Trends in Transport Engineering and Applications

Volume 1, Issue 2
www.stmjournals.com

# Development and Evaluation of Neural Network Model for Incident Detection on Urban Arterials using Simulated Database

***Tean C*[1], *Hossain M*[2]***
[1]Intel Malaysia Inc., Palau Penang, Malaysia University of Science and Technology, Malaysia
[2]Department of Civil Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh

***Abstract***

*Incident detection in urban arterial situation is more difficult than the similar job in freeway situation because of the presence of traffic signals and other intersections with associated recurrent queue. Most of the earlier automatic incident detection algorithms address mainly freeway situation. This study aims at development, calibration, validation and testing of an ANN model for incident detection in Kuala Lumpur (KL) arterials using simulated incident database. Database for the study is generated under incident and no-incident condition by simulating the traffic flow through the arterial network of Golden Triangle area of KL. Calibration efforts are divided into different tasks such as the time interval for input neuron, recalculation interval, location of the detector and the threshold values for the model. The calibrated model for optimum location of detector yields 98.5% of detection rate and 2.9% of false alarm rate for normal traffic demand situation. It is found that in case of link longer than 350 m data from two detectors are required for better performance of the ANN model but a single detector data is good enough for link length of around 220 m or less. Testing of the model on other link sites also yields similar results with more accurate detection in case of shorter links. While one cycle time was found to be long enough as a recalculation interval, further sensitivity analysis on this revealed that lower cycle time of around 60 s degrades the performance of the model in terms of false alarm rate. The results from this study provide useful insights for the design of AID system in urban arterials.*

***Keywords:*** *Incident detection, neural network, urban arterial, micro-simulation*

****Author for Correspondence*** E-mail: moazzem@ce.buet.ac.bd

## INTRODUCTION

Significant traffic delay in urban freeways and arterials is caused by non-recurring congestion initiated by incident of some sort. The incident type includes accident, vehicle disablements, flat tires, adverse weather condition, road maintenance activities and other events that obstruct the normal traffic flow and result in a capacity reduction. In order to reduce the adverse impact brought about by incidents, it is important to be able to know the occurrence of an incident as soon as possible. In large urban arterial network like Kuala Lumpur (KL) early detection of incidents can reduce the frequency of network wide flow breakdown. Automatic incident detection (AID) model on freeways has been developed since the early 1970s and their structure varies in the degree of sophistication, complexity and data requirements and can generally be grouped into categories, namely, simple filtering [1], comparative or pattern comparison algorithms [2–4], traffic model and theoretical algorithms [5], time series algorithms, statistical algorithms [6] and neural networks [7, 8]. Comparing among three types of neural network models, namely, the multi-layer feed forward (MLF), the self-organizing feature map (SOFM) and adaptive resonance theory 2 (ART2), it was claimed that MLF has the highest potential to achieve a better incident detection performance [9]. MLF was also reported to be outperforming California, McMaster and Minnesota algorithms in the same study. More recently, Hussein and Rose [10] developed and evaluated a neural network model on freeway incident detection using field data of 100 incidents. Their model used speed, flow and occupancy data as the input to neuron. The results obtained proved that neural network models can provide reliable incident detection on freeways. But incident characteristics and mechanism on urban arterials are significantly different from those on freeways due to differences in access, geometric constraints, control measures, operating conditions and surveillance infrastructures [11]. William and Koppelman investigated the effectiveness of individual vehicle movement measures in the detection of incidents on urban arterial road segments using vehicle-positioning data with only modest success [12]. Traffic dynamics in average speed, running time, speed and coefficient of variation of speed calculated from vehicle positioning data are used in the study for identifying incidents on link. Many practitioners find AID too problematic for implementation in large urban environments [13] due to high false alarm rate resulting from stop-and-go traffic in urban arterial situation. Using the simulation data, Sethi *et al.* developed two threshold parameters for incident detection on urban arterial: volume divided by occupancy from fixed traffic detector data and average speed using probe vehicle data [14]. Performing a case study on an arterial road in Melbourne, Australia, Luk and Chung [11] showed that detectors upstream of an incident are more useful for incident detection than downstream detectors. The study also identified occupancy and speed as the appropriate parameters to characterize and detect arterial incidents.

Khan and Richie compared the efficiency of neural network and other statistical classifiers in detecting incidents on urban arterials and claimed that modular neural classifier outperformed other statistical classifiers. It is also suggested that incident detection depends on factors such as operating conditions, configuration of sensors within the network, and block or link length. Hawas (2007) commented that detecting incidents on urban streets or arterials using loop detector data is quite challenging. The pattern of the incident might vary significantly due to several factors including the average flow rate, the link length, the detectors' locations, the incident location with respect to the detectors' locations, the degree of link blockage by the incident, the green time of the downstream signal and the cycle length. Using fuzzy logic for incident detection, the study suggested further research on issues of optimal number of detectors, and relationship between numbers of detectors and their locations vis-à-vis the effectiveness of the devised logic [15]. In absence of real incident data in most cases, development of simulated database is necessary to investigate the detailed characteristics of these incident detection model variables and the relevant factors influencing the performance of the model.

Considering the successful utilization of neural network approach to model the incident detection of freeways and arterials, this research will use the MLF neural classifiers in studying the factors influencing the incident detection performance of the model. While neural network is an extensive data-based modeling technique, authentic incident data and associated other data items are not available for the study area of KL. Microscopic traffic simulation model MITSIM developed at MIT, USA, can simulate incidents with stochastic characteristics close to real life incidents [16–18]. As a part of KL intelligent transport system (ITS) initiative, MITSIM is calibrated and validated for a portion of KL arterial network. Therefore, a simulated incident database can be created for developing and testing the ANN incident detection model under various situations. This paper presents development, calibration, validation and testing of the artificial neural network (ANN) model for incident detection on urban arterial as well as applications for detailed study of the factors influencing the performance of such ANN model.

## METHODOLOGY

This paper reports on the application of MLF ANN model to study the factors influencing the performance of incident detection model on urban arterials. In this section, ANN modeling scheme and model performance measures will be described.

