Incident-Detection Algorithms

Part 1. Off-Line Evaluation

Moshe Levin and Gerianne M. Krause, Bureau of Materials and Physical Research, Illinois Department of Transportation, Oak Park

Five incident-detection algorithms of the pattern-recognition type were evaluated off-line by using incident and incident-free data collected on Chicago's expressways under various traffic and environmental conditions. Algorithm efficiency was evaluated in terms of detection and false-alarm rates and mean-time-to-detect. Evaluated were a comparative analysis of algorithm efficiency, the effect of lateral detectorization on algorithm performance, a hierarchical analysis of threshold effectiveness, and the effect of incident severity on algorithm performance. Although no specific algorithm was found to be superior for levels of detection lower than 95 percent, for higher levels of detection one algorithm developed by Technology Services Corporation was found to be best. The algorithms did not differ statistically in mean-time-to-detect, which ranged from 2 to 4 min, rendering this parameter ineffective in algorithm selection. The relation between detection rate and false-alarm rate, however, was found to be the critical criterion for algorithm selection. Feature thresholds developed for detector-lane incidents were found to be less sensitive to traffic-flow disturbances than were thresholds developed for non-detector-lane incidents, thus yielding lower false-alarm rates. Analysis of algorithm performance under various traffic and environmental conditions revealed that thresholds developed for a representative sample of incidents were effective when used on the "rush wet", "nonrush dry", and "nonrush wet" traffic data. Therefore, less effort was needed to develop the set of thresholds. Thresholds developed for accidents occurring on the detector lane proved to be effective in detecting accidents and nonaccident incidents on both the detector and non-detector lanes.

Freeway incident-management systems that offer various levels of service to the motoring public have been in operation for quite some time. In essence, each such system provides some or all of the following system elements:

1. Detection of traffic-flow abnormalities,
2. Incident identification,
3. Traffic-management strategies and tactics to be implemented through driver communication and control subsystems, and
4. Early removal of incidents by motorist-aid subsystems.

The comprehensiveness of the incident-management system and the level of sophistication of its elements will determine the operational efficiency of the system.

A key element of such a system is the detection of traffic-flow abnormalities and their identification as capacity-reducing incidents. A positive identification will normally activate the control, driver communication, and incident-handling subsystems. Obviously, a missed incident or a false alarm will affect the efficiency of the management system and its credibility. But the incident-detection process uses algorithms that relate certain measured relations among traffic characteristics to calibrated thresholds to yield a decision with regard to the incident.

The Federal Highway Administration (FHWA) contracted with Technology Services Corporation (TSC) to evaluate existing algorithms (1) and to develop new ones (2). The evaluation included pattern recognitions (3, 4) and time-series algorithms (5, 6). The Illinois Department of Transportation (IDOT) has assumed the task of the off-line and on-line evaluations of the selected algorithms developed by TSC. The facilities of IDOT's Traffic Systems Center will be used for this.

The specific objectives of the research reported here were

1. To determine the efficiency of the selected TSC algorithms in detecting incidents on the Chicago-area expressway system for various traffic and environmental conditions,
2. To develop algorithm thresholds compatible with the traffic characteristics of the expressway system and various environmental conditions,
3. To determine the effect of the existing level of detectorization on algorithm performance,
4. To determine the effect of incident severity on algorithm performance, and
5. To compare the efficiency of TSC algorithms with a pattern-recognition algorithm developed locally.

ALGORITHM DESCRIPTION

This section describes the structure of the incident-detection algorithms evaluated in this research. They include five pattern-recognition algorithms, four of which were developed by TSC (2); the fifth was developed locally in the course of the research.

The research effort of TSC included the development of 10 incident-detection algorithms. Algorithms 1-7 are variations on the classic California algorithm (3), while 8 and 9 use, in addition to those elements of algorithm 7, a feature that suppresses incident detection at any station for 5 min after detection of a compression wave at the downstream station. Algorithm 10 attempts to detect incidents occurring in light-to-moderate traffic that do not lower capacity below the volume of oncoming traffic.

Of these 10 algorithms, 4 were selected for evaluation: algorithms 7, 8, 9, and 10. Preliminary investigation indicated algorithm 7 to be a superior form of the California algorithm. Algorithm 8, which is identical to algorithm 9 except for an added persistence check, was found to have, according to TSC's investigation, a slightly lower false-alarm rate (FAR) but a longer mean-time-to-detect (MTTD) than algorithm 9. Although algorithm 10 did not perform especially well, it was included in the off-line evaluation because it represents a first attempt to solve the problem of detecting incidents that do not produce marked traffic-flow discontinuities.

The TSC algorithms are in binary decision-tree form; at each node of the decision tree a feature value is compared with a user-specified threshold value to determine whether an incident is to be signalled. Obviously the effectiveness of the algorithm depends on the thresholds chosen.

TSC developed a program for optimizing threshold selection. This program, called CALB, uses a random-number generator that produces increments to be added to the current optimal threshold vector to produce a new threshold vector for evaluation. After a predetermined number of iterations, the threshold vector with the lowest false-alarm rate, given a certain level of detection, is termed the optimal threshold vector at that level of detection.

Before CALB was used to calibrate the algorithms for the off-line evaluation, a detailed study was performed to determine how best to set certain user-supplied parameters needed by CALB in the algorithm calibration process. The point was to ensure selection of the best threshold vectors for use in the algorithm evaluation.

Finally, the four TSC algorithms selected were compared with algorithm 16-14, one in a series of pattern-recognition algorithms developed in the course of this research (7).

Following is a detailed description of the above algorithms; the meanings of the features involved in each algorithm are given in the listing below.

Feature Name	Definition
OCC(t)	Minute average occupancy measured at upstream detector at time t
DOCC(t)	Minute average occupancy measured at downstream detector at time t
OCCDF(t)	OCC(t) - DOCC(t)
OCCRDF(t)	OCCDF(t)/OCC(t)
SPEED(t)	Minute average speed calculated at upstream detector at time t
DOCCTD(t)	[DOCC(t-2) - DOCC(t)] /DOCC(t-2)
SPDTDF(t)	[SPEED(t-2) - SPEED(t)] /SPEED(t-2)
OCCRDF(t-1)	[OCC(t-1) - DOCC(t-1)] /OCC(t-1)
UPDF(t)	OCC(t-1) - OCC(t-2)
UPRDF(t)	UPDF(t)/OCC(t-1)
DNDF(t)	DOCC(t-2) - DOCC(t-1)
DNRDF(t)	DNDF(t)/DOCC(t-2)
DPDNDF(t)	UPDF(t) = DNDF(t)
UPDNR1(t)	UPDNDF(t)/OCC(t-1)
UPDNR2(t)	UPDNDF(t)/[OCC(t-1) - DOCC(t-1)]
RDF(t)	OCCDF(t)/[OCC(t-1) - DOCC(t-1)]

Algorithm 7 differs from the classic California algorithm in the following three ways. Whereas the California algorithm produces an incident signal whenever OCCDF, OCCRDF, and DOCCTD are greater than associated thresholds, algorithm 7 replaces DOCCTD with DOCC, suppresses incident signals after the initial detection, and contains a persistence requirement that OCCRDF be greater than the threshold for two consecutive minutes (Figure 1).

