Anomaly Detection using Microscopic Traffic Variables on Freeway Segments

Suttipong Thajchayapong
st@imperial.ac.uk

Javier A Barria
j.barria@imperial.ac.uk

Intelligent Systems and Networks Research Group
Department of Electrical and Electronics Engineering
Imperial College London
London SW7 2AZ, United Kingdom
Phone: +44(0)2075946275
Fax: +44(0)2075946274

5247 Words + 6 Figures + 2 Tables

Transportation Research Board
89th Annual Meeting
January 10-14, 2010
Washington, D.C.
ABSTRACT

This paper proposes and assesses the effectiveness of monitoring vehicular traffic anomalies using microscopic traffic variables, namely relative speed and inter-vehicle spacing. We present an algorithm that detects transient changes in traffic patterns using microscopic traffic variables. In particular, we show that when applied to real-world scenarios, our algorithm can use the variance of statistics of relative speed to detect traffic anomalies and precursors to non-recurring traffic congestion. The performance of the proposed algorithm is also assessed using a microscopic traffic simulation environment, where we show that with minimum prior knowledge, the proposed algorithm has comparable performance to an ideally placed loop detector monitoring the standard deviation of speed. The algorithm also performs very well even when the microscopic traffic variables are available only from a fraction of the complete population of vehicles.

Keywords: Traffic Monitoring, Anomaly Detection, Incident Precursors, Microscopic Traffic Variables
INTRODUCTION

It is well known that road traffic incidents are still the cause of billions of dollars in extra hours of travel and extra fuel (1). A large number of automatic incident detection algorithms has emerged to address the problems of traffic incident detection while the majority of them focus on detection after a major disruption of traffic has occurred (2, 3, 4). Since traffic congestion that may have a significant impact on travel time is normally related to traffic incidents, it is important to proactively assess the occurrence of anomalous traffic patterns, as an early warning incident precursor signal, that might prevent or minimize the impact and duration of a traffic incident.

The first step to proactively assess the occurrence of traffic incidents is to detect a deviation from normal traffic patterns, which we refer to as traffic anomaly in this study (5, 6). It is important to note that in this paper, we are primarily interested in transient anomaly which is an early stage of deviation of traffic patterns, e.g. drops in individual speeds caused by a distraction on a freeway shoulder. This type of anomalies usually receive less attention in literature as there is uncertainty if they will lead to traffic incidents. Nevertheless, detection of transient anomalies still is challenging and important as they could be the sign of traffic developing into major traffic congestion. Therefore, it is highly relevant to traffic control centres to automate the process of early warning detection of traffic anomalies and infer the likelihood of evolving into a traffic incident.

The characteristics of traffic anomalies that lead to traffic incidents, commonly known as incident precursors, have already been thoroughly studied based on macroscopic traffic variables derived from road-side infrastructure, e.g. loop detectors (5, 6, 7). The majority of these studies have shown that the variation of speed is often associated with the deviation of traffic patterns and hence a signal of a probable anomalous condition. However, the effectiveness of those detection algorithms largely depends on the relative location of the anomaly in respect to the loop detectors. If a disruption takes place far away from the loop detector location, the anomaly may not be detected and/or a long delay may be present before the anomaly is eventually identified.

In this study, we advocate that non-recurring events are expected to have early impact on behaviour of individual vehicles causing early deviations of the measured microscopic traffic variables, namely relative speed and inter-vehicle spacing. Even though relative speed and inter-vehicle spacing have been used for analyzing road traffic characteristics by some studies (8, 9, 10, 11), they have not yet been an anomaly detection algorithm that can explicitly utilising the variation of these microscopic traffic variables under dynamic availability of individual vehicle information. The study in (8) is among the first to employ these microscopic traffic variables to analyze road traffic characteristics, where the proposed model of reliability of freeway traffic flow is derived using relative speed and inter-vehicle spacing. However, the model itself was derived to be used with macroscopic traffic variables, e.g. flow and density, and author did not further propose an algorithm for anomaly detection.

