Transportmetrica, Vol. 2, No. 1 (2006), 11-29

A ROLLING-TRAINED FUZZY NEURAL NETWORK APPROACH FOR FREEWAY INCIDENT DETECTION

LAWRENCE W. LAN[1] AND YEH-CHIEH HUANG[2]

Received 23 March 2005; received in revised form 6 May 2005; accepted 10 May 2005

This paper develops a rolling-trained fuzzy neural network (RTFNN) approach for freeway incident detection. The core logic of this approach is to establish a fuzzy neural network and to update the network parameters in response to the prevailing traffic conditions through a rolling-trained procedure. The simulation results of some thirty-six incident scenarios in a two-lane freeway mainline case study show that the proposed RTFNN approach can improve the detection performance over the fuzzy neural network approach, which is based on the same network structure but without updating the parameters through a rolling-trained procedure. The highest detection rate is found at a rolling horizon of 45 minutes and a training sample size of 90 samples in this case study.

KEYWORDS: Freeway incident detection, fuzzy neural network, rolling-trained fuzzy neural network

1. INTRODUCTION

In the past decades, substantial research has been done in developing automatic incident detection (AID) algorithms to diagnose roadway traffic incidents (for instance, Cook and Cleveland, 1974; Dudek et al., 1974; Levin and Krause, 1978; Payne and Tignor, 1978; Willsky et al., 1980; Ahmed and Cook, 1982; Persaud and Hall, 1989; Steed and Clowes, 1989; Luk and Sin, 1992; Stephanedes et al., 1992; Ivan et al., 1993; Parkany and Bernstein, 1993; Ritchie and Cheu, 1993; Stephanedes and Chassiakos, 1993; Hsiao et al., 1994; Cheu and Ritchie, 1995; Sethi et al., 1995; Stephanedes and Liu, 1995; Dia and Rose, 1997; Lee et al., 1998; Lin and Chang, 1998; Sheu and Ritchie, 1998; Xu et al., 1998; Srinivasan et al., 2000; Lan et al., 2003b; Sheu, 2004). The detection performance in terms of detection rate and false alarm could be sensitive to the chosen traffic parameters, their designated criteria for judging the incident occurrence, and the detection locations. It can also be sensitive to the changes in prevailing traffic conditions. In practice, the complexity of traffic dynamics is characterized with uncertain and nonlinear nature. Most previous AID algorithms, however, subjectively set the parameters and use crisp criteria in distinguishing the abnormal traffic (incident-occurrence) from the normal one (incident-free), thus they may result in poor detection performance as the traffic conditions alter drastically.

In dealing with the uncertain contexts (unclear input-output relationships or imprecise input values), both neural networks (NN) and fuzzy systems (FS) have been proven as powerful tools. NN generally represents a complex system with precise inputs and outputs used for training the generic model to formulate a good approximation of the unclear relationship. FS, in contrast, addresses the imprecision of the input and output variables (often defined with fuzzy numbers) but their interrelationships take the form of well-defined if-then rules (Tsoukalas and Uhrig, 1997). Each of these two tools has its own advantages and disadvantages. For instance, the NN approaches have the advantages of learning capability to avoid subjectively setting of the parameters and

[1] Institute of Traffic and Transportation, National Chiao Tung University, 4F, 114 Section 1, Chung Hsiao W. Rd., Taipei, Taiwan 10012. Corresponding author (E-mail: lawrencelan@mail.nctu.edu.tw).

[2] Institute of Traffic and Transportation, National Chiao Tung University, 4F, 114 Section 1, Chung Hsiao W. Rd., Taipei, Taiwan 10012.

possessing high fault tolerance due to the distributed memory of parameters separately stored on each link of the network. However, NN approaches usually require long training time, especially when such network parameters as training rate, momentum and initial weights are not appropriately chosen (Wasserman, 1993; Shepherd, 1997). This may preclude the online training procedure for some advanced applications where real-time adjustments are required. The crisp criteria to judge for any event occurrence may be too sensitive, leading the NN approaches to misjudge easily.

Taking traffic incident detection as an example, the distributed memory of parameters separately stored on each link of a NN will have the advantage of high fault tolerance. Consequently, reducing the number of input nodes or poor quality of few input data will not remarkably influence the output results. However, the sensitivity of crisp thresholds or criteria used in NN to judge if an incident occurs or not may lead to poor detection performance. Incorporating fuzzy inference into the NN (called FNN) may avoid the sensitivity problem but still keeps the good features of self-learning capability with high fault tolerance. Consequently, FNN approaches have been commonly employed in traffic engineering, ranging from pavement diagnosis (Lan and Chiou, 1997), vehicular count and classification (Lan and Kuo, 2002; Lan et al., 2003a) to traffic prediction (Abdulhai et al., 2002; Yin et al., 2002). More recently, Lan et al. (2004) developed incident detection algorithms with various FNN structures based on the averages of traffic parameters across all lanes. Off-line tests have validated that their proposed FNN system was capable of detecting the freeway incidents with rather high accuracy. Sensitivity analysis further showed that alternating the FNN structures by reducing the number of detectors or number of input traffic parameters only slightly deteriorated the detection performance, implying the high fault tolerance of the FNN incident detection system. However, their FNN approach did not adaptively adjust the network parameters in response to the prevailing traffic conditions, leaving some room for improvement.

To capture the change in traffic dynamics through network training, Yin et al. (2002) developed a FNN-type model with online rolling-trained procedure to predict the traffic flows in an urban street network. Their FNN model consists of two modules: a gate network and an expert network. The gate network classifies the inputs into several clusters using a fuzzy approach and the expert network specifies the input-output relationship as in a conventional NN approach. Both simulation and real observation data demonstrated that the prediction power can be enhanced through the online rolling-trained procedure in response to the prevailing traffic conditions. Jin et al. (2002) developed constructive probabilistic neural network (CPNN) model to detect the freeway incidents. They found that the CPNN approach has three main advantages over conventional basic probabilistic neural network (BPNN) approach: (1) CPNN has clustering ability and thus could achieve similarly good incident-detection performance with a much smaller network size; (2) each Gaussian component in CPNN has its own smoothing parameter that can be obtained by the dynamic decay adjustment algorithm with a few epochs of training; and (3) the CPNN adaptation methods have the ability to prune obsolete Gaussian components and therefore the size of the network is always within control. The logics of dynamic updating and network pruning of CPNN are similar to rolling-trained procedure which can capture the change in traffic dynamics through network training.

Inspired by Yin et al. (2002) and Jin et al. (2002) recent works, this study presumes that the rolling-trained procedure in FNN might be imperative in augmenting the incident detection performance. Thus, the present paper attempts to develop a rolling-trained fuzzy neural network (RTFNN) approach for freeway incident detection. Its

underlying logic is to establish a proper fuzzy neural network and then adaptively adjust the network parameters using the most up-to-date traffic data in response to the prevailing traffic conditions so as to improve the detection performance over the conventional FNN approach.

Section 2 details the proposed fuzzy neural network structure. Section 3 introduces the rolling-trained procedure for the RTFNN approach. Section 4 conducts a case study with 36 experimental incident scenarios and further undertakes sensitivity analyses by varying the rolling horizons, training sample sizes and a combination of both. Section 5 summarizes the findings and proposes future research directions.

2. THE FUZZY NEURAL NETWORK STRUCTURE

The FNN structure of the proposed RTFNN approach is established with four layers (Figure 1). The first layer is the input layer, which processes all of the traffic flow information. The second layer is the membership layer, which processes the original traffic flow data through the corresponding relationship of membership functions and calculates its fuzzy membership. The third layer is the rule layer composed of three categories of fuzzy inference rules: time-specific, lane-specific and space-specific. The fourth layer is the output layer. The components of each layer and their relationships are detailed as follows.