Algorithm 9 consists of algorithm 4 (a variant of the California algorithm) coupled with a compression-wave check and uses features DOCC and DOCCTD. It works as follows. First, a compression-wave check is made. If it succeeds, then algorithm 4 is not applied until five consecutive minutes have passed without a compression wave. If it fails then algorithm 4 is immediately applied.

Algorithm 8 is algorithm 9 with an OCCRDF-persistence requirement added. It can also be thought of as algorithm 7 incorporated with the 5-min compression-wave check (Figure 2).

Algorithm 10 separates traffic data into light, moderate, and heavy traffic by using the feature OCC. No incident check is applied to light-traffic data. Algorithm 7 is used under heavy-traffic conditions, and under moderate conditions OCCRDF and SPDTDF, a temporal speed change feature, are applied (Figure 3).

Algorithm 16-14 is a pattern-recognition algorithm developed locally by using occupancy-based features obtained through intensive observations of traffic behavior on different parts of the Chicago-area expressway system (Figure 4).

DEVELOPMENT OF DATA BASE

The data base was divided into two parts: incident data used to compute an algorithm's detection rate (DR) and MTTD and incident-free data used to calculate an algorithm's FAR. The surveillance data that make up each set consist of 20-s occupancies and volumes from each main-line detector on the relevant directional freeway. The data base includes a total of 100 incident and 14 incident-free data sets.

In the collection of incident data sets, "incident" was limited to mean unplanned physical obstructions of the traveled lanes. The incident data were collected by monitors at IDOT's Traffic Systems Center. Indications of a potential incident came in two ways. In the most common case, the data collector would spot a disturbance in the traffic-stream variables by monitoring the expressway system map panel, occupancy maps on the display, or typer output of the surveillance system. In these cases, the monitor would activate a program for saving the surveillance data from the affected directional expressway (the data-collection program kept a 30-min historical file of surveillance data that enabled the requisite 15 min of pre-incident data to be saved, if an incident was detected by the monitor within 15 min of its occurrence). The monitor then requested the IDOT Communication Center to dispatch an emergency patrol vehicle (EPV) to the area for confirmation and identification. In other cases, an incident would be reported by a field unit before signs of it appeared in the surveillance data. When traffic-stream measurements began to manifest signs of the incident's effect on traffic operations, data saving was initiated.

The incident data were collected to represent the following factors:

1. Rush or nonrush traffic conditions (R, NR),
2. Wet or dry pavement conditions (W, D),
3. Accident (10-50) or nonaccident incident (10-46) incident type (AI, NAI) according to Illinois State Police code, and
4. Detector lane or non-detector-lane (DL, NDL) incident lateral location.

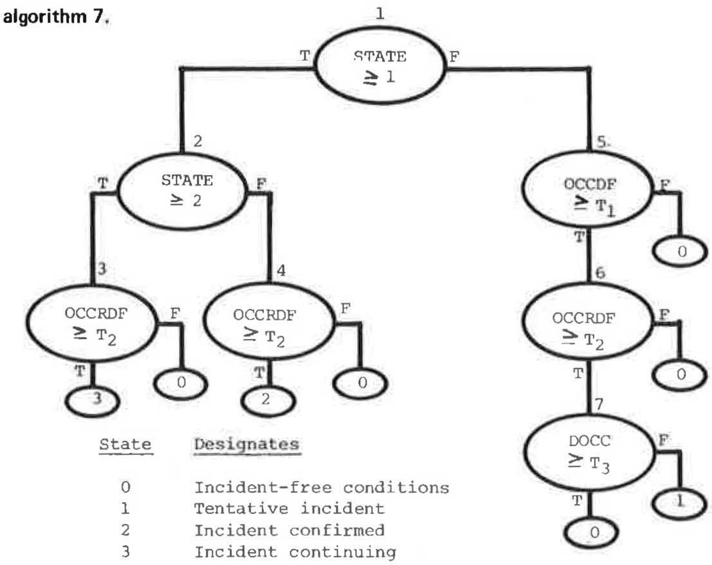
Figure 5 shows the stratification of the incident data and the code of each stratum. The meanings of the codes are explained below.

Code	Interpretation
R	Rush
RW	Rush wet
RD	Rush dry
RD-0	Incident occurring on nondectector lanes during rush dry period
RD-1	Incident occurring on detector lane during rush dry period
RD-50-1	Accident occurring on detector lane during rush dry period
RD-50-0	Accident occurring on nondetector lanes during rush dry period
RD-46-1	Nonaccident incident occurring on detector lane during rush dry period
RD-46-0	Nonaccident incident occurring on nondetector lanes during rush dry period
RD-50	Accident occurring during rush dry period
RD-46	Nonaccident incident occurring during rush dry period

NR, NRW, NRD, NRD-0, NRD-1, NRD-50-0, NRD-46-0, NRD-50-1, and NRD-46-1 have the same interpretation as above except that they refer to the nonrush period.

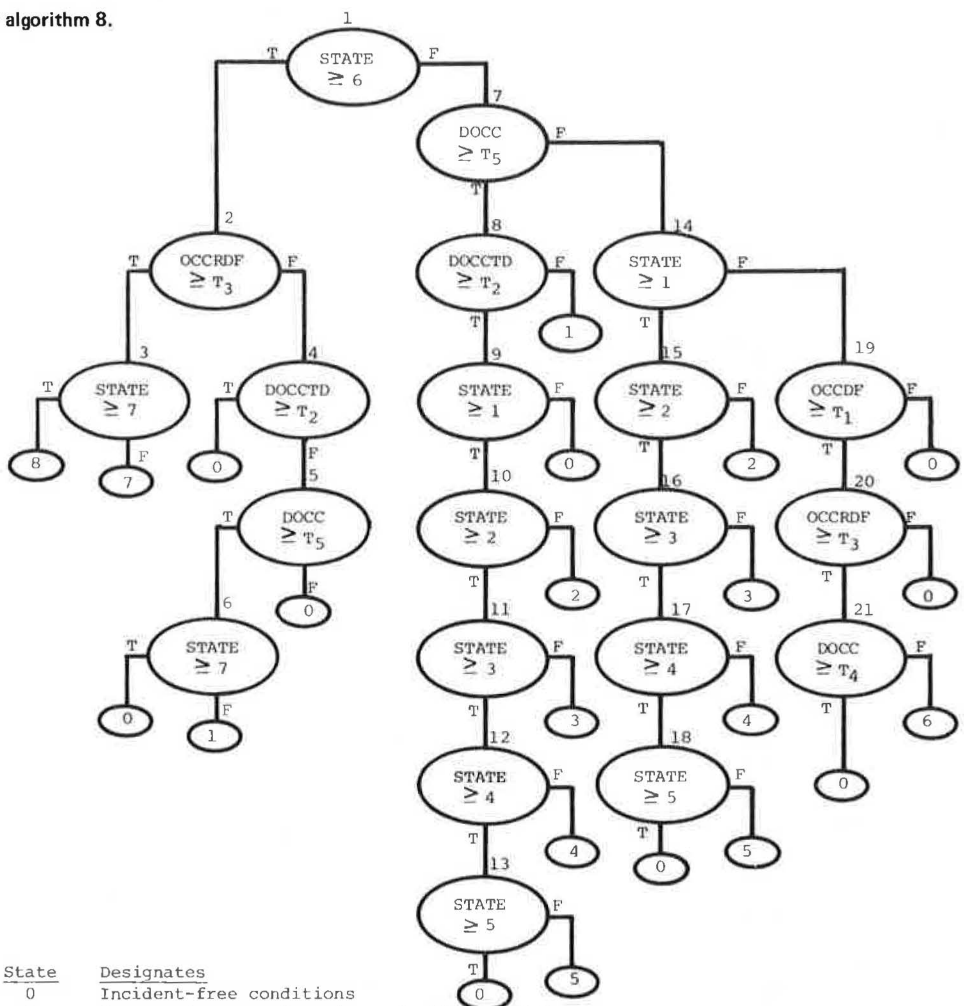
The collection of incident-free data sets involved the use of the same data-saving software as employed in the incident data collection. Verification of these data as incident-free was carried out with the use of a helicopter. Nearly 30 h of incident-free data were collected to appropriately represent rush, nonrush, wet, and dry conditions.