## ANN Model

Multi-layer feed-forward neural network was successfully used for incident detection models in freeway situations using both simulated and real field data. In this study, similar MLF neural net model is developed for the more complex traffic, geometric and control environment of urban arterial network in comparison to the freeway. The MLF consists of processing elements (PEs) arranged in three layers: the input layer, a hidden layer and the output layer. The input layer consists of data from loop detectors on the roads, hidden layer processes data and output layer gives an incident or incident-free signal. Supervised training which involves letting the network know if its output matches the correct condition is carried out in order to obtain a stable network. Through training, the network learns the appropriate weights to apply to the inputs and outputs. As most of the existing detection tools report three traffic variables of speed, flow and occupancy, all of them are used as input variables in the present study considering the complexity of identification in urban arterial situations. As detectors are lane based each lane will introduce three input neurons at the input layer of the network. The number of neurons in hidden layer and their connection architecture with input and output layer can be calibrated towards a suitable figure after experimenting with different numbers and architectures. For the present task of incident detection the output layer can be consisted of one neuron indicating incident or no-incident condition.

The ANN architecture taken for the present research study is shown in Figure 1. The network consists of lane-based input data such as count, speed and occupancy at the input layer requiring three input neurons per lane per location considered. This means, for a three lane road and inclusion of data from a pair of detectors in the link, will result 18 input neurons as in the case of present study. After a few experiments, it is found that number of neurons in excess of 80 in the hidden layer only result little or no improvement in network estimation efficiency. The output layer consisting of one neuron is required to provide a state value indicating incident or no-incident condition.

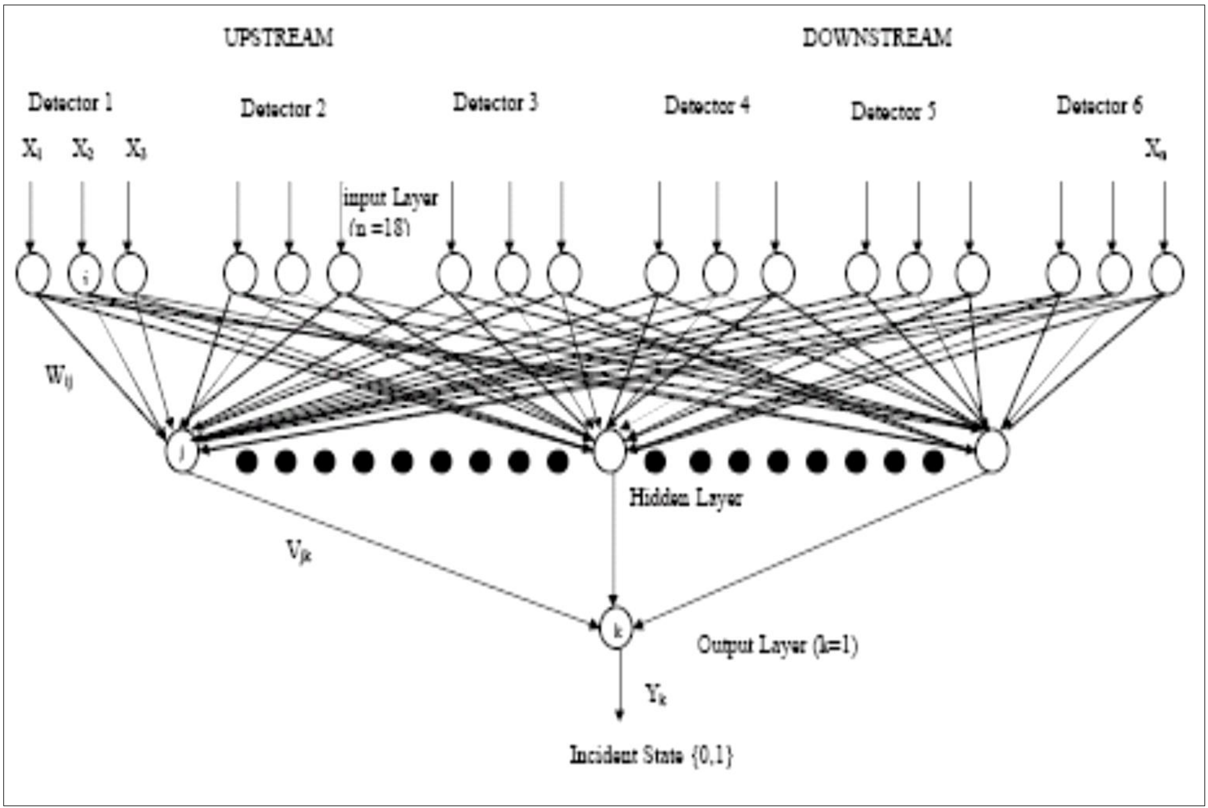



***Fig. 1:*** *Schematic of MLF Type ANN Model Used in the Study.*

**Model Performance Measures**

Generally, the performance of an AID algorithm can be measured by using three basic measures of effectiveness (MOEs): false alarm rate (FAR), detection rate (DR) and mean time to detect (MTTD). DR and FAR measure the effectiveness of an algorithm, whereas MTTD measures the efficiency of the algorithm (Thomas and Dia, 2000; Rakha and Aerde, 1996). They are defined as follows:

False alarm rate: FAR may be defined as the percentage of incorrect detections to the total number of algorithm applications to incident-free data set. However, in this project, the FAR is defined as:

$$FAR = \frac{No.OfIncidentFreeIntervalsWhichGivesFalseAlarms}{TotalNo.OfAlarmsRaise} x100\%$$

Detection rate: DR is defined as the percentage of number of detected incidents to the actual number of incident in the data set. In other words, it can also be defined as the percentage of actual incidents which are detected by the algorithm out of all the true incidents that have been reported during a time period.

$$DR = \left( \frac{No.OfActualIncidentsDetected}{TotalActualIncidents} \right) x100$$

Mean time-to-detection: Mean time-to-detection (MTTD) is computed as the average time elapsed between the time the incident started and was detected. Since the reporting interval is cyclic, TTD is expressed in cycles. This interval could vary depending on the base cycle length in place. However, in this research, MTTD would not be used as one of the performance measures because for arterials the recalculation interval would be a function of traffic signal cycle time and hence would affect the MTTD without any relevance to the algorithm efficiency.

**Data Generation**

This section describes the street network under study and application of MITSIMLab simulation model to the study network for generating the data required for the present modeling study.