2.1 The first layer

Twenty-four nodes are designed in this layer to input the lane-specific traffic parameters at upstream and downstream detectors. These nodes represent the speeds of previous time step (S_{u10}^1) and current time step (S_{u11}^1), flows (F_{u10}^1 and F_{u11}^1) and densities (D_{u10}^1 and D_{u11}^1) of the upstream inner lane, speeds (S_{u20}^1, S_{u21}^1), flows (F_{u20}^1, F_{u21}^1) and densities (D_{u20}^1, D_{u21}^1) of the upstream outer lane, speeds (S_{d10}^1, S_{d11}^1), flows (F_{d10}^1, F_{d11}^1) and densities (D_{d10}^1, D_{d11}^1) of the downstream inner lane, and speeds (S_{d20}^1, S_{d21}^1), flows (F_{d20}^1, F_{d21}^1) and densities (D_{d20}^1, D_{d21}^1) of the downstream outer lane. Note that the above densities are not directly measured from the detectors, but indirectly calculated from the detected flows and speeds.

The weighted values w_i in this layer are set equal to one and there is no need for adjustment. The output values o_i are expressed as:

$$o_i = f(u_i) = u_i, \quad i = 1 \sim 24 , \tag{1}$$

$$o_i = x_i \cdot w_i = x_i = u_i , \tag{2}$$

where u_i are the input values.

2.2 The second layer

A trapezoid membership function as shown in Figure 2 is used. The nodes in the second layer fall into three categories. The time-specific category (12 nodes) compares the upstream and downstream lane speeds, flows and densities at present time with those at the previous time step (upstream: S_{u1}^2, F_{u1}^2, D_{u1}^2, S_{u2}^2, F_{u2}^2, D_{u2}^2; downstream: S_{d1}^2, F_{d1}^2, D_{d1}^2, S_{d2}^2, F_{d2}^2, D_{d2}^2). The lane-specific category (6 nodes) calculates the membership degrees of the difference of speeds, flows and densities between upstream lanes (S_{u12}^2, F_{u12}^2, D_{u12}^2) and downstream lanes (S_{d12}^2, F_{d12}^2, D_{d12}^2). The space-specific

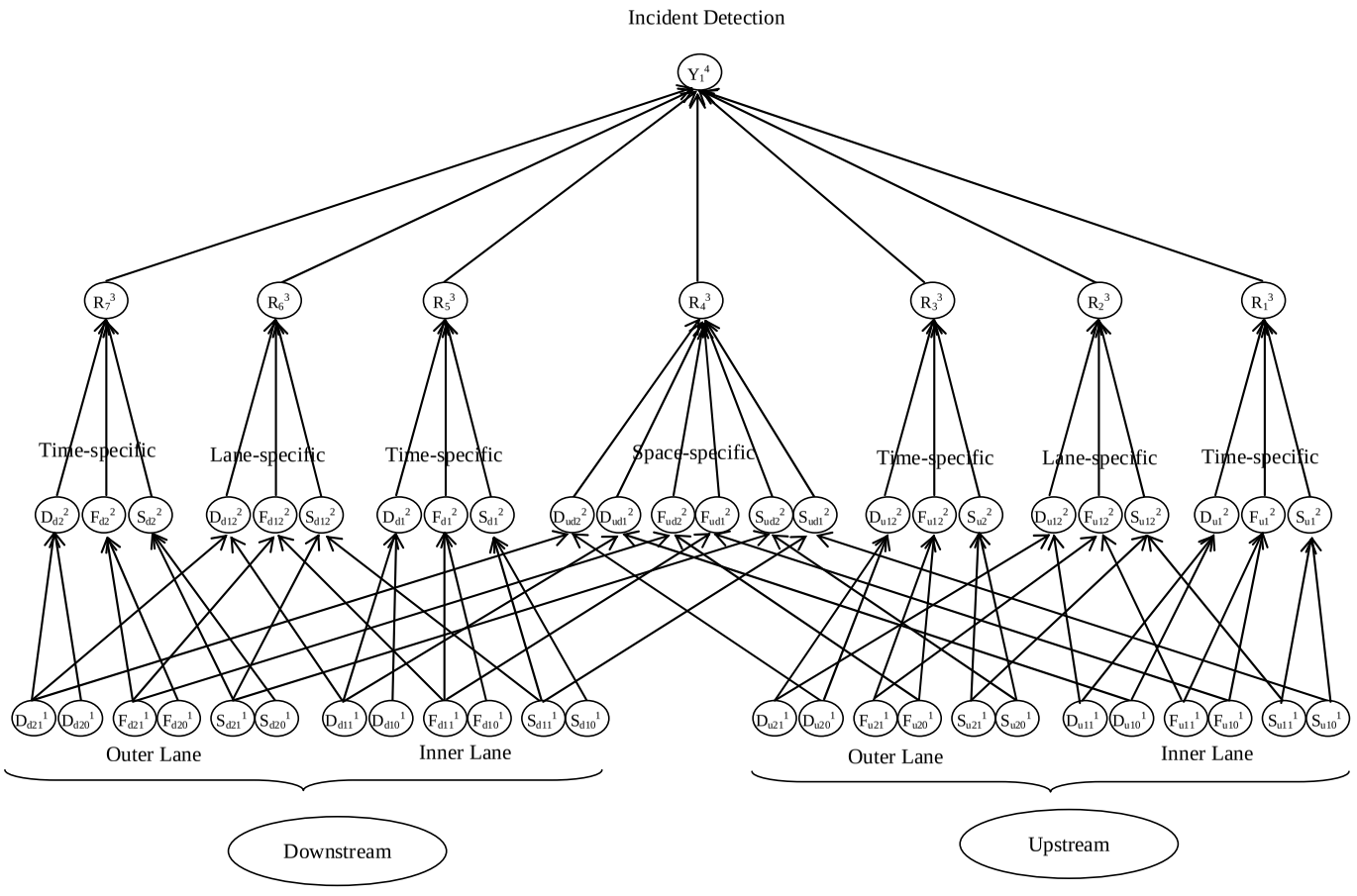

FIGURE 1: The fuzzy neural network (FNN) structure

category (6 nodes) compares the flows, speeds and densities between upstream at time t and downstream at time $t+\tau$, where τ is the time lag measured by the time for vehicles traveling from upstream detecting point to downstream detecting point. These space-specific nodes calculate the membership degrees of the difference of speeds, flows and densities between upstream and downstream (inner lane: S_{ud1}^2, F_{ud1}^2, D_{ud1}^2; outer lane: S_{ud2}^2, F_{ud2}^2, D_{ud2}^2).

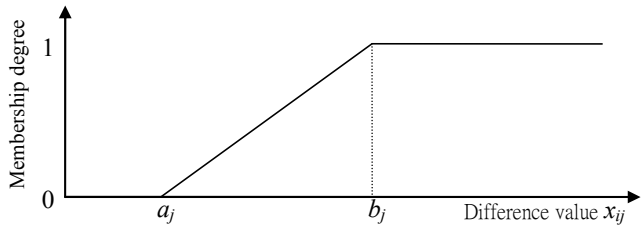

FIGURE 2: The membership function

The weighted values in the second layer w_{ij} are also set equal to one and there is no need for further adjustment. Both a_j and b_j are parameters of the trapezoid membership function, whose output values o_j can be written as:

$$o_j = f_j(u_j) = \mu_j(x_{ij}) = \begin{cases} 0 & \text{for } x_{ij} \le a_j \\ \dfrac{x_{ij} - a_j}{b_j - a_j} & \text{for } a_j < x_{ij} \le b_j \\ 1 & \text{for } x_{ij} > b_j \end{cases} . \tag{3}$$

In the time-specific category $x_{ij} = |u_i - u_{i+1}|$, where i=1,3,5 and j=1~3 for upstream inner lane, i=7,9,11 and j=7~9 for upstream outer lane, i=13,15,17 and j=16~18 for downstream inner lane, and i=19,21,23 and j=22~24 for downstream outer lane. In the lane-specific category $x_{ij} = |u_i - u_{i+6}|$, where i=2,4,6 and j=4~6 for upstream, i=14,16,18 and j=19~21 for downstream. In the space-specific category $x_{ij} = |u_i - u_{i+13}|$, where i=1,3,5 and j=10,12,14 for inner lane, i=7,9,11 and j=11,13,15 for outer lane.

