CHAPTER 2 LITERATURE REVIEW

The overview of the automatic incident detection (AID) is made in this chapter. The AID algorithms are in general clustered into two categories based on macroscopic or microscopic traffic flow parameters being used. Section 2.1 and 2.2 briefly investigate the literature review and Section 2.3 provides with some comments.

2.1 Macroscopic AID Algorithms

Macroscopic AID algorithms are composed into pattern recognition algorithm (Payne and Tignor, 1978; Luk and Sin, 1992; Stephanedes and Chassiakos, 1993; Steed and Clowes, 1989; Stephanedes, *et al.*, 1992), statistical approach (Ahemd and Cook, 1982; Cook and Cleveland, 1974; Levin and krause, 1978; Dudek, *et al.*, 1974; Sheu and Ritchie, 1998; Sheu, 2004), catastrophe theory (Persaud and Hall, 1989), artificial intelligent based approach (Lin and Chang, 1998; Ritchie and Cheu, 1993; Stephanedes and Liu, 1995; Dia and Rose, 1997), and fuzzy set theory (Lee, *et al.*, 1998; Xu, *et al.*, 1998).

2.1.1 Pattern Recognition Algorithm

Most pattern recognition algorithms for incident detection were developed based on the observation that incidents would result in a distinctive spatial pattern in surveillance data, reflective of the discontinuity in the traffic stream caused by the incident. Hence, the core logic of pattern recognition algorithms can be stated as follows: Compare the upstream and/or downstream traffic pattern within and between lanes with normal traffic patterns to ascertain the occurrence of an incident.

The following algorithms reported in the literature are classified into the category of pattern recognition. The original California-type algorithm and its modified versions (Payne *et al.*, 1976, and 1978; Tignor and Payne, 1977); Incident detection algorithms in

COMPASS (Masters *et al.*, 1991); The UK High Occupancy (HIOCC) algorithm (Collins, 1983; Steed and Clowes, 1989); Pattern Recognition (PATREG) algorithm (Collins *et al.*, 1979; Collins and Martin, 1991); Incident Detection in Australia (Luk and Sin, 1992); and Minnesota Algorithm (Stephanedes and Chassiakos, 1991, 1993).

Their common feature is that an incident decision is made on a sequence of efficient comparisons between the detected and predicted states, and is thus well suited for on-line implementation. Some algorithms can perform reasonable well under the simple single-loop detection system. Hence, although a consider number of incident detection algorithms has been surfaced in the literature, this category of simple detection algorithms, especially the California Algorithms, remains quite popular, and often serves as the benchmark for comparison.

The following summarizes all operational as well as implementation issues related to the application of pattern recognition algorithms, including their applicability, data requirements, threshold or parameter calibration, required surveillance systems, computational as well as theoretical limitations. Such issues are critical for their successful implementation, and deserve further exploration for potential improvements. (1) Surveillance system: This class of algorithms can usually be used on freeways equipped with single-loop detectors, except for Algorithm #10, the Australia and APID algorithms, because speed measurements are required for their incident detection. Hence, double-loop detectors or any other types of detectors capable of measuring speed, volume and occupancy are acceptable surveillance systems, HIOCC and PATREG algorithms also require single-loop detectors in addition to a special computer system, which can sample occupancy at every 1/10 second and output i-second instantaneous occupancy. Thus, it may be relatively difficult to transfer these two algorithms to different locations because the i-second occupancy measurements are not usually available on many freeway surveillance systems. (2) Performance under light traffic conditions: Most algorithms' performance deteriorates under light traffic conditions, including the California Algorithm #10 and APID. The reason is that incidents occurring under light traffic conditions may not have obvious impacts on the traffic flow pattern, and thus may not reflect in their measurements. The Australia algorithm and PATRETG seem to work reasonably well under light traffic conditions, but test results regarding this issue have not been found in the literature. (3) Historical data requirements: Historical data are needed for the threshold calibration. A software package is available for computing the optimal thresholds for the California-type algorithms. All other algorithms' thresholds have to be calibrated by the trial-and-error method. (4) Parameter updating mechanism: All algorithms in this category do not have a parameter updating mechanism and their thresholds are not responsive to variations in traffic patterns, which thus limit their performance robustness. (5) Sensitivity to environmental changes: All algorithms are sensitive to changes in geometric conditions and traffic compositions, since the former may result in the need to re-calibrate their thresholds values and the latter may affect the accuracy of the occupancy measurements. (6) Performance robustness: Most algorithms' performance is robust against moderate variations in weather conditions, since incident patterns under such circumstances may not differ significantly from those under normal weather conditions. (7) Detector spacing: Their performance is sensitive to sensor spacing to some extent. Hence, these algorithms usually perform better under shorter detector spacing. (8) End of incidents: The California type algorithms and the HIOCC algorithm can identify the end of an incident. (9) Severity detection: Theoretically, the severity of a detected incident may be estimated by further examining the change-rate of the detection variables. However, this issue has not been discussed in most existing literature reviewed in this section. (10) Data quality improvement: The Minnesota, Australia, and APID algorithms detect incidents with smoothed data, which may be one of the key contributing factors that results in their reduced false-alarm rate. Other algorithms do not require the use of smoothed data. (11) Detection time: This class of algorithms usually take longer time (minutes) to detect an incident because the rely on a sequence of pattern comparisons. (12) Threshold calibration: Most algorithms in this category calibrate their thresholds and parameters by the trial-and-error.