Recent emergence of vehicle-to-vehicle and vehicle-to-infrastructure communications have increased the possibility of incorporate microscopic traffic variables for road traffic monitoring (12, 13). Inter-vehicle spacing has been used to analyze queueing behaviour at a signalised intersection (10) and also to determine the position of each vehicle for traffic information aggregation and dissemination purposes (9). In (11), each vehicle uses the distance between itself and the vehicle in front to determine whether it is in a queue or near an incident location. However, the purpose of using inter-vehicle spacing in (11) is primarily for adjustment of vehicle mobility and smoothing traffic flow. Inter-vehicle spacing is not directly used as an indicator of traffic anomalies as the
authors assume that the information is provided and hence, did not investigate how the occurrence of anomalies can be identified correctly with inter-vehicle spacing. In contrast, in this paper, we assess relative speed and inter-vehicle spacing when used for anomaly detection.

In this paper, we propose an algorithm that can detect traffic anomalies by observing the change of variance of microscopic traffic variable, which can be measured by equipped vehicles sharing information with one another. Furthermore, the performance of the proposed algorithm is assessed under different availability of individual vehicles information.

The paper is organized as follows. The second section describes the framework of the analysis. The proposed detection methodology is presented in the third section. Then, in the forth section, the preliminary analysis of relative speed and inter-vehicle spacing in the context of anomaly detection is discussed through simulation. In the fifth section, we demonstrate through real-world data that traffic anomalies detected by the change of variances of relative speeds are associated with non-recurring events. Finally, the last section concludes this paper.

ANALYSIS FRAMEWORK

Our framework is based on a distributed traffic monitoring system that could rely on locally shared information amongst neighbours vehicles. The shared information will allow the calculation of microscopic traffic variables, to assess road traffic condition on a freeway segment. The relevant information would be measured by vehicles sharing information through automotive navigation systems and wireless communications. It is assumed that each vehicle is equipped with a capability to measure its instantaneous speed and position through automotive navigation systems (14, 15). Furthermore, vehicles are capable of sharing information with one another using wireless communications (16) through their neighboring vehicles (11, 17), bus ferries (18) or road-side infrastructure (11). Alternatively, when such information measuring and sharing capability does not exist in individual vehicles, the microscopic traffic information could instead be inferred form currently available road-side infrastructure, e.g. video surveillance cameras (19).

The statistics of microscopic traffic variables are calculated from \( I = P_{OA} \times I_{total} \) vehicles, where \( P_{OA} \) (Percentage of Availability) denotes the percentage of vehicles whose speed and position information can be measured on the road segment of interest, and \( I_{total} \) is the total number of vehicles on the segment.

DETECTION ALGORITHM

Our proposed detection algorithm is based on an observation that a disruption causes transient behaviors of individual vehicles (e.g. deceleration and lane change) that affects the variability in microscopic traffic variables. Therefore, we formulate the problem of detecting anomalies using microscopic traffic variables as a variance change point detection in which we adopt a methodology based on Bayes’ theorem and sliding windows (20). Let \( y_n \) denote the statistics of a microscopic traffic variable of interest at time \( n, n = 1, 2, \ldots, N \). We model \( y_n \) as Gaussian process \( N(\mu_n, \sigma^2_n) \), where \( \sigma^2_n \) denotes the changing variance of \( y_n \). In the proposed algorithm, the sliding window size \( L \) determines the number of temporal samples of \( y_n \) to be used for change detection. The algorithm compares the variances of \( y_n \) in two adjacent sliding windows \( W1 \) and \( W2 \), where the comparison is initiated only if there are at least \( L \) samples of \( y_n \) in both windows, i.e. \( W1 \geq L \) and \( W2 = L \).

