A fuzzy clustering-based approach to automatic freeway incident detection and characterization

Jiuh-Biing Sheu [*]

Institute of Traffic and Transportation, National Chiao Tung University, 4F, 114 Chung Hsiao W. Rd., Sec. 1, Taipei, Taiwan 10012, ROC

Abstract

Automatic incident detection and characterization is urgently required in the development of advanced technologies used for reducing non-recurrent traffic congestion on freeways. This paper presents a new method which is constructed primarily on the basis of the fuzzy clustering theories to identify automatically freeway incidents. The proposed approach is capable of distinguishing the time-varying patterns of incident-induced traffic states from the patterns of incident-free traffic states, and characterizing incidents with respect to the onset and end time steps of incidents, incident location, the temporal and spatial change patterns of incident-related traffic variables in response to the impacts of incidents on freeway traffic flows in real time. Lane traffic count and density are the two major types of input data, which can be readily collected from point detectors. Based on the spatial and temporal relationships of the collected raw traffic data, several time-varying state variables are defined, and then evaluated quantitatively and qualitatively to determine the decision variables used for real-time incident characterization. Utilizing the specified decision variables, the proposed fuzzy clustering-based algorithm executes recurrently three major procedures: (1) identification of traffic flow conditions, (2) recognition of incident occurrence, and (3) incident characterization. In this study, data used for model tests are generated from the CORSIM traffic simulator. Our preliminary test results indicate that the proposed approach is promising, and, in expectation, can be integrated with any published real-time incident detection technologies. Importantly, this study may contribute significantly to the applications of fuzzy clustering techniques, and stimulate more related research. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Fuzzy clustering; Pattern recognition; Automatic incident detection and characterization

1. Introduction

Automatic incident detection and characterization is urgently required in the development of advanced technologies used for reducing non-recurrent freeway traffic congestion effectively and efficiently. It has been illustrated that in urban areas, non-recurrent congestion caused by incidents results nearly in 60% of the

_____
[*] Tel.: +886-2-2349-4963; fax: +886-2-2349-4953.

E-mail address: jbsheu@cc.nctu.edu.tw (J.-B. Sheu).

urban freeway delay, and the impact continues to increase dramatically with years [9]. The preliminary test results of our related studies showed that compared to recurrent congestion in incident-free cases, the incident effects on freeway traffic congestion either in the spatial domain or in the temporal domain exhibit unusual change patterns which could be the major cause of bottlenecks or even secondary accidents on freeways [16]. Moreover, it has been suggested that the risk of secondary accidents can be significantly reduced by earlier detection and warning [4]. It is, therefore, understandable that real-time information with respect to incident characteristics is vital for reducing traffic impact during freeway incidents.

Despite the significance of identifying incident-related traffic characteristics in either analyzing or predicting non-recurrent traffic congestion on freeways, real-time incident characterization does not appear to draw much attention in early research related to automatic incident detection (AID) and management. That is, the focus of early AID-related studies is merely on recognizing incident occurrence. For better AID performance on freeways, the California-based algorithms execute decision making of incident detection primarily utilizing direct comparison technologies [11]. Nevertheless, a common problem with comparison approaches is the determination of thresholds employed for identification of incident occurrence. The determination of proper threshold values requires a large amount of data and tests which may limit the practical utility of comparison technologies. Levin et al. [8] proposed the use of Bayesian techniques for incident detection, and similarly, Tsai et al. [18] attempted to utilize Bayes' optimum decision rule in developing a freeway lane-blocking incident detection algorithm. Since the aforementioned Bayesian-theory-based algorithms are probabilistic approaches, the issue associated with the determination of probability density functions in these algorithms is critical.

Recently, researchers interested in freeway incident detection have focused greater attention on techniques related to raw data processing than to AID methods. Some researchers, for example, have proposed new approaches using pre-processed data for freeway incident detection rather than raw data used in early methods. According to a study by Persaud et al. [12], freeway incidents could be inferred from specific patterns exhibited by 30-s data sets on a catastrophe theory surface. However, some critical issues, such as incident detection during congestion, noisy data patterns on shoulder lanes, model testing under conditions complicated by weather or geometry, were not addressed in Persaud's study. The advantage of utilizing the data trained by artificial neural network technologies for freeway incident detection has also been suggested in several studies [1,5,7]. However, the model retraining problem which requires a large amount of incident data for model transferability remains in any Artificial Neural Networks-based algorithms. In addition, the application of video image processing to AID can be regarded as a technology of pattern recognition [3,10,19]. Limitations including the detection scope from a camera view and the time required for data processing, however, remain the main issues of image-processing systems used for real-time incident detection and characterization.

