U.S. Department
of Transportation

**Federal Highway
Administration**

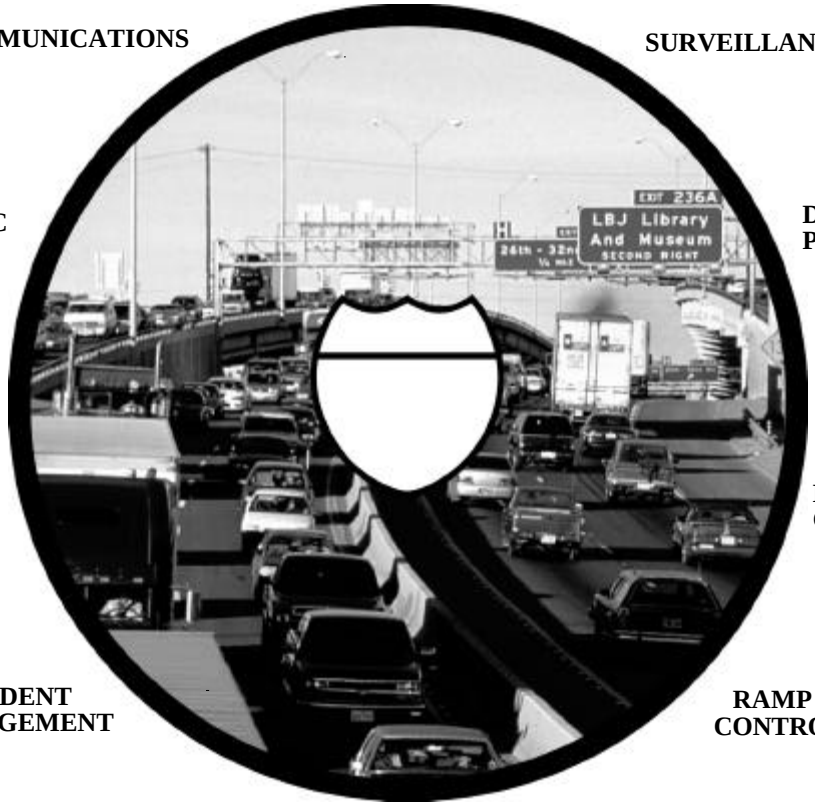
FREEWAY MANAGEMENT HANDBOOK

FREEWAY MANAGEMENT CONCEPTS

COMMUNICATIONS

SURVEILLANCE

ECONOMIC ANALYSIS



DECISION PROCESS

CONTROL CENTERS

LANE USE CONTROL

INCIDENT MANAGEMENT

RAMP CONTROL

INFORMATION DISSEMINATION

HOV CONCEPTS

Report No. FHWA-SA-97-064

August 1997

1. Report No. FHWA-SA-97-064	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle FREEWAY MANAGEMENT HANDBOOK		5. Report Date August 1997 6. Performing Organization Code 472810
Author(s) James D. Carvell, Jr., Kevin Balke, Jerry Ullman, Katherine Fitzpatrick, Lewis Nowlin, Christopher Brehmer		8. Performing Organization Report No. 472810-F
9. Performing Organization Name and Address Texas Transportation Institute		10. Work Unit No. (TRAIS) 11. Contract or Grant No. FHWA/TxDOT-96/591-1F
12. Sponsoring Agency Name and Address Department of Transportation Federal Highway Administration		13. Type of Report and Period Covered Draft: September 1, 1995-August 30, 1997 14. Sponsoring Agency Code

15. Supplementary Notes
FHWA Contracting Officer's Technical Representative(COTR): Morris Oliver (HTA-30)

Abstract:
This handbook, 1997 *Freeway Management Handbook*, is an update of the 1983 *Freeway Management Handbook* and reflects the tremendous developments in computing and communications technology. It also reflects the importance of *Integrated* Transportation Management Systems and the development of the concept of Intelligent Transportation Systems (ITS). The handbook development began with a survey of current practice, including site visits and interviews, of ten Freeway Management Systems throughout the country. It was developed under the advice of a panel of freeway management practitioners' panel.

The 1997 *Freeway Management Handbook* is organized in modular fashion with each module addressing a particular aspect or technology of the freeway management task. The modules are stand-alone treatments of particular areas of freeway management but are cross-referenced to reflect their interdependence. Each module is organized as follows:
INTRODUCTION- Including Module Objective and Scope.
DECISION PROCESS- Partners and Consensus Building, Establishing Goals and Objectives, Performance Criteria, Functional. Requirements, System Architectures, Identification and Screening of Technologies, and Implementation.
TECHNIQUES AND TECHNOLOGIES - Applications specific to the module.
LESSONS LEARNED - Experiences and observations from operating systems.
REFERENCES - Comprehensive list of references used in module preparation.

Specific modules are as follows:

INTRODUCTION AND EXECUTIVE SUMMARY
1. FREEWAY MANAGEMENT CONCEPTS
2. DECISION PROCESS
3. SURVEILLANCE
4. LANE CONTROL
5. RAMP CONTROL
6. HOV CONCEPTS
7. INFORMATION DISSEMINATION
8. INCIDENT MANAGEMENT
9. COMMUNICATIONS
10. CONTROL CENTERS
11. ECONOMIC ANALYSIS

17. Key Words	18. Distribution Statement
freeway management systems, freeway operations, ramp control, incident management, surveillance, information dissemination, communications, HOV, lane use control, ITS	No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	628	

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

ACKNOWLEDGMENTS

The *1997 Freeway Operations Handbook* was developed by Texas Transportation Institute with the guidance and assistance of numerous individuals and agencies. The authors would like to specifically acknowledge the following groups and individuals who made significant contributions to the development of the handbook.

Federal Highway Administration
Office of Technology Applications - development of Handbook requirements, draft reviews, and contract management,

- Mr. Morris Oliver, COTR

Office of Acquistion Management - contractual negotiations and administration.

- Ms. Lorraine Johnson

Practitioners' Panel
This panel consisted of persons who have been actively engaged in the practice of freeway management for a number of years. They provided reviews of the State of the Practice Document, the detailed outline for development of the handbook, and a review of the final draft handbook. The panel consisted of the following individuals:

- Mr. Pete Briglia, ITS Program Manager, Washington State Department of Transportation
- Mr. Glen Carlson, Manager, Traffic Management Center, Minnesota Department of Transportation
- Mr. Pat Irwin, Director of Transportation Operations, Texas Department of Transportation
- Mr. Stephen Leung, Chief, Office of Traffic Investigations, State of California Department of Transportation
- Mr. Joe McDermott, Chief, Bureau of Traffic, District 01, Illinois Department of Transportation
- Mr. Tom Werner, Director, Traffic Engineering and Safety Division, State of New York Department of Transportation

Texas Transportation Institute Internal Review Panel
This panel of TTI researchers provided periodic review and guidance during the development of the handbook in their specific fields of expertise. The panel included the following individuals:

- Mr. Charles Blumentritt, Research Scientist
- Dr. Conrad Dudek, Research Engineer
- Mr. William McCasland, Research Engineer
- Dr. Ed Seymour, Research Engineer
- Dr. Katie Turnbull, Research Scientist
- Dr. Thomas Urbanik II, Research Engineer
- Ms. Carol H. Walters, Research Engineer

Site Visit Facilitators

As part of the development of the State of the Practice Report, the following freeway management systems were visited by the project team. These site visits were facilitated or hosted by the individuals named below.

Baltimore, Maryland

- Mr. Lee McMichael, SOC Team Manager, Maryland State Highway Administration

Phoenix, Arizona

- Mr. Jim Shea, P.E., Manager, Intermodal Transportation Division,, Arizona Department of Transportation

Atlanta, Georgia

- Mr. Bert Riddle, Assistant State Traffic Engineer, Georgia Department of Transportation
- Mr. Joe Stapleton, Assistant State Traffic Engineer, Georgia Department of Transportation

Boston, Massachusetts

- Mr. Michael J. Costa, Director, ITS Programs, Massachusetts Highway Department
- Mr. Sergiu. F. Luchian, Project Engineer, Central Artery/Tunnel, Massachusetts Highway Department

Northern Virginia

- Mr. Jimmy Chu, Transportation Center Manager, Virginia Department of Transporatation
- Mr. Pete Todd, Incident Management Team Leader, Virginia Department of Transportation

Detroit, Michigan

- Mr. Ross Bremer, Operations Manager, Michigan Intelligent Transportation Systems Center
- Mr. Ray Klucens, Operations Engineer, Michigan Intelligent Transportation Systems Center

Seattle, Washington

- Mr. Mark Leth, Transportation Management Center, Washington State Department of Transportation
- Mr. Mark Bandy, Transportation Management Center, Washington State Department of Transportation

San Francisco, California

- Mr. Jim McCrank, Traffic Systems Chief, CALTRANS District 4

New Jersey Turnpike Authority

- Mr. Robert W. Dale, P.E., Director of Operations, New Jersey Turnpike Authority
- Mr. Leo Jackson, Manager, Traffic Operations Center, New Jersey Turnpike Authority

San Antonio, Texas

- Mr. Pat Irwin, District Traffic Engineer, Texas Department of Transportation
- Mr. Mark Olson, Federal Highway Administration
- Mr. Pat McGowan, Traffic Management Engineer, Texas Department of Transportation

Houston, Texas

- Mr. Carlton Allen, Freeway Operations Supervisor, Texas Department of Transportation
- Mr. Doug Wiersig, Executive Director, Traffic Management Center, City of Houston

TRANSCOM

- Mr. Tom Batz, Manager of Technology Development, TRANSCOM
- Mr. K. R. Marshall, Manager of Traffic Engineering, Farradyne Systems, consultant to the TRANSMIT project
- Mr. Bernie Wagenblast, Operations Manager, TRANSCOM

I-95 Coalition

- Mr. Dave Gehr, Chair, I-95 Corridor Coalition Executive Board and Administrator, Virginia Department of Transportation
- Mr. Stephen R. Kuciemba, Co-Chair, Smart Routes, Boston, Massachusetts

Previous Handbook Developers

Previous freeway management related handbooks facilitated and funded by the Federal Highway Administration provided an excellent baseline for the development of the *1997 Freeway Management Handbook* and the authors would like to acknowledge those handbooks and the performing organizations that developed them.

- ***1979 Traffic Control Systems Handbook*** - Pinnell, Anderson, and Wilshire and Assoc**.**
- ***1983 Freeway Management Handbook*** -Allen M. Voorhees, and Assoc.
- ***1985 Traffic Control Systems Handbook*** - Pinnell, Anderson, and Wilshire and Assoc.
- ***1991 Freeway Incident Management Handbook*** - Dunn Engineering and Assoc**.**
- ***1993 Communications Handbook for Traffic Control Systems*** - Dunn Engineering and Assoc.
- ***1996 Traffic Control Systems Handbook*** - Dunn Engineering and Assoc.

Page Down
to
The Table of Contents

TABLE OF CONTENTS

LIST OF FIGURES xxix

LIST OF TABLES xxxiii

EXECUTIVE SUMMARY xxxvii

MODULE 1. FREEWAY MANAGEMENT CONCEPTS

1.1 INTRODUCTION 1-3
CONGESTION 1-4
Measuring Congestion 1-5
Causes of Congestion 1-5
Geometric Design 1-7
Traffic Operations 1-7
Incidents 1-7
Maintenance and Construction 1-7
Weather 1-7

1.2 FREEWAY MANAGEMENT 1-9
WHAT IS FREEWAY MANAGEMENT? 1-9
WHAT IS A FREEWAY MANAGEMENT SYSTEM? 1-10
OBJECTIVES OF A FREEWAY MANAGEMENT SYSTEM 1-10
HISTORY OF FREEWAY MANAGEMENT 1-11
Functions of Freeway Management 1-11
Surveillance and Incident Detection 1-12
Lane Use Control 1-13
Ramp Control 1-13
High-Occupancy Vehicle (HOV) Priority Treatments 1-13
Information Dissemination 1-13
Freeway Management System 1-13
Incident Management 1-13
Elements of a Freeway Management System 1-14
Field Elements 1-14
Communications Elements 1-14
Control Center Elements 1-14
Toll Roads Versus Freeways 1-15
Relation to National ITS Architecture 1-15
Freeway Management 1-16
Incident Management 1-16
Integration with Other ITS Components 1-16

OBJECTIVE OF THE HANDBOOK 1-18
SCOPE OF THE HANDBOOK 1-18
STRUCTURE OF THE HANDBOOK 1-19
Module 1. Freeway Management Concepts 1-19
Module 2. Decision Process 1-19
Module 3. Surveillance 1-20
Module 4. Lane Use Control 1-20
Module 5. Ramp Control 1-20
Module 6. HOV Treatments 1-20
Module 7. Information Dissemination 1-20
Module 8. Incident Management 1-20
Module 9. Communications 1-21
Module 10. Control Center 1-21
Module 11. Economic Analysis 1-21

1.3 REFERENCES 1-23

MODULE 2. DECISION PROCESS

2.1 INTRODUCTION 2-3
MODULE OBJECTIVES 2-4
MODULE SCOPE 2-4

2.2 DESIGN PROCESS 2-4
DEFINE PROBLEM 2-6
ESTABLISH INSTITUTIONAL FRAMEWORKS AND BUILD COALITIONS . . . 2-7
ESTABLISH SYSTEM GOALS AND OBJECTIVES 2-9
ESTABLISH PERFORMANCE CRITERIA 2-10
DEFINE FUNCTIONAL REQUIREMENTS 2-12
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 2-13
IDENTIFY AND SCREEN TECHNOLOGIES 2-13
DEVELOP IMPLEMENTATION PLAN 2-15
DEPLOY PROJECTS 2-16
EVALUATION 2-18

2.3 TECHNIQUES AND TECHNOLOGIES 2-20
PERFORMANCE MEASURES 2-20
Systemwide 2-21
Total Travel Time 2-21
Total Travel 2-22
Vehicle Delay 2-22
Total Minute-Miles of Congestion 2-22
Isolated Locations 2-23

METHODS TO QUANTIFY OPERATIONAL PROBLEMS ON FREEWAYS . . . 2-23
Freeway Capacity Analysis and LOS 2-23
Definition of Freeway Subareas for Capacity Analysis 2-23
Basic Freeway Segments 2-24
Definition of Ideal Freeway Capacity 2-24
Factors Affecting Capacity 2-24
Level of Service 2-25
Weaving Areas 2-26
Definition of Weaving Area Capacity 2-26
Types of Weaving Areas 2-26
Operational Characteristics of Weaving Areas 2-26
Capacity and Level of Service for Weaving Areas 2-28
Ramp Junctions 2-28
Definition of Ramp Junction Capacity 2-28
Operational Characteristics of Ramp Influence Areas. 2-31
Level of Service of Freeway-Ramp Junctions. 2-31
Freeway Systems 2-31
Highway Capacity Software (HCS) 2-33
BOTTLENECK/QUEUING ANALYSES 2-34
FREEWAY SIMULATION/ OPTIMIZATION MODELS 2-37
Macroscopic Freeway Simulation/Optimization Models 2-38
FREQ 2-39
CORFLO 2-40
QUEWZ-92 2-41
Microscopic Freeway Simulation Models 2-41
FRESIM 2-42
INTEGRATION 2-43
Strengths and Weaknesses of the Available Models 2-44
FIELD MEASUREMENTS 2-44
Travel Time and Delay Studies 2-44
Origin/Destination Volume Studies 2-46

2.4 LESSONS LEARNED 2-46
SYSTEM MAINTENANCE 2-46
Types of Maintenance Requirements 2-47
Functional 2-47
Hardware 2-48
Software 2-48
MARKETING SYSTEMS 2-51
Target Audiences 2-51
Operating Agency 2-51
Affected Agencies 2-51
Elected Officials 2-52
General Public 2-52
Marketing Strategies 2-53

2.5 REFERENCES 2-55

APPENDIX A: SUBCHAPTER G - ENGINEERING AND TRAFFIC OPERATIONS 2-57

MODULE 3. SURVEILLANCE

3.1 INTRODUCTION 3-5
MODULE OBJECTIVES 3-5
MODULE SCOPE 3-6

3.2 DESIGN PROCESS 3-6
PROBLEM IDENTIFICATION 3-6
Identify and Locate Operational Problems 3-6
Determine Functions of Surveillance System 3-7
Type of Data 3-8
Traffic Measures. 3-8
Real-Time and Historical Data 3-9
Importance of Data 3-9
Inventory Existing Surveillance Capabilities 3-10
IDENTIFICATION OF PARTNERS 3-10
Intra-agency 3-10
Interagency (Information Exchange) 3-11
Additional Resources 3-12
ESTABLISH GOALS AND OBJECTIVES 3-12
ESTABLISH PERFORMANCE CRITERIA 3-13
DEFINE FUNCTIONAL REQUIREMENTS 3-13
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 3-14
IDENTIFY AND SCREEN TECHNOLOGIES 3-15
Identify Alternative Technologies 3-15
Evaluate Alternative Technologies 3-16
Select Appropriate Technology 3-16
PLAN DEVELOPMENT 3-17
Design Plans 3-17
Specifications 3-17
IDENTIFY FUNDING SOURCES 3-17
IMPLEMENTATION 3-19
EVALUATION 3-19
Monitoring Traffic Operations 3-20
Detecting Incidents 3-20
Implementing Control Strategies 3-20
Monitoring Environmental Conditions 3-21

3.3 TECHNIQUES AND TECHNOLOGIES 3-21
SYSTEM COMPONENTS 3-21
Detection Methods 3-22
Hardware 3-24
Software 3-25
Incident Detection Algorithms 3-25
Real-Time Expert System 3-25
Communications 3-26
SURVEILLANCE TECHNOLOGIES 3-27
EMBEDDED DETECTORS 3-27
Inductive Loop Detector 3-28
Characteristics 3-28
Applications 3-28
Installation Requirements 3-32
Magnetometer 3-34
Characteristics 3-34
Applications 3-34
Installation Requirements 3-34
NON-INTRUSIVE DETECTORS 3-35
Microwave Radar 3-37
Characteristics 3-37
Continuous Wave. 3-37
Frequency Modulated Continuous Wave (FMCW) 3-37
Pulsing Waveform 3-37
Wide Beam. 3-38
Narrow Beam 3-38
Long-Range 3-38
Applications 3-39
Installation Requirements 3-39
Infrared 3-40
Characteristics 3-40
Active. 3-40
Passive 3-40
Applications 3-40
Active 3-40
Passive 3-42
Installation Requirements 3-42
Ultrasonic 3-44
Characteristics 3-44
Applications 3-44
Installation Requirements 3-45
Acoustic 3-45
Characteristics 3-45
Applications 3-45
Installation Requirements 3-45
Video Image Processing 3-47

Characteristics 3-47
Classes of VIP Systems 3-47
VIP Detector Performance 3-48
Advantages of VIP Systems 3-49
CCTV Applications 3-50
Applications 3-50
Installation Requirements 3-51
COMPARISON OF EMBEDDED AND NON-INTRUSIVE DETECTORS 3-51
VEHICLE PROBES 3-51
Types of Vehicle Probes 3-53
Automatic Vehicle Identification (AVI) 3-53
Active 3-53
Passive 3-54
Semi-Active 3-54
Type I 3-54
Type II 3-54
Type III 3-54
Automatic Vehicle Location (AVL) 3-54
Dead-Reckoning and Map-Matching 3-55
Signpost 3-55
Ground-Based Radio-Navigation 3-55
LORAN-C 3-56
Global Positioning System 3-56
Differential GPS 3-56
Cellular Telephones 3-56
MOBILE REPORTS 3-57
Cellular Telephones 3-57
Freeway Service Patrols 3-58
Call Boxes/Emergency Telephones 3-58
CLOSED-CIRCUIT TELEVISION (CCTV) 3-58
Fixed Location 3-59
Full Motion Video 3-60
Compressed Video 3-60
Portable 3-61
SUMMARY OF TRAFFIC SURVEILLANCE APPLICATIONS 3-61
ENVIRONMENTAL DETECTORS 3-62
Freeways 3-62
Tunnels 3-65

3.4 LESSONS LEARNED 3-65
OPERATIONS AND MAINTENANCE 3-65
Operations 3-65
Maintenance 3-66
Operations and Maintenance Costs 3-66
PRIVACY CONCERNS 3-67
SPACING AND PLACEMENT ISSUES 3-70

Mainline Traffic Detectors 3-70
Spacing 3-70
Placement 3-70
Non-Intrusive Detectors 3-71
CCTV Cameras 3-71
Site Selection Criteria 3-71
Preliminary Site Selection 3-72
Final Site Selection 3-72

3.5 EXAMPLES 3-73
MIAMI AUTOMATIC VEHICLE LOCATION (AVL) 3-73
TRANSCOM ELECTRONIC TOLL AND TRAFFIC MANAGEMENT (ETTM) . 3-74
HOUSTON AUTOMATIC VEHICLE IDENTIFICATION (AVI) 3-75
CONNECTICUT MICROWAVE RADAR 3-75

3.6 REFERENCES 3-77

MODULE 4. LANE USE CONTROL

4.1 INTRODUCTION 4-3
BACKGROUND 4-3
TREATMENTS 4-3
MODULE OBJECTIVES 4-4
MODULE SCOPE 4-4

4.2 DESIGN PROCESS 4-4
PROBLEM IDENTIFICATION 4-4
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 4-5
Potential Partners 4-5
Consensus Building 4-6
ESTABLISH GOALS AND OBJECTIVES FOR LANE USE CONTROL 4-6
ESTABLISH PERFORMANCE CRITERIA 4-6
DEFINE FUNCTIONAL REQUIREMENTS 4-7
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 4-7
IDENTIFY AND SCREEN TECHNOLOGY 4-8
PLAN DEVELOPMENT 4-9
FUNDING SOURCE IDENTIFICATION 4-9
IMPLEMENTATION 4-9
EVALUATION 4-10

4.3 TECHNIQUES AND TECHNOLOGIES 4-10
STATIC SIGNING 4-10
Truck Restrictions 4-10
HOV Restrictions 4-11

INFORMATION DISSEMINATION SYSTEMS 4-11
USE OF NARROW LANES AND SHOULDERS 4-15
MAINLINE METERING 4-15
FREEWAY-TO-FREEWAY RAMP METERING 4-16
CHANNELIZING DEVICES FOR WORK ZONES 4-17
TOLL FACILITIES/CONGESTION PRICING 4-17
AUTOMATED HIGHWAY SYSTEMS 4-19
Concept 4-20
Vehicle Characteristics 4-21
Roadway Infrastructure 4-21
Command and Control 4-21
Entry and Exit Infrastructure 4-21
Goals 4-22

4.4 LESSONS LEARNED 4-22
PLANNING 4-22
Freeway-to-Freeway Ramp Metering 4-22
Toll Facilities 4-23
Congestion Pricing 4-23
Channelizing Devices for Work Zones 4-24
DESIGN/CONSTRUCTION 4-26
LCS and DMS Placement 4-26
LCS Visibility 4-26
LCS Spacing and Mounting Locations 4-26
DMS Placement 4-26
Use of Narrow Lane and Shoulders 4-27

4.5 EXAMPLES 4-27
SAN ANTONIO (TRANSGUIDE) 4-27
Description of System 4-27
Effects of Information 4-28
SR 91, ORANGE COUNTY, CALIFORNIA CONGESTION PRICING 4-29

4.6 REFERENCES AND SUGGESTED READINGS 4-31
REFERENCES 4-31
SUGGESTED READINGS 4-34

MODULE 5. RAMP CONTROL

5.1 INTRODUCTION 5-5
DEFINITION OF RAMP CONTROL 5-5
APPLICATION OF RAMP CONTROL 5-6
Entrance Ramp Metering 5-6
Entrance Ramp Closure 5-6
Exit Ramp Closure 5-6

Systemwide Ramp Control 5-6
RELATION TO OTHER FREEWAY MANAGEMENT FUNCTIONS 5-6
Surveillance 5-7
Vehicle Detection 5-7
Closed-Circuit Television 5-7
Environmental Sensors 5-7
HOV Treatments 5-7
Information Dissemination 5-7
Communication 5-7
Control Center 5-7
BENEFITS OF RAMP CONTROL 5-8
Improved System Operation 5-8
Improved Safety 5-8
Reduced Vehicle Operating Expense 5-8
Means for Positive Freeway Traffic Control/Management 5-8
Reduction in Vehicle Emissions and Fossil Fuel Consumption 5-9
Coordination With Other Corridor Management Elements 5-9
Promotion of Multimodal Operation 5-9
MODULE OBJECTIVE 5-9
MODULE SCOPE 5-9

5.2 DECISION PROCESS 5-9
PROBLEM IDENTIFICATION 5-10
Level of Service / Capacity Deficiency / Bottlenecks 5-10
Vehicle Crash Experience 5-10
Inventory of Infrastructure 5-11
Freeway System 5-11
Existing Freeway Management Systems 5-11
Existing Ramp Control Systems 5-12
Other Relevant Field Systems 5-12
Inventory of Traffic Characteristics 5-12
Traffic Composition 5-12
Traffic Flow 5-12
Other Traffic Parameters 5-12
Temporal Variations 5-13
Ramp Geometric Limitations 5-13
Cross Streets 5-14
Service Roads 5-14
Summary of Problem Definition 5-14
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 5-14
Relation to Other Agencies 5-14
City/County Traffic Operations Agencies 5-14
Enforcement Agencies 5-14
Emergency Management Agencies 5-15
Public Transportation Agencies 5-15
Relationship to Elected Official / Political Environment 5-15

Importance of Enforcement / Judicial System 5-15
Relationship With Media 5-15
ESTABLISHING GOALS AND OBJECTIVES 5-16
ESTABLISH PERFORMANCE CRITERIA / MEASURES OF EFFECTIVENESS 5-16
DEFINE FUNCTIONAL REQUIREMENTS 5-17
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 5-17

5.3 TECHNIQUES AND TECNOLOGIES 5-18
ENTRANCE RAMP CONTROL 5-18
Ramp Closure 5-18
Application 5-19
Methods 5-19
Ramp Metering 5-20
Metering Rates 5-20
Congestion. 5-20
Safety. 5-22
Pretimed Metering 5-22
System Components 5-23
System Operation 5-24
Single-entry metering 5-24
Platoon metering 5-24
Tandem Metering. 5-24
Two-abreast Metering 5-25
Traffic-Responsive Metering 5-26
Fundamental Traffic Flow Relationships 5-26
Basic Strategy 5-26
Demand-Capacity Control 5-28
Occupancy Control 5-28
System Components 5-29
System Operation 5-30
Gap-Acceptance Merge Control 5-31
Basic Concepts 5-32
System Components 5-32
System Operation 5-33
System Ramp Control 5-35
System Pretimed Metering 5-35
Example No. 1 5-36
Example No. 2 5-38
Example No. 3 5-38
Practical Considerations 5-40
Example No. 4 5-41
System Traffic-Responsive Metering 5-43
System Operation 5-43
Metering Rates 5-43

System vs. Independent Ramp Control 5-43
Controller Interconnection 5-43
Incremental Benefits of Various Levels of Control 5-44
Variations in the Ratio of Mainline to Entrance Ramp Demand 5-44
Variations in Overall Traffic Demand Pattern 5-44
Mainline Capacity Reductions 5-44
EXIT RAMP CONTROL 5-45
EMERGING TECHNOLOGIES 5-45

5.4 LESSONS LEARNED 5-46
IMPLEMENTATION 5-47
Public Relations 5-47
Media Relations 5-47
Implementation Strategies 5-48
Implementation Summary 5-48
OPERATIONS AND MAINTENANCE 5-49
DIVERSION OF TRAFFIC 5-50
ENFORCEMENT 5-51
EQUITY 5-51

5.5 EXAMPLES IN RAMP CONTROL 5-52
HISTORY OF RAMP CONTROL 5-52
ENTRANCE RAMP METERING CASE STUDIES 5-52
Portland, Oregon 5-52
Minneapolis/St. Paul, Minnesota 5-53
Seattle, Washington 5-54
Denver, Colorado 5-55
Detroit, Michigan 5-56
Austin, Texas 5-56
Long Island, New York 5-56
San Diego, California 5-57
SUMMARY OF ENTRANCE RAMP METERING BENEFITS 5-57

5.6 REFERENCES 5-59

MODULE 6. HOV TREATMENTS

6.1 INTRODUCTION 6-5
TREATMENTS 6-6
MODULE OBJECTIVES 6-6
MODULE SCOPE 6-7

6.2 DESIGN PROCESS 6-7
IDENTIFY PROBLEMS/NEEDS 6-7
IDENTIFICATION OF PARTNERS 6-8

CONSENSUS BUILDING 6-10
ESTABLISH GOALS AND OBJECTIVES 6-10
ESTABLISH PERFORMANCE CRITERIA / MOES 6-10
DEFINE FUNCTIONAL REQUIREMENTS 6-11
IDENTIFY AND SCREEN TECHNOLOGY 6-11
System Planning 6-13
IMPLEMENTATION 6-15
EVALUATION 6-16

6.3 TECHNIQUES AND TECHNOLOGIES 6-18
HOV FACILITIES 6-18
Operational Considerations 6-18
HOV Roadway Operations 6-20
Operating Efficiency 6-20
Considerations for 2+ Versus 3+ Occupancy Requirement 6-20
Hours of Operations 6-22
Design Elements 6-22
Design Speed 6-22
Cross Section Width 6-22
Vertical Clearance 6-23
Signs and Markings 6-23
Signs 6-23
Markings 6-23
Enforcement 6-25
Goals and Objectives 6-25
Methods 6-27
Enforcement Area Design 6-28
ACCESS 6-29
Ramp Connections 6-29
Design 6-29
PRIORITY ACCESS 6-34
Ramp Metering 6-34
Bypass Lanes 6-34
Signs 6-35
Markings 6-35
PARK-AND-RIDE FACILITIES 6-35
Location of Facilities 6-36
Establishment of Site Selection Criteria 6-36
Identification of Alternative Sites 6-38
Evaluation of Alternatives 6-38
Determining Size of Facility 6-38
Typical Designs 6-39
Internal Lot Layout 6-39
Functional Area Design. 6-39
Internal Circulation. 6-40
Amenities. 6-40

Pavement and Drainage. 6-40
Landscaping 6-41
Lighting 6-41
Security 6-41
Environmental Considerations. 6-41
Roadway Interface 6-41
General Access and Egress Considerations 6-42
Access Points 6-42
Access Roadways 6-42
Traffic Control Devices and Traffic Signals 6-42
Handicapped Considerations 6-42
Parking 6-42
Capacity 6-42
Location 6-42
Design 6-43
Signing and Marking 6-43
Bus Loading Area 6-43
PARKING INCENTIVES 6-44
Preferential Parking 6-44
Free/Low-Cost Parking 6-44
Differential Parking Rates 6-44
Benefits of Parking Incentives 6-44
Barriers to Implementing Parking Incentives 6-45
EMERGING TECHNOLOGIES 6-46
Enforcement 6-46
Movable Barrier 6-46

6.4 LESSONS LEARNED 6-48
PLANNING 6-48
Public and Political Support and Consensus Building 6-50
Public Involvement 6-51
System Approach to HOV Planning 6-52
Possible Pitfalls in Implementation 6-52
Development Concerns for Park-and-Ride Facilities 6-53
Program Planning 6-53
Design Requirements 6-55
Operation and Maintenance 6-55
Program Marketing 6-56
DESIGN/CONSTRUCTION 6-57
Tradeoffs in Retrofitting Cross Sections 6-57
Impacts of Ingress/Egress on Adjacent Freeway and Arterial System 6-57
Signing Guidelines 6-59
Pavement Marking Guidelines 6-60
OPERATIONS 6-61
Surveillance, Communication, and Control 6-61
Incident Response 6-62

Enforcement Strategies 6-62
Other Users 6-63
MAINTENANCE 6-64
CONCURRENT FLOW LANES 6-65

6.5 EXAMPLES OF HOV TREATMENTS 6-66
HOUSTON 6-66
SEATTLE 6-71

6.6 REFERENCES AND SUGGESTED READINGS 6-75
REFERENCES 6-75
SUGGESTED READING 6-77

MODULE 7. INFORMATION DISSEMINATION

7.1 INTRODUCTION 7-3
MODULE OBJECTIVES 7-4
MODULE SCOPE 7-4

7.2 DESIGN PROCESS 7-4
IDENTIFY NEEDS 7-4
IDENTIFY INFORMATION DISSEMINATION PARTNERS 7-6
BUILD CONSENSUS AMONG PARTNERS 7-6
Identify Differences in Operating Philosophies Among Partners 7-6
Establish Common Ground Between Partners 7-7
ESTABLISH GOALS AND OBJECTIVES 7-7
ESTABLISH PERFORMANCE CRITERIA AND MEASURES 7-8
Information Credibility 7-8
Market Penetration 7-8
Traveler Response 7-9
FUNCTIONAL REQUIREMENTS 7-10
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 7-11
IDENTIFY AND SCREEN TECHNOLOGIES 7-11
DEVELOP IMPLEMENTATION PLAN 7-12
IDENTIFY FUNDING SOURCES 7-13
IMPLEMENT 7-14
EVALUATE 7-14

7.3 TECHNIQUES AND TECHNOLOGIES 7-15
ON-ROADWAY INFORMATION TECHNOLOGIES 7-15
Types of DMS 7-16
Light-Reflecting DMSs 7-17
Fold-out. 7-18

Scroll (or Tape). 7-18
Rotating Drum. 7-19
Reflective Disk Matrix. 7-19
Light-Emitting DMSs 7-19
Neon or Blank-Out Signs. 7-19
Lamp (Incandescent Bulb) Matrix Signs. 7-19
Fixed-Grid or Shuttered Matrix Fiber-optic Signs. 7-19
Fixed-Grid or Matrix Light-Emitting Diode (LED) Signs. 7-20
Hybrid DMSs 7-20
Reflective Disks with Fiberoptics/LEDs. 7-20
Static/DMS. 7-20
Advantages and Disadvantages of the DMS Technologies 7-20
IN-VEHICLE INFORMATION 7-21
Auditory In-Vehicle Information Technologies 7-21
Highway Advisory Radio 7-21
Automated HAR (AHAR) 7-24
Cellular Telephone "Hotlines" 7-25
Commercial Radio 7-26
Citizen-Band Radio 7-26
Visual In-vehicle Information Technologies 7-26
Video Display Terminals 7-26
Head-Up Displays 7-27
OFF-ROADWAY INFORMATION 7-27
Television 7-28
Telephones 7-28
Pagers 7-28
Personal Data Assistants (PDAs) 7-29
Computers (Internet) 7-29
Kiosks 7-29

7.4 LESSONS LEARNED 7-30
NTCIP STANDARDS 7-30
INFORMATION MESSAGE DESIGN 7-30
ENVIRONMENTAL ISSUES CONCERNING DMSs 7-31
Sun Position 7-31
Weathering 7-31
"AT-REST" DISPLAY CONDITIONS 7-31

7.5 EXAMPLE OF AN INFORMATION DISSEMINATION SYSTEM: TRANSCOM 7-31
ORGANIZATION 7-31
OPERATIONS 7-32
FUTURE ACTIVITIES 7-33