2.1.2 Statistical Approach

The detection algorithms based on statistical approaches include time-series (Ahmed and Cook, 1979, 1980, 1982; Ahmed, 1983), double exponential smoothing (DES) (Cook and Cleveland, 1974), Bayesian (Levin and Krause, 1978 and 1979; Dudek and Messer, 1974; Tsai and Case, 1979), standard normal deviation algorithms -- SND (Dudek and Messer, 1974), Kalman filter (Chui and Chen, 1987) and modified sequential probability ration tests – MSPRT (Sheu and Ritchie, 1998; Sheu, 2004).

Sheu (2004) developed a new methodology for real-time detection and characterization of freeway incidents. The proposed technology is capable of detecting freeway incidents in real time as well as characterizing incidents in terms of time-varying lane-changing fractions and queue lengths in blocked lanes, the lanes blocked due to incidents, and duration of incident, etc. The architecture of the proposed incident detection approach consists of three sequential procedures: (1) symptom identification for identification of anomalous changes in traffic characteristics probably caused by incidents, (2) signal processing for stochastic estimation of incident-related lane traffic characteristics, and (3) pattern recognition for incident detection. Lane traffic count and occupancy are two major types of input data, which can be readily collected from point detectors. The primary techniques utilized to develop the proposed method include: (a) discrete-time, nonlinear, stochastic system modeling used in the signal processing procedure, and (b) modified sequential probability ratio tests employed in the pattern recognition procedure.

A common feature of these algorithms is that they rely on statistical methods to model the normal traffic pattern, rather than the use of established dynamic traffic flow relations to capture the complex traffic behavior. However, a number of statistical methods employed in these algorithms are static in nature and thus can not dynamically adjust their parameters in response to any systematic changes in traffic behavior. Hence, the effectiveness of such algorithms mainly depends on the available historical data. For instance, the incident detection algorithm with the time-series approach is based on Box-Jenkins' ARIMA model, and is calibrated with the observed occupancy data. The 95 confidence interval of the predicated value are used to perform the traffic state monitoring. An incident signal will be alarmed if the measured occupancy lies outside the confidence limits. The effectiveness of such an algorithm depends on how well the proposed ARIMA model captures the time-varying traffic flow patterns. Similarly, both DES (Cook and Cleveland, 1974) and SND (Dudek and Messer, 1974) algorithms are based on the assumption that the traffic pattern under normal conditions can be represented with some well calibrated function.

