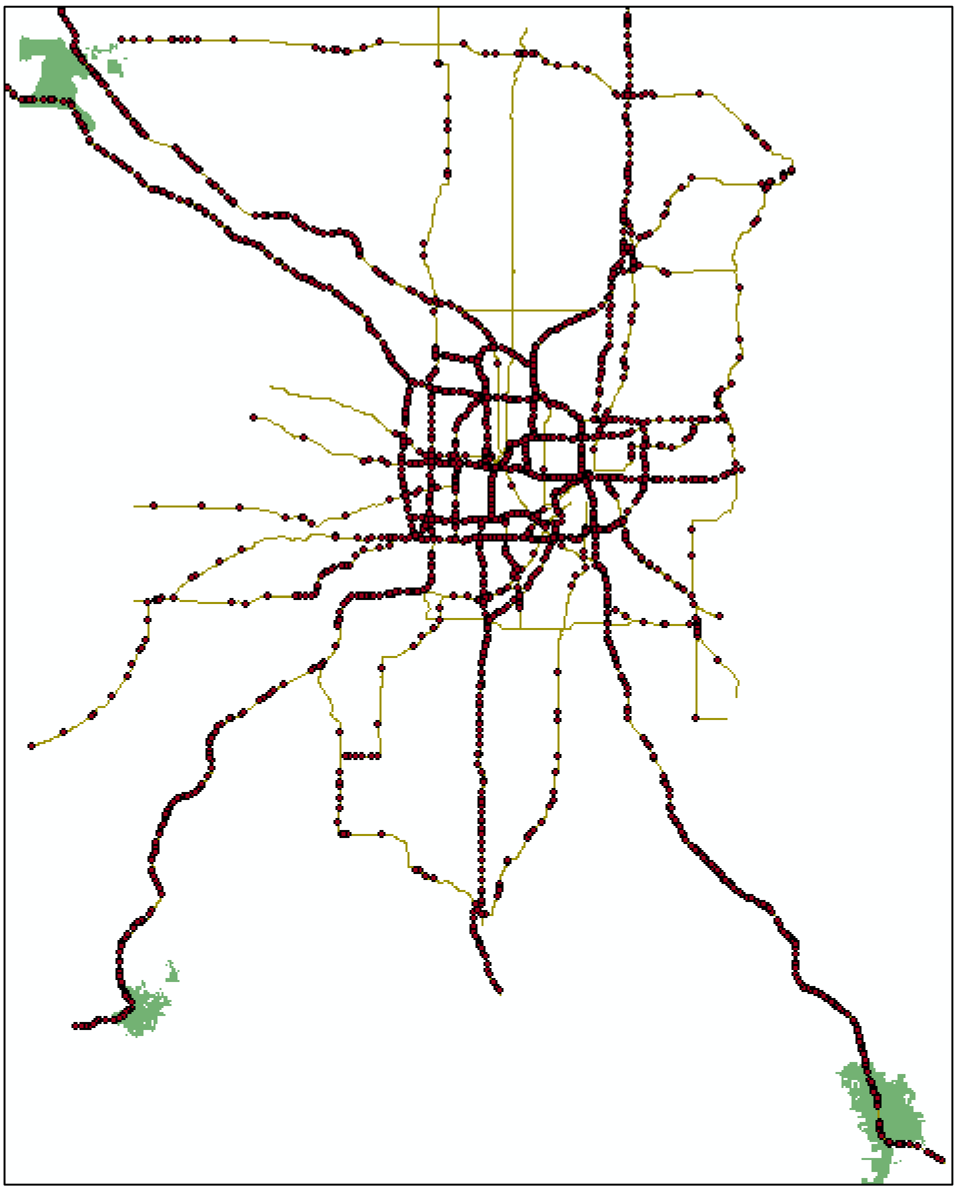


Figure 3-13 Incidents Related to Crash in Twin Cities Metro Area (Oct. 2012 – Sep. 2013)

Table 3-4 Summary of Incident Duration by Incident Type

Statistics	Crash	Disabled vehicle	Incident	Jackknifed semi-trailer	Stalled vehicle	Vehicle spun out
Mean (min)	35.0	27.8	22.1	55.3	17.8	28.6
Standard Deviation (min)	42.1	54.5	44.8	44.3	35.5	38.5
Size (N)	13,460	3,731	975	206	3,205	5,001
AM Peak (5-9 AM) Mean (min)	33.7	27.4	35.8	55.5	19.6	28.8
AM Data Size (N1)	2812	760	100	43	498	1090
PM Peak (3-7 PM) Mean (min)	33.7	24.5	22.6	65.6	17.1	25.6
PM Data Size (N2)	4443	725	247	40	950	987
Total Peak Size (N1+N2)	7255	1485	347	83	1448	2077
AM & PM Data Size Percentage (N1+N2)×100% /N	53.9%	39.8%	35.6%	40.3%	45.2%	41.5%

Duration and standard deviation of crash related incidents are plotted in Figure 3-14. Crash involved with a semi-truck has the highest duration in average. Stalled vehicle has the shortest average duration.

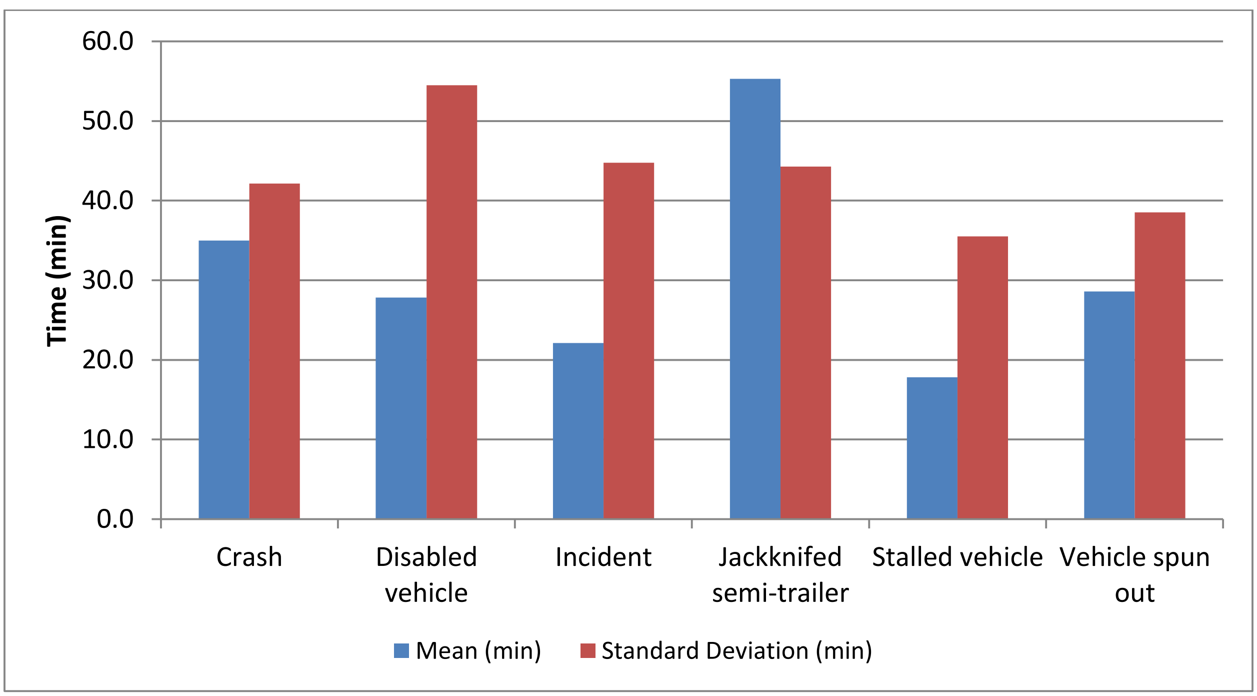


Figure 3-14 Incident Durations

The incident data are spatially joined with road network, traffic volume (AADT) and number of lanes to associate each incident location with roadway geometry information. A sample of the spatially joined data is listed in Table 3-5 as follows.

Table 3-5 Sample of Spatially Joined Incident Data

Duration (min)	Mon	Hour	Crash Type	Route	SEG ID	Highway Type	Route ID	AADT	Lane	AADT_LN
40	9	16	1	94	35	1	24	56000	2	28000
42	9	16	1	494	33	1	6	105000	2	52500
44	9	16	1	94	37	1	24	39000	3	13000
50	9	16	1	494	18	1	6	125000	3	41666.67
54	9	16	1	100	9	3	7	105000	3	35000
54	9	16	1	35	49	1	34	104000	3	34666.67
81	9	16	1	94	24	1	24	149000	3	49666.67

Where,

Crash Type:

7 – [3818] All other misc. crash (Lane reduced)
6 – [3818] vehicle spun out
5 – [3818] stalled vehicle [3819] blocked
4 – [3818] jackknifed semi-trailer
3 – [3818] incident
2 – [3818] disabled vehicle
1 – [3818] crash

Highway Type

1 – Interstate
2 – US Highway
3 – State Highway
4 – County Road

Traffic Direction

1 – Positive Direction
2 – Negative Direction
3 – Not Directional

3.3.3 Incident Duration Probability Distribution

The histogram of all incident durations calculated from MN511 data was analyzed, as displayed in Figure 3-15, using the R statistical software package. The distribution was multimodal with at least 3 distinctive peaks at around 4, 20, and 34 minutes. Further analysis was performed to examine the histogram for each crash type. The results were listed in Table 3-6, 11% of the incidents were related to vehicle “spin-out”. Disabled vehicles and stalled vehicles consist of 9% and 14 % of all incidents, respectively. The majority of incidents (63%) were classified as crashes. The type 1 incident (Crash) has a similar histogram as compared to the histogram of all incident types shown in Figure 3-15. There are at least 3 modes in the duration distribution of the crash incidents. Information on additional parameters (such as number of vehicles involved, fatality, and others) is needed to separate crash types and to model the duration.

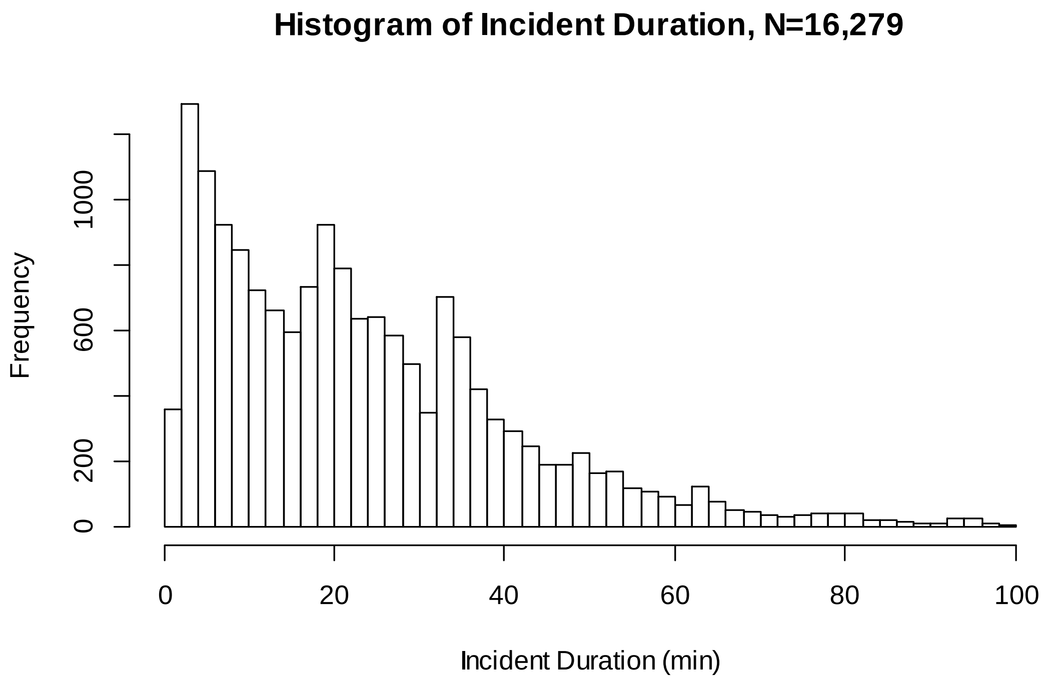


Figure 3-15 Histogram of All Incident Durations from MN511 Data

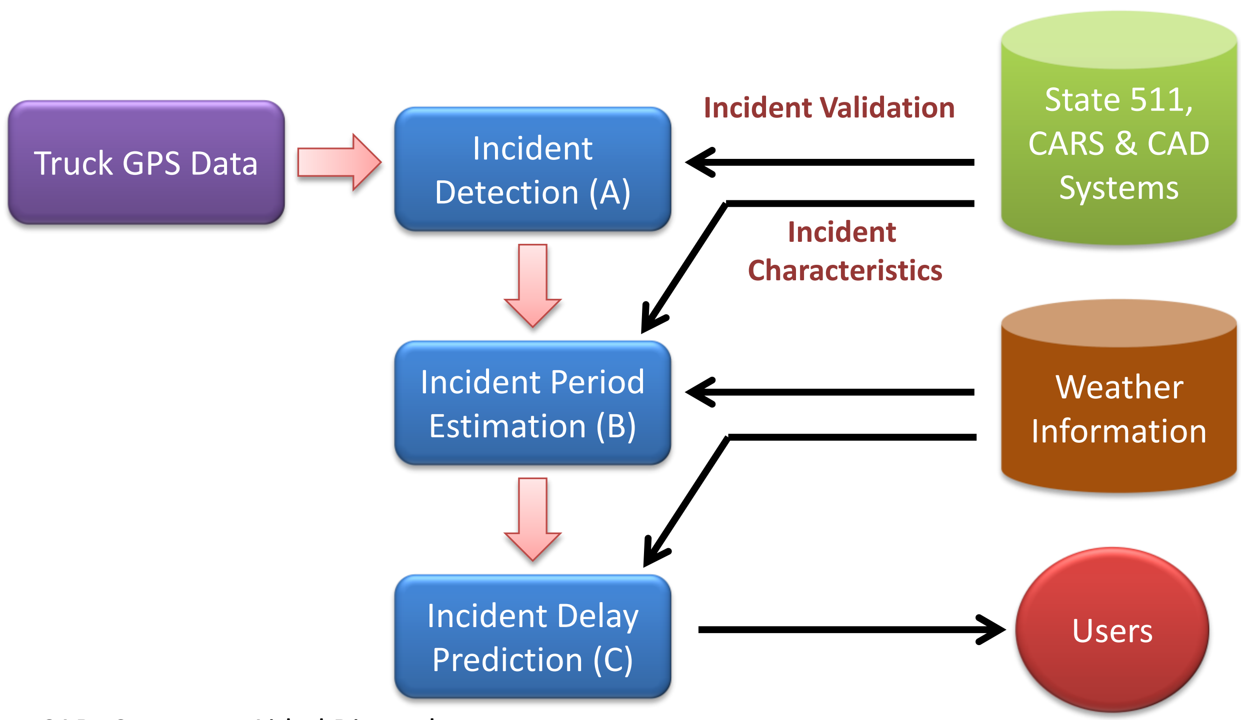
Table 3-6 Incident Duration Statistics

Incident Type	Count	Percentage	Mean	Median	SD
1 – crash	10,329	63.4%	27.3	25.0	18.1
2 – disabled vehicle	1,423	8.7%	20.4	15.0	18.3
3 – incident	467	2.9%	16.2	12.0	14.1
4 – jackknifed semi-trailer	30	0.2%	38.0	33.5	25.0
5 – stalled vehicle	2,228	13.7%	14.1	11.0	12.0
6 – vehicle spun out	1,792	11.0%	22.4	17.0	19.0
7 – All other misc. crash	10	0.1%	32.5	20.0	27.6

4. Methodologies

4.1 Incident Detection Algorithm

The incident decision support system proposed in this study is shown in Figure 4-1. The incident detection algorithm using a bivariate analysis model as discussed in this section is illustrated in block (A) in Figure 4-1. Methodologies discussed in section 4.1.1 and 4.1.2 were implemented to detect incidents using probe truck GPS data. Incident delay estimation illustrated in block (B) in Figure 4-1 is included in section 4.2. Incident delay prediction results described at the end of section 4.2 using the Random Forest algorithm performed better than the other algorithms. However, it is inadequate.