In view of the limitations of the published AID technologies and the significance of incident characterization, we have, in reality, conducted some related research which can be regarded as the preliminary tasks of the study [13–16]. According to our preliminary test results, real-time characterization of incidents in terms of the lanes blocked and incident-induced section-wide lane traffic variables seems promising by utilizing stochastic estimation approaches. On the other hand, some issues including the applicability of the published AID method to freeway incident cases, the capability of characterizing incident location in any given blocked lane, and the identification of incident termination remain in our earlier research.

The study, therefore, presents an advanced method which is proposed particularly for automatic freeway incident detection and characterization. The proposed method is developed primarily on the basis of fuzzy clustering techniques which have been applied to a variety of study areas for pattern recognition. The fundamentals and concepts of fuzzy clustering were utilized and improved in developing the proposed automatic incident detection and characterization algorithm. In addition, the decision variables used for incident characterization were analyzed, and then specified in the methodology development. The proposed fuzzy clustering-based method permits distinguishing non-recurrent congestion from recurrent congestion, and more importantly, recognizing freeway incident characteristics in real time by means of mapping the time-varying patterns of decision variables to the categories of traffic congestion with the highest degree of belongingness. Moreover,

real-time information including incident location, the onset and end time steps of incidents is provided in the procedure of incident characterization.

The remainder of the paper is organized as follows. Section 2 presents the methodology development which primarily includes: (1) fuzzy clustering, (2) determination of decision variables, and (3) algorithm development. Section 3 depicts model testing together with a brief summary of test results. Finally, conclusions and recommendations are summarized in Section 4.

2. Methodology development

The system investigated in this study is bounded by any given pair of the upstream and downstream detector stations on a mainline segment of a freeway. The area within the upstream and downstream detector stations is defined as a detection zone in this study.

The fundamental of the proposed method is to characterize incidents within any given detection zone utilizing the developed fuzzy clustering-based algorithm. Conveniently, raw traffic data collected from the upstream and downstream detector stations can be used to specify the decision variables which are the target variables input to the proposed algorithm for real-time characterization. Note that the proposed method also apply to the case of sequential detector station layouts on a mainline segment of a freeway, where the upstream detector station in a given detection zone can be regarded as the upstream detector station of the preceding detection zone.

The section primarily comprises three parts. In order to explicate the potential applicability of fuzzy clustering to an automatic incident detection and characterization in the study, the features of fuzzy clustering techniques are introduced in Section 2.1. Following this, the decision variables employed in the proposed algorithm for a real-time incident characterization are explored, and then, a brief description of the proposed real-time incident characterization algorithm is provided.

2.1. Fuzzy clustering

Before describing the proposed fuzzy clustering method, the fundamentals of fuzzy clustering are briefly introduced as an aid to understand the applicability of fuzzy clustering to an automatic freeway incident detection and characterization in the study.

Fuzzy clustering is a part of fuzzy data analysis, and can also be viewed as an improved clustering technique which has been used successfully not only for data compression but also for data categorization [2,6,17]. Compared to classical clustering techniques, the relative feature exhibited by fuzzy clustering is that fuzzy clustering utilizes fuzzy partitioning such that a given data point can belong to several groups with the degree of belongingness bounded within the range of 0 and 1. In classical clustering procedure, namely a partition of a set of S objects into T mutually exclusive clusters, the state of clustering can be expressed by an $S \times T$ matrix $U = [\mu_{st}]$, where $u_{st} = 1$ if object s belongs to the cluster t; otherwise $u_{st} = 0$. To ensure that the clusters are disjointed and non-empty, u_{st} must satisfy the following conditions:

$$\sum_{t=1}^{T} u_{st} = 1, \qquad s = 1, \ldots, S, \tag{1}$$

$$u_{st} \in \{0, 1\}, \qquad s = 1, \ldots, S; \quad t = 1, \ldots, T. \tag{2}$$

In reality, there are many cases in which the exclusive clusters are not suitable for natural subgroups. Therefore, fuzzy clustering concepts have been proposed. The fundamental of fuzzy clustering is to replace Eq. (2) with

$$u_{st} \in [0, 1], \qquad s = 1, \ldots, S; \quad t = 1, \ldots, T. \tag{3}$$

This means that the natural subgroup is considered as a fuzzy subset on a set of objects, where u_{st} represents the degree of belongingness of the sth object to the tth cluster.

In contrast with classical clustering techniques, the above feature can make fuzzy clustering more flexible for real-world applications, especially in cases where patterns of data attributes change irregularly with time such as incident events. Given a lane-blocking incident, the use of fuzzy clustering concepts could permit grouping patterns of lane traffic variables into relatively similar clusters in response to the incident effect on lane traffic flows in real time. Once an incident is detected, incident characteristics can be further identified on the basis of direct comparison of the degree of belongingness.