7.6 REFERENCES 7-35

MODULE 8. INCIDENT MANAGEMENT

8.1 INTRODUCTION 8-3
BACKGROUND 8-3
MODULE OBJECTIVES 8-4
MODULE SCOPE 8-4

8.2 DECISION PROCESS 8-4
IDENTIFY NEEDS 8-5
IDENTIFY INCIDENT MANAGEMENT PARTNERS 8-7
BUILD CONSENSUS AMONG PARTNERS 8-8
ESTABLISH INCIDENT MANAGEMENT GOALS AND OBJECTIVES 8-9
ESTABLISH PERFORMANCE CRITERIA 8-9
DEFINE FUNCTIONAL REQUIREMENTS 8-10
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 8-10
IDENTIFY AND SCREEN TECHNOLOGY 8-11
Incident Detection and Verification 8-11
Incident Response and Clearance 8-12
DEVELOP PLANS 8-13
IDENTIFY FUNDING SOURCES 8-14
IMPLEMENT 8-15
EVALUATE 8-15

8.3 TECHNIQUES AND TECHNOLOGIES 8-15
INCIDENT DETECTION AND VERIFICATION 8-15
Non-Automated Detection Techniques 8-15
Automated Surveillance Techniques 8-16
Technologies 8-16
Detection Algorithms 8-19
INCIDENT RESPONSE AND CLEARANCE 8-21
On-Site Response Techniques 8-21
Equipment 8-21
Identification Arm Bands and Vests 8-22
Incident Response and Hazardous Materials Manuals. 8-22
Total Stations Surveying Equipment. 8-22
Inflatable Air Bag Systems. 8-25
Equipment Storage Sites. 8-25
Tow Truck/Removal Crane Contracts. 8-25
Management Strategies/Operations 8-26
Emergency Response Vehicle Parking Coordination. 8-26
On-Site Traffic Control. 8-26
Incident Response Teams. 8-27
Vehicle Removal Laws. 8-28
Techniques to Improve Emergency Vehicle Access and Traffic Flow 8-28

Equipment 8-28
Barrier Openings. 8-28
Barrier Gates. 8-29
Emergency Ramps. 8-29
Management Strategies/Operations 8-29
Police Escorts and Wrong-Way Entrance of Non-Emergency Response Vehicles. 8-29
Shoulder Utilization. 8-30
Contraflow Diversion. 8-30
Accident Investigation Sites 8-31
Equipment 8-31
Management Strategies/Operations 8-31
Freeway Service Patrols 8-33
Equipment 8-33
Management Strategies/Operations 8-33
Electronic Traffic Management and Control Techniques 8-34
Equipment 8-34
Ramp Metering. 8-34
Information Dissemination Components. 8-34
Management Strategies/Operations 8-34
Alternative Route Diversion Techniques 8-34
Equipment 8-35
Management Strategies/Operations 8-37
Work Zone and Special Event Traffic Management 8-37
Equipment 8-37
Management Strategies/Operations 8-37

8.4 LESSONS LEARNED 8-37
WATER ACCESS 8-37
SERVICE PATROL CHARGES FOR ASSISTANCE 8-39
TRAINING 8-39

8.5 EXAMPLES OF FREEWAY SERVICE PATROLS 8-39
CHICAGO EMERGENCY TRAFFIC PATROL 8-39
MINNEAPOLIS HIGHWAY HELPER PROGRAM 8-40
LOS ANGELES FREEWAY SERVICE PATROL 8-40
HOUSTON MOTORIST ASSISTANCE PATROL 8-41
SAMARITANIA, INC. 8-41

8.6 REFERENCES 8-43

MODULE 9. COMMUNICATIONS

9.1 INTRODUCTION 9-3
MODULE OBJECTIVES 9-4
MODULE SCOPE 9-4

9.2 DECISION PROCESS 9-4
PROBLEM IDENTIFICATION 9-5
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 9-8
ESTABLISH SYSTEM GOALS AND OBJECTIVES 9-9
ESTABLISH PERFORMANCE CRITERIA 9-9
FUNCTIONAL REQUIREMENTS 9-10
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS,
AND INFORMATION FLOWS 9-12
Logical 9-12
Physical 9-12
Central 9-14
Distributed 9-16
Trunking 9-16
Backbone 9-16
Multimedia 9-18
IDENTIFY AND SCREEN TECHNOLOGIES 9-18
IMPLEMENTATION 9-23
National ITS Standards 9-23
Installation 9-26

9.3 TECHNIQUES AND TECHNOLOGIES 9-26
LAND-LINES 9-26
Twisted-Wire Pairs 9-27
Coaxial Cable 9-31
Fiber Optic 9-35
WIRELESS 9-39
Area Radio Network 9-39
Microwave 9-42
Spread Spectrum Radio 9-43
Cellular Radio 9-45
Packet Radio 9-48
Satellite Transmission 9-50

9.4 LESSONS LEARNED 9-51
DESIGN PRINCIPLES 9-51
OWNED VERSUS LEASED COMMUNICATIONS 9-53
NTCIP 9-56
SHARED RESOURCE PROJECTS 9-59

TRANSIENT PROTECTION 9-63
COMMUNICATIONS WITHIN THE CONTROL CENTER 9-64

9.5 REFERENCES 9-65

MODULE 10. CONTROL CENTER

10.1 INTRODUCTION 10-3
MODULE OBJECTIVE 10-5
MODULE SCOPE 10-5

10.2 DECISION PROCESS 10-5
PROBLEM IDENTIFICATION 10-5
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 10-7
ESTABLISH GOALS AND OBJECTIVES 10-7
DEFINE FUNCTIONAL REQUIREMENTS 10-8
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS,
AND INFORMATION FLOWS 10-11
Function Allocation 10-11
Task Analysis 10-13
Data Flows 10-15
IDENTIFY AND SCREEN TECHNOLOGIES 10-15
Physical Environment 10-15
Workstations 10-17
Controls and Displays 10-17
User Interfaces 10-18
EVALUATION 10-19

10.3 TECHNIQUES AND TECHNOLOGIES 10-20
PHYSICAL DESIGN 10-20
Lighting 10-24
Acoustics 10-25
Heating and Cooling 10-25
INFORMATION DISPLAYS 10-26
Static Wall Maps 10-26
Projection Television Systems 10-26
Video Walls 10-26
Design Guidelines and Recommendations 10-28
OPERATOR WORKSTATION DESIGN 10-28
COMMUNICATIONS SYSTEMS 10-29
Local Area Networks 10-30
Components 10-30
Planning for an LAN 10-31
Voice Communications 10-32

STAFFING 10-34
Qualifications 10-36
Training 10-36
Staffing Sources — Agency versus Outsourcing 10-38
SECURITY 10-38
AUTOMATION 10-39
JOINT OPERATIONS 10-40

10.4 LESSONS LEARNED 10-42
SYSTEM CAPABILITIES 10-42
OPERATIONS MANUAL 10-42
EFFECTIVE USE OF SYSTEM OPERATORS 10-42
SYSTEM MAINTENANCE 10-43

10.5 REFERENCES 10-45

MODULE 11. ECONOMIC ANALYSIS

11.1 INTRODUCTION 11-3
PURPOSE AND SCOPE 11-3
SYSTEM GOALS AND OBJECTIVES 11-4
INTERRELATIONSHIP OF SYSTEM ELEMENTS 11-4
STRUCTURE OF MODULE 11-4

11.2 ESTIMATING COSTS AND BENEFITS 11-4
SYSTEM COSTS 11-4
Types of Costs 11-4
Sources of Cost Information 11-5
SYSTEM BENEFITS 11-5
Types of Benefits 11-5
Sources of Benefit Information 11-6
IMPORTANCE OF LIFE CYCLE COST CONSIDERATIONS 11-6

11.3 SYSTEM EVALUATION 11-7
DEFINITION OF SYSTEM ALTERNATIVES/COMPONENTS 11-7
ASSIGNING BENEFITS TO ALTERNATIVES 11-8
ESTIMATING LIFE-CYCLE COSTS 11-9
ANALYSIS TECHNIQUES 11-9
Net Present Worth 11-9
Description 11-9
Example 11-9
Other Considerations 11-9
Benefit-Cost and Incremental Benefit-Cost Analysis 11-11
Description 11-11

Example 11-11
Other Considerations 11-12
Utility-Cost Analysis 11-13
Description 11-13
Example 11-13
Other Considerations 11-14
Value Engineering 11-14
Description 11-14
Example 11-14
Other Considerations 11-16
Sensitivity Analysis 11-17
Description 11-17
Example 11-17
Other Considerations 11-18

11. 4 REFERENCES 11-19
REFERENCES 11-19
SUGGESTED READING 11-20

GLOSSARY G-1

LIST OF FIGURES

Figure 1-1. I-35 Freeway in Austin, TX. 1-3
Figure 1-2. Congestion and the Fundamental Relationship Between Flow Rate and Density. 1-6
Figure 1-3. Generalized Traffic Flow Relationships. 1-6
Figure 1-4. Functions and Elements of a Freeway Management System. 1-12
Figure 1-5. Freeway Management as Part of an Integrated Intelligent Transportation System 1-17
Figure 2-1. Elements of a Freeway Management System. 2-3
Figure 2-2. Steps Performed in a Typical Systems Engineering Analysis. 2-5
Figure 2-3. Systems Engineering Process Used to Develop National ITS Architecture. 2-6
Figure 2-4. Example of System Goal and Objectives to Address an Identified Problem. 2-11
Figure 2-5. Elements of a Typical Implementation Plan. 2-17
Figure 2-6. Conceptual Illustration of Before-and-After Evaluation with Control Conditions. 2-20
Figure 2-7. Illustration of Freeway Components for Capacity Analysis. 2-25
Figure 2-8. Categories of Weaving Areas. 2-29
Figure 2-9. Influence Areas and Critical Flow Values for Freeway-Ramp Junctions. . 2-30
Figure 2-10. Graphic Representation of Level of Service for Freeway Systems. 2-34
Figure 2-11. Fundamentals of Shock Wave Analysis. 2-35
Figure 2-12. Shock Wave Phenomena at a Freeway Bottleneck During a Peak Period. 2-35
Figure 2-13. Use of Queuing Analysis to Analyze Temporary Blockage. 2-38
Figure 3-1. Closed-Circuit Television Camera 3-5
Figure 3-2. Control Center Components Typically Included in a Surveillance System 3-22
Figure 3-3. External System Overview for Real-Time Expert System 3-26
Figure 3-4. Inductive Loop Detectors Used for Freeway Surveillance 3-32
Figure 3-5. Typical Loop Installation for Measuring Speed 3-33
Figure 3-6. Typical Layout for Magnetometer Installation 3-35
Figure 3-7. Typical Bridge Deck Installation for Magnetometer 3-36
Figure 3-8. Example of Non-Intrusive Detector 3-36
Figure 3-9. Waveforms Used with Microwave Detectors 3-38
Figure 3-10. Alternative Radar Beam Widths 3-39
Figure 3-11. Detector in Side-Fire Position 3-41
Figure 3-12. Detector in Overhead Position 3-41
Figure 3-13. Active Infrared Detector 3-42
Figure 3-14. Passive Infrared Detector 3-43
Figure 3-15. Active Infrared Detector Measuring Vehicle Classification 3-43
Figure 3-16. Example of Ultrasonic Detector Installation 3-45
Figure 3-17. Acoustic Detector 3-46
Figure 3-18. Time Delay Versus Time Curve 3-46
Figure 3-19. Components of Video Image Processing System 3-48
Figure 3-20. Identification of Detection Lines in Field of View 3-51

Figure 3-21. Freeway Service Patrol Vehicle and Equipment 3-59
Figure 3-22. Call Box 3-60
Figure 3-23. Fixed-Location Closed-Circuit Television Camera 3-60
Figure 3-24. Portable Closed-Circuit Television System 3-62
Figure 3-25. Roadside Environmental Monitoring System 3-64
Figure 4-1. Proper Lane Use Control Can Facilitate Safe and Efficient Operations . . . 4-3
Figure 4-2. Example of Possible Functional Requirements of a Lane Use Control Objective 4-7
Figure 4-3. Typical Freeway Lane Control Signal 4-14
Figure 4-4. Schematic of an AVI Toll Plaza 4-19
Figure 5-1. Ramp Meter with HOV By-Pass in Minneapolis, MN. 5-5
Figure 5-2. Example of Pretimed Entrance Ramp Metering Rate Calculation. 5-21
Figure 5-3. Pretimed Ramp Metering Layout. 5-22
Figure 5-4. Typical Field Displays for Ramp Meter Installations. 5-25
Figure 5-5. Fundamental Flow Rate-Speed-Density Relationships 5-27
Figure 5-6. Typical Volume-Capacity Plot Related to Ramp Metering Rate. 5-30
Figure 5-7. Traffic Responsive Ramp Metering Layout. 5-31
Figure 5-8. Gap Acceptance / Traffic Responsive Ramp Metering Layout. 5-34
Figure 5-9. Integrated Entrance Ramp Control: Example No. 1 Calculation of Pretimed Metering Rates. 5-37
Figure 5-10. Integrated Entrance Ramp Control: Example No. 2 Calculation of Pretimed Metering Rates. 5-39
Figure 5-11. Integrated Entrance Ramp Control: Example No. 4 Calculation of Pretimed Metering Rates. 5-42
Figure 6-1. HOV Lane in San Diego, CA. 6-5
Figure 6-2. Multimodal System Planning Technique—Steps in the Process. 6-14
Figure 6-3. Examples of Exclusive HOV Facility - Separate Right-of-Way. 6-18
Figure 6-4. Examples of Exclusive HOV Facility - Freeway Right-of-Way. 6-19
Figure 6-5. Examples of Concurrent Flow Lane. 6-19
Figure 6-6. Examples of Contraflow Lanes. 6-20
Figure 6-7. Example of Delay on an HOV Facility Due to a Total Blockage Incident. 6-21
Figure 6-8. Sample Signing and Pavement Markings for HOV Facilities. 6-24
Figure 6-9. Typical Contiguous Concurrent HOV Lane w/o Buffer Signing and Pavement Markings. 6-26
Figure 6-10. Median Enforcement Area for Median 6.7 to 8.8 m (22 to 29 ft) Wide. . . 6-28
Figure 6-11. Median Enforcement Area for Median Less than 6.7 m (22 ft) Wide. . . . 6-30
Figure 6-12. Slip Ramp. 6-30
Figure 6-13. Elevated Slip Ramp. 6-30
Figure 6-14. Flyover Ramp. 6-31
Figure 6-15. Grade Separated Interchange. 6-31
Figure 6-16. Terminal Flyover Connection to Freeway Mainline. 6-32
Figure 6-17. Terminal Slip Ramp Connection. 6-32
Figure 6-18. Terminal Flyover Connection to Surface Street System. 6-33
Figure 6-19. Intermediate Flyover to Frontage Road. 6-33
Figure 6-20. Two-Lane Entrance Ramp with HOV Bypass. 6-35
Figure 6-21. Park-and-Ride Lots with HOV Lane. 6-36

Figure 6-22. Example of Park-and-Ride Layout. 6-40
Figure 6-23. Movable Barrier. 6-47
Figure 6-24. Movable Barrier Used in Dallas. 6-47
Figure 6-25. Typical Sections, Before and After East RLT Contraflow Lane Construction. 6-49
Figure 6-26. Houston HOV System. 6-67
Figure 6-27. Typical Houston HOV Lane. 6-68
Figure 6-28. Houston Grade Separated HOV Lane Access Ramps. 6-68
Figure 6-29. Houston HOV Slip Ramp. 6-69
Figure 6-30. Example of Variable Message Sign Showing New Occupancy Requirement. 6-69
Figure 6-31. Examples of Bus Station and Park-and-Ride Lot. 6-70
Figure 6-32. Seattle Area HOV System. 6-72
Figure 6-33. I-5 Reversible-Flow HOV Lanes. 6-73
Figure 6-34. I-5, Seattle Concurrent Flow HOV Lanes (Median Location). 6-73
Figure 6-35. I-405, Seattle Concurrent Flow HOV Lanes (Shoulder Location). 6-73
Figure 7-1. Dynamic Message Signing on New Jersey Turnpike 7-3
Figure 7-2. Examples of Light-Reflecting DMS Technologies 7-17
Figure 7-3. Examples of Light-Emitting DMS Technologies 7-18
Figure 8-1. Example of an Incident Response Site 8-3
Figure 8-2. Examples of Incident Management Functions. 8-11
Figure 8-3. Relationship between Detection Rate, False Alarm Rate and Detection Time. 8-20
Figure 8-4. Air Bag Use to Overturn Tractor Trailer. 8-26
Figure 8-5. Typical Set-up for Traffic Flow Past a Freeway Incident. 8-27
Figure 8-6. Example of a Typical Total Freeway Closure at an Incident. 8-27
Figure 8-7. Example of Crash Worthy Barrier Gate. 8-29
Figure 8-8. Example Set-Up for Shoulder Utilization at a Freeway Incident Site. . . . 8-30
Figure 8-9. Example of a Contraflow Diversion Set-Up at a Freeway Incident. 8-31
Figure 8-10. Examples of Accident Investigation Site Locations. 8-32
Figure 8-11. Alternative Route Plan. 8-36
Figure 9-1. Fiber Optic Cable Installation in Fort Worth, TX 9-3
Figure 9-2. High Level Logical Architecture for a Communication System 9-13
Figure 9-3. Illustration of a Central Communications Architecture 9-16
Figure 9-4. Illustration of a Distributed Communications Architecture 9-17
Figure 9-5. Illustration of a Trunked Communications Architecture 9-17
Figure 9-6. Illustration of a Backbone Communications Architecture 9-18
Figure 9-7. Traffic Control Communications Network Topologies 9-19
Figure 9-8. Illustration of a Multimedia Communications Channel 9-19
Figure 9-9. Illustrations of Generic Communication Links 9-20
Figure 9-10. Illustration of a Twisted-Wire Pair Cable 9-28
Figure 9-11. Crosstalk Reduction by a Twisted-Wire Pair 9-29
Figure 9-12. Illustration of a Coaxial Cable Design 9-33
Figure 9-13. Key Decisions and Supporting Information in Evaluating the Need for a Shared Resource Project 9-62
Figure 10-1. The Control Center of the TranStar Freeway Management System in Houston, TX 10-3

Figure 10-2. A User-Centered Design Approach 10-6
Figure 10-3. Example of Operational Flow Diagram for Managing an Incident 10-14
Figure 10-4. Data Flow Diagram of the TranStar System in Houston, TX 10-16
Figure 10-5. Typical Operator Workstation in the Houston TranStar Freeway Management Control Center 10-18
Figure 10-6. Floor Plan of the TransGuide Control Center in San Antonio, TX 10-22
Figure 10-7. Example of a Computer/Peripheral Area in a Control Center 10-23
Figure 10-8. Example of Media / War Room in a Freeway Management Center 10-23
Figure 10-9. Media Cabinet in Houston TranStar Center 10-24
Figure 10-10. Example of Static Wall Map 10-27
Figure 10-11. Example of a Projection Television Display 10-27
Figure 10-12. Examples of Popular LAN Topologies 10-32
Figure 10-13. Typical Communications Devices Used by Operators in a Control Center 10-34
Figure 10-14. Continuum of Operator Roles in Relation to Automation 10-41
Figure 10-15. Recommended Topics to be Covered in Operations Manual 10-43
Figure 11-1. An Economic Analysis Has Many Components 11-3
Figure 11-2. Example of Utility Cost Analysis 11-15

LIST OF TABLES

Table 1-1. Geometric Design Factors Causing Capacity Reduction. 1-8
Table 1-2. Percentage of Freeway Section Capacity Available Under Incident Conditions. 1-9
Table 2-1. Criteria for Developing Measures of Effectiveness (MOEs). 2-12
Table 2-2. Example of Functional Requirements of an Incident Management System. 2-14
Table 2-3. Level of Service (LOS) Descriptions. 2-27
Table 2-4. Level-of-Service Criteria for Weaving Sections. 2-30
Table 2-5. Level of Service (LOS) Descriptions for Freeway-Ramp Junction Influence Areas. 2-32
Table 2-6. Queuing Performance Equations for Freeway Lane Blockage. 2-39
Table 2-7. Strengths and Weaknesses of Freeway Traffic Simulation/Optimization Models. 2-45
Table 2-8. Items to be Included in Hardware Documentation. 2-50
Table 2-9. Elements of Software Documentation. 2-52
Table 2-10. Guidelines for Developing a Marketing Plan for Freeway Management Systems. 2-54
Table 3-1. Tasks Involved in Problem Identification 3-7
Table 3-2. Methods for Meeting Typical Surveillance Needs 3-8
Table 3-3. Examples of Performance Measures for a Surveillance System 3-14
Table 3-4. Examples of Functional Requirements Related to Monitoring System Performance 3-15
Table 3-5. Typical Information in Design Plan 3-18
Table 3-6. Guidelines for Specifications 3-18
Table 3-7. Typical Components for a Surveillance System 3-23
Table 3-8. Communication Media 3-28
Table 3-9. Summary of Detectors 3-29
Table 3-10. Summary of Traffic Detectors 3-30
Table 3-11. Results of Oakland County Evaluation of VIP System 3-49
Table 3-12. Characteristics of Traffic Detectors 3-52
Table 3-13. Traffic Surveillance Applications 3-63
Table 3-14. Estimated Operating and Maintenance Costs for Surveillance Systems . . . 3-68
Table 4-1. Lane Restrictions Experiences 4-12
Table 4-2. Summary of Impacts from Truck Restrictions 4-13
Table 4-3. Primary Advantages and Disadvantages of Shoulder Removal 4-16
Table 4-4. Typical User-Fee Schedule 4-20
Table 4-5. Advantages and Disadvantages of Toll Road Financing 4-24
Table 4-6. Barriers to Congestion Pricing Implementation 4-25
Table 5-1. Local Actuated Metering Rates as Function of Occupancy. 5-29
Table 5-2. Allowable Entrance Ramp Volumes for Example 3. 5-40
Table 6-1. Information Needs for Successful HOV Operations. 6-9
Table 6-2. Examples of Goals and Objectives for HOV Treatments. 6-11
Table 6-3. Suggested Objectives and Measures of Effectiveness for HOV Facilities. . 6-12

Table 6-4. Examples of Functional Requirements for HOV Treatments. 6-13
Table 6-5. Suggested Objectives, Data Collection Efforts, and Measures of Effectiveness for Evaluating HOV Facilities. 6-17
Table 6-6. Goals and Objectives of Enforcement Personnel. 6-27
Table 6-7. Accessability Standards. 6-43
Table 6-8. Example of Differential Parking Rates. 6-45
Table 6-9. Successful Characteristics of HOV Facility Implementation. 6-54
Table 6-10. Potential Pitfalls to Successful HOV Facility Implementation. 6-55
Table 6-11. Tradeoffs in Reducing Cross Section Elements 6-58
Table 7-1. Information Need Categories 7-5
Table 7-2. Examples of Goals and Objectives for Information Dissemination 7-7
Table 7-3. Examples of Performance Measures for Information Dissemination 7-9
Table 7-4. Examples of Functional Requirements for Information Dissemination Objectives 7-10
Table 7-5. Applications of DMSs 7-16
Table 7-6. Advantages and Disadvantages of Light-Reflecting DMSs 7-21
Table 7-7. Advantages and Disadvantages of Light-Emitting DMSs 7-22
Table 7-8. Advantages and Disadvantages of Hybrid DMSs 7-23
Table 8-1. Summary of Reported Freeway Incident Rates 8-5
Table 8-2. Typical Capacity Reductions During Incident Conditions 8-6
Table 8-3. Typical Capacities During Work Zone Activities 8-7
Table 8-4. Examples of Goals and Objectives of Incident Management 8-9
Table 8-5. Examples of Incident Management Performance Criteria 8-10
Table 8-6. Information in an Incident Response Plan 8-14
Table 8-7. Comparison of Non-Automated Incident Detection Technologies 8-17
Table 8-8. List of Available Incident Detection Algorithms 8-21
Table 8-9. Example of Resource Material Included in Response Manual 8-23
Table 8-10. Example of Material to be Included in a Hazardous Material Manual 8-24
Table 8-11. Advance Information Dissemination Planning for Incident Response 8-35
Table 8-12. Traffic Management Actions During Major Freeway Construction and Special Events 8-38
Table 9-1. Resources and Constraints Considerations in Communication Systems 9-6
Table 9-2. Advantages and Disadvantages of Longer Polling Periods 9-8
Table 9-3. Commonly Used Error Detection Techniques 9-11
Table 9-4. Characteristics of Transmission Modes 9-14
Table 9-5. Common Applications of Communications Architectures 9-15
Table 9-6. Attributes of Generic Links in a Communications System 9-21
Table 9-7. Relation of Communication Technology to Generic Communication Link for Data Transmission 9-22
Table 9-8. Relation of Communication Technology to Generic Communication Link for Video Transmission 9-23
Table 9-9. Advantages and Disadvantages of Using Communications Standards 9-25
Table 9-10. Advantages and Disadvantages of Installation Methods for Land-Line Communication Media 9-27
Table 9-11. Summary of Land-Line Technologies and Their Features 9-28
Table 9-12. Advantages and Disadvantages of Twisted-Wire Pairs 9-30
Table 9-13. Design and Installation Considerations Relating to Twisted-Wire Pairs . . . 9-32

Table 9-14. Advantages and Disadvantages of Coaxial Cable 9-33
Table 9-15. Advantages and Disadvantages of Fiber Optic Communications 9-36
Table 9-16. Wireless Communication Technologies and Their Associated Properties . . 9-40
Table 9-17. Advantages and Disadvantages of Owned Radio Communications 9-41
Table 9-18. Advantages and Disadvantages of Area Radio Networks 9-42
Table 9-19. Advantages and Disadvantages of Terrestrial Microwave Links 9-43
Table 9-20. Microwave Communication System Design and Installation Considerations 9-44
Table 9-21. Advantages and Disadvantages of Spread Spectrum Radio 9-45
Table 9-22. Spread Spectrum Radio Design and Installation Considerations 9-46
Table 9-23. Advantages and Disadvantages of Cellular Radio 9-48
Table 9-24. Advantages and Disadvantages of Packet Radio Service 9-49
Table 9-25. Advantages and Disadvantages of Satellite Communications 9-50
Table 9-26. Satellite Communication System Design and Installation Considerations . . 9-52
Table 9-27. Comparison of User-Owned and Leased Wire-pair Cable 9-54
Table 9-28. Advantages and Disadvantages of CATV 9-55
Table 9-29. Leased Communications Service Design and Installation Considerations . . 9-57
Table 9-30. Layers of Open Systems Interconnection Reference Model 9-59
Table 9-31. Specified Protocols for NTCIP Profile Classes 9-60
Table 9-32. Issues Associated with Shared Resource Project Development 9-61
Table 10-1. List of Possible Generic Functions Performed in Freeway Management Control Center 10-9
Table 10-2. Criteria of Assigning Functions to Humans and Machines 10-12
Table 10-3. Advantages and Disadvantages of Common LAN Topologies 10-33
Table 10-4. Staffing Plan for Orange County Traffic Operations Center 10-35
Table 11-1. Selected Freeway Management System Component Service Lives 11-8
Table 11-2. Example of a Net Present Worth Evaluation 11-10
Table 11-3. Example of a Benefit-Cost Analysis 11-12
Table 11-4. Example of an Incremental Benefit-Cost Analysis 11-13
Table 11-5. Mean Detection and Lane Change Distances for Channelizing Devices . . 11-16
Table 11-6. Example of Alternative Channelizing Device Values 11-17
Table 11-7. Example of Sensitivity Analysis of Benefit-Cost Ratios 11-18

MODULE 1. FREEWAY MANAGEMENT CONCEPTS

TABLE OF CONTENTS

1.1 INTRODUCTION 1-3
CONGESTION 1-4
Measuring Congestion 1-5
Causes of Congestion 1-5
Geometric Design 1-7
Traffic Operations 1-7
Incidents 1-7
Maintenance and Construction 1-7
Weather 1-7

1.2 FREEWAY MANAGEMENT 1-9
WHAT IS FREEWAY MANAGEMENT? 1-9
WHAT IS A FREEWAY MANAGEMENT SYSTEM? 1-10
OBJECTIVES OF A FREEWAY MANAGEMENT SYSTEM 1-10
HISTORY OF FREEWAY MANAGEMENT 1-11
Functions of Freeway Management 1-11
Surveillance and Incident Detection 1-12
Lane Use Control 1-13
Ramp Control 1-13
High-Occupancy Vehicle (HOV) Priority Treatments 1-13
Information Dissemination 1-13
Freeway Management System 1-13
Incident Management 1-13
Elements of a Freeway Management System 1-14
Field Elements 1-14
Communications Element 1-14
Control Center Elements 1-14
Toll Roads Versus Freeways 1-15
Relation to National ITS Architecture 1-15
Freeway Management 1-16
Incident Management 1-16
Integration with Other ITS Infrastructure Components 1-16
OBJECTIVE OF THE HANDBOOK 1-18
SCOPE OF THE HANDBOOK 1-18
STRUCTURE OF THE HANDBOOK 1-19
Module 1. Freeway Management Concepts 1-19
Module 2. Decision Process 1-19
Module 3. Surveillance 1-20
Module 4. Lane Use Control 1-20

Module 5. Ramp Control 1-20
Module 6. HOV Treatments 1-20
Module 7. Information Dissemination 1-20
Module 8. Incident Management 1-20
Module 9. Communications 1-21
Module 10. Control Center 1-21
Module 11. Economic Analysis 1-21

1.3 REFERENCES 1-23

MODULE 1. FREEWAY MANAGEMENT CONCEPTS

Figure 1-1. I-35 Freeway in Austin, TX.

1.1 INTRODUCTION

Freeways were originally conceived and designed to provide continuous, free-flow, high-speed movement of traffic on limited-access facilities. In their original design, little thought was given to providing for the needs of traffic management and control systems to maintain a high level of mobility on these facilities. However, as urban areas continued to grow, the freeway system became more congested. Today, the previous approach of constructing more freeway lane-miles to relieve congestion is often politically and socially unacceptable and economically infeasible. It is incumbent on transportation agencies planning, constructing, operating, and maintaining freeway infrastructure to make the best possible use of available capacity.

Freeway management systems are a primary means that transportation agencies can use to manage traffic flow and make better use of the existing freeway system. Freeway management systems make use of control strategies, and operational activities such as incident management and information dissemination to 1) keep congestion from occurring in the first place, and 2) lessen the duration and extent of congestion when it does occur.

While the term *freeway management system* was probably not applied to the tasks, early freeway management systems consisted primarily of fixed signs on the roadway providing regulatory and directional information and police officers handling traffic during incidents. As freeway systems have matured and traffic demands have grown in urban areas over the past half century, freeway management systems have

evolved and developed to provide many and varied services.

This handbook discusses the concepts and functional requirements for a freeway management system. It discusses some of the critical issues associated with planning, designing, operating, and maintaining each of the individual elements of a freeway management system. However, prior to comprehensive treatment of the various freeway management elements, it may be helpful to define management systems, the problems they address, and why they are important.

CONGESTION

Congestion on a freeway occurs when demand exceeds capacity. A section of freeway where traffic demand exceeds freeway capacity is called a bottleneck. Bottleneck (or congested) conditions occur either when demand has increased to a level greater than capacity or when capacity has decreased below a level that can accommodate the demand.[1]

Bottleneck conditions are commonplace on the freeway system in many urban areas. Generally, congestion can be classified as either recurring or nonrecurring. Recurring congestion is usually caused when the amount of traffic wanting to use the freeway exceeds the available traffic-carrying capabilities of the system. Generally, recurring congestion occurs at a predictable location during specific periods of the day. Nonrecurring congestion, on the other hand, is less predictable. It is generally caused by random or less predictable events, such as vehicle crashes and incidents, that temporarily reduce the capacity of the freeway, or by special situations (such as sporting events, construction and maintenance activities, inclement weather, etc.) that temporarily increase the demand on the freeway.

Understanding the type and primary causes of congestion in the freeway network is the first step in determining and evaluating options for addressing identified problems with a freeway management system. The strategies that might be implemented to address the recurring congestion problems are not necessarily the same as those that are effective in mitigating the impacts of nonrecurring congestion. From a driver's perspective, however, the impacts of congestion, regardless of whether it is recurring or nonrecurring, are the same:

- Reduced travel speeds.
- Erratic travel speeds characteristic of stop-and-go movement.
- Increased and inconsistent travel times.
- Increased potential for vehicle crashes.
- User dissatisfaction and frustration.

The inability to provide a reliable, albeit sometimes lower, level of service is perhaps a more severe problem than the inability to eliminate congestion altogether. If users know to expect a certain level of congestion during a travel period, they can plan their trip accordingly. If, on the other hand, drivers are unaware of the extent or nature of congestion, they cannot make accommodations to adjust their transportation mode, departure time, or route choices. From this standpoint, one objective of a freeway management system is to provide travelers a consistent level of operation on the freeway.

Measuring Congestion

A motorist usually thinks of congestion in terms of overcrowded freeways, freeway crashes, stop-and-go driving conditions, and the frustration and discomfort of restricted maneuverability. The transportation professional, on the other hand, often expresses congestion in terms of traffic variables such as flow rate, density (or occupancy), and average space speed, together with the fundamental relationships surrounding them. For example, figure 1-2 illustrates the fundamental relationship between flow rate and density. As traffic density increases from zero to some value, k_1, traffic flow rate increases, and the resulting operation is defined as uncongested. As density increases from k_1 to k_2, however, though flow rates tend to increase, traffic operations become unstable, and the probability of serious breakdown increases. Further increases in density above k_2 result in a decrease in flow rate until, theoretically, it reaches zero at jam density, k_j. The traffic flow regimes occurring at densities greater than k_2, are classified as congested.

The specific value of density used to define congestion depends on a number of factors including the geometrics of the freeway section, the composition of the traffic, and local driving habits. The 1994 *Highway Capacity Manual* defines the density of the freeway at congestion (i.e., Level of Service F) to be between 24.8 and 29.9 passenger cars per kilometer per lane (39.7 to 47.9 pc/mi/ln), depending on the original freeway flow speed of the freeway.[2] McDermott has reported that lane occupancies (a surrogate measure for density) in the range of 0 to 20 percent, 20 to 30 percent, and 30 to 100 percent indicate uncongested, unstable (impending congestion) and congested operations, respectively.[3] Figure 1-3 shows generalized freeway traffic operations curves that relate lane occupancy to flow rate and average space speed.[2] As lane occupancy exceeds 20 percent, travel speeds on the freeway decrease because of the following:

- There are fewer and shorter gaps between vehicles.
- Drivers have greater difficulty in changing lanes.
- There are generally more restrictive traffic flow conditions.

