

**Technical Report documentation Page**

1. Report No. SWUTC/14/600451-00083-1		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle <b>REAL TIME FREEWAY INCIDENT DETECTION</b>			5. Report Date <b>April 2014</b>		
			6. Performing Organization Code		
7. Author(s) Moggan Motamed and Randy Machemehl			8. Performing Organization Report No. 600451-00083-1		
9. Performing Organization Name and Address Center for Transportation Research University of Texas at Austin 1616 Guadalupe Street, Suite 4.200 Austin, Texas 78701			10. Work Unit No. (TRAIS)		
			11. Contract or Grant No. DTRT12-G-UTC06		
12. Sponsoring Agency Name and Address Southwest Region University Transportation Center Texas A&M Transportation Institute Texas A&M University System College Station, Texas 77843-3135			13. Type of Report and Period Covered		
			14. Sponsoring Agency Code		
15. Supplementary Notes Supported by a grant from the U.S. Department of Transportation, University Transportation Centers Program and general revenues from the State of Texas.					
16. Abstract The US Department of Transportation (US-DOT) estimates that over half of all congestion events are caused by highway incidents rather than by rush-hour traffic in big cities. Real-time incident detection on freeways is an important part of any modern traffic control center operation because it offers an opportunity to maximize road system performance. An effective incident detection and management operation cannot prevent incidents, however, it can diminish the impacts of non-recurring congestion problems. The main purpose of real-time incident detection is to reduce delay and the number of secondary accidents, and to improve safety and travel information during unusual traffic conditions. The purpose of this project is to evaluate two recently developed automatic incident detection algorithms. The majority of automatic incident detection algorithms are focused on identifying traffic incident patterns but may not adequately investigate possible similarities in patterns observed under incident-free conditions. When traffic demand exceeds road capacity, the traffic speed decreases significantly and the traffic enters a highly unstable regime often referred to as "stop-and-go" conditions. The most challenging part of real-time incident detection is recognition of traffic pattern changes when incidents happen during stop-and-go conditions. This work describes a case study evaluation of two recently evolved incident detection methods using data from the Dallas, TX traffic control center.					
17. Key Words Incident Detection, Dynamic Time Warping, Support Vector Machine DTW, SVM			18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161		
19. Security Classif.(of this report) Unclassified		20. Security Classif.(of this page) Unclassified		21. No. of Pages 42	
				22. Price	



# **REAL TIME FREEWAY INCIDENT DETECTION**

by

Moggan Motamed, M.S.  
Graduate Research Assistant  
The University of Texas at Austin

and

Rand Machemehl, P.E.  
Professor  
The University of Texas at Austin

Report SWUTC/14/600451-00083-1  
Project 600451-00083-1

Performed in cooperation with the  
Southwest Region University Transportation Center

Center for Transportation Research  
1616 Guadalupe Street, Suite 4.202  
The University of Texas at Austin  
Austin, Texas 78712

April 2014



## EXECUTIVE SUMMARY

Under medium to heavy traffic conditions, the promptness of response after an incident is a direct function of the detection time. Accurate and fast incident detection is essential for subsequent management plans that aim to reduce incident based congestion. Highway traffic surveillance systems are widely used for incident management, real-time traffic management, traveler information, and hazard evacuation. Some of the most widely used methods are closed circuit television systems, driver reports processing, highway crew patrols, and automatic incident Detection (AID) systems.

Based on our literature review, most traditional automated incident detection algorithms use roadway-based point data. There are several disadvantages to using point data for incident detection. The algorithms using loop data suffer from high rates of false alarms.

To improve incident detection data, it is reasonable to expect that using multiple data sources, e.g., fixed detectors (collecting point data) and probe vehicles (collecting spatial data), could enhance the input data reliability and completeness and hence improve the performance of an incident detection system.

Recently, a short-term congestion detection algorithm for freeway sections has been proposed using dynamic time warping (DTW) and support vector machine (SVM). Some studies show a higher detection rate than Artificial Intelligence algorithms with lower false alarm rates. The proposed methods are a data mining and time series classification that represent an interdisciplinary confluence of a set of disciplines, including statistics, machine learning, Artificial Intelligent, and information science. The literature review, methodology, and application of these two models will be presented in the following sections.

Dallas traffic data are being used to develop the incident detection model. The Dallas Transportation Management Center (DalTrans) is the nerve center for urban freeway and highway systems in Dallas. Real time information is gathered from many sources such as electronic sensors in the pavement, freeway call boxes, video cameras, 911 calls, officers on patrol, highway crews, motorist cellular calls, and commercial traffic reporters. The information is sent to the DalTrans 24-hours a day, seven days a week and traffic information (speed and flow) are gathered every 5 minutes on a per lane basis.

In this research two groups of experiments were performed to evaluate the two most robust incident detection algorithms: DTW and SVM. Evaluation of the DTW and SVM

algorithms revealed that both can successfully classify traffic conditions into two categories – incident, non-incident – during peak hours. As algorithm speed is a concern for real-time incident detection, both model predictions proved to be fast. Models were trained on a network based on freeway segments in Dallas, TX. Comparing these two methods, application of DTW in the field of transportation is quit new. The concept of DTW is simpler than SVM. In regards to the time, available data, and nature of DTW we did not have enough data to totally validate our DTW model however, that work is continuing. The advantage of using SVM is that it does not require a large dataset to train and validate the model. On the other hand, the accuracy of SVM is highly depend on the kernel function and the choice of parameters which requires solution of an optimization problem (in this research grid search has been applied to find C and Y). The work completed and described here indicates that a more accurate comparison of these two methods could be developed through generation of simulation data after using real world data for model calibration.

## **DISCLAIMER**

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Center Program, in the interest of information exchange. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

## **ACKNOWLEDGEMENTS**

The authors recognize that support for this research was provided by a grant from the U.S. Department of Transportation, University Transportation Centers Program to the Southwest Region University Transportation Center which is funded, in part, with general revenue funds from the State of Texas.



**CONTENTS**

Introduction..... 1

Review of current incident detection ..... 3

Evaluation of current incident detection algorithms ..... 5

Algorithm Development ..... 7

    DTW Literature ..... 7

    DTW using Dallas Data ..... 8

        Data..... 8

        DTW Concept..... 9

        Modifications and constraints for DTW ..... 10

        Methodology and results ..... 10

    SVM Literature ..... 20

    SVM Concept..... 20

    Modifications and constraints in SVM..... 22

    SVM Case Generation and Methodology ..... 22

    Results ..... 24

Conclusions..... 27

References..... 29

**List of Figures**

Figure 1. Locations, time, and impact of incidents on US 75..... 8  
Figure 2. A series of hypothetical sequences of two inputs, x and y. .... 9  
Figure 3. Sequences of warping map matrix to find optimal warping path..... 10  
Figure 4. Incident detection algorithm (Hi-ri-o-tappa, 2011). .... 11  
Figure 5. Incident detection pattern (Hi-ri-o-tappa, 2011). .... 11  
Figure 6. Typical Speed profile for 5 selected locations. .... 12  
Figure 7. Incident Speed profile for 5 selected locations..... 13  
Figure 8. DTW for non-incident. .... 14  
Figure 9. Graphical view of DTW output for a non-incident. .... 15  
Figure 10. DTW cost Matrix for Incident. .... 16  
Figure 11. Graphical view of DTW output for incident. .... 17  
Figure 12. K-means thresholds. .... 19  
Figure 13. Hyperplane through two linearly separable classes..... 20  
Figure 14. Linear separating hyperplanes for the non-separable case. Source:..... 21

**List of Tables**

Table 1. SVM sample input data and result..... 23  
Table 2. Input data counts..... 23  
Table 3. Input variables summary..... 24  
Table 4. Base model output. .... 25  
Table 5. Base model false detection details. .... 25  
Table 6. Different scenarios comparison. .... 26

## **INTRODUCTION**

Under medium to heavy traffic conditions, the promptness of response after an incident is a direct function of the detection time. Accurate and fast incident detection is essential for subsequent management plans that aim to reduce incident based congestion. Highway traffic surveillance systems are widely used for incident management, real-time traffic management, traveler information, and hazard evacuation. Some of the most widely used methods are closed circuit television systems, driver reports processing, highway crew patrols, and automatic incident Detection (AID) systems.

Generally incident detection includes two sequential components: collecting traffic data, and analyzing the collected data. For many locations a large investment has been made in detection hardware so there is no option of choosing type of detector hardware. In this case, the best analysis model may be determined by product of the detection system, that is, the available data. If one has the option of choosing detector hardware as well as analysis or detection algorithm(s) a great deal more flexibility may evolve into a potentially more robust system.

Based on our literature review, most traditional automated incident detection algorithms use roadway-based point data. There are several disadvantages to using point data for incident detection among these is the fact that the algorithms using point inductance loop detector data suffer from high rates of false alarms.

To improve incident detection data, it is reasonable to expect that using multiple data sources, e.g., fixed detectors (collecting point data) and probe vehicles (collecting spatial data), could enhance the input data reliability and completeness and hence improve the performance of an incident detection system.

In terms of incident detection algorithm, most studies have focused on machine learning methods applied to single point detector data due to the uncertainty of incident patterns. The machine learning or artificial intelligence algorithms are a set of procedures that apply inexact reasoning and uncertainty in complex decision-making as well as data analysis processes where all decisions are made by machine.

Recently, a short-term congestion detection algorithm for freeway sections has been proposed using dynamic time warping (DTW) and support vector machine (SVM). Some studies show a higher detection rate than Artificial Intelligence algorithms with a lower false alarm rate. These two proposed methods are a data mining and time series classification that represent the

confluence of a set of disciplines, including statistics, machine learning, Artificial Intelligent, and information science. The Literature review, methodology, and application of these two models will be presented in the following sections.

## REVIEW OF CURRENT INCIDENT DETECTION

Incident detection algorithms are typically categorized into five major groups depending on the type of operations data analysis they employ.

- Comparative algorithms
- Statistical algorithms
- Time-series and filtering based algorithms
- Traffic theory based algorithms
- Advanced algorithms.

