Development and Testing of Operational Incident Detection Algorithms:

Technical Report

NOTE TO READER:

THIS IS A LARGE DOCUMENT

Due to its large size, this document has been segmented into multiple files. All files separate from this main document file are accessible from links (blue type) in the table of contents or the body of the document.

U.S. Department
of Transportation

Federal Highway Administration

Development and Testing of Operational Incident Detection Algorithms: Technical Report

Report No. TBD

September 1997

DOT/FHWA
150291

FOREWORD

This report describes the development of operational surveillance data processing algorithms and software for application to urban freeway systems, conforming to a framework in which data processing is performed in stages: sensor malfunction detection, data repair, calibration, qualitative modeling of traffic conditions, and, finally, discrimination of incidents from recurrent congestion. Development and testing used real data obtained from freeway systems in Oakland, CA, San Diego, CA and the Twin Cities in Minnesota. Statistical pattern recognition techniques, including optimal decisions trees and neural nets were used. The algorithms and software produced are designed for integration into real-time traffic management systems. This report focuses on the development and testing of the surveillance data processing algorithms.

This document is intended for the technical staffs of local traffic management authorities responsible for operating and/or planning freeway incident and response capabilities, at locations where such capabilities are in place or are being considered for deployment; for developers of software systems that support freeway traffic management; and for researchers who are focused on development of surveillance data processing algorithms.

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The contents of this report reflect the views of the contractor who is responsible for the accuracy of the data presented herein. The contents do not necessarily reflect the official policy of the Department of Transportation.

This report does not constitute a standard, specification, or regulation.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein only because they are considered essential to the object of this document.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipients Catalog No.
4. Title and subtitle Development and Testing of Operational Incident Detection Algorithms: Technical Report		5. Report Date. September 1997
		6. Performing Organization Report No.
		8. Performing Organization Report No. R-009-97
7. Author(s) Harold J. Payne and Stephen M. Thompson		10. Work Unit No. (TRAIS)
9. Performing Organization Name and Address Ball Aerospace & Technologies Corp., Systems Engineering Operations, 9 179 Aero Drive, San Diego CA, 92123		11. Contract or Grant No. DTFH6 1-93-C-000 15
		13. Type of Report and Period Covered Final Report April 93 to Sent 97
12. Sponsoring Agency Name and Address Federal Highway Administration, ITS Research Division Turner-Fairbank Highway Research Center 6300 Georgetown Pike, McLean, VA 22101		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

New operational surveillance data processing algorithms and software for application to urban freeway systems have been developed and tested, conforming to a framework in which data processing is performed in stages: sensor malfunction detection, data repair, calibration, qualitative modeling of traffic conditions, and, finally, discrimination of incidents from recurrent congestion. Development and testing used real data obtained from freeway systems in Oakland, CA and the Twin Cities in Minnesota. Statistical pattern recognition techniques, including optimal decisions trees and neural nets were used.
This volume presents details of the research methodology, analyses and conclusions, and recommendations for future research.
The other volumes of this report are:
Development and Testing of Operational Incident Detection Algorithms: Executive Summary
Development and Testing of Operational Incident Detection Algorithms: Software Documentation
Modification and Application of FRESIM for Modeling Congestion and Incident Scenarios

In addition, this report is accompanied by a CD-ROM containing the source code documented, and various data bases developed and used in this project.

17. Key Words	18. Distribution Statement
Freeway operations, ATMS, freeway surveillance, incident detection	Unlimited

19. Security Classif. (of this report)	19. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

Development and Testing of Operational Incident Detection Algorithms: Technical Report

BSEO Report No. R-009-97

Prepared by

Ball Aerospace & Technologies Corp.
Systems Engineering Operations
9179 Aero Drive
San Diego, CA 92123

Prepared for

Federal Highway Administration
ITS Research Division
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101

September 1997

TABLE OF CONTENTS

Section		Page
1.0	**INTRODUCTION..**	**1**
1.1	Project Objectives	1
1.2	Project Documents	2
1.3	Overview of Project Work	2
1.4	Organization of this Report	3
2.0	**FREEWAY SURVEILLANCE SYSTEMS AND PRACTICES..**	**5**
2.1	The State of Freeway Surveillance	5
2.2	Impacts for Incident Detection Algorithms	7
2:2.1	Malfunction Detection and Data Repair	7
2.2.2	Calibration of Loop Measurements	8
2.2.3	Sensor Spacing and Placement	8
3.0	**INCIDENT DETECTION TECHNIQUES AND PRACTICES**	**11**
3.1	Operational Incident Detection Practices	11
3.2	An Overview of Proposed Algorithms for Automated Incident Detection	12
4.0	**DEVELOPMENT APPROACH**	**17**
4.1	Performance Requirements	17
4.2	A Framework for Surveillance Data Processing	19
4.3	Development Guidelines for Operationally Effective Algorithms	23
4.3.1	Insights from Previous Research	24
4.3.2	Feature Selection	25
4.3.3	Classifier Construction	27
4.3.4	Algorithm Evaluation	28
4.4	Software Subsystem Requirements	29
5.0	**PROBABILISTIC ALGORITHMS AND ASSOCIATED EVALUATION PROCEDURES**	**31**
5.1	The Probabilistic Approach to Data Classification	31
5.2	Evaluation Methodology	33
6.0	**DEVELOPMENT SITES AND ASSOCIATED FREEWAY MODELS**	**.39**
6.1	The I-880 Site	39
6.1.1	I-880 Freeway Geometry & Surveillance System Description	41
6.1.2	I-880 Data Characteristics	41
6.1.3	Modeling of I-880 Site	41
6.2	The Twin Cities Site	41
6.2.1	Modeling of Twin Cities Freeway System	42
6.2.1	Twin Cities Freeway Geometry & Surveillance System Description.	42
6.2.2	Twin Cities Data Characteristics	43
6.2.3	Modeling of the Twin Cities Site	44
6.3	The San Diego Site	44
6.3.1	San Diego Freeway Geometry & Surveillance System Description	45
6.3.2	San Diego Data Characteristics	46
6.3.3	Modeling of the San Diego Site	46

TABLE OF CONTENTS
(Continued)

Section **Page**

7.0 MALFUNCTION DETECTION AND DATA REPAIR TECHNIQUES 47
7.1 Malfunction Detection Algorithms 47
7.1.1 Overview 48
7.1.2 Data Validation for the San Diego and I-880 Sites 49
7.1.3 Data Validation for the Twin Cities Site 53
7.2 Data Repair Algorithms 56
7.3 Application to the I-880 Site 61

8.0 MEASUREMENT CALIBRATION FOR INDUCTION LOOP SENSORS 69
8.1 Calibration Procedures 69
8.2 Calibration Results for the Twin Cities Site 72
8.3 Calibration Results for the San Diego Site 75

9.0 QUEUE DETECTION AND TRACKING 77
9.1 Identification of Traffic Queues 77
9.2 Tracking Queue Movements 81

10.0 DATA COLLECTION AND LABELING 87
10.1 Data Sets for the I-880 Development Site 87
10.1.1 Identification and Selection of Incident Traffic Data 87
10.1.2 Identification and Selection of Incident-Free Traffic Data 93
10.1.2.1 Categories of Incident-Free Data. 93
10.1.2.2 Data Thinning Techniques. 95
10.2 Data Sets for the Twin Cities Development Site 97
10.2.1 Remote Data Labeling 97
10.2.2 On-site Data Labeling 98
10.2.3 Assembling the Data Sets 102

11.0 DEVELOPMENT OF OPERATIONAL INCIDENT DETECTION ALGORITHMS.. 105
11.1 Algorithm Development Using Data from the I-880 Site 105
11.1.1 Preliminary Data Analysis 105
11.1.2 Construction of Incident Detection Algorithms 109
11.2 Algorithm Development Using Data from the Twin Cities Site 113
11.2.1 Preliminary Development 113
11.2.2 The Bayesian Classification Algorithm 114
11.2.3 Neural Network Algorithms 118

12.0 ALGORITHM PERFORMANCE EVALUATIONS 121
12.1 Algorithm Evaluations Using the I-880 Data Sets 121
12.2 Algorithm Evaluations Using the Twin Cities Data Sets 126
12.2.1 Definitions of Applicable MOEs 127
12.2.2 Evaluation of the Bayesian Classification Algorithm 129
12.2.3 Evaluation of Neural Network Algorithms 130
12.2.4 Evaluation of the California Algorithm. 131
12.2.5 Algorithm Performance Comparisons 134

TABLE OF CONTENTS
(Continued)

Section		Page
12.2.6	The Contribution of Queue Tracking	135
12.3	Planned Evaluations in San Diego	137
13.0	**TECHNIQUES FOR EXPLOITING MOTORIST REPORTS**	**139**
13.1	The Fusion Method	139
13.1.1	Derivation of the Fused Estimate	139
13.1.2	Interpretation of the Result	142
13.2	The Cueing Method	144
13.2.1	Algorithm Development	144
13.2.2	Algorithm Evaluation	148
14.0	**DYNAMIC MODEL-BASED ALGORITHMS**	151
14.1	Overview of Kalman Filter Processing	15 1
14.2	On-Ramp Volume Estimation	154
14.3	Off-Ramp Volume Estimation	155
14.4	G-Factor Estimation	155
14.5	Density Estimation	158
14.6	Speed Estimation	16 1
14.7	Possible Future Enhancements	162
15.0	**CONCLUSIONS AND RECOMMENDATIONS**	**163**
15.1	Conclusions	163
15.1.1	Algorithms and Software	163
15.1.2	Applications for Operational Implementation	164
15.2	Recommendations	165
15.2.1	Further Research	165
15.2.2	Measures to Foster Implementation	166
16.0	**REFERENCES**	169
16.1	Project Documents	169
16.1.1	Interim Project Reports	169
16.1.2	Papers Presented at Conferences	170
16.1.3	Final Report Documents Pertaining to the Assessment of Impacts of Incidents	17 1
16.1.4	Final Report Documents Pertaining to the Development and Testing of Operational Incident Detection Algorithms	17 1
16.2	Other Applicable Documents	172
APPENDIX A:	**THE I-880 INCIDENT DATABASE**	**175**
APPENDIX B:	**THE TWIN CITIES INCIDENT DATABASE**	**185**
APPENDIX C:	**DETAILED FREEWAY MODELS FOR MINNEAPOLIS**	191
APPENDIX D:	**DETAILED FREEWAY MODELS FOR SAN DIEGO**	**223**
APPENDIX E:	**FEATURE DEFINITIONS**	**313**

LIST OF FIGURES

Figure		Page
3-l	Modular Approach to Incident Detection	14
4-l	Functional Design of the Operational Algorithm	20
4-2	Steps in IOD/ITD Development	26
5-l	Illustration of the California Algorithm	32
5-2	Probabilistic Algorithms Using Binary Decision Trees	32
5-3	MOEs as Functions of Detection Threshold	35
5-4	California Algorithm # 8	36
5-5	Detection Rate vs. Threshold Set # for California Algorithm #8	37
6-l	Overview of the I-880 Surveillance System	40
6-2	Overview of the Twin Cities Surveillance System	42
6-3	A typical Placement of Surveillance Loops in the Twin Cities Freeway System	43
6-4	Overview of the San Diego Metropolitan Area	44
6-5	Segments of the San Diego Freeway System Instrumented with Loop Surveillance	45
7-l	Measured Volume vs. Steady-State Volume (Lane 2)	60
7-2	Error in Measurement Estimate (Lane 2)	60
7-3	Distribution of Malfunction Rate Among I-880 Sensors	63
7-4	Distribution of Malfunction Rate Among I-880 Stations	64
7-5	Malfunction Rates of Individual I-880 Stations (Northbound)	65
7-6	Malfunction Rates of Individual I-880 Stations (Northbound)	66
7-7	Distribution of Malfunction Rate Among I-880 Stations - 10 Minute Wait	67
7-8	Malfunction Rates of Individual I-880 Stations (Northbound) - 10 Minute Wait	68
7-9	Malfunction Rates of Individual I-880 Stations (Northbound) - 10 Minute Wait	68
8-l	Measured and Filtered G-Factors	71
8-2	Examples of G-Factor Calibration for the Twin Cities	72
9-l	Sample Head-of-Queue Scenarios	80
9-2	Queue Boundary Requirementss	81
9-3	Examples of Dynamic Traffic Queues	83
9-4	Association of Traffic Queues with Congestion Events	85
9-5	Merging Congestion Events	85
10-l	Steady State Incident Conditions	90
10-2	Display Capability Used in Labeling Traffic Conditions	99
10-3	Recorded Event Information	100
1 l-l	Single Lane Occupancy During Congestion	106
11-2	Examples of Binary Decision Rules	110
1 l-3	Binary Decision Tree Developed from I-880 Traffic Data.	112
11-4	Incident-Free Density Function	116
11-5	Incident Density Function	117
11-6	Summary of Bayesian Classification Results	118
11-7	Structure of Neural Network Classifiers	119
12-1	Binary Decision Tree for I-880	123
12-2	Sample Evaluation Corresponding to Incident Conditions	124
12-3	Sample Evaluation Corresponding to Incident-Free Conditions	125
12-4	Performance of the Bayesian Classification Algorithm	129
12-5	Performance of the Neural Network ID4d Algorithm	130
12-6	Performance of the Neural Network IDS Algorithm.	13 1
12-7	California Algorithm # 8	133
12-8	Performance of the California #8 Algorithm	134
12-9	Algorithm Performance Comparison	135

LIST OF FIGURES
(Continued)

Figure **Page**

12- 10 Contribution of the Queue Tracking Algorithm.. 136
12- 11 Application of Algorithms to the San Diego Site 137
13- 1 Features Used to Detect Temporal Speed Drop 146
13-2 Identification of the Incident Zone.. 147
13-3 Evaluation of the Motorist Report Algorithm 148
14-2 Aggregate Variables for a Freeway Link

LIST OF TABLES

Table **Page**

3-l Summary of Previous Algorithms 14
4-l Processing Objectives for Various Incident Stages. 18
5-l Threshold Sets for California Algorithm #8, Calibrated for Minneapolis [PAYN 76] . . . . . .
7-1 Relative Frequency of Malfunction Types . . . . . .
7-2 Station Malfunction Rates 64
8-l Calibrated G-Factors for Detector Stations in the Twin Cities 73
8-2 Calibrated G-Factors for Detector Stations in San Diego 76
10-l Intervals in the I-880 Data Base with Large Sections of Incident-Free Conditions 94
10-2 Labeled Events from the Twin Cities Freeway System 101
10-3 Learning and Evaluation Sets for the Twin Cities Site 102
12-1 Threshold Sets for California Algorithm #8, Calibrated for Minneapolis [PAYN 76] 132

1.0 INTRODUCTION

Freeway congestion due to the occurrence of incidents is a major cause of traffic delays in the United States and around the world. Mitigation of such delays through rapid and reliable incident detection is a vital traffic management objective.

Responding to the increasing need to improve means for managing non-recurrent congestion, the Ball Aerospace & Technology Corp. was contracted by the Federal Highway Administration to develop and evaluate incident detection algorithms for use in Intelligent Transportation Systems, and to develop a procedure to report and quantify the operational effects of freeway incidents, in a study entitled "Incident Detection Issues."

1.1 Project Objectives

This study had three objectives:

1. Examine and evaluate the entire issue of incident detection and verification, existing incident detection systems, related surveillance requirements for traffic control systems, and determine the deficiencies in current incident detection and surveillance practices.

2. Develop and field verify three approaches to freeway incident detection which overcome the deficiencies identified in Objective 1.

3. Develop a methodology for the uniform measurement and reporting of incidents, and the necessary data base of incident frequencies, types, duration, capacity reduction, and other data needed to predict the occurrence and impacts of incidents on freeways.

The study was conducted over the period April 1993 through September 1997. The third objective was addressed by Ball and its subcontractor, California Polytechnic State University at San Luis Obispo, in work that concluded in April of 1995. The remaining work was conducted by Ball and its subcontractor, the University of Maryland.

Work on the third objective is reported in a separate pair of documents identified in the Section 16. This Technical Report and the other volumes of this final report are limited to

the remainder of the project work, that is, the development and testing of approaches that support automatic incident detection.

1.2 Project Documents

This project produced a series of interim technical reports, several papers presented at conferences, a final report pertaining to the assessment of the impacts of incidents, and a final report pertaining to the development and testing of incident detection algorithms. A complete listing of all documents produced from this project is provided in Section 16. Other applicable documents, such as papers describing previous research, are also listed in Section 16 and are referenced in the relevant discussions of this text.

1.3 Overview of Project Work

The purpose of this project is to investigate and develop operationally effective, surveillance-based incident detection algorithms for use in urban freeway systems. For our purposes, an incident is defined as any event which restricts the normal flow of traffic, such as motorist collisions, disabled vehicles and debris in the roadway. The primary algorithm objective is to detect incidents promptly and reliably, thereby minimizing incident-related traffic delays.

The thrust of our efforts has been to consolidate and build upon the various algorithms and methodologies used in this regard since the early 1970's. To this end, certain high-level objectives of incident detection processing were identified, and a modular and comprehensive algorithm architecture was selected to facilitate the realization of these objectives. Furthermore, specific development procedures have been identified in the interest of continued and systematic development of the selected architecture.

The work is characterized by these features:

- <u>All of the work is within a framework for operational auality incident detection for urban freewavs in which a process of surveillance data processing is performed in stages: sensor malfunction detection. data repair. calibration, qualitative modeling: of traffic conditions, and, finallv. discrimination of incidents from recurrent congestion</u>. Each stage has distinct objectives and needs to reflect the specific characteristics of the surveillance system, including

the particular sensors and their deployment. While our work has been restricted to the most prevalent sensor, the induction loop, this framework should be effective for other sensors as well, while individual algorithms will need to be specific to those sensors and how they are deployed. Our work has demonstrated the crucial role played by each of these steps in ultimately arriving at an operationally useful capability for incident detection. At the same time, each of these steps has great operational utility in themselves.

- Algorithm development and validation is based entirely on real data. This feature derives from an effort to use both simulated and real data, and reflects the judgment that simulation alone cannot support credible development and testing. We were unable to convince ourselves that the simulation we explored (FRESlM) in fact had a sufficiently realistic representation of either congestion or incidents to be trusted as the basis for development and testing. On the other hand, we did construct a capability for acquiring real data and attaching labels of incident or incident-free conditions, and successfully used it to obtain a very substantial body of useful data for both training and evaluation.

- Algorithms were developed that produced a probabilistic statement of the likelihood of an incident. Instead of declaring that an incident is definitely present or definitely not present, as is the case for many incident detection algorithms, our development focused on creating algorithms that produce an estimate of the probability that an incident is present. It is our belief that this probabilistic approach to classification reflects the input data characteristics more accurately than the conventional binary output, and also allows for a degree of certainty to be incorporated into the results, thereby indicating the appropriate level of attention required by TMC operators.

1.4 Organization of this Report

This report is organized into sections that form a logical chronology of the issues addressed under this project. First addressed is the preliminary research undertaken to assess current freeway surveillance and incident detection practices. This is followed by a description of the development approach selected for use in devising improved incident detection

algorithms, including specific discussions related to algorithms that produce a probabilistic statement of incident occurrence.

The sites chosen for algorithm development and testing are presented next, followed by descriptions of the various algorithms developed as necessary pre-processing steps in the sequence of surveillance data processing: malfunction detection and data repair, the calibration of loop measurement data, identification and tracking of traffic queues, and the construction of real data sets to be used for empirical algorithm development and testing.

The development and testing of incident detection algorithms is subsequently addressed, followed by potential methods for enhancing algorithm results: incorporating information received from freeway motorists and developing algorithms that utilize dynamic traffic models. Finally, conclusions reached from this project are presented, and recommendations are made for further research in the area of automated incident detection.

2.0 FREEWAY SURVEILLANCE SYSTEMS AND PRACTICES

Automated incident detection systems depend on a host freeway surveillance system for input. The data provided by the surveillance system may come in many forms, depending on the sensor technology employed, and on the physical deployment of the sensors.

In many locations, the surveillance system is already in place, and incident detection capability must be designed according to the characteristics of the existing system. For new systems, an opportunity exists to design the surveillance system specifically to support incident detection. At present, however, there is only limited information on which to base such guidelines (this issue was addressed by Payne and is documented in [PAYN 76]).

This section identifies the sensor types and deployment practices currently used for freeway surveillance. Potential new sources of traffic information are also addressed in the interest of anticipating the data types which may become available for use in incident detection systems.

2.1 The State of Freeway Surveillance

By far, the sensor most commonly used for freeway surveillance is the induction loop. These sensors are used exclusively in many areas, and are deployed as either “single” loops or “double” loops. Both types measure traffic volume (the number of vehicles passing the detector) and occupancy (the percentage of time that a vehicle resides directly over the sensor). Double loops also provide a measure of traffic speed. These sensors are typically deployed in mainline lanes, and also for on-ramps and off-ramps.

Mainline installations typically span all lanes of the freeway to form a detector station. However, not all facilities instrument all mainline lanes. In Chicago, for example, a typical deployment for three-lane freeways is to instrument only the center lane at approximately one-mile intervals, with full count stations at approximately three-mile intervals. Additionally, induction loops are subject to failure, so that full counts are not always available even where they have been installed.

Deployment practices for ramps vary considerably from one location to the next. Often, ramps are either not instrumented at all (particularly with regard to off ramps), or not all lanes of the ramp are instrumented.

New sources of surveillance data are becoming available, and therefore present new opportunities for incident detection. Among these new sources are imaging sensors, microwave radars, acoustic sensors and microloops, a new variant of the magnetometer. Each of these new sources are capable of producing measurements of the same type as induction loops; the imaging sensors have the potential to produce a much richer set of measurements. In fact, using "image-understanding " techniques, these sensors may be capable of directly detect incidents.

Additional sources of information that may prove useful for incident detection include vehicles that can communicate individual vehicle speeds and/or link travel times. There are two variants of vehicle instrumentation that can be distinguished: probe vehicles, and vehicles with in-vehicle navigation equipment:

- Probe vehicles: These vehicles have Automatic Vehicle Location (AVL) equipment which provides position and possibly speed (e.g., GPS provides both), and some means of communicating digitally with a central facility. Examples include the Freeway Service Patrol fleets in Los Angeles and the San Francisco Bay Area. Proper processing of this information can provide information on the traffic conditions for the links traveled by the probe vehicles.

- In-vehicle navigation vehicles: These vehicles have means to associate their position to an in-vehicle mapping system. They may use AVL technology, but they also have digital maps and some means of communicating digitally with a central facility. The Advance project in the Chicago area is an example of this type of capability.

In many urban areas, the growing prevalence of drivers with cellular phones has been exploited by private and public agencies to collect traffic information, particularly reports of incidents. The drivers are typically provided with a simple code (e.g., *99) that gives them free access to a reporting center. This center may be a commercial entity (e .g., MetroTraffic in many metropolitan areas) or a public agency set up specifically for this purpose. In the Chicago area, the *99 code is answered by a publicly funded agency, which then routes the call to the appropriate public emergency facility.

2.2 Impacts for Incident Detection Algorithms

The characteristics of the host surveillance system can profoundly affect the performance of incident detection algorithms. The primary issues involve the quality and availability of traffic measurement data, as addressed in the subsections that follow.

For this project, the induction loop is the only sensor explicitly considered for use in incident detection processing. This decision reflects the observation that the induction loop is currently the only sensor widely used for freeway surveillance. Specifically, this is the sensor type used by the development sites presented in Section 6, thereby necessitating use of induction loop data to support our approach of empirical algorithm development and testing, as discussed in Section 4.

2.2.1 Malfunction Detection and Data Repair

All induction loop sensors are subject to malfunction, and such malfunctions can dramatically and adversely affect subsequent processing of the sensor data. With regard to incident detection processing, faulty data can be misinterpreted as valid measurements of incident-caused conditions, resulting in “false alarms” (where an incident is declared to be present when, in fact, no incident exists). In poorly maintained systems, inadequate data quality and the resulting false alarms can render the incident detection capability useless.

Clearly, there is a need to mitigate the adverse impacts of sensor malfunctions on incident detection processing. In addressing this issue, some researchers have proposed algorithms that employ data filtering in an attempt to protect the algorithms from faulty data. This is not the desired approach. Data validation should be an explicit part of surveillance data processing, and the incident detection algorithm should be able to rely on the quality of the input data.

Acquired surveillance data must be routinely examined to identify faulty data in a manner that allows subsequent processing to avoid its use. Typically, field surveillance equipment will do some of this data validation task (as in the Type 170 controllers found in use in California and in many other States); additional software in the Traffic Management System may do further data validation [PAYN 76]. In order to implement an incident detection capability for a particular freeway system, it will be necessary to understand what data

validation techniques are presently used, and where necessary, to augment these techniques to yield the necessary data quality before applying the incident detection algorithms.

In the context of this project, data validation is addressed as an explicit processing step that is carried out prior to the application of incident detection algorithms. The developed malfunction detection algorithms and examples of their application are presented in Section 7. Additionally, methods are presented whereby faulty data may be "repaired" under certain circumstances in order to maintain the scope of incident detection processing to the extent possible.

2.2.2 Calibration of Loop Measurements

The measure of occupancy produced by induction loop sensors is computed as the percentage of time that a vehicle "occupies" the space directly above the sensor. A vehicle is said to occupy this space when the associated inductance rises above a given threshold value. Inconsistencies in loop length and the value of this threshold can cause occupancy measurements to be inconsistent from one sensor to the next.

These inconsistencies particularly affect the computation of traffic speed using measurements of volume and occupancy from "single" loop sensors (and stations). As a result, algorithms that use computed traffic speed as input are subject to error. These adverse impacts can be largely negated through calibration of the loop measurements.

For this project, a procedure was adopted whereby the "G-Factor" used in the computation of traffic speed was systematically calibrated for each detector station included in the roadway network. The calibration procedure is presented in Section 8, along with specific results for both the Twin Cities and San Diego freeway systems.

2.2.3 Sensor Spacing and Placement

For surveillance-based algorithms to function, main-line detector stations must be located in sufficient proximity so that an incident located between adjacent stations significantly affects their respective traffic measurements. Additionally, the distance between adjacent stations is directly related to the time required to detect an incident. Hence, the performance of incident detection algorithms can be generally improved through the installation of additional main line stations.

Certain types of algorithms, such as the Dynamic Model-based algorithms presented in Section 14, require as input traffic measurements for both on ramps and off ramps. Hence, locations where ramps are not instrumented or are only partially instrumented may preclude the use of such algorithms.

3.0 INCIDENT DETECTION TECHNIQUES AND PRACTICES

We have found that the methods used to perform incident detection vary significantly in both structure and effectiveness in urban locations across the country. In most instances, agencies with vehicles patrolling the freeways will detect incidents by discovery. Several locations also make use of cellular phone calls from freeway motorists to identify incidents. Conversely, relatively few sites perform incident detection based on electronic surveillance data, and such techniques are frequently employed only to detect very severe incidents. Nevertheless, automated surveillance-based incident detection systems have demonstrated significant abilities and hold the potential to greatly facilitate rapid incident detection and the subsequent deployment of appropriate response forces.

The discussion of this section is divided into two parts. First addressed are the methods currently employed by operational agencies for the purposes of incident detection. Secondly, a review is presented of proposed incident detection algorithms which we feel hold significant promise for effective detection of freeway incidents.

3.1 Operational Incident Detection Practices

The operational use of surveillance-based methods for incident detection is extremely limited. This issue was addressed during the preliminary stages of this project, and an assessment was made of the incident detection practices used in representative urban locations throughout the United States. Of the eleven sites visited, regular use of surveillance-based techniques is made only in Los Angeles, Seattle and Orlando.

More commonly, agencies with vehicles patrolling the freeways will detect incidents by discovery. For minor incidents, the discovering unit may provide all the on-site incident management necessary. Otherwise, appropriate authorities are contacted as necessary.

Most areas are patrolled by a police agency. A few areas also provide service patrols, such as the Freeway Service Patrols in California, the Highway Helpers in the Twin Cities area and the Minutemen in Chicago. These units often provide the first notice of an incident, and, in many cases, they can provide all aspects of required incident management.

Cellular phone calls placed by motorists are increasingly becoming a major source of rapid incident detection. These calls are often placed within a minute or so of incident onset, and

multiple calls tend to provide a degree of confidence in the information's accuracy. However, as motorists are not generally trained in these matters, caller information is frequently ambiguous and/or inaccurate. Hence, verification of motorist reports is an important issue.

3.2 An Overview of Proposed Algorithms for Automated Incident Detection

Since the mid-1970's, various surveillance-based incident detection algorithms have been developed. Unfortunately, the algorithms developed to date have met with only limited operational success, and it is clear that improved algorithms are needed to make surveillance-based incident detection technology operationally effective. Specifically, existing algorithms have been largely unable to maintain the high degree of reliability required in practice (e.g., high detection rate and low false alarm rate).

Despite this, several proposed algorithms clearly warrant further development and evaluation. The issue of identifying and characterizing proposed incident detection algorithms was addressed during the early stages of this project. The algorithms that were deemed to hold the most promise for effective incident detection are described briefly here. These concepts employed by these algorithms were used as the basis for our continued development.

The most well-known and widely-used incident detection algorithm is the California Algorithm(s) [PAYN 76]. This class of algorithm detects incidents based on comparisons of traffic measurements from adjacent detector stations. These stations are typically comprised of induction loops measuring traffic volume (vehicles per hour) and occupancy (percentage of time that a vehicle is directly above the station) at 30-second intervals. The underlying principle is that a capacity-reducing incident causes upstream occupancy to increase and downstream occupancy to decrease. Accordingly, an incident is declared when the difference between upstream and downstream occupancy readings is sufficiently high. The California Algorithm(s) also make use of "persistence tests", which prevent temporary traffic fluctuations from being misinterpreted as incidents.

Another well-known class of algorithm is the McMaster Algorithm(s) (see [GALL 89] and [PERS 89]). An important aspect of this algorithm is that incidents are detected in two distinct processing phases: (1) detect the existence of traffic congestion, and (2) determine

the cause of the congestion. For any given detector station, congestion is detected when occupancy and volume readings rise above established thresholds. The cause of congestion is then determined based on readings from the adjacent downstream station. In essence, the cause of congestion is deemed to be a capacity-reducing incident if the volume and occupancy readings from the downstream station are sufficiently low. Otherwise, the congestion is deemed to be of the "recurrent" variety, which arises when traffic demand exceeds freeway capacity.

In recent years, several new methods of performing surveillance-based incident detection have been proposed. Two algorithm categories were identified as holding significant promise for improved incident detection: (1) algorithms that utilize neural network technology, and (2) algorithms that are based on dynamic traffic flow models.

One of the principle applications of neural network technology is to pattern recognition problems. Hence, neural networks hold considerable potential for recognizing and classifying spatial and temporal patterns in traffic data, and therefore for detecting incidents. Various studies have been conducted which support this assertion. The findings indicate that neural network algorithms have the potential to achieve significant operational improvements in real-time incident detection over more conventional algorithms [RITC 93].

Algorithms that are based on models of traffic dynamics detect incidents by explicitly modeling unrestricted traffic flow on freeway sections of interest. The traffic model can be used as the basis of an extended Kalman filter that produces traffic state estimates and "residuals" (residuals indicate the disparity between the input measurements and produced state estimates) using discrete traffic measurements. Incidents can then be detected by classifying filter state estimates and residual values. The underlying principle is that large residual values indicate the model's assumption of incident-free traffic conditions is incorrect.

In summary, it should be noted that the algorithm types addressed in this section share a common basic approach: (1) manipulate raw surveillance data to form useful traffic features, and (2) classify these features according to pre-determined categories of traffic conditions. This process is illustrated in Figure 3- 1.

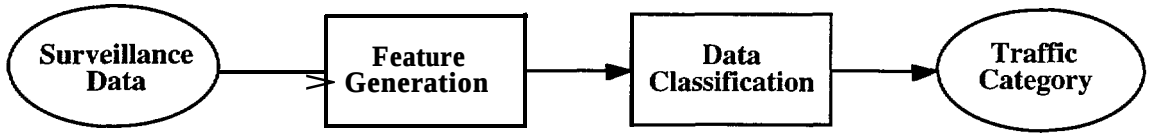


Figure 3-1. Modular Approach to Incident Detection

From this perspective, the algorithms deemed to hold the most promise for incident detection differ only in the types of features that are generated, the traffic categories that are supported, and the selected method of classification. This is summarized in Table 3- 1.

Table 3-1. Summary of Previous Algorithms

Algorithm	Primary Features	Traffic Categories	Method of Classification
California	• Spatial Occupancy Differences • Spatial Volume Differences	• Incident • Incident-Free	Binary Decision Tree
McMaster	• Downstream Volume • Downstream occupancy	• No Congestion . Recurrent Congestion • Incident Congestion	Binary Decision Tree
Neural Networks	• Raw Surveillance Measurements	• Incident . Incident-Free	Neural Network
Dynamic Models	. State Estimates • Filter Residuals	• Incident (+Type) • Incident-Free	Binary Decision Tree

The success of these algorithms demonstrates that the basic approach of generating and classifying traffic features is a potentially effective means of detecting incidents. This approach has therefore been adopted as the basis of our development. The primary objectives for developing improved algorithms are to consolidate the respective benefits of these approaches by selecting a comprehensive algorithm architecture and to identify a systematic means for continued development.

4.0 DEVELOPMENT APPROACH

Our approach to algorithm development is to consolidate the benefits of the most effective incident detection algorithms proposed to date and to devise improved algorithms through systematic development. To support this goal, a comprehensive framework for surveillance data processing was developed that is compatible with the general approach taken by the most effective algorithms investigated in early project work.

Specific performance requirements for the algorithm are addressed in Section 4.1. This is followed by a description of the selected algorithm architecture in Section 4.2. Guidelines for algorithm development are presented in Section 4.3, and the software subsystem used to implement the algorithms is addressed in Section 4.4.

4.1 Performance Requirements

Our objective was to develop a practical and reliable incident management tool dealing with all relevant issues from the acquisition of raw measurement data to the presentation of operationally useful results. The scope of such a tool should include much more than simply the detection of incident onsets, although this is clearly the primary objective of an incident detection system. To maximize effectiveness, detailed information should be provided during all stages of an incident, as illustrated in Table 4-l.

Since the processing objectives of Table 4-l reflect operational considerations, the algorithm which satisfies these objectives will be termed *the operational algorithm. The* high-level objectives of the operational algorithm are as follows:

(1) ESTIMATE INCIDENT RATES - Incident rates for freeway sections of interest can be estimated based on gross traffic indicators such as freeway geometry, weather conditions and average traffic volumes. These rates can be expressed as the probability of incident occurrence over a given time interval. This capability is included in the operational algorithm to provide a means of monitoring current traffic conditions. For example, a display could be devised that indicates incident likelihoods by using colored freeway segments.

Table 4-1. Processing Objectives for Various Incident Stages

Incident Stage	Preceding Incident Onset	Incident Begins	Incident Continues	Incident Ends	Following Incident Termination
Processing Objective	Incident Prediction • Predict Incident Rates	Onset Detection • Detect Existence of the Incident Incident Characterization • Estimate Incident Type • Estimate Incident Severity • Estimate Incident Duration	Process Motorist Reports • "Fuse" with Surveillance-based Results Incident Monitoring • Periodically Update Estimates	Termination Detection • Detect Absence of the Incident	Incident Prediction • Predict Incident Rates

(2) DETECT THE ONSET OF INCIDENTS SURELY AND QUICKLY - The purpose of incident detection is to facilitate rapid deployment of emergency response services and, possibly, of traffic control measures. Hence, the time between the occurrence of an incident and its detection is an important measure of algorithm performance. In addition, the number of incidents left undetected should clearly be minimized, and false alarms should be avoided to the extent possible. This necessitates monitoring the input data quality in order to detect malfunctioning sensors and prevent faulty data from inducing false alarms. Operational algorithms need to reflect the reality of sensor failures, and must have mechanisms for using the remaining sensor data effectively.

(3) IDENTIFY INCIDENT TYPES - This objective includes identifying the lane or lanes involved to assist emergency vehicles in choosing their approach to the incident scene. To the extent possible, additional incident characteristics should also be provided.

(4) PROVIDE A MEASURE OF INCIDENT SEVERITY - .In terms of traffic management, the severity of an incident is measured by the corresponding reduction of traffic flow capacity. Such a measure would be useful in prioritizing incidents, dispatching an appropriate response, and, possibly, in devising traffic control responses.

(5) PROVIDE AN ESTIMATE OF INCIDENT DURATION - Knowledge of how long an incident is expected to last is useful in applying traffic controls and for informing the public.

(6) DETECT INCIDENT TERMINATION - Once again, this information is useful for traffic control purposes, such as changing warning signs.

(7) UTILIZE MOTORIST REPORTS - Cellular phone calls are increasingly being used to aid rapid incident detection. In addition to aiding or verifying surveillance-based detection, motorist reports present an opportunity to enhance surveillance-based results by providing detailed information regarding the nature of an incident. Frequently, this type of information includes useful incident qualities (e.g., incident type data) that cannot be deduced from surveillance measurements. A desirable feature of any new incident detection system is the ability to use this emerging source of traffic information.

As stated in Section 2.2, the surveillance data inputs to the operational algorithm consist of traffic measurements from induction loop sensors only. This is due to the fact that algorithms have been empirically developed and evaluated, as discussed in Section 4.3, using actual traffic data obtained from the development sites listed in Section 6. These sites utilize induction loop sensors exclusively for the purposes of freeway surveillance.

4.2 A Framework for Surveillance Data Processing

In order to satisfy the processing objectives identified in the previous section, a comprehensive architecture for the operational algorithm was identified which consolidates the approaches of the most effective incident detection algorithms developed to date. In addition, the selected architecture significantly expands conventional functionality and is highly modular in order to facilitate evolutionary development while providing the

flexibility needed to pursue new ideas. Furthermore, objective and quantitative performance evaluations are possible using this architecture since all developed algorithms share a uniform set of processing objectives and conform to the structure of the operational algorithm.

The operational algorithm consists of eight functional components, as shown in Figure 4- 1. The functional design of the operational algorithm is illustrated in Figure 4-l as the data flow between these components.

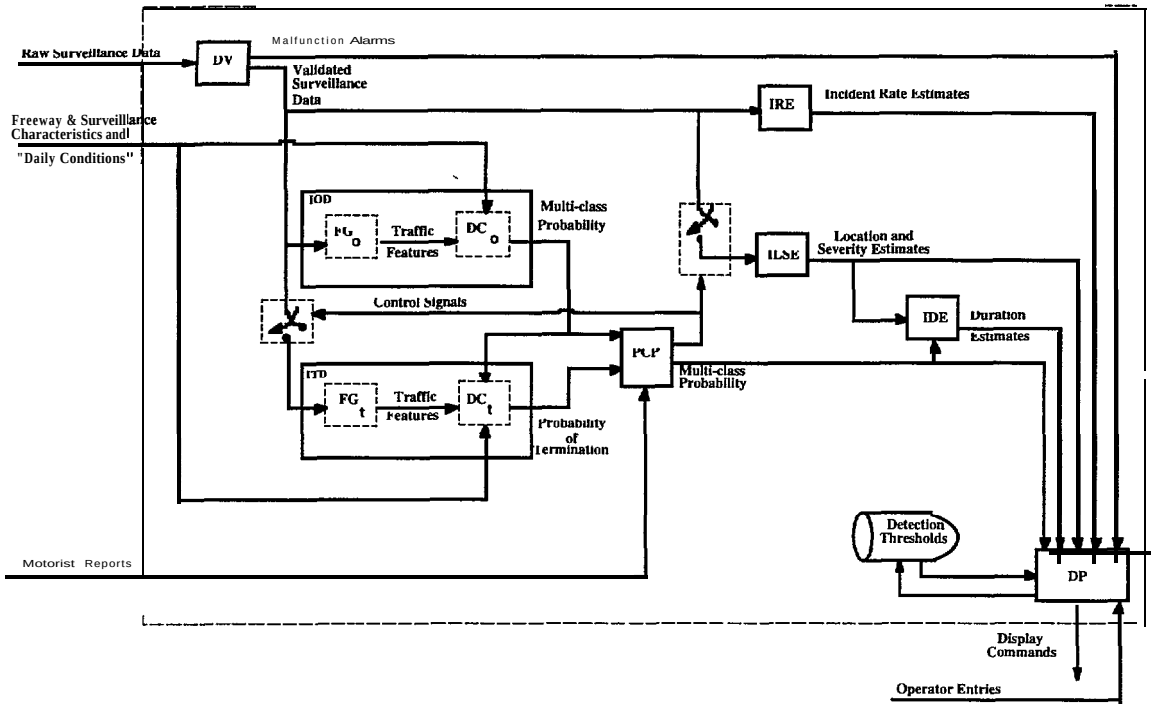


Figure 4-1. Functional Design of the Operational Algorithm

The individual components of the operational algorithm and their respective functions are described below.

(1) DATA VALIDATION (DV) - Identifies malfunctioning sensors and screens faulty data from other processing elements. This serves to identify faulty equipment and to reduce the number of false alarms.

(2) INCIDENT RATE ESTIMATION (IRE) - Estimates the likelihood of future incidents based on aggregate surveillance measurements.

(3) INCIDENT ONSET DETECTION (IOD) - Detects incident onsets based primarily on validated surveillance measurements. This component of the operational algorithm is comprised of two distinct functions: Feature Generation (FGo) and Data Classification (DCo). The FGo function manipulates the raw surveillance data to generate traffic features. These features are then used by the DCo function as indicators of the current traffic condition. The DCo function classifies the input traffic features to form a multi-class probability estimate of incident onset conditions.

(4) INCIDENT TERMINATION DETECTION (ITD) - Detects incident termination based primarily on validated surveillance data. This component of the operational algorithm is comprised of two distinct functions: Feature Generation (FGt) and Data Classification (DCt). The FGt function manipulates the raw surveillance data to generate traffic features. These features are then used by the DCt function as indicators of the current traffic condition. Separate feature generators are supported for onset detection and termination detection due to the expectation that different types of features will prove useful for these purposes. The DCt function classifies the input traffic features to produce a probability estimate of incident termination.

(5) INCIDENT LOCATION AND SEVERITY ESTIMATION (ILSE) - This component estimates the location and severity of suspected incidents based on surveillance data and classification results. Incident location is measured relative to the associated upstream and downstream sensor locations. The severity of an incident is measured by the corresponding reduction in traffic flow capacity.

(6) INCIDENT DURATION ESTIMATION (IDE) - This component estimates the duration of suspected incidents. Duration estimates are highly correlated with the type and severity of the associated incident.