2.2. Determination of decision variables

Determination of decision variables is the first step conducted in the methodology development scenario, and herein decision variables are utilized, in the proposed fuzzy clustering algorithm, for a congestion categorization as well as an incident characterization. Three procedures are primarily involved in this scenario. Incident-related traffic patterns were first analyzed and categorized using raw traffic data generated via the CORSIM simulation model. Several candidates of decision variables were then pre-specified on the basis of the spatial and temporal changes in the incident-related traffic patterns, and evaluated. The decision variables used in the fuzzy-clustering algorithm were finally determined via comparing their capabilities in terms of characterizing the specific change patterns of raw traffic data under conditions of lane-blocking incidents. The description of these critical procedures is summarized below.

To capture incident-induced traffic patterns, 27 types of incident cases associated with different traffic flow conditions and incident locations on a three-lane mainline segment of a freeway were simulated using the CORSIM microscopic simulation model. In the simulations, three types of traffic flow cases were considered: (1) high-volume, (2) medium-volume, and (3) low-volume. Under given traffic flow conditions, lane-blocking incidents associated with different location, including upstream, middle-stream and downstream, were simulated on a given lane within the detection zone. In each incident case, the simulation time was set to be 30 min which includes: (1) a warm-up period of 10min preceding simulated incident generation, (2) 10min for incident duration, and (3) 10 min following the period of incident occurrence for traffic recovery. The simulated lane traffic counts and occupancies collected from the upstream and downstream detector stations were recorded and graphically displayed. The incident-related traffic patterns exhibited by lane traffic counts and occupancies were then specified and categorized.

The next procedure is the specification of decision variables. Theoretically, a decision variable must possess the capability of characterizing the change patterns of traffic flows either in the temporal domain or in the spatial domain under conditions of lane-blocking incidents on freeways, three groups of decision variables including time-based, space-based, and composite variables were therefore considered. Totally, there are ten candidates of decision variables are pre-specified in the study for further analysis at this stage, and their notations are expressed as follows:

$$v_i^1(k) = \frac{\sum_{j=1}^{J} o_j^d(k)}{J} - o_i^d(k), \tag{4}$$

$$v_i^2(k) = \frac{v_i^1(k)}{o_i^d(k)}, \tag{5}$$

$$v_i^3(k) = \frac{\{[\sum_{j=1}^{J} f_j^d(k)]/J\} - f_i^d(k)}{f_i^d(k)}, \tag{6}$$

$$v_i^4(k) = \left\{ \left[\left[\sum_{j=1}^{J} f_j^d(k) \right] \middle/ J \right] - f_i^d(k) \right\} \left\{ \left[\sum_{j=1}^{J} f_j^d(k) \right] \middle/ J \right\}, \tag{7}$$

$$v_i^5(k) = \frac{[o_i^u(k) - o_i^d(k)]}{o_i^u(k)}, \tag{8}$$

$$v_i^6(k) = [f_i^d(k) - f_i^u(k)] + \left[f_i^d(k) - \left[\sum_{j=1}^J f_j^u(k)\right]\middle/J\right], \tag{9}$$

$$v_i^7(k) = [o_i^d(k) - o_i^u(k)] + \left[o_i^d(k) - \left[\sum_{j=1}^J o_j^u(k)\right]\middle/J\right], \tag{10}$$

$$v_i^8(k) = \frac{\left[\left[\sum_{\tau=0}^{T-1} f_i^d(k - \tau)\right] - f_i^u(k - T)\right]}{\sum_{\tau=0}^{T-1} f_i^d(k - \tau)}, \tag{11}$$

$$v_i^9(k) = \frac{v_i^8(k)}{\left\{\left[\sum_{j=1}^J \left(\left[\left[\sum_{\tau=0}^{T-1} f_j^d(k - \tau)\right] - f_j^u(k - T)\right]\middle/\sum_{\tau=0}^{T-1} f_j^d(k - \tau)\right)\right]\middle/J\right\}}, \tag{12}$$

$$v_i^{10}(k) = \left\{\sum_{\tau=0}^T f_i^d(k - \tau)\right\} - f_i^u(k - T) - \frac{\sum_{j=1}^J f_j^d(k)}{J}. \tag{13}$$

In Eqs. (4)–(13), $f_i^u(k)$ and $f_i^d(k)$ represent, respectively, the upstream and downstream traffic counts collected on the target lane i at time step k, and similarly, $o_i^u(k)$ and $o_i^d(k)$ correspond to the collected occupancies; J is the total number of the adjacent lanes; T represents the maximum time-lag pre-determined for consideration of the effect of the travel time taken from the upstream detector station to the downstream detector station on the relationship between the upstream and downstream traffic data; τ is a time-lag index.