The most challenging part of real-time incident detection is recognition of traffic pattern changes when incidents happen during stop-and-go conditions. Unfortunately “stop and go” conditions tend to develop on most urban freeways on most week days during both morning and evening rush hours. The Institute of Physics (2005) research shows that even small fluctuations in car-road density cause a chain reaction that can lead to a jam. It is practically impossible to obtain coherent predictions from a macroscopic traffic flow model due to occurrence of the huge number of small transient shockwaves under these conditions. Incident detection efforts are abandoned under such conditions.

Among the different traffic surveillance methods, closed circuit television (CCTV) systems, driver reports processing, highway crew patrols, and AID systems are the most widely used (Parkany and Xie, 2005). However, CCTV systems and sensor networks for AID require extensive infrastructure support.

Although many studies argue that driver-based incident detection systems (e.g., enhanced 911 services) can provide quick and accurate detection with less capital, maintenance, and operational costs, these systems do not perform well in areas with low cell phone usage or bad signal (Xie and Parkany, 2002; Mussa and Upchurch, 2000 and 1999; Walters *et al.*, 1999; Skabardonis, 1998; and Mussa, 1997). There is also always the risk of the phone call processing system becoming jammed during a severe incident. The labor intensive nature of highway crew patrols tends to limit their wide spread deployment. Consequently, highway traffic surveillance is currently limited to major highways and urban areas.

Most traditional automated incident detection algorithms use roadway-based single point data. Using single point data has several disadvantages for incident detection. Most of the time space-mean speed is not available which decreases the accuracy of the AID algorithm. These

algorithms are simple in theory and practical in operation, but would usually fail to deliver high detection rates and low rates of false alarm.

## **EVALUATION OF CURRENT INCIDENT DETECTION ALGORITHMS**

During incident occurrence, traffic measurements change. Occupancy (or density) increases upstream and decreases downstream while speed and volume decrease upstream. These differences between up- and downstream traffic measurements have been the basis of most freeway AID algorithms such as the California (Tan, 2011) and the Minnesota (Tan, 2011) ones. The California algorithm only utilizes current time occupancy information, which may produce high false alarm rates (FAR) because of dynamic traffic fluctuations. To decrease the high FAR, the Minnesota algorithm employs a cumulative sum of differences between up and downstream conditions.

AID algorithms can be categorized as macroscopic and microscopic, however most are macroscopic and most use point data. To improve incident detection data, it is reasonable to expect that using multiple data sources, e.g., fixed detectors (collecting point data) and probe vehicles (collecting spatial data), could enhance the input data reliability and completeness and hence improve the performance of an incident detection system. Building a microscopic model to mimic driving behavior would be extremely difficult, since drivers consider the current movement of the traffic stream ahead, not just one vehicle in front of them. To make microscopic modeling possible, vehicle trajectories data are required.

The trajectory data generated by vehicles in a vehicle-infrastructure integration (VII) (Ma, 2008) network have the potential to provide faster traffic condition detection and lower false alarm rates than existing infrastructure-based incident detection systems, such as inductive loop detectors, magnetometers and magnetic detectors, microwave radar, infrared, ultrasonic, acoustics and video image processing. Since 2003, the Federal Highway Administration (FHWA) has sponsored a variety of efforts that have led to the national development of the VII architecture and functional requirements (FHWA, 2005). Two large states, California (PATH, 2006) and Michigan (MDOT, 2005) are also testing various methods for implementing these programs (ITS America, 2007).

As noted, most of the traditional automated incident detection algorithms use roadway-based single point data that generally means space-mean speed is not available, which decreases the accuracy of the AID algorithm.

Based on recent studies [Hi-ri-o-tappa 2011, Oh 2005] mean speed, standard deviation of speed, headways, and flow are the best indicators for incident prediction.. According to Hi-ri-o-

tappa [2011] who conducted an evaluation of best indicators through a statistical comparison of the differences between two datasets, mean speed and standard deviation of speed are the best indicators followed by occupancy and traffic flow rate.

The literature indicates that evaluations of existing incident detection algorithms have traditionally been done through comparison of detection rates, false alarm rates and times to detection. Even though some algorithms produce detection rates of 100% (ARIMA, Bayesian and SSID), they may have either a high false alarm rate or a long time to detection (MARTIN, 2001). Other algorithms that provide low false alarm rates and time to detection (less than 1 min) included artificial neural networks (ANN), which projected a detection rate of 89% (MARTIN, 2001). Video image processing was found to have an incident detection rate of 90% and a false alarm rate of 3% (MARTIN, 2001); detection rates will improve with technological advancements.

Incident detection is a pattern classification problem. Any good classifier is a potential tool for the incident detection problem. Most recent research has focused on:

- Dynamic time warping algorithm (Hi-ri-o-tappa, 2012).
- Support vector machine (Ruey Long Cheu 2003, Ma 2010, Xiao *et al.* (2012))

## ALGORITHM DEVELOPMENT

Based on our literature review presented in the previous section, the two most recent robust incident detections are Dynamic Time Warping (DTW) and Support Vector Machine (SVM).

### DTW Literature

DTW algorithms were proposed around 1970 in the context of speech recognition, to account for differences in speaking rates between speakers and utterances. An example of this method application in this field is when one wants to find a low distance score between the sound signals corresponding to utterances “now” and “nooow” without being sensitive to the prolonged duration of the ‘o’ sound. Other applications have been in genetics—for gene sequencing and detection. DTW has also been applied for clustering and classification in the follow areas:

- Electro-cardiogram analysis (Huang and Kinsner 2002; Syeda-Mahmood *et al.* 2007; Tuzcu and Nas 2005),
- Clustering of gene expression profiles (Aach and Church 2001; Hermans and Tsiporkova 2007),
- Biometrics (Faundez-Zanuy 2007; Rath and Manmatha 2003),
- Process monitoring (Gollmer and Posten 1996).

The interesting fact about this algorithm is that it can warp other dimensions than time, for example an angle for shape recognition (Kartikeyan and Sarkar 1989; Wei *et al.* 2006; Tak 2007). Therefore, the term “time series” may even be misleading. DTW has not been widely used in the transportation field yet. Chandrasekaran *et al.* brought the concept of DTW to the transportation field for the first time to track vehicular speed variation. Since then, there have been a few more studies investigating its use in transportation:

- “Tracking vehicular speed variations by warping mobile phone signal strengths” Chandrasekaran, G., Tam Vu, Varshavsky, A., Gruteser, M., Martin, R.P., Jie Yang, Yingying Chen (2011)
- “Traffic Event Automatic Detection Based on OGS-DTW Algorithm” Zhang, N., Shi, Y., and Huang, W. (2012)
- “Traffic incident detection system using series of point detectors” Hi-ri-o-tappa, K., Likitkhajorn, C., Poolsawat, A., Thajchayapong, S. (2012)

Some studies show that the procedure may produce higher detection rates than Artificial

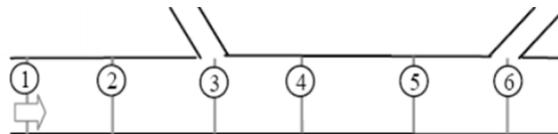
Intelligence algorithms with lower false alarm rates. For example, Hi-ri-o-toppa (2012) used upstream and downstream site changes to develop a DTW incident detection algorithm which achieved a 94% detection rate and a low false alarm rate. The proposed method uses data mining and time series classification and is the confluence of a set of disciplines, including statistics, machine learning, Artificial Intelligent, and information science.

## DTW using Dallas Data

### Data

Dallas traffic data were used to develop the incident detection model. The Dallas Transportation Management Center (DalTrans) is the nerve center for urban freeway and highway systems in Dallas. Real time information is gathered from many sources including electronic sensors in the pavement, freeway call boxes, video cameras, 911 calls, enforcement officers, highway crews, motorist cellular calls, and commercial traffic reporters. These information sources are provided to the DalTrans 24-hours a day, seven days a week. Traffic information (speed and flow) are gathered every 5 minutes on a per lane basis.

To develop a real sense of how the model might work for incident detection, we extracted five incident cases from the Dallas incident database, as shown in Figure 1. The incidents occurred at different locations along US75.

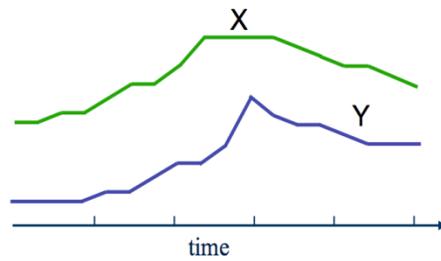


No.	RoadName	Road Direction	CrossStreet Name	DetectedTime	ClearedTime	AffectedLanes	Type
1	US 75	North	Monticello Ave	8/10/12 17:06	8/10/12 17:34	Lane1	DisabledVehicle
2	US 75	North	McCommas Blvd	8/31/12 23:23	9/1/12 1:47	Lane1	Accident
3	US 75	North	Mockingbird Ln	9/25/12 22:36	9/25/12 23:15	Lane1, Lane2	Accident
4	US 75	North	Caruth Haven Ln	9/14/12 17:20	9/14/12 17:57	Lane1, Lane2	Accident
5	US 75	North	Walnut Hill Ln	7/5/12 7:15	7/5/12 7:25	Lane1	Debris

Figure 1. Locations, time, and impact of incidents on US 75.

## DTW Concept

To illustrate the DTW concept one might assume two hypothetical sequences of X and Y as shown in Figure 2. In this example X could represent the observed query to be tested and Y might represent the reference series that we knew something about. To compare the two datasets, we measure and align their similarity or likeness into the closest matches by \*locally\* stretching or compressing portions of the series. As noted, the series are hypothetical sequences of X and Y, as shown in Figure 2, with the x axes showing the time index and the y-axis showing the outcome measure. The series may be different lengths, but measurements are taken at equidistant time points.