CAD: Computer Aided Dispatch
CARS: Condition Acquisition and Reporting System

Figure 4-1 Schematic Diagram of Incident Decision Support System

4.1.1 Bivariate Analysis Model for Incident Detection

Li & McDonald (2005) discovered that the joint distribution of travel time and travel time difference is bivariate normal in non-incident conditions. They developed a probe vehicle based algorithm using a bivariate analysis model to analyze travel time data for incident detection along four segments of motorways in UK. Travel time (T_i) and travel time difference ($\Delta T_i = T_i - T_{i-1}$) between adjacent time intervals are used, where *i* is the current time interval. The joint density function of a bivariate Gaussian distribution, $f(T_i,\ \Delta T_i)$, can be expressed as follows (Wilks, 2006).

$$f(T_i,\ \Delta T_i) = \left(2\pi\sigma_{T_i}\sigma_{\Delta T_i}\sqrt{(1-\rho_i^2)}\right)^{-1} e^{-Q/2}$$

Eq. (4-1)

$$Q(T_i,\ \Delta T_i) = \frac{1}{1-\rho_i^2}\left[\left(\frac{T_i-\mu_{T_i}}{\sigma_{T_i}}\right)^2 + \left(\frac{\Delta T_i-\mu_{\Delta T_i}}{\sigma_{\Delta T_i}}\right)^2 - 2\rho_i\left(\frac{T_i-\mu_{T_i}}{\sigma_{T_i}}\right)\left(\frac{\Delta T_i-\mu_{\Delta T_i}}{\sigma_{\Delta T_i}}\right)\right]$$

Eq. (4-2)

Or, in quadratic form, $Q(T_i,\ \Delta T_i) = X^T\ \Sigma^{-1}X$

Where,

Error Vector, $X = \begin{bmatrix} (T_i-\mu_{T_i}) \\ (\Delta T_i-\mu_{\Delta T_i}) \end{bmatrix}$

Eq. (4-3)

Variance-Covariance Matrix, $\Sigma = \begin{bmatrix} \sigma_{T_i}^2 & \rho_i\sigma_{T_i}\sigma_{\Delta T_i} \\ \rho_i\sigma_{T_i}\sigma_{\Delta T_i} & \sigma_{\Delta T_i}^2 \end{bmatrix}$

Eq. (4-4)

Correlation Coefficient, $\rho_i = \frac{Cov(T_i,\Delta T_i)}{\sigma_{T_i}\sigma_{\Delta T_i}} = \frac{\boldsymbol{E}\left[(T_i-\mu_{T_i})(\Delta T_i-\mu_{\Delta T_i})\right]}{\sigma_{T_i}\sigma_{\Delta T_i}}$

Eq. (4-5)

σ_{T_i} and $\sigma_{\Delta T_i}$are the standard deviations of T_i and ΔT_i, respectively.

The *k*-value, $(T_i,\ \Delta T_i)$, describes an ellipse in the $(T_i, \Delta T_i)$ plane with center at$\left(\mu_{T_i}, \mu_{\Delta T_i}\right)$. The *k* value is equal to Chi-Square value, $\chi^2(\alpha, 2)$. The elliptic contour will contain $100(1-\alpha)$% of the sample points on average. When a sample set of $(T_i, \Delta T_i)$ lies inside the contour, the following equation should be fulfilled and considered as non-incident data using 99% coverage (i.e., $\alpha = 0.01$).

$$k\text{-}value\text{: } X^T\ \Sigma^{-1}X \leq \chi^2(0.01, 2)$$

Eq. (4-6)

The k-value defined in equation (4-6) represents an index of abnormality where a measurement is located with respect to a two dimensional normal distribution function. A measurement sample with k-value greater than 9.21is considered abnormal using 99% confidence interval.

Similarly, the above model can be expanded to analyze travel time data for incident detection along equally divided roadway segments (e.g., 1 mile). We used the segment travel time (T_n) and segment travel time difference ($\Delta T_n = T_n - T_{n-1}$) between adjacent roadway segments for incident detection, where n is the n^{th} roadway segment.

Table 4-1 Chi-Square Distribution Table

Chi-Square	**Probability**					
Degrees of Freedom (DOF)	**0.01**	**0.05**	**0.1**	**0.9**	**0.95**	**0.99**
1	0.00	0.00	0.02	2.71	3.84	6.63
2	0.02	0.10	0.21	4.61	5.99	**9.21**
3	0.11	0.35	0.58	6.25	7.81	11.34
4	0.30	0.71	1.06	7.78	9.49	13.28
5	0.55	1.15	1.61	9.24	11.07	15.09
6	0.87	1.64	2.20	10.64	12.59	16.81
7	1.24	2.17	2.83	12.02	14.07	18.48
8	1.65	2.73	3.49	13.36	15.51	20.09
9	2.09	3.33	4.17	14.68	16.92	21.67
10	2.56	3.94	4.87	15.99	18.31	23.21

4.1.2 Outlier Detection Using Hampel Identifier

Hampel (1971) introduced the concept of the breakdown point. The breakdown point is the smallest percentage of contaminated data (or outliers) that can cause an estimator to take arbitrary large aberrant values. The median and the median absolute deviation (MAD) are often recommended for robust estimations. For example, consider a data series, $\{x_i\}$, where $i = 1$ to n. The MAD scale estimate (S) is defined as,

$$S = 1.482602\ median\{\ |x_i - x^m|\ \}$$

Eq. (4-7)

Where,

x^m is the median of a data series $\{x_i\}$,

Factor 1.4826 was chosen so the expected value of MAD is equal to the standard deviation (σ) for normally distributed data. That is,

$$\boldsymbol{E}[S(x_1 \ldots x_n)] = \sigma \tag{4-8}$$

Eq. (4-8)

For $\{x_i\}$ distributed as N (μ, σ^2) and large *n*.

If $|x_i - x^m| > t\,S$, x_i is considered as an outlier, where t is the rejection threshold often suggested to be around 2 to 5 as proposed by Pearson (2002).

The Hampel identifier is used to remove potential speed outliers from the GPS data processing.

4.1.3 Incident Analysis Examples

Two incident analyses were performed to evaluate the incident detection methodology. Section A describes an incident that occurred on 1/23/2013 on the I-94 eastbound west of Lowry tunnel in Minneapolis. Section B presents the analysis results from an incident occurring on 2/5/2013 along I-494 southbound in Minnetonka.

A. I-94 near Lowry Tunnel

Figure 4-2 and 4-3 illustrate the hourly average travel speed of trucks on weekdays in January 2013 from milepost 24 to 36 near the Lowry tunnel in Minneapolis for both westbound and eastbound traffic, respectively. Each line in the graph represents an hourly speed average. For example, the purple line with the x mark represents the average truck speed at 7AM, the blue line with the diamond mark represents the average truck speed at 8AM, the red line with the square mark represents the average truck speed at 9AM, and the light green line with the triangle mark represents the average truck speed at 10AM. In general, the average truck speed drops as vehicles approach the Lowry tunnel during the AM peak hours for both directions.

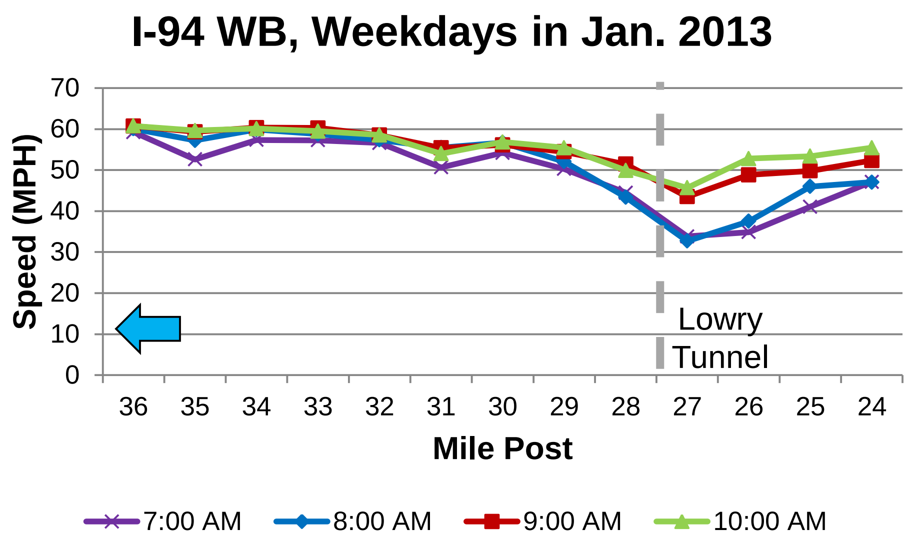


Figure 4-2 Average Weekday Travel Speed (I-94 WB near Lowry Tunnel)

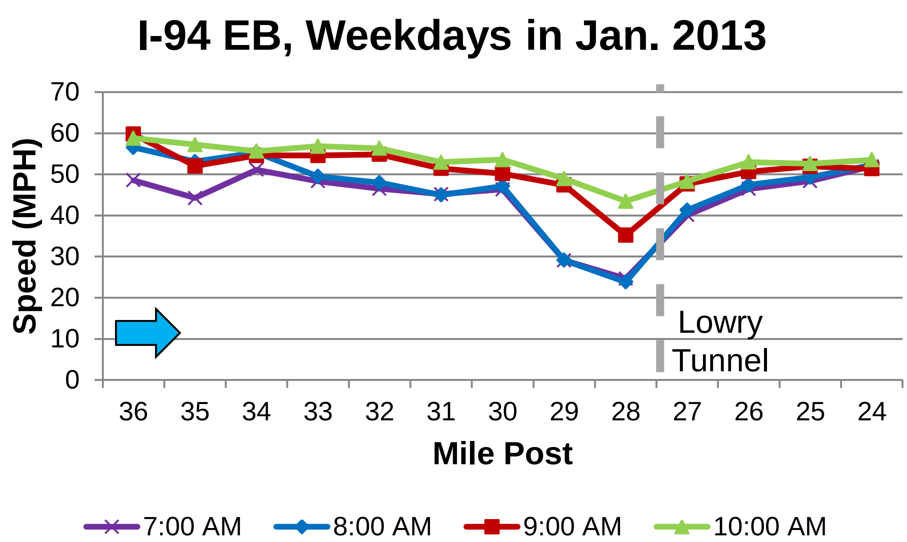


Figure 4-3 Average Weekday Travel Speed (I-94 EB near Lowry Tunnel)

Figure 4-4 and 4-5 display the hourly average travel speed of trucks on January 23rd, 2013 in both directions at the same roadway segments as described in Figure 4-2 and 4-3. In Figure 4-4, the westbound traffic maintained a similar traffic pattern as compared to the weekday hourly average shown in Figure 4-2. However, the hourly speed curve in the eastbound direction, shown in Figure 4-5, illustrates a significant disturbance over the AM peak hours along the roadway segment (milepost 36 to 24) due to a crash that occurred near milepost 29.

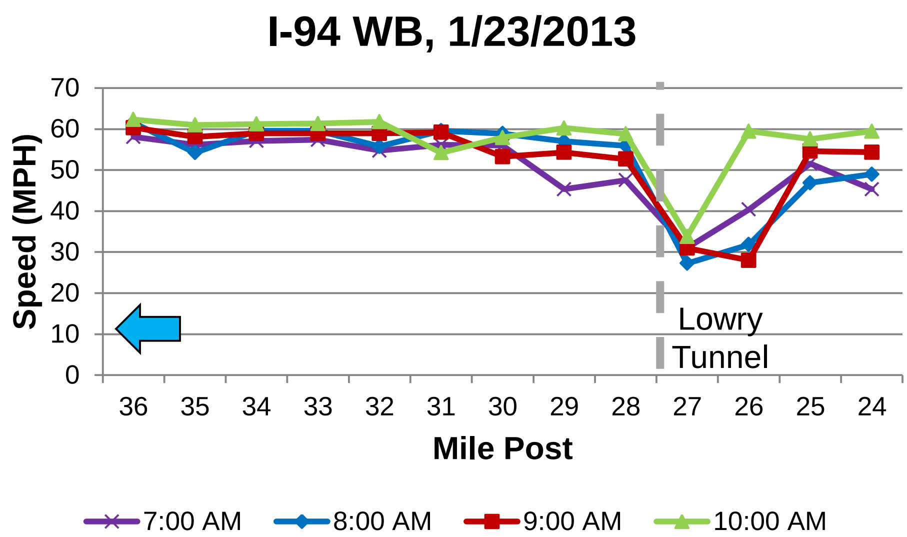


Figure 4-4 Average Travel Speed (I-94 WB near Lowry Tunnel, Jan. 23, 2013)

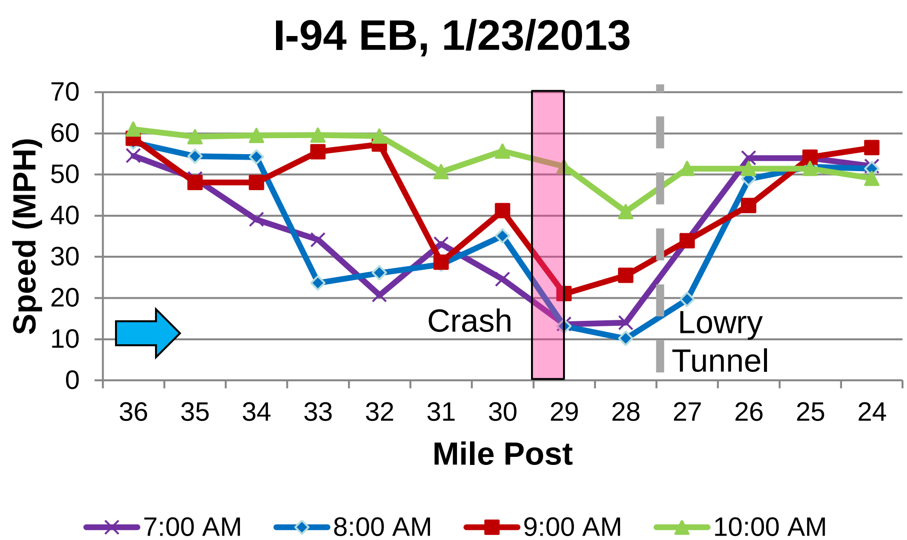


Figure 4-5 Average Travel Speed (I-94 EB near Lowry Tunnel, Jan. 23, 2013)

Figure 4-6 illustrates the computed k-value (defined in equation 4-6) from the bivariate analysis. The graph indicated a significant traffic delay (k-value greater than 9.21) approximately between 7:00AM and 8:00AM as a result of a crash and the AM peak hour traffic near the I-94 at Lowry tunnel.

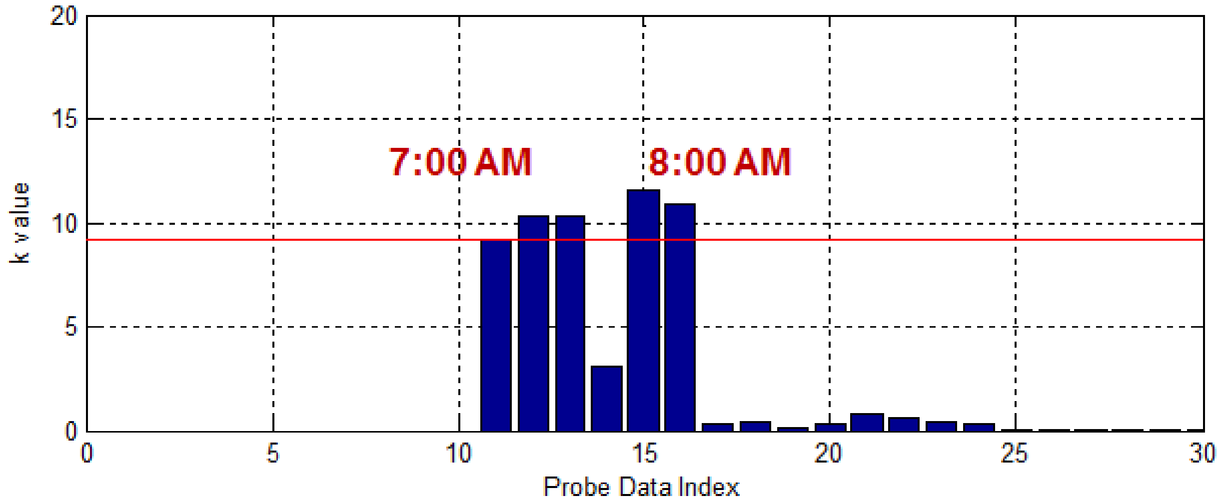


Figure 4-6 Distribution of k-value from bivariate analysis (I-94 EB near I-394, Jan. 23, 2013)

B. I-494 near I-394

Figure 4-7 and 4-8 illustrate the hourly average travel speed of trucks on weekdays in February 2013 from milepost 29 to 41 near the I-394 and I-494 interchange in Minnetonka for both northbound and southbound traffic, respectively. Each line in the graph represents an hourly speed average. For example, the purple line with the x mark represents the average truck speed at 7AM, the blue line with the diamond mark represents the average truck speed at 8AM, the red line with the square mark represents the average truck speed at 9AM, and the light green line with the triangle mark represents the average truck speed at 10AM. In general, the average truck speed drops as vehicles approach the Lowry tunnel during the AM peak hours for both directions.