Causes of Congestion

Freeway congestion may be characterized as either recurring (generally caused by demand exceeding fixed capacity or fixed geometric restrictions) or nonrecurring (where fixed capacity is reduced due to non-normal events such as vehicle crashes, stalls, spilled loads, weather, or maintenance activities). FHWA estimates that nonrecurring congestion accounts for approximately 60 percent of all congestion.[4]

Factors and situations that may cause a freeway segment to become congested include the following:[5]

- Geometric design.
- Traffic operations (including capacity deficiency).
- Incidents.
- Maintenance and construction.
- Weather.

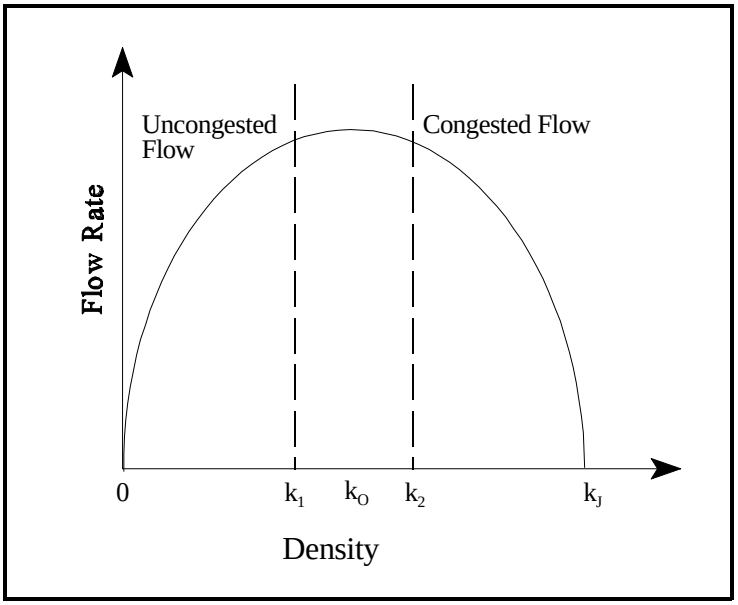


Figure 1-2. Congestion and the Fundamental Relationship Between Flow Rate and Density. [4]

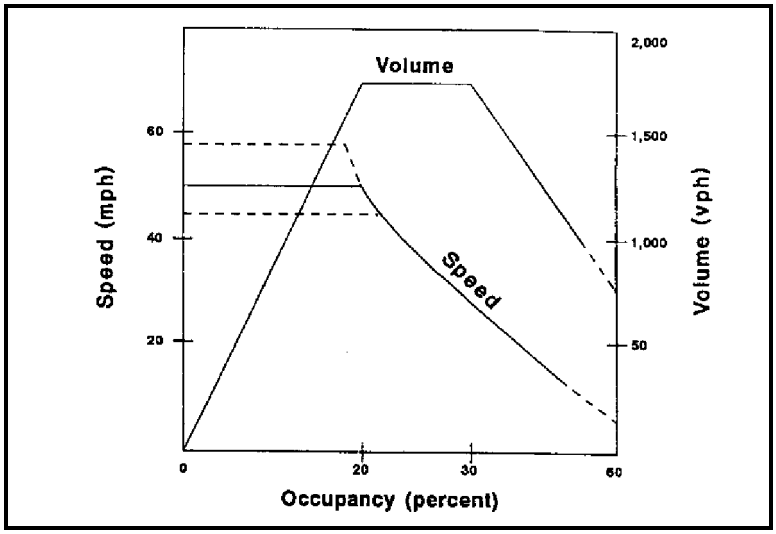


Figure 1-3. Generalized Traffic Flow Relationships. [3]

Geometric Design

The capacity of a freeway usually does not remain constant along its entire length. Certain physical features result in capacity restrictions. Upstream and downstream from the location of these features, capacity usually proves slightly higher, resulting in bottleneck conditions. Table 1-1 lists some of the features that restrict capacity of the freeway and are a source of congestion.

Traffic Operations

Traffic operations can also be a source of congestion—anything from too much demand on the freeway to the presence of heavy vehicles in the traffic stream. Some of the traffic operations factors that can cause congestion on a freeway include the following:

- True demands in excess of available capacity.
- Unrestrained ramp access where ramp traffic causes demand for the freeway to exceed the available capacity.
- Exit ramp queues where demand exceeds the ability of the merge area or downstream intersection to process it, causing queues to backup onto the freeway.
- Heavy weaving and merging movements between ramps and the freeway main lanes.

Incidents

Vehicle crashes and other traffic incidents are major sources of nonrecurring congestion. The factors that affect the amount of congestion resulting from an incident include the following:[5]

- The duration of the incident.
- The amount of capacity reduction caused by the incident.
- The amount of travel demand on the freeway at the incident location.

During an incident, the capacity of the freeway is reduced disproportionately to the physical reduction in width of the travel lanes. Table 1-2 shows the amount of capacity reduction imposed on freeways of different sizes by different incident types.

Maintenance and Construction

Maintenance and construction cause congestion with significant accompanying delays. Like incidents, construction and maintenance activities reduce the amount of capacity available to service demand. The impacts of construction and maintenance activities on traffic operations can be lessened by implementing the following measures:[4]

- Schedule activities for times when demand is anticipated to be low (i.e., at night or on weekends).
- Complete all activities in a given section and lane at one time.
- Provide adequate on-site traffic control and advance information to travelers.
- Report expected and current delays in activities.

Weather

The weather can also be a source of traffic congestion. Both rain and snow can significantly reduce the capacity of the freeway. For example, even trace amounts of rain can reduce the capacity of the

Table 1-1. Geometric Design Factors Causing Capacity Reduction. [4]

Design Factor	Result
Reduction in lanes	• May cause congestion where number of lanes is reduced. • Even though lane drop occurs at exit ramp, through volume may exceed remaining capacity. • Weaving out of dropped lanes may create turbulence, cause speed to decrease, and decrease capacity. (May cause problem where multiple freeways merge without maintaining the same number of lanes.)
Horizontal curvature	• Moderately sharp horizontal curve may reduce capacity. • In heavy flow, vehicle may cross into next lane, causing hesitation and speed decreases in adjacent lanes. Even momentary speed decreases during periods of unstable flow can result in congestion.
Vertical alignment	• Grades reduce capacity, particularly with trucks present. • Although design standards impose limitations on acceptable grades, vertical alignment may cause small, but imperceptible, speed changes that affect following traffic and can result in congestion. • Upgrades in tunnels limit capacity.
Other physical features: • Lane widths • Lateral clearance • Ramp design • Surface conditions	• Older freeways may have lanes narrower than 3.6 m (12 ft) standard, resulting in capacity reduction. • Lateral obstructions may reduce capacity if located closer than 1.8 m (6 ft) to a traveled lane. Examples include: - bridge abutments - retaining walls - illumination poles - sign supports • Lack of full shoulder-width bridges reported to result in a point reduction in capacity. • Weaving movements are restricted at access points and exits.

freeways in Houston, Texas by 14 to 19 percent. In Minneapolis, trace amounts of snow have been reported to reduce capacity by 8 percent and for each 0.02 cm/hr (0.01 in/hr) increase, capacity is reduced by an additional 0.6 percent.[4]

Table 1-2. Percentage of Freeway Section Capacity Available Under Incident Conditions. [4]

Number of Freeway Lanes in Each Direction	Shoulder Disablement	Shoulder Accident	Lanes Blocked		
			One	Two	Three
2	0.95	0.81	0.35	0	N/A
3	0.99	0.83	0.49	0.17	0
4	0.99	0.85	0.58	0.25	0.13
5	0.99	0.87	0.65	0.40	0.20
6	0.99	0.89	0.71	0.50	0.25
7	0.99	0.91	0.75	0.57	0.36
8	0.99	0.93	0.78	0.63	0.41

1.2 FREEWAY MANAGEMENT

WHAT IS FREEWAY MANAGEMENT?

The 1983 *Freeway Management Handbook* defined freeway management as the " . . . control, guidance and warning of traffic in order to improve the flow of people and goods on these limited access facilities."[6] Today, the definition of freeway management should be expanded to encompass all activities undertaken to operate a freeway facility in a manner consistent with predetermined goals and objectives of that facility including those related to the impacts on and the influence of surrounding communities and jurisdictions. Certainly, the efficient movement of goods and people continues to be one of the major goals of freeway management. However, legislative mandates and public pressure regarding environmental concerns, adjacent landowner and homeowner rights, and other issues can further influence the operation of the freeway system.

In many cases, freeway management requires striking a balance between competing goals and objectives. For example, a goal to minimize recurrent congestion may not necessarily be compatible with a goal to maximize the movement of people and goods on a facility, especially if the goal to maximize people movement is accomplished through the use of high-occupancy vehicle lanes. High-occupancy vehicle lanes must offer the user a significant travel time advantage over regular-use lanes to be attractive to travelers using the freeway. In terms of this example, it may be appropriate to allow some level of congestion in the regular-use lanes in order to achieve the best overall operation of the freeway system.

WHAT IS A FREEWAY MANAGEMENT SYSTEM?

In general terms, a system is defined as a set of components or elements that could be seen as working together for the overall objective of the whole.[7] These components or elements can be many different things, depending on the perspective of the person who defines the system. System components can be objects, concepts, processes, or even people. A system has a set of boundaries that define the interactions between system components and the environment. The environment consists of all those factors external to the system that influence the behavior of the system, but which cannot be controlled. There are interfaces between the components of the system which control how they interact with each other. Finally, the components themselves can be subsystems consisting of smaller elements or components.

A freeway management system, then, consists of the infrastructure elements utilized to accomplish the goals and objectives of freeway management. These things (components) include field hardware (cameras, variable message signs, electronic toll tag readers, etc.), communications equipment, a traffic management center (with associated hardware and software), the people who staff the center, and the policies and procedures established to deal with various transportation-related events that impact the freeway system. The freeway management system comprises several infrastructure subsystems (the motorist information system, the ramp metering system, etc.) that interface with each other to accomplish specific objectives as the need for them arises.

It is important to recognize that each freeway management system is a unique consolidation of components and interfaces that reflect the location-specific freeway, geographic, and political characteristics and needs of the region. However, most systems can be described in terms of which of the possible functions of freeway traffic management they perform. The following section presents a brief overview of these functions.

OBJECTIVES OF A FREEWAY MANAGEMENT SYSTEM

The goals and objectives of any one freeway management system are specific to the social and political desires of the community; however, there are goals and objectives for freeway management systems that are universal across all systems. These include the following:[8]

- To reduce the impacts and occurrence of recurring congestions on the freeway system.
- To minimize the duration and effects of nonrecurring congestion on the freeway system.
- To maximize the operational safety and efficiency of the traveling public while using the freeway system.
- To provide facility users with information necessary to aid them in making effective use of the freeway facilities and to reduce their mental and physical stress.
- To provide a means of aiding users who have encountered problems (crashes, breakdowns, confusion, etc.) while traveling on the freeway system.

HISTORY OF FREEWAY MANAGEMENT

Pioneering efforts in "freeway surveillance and control" as systems were generally known, took place in Detroit, Chicago, Houston, Los Angeles, Seattle, and Dallas. Highlights of some these early efforts are listed below:[9]

- **1960 Chicago** - Freeway service patrols.
- **1960 Detroit** - CCTV on freeways.
- **1962 Detroit** - Freeway lane control signs and variable speed signs.
- **1962 Chicago** - Surveillance by freeway loop detectors.
- **1963 Chicago** - Demand-capacity / occupancy metering.
- **1965 Houston** - Demand-capacity / gap-acceptance metering and surveillance by closed-circuit television and freeway loops.
- **1967 Los Angeles** - Fixed time metering and ramp closure.
- **1967 Seattle** - Reversible roadway control and closed circuit television surveillance.
- **1971 Dallas** - Corridor control integrating ramp metering, frontage road and arterial intersection control with preferential treatment for buses at ramps and intersections.
- **1972 Minneapolis** - Bus bypass ramps at metered ramps.
- **1990** - Intelligent Vehicle/Highway (IVHS) concept introduced including freeway management and related systems.
- **1992** - Intelligent Transportation Systems (ITS) succeeds IVHS, broadening scope of concept.
- **1993** - National ITS Architecture Initiative begun to address common interfaces and protocols for ITS systems.

As systems were upgraded and expanded, they began to be more properly termed "Freeway Management Systems." New and upgraded systems have incorporated many of the basic techniques developed on those earlier systems. The phenomenal developments in computer and communications technology have further enhanced capabilities for various freeway management tasks.

Functions of Freeway Management

Freeway management systems combine personnel, operational strategies and technologies together to control and manage traffic on the freeway more effectively. Figure 1-4 highlights the basic functions of freeway management. The functions that can be performed by a freeway management system include the following:

- Surveillance and incident detection.
- Lane use control.
- Ramp control.
- Priority treatment and control for high-occupancy vehicles.
- Information dissemination.
- Incident management.

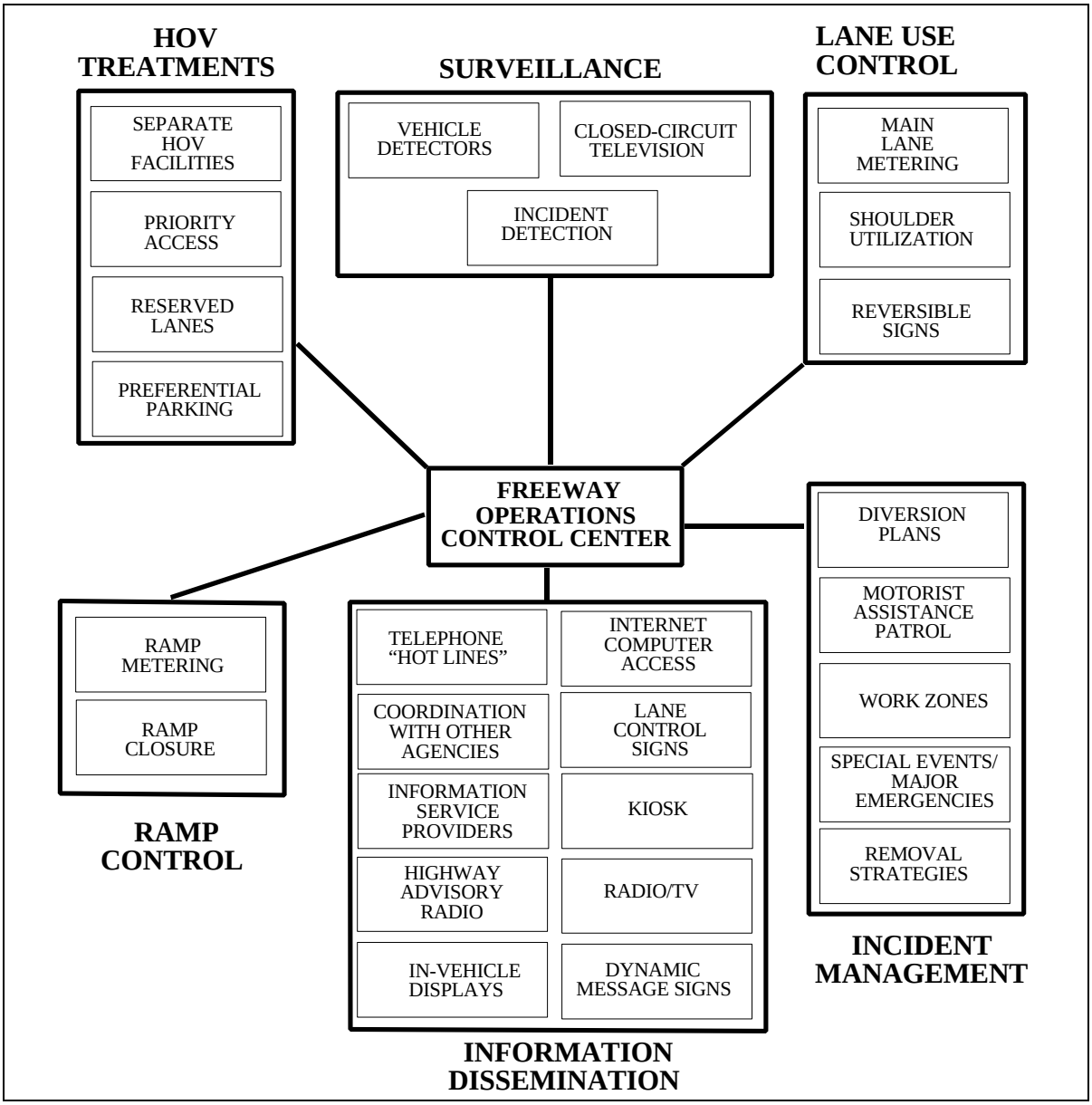


Figure 1-4. Functions and Elements of a Freeway Management System.

A brief review of each of these functions is provided below. A more detailed discussion of the pertinent issues affecting planning, design, operations, and maintenance is provided in subsequent modules.

Surveillance and Incident Detection

Traditionally, one of the primary functions of most freeway management systems is to provide traffic surveillance.[10] With surveillance, traffic conditions can be monitored and the location and causes of any operational problems that occur can be pinpointed. As traffic management systems have become increasingly multijurisdictional and multimodal, the surveillance function has expanded to include such things as transit and emergency fleet vehicle locations, weather and pavement conditions, and parking lot status.

Lane Use Control

The function of lane use control is to maximize the efficient use of existing pavement within the right-of-way. Mainline metering, temporary shoulder utilization, reversible lane operations, and large truck restrictions are all examples of lane use control.

Ramp Control

Managing the amount of traffic that can enter or exit the freeway is another function of freeway management systems. Both ramp metering and temporary ramp closures (during peak periods, for instance) are examples of the ramp control function. Controlling freeway access and egress reduces turbulence in freeway flow near the ramps and improves overall freeway operations.

High-Occupancy Vehicle (HOV) Priority Treatments

Another function of freeway management is to provide preferential treatment to buses, carpools, and vanpools on a freeway to generate a travel time advantage for the vehicle occupants and review the number of vehicles on the roadways. These treatments can include special lanes, priority access ramp controls for HOV vehicles, special surveillance of lanes to detect and remove incidents quickly, and occupant restriction enforcement.

Information Dissemination

From a traveler's perspective, information dissemination is one of the most important functions of freeway management. Research and experience have shown that travelers want and use real-time information about traffic conditions on the freeway and alternative routes; adverse weather and driving conditions; construction and maintenance activities; and special lane use and roadway control measures that are enacted. With current and accurate information, travelers are better able to make mode, departure time, and route choice decisions. A number of existing and emerging technologies are available to facilitate information transfer to travelers.

Freeway Management System

A freeway management system is generally the common interface for multiple transportation agencies in a given region. ITS infrastructure includes a "regional multimodal traveler information" element. This may be co-housed with the freeway management center. In that case, the freeway management center is the focal point for receiving, processing, and transferring various types of traffic and transportation data, including digital, video, and voice. Regional multiagency coordination will logically take place in the freeway management center.

Incident Management

Of all the major freeway management functions, incident management probably offers the greatest potential operational and safety benefit to freeway motorists.[5] Incident management requires the coordinated operation and preplanned use of human and technological resources to restore the freeway to full capacity quickly and efficiently after an incident occurs. Typically, incident management includes other freeway management functions such as the following:

- Control center operation.
- Surveillance.
- Information dissemination.

- Freeway service patrols.
- Ramp control.
- Decisions process control strategies.
- Lane use control signals.

Elements of a Freeway Management System

Each of the functions of a freeway management system requires one or more elements in order to successfully carry out that function. A number of these elements are pieces of hardware that must be in place, whereas others are procedures and strategies that are enacted. Generally, these elements can be categorized as existing or occurring either in the field, as part of the communication linkage, or within the control center.

Field Elements

A number of different technologies can be utilized in the field to monitor traffic conditions, control traffic access and lane utilization, manage incidents, and communicate with motorists. Most of the major field elements are illustrated in figure 1-4 (within each major freeway management function). Within certain elements, a variety of hardware technologies can be employed. For example, vehicle detection within the surveillance function can be accomplished through inductive loop detection, microwave or radar sensors, video imaging, automatic vehicle identification or location, or other technologies, depending on the type of data desired. Similar alternatives exist for information dissemination, lane use control, and incident management functions.

Communications Element

Communications can be thought of as the backbone of freeway management.[11] The communications element transfers information from the field elements back to the traffic control center. In some cases, instructions and other data are transferred out from the center to the field elements. Reliability and performance are critical issues related to communications system design.

The communications element consists of a number of hardware components and supporting software. Again, a number of hardware technologies are available upon which to base communications. The two basic categories that exist are buried or aerial cable (fiberoptic, coaxial, twisted pair) or airwave transmission (microwave, radio [including narrow band or spread spectrum], cellular telephone, and citizen-band radio).

Communications infrastructures (and continuing system operational and maintenance costs) are typically the most expensive part of a freeway management system. FMS functional design must include the communications system concurrently with field devices and central control elements to ensure compatibility and economies of system deployment.

Control Center Elements

The traffic control center is the hub or nerve center of a freeway management system. It is where information about the freeway system is collected, processed, and collated. It can also be the location where decisions about control strategies are made, coordinated with other agencies, and implemented. Information dissemination is also typically coordinated and implemented from the center.

The major elements of the control center are the display and control interfaces that link the center to the various field elements through the communications element. The other major elements of a center are the human operators who monitor conditions and make the adjustments in control and management strategies needed to maximize the efficiency of the freeway system. Ultimately, the degree to which a traffic control center meets the objectives of the freeway management system depends on how well the human operators are able to interface with the system devices.[12]

Toll Roads Versus Freeways

It is important to note that any of the elements of a freeway management system can also be applied to the toll roads. Toll roads exhibit many of the same design features and operating characteristics as freeways. For example, access to both types of facilities is controlled. Both try to achieve high operating speeds. Both need to communicate trouble locations to users of the facilities.

The primary operating difference between toll facilities and freeways is the presence of toll booths. Toll booths are necessary interruptions to the free flow of vehicles on the facility, and need to be considered in developing control strategies.

Relation to National ITS Architecture

It is beyond the scope of this handbook to describe the National ITS Architecture initiative. However, the reader is directed to the extensive documentation that has been developed under that FHWA sponsored project. Quoting from the *1996 ITS Architecture Executive Summary*:

> *The National ITS Architecture provides a common structure for the design of intelligent transportation systems. It is not a system design nor is it a design concept.*

Agencies should seek to ensure that Freeway Management System design is compliant with the National ITS Architecture and applicable National Traffic Control/ITS Protocol.

Growing out of the National ITS Architecture development is the concept of Intelligent Transportation System infrastructure. The term “infrastructure” has been in common use for a number of years to describe the supporting roadways, bridges, water and sewer lines, and other public works structural items that allow movement of persons and goods. It is not the truck or car; it is the roadway and bridge. It is not the water; it is the pipes or conduits that carry the water. A logical extension of the infrastructure terminology is its application to flow and delivery of information in Intelligent Transportation Systems.

Many of the functions needed for ITS implementation are already being provided or supported by a broad variety of ITS infrastructure features, which can serve as the building blocks of a full ITS implementation. ITS infrastructure refers to those portions of ITS-related hardware, software, etc. that today, and increasingly in the future will manage and support the transportation-related activities. This is typically happening first in the metropolitan areas, but is expanding to include commercial vehicle and rural needs.[13]

An intregrated transportation management system contains two or more of the following nine components:

- Traffic Signal Control.

- Freeway Management.
- Transit Management.
- Incident Management.
- Electronic Fare Payment.
- Electronic Toll Collection.
- Railroad Grade Crossings.
- Emergency Services.
- Regional Multimodal Traveler Information.

Although freeway management and incident management are the components included in this handbook, the other components are often noted as they relate to those two components.

Freeway Management

Real-time information describing flow conditions is essential to managing the freeway system. Such information is repackaged for dissemination to the traveler, and is also used to make control decisions such as in-ramp metering or lane use control. Deployment objectives for the freeway management element of an ITS infrastructure are as follows:[13]

- Provide critical information to travelers through infrastructure-based dissemination methods such as variable message signs and highway advisory radio.
- Monitor traffic and other environmental conditions on the freeway system.
- Identify recurring and nonrecurring impediments so that short-term and long-term actions can be taken to alleviate congestion.
- Implement various control and management strategies (such as ramp metering and/or lane control, or traffic diversion).
- Use probe vehicles for additional sensors for collecting real-time traffic information.

Incident Management

Rapid detection, response, clearing the roadway, and restoring capacity are the essential elements of incident management. Reduction of secondary accidents and delays to the traveler are the result of an efficient incident management program. Deployment objectives for the freeway management element of an intelligent transportation management system are as follows:[13]

- Coordinate incident management across regional boundaries to ensure efficient and sufficient response.
- Use traffic management capabilities to improve response times.
- Use onboard moving map route guidance equipment to assist incident response vehicles (e.g., ambulances and tow trucks).
- Reduce traveler delay due to incidents.

Integration with Other ITS Infrastructure Components

As shown in Figure 1-5, freeway management is just one component of an integrated transportation management system. Freeway management, along with traffic signal control, form the hub of an

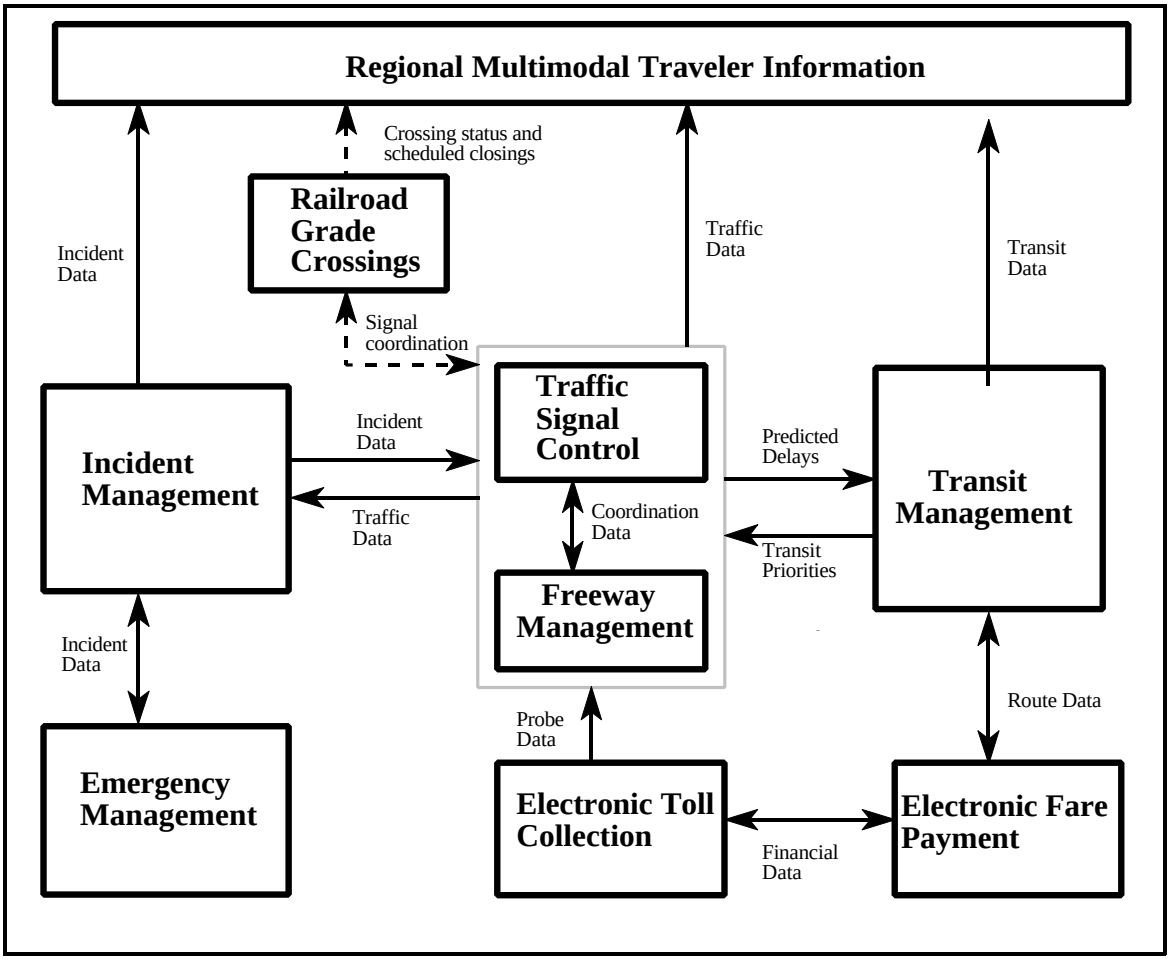


Figure 1-5. Freeway Management as Part of an Integrated Intelligent Transportation System. (14)

integrated transportation system. As shown inside the larger box, the traffic signal control and freeway management components exchange data with each other, allowing the traffic management strategies on the freeways, the freeway entrance ramps, and the surface street network to act as an integrated system. These data define the actions to be taken by the system when a particular signal timing plan is in effect on the road network, and when a particular sign plan is in effect on the highways. For example, ramp meter timings and traffic signal controls can be coordinated to ensure that queues do not back into intersections. When traffic is diverted off the freeway to bypass an incident, the traffic signals timings can be adjusted to handle the increased flow on the arterials. At the same time, dynamic message signs can be updated along the arterials to give directions on how to return to the freeway beyond the incident location.

In addition, both the freeway management and traffic signal control components can provide critical information to other components of the transportation system. Both the freeway management and traffic signal control components are responsible for the surveillance, monitoring, device control, and management of the road network. These components provide information on the status of the roadway system, including link travel times, traffic

volumes, and speeds currently flowing on the road and highway network. These data can be passed to the other components in the overall transportation system, such as the incident management and transit management components, and information from these other systems can be channeled through the freeway management and traffic control components. Traffic data can also be output to a regional traveler information system for dissemination to the public for trip planning and other purposes. Because the freeway management and traffic signal control components are monitoring traffic conditions on the roadways, information about predicted delays and incidents can be passed to the transit management component, which in turn can use the information to establish transit priorities and operating strategies, which are implemented by the freeway management and traffic signal control components.

Although it may be possible for some transportation management subsystems to operate in an isolated or independent manner, their operation and functions will affect and be affected by the other transportation subsystems. This will be particularly evident in an urban environment. For optimal efficiency, the various transportation subsystems must operate in a cohesive, integrated manner. Multiple agencies must maintain open communication and cooperation to achieve the goal of a truly integrated transportation system.

For a description of the other infrastructure components and further information on the intelligent transportation infrastructure requirements for freeway management and incident management, the reader is referred to *Building the ITI: Putting the National Architecture into Action* and *Operation TimeSaver-Building the Intelligent Transportation Infrastructure.* [13, 14]

OBJECTIVE OF THE HANDBOOK

The objectives of this handbook are as follows:

- Define the process of planning, designing, operating, and maintaining a freeway management system.
- Provide an overview of the techniques and technologies available for accomplishing the functions of a freeway management system.
- Identify critical issues associated with planning, designing, operating, and maintaining a freeway management system.
- Provide guidance in developing a freeway management system that is compliant with the ITS National Architecture and applicable NTCIP and national standards.
- Identify points in the design development process where integration and interface with other systems should take place.

SCOPE OF THE HANDBOOK

This handbook is intended to be a “How To” manual for planning, designing, operating, and maintaining a freeway management system. This manual is directed at a mid-level administrator or engineer for a State or local agency that is responsible for planning, designing, operating, or maintaining all or portions of a freeway management system. As a result, it primarily focuses on the issues associated with each element of a freeway management system. For many of the elements and functions, there are excellent reference materials that provide detailed information about the technologies and techniques. For example, the *Communications Handbook for Traffic*

Control Devices and the *Freeway Incident Management Handbook* provide many technical details on communications and incident management systems, respectively.[5,11] In those cases where there are other manuals and reference materials to provide technical details, the details have not been reproduced in this manual. Instead, the reader has been provided with references to these materials as a source of more detailed information than that presented in this handbook is desired.

STRUCTURE OF THE HANDBOOK

A modular approach has been used to prepare this manual. Each functional element of a freeway management system is covered in a separate module. Each module presents information specific to the planning, design, construction, operation, and maintenance of that particular functional element. This approach was used when preparing this handbook so that system designers and operators can turn to the modules of interest and find all the information pertinent to the planning, design, operation, and maintenance of those specific elements.

In addition, a decision process module has been provided. The decision process module describes the systems engineering approach that can be used to plan and design a freeway management system. The systems engineering approach can be used regardless of whether a system is being designed from scratch, modifications are being made to an existing system, or upgrades are being planned to an outdated system. The decision process module also serves as a model of each of the remaining modules.

The same structure has been applied to each of the modules. Each module contains an overview section that provides some background into the nature of the functional elements, and the objectives and scope of the information presented in each module. Following the overview section, a section describes a system engineering process for planning, designing, and implementing the functional element within a freeway management system. Another section describes the techniques and technologies commonly used in each functional element. Each module also has a section devoted to covering any special issues that could not be adequately covered in the previous sections. Wherever possible, examples of state-of-the-art facilities are provided in each module. The module concludes with a section showing the references and suggested readings for the module.

A summary of the type of material to be covered in each of the proposed remaining modules is provided below.

Module 1. Freeway Management Concepts

This module provides an introduction to the handbook, including objectives, scope, and structure. Also included is an overview of freeway management concepts as a primer or preface to the 10 modules addressing technical issues and management issues.

Module 2. Decision Process

The second module is devoted to discussing the process for determining functions and technologies to be included in a freeway management system. It provides the steps to be taken for conducting a system engineering analysis for planning, designing, and implementing a freeway management system. This module focuses on the issues associated with integrating each of the functional elements to form a comprehensive freeway management system.

Module 3. Surveillance

This module contains information related to planning, designing, and implementing the surveillance subsystem of a freeway management system. Using the structure for the decision process discussed in Module 2, this module provides a process for determining the appropriate elements to be included in a surveillance subsystem. It provides an overview of the various technologies and techniques that are available for performing the surveillance functions in a freeway management system.

Module 4. Lane Use Control

This module focuses on the process for planning, designing, operating, and maintaining systems for controlling the use of the freeway main lanes. This module provides information related to the design and operations of freeway main lane meters, techniques for using the shoulder, and providing reversible lane control on the freeway. It also provides insight into the special issues related to operating and maintaining these control systems.

Module 5. Ramp Control

This module presents information needed to plan, design, operate, and maintain freeway ramp control systems. It provides an overview of the theory and principles of ramp control, including ramp metering and ramp closures. It also presents information on the different types of ramp metering strategies. The module concludes with a section on the special issues associated with ramp control systems, including maintenance, enforcement, and public perceptions and acceptance of various ramp control strategies.

Module 6. HOV Treatments

This module presents information needed to include high-occupancy vehicle (HOV) priority treatments in a freeway management system. It provides a summary of the planning, design, and operations issues associated with including an HOV treatment in a freeway management system. It also provides an overview of the decision process for determining which type of priority treatment is appropriate for a specific system. It also discusses some of the special design, operating, and maintenance issues associated with HOV treatments.

Module 7. Information Dissemination

This module describes the process for establishing a cohesive information dissemination subsystem in a freeway management system. It identifies the existing and emerging technologies available to facilitate information dissemination to travelers. It also illustrates how the information dissemination subsystem integrates with the other subsystems of a freeway management system.