**Figure 2. A series of hypothetical sequences of two inputs, x and y.**

Furthermore, an optimal warping path between X and Y is a warping path  $p^*$  having minimal total cost among all possible warping paths. The total cost  $C_p(X, Y)$  of a warping path  $p$  between X and Y with respect to the local cost measure can be written as:

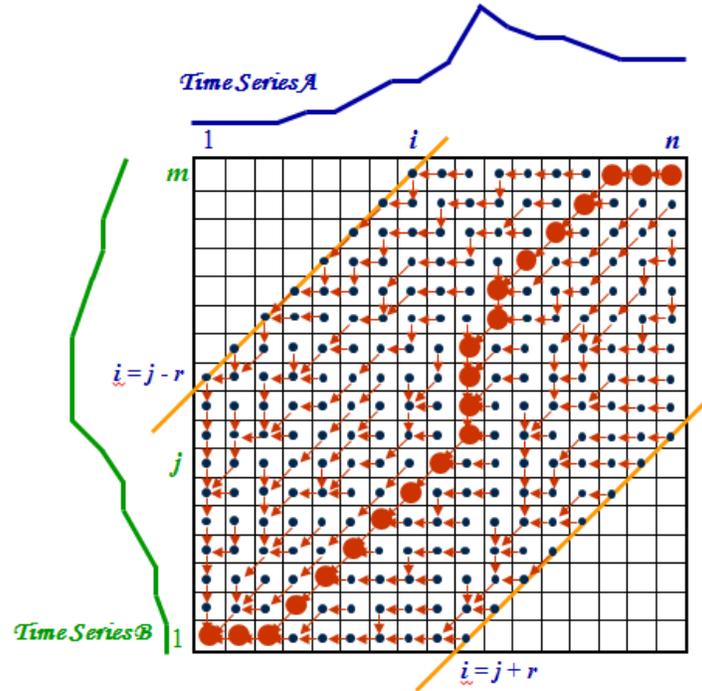
$$c_p(X, Y) := \sum_{\ell=1}^L c(x_{n_\ell}, y_{m_\ell})$$

The DTW distance  $DTW(X, Y)$  between X and Y is then defined as the total cost of  $p^*$ :

$$\begin{aligned} DTW(X, Y) &:= c_{p^*}(X, Y) \\ &= \min\{c_p(X, Y) \mid p \text{ is an } (N, M)\text{-warping path}\} \end{aligned}$$

The optimal warping path between X and Y is represented graphically in Figure 3 below as Time Series A and Time Series B. The orange "diagonal" goes from one corner to the other of the possibly rectangular cost matrix, therefore having a slope of  $M/N$ , not 1, as in the slanted Band Window. The computation is approximate: points having multiple correspondences are averaged, and points without a match are interpolated. This average between Time Series A and

B is graphically represented by large red dots in Figure 3 and interpolated points without a map are represented by blue dots and a red directional arrow. Please note that the area is not normalized by path length.



**Figure 3. Sequences of warping map matrix to find optimal warping path. Source: [http://homepages.inf.ed.ac.uk/group/sli\\_archive/slip0809\\_c/s0562005/theory.html](http://homepages.inf.ed.ac.uk/group/sli_archive/slip0809_c/s0562005/theory.html)**

### Modifications and constraints for DTW

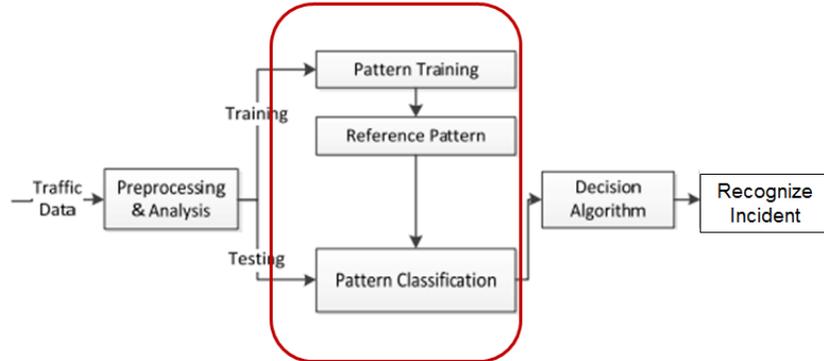
Additional constraints can be applied to the model to produce specific results, such as introducing an additional weight vector to favor the vertical, horizontal, or diagonal direction ( $w_d$ ,  $w_h$ ,  $w_v$ ) in the alignment. To constrain the slope of the admissible warping paths, we can modify the step size condition. To accelerate the speed with which the model can find the feasible solution one can put constraints on search windows.

### Methodology and results

Based on recent studies (Hi-ri-o-tappa 2011, Oh 2005) mean speed, standard deviation of speed, headway, and flow are the best indicators for incident prediction. An evaluation of the best indicators based on the statistical difference between datasets indicated that mean speed and standard deviation of speed are the best indicators followed by occupancy and traffic flow rate

(Hi-ri-o-tappa, 2011).

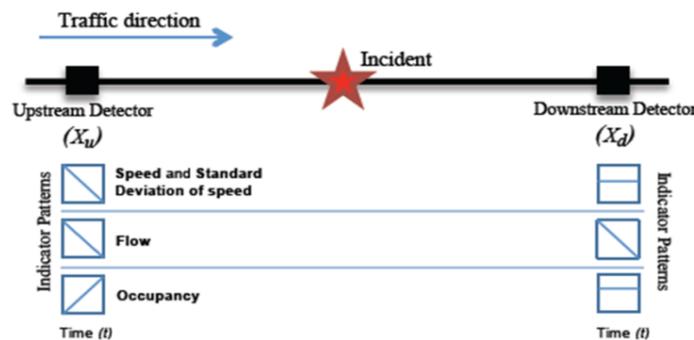
The following figure shows how the incident detection algorithm has been configured. In this algorithm DTW is used for pattern training and is followed by pattern classification and the



decision algorithm.

**Figure 4. Incident detection algorithm (Hi-ri-o-tappa, 2011).**

In the training process, the proposed system captures patterns associated with incidents in the training dataset. The patterns captured have common trends that can be described by categorizing each type of indicator. To illustrate that, we assumed the incident occurring at the location marked by the star sign marked as “incident” in Figure 5. We are expecting that Speed and Standard deviation of speed have the same trend which decrease upstream from the incident while remaining constant downstream of the incident. We do not expect to immediately see flow changes comparing up- and down-stream flows, however, occupancy has an increasing pattern at the upstream detector and a constant pattern in the downstream detector as shown in the last line of Figure 5.



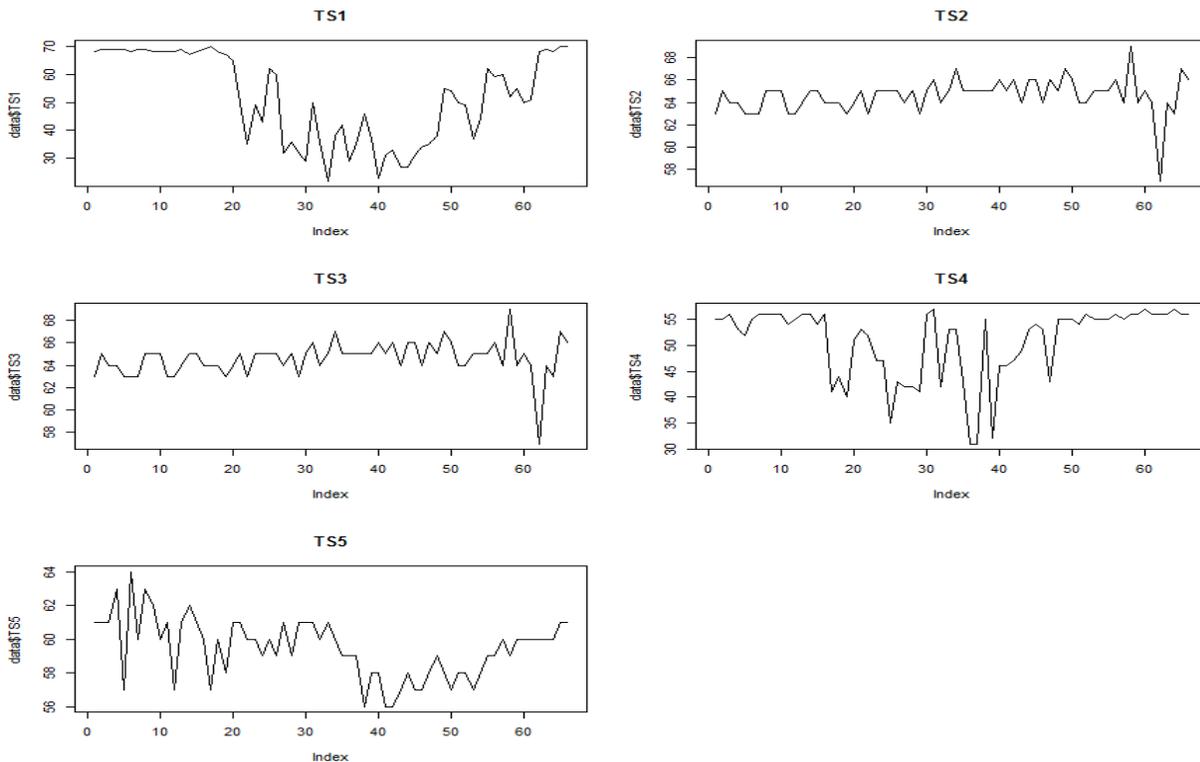
**Figure 5. Incident detection pattern (Hi-ri-o-tappa, 2011).**

As mentioned earlier, the Dallas database was used for this research. We started to