2.3 The third layer

The rules in the time-specific category are stated as: IF there is a remarkable difference of speeds, flows or densities between the present time step and the previous one, upstream or downstream, inner lane or outer lane, THEN an incident occurrence is inferred with membership degrees R_1^3 (upstream inner lane), R_3^3 (upstream outer lane), R_5^3 (downstream inner lane), and R_7^3 (downstream outer lane). The output values o_k can be expressed as:

$$o_k = f(u_k) = (w_{jk} \cdot x_{jk}) * (w_{(j+1)k} \cdot x_{(j+1)k}) * (w_{(j+2)k} \cdot x_{(j+2)k}), \tag{4}$$

where j=1 and k=1 for upstream inner lane, j=7 and k=3 for upstream outer lane, j=16 and k=5 for downstream inner lane, j=22 and k=7 for downstream outer lane.

The rules in the lane-specific category are stated as: IF there is a remarkable difference of speeds, flows or densities between the inner lane and the outer lane, upstream or

downstream, THEN an incident occurrence is inferred with membership degrees R_2^3 (upstream) and R_6^3 (downstream). The output values o_k are represented as:

$$o_k = f(u_k) = (w_{jk} \cdot x_{jk}) * (w_{(j+1)k} \cdot x_{(j+1)k}) * (w_{(j+2)k} \cdot x_{(j+2)k}) \tag{5}$$

for upstream: j=4 and k=2; downstream: j=19 and k=6.

The rules in the space-specific category are stated as: IF there is a remarkable difference of speeds, flows or densities between the upstream at time t and downstream at time $t+\tau$, inner lane or outer lane, THEN an incident occurrence is inferred with membership degrees R_4^3. The output values o_k can be expressed as:

$$\begin{aligned} o_k = f(u_k) = {} & (w_{jk} \cdot x_{jk}) * (w_{(j+1)k} \cdot x_{(j+1)k}) * (w_{(j+2)k} \cdot x_{(j+2)k}) \\ & * (w_{(j+3)k} \cdot x_{(j+3)k}) * (w_{(j+4)k} \cdot x_{(j+4)k}) * (w_{(j+5)k} \cdot x_{(j+5)k}) \end{aligned} \tag{6}$$

for j=10 and k=4.

2.4 The fourth layer

The fourth layer is the output layer, which contains one node Y_1^4 for this two-lane freeway. The center of area method is employed to defuzzify the fuzzy number to a crisp binary value (Y_1^4=0 indicates incident-free; Y_1^4=1 represents incident-occurrence). It is essential to set the initial weighted values for this layer w_{km} and then adjust them through the network training. The output values o_m are:

$$o_m = f(u_m) = \sum_{k=1}^{7} w_{km} \cdot x_{km} \tag{7}$$

3. THE ROLLING-TRAINED PROCEDURE

A backpropagation technique that minimizes the total error function with gradient steepest descent method is used for the network training. To capture the fluctuations of traffic, we further develop a rolling-trained procedure. The most up-to-date flow parameters are used to distinguish the traffic characteristics of one time interval from another. The proposed rolling-trained procedure is depicted in Figure 3 and detailed as follows.

Step 1: Gather the initial training data.

Step 2: Train and update the network parameters by backpropagation algorithm.
The backpropagation algorithm (see Appendix 1) may not have appropriate initial values, thus its network parameters need to be updated through the initial training process.

Step 3: Collect new training data.
Gather input data and conduct incident detection through the FNN algorithm. In the meantime, save both input data and output results in the training sample dataset for the follow-up training.

Step 4: Verify the incident by persistence tests.
As an incident can last for a while, it is necessary to conduct the persistence tests to avoid including the incorrect (misjudged) data in the training dataset. The underlying philosophy of a persistence test is that if there are no continuous detections of an incident occurrence, then that detected incident should be considered as a false alarm. In this case we should discard the training sample and go back to Step 3.

Step 5: Update the training dataset.
Put into the training dataset the data that have passed the persistence tests in Step 4. Replace the most distant data with the most updated data so that the training dataset can be maintained at a predetermined training sample size. Go back to Step 2 every fixed time interval (e.g., every hour; hereafter, rolling horizon) to keep renewing the network parameters to the latest traffic conditions. The idea of RTFNN is to capture the change in traffic dynamics through rolling-trained procedure so as to adaptively adjust the network parameters. Therefore, the potential advantage for fixing the training sample size is to avoid too many obsolete traffic data which may significantly differ from the prevailing traffic conditions.

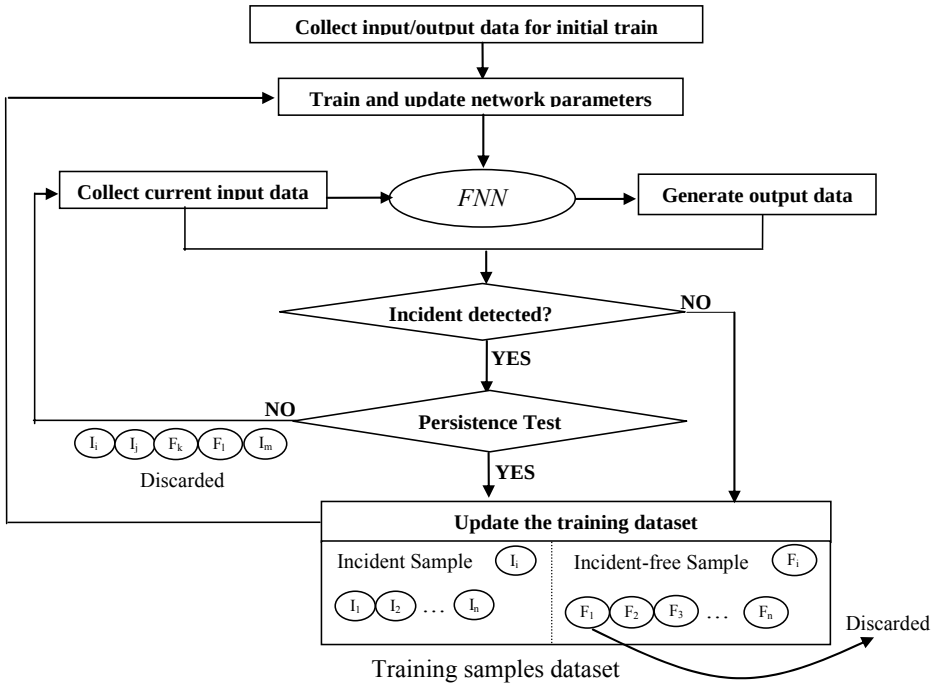

FIGURE 3: The rolling-trained procedure

4. CASE STUDY

4.1 Data

It is very difficult to generate sufficient real traffic incidents to validate any AID algorithm, thus most previous algorithms adopt off-line validations by simulation or incident database. In the present paper, a calibrated Paramics carried out on the Taiwan Freeway two-lane mainline is employed to produce enough training samples and off-line test samples required by the RTFNN and FNN algorithms. The calibration is briefly narrated in Appendix 2. Thirty-second traffic flow data are collected from 6:00 am to 12:00 pm at the same site where the deliberate incident was experimented. The data cover a typical morning peak hours and two off-peak periods before and after that peak.