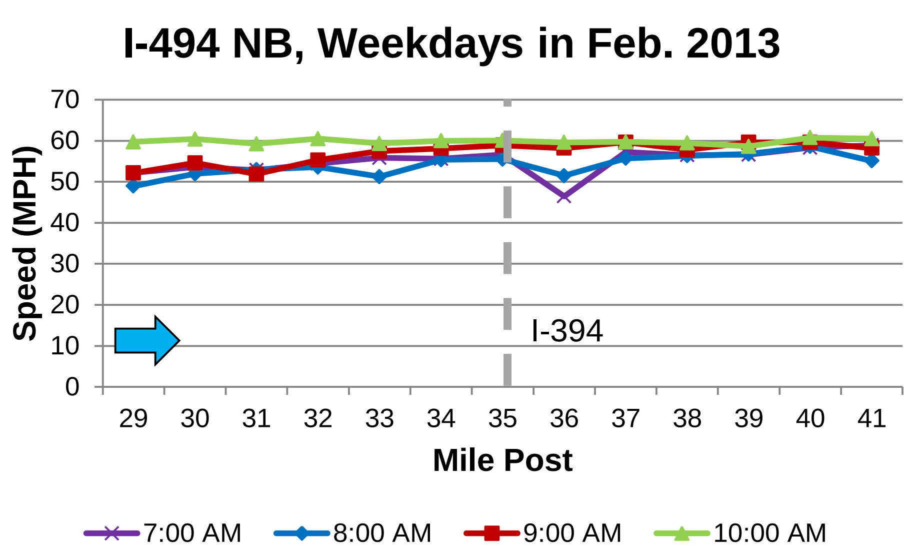


Figure 4-7 Average Weekday Travel Speed (I-494 NB near I-394)

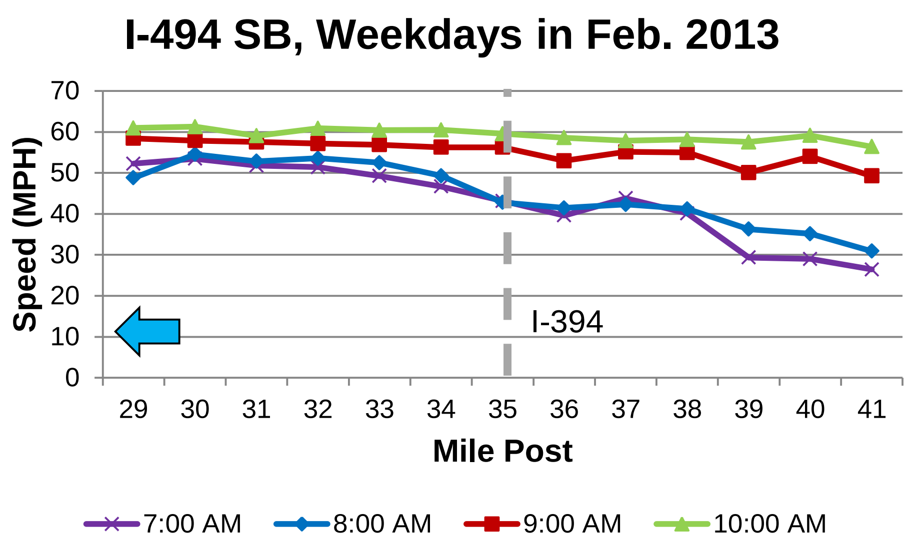


Figure 4-8 Average Weekday Travel Speed (I-494 SB near I-394)

Figure 4-9 and 4-10 display the hourly average travel speed of trucks on February 5th, 2013 in both directions at the same roadway segments as described in Figure 4-7 and 4-8. In Figure 4-9, the northbound traffic during 7 to 8AM period is slightly deviated from the weekday traffic pattern as compared to the weekday hourly average shown in Figure 4-7. However, the hourly speed curve in the southbound direction, shown in Figure 4-10, illustrates significantly slower speed during the 7 to 8AM period along the roadway segment (milepost 41 to 29) due to a crash that occurred near milepost 35.

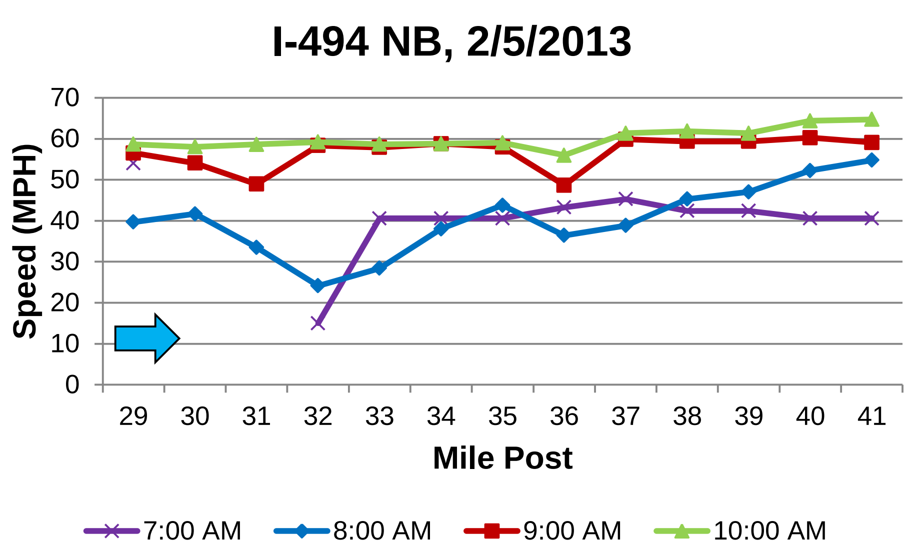


Figure 4-9 Average Travel Speed (I-494 NB near I-394, Feb. 5, 2013)

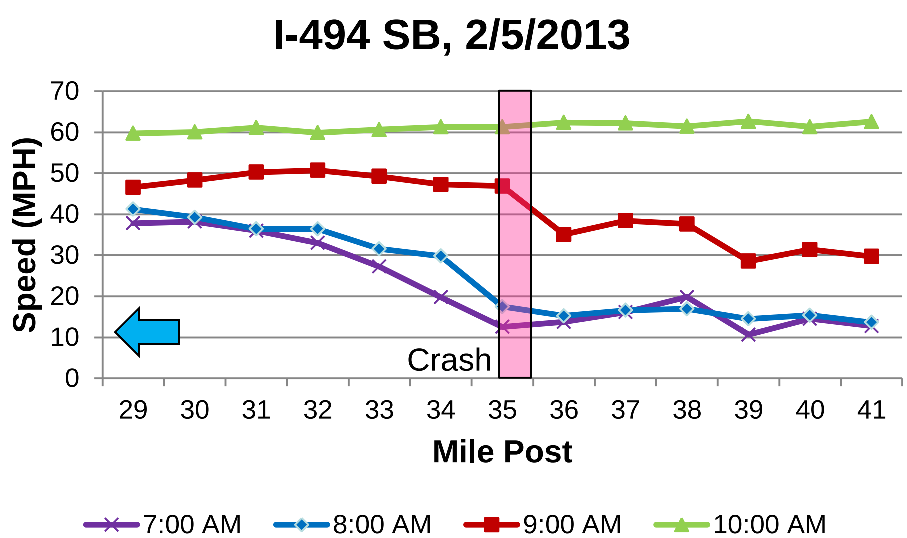


Figure 4-10 Average Travel Speed (I-494 SB near I-394, Feb. 5, 2013)

Figure 4-11 illustrates the computed k-value from the bivariate analysis. There were many measurements with k-value greater than 9.21 threshold. The graph indicated a significant traffic delay approximately between 5:30AM and 7:45AM as a result of a crash and the AM peak hour traffic near the I-494 and I-394 interchange.

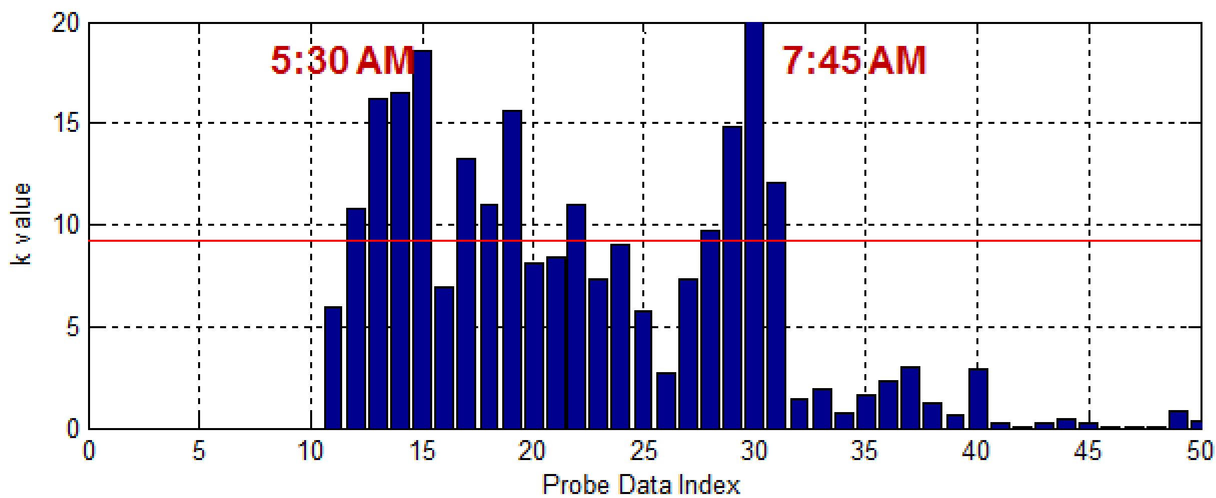


Figure 4-11 Distribution of k-value from bivariate analysis (I-494 SB near I-394, Feb. 5, 2013)

4.2 Incident Delay Estimation

The goal is to develop a model based on incident type, time, weather info, number of vehicles involved, roadway type, and other related parameters to predict incident delay.

Estimated incident duration (T_1) from MN511 data can be described as equation (4-9). It was used to validate the duration estimation derived from GPS data as described in equation (4-10) using the bivariate analysis described in section 4.1.1.

$$Duration\ Estimate\ (T_1) = f_1(MN511\ Data)$$

Eq. (4-9)

$$Duration\ Estimate\ (T_2) = t_i(Bivariate\ speed\ recovery) - t_i(Bivariate\ speed\ drop)$$

Where,

t_i is the time of an event occurs at roadway segment i.

Eq. (4-10)

In addition, vehicle delay estimation from the processed GPS data can be described as,

$$Length\ of\ Congestion\ (L_n) = S_{i-n}(Bivariate\ speed\ drop) - S_i(Bivariate\ speed\ drop)$$

Eq. (4-11)

Where,

S_i is the roadway segment i when incident occurred,

S_{i-n} is the last segment upstream from segment I where speed drops lower than normal, and

L_n is the length of congestion.

$$Vehicle\ Delay\ Estimate\ (T_3) = {L_n}/{Speed_c}$$

Eq. (4-12)

Where,

$Speed_c$ is the average congestion speed between segment i and i-n, and

L_n is the length of congestion.

Predicting traffic incident delay is a challenging task in Advanced Traffic Incident Management (ATIM). The duration of an incident delay consists of the incident time period (detection, response, and clearance time periods) and the recovery time as illustrated in Figure 4-12.

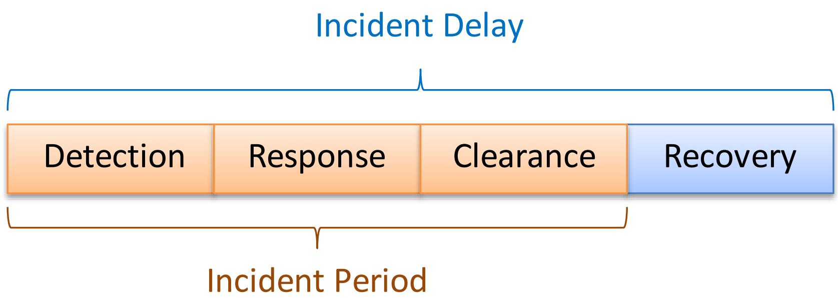


Figure 4-12 Illustration of Incident Period and Incident Delay

The total incident delay time can be described as follows.

$$Incident\ delay\ (T_4) = t_{i-n}(Bivariate\ speed\ recovery) - t_i(Bivariate\ speed\ drop)$$

Where,

t_i is the time of the speed drop at roadway segment i when incident occurred, and
t_{i-n} is the time of speed recovery at n^{th} roadway upstream from i^{th} segment.

Eq. (4-13)

Incident Delay and Congestion Estimation Using Bivariate Analysis

The bivariate analyses were performed at roadway segments +/- 5 miles from where the incident occurred and 3 hours before and after the incident. Figure 4-13 illustrates the 3-D bivariate analysis results of an incident occurred on 11/16/2012 in I-94 Eastbound. Traffic was slower than normal condition for about 2 hours and 4 miles upstream from where the incident occurred (segment 9) according to the computed k values displayed in Figure 4-13. Figure 4-14 shows the hourly speed distribution of each 1-mile roadway segment near where the incident occurred. Hour "0" in Figure 13 & 14 means the hour (17:00) when the incident occurred. K-value exceeding 9.21 threshold represents the traffic condition is abnormal using 99% confidence interval. Figure 4-15 and 4-16 display the corresponding bivariate analysis results and speed distribution of another incident which occurred at the same segment on I-94 on 8/2/2013. Hour "0" in Figure 15 & 16 means the hour (17:00) when the incident occurred. K-value exceeding 9.21 threshold represents the traffic condition is abnormal using 99% confidence interval.

Estimated incident delay duration and congestion length for a few incidents on I-94, I-35W, I-494 and I-694 in the Twin Cities metro area are listed in Table 4-2. The results from the bivariate analysis in incident #9 (Friday 1/26/2013 on I-94 WB, in Maple Grove) did not indicate any delay or congestion around 6PM toward the end of rush hours. I-94 WB in Maple Grove toward Roger has recurring congestion in Friday afternoon. The bivariate analysis found no significant delay but 3 miles of congestion for incident #16 at I-696 EB before I-35E.