(7) POST-CLASSIFICATION PROCESSING (PCP) - This component controls application of the incident termination logic by restricting FGt and DCt processing to suspected incident locations. When the termination classifier is

active, the PCP consolidates the outputs of the onset and termination classifiers. In addition, the PCP is also be responsible for the incorporation of received motorist information with surveillance-based results.

(8) DISPLAY PROCESSING (DP) - Displays results to TMC operators in a user-friendly manner. This component provides the interface between the remaining components and the Graphical User Interface (GUI).

The most important characteristics of this architecture are described briefly below.

First, all processing is preceded by a data validation phase. A significant problem affecting previous efforts was that faulty data frequently induced false alarms, which significantly reduced the operational effectiveness of the developed algorithms. This problem is largely negated in the current effort by requiring that all data presented to the incident detection algorithms first be tested for validity. Furthermore, faulty measurement data will be "repaired" (replaced by measurement estimates) when possible, further increasing the performance of the operational algorithm as a whole.

Second, the Incident Onset Detection (IOD) component of the operational algorithm was designed to incorporate the approach used by the most effective incident detection algorithms developed to date. This approach is as follows: (1) manipulate the raw surveillance data to form useful traffic features, and (2) classify these features into one of several pre-determined traffic categories. This architecture also allows a great deal of flexibility in IOD processing. For instance, the generated features could be as simple as the difference in occupancy between adjacent loop detectors, or as complex as the residuals generated by an extended Kalman filter that estimates the current traffic state based on intricate models of traffic dynamics. Furthermore, the features that are generated can be classified by one of several possible methods. For example, the classification portion of IOD processing may be carried out by a neural network, a binary decision tree, or some other heuristic method. Further still, the relative merits of these candidate IOD components can be evaluated in an objective and quantitative manner since the format of IOD output is the same for each candidate.

Third, the output of the IOD component is probabilistic in nature. This allows for a degree of certainty to be incorporated into the results, thereby indicating the appropriate level of attention required by an operator. It is our belief that this probabilistic approach to

classification reflects the input data characteristics more accurately than conventional binary output. For example, an incident probability of 95% is a very strong indication that action is required by an operator. However, an incident probability of 60% is a much weaker indication that an incident actually exists, and an operator may choose to investigate further by some independent means such as closed-circuit television. In conventional incident detection algorithms, the operator would not have access to this type of information since the algorithm output would indicate the existence of an incident equally in both instances.

Fourth, the PCP component allows surveillance-based results to be "fused" with received motorist information, as discussed in detail in Section 13. This process is facilitated by the fact that IOD output is probabilistic in nature. In addition, the PCP component is responsible for the application of "persistence tests" to the final output of the operational algorithm. This type of processing makes use of the fact that an actual incident persists for several minutes and was very effective in discerning actual incidents from temporary traffic fluctuations in the California algorithm.

Lastly, the ILSE, IDE, ITD and IRE components of the operational algorithm greatly expand the functionality of conventional algorithms, thereby aiding the process of effective incident management. These components are included to allow operators access to estimates of precise incident locations and durations, incident termination information, as well as estimates of overall incident likelihoods for freeway sections of interest.

It should be noted that the ambitious scope of this effort is made possible in large part by today's more powerful computers, which are capable of supporting sophisticated algorithms and remove several of the limitations imposed on previous efforts.

4.3 Development Guidelines for Operationally Effective Algorithms

Our development approach is to separately address the processing for each of the functional components of the operational algorithm with the objective of maximizing the operational effectiveness of the system as a whole. This section describes methods and procedures employed in developing the IOD and ITD components.

The overriding principle guiding our development is that the resulting algorithms should be ouerationallv useful. The primary elements of our development approach derive from this requirement and are discussed in detail in the subsections that follow.

4.3.1 Insights from Previous Research

During the literature review, several high-level characteristics of an effective incident detection algorithm were identified in the interest of retaining in the current development attributes that have been shown to be effective. These include: (1) the use of "difference-type" features and persistence tests from the California algorithm, (2) the empirical classifier construction methods used for neural network algorithms, where classifiers are "trained" using actual and/or simulated traffic data, (3) the McMaster algorithm's division of processing into two distinct phases wherein congestion is identified first and the cause of congestion is subsequently determined, (4) the McMaster algorithm's use of downstream measurements to determine the cause of congestion, and (5) the use of dynamic traffic models to generate useful traffic features.

Certain deficiencies in previous research were also noted in the interest of avoiding similar inadequacies in the development of our algorithms. These deficiencies are addressed individually below.

First, we found that many researchers used only simulations for the development and testing of their algorithms. It is our observation that traffic simulators frequently fail to capture the complexities encountered in actual traffic, and it is therefore expected that the performance of such algorithms in an operational setting will be significantly reduced from the reported performance, particularly with regard to false alarms.

Second, many researchers utilized actual traffic data but only for relatively short and/or uniform sections of freeway. Effective operational algorithms must be capable of dealing with the high degree of diversity encountered in practice, particularly with respect to freeway geometry, traffic levels and the deployment of surveillance sensors.

Lastly, many of the empirically developed algorithms researched utilize the same data set for both training and evaluation. In such instances, it is possible for the algorithm to "memorize" the specific characteristics of the training data set. Consequently, the performance of such algorithms in an operational setting can not be objectively assessed using the training data set.

AU of the deficiencies listed above generally cause reported algorithm performance to be overly optimistic with regard to use in an operational setting. For this project, these issues were specifically addressed by establishing the following guidelines for algorithm development and testing: (1) development and testing of incident detection algorithms must include use of actual traffic data, (2) the geographic extent of the actual traffic data must be sufficiently large to ensure that the resulting data sets are representative of the data to be encountered in practice, and (3) distinct data sets must be used for the purposes of algorithm training and evaluation.

4.3.2 Feature Selection

All surveillance-based incident detection algorithms utilize traffic features as indicators of the current traffic condition. For our purposes, a traffic feature will be defined as any function of the raw surveillance data. Examples include:

(1) The raw surveillance data itself, such as the number of vehicles which pass a detector station;

(2) Simple functions of the raw surveillance data, such as the difference between upstream and downstream occupancy measurements;

(3) Time series functions, such as a three-minute moving average of occupancy measurements from a single detector station; and,

(4) Functions based on models of traffic dynamics. We believe that this type of feature, as discussed in Section 14, may prove particularly useful in detecting incidents.

The purpose of the FGo and FGt modules of the operational algorithm is to generate the traffic features that form the input .to the DCo and DCt modules, respectively. Hence, a necessary first step in developing the IOD and ITD components of the operational algorithm is to identify those features which optimize the performance of the DCo and DCt modules, respectively. The process of determining optimal features will henceforth be termed feature *selection.* In the interest of clarity, the remaining discussion addresses IOD processing

only. Feature selection with regard to ITD processing is accomplished in an analogous manner.

The following process of feature selection was adopted as the basis of development for the IOD module: (1) identify candidate feature sets, (2) construct optimal classifiers to operate on these feature sets, (3) implement the candidate IOD modules in software, and (4) evaluate the performance of these candidate IOD modules. This process is illustrated in Figure 4-2.

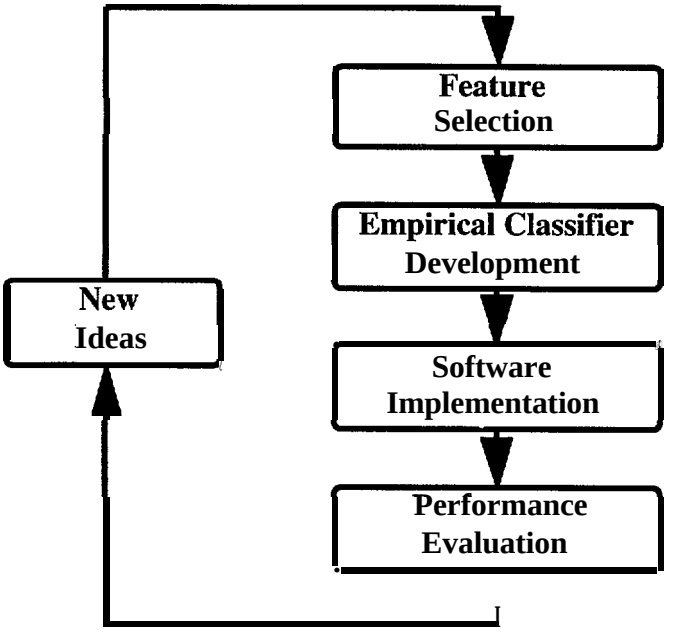


Figure 4-2. Steps in IOD/ITD Development

From this perspective, our development approach is largely a systematic methodology for determining optimal traffic features. In this context, "optimal" implies that the selected features correspond to those functions of the raw surveillance data that are most indicative of incident-related traffic conditions.

The candidate feature sets employed in this study are documented in Appendix E. These candidate features were inspired by the features utilized by existing and/or proposed algorithms, as well as various heuristic considerations. The methods used for classifier construction and evaluation are described in Sections 4.3.3 and 4.3.4, respectively. The

software subsystem required to support algorithm implementation is addressed in Section **4.4.**

4.3.3 Classifier Construction

A central task of incident detection processing is to perform data classification. Given surveillance-based traffic features, the objective is to characterize the underlying traffic condition according to various incident and incident-free classifications. Certain issues are fundamental to developing this capability. Foremost among these are selecting the classes to be supported and defining the method of classification.

In the course of our research, it has become apparent that certain software structures readily support our design objectives with regard to data classification. In particular, the binary decision tree and the neural network have been identified as the most promising architectures for classification of traffic data.

The ability of binary decision trees to perform data classification has been repeatedly demonstrated. As noted in Section 3.2, decision trees form the basis of the most widely-used incident detection algorithm, the California algorithm. Conversely, the use of neural networks in incident detection algorithms is a relatively new idea. However, one of the principle applications of neural network technology is to pattern recognition problems. There is a small but growing body of evidence suggesting that neural network algorithms hold the potential to achieve significant operational improvements in real-time incident detection [RITC 93].

For neural network classifiers, the construction process consists of training the network, and a variety of software tools are available to facilitate such training. For binary decision tree classifiers, the work by Breiman, Friedman, Olshen and Stone (see [BREI 84]) offers a systematic method of constructing optimal classification trees for a given feature set, and a software product is similarly available to implement this method.

For both structures, our approach to developing optimal classifiers is empirical in nature, and a necessary first step in empirical classifier construction is the development of a *learning set*.. Such a set consists of a wide variety of data samples and their corresponding true classifications. In essence, the process of classifier construction assumes that the learning set encompasses all types of data samples to be classified in the future. Hence, the

optimal learning set would consist of the entire sample space. Since this is clearly not possible, a learning set must be constructed that is representative of this space. Specifically, the learning set must include several data samples for each incident type, as well as several samples of incident-free traffic conditions.

Traffic simulators provide a means of constructing comprehensive learning sets. The primary benefit is that direct control of the data is possible, which allows for systematic development with regard to various traffic scenarios of interest. For this project, the use of simulated data in constructing learning sets was explored using the FRESIM traffic simulator.

Unfortunately, we encountered several difficulties in our efforts to utilize FRESIM in this regard, as documented in detail in [CHAN 97]. In summary, we were unable to convince ourselves that FRESIM is capable of adequately simulating actual traffic behavior, particularly with regard to complex phenomenon such as the formation recurrent traffic queues.

As a result, the learning sets employed in this study consisted entirely of real traffic data. We feel that the use of actual traffic data is a critical phase of classifier development and testing, since this data is representative of the data to be encountered in practice and frequently contains subtleties that simulations fail to capture.

The use of real traffic data necessitates conducting an off-line activity during which representative traffic scenarios are selected and the corresponding true classifications are identified (e.g., with the aid of closed-circuit television). Project work related to the collection and labeling of real traffic data is addressed in Section 10.

4.3.4 Algorithm Evaluation

The final step in the development approach of Figure 4-2 is to evaluate the performance of candidate IOD modules. This requires the identification of an objective set of performance Measures of Effectiveness (MOEs). The MOEs utilized to measure various performance aspects of the operational algorithm are presented in Section 5.2

A critical issue in algorithm evaluations involves the data that is employed in computing the selected MOEs. The data used in algorithm evaluations will henceforth be termed the

evaluation set.. Many of the issues involved in constructing good learning sets also apply to the process of constructing evaluation sets. In order to ensure that operational issues are addressed, all evaluation sets used in this study are comprised of actual traffic data. Furthermore, the data were taken from a wide range of traffic conditions and freeway geometry, as discussed in detail in Section 10. In order to ensure that evaluations are unbiased, the data sets used for classifier construction and evaluation in this project are strictly independent.

4.4 Software Subsystem Requirements

In order to develop and test incident detection algorithms using real surveillance data, a software subsystem is required to house the operational algorithm and to provide the support necessary for algorithm development and execution. This support includes interfacing to various sources of surveillance data, modeling freeway sections of interest and storing and displaying algorithm results. In addition, the software subsystems must provide data output functions for analysis purposes and for generating the learning and evaluation data sets required for empirical algorithm development and testing.

The software subsystem developed for this project is termed the Real-time Incident Detection Environment (RIDE). The software documentation for RIDE constitutes a separate volume of this final report [BOAZ 97].

5.0 PROBABILISTIC ALGORITHMS AND ASSOCIATED EVALUATION PROCEDURES

It is important to note that the algorithms developed for this project differ significantly from most incident detection algorithms in that algorithm outputs are probabilistic in nature. As a result, new methods of quantifying algorithm performance need to be devised.

This section addresses issues related to the probabilistic classification of traffic data. An overview of probabilistic algorithms is presented first in Section 5.1. This is followed by a description of the evaluation procedures to be employed in assessing algorithm performance in Section 5.2.

5.1 The Probabilistic Approach to Data Classification

Data classification for the purposes of incident detection can be pursued in two distinct ways. Consider the case where we are interested only in classifying the traffic condition as either incident or incident-free.

The objective of the first and most common approach is to select a single classification as an estimate of the corresponding true classification. In other words, we are interested in estimating whether the true classification is incident or incident-free. Conventional incident detection algorithms utilize this "class identification" approach. For example, the California algorithm implements this scheme via a binary decision tree, as illustrated in Figure 5- 1.

At each node in the figure, a branching decision is made based on measurements from available surveillance data. When an end node is reached, the corresponding traffic condition is labeled as either incident or incident-free.

The second approach pursues classification in a probabilistic sense. Rather than attempting to identify the true classification, the objective is to assign to each class a probability which reflects the likelihood that the classification is correct. The resulting vector of class probabilities is termed a *multi-class probability estimate*. The tree-based logic of the California algorithm can be readily adapted to this approach, as shown in Figure 5-2.

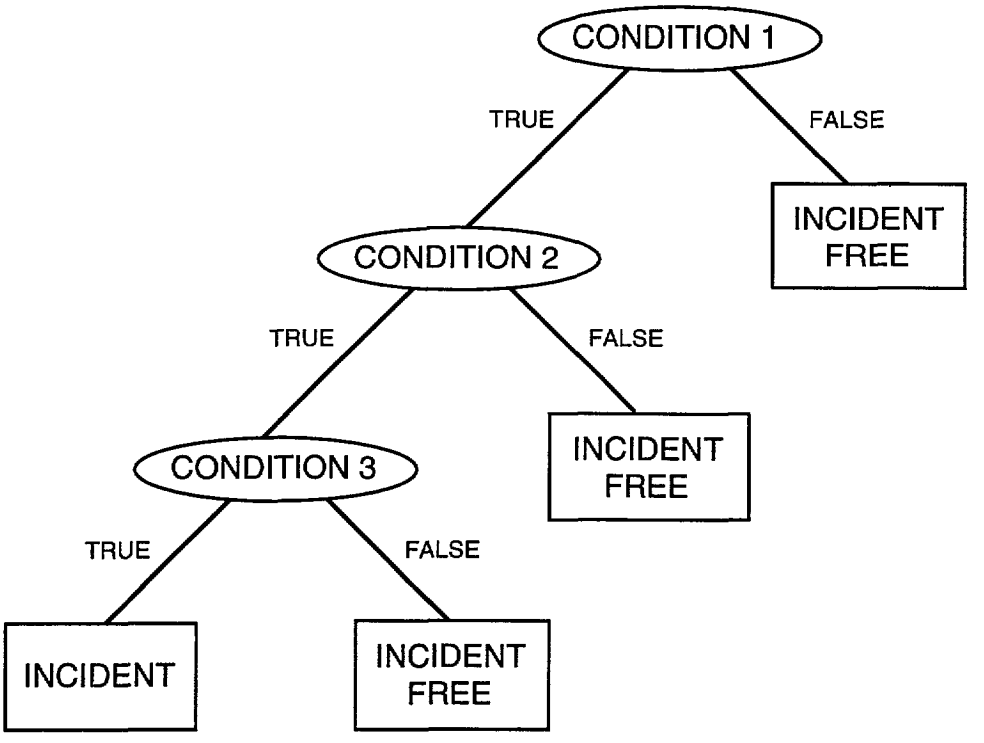


Figure 5-1. Illustration of the California Algorithm

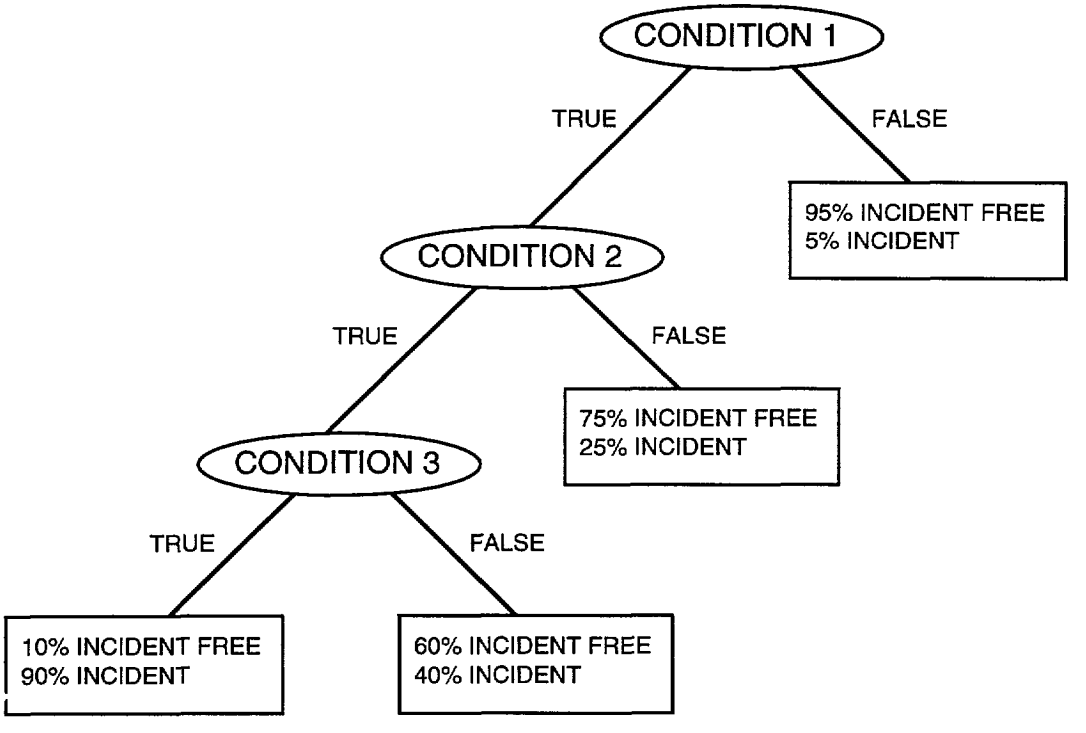


Figure 5-2. Probabilistic Algorithms Using Binary Decision Trees

The branching logic of this figure is identical to that of Figure 5-l ; however, the class labels associated with the end nodes have been replaced with multi-class probability estimates. The probability estimates of Figure 5-2 were assigned arbitrarily in order to illustrate a point. In practice, such estimates can be derived empirically for a given tree by running a large amount of data through the tree and observing the resulting class distributions of the terminal nodes.

We believe that algorithms that utilize the second approach reflect the input data characteristics more accurately than the conventional binary output. One should first note that the class identification results of Figure 5-l can be recovered from the multi-class probabilities of Figure 5-2 by simply designating the most likely classification of each end node as the best estimate of the true classification. Furthermore, the probabilistic approach of Figure 5-2 indicates a measure of confidence in classification results, and could therefore be used to indicate the appropriate level of attention required by an operator.

To clarity this point, consider the scenario in which an algorithm is to be developed that performs classification based on some traffic measurement X. Further, assume that traffic observations consistently indicated that when ($20 < X < 30$) an incident was somewhat likely, that when ($10 < X < 20$) an incident was very likely, and when ($X < 10$) an incident condition was definite. Algorithms that estimate class probabilities would be capable of making distinctions base on these “natural partitions” in the data. Conversely, conventional algorithms would neglect these observations and declare all cases with ($X < 30$) to be an incident.

We believe that an operator should have access to the type of information provided by probabilistic algorithms. Furthermore, the use of multi-class probabilities supports the algorithm objective of utilizing information received non-surveillance data sources, such as obtained via cellular phone calls from freeway motorists (see Section 13).

5.2 Evaluation Methodology

As a result of adopting the probabilistic approach to data classification, the conventional algorithm Measures of Effectiveness (MOEs) of detection rate, false alarm rate and mean time to detect do not directly apply to these algorithms.

In order to recover the conventional notion of an incident "alarm" from probabilistic results, all MOEs will be defined in terms of a *detection threshold.* For a given detection threshold, the algorithm will be said to produce an alarm when the algorithm output exceeds the threshold. For instance, if the detection threshold is set to 0.7, an incident alarm will be declared when the algorithm output is greater than 0.7. In order to quantify algorithm performance, all MOEs are defined as *functions* of detection threshold, and algorithm evaluations are therefore based on the value each MOE takes over the entire range of detection thresholds.

For example, consider the evaluation data set which consists of several incident and incident-free congestion events. For any given detection threshold, one can execute the algorithm to be evaluated and tabulate the number of true incident alarms and the number of false incident alarms. From this, one may compute the MOE we will term "operational detection rate," defined as the percentage of incident alarms which are true (this and other MOEs used in algorithm evaluations are described more fully Section 12). By repeating the procedure as detection threshold varies over its range, one can compute the operational detection rate of the algorithm in question *as a function* of detection threshold.

Defining MOEs in this manner allows one to investigate an algorithm's performance sensitivity with respect to various detection thresholds. Specifically, one can identify threshold "regions of performance" and subsequently select appropriate detection thresholds for operational use. For example, consider the plots of Figure 5-3, which show the number of detected incidents and the number of false detections as functions of detection threshold.

For the algorithm in question, it is clear that the optimal detection threshold, at least in terms of these MOEs, lies in the region (0.8, 0.9) This follows from the observation that nearly all incidents are detected in this range, and relatively few false alarms were produced. Conversely, a detection threshold that lies in the region (0.6, 0.7) would be a much poorer choice. For thresholds in this range, the number of detected incidents is only marginally higher than in the previously considered region, but the number of false alarms is considerably increased. To evaluate the algorithm's overall performance with regard to detection rate and false alarm rate, one could perform statistical analysis on the functions of Figure 5-3, such as tabulating the number of detections versus the number of false alarms for representative detection thresholds.

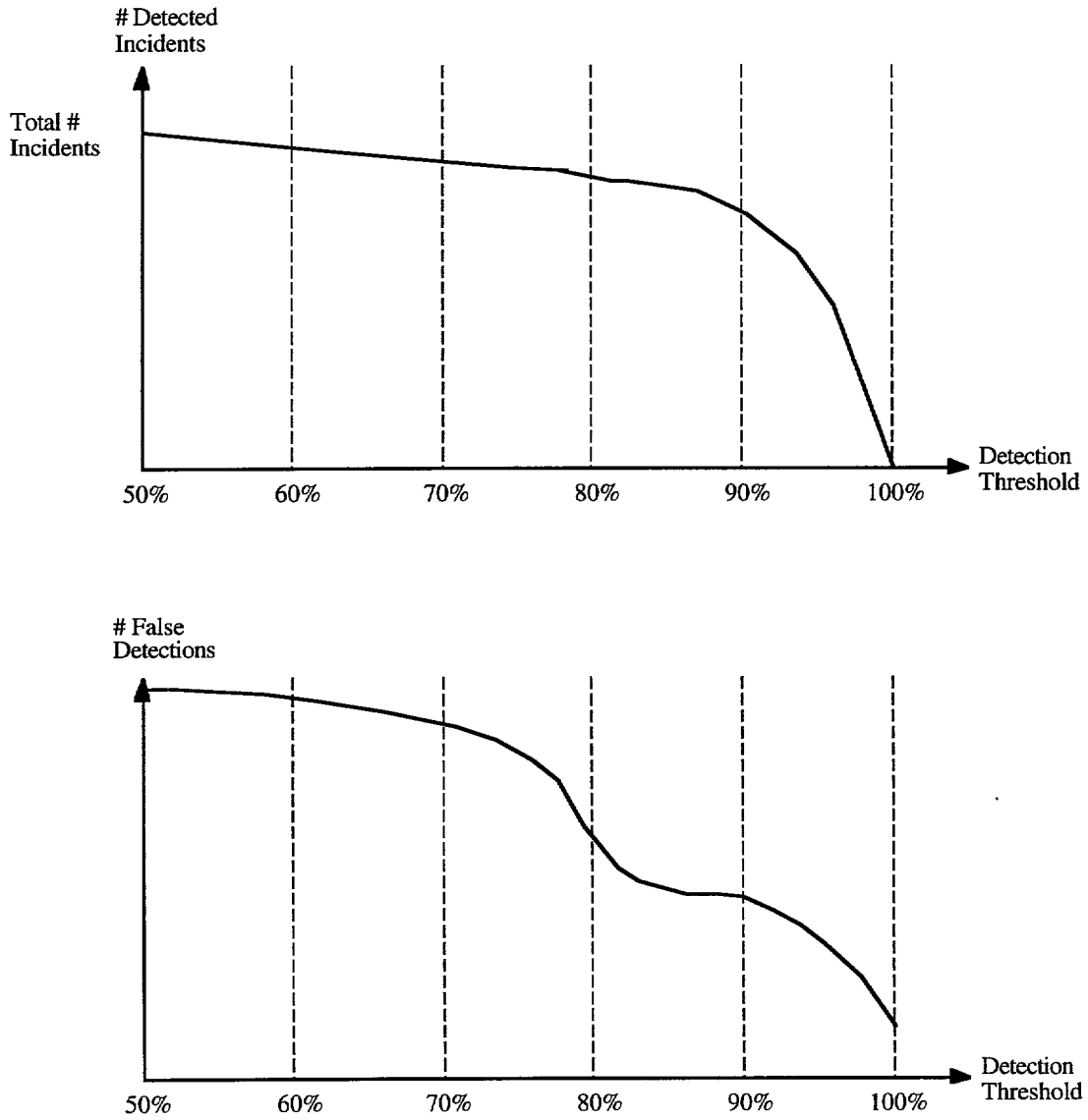


Figure 5-3. MOEs as Functions of Detection Threshold

The California Algorithm(s), each having a collection of threshold sets, can be treated in the same way by considering the index of the threshold sets at the "detection threshold." For example, consider California Algorithm # 8 [PAYN 76], illustrated in Figure 5-4, and with threshold sets presented in Table 5-1.

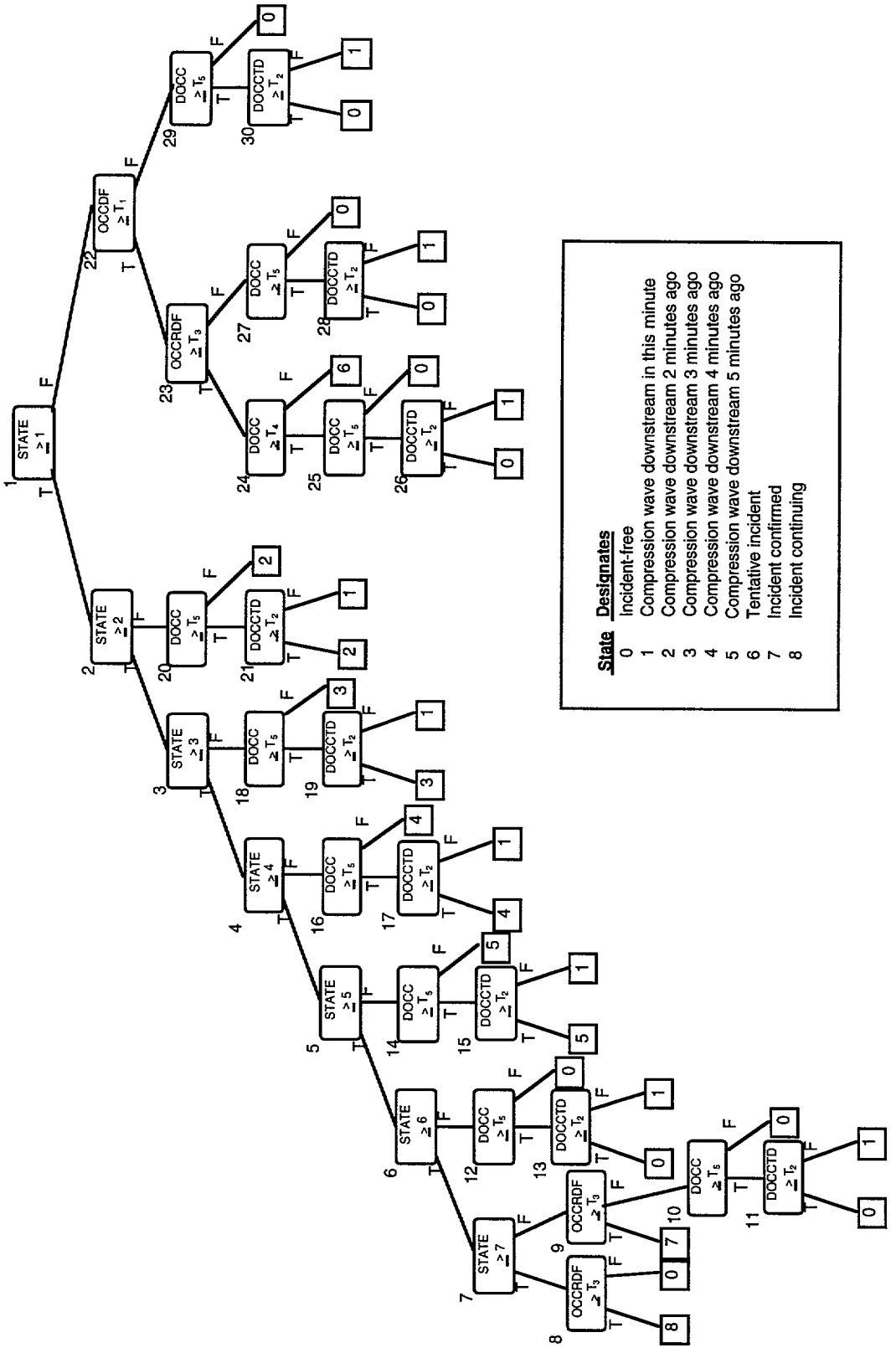


Figure 5-4. California Algorithm # 8

Table 5-1. Threshold Sets for California Algorithm #8, Calibrated for Minneapolis [PAYN 76]

Set #	Detection Rate (%)	T_1	T_2	T_3	T_4	T_4
1	80.6	7.4	-.259	.302	27.3	30.0
2	72.2	17.4	-.649	.391	25.2	30.0
3	61.1	27.8	-.320	.606	28.8	30.0
4	55.6	30.0	-.453	.508	15.4	30.0
5	41.7	30.0	-.677	.724	15.3	30.0
6	36.1	27.8	-.689	.750	11.6	30.0
7	27.8	28.0	-1.084	.792	10.7	30.0

The values of detection rate shown in Table 5-1 were obtained from the calibration results using data from Minneapolis. In operational use, this algorithm can be tested against each of the threshold sets in turn, with the probability of an incident being reported as the highest detection rate associated with a successful indication of an incident. If these values of detection rate are plotted against "Set #," the results are as shown in Figure 5-5.

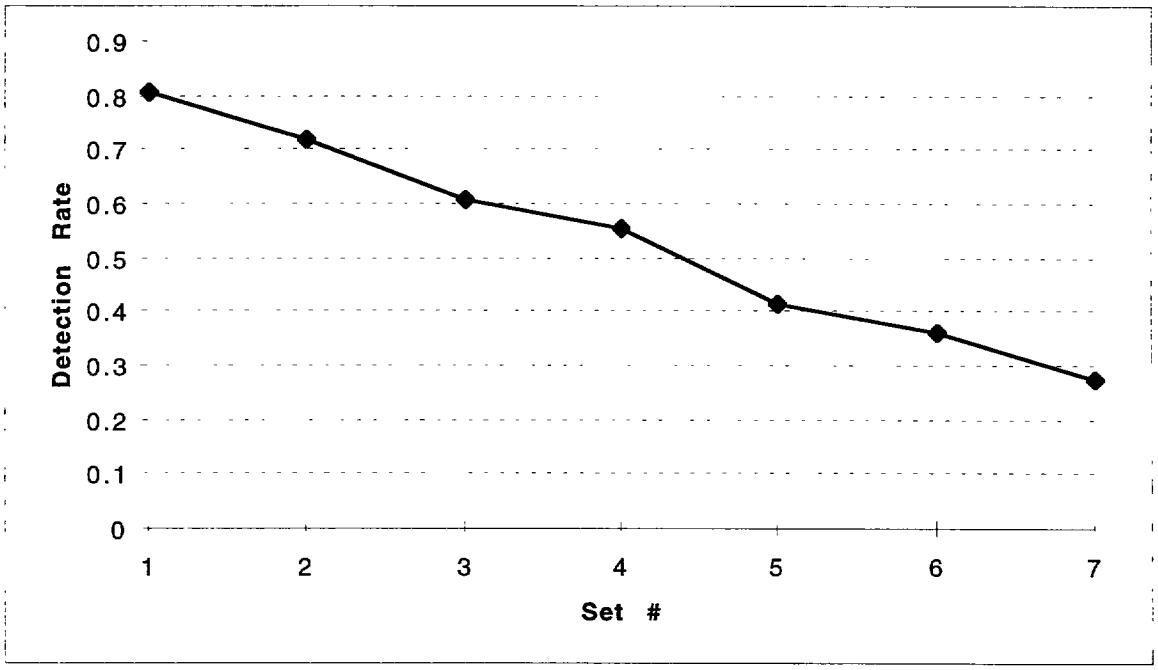


Figure 5-5. Detection Rate vs. Threshold Set # for California Algorithm #8

6.0 DEVELOPMENT SITES AND ASSOCIATED FREEWAY MODELS

In order to pursue empirical development and operational evaluation, real data from operational freeway surveillance systems are needed. The three sites used in our work are identified here.

6.1 The I-880 Site

The portion of the San Francisco Bay Area freeway system selected for real-time field tests consists of sections of I-880, as shown in Figure 6-l. This freeway sub-system was selected primarily due to the availability of an extensive collection of incident and traffic data for this section of I-880. The existence of this data in a readily usable format greatly facilitates the development of incident detection algorithms.

The data was recorded during a 1993 study that addressed the reduction of incident-induced traffic delays through the use of Freeway Service Patrols (FSP) [PETT 94]. The traffic database consists of loop detector data, probe vehicle data and incident data collected over a period of approximately two months. The database, as well as associated documentation and support software, is available to interested parties via the Internet's World Wide Web (http://www-path.eecs.berkeley.edu).

This section presents an overview of the data available from the I-880 database. The discussion is divided into two parts. First, general characterizations of the I-880 geometry and surveillance system are presented. This is followed by a description of the sensor data available from the FSP database.

6.1.1 I-880 Freeway Geometry & Surveillance System Description

The section of I-880 under consideration spans nine centerline miles of four and five lane freeway and supports traffic between Oakland and San Jose. The test section is bounded by three-lane freeway at each end and intersects both SR92 and SR238. An HOV lane is supported along the entire segment. Refer to Figure 6- 1.

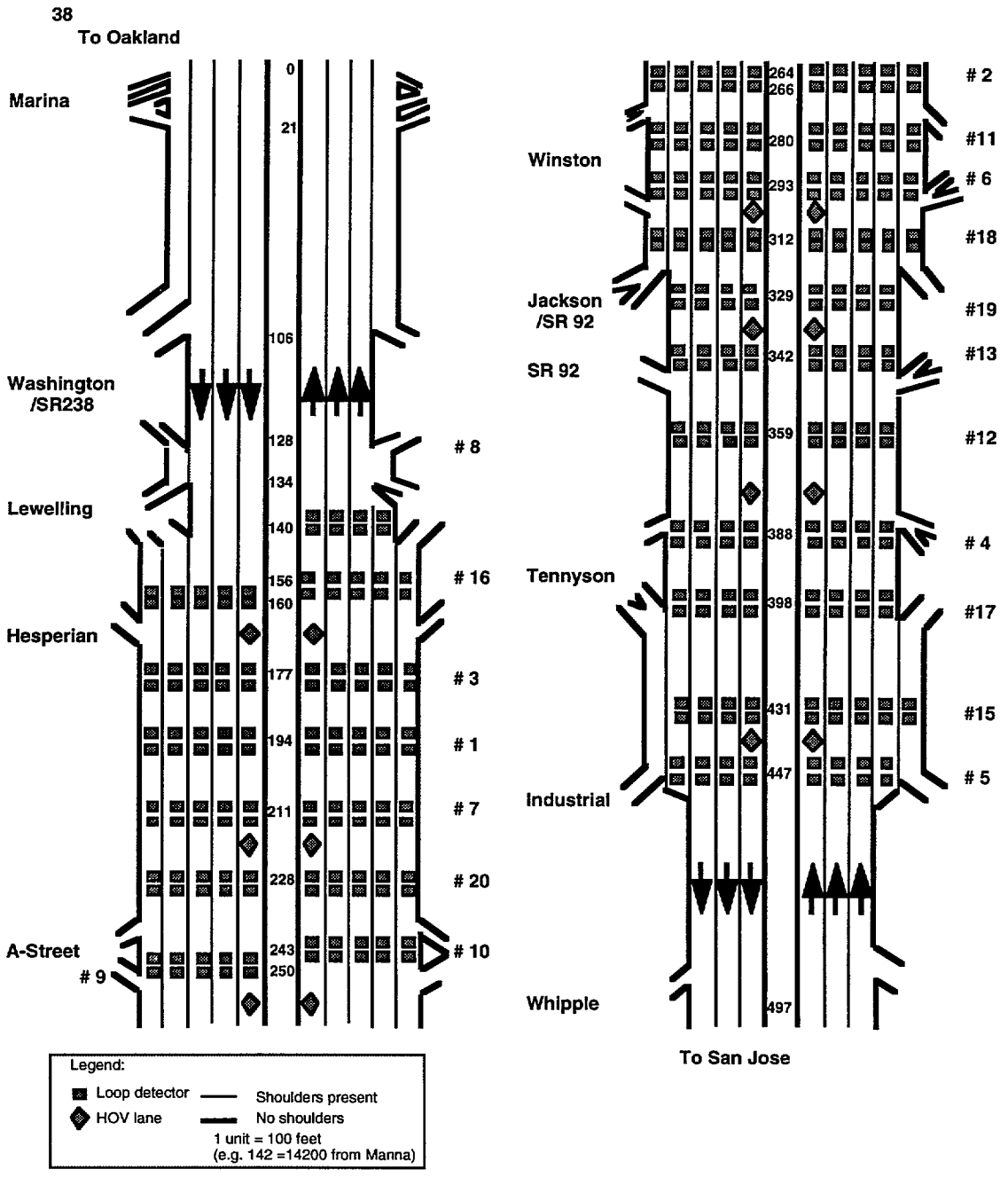


Figure 6-1. Overview of the I-880 Surveillance System

The I-880 surveillance system employs double induction loops in both mainline and ramp lanes and provides measurements of volume, occupancy, and speed for all sensors at 30-second intervals. The raw sensor data is fed to *controllers* for formatting and distribution.

Mainline sensors are grouped into logical ***stations*** that span all freeway lanes.Mainline stations are generally located near ramps at approximately 1/3 mile spacing. All mainline stations include a sensor in the HOV lane. It should be noted that certain stations also include a sensor in an auxiliary lane or in lanes that just started or soon will end. Hence, the distribution of traffic across a station is generally not uniform.

The naming convention for stations is as follows. Each controller listed in Figure 6-1 serves two mainline stations. The station identifiers consist of three parts: (a) "ST" for station, (b) the controller number, and (c) either a "1" to indicate a northbound station or a "0" to indicate a southbound station. Hence, station identifier ST1 10 corresponds to the southbound station for controller 11.

6.1.2 I-880 Data Characteristics

Sensor data is available for dates between 2/16/93 and 3/19/93 and also from 9/27/93 to 10/29/93. The data was collected on weekdays only during the hours of peak traffic levels. The AM peak period spans the hours between 5:00 and 10:00 in the morning, while the PM peak period includes data from 2:00 to 8:00 in the evening. The selected portion of I-880 generally experiences high levels of congestion during the morning and evening peak periods.

6.1.3 Modeling of I-880 Site

Modeling of the I-880 Site was based on zones bounded by mainline loop stations. There were 35 mainline stations, and 33 zones defined from them. The complete model as implemented in RIDE is documented as a file on the delivered CD-ROM.

6.2 The Twin Cities Site

The Twin Cities freeway system was selected for use for several reasons: (1) the freeway system exhibits significant recurrent congestion, (2) existing CCTV is available for the region to aid in labeling congestion as either incident or incident-free, and (3) the Minneapolis department of transportation exhibited a willingness to cooperate with our efforts.

6.2.1 Modeling of Twin Cities Freeway System

Modeling of Minneapolis freeways entailed acquisition of geometric and demand data from the Traffic Operation Department at the Minnesota Department of Transportation. Fortunately, previous modeling work made this data readily available.

Figure 6-2 shows an overview of the Minneapolis metropolitan area. (Also shown are the three segments simulated in FRESIM. The segments include parts of I-35W, I-94, and I-494.) Essentially all freeways have surveillance.