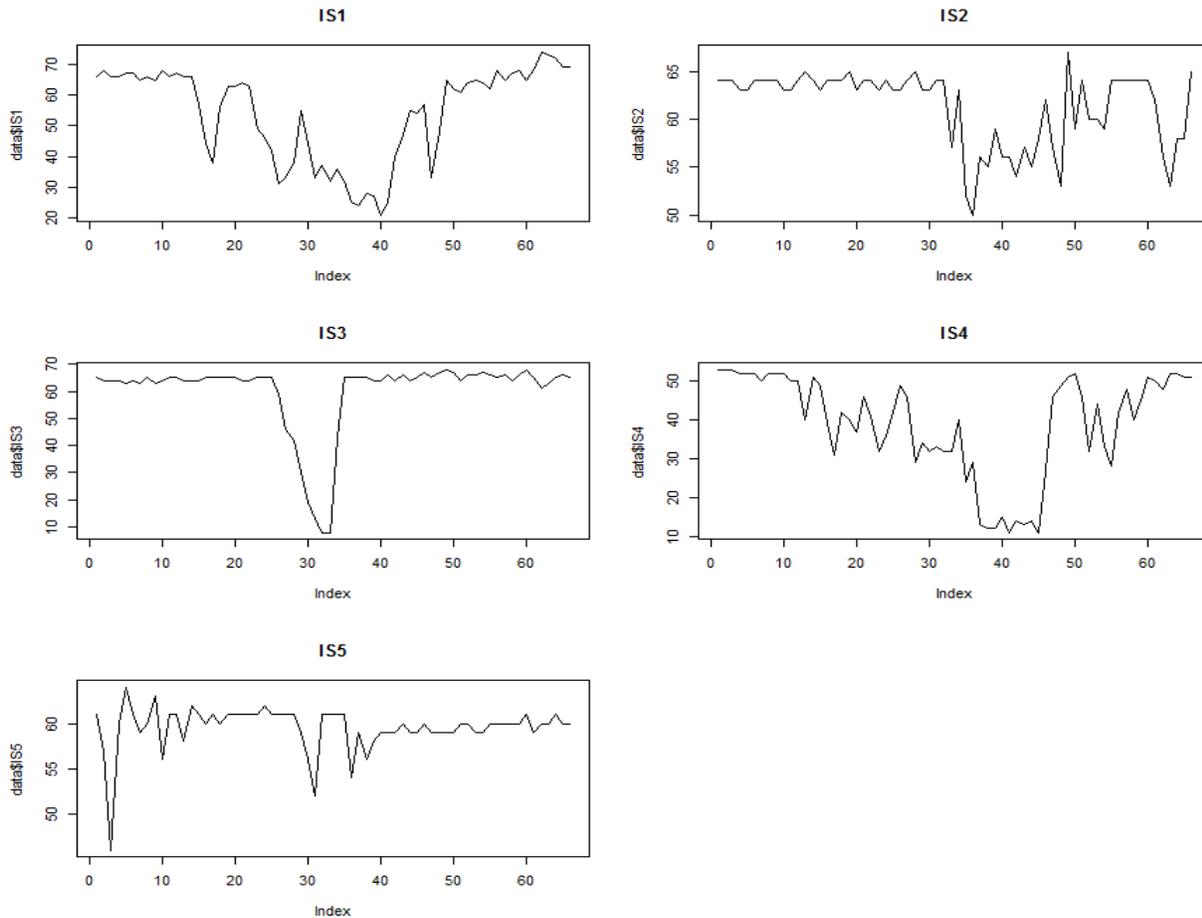
develop a base model using US-75, which is one of the most heavily used freeways in Dallas. Traffic data and incident data are provided from two different sources with different formats. Traffic data is a huge database in which for each specific detector data is recorded every 5 minutes for each lane.

From the incident data, five sequential locations were chosen (the same five locations as used in Figure 1). In all cases incidents happened northbound on lane 1 but during different times of day. To have a basis for comparing incident with non-incident situations, we chose to characterize the non-incident situation as having no incident within 5 miles before or after the specific incident location.

First we tried to find the speed pattern for each location. Then we attempt to use a smoothing technique to take some noise out. Figures 6 shows speed patterns for five locations during typical, non-incident flow and Figure 7 shows speed patterns for the same five locations during an incident. The y-axis shows the speed in miles per hour and the x-axis represents time and shows three hours before the incident and three hours after.



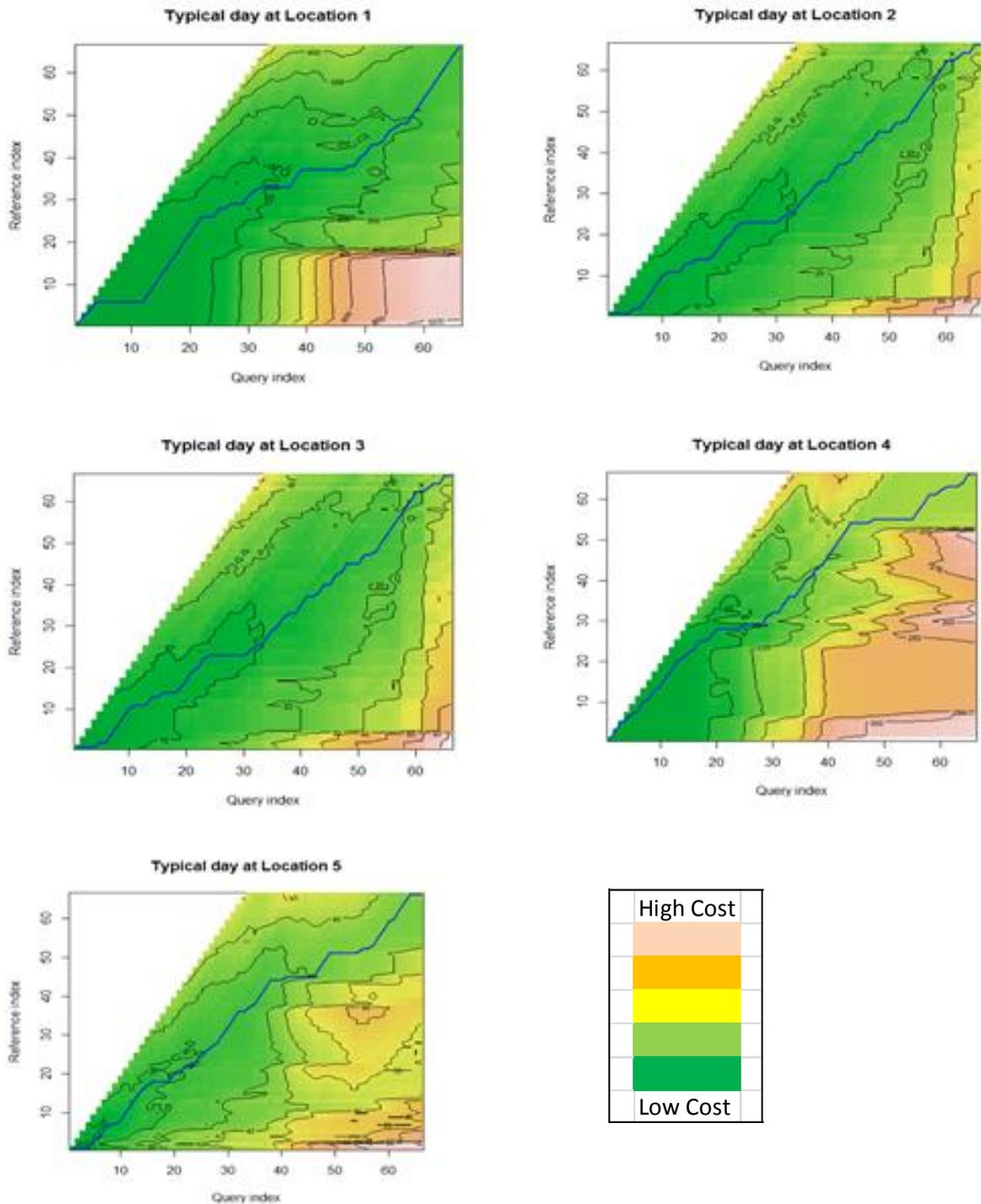
**Figure 6. Typical Speed profile for 5 selected locations.**



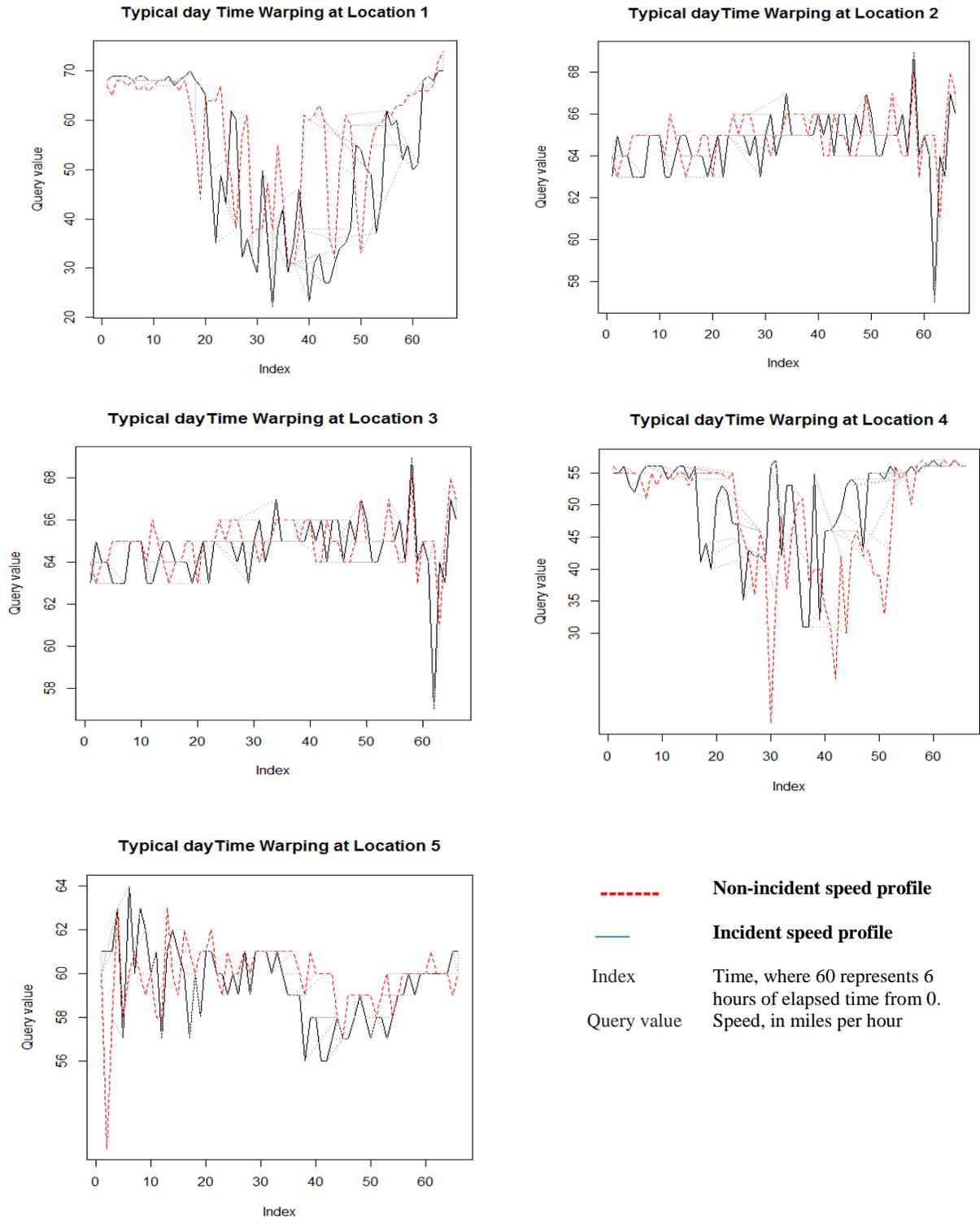
**Figure 7. Incident Speed profile for 5 selected locations.**