Figure 4 presents the 30-second flow variations for six hours, in which conspicuous fluctuations and manifest changes in traffic volume (from low to intermediate to heavy volumes and then from heavy to intermediate to low volumes) are found. Such drastic variations suggest the necessity of renewing the network parameters in response to the most up-to-date traffic conditions through the rolling trained procedure.

The experimental design of the following simulation places two detection points (upstream and downstream), one kilometer apart, in this two-lane freeway mainline. In between, various incident scenarios are generated, which include 36 incidents taking place in different lanes (inner and outer) at different locations (250, 500 and 750 meters from the upstream detectors) under six different traffic conditions (from light to intermediate to heavy traffic and then from heavy to intermediate to light traffic). As an example, Figure 5 demonstrates the 30-second flow variations of both incident-free and incident conditions for one scenario where the incident occurs at 7:30am and lasts for 15 minutes. In each scenario, Paramics generates 200 sets of simulation data--100 of which are used for network training (called training sets) and the rest 100 sets are used for off-line validation (called evaluation sets).

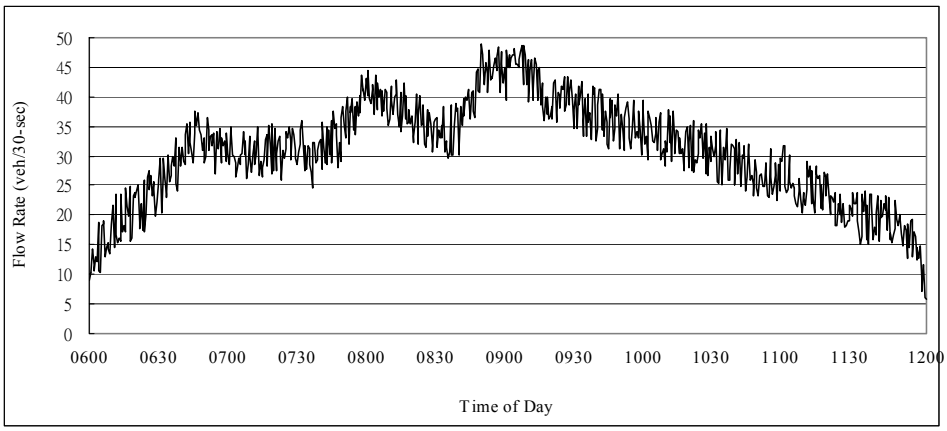

FIGURE 4: Observed 30-second flow rates on the studied freeway (two-lane mainline)

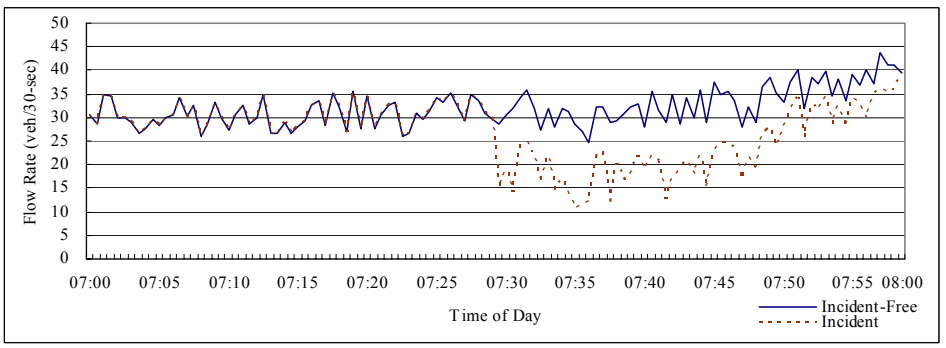

FIGURE 5: Flow variation for one incident scenario (incident duration 7:30–7:45 am)

4.2 Results

The detection performance is evaluated by three criteria: detection rate (DR), false alarm rate (FAR), and time to detection (TTD). DR is defined as a ratio of the number of detected incidents to the number of actual incidents for the overall lanes of flow direction. FAR is defined as a ratio of the number of detected incidents to the number of incident-free for the overall lanes of flow direction. TTD is the difference between the time an incident being detected and the time that incident actually taking place.

As a base for comparison, we set the rolling horizon with 60 minutes and 30-second traffic data as one sample, thus the training sample size is 120 samples. Table 1 and Figure 6 present the detection performance, based on the 100 evaluation sets, of 36 incident scenarios by two approaches--with rolling-trained (hereafter referred as RTFNN approach) and without rolling-trained (hereafter referred as FNN approach). Note that the six data rows in Table 1 and six points in Figure 6 represent six different incident locations within the same simulation hours. Initially, both RTFNN and FNN network parameters are based on the same trained results using all six-hour 100 training sets of simulation data, thus they have exactly the same detection performance in the first hour validation. However, after a few hours, RTFNN gradually outperforms over FNN because RTFNN updates the trained parameters in every 60 minutes, but FNN keeps using the initially trained parameters. From Table 1, we find that the detection performance for both approaches consistently depend on the location of the incident. In general, if the incident takes place near the detector, either upstream (the 250-meter scenarios) or downstream (the 750-meter scenarios), the DR is higher and the TTD is shorter than the one occurring farther away from the detector (the 500-meter scenarios). Figure 6 also demonstrates that the RTFNN approach has outperformed with higher DR, lower FAR and shorter TTD, compared with the FNN approach in various traffic conditions.

Table 2 reports the statistical difference of mean values (t-test) of detection performance between these two approaches. It is found that the overall DR for RTFNN is 93.95% and for FNN is 91.09%; both are quite high and have statistical difference at 5% significance level. The overall FAR for RTFNN is 0.0754% and for FNN is 0.0803%; both are quite low but have no significant difference. The overall TTD requires only about two time steps, 68.39 seconds for RTFNN and 74.19 seconds for FNN; both are statistically different. The high detection performance suggests that both FNN and RTFNN approaches are all satisfactory in freeway incident detections; but through the rolling-trained, the detection performance can be significantly enhanced. Specifically, as the traffic conditions changed from low to high and then from high to low, the detection performance (DR and TTD) for RTFNN is increased from 90% to about 96%; but the detection performance for FNN remains rather stable between 90% and 92%. As for the FAR, the overall performance shows that there is no significant difference between these two approaches (Table 2). Figure 7 further presents the interaction between DR and FAR for both approaches. In sum, the enhancement of DR (and TTD) without significantly deteriorating the FAR and the superior performance of RTFNN over FNN should be ascribed to the rolling-trained effects of adaptively adjusting the network parameters in response to the traffic variations.

TABLE 1: Detection performance for 36 incident scenarios (rolling horizon = 60 minutes, training sample size = 120)