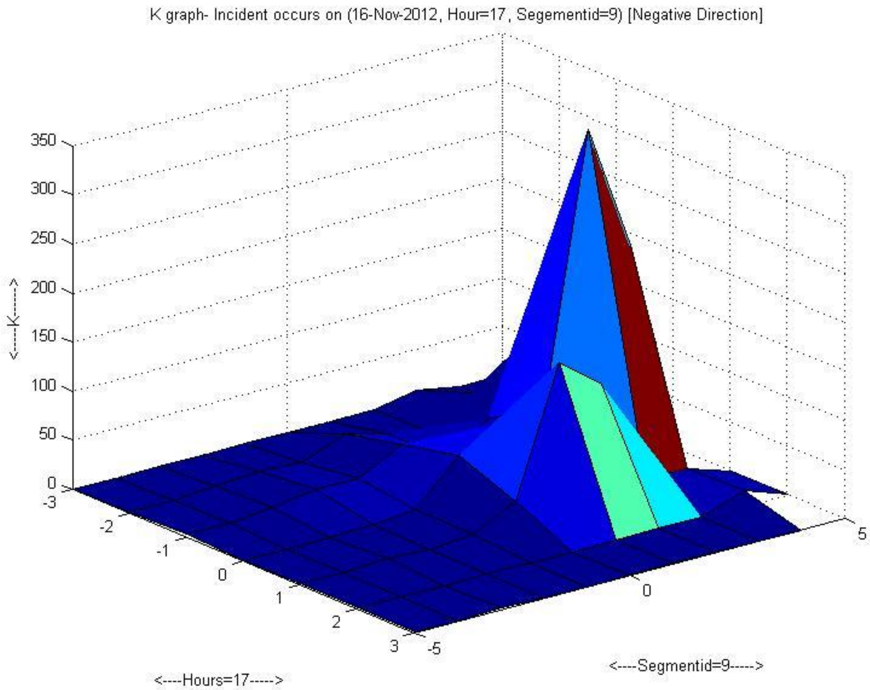


Figure 4-13 Bivariate K-plot of an Incident Occurred on 11/16/2012 at 5PM on I-94 EB Segment 9 in Woodbury

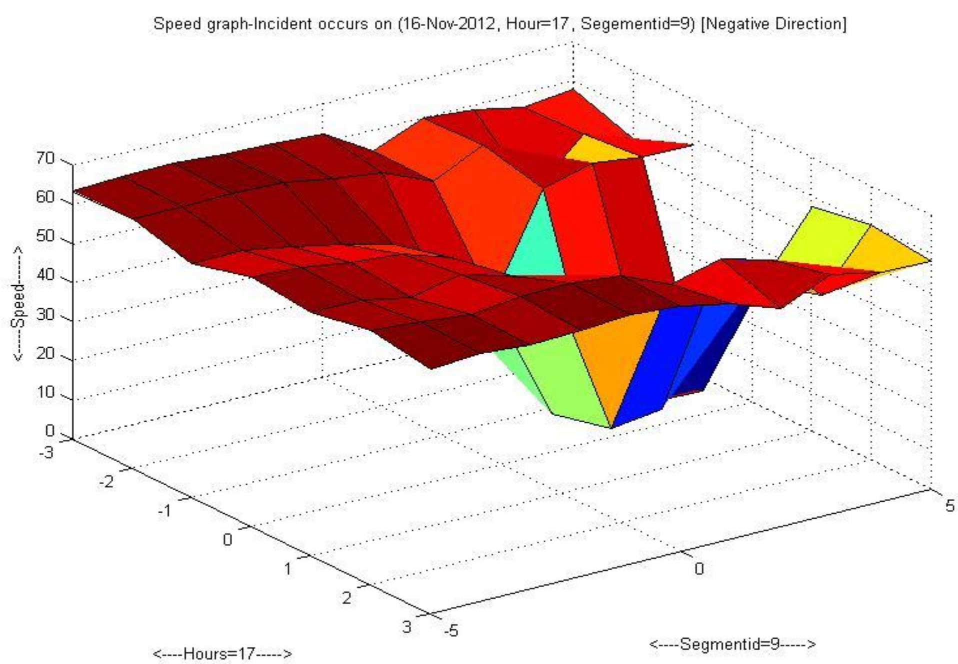


Figure 4-14 Average Speed of an Incident Occurred on 11/16/2012 at 5PM on I-94 EB Segment 9 in Woodbury

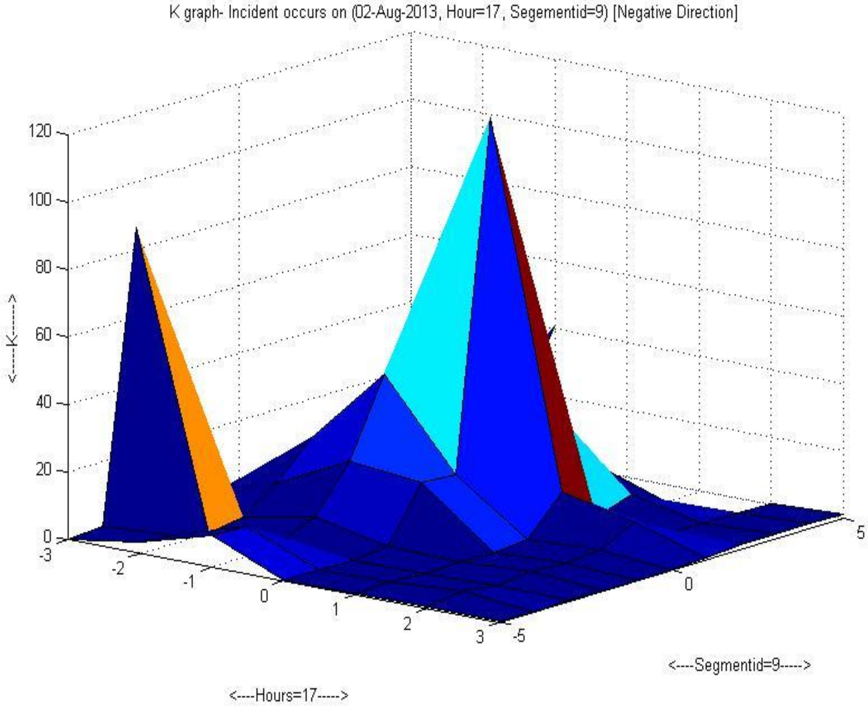


Figure 4-15 Bivariate K-plot of an Incident Occurred on 8/2/2013 at 5PM on I-94 EB Segment 9 in Woodbury

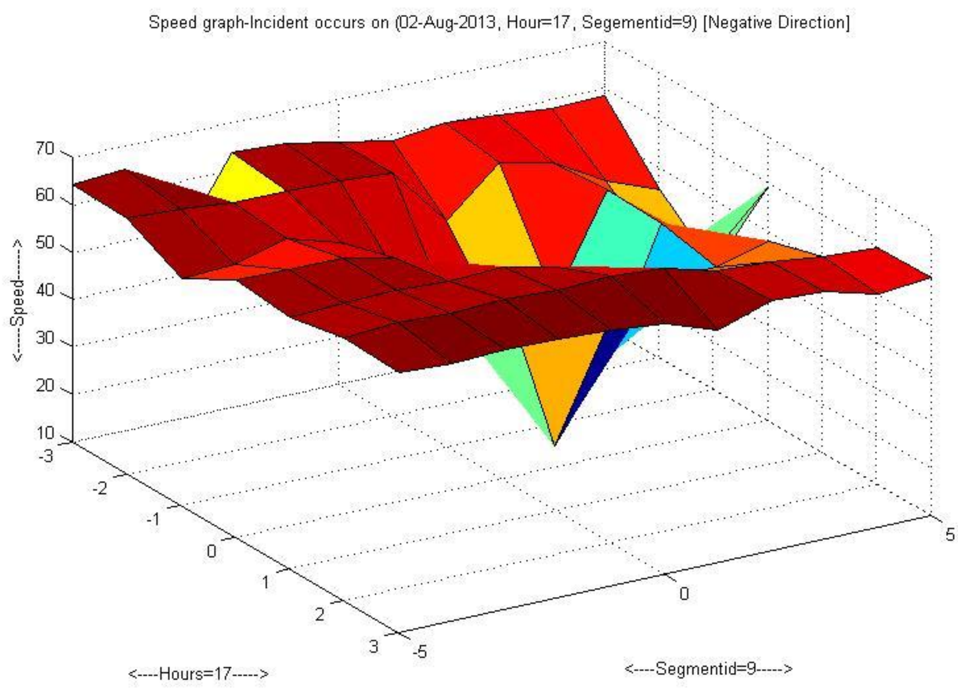


Figure 4-16 Average Speed of an Incident Occurred on 8/2/2013 at 5PM on I-94 EB Segment 9 in Woodbury

Table 4-2 Incident Delay and Congestion Estimation

Incident #	Route	Incident Date	Hour	Segment ID	Direction	Incident Duration MN511 (min)	Est. of Delay (min)	Queue Est. (mile)
1	I-94	11/16/2012	17	9	Negative (EB)	63	180	4
2	I-94	8/2/2013	17	9	Negative (EB)	57	120	3.5
3	I-94	7/19/2013	16	9	Negative (EB)	23	90	5
4	I-94	5/20/2013	16	12	Negative (EB)	19	180	5
5	I-94	6/25/2013	17	21	Negative (EB)	19	60	7.5
6	I-94	12/5/2012	10	25	Negative (EB)	5	120	1
7	I-94	7/16/2013	15	35	Positive (WB)	438	180	5
8	I-94	6/27/2013	9	35	Positive (WB)	225	240	5
9	I-94	1/25/2013	18	46	Positive (WB)	179	0	0
10	I-35W	12/2/2012	17	42	Positive (NB)	274	60	2
11	I-35W	12/1/2012	9	47	Positive (NB)	47	0	0
12	I-35W	2/10/2013	13	39	Positive (NB)	83	150	7
13	I-35W	1/23/2013	7	36	Positive (NB)	77	120	4
14	I-494	7/1/2013	20	6	Positive (WB)	111	120	3
15	I-494	8/27/2013	7	6	Positive (WB)	18	120	4
16	I-694	2/9/2013	9	10	Positive (EB)	11	0	3
17	I-694	2/14/2013	8	10	Positive (EB)	15	180	4

Model Features

Models are trained using a subset of the following features: incident duration computed from state patrol data (CAD) or MN511 traveler information system, traffic direction, HCADT traffic volume, interstate, number of lanes, month, number of vehicles involved in an incident, temperature, time of day, type of incident, visibility distance, average wind speed. During preprocessing phase, we converted the numerical features into appropriate features observing its distribution. Missing values are replaced or excluded based on the classification algorithm.

Preprocessing

Range Values of Features are as follows:

- CAD duration: 0-2670 minutes.
- HCADT 2012 traffic volume: 85-15100.
- Temperature in degree Celsius: -239-3276.7.
- Visual Distance in meters: -132767.
- Wind speed average: 0-255 km/hr.

Features are converted into categories in the following way:

- CAD duration or MN511 duration (min): 13 groups of size interval 10. Excluding samples with values greater than 2000.
- Direction: 2 types, positive and negative.
- HCADT 2012: 8 groups of size 2000.
- Interstate: 2 types.
- Number of lanes: 5 unique values.
- Month: 12 unique months.
- Number of vehicles involved: 11 unique values.
- Precipitation type: 9 types.
- Temperature: 24 groups of size interval 10.
- Time of day: 3 groups: 5-9AM, 3-6PM, and the other time of day.
- Type of incidents: 7 types.
- Type property damage: 13 types.
- Visual distance: 6 groups of size interval of 5000 value.
- Wind Speed Average: 2 groups, less or greater than 100 values.

Classification

Based on the data, it is preferable to convert the regression to a classification problem as more efficient algorithms are available in this domain. Since our training output consisted of ordinary values, it is suitable to perform classification. To remove irrelevant features during classification we adopted a process called the feature selection process. In this process, a subset of features are considered for creating the model and we add or remove the next feature based on the evaluation criteria (in our case training error) using this approach. Thus, the best model will contain some subset of features mentioned above.

For classification, we first considered the linear discriminant analysis also known as the Fisher Discriminant. In this, each class is described as a linear model. i.e. $y_c(x) = w^T x$, where w is the model coefficient and x is the input feature vector. The idea proposed by Fisher is to maximize a function, or find a hyper-plane in D dimension, that will give a large separation between the projected class means on this hyper-plane while also giving a small variance within each class, thereby minimizing the class overlap. Linear discriminant yields the following result: Training error= 66%, Testing error=66%. Testing error is calculated using 10-fold validation process in which 10 percent of data is separated out for testing. For quadratic discriminant, we also included the interaction term in our analysis. Results are as follows: Training error= 63%, Testing Error=63%.

After discriminant analysis, we considered multi nominal logistic regression. The posterior probabilities of classes are given by a softmax transformation of linear functions of the feature variables, so that $p(C_k|x) = y_k(x) = \exp(a_k) / \mathrm{Sum}[\exp(a_k)]$, where the `activations' a_k is given by

$a_k = w^T_k x$. Parameters w of model are obtained after training the data and results are as follows: training error= 59% and testing error= 60%.

We moved to more sophisticated classification algorithms such as using ensemble of classifiers. AdaBoost (adaptive boosting) is an ensemble learning algorithm that can be used for classification or regression. Although AdaBoost is more resistant to over fitting than many other machine learning algorithms, it is often sensitive to noisy data and outliers. AdaBoost is called adaptive because it uses multiple iterations to generate a single composite strong learner. AdaBoost creates the strong learner (a classifier that is well-correlated to the true classifier) by iteratively adding weak learners (a classifier that is only slightly correlated to the true classifier). During each round of training, a new weak learner is added to the ensemble and a weighting vector is adjusted to focus on examples that were misclassified in previous rounds. The result is a classifier that has a higher accuracy than the weak learners' classifiers. Using this algorithm and feature selection process, the best results were as follows: Training Error=37% and Testing Error=59%. All except air pressure, average wind speed, visual distance features are used in this process.

Lastly, we considered the most widely used Random Forest algorithm for real data model prediction. Random Forest is a trademark term for an ensemble of decision trees. Unlike single decision trees which are likely to suffer from high variance or high bias (depending on how they are tuned) Random Forest algorithms use averaging to find a natural balance between the two extremes. Since they have very few parameters to tune and can be used quite efficiently with default parameter settings (i.e. they are effectively non-parametric), Random Forests are good to use when the underlying model is unknown. Random Forest randomly selects subset of samples as well as features following the bootstrap aggregation method and trains the weak classifier (binary decision tree) using the above data. This process is repeated to create a forest of trees which is our prediction model. In the end, we take the majority vote to label the input.

In our case, we trained the random forest model with 100 binary decision trees along with input features from the feature selection process. This yielded the following results: training error=10% and Generalization error= 50%. All features expect air pressure were considered for the training of the above model.

Table 4-3 lists the results of various classification algorithms we tested using incident records from both CAD and MN511 data. The results indicated the Random Forest algorithm has the least amount of training and testing errors as compared to the other algorithms. However, the Random Forest algorithm is not effective in predicting the incident delay with additional information regarding the incident.

Table 4-3 List of Classification Algorithms Considered

Algorithms	**Training Error**		**Testing Error**	
	CAD Data	MN511 Data	CAD Data	MN511 Data
Linear Discriminant Classification	66%	68%	66%	68%
Quadratic Discriminant Classification	63%	66%	63%	66%
Multi-Nominal Logistic Regression	59%	63%	60%	63%
Boosting Algorithms: Adaboost	37%	43%	59%	63%
Ensemble Algorithms: Random Forest	10%	19%	50%	59%

5. Pilot Implementation

The objective of the pilot implementation is to demonstrate a concept of operations for the Incident Decision Support System (IDSS) to provide incident information to truck operators. The pilot implementation uses available archived data for demonstration purposes. A real-time data feed from ATRI can be incorporated into the developed IDSS later when real-time data becomes available.

5.1 Incident Detection System

The incident detection routines continuously processes and analyzes the incoming truck GPS data to detect abnormal traffic patterns based on the roadway segment, direction, and time of day.

5.1.1 System Architecture for pilot implementation

Figure 5-1 illustrates the system architecture of the IDSS system. The UMN IDSS geo-spatial database takes as input the truck GPS data, weather information, CAD and MN511 traffic data for performing the data processing analysis and incident detection. A web service, acting as a middleware with additional security protection, handles requests from mobile or in-vehicle devices to provide incident related information to the end users.