To develop our DTW algorithm, “R” programming language was used. First we developed a classic DTW model with our typical data to describe what we are expecting to see during non-incident times. We considered two random days in which no incident occurred within several miles radius of the data location. The y-axis represents the first day speed profile data over time in minutes and the x-axis represents the second day (both are non-incident cases). As seen in Figure 8, the model then used the data to find the shortest path between the two time series, which is represented by the blue line. The cost model shows the lowest cost or the best compatible match of the two time series. During a non-incident case, the two time series are compatible, so the cost matrix shows green colors and a diagonal path. The query index scale refers to time where 60 represents six hours of elapsed time from time zero. The query index represents three hours before the incident and three hours after with the incident occurring at query index 30. Figure 9 provides better visual expression of the result of time warping. Again,

the query index represents three hours before the incident and three hours after with the incident occurring at query index 30 and 60 represents six hours of elapsed time from time zero. The red and black lines represent traffic flow on two different typical days with the same time frame. The models are then warped to show how data points align.

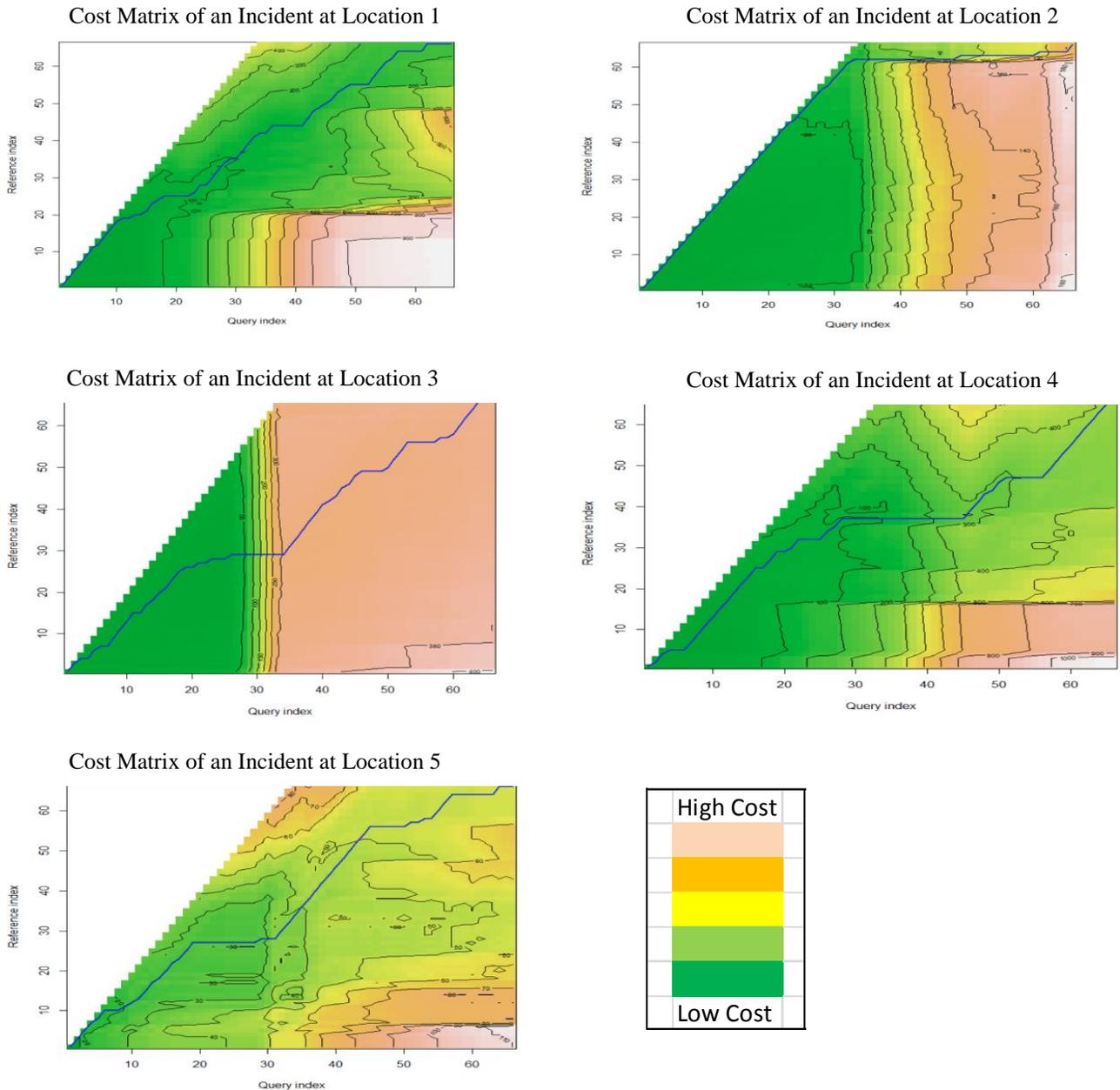


**Figure 8. DTW for non-incident.**

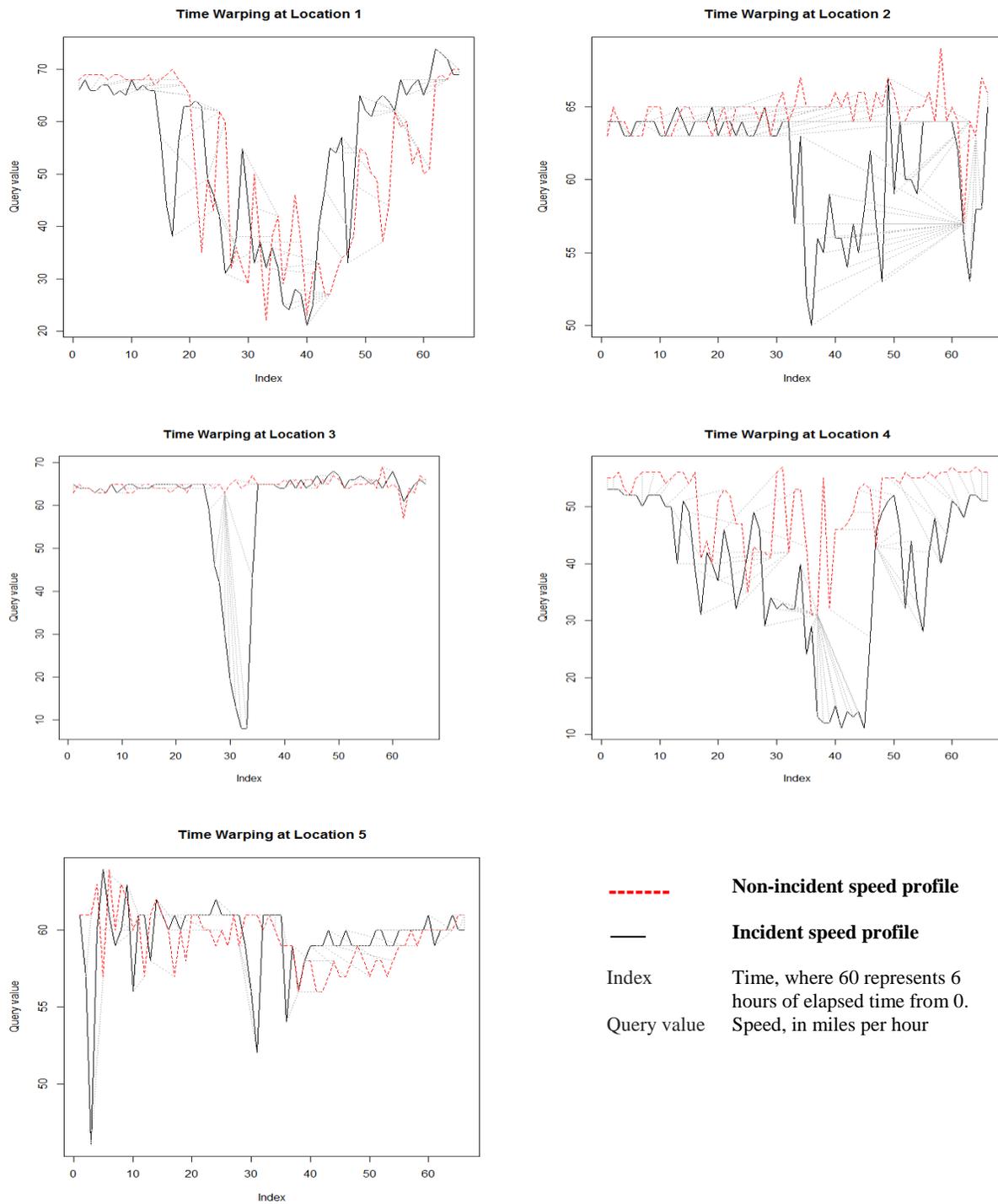


**Figure 9. Graphical view of DTW output for a non-incident.**

Next, we developed models for the same locations during an incident. Figures 10 shows the output of our DTW model using incident data. The blue line is the shortest path found by the model. Here, we see incident patterns like those shown in Figure 5. Deviation from the diagonal trend indicates higher cost to the user. During an incident, the two time series are not well matched, so the cost matrix shows more orange and yellow colors, which also indicate an incident.