The aforementioned decision variable candidates are specified on the basis of the effects of incidents potentially on the changes of the collected raw traffic data, either or both, in the spatial and temporal domains. $v_i^1(k)$ and $v_i^2(k)$, for example, are two variables designed in consideration of the spatial difference of lane occupancy data collected at the same detector station in a given time step. Similar to $v_i^1(k)$ and $v_i^2(k)$, $v_i^3(k)$ and $v_i^4(k)$ are donated primarily on the basis of the spatial relationship of lane traffic counts. Suppose that a lane is targeted in a given time step, the spatial relationship of lane occupancies collected at the upstream and downstream detector stations in the target lane can be utilized to specify variables such as $v_i^5(k)$ for incident characterization. $v_i^6(k)$ and $v_i^7(k)$ are also two spatial-comparison-based variables, which are more complicated. If both the spatial and temporal factors are taken into account in the specification of decision variables, $v_i^8(k)$, $v_i^9(k)$ and $v_i^{10}(k)$ are then specified.

Under given traffic flow conditions (e.g., high-volume, medium-volume, and low-volume cases), a decision variable candidate is selected as the decision variable if its time-varying patterns significantly exist in most of the simulated incident cases. To quantitatively evaluate the capability of a given decision variable candidate ($v_i^\theta(k)$) in terms of characterizing incident-induced traffic patterns, a specific performance measure is specified, and it is denoted by

$$\varepsilon_i^\theta = \frac{\sum_{k=1}^K [v_i^\theta(k) - \bar{v}_i^\theta]}{K}, \tag{14}$$

where θ represents the code of a given decision variable candidate; K is the incident duration represented in the unit of time step; $\bar{v}_i^\theta$ corresponds to the average value of $v_i^\theta(k)$ under incident-free conditions. Based on comparisons of the estimated performance measures associated with the pre-specified variable candidates, the candidate with the highest value of the performance measure was selected as a decision variable for each traffic-flow case, and in consequence, $v_i^5(k)$, $v_i^4(k)$ and $v_i^{10}(k)$ were determined finally as decision variables in the study. Moreover, their time-varying patterns associated with different types of incidents were recorded as specific clusters for use in characterizing high-volume, medium-volume and low-volume incidents, respectively.
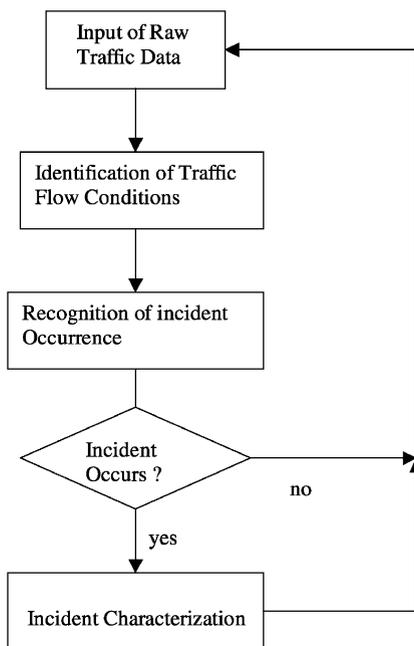
Fig. 1. Architecture of the proposed fuzzy clustering-based algorithm.

2.3. Algorithm development

This subsection presents a fuzzy clustering-based algorithm which is proposed to characterize freeway incidents in real time. The entire architecture of the proposed algorithm is as shown in Fig. 1, and it contains three major procedures which are sequentially executed at each time step: (1) identification of traffic flow conditions, (2) recognition of incident occurrence, and (3) incident characterization. The three sequential procedures are detailed in the following subsections.

2.3.1. Identification of traffic flow conditions

Identification of traffic flow conditions is a procedure executed to pre-classify traffic flow conditions using the occupancy data collected from the upstream detector station. Four types of traffic flow conditions are categorized in the study: (1) over-congestion, (2) high-volume, (3) medium-volume, and (4) low-volume. Each of the four traffic flow cases is assigned with an occupancy-based threshold which is pre-determined using historical data. At a given time step (time step index $= k$), the time-varying occupancy-based correlation value associated with each type of traffic flow conditions, $\gamma_m(k)$, is calculated using the lane occupancies collected from the upstream detector station together with the pre-determined occupancy-related thresholds (see Eq. (15)). Based on the comparison of the time-varying occupancy-based correlation values, the specific type of traffic flow conditions associated with the highest correlation value is then identified at the given time step.

$$\gamma_m(k) = 1 - \frac{1}{\alpha}\sqrt{\sum_{\tau=0}^{5}(\bar{o}^{\mathrm{u}}(k-\tau) - o_m)^2}, \tag{15}$$

where α is a pre-determined value which bounds the value of $\gamma_m(k)$ within the range 0–1; $\bar{o}^u(k - \tau)$ represents the average occupancy value of the upstream lane occupancies collected at time step $k - \tau$; o_m is the pre-determined occupancy-related threshold associated with a specific type "m" of traffic flow conditions.