Let \( n_0 = N - L \) denote the time where the variance of \( y_n \) changes. We compare the null
hypothesis \( \{ H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_{n_0-1}^2 = \sigma_{n_0}^2 = \ldots = \sigma_N^2 \} \) against the alternative hypothesis \( \{ H_1 : \sigma_1^2 = \ldots = \sigma_{n_0-1}^2 \neq \sigma_{n_0}^2 = \ldots = \sigma_N^2 \} \). Now, let \( \Theta_N^q \) be a vector of variances associated with hypothesis \( H_q \), i.e. \( \Theta_N^q = \{ \sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2 \} \). Using Bayes’ theorem, the probability of the change of variance at point \( n_0 \) is estimated by the posterior probabilities associated with the hypotheses:

\[
p (H_0 | y_n) = \frac{p(y_n | \Theta_N^0) p(\Theta_N^0)}{p(y_n)} \quad \text{and} \quad p (H_1 | y_n) = \frac{p(y_n | \Theta_N^1) p(\Theta_N^1)}{p(y_n)}
\]

where \( p(y_n | \Theta_N^q) \) denotes the likelihood function and \( p(\Theta_N^q) \) denotes the prior probability. The change point detection is then performed by comparing \( p (H_0 | y_n) \) and \( p (H_1 | y_n) \) and the alarm is raised when \( \frac{\log p(H_0 | y_n)}{\log p(H_1 | y_n)} > 1 \).

**PERFORMANCE EVALUATION USING SIMULATION**

Performance Evaluation Parameters

Let \( n_{q,i} \) be the time that the \( i^{th} \) alarm is raised for change point \( n_q \). The alarm is considered a true alarm if \( n_{q,i} \in [n_q, n_q + n_b] \) where \( n_b \) denote the detection bound, otherwise it is considered a false alarm. Given \( M \) experiments for each change point, the performance evaluation parameters we consider are Detection Rate of Change Point \( q \) (\( DR_q \)), Mean Time to Detection of Change Point \( q \) (\( MTTD_q \)) and False Alarm Rate (\( FAR \)), which are calculated as shown in equations (1), (2) and (3) respectively. In this paper, \( FAR \) is calculated by collectively taking into account any alarm that is neither in \([n_1, n_1 + n_b]\) nor in \([n_2, n_2 + n_b]\). Therefore, it is possible to have \( DR_q = 0 \) while \( FAR > 0 \).

\[
DR_q = \frac{\text{Number of Anomaly Detected that lie with in } [n_q, n_q + n_b]}{M}. \quad (1)
\]

\[
MTTD_q = \frac{\sum_{i=1}^{M} (n_{q,i} - n_q)}{M}, \quad n_{q,i} \in [n_q, n_q + n_b]. \quad (2)
\]

\[
FAR = \frac{\text{Number of Anomaly Detected that are not in } [n_q, n_q + n_b]}{\text{Total number of Detections}}. \quad (3)
\]

**Performance Evaluation Results**

*Benchmark Anomaly Detection Algorithm (5)*

As benchmark, we employ the traffic anomaly detection methodology proposed in (5) which uses the standard deviation of speed measured from loop detectors. This benchmark detection methodology examines if the probability that there is an anomaly on the measured standard deviation of speed \( X \), \( P(A|X) \), exceeds a threshold \( \delta \). \( P(A|X) \) is calculated as:
\[ p(A|X = x) = \frac{P_A f_A^A(x)}{P_A f_A^A(x) + P_N f_N^N(x)}, \]

where

\[
P_A = \frac{\text{Number of h-seconds standard deviations of speed associated with anomaly}}{\text{Total number of h-second standard deviations of speed}}\]

and \( P_N = 1 - P_A \). Also, \( f_A^A(x) \) is the empirical non-parametric probability density function of standard deviations of speed associated with traffic anomaly, and \( f_N^N(x) \) is the empirical non-parametric probability density function of standard deviations of speed associated with normal traffic.