Figure 1. Decision tree for algorithm 7.



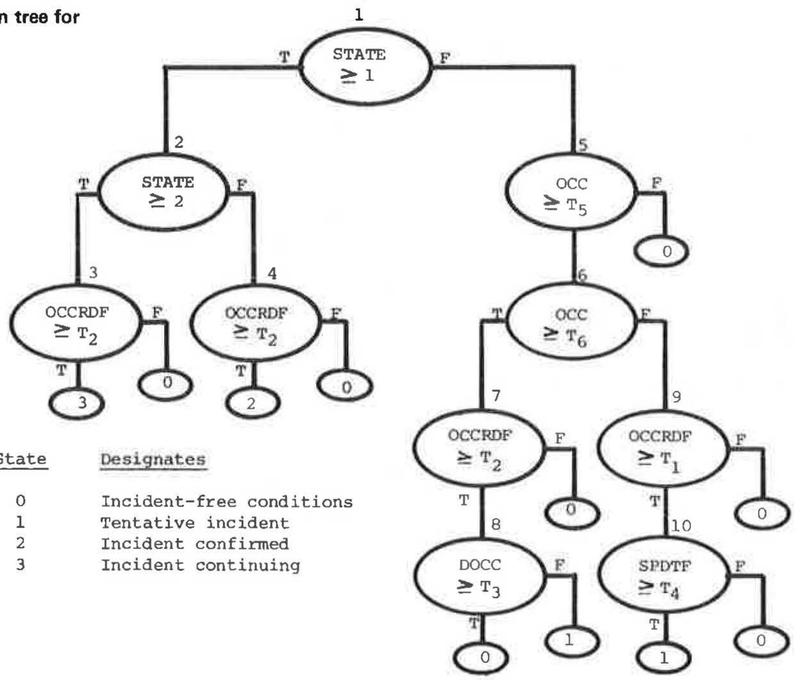
State	Designates
0	Incident-free conditions
1	Tentative incident
2	Incident confirmed
3	Incident continuing

Figure 2. Decision tree for algorithm 8.



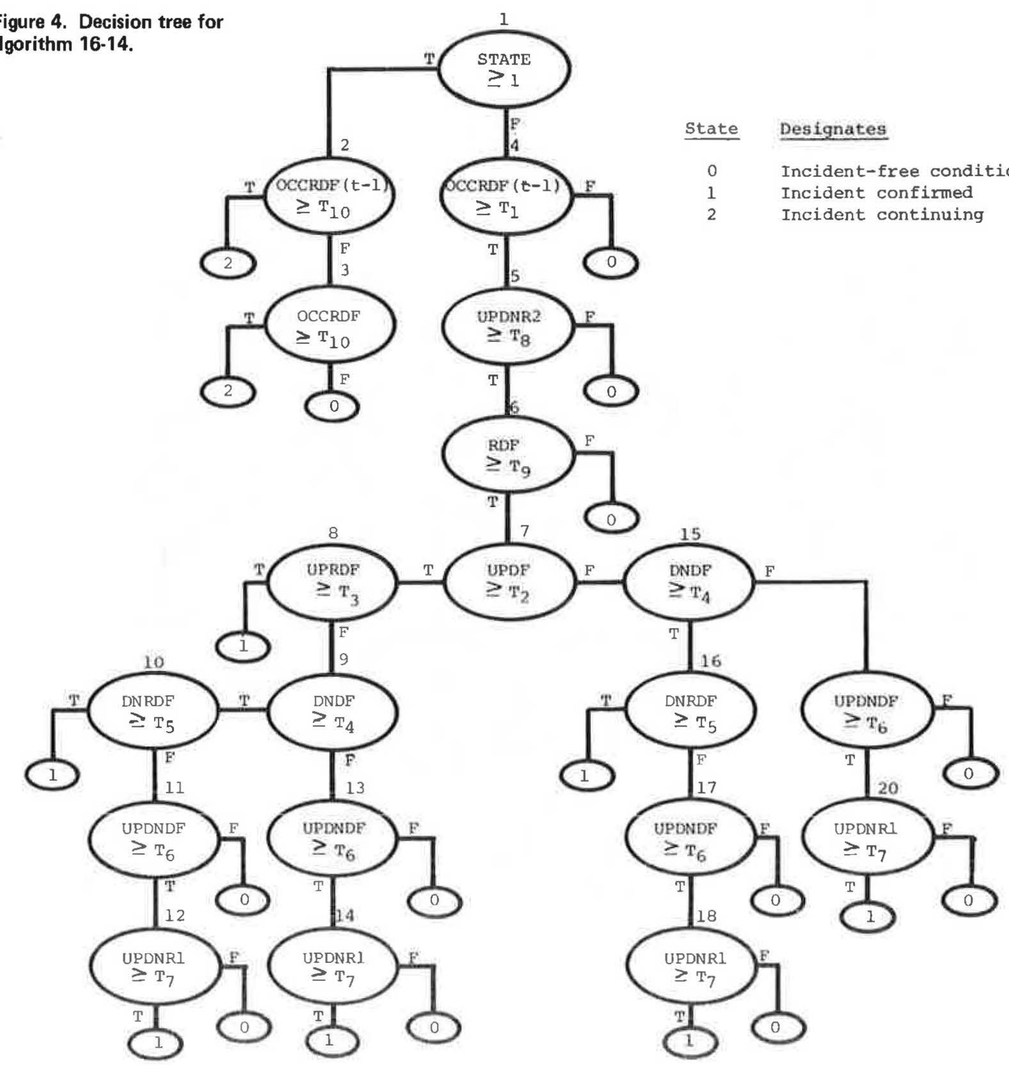
State	Designates
0	Incident-free conditions
1	Compression wave downstream in this minute
2	Compression wave downstream 2 minutes ago
3	Compression wave downstream 3 minutes ago
4	Compression wave downstream 4 minutes ago
5	Compression wave downstream 5 minutes ago
6	Tentative incident
7	Incident confirmed
8	Incident continuing

Figure 3. Decision tree for algorithm 10.