Simulation hour (time of day)	Hourly volume (vph)	Incident location		RTFNN approach			FNN approach		
		Lane position	Distance from upstream detector (meter)	DR (%)	FAR (%)	TTD (sec)	DR (%)	FAR (%)	TTD (sec)
1 (6:00-7:00)	2,403	Inner	250	92.78	0.08	73	92.78	0.08	73
			500	89.68	0.05	88	89.68	0.05	88
			750	91.11	0.09	74	91.11	0.09	74
		Outer	250	88.98	0.08	79	88.98	0.08	79
			500	87.05	0.07	83	87.05	0.07	83
			750	90.62	0.08	76	90.62	0.08	76
2 (7:00-8:00)	2,919	Inner	250	94.70	0.08	74	91.48	0.08	77
			500	90.01	0.07	78	89.05	0.06	81
			750	93.89	0.09	72	92.97	0.08	75
		Outer	250	92.56	0.08	71	91.77	0.07	76
			500	91.17	0.06	80	90.12	0.06	85
			750	93.58	0.08	73	91.52	0.07	77
3 (8:00-9:00)	3,664	Inner	250	95.66	0.09	66	92.56	0.07	73
			500	92.71	0.09	72	90.12	0.07	80
			750	94.76	0.08	61	92.36	0.08	68
		Outer	250	94.33	0.08	65	91.91	0.09	71
			500	91.58	0.07	69	89.94	0.07	79
			750	92.98	0.09	66	91.29	0.08	75
4 (9:00-10:00)	4,514	Inner	250	95.78	0.06	66	91.35	0.09	73
			500	92.77	0.07	65	87.32	0.07	72
			750	95.02	0.07	67	91.58	0.09	75
		Outer	250	96.53	0.06	68	92.41	0.08	75
			500	93.26	0.05	62	87.74	0.08	70
			750	94.97	0.08	67	91.23	0.09	74
5 (10:00-11:00)	3,310	Inner	250	98.12	0.08	63	91.76	0.09	70
			500	95.37	0.06	63	89.43	0.08	70
			750	97.81	0.08	64	91.67	0.09	72
		Outer	250	96.59	0.07	66	92.06	0.08	73
			500	94.83	0.07	61	88.61	0.08	67
			750	97.98	0.08	65	93.16	0.09	73
6 (11:00-12:00)	2,484	Inner	250	96.89	0.07	62	93.61	0.09	75
			500	94.31	0.05	62	90.47	0.07	76
			750	97.00	0.06	64	92.96	0.09	69
		Outer	250	96.28	0.08	64	94.21	0.08	75
			500	93.73	0.06	60	90.36	0.08	73
			750	96.69	0.07	59	93.75	0.09	72

Note: 1. The results for RTFNN and FNN approaches for the first-hour simulation are the same as it is the initial condition.
2. Distance of incident location is measured from the upstream detecting point.
3. Each scenario is simulated for 100 times. The values in this table are the average of 100 simulation runs.

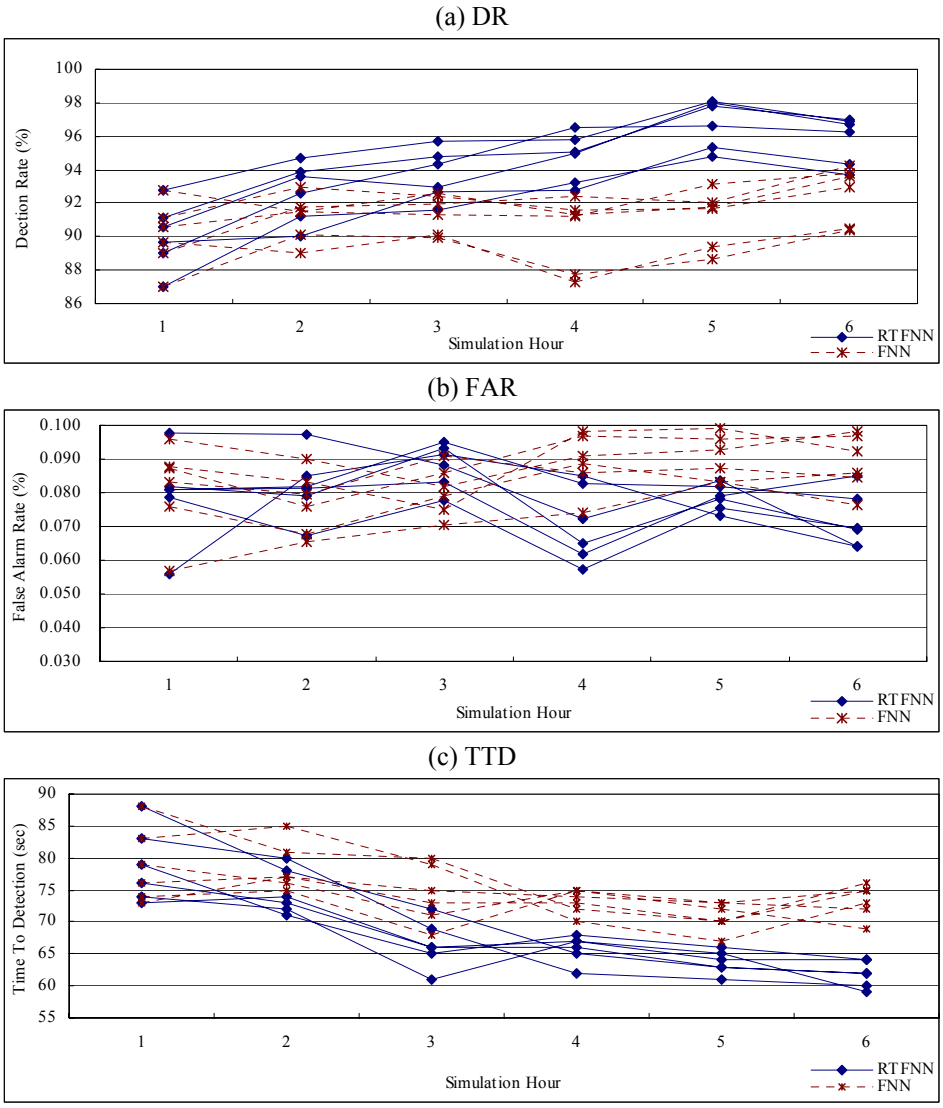

FIGURE 6: Comparison of detection performance for each simulation hour between RTFNN and FNN approaches (rolling horizon = 60 minutes, training sample size = 120)

4.3 Sensitivities of rolling horizon and training sample size

The above-mentioned results conclude that the RTFNN approach, based on the rolling horizon of 60 minutes and the training sample size of 120 samples, can improve the detection performance over the conventional FNN approach. One might wonder if there still exists some room for improvement of detection performance by changing the rolling horizons and/or training sample sizes. Thus, the following sensitivity analyses are further undertaken: case (I) altering the rolling horizons from 15, 20, 30, 45, 90 to 120 minutes, provided that the training sample size is remained as 120; case (II) altering the training sample sizes from 30, 40, 50, 60, 180 to 240, provided that the rolling horizon is

TABLE 2: Test for the difference of detection performance between RTFNN and FNN approaches (rolling horizon = 60 minutes, training sample size = 120)

Simulation hour (time of day)	Hourly volume (vph)	Detection approaches	DR		FAR		TTD	
			Avg. (%)[1]	Test result[2]	Avg. (%)[1]	Test result[2]	Avg. (sec)[1]	Test result[2]
1	2,403	RTFNN	90.04	Same	0.0794	Same	78.83	Same
(6:00-7:00)		FNN	90.04		0.0794		78.83	
2	2,919	RTFNN	92.65	NSD	0.0815	NSD	74.67	SD
(7:00-8:00)		FNN	91.15	(0.752)	0.0760	(0.891)	78.50	(0.046)
3	3,664	RTFNN	93.67	SD	0.0881	NSD	66.50	SD
(8:00-9:00)		FNN	91.36	(0.021)	0.0795	(0.701)	74.33	(0.009)
4	4,514	RTFNN	94.72	SD	0.0658	SD	65.99	SD
(9:00-10:00)		FNN	90.27	(0.003)	0.0801	(0.051)	73.36	(0.024)
5	3,310	RTFNN	96.78	SD	0.0742	NSD	62.62	SD
(10:00-11:00)		FNN	91.19	(0.015)	0.0807	(0.053)	71.19	(0.037)
6	2,484	RTFNN	95.82	SD	0.0634	SD	61.75	SD
(11:00-12:00)		FNN	92.55	(0.011)	0.0862	(0.027)	73.27	(0.049)
Overall		RTFNN	93.95	SD	0.0754	NSD	68.39	SD
		FNN	91.09	(0.033)	0.0803	(0.092)	74.19	(0.008)

Note: 1. The results for RTFNN and FNN approaches for the first-hour simulation are the same as it is the initial condition. Average represents the mean values of six incident scenarios, each of which undertakes 100 simulation runs.