**Study Network**

Golden Triangle road network of Kuala Lumpur (shown in Figure 2) is calibrated by the MITSIM research group at Malaysia University of Science and Technology (MUST) and it is taken as the network for present study. For the purpose of detailed modeling Jalan Sultan Ismail corridor has been chosen as the area of study. The details of Sultan Ismail corridor including the possible location of the detectors and incidents are shown in Figures 3 and 4. In a typical link, four locations of detectors and incidents are taken to be investigated in the present study with various combinations of detector data sets as shown in Tables 1 and 2.

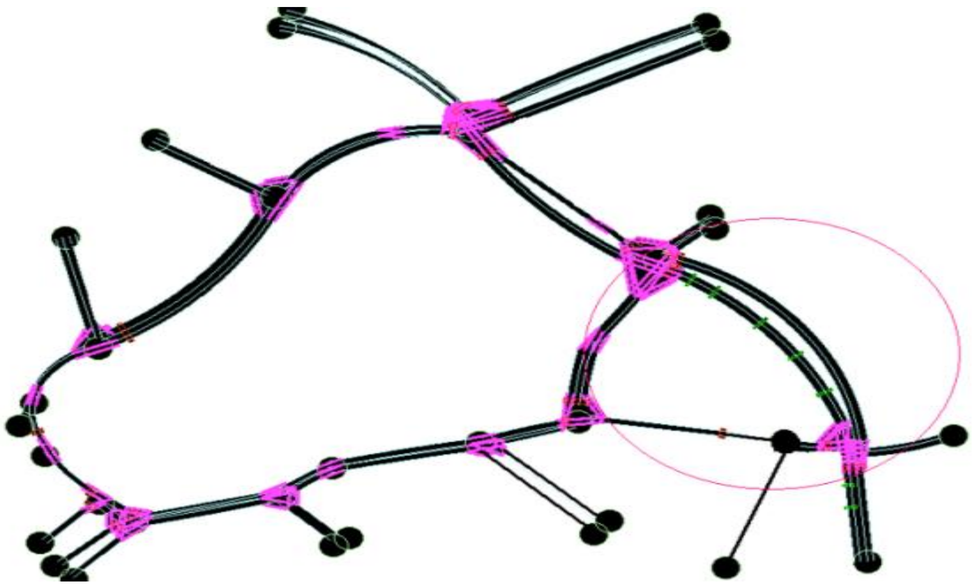


***Fig. 2:*** *Schematic of Golden Triangle Road Network from MITSIM Simulation Display (Sultan Ismail Corridor Marked in the Circle).*

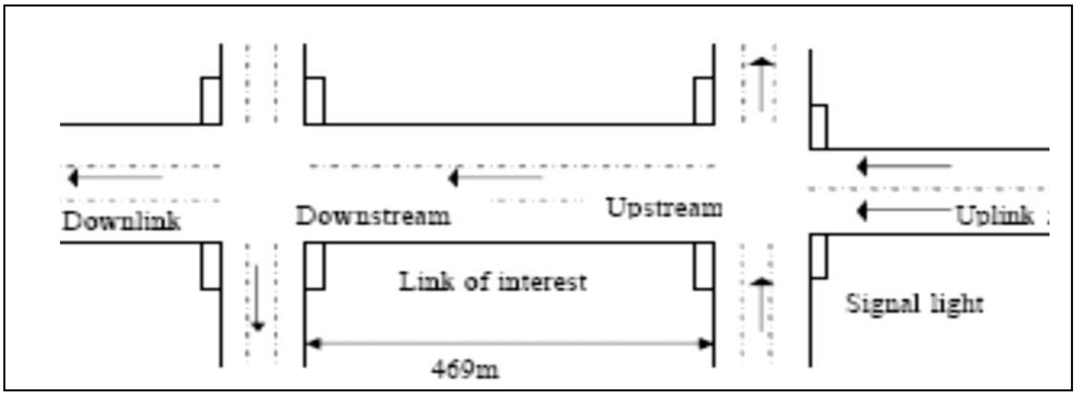



***Fig. 3:*** *Details of Jalan Sultan Ismail Corridor.*

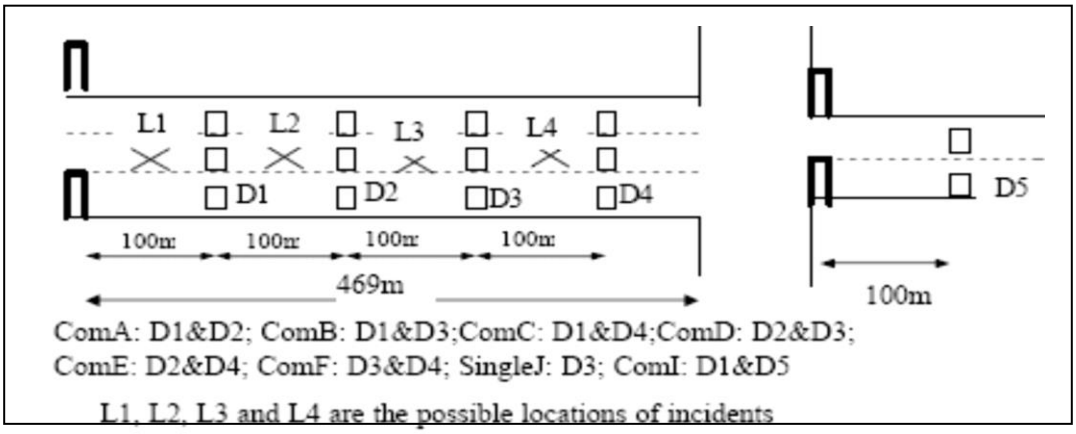



***Fig. 4:*** *Link of Detailed Study with Detector Pairs.*

***Table 1:*** *Characteristics of the Road in the Corridor of Jalan Sultan Ismail.*

| **Item** | **Link** | | |
|---|---|---|---|
| | **Uplink** | **Link of interest** | **Downlink** |
| Length | 220 m | 469 m | 370 m |
| Width | 2 lanes each way | 3 lanes each way | 3 lanes each way |
| Speed limit | 55 km/h | 55 km/h | 55 km/h |
| Mean free flow speed | 65 km/h | 65 km/h | 65 km/r |
| Location of incident | | Distributed equally at L1, L2, L3 and L4 | |
| Location of sensor | 100 m away from traffic signal | Every 100 m away from traffic signal | - |
| Type of sensor | Lane based | Lane based | Lane based |
| Accuracy of sensor | 100% | 100% | 100% |
| Data output from sensor | Count, speed, occupancy | Count, speed, occupancy | Count, speed, occupancy |
| Incident characteristic | - | Incident lane fully blocked while non-incident lane partially blocked | - |