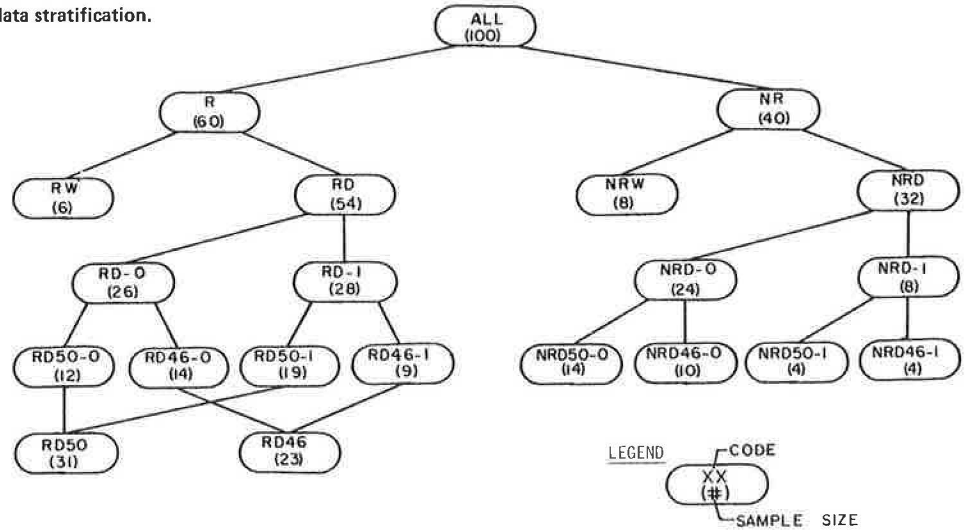
State	Designates
0	Incident-free conditions
1	Tentative incident
2	Incident confirmed
3	Incident continuing

Figure 4. Decision tree for algorithm 16-14.



State	Designates
0	Incident-free conditions
1	Incident confirmed
2	Incident continuing

Figure 5. Incident-data stratification.



OFF-LINE EVALUATION

The ultimate goal of the off-line evaluation was to obtain for the tested algorithms optimal sets of thresholds related to various traffic and environmental conditions. These sets could then be implemented in an operational on-line incident-response system. To achieve that goal off-line evaluation was divided into four major tasks:

1. Comparative analysis of algorithm efficiency,
2. Evaluation of the effect of lateral detectorization on algorithm performance,
3. Hierarchy analysis of the threshold effectiveness, and
4. Evaluation of the effect of incident severity on algorithm performance.

Algorithm efficiency could be determined by three related parameters:

1. DR: percentage of detected incidents out of all incidents that affect traffic and occur during a specified time period;
2. FAR (off-line definition): percentage of incident messages (1s) out of all messages (1s and 0s) where messages are produced at specific intervals (i.e., every 1 min) out of representative incident-free data; and
3. MTTD: the mean delay between the apparent occurrence of incidents, as estimated from changes in upstream and downstream occupancy values, and their detection time by the algorithms during a certain period of time.

Comparative Analysis of Algorithm Efficiency

The comparative analysis of the tested algorithms was performed by running each of the five algorithms—7, 8, 9, 10, and 16-14—through the various incident and incident-free data strata, by using TSC's CALB program, which had been modified for the Traffic Systems Center's computer. The CALB evaluation of these algorithms was performed for six nominal detection rates of 75, 80, 85, 90, 95, and 99 percent and for three incident data categories: ALL, RD, and NRD. The strata of RW and NRW included only six and eight incident cases, respectively, and were excluded from the detailed analysis.

A comparison of the DR-FAR relationships of algorithms 9 and 16-14 with those of algorithms 7, 8, and 10 indicated that algorithms 9 and 16-14 experienced relatively high FAR across the whole DR spectrum. At the same time, however, their DR-MTTD relationships seemed to be more favorable than those of the other algorithms. However, because in many cases the differences in MTTD for the various algorithms were not found to be statistically significant, the relatively poor DR-FAR relationships between algorithms 9 and 16-14 suggested their elimination from further analysis even though favorable results were indicated for algorithm 9 (2). However, for the sake of representative analysis and future on-line evaluation, it was decided to eliminate only algorithm 9.

Overall, the three algorithms (7, 8, 10) produced better DR-FAR relationships for the NRD category than for the RD category. Over the investigated range of the DR, the FAR for the NRD category ranged from 0.00 to 0.01 percent, while the range for the RD category was from 0.02 to 0.11 percent.

Within the RD category no single algorithm displaying invariably better FARs over the DR spectrum could be found. However, for the higher DRs (0.95 and above), algorithm 7 was the most efficient. Also, the same algorithm was found to yield the fewest FARs over the whole DR spectrum for the NRD category.

The time-to-detect analysis used the optimal sets of thresholds developed for the DR-FAR relationships. The MTTD for the RD and NRD categories ranges from 1.9 to 4.4 min and from 3.6 to 6.2 min, respectively. The results for the ALL category (2.2-4.6 min) represent, to a large extent, the combinations of the RD and NRD results.

Within the RD category, algorithm 7 displayed the lowest MTTD for DRs higher than 95 percent. For lower DRs, no single, most efficient algorithm could be found. Within the NRD category no single algorithm displayed invariably lower MTTD over the whole DR spectrum.

Further insight into the differences in MTTD between algorithms for the various incident data categories was gained from the Kolmogorov-Smirnov test and the Mann-Whitney U-test (8) for thresholds representing the 95 percent detection level. This level was selected for its assumed applicability to an operating on-line system. The results of the statistical analyses for algorithms 7, 8, 10, and 16-14 are presented in Table 1. From this table it can be seen that, as far as the MTTD is concerned, no statistically significant difference (0.05 level

Table 1. Comparison of algorithm performance at 95 percent detection rate for ALL, RD, RW, NRD, and NRW conditions.

Traffic Category	Sample Size	Algorithm No. 7	8	10	16-14	Apparent Best Algorithm	Statistically Best Algorithm (for MTTD)
ALL	99						
FAR, %		0.019	0.0297	0.0231	0.11	7	
MTTD, min		3.39	2.85	3.68	2.28	16-14	None[a]
SD, min		3.25	3.01	3.42	3.05		
RD	54						
FAR, %		0.056	0.0786	0.067	0.26	7	
MTTD, min		2.23	2.75	2.88	1.26	16-14	None[a]
SD, min		1.60	2.15	2.65	1.83		
RW	6						
FAR, %		0.0336	0.0	0.045	0.045	8	
MTTD, min		2.83	3.99	2.50	2.33	10	None[a,b]
SD, min		0.69	2.89	5.02	3.03		
NRD	32						
FAR, %		0.002	0.002	0.005	0.018	7	
MTTD, min		3.73	3.56	2.87	3.22	10	None[a]
SD, min		3.75	3.77	2.49	4.72		
NRW	8						
FAR, %		0.005	0.005	0.005	0.009	7, 8, 10	
MTTD, min		2.71	2.63	2.50	1.88	16-14	None[a,b]
SD, min		2.31	1.99	2.24	2.15		

[a]Kolmogorov-Smirnov test at the 0.05 level of significance. [b]Mann-Whitney U-test at the 0.05 level of significance.