Based on ten simulated realizations consisting of approximately 11,000 data points, we obtain \( f_A^A(x) \) and \( f_N^N(x) \) from a non-parametric estimation of empirical probability density functions of standard deviations of speed under normal and anomalous conditions. The probability density functions are estimated using kernel density estimator with Epanechnikov kernel as in (5). As it is shown in (5) that the selection of threshold is dependent on the data, we also consider a range of threshold and select the most appropriate one that reduces many false alarms based on our own data.

Since a well-known inherent problem of loop detectors is that if a disruption takes place far from the loop detector positions, the traffic anomalies are likely to be missed, we ideally place the loop detector stations at the position where the disruption is originated. In addition, we calculate the standard deviations of speeds every \( h \) seconds where \( h \) is always chosen to be ideally finer than the commonly used five-minute interval (5), and consequently should give better performance than the scenarios in (5) and any real-world deployment of loop detectors.

**Experimental Setups**

To model vehicle mobility more realistically, we have implemented Gipps safe-distance car following model with the parameters that were calibrated on a real freeway segment as in (21) into the microscopic traffic simulation environment Groovenet (22). On the two-lane freeway segment, each simulated vehicle recorded its timestamp, speed, position and bearing at every second. We consider a low vehicle density scenario where \( I_{\text{total}} \) is 16 vehicles/mile/lane on average (approximately 10% of the segment area) and hence, it would be more difficult to detect anomalies using only stationary loop detectors. Furthermore, we avoid the scenario of having congestion due to the exceeding of the freeway segment capacity by setting an average speed of 70mph and the volume to always be well below a breakdown threshold (23). This ensures that any change is caused by the simulated disruption only. We focus on the detection of short-term transient changes (\( \leq 5 \) minutes) as such kind of anomalies require methodology with high sensitivity making them usually difficult to detect and keep low false alarm rate at the same time. Nevertheless, it is important to identify this type of anomalies as they can further lead to traffic incidents.

Disruptions of traffic flow on a specific location on the freeway are generated by having a designated vehicle decelerated to a low speed of 10 mph between \( n_1 = 690s \) and \( n_2 = 750s \), and then accelerated back to normal speed of 70 mph. Such anomalous behavior causes the average traffic speed to slightly drop to approximately 50-60 mph and hence, causing a temporarily disruption to the traffic flow. Note that the disruption period we used is much shorter than a common minimum interval of five minutes used for analyzing traffic characteristics prior to incidents, and hence will be difficult to detect by currently deployed algorithms based on measurements from
FIGURE 1 Detections of Anomalies on Single Simulated Realizations of Statistics of Relative Speed for PoA 40%, L=30s; Dotted vertical lines denote the alarm times using the proposed algorithm.

The microscopic traffic variables preliminarily investigated are 1) inter-vehicle spacing: \( s_{i,n} = x_{i-1,n} - x_{i,n} \) where \( x_{i,n} \) is the position of a vehicle \( i \) at time \( n \), and 2) relative speed: \( v_{i,n} = w_{i-1,n} - w_{i,n} \) where \( w_{i,n} \) is the speed of a vehicle \( i \) at time \( n \). The statistics of microscopic traffic variables we consider for traffic anomaly detection are the variances of the sample averages and standard deviations of inter-vehicle spacing and relative speed. Furthermore, we select flat prior, \( p(\Theta^N_0) = 1 \) (24), to reflect a real-time application with minimum prior knowledge where performance depends largely on the analyzed microscopic traffic variables. In addition, for a given PoA, the vehicles that possess sensor and communication capability are randomly chosen from the vehicles on the freeway segment according to a uniform distribution.

Results and Discussions

Figure 1 shows the statistics of relative speed of a single simulated realization. The proposed algorithm utilizes the fact that the short-time transient anomalies disruption causes changes in individual speeds of the vehicles and exacerbates variation of the relative speeds. Note that such changes will be difficult to detect with the variation of speed calculated from loop detectors at a specific location since the anomaly is caused by a moving vehicle and the variations of locally measured speed will not be persistence once the vehicle moves pass the loop detector location.