***Table 2:*** *Traffic Signal Characteristics.*

| *Item* | *Uplink* | *Link of interest* | *Downlink* |
|---|---|---|---|
| Duration of total green phase | 45 s | 45 s | 50 s |
| Duration of total amber phase | 3 s | 3 s | 3 s |
| Duration of red phase | 105 s | 101 s | 101 s |
| Total cycle time | 153 s | 149 s | 154 s |
| Number of phases | 2 | 3 | 4 |

**Microscopic Simulation**

MITSIMLab is a microscopic traffic simulation laboratory that is developed to evaluate advanced traffic management systems (ATMS) and advanced traveler information systems (ATIS) at the operational level. A wide range of traffic management systems and the response of drivers to real-time traffic information and control can be modeled in MITSIMLab. By using microscopic traffic simulator in this study, a traffic network can be simulated under a controlled environment. Loop detectors, incident location and the time period where the incident occurs and also the severity and duration of the incidents can be assigned easily to the network. For training of the model, data has been generated for a total of 180 simulated hours with 100 incidents. But for calibration purpose, data set comprises a total of 540 simulated hours with 300 incidents. Also another data set comprising 72 simulated hours with 40 incidents is used for testing the model and studying the factors influencing the performance of the model. The design of data set generation from simulation lab is presented in Table 3. Different durations of incidents in the range of 15 to 30 min are simulated under different traffic demand conditions with each set of data generated from 18 h of simulation runs.

***Table 3:*** *Simulated Incident Characteristics for Training, Calibration and Validation Data Set (Each set Represents 18 h of Simulation Run).*

| **Data *division*** | **Traffic *demand*** | **Number of incidents** | **Duration of *inc*idents (min.)** | **No. *of* incidents at different durations** |
|---|---|---|---|---|
| Training data (10 sets and 100 incidents) | Low | 40 | 15 | 10 |
| | | | 20 | 10 |
| | | | 25 | 10 |
| | | | 30 | 10 |
| | Normal | 20 | 15 | 5 |
| | | | 20 | 5 |
| | | | 25 | 5 |
| | | | 30 | 5 |
| | High | 40 | 15 | 10 |
| | | | 20 | 10 |
| | | | 25 | 10 |
| | | | 30 | 10 |
| Calibration data (30 sets and 300 incidents) | Low | 116 | 15 | 46 |
| | | | 20 | 30 |
| | | | 25 | 22 |
| | | | 30 | 18 |
| | Normal | 68 | 15 | 28 |
| | | | 20 | 20 |
| | | | 25 | 12 |
| | | | 30 | 10 |
| | High | 116 | 15 | 42 |
| | | | 20 | 36 |
| | | | 25 | 26 |
| | | | 30 | 10 |
| Validation data (10sets and 100 incidents) | Low | 35 | 15 | 8 |
| | | | 20 | 9 |
| | | | 25 | 9 |
| | | | 30 | 9 |
| | Normal | 30 | 15 | 7 |
| | | | 20 | 7 |
| | | | 25 | 8 |
| | | | 30 | 8 |
| | High | 35 | 15 | 8 |
| | | | 20 | 9 |
| | | | 25 | 9 |
| | | | 30 | 9 |

## Model Development

This section studies the incident detection model parameters. It also discusses the model building, training and calibration using the above described MITSIMLab generated simulation database.

### *Incident Detection Parameters*

In the present study, speed, volume and occupancy data for each lane averaged over a signal cycle has been used as basic input data. The purpose of setting the data interval similar to the traffic light cycle time is to overcome the problem of the effect resulted by the traffic light. This is because, if the interval is not taken at least equal to a full cycle, it would happen that the data collected are under the red light conditions when all the vehicles need to stop/slow down without any regard to incident. The performance of the network would be greatly affected being confused between the red light stop and stop due to incident. However, the effect of other recalculation interval as multiple of the signal cycle will also be checked later. Two data input schemes will be employed: one comprising data points at time t only; another comprising data points at time t and t−1 intervals. Input data vector from pairs of detectors as well as from single detector would be tested which are illustrated as different combinations of input data source, i.e., detectors in the detailed corridor diagram (Figure 4). Persistence of incident state will also be checked in the subsequent intervals in order to be certain about the case of an incident and thus reducing the higher false alarm rate from possible traffic noises at upstream of a traffic signal.

### *ANN Training*

Under the supervised training, data generated from MITSIMlab under incident and non-incident conditions are presented to the network with corresponding desired state value. This supervised training is also called back propagation training because the information is going from input layer, passing through hidden layer, to output layer and the information reverse back from output to input layer as a function of estimation error. This process is repeated and the information used for training is propagated. After training, the network is stable and it should have the ability to produce the correct output vector which indicates incident or incident-free conditions in this case. Two output states are used to describe the traffic conditions within the section of link under consideration: state 1 with introduced state value of 0 to represent incident-free conditions and state 2 with introduced state value of 1 to represent incident conditions. In order to improve the robustness of the model in detecting incident under diverse condition, data set used for training are generated from a wide range of road conditions such as different flow, traffic periods, duration and severity of incidents. In the present study, the training data have been iteratively presented to the ANN for 450 cycles. During each of the iteration, sum of squared of errors (SSE) will be reduced gradually until it reaches a minimum point which is set at zero in the present case. When SSE becomes zero, it means that all the weights in the network have already converged and it is considered as stable and at optimal state. So the neural network has already learned and memorized the relationship between the input and output data. Activation functions are the mathematical transfer functions that determine which group of the data is given the most emphasis. In this research, logistic activation function is used because all the data is treated equally during training of neural network for incident detection. Also as a learning process during training scaled conjugate gradient is used in the present study considering the claim of its better performance (Bishop C, 1995). For details of the mathematical mechanisms of training procedure readers may refer to this study [9].