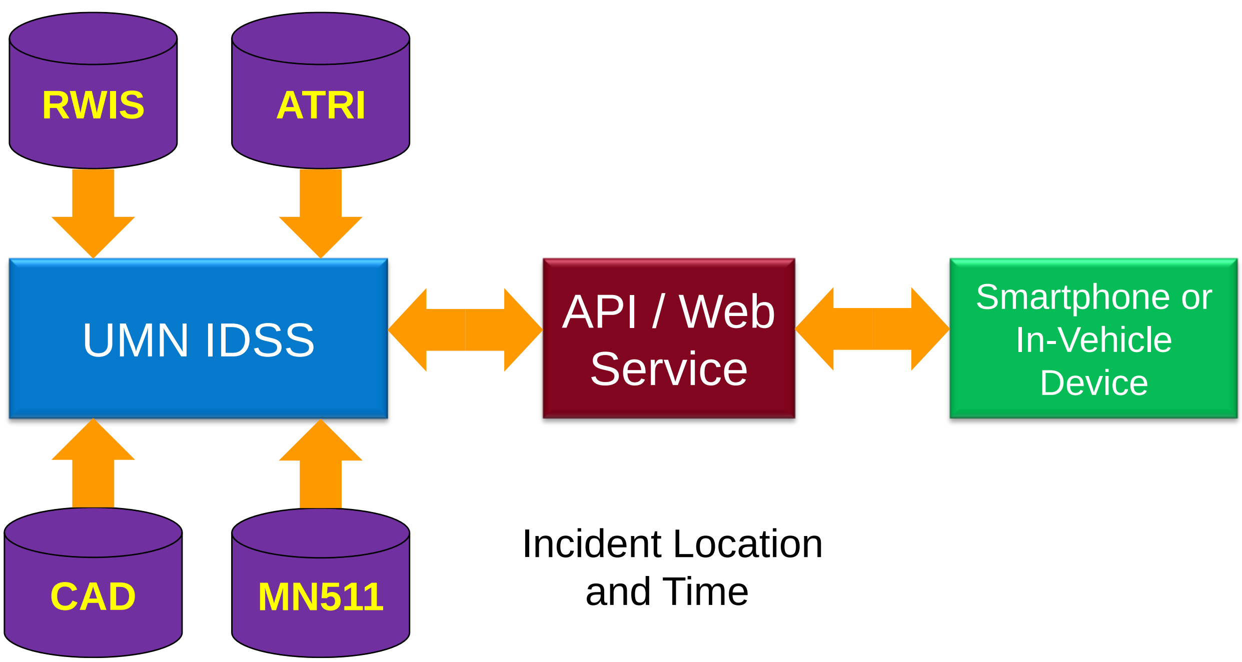


Figure 5-1 Overall IDSS System Architecture

5.1.2 Incident detection examples

<u>Example 1 #36, MSPCAD-P130556879</u>

Figure 5-2 illustrates an incident that occurred on Wed 10/30/2013 at 6:47 AM on US Highway 52 NB at 70th Street. This incident involved 6 vehicles and the duration between police arrival and departure was 33 minutes according to the archived CAD data. The top 2 graphs are the analysis results from loop detector data on the previous day (10/29/13) for both lanes. The blue lines represent the measured speed from loop detectors and the green lines are the incident indices (k-values) from the bivariate analyses determined by our algorithm. The top 2 figures show no incident.

The bottom 2 graphs display the loop detector speed (blue), loop detector bivariate index (green), the GPS speed (magenta), and bivariate index (k-value) from GPS speed (red) at the roadway segment where an incident occurred on 10/30/13 for both lanes. As indicated in both graphs in the bottom, the speeds as measured by loops and by the probe data dropped and the k-values spiked before 7AM when the incident occurred. A high k-value indicates an incident. A drop of k-value to a low value indicates that the incident has cleared.

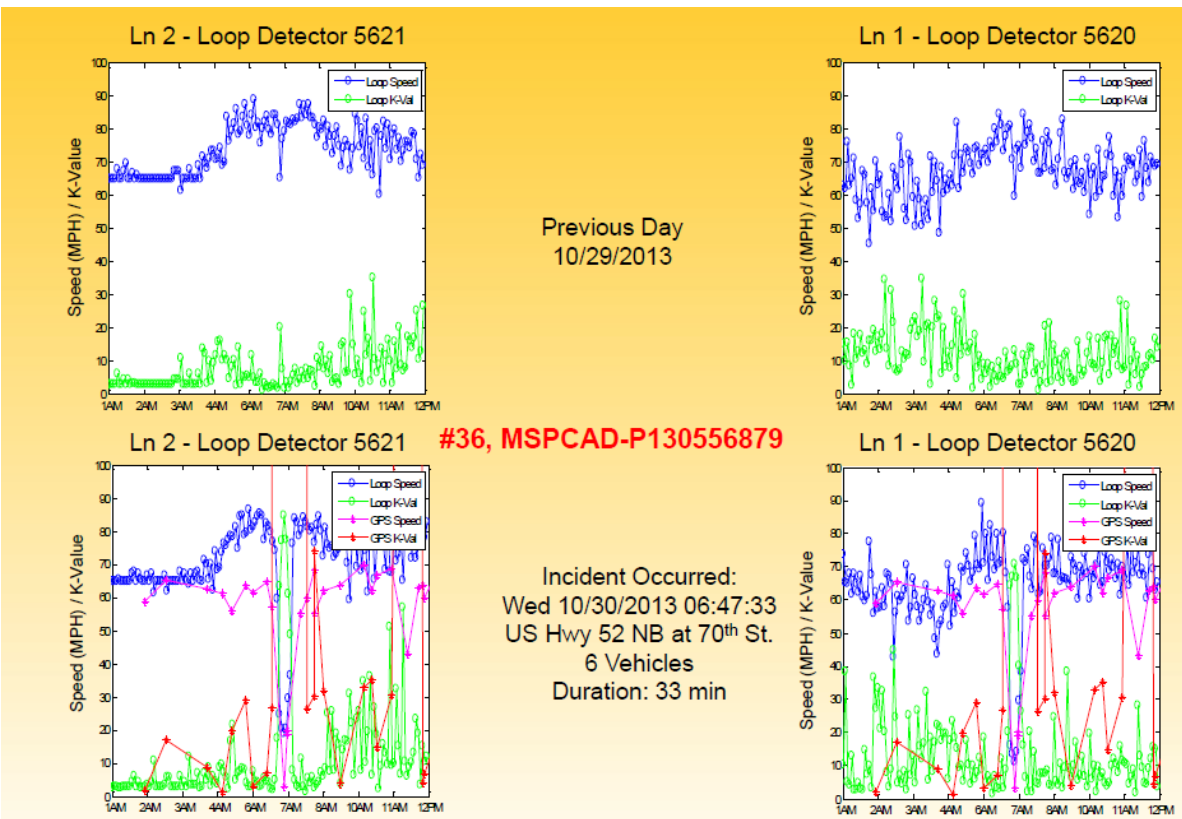


Figure 5-2 Incident Detection Analysis #36, MSPCAD-P130556879: 2 adjacent loops (different lanes) on US Hwy. 52 NB at 70th St.

Example 2 Incident #60, MSPCAD-P130556946

Figure 5-3 illustrated an incident that occurred on Wed 10/30/2013 at 7:34 AM on I-94 EB at Chicago Ave. This incident involved 2 vehicles and the duration between police arrival and departure was 36 minutes according to the archived CAD data. The top 2 graphs are the analysis results from the loop detector data on previous day (10/29/13) for both lanes. The blue lines represent the measured speed from adjacent loop detectors and the green lines are the indices (k-values) from the bivariate analyses.

The bottom 2 graphs display the loop detector speed (blue), loop detector bivariate index (green), the GPS speed (magenta), and the bivariate index (k-value) from GPS speed (red) at the roadway segment where an incident occurred on 10/30/13 for both lanes. As indicated in both graphs in the bottom, the speeds dropped and the k-values spiked around 7AM when the incident occurred. Again, a high k-value indicates an incident occurred. When the k-value drops, the incident has cleared.

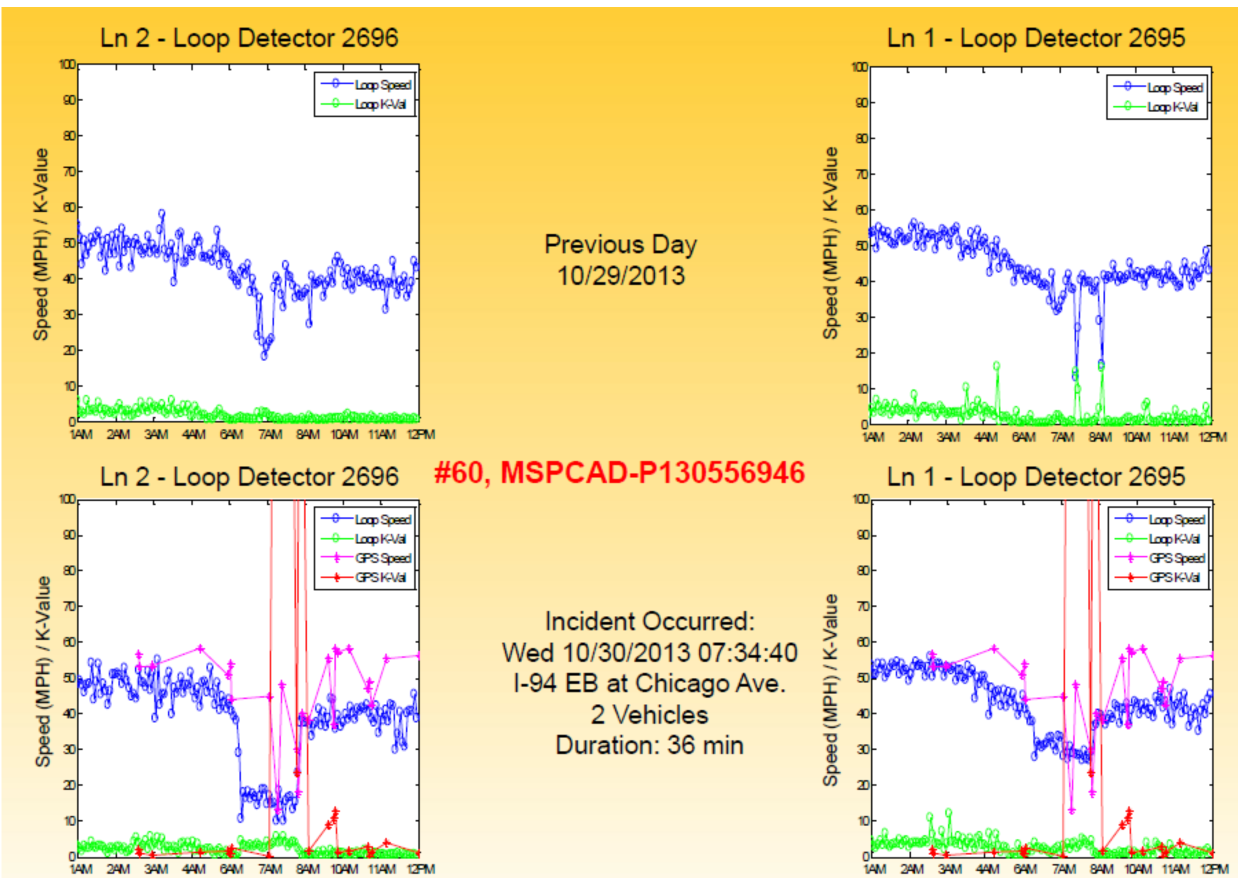


Figure 5-3 Incident Detection Analysis #60, MSPCAD-P130556946: 2 adjacent loops (different lanes) on I-94 EB at Chicago Ave.

Note that the 36 minutes associated with the arrival and departure of the state patrol represents a portion of the total incident duration. The loop detector data is the best measure of the incident duration and is used here to validate our incident detection algorithm. The algorithm would be used on rural roads when no loop detector is available.

5.2 Pilot Implementation System Design

Figure 5-4 illustrates the functionalities and data flow for the proposed pilot implementation. Block A represents a user's device (a smartphone or some other in-vehicle device) that has the capability to automatically determine the user's location and heading using GPS. Block B is a database and web server that hosts the incident information (timestamp, location, and heading) resulting from the truck GPS data processing and analysis as displayed in Block C. Block C will continuously process incoming GPS data for incident detection. If an incident is detected, the results will be sent to the incident database residing in Block B. The purpose of the middle tier web service design in Block B is to protect the truck GPS data in block C and provide web services for different applications such as the smartphone, web XML or other 3rd party applications. Data flow (1 → 2 → 3 → 4) represents a configurable cycle of information update as a truck operator travels down the roadway.

For an initial pilot implementation, block B and C will reside on the same PC at UMN. For a real-time application, the algorithm running in Block C will reside on a server receiving a real-time GPS raw data feed at ATRI (over 1 million data points per day for trucks travelling on all roads in MN) The system design architecture for the pilot implementation using archived data would work similarly for an application using a real-time data feed when it's ready.

This pilot implementation focuses on the mechanism for providing the information to the user. A simple text display and text-to-speech output was announced to the user in the initial phase. The research team is aware of the potential for distraction by providing travel information through the smartphone. The user interface and human factors issues will later need further investigation in order to determine the best way to ultimately present the incident information to the driver.

5.2.1 Incident Information Flow

1. The smartphone (equipped with GPS) determines its current latitude-longitude and heading information. The information together with the current timestamp is submitted to a web interface to determine the roadway and segment where the user is currently located.
2. Roadway ID, segment ID and direction information is used to query incident formation from the database (described in Block B).
3. The IDSS then checks the downstream traffic condition (e.g., 50 miles ahead) if there is any incident along the roadway on which the operator is traveling. (This information can

be integrated with a navigation application when a route is specified; the IDSS would check the incident information downstream along the route.)

4. A simple text message such as “Incident xxx, yyy miles ahead on zzz road” would be displayed on the smartphone or announced to the users.

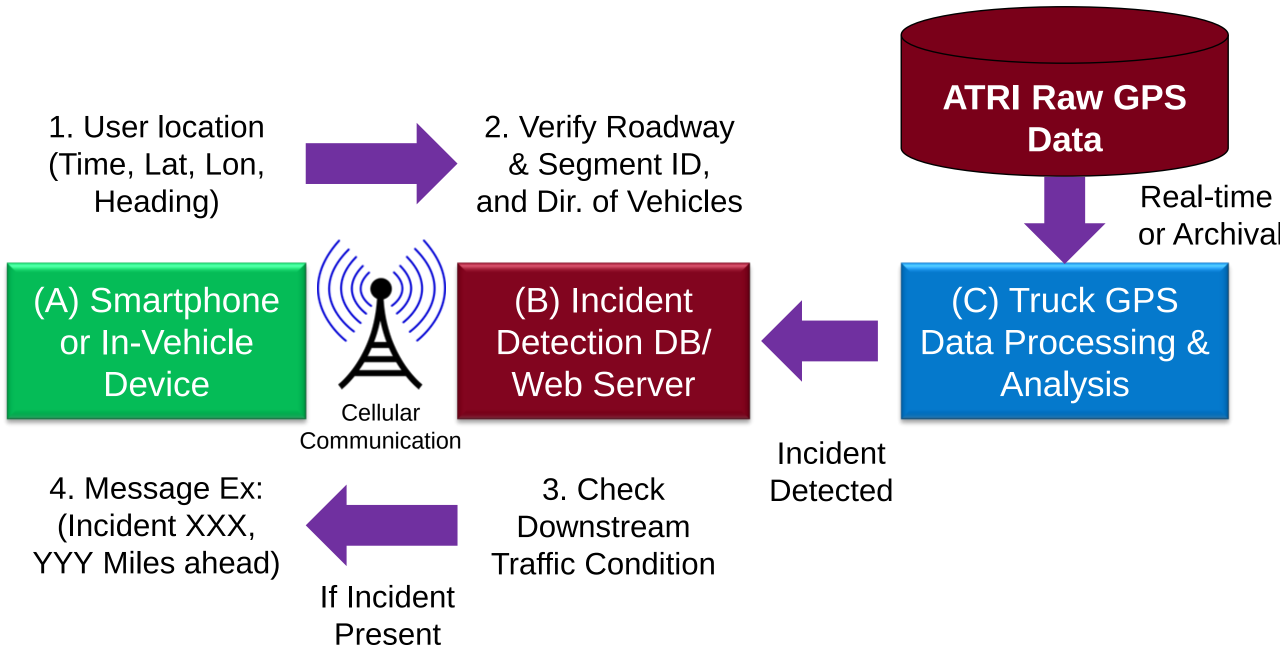


Figure 5-4 Functionality Block Diagram of Proposed Pilot Demonstration

5.2.2 Smartphone App

An Android application was developed to implement the Incident decision support system. The Android app shows the current location on a Google map (Google Maps Android API v2) and updates it positions as the current location of the user changes. As the location is updated, the application sends its location and date-time data to the server to get information about any incidents that might have taken place on the route that the user is travelling. The application provides settings to adjust date, time, look-ahead distance, test-to-speech settings etc.

Figure 5.5 illustrates the flowchart of the Android application. All the default values for the application settings have been considered. Assuming that the “enable voice” option is checked, the application speaks to the driver if there is any valid response from the incident detection web server. Detailed information about the smartphone app is discussed in Appendix C.