Table 2. Comparison of algorithm performance at 95 percent detection rate for RD and NRD conditions.

Traffic Category	Sample Size	Algorithm No. 7	8	10	16-14	Apparent Best Algorithm	Statistically Best Algorithm (for MTTD)
RD-1	28						
FAR, %		0.0449	0.0449	0.0336	0.112	10	
MTTD, min		2.96	2.69	3.18	1.26	16-14	16-14[a,b]
SD, min		1.93	1.93	3.49	1.14		
RD-0	26						
FAR, %		0.0561	0.0673	0.0673	0.112	7	
MTTD, min		2.28	2.32	3.07	2.56	7	None[a,b]
SD, min		1.84	1.91	3.09	2.22		
NRD-1	8						
FAR, %		0.0	0.0	0.0	0.014	None	
MTTD, min		2.12	2.12	2.12	2.75	None	None[a,b]
SD, min		1.89	1.89	1.89	2.5		
NRD-0	24						
FAR, %		0.0	0.0047	0.0094	0.014	7	
MTTD, min		4.08	4.04	4.08	4.13	None	None[a,b]
SD, min		4.06	4.16	2.74	5.69		

[a]Kolmogorov-Smirnov test. [b]Mann-Whitney test.

of significance) was found between the algorithms at the 95 percent detection level for all the incident categories.

It seems, then, that the DR-FAR relationship is more representative of the difference among algorithms than the DR-MTTD relationship and should be the major criterion for selecting algorithms.

Based on the results in Table 1, algorithm 7 was the apparent best for the ALL, RD, NRD, and NRW categories at the 95 percent detection level, while algorithm 8 was the apparent best for the RW category at the same detection level.

Evaluation of the Effect of Lateral Detectorization on Algorithm Performance

In the design process of a freeway surveillance and control system there is always the question of a trade-off between the level of detectorization (longitudinal and lateral) and the gains in terms of control and incident-detection effectiveness.

The Chicago expressway system under surveillance uses full detector stations every 4.8 km (3 miles) and single-detector stations, usually on lane 2 (lane 1 being the inner lane), every 0.8 km (0.5 mile). The analysis presented in this section compares the performance of algorithms 7, 8, 10, and 16-14 as related to incidents occurring on the detector lane (DL) versus those occurring on the nondetector lanes (NDL) under RD and NRD conditions. The results suggest that for both conditions the optimal thresholds obtained for incidents occurring on DL are less sensitive to discontinuities in traffic flow, as expressed in lower FAR, than those obtained for incidents occurring on NDLs.

This is explained by the fact that, generally, incidents occurring on DL have higher feature values that require less sensitive thresholds, which lower FAR. Incidents occurring on NDL have a somewhat attenuated impact when measured off another lane; this requires more sensitive thresholds (lower value) and risks a high FAR.

For the RD category, the relationship between the DR and MTTD is more favorable for incidents occurring on NDLs than for those occurring on DL. This trend could be explained by the fact that FAR increases with DR, while MTTD decreases with DR, which yields a decrease in MTTD with an increase in FAR. Thus, for a certain DR, the FAR on the DL is higher than the one experienced on NDL, which yields a higher MTTD. This, however, is not the case for the NRD category. The reason could be the small sample of incidents (eight) occurring on DL in the NR category.

In order to find out whether there was a statistically significant difference between the MTTD for incidents on DL and for those on the NDL for both RD and NRD categories, the Kolmogorov-Smirnov test was conducted (95 percent detection level). For RD and NRD categories, tests were made for algorithms 7 and 10, respectively, because each was the most efficient algorithm at that detection level. According to the Kolmogorov-Smirnov test, no significant differences between MTTD were

Table 3. Comparison of algorithm performance at 95 percent detection rate for RD conditions.

Traffic Category	Sample Size	Algorithm No. 7	8	10	16-14
RD-50-1	18				
FAR, %		0.0225	0.0225	0.0562	0.0337
MTTD, min		4.94	4.83	2.05	2.77
RD-50-0	12				
FAR, %		0.0562	0.0562	0.0786	0.1123
MTTD, min		2.92	3.83	3.08	2.41
RD-46-1	9				
FAR, %		0.0562	0.0562	0.0562	0.1123
MTTD, min		2.11	2.22	2.11	1.89
RD-46-0[a]	14				
FAR, %		0.1123	0.0786	0.0562	0.1235
MTTD, min		0.93	1.35	3.21	2.92

[a] This was the only category that displayed a significant difference.

Table 4. Effect of incident severity on algorithm performance.

Traffic Category	Sample Size	Algorithm No. 7	8	10	16-14
RD-46	23				
FAR, %		0.056	0.078	0.078	0.112
MTTD, min		2.31	2.36	2.36	2.27
RD-50	30				
FAR, %		0.056	0.078	0.078	0.112
MTTD, min		2.17	2.53	2.53	1.59
RD-46-1	9				
FAR, %		0.045	0.045	0.045	0.112
MTTD, min		3.34	3.44	2.89	1.89
RD-50-1	18				
FAR, %		0.022	0.056	0.045	0.033
MTTD, min		4.94	2.05	3.22	2.77
RD-46-0	14				
FAR, %		0.112	0.078	0.056	0.123
MTTD, min		0.93	1.35	3.21	2.92
RD-50-0	12				
FAR, %		0.056	0.056	0.078	0.112
MTTD, min		2.92	3.83	3.08	2.41

found for RD and NRD categories at the 0.10 level of significance.

The above analyses suggest that the relation between DR and FAR is more critical than that between DR and MTTD.

As to the relative performance of the individual algorithms within the various incident data categories, Table 2 presents, for the 95 percent level of detection, the MTTD, the standard deviation of the detection time, and the FAR for algorithms 7, 8, 10, and 16-14 and for the incident data categories RD-1, RD-0, and NRD-0. The Kolmogorov-Smirnov and Mann-Whitney tests were conducted for significant differences in MTTD. The results of these tests are also presented in terms of the statistically best algorithm compared with the apparent best. According to these tests, no single algorithm proved to be superior to the others with respect to MTTD for the RD-0, NRD-1, and NRD-0 categories. Algorithm 16-14, however, proved to be the best for the RD-1 category. Considering FAR, algorithm 10 seemed to be the best for the RD-1 category, while algorithm 7 excelled in the RD-0 and NRD-0 categories. No apparent best algorithm was found for the NRD-1 category.

Additional analysis was made of the differences in FAR and MTTD for accident and nonaccident incidents (AI and NAI) occurring on both DL (50-1, 46-1) and NDL (50-0, 46-0). Optimal thresholds were obtained for each particular situation. The analysis included tests for significant differences in MTTD among and within the RD for algorithms 7, 8, 10, and 16-14 at the 95 percent detection level using the Kolmogorov-Smirnov and Mann-Whitney tests at the 0.05 level of significance. The results of this analysis are presented in Table 3.