Table 1 shows performance evaluation results from applying the proposed algorithm and the
TABLE 1 Transient Anomaly: Performance Comparisons Results for $L = 30\, s$ with Detection Bound $n_b = 60\, s$, from 10 simulated realizations (AVG = Average, STD = Standard Deviation)

<table>
<thead>
<tr>
<th>Proposed Algorithm (PoA = 50%)</th>
<th>$DR_{n_1}$</th>
<th>$MTTD_{n_1}(s)$</th>
<th>$DR_{n_2}$</th>
<th>$MTTD_{n_2}(s)$</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG Relative Speed</td>
<td>1</td>
<td>31.6</td>
<td>1</td>
<td>34.6</td>
<td>0</td>
</tr>
<tr>
<td>STD Relative Speed</td>
<td>1</td>
<td>31.7</td>
<td>1</td>
<td>32.6</td>
<td>0</td>
</tr>
<tr>
<td>AVG Inter-vehicle Spacing</td>
<td>1</td>
<td>30.2</td>
<td>1</td>
<td>28.9</td>
<td>0.87</td>
</tr>
<tr>
<td>STD Inter-vehicle Spacing</td>
<td>1</td>
<td>30.2</td>
<td>1</td>
<td>28.9</td>
<td>0.87</td>
</tr>
<tr>
<td>Proposed Algorithm (PoA = 40%)</td>
<td>$DR_{n_1}$</td>
<td>$MTTD_{n_1}(s)$</td>
<td>$DR_{n_2}$</td>
<td>$MTTD_{n_2}(s)$</td>
<td>FAR</td>
</tr>
<tr>
<td>AVG Relative Speed</td>
<td>1</td>
<td>34.6</td>
<td>0.9</td>
<td>35.7</td>
<td>0</td>
</tr>
<tr>
<td>STD Relative Speed</td>
<td>1</td>
<td>31.7</td>
<td>0.9</td>
<td>33.3</td>
<td>0</td>
</tr>
<tr>
<td>AVG Inter-vehicle Spacing</td>
<td>1</td>
<td>26.9</td>
<td>1</td>
<td>28.9</td>
<td>0.87</td>
</tr>
<tr>
<td>STD Inter-vehicle Spacing</td>
<td>1</td>
<td>26.9</td>
<td>1</td>
<td>28.9</td>
<td>0.87</td>
</tr>
<tr>
<td>Benchmark Algorithm (5)</td>
<td>$DR_{n_1}$</td>
<td>$MTTD_{n_1}(s)$</td>
<td>$DR_{n_2}$</td>
<td>$MTTD_{n_2}(s)$</td>
<td>FAR</td>
</tr>
<tr>
<td>$h = 30s$</td>
<td>1</td>
<td>38.0</td>
<td>0</td>
<td>-</td>
<td>0.33</td>
</tr>
<tr>
<td>$h = 60s$</td>
<td>0.9</td>
<td>27.0</td>
<td>0.3</td>
<td>28.0</td>
<td>0.64</td>
</tr>
</tbody>
</table>

The benchmark algorithm (5) to detect change points $n_1$ and $n_2$, where detection rates (DR), false alarm rates (FAR) and the mean time to detections (MTTD) are averaged over 10 simulated realizations. The proposed algorithm with the statistics of relative speed detects relatively higher number of change points with smaller number of false alarms than the benchmark algorithm. However, the proposed algorithm also misses one change point $n_2$ using the statistics of relative speed with $PoA = 40\%$. Recall that $n_2$ is the discharge point when the vehicles start to accelerate back to the normal speed and it is likely that the speed information in that realization comes from the vehicles that are less affected causing the change of statistics of relative speed to be more gradual. To increase the detection rate of point $n_2$, we increase $PoA$ to 50$\%$ which improves the detection rate of point $n_2$. Increasing $PoA$ increases the probability of obtaining information from vehicles that are affected by the change and consequently, increases the variability of relative speed for the proposed algorithm to detect the change.