Module 8. Incident Management

Module 8 presents information related to the planning, design, operations, and maintenance of an incident management subsystem in a freeway management system. It is intended to illustrate how to develop or enhance the incident management capabilities of a freeway management system. It also provides insight into the key factors and issues to be considered in planning, designing, and implementing an incident management subsystem. It is intended to complement the material already presented in the *Freeway Incident Management Handbook*.[5]

Module 9. Communications

This module provides insight into the issues associated with planning, designing, operating, and maintaining a communications subsystem in a freeway management system. It focuses on the decision process for determining the appropriate technologies and system architecture for specific freeway management systems. It also provides a summary of existing and emerging communications technologies. It is intended to complement the material already presented in the *Communications Handbook for Traffic Control Systems.*[11]

Module 10. Control Center

This module provides information related to planning, design, operations, and maintenance of the control center. It focuses on both the physical design and human factor elements associated with freeway management control centers. It also provides insight into how the functions of the freeway management system affect the design and operations of the control center. Special issues such as privatization of operator functions and multi-agency control centers are also discussed in this module.

Module 11. Economic Analysis

This module presents information on some of the economic analyses that can be used in planning and evaluating alternative designs of freeway management systems. Included in this module is step-by-step instruction on how to conduct a life-cycle cost evaluation, a benefit-cost evaluation, and other pertinent economic analyses.

This module also provides guidance to the reader on when it is appropriate to use the different types of economic analyses.

1.3 REFERENCES

1. Wilshire, R., R. Black, R. Grochoske, and J. Higginbotham. *Traffic Control Systems Handbook*. Report No. FHWA-IP-85-12. FHWA, U.S. Department of Transportation, 1985.

2. *Highway Capacity Manual*. Special Report 209. TRB, National Research Council, Washington, DC, 1994.

3. McDermott, J.M., et. al. *Chicago Area Expressway Surveillance and Control, Final Report.* Report No. FHWA-IL-ES-27. FHWA, U.S. Department of Transportation, March 1979.

4. *Incident Management Workshop; Relieving Traffic Congestion Through Incident Management*. FHWA, U.S. Department of Transportation, 1994.

5. Gordon, R.L., et. al. *Traffic Control Systems Handbook.* Report No. FHWA-SA-95-032. FHWA, U.S. Department of Transportation, February 1996.

6. Reiss, R.A., and W.M. Dunn. *Freeway Incident Management Handbook.* Report No. FHWA-SA-91-056. FHWA, U.S. Department of Transportation, July 1991.

7. Sumner, R. , et. al. *Freeway Management Handbook, Volume 1: Overview*. Report No. FHWA-SA-88-026. FHWA, U.S. Department of Transportation, May 1983.

8. Athey, T.H. *Systematic Systems Approach*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982.

9. *Traffic Control Systems Handbook.* Implementation Package 76-10. FHWA, U.S. Department of Transportation, June 1976.

10. Everall, P.F., *Urban Freeway Surveillance and Control: The State of the Art.* FHWA, U.S. Department of Transportation, November 1972.

11. Sumner, R., et. al. *Freeway Management Handbook, Volume 3: Operations and Maintenance*. Report No. FHWA-SA-88-028. FHWA, U.S. Department of Transportation, May 1983.

12. Gordon,R., Reiss, R., Dunn, W., and Morehead, D. *Communications Handbook for Traffic Control Systems*. Report No. FHWA-SA-93-052. FHWA, U.S. Department of Transportation, January 1993.

13. *Human Factors Handbook for Advanced Traffic Management Center Design*. FHWA, U.S. Department of Transportation, October 1995.

14. *Building the ITI: Putting the National Architecture Into Action*. Mitretek Systems for the FHWA, U.S. Department of Transportation, April 1996.

15. *Operation TimeSaver-Building the Intelligent Transportation Infrastructure*. Informational Packet, FHWA, U.S. Department of Transportation, January 1996.

MODULE 2. DECISION PROCESS

TABLE OF CONTENTS

2.1 INTRODUCTION 2-3
MODULE OBJECTIVES 2-4
MODULE SCOPE 2-4

2.2 DESIGN PROCESS 2-4
DEFINE PROBLEMS 2-6
ESTABLISH INSTITUTIONAL FRAMEWORKS AND BUILD COALITIONS . . . 2-7
ESTABLISH SYSTEM GOALS AND OBJECTIVES 2-9
ESTABLISH PERFORMANCE CRITERIA 2-10
DEFINE FUNCTIONAL REQUIREMENTS 2-12
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 2-13
IDENTIFY AND SCREEN TECHNOLOGIES 2-13
DEVELOP IMPLEMENTATION PLAN 2-15
DEPLOY PROJECTS 2-16
EVALUATION 2-18

2.3 TECHNIQUES AND TECHNOLOGIES 2-20
PERFORMANCE MEASURES 2-20
Systemwide 2-21
Total Travel Time 2-21
Total Travel 2-22
Vehicle Delay 2-22
Total Minute-Miles of Congestion 2-22
Isolated Locations 2-23
METHODS TO QUANTIFY OPERATIONAL PROBLEMS ON FREEWAYS . . . 2-23
Freeway Capacity Analysis and LOS 2-23
Definition of Freeway Subareas for Capacity Analysis 2-23
Basic Freeway Segments 2-24
Definition of Ideal Freeway Capacity 2-24
Factors Affecting Capacity 2-24
Level of Service 2-25
Weaving Areas 2-26
Definition of Weaving Area Capacity 2-26
Types of Weaving Areas 2-26
Operational Characteristics of Weaving Areas 2-26
Capacity and Level of Service for Weaving Areas 2-28
Ramp Junctions 2-28
Definition of Ramp Junction Capacity 2-28
Operational Characteristics of Ramp Influence Areas. 2-31
Level of Service of Freeway-Ramp Junctions. 2-31

Freeway Systems 2-31
Highway Capacity Software (HCS) 2-33
BOTTLENECK/QUEUING ANALYSES 2-34
FREEWAY SIMULATION/ OPTIMIZATION MODELS 2-37
Macroscopic Freeway Simulation/Optimization Models 2-38
FREQ 2-39
CORFLO 2-40
QUEWZ-92 2-41
Microscopic Freeway Simulation Models 2-41
FRESIM 2-42
INTEGRATION 2-43
Strengths and Weaknesses of the Available Models 2-44
FIELD MEASUREMENTS 2-44
Travel Time and Delay Studies 2-44
Origin/Destination Volume Studies 2-46

2.4 LESSONS LEARNED 2-46
SYSTEM MAINTENANCE 2-46
Types of Maintenance Requirements 2-47
Functional 2-47
Hardware 2-48
Software 2-48
MARKETING SYSTEMS 2-51
Target Audiences 2-51
Operating Agency 2-51
Affected Agencies 2-51
Elected Officials 2-52
General Public 2-52
Marketing Strategies 2-53

2.5 REFERENCES 2-55

APPENDIX A: SUBCHAPTER G - ENGINEERING AND TRAFFIC OPERATIONS 2-57

MODULE 2. DECISION PROCESS

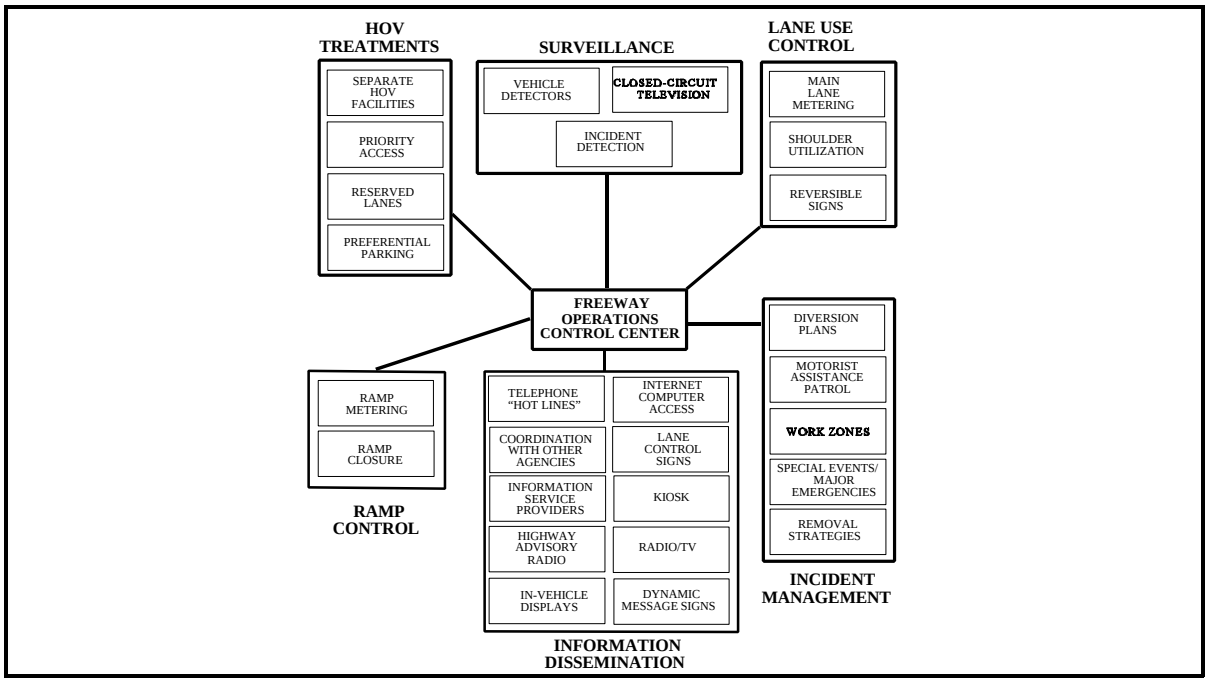


Figure 2-1. Elements of a Freeway Management System.

2.1 INTRODUCTION

Management of the freeway system has become increasingly popular as a tool for combating congestion problems in urban areas. A freeway management system is a collection of strategies, personnel, and technologies combined together to accomplish the following:

To promote the efficient and effective movement of people and goods, to improve the safety of the traveling public, and to improve the environment by reducing both the duration and extent of recurring and nonrecurring congestion on the freeway system.

A freeway management system includes one or more subsystems to accomplish the following tasks:

- Monitoring operations of the freeways and detecting sources of recurring and non-recurring congestion.
- Controlling use of existing lanes and shoulders to maximize the use of the existing pavement within the freeway right-of-way.
- Managing the amount of traffic that enters and exits the freeways via the ramps.
- Providing priority treatment of high-occupancy vehicles within the freeway.
- Disseminating both pre-trip and en route information to travelers so that they can make informed mode, route, and departure time decisions.
- Restoring the freeway to full capacity as quickly as possible after an incident.

Regardless of the size of the freeway system or the level of sophistication and extent of the existing freeway management system, a systems engineering approach to planning and designing a freeway management system should be followed. A systems engineering approach is an iterative process whereby operational problems are identified, system concepts and objectives are established, potential solutions are developed and evaluated, new solutions are identified, and system objectives are redefined. The approach is applied throughout the entire life cycle of the system (planning, design, operations, and maintenance), with the level of detail being progressively refined. The same systems engineering process can be used in each of the following applications:

- Planning and implementing a new system where no current freeway management system exists.
- Modifying an existing system to add new freeway management functions.
- Upgrading the technology in an existing freeway management system to current standards.

MODULE OBJECTIVES

The objectives of this module are as follows:

- Describe the systems engineering approach to planning and designing a freeway management system.
- Present techniques for quantifying operational problems and measuring performance of a freeway management system.
- Highlight some of the special issues that need to be considered in planning, designing, operating, and maintaining a freeway management system.

MODULE SCOPE

This module describes a systems engineering process for planning and designing a freeway management system. It highlights some of the major issues and decisions that planners and designers must address when developing a freeway management system. It also illustrates how decisions made during the planning and design phase affect the operations and maintenance of the system and how operational and maintenance issues can be addressed during the planning and design phases.

2.2 DESIGN PROCESS

A systems engineering approach examines how various tasks, components, and operational strategies can be combined to form a working and coherent unit to address specific goals, objectives, and user needs.[1] With the systems engineering approach, the components of the system are viewed as an entity rather than as an assembly of individual parts.

Figure 2-2 illustrates the steps involved in a typical systems engineering analysis. These same steps can be used to complete a systems engineering analysis, whatever the size of the system being designed and evaluated. The same process can be used to design an entire freeway management system, individual components of the system, or individual elements within each component: what varies is how much detail is used in the analysis and the data required to complete the analysis.

Also note that the process is iterative. After defining the problem, those internal and external organizations and agencies that must be brought on board if the system is to be effective can be identified. In cooperation with these organizations and agencies,

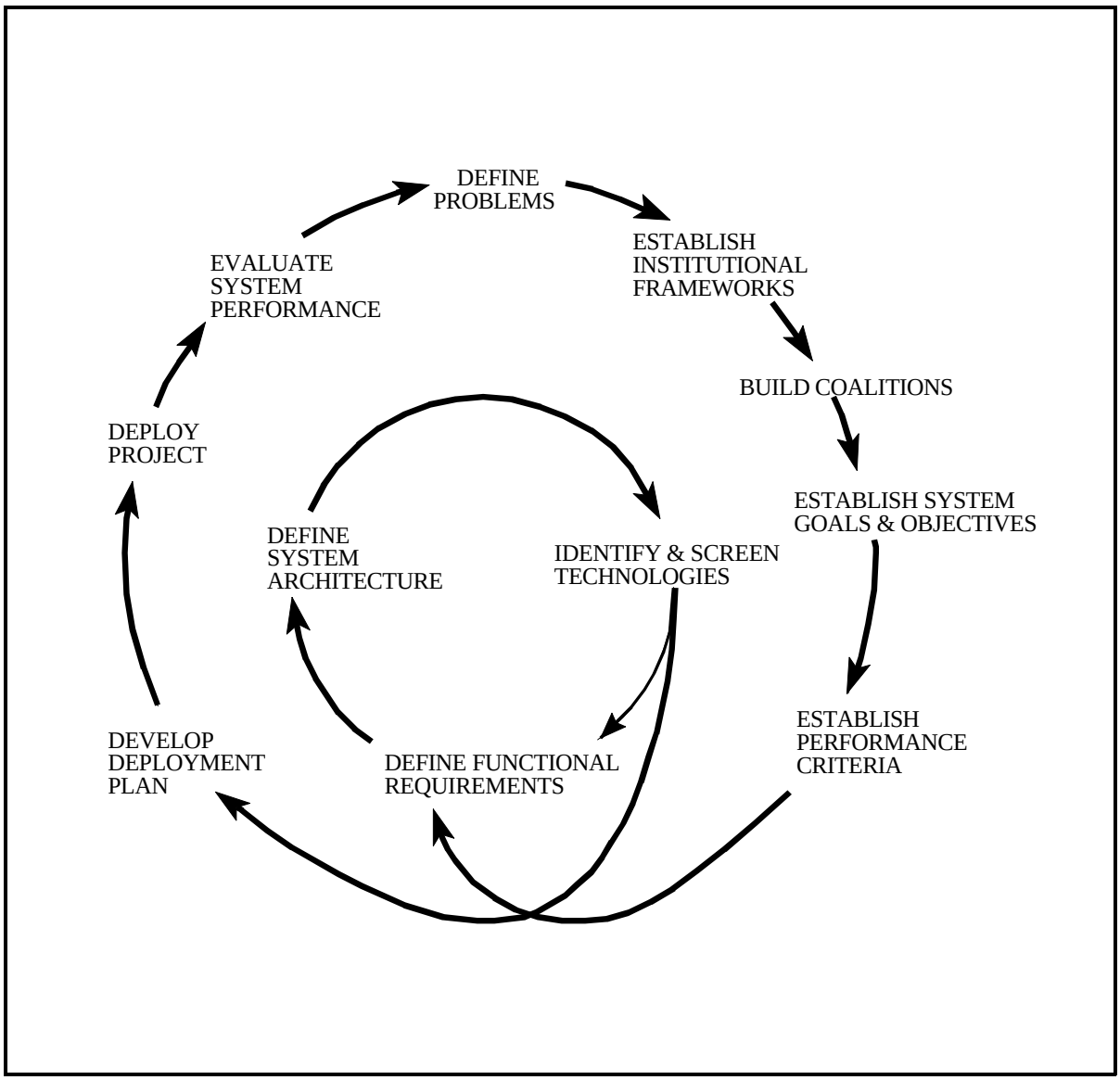


Figure 2-2. Steps Performed in a Typical Systems Engineering Analysis.

system goals and objectives can be identified and performance criteria established. Functional requirements can then be defined to meet the specified goals and objectives. Once the functional requirements have been identified, the system architecture can be defined, and technologies can be identified and screened. The process of defining functional requirements, defining system architecture, and identifying and screening technologies is reiterated at increased levels of detail until system plans and specifications can be prepared. Once the final design of the system is complete, the process of deploying the system can be planned. Using this plan, individual projects can be identified that will lead to the full deployment of the system. Once individual projects have been installed, the performance of the system can then be evaluated. This evaluation may lead to new operational problems being identified, whereby the systems analysis process can be used to identify new improvements to the system. Each step in a typical systems engineering analysis is discussed in the sections below.

A similar process was used to develop the National Intelligent Transportation System (ITS) Architecture. That process is shown in figure 2-3.[2] It is beyond the scope of this handbook to provide a detailed treatment of the National ITS Architecture development. The Architecture was being developed during the same time this handbook was prepared and, as of this date, has not been finalized. The reader is referred to numerous National ITS Architecture documents developed under a Department of Transportation contract with Loral Federal Systems and Rockwell International. In particular, the National ITS Architecture Implementation Strategy is of benefit because it summarizes many of the previous efforts and provides broad guidance in implementing not only freeway systems, but also other related systems.[2]

The National ITS Architecture approach began in 1993 with a concept of "user services" rather than specific hardware items or techniques to support the development of the architecture. The National Architecture work enhanced the user services concept by defining "market packages" that provide groupings of technologies and techniques to satisfy those user services. Reference is made to table 2.2-2 (page 2-29) in the Implementation Strategy document for a relationship of market packages to various services. Succeeding discussions in this module will refer the reader to the appropriate sections of the Implementation Strategy where appropriate.[2]

DEFINE PROBLEMS

The first step in the systems engineering approach is to adequately and properly define and quantify the problems to be addressed by the freeway management system. Properly defining the problems to be addressed by a freeway management system directly leads to the identification of the subsystems that can address those problems.

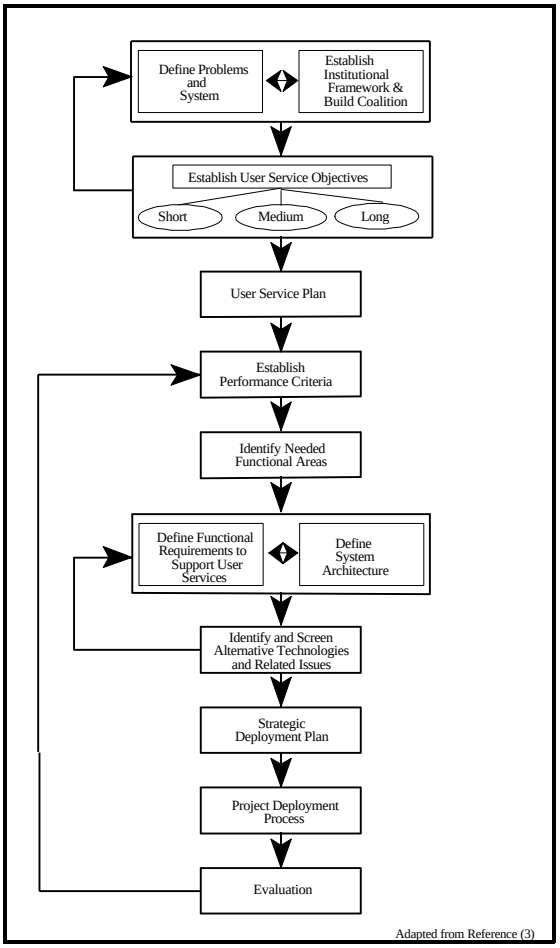


Figure 2-3. Systems Engineering Process Used to Develop National ITS Architecture. [2]

Several methods are available for quantifying the magnitude of the operational problems that can be potentially addressed by a freeway management system. Traditional analytical techniques that can be used to quantify operational problems in a freeway network include the following:

- Capacity and level of service analysis.
- Bottleneck analysis.

- Queuing analysis.
- Travel time and delay studies.
- Computer simulation modeling.
- Crash history analyses.
- Origin/destination studies.

The major tools available for conducting operational analyses of the freeways are discussed in Section 2.3 below.

The importance of coordinating with other transportation-related agencies to identify problems to be addressed should not be overlooked. For example, long-range transit plans can provide an indication of when high-occupancy vehicle facilities may be needed in the freeway system. Input from commercial industries can also be valuable in identifying specific areas of concerns or needs. Finally, and perhaps most importantly, input from local elected officials can also provide insight into the public's perception of problems with the roadway system.

Most of the techniques above help identify existing operational problems; however, in planning a freeway management system, it is important to consider the location and type of operational problems that might occur in the future. Potential sources of information that could be used to identify future operational problems include the following:

- Regional transportation and land use planning studies.
- Site impact analyses.
- Air quality assessments.

In addition to quantifying existing and future operational problems, another critical element in defining the problems that exist is to obtain an accurate and complete inventory of the entire existing transportation system. This inventory should include both physical and organizational components. Examples of physical components that should be identified in the inventory include the following:

- Roadway network.
- Existing surveillance and control systems.
- Existing information dissemination systems.

In addition to the existing physical elements in a system, how the system is envisioned to operate can also influence the design of the system. Examples of the operational components that might influence the design of a traffic management system and, thus, must be identified in an inventory include the following:

- Operating agencies.
- Funding sources.
- Political and agency jurisdictions.

ESTABLISH INSTITUTIONAL FRAMEWORKS AND BUILD COALITIONS

The most critical step in implementing a successful freeway management system is to build critical coalitions and institutional frameworks. Institutional frameworks and coalitions need to be established at the beginning of the planning and design process. These coalitions and institutional frameworks need to be established at three levels: between agencies (inter-agency), within agencies (intra-agency), and with

other stakeholders (including the private sector) affected by traffic operations.

Traffic congestion is not restricted by jurisdictional boundaries. When one part of the transportation system (e.g., the freeway network) is not functioning properly, it affects the operations of other parts of the system (e.g., the surface streets, the high-occupancy vehicles lanes, etc.). Therefore, there is a strong need to develop good working relationships and build coalitions among agencies responsible for managing traffic in an area. Examples of the types of agencies with which strong coalitions would help in implementing a freeway management system include the following:

- Metropolitan planning organizations (MPOs).
- Federal, State, and local traffic and transportation engineering agencies.
- Federal, State, and local transit agencies.
- State and local law enforcement agencies.
- Emergency services (fire, ambulance).
- Turnpike / toll road authorities.
- Port authorities.
- State and local emergency management authorities.

In addition, private sector needs should be considered in developing a freeway management system, and should be included in the consensus building process. Examples of these private sectors entities that agencies may want to include in the consensus building process include the following:

- Private transportation information providers and services.
- Private transportation companies.
- Commercial delivery services.
- Interstate and intrastate commercial trucking companies.

Cooperation and coalitions within an agency are also essential in establishing an effective freeway management system. Often, this type of cooperation is the hardest to obtain. Some sections within an agency may view a traffic management system as usurping some of their responsibilities and power. It is essential that all elements within an agency (e.g., planning, administrative, construction, design, operations, and maintenance) are committed to constructing, operating, and maintaining the system.

Finally, there may be others in the community that may be important allies when implementing freeway management systems. Perhaps the most important of these is the general public. Without the support of the general public, it will be extremely difficult to implement a traffic management system. Extensive public relations and media campaigns may be required to show the public how a traffic management system will directly benefit them. Strong public support makes it easier to secure funding and political support. Without public support, it will be extremely difficult to generate elected officials' and agency interest and support for freeway management.

Other stakeholders that may be useful in implementing traffic management systems include the following:

- Major traffic generators.

- Utility companies.
- Elected officials.
- Media.
- Private transportation providers.

Establishing institutional frameworks and coalitions can be difficult at times. The first step in building effective coalitions is to identify "champions" in those agencies responsible for transportation in a community (e.g., State highway agencies, MPOs, transit authorities, etc.). These individuals are likely to be top administrative officials in these organizations. Since traffic management systems often compete with other "traditional" agency activities and expenditures (e.g., pot-hole patching, construction, etc.), the support of top management is essential if agency resources are to be allocated to the operation and maintenance of the system. If congestion is widely recognized as a major issue in a community, upper management support may already exist within many organizations and propagate through the agencies in a "top-down" fashion; however, if it is not, support must be generated from the bottom up.

In addition to identifying "champions" for the system, it is also important to identify those individuals who will be critical to the success of the systems (e.g., the public, elected officials, major employers, etc.). The support of one or more local elected officials can be highly effective in securing funding for the system. Often this support is automatic if traffic conditions have already deteriorated or if the potential for major traffic problems looms on the near horizon.

Another effective way of building strong coalitions between and within agencies is to take advantage of institutional frameworks that already exist. Many locales have institutional frameworks to address freeway management concerns. For example, many locations use Traffic Management Teams and Incident Management Teams to address problems on freeways. Often these teams are a coalition between State and local transportation agencies, and law enforcement personnel. These coalitions can be expanded to encompass additional functions of a traffic management system.

Finally, in building effective coalitions, it is important to identify, within organizations, key individuals who have the appropriate level of knowledge and experience. These individuals must have the level of authority that is appropriate to the type and level of responsibilities that will be performed by the agency. In addition, it is also important to identify individuals who are likely to be present throughout the entire planning and design process for a traffic management system. A common thread among past successful transportation management systems is that key personnel committed to the system have remained with the agencies throughout the planning, implementation, and ongoing operation of the system.

ESTABLISH SYSTEM GOALS AND OBJECTIVES

Once coalitions have been formed, agencies should work together to define the goals and objectives of the system. The system goals and objectives should describe what it is the system is supposed to accomplish. The goals and objectives should be directly related to the specific problems to be addressed by the system. Typical system goals include the following:

- Optimize the utilization of capacity on the freeway.
- Maintain an acceptable level of service of the freeway.

- Provide a balance between demand and capacity in a freeway corridor.
- Provide for the rapid detection, response to, and clearance of freeway incidents.
- Reduce the number of vehicle crashes on the system.
- Improve vehicle operating costs, fuel consumption, and air quality by reducing the amount of travel delay on the freeway system.

Generally, system goals are used to define the long-range vision of the system. System goals also tend to be broad in scope. System objectives, on the other hand, define the level of performance that is expected to be obtained in the future. As such, system objectives are measurable. Often, more than one system objective is required to fulfill a system goal. Likewise, more than one system goal may be required to address an identified problem in a system.

Figure 2-4 provides an example of a system goal and objectives developed to address the problems of incidents in a system. Note that while the defined problems and system goals are broad in nature, the objectives of the system are specific and measurable. Also note that more than one objective is required to fulfill the goal of the system.

It is also important to note that system objectives are defined in terms of what services and functions a system is to provide — not in terms of technology. Notice in the example that no mention is made of the type of technology that will be employed to achieve a 2-minute detection time. Focusing on what the system is to achieve, instead of on how it is to achieve it, gives the designers flexibility in the way that they can combine components to build a system to achieve a desired outcome.

The National ITS Architecture Implementation Strategy provides a good overview of the connection between problems, solutions, and the National Architecture.[2] The approach identifies traditional problems (e.g., traffic congestion) and a solution in terms of both a conventional approach (e.g., roadway construction) and an advanced systems approach (e.g., computer traffic control), and links them to ITS marketing packages. The reader is referred to that document for problems and solutions in the context of the National ITS Architecture.

ESTABLISH PERFORMANCE CRITERIA

After establishing the system goals, the next step in the systems engineering approach is to establish the criteria for judging the performance of the system or subsystem. The performance criteria are used to determine whether the objectives of the system are being achieved. The criteria include both qualitative and quantitative measures of performance for the system. They also form the basis for evaluating the design and operations of the system. The criteria used to measure the performance of the system should correspond directly to the objectives of the system. Examples of the type of criteria that are commonly used to evaluate the effectiveness of freeway management systems include the following:[3]

- Changes in congestion levels and travel patterns.
- Changes in system operation costs, both to the user and to the system operator.
- Changes in other measures that effect the overall community (e.g., reduction in crash frequencies, vehicle emissions, and increased transit usage).

Identified Problem

- Incidents are the primary source of congestion on freeway

System Goal

- Reduce the impacts of incidents

System Objectives

- Detect all major incidents within 2 minutes of occurrence
- Reduce the time to clear an incident from the freeway by 5 minutes

Figure 2-4. Example of System Goal and Objectives to Address an Identified Problem.

- Improved accessibility as evidenced by decreased delays, increased economic activity, and reduced travel time for commuters benefiting from HOV improvements.

- Changes in vehicular demands on congested facilities.

In selecting criteria for evaluating the system performance, the measures of effectiveness must reflect the goals and desires of the community, and must be important and understandable to local elected officials, administrators, citizens, and other affected groups.[3] Table 2-1 lists some of the criteria that should be used in developing and selecting measures to evaluate the performance of a freeway management system. Common measures used to evaluate the effectiveness of freeway management systems include the following:

- Reduction in total travel time.

- Reduction in individual and overall delays.

- Reduction in crash frequencies and rates.

- Reduction in the duration of periods of congestion.

- Reduction in the number of congested freeway segments or length of congestion on a freeway section.

- Improvements in throughput in a corridor.

- Reduction in vehicular demands in a corridor.

- Improvements in network speed.

- Improvements in the consistency of travel times on individual trips.

- Reduction in the number of single-occupant vehicles in the freeway system.

- Improvements in the response and clearance times of freeway incidents.

- Improvements in air quality.

The National ITS Architecture Implementation Strategy considers benefits in the context of market packages, where the benefits are likely to occur, and where

Table 2-1. Criteria for Developing Measures of Effectiveness (MOEs). [4]

- **Relevancy to Objectives**

  Each MOE should have a clear and specific relationship to transportation objectives to assure the ability to explain changes in the condition of the transportation system

- **Simple and Understandable**

  Within the constraints of required precision and accuracy, each MOE should prove simple in application and interpretation

- **Quantitative**

  Specify MOEs in numerical terms whenever possible

- **Measurable**

  Each MOE should be suitable for application in pre-implementation simulation and evaluation (i.e., have well-defined mathematical properties and be easily modeled) and in post-implementation monitoring (i.e., require simple direct field measurement attainable within reasonable time, cost, and staffing budgets)

- **Broadly Applicable**

  Use MOEs applicable to many different types of strategies whenever possible

- **Responsive**

  Specify each MOE to reflect impacts on various groups, taking into account, as appropriate, geographic area and time period of application and influence

- **Sensitive**

  Each MOE should discriminate between relatively small changes in the nature or implementation of a control strategy

- **Not Redundant**

  Each MOE should avoid measuring an impact sufficiently measured by other MOEs

- **Appropriately Detailed**

  MOEs should be formulated at the proper level of detail for the analysis

the benefits may accrue.[2] The reader is referred to that document for a more detailed treatment of the approach as related to freeway systems, in particular tables 4.2.3, 4.2.4, and 4.2.5 (pp 4-17 and 4-18.)

DEFINE FUNCTIONAL REQUIREMENTS

The next step in the systems engineering approach in designing a system is to define all of the features or activities (commonly called functions) of the system that are necessary to achieve the identified objectives. The system functions need to be described, at least initially, independent of the technology or architecture to be employed in the system. In other words, this step focuses on describing *what* it is the system will be designed to do, not *how* the system will be doing it.

The functional requirements needed to achieve a system objective can often be outlined in hierarchial order. For example, the functional requirements of an incident management program might be described as shown in table 2-2. Note that each of the functional requirements defines an action or

activity that is to be performed by the system and that is independent of technology.

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

After defining *what* the system is supposed to accomplish, the next step in the systems approach is to define the functional relationships, data requirments, and information flows of the system. This provides a framework within which the system carries out the functions required to support the desired objectives. It describes the system elements and their relationship to one another.

The functional relationships, data requirements, and information flows of many freeway management systems in operation today have evolved as new functions were added to the system. However, there are real benefits to be achieved in planning how systems relate to one another in advance, even if the system will not be fully implemented immediately. Planning these relationships minimizes the number of redundant functions and efforts performed by the system. It also promotes the efficient use of equipment, staff, and resources. A well-planned system permits easy expansion and modernization of the system in the future. How the functions relate to one another also facilitates the sharing of information between jurisdictions, and leads to cost savings throughout the design, implementation, and operation of the system.

The design or architecture of a system consists of three elements: the functional requirements, the logical design, and the physical design. As discussed above, the functional requirements define what the system is supposed to do. The logical design identifies what information flows between the functions. The physical design identifies where functions occur and who is responsible for performing each function. The physical design permits like functions to be grouped into subsystems.

It is extremely important when defining the design of the system that it remain as open as possible. An “open” design is a system that has been designed with standard data interfaces so that equipment from multiple vendors can be used throughout the system. In addition, an open design helps to keep the system from becoming obsolete, because new functions and technologies can be easily added as they became available. Furthermore, an open design will make systems being developed today compatible with the National ITS Architecture as it emerges.

IDENTIFY AND SCREEN TECHNOLOGIES

Once the system functional relatioships, data requirements, and information flows have been defined, the next step in the systems engineering process is to identify alternative technologies whose performance and reliability meet the defined functional requirements. Methods for estimating the benefits derived by implementing a particular technology or system component include the following:

- Capacity and level of service analyses.
- Bottleneck and queuing analyses.
- Computer simulation.
- Field measurements.

In addition, evaluation studies from other systems, promotional documents from vendors, and demonstrations from

Table 2-2. Example of Functional Requirements of an Incident Management System.

- Detect incidents
  - Identify location of incident
  - Identify impacts of incident
  - Identify characteristics of incident
- Formulate response actions
  - Identify necessary emergency vehicle response
  - Select incident information for dissemination to travelers
  - Identify traffic control strategies
- Initiate and monitor response
  - Provide response procedures to agencies
    - Implement emergency vehicle response
    - Provide incident information to travelers
    - Implement traffic control strategies
  - Monitor response
    - Arrival emergency vehicles
    - Implementation of traffic control
    - Clearance of incidents
    - Clearance of congestion

manufacturers are excellent sources of information on the state-of-the-art of specific technologies. Site visits to existing systems that use specific types of technologies can also be an excellent tool for evaluating various technologies.

The interaction between alternative technologies and other elements within the system should be considered when evaluating different technologies. How the systems work together and what functions they perform can greatly influence how different technologies perform in a system. The impacts of different technologies on the physical configuration of the system and on the performance of other technologies and components in the system should also be considered. The expandability and flexibility of the technologies should also be considered.