2. NSD represents no significant difference and SD represents significant difference with P-value in parenthesis (α=0.05). The null hypothesis is that the mean values (DR, FAR, or TTD) between two approaches are the same.

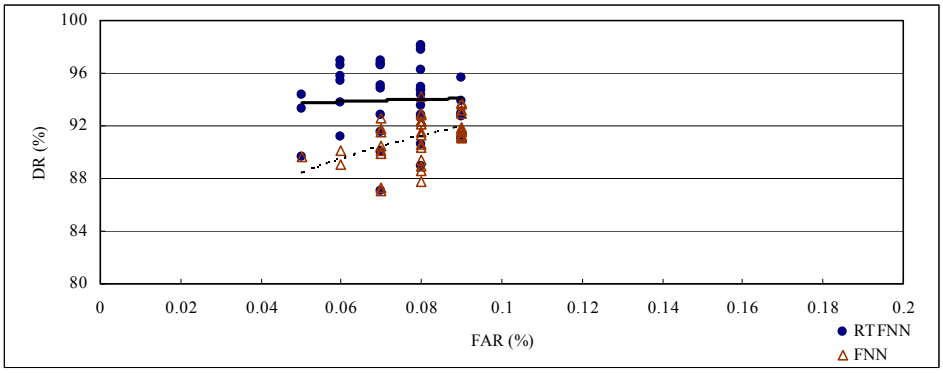

FIGURE 7: Graph of detection rate vs. false alarm rate for 36 incident scenarios

remained as 60 minutes; case (III) simultaneously altering both rolling horizons and training sample sizes. Figures 8, 9 and 10 respectively present the change in detection rates for these three cases and Tables 3, 4 and 5 respectively report the details of the change. The sensitivity analyses of these three cases consistently show that the highest average detection rate is at the 45-minute rolling horizon and 90 training sample sizes in this case study.

It is interesting to note that very short or very long rolling horizons can lower the detection rates, compared with the base with rolling horizon of 60 minutes. The main

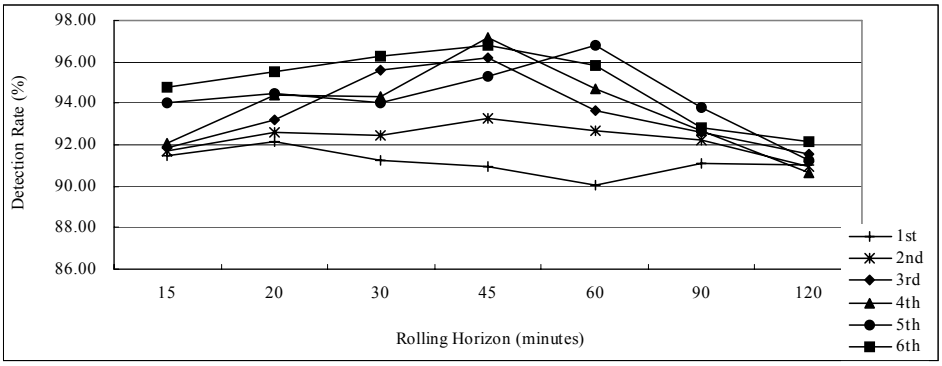

FIGURE 8: Detection rates in each simulation hour for Case (I) (training sample size fixed at 120)

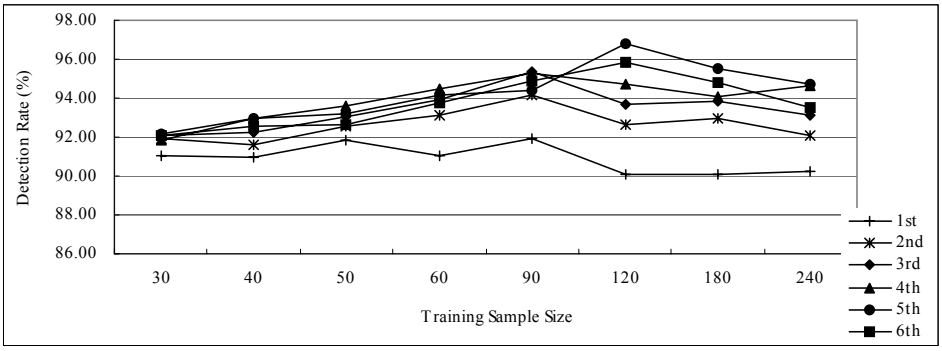

FIGURE 9: Detection rates in each simulation hour for Case (II) (rolling horizon fixed at 60 minutes)

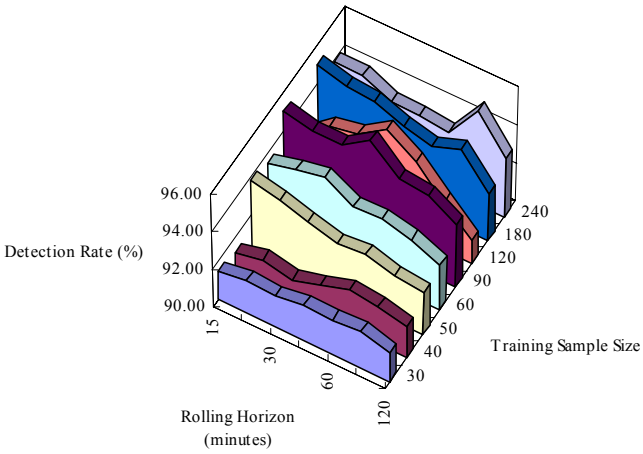

FIGURE 10: Average detection rates of 36 incident scenarios for Case (III) (rolling horizon and training sample size varied)

TABLE 3: Detection rates for Case (I) (training sample size fixed at 120)

Simulation hours	Hourly flow (veh/hr)	Rolling horizons (minutes)						
		15	20	30	45	60	90	120
1	2,403	91.50	92.13	91.28	90.96	90.04	91.08	91.02
2	2,919	91.68	92.61	92.47	93.24	92.65	92.21	90.93
3	3,664	91.87	93.17	95.61	96.22	93.67	92.59	91.54
4	4,514	92.08	94.37	94.33	97.15	94.72	92.70	90.63
5	3,310	94.02	94.49	94.02	95.29	96.78	93.79	91.26
6	2,484	94.74	95.56	96.25	96.81	95.82	92.84	92.15

Note: shadow indicates the base condition of rolling horizon

TABLE 4: Detection rates for Case (II) (rolling horizon fixed at 60 minutes)

Simulation hours	Hourly flow (veh/hr)	Training sample sizes							
		30	40	50	60	90	120	180	240
1	2,403	91.03	90.93	91.86	91.02	91.89	90.04	90.12	90.23
2	2,919	91.95	91.64	92.55	93.11	94.17	92.65	92.97	92.10
3	3,664	92.09	92.27	93.07	93.95	95.32	93.67	93.85	93.11
4	4,514	91.81	92.96	93.59	94.48	95.28	94.72	94.07	94.66
5	3,310	92.17	92.94	93.21	94.15	94.39	96.78	95.48	94.70
6	2,484	92.06	92.56	92.61	93.74	94.90	95.82	94.78	93.51

Note: shadow indicates the base condition of training sample size

TABLE 5: Average detection rates for Case (III) (rolling horizon and training sample size varied)

Training sample size	Rolling horizon (minutes)						
	15	20	30	45	60	90	120
30	91.51	91.72	91.64	91.92	91.85	91.97	91.55
40	91.22	91.60	91.26	91.73	92.22	92.03	91.67
50	93.86	93.49	93.07	92.66	92.82	92.36	92.20
60	93.48	93.93	94.20	93.35	93.41	93.06	92.46
90	95.04	94.67	94.70	95.63	94.33	94.15	93.28
120	92.65	93.72	93.99	94.95	93.95	92.54	91.26
180	95.02	94.63	94.48	93.87	93.54	94.07	92.47
240	93.96	93.98	93.32	93.25	93.05	94.87	93.10

reasons are insufficient updated training samples would be collected if the rolling horizon is too short and less capable of capturing the flow variations for longer rolling horizons. Similarly, small training sample sizes can also lower the detection rates, compared with the base with training sample size of 120. The main reason is the difficulty in reaching the convergence of total error function, should one select a training sample size as small as 30 or 40 samples. The sensitivity analyzes also find that heavier traffic conditions tend to have higher detection rates than lighter ones, regardless of the changes in rolling horizon and/or training sample size.