However, before the process of training, normalization of the data needs to be carried out. This is because count in number of vehicles, speed is km per hour and occupancy in percentage of time have different values which are different by several orders of magnitude and the typical sizes of the inputs may not reflect their relative importance in determining the required outputs. Through normalization or linear transformation, all of the inputs can be arranged to have similar values and the speed for training is increased. Normalization is performed by finding the mean and standard deviation for the input data as shown in equations below:

$$\bar{x}_i = \frac{1}{N}\sum_{n=1}^{N} x_i^n \tag{1}$$

$$\sigma_i^2 = \frac{1}{N-1}\sum_{n=1}^{N}\left(x_i^n - \bar{x}_i\right)^2 \tag{2}$$

Then each of the data can be transformed into new representation through equation below:

$$\bar{x}_i^n = \frac{x_i^n - \bar{x}_i}{\sigma_i} \tag{3}$$

where, n = 1, 2, 3,…N (number of the data) and i in this case is count/speed/occupancy.

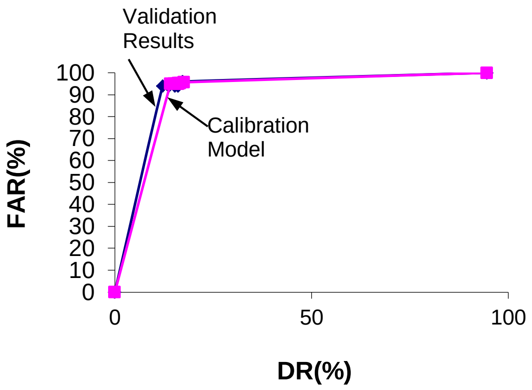



***Fig. 5:*** *Validation of the Calibrated Model.*

### *Model Calibration and Validation*

After development and training of the neural network for automatic incident detection on arterial, calibration needs to be done in order to obtain an optimal and mature network. Besides, the purpose of calibration is to explore the relationship of each parameter involved in the incident detection model so that it could be understood better. Calibration is done on the selection of input data, recalculation interval of data, different placement of detector and the threshold value for output evaluation.

The investigation reveals that input data to the model should comprise only the present data interval. The performance of the neural networks is worse when input data is taken from both previous and present intervals. The results as presented in Table 4 show that although for both input data schemes detection rates (DRs) are acceptable the false alarm rates (FARs) for input data including previous period are too high. It is probably due to the fact that data sets including the previous cycle data may make the network confused and hence it fails to find out an optimal relationship from them. Besides, in cases of freeway, input with previous data would be more useful because the data reporting interval would be as small as 30 to 60 s, so the data represent traffic pattern of short duration only, which in most cases remains similar. Therefore, in the present study, it is decided to use the data at the present time interval as input to the neuron for the neural network to identify incident or no-incident condition.

***Table 4:*** *Calibration on Data Set Selection.*

| **Input types** | **Detection outputs** | **Detector Combinations** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **A** | **B** | **C** | **D** | **E** | **F** | **I** | **Single _J** |
| Inputs of data interval t | Total incidents | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| | Total false alarms | 115 | 81 | 102 | 84 | 54 | 227 | 264 | 303 |
| | Total detections | 258 | 268 | 276 | 273 | 285 | 231 | 229 | 230 |
| | DR | 86.0 | 89.3 | 92.0 | 91.0 | 95.0 | 77.0 | 76.3 | 76.7 |
| | FAR | 30.8 | 23.2 | 27.0 | 23.5 | 15.9 | 49.6 | 53.5 | 56.8 |
| Inputs of data interval t, t−1 | Total incidents | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| | Total false alarms | 845 | 253 | 249 | 261 | 170 | 876 | 148 | 235 |
| | Total detections | 298 | 283 | 264 | 288 | 276 | 243 | 250 | 231 |
| | DR | 99.3 | 94.3 | 88.0 | 96.0 | 92.0 | 81.0 | 83.3 | 77.0 |
| | FAR | 73.9 | 47.2 | 48.5 | 47.5 | 38.1 | 78.3 | 37.2 | 50.4 |

Considering the detector location, combination E yields the best result among all the placements of detectors with 95.0% of detection rate and 15.9% of false alarm rate (Table 4). Taking input data at present time interval, one cycle of recalculation interval and detectors placed at combination E, the effect of output layer threshold value is studied. The performance of the model with different threshold is shown in Table 5.

From Table 5, it is found that the difference in threshold value from 0.1 to 0.9 does not bring much effect to the performance of the detection rate. The entire detection rate is around 95%. Therefore, a threshold value of 0.5 would be chosen since all the threshold value's performance beyond that value is quite close.

With all the optimal parameters, validation of the model is again carried out by using a new set of simulated data. Ten sets of new data are generated from the same Jalan Sultan Ismail corridor and the data with the parameters suggested are used. Table 6 and Figure 5 show the result of the validation using the suggested parameters which shows consistent performance of the network with the new data set even at all threshold value levels.

***Table 5:*** *Performance of the Model with Calibrated Parameters for Detector Combination E.*

| Threshold Value | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| Total incidents | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| Total false alarms | 61 | 57 | 55 | 55 | 54 | 52 | 50 | 50 | 47 |
| Total detections | 287 | 286 | 286 | 285 | 285 | 285 | 285 | 285 | 285 |
| Detection rate | 95.7 | 95.3 | 95.3 | 95.0 | 95.0 | 95.0 | 95.0 | 95.0 | 95.0 |
| False alarm rate | 17.5 | 16.6 | 16.1 | 16.2 | 15.9 | 15.4 | 14.9 | 14.9 | 14.2 |

***Table 6:*** *Result of the Validation Process Using Suggested Parameters.*

| Threshold value | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| Total incidents | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Total false alarms | 20 | 18 | 18 | 18 | 18 | 17 | 15 | 15 | 13 |
| Total detections | 96 | 95 | 95 | 94 | 94 | 94 | 94 | 94 | 94 |
| Detection rate | 96.0 | 95.0 | 95.0 | 94.0 | 94.0 | 94.0 | 94.0 | 94.0 | 94.0 |
| False alarm rate | 17.2 | 15.9 | 15.9 | 16.1 | 16.1 | 15.3 | 13.8 | 13.8 | 12.2 |

## Factors Influencing the Performance of ANN Model

This paper investigates the factors influencing the performance of incident detection model of ANN type. The factors considered to be affecting the performance of incident detector may include the average flow rate, the link length, the detectors' locations, the incident location with respect to the detectors' locations, the degree of link blockage by the incident, data interval and the signal cycle length.