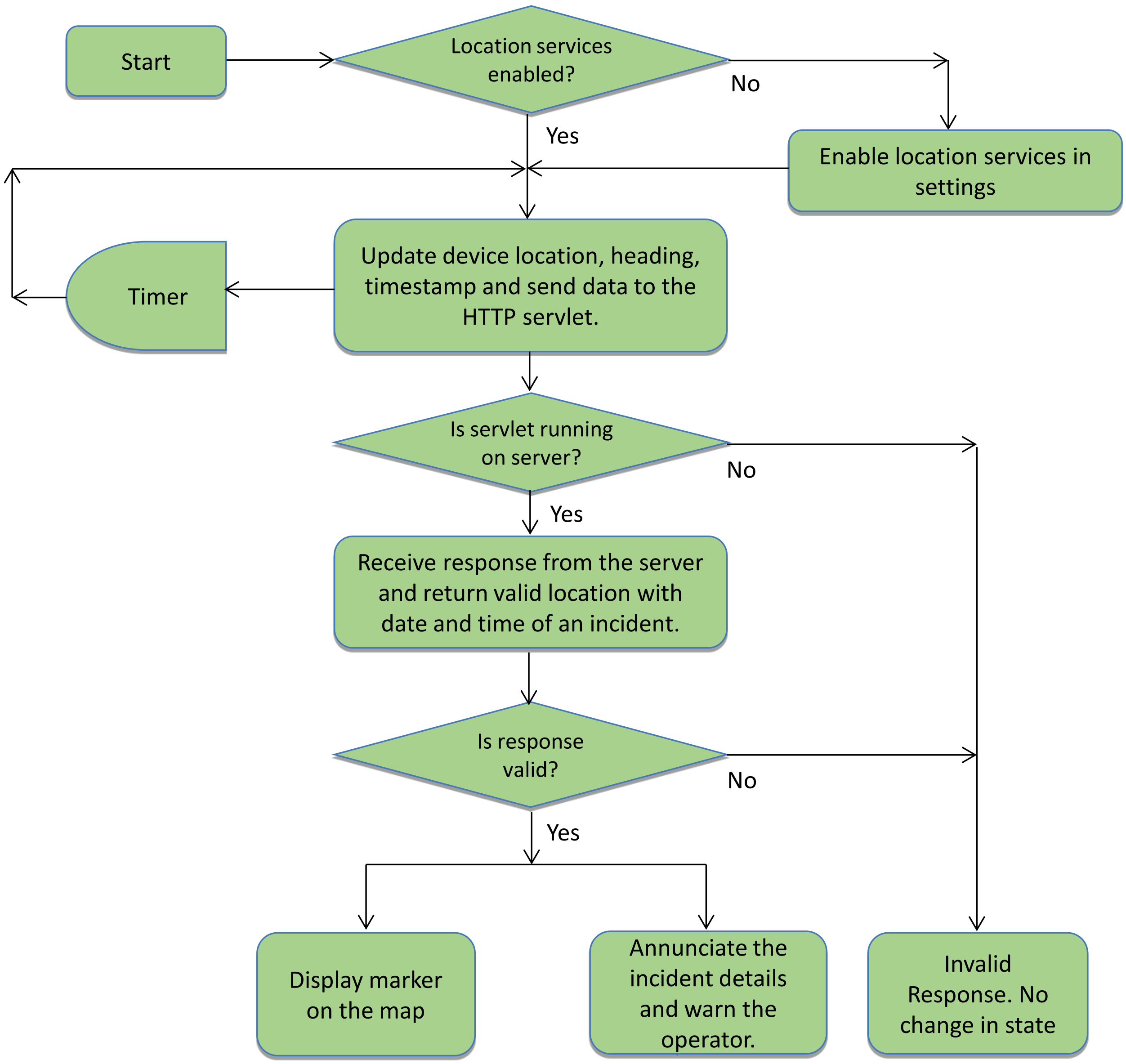


Figure 5-5 Design Flowchart of Smartphone App for IDSS

5.2.3 Web Server

A HTTP Servlet was developed which acts as a middleware between the Android application and the central database server. This HTTP Servlet reads the request from the Android application and then establishes a communication connection with the incident detection database server. The Java servlet connector object takes all the parameters sent by the Android application and then creates the query string in the required format. The query string is sent to the server and then the object waits for a response from the database server. As soon as the response arrives, it checks for the validity for the response. If the number of responses matches the expected number, it then passes the result to the smartphone app. The flowchart for the servlet is illustrated in Figure 5-6. Detailed information about the servlet development is discussed in Appendix D.

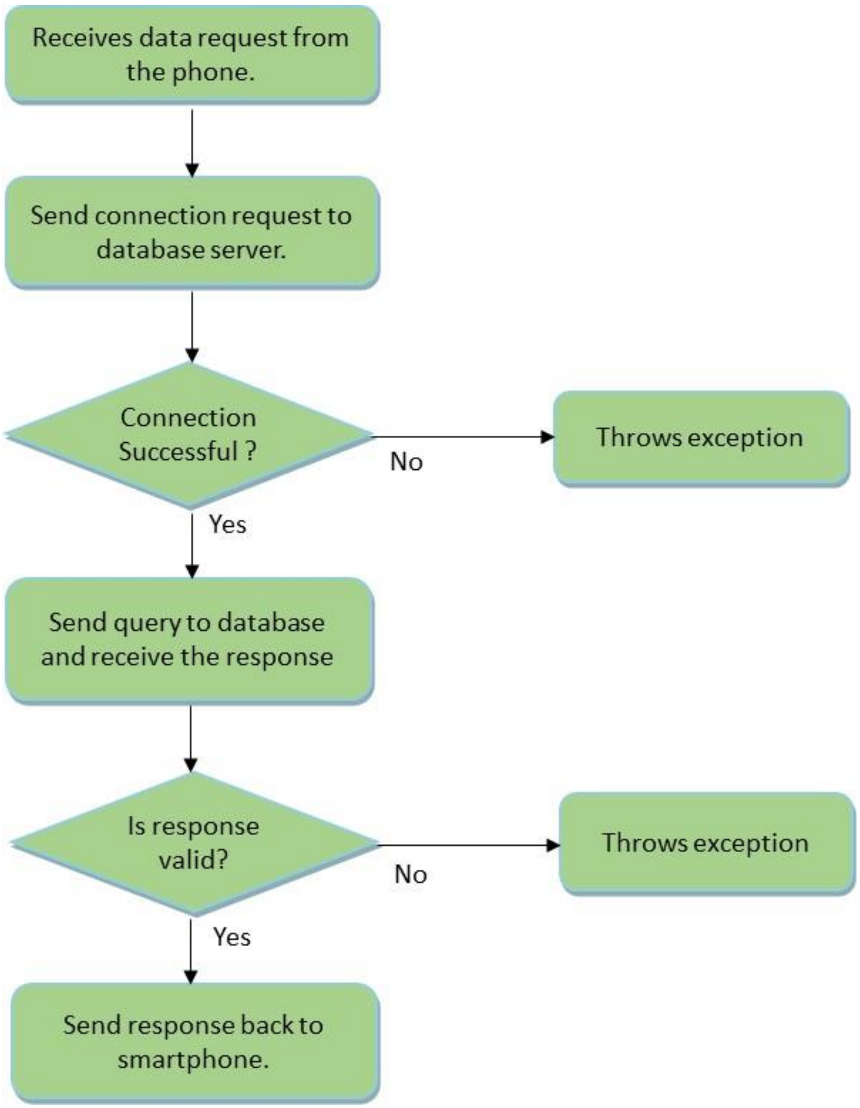


Figure 5-6 Design Flowchart of the Servlet

5.3 System Testing and Simulation

The smartphone app was tested on a passenger vehicle driving on I-94 WB near the University campus in Minneapolis where an incident occurred according to the historical data. A few historical incidents (Table 5-1) were identified for testing and validation purposes. We set the time on the smartphone app to a timestamp based on the archived data when an incident occurred. As the test vehicle drove toward a location where the incident occurred, the app successfully detected the available incident information through the web server and announced the corresponding information (where and when) to the driver during the experiments.

For example, Figure 5-7 illustrates that a test vehicle is driving toward an incident (ID 2 in table 5-1) which occurred about 5 miles ahead on interstate I-94 on 1/10/2013 around 6PM. When an incident is detected, the app will announce "*incident occurred 5 miles ahead 19 minutes ago on Interstate 94*".

Table 5-1 List of Sample Incidents

ID	Incident Latitude	Incident Longitude	Incident Date/Time	Route Name
1	44.953472	-93.089083	2/1/2013 19:31	Interstate 94
2	44.967203	-93.222714	1/10/2013 17:41	Interstate 94
3	44.960120	-93.206196	3/6/2013 16:59	Interstate 94
4	44.964364	-93.242091	7/15/2013 16:45	Interstate 94

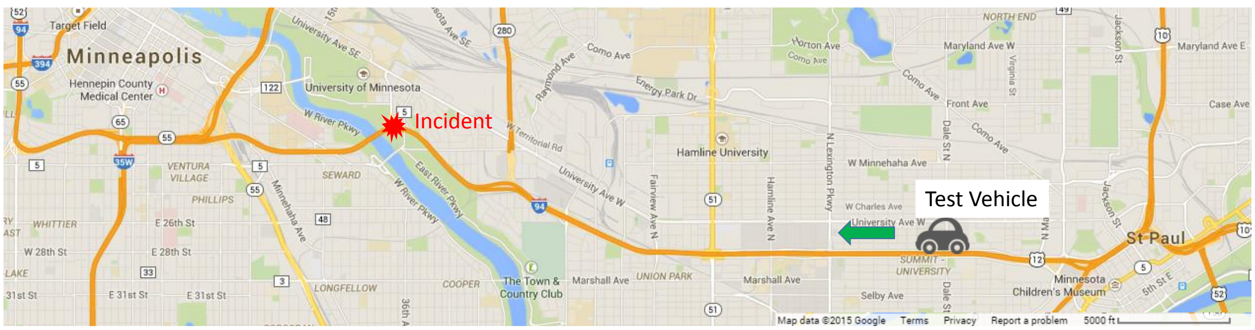


Figure 5-7 Illustration of a Test Vehicle Approaching an Incident
(Background Image from Map Data © Google 2015)

6. Summary and Conclusion

Building upon our previous effort, we developed an Incident Decision Support System (IDSS) that uses GPS-equipped commercial trucks as probe vehicles on key freight corridors on inter-regional corridors (IRC) and supplemental freight routes in Minnesota. An automatic incident detection algorithm was developed to support incident decision-making for state DOTs for traveler information purposes and for truck drivers and dispatchers who need to make necessary routing and operations decisions when incidents occur.

The objectives of this project are to (1) develop an Incident Decision Support System (IDSS) that uses GPS-equipped commercial trucks as probe vehicles on key freight corridors, and (2) provide incident information to truck operators.

Archived truck GPS data, MN511 traffic data, Computer Aided Dispatch (CAD) data from the Minnesota State Patrol, and weather data in the Twin Cities Metro Area (TCMA) were obtained for this study. The MN511 data provides basic incident information such as time, location, crash, incident description, stalled vehicle, presence of debris, etc. Currently, the incident descriptions are not standardized and not queryable. The Computer Aided Dispatch (CAD) data from the Minnesota State Patrol and incident management team contains more information than the MN511 data such as incident type, patrol arrival/leaving time, vehicle class, number of vehicles involved, fatality, etc.

A bivariate incident detection methodology was developed by including the travel time computed from raw GPS truck data and the travel time difference between roadway segments and time intervals as key parameters. The incident detection methodology was implemented through a Java program which was developed to process and analyze probe vehicle data for incident detection. The incident detection algorithm worked very well but the incident duration prediction estimator did not.

The individual incident duration was calculated using the time difference between the update time of the "last" message and the start time of the "first" message concerning the event. The research team manually computed the incident delay using a small sample dataset using the bivariate model to validate the results against the incident duration from MN511 data and CAD data. An automatic delay estimation process was developed and tested. In an attempt to develop an incident delay prediction model, the research team tried several algorithms (such as, linear discriminant classification, quadratic discriminant classification, multi-nominal logistic regression, boosting algorithms: Adaboost, and ensemble algorithms: random forest) with limited success. The delay estimation model using the Random Forest algorithm with 100 binary decision trees along with input features from the feature selection process yielded the best results (10% training error and 50% of generalization error) which was inadequate.

Validation on the incident detection algorithm was successfully performed by comparing the results to loop detector data obtained from MnDOT on freight corridors with sufficient GPS samples in the Twin Cities metropolitan area. Our incident detection algorithm was designed to focus on rural roadways where loop detector data is not available.

A pilot implementation was developed to demonstrate a concept of operations for the Incident Decision Support System (IDSS) and provide incident information to truck operators. The pilot implementation used available archived data for demonstration purposes. A real-time data feed from ATRI can be incorporated into a future version of IDSS later when real-time data becomes available.

To provide incident information to truck operators, an in-vehicle display through a smartphone device was developed for pilot implementation to provide auditory advanced alerts. The research team is aware of the potential for distraction by providing travel information through the smartphone. The user interface and human factors issues will later need further investigation in order to determine the best way to ultimately present the incident information to the driver.

When an incident is detected, the app displays incident information to travelers through an auditory text-to-speech (TTS) interface as they approach the incident. For example, if the distance of an incident is xx miles away, the app will announce "incident occurred xx miles ahead 'yy' minutes ago. The TTS speech engine also mentions the route name of the incident assuming that the driver is aware of the current route so that he or she can determine whether the incident is on the same route or nearby route.

The app was tested on a passenger vehicle driving on I-94 WB near the University campus in Minneapolis where an incident occurred according to historical data. For testing and validation purposes, we set the time on the smartphone app to a timestamp based on the archived data when an incident occurred. As the test vehicle drove toward a location where the incident occurred, the app successfully detected the incident information through the web server and announced the corresponding information (where and when) to the driver during the experiments.

References

Antoniades, C.N. and Stephanedes, Y. J., (1996). Single-Station Incident Detection Algorithm (SSID) for Sparsely Instrumented Freeway Sites. *Transportation Engineering*,

ATRI (2011). FPM Congestion Monitoring at 250 Freight Significant Highway Locations. http://atri-online.org/2011/10/01/fpm-congestion-monitoring-at-250-freight-significant-highway-locations/, accessed March 2012.

Boyles, S., S. T. Waller. (2007) A Stochastic Delay Prediction Model for Real-Time Incident Management, *ITE Journal* 77, 18-24.

Fenno D., and M. Ogden, (1998). Freeway Service Patrols: A State of the Practice, TRB 77th Annual Meeting, Compendium of Papers, Washington, DC.

FHWA, (2002). *Commercial Vehicle Travel Time and Delay at U.S. Border Crossings*, http://ops.fhwa.dot.gov/freight/freight_news/travel_time/travel_time_delay.htm, accessed November 2012.

FHWA (2008). Estimated Cost of Freight Involved in Highway Bottlenecks. http://www.fhwa.dot.gov/policy/otps/freight.pdf, accessed March 2012

FHWA, (2012). Evaluation of Utilizing Probe-based Freight Data to Support Congestion Monitoring on the Highway System, Results and Initial Conclusions Summary, Science Applications International Corporation (SAIC), Delcan Corporation and University of Virginia.

Fu, L., (2004). A Fuzzy Queuing Model for Real-Time, Adaptive Prediction of Incident Delay for ATMS/ATIS, *Transportation Planning and Technology*, Vol. 27, No. 1, pp.1-23

Giuliano, G. (1989) Incident characteristics, frequency, and duration on a high volume urban freeway. Transportation Research A 23, 387-396.

Golob, T., W. Recker, and J. Leonard. (1987) An analysis of the severity and incident duration of truck-involved freeway accidents. Accident Analysis and Prevention 19, 375-395.

Guo, J., Smith, B.L., and Zhang, G., (2008). Information Technology Used to Evaluate Nontraditional Traffic Monitoring Systems, *Journal of the Transportation Research Board*, No. 2024, pp. 1-7.

Hampel, F. R. (1971). A general qualitative definition of robustness. *Annals of Mathematics Statistics*, 42, 1887–1896.

Ji, Y., Zhang, X., and Sun, L., (2011). Estimation of Traffic Incident Delay and its Impact Analysis Based on cell transmission model, IEEE Intelligent Vehicles Symposium (IV), p. 54-59, Baden-Baden, Germany

Jones, B., Janssen, L., and Mannering, F.L., (1991). Analysis of the Frequency and Duration of Freeway Accidents in Seattle. *Accident Analysis and Prevention*, Vol. 23, pp. 239–255.