2.3.2. Recognition of incident occurrence

The procedure of incident occurrence recognition primarily serves the purpose of distinguishing incident-related traffic patterns from incident-free traffic patterns. To provide this functionality, the specified decision variables together with the fuzzy clustering approaches are used. Several steps are conducted in the procedure as follows:

Step 1: Under given traffic flow conditions, calculate the related decision variable. As mentioned previously, three types of traffic flow conditions including high-volume, medium-volume and low-volume cases are taken into account in the algorithm. Once a specific type of traffic flow conditions is determined in the previous procedure of the algorithm, the time-varying value of the related decision variable at the current time step is computed using raw traffic data.

Step 2: Compute, lane-by-lane, the time-varying decision-variable correlation for different incident location cases (i.e., upstream, middle-stream and downstream) as well as the incident-free cases. The time-varying decision-variable correlation measures the amount of association between the specific decision variables calculated in Step 1 and the related pattern pre-clustered utilizing historical data. In the study, the time-varying decision-variable correlation $\omega_{m,n}^{p,q}(k)$ associated with the type of traffic flow conditions m, lane n, location p, and incident signal can be expressed as

$$\omega_{m,n}^{p,q}(k) = 1 - \frac{1}{\beta}\sqrt{\sum_{\tau=0}^{3}(v_n^s(k-\tau) - \mu_{m,n}^{p,q})^2}, \tag{16}$$

where β is a pre-determined value which is set for the boundaries of $\omega_{m,n}^{p,q}(k)$; the subscript m is an integral parameter indicating four types of traffic flow conditions such as: $m = 1$ for low-volume, $m = 2$ for medium-volume, $m = 3$ for high-volume, and $m = 4$ for over-congestion; the subscript n corresponds to such a lane code as: $n = 1$ for the outside lane, $n = 2$ for the lane adjacent to the outside lane, etc.; the superscript p denotes a location-related parameter which indicates the location investigated in the detection zone such as $p = 1$ for upstream, $p = 2$ for middle-stream and $p = 3$ for downstream; the superscript q is a binary digit: $q = 1$ representing incident occurrence, and $q = 0$ non-incident; $\mu_{m,n}^{p,q}$ denotes the pattern of the decision variable pre-clustered on the basis of historical traffic data associated with the attributes m, n, p and q.

Step 3: Recognize lane-blocking incidents. Given the type of traffic flow condition m in a given lane n, a lane-blocking incident is recognized employing the following fuzzy clustering-based rules:

IF $\quad \{\text{Max}[\omega_{m,n}^{1,1}(k), \omega_{m,n}^{2,1}(k), \omega_{m,n}^{3,1}(k)] - \omega_{m,n}^{p,0}(k)\} > \lambda_1$

THEN a lane-blocking incident with the attributes m and n, is recognized at time step k $\qquad$ (17)

ELSE $\quad k = k+1$ and go back to the previous procedure (Identification of Traffic Flow Conditions)

where λ_1 denotes a pre-determined threshold for distinguishing incident-related traffic patterns from incident-free traffic patterns. Note that in Step 3, the mechanism of incident recognition is conducted lane-by-lane.

Step 4: Go to the following procedure for incident characterization.

2.3.3. Incident characterization

This procedure is triggered to execute incident characterization once a lane-blocking incident has been recognized in the previous procedure (Section 2.3.2). The primary function provided by this procedure is characterizing incidents in terms of incident location and the end time step of incident. To achieve this, two steps are involved in this procedure as follows:

Step 1: Given the blocked lane n and the type of traffic flow conditions m, characterize the incident location utilizing the following fuzzy-clustering rule:

IF $\text{Max}[\omega_{m,n}^{1,1}(k), \omega_{m,n}^{2,1}(k), \omega_{m,n}^{3,1}(k)] > \lambda_2$

THEN   a lane-blocking incident with the location attribute is determined at time step k (18)

where λ_2 is a pre-determined threshold for characterizing incident location.

Step 2: Determine if the lane-blocked incident terminates at time step k employing the following fuzzy-clustering rule:

IF $\{\text{Max}[\omega_{m,n}^{1,1}(k), \omega_{m,n}^{2,1}(k), \omega_{m,n}^{3,1}(k), -\omega_{m,n}^{p,0}(k)\} < \lambda_3$

THEN it is determined that the lane-blocking incident terminates at time step k,

          and stop the procedure of incident characterization.   (19)

ELSE it is determined that the lane-blocking incident remains at time step k

where λ_3 represents a pre-determined threshold for the determination of the incident ending time step.