**Figure 10. DTW cost Matrix for Incident.**

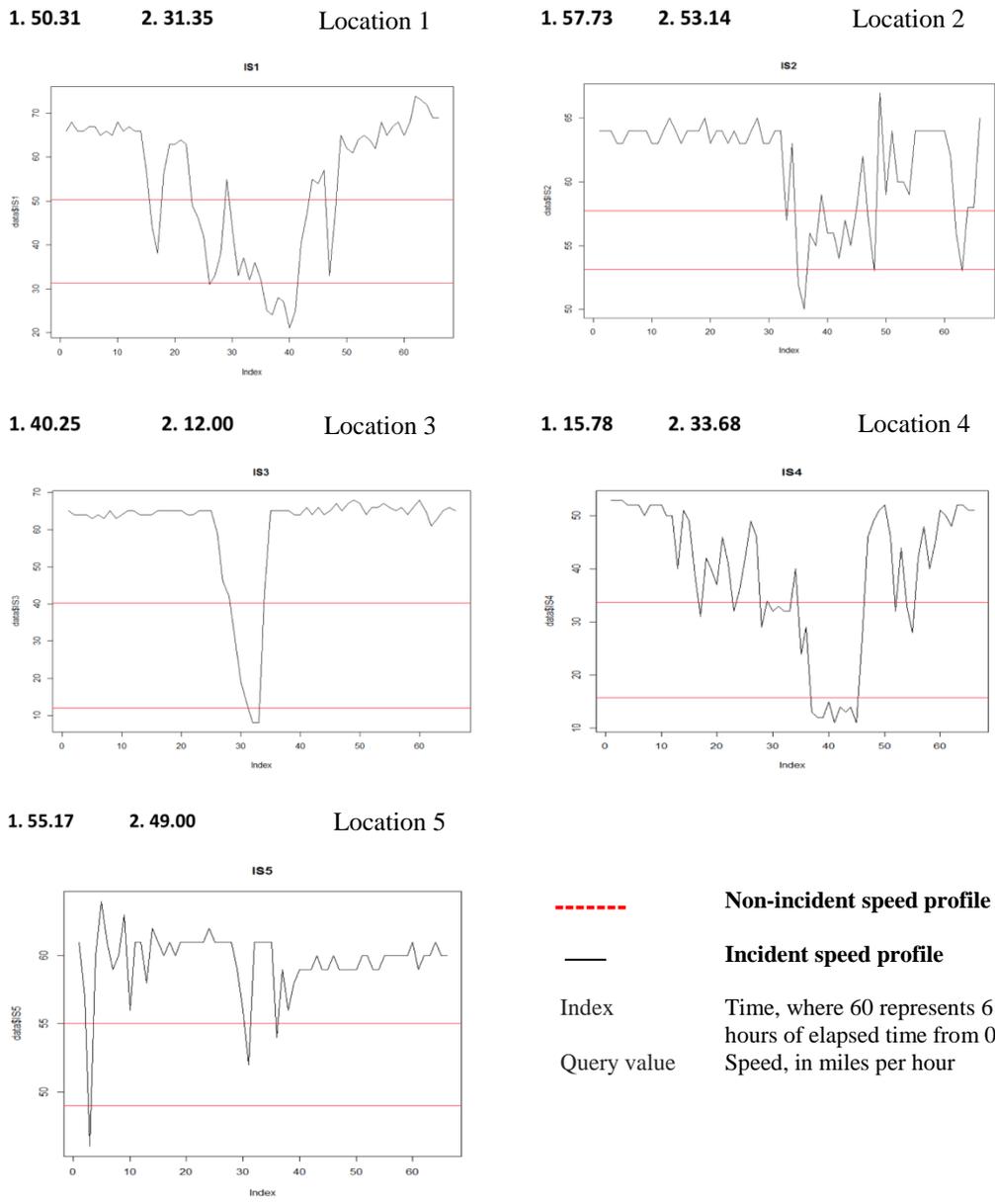


**Figure 11. Graphical view of DTW output for incident.**

Time warping does not necessarily show that an incident happened, but it does show patterns. Therefore applying a classifier is necessary. We applied a k-fold cross-validation

technique to obtain a reliable estimate of the classifier accuracy. Initially, the incident traffic dataset was used to evaluate the performance of the proposed algorithm.

Each test set is processed by DTW algorithm. In order to compare the similarity between normal and incident conditions, we used predefined incident patterns to classify the current pattern. Because running DTW takes a lot of memory, we used a threshold obtained from cross validation to alert the data collection program when to start scanning for an incident. The threshold was calculated from the k-mean and is represented by red lines in the graphs in Figure 12. In Figure 12, each chart has horizontal lines depicting threshold 1 and threshold 2. The values are shown above each figure with number 1 followed by the speed in miles per hour and numeral 2 identifying the second threshold followed by speed in miles per hour. For example, at location 1, threshold 1 is 50.31 mph and threshold 2 is 31.35. The higher speed threshold can be selected as the trigger to start collecting data. When the speed decreases until it meets the second threshold, the trigger starts scanning and collecting a data point backwards from current speed toward the free flow speed. It also starts collecting forward until it meets the second threshold. The second threshold is a trigger to stop collecting data and start the DTW algorithm.



**Figure 12. Using K-means thresholds to trigger DTW to search for incidents at five locations.**

Using dynamic thresholds based on historical traffic data, thereby accounting for typical variations of traffic throughout the day, can increase the accuracy of the algorithm. Therefore, this approach could recognize recurrent congestion and therefore reduce the incidence of false alarms. In the next step we will try a dynamic threshold to modify our model using Support Vector Machine (SVM).

## SVM Literature

The SVM algorithm first was proposed in 1993 by Cortes and Vapnik and published in 1995. SVM has had limited applications in the transportation field. Previous studies include use for travel time, traffic speed and traffic flow predictions, in the context of ITS applications (Bhavsar 2007, Ding 2002). SVM is a powerful, robust and computationally efficient tool in solving various transportation classification problems. Furthermore, SVM was successfully applied to detect highway incidents (Chen 2003, Yuan 2003). Chowdhury *et al.* (2006) and Bhavsar *et al.* (2007) used SVM for travel time prediction. They found it suitable for the hierarchical intelligence applications due to its low memory requirements and processing requirements. Xiao *et al.* (2012) and Cai *et al.* (2010) modified the standard SVM classifier to improve incident detection results. Nevertheless in their studies they either use simulation data or an I-880 database that is somewhat problematic since either the database is old or it suffers from unidentified actual incident data.

## SVM Concept

A classification task usually involves separating data into training and testing sets, as shown in Figure 13. Given a training set of instance-label pairs  $(x_i; y_i)$ ;  $i = 1, \dots, l$  where  $x_i$  is a training vector and  $y_i$  is our output which in this case is traffic state (-1 nonincident, 1 incident).

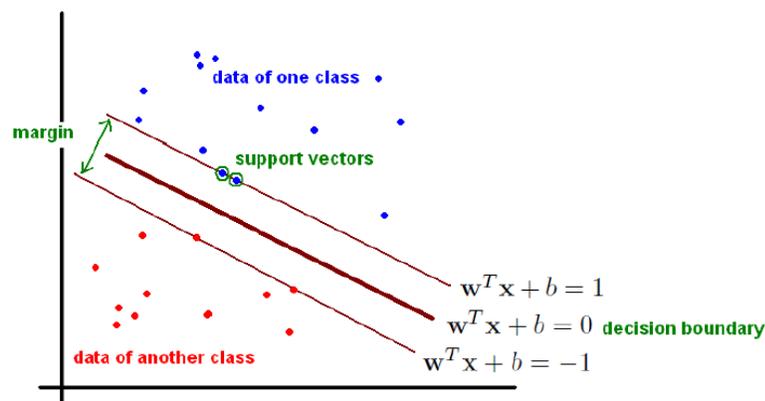


Figure 13. Hyperplane through two linearly separable classes.

According to Figure 13, implementing SVM simplifies training data to:

$$x_i \cdot w + b \geq +1 \text{ for } y_i = +1$$

$$x_i \cdot w + b \leq -1 \text{ for } y_i = -1$$

These equations can be combined into:

$$y_i(\mathbf{x}_i \cdot \mathbf{w} + b) - 1 \geq 0 \quad \forall_i$$

And  $d^+$  and  $d^-$  represent the shortest distance to the closest positive and negative point respectively. SVM requires the solution of the following optimization problem (Boser *et al.* 1992; Cortes and Vapnik, 1995):

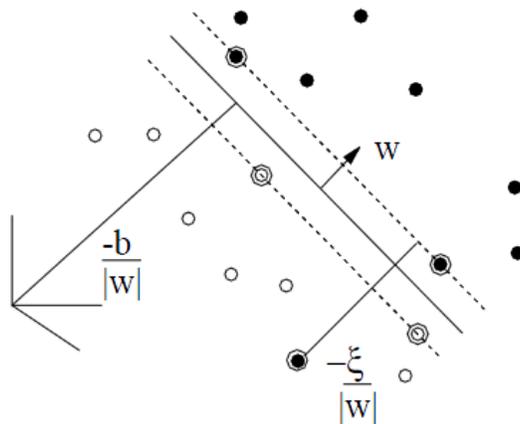
$$\min_{\mathbf{w}, b} \quad \frac{1}{2} \mathbf{w}^T \mathbf{w}$$

$$\text{subject to} \quad y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \quad i = 1, \dots, l$$

This model works well for separable data but gets more complicated when the data is not exactly separable, which commonly is the case of real world data. One way to make the model more realistic is introducing penalty ( $C$ ) and error ( $\xi$ ) terms. The objective of the prediction function can be achieved by solving the following optimization problem (Hsu *et al.* 2007):

$$\begin{aligned} \min_{\mathbf{w}, b, \xi} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^l \xi_i \\ \text{subject to} \quad & y_i(\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0. \end{aligned}$$

In Figure 14, training vectors  $\mathbf{x}_i$ 's are mapped into a higher dimensional space by the function  $\phi$ . SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space.  $C > 0$  is the penalty parameter of the error term. Furthermore,  $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$  is called the kernel function. The result of SVM is sensitive to the kernel function.