Jones, C., Murray, D. & Short, J., (2005). Methods of Travel Time Measurement in Freight-Significant Corridors. TRB 84th annual meeting, Compendium of Papers, Washington, D.C.

Khattak, A., J. Schofer, and M. Wang. (1995) A simple time-sequential procedure for predicting freeway incident duration. IVHS Journal 2, 113-138.

Krizek, K., Iacono, M., El-Geneidy, A.M., Liao, C.F., and Johns, R., (2009) Access to Destinations: Application of Accessibility Measures for Non-Auto Travel Modes, Final Report, Minnesota Department of Transportation, Report No. MnDOT 2009-24, St. Paul, MN.

Kuzmyak, J. (2008). Forecasting Metropolitan Commercial and Freight Travel, National Cooperative Highway Research Program (NCHRP), Synthesis 384. Washington, D.C.: Transportation Research Board.

Li, J., Lan, C., and Gu, X., (2006). Estimation of Incident Delay and Its Uncertainty on Freeway Networks, *Journal of the Transportation Research Board*, No.1956, p.37-45, Washington, DC.

Li, Y. and McDonald, M., (2005). Motorway incident detection using probe vehicles, *proceedings of the Institute of Civil Engineers*, Transport 158, 11-15.

Liao, C.F., (2008). Freight Performance Measure Systems (FPMS) System Evaluation and Data Analysis. Final Report, Center for Transportation Studies, Report No. CTS 08-01, University of Minnesota (UMN), Minneapolis, MN.

Liao, C.F., (2009). Using Archived Truck GPS Data for Freight Performance Analysis on I-94/I-90 from the Twin Cities to Chicago. Final Report, Minnesota Department of Transportation, Report No. CTS 09-27, St. Paul, MN.

Liao, C.F. and Liu, H.X. (2010). Development of a Data Processing Framework for Transit Performance Analysis, Transportation Research Record, *Journal of the Transportation Research Board*, No.2143, pp. 34-43.

Liao, C.F., (2011). Computing Auto Accessibility to Other Destinations, Final Report, Intelligent Transportation Systems Institute, Center for Transportation Studies (CTS 2011), University of Minnesota, Minneapolis, MN.

Lomax, T., Schrank, D., Turner, S., and Margiotta, R., (2003). *Selecting Travel Time Reliability Measures*. Final Report, Texas Transportation Institute, College Station, TX. http://tti.tamu.edu/documents/TTI-2003-3.pdf, accessed November 2012.

Lund, A.S., and Pack, M.L., (2010). Dynamic Wide-Area Congestion and Incident Monitoring Using Probe Data, Transportation Research Record, *Journal of the Transportation Research Board*, No.2174, pp. 1-9.

Martin, P.T., Perrin, J., and Hansen, B., (2001). Incident Detection Algorithm Evaluation, final report, Utah Department of Transportation, Salt Lake, UT

Maze, T.H., Albrecht, C.P., and Smadi, O., (2010). Mn/DOT Truck Parking Study: Phase 2, Final Report, Minnesota Department of Transportation, MN/RC 2010-34, St. Paul, MN. http://www.dot.state.mn.us/ofrw/PDF/truckparkingstudyphase2.pdf, accessed November 2012.

McCormack, E. and Hallenbeck, M.E., (2006). ITS Devices used to Collect Truck Data for Performance Benchmarks, *Journal of the Transportation Research Board*, No. 1957, pp. 43-50.

Minnesota Department of Transportation (2005), Minnesota Statewide Freight Plan, St. Paul, MN.

Minnesota Department of Transportation (1994). Highway Helper Summary Report—Twin Cities Metro Area, Minneapolis, MN, July 1994.

Morales, J.M., (1989). Analytical procedures for estimating freeway traffic congestion. TRB Research Circular, 344, pp. 38-46

Nam, D., and Mannering, F.L., (2000). An Exploratory Hazard-Based Analysis of Highway Incident Duration. *Transportation Research* Part A, Vol. 34, pp. 85–102.

Nanthawichit, C., Nakatsuji, T., and Suzuki, H., (2003). Application of Probe-Vehicle Data for Real-Time Traffic-State Estimation and Short-Term Travel-Time Prediction on a Freeway, *Journal of the Transportation Research Board*, No. 1855, pp. 49-59.

Ozbay K., and Kachroo, P., (1999). *Incident Management in Intelligent Transportation Systems*, Artech House, Boston, MA.

Pearson, R. K. (2002). Outliers in process modeling and identification. *IEEE Transactions on Control Systems Technology*, 10, 55–63.

Pu, W., (2011). Analytic Relationships Between Travel Time Reliability Measures, *Journal of the Transportation Research Board*, No. 2254, pp. 122-130.

Saka, A.A., Jeihani, M., and James, P.A., (2008). Estimating Traffic Recovery Time for Different Flow Regimes on Freeway. MD-09-SP708B4L. Maryland Department of Transportation, Baltimore, MD.

Schofield, M., and Harrison, R., (2007). Developing Appropriate Freight Performance Measures for Emerging Users. Final report, Southwest Regional University Transportation Center, SWUTC/07/473700-00073-1. http://d2dtl5nnlpfr0r.cloudfront.net/swutc.tamu.edu/publications/technicalreports/473700-00073-1.pdf, accessed November 2012

Short, J. and Jones, C., (2008). Utilization of Wireless Truck Position Data to Estimate Transportation System Demand, 10th International Conference on Applications of Advanced Technologies in Transportation, Athens Greece.

Short, J., and Murray, D., (2008). Assessment Of Interstate Truck Parking Needs Using Wireless Truck Position Data. 15th World Congress on Intelligent Transport Systems, New York, NY.

Short, J., Pickett, R., and Christianson, J., (2009). Freight Performance Measures Analysis of 30 Freight Bottlenecks, American Transportation Research Institute, Arlington, VA. http://atri-online.org/2008/01/01/freight-performance-measures-analysis-of-30-freight-bottlenecks/, accessed November 2012.

Skabardonis, A., Petty, K.F., Bertini, R.L., Varaiya, P.P., Noeimi, H., and Rydzewski, D. (1997). I-880 Field Experiment: Analysis of Incident Data, *Journal of the Transportation Research Board,* No. 1603, pp. 72-79.

Skabardonis, A., Petty, K.F., Noeimi, H., Rydzewski, D., and Varaiya, P.P., (1996). I-880 Field Experiment: Data-Base Development and Incident Delay Estimation Procedures, *Journal of the Transportation Research Board*, No. 1554, pp. 204-212.

Skabardonis, A., Petty, K.F., and Varaiya, P.P., (1999). Los Angeles I-10 Field Experiment: Incident Patterns, *Journal of the Transportation Research Board*, No. 1683, pp. 22-30.

Smith, B.L., (2006). Purchasing Travel Time Data: Investigation of Travel Time Data Service Requirements, *Journal of the Transportation Research Board*, No. 1978, pp. 178-183.

Sullivan, E. (1997) New model for predicting freeway incidents and incident delays. *ASCE Journal of Transportation Engineering* 123, 267-275.

Turner, S., Margiotta, R., & Lomax, T., (2004). Lessons Learned: Monitoring Highway Congestion and Reliability Using Archived Traffic Detector Data, FHWA.

Turner, S., Sadabadi, K.F., Haghani, A., Hamedi, M., Brydia, R., Santiago, S., & Kussy, E., (2011). *Private Sector Data for Performance Management*. Final Report, USDOT, FHWA, Washington DC. http://www.ops.fhwa.dot.gov/publications/fhwahop11029/index.htm, accessed November 2012.

Varma, A., (2008). Measurement Sources for Freight Performance Measures and Indicators. Final Report, MN/RC 2008-12, Minnesota Department of Transportation, St. Paul, MN. http://www.dot.state.mn.us/ofrw/PDF/2008PerfIndic.pdf, accessed November 2012.

Wilks, D. S. *Statistical methods in the atmospheric sciences*, 2nd ed., Academic Press, Boston, 2006.

Yu, L., Yu, L., Qi, Y., Wang, J., and Wen, H., (2008). Traffic Incident Detection Algorithm for Urban Expressways Based on Probe Vehicle Data, *Journal of Transportation Systems Engineering and Information Technology*, Vol.8(4), p.36-41

Zeng, X., and Songchitruksa, P., (2010). Empirical Method for Estimating Traffic Incident Recovery Time, *Journal of the Transportation Research Board*, No.2178, p.119-127, Washington, DC.

Appendix

Appendix A: Sample Incident Data

1. <u>An example of MN511 incident XML String</u>


```
<feu:full-event-update>
  <message-header>
  <sender>
    <organization-id>MNSEG</organization-id>
    <center-id>MNSEG</center-id>
  </sender>
  <message-type-version>1</message-type-version>
  <message-number>300271</message-number>
  <message-time-stamp>
    <date>20130102</date>
    <time>062928</time>
    <utc-offset>-0600</utc-offset>
  </message-time-stamp>
  <message-expiry-time>
    <date>20130102</date>
    <time>102928</time>
    <utc-offset>-0600</utc-offset>
  </message-expiry-time>
  </message-header>
  <event-reference>
    <event-id>MNSEG-300271</event-id>
    <update>1</update>
  </event-reference>
  <event-indicators>
    <event-indicator>
      <priority>10</priority>
    </event-indicator>
  </event-indicators>
  <headline>
    <headline>
      <winter-driving-index>driving conditions good</winter-driving-index>
    </headline>
  </headline>
  <details>
    <detail>
      <element-id>1</element-id>
      <descriptions>
        <description>
          <phrase>
            <winter-driving-index>driving conditions good</winter-driving-index>
          </phrase>
        </description>
      </descriptions>
```

```
  <locations>
    <location>
      <location-on-link>
        <link-ownership>Minnesota</link-ownership>
          <route-designator>MN 41</route-designator>
          <primary-location>
            <geo-location>
              <latitude>44765709</latitude>
              <longitude>-93578467</longitude>
            </geo-location>
            <linear-reference>0.0</linear-reference>
          </primary-location>
          <secondary-location>
            <geo-location>
              <latitude>44891806</latitude>
              <longitude>-93580601</longitude>
            </geo-location>
          <linear-reference>9.382</linear-reference>
          </secondary-location>
        <link-direction>not directional</link-direction>
        <linear-reference-version>0</linear-reference-version>
    </location-on-link>
  </location>
</locations>
<times>
  <update-time>
    <date>20130102</date>
    <time>062928</time>
    <utc-offset>-0600</utc-offset>
  </update-time>
<valid-period>
  <duration>2000000</duration>
</valid-period>
</times>
</detail>
</details>
</feu:full-event-update>
```


2. <u>SQL Table</u>

```
CREATE TABLE events
(
  send_org_id text, send_ctr_id text,
  msg_version integer,
  msg_number bigint,
  msg_timestamp timestamp with time zone,
  msg_expire_time timestamp with time zone,
  event_id text,
  event_update integer,
  event_status text,
```

```
event_priority integer,
headline_text text,
d_desc_element_id text,
d_desc_phrase_text text,
d_desc_cause_text text,
d_desc_advice_text text,
d_desc_qualifier_text text,
qty_length_affected integer,
qty_link_delay integer,
qty_link_headway integer,
qty_link_TT integer,
qty_veh_involved integer,
qty_car_involved integer,
qty_truck_involved integer,
qty_bus_involved integer,
qty_wind_dir integer, qty_wind_spd integer,
qty_air_temp integer, qty_humidity integer,
qty_visibility integer,
qty_park_space integer,
qty_park_occ integer,
qty_water_depth integer,
qty_adj_snow_depth integer,
qty_road_snow_depth integer,
qty_road_snow_pack_depth integer,
qty_ice_thickness integer,
qty_pavement_temp integer,
qty_spd_limit_advisory integer,
qty_spd_limit integer,
qty_spd_limit_truck integer,
qty_restrict_length integer, qty_restrict_height integer,
qty_restrict_width integer, qty_restrict_weight_veh integer,
qty_restrict_weight_axle integer,
qty_restrict_axle_count integer,
d_desc_additional_text text,
d_loc_area_id integer,
d_loc_link_owner text,
d_loc_route_designator text,
d_loc_primary_loc real[3],
d_loc_secondary_loc real[3],
d_loc_link_dir text,
d_loc_link_align text,
d_loc_link_lref_ver text,
d_times_update_time timestamp with time zone,
d_times_end_time timestamp with time zone,
d_times_valid_duration integer,
d_times_start_time timestamp with time zone,
d_times_recur_days text,
d_times_recur_schedules text,
d_times_recur_utc_offset text,
```

```
  source_org_id text, source_ctr_id text,
  PRIMARY KEY(event_id, event_update)
)
WITH (
  OIDS=FALSE
);
ALTER TABLE events OWNER TO postgres;
```

Appendix B: Java Application for Bivariate Analysis

A Java program was developed to perform hourly bivariate analysis for each roadway segment. This program was designed to handle real-time truck vehicle data for future deployment. The application classes are summarized as follows.

- **IDet_Main.java** – This is the main class that specifies which year, month, and route to perform bivariate analysis.
- **Process_IDet.java** – This class performs bivariate analysis based on selected roadway segment and time of day. It includes methods to process by hour, by day or by location.
    - process_IDet()
    - exeByHr(int, int, int, String)
    - exe1(int, int, int, String)
    - exeByLoc(int, int, int, String)
    - exeByDay(int, int, int, String)
    - getSegmentSize(int)
    - getRouteSize()
    - queryDailySpeedData(String)
    - querySpeedData(String)
    - init_JDBC()
    - writeStr2File(String)
    - write2CSVFile(String, String)
- **myBivariate.java** – This is bivariate model class that performs variances, correlations and k-value calculations.
    - myBiVariate(myArray, myArray)
    - calc_K_value(double, double)
    - matrix2by2Inverse()
    - getCorrelation()
- **myArray.java** – This is the data array class for travel time and travel time differences.
    - myArray(double[])
    - getSum()
    - getMean()
    - getVariance()
    - getStdDev()
    - median()
    - max()
    - min()

Appendix C: Smartphone App

Introduction

An Android application was developed to implement the idea of an Incident decision support system. The Android application shows the current location on a google map and changes its position as the current location of the user is updated. As the location is updated, the application sends its location and date-time data to the server to get information about any incidents that might have taken place on the route that the user is travelling. The application provides settings to adjust date, time, look-ahead distance, speech enabling etc.

Map Details

For the purpose of map implementation, Google Maps Android API v2 has been used. To use the map API, the google-play-services_lib had to be included as a library. This project can be found in the extras folder under Android-sdk. In addition, a license key is needed from Google API console to get permission to use the application. This key has to be written in the Android application manifest file, otherwise the map won't appear on the screen. User configurable options can be set by the user when the app is installed. These settings will be used as default to provide incident information to the driver. Drivers are not required and allowed (by law) to input any settings while driving.

Design and User Interface

- Screen 1

  As soon as we click the application icon, the application starts. The application checks the setting of location services. If the location services are not enabled, it takes users to the location settings screen where they can turn on the location settings as illustrated in Figure C-1.
- Screen 2

  If the location setting is enabled the application can now be started. Figure C-2 is the screenshot of the first screen which shows the current location of the user. The application uses Google Maps and the map type is normal. The two buttons in the action bar are to start the incident decision support system and stop it respectively. If the start button is not pressed the application will behave as a normal location tracking application where you can just see your current location. When the start button is pressed the application queries the database by sending the location.
- Screen 3

  Figure C-3 is the screen shot of the screen that appears after we press the "action overflow" (or more option) button at the upper right of the screen. All the options available are listed in the screen.
    - Choose Date – Opens Date-Picker Dialog
    - Choose Time – Opens Time-Picker Dialog
    - Other Settings – Opens Normal Settings dialog

- Reset Settings – resets Date, Time and Other Settings.

- Screen 4
  The screen as displayed in Figure C-4 allows users to choose the date. This appearance may vary in different Android versions.
- Screen 5
  Figure C-5 is the screen shot of the time selection view. If the user wants to change the time of the query he can do so by changing the time or else the application will use the system default time. The screen display is from Android 5.0. The appearance may vary depending on different Android versions.