Cost is another factor that should be considered when identifying and screening different technologies. The designers should consider the life-cycle costs (see Module 11) of each alternative technology. When examining the life-cycle costs of each technology, the engineer should consider the following costs:

- The costs associated with procuring, installing, and constructing each alternative technology.
- The costs associated with operating and maintaining each alternative technology within the overall system.
- The costs associated with replacing each technology at the end of its life.
- The costs associated with expanding the system, given that each alternative technology has been implemented.

Operation and maintenance are also important factors that must be considered when evaluating different technologies for use in a management system. Often, each technology requires unique operating and maintenance activities. The resource

requirements, in terms of the number and qualifications of the personnel, the equipment and facility needs, and the operating and maintenance costs, should be factored into the evaluation. In evaluating system alternatives, the following operations and maintenance requirements should be considered:[5]

- An assessment of the existing operation and maintenance capabilities in terms of personnel, skills, and equipment.
- A determination of the necessary skills and work load impact of each system alternative.
- A comparison of the existing conditions with what is required for the successful operation and maintenance of the various alternatives.
- An analysis of the operational and maintenance deficiencies that exist with each system alternative.
- An assessment of the feasibility of providing additional operational and maintenance capabilities to address the identified deficiencies.

The process of identifying and screening different technologies for inclusion in a system is often iterative. There are multiple ways that different technologies can be combined to achieve an objective. Because how different technologies interact with one another can affect system performance, each combination must be evaluated in an iterative fashion.

DEVELOP IMPLEMENTATION PLAN

After the technologies that will be used in the system have been selected, the next step in the process is to develop a plan for implementing the system. The implementation plan documents the results of the previous steps and identifies how the system will be implemented in the field. At a minimum, an implementation plan should include the following elements:

- A description of the transportation system problems and opportunities to be addressed by the system.
- The institutional arrangements (i.e., who, what, when, where, why, and how) needed to make the system work.
- The goals and objectives of the system (i.e., what the system will be expected to do).
- The functional requirements and architecture of the entire system.
- The technology options to be used in the system.

The implementation plan should also assess the phasing, procurement, and funding options available for implementing the system.

The purpose of an implementation plan is "to ensure that the system is designed, built, operated, and maintained so that it accomplishes its purpose in the most efficient manner possible, considering performance, cost, and schedule." An implementation plan is required when either a new traffic control system or an expansion of an existing system uses Federal funds. Because it completely describes how the system is going to be designed, operated, and maintained, it is highly recommended that an implementation plan be developed as well for those systems that do not use Federal funds. Figure 2-5 lists the elements of a typical implementation plan.[6]

The National ITS Architecture Implementation Strategy provides guidance for major traffic management design options in the context of its market packages.[2] The reader is referred to Section 4.5 of that document in particular tables 4.5.2 and 4.5.3 (pp. 4-59 to 4-63) for the National ITS Architecture approach to implementation.

DEPLOY PROJECTS

There are a number of approaches that are commonly used by agencies to deploy individual projects or systems. The more common types of procurement approaches include the following:[6]

- Sole-source.
- Engineer/contractor.
- Two-step approach.
- System management.
- Design-build.

With a sole-source project, a contract is awarded to a named supplier without any competition for the project. This type of procurement approach typically involves a standard off-the-shelf product that can be made only by one manufacturer. Such an approach should be used only if it can be justified to be more cost-effective than a competitive low-bid process.

With an engineer/contractor approach, a single contract is awarded to the lowest responsive bidder to a specific request by the highway agency. The contractor is then responsible for providing a complete and fully operational system. In designing and deploying the system, the contractor may elect to subcontract much of the work that is outside the area of expertise of the contractor. The highway agency should be careful to specify the appropriate level of qualifications for the prime contractor to avoid getting a contractor who is unable to complete the job.

A two-step procurement process helps to eliminate some of the problems associated with the standard engineer/contractor approach. With a two-step procurement approach, a formal technical prequalification process is added to the engineer/contractor approach. This helps ensure that the contract team has the appropriate skills and expertise for implementing the desired type of system.

A systems management approach is commonly used by many highway agencies to implement freeway management systems. With this approach, a system manager is hired to perform the system design, software development, and system integration activities. Separate contracts are then prepared and awarded for implementing the various subsystems as dictated by the design.

A design-build type of procurement approach is another common way that highway agencies deploy freeway management projects. In a design-build approach, a single entity is responsible for all of the work associated with deploying a system. This includes the system design, the contracts for and the construction of the system components, and the integration of the system elements. Upon completion of the project, the designer-builder turns over the system to the agency for operations and maintenance. This process is often used to fast-track projects, since it can significantly reduce the process time. Agency supervision is required, however, to ensure that a satisfactory quality of product is provided by the contractor.

The reader is referred to the *Traffic Control Systems Handbook* and Part 655-Traffic

- Needed Legislation
- System Design
  - System Designer
  - System Design Life
  - System Coverage
  - System Design and Operations/Maintenance Philosophies
  - System Architecture
  - Integration of Other Functions
  - System Components and Functions
  - Communication Subsystem Design Approach
  - Traffic Operations Center Design Features
  - Project Phasing/Scheduling
  - Design Review
- Procurement Methods
- Construction Management Procedures
  - Division of Responsibilities
  - Scheduling and Establishing Mileposts
  - Conflict Mitigation
  - Coordination with Other Projects
- System Start-up Plan
  - Software and System Acceptance Tests
  - Partial Acceptance
  - Documentation
  - Transition from Old to New Control
  - Operational Support and Warranty Period
  - Training
  - Coordination with Media
- Operations and Maintenance Plan
  - Evaluation
    - System evaluator
    - Method of evaluation
    - Cost of evaluation
  - Maintenance Plan
    - Maintenance policies for preventative maintenance, system malfunctions, etc.
    - Formal maintenance management programs
    - Initial inventory of spare parts and all necessary test equipment
    - Training in providing limited maintenance to software and equipment
- Institutional Arrangements
  - Contact Person/project Liaison Within Each Organization
  - Delineation of Organizational Responsibilities
  - Provisions for Periodic Project Updates
  - Utility Arrangements
  - Written Cooperative Agreements for Personnel-sharing, Cost-sharing, Metering, Traffic Diversion, Etc.
- Personnel and Budget Resources
  - Staffing Plan (including the number of persons and their functions per shift)
  - Contract Operations Staff Agreements (if used)
  - Provisions for Training New Staff
  - Sources of Budgetary Resources
  - Estimates of Annual Expenses by Category
  - Signatures of the Head of the Operating Agency, Head of State Highway Agency, and the FHWA Division Administrator (or their designates)

Figure 2-5. Elements of a Typical Implementation Plan.

Operations, Subchapter G-Engineering and Traffic Operations in the *Federal-Aid Policy Guide* (which has been regenerated in Appendix A at the end of this module) for more information on alternative procurement strategies.[3,6]

Staffing plays a critical role in the timing of the implementation of the system. Sufficient lead time must be allocated to allow system operators to establish new positions (if needed); develop position descriptions and salary classifications; recruit, hire, and train needed personnel. The agency's system supervisor should be employed (i.e., designated/hired) no later than the start of the implementation phase of the project. Other operations and maintenance personnel need to be employed during system construction so that they may receive training providing by the contractor/system manager, review system documentation, and participate in system testing. If these persons are to assist with construction inspection, they need to be hired even earlier.[5]

In addition, good preventative maintenance begins with and depends upon proper installation. Proper installation requires thorough inspection and testing of contractors' work. The system operators and maintainers should participate in the inspection process and during the acceptance testing. Such involvement is an excellent method of informal training and fosters the acceptance of the new system by these personnel. [5]

It is essential that all specified documentation and training be obtained during the installation phase of the system. Submitted materials (e.g., manuals, training course outlines, etc.) should be closely examined to ensure that they are complete and clear to the agency's personnel. Furthermore, the documentation must be kept up-to-date throughout the entire implementation phase to reflect any hardware or software modifications.[5]

EVALUATION

Evaluation is an ongoing process that occurs at all stages of system development and continues for the entire life of the system. Through the evaluation process, the system designers and operators are able to determine how well individual projects meet the previously established system objectives. The evaluation process also allows system managers to identify possible enhancements to the system. These enhancements can be to correct operational or design problems, expand the system either functionally or geographically, or incorporate additional systems into a regional architecture.

The most common method of evaluating the effectiveness of a freeway management system is *Before-and-After* studies. With Before-and-After studies, the performance of the freeway management system is measured before the system (or subsystem) being evaluated is implemented. The same performance measures are then taken again after the system has been implemented. The effectiveness of the system is then determined by comparing the performance of the freeway before the system (or subsystem) was implemented to the performance of the freeway after the system (or subsystem) was implemented.

Potential limitations of the *Before-and-After* approach for evaluating system performance include the following:[7]

- The effects of individual improvements are difficult to distinguish when more than one improvement is made at a time.

- Usually there is a long time lag between the "before" condition and the "after"

condition, which causes this approach to be susceptible to errors caused by time-related factors (such as changes in travel patterns, population growths, economic fluctuations, etc.).

- Usually, it takes some time for the drivers to adjust their travel behavior after the system has been implemented; therefore, depending upon when the "after" data are collected, true system performance may not be measured.

- Some performance measures (like the number of crashes, or demand) can fluctuate considerably over time. There is a tendency for these performance measures to return to more typical values after an extraordinary value has been observed. This tendency is called *regression to the mean*. It is possible that either the "before" condition or the "after" condition could fall at one of these extreme values, thereby, hiding the true performance of the system.

Before-and-After studies with a Control Condition is one way of mitigating the negative aspects of Before-and-After studies.[7] The concept of this evaluation methodology is illustrated in figure 2-6.

With this approach, the before-and-after system performance measures are compared to the same performance measures taken from a segment of freeway with similar characteristics as the first freeway, but not affected by the operation of the system. While this approach reduces history, maturation, and regression to the mean problems, it may be difficult to find a segment of freeway not impacted by the performance of the freeway management system.

In evaluating the effectiveness of the components of a freeway management system with field measured data, it is important to remember the following considerations:[3]

- The impacts of many alternative strategies are frequently small both in absolute and in percentage terms.

- The impacts of specific strategies may be confined to isolated or small geographic areas.

- It may not be possible to accurately estimate the specific impacts using commonly available data and estimation procedures.

In some limited situations, it may be appropriate to use simulation models to estimate the effectiveness of different elements of a freeway management system. Oftentimes, systemwide benefits of different freeway management elements are difficult to measure directly. In these cases, simulation could be used as a tool to estimate the impacts of isolated control elements on the entire system. Furthermore, the results of a simulation study should only be considered as estimates of the performance of the system, since the performance measures are often directly influenced by the assumptions inherent in the simulation model. Whenever simulation is used as an evaluation tool, limited field data should be collected to support the results of the simulation model. As a general rule, simulation should never be used when the same performance indicators can be measured directly in the field.

Another way to illustrate the benefits of a freeway management system is to use anecdotal information collected from travelers who have benefited from the system. For example, many locations keep records on the number of positive letters

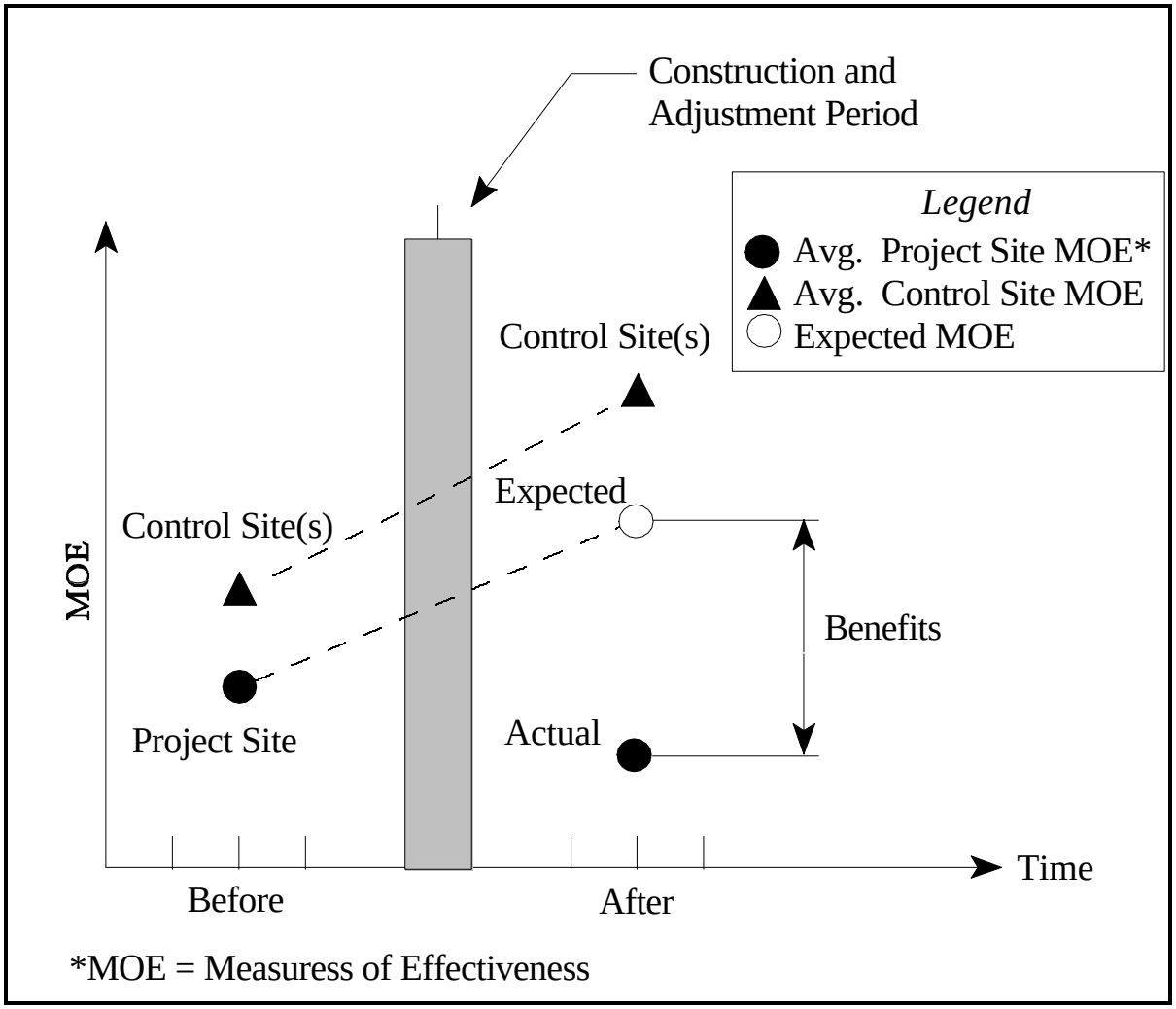


Figure 2-6. Conceptual Illustration of Before-and-After Evaluation with Control Conditions. [7]

they receive concerning their freeway service patrols.

Also, one of the most critical parts of the evaluation process is to document the lessons learned during the development and operations of the system. These lessons provide critical information to others who may be considering implementing a similar type of system. The lessons learned should not focus solely on the problems that were encountered during the development or operations of the system, but also describe the positive elements of a particular system architecture or technology.

2.3 TECHNIQUES AND TECHNOLOGIES

This section describes some of the measures of performance, techniques, and technologies used to identify and quantify operational problems on freeways.

PERFORMANCE MEASURES

Congestion occurs when the amount of traffic desiring to use a facility (demand) exceeds the traffic-carrying capabilities of the facility (capacity). There are two types

of congestion: recurring and nonrecurring. Recurring congestion occurs when normal, everyday demand exceeds the physical capacity of the freeway. Nonrecurring congestion is caused either by a temporary reduction in capacity (e.g., a crash or stalled vehicle blocking a lane) or by a temporary excess of demand (caused by a special event or other similar activity).

In planning and designing a freeway management system, it is important to quantify the magnitude of the congestion problem and the amount of improvement that can be made by implementing elements of a freeway management system. The following section provides some of the more common measures of performance used to quantify congestion levels both from a systemwide perspective and at isolated locations such as bottlenecks and incident sites.

Systemwide

There are a number of measures that have traditionally been used to quantify the performance of the freeway system, both with and without a freeway management system in place. These measures include the following:[8]

- Total Travel Time.
- Total Travel.
- Vehicle Delay.
- Total Minute-Miles of Congestion.

Each of the traditional measures of system performance is discussed below. Most of these measures can be computed using data routinely collected by freeway surveillance systems.

Total Travel Time

Total Travel Time is one measure used to quantify operational deficiency and measure the performance of the freeway system. Expressed in units of vehicle-hours, it is the product of the total number of vehicles using the roadway during a given time multiplied by the average travel time of the vehicles. The equation that can be used to compute the Total Travel Time is as follows:

$$TTT_j = N_j \, \overline{tt}_j = \frac{N_j \, X_J}{\overline{V}_j}$$

where,

TTT_j = Total Travel Time over freeway section j, in vehicle-hours.

N_j = Number of vehicles traveling over freeway section j, during time period T.

$\overline{tt}_j$ = Average travel time of vehicles over freeway section j, in hours.

X_j = Length of roadway section j, in kilometers.

$\overline{V}_j$ = Average speed of vehicle over freeway section, in kilometers per hour.

The Total Travel Time, TTT, in vehicle-hours, for all sections of a freeway can be computed using the following equation:

$$TTT = \sum_{j=1}^{K} TTT_j$$

where,

TTT = Total Travel Time for all sections of the freeway, in vehicle-kilometers.

TTT_j = Total Travel Time for section j, in vehicle-kilometers.

K = Total number of freeway sections.

Total Travel

Total travel is another measure commonly used to quantify operational problems and measure the performance of the freeway system. Total travel is the product of the total number of vehicles using the freeway during a given time interval multiplied by the average trip length of the vehicles. It can be determined using the following equation:

$$TT_j = X_j \, N_j$$

where,

TT_j = Total Travel over freeway section j, in vehicle-kilometers.

X_j = Length of freeway section j during time period T, in kilometers.

N_j = Number of vehicles traveling over section during time period T.

Total Travel (TT) for all sections of freeway can be computed using the following equation:

$$TT = \sum_{i=1}^{K} TT_j$$

where,

TT = Total Travel for all sections of the freeway, in vehicle-kilometers.

TT_j = Total travel for section j, in vehicle-kilometers.

K = Total number of freeway sections.

Vehicle Delay

Vehicle delay is another common measure used to quantify the level of performance of the freeway system. For freeways, vehicle delay is defined as the increase in travel time on a route over the free-flow travel time. Vehicle delay can be computed by subtracting the total free-flow travel time from the total travel time. Total free-flow travel time is the product of the total number of vehicles using the roadway during a given time interval multiplied by the average free-flow travel time.

Total Minute-Miles of Congestion

Originally developed to measure the effectiveness of the Chicago Area Freeway Surveillance Project, the total minute-miles of congestion method was used to indicate the extent of freeway congestion in both time and space.[9] (Minute-Miles was the measure of performance at the time of the study. Minute-kilometer is also appropriate.) To compute the total minute-miles of congestion, each mainline detector is assigned a section of freeway that covers half the distance between the adjacent mainline detectors on either side. The minute-miles of congestion at each detector is equal to the product of the minutes of congestion at the detector, multiplied by the distance, in miles, assigned to the detector. The sum of the minute-miles of congestion

of all detectors then would represent the total minute-miles of congestion.

Isolated Locations

Congestion caused by incidents and bottlenecks tends to be located at isolated locations. Engineers evaluating these locations are primarily concerned with the following issues:

- What is the cause of the congestion?
- How quickly can the congestion clear?
- What are the effects of the congestion on drivers through the congested area?
- How large is the effected area?

The types of measures that are commonly used to quantify the impacts of congestion at isolated locations include the following:

- Time until normal flow.
- Total vehicles delayed.
- Average delay per vehicle.
- Maximum time of queue.
- Maximum queue length.
- Fuel consumption.
- Air pollution.

Most of these measures can be computed by conducting a bottleneck/queuing analysis, which is discussed below.

METHODS TO QUANTIFY OPERATIONAL PROBLEMS ON FREEWAYS

The types of analytical methods used to quantify the freeway operational problems include the following:

- Capacity and Level of Service Analysis.
- Bottleneck and Queuing Analysis.
- Computer Simulation.
- Field Measurements.

How these methods are used to quantify operational problems on freeways is discussed in more detail below.

Freeway Capacity Analysis and LOS

One of the most basic types of analysis required for a section of freeway is a determination of the amount of traffic-carrying capacity that the section can provide. Capacity is defined as the maximum number of vehicles that can reasonably be expected to use the facility in a given time period under prevailing roadway, traffic, and control conditions.[10] Related to this definition is the concept of the operational quality or level of service (LOS) provided to users of a facility.

Definition of Freeway Subareas for Capacity Analysis

Freeways are composed of three types of component subsections. These are defined as follows:

- *Basic Freeway Segments* -- These are freeway sections that are unaffected by either merging or diverging traffic

movements at nearby ramps or by weaving movements.

- *Weaving Segments* -- These are freeway sections where two or more vehicle flows must cross each other's path. Weaving areas exist where merge areas are closely followed by diverge areas, and where a freeway on-ramp is followed closely by an off-ramp and an auxiliary lane is used to connect the two.

- *Ramp Junctions* -- These are freeway sections where on- and off-ramps join the freeway.

Figure 2-7 illustrates the various types of freeway components. Capacity analyses treat each one of these components with a separate computational procedure. After the components have been evaluated in isolation, the results are laid out together to assess how the freeway operates as a system. A system analysis perspective is particularly important in the evaluation of closely-spaced ramp effects upon the capacity and operating characteristics of the freeway.

Basic Freeway Segments

Definition of Ideal Freeway Capacity. According to the 1994 edition of the Highway Capacity Manual, the capacity of basic freeway segments under ideal roadway, traffic, and environmental conditions is assumed to be 2,200 passenger cars per hour per lane (pcphpl) on four-lane freeway sections and 2,300 pcphpl on six-or-more lane freeway sections.[10] These capacity values represent the maximum number of vehicles that can reasonably be expected to use the facility (an average over all lanes) in a given time period under prevailing roadway, traffic, and control conditions. Flow rates in excess of these values can sometimes occur in individual lanes, particularly in inside or median lanes where there is no influence of upstream or downstream ramps.

Conditions that are "ideal" from a freeway capacity perspective include the following:

- 3.6 m (12 ft) minimum lane widths.

- 1.8 m (6 ft) minimum lateral clearance between the edge of the travel lane and the nearest roadside or median obstacle influencing traffic behavior.

- All passenger cars in the vehicle stream (no trucks, buses, or recreational vehicles).

- Motorists who are regular and familiar users of the freeway.

Factors Affecting Capacity. Obviously, conditions more restrictive than those listed above reduce the traffic-carrying capacity of the freeway. For example, lane widths narrower than 3.6 m (12 ft) cause drivers to travel closer to one another laterally, and so they adjust by increasing the distance between themselves and the vehicle in front of them. Roadside obstructions closer than 1.8 m (6 ft) to the roadway cause drivers to shy away from the edge of the roadway, and they again respond by increasing the gap between their vehicles and the vehicles they are following. Trucks, buses, and recreational vehicles are larger than passenger cars and have more sluggish operating characteristics, which create larger gaps between them and passenger vehicles. Finally, the characteristics of the driving population have been shown in several studies to affect the capacity of a freeway segment. Nonfamiliar users of a facility are generally more cautious and therefore leave more room between themselves and vehicles they are following.

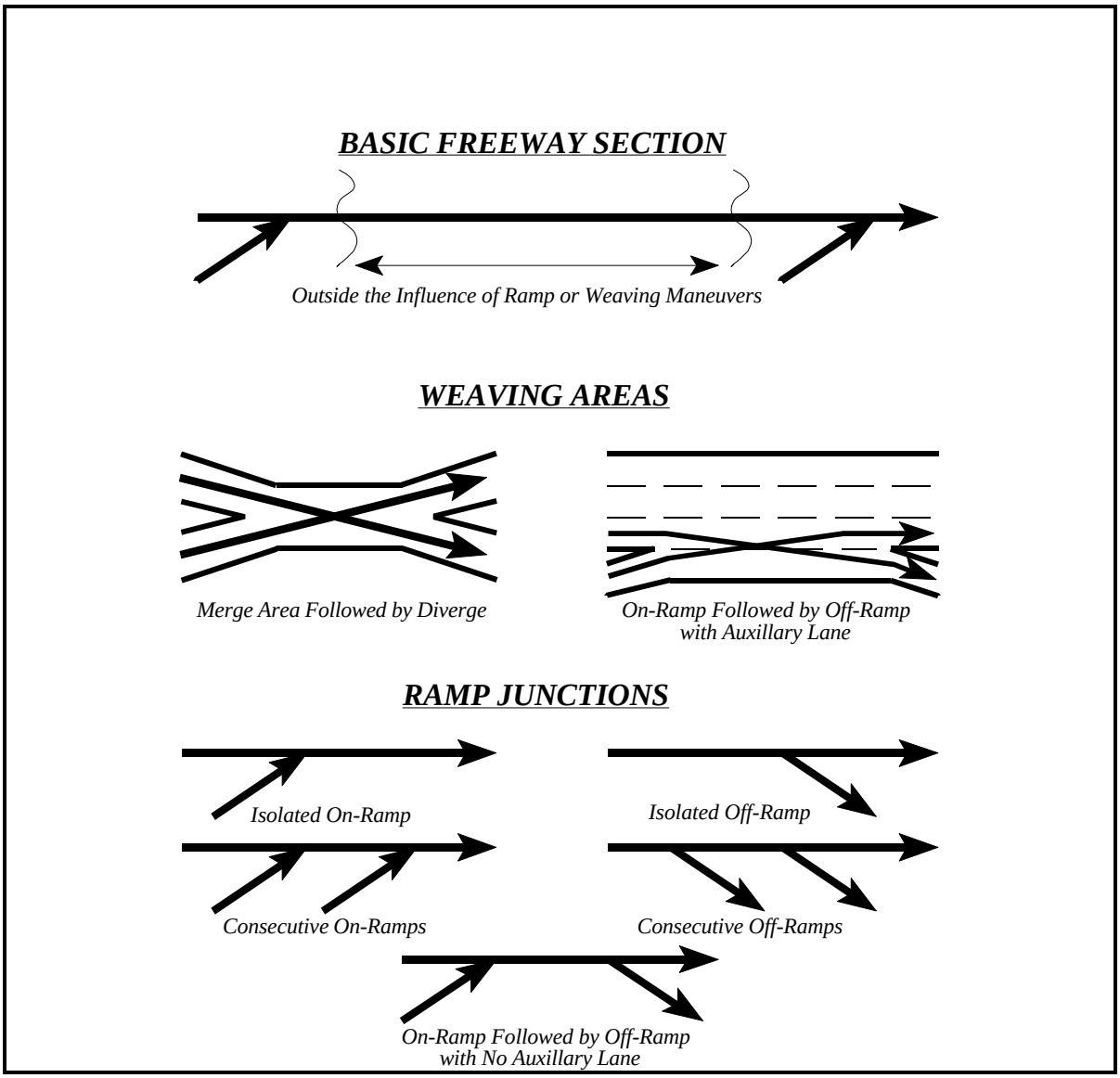


Figure 2-7. Illustration of Freeway Components for Capacity Analysis. [11]

Traffic operations and the capacity of basic freeway segments are also affected by horizontal and vertical alignment. Specifically, sharper horizontal curves and longer, steeper grades reduce the traffic-carrying capacity of the freeway segment.

Level of Service. As indicated previously, the concept of LOS was introduced into operational analysis as a means of quantifying the quality of operation attainable under a given roadway and traffic configuration. Although the traffic-carrying capacity of a basic freeway segment is defined in terms of vehicular flow, density (number of vehicles per lane per kilometer (mile)) is the parameter used to define LOS for basic freeway segments. LOS is defined by the letters A through E, with LOS F used to define breakdown, stop-and-go conditions. Table 2-3 gives a qualitative description of each level of service, and the density value that has been associated with each level.

Weaving Areas

Definition of Weaving Area Capacity. Weaving is defined as the crossing of two or more traffic streams traveling in the same general direction along a significant length of freeway. Weaving areas represent a region of intense lane-changing behavior and increased traffic flow turbulence as drivers attempt to access lanes to reach their desired exit points from the weaving area. Many different roadway and traffic factors can influence operations within these areas. Because of this complexity, the capacity of weaving areas as estimated by procedures in the *Highway Capacity Manual* is not defined in terms of maximum flow rates or vehicle densities, but in terms of estimated operating speeds.[10] Specifically, calculations are made to estimate the speeds expected of weaving and of nonweaving vehicles. These are then compared to tabular values to determine the LOS being provided by that weaving section.

Types of Weaving Areas. Because lane-changing is the critical operational feature of weaving areas, weaving sections are categorized according to the minimum number of lane changes that must be made by weaving vehicles.

Weaving areas are referred to as Type A, Type B, or Type C sections. Figure 2-8 illustrates each type of weaving area.

Type A weaving areas require that each weaving vehicle make one lane change in order to execute the desired movement. These areas may be ramp-weave areas created by adjacent entrance and exit ramps connected by an auxiliary lane, or major weaving areas characterized by three or more entry and exit roadways having multiple lanes.

Type B weaving areas may also be referred to as major weaving sections if they involve multilane entry and/or exit lanes. However, these areas differ from Type A areas in that one weaving movement can be accomplished without changing lanes, whereas the other weaving maneuver requires no more than one lane change.

Type C weaving areas are similar to Type B areas in that one weaving movement can be made without changing lanes. The major difference between these two types of areas is that Type C areas require the other weaving maneuver to make two or more lane changes. This can be an effective design if the second weaving flow is fairly small. However, it can have very adverse effects on operations if the second flow is large, the number of lane changes being made is large, and the length of the weaving area is fairly short.

Operational Characteristics of Weaving Areas. Traffic operations within weaving areas are affected by the following geometric features:

- Weaving length.
- Weaving area configuration.
- Number of lanes (weaving width).

Shorter weaving lengths create greater traffic flow turbulence as drivers are forced to make necessary lane changes over a limited distance. The configuration of a weaving area influences the number of lane changes that must be made and also the proportion of vehicles that must make a weaving maneuver. The number of lanes available within the weaving area defines the amount of space available to weaving vehicles to make the various types of lane changes.

Table 2-3. Level of Service (LOS) Descriptions. [11]

Level of Service	Status	Density Threshold, pc/km/ln (pc/mi/ln)	Description
A	Free Flow	6.25 (10)	Individual motorists are unaffected by the presence of others in the traffic stream. Freedom to select desired speeds and to maneuver within the traffic stream is high.
B	Stable Flow	10 (16)	Freedom to select desired speeds is relatively unaffected, but there is a slight decline in the freedom to maneuver within the traffic stream. The presence of others in the traffic stream begins to affect individual behavior.
C	Stable Flow	15 (24)	The behavior of individual motorists is significantly affected by interactions with others in the traffic stream. Selection of speed is affected by other vehicles and maneuvering within the traffic stream requires vigilance by the motorist.
D	High-Density Stable Flow	20 (32)	Speed and freedom to maneuver are severely restricted, and the driver experiences a generally poor level of comfort and convenience. Small increases in traffic flow will generally cause operational problems.
E	Unstable Capacity Flow	23-30 (36.7-47.9)[a]	Operations are at or near capacity. Speeds are lower, but relatively uniform among vehicles. Freedom to maneuver is extremely limited. Operations are unstable.
F	Forced or Breakdown Flow	>23-30 > (36.7-47.9)[a]	Amount of traffic approaching a point exceeds the amount which can traverse the point. Queues form and operations are characterized by unstable stop-and-go waves.

[a] Capacity occurs at different densities depending on the design speed of the freeway

If the weaving length, configuration, and width in combination with the traffic demand pattern permit the weaving and nonweaving vehicles to spread out evenly across the available lanes in the weaving section, the operation of the weaving area is more effective and is classified as **unconstrained**. Conversely, if the configuration and traffic

demand limit the ability of weaving vehicles to occupy their proportion of available lanes to maintain balanced operations, the operation is less effective and is classified as **constrained**. Imbalanced or constrained operations within weaving sections will result in weaving vehicles traveling at a lower average speed than nonweaving vehicles. Not only will the LOS being provided by the weaving section be lower because of this condition, but also the general safety of the section will be lower because of a mixing of lower and higher speed traffic flows.

Capacity and Level of Service for Weaving Areas. In weaving sections, operations are estimated to be at capacity when estimated weaving and nonweaving speeds are between 48 and 56 km/h (30 and 35 mph). Table 2-4 summarizes the weaving and nonweaving speed thresholds that have been defined for the other levels of service. Those average speeds are based on regression equations.

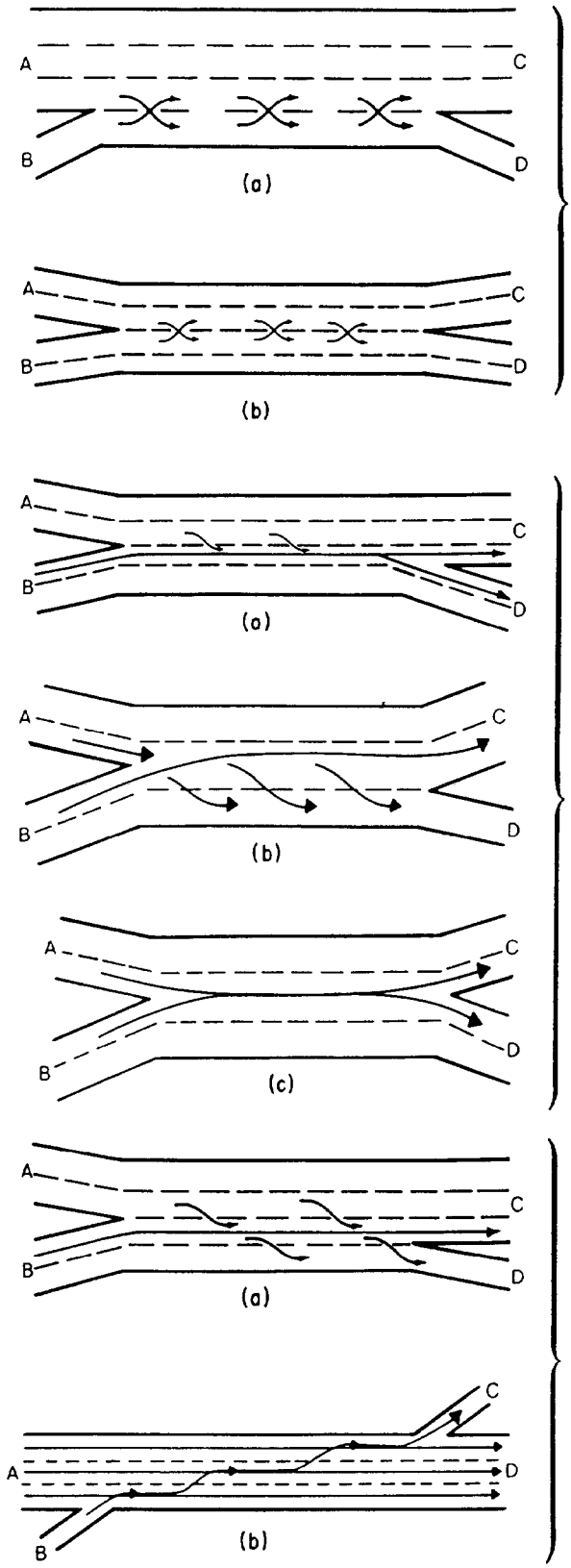
The regression constants vary according to the weaving configuration of the section and whether or not the weaving section operation is unconstrained or constrained. Both the weaving and nonweaving speeds must satisfy the criteria specified in table 2-4 for a given LOS. For example, the average weaving speed must exceed 80 km/h (50 mph) and the average nonweaving speed must exceed 90 km/h (54 mph) in order for the weaving section to be considered as operating at LOS B.