### *Flow Rate*

For testing the performance of the model under varying demand, traffic demand is classified into three groups: the volume below 600 vehicles per lane is categorized as low demand period; volume in the range of 600 to 1100 vehicles per lane is considered as normal demand period and volume of more than 1100 vehicles per lane is classified as high demand period.

***Table 7:*** *Sensitivity Analysis on Demand variation (1 cycle of data interval at ComE).*

| | | |
|---|---|---|
| Low Demand | Detection Rate | 89.6 |
| | False Alarm Rate | 29.7 |
| Normal demand | Detection rate | 98.5 |
| | False alarm rate | 2.9 |
| High demand | Detection rate | 98.3 |
| | False alarm rate | 6.6 |
| Overall result | Detection rate | 95 |
| | False alarm rate | 15.9 |

Table 7 shows the performance of the neural network with the best loop detector combination under different demand period by using one cycle of recalculation interval. During low demand, the detection rate is 89.65% with a false alarm rate of 29.73%. For normal demand, detection rate is 98.53% with false alarm rate 2.9% whereas for high demand, detection rate is 98.28% and false alarm rate is 6.577%. The algorithms raised more false alarms during low demand and this could be due to the relatively inconsistent queue length during low flow period. During high demand, the performances of the neural networks are better compared to low demand. This is due to the fact that when incident happens during high flow, the effect of that is more prominent in comparison to the similar situation under low demand. However, as predictable the performance of the ANN model is best under normal demand as this period is supposed to present relatively consistent data to the model.

***Link Length***

The calibrated ANN model is tested on different lengths of the links under normal traffic condition: one of 220 m length with location of the detectors as shown in Figure 6, where the downstream detectors, are 100 m away from traffic signal and the upstream detectors are 100 m apart as the link is short; one of 370 m length with detectors as indicated in Figure 7 and another of 469 m length as shown earlier in Figure 3 for the case of combination E.

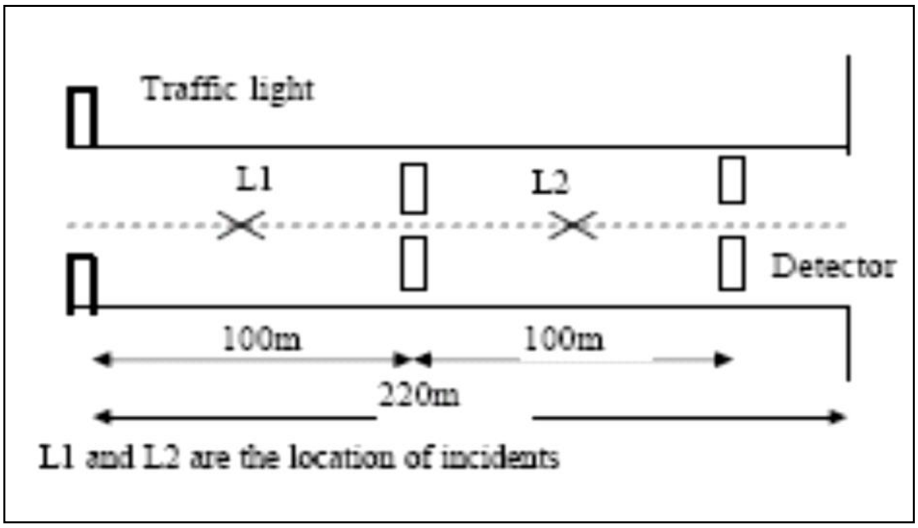



***Fig. 6:*** *Location of Detectors Pair in 220 m Link.*

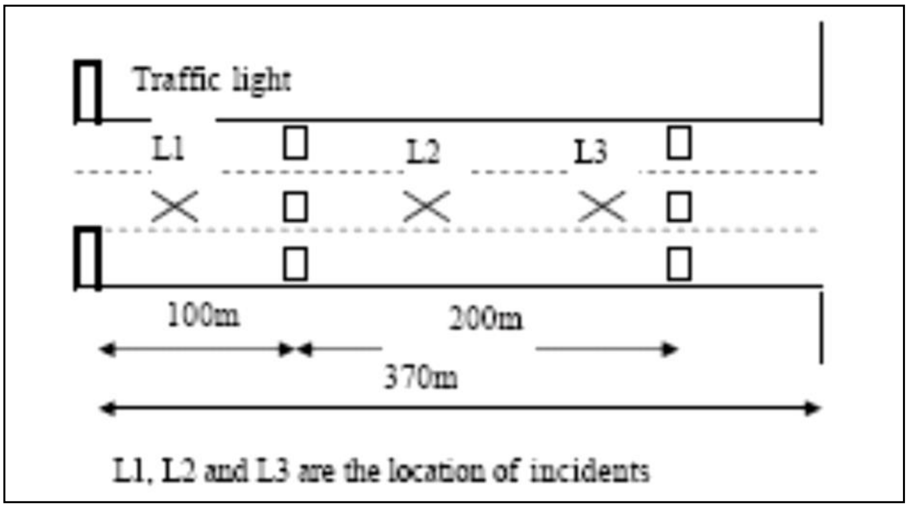



***Fig. 7:*** *Location of Detectors Pair in 370 m Link.*

The performances of the neural networks under different link lengths (typical of the study area) are shown in Table 8. DR and FAR are the best for shortest link of 220 m yielding 97.5% of detection rate and 2.5% of false alarm rate. This is because in shorter links incidents are more likely to be closer to the detectors, and hence more likely to be detected correctly. For the other two links of higher length, DR and FAR values are similar.

### *Recalculation Interval*

Recalculation interval is the time interval that the detectors report the field data to the neural network model. Investigation is carried out in order to find out the most suitable recalculation interval for the detectors to report the data – count, speed and occupancy to the neural network. In this study, one cycle, two cycles and three cycles of the traffic light timing have been experimented as the recalculation interval considering the fact that any interval less than a cycle will not be able to avoid the traffic light interferences.