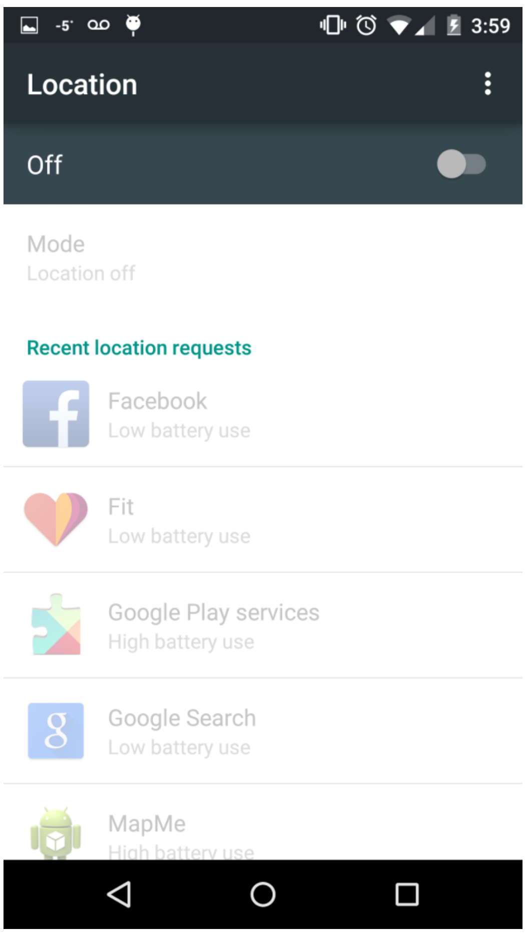


Figure C-1 Location Service Setting

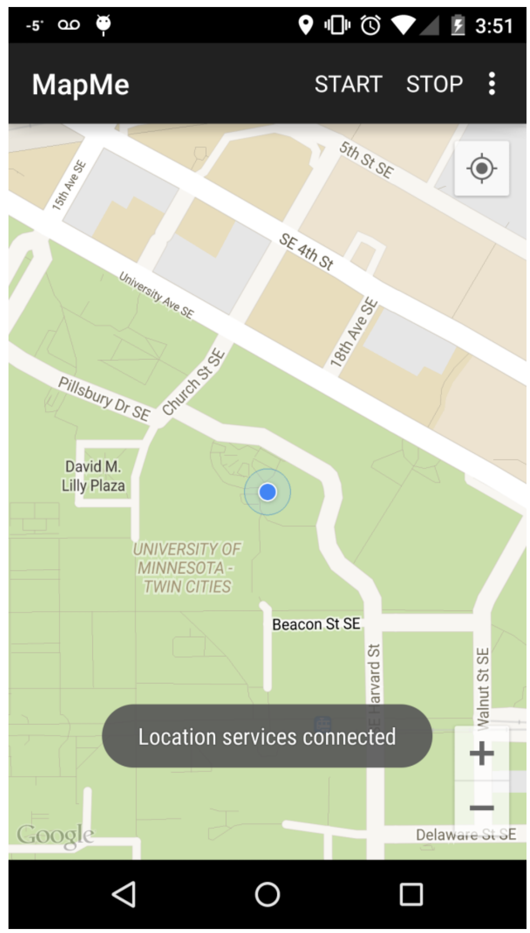


Figure C-2 Location Map

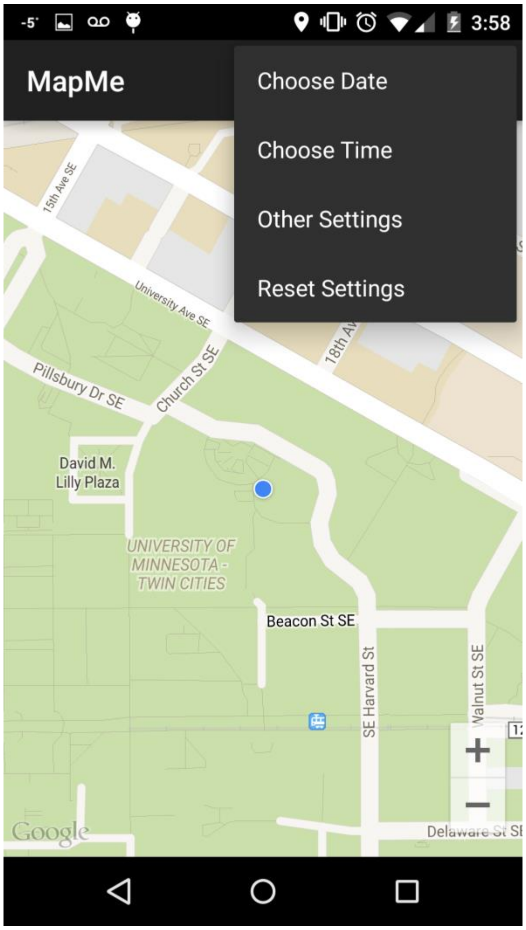


Figure C-3 Options & Settings

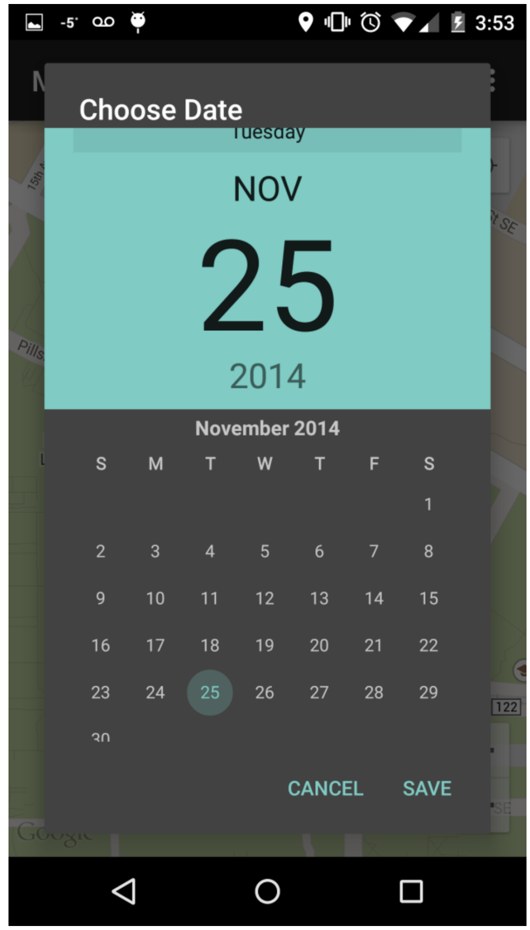


Figure C-4 Select Date

- Screen 6
  Figure C-6 is the screen shot attached for other settings. When users press the other settings option from the main screen menu options, a dialog appears which looks like this in Android 5.0 (may look a bit different on other Android version devices). The fields present in the dialog fragment are the following:
  - Look ahead miles: This option is set to determine the distance ahead for which you need the incident info from. For example if you set the option as 10, the application will query the database for incidents that might have occurred within 10 miles ahead in the driving direction.
  - Enable Voice: This option can be selected or deselected to enable or disable voice notifications. If the option is enabled, the application starts "speaking" an alert message if there is any incident on the current route. If this option is unchecked the "speak frequency" option is disabled.
  - Repeat after: This option will be enabled if the enable voice option is checked. This option is to select the time period with which the applications repeats the alert message. For example, if the selected option is "15 sec", the application

repeats the warning message every 15 seconds until the incident point is crossed.

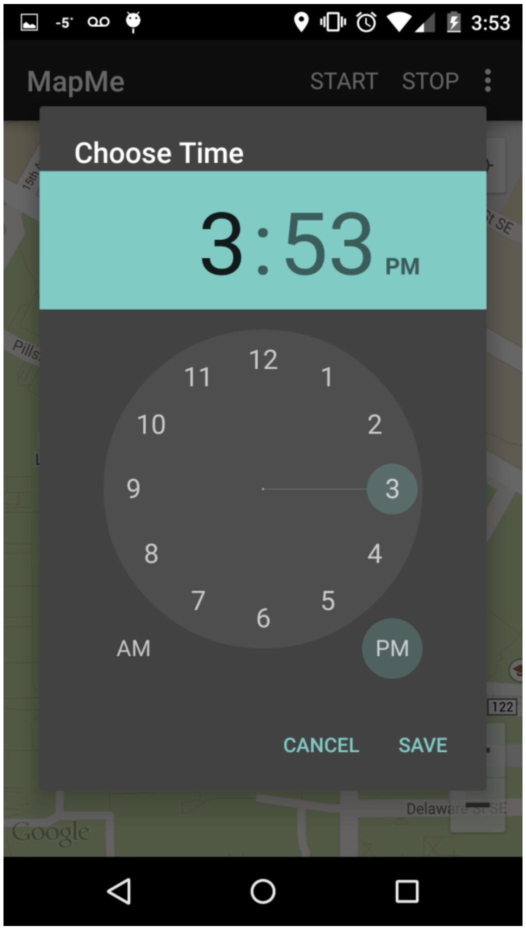


Figure C-5 Select Time

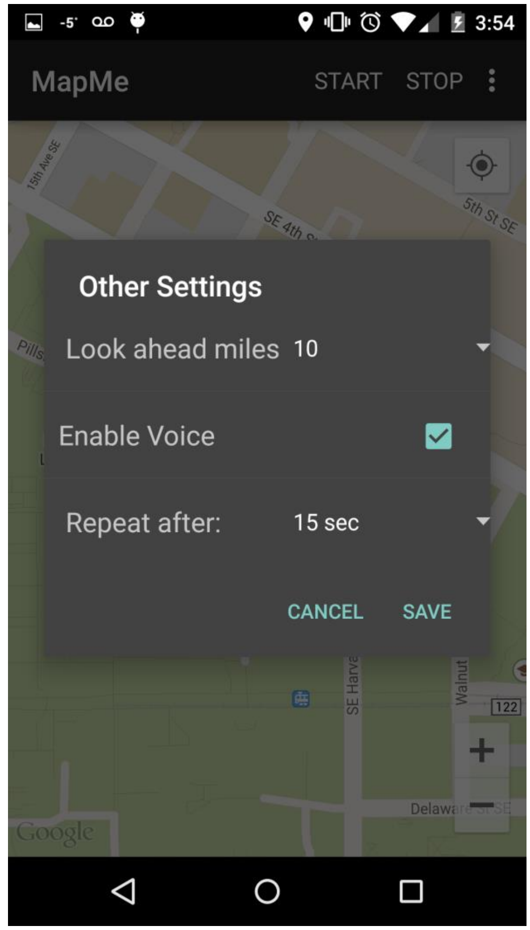


Figure C-6 Other Settings

Code Structure and Class Files

The application code is a normal Android project that contains resource files, source files, libraries and Android manifest file. The source folder consist of 5 java files.

- MainActivity.java – This is the main file of the application. This is the entry point of the application. The OnCreate() method of the activity initializes all the variables and inflates map layout. This class has the file which send the query to the database through ServerAsyncTask.java.
  - MainActivity implements OnMapClickListener, OnMapLongClickListener, OnMarkerClickListener, OnInfoWindowClickListener, OnInitListener, LocationListener, GooglePlayServicesClient.ConnectionCallbacks, GooglePlayServicesClient.OnConnectionFailedListener which ae all imported from the Google Map libraries.
  - MainActivity also contains the code for the settings dialog i.e. Date selection dialog, Time selection Dialog and the other settings dialog through which we can

change the settings manually. The mainActivity creates a ServerAsynctask after every minimum number of seconds that is set by the user in the settings

- ServerAsyncTask.java – This class file extends AsyncTask that runs in the background and sends the query to the server and receives response and then acts according to the response from the server. The serverAsyncTask collects all the information regarding location date, time etc. and sends the query and waits for the response from the server. As soon as the response comes. It parses the response and then validates it. If the response is worth creating an alarm it creates a marker on the map and then warns the user about the incident point ahead on the route. It keeps repeating the warning until the incident point is crossed by the user.
- FixedSizeLinkedQueue.java: This class is a queue implemented by using linked lists. The maximum nodes in this queue can be 10. Every time a location update is recorded it is sent to this queue and thus entered to this queue and the 10^{th} last node becomes obsolete and moves out of the queue.
- Node.java: This is a node of the LinkedList that is used in the queue. The value field of the Node contains the Latitude-Longitude value.
- Utils.java: This file contains all the utilities functions which are used in the settings. Utils class is used to retrieve the date, time, status of the speech option etc. Every time the settings are changed the variables in the Utils file is updated.

Speech Engine Response Interpretation

- If the distance of the incident is less than 1 mile, it says incident occurred ahead ‘xx’ minutes ago. If the distance of the incident is more than 1 mile, it says “incident occurred ‘yy’ miles ahead ‘xx’ minutes ago.
- If the incident occurred less than 15 minutes ago, then the engine says incident occurs a few minutes ago or else it says incident occurred ‘yy’ miles ahead ‘xx’ minutes ago.
- The speech engine mentions the route id of the incident assuming that the driver is aware of the current route so that he can find out whether the incident is on the same route or some other route.

Appendix D: HTTP Servlet Development

Introduction

A HTTP Servlet was developed which acts as the middleware between the Android application and the central database server. This HTTP Servlet reads the request from the Android application and then creates a connector object (a java class which has been developed to make the actual communication with the central database). The connector takes all the parameters sent by the Android application and then creates the query string in the required format. The query string is sent to the server and then the object waits for a response from the database server. As soon as the response arrives, it checks for the validity of the response. If the number of responses matches the expected number, it then passes the result to the 'Servlet' object on the webserver and then to the smartphone application.

Library

postgresql-9.1-903.jdbc3 – The Servlet Project contains the library postgresql-9.1-903.jdbc3.jar which provides the APIs for database connection. PostgreSQL is an object relational database management system. As a database server its main function is to store data securely and retrieve it later as requested by software applications. The Servlet code has this library to access the database.

Code Structure and Class Files

The Servlet code consists of a java Project that contains two java class files and a library. The class files are as follows:

- **MTOProject.**java – This class file extends HTTPServlet which provides overrides two methods.
    - doget() – In our project we are not using this method for exchange of information. This method is just used to check the server availability.
    - dopost() – This method is used to receive all the query details from the application and then passed on to the connector object. The query details that this method expects are:

      ```
      String date = request.getParameter("date");
      String curLon = request.getParameter("curLon");
      String curLat = request.getParameter("curLat");
      String prevLon = request.getParameter("prevLon");
      String prevLat = request.getParameter("prevLat");
      String look_ahead = request.getParameter("look_ahead");
      ```

      The string parameters are hardcoded keys that are paired with their respective values.
- DataBaseConnector.java:

  This class is created to execute all the database queries. It uses the APIs from the "postgresql-9.1-903.jdbc3.jar" library and gives the expected output response. The important methods of the class are as follows.

- ***connect_sql (String db_name)***
  This methods connects to the SQL JDBC. In our case the db_name is hardcoded as ”TCMA”. If its unable to establish the connection it throws a ClassNotFoundException . The object is successful in establishing a connection if Connection pconn is assigned a value:
  pConn =
  DriverManager.getConnection("jdbc:postgresql://your_ip_address:port_number/" +db_name, "username", "password");
  where the IP address belongs to the central database server and the user name and password are used to access Postgres database.

- ***create_SL_table(String table_name)***
  This method is called to create a table in the database and then save data to PostgreSQL database.

- ***get_table_data (String table_name, String date, String curLon, String CurLat, String prevLon, String prevLat, String look_ahead)***
  This is the most important function for this class. This method is used to query the database and get the response. The various parameters that it takes are:
  table_name - This is the table in the database from which trying to get information. In this case it is “get_incident_info_sorted”.

  date – This is generally the current date that the application sends, anyhow the date can be customized through the settings tab. Whatever the date is sent by the application the system queries for it.
  curLon – This is the current longitude of the user.
  CurLat – This is the current latitude of the user. This curLon and curLat determine the current location of the user. The aim is to check if an incident has occurred beyond this location.
  prev_Lon – This is the 10th last known longitude of the user. If 10 locations are not known then the first longitude that was known would be sent by the application.
  prevLat – This is the 10th last known latitude of the user. If 10 locations are not known then the first latitude that was known would be sent by the application. The previous location is required to interpret the direction of travel of the user. Direction is important to detect what incident points to return
  look_ahead – This parameter is used by the database server to determine its search distance while querying incident information based on a vehicle’s current position.