It is important to note that the above LOS criteria have been established in conjunction with the regression equations, and only provide a mathematical procedure for assessing the general operations at weaving areas. Considerable variation may exist between the estimates obtained via the mathematical procedures and the actual operations observed at a given weaving area. Research continues to improve the estimates of weaving area operations. The formulas and thresholds presented in this chapter should be supplemented with sound engineering judgement and actual field data as necessary until a better understanding and mathematical formulation are available.

Ramp Junctions

Definition of Ramp Junction Capacity. The points at which vehicles enter a freeway mainlane from an on-ramp or the point at which mainlane traffic diverges to an off-ramp are termed freeway-ramp junctions. A freeway-ramp junction is an area of competing traffic demand for space. Analyses of freeway-ramp junctions focus on the merging and diverging behaviors required in the outer two freeway lanes and the acceleration or deceleration lane in the vicinity of on- and off-ramps, respectively, as depicted in figure 2-9. Elements such as the length and type (e.g., taper, parallel) of acceleration and deceleration lanes, free-flow speed of the ramp in the immediate vicinity of the junction, and sight distances all influence traffic operations within these regions.

For freeway-ramp junctions, two capacity values are important. The first refers to the requirement that the capacity of the freeway lanes (plus the off-ramp lane in the case of diverging freeway-ramp junctions) where vehicles *exit* the region of ramp influence must not exceed the capacity of those lanes as defined for basic freeway segments. This value is 2,200 pcphpl for four-lane facilities or 2,300 pcphpl for freeway facilities with six or more lanes. The second capacity value of interest is the freeway traffic flow *entering* the ramp influence area. This cannot exceed 4,400 pcph on four-lane



Type A weaving areas: (a) ramp-weave/one-sided weave, and (b) major weave with crown line.

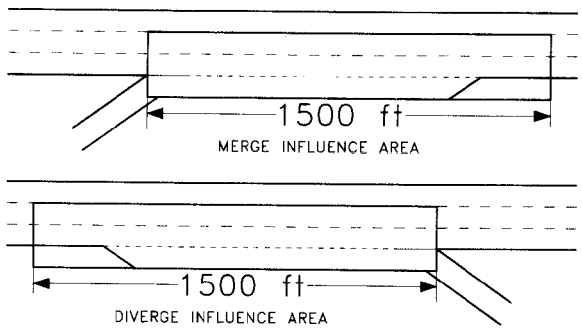
Type B weaving areas: (a) major weave with lane balance at exit gore, (b) major weave with merging at entrance gore, and (c) major weave with merging at entrance gore and lane balance at exit gore.

Type C weaving areas: (a) major weave without lane balance or merging, and (b) two-sided weave.

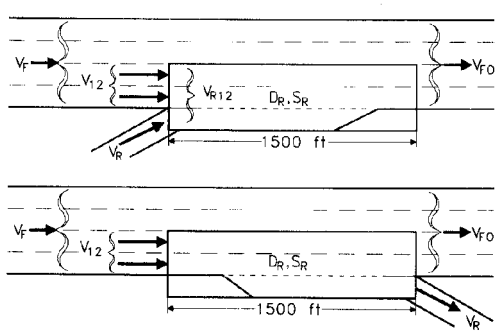
Figure 2-8. Categories of Weaving Areas. [13]

Table 2-4. Level-of-Service Criteria for Weaving Sections. [11]

Level Of Service	Minimum Average Weaving Speed, S_w km/h (mph)	Minimum Average Nonweaving Speed, S_{nw} km/h (mph)
A	90 (55)	100 (60)
B	80 (50)	90 (54)
C	70 (45)	77 (48)
D	60 (40)	67 (42)
E	50/60 (35/30)	50/60 (35/30)
F	<60/50 (<35/30)	<60/30(<35/30)



On- and Off-Ramp Influence Areas



Critical Ramp Junction Values

Figure 2-9. Influence Areas and Critical Flow Values for Freeway-Ramp Junctions. [11]

facilities or 4,600 pcph on six- or more lane facilities.

Operational Characteristics of Ramp Influence Areas. In the merge area at an on-ramp, individual merging vehicles from the ramp attempting to enter the outer freeway lane (lane 1) create turbulence in the traffic stream in the vicinity of the ramp. Approaching freeway vehicles therefore tend to move towards the left to avoid this turbulence. As ramp volumes increase, more and more freeway vehicles will move to the left as long as there is available capacity to do so. However, the interaction between freeway and ramp vehicles in the influence area is quite dynamic. Although the intensity of the ramp demand flows most generally affects what percentage of approaching freeway vehicles move to the left, any congestion that develops on the freeway itself affects the ability of ramp vehicles to enter the freeway and can cause diversion of some of the ramp traffic to other routes or ramps in the vicinity, or excessive ramp queuing.

In the diverge area of an off-ramp, the basic maneuver being performed is the separation of a single traffic stream into the through and exiting movements. Since exiting vehicles must occupy lane 1 in order to utilize the off-ramp, they begin moving towards that lane in the influence area. This in turn also causes through freeway vehicles to begin moving towards the left in order to avoid the turbulence of the diverge area.

These operational characteristics serve as the basis of current freeway-ramp junction analytical procedures included in the 1994 *Highway Capacity Manual*.[10] Regression equations are presented in that manual to estimate the percentage of freeway traffic that remains in the two lanes closest to the ramp in the influence area. The choice of the appropriate regression equation depends on the following factors:

- The number of freeway lanes in the influence area.
- The presence of nearby on- or off-ramps.

Level of Service of Freeway-Ramp Junctions. Although the capacity analysis procedures for freeway-ramp junctions focus on the flow rates entering and exiting the influence area, traffic flow is not the primary measure used to evaluate the LOS that exists within a ramp influence area. Rather, as is done for basic freeway segments, vehicle density is the primary measure used to define levels of service. Table 2-5 provides the density thresholds corresponding to each level of service in ramp influence areas. Regression equations presented in the Highway Capacity Manual are used to estimate volumes, densities, and speeds in the ramp influence area.

As shown in table 2-5, the density threshold values for each LOS are somewhat higher than for basic freeway segments. This is because drivers naturally expect some increased turbulence and closer proximity of other vehicles in a merge or diverge area, and because drivers are typically traveling at a slightly slower speed in any given lane in the influence area than they would on a similar section of open freeway.

Freeway Systems

The freeway is a complex facility made up of many component segments, each having a potential impact on operations in upstream and downstream segments. Capacity and LOS analyses discussed to this point have addressed each of the major types of freeway components. After the component analysis has been completed, the final step in such an analysis is to put the various components

Table 2-5. Level of Service (LOS) Descriptions for Freeway-Ramp Junction Influence Areas. [11]

Level of Service	Status	Density Threshold, pc/km/ln (pc/mi/ln)	Description
A	Unrestricted Operations	6.25 (10)	Merging and diverging maneuvers occur without disrupting freeway flow. No noticeable turbulence exists. Freeway speeds are nearly as high as in basic freeway sections.
B	Stable Operations	12.5 (20)	Merging and diverging maneuvers become noticeable to motorists, but only minimal levels of turbulence exist. Merging vehicles must adjust their speeds in order to move into available gaps in the freeway traffic stream.
C	Stable Operations	17.5 (28)	Average speeds in the ramp influence area begin to decline. Driving conditions are still relatively comfortable at this level.
D	High-Density Stable Operations	21.9 (35)	Turbulence levels become intrusive. All vehicles in influence area slow to accommodate those making merging and diverging maneuvers. Freeway operations tend to remain stable.
E	Capacity Operations	> 21.9 (> 35)	Operations are at or near capacity. Speeds are lower. Freedom to maneuver is extremely limited and intrusive to all drivers. Operations are unstable.
F	Breakdown Operations	[a]	Traffic attempting to use influence area exceeds the amount which can traverse the area.

[a] Density criteria is not valid or relevant under breakdown conditions

together to get an overall systemwide picture of anticipated freeway operations.

Generally speaking, capacity and LOS analyses can be oriented towards design (determining necessary number of lanes, locations of ramps, weaving area lengths, etc.) or operations. In this section, the basic steps for an operational analysis of a freeway system are outlined.

As the first step, the analysis of individual freeway components occurs as per the techniques described earlier.

1. Basic freeway segments are evaluated using the procedures discussed in chapter 3 of the *Highway Capacity Manual*.[10]

2. Ramp junctions are evaluated using the procedures in chapter 4 of the *Manual*, considering each ramp as follows:[10]

- As an isolated ramp.

- In conjunction with the adjacent downstream ramp.

- In conjunction with the adjacent upstream ramp.

  Ramps that are clearly operating as part of a weaving section would not be evaluated with these procedures.

3. Weaving sections are evaluated using the procedures documented in chapter 5 of the *Manual*.[10]

When a given component falls under more than one of these categories, the analysis category that results in the lowest LOS is used as the controlling factor defining operations within that component.

Once the analysis results of the various freeway components have been computed, a graphical procedure can then be used to obtain an overall perspective of the operating conditions of the freeway system as a whole. The general freeway alignment is plotted along with an indication of the level of service that has been predicted for each particular freeway component or segment. Figure 2-10 illustrates this pictorial layout of level of service for a hypothetical freeway section.

As the illustration shows, the weaving area in section 4 of this hypothetical situation is a candidate bottleneck situation that will likely be the segment that controls the operation of the overall system. Whereas the weaving section is anticipated to operate near level of service E, the upstream segments **could** operate at levels of service B or C as long as the weaving section does not break down and create queues and congestion that propagates upstream into these segments. However, if flows are slightly higher than anticipated, it is likely that the weaving segment would break down very quickly and lead to operational problems within those upstream segments as well.

Highway Capacity Software (HCS)

Capacity and LOS analyses for each type of freeway segment can be performed manually using the procedures, nomographs, tables, and worksheets that are provided in the *Highway Capacity Manual*. However, these computations can be quite repetitive and time-consuming for the analysis of many design or operational alternatives. Fortunately, these procedures have been automated, and can be performed quite easily on a personal computer using the *Highway Capacity Software (HCS)*, developed under FHWA sponsorship and maintained by the McTRANS Center at the University of Florida.

The HCS is a macroscopic, primarily empirical, deterministic simulation program used to quickly and easily evaluate traffic flow conditions at specific freeway features. The program contains separate modules for analyzing basic freeway segments, freeway-ramp junctions, weaving areas, and freeway systems. Modules for analyzing signalized and unsignalized intersections are also included (these may be needed when

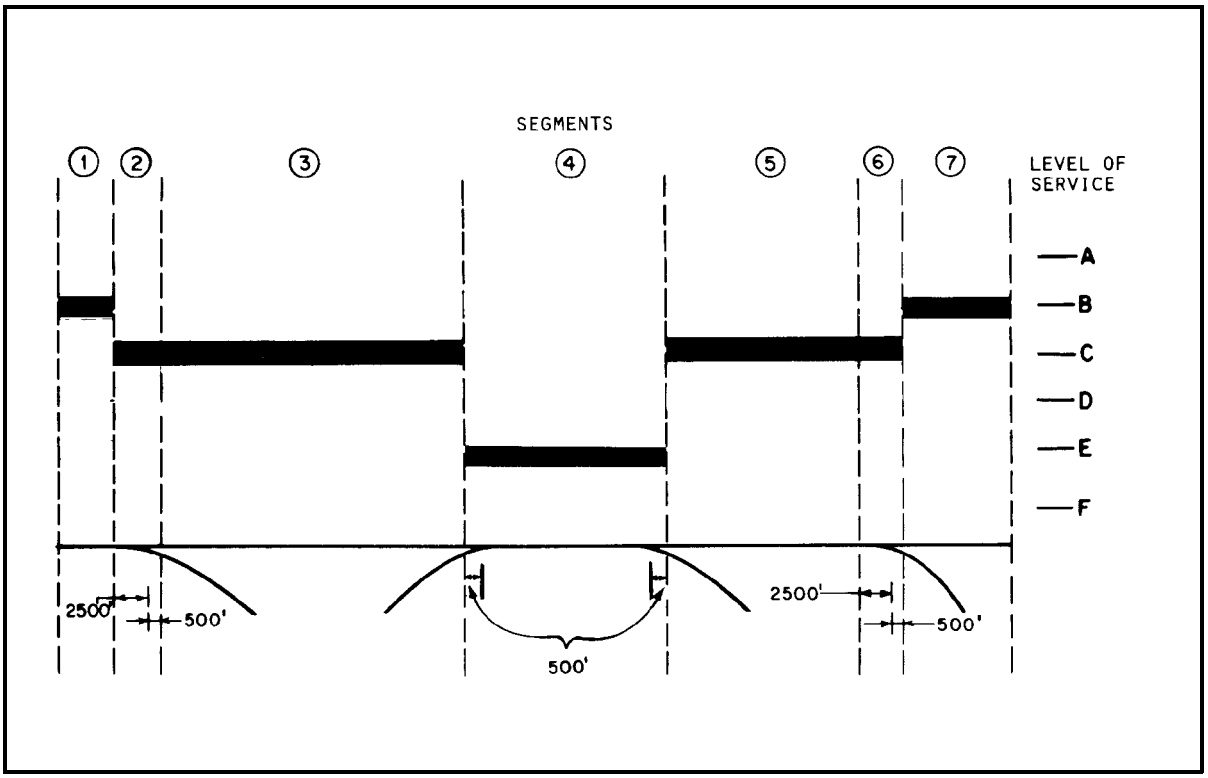


Figure 2-10. Graphic Representation of Level of Service for Freeway Systems. [11]

evaluating ramp-arterial street terminals at the ends of exit ramps).

The HCS (and the HCM procedures in general) are most appropriate for quick analyses of individual freeway components or for specific locations. If a more accurate and detailed evaluation of freeway operations as a system is required or desired, it is necessary to revert to the more sophisticated traffic simulation and optimization models in order to fully examine traffic operating conditions. These simulation and optimization models are addressed below.

BOTTLENECK/QUEUING ANALYSES

The state of the freeway (as determined by its flow-speed-density relationship) changes over space and time. When these changes of state occur, a boundary is established that denotes the time-space domain of one flow state as distinguishable from another. These boundaries are referred to as shock waves. As shown in figure 2-11, when a change in state occurs on a high-volume freeway, the queue and its resulting congestion begin to back upstream from the scene of the bottleneck. The boundary that denotes this change in state is a shock wave.

As shown in figure 2-12, there are four common types of shock waves. A *frontal stationary* shock wave always forms at a bottleneck location and occurs when demand exceeds capacity. It can be caused by either recurrent or nonrecurrent congestion situations. The term *frontal* implies that the shock wave is at the front (or downstream edge) of the congested region. It is called *stationary* because the shock wave is fixed by the location and does not change its location over time.

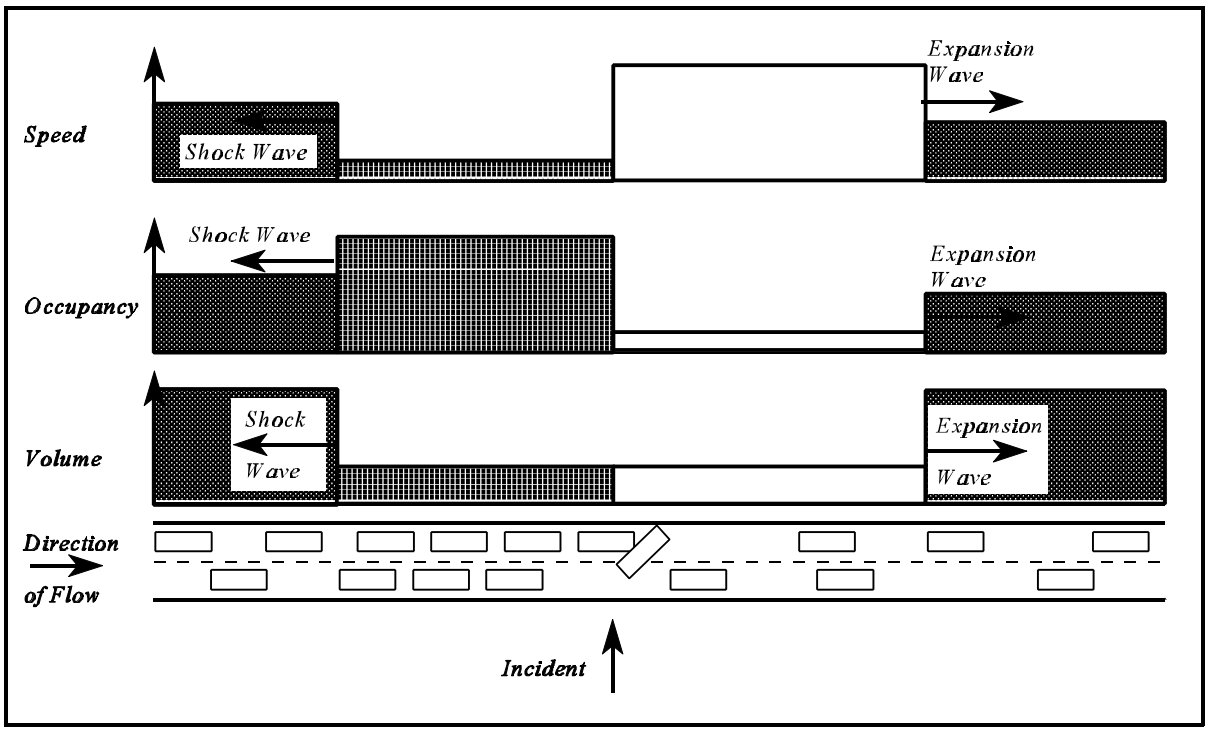


Figure 2-11. Fundamentals of Shock Wave Analysis. [15]

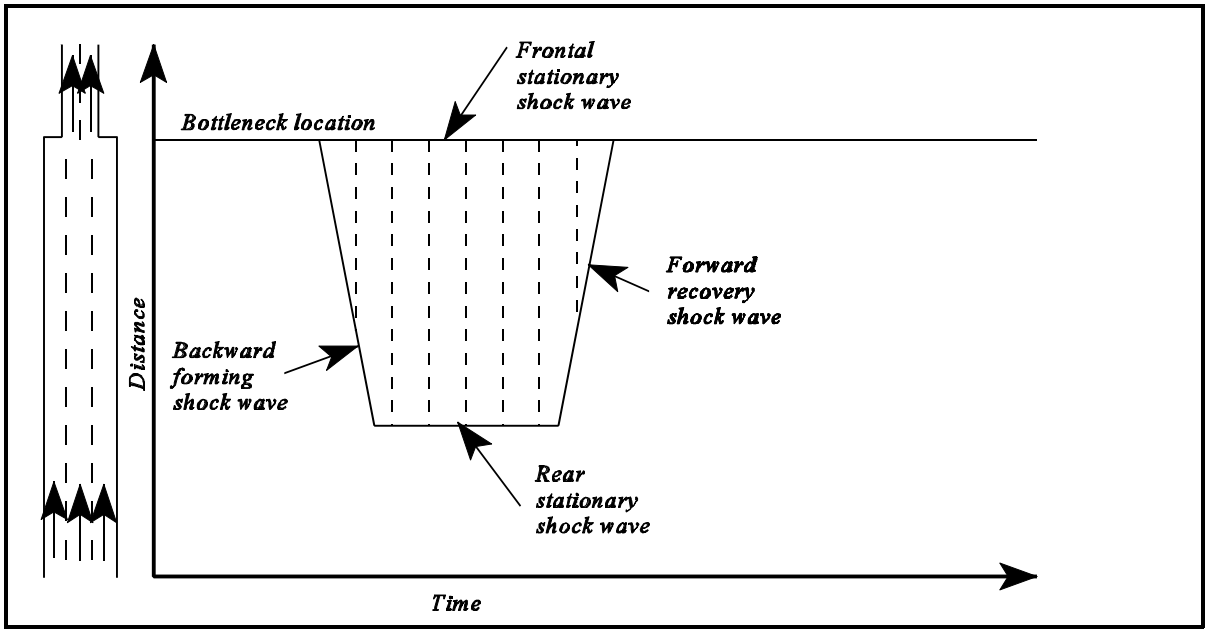


Figure 2-12. Shock Wave Phenomena at a Freeway Bottleneck During a Peak Period. [14]

A *backward forming* shock wave is also always present when congestion occurs. It is formed by the area where excess demands are being stored (i.e., the queue). It is termed *backward* because, over time, the shock wave moves backwards or upstream in the opposite direction of traffic. The term *forming* implies that, over time, the congestion gradually extends farther upstream. The slope of the wave represents how fast the shock wave is traveling upstream (e.g., a flatter slope implies that the shock wave is moving slowly upstream, while a steeper slope implies that the shock wave is moving rapidly upstream). A backward forming shock wave is the most commonly occurring type.

A *forward recovery* wave is the next most commonly encountered type of shock wave. It occurs when there has been congestion but demand decrease is below the bottleneck capacity. It forms as the length of congestion reduces. It is a *forward* wave because it moves downstream in the same direction as the flow of traffic. The term *recovery* implies that free-flow conditions are gradually occurring as the wave moves forward.

A *rear stationary* shock wave occurs when the arriving traffic demand is equal to flow in the congested region for some period (i.e., the queue is neither growing or dissipating). It is a *rear* wave because it occurs at the rear (or upstream edge) of the congested region with higher densities downstream and lower densities upstream. It is *stationary* because the shock wave does not change its location over time.

Another type of shock wave encountered on freeways is the *backward recovery* shock wave. This wave commonly occurs after the source of congestion (e.g., an incident) has been removed from the freeway. After the capacity has been restored, the discharge rate through the congested area exceeds the flow rate in the congested region. This causes the shock wave to move upstream (or *backwards*) in the opposite direction from traffic. It is a *recovery* wave because over time, free-flow conditions are extending further and futher upstream from the previous bottleneck location.

A final type of shock wave that may be encountered is the *forward forming* shock wave. This type of shock wave moves in the same direction of traffic. It is a *forming* wave because, over time, congestion is gradually extending further and further downstream. This type of wave is not commonly observed with freeway operations.

Queues form when demand on a freeway facility exceeds the capacity for an interval of time. To conduct a queuing analysis the following inputs are required: [14]

- The mean (or average) arrival flow rate (vph).
- The distribution of arriving vehicles.
- The mean (or average) service flow rate (vph).
- The distribution of vehicles being serviced.
- The queue discipline.

"First In First Out (FIFO)" is the commonly used queuing discipline used in transportation engineering. With FIFO, vehicles are serviced in the order in which they arrive. Another queuing discipline that can be used is "first in last out (FILO)." This type of queuing discipline would be used to model how queues might dissipate from behind a slow-moving vehicle on a multilane freeway.

Queuing analyses can be conducted at two levels of detail. With the first, queues are analyzed at a macroscopic level where arrival patterns and service patterns are considered to be continuous. The other level of detail is microscopic, and is conducted where arrival and service patterns are considered to be discrete.

One situation where queuing analysis is commonly used in freeways is in analyzing the impacts of temporary blockages (caused by incidents or work zone activities) on traffic flow. Figure 2-13 provides an illustration of this situation. Figure 2-13a provides all the input requirements needed to analyze this situation. The arrival rate (λ) is specified in vehicles per hour and is constant for the analysis period. The normal service rate (without the blockage) is indicated in the diagram as μ, and since it exceeds the arrival rate, no queuing would normally exist. However, an incident occurs that reduces the service rate to μ_R, which is below the arrival rate, and this lower service rate is maintained for t_R hours. As is the case in most freeway situations, a FIFO queue discipline is assumed.

In figure 2-13b, a cumulative vehicles versus time diagram is constructed. The arrivals are shown as a straight line passing through the origin with a slope up and to the right equivalent to the arrival rate (λ). For the first period, the service line follows the arrival line, until the blockage occurs. At that point, the service rate becomes equivalent to μ_R and maintains a flatter slope until the blockage is removed. When the service rate increases to μ, the service line asumes a steeper slope. This continues until the arrival line and service line intercept, at which time the service line once again overlays the arrival line. A triangle is formed with the cumulative arrival link forming the top side of the triangle and the cumulative service line forming the other two sides of the triangle. The triangle represents the congestion that is caused by the bottleneck. Table 2-6 summarizes the equations used to analyze this queuing situation.

FREEWAY SIMULATION/ OPTIMIZATION MODELS

Capacity and LOS analysis is a very useful tool for gauging the expected operating conditions at certain freeway locations, and for determining the magnitude-of-scale changes that would result in those operations if a major freeway improvement were made. However, the characteristics of the freeway system and/or the traffic demands on the freeway create conditions that cannot be easily or accurately evaluated with the capacity analysis techniques. The changes to the freeway may be too subtle in nature to be addressed by the capacity and LOS analysis. In addition, the capacity analysis procedures only indicate where traffic operations are expected to break down or the LOS being provided by a particular roadway and traffic configuration. Information on total and individual vehicle delays, stops, fuel consumption, vehicle emissions, and other measures are not provided by capacity analysis techniques. Consequently, it is often necessary to turn to more complicated but powerful traffic simulation or optimization analyses.

The following sections provide an overview of the major traffic simulation and optimization computer models available for freeway operations analysis. The models are discussed according to the analysis approach used to represent traffic flow in each. First, the macroscopic models are described. Then, models that provide a microscopic analysis of freeway operations are reviewed.

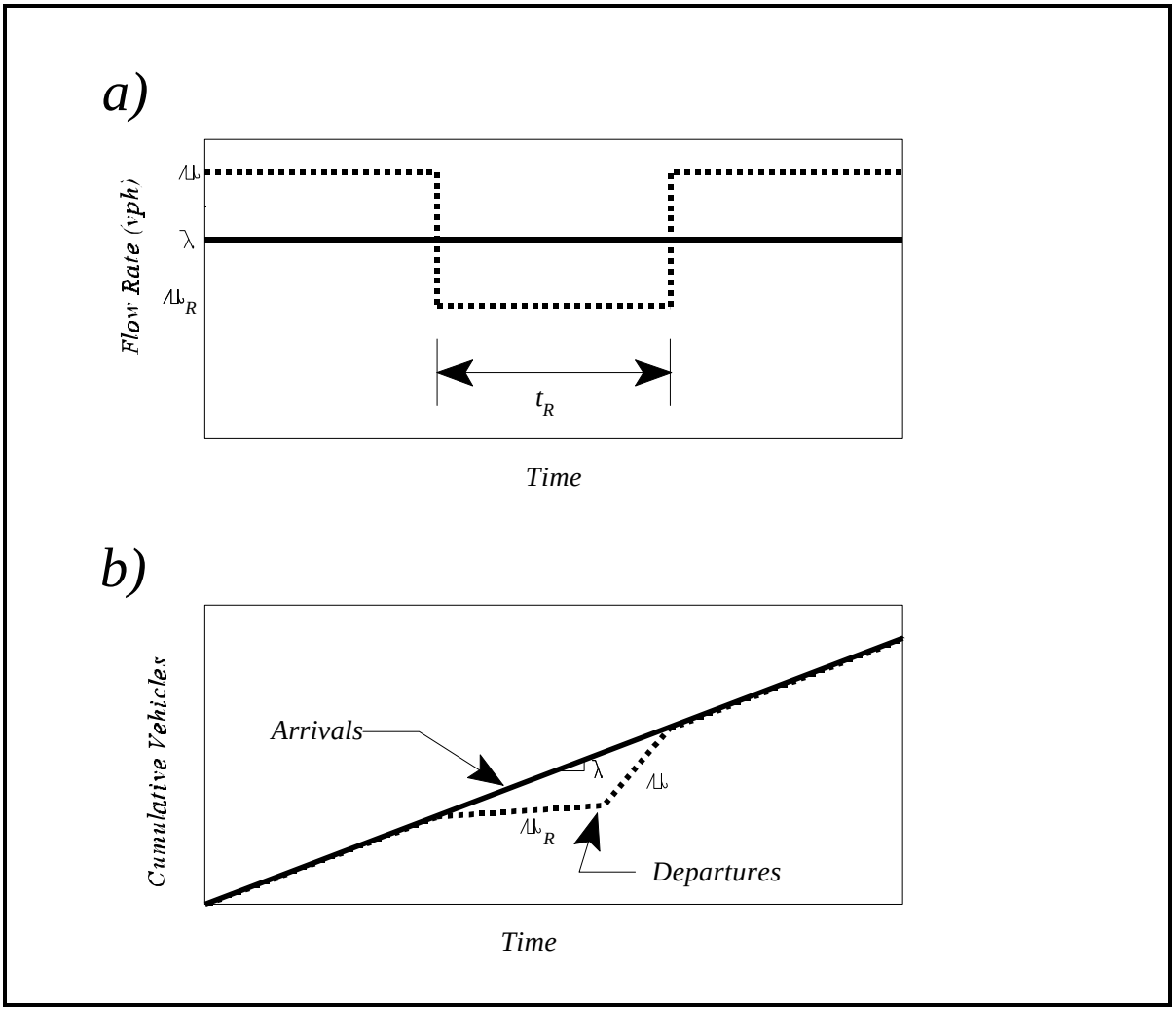


Figure 2-13. Use of Queuing Analysis to Analyze Temporary Blockage. [14]

Macroscopic Freeway Simulation/Optimization Models

One of the principal advantages of most freeway simulation/optimization models is their incorporation of the time element into the analysis. Typically, this is accomplished by dividing the duration of analysis into incremental units, or "slices," of time over which traffic demands and freeway characteristics are approximately constant. The time slices are then analyzed sequentially to allow the traffic impacts of previous slices to be considered in the evaluation of the current time slice. In a similar manner, the spatial relationships between various components on the freeway are specified by defining homogenous segments of freeway in the program and linking those segments together to form the freeway section of interest. These segments have uniform characteristics in terms of traffic demand, number of lanes, basic geometry, etc.

A number of simulation and optimization models have been developed over the years to assist in the analysis of macroscopic traffic flow behavior on freeways. Characteristics of the three such programs that are most widely used are discussed below:

Table 2-6. Queuing Performance Equations for Freeway Lane Blockage. [14]

Performance Measure	Equation
Time duration in queue, t_Q (hr)	$\frac{t_R(\mu - \mu_R)}{\mu - \lambda}$
Number of vehicles queued, N_Q (vehicles)	λt_Q
Maximum queue length, Q_M (vehicles)	$t_R(\lambda - \mu_R)$
Average queue length while queue present, $\overline{Q}_A$ (vehicles)	$\frac{t_R(\lambda - \mu_R)}{2}$
Maximum individual delay, d_M (minutes)	$\frac{60t_R(\lambda - \mu_R)}{\lambda}$
Average individual delay while queue present, d_A (minutes)	$\frac{30t_R(\lambda - \mu_R)}{\lambda}$
Total delay TD (vehicle-hours)	$\frac{t_R t_Q(\lambda - \mu_R)}{2}$

λ = the mean arrival flow rate (vph)
μ = Capacity flow rate (vph)
μ_R = Capacity flow rate during incident (vph)
t_R = duration of incident (hrs)

- FREQ.
- CORFLO.
- QUEWZ-92.

FREQ

FREQ is a macroscopic, analytic, deterministic, simulation and optimization program designed to evaluate freeway operations in a single direction of travel.

The program has been tested and validated extensively, and is widely used to evaluate the impacts of temporary freeway lane blockages, various freeway lane and ramp configurations, and high-occupancy vehicle treatments (dedicated lanes, priority treatments at entrance ramps).[15,16] FREQ can also develop optimized timing plans for ramp metering, based on one of several optimization objectives (e.g., maximize vehicle throughput, vehicle-kilometers of travel, passenger throughput, or passenger-kilometers of travel). However, the program is not capable of evaluating freeway-to-freeway interchange operations.

Input data requirements for FREQ include the following characteristics for each freeway segment within the study section:

- Segment length.
- Number of travel lanes.
- Grade.
- Capacity.
- Design speed.
- Ramp configuration.
- Ramp capacity.

Traffic demand data required by FREQ are as follows:

- Mainline volume entering the upstream end of the first freeway segment in section.
- Mainline volume exiting the downstream end of the last freeway segment in section.
- All entrance and exit volumes within the section being analyzed.

The major types of output of the FREQ analysis are listed below:

- Freeway travel time (vehicle- and passenger-hours).
- Ramp and freeway delays (vehicle- and passenger-hours).
- Total travel time (vehicle- and passenger-hours).
- Total travel distance (vehicle- and passenger-kilometers (miles)).
- Speed and density contour maps.
- Average speeds.
- Fuel consumption.
- Vehicle emissions by pollutant type.
- Optimized ramp metering rates.

CORFLO

CORFLO is an FHWA-sponsored package of macroscopic, analytical, deterministic simulation models that allows the analyst to evaluate traffic operations on a freeway, a freeway corridor, or an entire roadway network. CORFLO consists of a freeway simulation model, FREFLO, an arterial street analysis tool, NETFLO, and a user-equilibrium traffic assignment program, TRAFFIC. These programs have a common coding format that allows a unified, interactive analysis of combined freeway-arterial street operations.

The freeway component, FREFLO, is capable of evaluating freeway-to-freeway interchanges, or even an entire freeway network. The program is capable of simulating ramp metering impacts on operations, but is not capable of optimizing

the metering rates. Temporary lane blockages and HOV treatments can also be evaluated using the FREFLO program.[15,17]

Data input requirements for FREFLO are similar to those for FREQ. These requirements are listed below:

- Segment length.
- Number of travel lanes for regular use and HOV operations.
- Capacity.
- Free-flow speed.
- Ramp configuration.
- Volumes entering at the upstream segment and at on-ramps.
- Traffic volume exiting percentages at off-ramps.
- Percentage of volumes which are trucks or carpools.

The measures-of-effectiveness attainable from FREFLO are as follows:

- Travel distance measures (vehicle and person trips, vehicle- and person-kilometers [miles] traveled).
- Travel time measures (vehicle- and person-minutes).
- Average speed measures (vehicle- and person-kilometers per hour [miles per hour]).
- Space mean speeds by vehicle type and lane type.
- Lane density by vehicle type and lane type.

QUEWZ-92

QUEWZ-92 (Queue and User Cost Evaluation of Work Zones) is a simple macroscopic, analytical, deterministic simulation program designed to evaluate the traffic impacts and additional road user costs of freeway work zone lane closures. The program can also determine acceptable times of the day when lane closures can be allowed without causing freeway queues to become too extensive.[15,18]

QUEWZ-92 relies on simplistic assumptions about the freeway segment being analyzed, treating the approach as a basic freeway segment. The data requirements of the program are listed below:

- Configuration of the work zone (how many lanes are open, how many lanes are normally available, how long is the work zone section).
- The schedule of the lane closure activities.
- Hourly freeway traffic volumes approaching the lane closure.