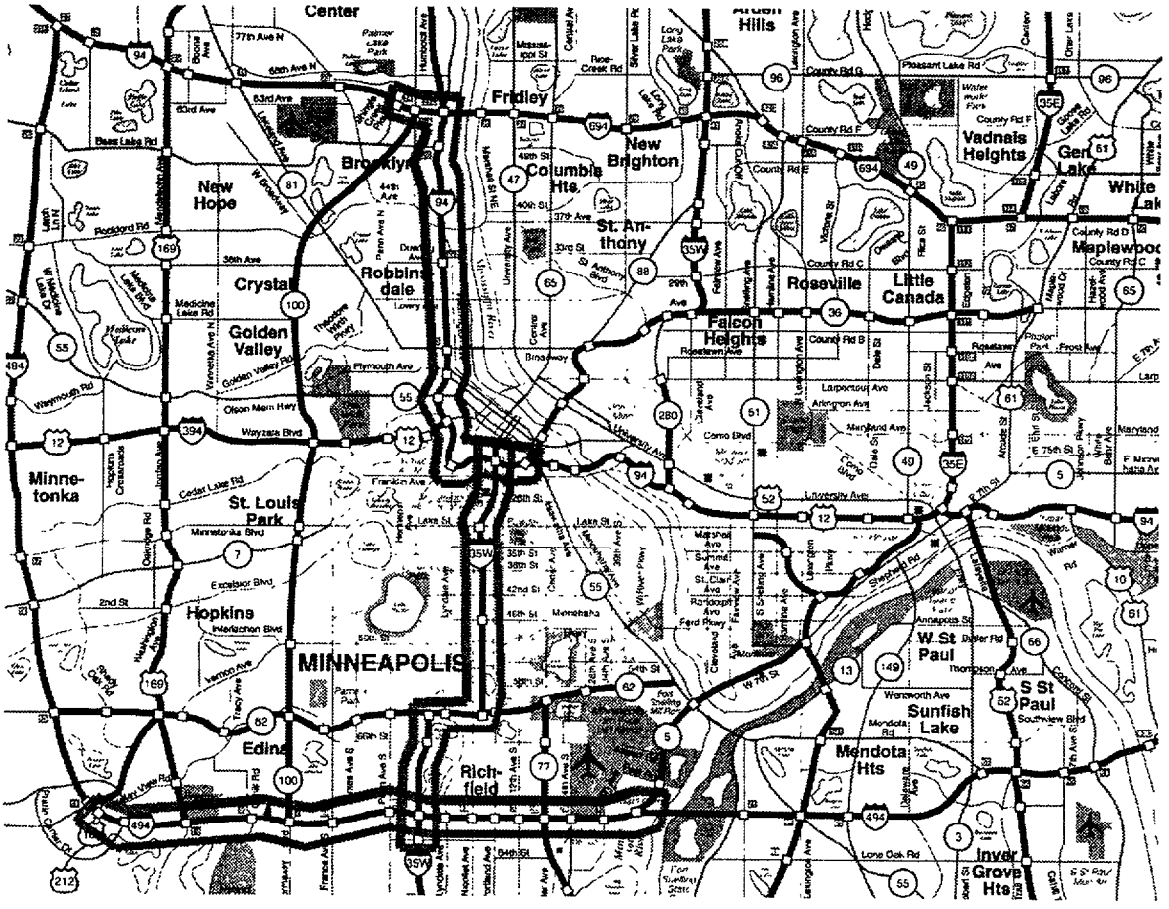


Figure 6-2. Overview of the Twin Cities Surveillance System

6.2.1 Twin Cities Freeway Geometry & Surveillance System Description

In the Twin Cities freeway system, mainline loop stations covering all lanes with a single loop are typically placed approximately midway between an on ramp and the next

downstream off ramp. This placement is intended to allow for measurements in areas away from the weaving associated with ramps. Figure 6-3 illustrates a typical placement. There were 677 mainline stations at the time in 1997 that we conducted our work. (Areas modeled by FRESIM are completely documented in the FRESIM report [CHAN 97].)

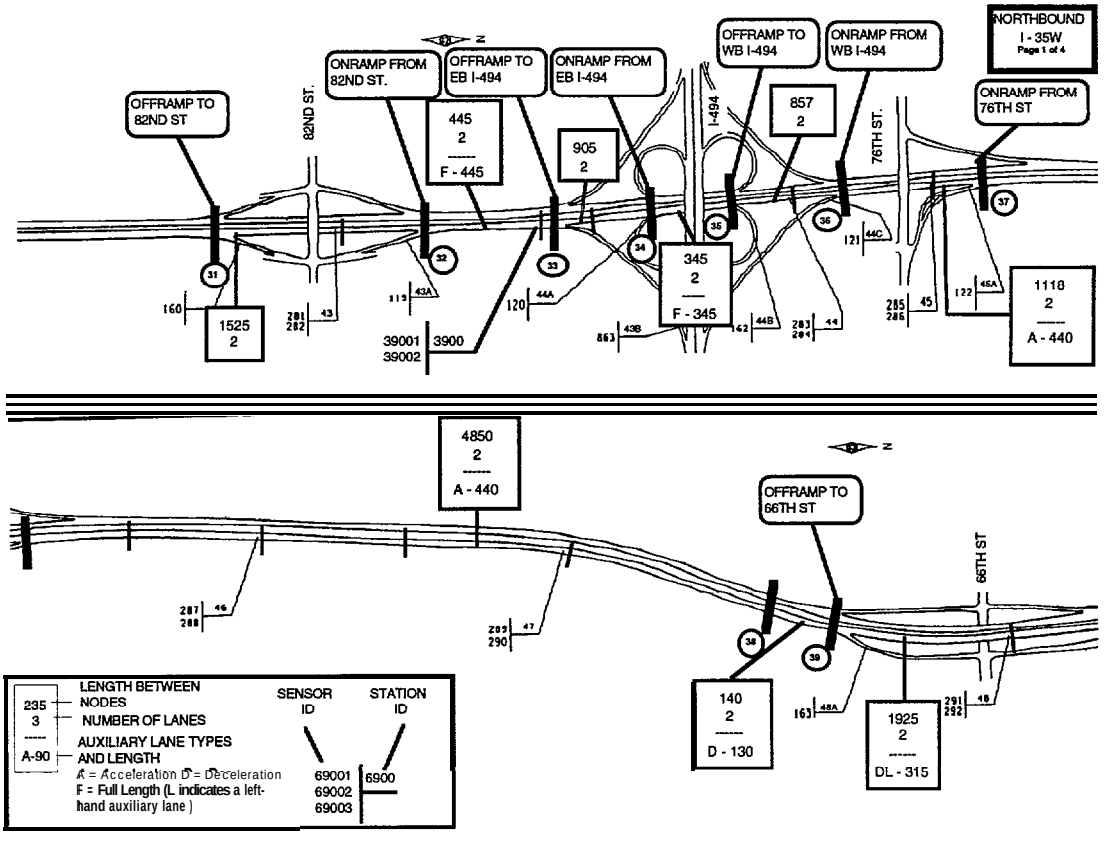


Figure 6-3. A typical Placement of Surveillance Loops in the Twin Cities Freeway System

6.2.2 Twin Cities Data Characteristics

Data for the Twin Cities Freeway System were available to us over the intemet, and locally while were in the TMC on a LAN, as 30-second station data, and 5-minute sensor data, each with associated data quality indicators. We did not make use of the individual loop data.

6.2.3 Modeling of the Twin Cities Site

Modeling of the Twin Cities Site was based on zones bounded by mainline loop stations. There were 677 mainline stations, and 641 zones defined from them. The complete model as implemented in RIDE is documented as a file on the delivered CD-ROM.

6.3 **The San Diego Site**

Figure 6-4 shows an overview of the San Diego metropolitan area. (Also shown are the freeway segments simulated in FRESIM. The simulated segments include parts of Hwy 94 EB, Hwy 94 WB, I-8 WB, and I-5 SB.)

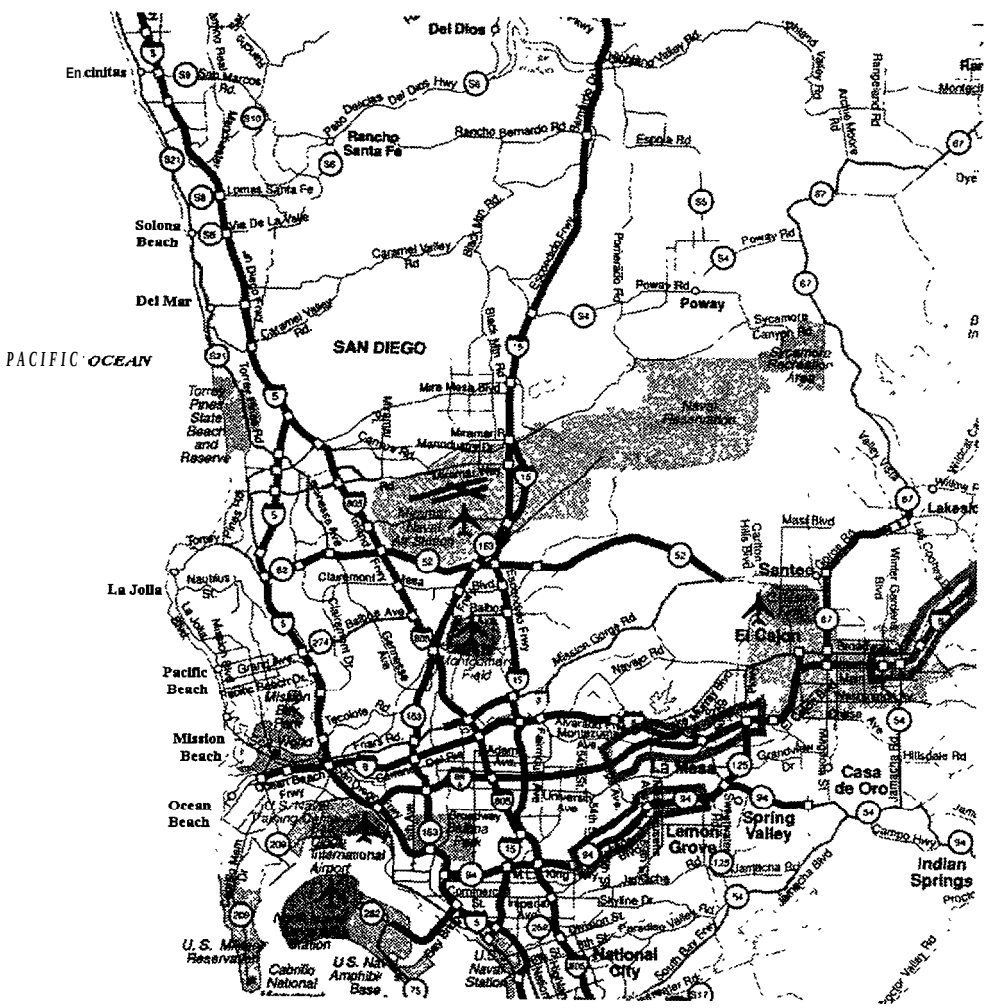


Figure 6-4. Overview of the San Diego Metropolitan Area

6.3.1 San Diego Freeway Geometry & Surveillance System Description

In the San Diego Freeway System, mainline loop stations covering all lanes with a single loop are typically placed in line with an on ramp stop line. This placement is intended to provide a screen-line count at each on ramp. There were 250 mainline stations at the time in 1997 that we conducted our work. Figure 6-5 shows those segments of the San Diego Freeway System that are instrumented (bounded by mainline loop stations), indicated as "Normal."

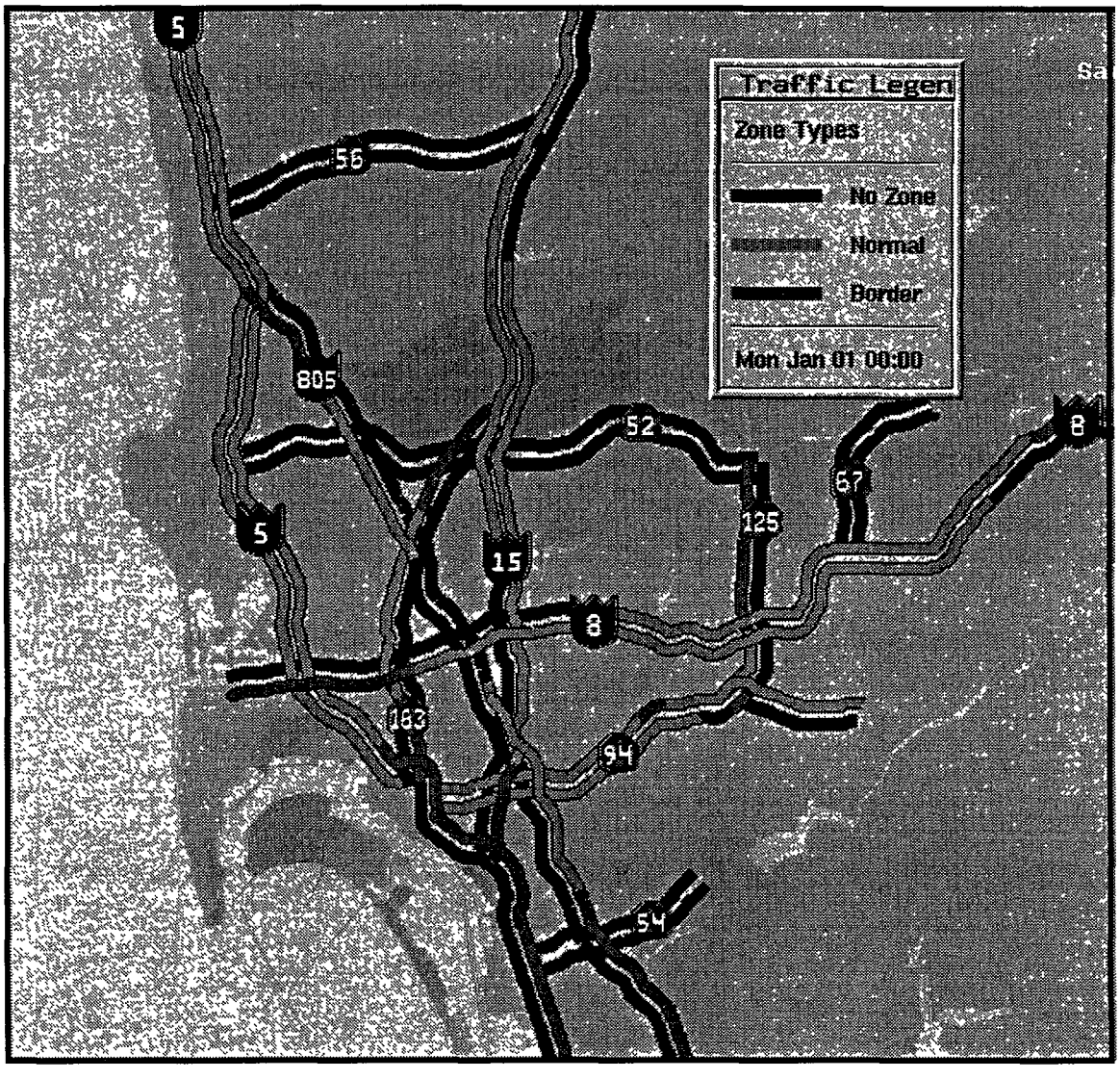


Figure 6-5. Segments of the San Diego Freeway System Instrumented with Loop Surveillance

6.3.2 San Diego Data Characteristics

Data for the San Diego Freeway System were available to us via a serial telephone line, or locally at the TMC on a serial connection, as 30-second individual loop data from mainlines, on ramps and off ramps, though data from off ramps was often not available. No malfunction indications were included in this data.

6.3.3 Modeling of the San Diego Site

Modeling of San Diego freeways was based on generally available geometric data, complemented by examination of sites. There were 250 mainline stations, and 226 zones defined from them. The complete model as implemented in RIDE is documented as a file on the delivered CD-ROM.

7.0 MALFUNCTION DETECTION AND DATA REPAIR TECHNIQUES

All induction-loop sensors are subject to malfunction, and such malfunctions can dramatically and adversely affect subsequent processing of the sensor data. As such, real-time detection of sensor malfunctions is a critical issue for many applications. Also of interest is the ability to "repair" faulty sensor data when possible by estimating traffic measurements for malfunctioning sensors.

In our case, the study of sensor malfunctions is motivated by the desire to develop advanced incident detection algorithms. A significant problem encountered in previous incident detection efforts was that faulty measurement data frequently induced false alarms and significantly reduced the performance of the resultant algorithms. Clearly, such adverse effects could be mitigated through effective malfunction detection and subsequent data repair.

7.1 Malfunction Detection Algorithms

The task of a malfunction detection processing is to inspect incoming sensor data for validity and to declare a sensor malfunction when the measurement data appears faulty. The malfunction detection algorithm must also address a malfunctioning sensor's return to normal operation.

Based on these design criteria, a malfunction detection algorithm was developed and tested using recorded sensor data from the I-880 database. The algorithm was then tested and refined using data from the San Diego surveillance system. As individual lane data is available for both of these sites, the developed algorithm includes validity tests for lane-specific measurement data. However, individual lane data is generally not available for the Twin Cities site. The malfunction detection algorithm was therefore modified for use in this location.

The organization of this section is to first present an overview of the processing objectives identified in developing the malfunction detection capability. This is followed by separate descriptions of the algorithms to be applied for each development site according to the availability of individual lane data.

7.1.1 Overview

It should be noted that the topic of malfunction detection for induction loop sensors has been explored previously. In general, two basic approaches have been pursued: (a) apply reliability checks to the aggregate traffic measurements of volume and occupancy, usually by establishing certain thresholds beyond which the data cannot be said to reflect actual traffic operations, and (b) process the raw signal from the loop directly, wherein the sensor "on" and "off' indications, from which volume and occupancy are computed, are inspected for credibility. Examples of algorithms that utilize the first approach can be found in [PAYN 76] and [NIHA 90]. The second approach was pursued by the Institute of Transportation Studies at the University of California, at Berkeley, and the resulting algorithms are described in [CHEN 86]. While algorithms that utilize a mixture of the two approaches hold the most promise for effective diagnosis of faulty sensor data, the algorithms described herein focus solely on the first approach, since only aggregate traffic measurements were available for the development sites used in this study.

Additionally, the malfunction detection logic presented here is intended only to identify obviously faulty data only. Hence, conservative validity tests are applied to the surveillance data in an effort to ensure that valid traffic data is not interpreted as a sensor malfunction.

Additional tests may be devised which compare traffic measurements from adjacent traffic lanes in order to detect more subtle malfunctions; however, such malfunctions are exceedingly difficult to diagnose using a universal set of validity tests. If one seeks to detect subtle malfunctions by contrasting sensor measurements from adjacent lanes, the nominal traffic characteristics of the station must be taken into account. Elements affecting the nominal traffic characteristics of a station include freeway geometry, time of day, special lanes such as HOV and auxiliary lanes, the proximity of ramps, the existence of incidents, construction projects, weather and many other factors. Due to the natural variation in these parameters, nominal traffic characteristics can vary greatly from station to station and also from one time to the next.

Nevertheless, subtle sensor malfunctions can be detected through careful inspection of the sensor data on a large temporal scale. For instance, if a particular mainline sensor reports a maximum occupancy measurement of 15% during peak traffic periods, while occupancy measurements from surrounding sensors consistently reach values of 25%, one might

conclude that the sensor reporting 15% occupancy is in error. In doing so, however, one must rule out all remaining explanations for the disparity between the sensor measurements, such as the existence of an upstream incident. Another possible explanation is that the lane reporting 15% occupancy is an HOV lane or auxiliary lane.

Due to the wide range of factors that must be taken into consideration, this type of analysis generally requires human interpretation of the data; however, software can clearly be written to aid in the analysis. In our case, the malfunction detection logic includes calculation of the mean and the standard deviation of volume and occupancy measurements for each sensor on a 30-minute basis. These statistics may then be reviewed in order to detect subtle sensor malfunctions.

7.1.2 Data Validation for the San Diego and I-880 Sites

Individual lane data is available for the San Diego and I-880 development sites, and the data validation algorithms developed for use in these locations therefore utilizes this information. Data validation for locations where individual lane measurements are not available is addressed in Section 7.1.3.

The malfunction detection algorithm utilizes thirteen distinct validity tests to identify faulty sensor data. A malfunction will be declared if a sensor fails any of these tests. In order to resume non-malfunction status, the sensor in question must pass all validity tests for ten consecutive minutes.

In many instances, both 30-second and 5-minute aggregate measurement values are employed. In such cases, the validation criteria associated with the 5-minute values are generally more restrictive. For example, the upper bound for valid 30-second occupancy is set to 100%, since it is conceivable that a vehicle may remain stationary above a given sensor for an entire 30-second time period. However, in the absence of incidents, the likelihood of a vehicle remaining stationary above a sensor for five consecutive minutes is extremely small. Therefore, the upper bound for valid 5-minute occupancy is significantly less than 100%.

The conditions that constitute faulty sensor data are as follows:

(0) Data Received out of Order

Sensor data must bc received in the correct temporal order. If decreasing data times are observed, a sensor malfunction will be declared. For instance, if a sensor reports the following sequence of data times -- 7:30:00, 7:30:30, 7:31:00, 7:28:30, 7:3 1:30 -- the data for time 7:28:30 was received out of order. Reception of late messages can be a somewhat common occurrence in real-time operations and was observed in the San Diego surveillance system.

(1) Data is Missing

This is by far the most common type of malfunction observed for all development sites and occurs when a sensor simply fails to report measurement data.

(2) Can't Compute Five-Minute Aggregate

A significant portion of the malfunction detection logic is based on five-minute aggregate sensor data. When these aggregates cannot be computed successfully (e.g., after missing data), a malfunction is declared until the aggregates can be computed.

(3) 30-Second Volume Fails Bounds Check

The bounds used in this study for valid 30-second lane volume are (0, 5000) vehicles per hour.

(4) 30-Second Occupancy Fails Bounds Check

The bounds used in this study for valid 30-second lane occupancy are (0, 100) percent.

(5) 30-Second Measurements Contradict -- Zero Occupancy, Non-zero Volume

A malfunction will be declared for any sensor that reports zero occupancy and non-zero volume, since this condition is clearly impossible. Any vehicle which passes the sensor must register a positive value of occupancy. Conversely, zero occupancy implies that no vehicles passed the detector station.

(6) **30-Second Measurements Contradict -- Zero Volume, Non-zero Occupancy**

A sensor reporting zero volume and non-zero occupancy over a 30-second period can be viewed as a potential malfunction. However, such data cannot be deemed faulty with a high degree of certainty. Consider the following examples of valid traffic measurements,

where the sensor is assumed to measure volume on the "leading edge" of the vehicle's signal (analogs exist for "trailing edge" logic): (a) A vehicle is stopped above of the sensor for the entire 30-second period, resulting in zero volume and 100% occupancy, (b) a vehicle begins the 30-second period positioned above the sensor, then moves away from the sensor without another vehicle passing the sensor in the 30-second period, resulting in zero volume and an occupancy greater than zero and less than 100. This second scenario tends to occur in very heavy traffic (e.g., upstream of an incident) or in very light traffic. Both of these scenarios become far less likely when the period of aggregation is increased from 30 seconds to five minutes. Hence, this type of validity check may be appropriate for longer aggregation intervals but cannot be used to reliably diagnose faulty sensor data using a 30-second aggregation period.

(7) 30-Second Measurements Contradict -- Unreasonable Speeds

A malfunction will be declared for any sensor reporting volume and occupancy data that implies unreasonable traffic speeds (e.g., the corresponding traffic speed would be unreasonably high). This validity test is only applied when traffic levels are sufficiently high -- at least four vehicles must be used in the speed calculation. The criteria employed in this test was developed as follows.

The objective is to establish a lower bound on occupancy as a function of volume. For each passing vehicle, the loop will be activated for a period of time (T) that can be defined in terms of the vehicle's speed (S) and effective length (L_e) as follows:

$$T = L_e/S$$

The value of L_e accounts for the vehicles true length, the length of the induction loop and the loop's activation threshold. Clearly, T is minimized for the minimum value of L_e and the maximum value of S. For this validity test, the minimum value of L_e was taken to be 14 feet (10 foot vehicle, 4 foot loop), and the maximum value of S was taken to be 80 mph (approximately 117.3 feet per second). For the I-880 surveillance measurement interval of 30 seconds, the occupancy attributable to such a vehicle is given by:

$$\text{Occupancy} = 100.0 * (T / 30.0) = 0.4\%$$

The malfunction logic therefore requires that the total measured occupancy of the sensor be at least 0.4*V, where V is the number of vehicles passing the sensor during the

measurement interval. Recall that V must be at least four vehicles for this test to be applicable. This minimum volume requirement was imposed to prevent the passage of a single motorcycle from being interpreted as a sensor malfunction.

(8) Five-Minute Volume Fails Bounds Check

The upper bound adopted for five-minute lane volume is 3000 vehicles per hour. The lower bound for five-minute lane volume depends on the type of lane being instrumented and the overall traffic level. For ramps and the mainline HOV lane, or when the average five-minute volume over all mainline stations is less than 360 vehicles per hour (three vehicles per 30-seconds), the minimum allowed five-minute lane volume is zero vehicles per hour. Otherwise, the minimum allowed five-minute lane volume is ten vehicles per hour.

(9) 5-Minute Occupancy Fails Bounds Check

The upper bound adopted for five-minute lane occupancy is 95%. As with volume, the lower bound for five-minute lane occupancy depends on lane type and the overall traffic level. For ramps and the mainline HOV lane, or when the average five-minute lane volume over all mainline stations is less than 360 vehicles per hour (three vehicles per 30-seconds), then the minimum allowed five-minute lane occupancy is 0%. Otherwise, the minimum allowed five-minute lane occupancy is 0.075%.

(10) Sensor is Known to be Faulty

A flag can be set manually to indicate to the malfunction detection logic that certain sensors consistently report faulty data. Malfunctions will be declared for such sensors indefinitely. This capability is intended for operational purposes. For example, if the malfunction detection logic indicates that a sensor is malfunctioning 70% of the time, an operator may conclude that the data from this sensor should be disregarded completely. The manual flag gives the operator a means to review all relevant data and make his/her own judgments regarding sensor malfunctions.

(11) Sensor Data Times Disagree

A malfunction will be declared for all sensors within a given station if the sensor data times do not match.

(12) Lane Deviation Tests -- Lane with Occupancy too Large

This test was suggested during development of the California algorithm [PAYN 76] and is intended to detect a single malfunctioning sensor out of an otherwise operational station. This test does not apply to the HOV lane or other lanes that are known to exhibit low volume levels. A malfunction is declared if: (a) the average five-minute occupancy of all applicable lanes is less than 20%, (b) the lane in question has a five-minute occupancy that is twice as large as the average five-minute station occupancy among applicable lanes, and (c) the lane in question has reported a 30-second occupancy value greater than 50% in the previous five minutes.

(13) Lane Deviation Tests -- Lane with Occupancy too Small

This test was suggested during development of the California algorithm [PAYN 76] and is intended to detect a single malfunctioning sensor out of an otherwise operational station. This test does not apply to the HOV lane or other lanes that are known to exhibit low occupancy levels. A malfunction is declared if: (a) the average five-minute occupancy of all applicable lanes is greater than 10%, (b) the lane in question has a five-minute occupancy that is more than three times smaller than the average five-minute station occupancy among applicable lanes, and (c) the lane in question has had a 30-second occupancy value less than 2% in the previous five minutes.

7.1.3 Data Validation for the Twin Cities Site

The malfunction detection logic described in the previous section needed to be modified for use in the Twin Cities Freeway System (see Section 6.2). For this surveillance system, individual lane data is available only for a very small portion of the freeway system, and traffic measurements are generally available only as station averages.

As a result, the validity tests of the preceding section needed to be modified to remove dependencies on individual lane data. When possible, the tests were adapted to utilize averaged station data. Some of the tests that depend on individual lane data (e.g., validity tests 12 and 13) could not be adapted for use with station averages and were simply not applied.

Another characteristic specific to the Twin Cities site was the availability of station validity information reported directly from the surveillance system. The surveillance system applies preliminary data validity processing and reports the validity information along with the

traffic measurements. This data was therefore utilized by our processing as described below.

If the surveillance system's data validity field for a given station indicates that the station data is faulty, a malfunction is declared. Categories of faulty data supported by this field include checksum failures, data time-out (e.g., late data), line failures and stations that are off-line. The validity field also indicates when a station fails rudimentary bounds checking applied by the surveillance system, but these bounds tests are replaced by the validity checks applied by our algorithm as identified below.

(1) Data is Missing

This occurs when a station simply fails to report measurement data.

(2) Can't Compute Five-Minute Aggregate

A significant portion of the malfunction detection logic is based on five-minute aggregate station data. When these aggregates cannot be computed successfully (e.g., after missing data), a malfunction is declared until the aggregates can be computed.

(3) 30-Second Volume Fails Bounds Check

The bounds used in this study for valid 30-second station volume are (0, 3000) vehicles per hour per lane.

(4) 30-Second Occupancy Fails Bounds Check

The bounds used in this study for valid 30-second station occupancy are (0, 100) percent.

(5) 30-Second Measurements Contradict -- Zero Occupancy, Non-zero Volume

A malfunction will be declared for any station that reports zero occupancy and non-zero volume, since this condition is clearly impossible. Any vehicle which passes the station must register a positive value of occupancy. Conversely, zero occupancy implies that no vehicles passed the detector station.

(6) Five-Minute Measurements Contradict -- Zero Volume, Non-zero Occupancy

A malfunction will be declared for any station reporting zero volume and non-zero occupancy over a five-minute period, since this condition is highly unlikely for actual traffic conditions.

(7) 30-Second Measurements Contradict -- Unreasonable Speeds

A malfunction will be declared for any station reporting volume and occupancy data that implies unreasonable traffic speeds (e.g., the corresponding traffic speed would be unreasonably high). This validity test is only applied when traffic levels are sufficiently high -- at least four vehicles must be used in the speed calculation. The criteria employed in this test is the same as described for San Diego and I-880.

(8) Five-Minute Volume Fails Bounds Check

The upper bound adopted for five-minute station volume is 2500 vehicles per hour per lane. The lower bound for live-minute station volume depends on the overall traffic level. When the average five-minute volume over all mainline stations is less than 360 vehicles per hour (three vehicles per 30-seconds), the minimum allowed five-minute station volume is zero vehicles per hour per lane. Otherwise, the minimum allowed five-minute station volume is ten vehicles per hour per lane.

(9) 5-Minute Occupancy Fails Bounds Check

The upper bound adopted for five-minute station occupancy is 95%. As with volume, the lower bound for five-minute station occupancy depends the overall traffic level. When the average five-minute volume over all mainline stations is less than 360 vehicles per hour per lane (three vehicles per 30-seconds), then the minimum allowed five-minute station occupancy is 0%. Otherwise, the minimum allowed five-minute station occupancy is 0.075%.

(10) Station is Known to be Faulty

A flag can be set manually to indicate to the malfunction detection logic that certain stations consistently report faulty data. Malfunctions will be declared for such stations indefinitely. This capability is intended for operational purposes. The manual flag gives the operator a means to review all relevant data and make his/her own judgments regarding long-term station malfunctions.

7.2 Data Repair Algorithms

Once a sensor malfunction has been detected, appropriate actions must be taken. One option is to simply screen the faulty data from subsequent processing. In this approach, the offending data is treated as "missing" and no further processing is attempted. A more ambitious approach is to "repair" the faulty data by estimating traffic measurements for the sensor in question. Such an approach attempts to minimize the impact of sensor malfunctions by maintaining the scope of freeway processing to the extent possible. For instance, the malfunction rate of a nominal I-880 station was found to be approximately 21%, as discussed in Section 7.3. If no corrective measures are taken, such a station would simply be unavailable for further processing 21% of the time. However, this rate can be mitigated by applying data repair algorithms.

The repairing algorithm developed in this study is applicable only to locations where individual lane data is available. Historical traffic distributions, as well as current measurements from lanes adjacent to the malfunctioning sensor, are utilized to estimate traffic measurements for the lane containing the malfunctioning sensor. Consider the case where measurement data from a single lane of a full-count detector station is unavailable due to a malfunction. An estimate of the volume attributable to this lane can be generated by extrapolating the volume measurements of adjacent lanes according to the historical distribution of volume across lanes for this detector station. Occupancy measurements can be estimated in an analogous manner. To avoid redundancy, the discussion which follows describes the estimation process in detail for volume data only.

The estimation scheme requires that the distribution of volume across all traffic lanes be known. In order to compute this quantity, a general characterization of traffic volume for each lane is required. This is obtained by estimating the volume for each lane over a period of time long enough to eliminate short-term traffic fluctuations. This long-term volume estimate is termed the steady-state volume. This quantity is estimated for each full-count station using a Kalman filter that models steady-state volume as a Markov process with a fairly long time constant. The estimation logic is described in detail below.

Assume one is given l, noisy measurements of lane volume for station j over time slice n modeled as:

$$z_{j,m}^{n} = q_{j,m}^{n} + \alpha_{j,m}^{n} \quad , \; m = 1,\ldots,\ell_j , \tag{7.1}$$

where

ℓ_j = Number of lanes for detector station j.

$q_{j,m}^{n}$ = mth lane exit volume, station j, time slice n, and

$\alpha_{j,m}^{n}$ = mth lane exit volume measurement error, station j, time slice n.

The Markov process model used to represent the steady-state volume is defined as:

$$q_{j,m}^{n+1} = e^{-(\Delta t/T_q)} q_{j,m}^{n} + \eta_{j,m}^{n} \tag{7.2}$$

where,

$\eta_{j,m}^{n}$ = the steady-state exit volume process modeling error for lane m at time n,

Δt = time period from time n to time n+1, and

T_q = Markov process time constant for the steady-state volume process model.

Note $\{\alpha_{j,m}^{n}\}$ and $\{\eta_{j,m}^{n}\}$ are assumed to be mutually independent sequences of zero mean, white, Gaussian noise independent of $\bar{q}_{j,m}^{0}$ with corresponding variances σ_α^2 and σ_η^2, respectively.

The steady-state lane volume estimator is solved for using the equations for the Kalman filter solution [ANDE 79]. The measurement update portion of the Kalman filter is defined by the equations:

$$\bar{q}_{j,m}^{n}(+) = \bar{q}_{j,m}^{n}(-) + K_q^n\left[z_{j,m}^{n} - \bar{q}_{j,m}^{n}(-)\right] \quad \text{, and} \tag{7.3}$$

$$P_q^n(+) = P_q^n(-)\sigma_\alpha^2 \,/\, \left(P_q^n(-) + \sigma_\alpha^2\right) \tag{7.4}$$

where,

$$K_q^n = P_q^n(+) / \sigma_\alpha^2$$

Notationally, terms appended by (−) are "a priori" terms and terms appended by (+) are "a posteriori" terms. A priori terms are values prior to the incorporation of the volume measurement at the current time while a posteriori terms are values following the incorporation of the measurement at that time.

The time update portion of the Kalman filter, used to predict the volume at the next time step, is defined by the equations:

$$\bar{q}_{j,m}^{n+1}(-) = e^{-(\Delta t/T_q)} \bar{q}_{j,m}^{n}(+), \text{ and} \tag{7.5}$$

$$P_q^{n+1}(-) = e^{-2(\Delta t/T_q)} P_q^n(+) + \sigma_\eta^2. \tag{7.6}$$

In order to initialize this Kalman filter, values for $\bar{q}_{j,m}^0$, $P_q^0(-)$, σ_α, σ_η, and T_q must be specified.

The fraction of total exit volume associated with lane m of station j can now be computed as:

$$d_{j,m} = \left(\frac{\bar{q}_{j,m}^n(+)}{\sum_{i=1}^{\ell_j} \bar{q}_{j,i}^n(+)} \right) \tag{7.7}$$

where ℓ_j is the number of lanes for station j. These parameters $(d_{j,m})$ can then be used to generate volume estimates for partial-count stations. Specifically, assuming one is given ℓ_j noisy measurements of lane volume $z_{j,m}^n$ for station j over time slice n, the total volume "measurement", $\hat{q}_j^n$, from a partial count station is estimated by:

$$\hat{q}_j^n = \sum_{m=1}^{\ell_j} a_{j,m} z_{j,m}^n \tag{7.8}$$

where

$a_{j,m}$	=	mth lane exit volume measurement weighting factor, station j, i.e.,	
$a_{j,m}$	=	0	(for m = lanes without a detector),
	=	$(1/\Gamma_j)$	(for m = lanes with a detector),
Γ_j	=	fraction of total steady-state flow measured by lanes with detectors, i.e., $\Gamma_j = \sum d_{j,m}$, and	
$d_{j,m}$	=	volume lane distribution value	

Once an estimate of the total station volume is determined, values for single-lane volume are obtained by scaling $\hat{q}_j^n$ with $d_{j,m}$ for malfunctioning lanes.

Sample Processing

To illustrate the performance of the algorithm described above, sample processing is provided for station ST191 from the I-880 data base. The initial values required by the Kalman filter were assigned as follows:

$\bar{q}_{j,m}^0$ = 500 (vehicles per hour)

$P_q^0(-)$ = 100 (vehicles per hour squared)

σ_α = 4.472 (vehicles per hour)

σ_η = 0.316 (vehicles per hour)

T_q = 1,000,000 (seconds)

The values of $\bar{q}_{j,m}^n(+)$ produced by the Kalman filter for lane 2 are shown in Figure 7-1, along with the raw measurement data for this lane (time is expressed in absolute hours). Using these steady-state volume estimates, the distribution $d_{j,m}$ was computed, and, for illustration purposes, volume estimates were generated for lane two of the station. The resulting error between the actual and estimated values of volume for lane two is shown in

Figure 7-2. As the figure shows, the volume estimate is unbiased, and the standard deviation of the error is approximately 300 (vph).

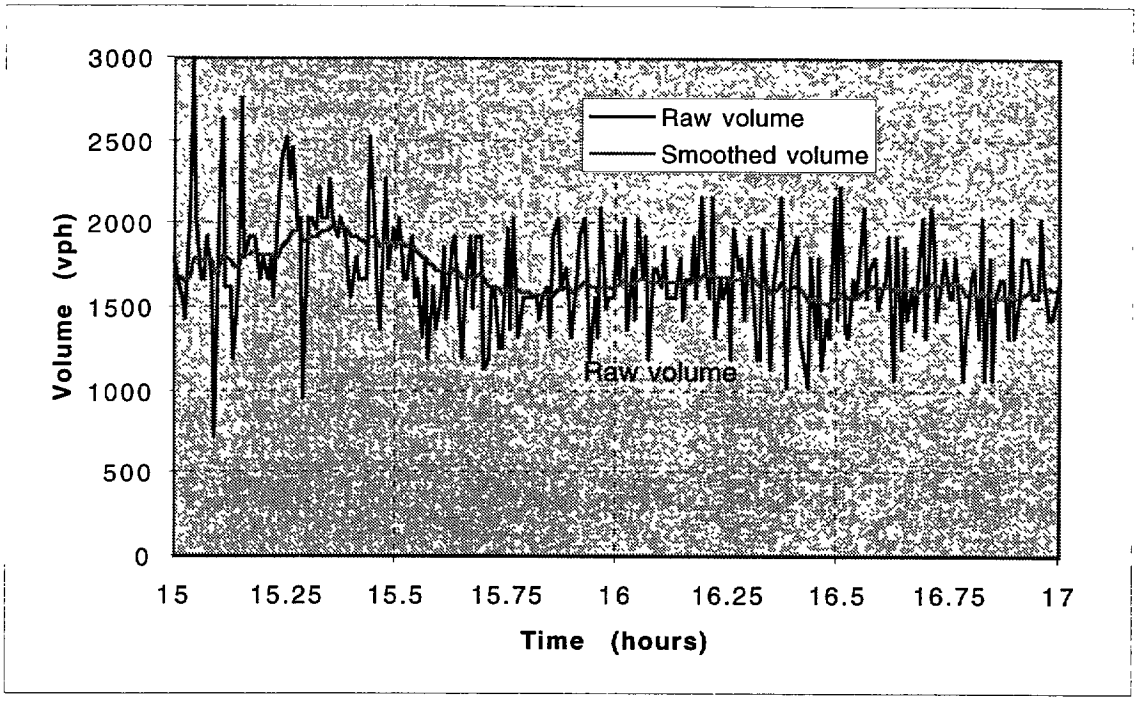


Figure 7-1. Measured Volume vs. Steady-State Volume (Lane 2)

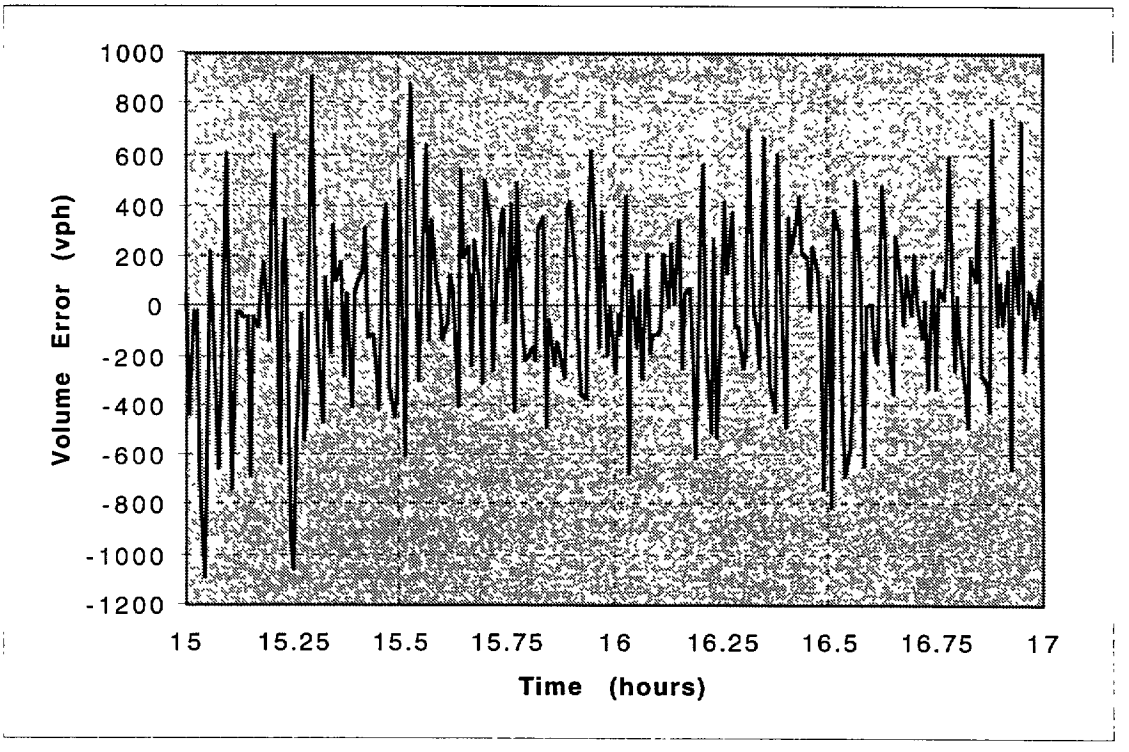


Figure 7-2. Error in Measurement Estimate (Lane 2)

Data Repair for Long-Term Malfunctions

It should be noted that the data repair algorithm described here is applicable to short-term malfunctions only, since traffic distributions can be computed for full-count stations only. If a malfunction persists for several hours, the traffic distributions computed prior to the start of the malfunction may no longer apply to the current traffic situation, resulting in inaccurate measurement estimates. To avoid this situation, a slightly modified approach is adopted for long-term malfunctions.