Step 3: $k = k + 1$, and then go back to the first procedure of the algorithm.

The sequence of major computational steps is shown in Fig. 2.

3. Numerical tests and results

The study site selected for the model tests was located at a 3-lane mainline segment of the southbound $N-1$ Freeway in Taiwan, which is 1.3 miles (2.0 km) in length. The average hourly volumes during afternoon-peak hours, off-peak daytime hours and off-peak nighttime hours at the study site were 5380, 3246 and 1275 vph, respectively. Pairs of loop detectors were situated upstream and downstream at the study site to collect raw traffic data including traffic count, speed and occupancy. The field data were collected over a period of 6 months, 7 days a week, 24 h a day with the assistance of National Freeway Bureau of Taiwan, and used primarily for model calibration.

Two groups of traffic data, including (1) incident-related data, and (2) incident-free data, were used in the model tests for different purposes. In contrast to incident-related data which were used to evaluate the model's performance concerning incident detection and characterization, the primary purpose of using incident-free data was to measure the model's capability in terms of recognition of traffic congestion patterns.

Considering the lack of real incident-related traffic data, the proposed method was tested using simulated data which were generated via the CORSIM microscopic simulation model, Version 1.03. To simulate incidents associated with different characteristics at the study site, a simulated network was built on the basis of the collected geometric and traffic data of the study site. The aforementioned average hourly volumes were used as the total inflows in CORSIM for simulating, respectively, high-volume, medium-volume and low-volume incidents as well as incident-free cases. Under given traffic flow conditions, lane-blocking incidents were simulated at the upstream, middle-stream and downstream segments of the study site for each lane, and thus, 27 types of incidents were totally simulated in these tests. Each simulation event in the study lasted for 30 min: the first 10 min for warming up, the next 10 min for incident duration, and the rest for traffic recovery.

In this study, the off-line tests were focused on testing the capability of the proposed method in terms of characterizing incident location and duration in real time. Detection rate (DR), false alarm rate (FAR), and time to detection (TTD) are the three major performance measures used for evaluating the performance of the proposed method in terms of determining the start and end time steps of incidents. In addition, an extra measure, the accuracy of incident characterization (AIC), is introduced particularly for testing model's
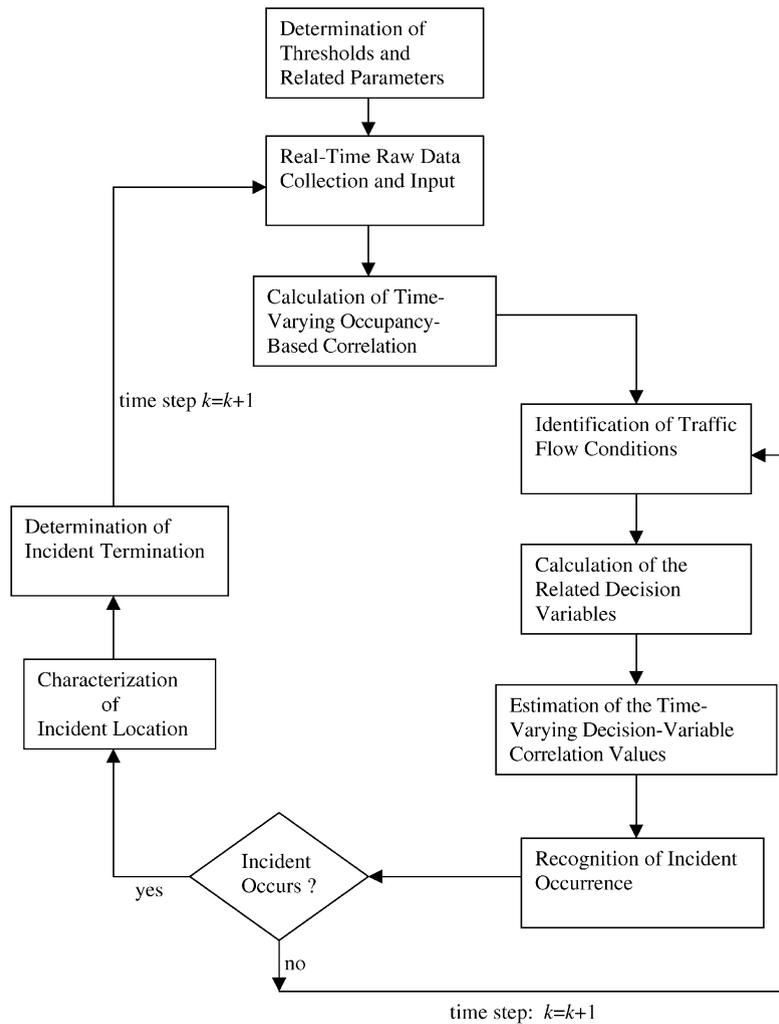
Fig. 2. Computational loop of the real-time incident characterization algorithm.

capability in terms of identifying incident location in the blocked lane. The donation of the accuracy of incident characterization is given by

$$AIC = \frac{\text{the number of incident events detected with correct incident characterization}}{\text{the total number of incident events}}.$$

The test results concerning the recognition of incident occurrence and termination are summarized in Tables 1 and 2, respectively.