From the results as shown in Table 9, it can be observed that with increase in recalculation interval DR is lowered marginally but significant improvement is achieved in lowering the FAR. This is due to the fact that with longer data interval, chances of small duration traffic noises appearing as incidents are reduced.

***Table 8:*** *Results under Different Lengths of Link.*

| Link length | 220m | 370m | 469m |
|---|---|---|---|
| Total incidents | 40 | 40 | 300 |
| Total false alarms | 1 | 6 | 54 |
| Total detections | 39 | 38 | 285 |
| DR | 97.5 | 95.0 | 95.0 |
| FAR | 2.5 | 13.6 | 15.9 |

***Table 9:*** *Results for Various Combinations of Detectors Pair at Different Recalculation Intervals.*

| ***Data interval*** | ***Detector pair*** | ***ComA*** | ***ComB*** | ***ComC*** | ***ComD*** | ***ComE*** | ***ComF*** | ***ComI*** | ***ComJ*** |
|---|---|---|---|---|---|---|---|---|---|
| 149 s (1 cycle) | Total incidents | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| | Total false alarms | 115 | 81 | 102 | 84 | 54 | 227 | 264 | 303 |
| | Total detections | 258 | 268 | 276 | 273 | 285 | 231 | 229 | 230 |
| | DR | 86 | 89.33 | 92 | 91 | 95 | 77 | 76.33 | 76.67 |
| | FAR | 30.8 | 23.2 | 27.0 | 23.5 | 15.9 | 49.6 | 53.5 | 56.8 |
| 298 s (2 cycles) | Total incidents | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| | Total false alarms | 27 | 16 | 17 | 13 | 4 | 89 | 61 | 93 |
| | Total detections | 258 | 279 | 278 | 276 | 282 | 216 | 223 | 222 |
| | DR | 86 | 93 | 92.67 | 92 | 94 | 72 | 74.33 | 74 |
| | FAR | 9.5 | 5.4 | 5.7 | 4.5 | 1.4 | 29.2 | 21.5 | 29.5 |
| 447 s (3 cycles) | Total Incidents | 300 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| | Total false alarms | 8 | 8 | 3 | 6 | 7 | 40 | 9 | 55 |
| | Total detections | 265 | 242 | 260 | 269 | 268 | 217 | 226 | 223 |
| | DR | 88.3 | 80.7 | 86.7 | 89.7 | 89.3 | 72.3 | 75.3 | 74.3 |
| | FAR | 2.9 | 3.2 | 1.4 | 2.2 | 2.5 | 15.6 | 3.8 | 19.8 |

### *Cycle Time*

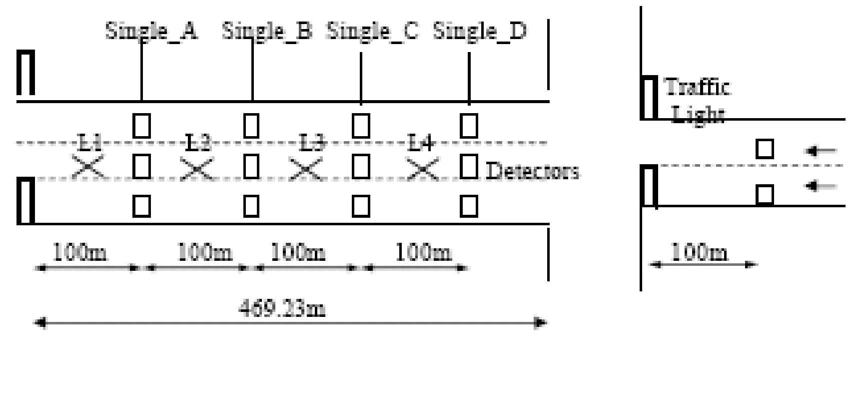
In order to test the ANN model under different signal cycle times, 14 sets of new data (as the earlier data sets are generated using existing signal cycles and hence do no provide a suitable data set for testing the cycle time variation) each representing 18 h of simulation run under normal demand flow are generated with signal cycles in the range of 60 to 180 s. From the data collected, 10 sets are used for training while 4 sets are used for testing for each cycle time. The results of the experiment are shown in Table 10, which reveals that the suggested neural network parameters are suitable to be used under different cycle timings as long as the network learns from the situation. In all the cycle timings, the neural networks have a very high detection rate and low false alarm rate in cases of 90, 120 and 180 s. The false alarm rate for 60 s is relatively higher which is a more difficult case for the ANN, due to the relatively frequent queuing and thus making the ANN confused in mapping out the pattern. Consequently, more false alarms are raised.

***Table 10:*** *Results under Different Cycle Timing.*

| *Cycle time* | *60 s* | *90 s* | *120 s* | *180 s* |
|---|---|---|---|---|
| Total incidents | 40 | 40 | 40 | 40 |
| Total false alarms | 16 | 4 | 0 | 0 |
| Total detections | 39 | 39 | 38 | 39 |
| DR | 97.5 | 97.5 | 95 | 97.5 |
| FAR | 29.1 | 9.3 | 0 | 0 |

***Detector Location***

When using two sets of detectors per link lane, there might be a number of combinations of detector locations for the relatively longer link length as shown in Figure 3 for link length of 469 m. For such a link, combination E yields the best result as shown in Table 9. The authors have also considered four different detector locations when using a single detector per link lane as shown in Figure 8. The results in case of the four detector locations are presented in Table 11. Therefore, for relatively longer link length (here 469 m), single detector is not effective in detecting incident. For shorter link lengths of 220 and 370 m, the authors have tested the single detector location at about the middle of the link. The relevant results in Table 12 show that single detector is only effective at about 220 m length. With two sets of detectors per link lane highest DR and lowest FAR are achieved in case of shortest links of about 200 m length, but for longer link length around or over 370 m, both DR and FAR achieved are similar and still within acceptable range.