The following summarizes the application of most statistical approach: (1) Surveillance system: All algorithms reviewed in this category can be transferred for use, in principle, to any roadways equipped with only single-loop detectors or other sensing devices capable of measuring occupancy. (2) Performance under light traffic conditions: Most of these reviewed algorithms face the problem of performance deterioration under light traffic conditions, because incidents may not have a significant impact on the traffic measurements under such conditions. Hence, these algorithms have the difficulty in recognizing incidents, and distinguishing them from false alarms. (3) Data requirements: Historical data are needed to calibrate models or thresholds for all algorithms. Extensive data of several kinds are required for the algorithms based on the Bayesian's approach. (4) Parameter updating mechanism: A few of the algorithms in this category have a parameter updating mechanism which thus limits their performance robustness. (5) Sensitivity to environmental changes: The performance of all the algorithms is sensitive to changes in geometric conditions. Models or thresholds must be re-calibrated in order to adapt to such changes. Variations in traffic composition also have some impacts on the performance of this class of algorithms. (6) Detector spacing: The performance of all these algorithms is sensitive to sensor spacing. Usually, they perform better at short sensor spacing. (7) Severity detection: No algorithm in this category can identify the severity of an incident. This issue has not been discussed in the literature, although all the algorithms, except Bayesian, have the potential to do so. (8) Data quality improvement: It is not necessary for the DES and SND algorithm to use smoothed data. The time-series and Bayesian algorithms do not use smoothed data either. It is not clear if the use of smoothed data would improve the performance of the time-series and Bayesian algorithms. (9) Detection time: The required detection time for most algorithms is around one to several minutes. The time-series algorithm offers a shorter detection time, which makes it more attractive to be integrated to a traffic management system.

2.1.3 Catastrophe Theory

The McMaster-type algorithms are the primary product of the number of researches who have made their efforts over the recent years on the application of the Catastrophe theory to the freeway traffic flow analysis, operation, and incident detection (Persaud and Hall, 1988, 1989; Forbes, 1992; Persaud et al., 1990; Gall and hall, 1989; Hall et al., 1991; Hall, 1987; Athol, 1965; and Forbes and Hall, 1990). The pioneer contributors in developing the McMaster algorithm are Persuad and Hall (1988), and Gall and Hall (1989) who proposed a systematic way for incident analysis and detection.

All the McMaster algorithms perform incident detection by either identifying the occurrence and dissipation of traffic congestion, or observing abrupt changes in the flow speed in addition to congestion pattern analyses.

As a common feature, all McMaster-type algorithms employ special templates on the volume-occupancy plan to assist in the identification of congestion patterns and their causes. In addition, they also use the abrupt change in speed and /or other traffic variables as indicators for incident detection. The McMaster algorithm and its modified version named the modified McMaster algorithm (Forbes, 1992), were reported to be capable of performing on-line template updating. Accordingly, their performance is expected to be robust against moderate changes in weather and other environmental conditions.

2.1.4 Artificial Intelligent Based Approach

Recently, various types of artificial intelligent based algorithms have been proposed. Cheu and Ritchie (1995) presented an incident detection technique based on artificial neural networks (ANNs). Three types of neural network models, namely the multi-layer feedforward (MLF), the self-organizing feature map (SOFM) and adaptive resonance theory 2 (ART 2), were developed to classify traffic surveillance data obtained from loop detectors, with the objective of using the classified output to detect lane-blocking freeway incidents. The models were developed with simulation data from a study site and tested with both simulation and field data at the same site. The MLF were found to have the highest potential, among the three ANNs, to achieve a better incident detection performance. The MLF was also tested with limited field data collected from three other freeway locations to explore its transferability.

Dia and Rose (1997) proposed a multi-layer feedforward (MLF) neural network incident detection model. The results of the comparative performance evaluation clearly demonstrate the substantial in incident detection performance obtained by the neural network model and also show how improvements in model performance can be achieved using variable decision threshold.

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.

Srinivasan, *et al*. (2000) developed a hybrid artificial intelligence technique, with fuzzy-logic and genetic-algorithm technique, for automatically detecting incidents on a traffic network. A cascaded framework of 11 fuzzy controllers takes in traffic indices such as occupancy and volume, to detect incidents along an expressway in California. The flexible and robust nature of the developed fuzzy controller allows it to model functions of arbitrary complexity, while at the same time being inherently highly tolerant of imprecise data. The maximizing capabilities of genetic algorithms, on the other hand, enable the fuzzy design parameters to be optimized to achieve optimal performance.