As seen in Table 1, the benchmark algorithm in (5) detects well the change point $n_1$, but also misses most of $n_2$ for $h \geq 30s$. The miss detection of $n_2$ is likely due to the fact that the loop detectors has been ideally placed at the location of point $n_1$ in which the anomaly originates. However, the anomaly is caused by a moving vehicle and consequently, the discharge point $n_2$ was spatially farther away from the loop detectors location. We note that the location where the change points $n_1$ and $n_2$ take place are random in a real-world scenario and hence, shows a limitation of loop detectors as the detectors will need to be placed closely to where the disruption takes place to be able to detect changes effectively.

We further note from Table 1 that anomaly detections using the statistics of inter-vehicle spacing perform poorly due to very high false alarm rates. Those false alarms were caused by periodic detections of variance change points because the implemented car-following model let every vehicle continuously adjust the distance between itself and its leader regardless of whether there was a disruption of traffic flow. Based on this preliminary analysis, we therefore do not
pursuit further the use of inter-vehicle spacing for anomaly detection in this paper. The usefulness of inter-vehicle spacing statistics is an ongoing investigation and we only focus on relative speed for further analysis in this paper.

The impact of PoA and window size L on the false alarm rate is also assessed. Figure 2 shows the false alarm rates for PoA ranges from 20% to 100%. With the average $I_{total}$ of 16 vehicles/mile/lane and PoA of 20%, the relative speed statistics are calculated from only four vehicles on average which result in certain alarms not associated with the simulated anomalies. Note that under a low density scenario, the vehicles have large room to maneuver and there still is a probability of the proposed algorithm obtaining relative speeds from pairs of vehicles with high variations and raising alarms that are neither in $[n_1, n_1 + n_b]$ nor in $[n_2, n_2 + n_b]$. With PoA = 100%, the proposed algorithm can utilize relative speeds from all vehicles on the segment to accurately detect the simulated traffic anomalies with zero false alarm rates.

When PoA is small, the number of false alarms can be further reduced by increasing the sliding window size to assess temporally more relative speed samples. As shown in Figure 2, increasing the sliding window size from 10s to 30s reduces the number of false alarms as the algorithm can incorporate more temporal relative speed samples in making a decision.

PERFORMANCE EVALUATION USING REAL WORLD DATA

Descriptions of the Data

The objective of this section is to assess the usefulness of our proposed algorithm when applied to microscopic traffic variables measured in real-world scenarios. In order to validate and assess the proposed algorithm, we analyze a freeway segment in which the microscopic traffic variables can also be obtained from a video surveillance camera (19). The freeway segment we studied is part of
the main route that links Bangkok to the Northern provinces of Thailand. As the density of vehicles on the segment can vary with time and the vehicle density can be very low at certain periods, we use \( PoA = 100\% \) to guarantee that there are always enough individual vehicle information for our analysis.

On the image frame of the camera, a virtual entrance and exit lines were drawn at the beginning and the end of the segment respectively. For a vehicle \( i \), \( \{ t_{in}^i, t_{out}^i, w_{emp}^i \} \) was record, where \( t_{in}^i \) is the time that the vehicle crossed the entrance line, \( t_{out}^i \) is the time that the vehicle crossed the exit line, and \( w_{emp}^i = \frac{\text{Segment Length}}{t_{out}^i - t_{in}^i} \). A relative speed observed by vehicle \( i \) to its leading vehicle \( i - 1 \) is calculated as \( v_{emp}^i = w_{emp}^i - w_{emp}^{i-1} \) for \( t_{out}^i > t_{out}^{i-1} \). Then, for a sampling interval of \( \tau \) seconds, the average \( v_{emp} \) and standard deviation \( \sigma_{emp} \) of relative speeds of the sampling interval \( m\tau \) can be calculated as: \( v_{emp} = \frac{\sum_{i=1}^{n} v_{emp}^i}{m} \) and \( \sigma_{emp} = \sqrt{\frac{\sum_{i=1}^{n} (v_{emp}^i - v_{emp})^2}{m}} \) for \( (m-1)\tau \leq t_{out}^i < m\tau \), where \( m = 1, 2, 3, \ldots \) We note that in this real-world data set, a different method to calculate relative speed from the one in the simulation environment is used. The method here employed allows us to obtain relative speed from the image frame of the camera.