In the event of a long-term malfunction, the traffic distributions utilized in computing measurement estimates may be taken from a nearby full-count station, usually immediately upstream or immediately downstream from the malfunctioning station. To implement this scheme, each station under consideration must be assigned such a "surrogate" station prior to execution of the algorithm. In doing so, care should be taken to ensure that the traffic distributions of the surrogate station are comparable to those of the station to which it is assigned.

7.3 Application to the I-880 Site

This section describes the malfunction characteristics of the I-880 database in terms of the validity tests identified in Section 7.1.2. The statistics presented herein are based on 277 hours of data collected from 22 six-hour morning periods and 29 five-hour evening periods.

The vast majority of malfunctions observed in the I-880 data set can be attributed to validity tests 1, 2, 8, and/or 9. A much smaller but significant number of malfunctions were detected using validity tests 6, 7, 12, and/or 13. The number of malfunctions that resulted from the remaining validity tests was zero or negligible for the I-880 data set.

It should be noted that a concerted effort was made to maintain the I-880 surveillance system in good condition for the duration of the 1993 FSP study. The malfunction characteristics of older and/or less well-maintained surveillance systems can be expected to be more extreme in nature. Furthermore, the malfunction rates presented here constitute a conservative estimate of the actual malfunction rates since the validity tests employed are generally quite conservative.

Malfunction Characteristks of Individual Sensors

By far, the most common type of malfunction for the I-880 database is missing data, where a sensor simply fails to report measurement data. The second most common type of malfunction occurs when a sensor reports zero volume and/or zero occupancy over an extended period of time. Several less-common, but not negligible, types of malfunctions were also observed and are documented in the discussion which follows. The relative frequency of the various malfunction types is shown in Table 7-l. As shown in the table, nearly 70% of all detected malfunctions resulted from sensors failing to report measurement data (e.g., validity test #l).

Table 7-1. Relative Frequency of Malfunction Types

Validity Test	Relative Freauency
1	66.8%
2	10.11%
3	0.00%
4	0.01%
5	0.03%
6	0.73%
7	2.21%
8	9.19%
9	9.02%
10	0.00%
11	0.00%
12	0.33%
13	1.57%

Based on the validity tests of Section 7.1.2, the overall malfunction rate for I-880 sensors was found to be approximately 11%. That is, any given sensor can be expected to fail one or more of these tests approximately 11% of the time. Equivalently, at any given time, 11% of the sensors can be expected to fail one or more validity test. Of course, this is an overly-simplified description since: (a) certain sensors are far more likely to malfunction that others, and (b) certain malfunctions tend to last for relatively long periods of time. This is illustrated in Figure 7-3.

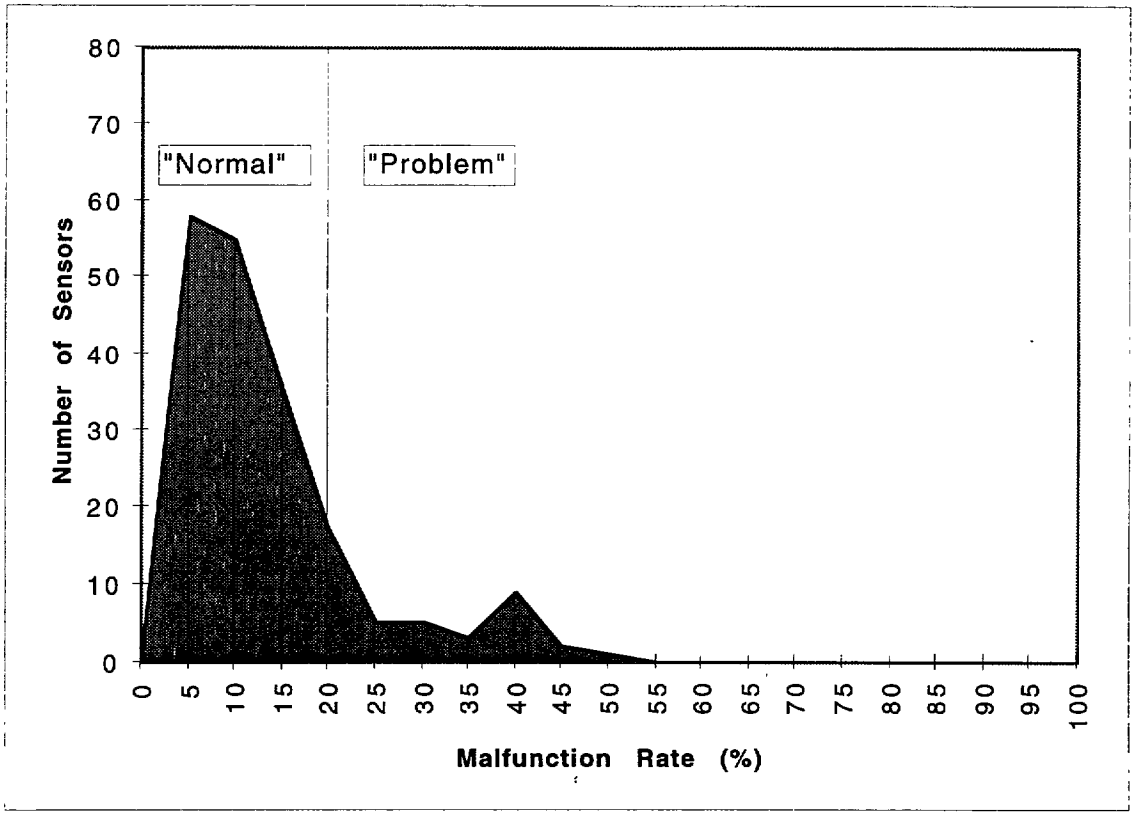


Figure 7-3. Distribution of Malfunction Rate Among I-880 Sensors

The figure shows the distribution of malfunction rate among all I-880 sensors. As shown in the figure, sensors can be divided into two major groups: (a) normal sensors that fail one or more validity test less than 20% of the time, and (b) problem sensors that fail one or more validity test between 20%-60% of the time.

Malfunction Impacts on Station-Based Processing

Almost all processing of sensor data, specifically with regard to traffic monitoring and incident detection, is based not on individual lane data but on *station* averages across all mainline lanes. If one defines a station malfunction as the failure to meet the validation criteria by one or more of the station's constituent sensors, the overall station malfunction rate for the I-880 data set is approximately 14%. A more detailed description of the malfunction rate for stations is presented in Table 7-2 in terms of the number of constituent sensors that failed the validation criteria.

Table 7-2. Station Malfunction Rates

Condition	Rate of Occurrence
Only one constituent sensor failed validity criteria	2%
All constituent sensors failed validity criteria	10%
More than one but not all constituent sensors failed validity criteria	2%
Station failed validity criteria (total)	14%

The distribution of malfunction rate among all I-880 stations is shown in Figure 7-4. Note the slight shift to the right from Figure 7-3.

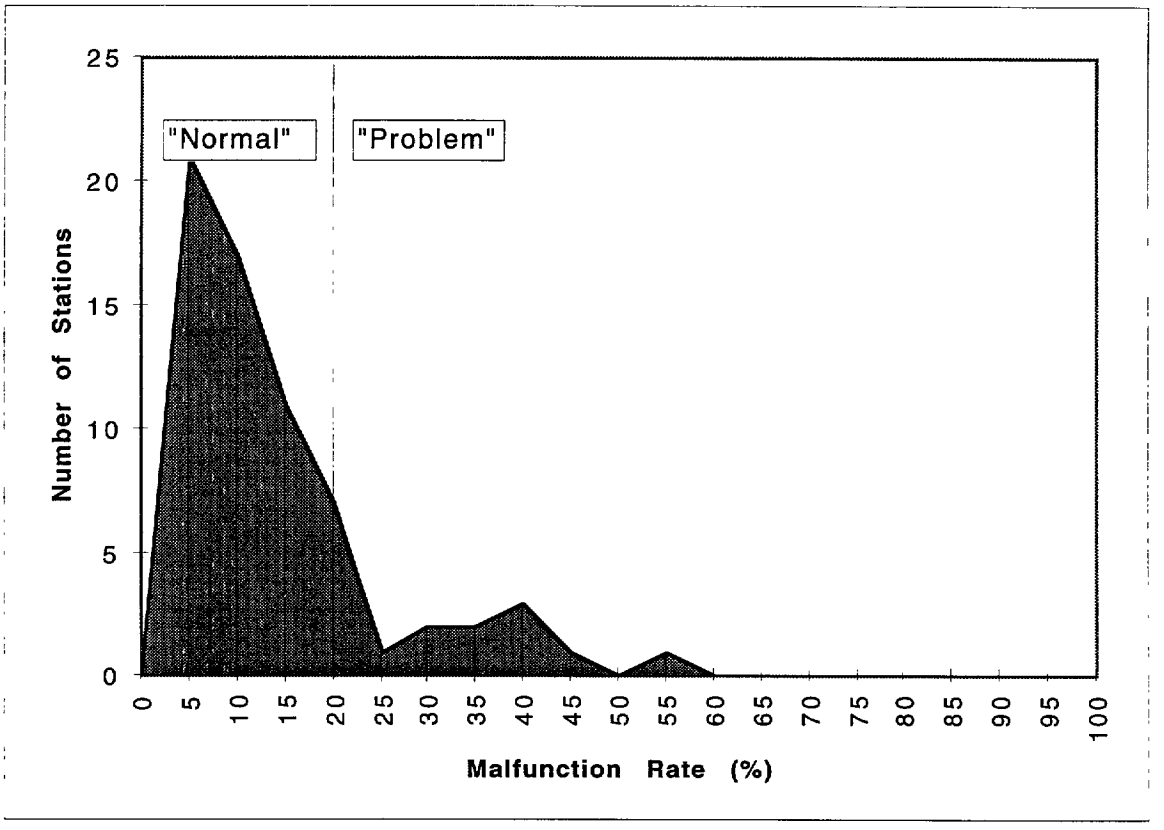


Figure 7-4. Distribution of Malfunction Rate Among I-880 Stations

The malfunction rates of individual I-880 stations are shown in Figures 7-5 and 7-6 for the northbound and southbound directions of travel, respectively. Note that eight out of the

total 34 stations malfunction more that 20% of the time, and four stations malfunction approximately 40% of the time or more.

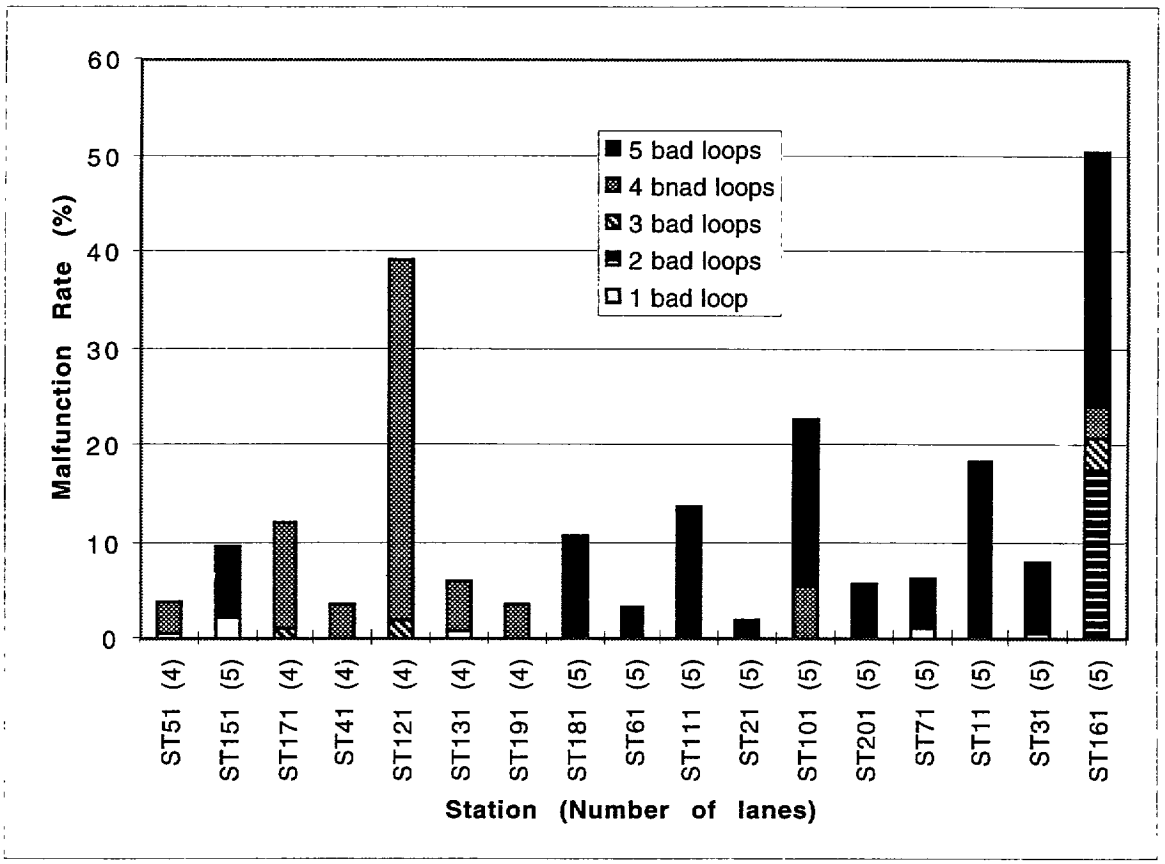


Figure 7-5. Malfunction Rates of Individual I-880 Stations (Northbound)

In the figures, each station's malfunction rate is presented along with the number of constituent sensors that failed the validity criteria. The number of constituent sensors for each station is indicated in parentheses. For instance, the malfunction rate of station ST110 was found to be approximately 31% and can be broken down as follows: all sensors failed the validation criteria at a rate of 13%; a single sensor only failed the validation criteria at a rate of 18%. As indicated in the figure, the majority of station malfunctions resulted from either all sensors failing the validation criteria or from a single sensor failing the validation criteria.

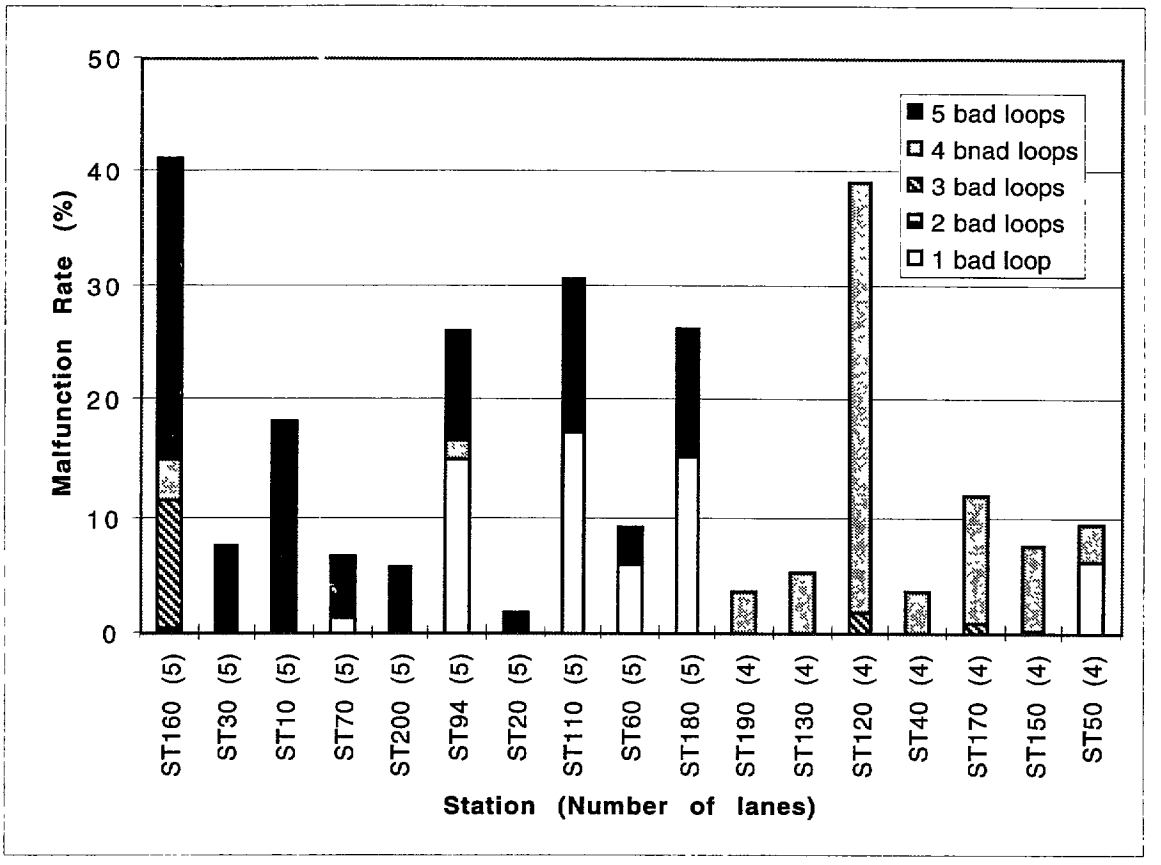


Figure 7-6. Malfunction Rates of Individual I-880 Stations (Northbound)

The naming convention for stations is as follows. Each controller of the I-880 surveillance system serves two mainline stations. The station identifiers consist of three parts: (a) "ST" for station, (b) the controller number, and (c) either a "1" to indicate a northbound station or a "0" to indicate a southbound station. Hence, station identifier ST110 corresponds to the southbound station for controller 11. (See Figure 6-1.)

Malfunction Impacts in Field Operations

All statistics presented to this point address "single-point" validation logic, where malfunction rate is defined simply as the percentage of data points which fail to meet the validation criteria. Actual field operations are likely to adopt more stringent validation requirements, wherein malfunctioning sensors must demonstrate a return to normal operation before once again becoming eligible for processing. The following discussion addresses the malfunction rates likely to be observed when such requirements are imposed.

Specifically, if one imposes the requirement that a sensor must pass all validity tests for ten consecutive minutes in order to resume non-malfunction status, the overall station malfunction rate increases to approximately 21% (13% all sensors bad, 6% only one sensor bad, 2% other). The distribution of this malfunction rate among all I-880 stations is as shown in Figure 7-7, and the malfunction rates of individual I-880 stations are shown in Figures 7-8 and 7-9.

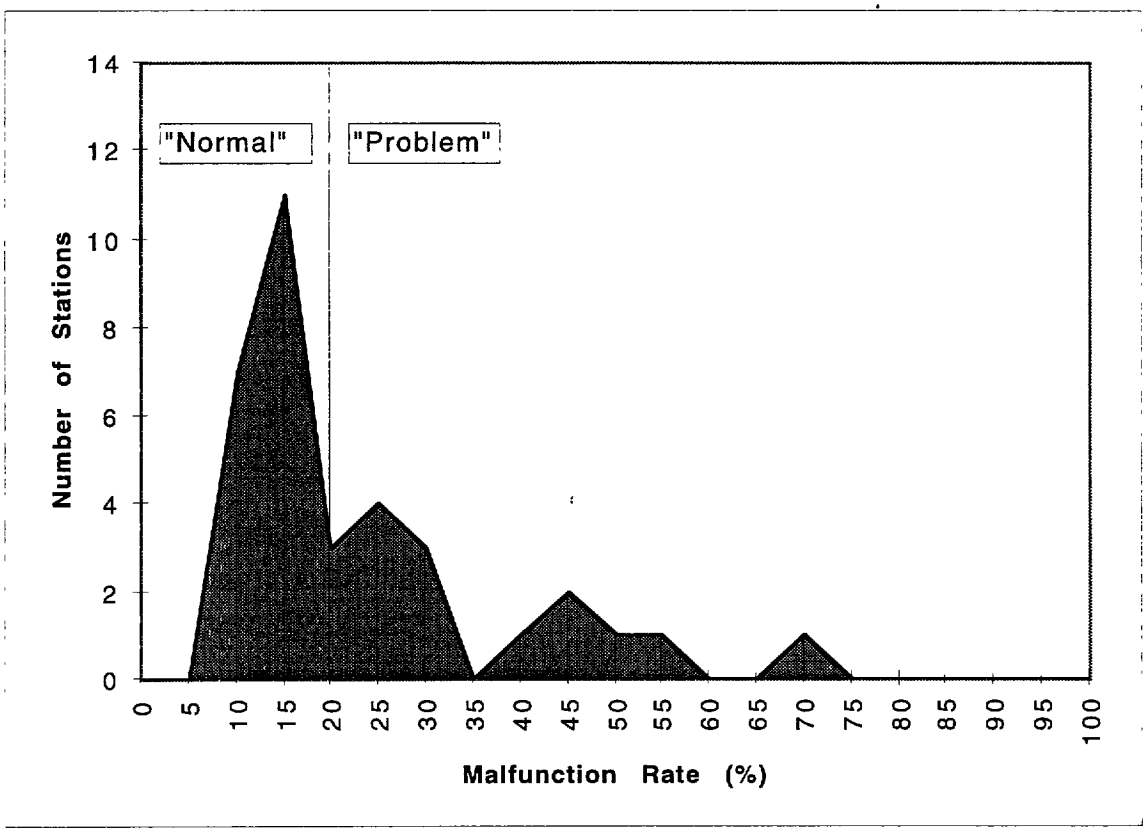


Figure 7-7. Distribution of Malfunction Rate Among I-880 Stations - 10 Minute Wait

By comparing the malfunction rates of Figures 7-8 and 7-9 to those of Figures 7-5 and 7-6, one may determine the portion of the malfunction rate that is attributable to the ten-minute wait requirement. Stations which exhibit a significantly higher malfunction rate in Figures 7-8 and 7-9 indicate that the station malfunctions were intermittent in nature. Conversely, stations that exhibited approximately the same malfunction rates in both figures indicate that the station malfunctions tended to occur in large blocks.

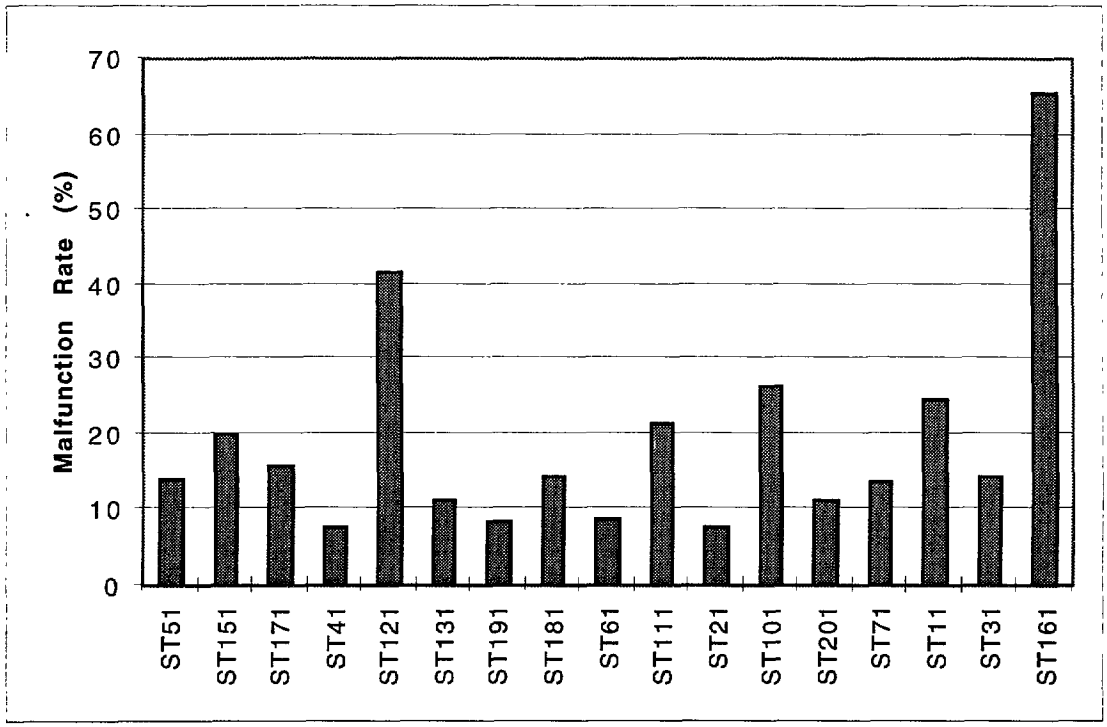


Figure 7-8. Malfunction Rates of Individual I-880 Stations (Northbound) - 10 Minute Wait

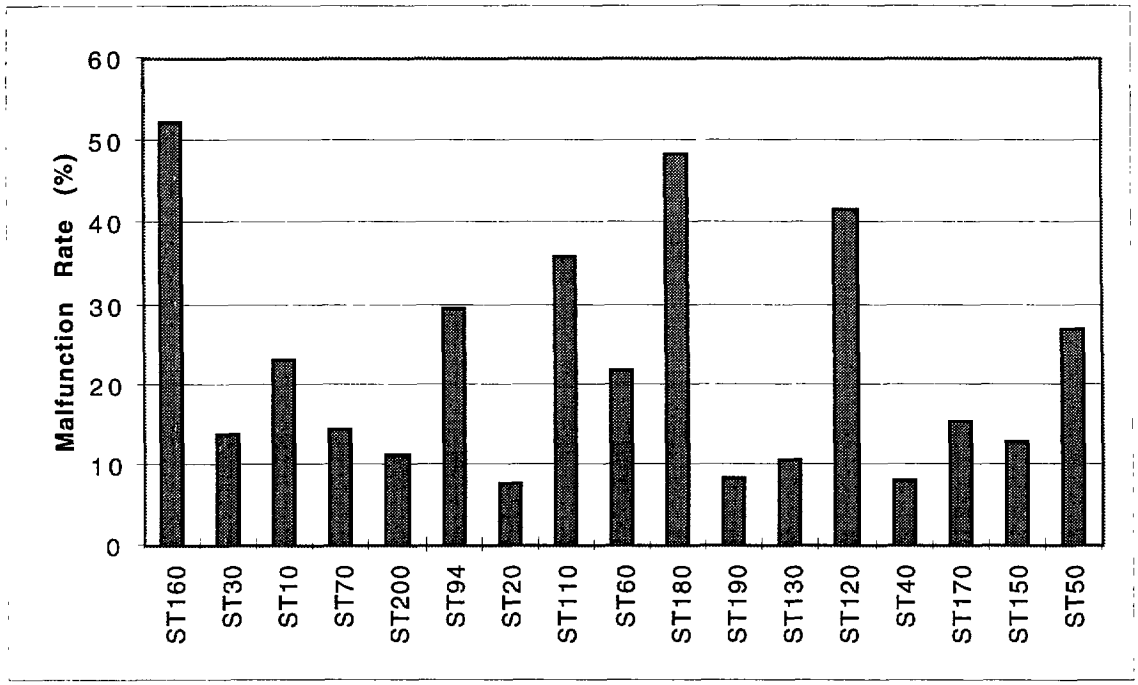


Figure 7-9. Malfunction Rates of Individual I-880 Stations (Northbound) - 10 Minute Wait

8.0 MEASUREMENT CALIBRATION FOR INDUCTION LOOP SENSORS

In the course of our research it has become apparent that sensor-specific calibration is a necessary activity in order to obtain sufficiently accurate measurements of the traffic state, particularly with respect to occupancy and traffic speed. This section documents the procedures employed in calibrating loop measurements for the Twin Cities and San Diego development sites. Specific results are also presented for these sites.

8.1 Calibration Procedures

Over any given time period, an estimate of the mean traffic speed passing a surveillance sensor can be obtained from the induction loop measurements of volume and occupancy as:

$$S = \frac{V}{0 . G} \tag{8-1}$$

where

V = Volume of traffic passing the station during the specified time period,
0 = Station average occupancy over the specified time period,
G = Station G-Factor, and
S = Average traffic speed during the time period.

The station G-Factor serves to convert speed to the appropriate units and also accounts for the average effective length of vehicles passing the detector station. Due primarily to inconsistencies in the computation of occupancy from one station to the next (e.g., stations use different "activation thresholds" for detecting the passage of a vehicle), the value of G must be calibrated for each detector station to arrive at a sufficiently accurate value of speed to be used for incident detection purposes.

Ideally, independent measurements of traffic speed would be utilized in the calibration process. Double induction loops and radar devices are suitable for this purpose. Unfortunately, no such independent measurements were available for the development sites used in this project. As a result, the following calibration process was adopted.

During time periods when freely flowing traffic can be expected, a constant traffic speed may be assumed, and the possibility of unusually slow traffic due to incidents or abnormal congestion can be accounted for by conducting calibrations over a period of several days and checking the G-Factor calibrated for each day for consistency.

For any given day, the calibration process employs Kalman filtering to arrive at the calibrated G-Factor value. To accomplish this, G-Factor "measurements" are computed for each volume/occupancy pair received from the station (at 30-second intervals) assuming a constant nominal traffic speed (our calibrations used 65 mph). We have:

$$\frac{V}{O \cdot G_{Measured}} = 65 \tag{8-2}$$

or

$$G_{Measured} = \frac{V}{0.65} \tag{8-3}$$

where $G_{measured}$ is the measured value of G-Factor. The values of $G_{Measured}$ can be viewed as noisy measurements of the true G-Factor. Accordingly, the calibration procedure uses a Kalman filter to estimate the true G-Factor according to a Markov model with a relatively large time constant. The equations used to implement the Kalman filter are the same as employed in the process of G-Factor estimation addressed in Section 14.4 with regard to traffic state estimation (e.g., in the filtering equations of Section 14.4, $S^n_{j,m}$ is replaced with $G_{Measured}$). The filtering process serves to "smooth" the measured G-Factor value, as illustrated in Figure 8-1, which shows the measured G-Factor along with the resulting filter estimate for an actual detector station from the Twin Cities freeway system.

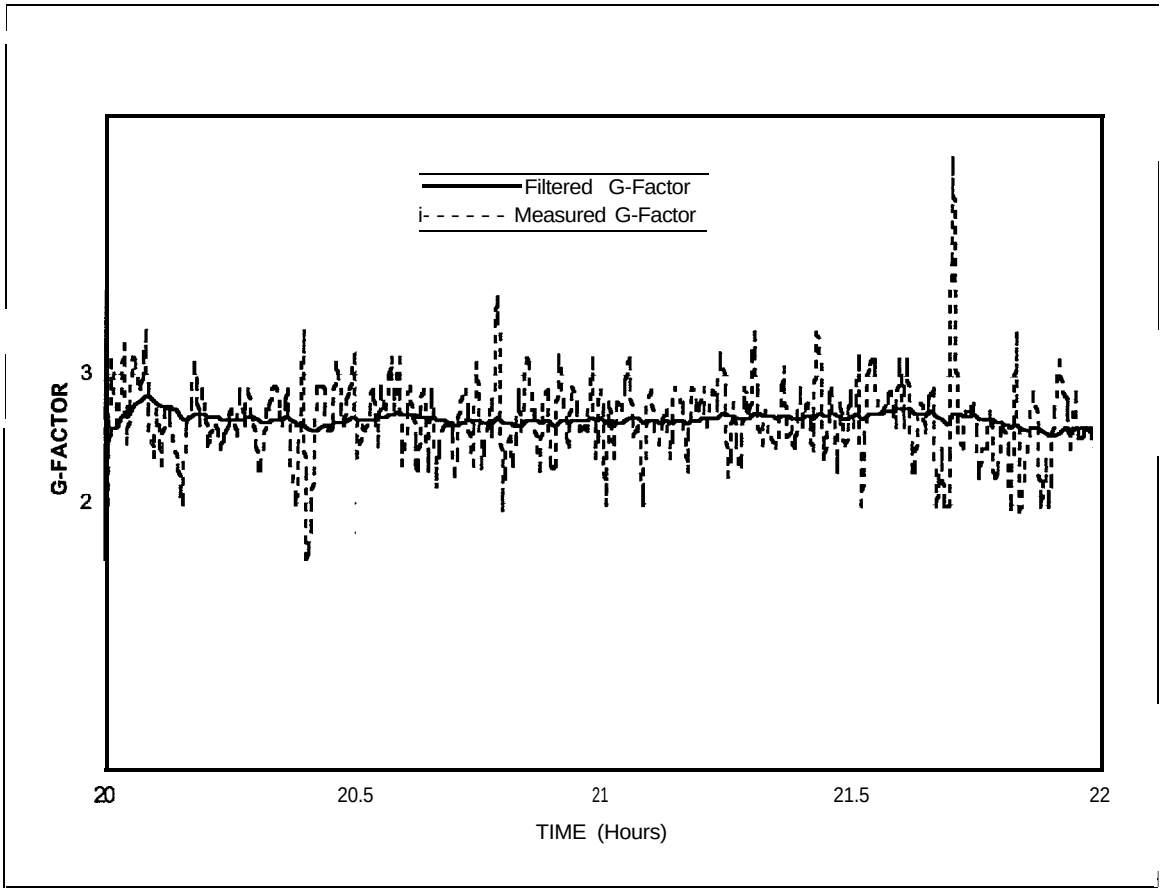


Figure 8-1. Measured and Filtered G-Factors

The calibrated value of G-Factor for any given day is taken as the final value of the filtered estimate for the time period in question. As previously stated, G-Factors are to be calibrated for each station over a period of several days, and the calibrated G-Factors for each day are then compared for consistency.

Stations which exhibit consistent calibration of G-Factors may be deemed acceptable, with the final calibrated G-Factor taken as the average of the G-Factors calibrated over the days in question. Stations where calibrated G-Factors are either inconsistent or where calibration is not possible due to recurring sensor malfunctions are deemed unacceptable. These stations are to be removed from consideration for use by incident detection algorithms. Examples of acceptable and unacceptable G-Factor calibrations are shown in Figure 8.2 for sample detector stations from the Twin Cities freeway system.

EXAMPLES OF SUCCESSFUL CALIBRATION

APR	(2.36	2.48	2.40	2.41	2.37	2.50		)	(2.42)	(-1.65	1.57	-0.54	-0.27	-1.37	2.08			)
OCT	(2.48	2.55	2.41	2.62	2.44	2.54	2.47	2.61)	(2.52)	(-0.92	0.89	-2.83	2.60	-2.00	0.64	-1.18	2.37)	
APR	(2.66	2.75	2.73	2.74	2.72	2.96		)	(2.76)	(-2.44	-0.24	-0.71	-0.47	-0.96	4.39			)
OCT	(2.74	2.72	2.62	2.88	2.65	2.76	2.73	2.75)	(2.73)	(0.21	-0.27	-2.76	3.36	-1.99	0.68	-0.03	0.44)	

EXAMPLES OF UNSUCCESSFUL CALIBRATION

APR	(1.75	x	x	x	x	x		)	(1.75)	(0.00	x	x	x	x	x			)	
OCT	(1.981.79	x	x	x	x	x	x)	(1.88)	(	3.12	-3.45	x	x	x	x	x	x)		
APR	(1.80	x	x	x	x	x		)(1.80)(0.00		x	x	x	x	x				)	
OCT	(1.97	2.21	1.72	x	x	x	x	x)	(1.97)	(0.11	7.16	-9.32	x	x	x	x	x)		
APR	(2.17	2.47	2.10	2.42	2.10	1.77		)	(2.17)	(-0.05	7.85	-2.22	6.67	-2.22	-14.75			)	
OCT	(2.18	2.30	2.26	x	2.15	1.76	x	x)	(2.13)	(1.49	4.80	3.74	x	0.60	-13.66	x	x)		

Figure 8-2. Examples of G-Factor Calibration for the Twin Cities

In the figure, the first set of numbers in parentheses are the calibrated G-Factors for each day. This is followed by the mean G-Factor value across these days. Following this are the variations in computed traffic speed for each G-Factor as compared to the speed associated with the mean G-Factor. For example, the first line shows calibrated G-Factors of 2.36, 2.48, . . . , 2.50 with a mean value of 2.42. The difference in speed that would result from using a G-Factor of 2.36 rather than the mean value of 2.42 is shown as negative 1.65 miles per hour, and so on. G-Factors marked with an "x" indicate that the station was malfunctioning and no calibration was possible for that day.

8.2 Calibration Results for the Twin Cities Site

The measurement calibration process employed for the Twin Cities site utilized data from six weekdays in October 1996 and eight weekday from April 1997 during the evening ours from 8:00 PM to 10:00 PM. As previously stated, this time interval was selected due to the expectation that traffic would be freely flowing and a constant traffic speed may be assumed (65 mph in our case).

For any given day, calibration was conducted for each station in the Twin Cities freeway system, and the calibrated values were compared for consistency. Out of the 621 stations under consideration in the Twin Cities freeway system, G-Factors were calibrated successfully for 534 stations (36%). The remaining stations either exhibited a high degree of variation in the calibrated G-Factor values or exhibited recurring malfunctions, or both.

The mean calibrated G-Factor value for detector stations in the Twin Cities surveillance system was 2.65. The maximum and minimum calibrated values were 3.93 and 1.98, respectively, and the standard deviation was 0.36. The calibrated G-Factors for Twin Cities detector stations are presented in Table 8- 1.

Table 8-1. Calibrated G-Factors for Detector Stations in the Twin Cities

itation	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor
1	3.02	70	2.281	183	2.43	253	3.04	317	2.90	388	2.51
2	2.74	72	2.47	184	2.12	254	2.91	318	2.74	389	2.47
3	2.34	73	2.23	185	3.05	256	2.90	319	2.45	391	2.74
4	2.81	74	2.25	186	2.57	258	2.91	320	2.61	392	2.67
5	2.77	76	2.19	187	3.18	259	2.90	322	2.34	394	2.32
6	2.81	77	2.10	188	3.09	260	2.82	323	2.13	397	2.49
8	3.45	78	2.07	189	3.15	262	2.32	324	2.47	398	2.18
10	3.05	79	2.17	190	3.12	264	2.62	325	2.55	401	2.31
11	2.89	80	2.63	191	2.94	265	3.09	326	2.72	403	2.63
13	2.50	81	2.47	192	3.40	266	2.59	327	2.85	404	2.64
14	2.96	82	2.47	193	2.88	267	2.53	330	2.82	405	2.71
16	2.72	86	2.04	194	2.66	269	2.39	331	3.05	407	3.05
19	3.14	87	2.17	195	3.01	270	2.76	333	2.27	408	2.65
20	3.09	109	2.32	196	3.21	271	2.68	334	2.28	409	2.45
21	2.50	110	2.41	197	2.70	272	2.69	335	3.08	410	2.43
22	2.32	124	3.37	198	2.89	273	2.64	336	2.19	412	2.77
23	2.59	125	2.82	199	2.21	274	2.92	337	2.45	413	3.13
24	2.53	126	2.76	200	3.21	275	2.60	338	2.15	414	2.97
25	2.37	127	2.26	201	3.93	276	2.22	339	2.63	416	2.89
26	2.45	128	2.42	202	3.45	277	2.76	340	3.11	417	3.10
27	2.46	132	2.57	204	2.23	279	2.76	341	2.76	418	2.53
28	2.51	134	2.18	207	2.54	281	2.70	342	2.57	420	2.74
29	2.34	135	2.65	208	2.47	282	2.46	343	2.61	421	2.36
30	2.32	136	2.74	209	2.21	283	2.42	344	2.70	426	2.17
31	2.21	137	2.99	211	2.88	284	2.44	345	2.42	427	2.24
32	3.45	138	3.26	221	2.18	286	3.00	347	2.76	428	2.75
36	2.24	145	2.27	224	2.17	287	2.84	348	2.94	430	2.41
37	2.62	146	3.19	227	2.93	288	2.77	349	2.83	431	2.46
38	2.41	147	2.67	228	2.8 1	289	2.70	3.51	2.53	432	2.23

Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor
39	2.39	148	2.66	229	2.47	290	2.74	352	3.28	433	2.28
40	2.20	151	2.39	230	2.44	291	2.50	355	2.52	434	2.26
41	2.72	152	2.28	231	2.53	292	2.21	356	2.80	435	2.00
42	2.24	153	2.64	232	2.55	294	2.34	358	2.84	437	2.47
46	2.94	154	2.24	233	2.73	295	3.33	359	2.48	438	2.56
47	2.65	156	2.64	234	2.51	296	2.30	361	2.28	439	2.49
48	2.92	159	2.19	235	2.51	297	2.16	363	2.49	441	2.16
49	3.06	163	2.37	236	2.37	298	2.19	366	2.27	442	2.18
50	2.98	165	2.10	238	2.68	299	3.60	368	2.39	443	2.39
57	3.48	166	2.40	239	2.30	301	2.23	369	2.35	446	2.41
58	3.58	167	3.38	240	2.43	302	2.43	375	2.56	447	2.24
60	2.43	168	2.67	241	2.68	304	2.36	376	3.04	448	2.43
61	2.56	170	2.14	242	3.05	305	2.59	378	2.63	450	2.16
62	2.41	171	2.33	243	3.50	306	2.51	379	3.27	451	2.49
63	2.96	172	2.42	244	3.63	308	2.32	380	2.92	452	2.59
64	2.82	175	2.59	246	2.95	310	2.78	381	2.93	453	2.93
65	2.57	176	2.69	248	2.89	311	2.63	382	2.93	455	2.20
66	2.61	177	2.14	249	3.13	312	2.74	383	3.47	456	2.22
67	2.63	179	2.27	250	2.99	313	2.96	384	2.72	457	2.28
68	2.10	180	2.38	251	2.72	314	2.62	386	2.27	459	2.32
69	2.341	181	3.321	252	2.781	315	2.531	387	2.361	460	2.361

Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor
461	2.76	520	2.82	583	3.28	583	3.28	652	2.37	726	2.43
462	2.59	521	2.86	584	2.33	584	2.33	653	2.70	728	2.40
463	2.39	522	2.92	585	2.32	585	2.32	654	2.88	729	2.99
465	2.05	523	3.04	587	3.16	587	3.16	655	2.88	730	2.41
468	3.04	524	3.01	590	2.99	590	2.99	656	2.37	731	2.70
469	2.39	525	3.23	591	3.01	591	3.01	657	2.41	736	2.93
470	2.30	526	3.29	592	3.13	592	3.13	658	2.34	737	2.90
472	2.91	528	2.88	593	2.94	593	2.94	664	2.40	738	2.73
473	2.26	529	2.86	594	3.14	594	3.14	665	2.27	739	2.99
475	2.24	530	2.86	595	2.62	595	2.62	667	2.57	740	2.91
476	2.45	531	2.30	596	3.14	596	3.14	668	2.26	742	3.05
477	2.53	533	2.41	597	2.86	597	2.86	669	2.03	743	2.80
478	2.30	534	2.24	599	3.17	599	3.17	670	2.39	744	3.03
479	2.57	535	2.74	600	3.22	600	3.22	671	2.68	745	3.50
480	2.26	536	2.86	601	2.71	601	2.71	672	2.26	746	2.97
481	2.32	537	2.83	602	2.42	602	2.42	678	2.49	747	2.68
483	2.42	538	3.13	603	2.47	603	2.47	679	2.37	748	2.88
484	2.93	539	2.90	605	2.79	605	2.79	680	2.32	750	2.63
485	2.23	540	3.01	606	2.34	606	2.34	681	2.39	751	3.11
486	2.47	541	2.84	607	3.14	607	3.14	682	2.54	755	2.78
487	2.04	542	2.91	608	2.78	608	2.78	683	2.72	756	2.63
488	2.45	543	2.88	609	2.84	609	2.84	68-k	2.45	757	2.60
489	3.801	545	2.88	610	3.46	610	3.461	685	2.271	759	2.861

Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor
490	2.47	546	2.11	611	2.82	611	2.82	691	2.27	760	2.82
491	2.32	547	2.69	613	2.76	613	2.76	692	2.36	761	3.10
492	2.55	548	2.45	615	2.80	615	2.80	693	2.30	762	2.80
493	2.60	550	2.39	616	3.30	616	3.30	694	2.43	763	3.10
494	2.85	551	2.45	617	3.00	617	3.00	695	2.38	764	3.30
497	2.76	553	2.40	618	3.22	618	3.22	696	2.54	765	2.80
498	3.22	556	2.43	619	2.47	619	2.47	697	2.18	766	2.90
500	2.37	557	2.57	620	3.10	620	3.10	700	2.57	767	2.80
501	2.75	559	2.35	621	2.58	621	2.58	701	2.18	768	2.90
502	2.55	560	3.28	622	3.35	622	3.35	702	2.30	769	2.70
503	3.03	565	2.25	624	2.50	624	2.50	703	2.26	770	3.00
504	2.86	566	2.63	625	3.67	625	3.67	706	2.20		
505	2.70	567	2.35	626	2.90	626	2.90	707	2.51		
506	2.50	568	2.66	627	2.75	627	2.75	708	3.12		
507	2.62	569	2.87	628	2.65	628	2.65	709	2.22		
508	2.57	570	2.83	629	2.84	629	2.84	710	2.46		
509	2.13	571	2.82	630	2.98	630	2.98	711	2.51		
510	2.29	572	2.53	633	3.45	633	3.45	712	2.45		
511	3.47	573	3.50	634	2.55	634	2.55	713	2.38		
512	2.22	574	2.78	635	2.66	635	2.66	718	3.30		
513	2.31	575	3.53	636	2.95	636	2.95	719	2.51		
514	2.33	577	2.97	638	3.23	638	3.23	720	2.42		
515	2.78	578	3.61	639	2.24	639	2.24	721	2.06		
516	2.35	579	2.74	641	3.06	641	3.06	722	2.53		
517	2.19	580	2.65	642	2.36	642	2.36	723	2.15		
518	2.41	581	2.78	643	3.24	643	3.24	724	2.62		
519	1.98	582	2.41	644	2.36	644	2.36	725	2.29		

8.3 Calibration Results for the San Diego Site

Due to time constraints, the loop calibration process used for detector stations in San Diego was not as thorough as that employed for the Twin Cities freeway system. However, a preliminary calibration was conducted for a single day in October 1996 during the hours of 8:00 PM to 10:00 PM. As with the Twin Cities site, a constant traffic speed of 65 mph was assumed. The calibrated G-Factor values are presented here, but care should be taken in interpreting these results as the necessary step of comparing independently calibrated values for several different days was not performed.