**Figure 14. Linear separating hyperplanes for the non-separable case.** Source:

[http://www.ce.rit.edu/research/projects/2004\\_winter/rt\\_objtrack/detect.htm](http://www.ce.rit.edu/research/projects/2004_winter/rt_objtrack/detect.htm)

There are four basic kernel functions: linear, polynomial, radial basis function (RBF), and sigmoid. In general, the RBF kernel performs well in many scenarios (Vanschoenwinkel and Manderick 2006). Because this kernel nonlinearly maps samples into a higher dimensional space, it can better find patterns when the relation between class labels and attributes is nonlinear. Furthermore, the linear kernel is a special case of RBF (Keerthi and Lin 2003) since the linear kernel with a penalty parameter  $C$  has the same performance as the RBF kernel with parameters  $(C, Y)$ . In addition, the sigmoid kernel behaves like RBF for certain parameters (Lin and Lin 2003). The second reason is the number of hyperparameters influences the complexity of model selection and the polynomial kernel has more hyperparameters than the RBF kernel. Finally, the RBF kernel presents fewer numerical difficulties.

### **Modifications and constraints in SVM**

Scaling before applying SVM is very important (Sarle 1997). The main advantage of attribute scaling is to represent all attributes with comparable number ranges. Another advantage is to avoid numerical difficulties during the calculation because kernel values usually depend on the inner products of feature vectors. The range  $[-1,1]$  or  $[0,1]$  is recommended. The same scaling method must be applied to both training and testing data.

$C$  and  $Y$  are unknown parameters that must be identified for each specific problem. Consequently some kind of model selection (parameter search) must be done. The goal is to identify best fitting  $(C, Y)$  so that the classifier can accurately predict unknown data (i.e., testing data). Note that it may not be useful to achieve high training accuracy. The prediction accuracy obtained from the “unknown” data set more precisely reflects the prediction accuracy. An improved version of this procedure is known as cross-validation. The cross-validation procedure can prevent the over fitting problem.

### **SVM Case Generation and Methodology**

This section presents a case study for application of SVM for incident detection. The same database used for the DTW model has been used for this section. The first step is generating the cases required for developing and evaluating the SVM incident detection model. The idea of incident detection is based on the concept that when an incident happens, the kinetics of passing vehicles would be affected: the speed drops upstream and increases downstream, lane changing

increases, and involved vehicles demonstrate large acceleration and deceleration rates. This study identified the speed profile and volume over a selected time step  $t_s$  (5 min for Dallas data) to recognize the patterns that indicate the incident occurrence. An array of five values for each time slice has been chosen as the input file for the model (Table 1).

**Table 1. SVM sample input data and result.**

Time Instant	Kinetics					Decision
	Down Stream real-time Speed (DSpeed)	Up Stream real-time Speed (USpeed)	Up Stream typical Speed (UTSpeed)	Down Stream real-time volume (DVolume)	Up Stream real-time volume (UVolume)	
$t$	55	49	55	118	145	-1
$t + t_s$	67	14	53	112	91	+1

The decision variable  $y_i$  can only have values of +1 representing an incident, or -1 representing a non-incident condition. During peak hour flow conditions, when vehicles pass the incident, their speed downstream of the incident will be increased significantly. In this study we followed classical SVM (Boser *et al.* 1992; Cortes and Vapnik 1995) for two-class classifications. To be consistent with the DTW model the time step is 5 minutes. Here we used 244 instances with 36 incident cases.

**Table 2. Input data counts.**

----- Training Data -----		
-----Actual-----		
Category	Count	
Non Incident	-1	208
Incident	1	36
Total	244	

The objective of training is to find the prediction function:

$$f(x_i) = w * x_i + b$$

This objective function optimizes the minimum distance between the classification hyper-plane and any sample of training data. Considering the complexity of traffic behavior, non-separable data must be allowed for training. As mentioned earlier scaling is important for the success of AI models such as ANN and SVM. (Hsu *et al.* 2007; Sarle 2007) Before training, all the data were linearly scaled to a range of [0, 1].

Here we used v-fold cross-validation to maximize the use of training data and search for

optimal parameters (C,  $\gamma$ ). First, data has been divided into  $v$  subsets of equal size. Sequentially one subset is tested using the classifier trained on the remaining ( $v - 1$ ) subsets. Thus, each instance of the whole training set is predicted once so the cross-validation accuracy is the percentage of data that are correctly classified. Different numbers of  $v$ -folds have been tried to find the sensitivity of the data to the number of folds, with 5 and 6 groups producing the best results.

The optimal parameters were identified through grid searching of many combinations in the range of  $[C, \gamma] = [2^{-5} : 2^2 : 2^6, 2^{-15} : 2^2 : 2^4]$ . The experiment was performed by increasing parameters in exponential order, i.e.  $2^n$ , in the range of -5 to 5 for C and -15 to 3 for  $\gamma$  within two steps. The identified optimal parameters were then used to train the entire training set to generate a trained SVM algorithm.

This study used LIBSVM (Chang and Lin 2007), an open source implementation routine for SVM to train and test the SVM model. The training time of the SVM model was less than five seconds in all the training cycles. The prediction time was quite short as well, which is a vital element for real-time applications.

## Results

Different kernel functions were tried to find the best fit model. As expected, the RBF kernel function was the fastest and had higher accuracy. The base model includes all variables introduced in the previous section. Table 3 shows a summary of input variables.

**Table 3. Input variables summary.**

===== Summary of Variables =====					
Number	Variable	Class	Type	Missing rows	Categories
1	class	Target	Categorical	0	2
2	Dspeed	Predictor	Continuous	0	
3	Uspeed	Predictor	Continuous	0	
4	UTspeed	Predictor	Continuous	0	
5	Dvolume	Predictor	Continuous	0	
6	Uvolume	Predictor	Continuous	0	

The Radial Basis Function (RBF) has been used as the SVM kernel function. The optimum values found for parameters using a grid search to minimize total error in the objective function are  $(C, \gamma) = (0.3125, 8)$ .

The overall accuracy of training and validation is represented in Table 4. The results suggest the robustness of the model to predict incidents during peak hours.

**Table 4. Base model output.**

```
===== Misclassification Tables =====
--- Training Data ---
-----Actual-----Misclassified-----
Category      Count      Count      Percent      Cost
-----
Non-Incident -1      208        1           0.481      0.005
Incident      1       36         3           8.333      0.083
-----
Total         244        4           1.639      0.016

overall accuracy = 98.36%
--- Validation Data ---
-----Actual-----Misclassified-----
Category      Count      Count      Percent      Cost
-----
Non-Incident -1      208        2           0.962      0.010
Incident      1       36         3           8.333      0.083
-----
Total         244        5           2.049      0.020

overall accuracy = 97.95%
=====
```

As the results show, the false alarm rate is pretty low (around 2%). The following table shows that we had only one case in which an incident happened and the model was not able to detect it and 3 cases in which non-incidents were falsely detected as incidents.

**Table 5. Base model false detection details.**

```
===== Confusion Matrix =====
----- Training Data -----
Actual : ----Predicted Category----
Category:  -1      1
-----
-1:        207      1
1:         3       33

----- Validation Data -----
Actual : ----Predicted Category----
Category:  -1      1
-----
-1:        206      2
1:         3       33
=====
```

In the next step different scenarios with different variables have been evaluated as shown in the Table 6.

**Table 6. Different scenarios comparison.**

Scenario	Model	Variable	Model Accuracy	False Alarm
1	Base-RBF	<ul style="list-style-type: none"> <li>• Down Stream real-time Speed</li> <li>• Up Stream real-time Speed</li> <li>• Up Stream typical Speed</li> <li>• Down Stream real-time volume</li> <li>• Up Stream real-time volume</li> </ul>	<ul style="list-style-type: none"> <li>• Training : 98.36%</li> <li>• Validation: 97.95%</li> </ul>	<ul style="list-style-type: none"> <li>• Training: 4 out of 244</li> <li>• Validation: 5 out of 244</li> </ul>
2	Sigmoid	<ul style="list-style-type: none"> <li>• Down Stream real-time Speed</li> <li>• Up Stream real-time Speed</li> <li>• Up Stream typical Speed</li> <li>• Down Stream real-time volume</li> <li>• Up Stream real-time volume</li> </ul>	<ul style="list-style-type: none"> <li>• Training : 97.54%</li> <li>• Validation: 97.13%</li> </ul>	<ul style="list-style-type: none"> <li>• Training: 6 out of 244</li> <li>• Validation: 7 out of 244</li> </ul>
3	RBF	<ul style="list-style-type: none"> <li>• Down Stream real-time Speed</li> <li>• Up Stream real-time Speed</li> <li>• Up Stream typical Speed</li> </ul>	<ul style="list-style-type: none"> <li>• Training : 97.54%</li> <li>• Validation: 96.72%</li> </ul>	<ul style="list-style-type: none"> <li>• Training: 6 out of 244</li> <li>• Validation: 8 out of 244</li> </ul>
4	RBF	<ul style="list-style-type: none"> <li>• Down Stream real-time Speed</li> <li>• Up Stream real-time Speed</li> </ul>	<ul style="list-style-type: none"> <li>• Training : 97.13%</li> <li>• Validation: 97.13%</li> </ul>	<ul style="list-style-type: none"> <li>• Training: 7 out of 244</li> <li>• Validation: 7out of 244</li> </ul>

The prediction accuracy of developed models was compared. First two different kernel functions were compared while all other variables were kept the same. The results show that applying the Base-RBF kernel function is not only faster it is more accurate also. Next the sensitivity of the model to chosen variables was evaluated. Interestingly we observed that using typical speed (non-incident speed at same location using historical data) was a robust choice. However the results are based upon comparing cases with the same geometry, congestion level, and data sources. Since all these factors could have significant impacts on the results, we might want to test this model on cases in which these conditions are varied.