The results shown in Table 1 indicate that the proposed fuzzy clustering-based algorithm seems to permit characterizing incidents as well as recognizing incident occurrence in real time. Using the specified decision variables, the proposed algorithm can recognize precisely a specific type of incident via fuzzy clustering technologies within 30 s in the tests. Moreover, incident characteristics including traffic congestion conditions, the lanes blocked by incidents, and incident location relative to the downstream detector station in a given blocked lane can also be identified together with the recognition of incident occurrence.

Table 1
Results of off-line tests for characterization of incident occurrence

Incident location	Upstream				Middle-stream				Downstream			
Traffic conditions	TTD (s)	DR (%)	FAR (%)	AIC (%)	TTD (s)	DR (%)	FAR (%)	AIC (%)	TTD (s)	DR (%)	FAR (%)	AIC (%)
High-volume	23.3	100	0	100	23.3	100	0	100	26.6	100	0	100
Medium-volume	30.0	100	0	100	30.0	100	0	100	27.5	100	0	100
Low-volume	30.0	100	0	100	30.0	100	0	100	30.0	100	0	100

Table 2
Results of off-line tests for characterization of incident termination

Incident location	Upstream				Middle-stream				Downstream			
Traffic conditions	TTD (s)	DR (%)	FAR (%)	AIC (%)	TTD (s)	DR (%)	FAR (%)	AIC (%)	TTD (s)	DR (%)	FAR (%)	AIC (%)
High-volume	150.	100	0	100	170.	100	0	100	163.	100	0	100
Medium-volume	93.3	100	3.7	100	95.0	100	3.7	100	83.3	100	0	100
Low-volume	40.0	100	0	100	40.0	100	0	100	23.3	100	0	100

The test results shown in Table 2 offer two generalizations as follows. First, the capability of the proposed method in terms of recognizing incident termination is as good as we expected although two false alarms associated with medium-volume incident cases were reported in this test scenario. Second, in contrast to the test results shown in Table 1, the time taken to confirm incident termination is relatively longer. This generalization is explainable from a theoretical point of view. It is worth noting that the methodology is constructed on the basis of fuzzy clustering theories. The termination of an incident can be confirmed only when the patterns of decision variables are mapped into incident-free categories. That is, in the characterization procedure of the proposed algorithm, the termination of an incident is not confirmed until the time-varying degree of belongingness associated with a given incident-free case is greater than the degree of belongingness associated with any type of incident at a given time step.

As a whole, our preliminary test results show the feasibility of achieving real-time freeway incident characterization using the proposed methodology which is never explored in early studies. It is worth noting that according to the test results in terms of the measures of *AIC* which we proposed particularly for the evaluation of the algorithm's performance with respect to incident characterization in the study, it seems promising that the proposed algorithm is capable of characterizing freeway incidents as well as recognizing incident occurrence at the current stage of testing.

4. Conclusions and recommendations

Despite a great deal of advances in freeway incident detection that have been made in early research, the issue of real-time incident characterization still remains. Without the functionality of providing real-time information related to incident characteristics, conventional AID methods seem to be restricted to recognition of incident occurrence. This significantly results in a defect in the advanced freeway traffic control and management systems, and continues to be a source of frustration in real-time freeway incident management.

This paper has presented a new methodology utilizing the fuzzy clustering approaches to address the above issue. To capture incident-induced traffic patterns, effectively and efficiently, several decision variables were

designed, and then evaluated. Based on the pre-determined decision variables, a total of 27 types of freeway incidents were categorized, and each type of incident was referred to as a cluster along with specific patterns of decision variables to distinguish incidents from incident-free cases. A fuzzy clustering-based algorithm, which consists of three major procedures including: (1) identification of traffic flow conditions, (2) recognition of incident occurrence, and (3) incident characterization, was then developed.

The off-line tests for evaluating the performance of the proposed algorithm were conducted on the basis of simulation data generated via CORSIM. The test results summarized in the testing scenario for recognition of either incident occurrence or incident termination have indicated the feasibility of achieving real-time incident characterization as well as incident recognition using the proposed method.