2.1.5 Fuzzy Set Theory

Many researchers have developed the incident detection algorithms using fuzzy set theory including fuzzy pattern recognition, fuzzy inference, fuzzy expert system, fuzzy logic incident patrol system (FLIPS), and so on.

Lin and Chang (1998) presented the exploratory results of using fuzzy expert systems for incident detection and classification. The proposed system functions to detect not only the occurrence of incidents, but also their located lanes, and the resulting type of severity. With such information, the traffic control center can better advise drivers to take necessary lane changes and take timely actions for minimizing the incident impacts on traffic conditions.

Lee, *et al*. (1998) proposed a fuzzy-logic-based incident detection algorithm for signalized urban diamond interchanges. The model is capable of detecting lane-blocking incidents whose effects are manifested by patterns of deterioration in traffic conditions that require adjustments in signal control strategies. As a component of a real-time traffic adaptive control system for signalized diamond interchanges, the algorithm feeds an incident report (i.e., the time, location, and severity of the incident) to the system's optimization manager, which uses that information to determine the appropriate signal control strategy.

Sheu (2002) presented 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.

2.1.6 Chaotic Theory

Chaos theory has been widely applied on the science of various fields ranging from meteorology, chemistry, biology to sociology. In the area of traffic flow theory, some previous researches have shown that the dynamics of traffic flow time series data exhibit chaotic phenomena (For instance, Addison and Low, 1996; Dendrinos, 1994; Disbro and Frame, 1989; Frison and Abarbanel, 1997; Lan and Lin, 2003; Zhang and Jarrett, 1998). More recently, some literatures have found that chaotic approaches such as fuzzy local reconstruction method (Iokibe, *et al.* 1994, 1995), phase space local approximation method (Farmer and Sidorowich, 1987; Lan and Chen, 1998; Lan and Lin, 2001) and confined space fuzzy neighboring difference method (Lan, *et al*. 2003c) can make short-term predictions on the traffic dynamics with satisfactory accuracy. In addition, some other studies have employed the change in chaotic parameters to diagnose the abnormality of automobile engines as well as to discriminate such diseases as diabetes, arrhythmia, and ventricular fibrillatory (Iokibe, *et al*. 1996, 1997).

Conventional traffic incident detection algorithms, including pattern recognition, statistical approach, catastrophe theory, artificial intelligent based approach and fuzzy set theory, are mainly based on the change in some macroscopic traffic parameters such as flow, speed, occupancy and density or microscopic traffic parameters such as headway and spacing. None of these algorithms have tried such chaotic parameters as largest Lyapunov exponent, capacity dimension, correlation dimension, relative Lz complexity, delay time, and Hurst exponent to diagnose the abnormality of traffic flow under incident circumstances. Lan, *et al*. 2003b attempts to use the change in chaotic traffic parameters to examine the existence of traffic incident. Takens' embedding theorem is used to reconstruct the traffic flow time series data in the phase space. If an incident occurs, some of the chaotic traffic parameters might change drastically; thus, we can select the chaotic parameters with significant change to discriminate the incident traffic flow from the normal one.

2.2 Microscopic AID Algorithms

Microscopic AID algorithms are divided into low volume approach (Fambro and Ritch, 1980), probe vehicle method (Sethi, *et al*., 1995; Parkany and Bernstein, 1993; Ivan, *et al*., 1993; Ivan, *et al*., 1995; Ivan, 1997) and image processing based approach (Huang, 1996).

2.2.1 Low Volume Approach

The low volume approach is proposed by Fambro and Ritch (1980) to determine that a vehicle has stopped anywhere between outstations. Should a vehicle stop over a loop then a simple timeout condition raises an alarm. Between loop sites the stopped vehicle algorithm uses an accounting procedure for consecutive sites which predicts a time period for a vehicle's arrival at the downstream site. If the vehicle fails to arrive in its allocated time period, then it is assumed to have stopped between the two sites. Predictions are based on the vehicle's speed and make allowance for a certain amount of acceleration and deceleration. In correlating a vehicle to its time period when it arrives at a site, a match is also sought for vehicle length and speed within predetermined tolerance, allowance is also made within the algorithm for lane changing. A separate algorithm at the outstation discards partial activations and discriminates as far as possible between the presences of one or two vehicles when all four loops are activated most simultaneously. The stopped vehicle algorithm at the in station caters for the few indeterminate cases hence avoiding "a phantom vehicle" at one site generating an alarm because it has not been detected at the next site downstream.