The output obtained from QUEWZ-92 consists of hourly and cumulative estimates of the measures listed below:

- Average speed through the work zone.
- Average queue length if one exists.
- Traffic volumes diverted because of excessive queuing.
- Additional road user costs.

Microscopic Freeway Simulation Models

In cases where macroscopic models do not provide the measures needed to properly

analyze a freeway situation, or the assumptions about traffic behavior inherent in the macroscopic models do not represent the condition being evaluated, it may be necessary to utilize a microscopic freeway simulation model for the analysis. Generally speaking, such models are extremely data intensive and time-consuming to set up, debug, and use in an analysis. However, they are very powerful tools for assessing the impacts of more subtle changes in freeway geometrics and for evaluating complicated freeway sections such as at weaving areas with unusual geometrics.

There are two main microscopic freeway simulation models available at the present time. These include FRESIM, an FHWA-sponsored model that is designed to fit within the TRAF system of computer models that FHWA has developed, and INTEGRATION, a proprietary model developed in Canada that links the freeway and arterial street systems together in a single modeling structure. A third model, INTRAS, was also developed under FHWA sponsorship and has been used in past microscopic analyses of freeway operations. However, the more recent FRESIM model is intended to replace INTRAS. Consequently, that model is not discussed in detail.

FRESIM

FRESIM is a microscopic, analytical, stochastic freeway simulation model. FRESIM is capable of simulating freeway mainlanes, ramps, freeway-to-freeway connectors, variations in grade, horizontal curvature and superelevation, lane additions and drops, temporary lane blockages, acceleration or deceleration lanes, and auxiliary lanes. The model can also simulate freeway surveillance via loop detectors, one of three incident detection algorithms based on the loop detector data, and ramp metering operations.[19] However, the model is not capable of optimizing these metering rates. Furthermore, the user has the ability to control very detailed driver characteristics such as the parameters of the lane-changing algorithm or the aggressiveness of the driving population (which affects the acceleration and deceleration rates and gap lengths that will be acceptable to drivers).

FRESIM requires extensive input data in order to fully analyze a given freeway section. Fortunately, default values are provided for many of the input requirements. However, the analyst should understand what these default values represent and be able to adjust them to local conditions if warranted. The following is a list of the data that can be entered into FRESIM on a segment-by-segment basis:

- Segment length and type (mainlane or ramp).
- Number of through and auxiliary lanes.
- Length of auxiliary lanes.
- Grades and superelevation.
- Pavement condition (dry, wet, asphalt, concrete).
- Free-flow speed.
- Average queue discharge headway.
- Lanes that are restricted to trucks.
- Turning/exiting percentages.
- Loop detector locations and characteristics.
- Incident characteristics (effect of lane blockage on capacity, effect of rubbernecking on capacity, duration of incidents).

- Location of lane additions or drops.
- Ramp metering settings and operating type (fixed, demand/capacity, speed control, gap acceptance).
- Traffic volumes entering and exiting the section.
- Car-following sensitivity factors.
- Pavement friction coefficients.
- Lane-changing characteristics (rates of compliance to warning signs, acceptable gap sizes, speed of lane change, acceptable acceleration rates).
- Operating characteristics by vehicle type.

Categories of output statistics computed by FRESIM for each segment and for the entire section include the following:

- Vehicles discharged from each segment.
- Vehicles making turning/exiting maneuvers.
- Average delay (actual travel time minus free-flow travel time) per vehicle.
- Average time spent moving.
- Average speed.
- Number of lane changes.
- Total travel distance (vehicle-kilometers [vehicle-miles]).
- Volume.
- Density.
- Speed.
- Vehicle emissions.

INTEGRATION

INTEGRATION is also a microscopic, analytical, stochastic traffic simulation model. The major difference between INTEGRATION and FRESIM is that the former includes dynamic routing capabilities that can change the intended trip path of vehicles as they progress through the network, depending on current estimates of travel time on the available paths. The program developers see this model as a way of evaluating the effect of advanced technologies (particularly advanced traffic management systems and advanced traveler information systems) upon traffic operations in urban areas. The program has been tested on networks in both the United States and Canada, with favorable results.[20,21]

Because the program models both freeway and arterial street segments within the analysis area, data requirements are fairly large. However, specific geometric details such as grades, superelevation, and pavement conditions are not modeled explicitly. Specific data items needed are listed below:

- Segment length.
- Number of “effective” through lanes.
- Saturation flow rate per lane.
- Free-flow speed.
- Traffic signal characteristics for ramp meters and signalized intersections.
- Loop detector locations and characteristics.
- Incident characteristics (effect of lane blockage on capacity, effect of

rubbernecking on capacity, duration of incidents).

- Zonal traffic origin-destination tables for each time period (work is underway to build these synthetically from freeway and ramp volumes).

The output of the INTEGRATION model includes the following categories of statistics:

- Trip times for each origin-destination pair (total and per vehicle).
- Total trips entering the network.
- Speeds on each segment.
- Volume-to-capacity ratio for each segment.
- Trip times on each segment (total and per vehicle).
- Number of stops.
- Queue sizes.
- Fuel consumption.
- Vehicle emissions.

Strengths and Weaknesses of the Available Models

Table 2-7 summarizes the strengths and weaknesses of each of the freeway simulation/optimization models discussed in this chapter. Generally speaking, the microscopic models offer more detailed analyses, but at the cost of more extensive data requirements. The inability to model freeway-to-freeway connectors is the major limitation of FREQ, whereas the FREFLO portion of the CORFLO computer package does not collect environmental statistics on the freeway segments that are evaluated.

QUEWZ-92 is a special-use tool appropriate only for the evaluation of alternative work zone lane closure alternatives. FRESIM does not offer the ability to model HOV lanes, nor can the influence of reduced lane widths be modeled. Finally, despite the microscopic analysis approach utilized by INTEGRATION, the model is not designed to offer detailed operational analyses of freeway geometrics. Rather, the model focuses attention on how operational control and driver information influence freeway and arterial street traffic conditions.

FIELD MEASUREMENTS

One way to quantify the magnitude of operational problems on freeways is through field studies. The following two types of field studies can be used to provide useful insight in identifying problem locations on a freeway and to identify methods for improving operations on the freeway.

Travel Time and Delay Studies

Travel time and delay studies can be used quantitatively for the following purposes, which can help identify and equate operational problems and solutions in a freeway management system:[22]

- Determining the efficiency of a route with respect to its ability to carry traffic, and relative to other routes, through the use of sufficiency ratings or congestion indices.
- Providing input to capacity analysis of roadway segments.
- Identifying problem locations as indicated by delay.

Table 2-7. Strengths and Weaknesses of Freeway Traffic Simulation/Optimization Models. [12]

	Macroscopic Models			Microscopic Models	
	FREQ	**CORFLO**	**QUEWZ-92**	**FRESIM**	**INTEGRATION**
Strengths	Widespread validation Easy to use Can optimize ramp metering rates	Can explicitly model freeway-arterial street interface Can simulate freeway-to-freeway connectors	Simple to use Tailored to work zone lane closure analysis Computes road user costs directly	Detailed analysis of geometrics are possible Can configure parameters to local driver and vehicle characteristics Can simulate freeway-to-freeway connectors	Explicitly models freeway-arterial street interface Can dynamically route vehicles to simulate real-time motorist information Can simulate freeway-to-freeway connectors
Weaknesses	Analysis limited to one direction of a single freeway Does not model freeway-to-freeway connectors	FREFLO does not provide fuel consumption or vehicle emission statistics	Does not evaluate restricted lane widths only Applications limited to work zone analyses	Does not model HOV lanes Does not evaluate impact of reduced lane widths Extensive geometric data requirements (curvature, superelevation rates, grade)	Detailed vehicle interactions and lane-changing behavior not explicitly modeled Extensive data requirements (i.e., zonal origin-destination data)

Source: Adapted from Reference 11.

- Providing input to transportation planning models, trip assignment models, and route diversion models.
- Providing input to economic analyses of alternatives.
- Generating travel time contour maps.
- Providing input to studies that evaluate trends in efficiency and level of service over time.

There are several techniques available that can be used to conduct travel time and delay studies, including the following:

- Test-car runs.
- License-plate observations.
- Toll-road electronic payment/automatic vehicle identification cards.
- Observation of vehicles from vantage points.

The reader is referred to ITE's *Manual of Traffic Engineering Studies* for specific details on how to perform travel time and delay studies.[22]

Origin/Destination Volume Studies

In a number of situations, simply knowing the number of vehicles at various locations is not sufficient to assess travel demands and patterns on the freeway system. Knowing where vehicles are coming from and where they are going is essential in many freeway applications. For example, simply counting the number of vehicles entering and exiting a certain ramp on the freeway does not provide the type of information that allows alternative routes and diversion strategies to be adequately planned and analyzed. For these applications, it is important to know how many vehicles entering the freeway at a specific ramp exit at another ramp downstream. This type of information can be collected using origin/destination studies. Situations where origin/ destination information would be useful include the following:

- Freeway interchanges.
- Weaving areas.
- Traffic at major activity centers.
- Regional planning studies.

There are a number of techniques that can be used to collect origin/destination information from travelers, including the following:

- Light-on studies.
- License-plate studies.
- Post-card studies.
- Interview studies.

The reader is referred to ITE's *Manual of Traffic Engineering Studies* for specific details on how to use these techniques to collect origin/destination information.[23]

2.4 LESSONS LEARNED

SYSTEM MAINTENANCE

System maintenance should be considered at all levels of the planning, design, and implementation process of the freeway management system. Without proper maintenance, the effectiveness of the system is significantly reduced. Equipment failures can lead to vehicle crashes that not only affect the overall level of congestion on the freeway system, but also expose the agency

to liability concerns if the malfunction is not corrected in a timely manner. Inadequate maintenance also affects the ability of the control system to perform at an optimal level. Furthermore, inadequate maintenance can also reduce the service life of system components, leading to higher overall life-cycle costs in the system.

To be effective, a maintenance program must have the following elements:[5]

- Adequate staff of well-trained personnel.
- Up-to-date documentation on all system components.
- Adequate budget for spare parts and expendables.
- A long-term commitment on the part of the agency to utilize the system to its full potential including keeping the system "up-to-date" on a continual basis.

Ways in which maintenance can be supported during the design of the system include the following:[5]

- Specify modular components so that components can be swapped in the field and repairs handled in the shop, where they will not interfere with field operations.
- Specify environmentally-hardened components that meet recognized standards.
- Use standardized equipment makes and models to reduce the number of spare parts and different maintenance techniques.
- Specify self-diagnostic capabilities.
- Specify transient protection (e.g., power surges and lightning) and equipment grounding.
- Locate components to minimize vulnerability to damage (e.g., knockdowns of field equipment).
- Provide for safe and convenient access to the hardware for maintenance personnel (e.g., will traffic lanes need to be closed when variable message signs are being repaired or relamped, etc.).

Types of Maintenance Requirements

The maintenance of hardware elements in a system traditionally falls under the maintenance heading. In fact, there are three major types of maintenance activities that are performed in a freeway management system:[3]

- Functional.
- Hardware.
- Software.

Functional

Traffic conditions, travel patterns, land-use patterns, and the political environment in many communities are continuously changing. Therefore, the control strategies and functions performed by a freeway management system need to be continuously reviewed, revised, and updated to keep pace with the change in operating environments. The types of activities included in this maintenance category include the following:

- Relocating or repositioning system detectors and surveillance equipment.
- Reconfiguring subsystems.

- Expanding the system to cover a greater geographic area.
- Expanding the system to provide additional functionality.
- Changing detection and/or control technologies to keep pace with technological changes.
- Updating or improving control strategies.
- Updating, revising, and improving incident response plans.

Hardware

Hardware maintenance can be divided into three categories, each of which must be planned for in designing the system and accounted for in the preparation of annual operating budgets:[3]

- Remedial.
- Preventative.
- Modification.

Remedial maintenance refers to the type of maintenance activities performed to correct or replace malfunctioning or failed equipment. Because this type of maintenance activity includes maintenance performed on an emergency basis, it usually commands the highest priority.

Preventative maintenance refers to those activities that are performed to keep equipment failures from occurring. Examples of the types of activities performed in this category include the following:

- Relamping and cleaning of variable message signs, ramp meters, etc.
- Inspection of poles, foundations, and wiring.
- Re-tuning of detectors.

Good preventative maintenance begins with and depends upon proper installation, and proper installation requires thorough inspection and testing throughout the design and installation process.[5] System operators and maintainers should participate in the inspection process and acceptance testing. Such involvement is an excellent source of informal training and fosters a sense of ownership of the new system, essential ingredients to developing an effective maintenance program.

System modification or reconstruction is often included as a maintenance item because, in many cases, maintenance personnel are used to implement the changes. Generally, system modification or reconstruction becomes necessary when the following situations occur:

- A manufacturing or design flaw is identified.
- Changes are needed to improve the performance characteristics of the equipment.

Good maintenance practices require that hardware functions and specifications be well documented. Table 2-8 lists some of the items that should be included in the documentation of the hardware to support hardware maintenance activities.

Software

The maintenance of software is often overlooked in freeway management systems. Most computer software undergoes (or should undergo) complete and extensive testing and debugging before it is accepted

by the operating agency; however, because of the multitude of different possible computational combinations and operational circumstances, it is impossible to discover every flaw or bug during the testing and acceptance period. Additionally, as operators gain experience with the system, they often discover additional procedures or features that could improve operations. Agencies need to have a mechanism for making corrections and modifications to software after the initial warranty period has expired.

Operating agencies must be committed to providing adequate funding and staffing resources to effectively maintain software. Traditionally, two mechanisms exist for maintaining system software:

- In-house.
- Maintenance contract.

If an agency elects to maintain the software system in-house, it must provide a sufficient number of qualified staff to perform these functions. If software maintenance is to be performed in-house, most agencies devote one or more full-time positions solely to developing and maintaining system software. Many agencies use engineering staff that also happen to have the requisite software skills to maintain their system software.[22] Because highly qualified programmers and software engineers are in great demand, agencies may find it difficult to attract and retain qualified programmers. In these situations, an agency can contract its software maintenance. Under this type of agreement, the contractor can correct any latent bugs and make minor enhancements based on the experience of the operations personnel.[5] Oftentimes, software maintenance can be accomplished over a phone link and modem.

One way software maintenance can be facilitated at the design and implementation phase of the system is through the adoption of coding standards.[24] Coding standards should be adopted prior to beginning work on the system. The purpose of the coding standards is to ensure that programmers use the same style and format when developing their software. The coding standards are to be used by every programmer working on the system. Coding standards are important because many routines written by one programmer may be used by other programmers in developing their routines. Furthermore, since all the programmers use the same style, the overall package has a consistent “look and feel” about it. Coding standards also lead to better documentation and enforce good programming practices. All of these aspects of coding standards can save considerable time in the initial programming of the system, plus making it easier for unfamiliar programmers to make modifications and enhancements after the original programmers are no longer working on the system.

Another way software maintenance can be fostered in the initial design and implementation phase is through good documentation.[24] Documentation is one of the most valuable and yet variable deliverables associated with any software development effort. It can be very difficult to obtain the necessary level of detail when project funds run short and the programmers are disinterested in the documentation effort. Documentation standards should be established as a guide for both in-house and contract programmers. Table 2-9 lists some of the elements that should be submitted as part of the software documentation.

Table 2-8. Items to be Included in Hardware Documentation. [5]

- General description — General description of the component.
- Theory of operation — Detailed description of the operation of the component.
- Normal operating procedure — Description of the procedure for the routine operations of the component, including normal operating characteristics, voltage levels, and waveforms measured at test points.
- Parts list — Listing and identification of various parts of the component.
- Schematic drawings — Complete and accurate electronic schematics that specify component interconnections, component values, voltage levels, and component locations.
- Drawings of cabinet layouts, wiring diagrams, and lightning protection. Wiring and cabling lists describing interconnection of all plugs, chassis, and other components. In addition, they must identify wire type, size, and color code. They should also identify connector type, pin numbers or terminal strip numbers, and test points.
- Mechanical details — Equipment layouts, physical dimensions, access points, and test point locations.
- Power supply cabling — A description of the power supply, the power distribution system, and the characteristics of the power supply. The power source and all protective devices in the power system must also be described.
- Environmental controls — Power, heating, cooling, and humidifying.
- Descriptions of all preventative maintenance activities for all system components (e.g., computer, communications units, controllers, etc.). This information must include both the procedures and the frequency with which these activities are to be performed.
- Description of emergency maintenance trouble-shooting and diagnostic procedures. This documentation begins with a list of symptoms and proceeds through a series of analyses until the most common cause of the symptom is identified.
- A recommended set of spares and test equipment to be purchased by the local agency or supplied under the contract.
- Instructions on the use of computer diagnostic software furnished by the computer manufacturers for evaluation of the computer and peripheral equipment operation.

MARKETING SYSTEMS

The importance of marketing freeway management systems should not be overlooked. Public acceptance and cooperation are essential to achieving the goals and objectives of a freeway management system. Without public acceptance and support, an agency will find it difficult, if not impossible, to maintain a secure and stable political support and funding base for operating and maintaining a freeway management system. Marketing the system — highlighting what it does and how it will directly benefit the traveling public — is crucial to building public support and acceptance.

Target Audiences

Transportation agencies should develop a plan for marketing their freeway management system. The market plan should address how the system will be marketed to the following four target audiences:[26]

- The operating agency itself.
- Other agencies that are affected by the system.
- Important public and elected officials.
- The general public and users of the system.

Operating Agency

A freeway management system will not be successful or effective without strong support from those individuals who will be funding, operating, and maintaining the system. Decision-makers within the agency need to be aware of the costs and benefits of the system and where the freeway management system fits into the overall future of the organization as a whole.[25] Often, these decision-makers can be converted to being champions for the system. Marketing strategies need to be employed internally within an organization so that champions can be identified, and pride and ownership in the system can be developed.

Affected Agencies

The marketing plan should also include those agencies that are not directly responsible for but directly affected by a freeway management system. Examples of affected agencies that should be addressed in a marketing campaign include the following:

- Law enforcement agencies.
- Emergency service providers.
- Local and regional transportation agencies.
- Transit providers.

These agencies are concerned primarily with how the system will affect their operations. Law enforcement agencies will want to know whether the system will require additional enforcement emphasis, reduce their enforcement burden, or have no effect on their service. Emergency service providers will want to know whether the system will enhance or hinder their response time and access to facilities. Local transportation agencies and transit providers will want to know whether the system will increase or decrease the demands on their systems.[25] A marketing strategy should be developed that addresses these concerns.

Table 2-9. Elements of Software Documentation. [5]

- Source listings of the programs. These listings must be well-annotated to describe input/output, variables, purpose of subroutines, other lower level subroutines called.
- Flowcharts or HIPO's indicating the processing steps taken by every functional grouping of instructions. The flowcharts should be keyed to the software listings using instruction addresses and subroutine names. Each flowchart must be preceded by a functional description of the subroutine's purpose, storage requirements, and execution time.
- Descriptive discussion of each routine and subroutine in the application program and the database update program.
- Lists that define every variable used in the software. These definitions must include variable name, its purpose, structure, and a list of the routines in which it is used.
- Maps showing the layouts of data and programs within the computers and storage facilities.
- A summary of system timing, including average execution times of all routines and average execution times of each priority level.
- All computer inputs and outputs, including addresses, command structures, data transfer rates, source/destination equipment, and printer and CRT formats.
- Backup and disk-resident copies of the source programs, along with hard-copy listings.

Elected Officials

Elected officials are primarily interested in cost-effective systems that support the goals of their jurisdiction or district. They are interested in secondary effects of technologically advanced systems, such as the effects on the environment, privacy, and land use. They are also interested in the views of constituents, and whether the systems will be perceived as having a positive or negative impact on their constituency. To enable the system to be effectively marketed to elected officials, the information about the system should be concise and easy to understand.[25]

General Public

The ultimate success of any advanced transportation system depends on the perceptions of the users of the system. If the general public does not see benefits that outweigh any perceived costs or disadvantages, the system will not be successful, even if there are positive benefits being achieved by the system. It is critical that public agencies remember that the ultimate customers of the system are the people who will be traveling on the freeway. The marketing campaign should illustrate how these travelers will directly benefit from the system.[25]

Marketing Strategies

There are numerous methods available for marketing freeway management systems. Determining which is the most appropriate and effective method depends on the target. Strategies commonly employed to market freeway management systems included the following:

- Logos.
- Brochures.
- Slide shows/videos.
- Meetings with agency boards, citizen advisory committees, and operating staffs.
- Meetings with community groups, professional organizations, high schools, college groups, employers, etc.
- Workshops and meetings with City Council members, public works departments, and transportation advisory committees.
- Press releases and media events.
- Displays and billboards.
- Newsletters and fact sheets.
- Internet home pages.
- Tours and site visits.

Table 2-10 lists some of the key items to be remembered and included in developing a marketing plan for freeway management systems.

Table 2-10. Guidelines for Developing a Marketing Plan for Freeway Management Systems. [27]

- While it is critical and necessary to involve members of operating and affected agencies early in the marketing of the system, be careful about involving the public too early. Very early involvement of the public can lead to high and unrealistic expectations.
- Identify the target audiences in the beginning stages of the program. Have a clear understanding of what it is the audience is supposed to get from the media campaign and how they are likely to respond to having knowledge of the system.
- A good marketing plan supports the strategic plan for project implementation. The marketing plan should support the goals and expectations of the project and clearly convey them to the target audiences.
- Do not overexaggerate the impact the system will have on the transportation system. Encouraging high expectations may lead to disappointment and the immediate loss of public trust. Moderate goals are most appropriate when implementing new technologies. Do not let the public believe that the system will solve all their transportation problems.
- Keep communications as open and honest as possible throughout the program implementation. If something goes wrong, let the public know what happened in terms they can understand. Chances are the public will appreciate the explanation and understand the circumstances, and will be more likely to give you the benefit of the doubt if other problems arise.
- If you do not have the qualified staff to develop a marketing plan, hire an outside firm that is familiar with transportation projects to assist you. Human behavior is an unpredictable phenomenon. It may be beneficial to consult with individuals who are experienced with dealing with the public in formulating your marketing plan.
- When promoting the system, whether verbally, visually, or in written form, use terms that are understandable and tangible to the general public. Highlight the positive aspects of the system. For example, indicate that the ramp metering system is intended to maintain speed above 40 mph on the freeway. Take car to mak the wording direct so that it cannot be misinterpreted easily.
- Have something to show. A tangible product is easier to relate to than a drawing or report. The public will be able to see how it works and the benefits that can be derived. They want to see it, not just hear about it.
- Throughout the implementation of the system, survey user needs to determine whether the needs of the public are being met. If not, adjust the program accordingly. Monitoring of user acceptance and attitudes should continue after the system has been implemented to ensure that it is working as planned.

2.5 REFERENCES

1. Athey, T.H. *Systematic Systems Approach.* Prentice Hall, Englewood Cliffs, NJ, 1982.

2. Intelligent Vehicle Highway Systems (IVHS): "What, Why, & How." Seminar by FHWA, U.S. Department of Transportation, 1993.

3. Gordon, R.L., et. al. *Traffic Control System Handbook*. Report FHWA-SA-95-032. FHWA, U.S. Department of Transportation, 1996.

4. Abrams, C.M. and Direnzo, J.F. *Volume 2: Measures of Effectiveness for Multimodal Urban Management.* 1979.

5. Neudorff, L.G. *Guidelines for Successful Traffic Control Systems: Final Report*. Report FHWA-RD-88-014. FHWA, U.S. Department of Transportation, 1988.

6. Part 655-Traffic Operations, Subchapter G-Engineering and Traffic Operations. In *Federal-Aid Policy Guide*, FHWA, U.S. Department of Transportation.

7. Hummer, J.E. Appendix A. Experimental Design. In *Manual of Transportation Engineering Studies.* Institute of Transportation Engineers. Prentice Hall, Englewood Cliff, NJ, 1994.

8. Lindley, J.A. and Capelle, D.G. Freeway Surveillance and Control. In *Transportation Engineering Handbook*. Fourth Edition. Institute of Transportation Engineers. Prentice Hall, Englewood Cliffs, NJ, 1992.

9. McDermott, J.M., Kolendo, S.J. , and Wojcik, R.J . *Chicago Area Expressway Surveillance and Control: An Interim Report*. Illinois Department of Transportation, 1973.

10. *Highway Capacity Manual*. Special Report 209, TRB, National Research Council, Washington, DC, 1994.

11. *Texas Highway Operations Manual.* Report FHWA/TX-94/1232-3. Texas Transportation Institute, College Station, TX, 1994.

12. Reilly, W.R. Operational Aspects of Highway Capacity. In *Traffic Engineering Handbook*. Fourth Edition. Institute of Transportation Engineers. Prentice Hall, Englewood Cliffs, NJ, 1992.

13. May, A.D. *Traffic Flow Fundamentals*. Prentice Hall, Englewood Cliffs, NJ, 1990.

14. Wilshire, R.,Black, R., Grochoske, R., and Higinbotham, J. *Traffic Control System Handbook*. Report FHWA-IP-85-12. FHWA, U.S. Department of Transportation, 1985.

15. Krammes, R.A., Ullman, G.L., Dresser, G.B., and. Davis, N.R. *Application of Analysis Tools to Evaluate the Travel Impacts of Highway Reconstruction with Emphasis on Microcomputer Applications*. Report FHWA-ED-89-023. FHWA, U.S. Department of Transportation, 1989.

16. Singh, R. Traffic Congestion, HOV Lanes, and Ramp Metering — A Case Study. In *Compendium of Technical Papers*, 62nd Annual Meeting of the Institute of Transportation Engineers, Washington, DC, 1992.

17. Liv, D.D. and Kanaan, A. Evaluation of Freeway Alternatives Using CORFLO. In *Proceedings*, Fourth International Conference of Microcomputers in Transportation, American Society of Civil Engineers, Baltimore, MD, 1992.

18. Krammes, R.A., Ullman, G.L., Memmott, J.L., and Dudek, C.L. *User's Manual for QUEWZ-92*. Texas Transportation Institute, College Station, TX, 1993.

19. *TRAF User Guide*. Report FHWA-RD-92-060. FHWA, U.S. Department of Transportation, 1992.

20. Bacon, V.W. Jr., Lovell, D.J., May, A.D., and Van Aerole, M.. Use of Integration Model to Study High-Occupancy Vehicle Facilities. In *Transportation Research Record 1446.* TRB, National Research Council, Washington, DC, 1994.

21. Van Aerde, M.,Voss, J., and Blum, Y. Modeling the Burlinton Skyway FTMS During Recurring and Non-Recurring Traffic Congestion. In *Traffic Engineering and Control*, London, England, 1989.

22. Robertson, H.D. Travel-Time and Delay Studies. In *Manual of Transportation Engineering Studies*. Institute of Transportation Engineers. Prentice Hall, Englewood Cliffs, NJ, 1994.

23. Nelson, D.C. Transportation Planning Data. In *Manual of Transportation Engineering Studies.* Institute of Transportation Engineers. Prentice Hall, Englewood Cliffs, NJ, 1994.

24. Nordby, A. Software. In *Operating and Maintaining Advanced Traffic Management Systems Centers — White Papers*. Prepared for the National Conference on Operating and Maintaining Advance Traffic Management Systems (ATMS) Centers. Institute of Transportation Engineers, 1996.

25. Jacobson, L.N. Marketing ATMS and Measuring the Benefits. *Surface Transportation and the Information Age*. Volume 2. Proceedings of the IVHS AMERICA 1992 Annual Meeting. IVHS AMERICA, Washington, DC, 1992.

26. Ricci, J. Marketing Intelligent Transportation Systems (ITS). In *Graduate Student Papers on Advanced Surface Transportation Systems*. Report SWUTC/94/72194-2. Southwest Region University Transportation Center, Texas A&M University System, College Station, TX, 1994.

APPENDIX A: SUBCHAPTER G - ENGINEERING AND TRAFFIC OPERATIONS

PART 655 - TRAFFIC OPERATIONS

Subpart D - Traffic Surveillance and Control

Sec. 655.401 Purpose.
655.403 Traffic surveillance and control systems.
655.405 Policy.
655.407 Eligibility
655.409 Traffic engineering analysis.
655.411 Project administration.

Subpart E - [Reserved]

Authority: 23 U.S.C. 101(a), 104, 105, 109(d), 114(a), 135, 217, 307, 315, and 402(a); 23 CFR 1.32 and 1204.4; and 48 CFR 1.48(b).

Source: 49 FR 8436, Mar. 7, 1984, unless otherwise noted.

Sec. 655.401 Purpose.

The purpose of this regulation is to provide policies and procedures relating to Federal-aid requirements of traffic surveillance and control system projects.

Sec. 655.403 Traffic surveillance and control systems.

(a) A traffic surveillance and control system is an array of human, institutional, hardware and software components designed to monitor and control traffic, and to manage transportation on streets and highways and thereby improve transportation performance, safety, and fuel efficiency.

(b) Systems may have various degrees of sophistication. Examples include, but are not limited to, the following systems: traffic signal control, freeway surveillance and control, and highway advisory radio, reversible lane control, tunnel and bridge control, adverse weather advisory, remote control of movable bridges, and priority lane control.

(c) Systems start-up is the process necessary to assure the surveillance and control project operates effectively. The start-up process is accomplished in a limited time period immediately after the system is functioning and consists of activities to achieve optimal performance. These activities include evaluation of the hardware, software and system performance on traffic; completion and updating of basic data needed to operate the system; and any modifications or corrections needed to improve system performance.

Sec. 655.405 Policy.

Implementation and efficient utilization of traffic surveillance and control systems are essential to optimize transportation systems efficiency, fuel conservation, safety, and environmental quality.

Sec. 655.407 Eligibility.

Traffic surveillance and control system projects are an integral part of Federal-aid highway construction and all phases of these projects are eligible for funding with appropriate Federal-aid highway funds. The degree of sophistication of any system must be in scale with needs and with the availability of personnel and budget resources to operate and maintain the system.

Sec.655.409 Traffic engineering analysis.

Traffic surveillance and control system projects shall be based on a traffic engineering analysis. The analysis should be on a scale commensurate with the project scope. The basic elements of the analysis are:

(a) Preliminary analysis. The Preliminary Traffic Engineering Analysis should determine: The area to be controlled; transportation characteristics; objectives of the system; existing systems resources (including communications); existing personnel and budget resources for the maintenance and operations of the system.

(b) Alternative systems analysis. Alternative systems should be analyzed as applicable. For the alternatives considered, the analysis should encompass incremental initial costs; required maintenance and operating budget and personnel resources; and expected benefits. Improved use of existing resources, as applicable, should be considered also.

(c) Procurement and system start-up analysis. Procurement and system start-up methods should be considered in the analysis. Federal-aid laws, regulations, policies, and procedures provide considerable flexibility to accommodate the special needs of systems procurement.

(d) Special features analysis. Unique or special features including special components and functions (such as emergency vehicle priority control, redundant hardware, closed circuit television, etc.) should be specifically evaluated in relation to the objectives of the system and incremental initial costs, operating costs, and resource requirements.

(e) Analysis of laws and ordinances. Existing traffic laws, ordinances, and regulations relevant to the effective operation of the proposed system shall be reviewed to ensure compatibility.

(f) Operations plan. The final element in the traffic engineering analysis shall be an operations plan. It shall include needed legislation, systems design, procurement methods, construction management procedures including acceptance testing, system start-up plan, operation and maintenance plan. It shall include necessary institutional arrangements and the dedication of needed personnel and budget resources required for the proposed system.

(Approved by the Office of Management and Budget under control number 2125-0512)

Sec. 655.411 Project administration.

(a) Prior to authorization of Federal-aid highway funds for construction, there should be a commitment to the operations plan (see Sec.655.409 (f)).

(b) The plans, specifications, and estimates submittal shall include a total system acceptance plan.

(c) Project approval actions are delegated to the Division Administrator. Approval actions for traffic surveillance and control system projects costing over $1,000,000 are subject to review by the Regional Administrator prior to approval of plans, specifications, and estimates.

(d) System start-up is an integral part of a surveillance and control project.

(1) Costs for system start-up, over and above those attributable to routine maintenance and operation, are eligible for Federal-aid funding.

(2) Final project acceptance should not occur until after completion of the start-up phase.

1. ELIGIBILITY OF POLICE ENFORCEMENT AND SURVEILLANCE ACTIVITIES IN TRAFFIC MANAGEMENT DURING MAJOR HIGHWAY RECONSTRUCTION (23 CFR 655.403)

a. In December 1986, the Washington Headquarters office provided guidance on traffic management actions that may be eligible for Federal-aid (FA) funding during major highway reconstruction. See NS 23 CFR 655A, paragraph 1, for activities discussed involving police surveillance and enforcement necessary to mitigate congestion and/or improve the safety of motorists and workers within the highway corridor. General guidance for these particular activities are contained in 23 CFR 655A, Traffic Operations Improvement Programs (TOPICS), and 23 CFR 655D, Traffic Surveillance and Control.

b. Police surveillance and enforcement activities are often essential to safety and efficient traffic operations during major highway reconstruction. To avoid any possible misunderstanding, this supplement is intended to further define FHWA's policy regarding the eligibility of FA funding. The following criteria are provided.

(1) Police enforcement and surveillance activities that normally would be expected in and around highway problem areas requiring management of traffic are not eligible for Federal-aid funding.

(2) On projects where normal police enforcement and surveillance practices within construction zones may not be adequate, the traffic management plan should address the appropriate extra police activities required.

(3) The primary purpose of extra police enforcement and surveillance activities must be to control traffic in order to maintain safe travel and efficient operations throughout the highway corridor.

(4) The State and FHWA division office must agree that extra police activities would be an effective and appropriate means to maintain safe and efficient travel and adequately protect workers. The traffic management plan should set forth the justification and details of the proposed activities. The costs of extra police enforcement and surveillance activities documented in the approved traffic management plan are eligible for Federal-aid reimbursement.

(5) Extra police activities should not be limited to passive monitoring and could also include appropriate positive guidance of traffic and enforcement of regulations.