The mean calibrated G-Factor value for detector stations in the San Diego surveillance system was 2.22. The maximum and minimum calibrated values were 2.95 and 1.7 1,

respectively, and the standard deviation was 0.21. The calibrated G-Factors for San Diego detector stations are presented in Table 8-2.

Table 8-2. Calibrated G-Factors for Detector Stations in San Diego

Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor	Station	G-Factor
1	1.94	29	2.12	56	2.07	84	1.76	117	2.33	151	2.84
2	2.08	30	2.121	57	2.08	86	2.76	118	2.21	155	2.12
3	1.90	31	2.10	59	2.34	87	1.97	121	2.30	156	2.12
4	2.02	32	2.10	61	2.12	88	2.45	123	2.38	157	2.30
5	2.42	35	2.10	62	2.14	90	1.77	125	2.27	181	2.15
6	2.18	36	2.12	64	2.14	91	2.51	126	2.24	182	2.04
7	2.14	38	1.78	66	2.3 1	92	2.09	127	2.06	184	2.08
8	1.97	39	1.79	67	1.71	93	2.24	128	2.19	185	2.16
10	1.95	40	2.12	68	2.02	94	2.67	130	2.11	186	2.17
11	2.27	41	2.09	69	2.02	96	2.36	132	2.13	187	2.40
12	2.28	42	2.17	70	2.22	97	2.52	134	2.28	188	2.29
13	2.26	43	2.38	71	2.17	101	2.28	136	2.19	223	2.40
14	2.12	44	2.21	73	2.17	102	2.26	137	2.25	241	2.26
15	2.21	46	2.18	74	2.23	105	2.43	138	2.36	244	2.32
16	2.25	47	2.17	75	1.95	107	2.22	140	2.24	276	1.95
17	2.22	48	2.33	77	2.22	108	2.25	141	2.5 1	277	2.29
19	2.23	49	2.01	78	2.41	109	2.36	143	2.31	278	2.12
21	2.23	50	2.21	79	2.21	110	2.29	144	2.56	282	2.12
23	2.24	52	2.08	80	2.95	111	2.28	146	2.31	283	2.48
26	1.95	53	2.10	81	2.66	112	2.28	148	2.76		
27	2.26	54	1.97	82	2.64	114	2.18	149	2.38		
28	2.561	55	2.441	83	2.051	115	2.43	150	2.281		

9.0 QUEUE DETECTION AND TRACKING

A traffic queue is the contiguous region of congested traffic that results when traffic demand exceeds freeway capacity. In many locations, queues develop on a recurrent basis due to heavy traffic volume. Queues also occur upstream of incidents as a result of reduced freeway capacity.

Automatic queue detection and tracking is useful for traffic monitoring purposes and also for determining the extent of an incident's impact on traffic flow. For our purposes, this functionality constitutes an important preliminary step in incident detection processing.

Processing is conducted in two distinct steps. First, a queue identification algorithm is applied to the sensor data at each 30-second reporting interval to arrive at a set of detected traffic queues. These queues are then presented to a queue tracking algorithm, which reconciles the detected queues with previously identified queues in order to distinguish between queue movements and the onset of new congestion. These two processing steps are addressed separately in the subsections that follow.

9.1 Identification of Traffic Queues

This section presents algorithms for identifying traffic queues using measurements from induction loop sensors. The algorithms rely primarily on measurements of traffic speed, which can be computed from the induction loop measurements of volume (vehicles per hour) and occupancy (percentage of time that a vehicle is directly over the sensor). The speed computations make use of a calibrated G-Factor for each individual station. The calibration process is necessary in order to obtain sufficiently accurate traffic speeds, as addressed in Section 8.

The queue identification algorithm operates on freeway zones defined between adjacent detector stations. The objective of the algorithm is to label each zone according to the following four categories of traffic conditions: (1) no queue, (2) head-of-queue, (3) in-queue, and (4) tail-of-queue. A head-of-queue zone represents the location downstream of the queue body where traffic returns to uncongested conditions. Similarly, a tail-of-queue zone represents the location upstream of the queue body where traffic first becomes congested. All zones, if any, lying between the head-of-queue and tail-of-queue zones are categorized as in-queue zones.

The queue identification algorithm utilizes five-minute average traffic speed as the basis for all computations. This quantity is defined for each station in terms of the raw measurements of volume and occupancy as follows:

$$\boldsymbol{Speed_5} = \frac{\boldsymbol{Vol_5}}{\boldsymbol{G \cdot Occ_5}}$$

where

$\boldsymbol{Speed_5}$	=	five-minute average station speed ,
Vol_5	=	five-minute average station volume,
Occ_5	=	five-minute average station occupancy , and
G	=	calibrated station G-Factor .

The queue identification algorithm utilizes three categories of traffic conditions that are defined in terms of three user-adjustable thresholds for five-minute speed: FREE, with a default value of 40 mph; BREAK, with a default value of 32 mph; and EPSILON, with a default value of 3 mph. The threshold FREE is used to identify free-flowing traffic, while BREAK and EPSILON are used to identify congested traffic conditions, also termed traffic ***breakdown.*** This is summarized below:

Traffic Category	Criteria
Free Flow	$\boldsymbol{Speed_5}$ > FREE
Breakdown	$\boldsymbol{Speed_5}$ < BREAK*
Undefined	Otherwise

* - substitute BREAK+EPSILON if congestion previously diagnosed

Speeds greater than FREE constitute free-flowing traffic. Similarly, speeds less than BREAK constitute traffic breakdown. The parameter EPSILON is used to maintain a degree of temporal continuity for congested zones - once breakdown has been diagnosed, the threshold for breakdown becomes BREAK+EPSILON Hence, the thresholds for the onset of congestion and for the termination of congestion are different (by EPSILON mph).

Also, one should note that the "Undefined" traffic category represents a "gray area" between congested and free flow traffic conditions. Zones exhibiting speeds in this region are labeled according to the traffic conditions of neighboring zones, as described in detail below.

The high-level logic of the queue identification algorithm is rather intuitive - identify freeway regions where speeds are consistently low. The algorithm can be summarized as follows: (1) initialize all zones to "No Queue", (2) identify all Head-of-Queue conditions, and (3) for each zone that has been identified as a Head-of-Queue, find the associated Tail-of-Queue, and label all zones in between, if any, as In-Queue.

The last two algorithm steps are described in detail separately below. Implementing these steps requires that a representation of the freeway topology be established. This is accomplished by defining freeway sections that are comprised of one or more zones linked together according to their geometric configuration. For example, a simple stretch of freeway with no branching (e.g., no merges with other freeways) has two freeway sections - one for each direction of travel. Within a freeway section, a constraint is imposed that no zone may be connected to more than one upstream zone or more than one downstream zone. In other words, freeway merges within a section are <u>not allowed</u>. Rather, freeway merges correspond to section boundaries.

<u>Identifying: Head-of-Queue Zones</u>

For each freeway section, beginning with the most upstream zone and moving to the most downstream zone , a zone is labeled Head-of-Queue if:

(1) Upstream traffic is "Breakdown," and

(2) Downstream traffic is not "Breakdown", and one of the following three conditions apply:

   (a) Downstream traffic is "Free Flow".

   (b) Downstream traffic is "Undefined" and remains so farther downstream until "Free Flow" traffic is encountered.

   (c) Downstream traffic is "Undefined" and remains so farther downstream until the end of the segment is encountered.

This is illustrated in Figure 9-1. One should note that special processing is required for the most downstream zone in the section. Specifically, the existence of traffic breakdown in this zone will cause the zone to be labeled temporarily as a "pseudo" Head-of-Queue so that the adjacent upstream zones will be labeled appropriately (this zone will subsequently be labeled correctly as In Queue).

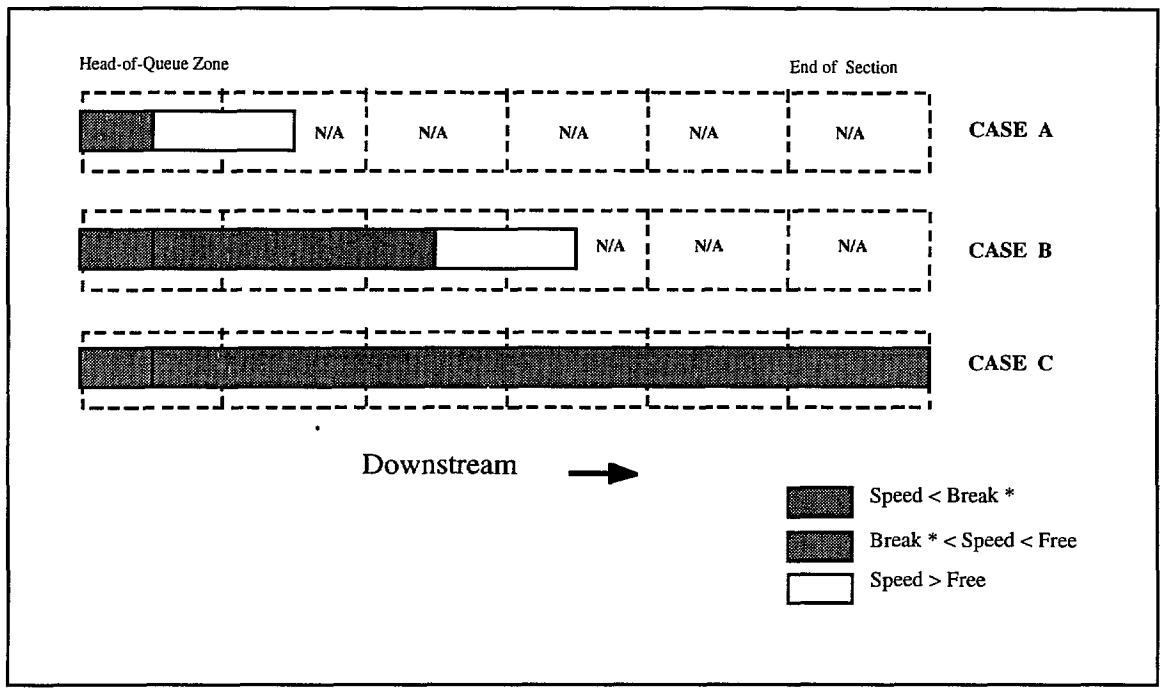


Figure 9-1. Sample Head-of-Queue Scenarios

Identifying Tail-of-Queue Zones and In-Queue Zones

For each zone diagnosed as Head-of-Queue, the algorithm identifies the associated Tail-of-Queue and In-Queue zones as follows:

(1) Moving upstream from the Head-of-Queue zone, look for "Free Flow" traffic.

(2) If not found, the Tail-of-Queue does not exist. Label all upstream zones in the section as In-Queue.

(3) Otherwise the Tail-of-Queue exists - identify it's location. Moving downstream from the zone with "Free Flow" traffic, look for the first instance of traffic "breakdown". This is the Tail-of-Queue zone.

(4) Label all zones between the Head-of-Queue and Tail-of-Queue zones, if any, as In-Queue.

With regard to the "Undefined" traffic category, one should note that upstream traffic breakdown is required to constitute a Head-of-Queue zone and downstream traffic breakdown is required to constitute a Tail-of-Queue zone. This is illustrated Figure 9-2. It should also be noted that all zone types are allowed to exhibit "Undefined" traffic conditions.

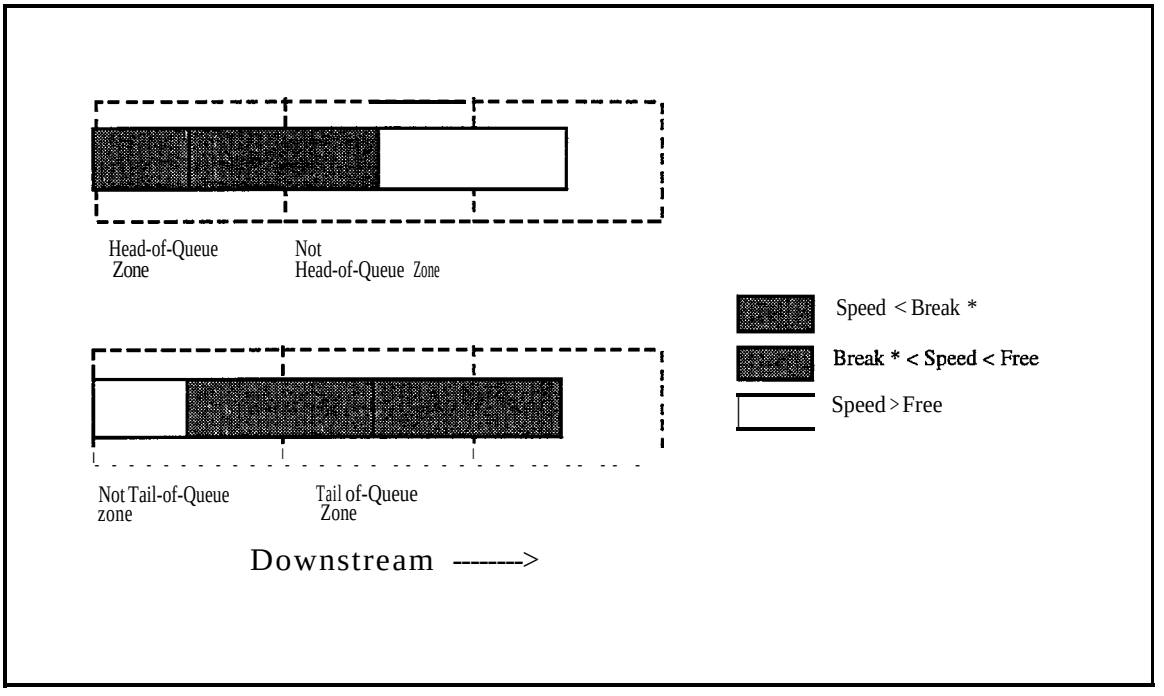


Figure 9-2. Queue Boundary Requirements

9.2 Tracking Queue Movements

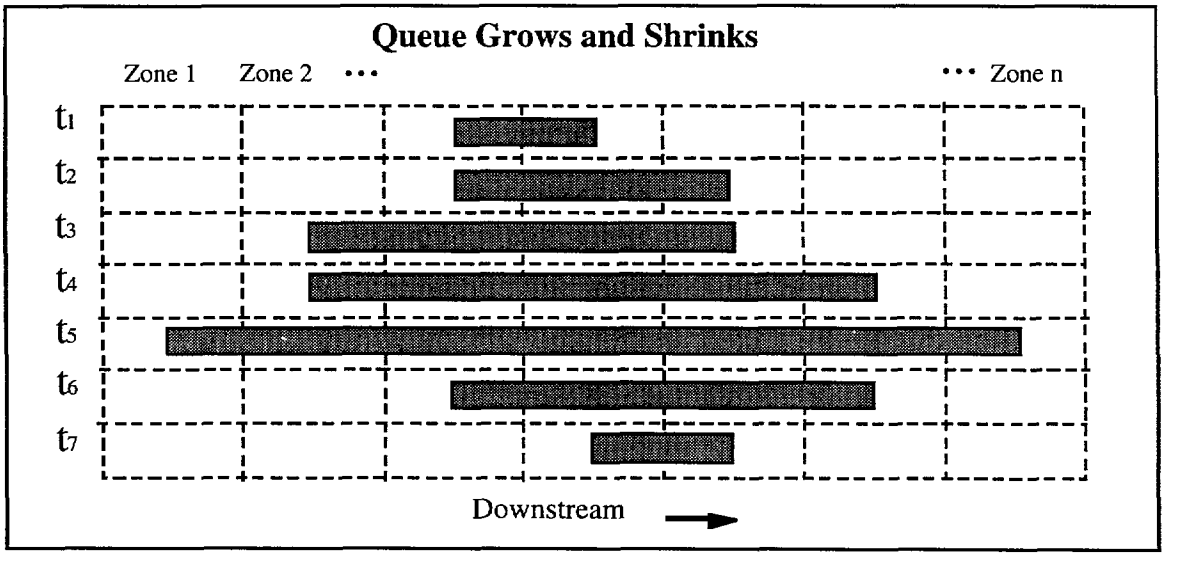
The discussion of the previous section dealt with identifying traffic queues at any given time using freeway surveillance data. This section addresses the issue of tracking queue movements from one time to the next. The intent is to distinguish between the movement of traffic queues and the onset of entirely new congestion.

The algorithm presented herein was developed specifically to be used in conjunction with incident detection algorithms. The purpose of this effort was to develop an algorithm that would disregard in-queue incidents and incidents that do not cause significant upstream

congestion, and to focus instead on detecting incidents which occur in previously free flowing traffic and result in the formation of a traffic queue. The queue tracking algorithm is a direct result of this effort.

As an example, consider a traffic queue that has been identified extending from tail-of-queue zone Z_T to head-of-queue zone Z_H. If at some later time a queue is identified that extends from tail-of-queue zone Z_T to head-of-queue zone Z_{H+1}, which is immediately downstream from zone Z_H, one may conclude that the newly identified queue is the <u>same queue</u> as identified previously, but that the head-of-queue has moved downstream. Furthermore, if the previously identified queue has been subjected to an incident detection algorithm and found to be incident-free, one can argue that there is <u>no need to re-classify the newly detected queue</u>. In fact, the observation that the queue has moved is actually evidence of incident-free conditions, since incidents are stationary events. Similar examples can be presented for congestion which dissipates only to return a short while later. Our experience with the development sites addressed in Section 6 strongly indicates that this type of behavior is very common for freeways subject to recurrent congestion.

Traffic queues are dynamic in three fundamental respects: (1) queues can grow and shrink, (2) queues can split and merge, and (3) queues can move upstream or downstream. This is illustrated in Figure 9-3. The examples of Figure 9-3 are representative of the queuing behavior encountered during the course of our research. Once again, this type of behavior was found to be very common.



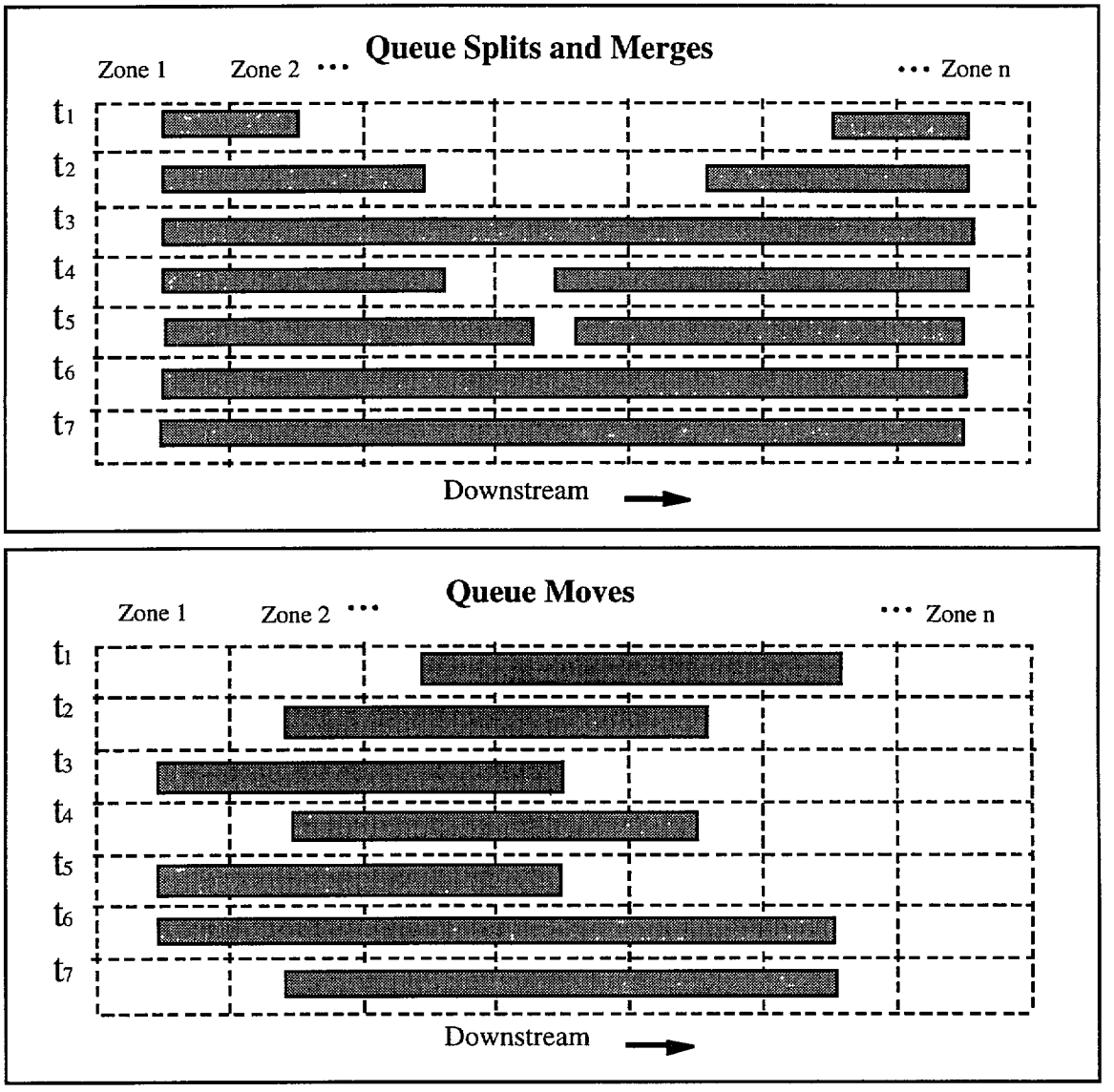


Figure 9-3. Examples of Dynamic Traffic Queues

The purpose of queue tracking is to reconcile detected queues from one time to the next. For instance, we would like the algorithm to recognize the "queue grows and shrinks" example of Figure 9-3 as the same *congestion event*, thereby simplifying the task of incident detection processing. Equivalently, we would like to identify those queues that have not been diagnosed at a previous time - to determine which of the queues identified at the current time step are entirely new congestion events.

The queue tracking algorithm operates by maintaining a list of active congestion events. At each time step, a new congestion event is declared only when a detected queue cannot be reconciled with previously diagnosed congestion.

The queue tracking algorithm has the following qualitative features: (1) queues are allowed to come and go in time (within a specified time limit), (2) queues are allowed to move upstream and downstream (constrained to certain rates of movement), and (3) queues are allowed to merge and split.

Implementation of the algorithm is relatively simple. A list of existing congestion events is maintained, and, at each time step, all identified queues are compared to the list. If an existing congestion event can be associated with a particular queue, the extent of the congestion event may be updated based on the characteristics of the queue. If no congestion event can be associated with a particular queue, a new congestion event is declared. Congestion events are removed from the list only after no queue has been associated with the event for an extended period of time (we used one hour).

The extent of a congestion event is updated only when warranted by the extent of the associated queue. Specifically, if the extent of a queue lies entirely within the extent of the associated congestion event, no update to the extent of the congestion event is made. The reason for this is to prevent a queue with a receding and subsequently advancing head-of-queue from being declared as a new congestion event.

The association scheme entails comparing the extent of the identified queue to the extent of each existing congestion event. If the extents overlap, or if the extents are separated by only a single zone, an association is made between the congestion event and the queue in question. This is illustrated in Figure 9-4. Special processing is required in the case where more than one existing congestion event is associated with a particular queue. In this instance, the congestion events are merged into a single event, as illustrated in Figure 9-5.

The simplification of incident detection processing through use of the queue tracking algorithm is best illustrated through example (this topic is addressed in detail in Section 12.2.5). The queue tracking algorithm was applied to the entire Twin Cities freeway system on 10/l/96 from 6:00 AM to 9:00 AM. During this time period (360 time steps at 30 second intervals), a total of 3959 head-of-queue conditions were identified (approximately 11 per time step) by the queue identification logic described in Section 9.1. The queue tracking algorithm reduced these to only 39 congestion events. Hence, a 100-fold reduction in the number of items to be classified by the incident detection algorithm was achieved.

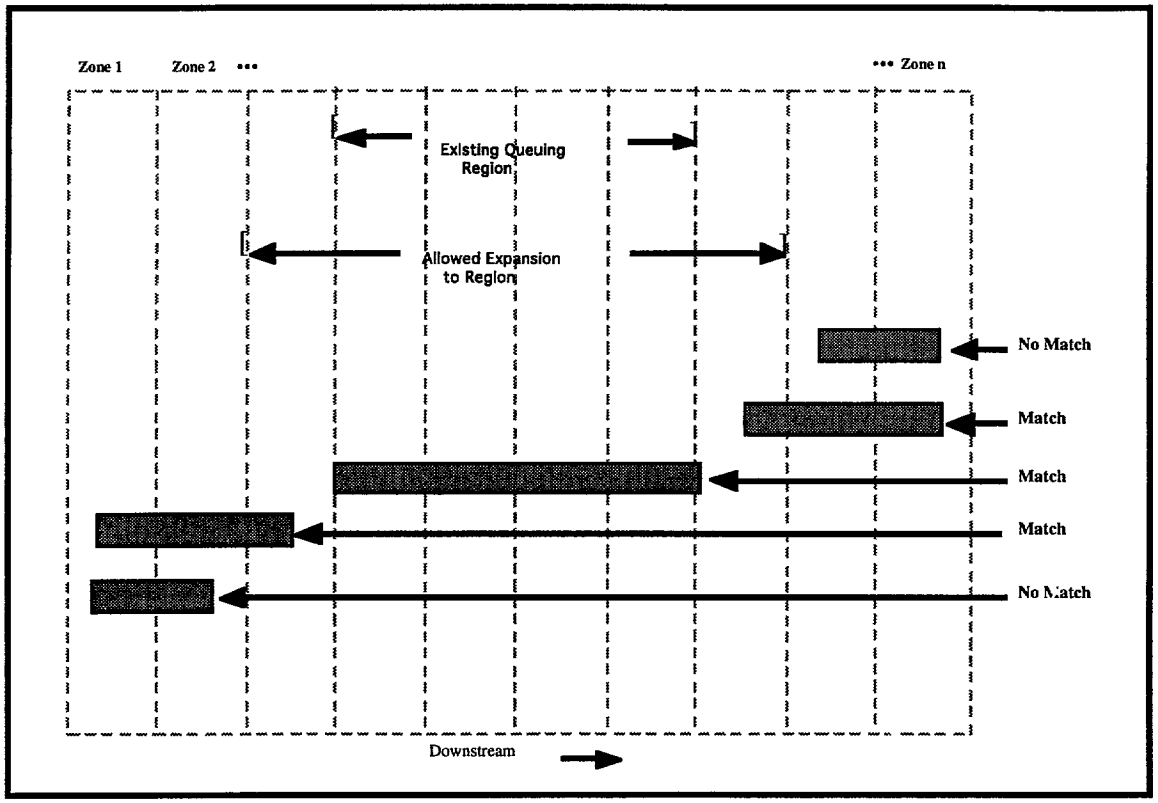


Figure 9-4. Association of Traffic Queues with Congestion Events

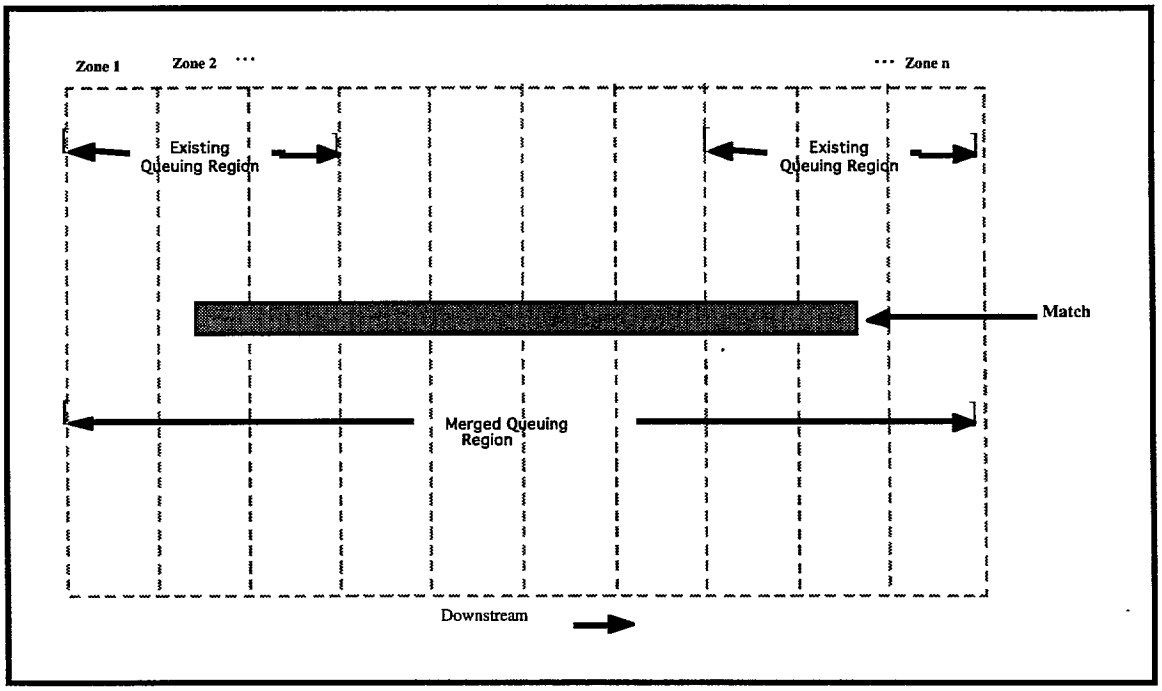


Figure 9-5. Merging Congestion Events

10.0 DATA COLLECTION AND LABELING

The development strategy employed in this project is one of empirical classifier development - collect representative traffic data, then process this data using various software tools in order to generate optimal classifiers. This section addresses relevant issues with regard to the generation of training and evaluation data sets for the development sites identified in Section 6.

10.1 Data Sets for the I-880 Development Site

The first use of actual traffic data in this project involved data from I-880 in the San Francisco Bay Area. For a portion of this freeway, traffic and incident data are available in a readily usable format from a 1993 study investigating the impacts of Freeway Service Patrols (FSP) on traffic delays (refer to Section 6.1 for a complete description of the I-880 development site). With regard to the labeling of traffic conditions, the primary concern is identifying the times and locations of freeway incidents, and information of this type is available from the FSP database.

The portion of I-880 under consideration is a heavily traveled section of freeway, and data is included in the FSP database for weekdays only during peak traffic periods. Hence, there is little to no opportunity to develop algorithms for use in light traffic conditions using this database.

The discussion of this section is divided into two parts. First, the labeling of incident traffic conditions is addressed. This is followed by a discussion of the methods employed to identify and select appropriate incident-free data for use in classifier development and testing.

10.1.1 Identification and Selection of Incident Traffic Data

The I-880 database was used to identify incident traffic scenarios for use in classifier training and evaluation. The relevant issues involved in this effort are addressed below.

Available Information Regarding Incidents

Each incident in the FSP database is defined by an incident "record," which consists of several fields corresponding to relevant incident characteristics. There are approximately 1500 such incidents in the database. Many of these incidents appear to be quite minor and have little to no effect on traffic.

The task at hand was to sort through this large data set and select those incidents which are appropriate for training and evaluation of incident detection algorithms. A software tool entitled "FSP" is available to facilitate this task by selecting incident records from the database according to user-specified incident characteristics (such as incident type).

In the FSP database, incident location data is available in the following form. The freeway was divided into sections, and each section spans an area of roadway between major cross streets. The section boundaries are defined as follows (refer to the schematic of I-880 presented in Figure 6.1):

Section 1 - Marina to Washington/SR238
Section 2 - Washington/SR238 to Hesperian
Section 3 - Hesperian to A-Street
Section 4 - A-Street to Winton
Section 5 - Winton to Jackson/SR92
Section 6 - Jackson/SR92 Tennyson
Section 7 - Tennyson to Industrial
Section 8 - Industrial to Whipple

In the FSP database, incident location data is presented in terms of the freeway section containing the incident, the direction of travel, and the distance relative to the section boundary. For example, the database may indicate that an incident occurred in Section 3 northbound at 8:30 in the morning of 10/l/93 a distance of one-half mile north of A-Street.

Overview of Labeling: Process

For our purposes, incident location must be known relative to the locations of surveillance detector stations. This is necessary so that features can be defined in terms of the traffic measurements of the respective upstream and downstream detector stations.

As can be seen from Figure 6-1, each defined section of I-880 contains several detector stations. In theory, the location of a reported incident in the I-880 database could be identified relative to the upstream and downstream detector stations by making use of the reported distance from the section boundary. To accomplish this, however, the reported distances would need to be sufficiently accurate.

Unfortunately, this was not found to be the case. First, the reported distances are given only in intervals of l/4 mile. Also, it was apparent that the student drivers used to identify the incidents were frequently inaccurate in assessing the location of the incident relative to the section boundaries. Furthermore, the reporting methodology is inherently imprecise. Each cross-street for I-880 is typically served by an on-ramp and an off-ramp on each side of the freeway, and the distance between the ramps is on the order of 100-200 feet. Since the location of the cross street with respect to ramps (or detector stations for that matter) is not specified, the boundaries of the I-880 sections are not well defined.

For these reasons, the specification of incident location available in the I-880 database is not sufficiently accurate for our purposes (in fact, we encountered several instances where the reported location of an incident clearly appeared to be in error). As a result, the incident specifications in the FSP database could not be used directly, and it was necessary to further isolate the exact time and location of the incidents. Consequently, the process of data labeling for I-880 was largely one of verifying the incident reports in the FSP database and enhancing the accuracy of such reports.

Unfortunately, the only source of information at our disposal was the loop data itself. To determine the location of the reported incidents relative to surveillance detector stations, the associated traffic data was inspected near the reported time and location of the incident in question, and incidents were identified when a significant discontinuity was observed in the traffic measurements of the surrounding detector stations.

This was a large and tedious task that necessarily restricted our attention to only those incidents that are relatively severe and created significant disturbances in the traffic stream (e.g., a significant increase in upstream occupancy and decrease in downstream volume). If no discontinuity in the traffic measurements was observed in the neighborhood of a reported incident, the traffic data was discarded and not utilized for algorithm training and

evaluation. As a consequence, there is no opportunity to develop algorithms for detecting such minor incident; using this data set.

Specific Labeling Techniques

For I-880, two distinct methods were employed to identify and select incident traffic data for the purposes of algorithm training and evaluation. The first method was utilized early in the development process and constituted a relatively conservative approach. The objective was to be sure of the labeling. Due to uncertainties involved in identifying precise times of incident onset and termination, no attempt was made to explicitly label these conditions. Rather, only "steady state" incident conditions were labeled by identifying a period of time following the incident onset when congestion at the upstream detector station is no longer increasing and remains approximately constant. This is illustrated in Figure 10-l.

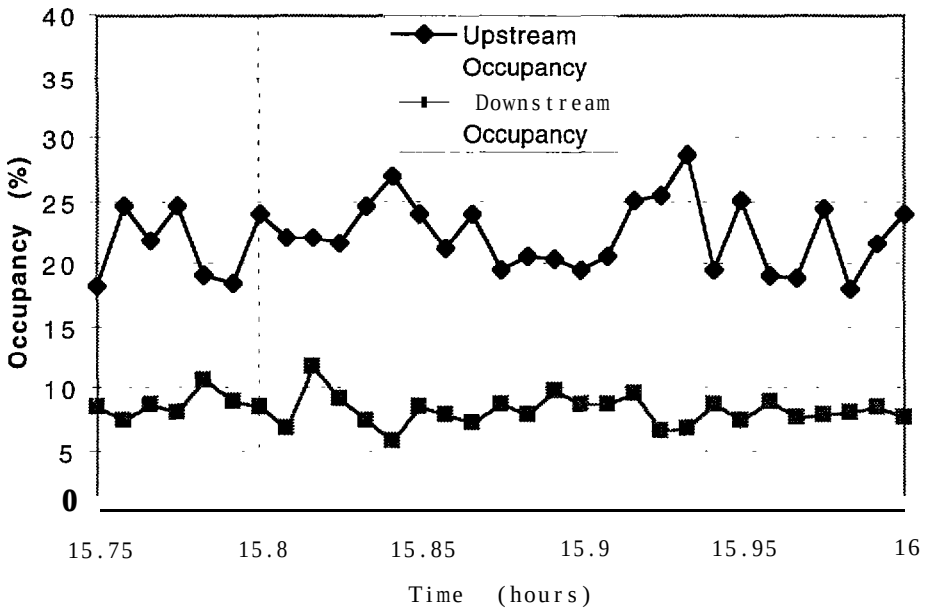


Figure 10-l. Steady State Incident Conditions

This method of labeling was successful in isolating 59 incidents from the I-880 database. For several of the identified incidents, the distance between the nearest functioning upstream and downstream detector stations was large due to the occurrence of sensor malfunctions. Due to this, these incidents were removed from consideration in the training

of incident detection algorithms. Consequently, 39 incidents obtained from the first labeling method were used for development of incident detection algorithms. The definitions of the training data for these incidents is presented in Appendix A.

For the first labeling method, the overriding principle was to utilize only steady state incident conditions in order to be sure of the labeling (e.g., to avoid labeling incident-free traffic conditions near the time of incident onset as incident conditions). The second method employed in labeling I-880 data was much more ambitious and attempted to explicitly identify the periods of incident onset and incident termination. In pursuing this goal, the following model was adopted concerning the various stages related to a freeway incident:

(1) The incident occurs.
(2) The incident begins to affect measurements at one of the adjacent detector stations (usually the downstream station).
(3) The incident begins to affect measurements at the other adjacent detector station (usually the upstream station).
(4) Traffic at the upstream station reaches a “steady-state” condition.
(5) The incident ends.
(6) Traffic begins to disperse at the upstream station.
(7) Traffic dispersal at the upstream station is complete.

Using this model, time periods were identified corresponding to the incident onset, the continuation of the incident, and finally the termination of the incident. These time intervals were determined as follows.

Incident Onset

Due primarily to the length of zones, the time when an incident occurs can be significantly prior to the time at which the existence of the incident becomes apparent in the surveillance data. In order to best estimate the time of incident onset, the following steps were taken:

(1) Identify a time following the incident onset when it is clear that an incident condition exists. This is usually marked by large occupancy differences between the upstream and downstream stations and can be readily identified according to the results of the first labeling method.

(2) Moving backward in time, determine the time at which the incident's existence first affected measurements at the upstream station.

(3) If possible, determine the time at which the incident's existence first affected measurements at the downstream station.

Of the times identified in steps (2) and (3), the earlier time was used to mark the beginning of the incident onset period. Note that it is conceivable, if the incident is in sufficient proximity to the upstream station, that the existence of the incident may affect the upstream station's measurements prior to it affecting the downstream station's measurements. In this case (which was found to be quite rare), or if no discernible effect was found at the downstream station, the time at which the incident's existence first affected measurements from the upstream station was used as the beginning of incident onset period.