In contrast to conventional AID algorithms, the proposed method may offer two relative advantages:

1. In addition to recognition of incident occurrence, the proposed method is capable of providing incident-related information in real time including: (1) the lanes blocked, (2) the start and end time steps of incidents, (3) the location of an incident relative to the downstream detector station in a given blocked lane. The functionality provided by the algorithm is vital to the advanced freeway traffic management systems because it may help us better understand incident characteristics, and thus, facilitate the optimal decision making in response to freeway incident occurrence via advanced technologies such as real-time incident-responsive traffic control and management systems.
2. The proposed automatic incident detection and characterization algorithm is compatible with automatic incident management systems. In addition to the function of AID provided via the second procedure of the proposed algorithm, the third procedure, incident characterization, can be regarded as an assistant mechanism for automatic incident management which rely heavily on the capacity of recognizing time-varying incident-induced traffic patterns in response to the variety of incident impacts in real time.

In addition to the aforementioned advantages, it is worth noting that our latest attempt is the integration of the proposed fuzzy clustering-based algorithm with our related research in an effort to achieve the goal of real-time incident-responsive traffic control [16]. This may help illustrate the flexibility of the proposed method as well. More importantly, we really hope this study can be viewed as a striking example for the applications of fuzzy clustering techniques, and stimulate more related research.

Acknowledgements

References

[1] B. Abdulhai, S.G. Ritchie, Development of a universally transferable freeway incident detection framework, Proceedings of Transportation Research Board, 76th Annual Meeting, Washington, DC, 1997.

[2] J.C. Bezdek, Fuzzy mathematics in pattern classification, Ph.D. Thesis, Cornell University, USA, 1973.

[3] J.M. Blosseville, V. Motyka, N. Djemame, F. Lenoir, Automatic incident detection using image processing techniques: a special system used in INVAID, applications of advanced technologies in transportation engineering, Proceedings of the Second International Conference, Pacific Rim Transtech Conference, 1991, pp. 383–387.

[4] F. Busch, Incident detection, in: M. Papageorgiou (Ed.), Concise Encyclopedia of Traffic & Transportation Systems, Pergamon Press, Oxford, 1987, pp. 219–225.

[5] R.L. Cheu, S.G. Ritchie, Automatic detection of lane-blocking freeway incidents using artificial neural networks, Transp. Res. 3C (6) (1995) 371–388.

[6] L.O. Hall, A.M. Bensaid, L.P. Clarke, R.P. Velthuizen, M.S. Silbiger, J.C. Bezdek, A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain, IEEE Trans. Neural Networks 3 (5) (1992) 672–682.

[7] S. Ishak, H.M. Al-Deek, Freeway incident detection using fuzzy ART, Proceedings of the Fifth International Conference of Applications of Advanced Technologies in Transportation, Newport Beach, CA, 1998.

[8] M. Levin, G.M. Krause, Incident detection: a Bayesian approach, Transp. Res. Rec. 682 (1978) 52–58.

[9] J.A. Lindley, Urban freeway congestion: quantification of the problem and effectiveness of potential solutions, Inst. Transp. Eng. 57 (1987) 27–32.

[10] P.G. Michalopoulos, R.D. Jacobson, Automatic incident detection through video image processing, Traffic Eng. Control 34 (2) (1993) 66–75.

[11] H. Payne, S. Tignor, Freeway incident detection algorithms based on decision trees with states, Transp. Res. Rec. 682 (1978) 30–37.

[12] B.N. Persaud, F.L. Hall, Catastrophe theory and patterns in 30-s freeway traffic data-implications for incident detection, Transp. Res. 23A (2) (1989) 103–113.

[13] J.-B. Sheu, Stochastic estimation of lane-changing probabilities and its application to incident detection, Doctoral Dissertation, University of California, Irvine, USA, 1997.

[14] J.-B. Sheu, S.G. Ritchie, A new methodology for incident detection and characterization on surface streets, Transp. Res. Part C 6C (5–6) (1998) 315–335.

[15] J.-B. Sheu, A stochastic modeling approach to dynamic prediction of section-wide inter-lane and intra-lane traffic variables using point detector data, Transp. Res. 33A (2) (1999) 79–100.

[16] J.-B., Sheu, Y.-H. Chow, L.-J. Shen, A stochastic estimation approach to real-time prediction of incident effects on freeway traffic congestion, Transp. Res. 35B (6) (2001) 575–592.

[17] M. Sugeno, T. Yasukawa, A fuzzy-logic-based approach to qualitative modeling, IEEE Trans. Systems 24 (1993) 1279–1284.

[18] J. Tsai, R. Case, Development of freeway incident-detection algorithms by using pattern-recognition techniques, Transp. Res. Rec. 722 (1979) 113–116.

[19] J. Versavel, B. Boucke, Video for traffic data and incident detection by Traficon, Proceedings of the Fifth International Conference of Applications of Advanced Technologies in Transportation, Newport Beach, CA, 1998.