2.2.2 Probe Vehicle Method

Sethi, *et al.* (1995) proposed the incident detection algorithms for urban arterial streets using two distinct data sources: fixed traffic detectors (volume and occupancy) and probe vehicles (travel time). The data sources are used independently to obtain two distinct algorithms. This approach is undertaken to increase the overall coverage of incident detection capabilities as easily implementation will result in relatively few cases when data is available from both fixed detectors and probe vehicles on the same link and during the same time period. The performance of the algorithms was evaluated using detection rates and false alarm rates, which were found to be in the same range for both the algorithms. The fixed detector algorithm showed better detection ability, but its use is limited by the number of detectorized links in the network, while performance of the probe vehicle algorithm was dependent on the number of reports available per time period.

Ivan (1997) developed a new technique based on data fusion methods using multiple data sources: inductive loop detectors, and travel times collected from probe vehicle traveling through the street network.

2.2.3 Image Processing Based Approach

Huang (1996) focused on real-time area wide vehicle detection and tracking by using computer vision with PC architecture to develop real-time area wide vehicle detection systems by using a PC-based image processing technique. In comparison with point or line detection of vehicles, area wide detection has the capability of monitoring additional traffic parameters. A parallel processing framework named TVS is established to meet the above-mentioned requirement. Besides TVS, another system named HAWK is also established. The objective of HAWK is to do vehicle tracking by using correlation coefficient method with Pentium framework. In comparison with TVS, HAWK uses more advanced technology and algorithm. Finally, HAWK system is tested for automatic instant incident detection. By locking vehicles and tracking their movements, a new technique is developed for detecting traffic incidents. Because of locking, this research can evaluate the traffic performance by the traveling time, including moving and delay time, of locked target. The research has built TIGER (Traffic Image interGrated Evaluation and Response) procedure to evaluate performance by using locked vehicle's traveling time and confirm incident by using locked vehicle's trajectory in area-wide camera detection field.

2.3 Some Comments

Early incident detection algorithms focused on pattern comparison methods using raw traffic data gathered by loop detectors. Recently, various advanced techniques, such as artificial intelligent based algorithms, fuzzy logic, fuzzy expert systems, genetic algorithms, and image processing technologies, have been proposed and tested to achieve real-time and more precise incident detection. However, the above-mentioned incident detection algorithms may have encountered one or some of the following problems: (a) The detection performance is subject to the algorithm thresholds, traffic flow conditions (medium and heavy traffic flows normally have higher detection rate), the distance between two adjacent detectors and roadway geometry. Additionally, the 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 pre-defined crisp thresholds and do not permit the utilization of time-varying parameters in distinguishing incident traffic from the normal traffic, thus they may result in poor detection performance as the traffic conditions alter drastically. (b) Most detection algorithms are not transferable in that parameters and thresholds of an algorithm must be re-calibrated and re-validated to be valid for different locations or times. Notice that these algorithms are strong site-related. (c) Quantity and quality of traffic flow data are subject to detector types (some detector data sets are unrealistic and unusable due to detector malfunction and some interferes on sensors and communication media). Flow and occupancy parameters collected by inductive loop detectors, for instance, are often limited with lower accuracy. In addition, it is difficult to reflect the section-related traffic flow behavior only depending on spot traffic parameters obtained from loop detectors. (d) Most algorithms, which only focus on detecting whether there is an incident without further identifying the location or severity of an incident, are passive incident detection in nature. The merely incident occurrence output have proved insufficient information and can not satisfy the demand in term of incident-related traffic conditions. It is essential for incident management to predict the incident-related traffic characteristics. (e) Because of the difficulty in gathering actual incident data, most algorithms adopt off-line validation with simulation data or incident database. The following proposed incident detection algorithms are introduced to challenge some of these issues.