In identifying the data times addressed in steps (2) and (3), individual lane data was used in addition to station average data (conversely, only station average data was used in the first labeling method). This data was used due to the observation that **the effects of an incident were frequently more pronounced in the lane data, and these effects were, on occasion, observable prior to any conclusive indications of an incident observed in the station average data.**

Incident onset was said to continue until such time that the measurements from the upstream station reached some type of "steady state" condition (e.g., the resultant queue engulfed the station). One should note that the time between an incident's occurrence and the time that the resultant queue reaches and engulfs the upstream station can last several minutes. The duration of this event clearly depends on the level of traffic, the severity of the incident and physical characteristics of the zone such as it's length, number of lanes and the configuration of on-ramps and off-ramps.

Incident Termination

Incident termination was said to begin at the time when congestion begins to disperse at the upstream detector station (e.g., occupancy begins to fall), and was said to end at the time when upstream traffic once again reached a steady-state condition. As with incident onset, the incident termination period can last for several minutes.

Incident Body

The body of the incident was said comprise all times between the end of incident onset and the beginning of incident termination.

Using the second labeling method, a total of 32 incidents were identified for use in developing incident detection algorithms (exact times for incident onset and termination could not be conclusively identified for certain of the incidents identified by the first labeling method). The definitions of the training data for these incidents is presented in Appendix A.

10.1.2 Identification and Selection of Incident-Free Traffic Data

Selection of the incident-free traffic data to be used in algorithm training is a critical issue in developing effective classifiers. It is important to note that this data is just as important as the selected incident traffic data, since the task of the classifier is to distinguish between the two categories of traffic conditions.

The objective in building learning sets is to present the classifier training mechanism (e.g., CART) with data samples that are representative of all categories of incident and incident-free data to be encountered in practice. Clearly, there is far too much incident-free traffic data available from the I-880 database to utilize all of it, or even a reasonable percentage of it, in training classifiers. Hence there is a need to thin the data.

We explored several categories of incident-free traffic data that were used in classifier training. We also utilized several methods of thinning the data. Various learning sets were constructed using combinations of the data selection and thinning methods addressed in the subsections which follow.

10.1.2.1 Categories of Incident-Free Data

The following categories of incident-free traffic conditions were utilized for the I-880 database. Definitions of the incident-free data used for algorithm development corresponding to these categories is presented in Appendix A.

No Incidents in the Atea

This is the category of incident-free traffic data typically considered - large sections of freeway that are entirely absent of incidents. To identify such data in the I-880 database, we first identified all days in the study where large sections of the freeway had no reported incidents. These days are identified in Table 10- 1.

Table 10-l. Intervals in the I-880 Data Base with Large Sections of Incident-Free Conditions

Spring 1993		Fall 1993	
Date	Peak Period	Date	Peak Period
02/16/93	AM	09/27/93	PM
02/l 9/93	PM	09/30/93	PM
02/22/93	AM	1 0/06/93	AM
02/25/93	AM	1 0/07/93	PM
02/26/93	PM	1 0/08/93	AM
03/0 l/93	AM	10/12/93	AM
03/02/93	AM	10/15/93	AM
03/02/93	PM	10/2 l/93	AM
03/05/93	AM	1 0/22/93	PM
03/08/93	PM		
03/09/93	AM		
03/09/93	PM		
03/16/93	AM		
03/17/93	PM		
03/18/93	AM		
03/19/93	AM		

Data was then extracted from this set corresponding to the time intervals and lengths of freeway where no incidents were reported. In doing so, care was taken to ensure that each zone of the freeway was represented. Additionally, an effort was made to include approximately equal portions of data corresponding to fair and rainy weather, and also of AM and PM shifts for each zone. Again, only long sections of freeway where no incidents were reported over a period of several hours were used for this category of incident-free data.

Incident-Free Traffic in the Neighborhood of Incidents

The incident-free data surrounding incidents should also by utilized in constructing learning sets. In operational circumstances, the classifier will have to work on all the data,

including that which surrounds incidents, both temporally and spatially. To avoid false alarms in the zones adjacent to the incident location, the incident-free data of these zones was included in classifier training sets during times when the incident condition existed. In addition, the incident-free data both preceding and following the incident condition was also utilized for both the incident zone and also the adjacent upstream and downstream zones.

Congested Incident-Free Traffic Conditions

One of the primary causes of false alarms in incident detection algorithms is the onset of recurrent congestion, particularly with regard to head-of-queue traffic conditions. To account for this, the classifier training set should include examples of this category of incident-free traffic data.

To identify these traffic conditions, the measurement data from the I-880 database was inspected corresponding to times and locations where no incidents were reported. When congested traffic conditions were observed (e.g., head of queue), or when complex traffic behavior was noted (e.g., significant fluctuations in occupancy and volume readings), the data was extracted for use in constructing learning sets. During this process, it was noted that certain zones are much more susceptible to this type of traffic behavior than other zones.

10.1.2.2 Data Thinning Techniques

As previously stated, there was a need to reduce the size of identified incident-free data due to the vast quantities of this data available from the I-880 database. The method by which this is accomplished is very important, since developed learning sets should be representative of all types of data to be classified by the algorithm, and unwarranted redundancy should be avoided. The following techniques were employed to thin the data. Various combinations of these techniques were employed in constructing training data sets for this project

Uniform Data Thinning

Two distinct approaches were taken for uniform thinning of incident-free data. In the first approach, specific time intervals were identified to be included in the learning sets. For

example, one method was to select two minutes out of every five-minute interval for inclusion (e.g., select 7:00-7:02,7:05-7:07,7: 10-7: 12, etc.).

In the second method, the 30-second data records were thinned by simply selecting every "Nth" data record, where the value of N varied (e.g., one method was to select every 5th data record for inclusion in the learning set).

Data Filtering

In this approach, the objective is to reduce the high degree of redundancy common in incident-free data by removing data records that clearly correspond to incident-free traffic conditions. For example, there is no need to present to the classifier training mechanism an excessive number of data records where congestion does not exist at the upstream or downstream detector stations (recall that only incidents with upstream congestion were identified in the labeling process).

It was our experience that inclusion of such "easy incident-free" data in the learning sets was, in fact, counter-productive, as the resulting classifiers tended to focus on the characteristics of this data rather addressing the discernment of recurrent congestion from incident-induced congestion. We therefore concluded that "easy incident-free" data should be removed from the learning set. In operation, this data can be identified as incident-free prior to application of the classifier.

To pursue this approach, we identified certain criteria which, when met, indicate that the data is incident-free with a high degree of accuracy (with regard to the types of incidents identified in the labeling process). Specifically, it was found that a high percentage of the incident-free data satisfied the following criteria and that essentially none of the identified incidents exhibited these traffic conditions. The criteria is as follows:

$$((OCCSDFl < 3) \;\&\&\; (OCCUPl < 12))$$

The variable OCCUPl is one-minute average station occupancy for the upstream station, and OCCSDFl is the "spatial difference" of one-minute station occupancy between the upstream and downstream stations (e.g., OCCSDFl = OCCUPl - OCCDNl). These and all other features used in this study are defined in Appendix E.

The filtering criteria basically states that if there is no appreciable upstream congestion (OCCUPl < 12), and there is no appreciable reduction in downstream occupancy (OCCSDFl < 3), then the data is almost certainly incident-free and therefore doesn't need to be represented in classifier training sets.

One should note that this is only one such filtering criteria that can be used in this regard, and analogous criteria may be explored for potential performance benefits.

10.2 Data Sets for the Twin Cities Development Site

Subsequent to our efforts involving the I-880 development site, we constructed a capability for acquiring real traffic data and attaching labels of incident or incident-free conditions, and successfully used it with the Twin Cities freeway system to obtain a very substantial body of useful data for both training and evaluation. The scope of this data acquisition effort was the entire Twin Cities freeway system.

The first phase of this effort involved remote labeling of the traffic data, such as employed for I-880. After developing various algorithms using this data set, we concluded that a more comprehensive data set was required. To support this goal, we visited the Minneapolis TMC and collected and labeled traffic data using an available closed-circuit television system to monitor the traffic conditions.

10.2.1 Remote Data Labeling

The preliminary efforts in labeling incident data for the Twin Cities development site involved inspection of surveillance data for reported incident locations. The incident reports were obtained from the Minneapolis Department of Transportation, and the surveillance data was inspected remotely (in San Diego) using the data collection capability of RIDE [BOAZ 97].

The labeling methods used in this regard are the same as those described for the first method utilized in the I-880 development site. A total of 37 incidents were identified for algorithm training and evaluations.

The algorithms developed from this data are designed to be used in conjunction with the queue identification algorithm presented in Section 9. Classification of the traffic category

is attempted by these algorithms only at the time that the associated head of queue is detected. As a result, no attempt was made to explicitly label the time of incident onset. Rather, the effective onset time was taken to be the time of detection of the associated head of queue.

The definitions of the training intervals used for algorithm development are presented in Appendix B. The times given represent the time that the associated head of queue was detected by the queue identification algorithm. Hence, the time of incident onset precedes the times listed in Appendix A due to the time which elapsed during formation of the queue.

10.2.2 On-site Data Labeling

After exploring algorithm development using the incident data obtained and labeled remotely, we concluded that a more comprehensive data set was needed. To construct such a set, visits to the Minneapolis TMC were made in order to utilize existing closed-circuit television in labeling data corresponding to head-of-queue traffic conditions only.

The queue identification and tracking logic described in Section 9 was incorporated into RIDE and applied to the Twin Cities freeway system over a period of two weeks. The primary purpose of this effort was to generate a labeled data set to be used in developing and testing of incident detection algorithms.

In order to facilitate generation of the data set, support software was written to cue the user for every new congestion event identified by the tracking algorithm. This was accomplished through use of Ball's MetroView data visualization capability, which displayed new events (shown as white boxes with a "?") on a map of the region, as shown in Figure 10-2.

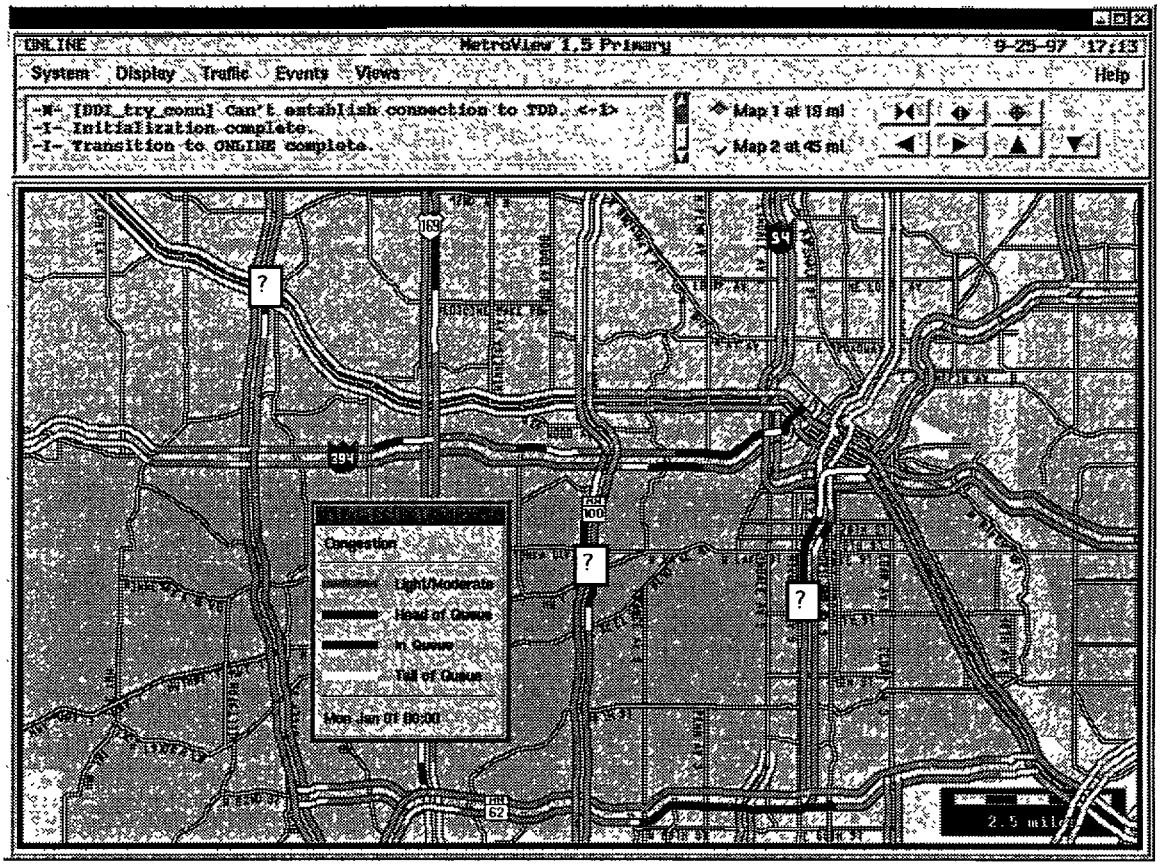


Figure 10-2. Display Capability Used in Labeling Traffic Conditions

For each congestion event, the operator was allowed to examine a CCTV display for the freeway section of interest and to label the event as either incident, incident-free or unknown. The unknown category was necessary since not all congestion events could be conclusively labeled as either incident or incident-free. The objective was to label an event as incident or incident-free only when the operator was relatively certain of the labeling.

Once an event was labeled, MetroView recorded the relevant data for subsequent analysis. This data consisted primarily of the time and location of the event and the associated labeling. A sample of the recorded information is presented in Figure 10-3. The figure shows the time of the event and name of the freeway on which it occurred, followed by the defmition of the associated queue in square brackets. Station identifiers shown in parentheses indicate that the station was malfunctioning.

/96 Incident Events

7:11 { I94-I EB} [ST246 <-- ST135] : real bad incident...bus, couple of cars crashed
15:03 { I35E-I NB} [ST624 <-- (ST623)<-- ST622] : cage in middle of fwy...essentially blocking 1 lane out of 3 lanes
15:24 { I35W-I NB} [ST667 <-- (ST666)<-- ST665] : major incident...traffic moving by it VERY slowly...ambulance
15:48 { I35W-D SB} [ST2 <-- (ST586)<-- ST585] : emergency vehicle with lights flashing on LHS
16:02 { I35W-I NB} [ST670 <-- ST669] : overturned vehicle on middle divided...Hwy Helper there
16:21 { I394-D WB} [ST338 <-- ST337] : Cop pulled someone over.
17:01 { I94-D WB} [ST553 <-- (ST552)<-- ST551] : van with hood up & flashers on...stalled

5/14/96 Incident Free Events

6:56 { MN62-D WB} [ST127 <-- (ST75) <-- ST335]
6:56 { I35E-D SB} [ST639 <-- ST638]
7:02 { I494-I WB} [ST186 <-- ST185]
7:09 { MN36-D WB} [ST616 <-- ST615]
7:17 { I494-D EB} [ST194 <-- ST193]
7:20 { I35W-I NB} [ST77 <-- (ST35) <-- (ST33) <-- ST32]
7:34 { TH100-D SB} [ST410 <-- ST409]
7:36 { I494-I WB} [ST189 <-- ST188]
7:44 { I35W-D SB} [ST578 <-- ST577]
7:48 { TH100-I NB} [ST388 <-- (ST387) <-- ST386]
7:49 { I394-D WB} [ST337 <-- ST336]
7:53 { I494-D SB} [ST701 <-- ST700]
15:28 { I494-D EB} [ST194 <-- ST193]
15:32 { I94-D WB} [ST230 <-- ST232]
15:58 { I35W-I NB} [ST657 <-- ST656]
16:41 { MN36-I EB} [ST592 <-- ST591]
16:49 { I94-I EB} [ST110 <-- ST109]
16:50 { I94-I EB} [ST468 <-- (ST467) <-- (ST466) <-- ST465]
16:53 { I694-D WB} [ST176 <-- (ST178) <-- ST180]
17:16 { I494-I WB} [ST515 <-- ST514]

Figure 10-3 Recorded Event Information

Characteristics of the Data Set

The data set obtained from this effort is summarized in Table 10-2. The labeled data set, along with surveillance data for the entire Twin Cities freeway system over the interval of collection, is available in electronic format and was delivered on a CD-ROM along with this report.

No incident-free events were obtained on 4/21/97 (noted with an asterisk in Table 10-2) as the first days efforts focused on setting up the software and refining the logistics of the labeling processes. However, since three large incidents were observed, these were included in the data set.

Table 10-2. Labeled Events from the Twin Cities Freeway System

	Number of Events	
Date	Incident	Incident-Fret
4/2 l/97	3	0”
4/22/97	14	40
4/23/97	6	37
4/24/97	11	33
4/25/97 (AM)	2	19
5/12/97	5	21
5/l 3/97	12	35
5/14/97	7	27
5/15/97	2	18
5/16/97 (AM)	2	9
5/l 2/97	4	17
5/l 3/97	11	32
5/14/97	6	20
5/l 5/97	2	5
Total	64	239

One should note that the data set is comprehensive in geographic scope. The entire Twin Cities freeway system, consisting of over 300 miles of freeway, was used for data collection during both the AM and the PM peak traffic periods. By encompassing a wide range of freeway geometry, the comprehensive scope of the data set ensures that real operational issues are addressed.

From Table 10-2, one *can* estimate the *a priori* probabilities of incident and incident-free congestion events in the Twin Cities freeway system as 0.2 and 0.8, respectively. Hence, development of an operationally useful incident detection algorithm will require development of a highly accurate classification scheme. Refer to Section 11.2 for a description of the incident detection algorithms developed from this data set. Corresponding assessments of algorithm performance are addressed in Section 12.2.

10.2.3 Assembling the Data Sets

The labeled traffic data obtained from the Minneapolis data collection effort was partitioned into two distinct data sets for use with incident detection algorithms - one set for algorithm development and one set for algorithm evaluations. The definitions of these data sets are presented in Table 10-3.

Table 10-3. Learning and Evaluation Sets for the Twin Cities Site

Evaluation Set	Number of Events	
Date	Incident	Incident-Free
5/12/97	4	17
5/13/97	11	32
5/14/97	6	20
5/15/97	2	5
5/16/97 (AM)	1	6
10/l/96	4	0
1 O/2/96	4	0
1 O/3/96	4	0
10/4/96	1	0
10130196	2	0
10/3 l/96	4	0
1 l/l/96	1	0
Total	44	80

Evaluation Set	Number of Events	
Date	Incident	Incident-Fre
4/2 l/97	3	0
4/22/97	10	16
4/23/97	2	18
4/24/97	5	21
4/25/97	1	13
Total	21	68

The data set definitions of Table 10-3 differ somewhat from those presented in Table 10-2. The reasons for this are two-fold: first, of the identified congestion events, certain events were deemed to be inappropriate for training and evaluation; and second, several incidents from the Twin Cities freeway system that were identified remotely, as discussed in Section 10.2.1, were included in the learning set in order to make it as rich as possible with respect to incident events. Examples of congestion events that were deemed inappropriate for training include zones which were exceedingly long due to the occurrence of sensor malfunctions, and incident-related queues that developed on the opposite side of the freeway from the actual incident location.

11.0 DEVELOPMENT OF OPERATIONAL INCIDENT DETECTION ALGORITHMS

This section addresses the development of algorithms for the purposes of incident detection. Several algorithms were devised using data from the development sites identified in Section 6. No algorithms were developed using data from the San Diego site. Performance evaluations for the algorithms addressed in this section are presented in Section 11.

11.1 Algorithm Development Using Data from the I-880 Site

The algorithms presented in this section were developed using the data sets addressed in Section 10.1. The use of I-880 data constituted our preliminary efforts in developing incident detection algorithms. As such this was largely a learning experience, and resulted in the formation of several important ideas that were successfully applied to the Twin Cities development site.

It should be noted that one of the principle results obtained through use of data from the I-880 development site was the development of algorithms for detecting sensor malfunctions. This is addressed in detail in Section 7.1.

11.1.1 Preliminary Data Analysis

The first work conducted for I-880 involved statistical analysis of the incident and incident-free data sets obtained from the data labeling process described in Section 10. Several very useful insights into the nature of the data were obtained from this effort, as addressed below.

<u>Characteristics of I-880 Surveillance Data</u>

The first observation made was that variations associated with the traffic measurement data for I-880 were significantly more extreme than expected. It is clear that the number of vehicles that pass a particular detector sensor in a 30-second period (e.g., volume) can vary greatly from one time step to another. As the measured value of occupancy is computed by summing the occupancies associated with individual vehicles, this quantity can also vary

greatly from one time step to the next. A typical incident-free plot of single-lane occupancy as a function of time is shown in Figure 1 l-l for congested traffic conditions.

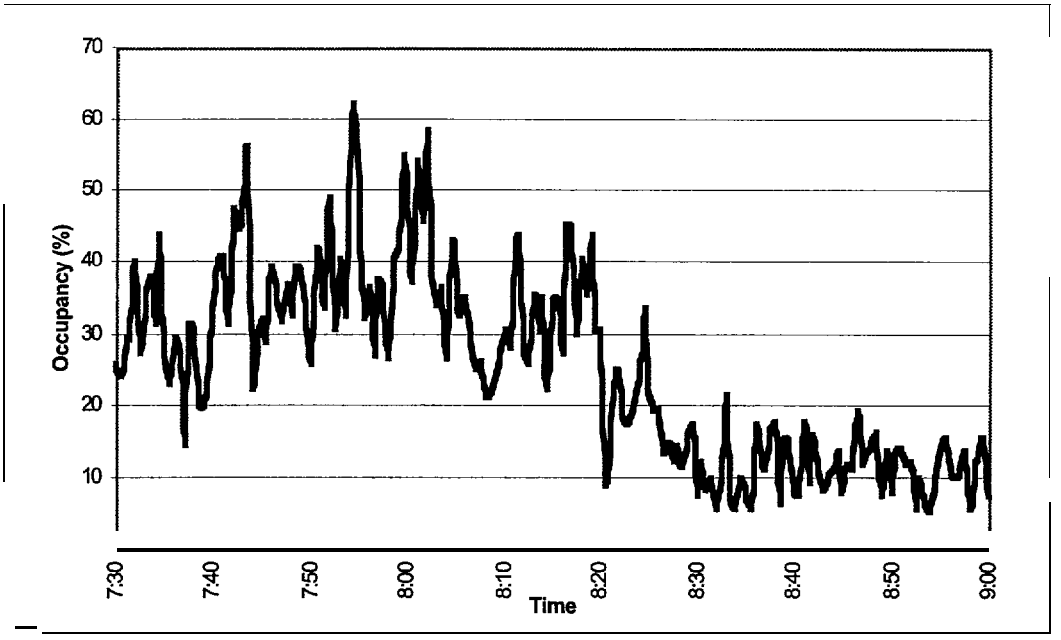


Figure 11-l. Single Lane Occupancy During Congestion

This data was not selected for presentation because of it's extreme nature. On the contrary, this type of behavior is representative of I-880 traffic data in general (traffic data from the other development sites also exhibited similar characteristics).

While traffic measurements were found to be somewhat less variable during uncongested traffic conditions, data characteristics during periods of congestion are of primary concern in developing incident detection algorithms. The reasons for this are three-fold: (1) incident detection algorithms are to be deployed in urban areas where congestion is common, (2) most incidents occur during peak traffic periods, and (3) congested traffic conditions pose the greatest problems for incident detection algorithms and therefore should be the focus of development. Furthermore, with regard to I-880 database, there is very little uncongested traffic data available since this is a heavily-traveled freeway and data is available only for the AM and PM peak traffic periods.

Note that the measurement data shown in Figure 11- 1 corresponds to a period of decreasing congestion (e.g., decreasing occupancy), and there are some rather large jumps. The reverse is also very common - the onset of recurrent congestion can be rapid and

extreme. Specifically, large positive jumps in occupancy are quite common in the incident free traffic examined for I-880.

The existence of such extreme traffic variations in incident-free traffic poses a significant challenge with regard to rapid and reliable incident detection, since the primary symptom of a lane blockage is a marked increase in upstream occupancy. Data averaging and/or filtering may be employed in an effort to reduce false alarms; however, detection times are generally sacrificed in adopting such an approach. The tradeoff between detection time and false alarm rate is fundamental issue in developing incident detection algorithms.

Recurrent Head-of-Queue Traffic Conditions

Another significant problem for incident detection algorithms involves head-of-queue traffic conditions. A head of queue is defined as the freeway region downstream of the queue body where congestion begins to dissipate. As the traffic measurements produced for such an area greatly resemble incident-induced traffic conditions (e.g., high upstream occupancy and low downstream occupancy), recurrent head-of-queue conditions need to be specifically addressed in developing incident detection algorithms.

As previously noted, heavy congestion regularly occurs at several locations along I-880 in both the morning and evening peak periods. In addition, head-of-queue traffic conditions are frequently encountered. This is particularly true near the SR92 junction and at both ends of the freeway segment studied, where the number of lanes is reduced (five lanes to three lanes traveling northbound, five lanes to four lanes to three lanes traveling southbound).

Upon examining the queuing characteristics of I-880, it became apparent that such effects could not be neglected in devising incident detection algorithms. In fact, explicit processing was devised in this regard, and various methods were employed in an attempt to statistically capture the recurrence of congestion.

Our objective was to devise algorithms to identify times and locations where recurrent queuing existed, and to remove the data corresponding to recurrent head of queue conditions from consideration by the incident detection algorithms. In doing so, we deferred the issue of detecting in-queue incidents to algorithms designed specifically for this purpose.

The development effort in this regard used data from the I-880 development site to devise the initial algorithms, and these algorithms were subsequently tested and refined using data from the Twin Cities development site. The algorithms focused entirely on statistical methods and utilized historical observations of traffic speeds for each I-880 zone. Statistics were collected for upstream speed, downstream speed and speed difference during an extended period of incident-free traffic. All speed values refer to five-minute average station speed. As the speed characteristics of zones naturally vary by time of day, the statistics were also kept by time of day using ten-minute time periods.

Qualitatively, the following question was addressed for each encountered head of queue: "Based on incident-free traffic observations near the current time of day, what is the probability of significant speed increase in this zone given that upstream congestion exists?" This quantity was termed the "historical probability" (HP) of speed increase. Quantitatively, we have:

$$HP = P[Speed_{Down} > Speed_{FreeFlow} / Speed\ Up < Speed_{Congestion}]$$

where

$Speed_{Down}$	=	speed at the downstream station
$Speed_{Up}$	=	speed at the upstream station
$Speed_{FreeFlow}$	=	free-flow speed (we used 40 mph (68 kph)), and
$Speed_{Congestion}$	=	congestion speed (we used 32 mph (50 kph)).

This probability was estimated for each I-880 zone as a function of time using a very large data set of incident-free traffic conditions. We also generated similar probability estimates with data grouped according to values of a segment-wide average occupancy (rather than time), in an attempt to capture the dependence of the development of the head-of-queue on the general level of traffic.

Unfortunately, we concluded that the recurrence of congestion is not sufficiently regular for a purely statistical approach to be effective in discriminating recurrent congestion. To account for the observed variability, a more sophisticated approach is required wherein traffic demand and freeway capacity are explicitly modeled for freeway sections of interest. Such an approach may be pursued through the use of dynamic traffic models, as described in Section 14.

Calibration of Loop Measurements

In the analysis of traffic speed data for I-880, it became apparent that measured speed values varied considerably from one station to the next during periods of uncongested traffic conditions. Furthermore, these speed differences were frequently found to be quite consistent from one day to the next.

Two conclusions can be reached from these observations: (1) traffic speeds actually differ on a consistent basis for detector station locations, or (2) a measurement bias exists among detector stations. The first conclusion may be supportable for stations corresponding to significant geometric differences, such as steep grades or curves. However, as the observed differences were so widespread, we concluded that the second possibility is actually the case.

Hence, to utilize speed data effectively, loop measurement data needs to be calibrated on a station-specific basis. This was done for the Twin Cities and San Diego development sites, as addressed in Section 8. The calibration results serve to reinforce our assertion that the observed speed differences are due to measurement bias.

11.1.2 Construction of Incident Detection Algorithms

The classifier construction techniques used with regard to data from the I-880 development site involved a software tool entitled CART (Classification and Regression Trees), which constructs "optimal" binary decision trees for a particular training data set. The use of binary decision trees allows one to explicitly address characteristics of the resulting classifiers (as opposed to neural network algorithms, which offer very little insight into the nature of the problem or the resulting algorithms). Due to this characteristic, binary decision trees were utilized exclusively in the algorithm development efforts for I-880.

The Use of CART in Developing Binary Decision Tree Classifiers

The CART software applies statistical pattern recognition techniques to the learning set to empirically deduce an optimal decision tree as described below. The methodology is presented in very general terms, and interested parties should refer to [BREI 84] for a complete technical description.

A binary decision tree consists of a set of "decision rules" whereby data is partitioned at each node of the tree. The decision rule of the top node in the tree can be viewed as dividing the space of all possible data samples into two distinct sub-spaces. One can similarly envision further partitioning of the sample space by the decision rules of the lower nodes in the tree.

Starting at the top node of the tree, the CART software uses the learning set to systematically select decision rules. The selection criteria is that the "purity" of the data partition should be maximized. In this sense, purity is a measure of the disparity in the resulting class distributions of the created sub-spaces. To illustrate this process, consider the single-node example of Figure 1 l-2.

The learning set for this example consists of 100 labeled samples, 68 of which belong to class A and 32 of which belong to class B. Decision rule #1 partitions the data set so that 62 of the 68 class A samples are associated with the true branch, while only six class A samples as associated with the false branch. Similarly, only two class B samples are associated with the true branch, while 30 class B samples are associated with the false branch.

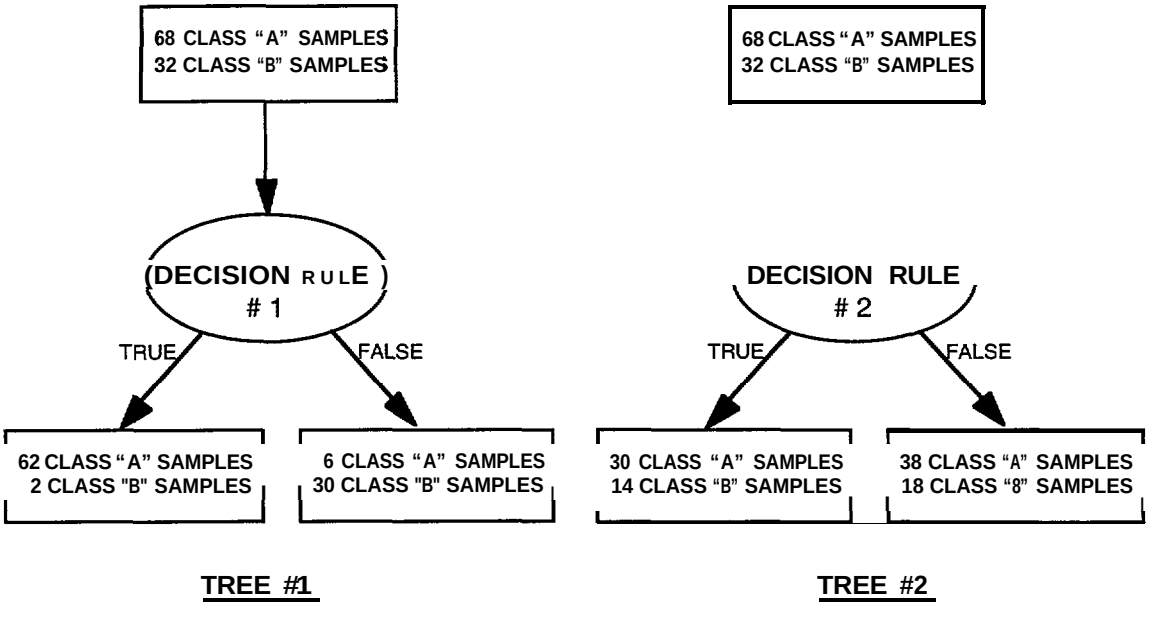


Figure 11-2. Examples of Binary Decision Rules

Since the objective is to accurately classify data samples, it should be clear that decision rule #l results in a more favorable partition of the data than decision rule #2. In other words, if a sample in the learning set were of unknown class and satisfied the true branch of decision rule #l, one could estimate with a fair degree of confidence that the sample belongs to class A. Conversely, class estimation for the same sample based on decision rule #2 would be much more ambiguous. One should realize, however, that we are not interested in classifying data samples from the learning set. Our intent is to apply this type of result to an arbitrary data sample of unknown class. This requires careful construction of the learning set, as described in Section 4.

The CART documentation describes in detail the criteria which may be employed to evaluate a given decision rule's "goodness of split;" however, such descriptions are beyond the scope of this document. Suffice to say that for a problem of fixed dimensionality (such as ours), the set of possible decision rules is finite, and a decision rule for each node of the tree can therefore be selected systematically by observing and evaluating partitions of the learning set.

Also discussed in detail in the CART documentation, but only noted here, is the issue of when to stop partitioning the data and designate a node as a "terminal node" in the tree. The method employed by CART is to "grow" a very large tree that results in a near perfect partitioning of the data and to subsequently "prune" back terminal nodes based on various theoretical considerations. Once again, the interested reader should refer to [BREI 84].

CART software supports assignment of relative costs to classification errors, giving rise to tree construction methods that minimize the overall misclassification cost incurred by the tree. In this manner, for instance, one may specify that classifying incident-free data as incident (e.g., a false alarm) is more costly that classifying incident data as incident-free (e.g., a missed detection).

Another important aspect of CART is that it supports use of a *priori* class probabilities to account for class distributions that are misrepresented in the learning set. For example, if 70% of the samples in the learning set belong to class A, but an arbitrary sample in the real world only has a 3% chance of belonging to class A, a classifier trained on the learning set may place undue emphasis on classifying type A data samples. A more appropriate classifier can be constructed by accounting for the knowledge that the a priori class probability for type A samples is only 3%.

Developed Incident Detection Algorithms

Several incident detection algorithms were developed using CART and the learning sets generated from the I-880 database. The various algorithms were developed using several combinations of learning data (identified in Section 10) and development methodologies, specifically with regard to the training options available for use in the CART software (e.g., priors, costs, etc.).

A general characteristic of all developed algorithms is that the resulting trees were rather large, resulting in a large number of terminal nodes. Furthermore, the branching decisions were generally not intuitive. A representative tree is shown in Figure 1 l-3 (All feature definitions are presented in Appendix E).

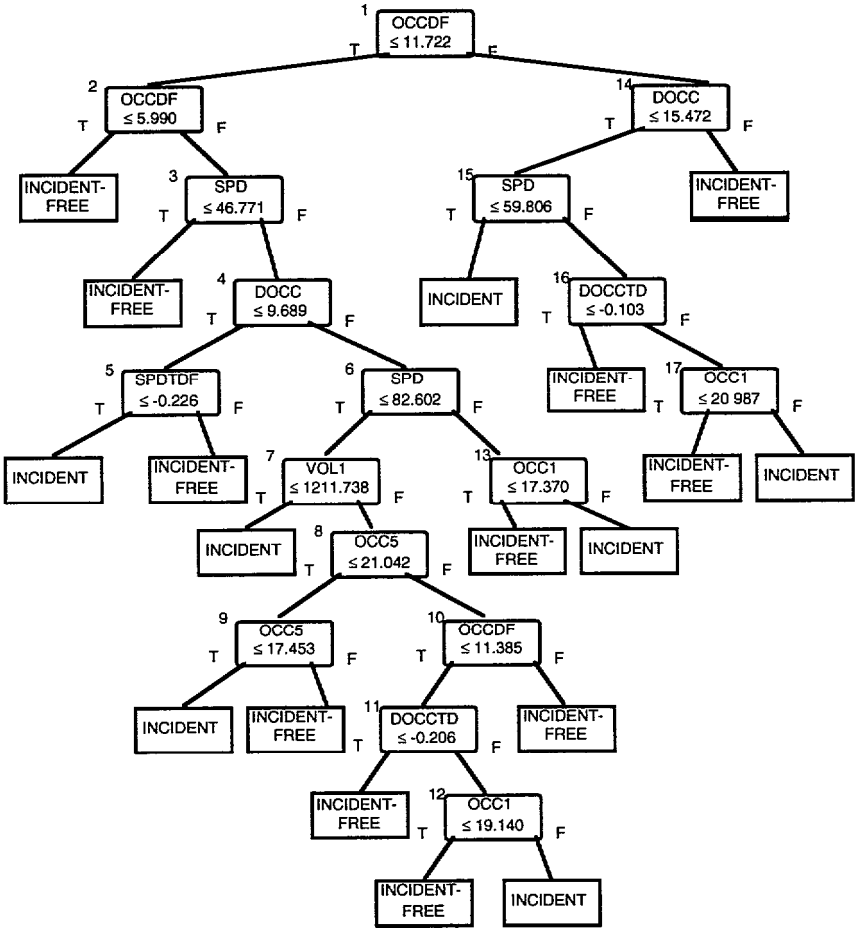


Figure 11-3. Binary Decision Tree Developed from I-880 Traffic Data

It was apparent from this that CART software was not very successful in generalizing the characteristics of incident and incident-free traffic data. On the contrary, the large number of terminal nodes indicates that the resulting trees were specific to the characteristics of the data in the learning set.

11.2 Algorithm Development Using Data from the Twin Cities Site

This section describes the algorithms that were developed using data from the Twin Cities development site. The collection and labeling of this data is addressed in Section 10.2. All algorithms described here were developed to distinguish between recurrent and incident-induced head of queue traffic conditions.

11.2.1 Preliminary Development

Preliminary algorithm development consisted primarily of exploring the performance benefits of various traffic features, defined as any function of the raw surveillance data. The features used in this regard are explicitly defined in Appendix E. In order to identify the features most useful in classifying traffic queues, CART was again used. Although our original intent in using CART was to arrive at a binary decision tree classifier to be used for incident detection, the simplicity of the resulting trees provided motivation for exploring other approaches.

The CART software was executed using a wide variety of possible traffic features. However, the produced algorithms clearly focused on two prominent features while essentially disregarding the rest. These features, downstream volume and the ratio of upstream speed to downstream speed, clearly demonstrated the most predictive power in discriminating incident congestion from incident-free congestion.

Subsequent analysis verified that these features were by far the most important indicators of the cause of congestion. It was found that congestion events with low downstream volumes very frequently resulted from an incident condition. Similarly, congestion events with high downstream volumes were much more likely to be incident free. This is consistent with certain features of the McMaster algorithm [Gall 89]. It was also noted that classification of congestion events with moderate downstream volumes was possible through use of the speed ratio feature.

For moderate values of downstream volume, a low value of speed ratio (upstream speed much lower than downstream speed) proved to be a good indicator of an incident. For congestion events in the learning set with downstream volumes in the range of 1600-1800 vehicles per hour per lane, all events with speed ratio less than 0.3 were found to be incidents. This clearly demonstrates the discriminating power of this feature for congestion events with moderate downstream volume.

By determining that the optimal classifier needs to make use of only two input variables, it was noted that a theoretical approach to probabilistic classification was feasible to pursue. This lead to the development of the Bayesian classification algorithm presented in the following section.

11.2.2 The Bayesian Classification Algorithm

The Bayesian classification algorithm presented in this section makes use of a well known result from probability theory. The desired output of the classifier can be written as the conditional probability P[I/M] , where

l = Incident Event, and
M = Surveillance Measurements

Given the a *priori* probabilities of incident and incident-free traffic conditions, one can compute the desired a *posteriori* probability making use of Bayes Theorem [ANDE 79]:

$$\mathbf{P[I/M] = \frac{P[I] \cdot P[M/I]}{P[I]* P[M/l]+ P[F]. P[M/IF]}}$$

where

1 = incident event,
I F = incident-free event,
M = surveillance measurements,
P[I] = a priori probability of an incident event, and
P[IF] = a priori probability of an incident-free event.

Implementation of this result is feasible only when the dimension of M (i.e., the number of surveillance measurement features) is small. Since preliminary algorithm development indicated that the features of primary importance were downstream volume and the ratio of upstream to downstream speeds, these features were selected for use in this algorithm. Hence, the desired algorithm output can be written as:

$$P[I/S,V] = \frac{P[I] \cdot P[S,V/I]}{P[I] \cdot P[S,V/I] + P[IF] \cdot P[S,V/IF]} \tag{11.1}$$

where V is the downstream volume feature and S is the ratio of upstream speed to downstream speed feature.

The unknowns on the right side of Equation 11.1 are: (1) the a priori probabilities $P[I]$ and $P[IF]$, and (2) the conditional probabilities $P[S,V/I]$ and $P[S,V/IF]$. Estimating the a priori probabilities was accomplished by inspecting the relative numbers of incident and incident-free events found in the learning set. Using this approach, a reasonable estimate of prior probabilities for the Twin Cities freeway system was obtained as: $P[I] = 0.2$, and $P[IF] = 0.8$.

Estimating the probabilities $P[S,V/I]$ and $P[S,V/IF]$ involved computing the joint density functions for downstream volume and speed ratio under both incident and incident-free traffic conditions. If we denote these densities by $f(S,V/I)$ and $f(S,V/IF)$, the desired a posteriori probability can be expressed as follows:

$$P[I/S,V] = \frac{P[I] \cdot f(S,V/I)}{P[I] \cdot f(S,V/I) + P[IF] \cdot f(S,V/IF)} \tag{11.2}$$

Estimation of the joint density functions $f(S,V/I)$ and $f(S,V/IF)$ required to implement Equation 11.2 was accomplished using data from the learning set and a technique termed Parzen Estimation [ANDE 79]. In essence, an estimate of the joint density function is obtained through a summation of Gaussian "kernels" corresponding to each (S,V) sample point in the learning set.

The resulting density estimates for incident-free and incident traffic conditions are shown in Figure 1 l-4 and 1 l-5, respectively. Utilizing the Parzen Estimation technique required that the range of the dependent variables be approximately equal. For this reason, downstream volume is scaled by a factor of 1000 in the figure and in all subsequent discussions.