5. CONCLUSION

The main advantage of the proposed RTFNN approach is to adaptively adjust the network parameters using the most up-to-date traffic data in response to the prevailing traffic conditions. The case study has shown that as the traffic volumes vary from low to high and then change from high to low, the detection performance for RTFNN is getting better (but not for the FNN approach), which can be ascribed to the rolling-trained effects of adaptively adjusting the network parameters in response to the traffic variations.

The findings are limited to some thirty-six incident scenarios in the two-lane freeway contexts. One might argue that the lane-specific traffic data used in the input layer of the neural network would limit the transferability potential of the proposed algorithm to other freeway facilities with three or more lanes. A number of studies in the literature (e.g., Ritchie and Cheu, 1993; Lan et al., 2004) have shown that using the station averages across all lanes rather than lane-specific traffic data does not substantially reduce the accuracy of the incident detection model. Of course, future study can further examine the transferability of the proposed RTFNN approach basing on the station average data.

Paramics is employed in the present paper for generating speed, flow and density data which are used for training and off-line validations. Future work can attempt other different micro traffic simulators. Additionally, the density is difficult to measure in the field, a surrogate of it, such as percent occupancy that can be readily provided by the field detectors, should be used in the proposed RTFNN approach from practical perspectives.

The robustness of RTFNN approach at different places (e.g., freeway mainline sections with three or more lanes) with different scenarios (e.g., incidents at different locations and affecting more than one lane) can also be examined. Development of new methods to determine the optimal rolling horizon and/or training sample size deserves further exploration.

ACKNOWLEDGEMENTS

This study is part of the research project granted by National Science Council of the Republic of China, under contract NSC 90-2211-E-009-058. The authors wish to thank Y.C. Chiou and April Y. Kuo who provide helpful comments on the initial work. Special thanks go to four anonymous referees who give insightful comments and constructive suggestions to improve the revised work.

REFERENCES

Abdulhai, B, Porwal, H. and Recker, W. (2002) Short-term traffic flow prediction using neuro-genetic algorithms. ITS Journal, 7, 3-41.

Ahmed, S.R. and Cook, A.R. (1982) Application of time-series analysis techniques to freeway incident detection. Transportation Research Record, 841, 19-21.

Cheu, R.L. and Ritchie, S.G. (1995) Automated detection of lane-blocking freeway incidents using artificial neural networks. Transportation Research Part C, 3, 371-388.

Cook, A.R. and Cleveland, D.E. (1974) Detection freeway capacity-reducing incidents by traffic steam measurements. Transportation Research Record, 495, 1-11.

Dia, H. and Rose, G. (1997) Development and evaluation of neural network freeway incident detection models using field data. Transportation Research Part C, 5, 313-331.

Dudek, C.L., Messer, C.J. and Nuckles, N.B. (1974) Incident detection on urban freeways. Transportation Research Record, 495, 12-24.

Hsiao, C.H., Lin, C.T. and Cassidy, M. (1994) Application of fuzzy logic and neural networks to automatically detect freeway traffic incidents. Journal of Transportation Engineering, ASCE, 120, 753-771.

Ivan, J.N., Schofer, J.L., Bhat, C.R., Liu, P.C., Koppelman, F.S. and Rodriguez, A. (1993) Arterial street incident detection using multiple data sources: plans for advance. Proceedings of the Pacific Rim Trans Tech Conference, The 3rd International

Conference on Applications of Advanced Technologies in Transportation Engineering, Seattle, Washington, pp. 429-435.

Jin, X., Cheu, R.L. and Srinivasan, D. (2002) Development and adaptation of constructive probabilistic neural network in freeway incident detection. Transportation Research Part C, 10, 121-147.

Lan, L.W. and Chiou, Y.C. (1997) Development and applications of learning algorithms for fuzzy neural networks: a case of pavement diagnosis. Transportation Planning Journal, 26, 233-252 (in Chinese).

Lan, L.W. and Kuo, A.Y. (2002) Development of a fuzzy neural network color image vehicular detection (FNNCIVD) system. 5th IEEE International Conference on Intelligent Transportation Systems, Singapore, pp. 88-93.

Lan, L.W., Kuo, A.Y. and Huang, Y.C. (2003a) Color image vehicular detection systems with and without fuzzy neural network: a comparison. Journal of the Chinese Institute of Engineers, 26, 659-670.

Lan, L.W., Lin, F.Y. and Huang, Y.C. (2003b) Diagnosis of freeway traffic incidents with chaos theory. Journal of the Eastern Asia Society for Transportation Studies, 5, 2025-2038.

Lan, L.W., Huang, Y.C. and Kuo, A.Y. (2004) Development of a fuzzy neural network (FNN) incident detection system. Journal of the Chinese Institute of Civil and Hydraulic Engineering, 16, 499-512 (in Chinese).

Lee, S., Krammes, R.A. and Yen, J. (1998) Fuzzy-logic-based incident detection for signalized diamond interchanges. Transportation Research Part C, 6, 359-377.

Levin, M. and Krause, G.M. (1978) Incident detection: a Bayesian approach, Transportation Research Record, 682, 52-58.

Lin, C.K. and Chang, G.L. (1998) Development of a fuzzy-expert system for incident detection and classification. Mathematical and Computer Modelling, 27, 9-25.

Luk, J.Y.K. and Sin, F.Y.C. (1992) The calibration of freeway incident detection algorithms. Proceedings of 7th Road Engineering Association of Asia and Australasia conference, Singapore, pp. 347-355.

Parkany, A.E. and Bernstein, D. (1993) Using VRC data for incident detection. Proceedings of the Pacific Rim Trans Tech Conference, The 3rd International Conference on Applications of Advanced Technologies in Transportation Engineering, Seattle, Washington, pp. 25-28.

Payne, H.J. and Tignor, S.C. (1978) Freeway incident detection algorithms based on decision trees with states. Transportation Research Record, 682, 30-37.

Persaud, B.N. and Hall, F.L. (1989) Catastrophe theory and patterns in 30-second freeway traffic data—implications for incident detection. Transportation Research Part A, 23, 103-113.

Ritchie, S.G. and Cheu, R.L. (1993) Simulation of freeway incident detection using artificial neural networks. Transportation Research Part C, 1, 203-217.

Sethi, V., Bhandari, N., Koppelman, F.S. and Schofer, J.L. (1995) Arterial incident detection using fixed detector and probe vehicle data, Transportation Research Part C, 3, 99-112.

Shepherd, A.J. (1997) Second-order Methods for Neural Networks: Fast and Reliable Training Methods for Multilayer Perceptrons. Springer, London.

Sheu, J.B. (2004) A sequential detection approach to real-time freeway incident detection and characterization. European Journal of Operational Research, 157, 451-485.