L1, L2, L3 and L4 are the location of incidents

***Fig. 8:*** *Possible Locations of Single Detector Set.*

***Table 11:*** *Performance of Single Location Model for 469 m Link.*

| **Item** | **Single A** | **Single B** | **Single C** | **Single D** |
|---|---|---|---|---|
| Total incidents | 300 | 300 | 300 | 300 |
| Total false alarms | 416 | 186 | 303 | 456 |
| Total detections | 201 | 253 | 230 | 186 |
| Detection rate | 67.0 | 84.3 | 76.7 | 62.0 |
| False alarm rate | 67.4 | 42.7 | 56.8 | 71.3 |

***Table 12:*** *Result for Single Location of Detectors in 220 m and 370 m Link Length.*

| Item | 220 m | 370 m |
|---|---|---|
| Total incidents | 40 | 40 |
| Total false alarms | 0 | 40 |
| Total detections | 40 | 34 |
| DR | 100.0 | 85.0 |
| FAR | 0 | 54.0 |

## CONCLUDING REMARKS

This study has identified the factors influencing the performance of neural network incident detection model and demonstrated the influence range of those factors on the performance of such incident detection model. Although the calibrated model for optimum location of detector may yield 98.5% of detection rate and 2.9% of false alarm rate for normal traffic demand situation, its performance may be down drastically under a different setting of influencing factors. Under all optimum setting but at a low demand period, ANN model results in high FAR. Similarly, FAR is also on the higher side during high demand period although it is considerably lower than that of low demand period. Therefore, some verification aid like traffic video wall on top of AID system is required to be active during these two periods to avoid attending to false alarm cases. With two detector sets per link lane, all the link lengths under study can be covered for incident detection with acceptable results; however, highest DR and lowest FAR are achieved in case of shortest link. Recalculation, i.e., data reporting interval may be as high as three times of traffic light cycle with marginally lower DR and considerably lower FAR for increasing data interval. This will offer leverage to the AID system design in terms of selecting different data communication technology with different cost implications. Experiment with varying cycle time data shows that ANN model has a very high detection rate and low false alarm rate in cases of 90, 120 and 180 s of cycle time. However, FAR in case of 60 s cycle time is relatively higher, as frequent queuing might confuse ANN in mapping out the pattern with resulting higher false alarm rate. Experimentation with detector location suggests that detector should be placed at places outside of the consistent queuing area and in case of more than one detector per link lane they should be evenly distributed over the length outside of the queuing area. When using single detector, only the shorter link length of about 220 m can be detected very effectively; however, in case of link length increase to about 370 m with single detector FAR increases to unacceptable proportion while DR is still acceptable. Based on the average of these two length cases, it may be suggested that effective detectable range for a detector is about 150 m.

Detailed understanding of the factors influencing the performance of incident detection model might be useful in designing the AID system under a particular road geometry and traffic control arrangements. Also the same knowledge base may be helpful in incident assessment and subsequent incident management planning. While this study has investigated the case of ANN model, similar detailed investigations should be carried out for other AID models and algorithms.

## ACKNOWLEDGEMENTS

The authors acknowledge their sincere gratitude to the MITSIM research group of Malaysia University of Science and Technology (MUST) from which the calibrated simulation model of Golden Triangle, KL network was made available for this research. Thanks are also due to Kuala Lumpur City Hall and ITS consortium Bhd of Malaysia for supporting MITSIM project and facilitating the MITSIMlab.

## REFERENCES

1. Stephanedes YJ, Chassiakos. Application of filtering techniques for incident detection. *Journal of Transportation Engineering*. January/February1993; 119: 1: 13–26p.
2. Payne HJ, Helfenbein ED, Knobel HC. Development and testing of incident-detection algorithms: Vol. 2, research

methodology and detailed results. *Technical Report* FHWA-RD-76-12, Federal Highway Administration, Washington, DC. 1976.
3. Payne HJ, Tignor SC. Freeway incident-detection algorithms based on decision tree with states. *Transportation Research Record.* 1978; 682: 30–37p.
4. Tsai J, Case R. Development of freeway incident-detection algorithms by using pattern-recognition techniques. *Transp. Res. Rec.* 1979; 722 (1979): 113–6p.
5. Abdulhai B, Ritchie SG. Development of a universally transferable freeway incident detection framework. *Proceedings of Transportation Research Board*, 76th Annual Meeting, Washington, DC. 1997.
6. Solomon M. A review of automatic incident detection techniques. *Advance Program Technical Report*, Transportation Center, Northwestern University. 1991.
7. Ritchie SG, Cheu RL. Simulation of freeway incident detection using artificial neural networks. *Transportation Research*. 1993; C1 (3), 203–17p.
8. Hsiao CH, Lin CT, Michael C. Application of fuzzy logic and neural networks to automatically detect freeway traffic incidents. *J. Transp. Engg.* 1994; 120(5): 753–72p.
9. Cheu RL, Ritchie SG, Automated detection of lane blocking freeway incidents using artificial neural networks. *Transportation Research C.* 1995; 3(6): 371–88p.
10. Hussein D, Rose G. Development and evaluation of neural network freeway incident detection models using field data. *Transportation Research C.* 1997; 5: 313–31p.
11. Luk JYK, Chung ECS. Characterization of incidents on an urban arterial road. *Journal of Advanced Transportation*. 2001; 35(1): 67–92p.
12. William SM, Koppelman FS. Use of vehicle positioning data for arterial incident detection." *Transportation Research C.* 1996; 4(2): 87–96p.
13. Balke KN. An evaluation of existing incident detection algorithms. *Technical Report.* 1993.
14. Sethi V, Bhandari FS, Koppelman, et al. Arterial incident detection using fixed detector and probe vehicle data. *Transportation Research C.* 1995; 3(2): 99–112p.
15. Hawas YE. A Fuzzy-based system for incident detection in urban street network. *J. of Trans Res. C*. 2007.
16. Ben-Akiva M, David C, Masroor H, et al. Evaluation of freeway control using a microscopic simulation laboratory. *Transportation Research C.* 2003; 11: 29–60p.
17. Masroor H, Mithilesh J, Ben-Akiva M. Evaluation of ramp control algorithms using microscopic traffic simulation. *Transportation Research C.* 2002; 10: 229–56p.
18. Yang Q, Koutsopoulos H. A microscopic traffic simulator for evaluation of dynamic traffic management systems. *Transportation Research C.* 1996; 4(3): 113–29p.