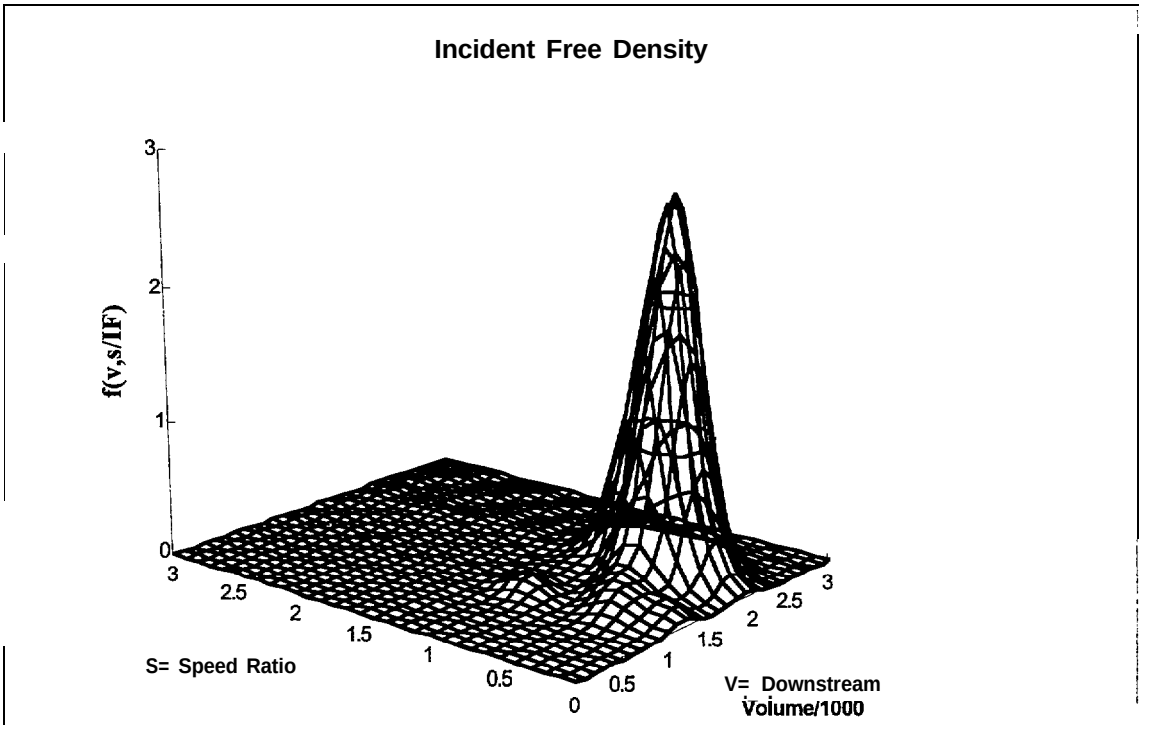


Figure 11-4. Incident-Free Density Function

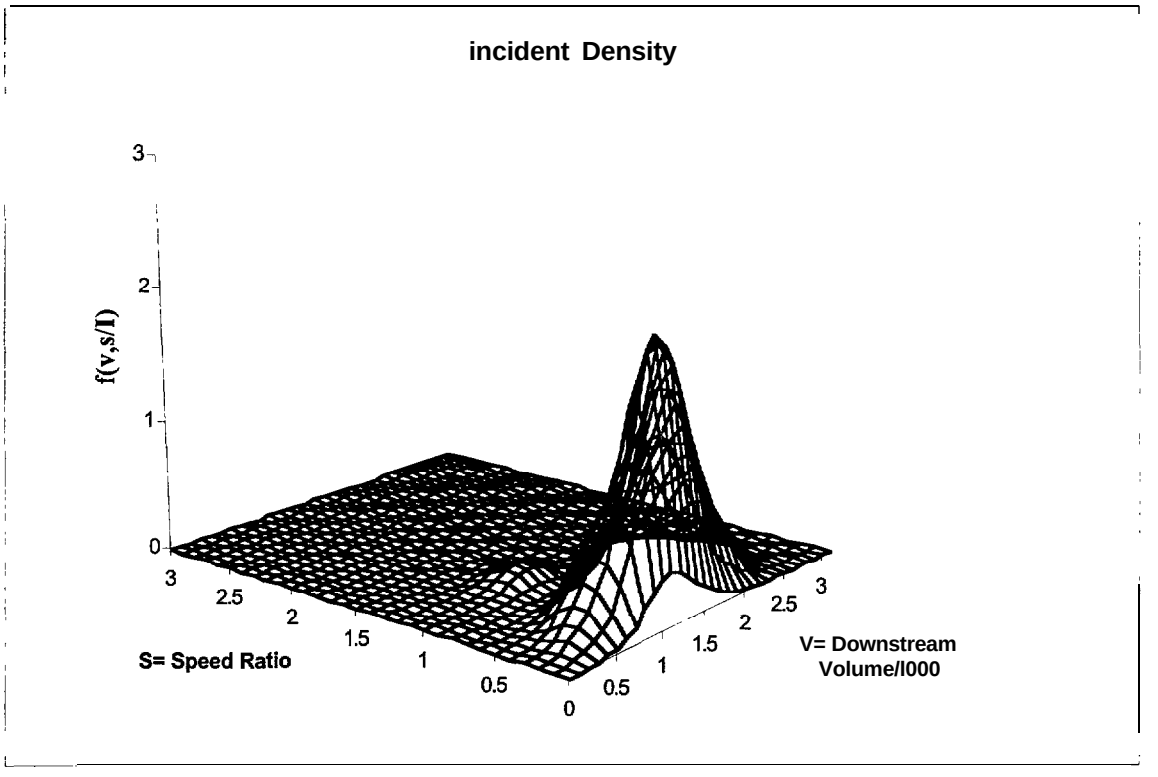


Figure 11-5. Incident Density Function

Given the joint densities for incident and incident-free traffic conditions, one can apply Equation 11.2 to compute the associated a posteriori probabilities. Figure 1 l-6 shows P[I/S, *V]* for various values of prior probabilities and values of *V* and S in the range (1.5 < *V* < 2.0) and (0.2 < S < 0.8), respectively. This range of inputs was selected to ensure that the estimated values of the joint density functions are sufficiently accurate to apply Equation 11.2. Outside this range, the number of congestion events in the learning set was insufficient to warrant use of the joint densities in Equation 11.2.

As shown in Figure 11-6, the prior probabilities of incident and incident-free congestion events has a significant impact on the algorithm's output. Specifkally, one can expect algorithm performance to increase as the prior probability of incident-free congestion decreases (e.g., in areas subject to little or no recurrent congestion).

In order to use these results operationally, an interpolation scheme is required to account for values of *V* and/or S that lie outside the range of input values selected for use with Equation 11.2. For the algorithm evaluations presented in Section 12.2, the algorithm was implemented so that values of *V* greater than 2.0 were evaluated in Equation 11.2 using

$V = 2.0$. Similarly, values of V less than 1.5 were evaluated at $V = 1.5$, values of S less than 0.2 were evaluated at $S = 0.2$, and values of S greater than 0.8 were evaluated at $S = 0.8$.

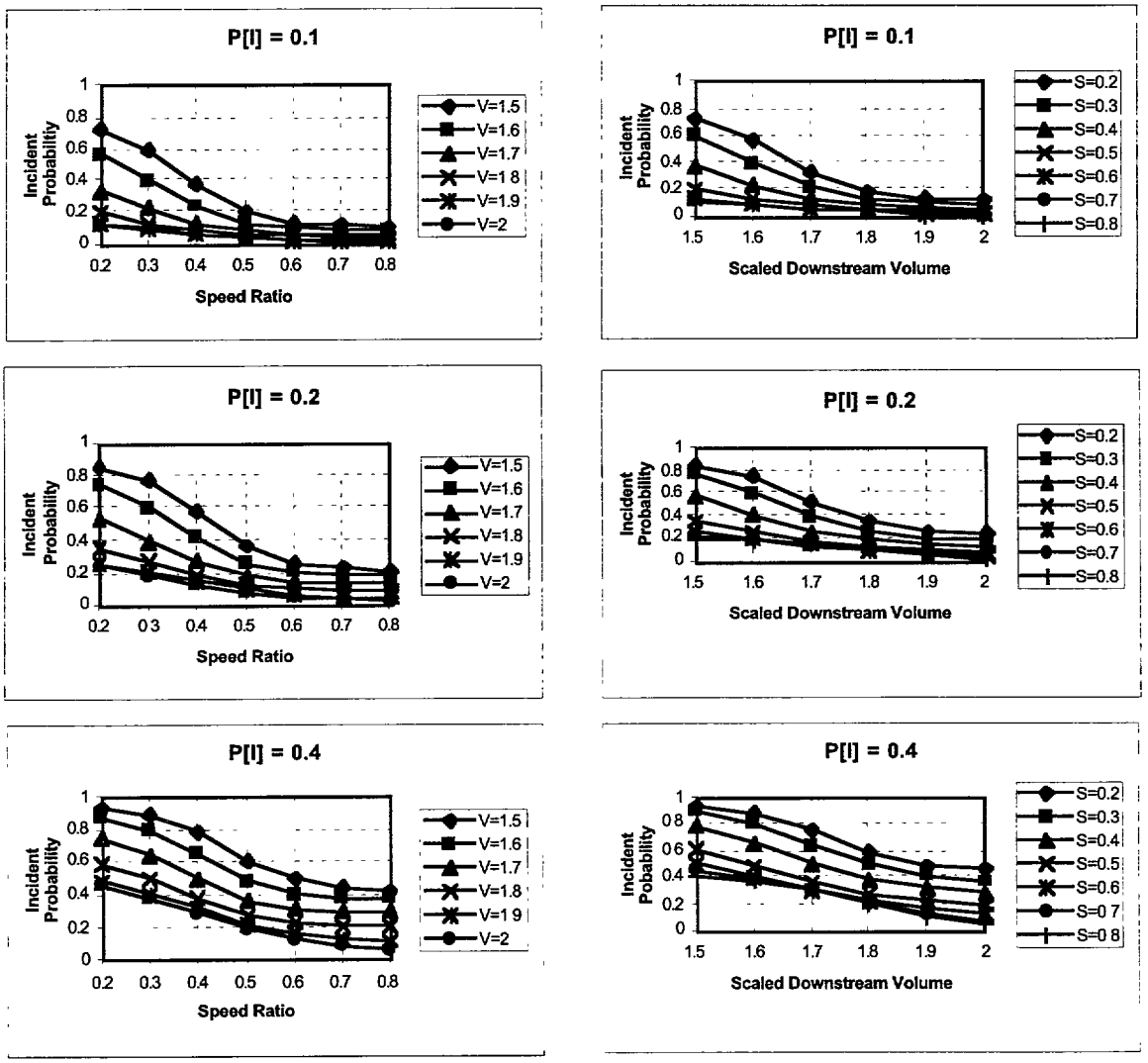


Figure 11-6. Summary of Bayesian Classification Results

11.2.3 Neural Network Algorithms

Neural networks have repeatedly demonstrated a high potential in solving statistical pattern recognition problems. As such, various neural networks were developed for use in classifying traffic queues. The network architecture and input features proposed by Ritchie [RITC 93] was used as a baseline. Other network inputs were also explored. All

developed networks are multi-layer feed-forward neural networks, also known as back propagation networks, of the form shown in Figure 1 l-7.

Neural network training was performed using a software package which allows the user to both train and test networks under various learning paradigms and network architectures. Each training session consisted of presenting the network with a sufficient number of input vectors (approximately lOO,OOO-150,000) so that the network RMS error reached a steady-state condition and it was apparent that no more learning could take place.

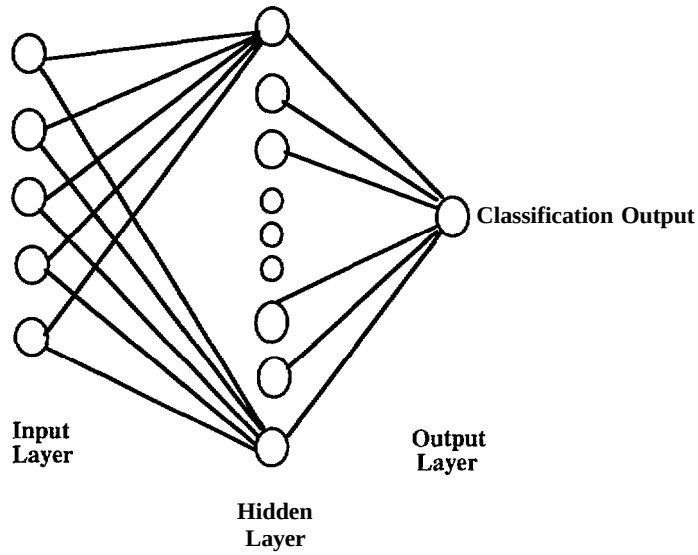


Figure 11-7. Structure of Neural Network Classifiers

Each developed network was subsequently implemented in the C++ programming language for evaluation purposes (the developed networks were delivered along with this final report in electronic format on CD-ROM). The evaluation procedure described in Section 12 was conducted in parallel with network development and served to guide the development process. Network inputs and the number of processing elements (PE's) in the hidden layer were systematically varied in order to arrive at a network architecture that performed best.

Two of the developed networks are highlighted here. Although many networks performed somewhat similarly, these networks exhibited the best overall performance.

The first network will be subsequently referenced as network ID4d in the algorithm evaluations presented in Section 12. This network has 10 PEs in the hidden layer and takes

as input upstream and downstream values of one-minute volume, occupancy and speed. Again, networks were explored that utilized a much richer input set, but the performance of these networks was generally inferior to the performance of network ID4d.

The second network to be presented will be subsequently referenced as network IDS in the algorithm evaluations presented in Section 12. This network also has 10 PEs in the hidden layer and takes as input a time series of the raw upstream and downstream surveillance measurements of 30-second volume and occupancy. The time series consists of the measurement value at the current time as well as the measurements of the preceding two minutes. This is the network architecture and input data proposed by Ritchie [RJTC 93].

12.0 ALGORITHM PERFORMANCE EVALUATIONS

This section presents performance evaluations for the algorithms described in Section 11. All algorithm evaluations are based on real traffic data from the development sites identified in Section 6.

12.1 Algorithm Evaluations Using the I-880 Data Sets

The performance evaluations conducted for the algorithms developed from the I-880 database indicated that these algorithms were generally ineffective for the purposes of operational incident detection. When executed in an operational setting, these algorithms were successful in identifying the great majority of the incidents contained in the learning set, as well as several incidents that were reported but not included in the learning set due to uncertainties in the labeling (refer to Section 10.1). However, the algorithms generally produced an excessive number of false alarms when used during periods of congestion.

Frequently, the binary decision tree produced by CART successfully partitioned the learning set so that the distribution of data samples in the terminal nodes clearly indicated either an incident or an incident-free classification. Nevertheless, when the algorithms were run operationally against the I-880 surveillance data (e.g., in continuous time for each zone), a large number of false alarms were produced that apparently corresponded to minor fluctuations in the traffic measurements.

Formal evaluations involving the computation of sophisticated MOEs were not conducted for the I-880 algorithms. The general performance characteristics of the algorithms, as determined by the methods described below, did not warrant such formal evaluations.

The evaluation methodology used for I-880 algorithms consisted primarily of executing the algorithm in an operational setting and tabulating the results in terms of identifying the terminal node of the binary decision tree that was reached for each data sample. The algorithms were seen to perforrn poorly when incident-free data samples reached terminal nodes associated with the incident classification. In this manner, an assessment was made regarding the true and false incident alarms that would result for the algorithm using a nominal value of detection threshold.

Additionally, the branching logic of the trees were qualitatively examined in an attempt to ascertain certain underlying theoretical principles involved in discriminating between incident and incident-free traffic conditions. Frequently, the branching logic was found to be counter-intuitive and could not be reconciled with theoretical considerations. The lack of understandable results indicates that branching logic was selected by CART to match the specific characteristics of the learning set, and that the classifier would be ineffective when applied operationally.

Furthermore, the size of the developed trees was frequently large, resulting in a large number of terminal nodes. Although CART was somewhat successful in partitioning the learning sets into the terminal nodes representing incident and incident-free conditions, the "size" of the terminal nodes (e.g., the number of data samples which satisfy the node's branching criteria) was frequently rather small. This provides further evidence that the classifier was unsuccessful in identifying general incident characteristics from the learning set.

Consider the classifier shown in Figure 12-1 that was developed using data from the I-880 data set. This classifier is representative of the developed I-880 algorithms in general. Definitions of the features used in the branching logic of this tree are presented in Appendix E. For each terminal node of the tree, the number of incident and incident-free data samples from the learning set that satisfied the branching criteria leading to the node are shown. If this classifier were to be used operationally, the ratio of incident to incident-free data samples would be used to form a multi-class probability estimate associated with each terminal node. To investigate algorithm performance, terminal nodes are designated as either incident or incident-free based on the distribution of data samples in that node. For example, terminal node 10 clearly corresponds to an incident classification.

As shown in Figure 12-1, several terminal nodes contain only a few data samples. Also, the branching logic is generally not intuitive. This leads one to believe that this branching logic reflects specific characteristics of the learning set and is not representative of incident and incident-free traffic characteristics in general (in spite of the fact that the samples in learning set are classified effectively).

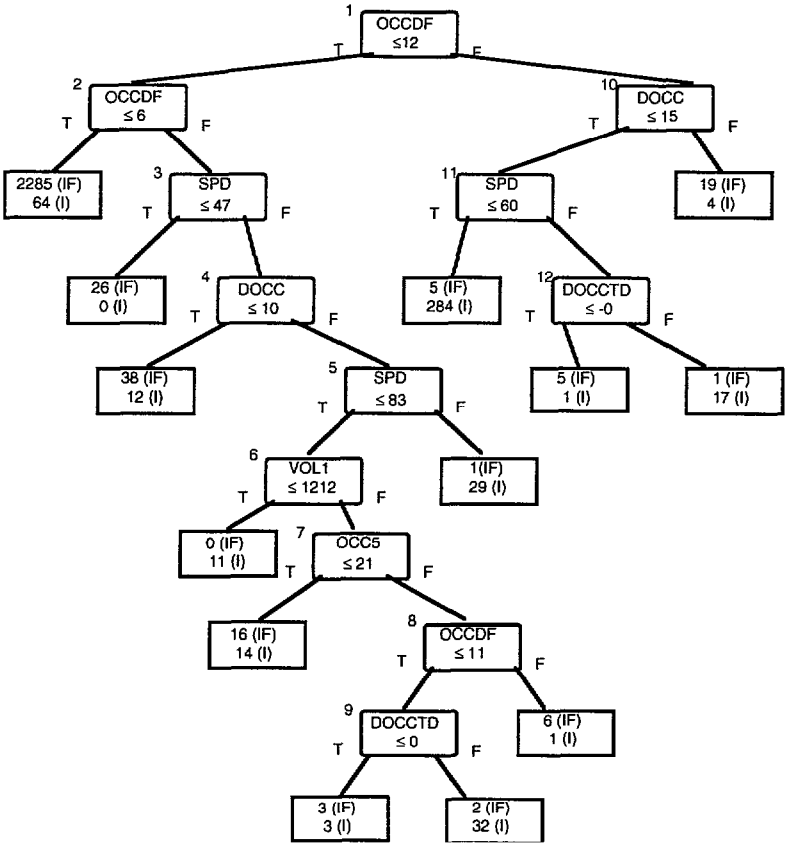


Figure 12-1. Binary Decision Tree for I-880

This assertion is verified by conducting operational evaluations. The results of such an evaluation are illustrated in Figure 12-2, which shows the terminal node reached for data samples taken from the southbound segment of I-880.

Zones 2160 and 2180 are not shown in the figure as they were not available due to sensor malfunctions. Terminal node numbers shown in square brackets correspond to terminal nodes with incident designations. The incident alarm in zone Z70 at time 15:58 corresponds to an actual incident. This was true of I-880 algorithms in general - the incidents included in the training set were generally detected.

Note however that there are multiple false alarms in zones Z30 and Z40 - no incidents were reported in these zones. The propensity of the I-880 algorithms to produce false alarms during periods of congestion render them operationally ineffective. This is further illustrated by Figure 12-3, which shows the terminal nodes reached for the classifier of Figure 12- 1 during a period when no incidents were reported.

Time	Z30	Z10	Z70	200	Z94	Z20	Z110	Z190	Z130	Z120	Z40
15:55:00	[10]	1	1	1	1	1	1	1	1	1	[12]
15:55:30	[10]	1	1	1	1	1	1	1	1	1	11
15:56:00	[12]	1	1	1	1	1	1	1	1	1	[12]
15:56:30	[7]	1	1	1	1	1	1	1	1	1	[12]
15:57:00	[10]	1	1	1	1	1	1	1	1	1	[12]
15:57:30	[10]	1	1	1	1	1	1	1	1	[7]	[9]
15:58:00	[10]	1	[9]	1	1	1	1	1	1	1	[9]
15:58:30	[10]	1	3	1	1	1	3	1	1	1	[12]
15:59:00	[10]	1	5	1	1	1	1	1	1	1	8
15:59:30	[7]	1	1	1	1	1	1	1	1	1	[12]
16:00:00	[12]	1	5	1	1	1	1	1	1	1	[12]
16:00:30	[12]	1	[10]	1	1	1	1	1	1	1	11
16:01:00	[12]	1	[10]	1	1	1	1	1	1	1	[12]
16:01:30	[10]	1	[10]	1	1	1	1	1	1	1	[9]
16:02:00	[10]	1	[10]	1	1	1	1	1	1	1	11
16:02:30	[10]	1	[10]	1	1	1	1	1	1	1	[9]
16:03:00	[10]	1	[10]	1	1	1	1	1	1	1	[9]
16:03:30	[10]	1	[10]	1	1	1	1	1	1	1	[9]
16:04:00	[10]	1	[10]	1	1	1	1	1	1	1	[9]
16:04:30	[7]	1	[10]	1	1	1	1	1	1	1	[9]
16:05:00	[11]	1	[10]	1	1	1	1	1	1	1	[9]
16:05:30	[7]	1	[10]	1	1	1	1	1	1	1	[9]
16:06:00	3	1	[10]	1	1	1	1	1	1	1	[9]
16:06:30	[12]	1	[10]	1	1	1	1	1	1	1	11
16:07:00	[10]	1	[12]	1	1	1	1	1	1	1	[12]
16:07:30	[10]	1	3	1	1	1	1	1	1	1	[12]
16:08:00	[10]	1	1	1	1	1	1	1	1	1	[12]
16:08:30	[10]	1	1	1	1	1	1	1	1	1	11
16:09:00	[10]	1	1	1	1	1	1	1	1	1	6
16:09:30	[10]	1	[7]	1	1	1	1	1	1	1	6
16:10:00	[10]	1	[10]	1	1	1	3	1	1	1	[12]

Figure 12-2. Sample Evaluation Corresponding to Incident Conditions

For this time period, zones Z10 and Z180 were not available due to sensor malfunctions and are therefore not shown in the figure. Again, terminal node numbers shown in square brackets correspond to terminal nodes with incident designations. Although the time period for this execution of the algorithm corresponded to congested traffic conditions, the number of false alarms produced by the algorithm is excessive.

Time	Z30	Z70	200	Z94	Z20	Z110	Z60	Z190	Z130	Z120	Z40
7:40:00	1	1	1	1	1	3	1	1	1	2	[10]
7:40:30	[7]	1	1	1	1	3	1	1	1	2	[10]
7:41:00	[10]	1	1	1	1	1	1	1	1	1	[10]
7:41:30	[10]	1	1	1	1	3	1	1	1	2	[10]
7:42:00	[10]	1	1	1	1	3	1	1	1	2	[10]
7:42:30	[10]	1	1	1	1	1	1	1	1	13	[10]
7:43:00	[10]	1	1	1	1	1	1	1	1	13	[10]
7:43:30	[10]	1	1	1	1	1	1	1	1	13	[10]
7:44:00	[10]	1	1	1	1	1	1	1	2	1	[10]
7:44:30	[10]	1	1	[9]	1	1	1	1	13	1	[10]
7:45:00	[10]	1	1	[9]	1	1	1	1	2	1	[10]
7:45:30	[10]	1	1	5	1	5	1	1	1	1	[10]
7:46:00	[10]	1	1	1	1	[10]	1	1	1	1	[10]
7:46:30	[12]	1	1	1	1	[10]	1	2	1	1	[10]
7:47:00	[10]	1	1	1	1	5	1	2	1	2	[10]
7:47:30	[10]	1	1	1	1	[9]	1	1	1	2	[10]
7:48:00	[10]	1	1	5	1	[9]	1	1	1	13	[10]
7:48:30	[10]	1	1	5	1	[7]	1	1	1	2	[10]
7:49:00	13	5	1	5	1	[12]	1	1	1	13	[10]
7:49:30	6	5	1	1	1	[10]	1	1	1	13	[10]
7:50:00	13	[9]	1	1	1	[10]	1	1	1	2	[10]
7:50:30	13	1	1	1	1	[7]	1	1	1	[7]	[10]
7:51:00	[10]	1	1	1	1	[7]	1	1	2	[7]	[10]
7:51:30	13	1	1	1	1	[7]	1	1	1	1	[10]
7:52:00	13	1	1	1	1	1	1	1	1	1	[10]
7:52:30	[10]	1	1	1	1	[9]	1	1	1	1	[10]
7:53:00	[10]	1	1	1	1	[9]	1	2	1	1	[10]
7:53:30	[10]	1	1	1	1	[9]	1	2	1	1	[10]
7:54:00	[10]	1	1	1	1	[9]	1	1	1	1	[10]
7:54:30	[10]	1	1	1	1	5	1	[7]	1	2	[10]
7:55:00	[10]	1	1	1	1	[10]	1	[7]	1	2	[10]

Figure 12-3. Sample Evaluation Corresponding to Incident-Free Conditions

As the analysis of the type shown above clearly demonstrated that the developed algorithms would not be operationally effective, no formal evaluations were conducted, and new algorithms were developed to address the observed algorithm deficiencies. The example presented here is indicative of the performance of the I-880 algorithms in general (e.g., too many false alarms). However, the development effort for I-880 constituted our first use of actual traffic data and was indispensable as a preliminary step in devising effective algorithms.

The effort to develop effective incident detection algorithms using data from the I-880 surveillance system was largely a learning experience whereby new ideas were postulated

and the corresponding algorithms subsequently developed and tested. Many of the concepts explored appeared to hold significant promise for improving a particular aspect of algorithm performance. However, new problems were encountered at each stage of development, leading to further innovations in our approach.

In this manner, we were able to identify several important characteristics of traffic measurement data as well as their associated adverse impacts on incident detection algorithms. As a direct result, several important pre-processing steps with regard to incident detection were identified, and algorithms were developed in subsequent phases of this project to address these issues.

These efforts ultimately lead to the development of methods for characterizing congestion, such as the queue identification and tracking algorithm described in Section 9. These methods were subsequently used in the algorithm development effort for the Twin Cities site.

12.2 Algorithm Evaluations Using the Twin Cities Data Sets

This section presents performance evaluations of the algorithms addressed Section 11.2. All evaluated algorithms address the discernment of incident-induced queuing from recurrent head-of-queue traffic conditions.

These algorithms were developed using data from the learning set identified in Table 10.2. Consequently, evaluations based on the learning set tend to be unrealistically optimistic. For comparison purposes, evaluations are presented in terms of algorithm performance for both the learning and evaluation data sets. However, objective estimates of algorithm performance can be obtained only through use of the independent evaluation data set.

The approach to evaluation is described in detail in Section 5. Note that the interpretation of detection threshold is different for each of the presented algorithms. For the Bayesian Classification Algorithm, the detection threshold is the a posteriori probability of incident occurrence; for the neural networks, it is the network output, in the range of 0 to 1; and for the California algorithm, it is the index of the threshold set being applied.

12.2.1 Definitions of Applicable MOEs

The MOEs defined here were generated for each algorithm as a function of detection threshold for both the learning set and the evaluation data set. All MOEs are defined in terms of an algorithm's performance with respect to congestion *events.*

Each congestion event in the evaluation data set consists of several measurement records. The algorithms to be evaluated produce a probability of incident occurrence for each measurement record. In order to determine an algorithm's performance with respect to congestion events, the algorithm output is processed to form *an event probability,* defined as the highest probability associated with the event's measurement records within a window spanning from one minute prior to the onset of the head-of-queue condition to two minutes following the onset of the head-of-queue condition.

The MOEs selected for use in algorithm evaluations are Detection Rate (DR), False Alarm Rate (FAR) and Operational Detection Rate (ODR). For anv given detection threshold, DR is defined as the number of detected incident events (e.g., those events with probabilities exceeding the detection threshold) divided by the total number of incident events in the evaluation data set. Once again, this is not the same as the percentage of incident records classified correctly divided by the total number of incident records, as some researchers define it. We feel that our definition more accurately reflects operational concerns (e.g., was the incident detected or not).

Similar to DR, for any given detection threshold, FAR is defined as the number of incident-free events that resulted in an incident alarm (e.g., those events with probabilities exceeding the detection threshold) divided by the total number of incident-free events in the evaluation data set. Again, we feel this reflects operational concerns more accurately than FAR defined as the percentage of incident-free records that were misclassified.

The ODR measure is defined in terms of DR, FAR and also the a priori probabilities of incident and incident-free congestion events. The a priori probabilities must be taken into account since the relative frequency of these events as encountered in practice may not be accurately reflected in the evaluation data set. For the Twin Cities freeway system, we estimate that the prior probability of an incident event is 0.2 and the prior probability of an incident-free event is 0.8. These probabilities will vary for different sites.

For a given value of detection threshold, ODR is defined as the number of true incident alarms divided by the total number of incident alarms (i. e., the percentage of incident alarms which are true).

$$ODR = \frac{Alarms_{True}}{Alarms_{True} + Alarms_{False}}$$

If the number of congestion events to be classified is large, we can accurately estimate the number of true and false incident alarms produced by a given algorithm as

$$Alarms_{True} = DR \cdot Events_{Incident} = DR \cdot Events_{Total} \cdot P[I]$$
$$Alarms_{False} = FAR \cdot Events_{Incident\text{-}Free} = FAR \cdot Events_{Total} \cdot P[IF]$$

where

$Events_{Incident}$	= Number of incident events,
$Events_{Incident\text{-}Free}$	= Number of incident-free events,
$Events_{Total}$	= Total number of congestion events,
$P[I]$	= a priori probability of an incident event, and
$P[IF]$	= a priori probability of an incident-free event.

Hence, the ODR measure can be written as

$$ODR = \frac{DR \cdot Events_{Total} \cdot P[I]}{DR \cdot Events_{Total} \cdot P[I] + FAR \cdot Events_{Total} \cdot P[IF]}$$

or,

$$ODR = \frac{DR \cdot P[I]}{DR \cdot P[I] + FAR \cdot P[IF]}$$

12.2.2 Evaluation of the Bayesian Classification Algorithm

The performance of the Bayesian Classification Algorithm (see Section 11.2.2) is shown in Figure 12-4 in terms of the selected MOEs of DR, FAR and ODR.

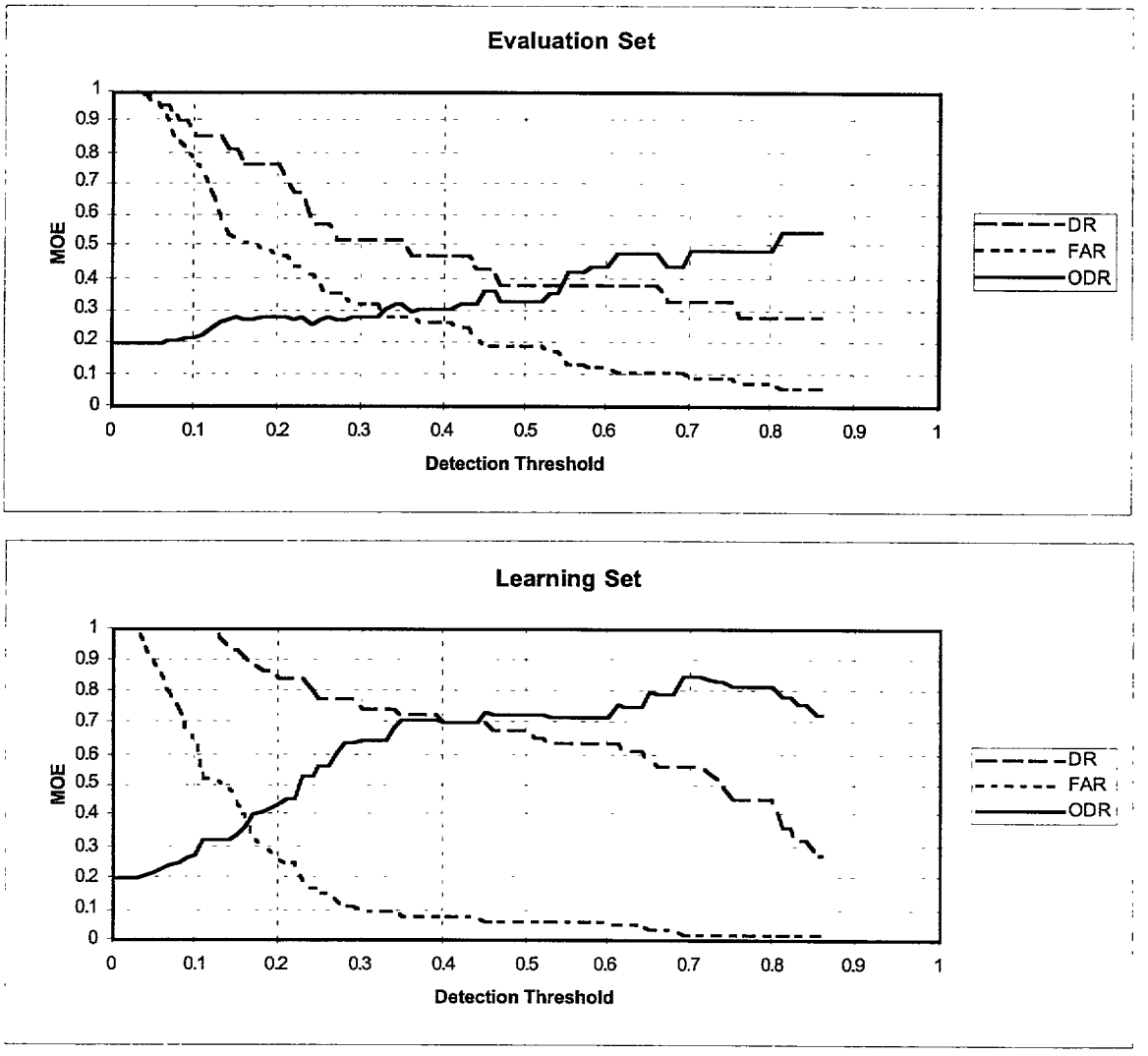


Figure 12-4. Performance of the Bayesian Classification Algorithm

Each MOE shown in the figure is presented as a function of detection threshold for both the learning and evaluation data sets. Note that the performance of the algorithm with respect to the learning set is markedly increased from the performance relative to the evaluation set. For example, using a nominal detection threshold of 50%, one can see that approximately 70% of the congestion events in the learning set were detected by the algorithm. Conversely, approximately 40% of the incidents in the evaluation data set were detected at

this threshold value. Similar observations can be made regarding the ODR and FAR measures.

In general, the algorithm's performance with respect to the evaluation data set is much poorer that the performance relative to the learning set. The disparity between these evaluations arises due to the fact that the algorithm was trained using data from the learning set. This result is true in general. Although certain researchers present algorithm evaluations in terms of performance relative to the learning set, objective assessments of field performance can only be obtained through use of an independent evaluation data set.

12.2.3 Evaluation of Neural Network Algorithms

Figures 12-5 and 12-6 show the performance evaluations for the neural network algorithms addressed in Section 11.2.3. Although several additional networks were developed in the course of this research, the networks presented here were found to perform best.

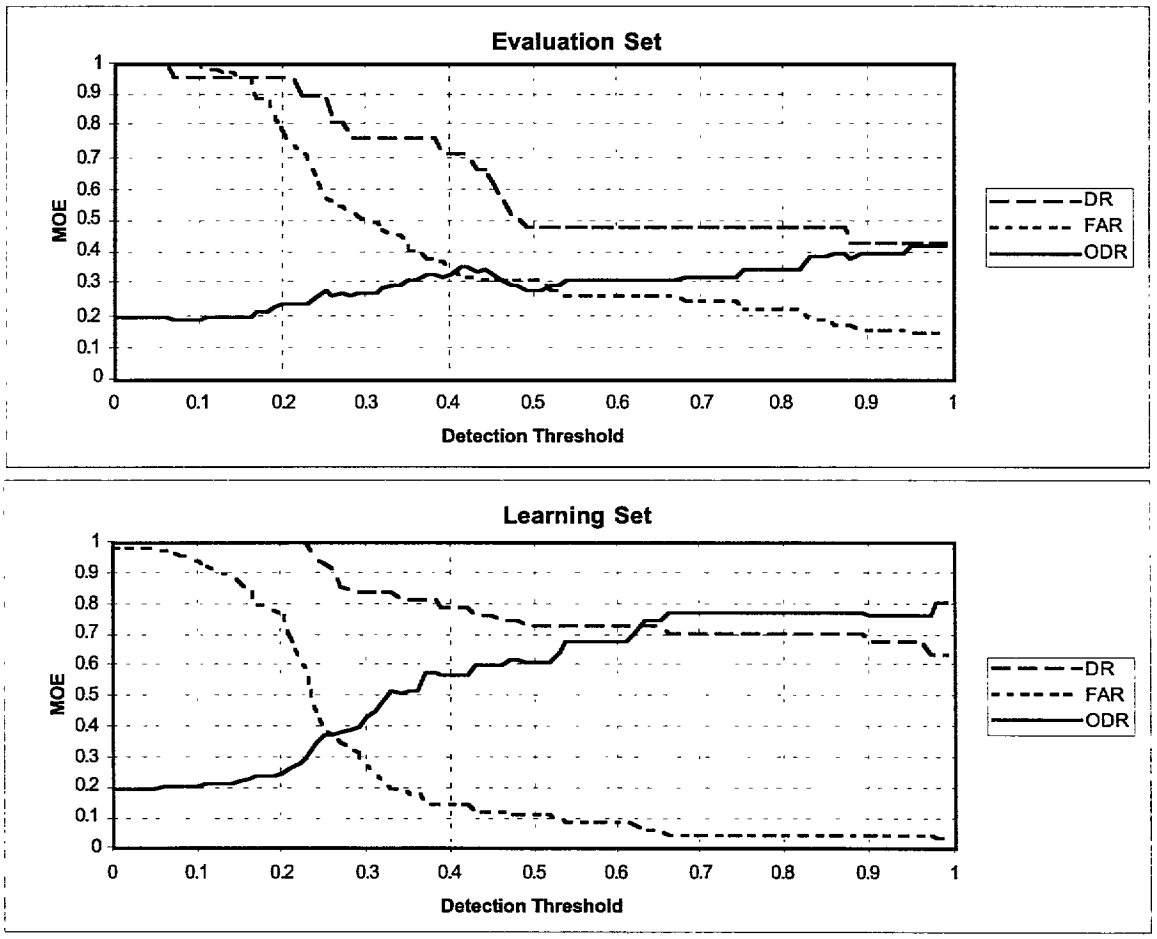


Figure 12-5. Performance of the Neural Network ID4d Algorithm

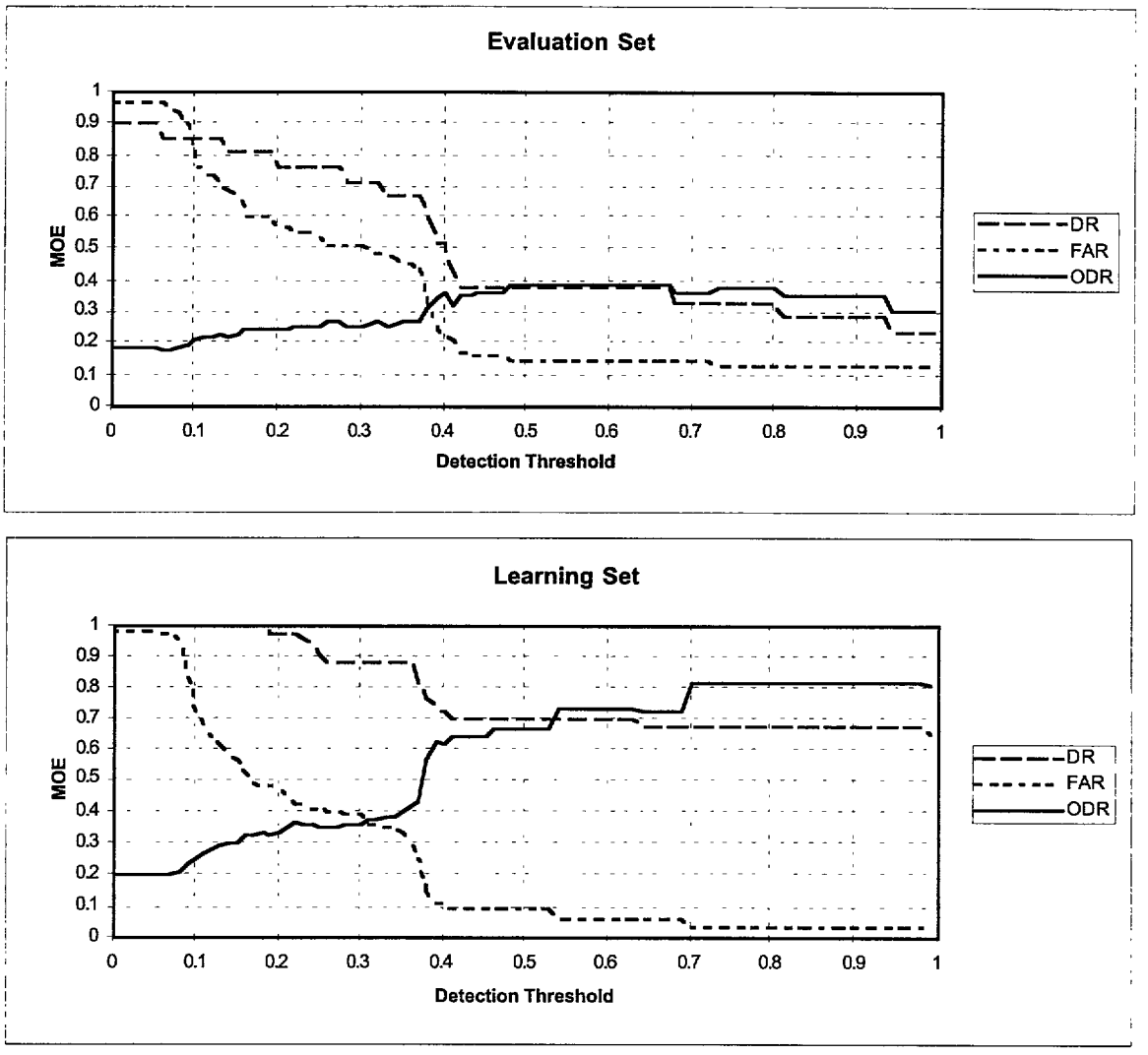


Figure 12-6. Performance of the Neural Network IDS Algorithm

The figures show the selected MOEs of DR, FAR and ODR as functions of detection threshold for both the learning and evaluation data sets. Once again, it is clear that the performance of the algorithms with respect to the learning set is markedly increased from the performance relative to the evaluation set.