From this table it can be seen that, as far as MTTD was concerned, there was no significant difference for AI and NAI that occurred on either DL or NDL for each of the tested algorithms. Also, no significant differences in MTTD were found among algorithms within the categories RD-50-1, RD-50-0, and RD-46-1. Algorithm 7, however, was found to be the best within the RD-46-0 category.

As far as FAR was concerned, thresholds that were developed for AI and NAI occurring on DL yielded equal or better results than thresholds developed for AI and NAI occurring on NDL for all the tested algorithms. This is to be expected, because thresholds for detecting incidents on DL could be less sensitive to discontinuities in traffic flow than thresholds for incidents on NDL.

With regard to the individual categories, algorithms 7 and 8 performed the best for RD-50-1, RD-50-0, and RD-46-1, whereas algorithm 10 excelled in the RD-46-1 category. The local algorithm 16-14 yielded relatively high FAR for all categories tested.

The above results indicate that MTTD, unlike the FAR, did not prove to be a major criterion in the selection of algorithms.

It seems that, in order to generate low FAR, thresholds developed for incidents on DL should be used even though the probability of incident occurrence is naturally higher on NDL than on DL. However, these less sensitive thresholds would reduce the rate of detection of incidents occurring on the NDL.

Evaluation of the Effect of Incident Severity on Algorithm Performance

One of the considerations in selecting a particular set of thresholds for the operation of a certain algorithm could be its relative effectiveness in detecting AI and NAI, which usually differ in their impact on traffic flow. As shown previously, thresholds for incidents occurring on DL are less sensitive in terms of FAR than those for incidents occurring on NDL. However, the effectiveness and efficacy of thresholds developed separately for AI and NAI are yet to be evaluated.

Table 4 presents a comparison of MTTD and FAR, at the 95 percent detection level, for algorithms 7, 8, 10, and 16-14, between AI and NAI occurring either on DL or NDL or on both. As can be seen from Table 4, as far as MTTD was concerned, the Kolmogorov-Smirnov and Mann-Whitney tests did not show any significant difference at the 0.05 level. As far as FAR was concerned, thresholds that were developed for the accident data performed better than those developed for the nonaccident data in all cases. This, of course, is predictable, because AI would have a greater disruptive impact on traffic flow than NAI would.

The question that remains to be answered concerns the effectiveness of thresholds developed for AI in detecting NAI. Analysis showed that thresholds developed for accident data on DL at the 95 percent detection level detected only 78 percent of NAI on that lane for algorithms 7 and 8 (FAR = 0.22 percent) and detected all NAI for algorithm 10 (FAR = 0.56 percent). It seems that, if FAR is the major criterion, then thresholds developed for accidents (RD-50-1) could be used to detect other incidents (RD-46-1). This also holds true for RD-46-0 and RD-50-0 for algorithms 7, 8, and 10.

Hierarchy Analysis of Threshold Effectiveness

The effort involved in developing the input necessary for

Table 5. Threshold hierarchy analysis.

	Algorithm No.								
	7			8			10		
Thresholds Compared	DR	FAR	MTTD	DR	FAR	MTTD	DR	FAR	MTTD
ALL on RD v. RD on RD	0.92	0.056	3.40	0.90	0.067	2.52	0.93	0.056	3.25
	0.96	0.056	2.23	0.96	0.078	2.75	0.95	0.067	2.88
RW on RD v. RD on RD	0.85	0.034	3.63	0.64	0.000	5.43	0.81	0.045	4.21
	0.96	0.056	2.23	0.96	0.078	2.75	0.95	0.067	2.88
NRD on RD v. RD on RD	0.92	0.056	3.36	0.83	0.045	2.86	0.87	0.056	2.21
	0.96	0.056	2.23	0.96	0.078	2.75	0.95	0.067	2.88
NRW on RD v. RD on	0.92	0.056	3.40	0.83	0.045	2.86	0.87	0.056	2.21
RD	0.96	0.056	2.23	0.96	0.078	2.75	0.95	0.067	2.88
ALL on RW v. RW on	1.00	0.056	2.33	1.00	0.067	2.33	1.0	0.056	2.50
RW	1.00	0.034	2.83	1.00	0.000	3.99	1.0	0.045	2.50
RD on RW v. RW on	1.00	0.056	2.16	1.00	0.078	2.16	1.0	0.067	2.50
RW	1.00	0.034	2.83	1.00	0.000	3.99	1.0	0.045	2.50
NRD on RW v. RW on	1.00	0.056	2.21	0.84	0.044	2.80	1.0	0.056	1.99
RW	1.00	0.034	2.83	1.00	0.000	3.99	1.0	0.045	2.50
NRW on RW v. RW on	1.00	0.056	2.33	0.84	0.044	2.80	1.0	0.056	1.99
RW	1.00	0.034	2.83	1.00	0.000	3.99	1.0	0.045	2.50
ALL on NRD v. NRD on	1.00	0.005	3.78	0.96	0.014	3.61	1.0	0.009	4.28
NRD	0.96	0.005	3.93	0.96	0.005	3.67	0.96	0.005	2.87
RD on NRD v. NRD on	1.00	0.014	3.46	1.00	0.023	3.46	0.93	0.019	4.23
NRD	0.96	0.005	3.93	0.96	0.005	3.67	0.96	0.005	2.87
RW on NRD v. NRD on	0.90	0.009	4.38	0.68	0.009	4.27	0.97	0.005	4.71
NRD	0.96	0.005	3.93	0.96	0.005	3.67	0.96	0.005	2.87
NRW on NRD v. NRD on	1.00	0.005	3.78	0.68	0.009	4.27	0.96	0.005	2.87
NRD	0.96	0.005	3.93	0.96	0.005	3.67	0.96	0.005	2.87
ALL on NRW v. NRW on	0.87	0.005	2.71	1.00	0.014	2.62	0.87	0.009	5.14
NRW	0.87	0.005	2.71	1.00	0.005	2.63	1.0	0.005	2.87
RD on NRW v. NRW on	1.0	0.014	2.50	1.00	0.023	2.50	1.0	0.019	6.00
NRW	0.87	0.005	2.71	1.00	0.005	2.63	1.0	0.005	2.50
RW on NRW v. NRW on	0.75	0.009	7.14	0.75	0.009	3.34	0.87	0.009	5.14
NRW	0.87	0.005	2.71	1.00	0.005	2.63	1.0	0.005	2.50
NRD on NRW v. NRW on	0.87	0.005	2.85	1.00	0.005	2.63	1.0	0.005	2.50
NRW	0.87	0.005	2.71	1.00	0.005	2.63	1.0	0.005	2.50
RD on RD-0 v. RD-0 on	0.96	0.056	2.47	0.96	0.078	2.04	0.92	0.067	3.04
RD-0	0.96	0.056	2.28	0.96	0.067	2.32	0.96	0.067	3.07
RD-1 on RD-0 v. RD-0	0.77	0.045	2.45	0.77	0.045	3.15	0.69	0.033	3.72
on RD-0	0.96	0.056	2.28	0.96	0.067	2.32	0.96	0.067	3.07
RD on RD-1 v. RD-1 on	0.96	0.056	1.99	0.96	0.078	1.81	0.96	0.067	2.74
RD-1	0.96	0.045	2.96	0.96	0.045	3.77	0.96	0.033	3.18
RD-0 on RD-1 v. RD-1	0.96	0.056	1.84	0.96	0.067	1.88	0.96	0.067	2.77
on RD-1	0.96	0.045	2.96	0.96	0.045	3.77	0.96	0.033	3.18
RD on RD-46 v. RD-46	0.96	0.056	2.31	0.96	0.078	2.05	0.96	0.067	2.86
on RD-46	0.96	0.056	2.31	0.96	0.078	2.18	0.96	0.078	2.36
RD-50 on RD-46 v.	0.96	0.056	2.31	0.87	0.067	2.44	0.91	0.078	2.56
RD-46 on RD-46	0.96	0.056	2.31	0.96	0.078	2.18	0.96	0.078	2.36
RD on RD-50 v. RD-50	0.97	0.056	2.17	0.97	0.078	1.83	0.93	0.067	2.89
on RD-50	0.97	0.056	2.17	0.97	0.067	2.56	0.97	0.078	2.53
RD-46 on RD-50 v.	0.97	0.056	2.17	0.97	0.078	2.03	0.94	0.078	2.03
RD-50 on RD-50	0.97	0.056	2.17	0.97	0.067	2.56	0.97	0.078	2.53
RD-0 on RD-46-0 v.	0.93	0.056	1.84	0.93	0.067	2.00	1.0	0.067	2.43
RD-46-0 on RD-46-0	0.93	0.056	1.84	1.0	0.078	1.35	1.0	0.056	3.21
RD-46 on RD-46-0 v.	1.0	0.078	1.35	0.93	0.078	1.77	0.93	0.078	2.00
RD-46-0 on RD-46-0	0.93	0.056	1.84	1.0	0.078	1.35	1.0	0.056	3.21
RD-1 on RD-46-1 v.	1.0	0.045	3.34	1.0	0.045	3.44	1.0	0.045	2.89
RD-46-1 on RD-46-1	1.0	0.045	3.34	1.0	0.045	3.44	1.0	0.045	2.89
RD-46 on RD-46-1 v.	1.0	0.056	2.11	1.0	0.078	2.78	1.0	0.078	2.89
RD-46-1 on RD-46-1	1.0	0.045	3.34	1.0	0.045	3.44	1.0	0.045	2.89
RD-0 on RD-50-0 v.	1.0	0.056	2.65	1.0	0.067	2.67	0.92	0.067	3.30
RD-50-0 on RD-50-0	1.0	0.056	2.92	1.0	0.056	3.83	1.0	0.078	3.08
RD-50 on RD-50-0 v.	1.0	0.078	2.58	1.0	0.067	3.42	1.0	0.078	3.08
RD-50-0 on RD-50-0	1.0	0.056	2.92	1.0	0.056	3.83	1.0	0.078	3.08
RD-1 on RD-50-1 v.	0.95	0.045	2.76	0.95	0.045	2.85	0.95	0.045	3.22
RD-50-1 on RD-50-1	0.95	0.022	4.94	0.95	0.022	4.83	0.95	0.045	3.22
RD-50 on RD-50-1 v.	0.95	0.078	1.49	0.95	0.067	1.77	0.95	0.078	2.11
RD-50-1 on RD-50-1	0.95	0.022	4.94	0.95	0.022	4.83	0.95	0.045	3.22