Sheu, J.B. and Ritchie, S.G. (1998) A new methodology for incident detection and characterization on surface streets. Transportation Research Part C, 6, 315-335.
Srinivasan, D., Cheu, R.L., Poh, Y.P. and Ng, A.K.C. (2000) Development of an intelligent technique for traffic network incident detection. Engineering Applications of Artificial Intelligence, 13, 311-322.
Steed, J.J. and Clowes, D.S. (1989) Detection of road hazards – problems and practicalities, Proceedings of 2nd International Conference on Road Traffic Monitoring, No. 299, London, pp. 161-166.
Stephanedes, Y.J. and Chassiakos, A.P. (1993) Application of filtering techniques for incident detection. Journal of Transportation Engineering, ASCE, 119, 13-26.
Stephanedes, Y.J. and Liu, X. (1995) Artificial neural networks for freeway incident detection. Transportation Research Record, 1494, 91-97.
Stephanedes, Y.J., Chassiakos, A.P. and Michalopoulos, P.G. (1992) Comparative performance evaluation of incident detection algorithms. Transportation Research Record, 1360, 50-57.
Tsoukalas, L.H. and Uhrig, R.E. (1997) Fuzzy and Neural Approaches in Engineering. John Wiley & Sons, Inc. New York.
Wasserman, P.D. (1993) Advanced Methods in Neural Computing. Van Nostrand Reinhold, New York.
Willsky, A.S., Chow, E.Y., Gershwin, S.B., Greene, C.S., Houpt, P.K. and Kurkjian, A.L. (1980) Dynamic model-based techniques for the detection of incidents on freeways. IEEE Transactions on Automatic Control, 25, 347-360.
Xu, H., Kwan, C.M., Haynes, L. and Pryor, J.D. (1998) Real-time adaptive on-line traffic incident detection. Fuzzy Sets and Systems, 93, 173-183.
Yin, H., Wong, S.C., Xu, J. and Wong, C.K. (2002) Urban traffic flow prediction using a fuzzy-neural approach. Transportation Research Part C, 10, 85-98.

APPENDIX 1. THE BACKPROPAGATION ALGORITHM

Step 1: Initialize the network parameters, including weighted value (w_{km}), membership function parameters (a_j, b_j), momentum term α, and learning rate η.

Step 2: Obtain the output value (o_m^4) with the above-mentioned FNN by inputting a training sample using the existing network parameters.

Step 3: Calculate the error of the fourth layer (δ_m^4):

$$\delta_m^4 = d_m^4 - o_m^4 , \tag{A1}$$

where d_m^4 is the observed output value of training sample.

Step 4: Update the weighted values between the third and fourth layers (w_{km}):

$$w_{km}(t+1) = w_{km}(t) + \eta \cdot \delta_m^4 \cdot x_{km} + \alpha \cdot [w_{km}(t) - w_{km}(t-1)] . \tag{A2}$$

Step 5: Calculate the errors of the third and second layers (δ_k^3, δ_j^2):

$$\delta_k^3 = \delta_m^4 \cdot w_{km}(t+1) , \tag{A3}$$

$$\delta_j^2 = \delta_k^3 \cdot w_{jk} \cdot x_{jk} . \tag{A4}$$

Step 6: Adjust the parameters of the membership function in the second layer (a_j, b_j):

$$a_j(t+1) = a_j(t) + \eta \cdot \delta_j^2 \cdot x_k \cdot \frac{x_j - b_j(t)}{[b_j(t) - a_j(t)]^2} + \alpha \cdot [a_j(t) - a_j(t-1)] , \tag{A5}$$

$$b_j(t+1) = b_j(t) + \eta \cdot \delta_j^2 \cdot x_k \cdot \frac{a_j(t) - x_j}{[b_j(t) - a_j(t)]^2} + \alpha \cdot [b_j(t) - b_j(t-1)] \,. \tag{A6}$$

Step 7: Repeat Steps 2 through 6 until all training samples have been inputted.

Step 8: Calculate the value of total error function of ith epoch (TE_i):

$$TE_i = \frac{1}{2}\sum_{t=1}^{N}[d_m(t) - o_m(t)]^2 \,, \tag{A7}$$

where N is the total number of training samples.

Step 9: Test the stop condition. Training can be terminated when the predetermined number of training epochs reaches or the total error function converges; otherwise, go to Step 2. In this paper, the later condition is used. Namely,

$$\left|TE_n - TE_{n-1}\right| \le \varepsilon \,, \tag{A8}$$

where ε is an arbitrary small number. We stop the network training as TE_i decreases smoothly.

APPENDIX 2. CALIBRATION OF PARAMICS WITH A REAL INCIDENT

Under special permission from the Taiwan Freeway Authority, Lan et al. (2004) deliberately generated a real traffic incident by placing two cars at 19K+400, northbound of Taiwan Freeway No. 1, so as to block the shoulder and outer lane of the two-lane freeway mainline section by allowing only one lane (inner) for traffic passing through. The geometry of the experimental freeway segment is schemed in Figure A1, in which no on-ramp or off-ramp exists between the nearest upstream (21K+300) and downstream (18K+400) CCTV cameras in situ. The upstream and downstream traffic image data are concurrently recorded by the video cameras, 15 minutes prior to, during, and after this experimented incident, respectively. Such recorded traffic images are then analyzed frame by frame (per 0.1 second) and converted to 30-second traffic data (speed, flow and density) and used for Paramics calibration. The origin-destination pattern, arrival distribution, mixed traffic ratio, driver familiarity, and so on are fine-tuned until the Paramics simulation results can best fit the observed data for each 15-minute period. As indicated in Table A1, the statistical tests conclude that there is no significant difference between the observed and simulated data for each 15-minute period, suggesting that the Paramics has been calibrated and can be employed for further simulation.

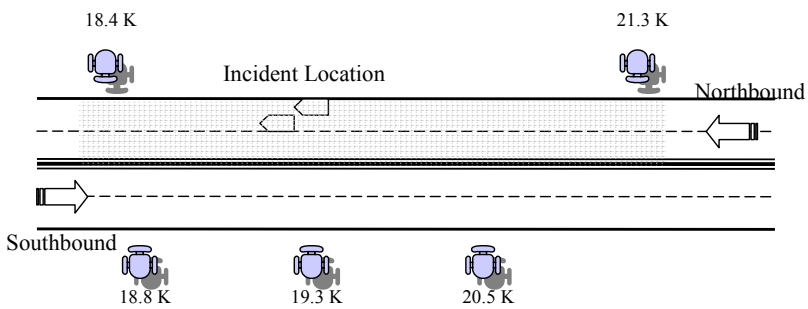

FIGURE A1: The schematic of Taiwan Freeway No. 1 (shadow indicates the experimental section where incident took place at 19.3K and traffic data were observed at downstream 18.4K and upstream 21.3K), Source: Lan et al. (2004)

TABLE A1: Test for difference between observed and Paramics simulated data

Time	Parameter[1]	Observed	Simulated[2]	t value	P value	Test result[3]
15 mins before incident	Speed	80.47	83.70	1.3077	0.1997	No significant difference
	Flow	3,304	3,216	-1.8728	0.0869	No significant difference
	Density	20.54	19.36	-1.5811	0.1890	No significant difference
15 mins during incident	Speed	62.27	64.98	1.6912	0.1039	No significant difference
	Flow	2,656	2,732	-0.2202	0.9609	No significant difference
	Density	21.43	21.02	0.1414	0.8884	No significant difference
15 mins after incident	Speed	68.23	72.16	1.7797	0.0892	No significant difference
	Flow	3,272	3,192	-1.5624	0.1091	No significant difference
	Density	23.95	22.12	-1.0389	0.1563	No significant difference
Total experiment 45 mins	Speed	70.30	73.61	1.6867	0.1080	No significant difference
	Flow	3,078	2,964	-0.0822	0.9354	No significant difference
	Density	21.97	20.11	-1.4554	0.1619	No significant difference

Note: (1) The data represent average values of downstream traffic flow parameters. Unit of speed: kilometer/hour, flow: vehicle/hour and density: vehicle/kilometer.
(2) Paramics simulated data is the average of 10 simulation runs.
(3) Null hypothesis H_0: $\mu_{observed}=\mu_{simulated}$ and significance level $\alpha=0.05$.

Source: Lan et al. (2004)