## **CONCLUSIONS**

In this research two groups of experiments were performed to evaluate two incident detection algorithms: DTW and SVM. Evaluation of the DTW and SVM algorithms revealed that both can successfully classify traffic conditions into two categories – incident, non-incident – during peak hours. Since both models produce very fast responses, both are candidates for real-time incident detection applications. Models were trained on a network based on freeway segments in Dallas, TX. Comparing these two methods, application of DTW in the field of transportation is quite new and the concept of DTW is simpler than SVM. The advantage of using SVM is that it does not require a large dataset to train and validate the model. On the other hand, the accuracy of SVM is highly dependent on the kernel function chosen and the value(s) of parameters which are best estimated through the solution of an optimization problem (in this research grid search was applied to find C and  $\gamma$ ). We recommend a more comprehensive comparison of these two using field data for model calibration and simulation to produce a wide range of condition specific data sets for model testing.



## REFERENCES

- Aach J, Church GM. "Aligning Gene Expression Time Series with Time Warping Algorithms." *Bioinformatics*, 17(6), pp. 495-508, 2001.
- Bhavsar, P., Chowdhury, M. A., Sadek, A., Sarasua, W., and Ogle, J. (2007). "Decision support system for predicting traffic diversion impacts across transportation networks using support vector regression." *Transportation Research Board Annual Meeting (CD-ROM)*, Washington D.C., 2007.
- Boser B. E., Guyon I., and Vapnik V. "A training algorithm for optimal margin classifiers." In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pages 144-152. ACM Press, 1992.
- CHEN R.L., SRINIVASAN D., TIAN E.: 'Support vector machine models for freeway incident detection', *IEEE Proc. Intell. Transp. Syst.*, 1, pp. 238–243, 2003.
- Chang, C-C., and Lin, C-J. "LIBSVM: a library for support vector machines." 2007. <<http://www.csie.ntu.edu.tw/~cjlin/libsvm>>, (December 2013)
- Chowdhury, M., A. Sadek, Y. Ma, N. Kanhere and P. Bhavsar. "Applications of artificial intelligence paradigms to decision support to real-time traffic management." *Transportation Research Record*, Transportation Research Board, Washington D.C., pp. 92-98, 1968.
- Corinna Cortes, Vladimir Vapnik, "Support-Vector Networks", *Machine Learning*, 20, pp.273-297, 1995.
- Cortes, C., V. Vapnik. "Support vector networks." *Machine Learning*, 20:273 – 297, 1995.
- Ding A., Zhao X., Jiao L.: 'Traffic flow time series prediction based on statistics learning theory'. *Proc. IEEE Fifth Int. Conf. on Intelligent Transportation System*, pp. 727–730, 2002.
- Faundez-Zanuy M. "On-Line Signature Recognition Based on VQ-DTW." *Pattern Recognition*, 40(3), pp. 981-992, 2007.
- FHWA: 'VII architecture and functional requirements version 1.1'. ITS Joint Program Office, US DOT, 2005.
- Gollmer K, Posten C. "Supervision of Bioprocesses Using a Dynamic Time Warping Algorithm." *Control Engineering Practice*, 4(9), pp. 1287-1295, 1996.
- Hermans F, Tsiorkova E. "Merging Microarray Cell Synchronization Experiments Through

- Curve Alignment." *Bioinformatics*, 23(2), pp. 64-70, 2007.
- Hi-ri-o-tappa, K., et.al, "A Novel Approach of Dynamic Time Warping for Short-Term Traffic Congestion Prediction," *Transportation Research Board of the National Academies*, Washington, D.C., 2011.
- Hi-ri-o-tappa, K. ; NECTEC, Nat. Sci. & Technol. Dev. Agency, Bangkok, Thailand ; Likitkhajorn, C. ; Poolsawat, A. ; Thajchayapong, S., "Traffic incident detection system using series of point detectors", *Intelligent Transportation Systems (ITSC)*, 15th International IEEE Conference, pp. 182-187, 2012.
- Huang B, Kinsner W. "ECG Frame Classification Using Dynamic Time Warping." In W Kinsner, A Sebak, K Ferens (eds.), *Proceedings of the Canadian Conference on Electrical and Computer Engineering- IEEE CCECE 2002*, volume 2, pp. 1105-1110, 2002.
- Hsu, C-W., Chang, C-C., and Lin, C-J. "A Practical Guide to Support Vector Classification." (2007) <<http://www.csie.ntu.edu.tw/~cjlin/guide.pdf>>.
- ITS America: 'Primer on VII', <http://www.itsa.org/itsa/files/pdf/VIIPrimer.pdf>, accessed 11 July 2007.
- Kartikeyan B, Sarkar A. "Shape Description by Time Series." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(9), 977-984, 1989.
- Keerthi, S. S. and C.-J. Lin. "Asymptotic behaviors of support vector machines with Gaussian kernel." *Neural Computation* 15 (7), pp. 1667-1689, 2003.
- Lin, K.-M. and C.-J. Lin. "A study on reduced support vector machines." *IEEE Transactions on Neural Networks* 2003.
- MARTIN P.T., PERRIN J., HANSEN B.: 'Incident detection algorithm evaluation', Prepared for Utah Department of Transportation, March 2001.
- Ma Y. "A Real-Time Traffic Condition Assessment and Prediction Framework Using Vehicle-Infrastructure Integration (Vii) with Computational Intelligence." Ph.D. Dissertation, Clemson University, Clemson, SC, USA. Advisor(s) Mashrur A. Chowdhury, 2008.
- Ma, Y, M. Chowdhury, M. Jehani, R. Fries, "Accelerated incident detection across transportation networks using vehicle kinetics and support vector machine in cooperation with infrastructure agents" *IET Intelligent Transport Systems*, Vol. 4, pp. 328-337, 2010.
- Mussa R. N. and J. E. Upchurch, "Modeling incident detection using vehicle-to-roadside communication system." *Journal of the Transportation Research Forum*, vol. 39, no. 4,

- pp. 117-127, 2000.
- Mussa R. N. and J. E. Upchurch, "Simulation assessment of incident detection by cellular phone call-in programs." *Transportation*, vol. 26, no. 4, pp. 399-416, 1999.
- Mussa R. N., "Evaluation of driver-based freeway incident detection," *Journal of ITE*, vol. 67, no. 3, pp. 33-40, 1997.
- Oh, J.-S., C. Oh, S. G. Ritchie, and M. Chang, "Real-Time Estimation of Accident Likelihood for Safety Enhancement," *ASCE Journal of Transportation Engineering*, pp. 358–363, 2005.
- PATH: 'VII California bay area test bed development plan', 4 April 2006.
- Rath TM, Manmatha R. "Word Image Matching Using Dynamic Time Warping." In R Manmatha (ed.), *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pp. II-521-II-527, 2003.
- Ruey Long Cheu, Srinivasan, D. ; Eng Tian Teh, "Support vector machine models for freeway incident detection." *Intelligent Transportation Systems, Proceedings. 2003 IEEE (Volume:1 )*, 2003, pp. 238-243.
- Sarle W. S., "Neural Network FAQ," 1997. Periodic posting to the Usenet newsgroup comp.ai.neural-nets. <ftp://ftp.sas.com/pub/neural/FAQ.html> (December 2013)
- Skabardonis A., T. C. Chavala, and D. Rydzewski, "The I-880 field experiment: effectiveness of incident detection using cellular phones," *Institute of Transportation Studies, University of California at Berkeley, CA, PATH Report UCB-ITS-PRR-98-1*, 1998.
- Syeda-Mahmood T, Beymer D, Wang F. "Shape-Based Matching of ECG Recordings." In A Dittmar, J Clark, E McAdams, N Lovell (eds.), *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 2012-2018, 2007.
- TAN Z., LU X., "A Combination Algorithm of Freeway Traffic Automatic Incident Detection," *ASCE First International Conference on Transportation Information and Safety (ICTIS), Volume I: Highway Transportation*, pp. 1106-1112, 2011.
- Tak YS. "A Leaf Image Retrieval Scheme Based on Partial Dynamic Time Warping and Two-Level Filtering." In D Wei, T Miyazaki, I Paik (eds.), *Proceedings of the 7th IEEE International Conference on Computer and Information Technology - CIT 2007*, pp. 633-638. IEEE Computer Society, Los Alamitos, CA, USA 2007.

- Vanschoenwinkel, B., and Manderick, B. "Context-sensitive Kernel Functions: A Comparison between Different Context Weights." *Lecture Notes in Computer Science*, 3930, 861-870, 2006.
- Walters C. H., P. B. Wiles, and S. A. Cooner. "Incident detection primarily by cellular phones—an evaluation of a system for Dallas, Texas," in *Proc. 78th Annual Meeting of the Transportation Research Board*, Washington D.C., Jan. 1999.
- Wei L, Keogh E, Xi X. "SAXually Explicit Images: Finding Unusual Shapes." In CW Clifton, N Zhong, J Liu, BW Wah, X Wu (eds.), *Sixth International Conference on Data Mining 2006-ICDM '06*, pp. 711-720. IEEE Computer Society, Los Alamitos, CA, USA 2006.
- Xiao J., Liu Y., "Traffic Incident Detection Using Multiple-Kernel Support Vector Machine," *Transportation Research Record: Journal of the Transportation Research Board*, Volume 2324, , pp. 44 – 52, 2012.
- Xie C. and E. Parkany, "Use of driver-based data for incident detection," in *Proc. 7th International Conference on Application of Advanced Technologies in Transportation Engineering*, Cambridge, MA, pp. 143-150, 2002.