With the availability of video images, it is possible to determine types of anomalies and the times when traffic anomalies and the following traffic incidents took place. We also classify the real-world cases according to whether they lead to non-recurring congestion and assess the performance of our algorithm. We refer to the anomaly cases that lead to non-recurring congestion as non-recurring congestion precursors, while the anomaly cases that do not lead to congestion are simply referred to as transient anomalies.

There are totally of 22 cases of real-world traffic anomalies where 7 of them are transient anomalies and 15 are non-recurring congestion precursors. Each anomaly case consists of three timestamps: \( \{ T_s^a, T_i^a, T_e^a \} \), where \( T_s^a \) denotes the time when traffic anomaly was observed, \( T_i^a \) denotes the times when a traffic incident (e.g. congestions) took place and \( T_e^a \) denotes the end time of traffic incident, i.e. when traffic starts to recover. For an evaluation of an anomaly detection algorithm, an anomaly case is considered detected if an alarm that is raised by the algorithm within \( [T_s^a, T_e^a] \). Then, the mean time to detect (MTTD) is calculated as the average of the difference between the alarm time and \( T_s^a \).

### Results and Discussions

| TABLE 2 Evaluation Results of Anomaly Detection using the Proposed Algorithm with Relative Speed on Real-World Data. |
|-------------------------------|------------------|------------------|------------------|
| **Transient Anomalies**       | **Number of Cases** | **Detected Cases (s)** | **MTTD (s)**    |
| AVG Relative Speed            | 7                | 7                | 390             |
| STD Relative Speed            | 7                | 7                | 156             |
| **Non-recurring Congestion Precursors** | **Number of Cases** | **Detected Cases (s)** | **MTTD (s)** |
| AVG Relative Speed            | 15               | 12               | 300             |
| STD Relative Speed            | 15               | 14               | 210             |

In this section, for the purpose of assessing the statistics of relative speed, we show and discuss the performance evaluations of the proposed algorithm using both the average and stan-
Figure 3 shows a detection time of variance change point of relative speed associated with non-recurring congestion precursors on December 27, 2008. Furthermore, we also show the corresponding video snapshot of the freeway segment in Figure 4. As illustrated in Figure 4, the disruption was initially caused by two big trucks which evolved into a state of congestion.

Table 2 shows the results from using the proposed algorithm with the variance of statistics of relative speed to detect anomalies on the real world data with MTTDs less than seven minutes which is fast enough for some following actions to be taken. There is one case of precursors to non-recurring congestion that are missed by the proposed algorithm using the standard deviation of relative speed. This is a special case which actually took place between two consecutive periods of non-recurring congestion. The proposed algorithm in fact detects a precursor to the first period of congestion which lasted for approximately 45 minutes. Then, there was a discharge period.
of approximately 30 minutes before the second period of congestion took place again, and this 30-minutes discharge period was recorded as a precursor to the second non-recurring congestion. However, since there have already been a large number of vehicles on the segment due to the first congestion period, individual vehicles could only gradually increase their speed and consequently, there was not enough variability of relative speed for the algorithm to detect.

Similar phenomenon can be used to describe the other three undetected cases of precursors to non-recurring congestion using the average relative speed statistic. Even though these three anomaly cases were not discharge periods, they also took place while there was already high volume of vehicles ($\geq 2000$ vehicles/hour) on the segment. In fact, two cases occurred during the first days of long public holidays while another case occurred on a Saturday, which were the times that this route was heavily used by many Bangkok residents to go to the northern provinces. With such high volume of vehicles, individual vehicles did not have much room to maneuver and the variability of relative speed was reduced causing these precursors to be missed by the proposed algorithm.