an optimal on-line incident-detection system could be enormous. Part of this effort lies in developing thresholds appropriate for various environmental, geometric, and traffic conditions. In addition, for freeway systems that have low levels of detectorization, the question exists as to whether thresholds representing AI or NAI on either DL or NDL should be used.

This section evaluates the efficiency, in terms of DR, FAR, and MTTD, of applying lower-level thresholds to higher-level incident data categories (i.e., thresholds developed for the ALL category are tested on the RD category). The object of such an analysis is to investigate the possibilities of reducing the amount of effort required to develop the optimal sets of thresholds.

The thresholds for each lower-level incident category were obtained for the 95 percent nominal DR and were applied to a higher-level incident category to yield appropriate values for the other measures of effectiveness. The Mann-Whitney V-test was applied to establish the significance of the difference between MTTD of each two compared incident categories. Table 5 presents the results of this analysis.

As it can be seen from this table, thresholds developed for ALL could be used during the RW period by all algorithms. On the other hand, when used during the RD period, the ALL thresholds yielded reduced DR (algorithms 7, 8, and 10) and also equal or reduced FAR. As far as the MTTD was concerned, the ALL thresholds yielded larger values, which were significantly different, however, for algorithm 10 only. It was also indicated that during the NRD period, as well as during the NRW period, the ALL thresholds could be used quite effectively in algorithms 7 and 10.

The RD thresholds were found to be generally inferior in terms of FAR to those developed for ALL when they were used during the RW, NRD, and NRW periods.

Thresholds developed for the RD category were applied to both the RD-1 and RD-0 categories. In both cases these thresholds were found to be inferior to the thresholds representing the two categories. When RD-1 thresholds were applied to the RD-0 category, FAR improved but DR decreased for all algorithms. When RD-1 thresholds were applied to the RD-50-1 category, there was no change in DR and no significant difference in MTTD. Other threshold hierarchy relations could be easily obtained from Table 5. The few significant differences that appeared in the threshold comparisons are shown below.

Thresholds Compared	Significant Difference (algorithm no.)
RW on RD v. RD on RD	7, 8
RD on NRW v. NRW on NRW	10
RD-50 on RD-50-1 v. RD-50-1 on RD-50-1	7, 10

FINDINGS, OBSERVATIONS, AND RECOMMENDATIONS

From the data collected and the various analyses, the following are the major findings.

1. Algorithm 9, which was found to yield favorable results in previous studies (1), displayed a poor DR-FAR relationship compared with algorithms 7, 8, and 10.
2. For the RD period and for detection levels lower than 95 percent, no best algorithm with respect to FAR could be found.
3. For detection levels of 95 percent and above, algorithm 7 was found to have the lowest FAR for the RD period.
4. No significant differences in MTTD among algorithms were found at the 95 percent detection level at the 0.05 level of significance.
5. In order to be detected, incidents that occurred on DL required less sensitive thresholds than those on NDL.
6. For the most efficient algorithms, 7 and 10, for RD and NRD, respectively, no significant difference in MTTD was found for the 95 percent detection level at the 0.10 level of significance.
7. For incidents occurring on DL and NDL during the RD period, algorithms 10 and 7, respectively, were found to be the most efficient as far as FAR was concerned at the 95 percent detection level.
8. No significant differences in MTTD among algorithms were found to exist for AI and NAI on either DL or NDL, at the 95 percent detection level for the RD category.
9. At the 95 percent detection level, thresholds developed for AI and NAI on DL are less sensitive to false alarms than those developed for the above incident data on NDL for all algorithms during the RD period.
10. Thresholds developed for AI yielded lower FAR than thresholds developed for NAI for both DL and NDL at the 95 percent detection level and for the RD period.
11. Thresholds developed at the 95 percent detection level for AI occurring on DL detected only 78 percent of the NAI on that lane, for algorithms 7 and 8, and all such incidents for algorithm 10.
12. Thresholds developed at the 95 percent detection level for a representative sample of incidents (ALL) could effectively be used during RW, NRD, and NRW periods.