12.2.4 Evaluation of the California Algorithm

For comparison purposes, the California #8 algorithm was also selected for evaluation. The reasons for selecting this algorithm are as follows: (1) the California algorithm is well known and is often used for performance comparisons, (2) calibrated threshold sets from

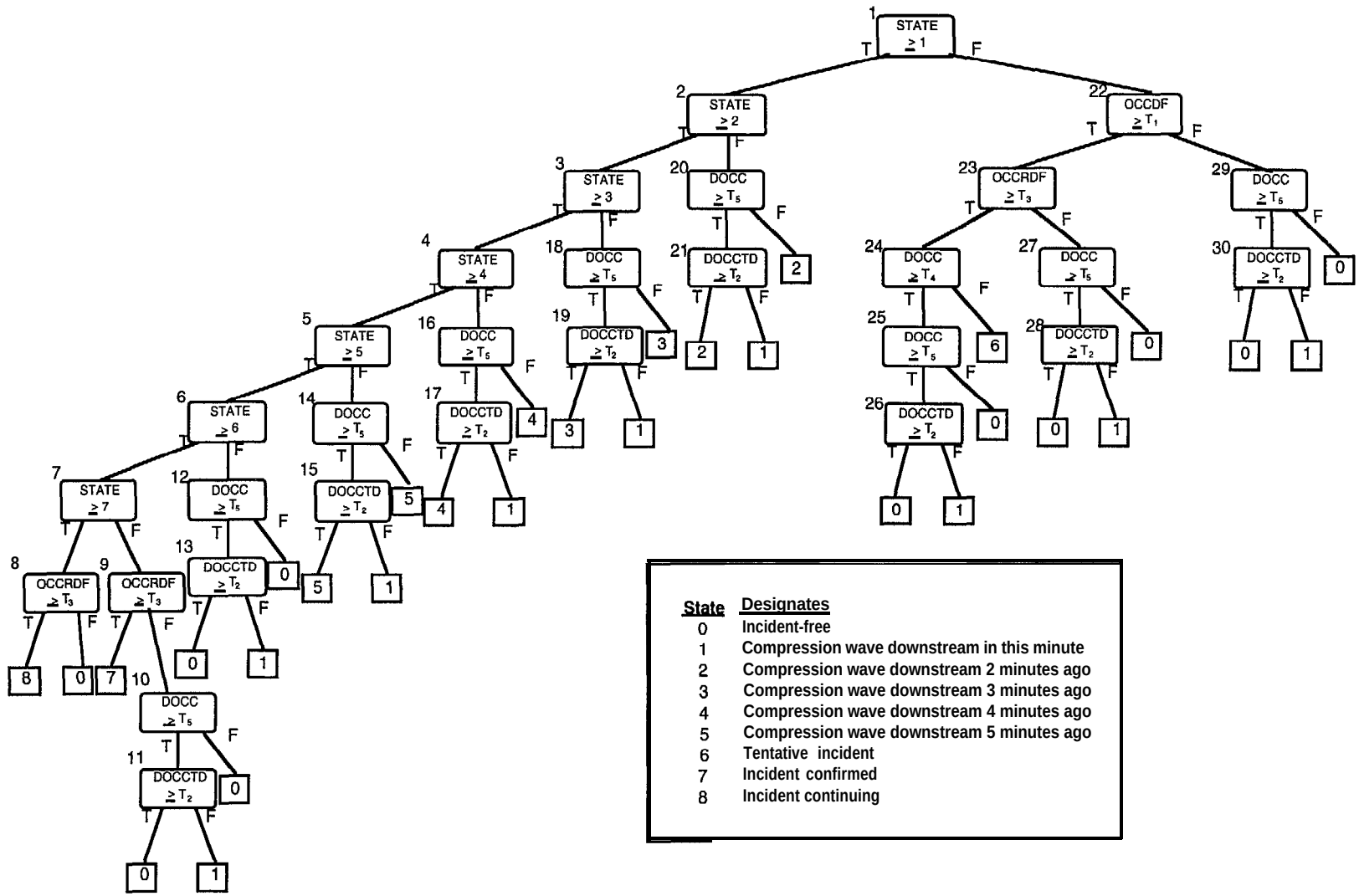
the Twin Cities freeway system are available for this algorithm, and (3) the California #8 algorithm specifically addresses the issue of distinguishing between recurrent congestion and incident-induced congestion.

The report [PAYN 76] provides a complete description of the algorithm and the associated threshold sets calibrated for the Twin Cities freeway system during original development of the algorithm. These results are summarized here in Figure 12-7, which shows the branching logic of the algorithm, and in Table 12-1, which shows the algorithm threshold sets calibrated for the Twin Cities freeway system.

Table 12-1. Threshold Sets for California Algorithm #8, Calibrated for Minneapolis [PAYN 76]

Set #	Detection Rate (%)	T_1	T_2	T_3	T_4	T_4
1	80.6	7.4	-.259	.302	27.3	30.0
2	72.2	17.4	-.649	.391	25.2	30.0
3	61.1	27.8	-.320	.606	28.8	30.0
4	55.6	30.0	-.453	.508	15.4	30.0
5	41.7	30.0	-.677	.724	15.3	30.0
6	36.1	27.8	-.689	.750	11.6	30.0
7	27.8	28.0	-1.084	.792	10.7	30.0

Our implementation of the algorithm consisted of executing the algorithm in parallel for <u>each</u> of the seven calibrated threshold sets listed in Table 12-1. In doing so, we recover the notion of a probabilistic algorithm output (see Section 5). Specifically, if the algorithm produces an alarm for the most stringent threshold set (e.g., the threshold set least likely to produce a false alarm), then a users degree of confidence in the result would be higher than if the algorithm produced an alarm using the least stringent threshold set (e.g., the threshold set most likely to produce a false alarm).



State	Designates
0	Incident-free
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 12-7. California Algorithm # 8

The performance of the California Algorithm #8, as defined above, is presented in Figure 12-8.

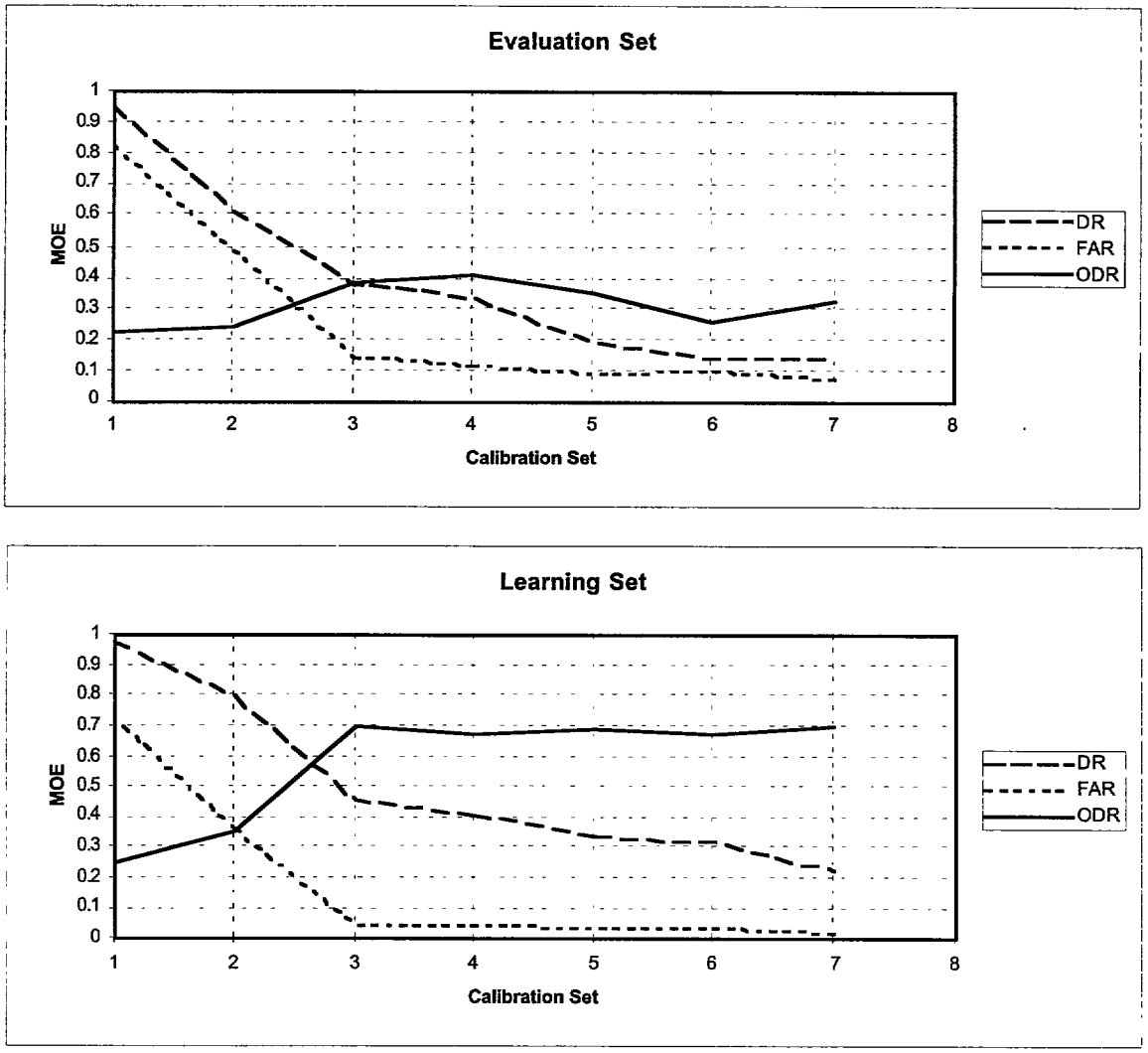


Figure 12-8. Performance of the California #8 Algorithm

12.2.5 Algorithm Performance Comparisons

All algorithms addressed in the preceding sections have the same general performance characteristics. Specifically, increased detection rate (e.g., for low detection thresholds) can be obtained only at the expense of increased false alarm rate.

Since the relative number of true and false incident alarms is reflected in the ODR measure, this tradeoff can be quantified by plotting DR versus ODR, as shown in Figure 12-9.

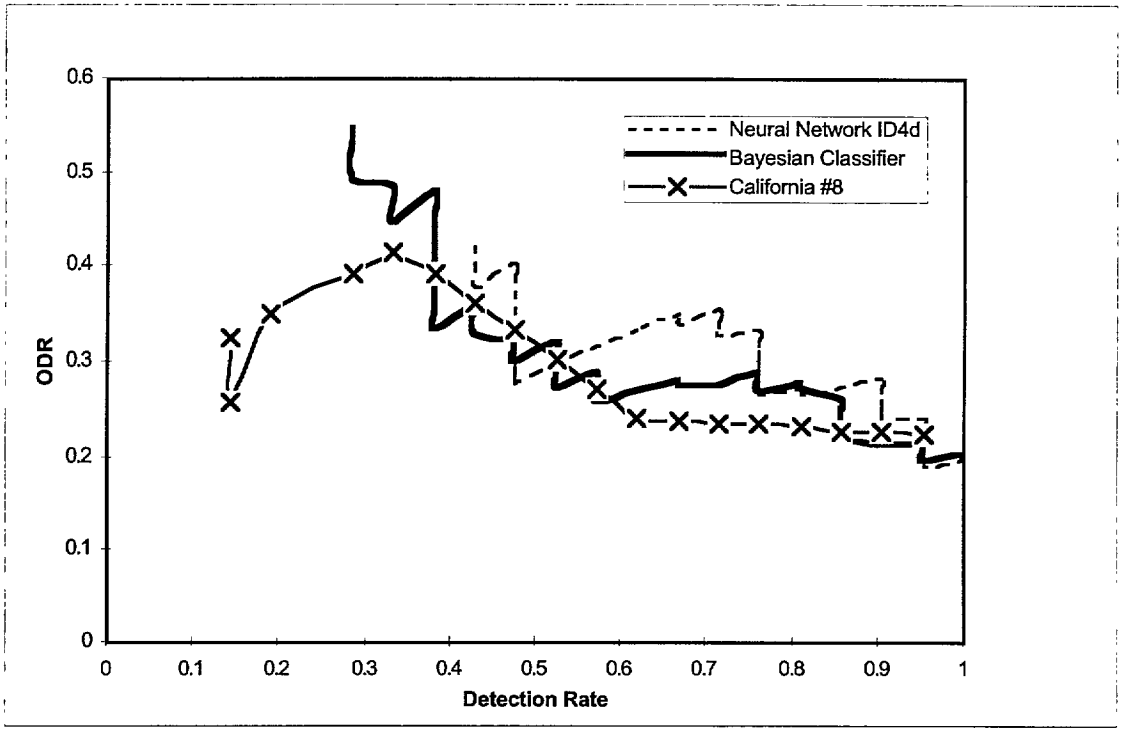


Figure 12-9. Algorithm Performance Comparison

In the interest of clarity, only one of the neural network algorithms are shown in the figure. Clearly, the selected algorithms can be seen to perform similarly.

There is some indication (though not statistically validated) that the neural network approach produced somewhat better results. As can be seen from Figure 12-9, all of the algorithms produced only modest ability to discriminate incidents as the cause of new congestion events, with the operational detection rate being in the range of 20 to 40 per cent.

12.2.6 The Contribution of Queue Tracking

The evaluations presented in the preceding sections indicate that the respective algorithms perform quite similarly when applied to head-of-queue traffic conditions. Remaining to be explored is the performance benefit associated with the queue tracking algorithm. The California algorithm was selected to pursue this issue since it was developed to operate on

all the data, while the remaining algorithms are intended to operate only on head-of-queue traffic conditions.

The methodology used to explore the benefits of the queue tracking approach is as follows. The California algorithm was executed against the entire evaluation data set (as opposed to just the onset of congestion events, as presented in the previous section), and the number of incident alarms was noted. By disregarding alarms that corresponded to congestion events labeled as either "incident" or "unknown" during the Minneapolis data collection effort (see Section 10.2.2) one may obtain an estimate of the number of alarms that were either false or did not correspond to a congestion event. If one asserts that alarms not corresponding to a congestion event are "false" in the sense that no operator action is required at the TMC, one can compute the Operational False Alarm rate (OFA) of the algorithm (defined as 1.0 - ODR), as shown in Figure 12-10.

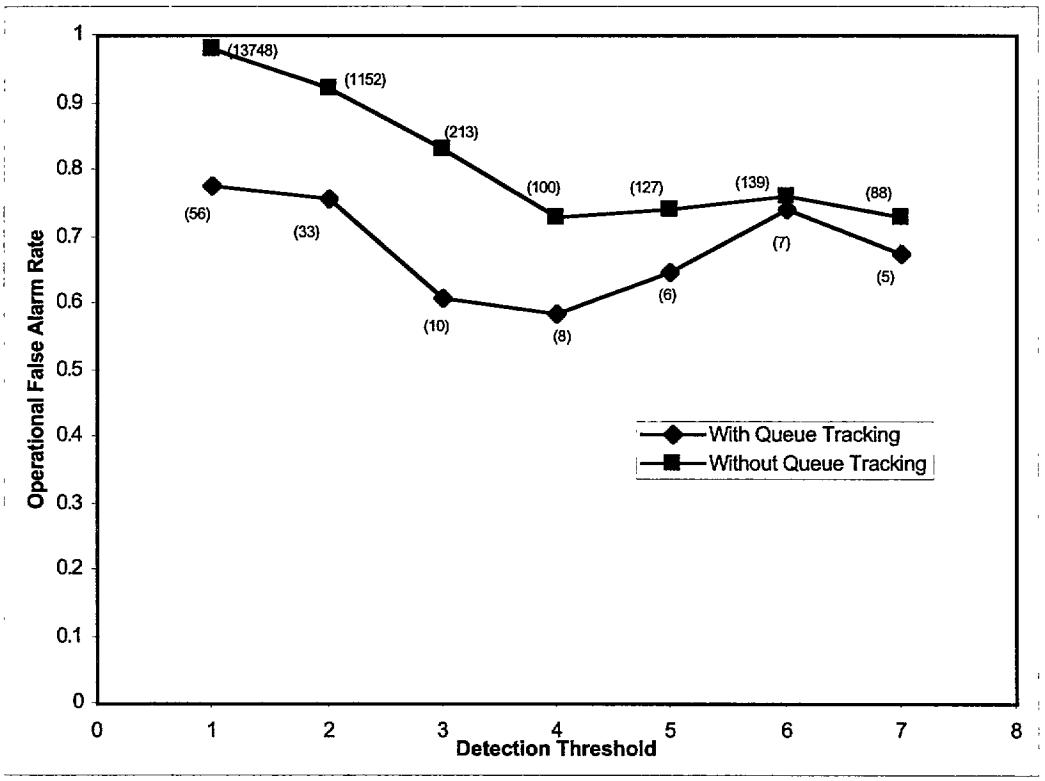


Figure 12-10. Contribution of the Queue Tracking Algorithm

In the figure, the OFA of the California #8 algorithm used in conjunction with queue tracking is reproduced from Figure 12-8. Note that queue tracking serves to reduce the OFA of the algorithm for all threshold sets. Perhaps more importantly, use of the queue

tracking algorithm greatly reduces the total number of false alarms generated by the algorithm. These are shown in parentheses on the figure for each threshold set.

For threshold set 2 (refer to Table 12-l), and when no queue tracking was applied, the algorithm produced a total of 1152 alarms that were “false” during the week-long evaluation interval. Upon the application of queue tracking, the algorithm produced only 33 false alarms during the same interval. This reduction in false alarms is quite dramatic. In fact, the high number of alarms produced at certain thresholds by the California algorithm would render it essentially useless from an operational perspective.

12.3 Planned Evaluations in San Diego

An important issue for incident detection algorithms is the transferability of results to different freeway systems. Partial work was done in this regard in order to explore the transferability of the presented algorithms to the San Diego development site. Specifically, the GUI display capability used in the Minneapolis data collection effort (see Section 10.2.2) was implemented for the San Diego site to support real-time field tests of the algorithms. Sample output from this display is shown in Figure 12-11.

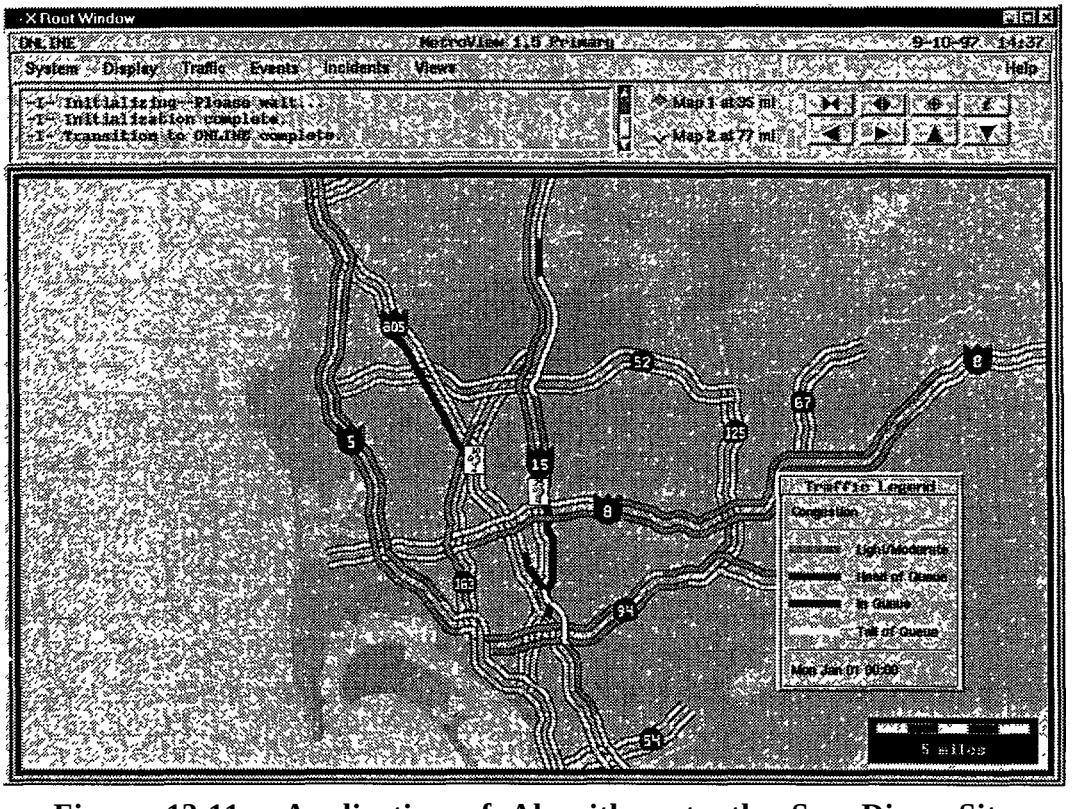


Figure 12-11. Application of Algorithms to the San Diego Site

As seen in the figure, icons appear on the freeway map when the queue identification and tracking algorithm identifies a congestion event. The a posteriori incident probability, as computed by a user-selected incident detection algorithm, is shown along with the icon.

The user is then allowed to either confii or deny the existence of the incident by using some independent means (e.g., closed-circuit television, police and motorist reports, etc.). As described in Section 10.2.2, the event information, including the operator's assessment of the actual traffic conditions, is recorded by the software for subsequent off-line analysis. In this manner, accurate performance evaluations may be conducted for the selected incident detection algorithms.

13.0 TECHNIQUES FOR EXPLOITING MOTORIST REPORTS

Cellular phones are increasingly being used to report freeway incidents, and a desirable feature of any new incident detection system is the ability to utilize this emerging source of traffic information. We note two distinct approaches to incorporating motorist reports with surveillance-based traffic data.

The first approach can be viewed as "fusing" received motorist information with surveillance-based incident detection results. In this approach, the motorist report is used to create a multi-class probability estimate of incident occurrence of the same form as generated from surveillance-based algorithms. The two estimates can then be combined according to relevant techniques from probability theory. An algorithm that utilizes this approach is described in Section 13.1.

The second approach is to use the information received from freeway motorists as direct inputs to the incident detection logic. In essence, reception of a motorist report would signal the algorithm that an incident was reported, and surveillance data would then be used to verify the existence of the incident and to isolate it's exact location. This approach was pursued using data from the Twin Cities freeway system. The developed algorithm, as well as a preliminary assessment of it's performance, are presented in Section 13.2.

13.1 The Fusion Method

In this approach, received motorist information is used to generate a multi-class probability estimate of incident occurrence which is of the same format as that output from the DCo classifier of the operational algorithm (see Section 4.2). Once this is accomplished, the objective is to "fuse" the motorist-based probability estimate with the corresponding surveillance-based probability estimate. The process of fusing the two estimates is described in detail below.

13.1.1 Derivation of the Fused Estimate

Denote the surveillance-based probability estimate by $\{ P[k/s], k = 1, N\}$ for surveillance data S and traffic category k. Similarly, denote the motorist-based probability estimate by $\{P[k/M], k=1,N\}$ for motorist information M and traffic category k. The goal of the

fusion process is to determine the combined probability $\{P[k/S,M],\ k=1,N\}$. This will be accomplished as described below.

The following relation can be taken as the definition of conditional probability for arbitrary events A and B:

$$P[A,B]=P[A/B]P[B]=P[B/A]P[A] \tag{13.1}$$

Hence, the combined probability $P[k/S,M]$ can be expressed as follows:

$$P[k/S,M]=\frac{P[k,S,M]}{P[S,M]}=\frac{P[S,M/k]P[k]}{P[S,M]} \tag{13.2}$$

If we make the reasonable assumption that S is independent of M, $P[S,M/k]$ can be factored as:

$$P[S,M/k]=P[S/k]P[M/k]$$

and we have:

$$P[k/S,M]=\frac{P[S/k]P[M/k]P[k]}{P[S,M]} \tag{13.3}$$

Again utilizing the definition of conditional probability, we can write:

$$P[S/k]=\frac{P[S,k]}{P[k]}=\frac{P[k/S]P[S]}{P[k]}$$

and similarly:

$$P[M/k]=\frac{P[M,k]}{P[k]}=\frac{P[k/M]P[M]}{P[k]}$$

Therefore, equation (13.3) can be written as:

$$P[k/S,M] = \frac{P[k/S]P[S]}{P[k]} \; \frac{P[k/M]P[M]}{P[k]} \; \frac{P[k]}{P[S,M]}$$

or equivalently:

$$P[k/S,M] = \frac{P[k/S]P[k/M]}{P[k]} \; \frac{P[S]P[M]}{P[S,M]} \tag{13.4}$$

It should be noted at this point that all of the quantities in the first term of equation (13.4) are known, since the value of $P[k]$ for any given freeway segment may be computed according to the methods developed in this project [SULL 94b]. Hence the objective is to determine an expression for the second term in equation (13.4). This may be accomplished as follows.

Since the summation of $P[k/S,M]$ over all values of k must equal unity, we have the following expression:

$$\sum_k P[k/S,M] = \sum_k \frac{P[k/S]P[k/M]}{P[k]} \; \frac{P[S]P[M]}{P[S,M]} = 1$$

or equivalently:

$$\frac{P[S]P[M]}{P[S,M]} \sum_k \frac{P[k/S]P[k/M]}{P[k]} = 1 \tag{13.5}$$

Solving for the unknown term and inserting into equation (13.4), we obtain the desired result:

$$P[k/S,M] = \frac{P[k/S]P[k/M]}{P[k]} \; \frac{1}{\sum_k \frac{P[k/S]P[k/M]}{P[k]}} \tag{13.6}$$

13.1.2 Interpretation of the Result

Equation (13.6) provides a practical means of combining motorist information with surveillance-based incident detection results. Certain observations can be made regarding the values taken by the variables involved in this relation. First, for any given section of freeway, an incident condition is an exceedingly rare event. Hence, $P[k]$ will generally be very small for incident categories and approximately unity for the incident-free category. Second, the value of $P[k/M]$ will be determined according to the source of the motorist report, and incident reports from police agencies and freeway service patrols will be viewed with significantly higher confidence than reports obtained from arbitrary motorists.

Perhaps the most important consequence of these findings is that a high surveillance-based incident probability is not required in order to verify a motorist-based incident report. This is best illustrated through example. Consider the case where we are interested only in determining whether the traffic condition on a given freeway segment is "incident" (I) or "incident-free" (IF). We could then assign the following nominal values to the two-class probability estimates in the case where an incident both exists and is reported by an arbitrary motorist.

$$P[I]=10^{-6} \qquad P[I/S]=0.01 \qquad P[I/M]=0.10$$
$$P[IF]=1-10^{-6} \qquad P[IF/S]=0.99 \qquad P[IF/M]=0.90$$

The low value of $P[I/S]$ indicates that the available surveillance data is insufficient, when viewed alone, to conclude that an incident condition is likely. However, when this information is combined with the relatively modest probability of an incident as determined from the motorist report, it will be seen that there is overwhelming evidence that an incident exists. Substituting the above values into equation (13.6), we obtain:

$$P[I/S,M]=\frac{(0.01)(0.10)}{(10^{-6})}\;\frac{1}{\left[\dfrac{(0.01)(0.10)}{(10^{-6})}+\dfrac{(.99)(.90)}{(1-10^{-6})}\right]}$$

At this point, it can be seen from inspection that the value of $P[I/S,M]$ will be approximately equal to unity since the value of $\dfrac{(.99)(.90)}{(1-10^{-6})}$ is negligible when compared to

the value of $\frac{(0.01)(0.10)}{(10^{-6})}$. This assertion is verified by carrying out the calculations. We have:

$$P[I/S,M] = 100.0\frac{1}{[100.0+0.891]} = 0.991$$

Although this result may be surprising at first, further consideration reveals that it is indeed reasonable. Specifically, the values of $P[I/S]$ and $P[I/M]$ constitute two independent sources of information which indicate that the probability of a given event (e.g., an incident) is significantly higher that the event's apriori probability, $P[I]$. Since an incident event is very rare, the values of $P[I/S]$ and $P[I/M]$ need not be high in order to conclude that the existence of an incident is very probable.

One should not conclude, however, that the combined probability will always indicate the existence of an incident. For instance, if the value of $P[I/S]$ is approximately equal to the value of $P[I]$ (e.g., the surveillance data reveals no indication of an incident), the combined probability will be approximately equal to $P[I/M]$. For such a scenario we have:

$$P[I/S,M] \approx P[I/M]\ \frac{1}{[P[I/M]+P[IF/M]]} = P[I/M]$$

These findings indicate that surveillance-based incident detection algorithms hold considerable potential toward validating motorist-based incident reports. Furthermore, such algorithms need not attempt to independently detect the suspected incident condition, but need only provide a meaningful probability of incident occurrence relative to the incident's apriori probability. Therefore, single-station algorithms may be effectively employed for the purpose of verifying motorist reports, and the fast response times exhibited by these algorithms can be utilized in order to verify such reports with negligible delays.

13.2 The Cueing Method

The approach presented in this section utilizes received motorist information to "cue" special processing which examines surveillance data in the neighborhood of the reported incident location and attempts to isolate the exact location of the incident.

The incident database obtained from the Twin Cities data collection effort was used as the basis for developing algorithms of this type. This database is described in detail in Section 10.2. All incidents in this database created significant upstream congestion, resulting in a head-of-queue condition developing in the incident zone. Hence, if one waits a sufficient amount of time after receiving the motorist report, the incident zone can be distinguished from it's neighboring zones by identifying the head-of-queue. Unfortunately, the head-of-queue condition frequently takes several minutes to develop, particularly for long zones.

Since motorist reports are expected to arrive in a timely manner, it is desirable to develop algorithms that are capable of distinguishing the incident zone prior to the onset of the head-of-queue condition. Development of algorithms for this purpose is addressed in the following subsection. Preliminary evaluations of algorithm performance are presented in Section 13.2.2.

13.2.1 Algorithm Development

Our approach to development was as follows. For each of the incidents in our database, data was extracted for the incident zone as well as the two adjacent downstream zones and the two adjacent upstream zones. The temporal scope of the data was several minutes prior to the onset of the head of queue condition caused by the incident. This data was then examined to determine if the incident zone could be distinguished from the surrounding zones prior to the onset of the head of queue.

Specifically addressed were the existence of various theoretical effects, including: (1) an increase in downstream speed, and (2) a decrease in downstream volume. Theoretical effects at the downstream end of the zone were considered due to the expectation that algorithms utilizing these effects would be relatively fast. Unfortunately, inspection of the data indicated that these effects are not effective discriminators for identifying the incident zone for the incident cases in the Twin Cities database.

For all incidents in the Twin Cities database, downstream traffic was freely flowing prior to the onset of the incident condition (this database was created by labeling the onset of new congestion events - see 10.2.2). Consequently, no increase in downstream speed at the time of incident onset is observable for this data set. Regarding the expected decrease in downstream volume at the time of incident onset, use of this feature is precluded due to the excessive measurement noise associated with the volume data.

Inspection of the data did reveal, however, that a significant speed drop at the station immediately upstream of the incident is commonly observed prior to the onset of the head of queue condition. Furthermore, such a drop in speed is generally absent for the neighboring stations prior to the onset of the head-of-queue condition. A possible explanation for this effect is that motorists approaching the incident zone observe congestion forming downstream and begin to slow down in anticipation of entering the congested region.

For the incidents in the Twin Cities database, the congestion will move upstream and eventually engulf the upstream station, resulting in a head-of-queue condition in the incident zone. Recall, however, that our objective is to identify the incident zone prior to the onset of the head-of-queue condition. Having noted that upstream speed decreases significantly before the head-of-queue develops, the following algorithm was postulated.

In order to detect a drop in speed at a given station, a five-minute moving average speed is defined as:

$$S_{MA}(t) = \frac{1}{10}\sum_{k=t-9}^{t} S(k)$$

where $S(k)$ is a filtered value of 30-second speed (refer to Appendix E regarding explicit feature definitions). A temporal drop in speed at a given station can then be characterized as:

$$\Delta S(t) = S(t) - S_{MA}(t-m),$$

where m is a parameter taking positive values. In evaluating algorithm performance, various values of m were investigated for possible performance benefits. Figure 13-1

shows the values of S, S_{MA} and ΔS for an actual station in the Twin Cities freeway system.

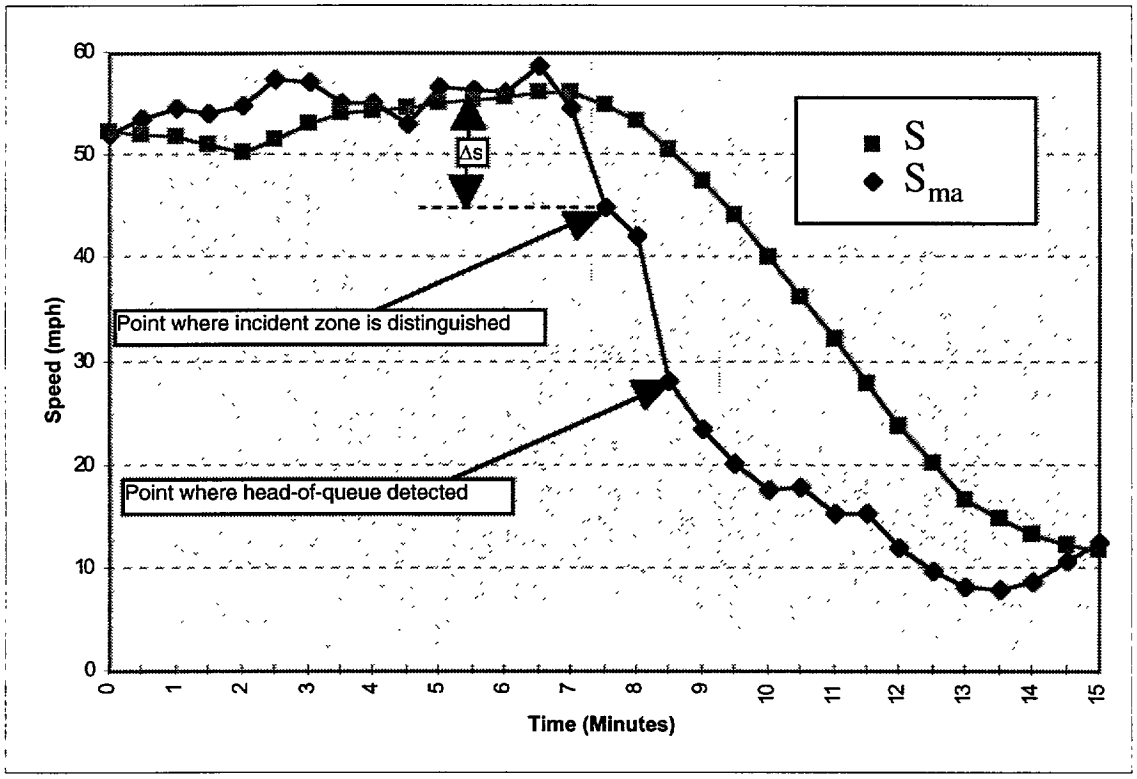


Figure 13-1. Features Used to Detect Temporal Speed Drop

The motorist report algorithm localizes the position of the reported incident by identifying the zone that first exhibits a value of ΔS greater than some threshold, T. Specifically, the algorithm is as follows:

(1) Upon receiving a motorist report, gather data from the appropriate zones surrounding the reported incident location and compute ΔS for values of time five minutes into the past.

(2) Starting at the first time (five minutes prior to receiving the motorist report in our case) and moving to the time that the report was received, identify the zone, if any, that first exhibits a value of ΔS greater than the detection threshold T. This is the incident zone. In the event that two zones simultaneously exhibit a value of ΔS greater than the detection threshold, the zone with the larger value will be declared as the incident zone.

(3) If no zone exhibits sufficiently large values of ΔS prior to receiving the motorist report, continue monitoring the zones in question as time proceeds.

This process is illustrated in Figure 13-2 for a detection threshold of ten miles-per-hour (sixteen -per hour). The data shown in this figure was taken from an actual incident in the Twin Cities database.

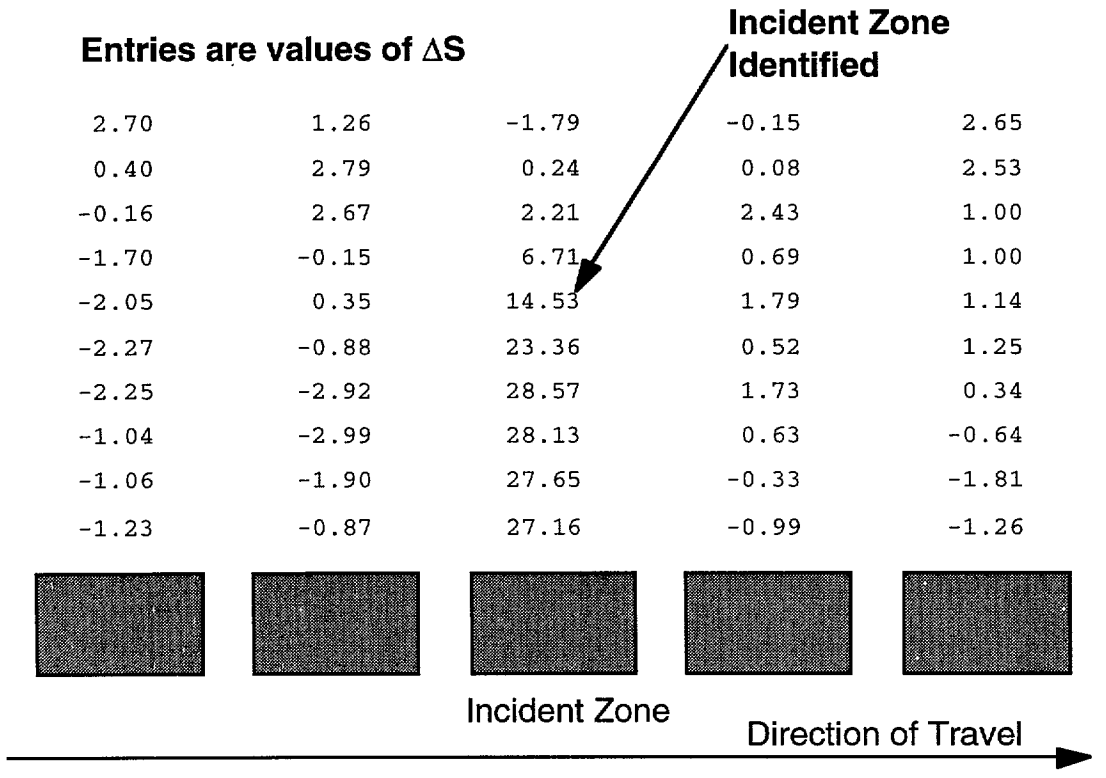


Figure 13-2. Identification of the Incident Zone

Other considerations can be taken into account depending on the time that the motorist report is received. For instance, if one of the zones in question already exhibits a head-of-queue condition at the time that the motorist report is received, there would be no need to exercise the above algorithm. Similarly, if the algorithm identifies a certain zone as containing the incident, and a head-of-queue condition subsequently develops in a different zone, the original diagnosis should be overridden.

It should also be noted that since this is a single-station algorithm, we don't really know if the incident is upstream or downstream of the station - just that it is in the vicinity of the

the incident in a two-zone region, and localization of the incident into a single zone must wait until the head-of-queue condition develops.

13.2.2 Algorithm Evaluation

Since the motorist report algorithm was developed heuristically (e.g., no training was involved), there is no need to employ distinct training and evaluation sets, as required in evaluating the onset detection algorithms. In order to assess algorithm performance, the algorithm was implemented in the C++ programming language and executed for 43 incident scenarios that were identified in the Twin Cities data collection effort.

For each incident scenario (e.g., group of zones), the algorithm was executed using various values of m and T. Results were compiled for each scenario, and it was noted whether the correct incident zone was identified, as well as the time savings achieved relative to the time of onset of the head-of-queue.

These data were then averaged over the 43 incidents in order to estimate the algorithm's detection rate (percent of time that the incident zone was identified correctly) and false alarm rate (percent of time that the incident zone was identified incorrectly) for each threshold value. These measures of effectiveness are shown in Figure 13-3 for a value of m corresponding to two minutes (this value of m was found to be optimal).

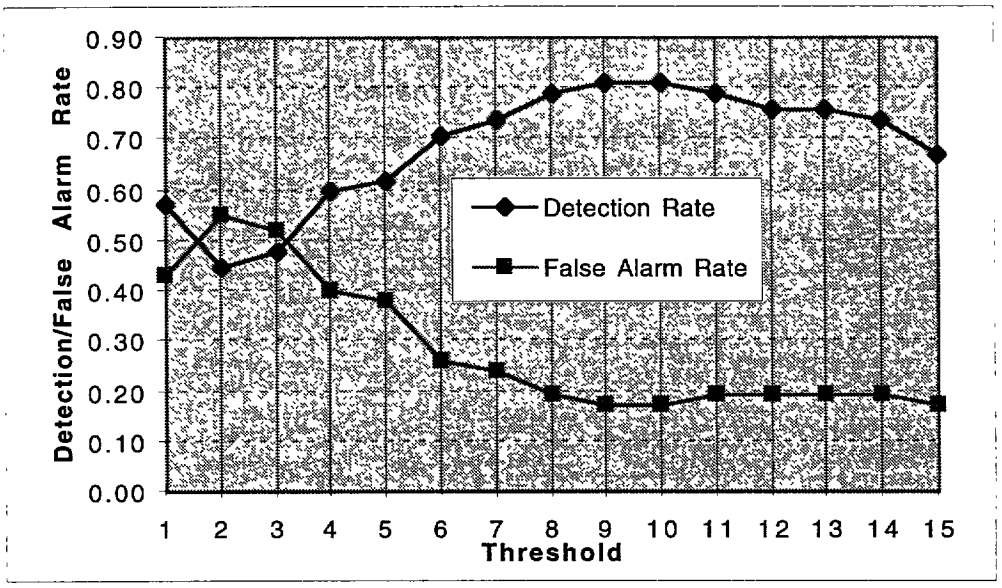


Figure 13-3. Evaluation of the Motorist Report Algorithm

As one can see, the detection rate is approximately 80% for a value of ***T*** equaling ten miles-per-hour (sixteen km-per-hour). The associated average time savings for this value of ***T*** was found to be approximately 2.5 minutes prior to the onset of the head-of-queue condition.