As the undetected anomaly cases are mainly due to high density of vehicles on the freeway segment that reduces the variability of relative speed, further improvement could be obtained by employing multi-resolution models (24, 25), to extract and assess variability on different time scales. Also inter-vehicle spacing statistics could be used for detection and classification of transient anomalies and non-recurring congestion precursors. This is our ongoing investigation.

Another possible application of the proposed algorithm is in the classification of different traffic regimes. Figure 5 shows preliminary results in which transient anomalies are detected at 7:26:00, 10:49:00 and 14:48:00 on the inbound route. In Figure 6, we use these detection times (7:26:00, 10:49:00 and 14:48:00) to cluster the traffic flow and average speed data into four groups, and found that those detection times can approximate the times when traffic evolves. The first group belongs to the period 6:00:00-7:26:00 and consists mostly of low flow rate with scattering average speed as the traffic was light and individual vehicles had freedom to maneuver. The second and third groups, 7:26:00-10:49:00 and 10:49:00-14:48:00, consist mostly of higher speeds with increasing flow indicating the increasing number of inbound vehicles. Finally, the forth group

---

FIGURE 4 Video Snapshot of Non-Recurring Congestion on December 27 2008.
FIGURE 5 Average Relative Speed on August 12, 2008: Dotted vertical lines denote points where the variance changes.

consists of traffic data from 14:48:00-18:00:00 when more vehicles had arrived and the capacity of the segment had been reached as the traffic flow and speed start to decline.

FINAL REMARKS

We investigate the suitability of the proposed anomaly detection algorithm using the variance of microscopic traffic variables and the results so far are very encouraging. Based on the simulation results, the statistics of relative speed perform very well with minimal prior knowledge in detecting traffic anomalies compared to a well-known algorithm (5) running in an ideal setting. The algorithm is also shown to perform well under different availability of individual vehicle information. Furthermore, the analysis on real world data shows that our algorithm with the statistics of relative speed can be used to detect transient anomalies and also the precursors to non-recurring traffic congestion. Subsequent analysis also shows another possible application of our algorithm to identify different traffic regimes. We note that the early warning alarms points that our algorithm detects are not the points where any action/response (e.g. dispatching tow trucks) should be taken. Instead, these are the time points where the road traffic should be more closely monitored, running further classification algorithm and eventually, triggering a more comprehensive set of actions like, e.g. activating automatic incident detection algorithms (2, 3) or disseminating warning messages to other vehicles (11, 12, 13).

One aspect in which we are expanding this work is the classification of anomalies to determine whether they could lead to traffic congestion. We are at present in the process of assessing other microscopic variables including inter-vehicle time gaps and lane changing characteristics that will permit us to enhance the classifications of traffic anomalies. The methodology will incorporate both temporal and spatial changes in the variability of microscopic traffic variables. The effectiveness of the classification algorithm will be assessed using real-world data as well as a simulation
As it is found in this paper that anomalies are likely to be missed under high vehicle density scenarios, another aspect is to derive a model that can incorporates statistics of inter-vehicle spacing to improve detection and classification. We are also investigating more sophisticated signal processing techniques such as wavelet transform and multi-resolution models (24, 25) that will permit us to extract microscopic changes on different scales. These models will be used to analyze temporal and spatial changes of microscopic traffic variables, while changes can be identified on different resolution components of each variable. We expect that incorporating multi-resolution models will enhance the detection and classification especially under extremely low and high vehicle density scenarios. The assessment of these aspects will be presented in a forthcoming paper.

ACKNOWLEDGMENT

We would like to thank Prof. John Polak of the Centre for Transport Studies (CTS) at the Imperial College London for useful discussions. We also would like to thank Dr. Wasan and Dr. Supakorn from the National Electronic and Computer Technology Center (NECTEC), Thailand for their assistance in providing the real-world data for our analysis.

REFERENCES