Based on the major findings of this study, the following observations could be made.

1. MTTD should not be a critical criterion for selecting an operational algorithm because no significant differences in this parameter were found among the tested algorithms for desired detection levels.
2. The DR-FAR relationship should be a critical criterion in the process of selecting incident-detection algorithms.
3. On the whole, algorithm 7 seemed to yield the most favorable results of all the algorithms tested in this study.
4. Thresholds developed for accidents on DL could be used to guarantee the lowest FAR.
5. The level of lateral detectorization is not a critical issue as far as detection time for incidents on various lanes is concerned.
6. If a high level of lateral detectorization (fully detectorized lanes) exists, algorithms should be applied to each lane in the detection process to yield low FAR and high DR.
7. The effort in developing thresholds for the RW, NRD, and NRW periods could be avoided by using thresholds developed for a representative sample of incidents (ALL).
8. Complicated algorithms are not necessarily the best ones.

The following recommendations are made.

1. Conduct an on-line evaluation of the above algorithms.
2. Conduct a discriminant analysis of traffic features to find the best combination of features to be used in an algorithm.
3. Develop algorithms based on speed-related features.
4. Investigate traffic-feature characteristics in bottlenecks during incidents to improve detection and false-alarm rates.
5. Because there are some differences between the results of this study and those of TSC, evaluating other non-pattern-recognition algorithms with the above data ought to be considered.

ACKNOWLEDGMENT

The research reported in this paper was conducted within the framework of IDOT and with the cooperation of FHWA. We wish to acknowledge the support of the staffs of the Bureau of Air Operations and the Traffic Systems Center in the data-collection and programming phases of this study. The contents of this report reflect our views, and we alone are responsible for the facts and the accuracy of the data presented here. The contents do not necessarily reflect the official views or policies of IDOT or FHWA. This report does not constitute a standard, specification, or regulation.

REFERENCES

1. H. J. Payne, D. N. Goodwin, and M. D. Teener. Evaluation of Existing Incident Detection Algorithms. Federal Highway Administration, Rept. FHWA-RD-75-39, Feb. 1975, p. 118.
2. H. J. Payne, E. D. Helfenbein, and H. C. Knobel.

Development and Testing of Incident Detection Algorithms: Vol. 2—Research Methodology and Detailed Results. Federal Highway Administration, Rept. FHWA-RD-76-20, April 1976.
3. R. Kahn. Interim Report on Incident Detection Logics for the Los Angeles Freeway Surveillance and Control Project. California Department of Transportation, Sacramento, Sept. 1972.
4. K. G. Courage and M. Levin. A Freeway Corridor Surveillance Information and Control System. Texas Transportation Institute, College Station, Research Rept. 488-8, Dec. 1978.
5. A. R. Cook and D. E. Cleveland. Detection of Freeway Capacity-Reducing Incidents by Traffic-Stream Measurements. TRB, Transportation Research Record 495, 1974, pp. 1-11.
6. C. L. Dudek, C. J. Messer, and N. B. Nuckles. Incident Detection on Urban Freeways. TRB, Transportation Research Record 495, 1974, pp. 12-24.
7. F. C. Bond. The Development of Computer Logic for Automatic Incident Detection. CAESP Internal Rept., 1974.
8. S. Siegal. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York, 1956.

Part 2. On-Line Evaluation

Moshe Levin, Gerianne M. Krause, and James A. Budrick, Bureau of Materials and Physical Research, Illinois Department of Transportation, Oak Park

Five algorithms were evaluated on-line by using the facilities of the Traffic Systems Center of the Illinois Department of Transportation. Three of the algorithms developed by Technology Services Corporation (TSC), were of a pattern-recognition nature. The other two—a pattern-recognition and a probabilistic or Bayesian algorithm—were developed locally. Thresholds for the features used in each of the pattern-recognition algorithms were developed by TSC. The thresholds for the probabilistic algorithm were developed by using accident data on the Eisenhower Expressway. The measures of effectiveness in the evaluation were detection rate, false-alarm rate, and mean-time-to-detect. The three TSC algorithms were evaluated twice on the Eisenhower Expressway at the 80 and 90 percent levels of detection thresholds, and then problem areas showing high false-alarm rates were represented by the 50 percent level. The three TSC algorithms were then evaluated on a section of the Dan Ryan Expressway that was free of geometric problems, for comparison purposes. Statistical analysis showed no difference in detection rate, false-alarm rate, and mean-time-to-detect among the three TSC algorithms at any of the evaluated detection levels. Introduction of the 50 percent level improved certain measures of effectiveness. Algorithm 7, the best of the TSC algorithms, showed overall superiority to the two local algorithms. The false-alarm rate was shown to be related to geometric and other features of the problem areas and yielded algorithm 8, which uses a shockwave-suppressor mechanism and requires the least effort in developing appropriate thresholds.

This paper discusses the on-line evaluation of five incident-detection algorithms that were all evaluated off-line in the preceding paper to obtain the optimal threshold sets used in the on-line evaluation.

The specific goals of this research were

1. To determine the on-line efficiency of algorithms proved effective in the off-line evaluation,
2. To correlate algorithm efficiency parameters derived from the on-line evaluation with those derived from the off-line evaluation, and
3. To evaluate combinations of thresholds with respect to geometric conditions on the freeway.

ALGORITHM DESCRIPTION

Consider an n-lane freeway section of length L between two fully detectorized stations. At each station a set of flow characteristics for occupancy, volume, and speed is measured at specific time intervals.

Suppose that at time t_0 an incident occurs at a certain point on one of the lanes in section L. A shock wave will develop and travel upstream of the incident with an intensity that is dictated by the severity and lateral location of the incident and by environmental and geometric conditions. At time $t_0 + dt$ an incident-detection algorithm, by continuously measuring and comparing the flow characteristics upstream and downstream of the incident with predetermined thresholds, will detect the incident.

This section describes the structure of the incident-detection algorithms evaluated in this research. Of the five algorithms evaluated, three of the pattern-recognition type were developed by TSC (2) and the other two, one pattern-recognition and one probabilistic (7), were developed locally in the course of this research.

The research effort of TSC included the development of 10 incident-detection algorithms that could be grouped into three categories. The first, comprising algorithms 1-7, is composed of variations on the classic California algorithm (2). The second consists of algorithms 8 and 9, which are characterized by suppression of incident detection after a compression wave is detected. Finally, algorithm 10 represents an attempt to detect those incidents that occur in light-to-moderate traffic but do not lower capacity below the volume of oncoming traffic.

Of these 10 algorithms 3 were selected for evaluation, 1 from each category. The algorithms selected (7, 8, and 10) were chosen for a number of reasons. Preliminary investigation by TSC had indicated algorithm 7 to be a superior form of the California algorithm. Algorithm 8 is identical to algorithm 9 except for an added persistence check. According to TSC's preliminary investigation, algorithm 8 has a slightly lower FAR but a longer MTTD than algorithm 9. Although algorithm 10 did not perform especially well in TSC's view, it was included in the on-line evaluation because it represents a first attempt to solve the problem of detecting incidents