2. IMPLEMENTATION PLAN GUIDANCE (23 CFR 655.409)

a. An Operations Plan is the final element of a traffic engineering analysis according to 23 Code of Federal Regulations (CFR) 655.409(f). Since an operations plan covers such a wide range of activities both prior to and after construction such as system design, procurement, personnel, operations and maintenance, the name **"Operations Plan"** is being changed to **"Implementation Plan"** to more accurately reflect its contents. The name change is being made in the CFR and will be used throughout the remainder of this guidance. The following is issued to provide State and other agencies, which are utilizing Federal funds, guidance traffic control systems, ensure adequate planning by the sponsoring agency, and commit the sponsoring agency to use Federal funds efficiently.

b. Implementation plans are **required** both for new traffic control systems, as well as expansions of existing systems, which use Federal funds and are encouraged for those systems which do not use Federal funds. Traffic control systems are defined as systems which contain elements to monitor, guide, control, and/or process forms of traffic along the surface streets and/or freeways. Implementation plans can be for individual projects (i.e. stand-alone), or as a part of a larger system. For expansion projects, if an implementation plan had not previously been prepared, one **must be prepared** and include the expansion as well. The plan should be completed prior to authorization of construction. This will ensure

that the system is designed, built, operated, and maintained so that it accomplishes its purpose in the most efficient manner possible, considering performance, cost, and schedule. Too often in the past, plans were developed after the system was operational and did not include the design approach and other information which should have previously been addressed and documented.

c. An implementation plan need not be a legal document; however, if it is to be effective, it must carry the weight of a memorandum of agreement (understanding) and should be signed by the head of the operating agency, State highway official, and Federal Highway Division Administrator, or their designates.

d. Before the guidance is explained, a few words need to be mentioned in regards to conformity and the planning process. Transportation Improvement Programs (TIPs) and Statewide Improvement Plans (STIPs) are area wide programs, while implementation plans are project specific. Hence, projects for which implementation plans are being developed have already been approved in the related TIPs and STIPs and the related conformity and management systems issues have been addressed. In non-attainment areas, the traffic control system being proposed for implementation must be consistent with what was proposed in a conforming transportation plan. If the traffic control system deviates from that design concept and scope, it may trigger a new conformity determination.

e. The following sections correspond to the implementation plan **requirements** listed in 23 CFR 655.409 (f) and provide discussion for each. The level of detail of the implementation plan will depend on the type and size of the system. Since some of the items required in an implementation plan will have been covered in other contract documents and other elements of the traffic engineering analysis (23 CFR 655.409), these items may be summarized and referenced in the plan.

(1) Legislation. This section includes the legal considerations, if any, for the project. **Existing** laws, regulations, and policies affecting the project need to be reviewed and assessed. In addition, State or local legislative changes such as authority for metering and HOV facilities, enforcement authority, and roadway clearance policies should be addressed if applicable to the project. Also, the operating procedures for the system may need to be defined to be sure that there are no potential legal problems.

(2) System design. A system contains elements which may monitor, guide, control, and/or process forms of traffic along the surface streets and/or freeways. System design consists of taking the recommendations from the planning phase, converting those needs into hardware/software requirements, and formulating the equipment needs into contract documents. The system design may be based on off-the-shelf, customized, or experimental technologies. Actual systems vary greatly in practice. For example, a system may contain several like devices such as

an expansion of a traffic signal system, or it may consist of a traffic management center and its associated hardware/software. For the purpose of this guidance, system operation and maintenance must be the responsibility of a public agency. The conduct of the system operation and maintenance may be carried out either by (1) the public agency (2) contract, or (3) franchise operation. An implementation plan should include the following elements for the system design portion:

(a) System Designer: Depending upon the complexity of the system and in-house expertise, consultant services are usually needed to design a system. The designer needs to be identified in order to resolve any conflicts.

(b) System Design Life: The functional operating life of the system should be identified. The design life and the costs can be used to perform an economic analysis to identify the return on their investment. The system design life will be helpful for a Life-Cycle Cost Analysis (LCCA).

(c) System Coverage: This should address the area that the system covers. The coverage related to the future expansibility of the system should also be addressed. Ideally, the expansibility should be commensurate with the system's design life.

(d) System Design and Operations/Maintenance Philosophies: System operations philosophies have a significant impact on the system design. For example, system operations centers that are staffed only during rush hours do not require kitchen and/or shower facilities. However, operations centers that are staffed during the majority of the day, especially during special events and inclement weather, do require extra amenities. Ideally, system operations and maintenance functions, as well as facilities, should be close to each other to facilitate coordination.

(e) System Architecture: A discussion of the overall system architecture (i.e. central, distributed, or hybrid) should be addressed.

(f) Integration with Other Functions: Ideally, consideration should be given to integrating a traffic control system with other systems to provide for data base exchange and other strategies so that the entire metropolitan area is covered and coordinated.

(g) System Components and Functions: Hardware components needed to perform system functions such as, surveillance, control, and coordination should be identified.

(h) Communication Subsystem Design Approach: Typically, the communication portion of the system, because of the necessary redundancy, represents a large portion of the system budget. Great care should be given to the subsystem design approach. An economic analysis of the design approach, should be a key consideration.

(i) Traffic Operations Center Design Features: The design of a control center is largely dependent upon the agency's operating philosophies (time of operation, special event operation, tour accessibility; media facilities etc.) The size of the system will also affect the design (As an example, agencies utilizing large numbers of closed circuit television (CCTV) will need more space for wall monitors.)

(j) Project Phasing/Scheduling: A formalized tracking system should be used to manage the project. Many common methods utilize critical path analysis. Depending upon the approach used, these management tools don't necessarily have to be developed during the design phase but should be in place prior to any construction scheduling.

(k) Design Review: The system design is reviewed and the problems and concerns are addressed and documented. (The system design should be checked for consistency with the statewide and metropolitan plans, if applicable.)

(3) Procurement methods. An important element of the implementation plan is the method used for procuring and implementing the system (23 CFR 172). Regardless of the method used, the implementation plan should include the following procurement related items: (1)Method, (2) Schedule, and (3)Funding. A brief description of common procurement methods follows:

(a) Sole-Source - a single manufacturer's specifications are openly used, or they serve as the basis for contract negotiations between the owner and the supplier. The contract is then awarded without competition. Sole-source contracts can be used in Federal-aid projects, but only if there has been a finding that it is more cost-effective than a competitive low-bid process. This method is most common for system expansions.

(b) Engineer/Contractor (turn-key) - an engineer prepares a single set of contract documents (i.e., plans, specifications, and estimates (PS&E) for the proposed system), the contract documents go through the procurement channels, and the contract is awarded to the lowest responsive bidder. The winning contractor is

responsible for providing a complete and fully operational system, including furnishing and installing all hardware/software, system integration efforts, and training and documentation. This method is the traditional low-bid process. However, there may be some significant potential problems with this method as it relates to traffic control systems: No single contractor may process the necessary experience and qualifications to perform all of the work; administering multiple layers of subcontractors and suppliers is difficult; and the prime contractor may not have sufficient knowledge of some of the elements of a traffic control system to select appropriate or qualified subcontractors.

(c) Two-Step Engineer/Contractor - in the first step, the plans and functional specifications, along with a Request for Proposals (RFP), are submitted to contractors. The submitted proposals are evaluated and the qualified proposals go to the second step. In the second step, a formal request for bid is issued. From this point on, the standard bid/award process of the engineer/contractor approach is used.

(d) Systems Manager - instead of a single turn-key contract in which all of the work is outlined, several contracts for the various subsystems are prepared. The agency's normal procurement process is utilized to obtain the equipment, but the systems manager administers the contracts and is responsible for integrating the various subsystems into an operating system.

(e) Design/Build- this concept involves awarding a single contract to provide for both the design and construction of a project. For certain circumstances, design/build has the potential for improving the contracting process by allowing contractors the maximum flexibility in the selection of innovative designs, materials, and construction techniques. Under current statutes and regulations, the design/build concept is a viable option for Federal-aid highway projects, as long as the following **requirements** are met:

1 The contracts are awarded following competitive bidding procedures;

2 If a warranty requirement is included, the period of coverage should only be sufficient in length (i.e., 1 - 5 years) to allow defects in materials and workmanship to become evident. Ordinary wear and tear, damage caused by others, and routing service maintenance should remain the responsibility of the State; and

3 Federal-aid projects which provide for evaluation of either the design/build or warranty concepts must be approved, under Special Experimental Project No. 14 (SEP 14), by FHWA Headquarters Office of Engineering (HNG-22), prior to project approval.

(4) Construction management procedures. Procedures which will be used for the particular system should be specified in the implementation plan. Construction management procedures provide the necessary framework for coordinating construction and installation activities to ensure the system is built in accordance with the contract documents. Implementation plan construction management procedures that can be addressed include, but are not limited to:

(a) Division of Responsibilities (identifying who is involved and their associated responsibilities).

(b) Scheduling and establishing mileposts (developing a construction schedule to keep track of system installation). This will also ensure a mechanism for monitoring progress, cost, and quality assurance.

(c) Conflict Mitigation (developing a procedure or mechanism for resolving contract disputes).

(d) Coordination with other projects (defining project's relationship with other projects).

(5) System start-up plan. Integration is the "glue" that binds components together to form the system. Components are physically tied based on interfaces defined by the system architecture and tests are performed to verify and validate whether or not system requirements are met. Verification of a component or subsystem determines if the components or subsystems are interfaced as per design and are working properly. Validation consists of ensuring (through acceptance tests) that all interfaced components or subsystems meet system requirements. Software coding and database development are also important elements of this phase. The start-up process is typically performed in a limited time period immediately after system integration. A start-up plan is necessary to document the validation process (software and system evaluation). An implementation plan should include, but is not limited to, the following:

(a) Software acceptance tests (responsibilities of those involved, test procedures, equipment involved, test criteria, verification of specific software features, methods to correct errors, etc.);

(b) System acceptance tests (responsibilities of those involved, test procedures, equipment involved, test criteria, verification that system performs required functions, methods to correct errors, final acceptance, etc.);

(c) Partial acceptance (provisions for accepting a partially completed system);

(d) Documentation (detailed documentation pertaining to hardware and software should be discussed as well as references to operating manuals for the system);

(e) Transition from old to new control (procedures for transitioning from a previously functioning system to a system with new features and functions);

(f) Operational support and warranty period (provisions for initial or continuing operational support and a system warranty period). Federal regulations on guaranty and warranty clauses are defined in 23 CFR 635.413;

(g) Training (provided to system operators and maintenance technicians prior to system acceptance);

(h) Coordination with the media is very important and should be included in the system start-up plan. Public support is critical to the success and ongoing operations of the system.

(6) Operations and maintenance plan. Traffic control systems require active management to be effective, including periodic reassessment of the control strategies used. In order to have a system that is operated and maintained properly, there must be a staff and budget commitment by the operating agency. The resources required to effectively operate and maintain a traffic control system may represent a significant continuing investment, particularly if the agency responsible for the system is relatively small or is implementing a traffic control system for the first time. The process of defining system operations and maintenance activities during the preparation of the implementation plan can expose these issues and allow time for their resolution prior to system implementation. The operations and maintenance plan may include a section for evaluation and applicable maintenance policies:

(a) Evaluation. Federal-aid highway funds may be used for evaluation activities (23 CFR 655.403 (c) (Systems Start-Up); (23 U.S.C.307 (c) (1) (e) (State Planning and Research); 23 U.S.C. 133 (b)(6) (Surface Transportation Program); and 23 U.S.C. 103 (i) (8) (National Highway System). A comprehensive evaluation

of a traffic control system determines if the system meets the goals and objectives established for it. A formal evaluation is recommended at appropriate stages. The evaluation should be completed as soon after the implementation of the system as possible, after traffic patterns have stabilized. Regular system re-evaluating should subsequently be planned every few years and should be executed by the operations and maintenance personnel. Key evaluation issues to be described in the implementation plan include:

1 The system evaluator (Preferably, this should be an independent third party, **not** the system installer.) The system evaluator should be selected prior to the implementation of the system in order to properly perform the evaluation.

2 The method of evaluation (This should also include time period for evaluation.)

3 The cost of evaluation.

(b) Maintenance Plans. Development of maintenance plans cannot be performed by designers alone. Maintenance persons must be consulted. In addition, a system may require a higher and more responsive degree of maintenance than an agency may be accustomed to. Some agencies may choose to use contract maintenance as opposed to in-house staff. Whatever method of maintenance is selected, the following implementation plan issues will help the operating agency to determine the necessary maintenance resources (budget and staff) :

1 Maintenance policies for preventative maintenance, system malfunctions (response times), etc. There should be a documentation of the policies, possibly as an attachment.

2 Formal maintenance management programs (software and hardware agreements with the developers). There should be a documentation of the programs, possibly as an attachment.

3 Initial inventory of spare parts and all necessary test equipment.

4 Training in providing limited maintenance to software and equipment.

(7) Institutional arrangements. Nearly all projects involve numerous organizations and multiple levels of government, all of which approach the project from various perspectives. However, the institutional aspects of a system are likely to be even more complex because of the additional governmental entities and organizations (e.g., FHWA, regional organizations, State and local governments, traffic engineering departments, MPOs, fire, police, transit, private sector groups, media utility companies, etc.) which are typically involved. The complex mix of governmental and private sector interests has the potential for difficulties: overlapping responsibilities, lack of understanding, and conflicting priorities and policies. To avoid these problems, it is important that close coordination be established during the early stages of planning. This will permit the various agencies to develop a better understanding of the system alternatives and the recommended system's features and functions; to identify overlapping responsibilities and determine which agency will take the lead in various areas; and to work harmoniously so that each agency can better fulfill its role. Developing a good, early working relationship with each involved organization and then maintaining this cooperation throughout the system process will help ensure that the system effectively meets the needs and expectations of each agency. An implementation plan should include, but is not limited to, the following institutional arrangement issues:

(a) A contact person/object liaison within each organization should be identified.

(b) Delineation of organizational responsibilities and the lead organization for the various elements of the system.

(c) Provisions for periodic project updates to be given to upper management to keep them informed.

(d) Utility arrangements.

(e) Written cooperative agreements for: personnel-sharing, cost-sharing, metering, traffic diversion, etc.

(f) Consideration should be given to the formation of an "Advisory Committee" which will meet to discuss and resolve system issues and to acquaint participants with the overall project goals, schedule, and work plan. All agencies involved in the project should be represented on this committee and should be involved throughout the entire project.

(8) Personnel and budget resources. Staffing for operations and maintenance of systems is a function of system complexity, hours of operation, and activities supported by the system. Ideally, staffing responsibility for

operating and maintaining the system should be integrated into the operating agency's existing organizational structure. It is understood that institutional agreements may need to be developed for personnel/cost-sharing purposes. The following personnel and budget items, as a minimum, should be addressed:

(a) Staffing plan (listing of the job functions supported by the system and the number of persons who fulfill those functions).

(b) If shifts are to be used, the number of persons and their functions per shift.

(c) Contract operations staff agreement (if used).

(d) Provisions for training new staff on the system.

(e) Sources of budgetary resources, including Federal, and their committed contributions.

(f) Estimates of annual expenses by category (operations, maintenance)

(g) The last page should have a section for the signatures of the head of the operating agency, head of the State highway agency, and FHWA Division Administrator or their designates. This concurrence ensures that the necessary agencies are committed to the implementation plan.

MODULE 3. SURVEILLANCE

TABLE OF CONTENTS

3.1 INTRODUCTION 3-5
MODULE OBJECTIVES 3-5
MODULE SCOPE 3-6

3.2 DESIGN PROCESS 3-6
PROBLEM IDENTIFICATION 3-6
Identify and Locate Operational Problems 3-6
Determine Functions of Surveillance System 3-7
Type of Data 3-8
Traffic Measures. 3-8
Real-Time and Historical Data 3-9
Importance of Data 3-9
Inventory Existing Surveillance Capabilities 3-10
IDENTIFICATION OF PARTNERS 3-10
Intra-agency 3-10
Interagency (Information Exchange) 3-11
Additional Resources 3-12
ESTABLISH GOALS AND OBJECTIVES 3-12
ESTABLISH PERFORMANCE CRITERIA 3-13
DEFINE FUNCTIONAL REQUIREMENTS 3-13
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 3-14
IDENTIFY AND SCREEN TECHNOLOGIES 3-15
Identify Alternative Technologies 3-15
Evaluate Alternative Technologies 3-16
Select Appropriate Technology 3-16
PLAN DEVELOPMENT 3-17
Design Plans 3-17
Specifications 3-17
IDENTIFY FUNDING SOURCES 3-17
IMPLEMENTATION 3-19
EVALUATION 3-19
Monitoring Traffic Operations 3-20
Detecting Incidents 3-20
Implementing Control Strategies 3-20
Monitoring Environmental Conditions 3-21

3.3 TECHNIQUES AND TECHNOLOGIES 3-21
SYSTEM COMPONENTS 3-21
Detection Methods 3-22
Hardware 3-24

Software 3-25
Incident Detection Algorithms 3-25
Real-Time Expert Systems 3-25
Communications 3-26
SURVEILLANCE TECHNOLOGIES 3-27
EMBEDDED DETECTORS 3-27
Inductive Loop Detector 3-28
Characteristics 3-28
Applications 3-28
Installation Requirements 3-32
Magnetometer 3-34
Characteristics 3-34
Applications 3-34
Installation Requirements 3-34
NON-INTRUSIVE DETECTORS 3-35
Microwave Radar 3-37
Characteristics 3-37
Continuous Wave. 3-37
Frequency Modulated Continuous Wave (FMCW) 3-37
Pulsing Waveform 3-37
Wide Beam. 3-38
Narrow Beam 3-38
Long-Range 3-38
Applications 3-39
Installation Requirements 3-39
Infrared 3-40
Characteristics 3-40
Active. 3-40
Passive 3-40
Applications 3-40
Active 3-40
Passive 3-42
Installation Requirements 3-42
Ultrasonic 3-44
Characteristics 3-44
Applications 3-44
Installation Requirements 3-45
Acoustic 3-45
Characteristics 3-45
Applications 3-45
Installation Requirements 3-45
Video Image Processing 3-47
Characteristics 3-47
Classes of VIP Systems 3-47
VIP Detector Performance 3-48
Advantages of VIP Systems 3-49
CCTV Applications 3-50

Applications 3-50
Installation Requirements 3-51
COMPARISON OF EMBEDDED AND NON-INTRUSIVE DETECTORS 3-51
VEHICLE PROBES 3-51
Types of Vehicle Probes 3-53
Automatic Vehicle Identification (AVI) 3-53
Active 3-53
Passive 3-54
Semi-Active 3-54
Type I 3-54
Type II 3-54
Type III 3-54
Automatic Vehicle Location (AVL) 3-54
Dead-Reckoning and Map-Matching 3-55
Signpost 3-55
Ground-Based Radio-Navigation 3-55
LORAN-C 3-56
Global Positioning System 3-56
Differential GPS 3-56
Cellular Telephones 3-56
MOBILE REPORTS 3-57
Cellular Telephones 3-57
Freeway Service Patrols 3-58
Call Boxes/Emergency Telephones 3-58
CLOSED-CIRCUIT TELEVISION (CCTV) 3-58
Fixed Location 3-59
Full Motion Video 3-60
Compressed Video 3-60
Portable 3-61
SUMMARY OF TRAFFIC SURVEILLANCE APPLICATIONS 3-61
ENVIRONMENTAL DETECTORS 3-62
Freeways 3-62
Tunnels 3-65

3.4 LESSONS LEARNED 3-65
OPERATIONS AND MAINTENANCE 3-65
Operations 3-65
Maintenance 3-66
Operations and Maintenance Costs 3-66
PRIVACY CONCERNS 3-67
SPACING AND PLACEMENT ISSUES 3-70
Mainline Traffic Detectors 3-70
Spacing 3-70
Placement 3-70
Non-Intrusive Detectors 3-71
CCTV Cameras 3-71

Site Selection Criteria 3-71
Preliminary Site Selection 3-72
Final Site Selection 3-72

3.5 EXAMPLES 3-73
MIAMI AUTOMATIC VEHICLE LOCATION (AVL) 3-73
TRANSCOM ELECTRONIC TOLL AND TRAFFIC MANAGEMENT (ETTM) 3-74
HOUSTON AUTOMATIC VEHICLE IDENTIFICATION (AVI) 3-75
CONNECTICUT MICROWAVE RADAR 3-75

3.6 REFERENCES 3-77

MODULE 3. SURVEILLANCE

Figure 3-1. Closed-Circuit Television Camera.

3.1 INTRODUCTION

In a traffic management system, the surveillance component is the process in which data is collected in the field. This data is used to supply information about conditions in the field to other system components. Surveillance provides the information needed to perform the following functions:

- Measure traffic and environmental conditions.
- Make control decisions.
- Monitor system performance.

Surveillance is intended to provide support for other elements in the system (e.g., incident detection, information dissemination, ramp metering, etc.), not to drive the decisions about what system elements should be included. In other words, the goals and objectives of a surveillance system should be defined first, and then the system should be designed to meet these goals and objectives. A common mistake to be avoided is to first install a surveillance system, then ask the question "What can this system do for me?". It is essential to determine what system elements are to be supported before selecting and designing a surveillance system.

MODULE OBJECTIVES

The objectives of this module are as follows:

- Provide insight into the issues associated with planning, designing, installing, operating, and maintaining a surveillance system.
- Provide a summary and description of available and emerging surveillance technologies.

MODULE SCOPE

The intent of this module is to help engineers/planners in the decision process involved in implementing an appropriate surveillance system. Included in this module are a description of the decision process, a summary of the components of a surveillance system, a discussion of available and emerging surveillance technologies, and examples of existing traffic management centers that are using various surveillance technologies. In addition, special issues such as privacy concerns and spacing requirements for sensors are addressed.

3.2 DESIGN PROCESS

This section describes the process for planning, designing, and installing a surveillance system. Issues associated with operations and maintenance are also addressed. The focus of this section is on the decision process involved in implementing a surveillance system using a systems engineering approach. Instead of focusing on technological solutions to perceived problems, the system engineering approach described in this section involves identifying user needs and system requirements first, and then designing a system to meet these needs and requirements.

PROBLEM IDENTIFICATION

The first step in the decision process is to identify the problems to be addressed by the system. Issues that should be addressed in the problem identification stage include the following:

- Identifying and locating operational deficiencies.
- Determining functions to be performed by surveillance system.
- Inventorying existing surveillance capabilities.

Table 3-1 describes the objectives for conducting each of the above tasks.

Identify and Locate Operational Problems

This process involves identifying those freeways that would greatly benefit from the use of surveillance and traffic management. Areas benefiting the most would be those with significant amounts of congestion. There are two types of congestion:

- **Recurring.** Typically predictable and occurs at locations where demand exceeds capacity or at geometric bottlenecks (e.g., lane drops, high-volume entrance ramps, etc.).
- **Nonrecurring.** Caused by a random event (e.g., incident, maintenance activity, special event, etc.) and has the effect of reducing the capacity on a specific section of freeway.

Whether the congestion is recurring or non-recurring, the effects on traffic operations include:[1]

- Increased delay.
- Slower and inconsistent travel speeds.
- Increased accident potential.
- Other undesirable characteristics.

Table 3-1. Tasks Involved in Problem Identification.

Task	Objective
Identify Operational Problems	• Identify freeway sections to receive surveillance system
Determine Functions of Surveillance System	• Evaluate freeway system to determine surveillance needs • Determine type of data needed • Determine importance of data • Develop criteria for selecting detection technology
Inventory Existing Surveillance Capabilities	• Determine if existing system needs to be replaced/expanded based on cost and needs • Estimate how long the existing system will be able to meet the needs • Identify advantages and disadvantages of existing system to aid in the process of selecting new equipment

Other freeways that might be considered for surveillance include the following:

- Freeways in areas in which significant increases in traffic demand are expected.
- Freeways in areas with significant amounts of maintenance or construction activities.
- Freeways in areas with high frequencies of traffic incidents.

These areas can expect an increase in congestion, and surveillance in combination with traffic management may be used as an alternative to the very expensive solution of building additional lanes.

Determine Functions of Surveillance System

A surveillance system can serve several purposes, including the following:

- Detect incidents that have an impact on traffic operations.
- Monitor incidence clearance.
- Monitor traffic operations and support the implementation of control strategies (e.g., lane control, ramp metering, etc.).
- Monitor environmental and pavement conditions (e.g., flood, ice, winds, fog, etc.).

Table 3-2 lists several scenarios in which surveillance is used and typical methods that meet the surveillance needs.

The type of surveillance system needed depends not only on the purpose(s) that it will serve, but also the type and importance of data to be collected. The following sections discuss the types of data typically collected through surveillance systems and factors that should be considered when determining the importance of the data.

Table 3-2. Methods for Meeting Typical Surveillance Needs.

Need	Method of Surveillance
Detect incidents and monitor incident removal	Mainline detectors, vehicle probes, mobile reports, closed-circuit TV
Meter ramps	Mainline detectors, ramp detectors
Disseminate travel information to motorists	Mainline detectors, vehicle probes, mobile reports, closed-circuit TV
Monitor traffic during special event	Closed-circuit TV
Inform motorists and/or maintenance personnel of icy conditions on freeways	Environmental sensors, weather stations, probe reports

Type of Data

Traffic Measures. The measurements that have traditionally been used to monitor traffic operations on freeways include the following:

- Volume.
- Speed.
- Occupancy.

Volume is used to measure the quantity of traffic and is defined as the number of vehicles traveling a given section of freeway over a period of time. The capacity of a freeway represents the maximum volume of traffic that can pass a given roadway section under prevailing roadway and traffic conditions. [(2)] As volume increases and approaches capacity, congestion will occur. Volume is typically used to track historical trends and to predict the future occurrence of congestion on given freeway sections.

Speed is an important measurement in determining the quality of traffic operations. Speed is frequently used to describe traffic operations because it is easy to measure in the field, and it is easy to explain and understand. Speed measurements are typically taken for individual vehicles and averaged to characterize the traffic stream as a whole. Measured speeds can be compared to optimum values to estimate the level of operations for a freeway or to detect incidents. For example, an alarm for an incident detection system might be set to go off if average speeds fall below a target value. Taking speed measurements at different points along a freeway can help in determining where congestion might exist.[(1)]

Occupancy is defined as the percent of time a given section of roadway is occupied and can be used as a surrogate for density. Occupancy is measured using presence detectors, and is much easier to obtain than density. When occupancy is being measured, a single lane is usually considered, with occupancy ranging from 0 percent (no vehicles passing over a section of roadway) to 100 percent (vehicles stopped over a section of roadway).

Although volume, speed, and occupancy have been the traditional types of data collected by a surveillance system, today's traffic management centers also rely on other

types of data for traffic management purposes. Examples of other data include the following:

- Vehicle travel times.
- Bus location.
- Emergency vehicle location.
- Queue length.
- Pavement condition.

Until recently, most of the data listed above was difficult to measure in the field; however, due to improvements in detector technologies, these measurements can now be obtained.

Real-Time and Historical Data. Both real-time and historical data may be used for traffic management purposes. Real-time data is needed for the following purposes in a freeway management system:

- Monitoring current traffic operating and environmental conditions.
- Detecting incidents.
- Implementing control strategies.

Historical data refers to past traffic conditions on a given section of freeway. Historical data can be used for several purposes, including the following:

- To establish a record of past traffic conditions on a certain freeway section.
- To compare real-time data to historical data to determine irregular traffic patterns; results of this comparison can be used to detect traffic congestion and incidents.
- To perform before/after analyses to determine the effects of implementing certain traffic management techniques.
- To create simulation models for analyzing potential improvements.
- To create planning models for establishing priorities for deployment.

Importance of Data

In the decision process involved in selecting an appropriate surveillance system, the importance of the data to be collected must be identified to establish the data requirements. The following factors should be considered when determining the data requirements:

- Speed.
- Accuracy.
- Cost.

The speed of a surveillance system relates to the frequency in which information about field conditions is relayed to the traffic management center. Speed plays an important role for some applications. For example, in order to minimize the effects of an incident on freeway operations, it is important to minimize the detection time. In addition, the speed of data collection is important when the data is used to implement a control strategy to reduce or prevent the formation of congestion.

Other factors must be considered when determining the speed of a surveillance system. The speed of the data collection determines the amount of data to be transmitted to the traffic control center. Therefore, operator overload should be taken into consideration. In addition, the amount and speed of data collection affects

the type of communication system required. As the amount of data to be transmitted increases, the communication requirements increase.

Data accuracy requirements are also dependent upon the elements that the surveillance system is to support. For example, the accuracy of the data is important for incident detection systems to avoid false alarms. Accuracy, however, may not be as crucial when collecting traffic data for traveler information systems. Typically, the faster and more accurate a surveillance system is, the more it is going to cost; therefore, it is important to balance speed, accuracy, and cost when choosing a system.[3]

Inventory Existing Surveillance Capabilities

The existing surveillance resources should be identified and evaluated to determine if they are suitable for continued use. The evaluation should include items such as the following:[4]

- Detectors.
- Controllers.
- Communication media.

It is important to determine if these components can meet existing needs and if they can accommodate changes in system requirements. The existing surveillance system should be evaluated to determine if the required speed and accuracy of data collection are attainable. Other factors that should be considered are the reliability and required maintenance of the existing system. It may be more cost effective in the long run to replace a system that requires extensive maintenance with a more reliable, low maintenance system.

The capabilities of the existing communications system should be evaluated, because the various detection technologies available have different communications requirements. For example, transmitting full-motion video requires a wide communication bandwidth (such as that provided by fiber optic cable); however, transmitting only data requires considerably less bandwidth than can be met by most communication media.

Existing infrastructure on which non-intrusive detectors and CCTV cameras may be mounted should be identified at this stage. In addition, existing conduit for the communication system should also be noted.

IDENTIFICATION OF PARTNERS

Another important step in implementing a successful surveillance system is to identify partners to be involved. Partners should be considered in the following three areas:

- Intra-agency (within agency).
.
- Interagency (between agencies).
- Additional resources.

Intra-agency

During the planning and design stages, all interested groups and individuals that will be involved in the surveillance system should be identified and included in the decision making process. The project team should include representatives from the following areas:[4]

- Management.
- Planning.
- Design.

- Operations.
- Maintenance.

Including representatives from each of these areas will help ensure the success of the system. For example, it is important to include persons from management on the team to help gain support for the surveillance system. Since freeway management systems often compete for funding with other agency expenditures, the support from top management is essential if agency resources are to be allocated to the operation and maintenance of the system. By including representatives from operations and maintenance, the following issues may be addressed:

- Support staff required for surveillance system.
- Number and qualifications of existing support staff.
- Training required to operate and maintain specific surveillance systems.

Using this team approach, specific concerns and requirements from each area can be addressed from the beginning. The team should then prioritize the requirements to determine which are most important and which are desirable but not needed.

Interagency (Information Exchange)

During the operation of a traffic management system, it is important for public agencies to exchange certain information on a continual basis. Information to be exchanged includes scheduled maintenance activities and special events. This coordination between agencies will ensure that proper measures are taken to minimize the effects of the event on overall traffic operations. Data exchange should take place between the following public transportation agencies:

- State.
- City.
- County.
- Transit.
- MPO.

Sharing information about the occurrence of incidents between agencies permits joint incident management by two or more agencies.[(4)] This allows more than one agency to be involved in responding to and clearing incidents. Data exchange may take place between public transportation agencies and the following enforcement and emergency agencies and companies:

- Police.
- Fire.
- Medical.
- Wrecker operators.

Real-time traffic information may be relayed to the motorists through traveler information systems such as dynamic message signs, highway advisory radio, or in-vehicle information systems. In addition, motorists may be informed about traffic conditions by providing real-time traffic data to the following sources:

- Media.
- Cable television companies.
- Public kiosks.
- Other traffic advisory services.

Another private sector entity that benefits from real-time traffic data is commercial vehicle operators. For example, dispatchers can use information about current traffic conditions to re-route commercial vehicles in an effort to minimize delay for the commercial drivers. This not only benefits operators of the commercial vehicles, but it also benefits other vehicles in areas of heavy congestion by directing the commercial vehicles away from the congested areas.

Additional Resources

During the process of selecting the appropriate equipment to be used in the surveillance system, it is important to identify and evaluate all of the alternatives. Because of the constant change in available systems, the following groups should be considered as resources during the planning and design of a surveillance system:[4]

- Manufacturers.
.
- Suppliers.
- Users.
- Researchers.
- Consultants.
- Other interested groups or individuals.

Manufacturers continually develop and improve system capabilities and therefore can provide information on the state-of-the-art in surveillance technology. Information on the equipment specifications, functional and design features, and costs may be obtained from the manufacturers and suppliers. Users of available systems develop unique approaches for some systems and can provide evaluations for certain technologies. Researchers and consultants test the available technologies to determine strengths and weaknesses. In addition, researchers produce technological advances in surveillance systems.[4]

ESTABLISH GOALS AND OBJECTIVES

To establish the goals and objectives of a system, it is important to identify what the system is to accomplish. Goals are used to define the long-range desires for the system. Objectives define the level of performance that is to be expected in the future. At this stage, it is important to note that system objectives are defined in terms of what services and functions the system is to provide — not in terms of technology. The focus should be on *what* the system is to achieve instead of on *how* it is to achieve it.

As discussed earlier, the surveillance system provides support for other elements of a traffic management system (such as incident management, information dissemination, ramp control, etc.). Therefore, the goals and objectives of the surveillance system must relate to the goals and objectives of the elements that it is supporting. For example, a goal of an incident management system is to reduce the impact of incidents on traffic operations. The objectives of the system might be to detect an incident in less than two minutes and reduce the incident clearance time by five minutes. The surveillance system can be evaluated by determining the system's ability to meet these established objectives.

Since the goals and objectives of a surveillance system relate to those of the element that the surveillance system is supporting, the reader is referred to the specific modules within this report that address each element. Additional goals and objectives of a surveillance system might include those that relate to monitoring the performance of a certain system. For

example, a goal might be to ensure that the proper message is being displayed on a dynamic message sign. Objectives would be to determine if the sign was operational and to identify the message being displayed. Other goals and objectives of a surveillance system might relate to the effects of implementing certain control strategies.

ESTABLISH PERFORMANCE CRITERIA

There is currently a wide range of traffic detectors from which to choose, and with advancements in technology, the number of alternatives is becoming even greater. It is important, therefore, to establish performance criteria to aid in the selection of an optimum system. Establishing performance criteria allows alternative systems to be compared against these criteria in a later task.

The established performance criteria should be related to the ability of the system to meet the pre-established goals and objectives. Criteria that may be used to measure the performance of a surveillance system include the following:

- Reliability of system.
- Accuracy of data.
- Timeliness of data.

Each of the above criteria is important in measuring the performance of a system. For example, a system is not effective if it provides accurate data but produces it 30 minutes after it is needed.

The above criteria should be used to establish parameters by which to evaluate the system. The desired performance of the system together with a selected range of tolerance should be used to develop quantifiable measures. The measures identified will be based on the elements of the traffic management system that the surveillance component will support. In addition, the established performance measures will be based on local concerns and policies. Table 3-3 provides some examples of performance measures that may be used to evaluate a surveillance system.

DEFINE FUNCTIONAL REQUIREMENTS

The next step in the decision process involves defining all of the functions of a system that are necessary to achieve the established objectives. At this stage, the focus should still be on *what* the system will be designed to do, not *how* the system will do it. Therefore, the functions should be defined independent of the available technology.

Again, the surveillance system provides support to other elements in a traffic management system. Therefore, the functional requirements of a surveillance system are dependent upon the element that it will support. For a surveillance system, the functional requirements typically relate to the type, frequency, and quantity of data required. The data typically used for freeway surveillance have included measures such as volume, speed, and occupancy; however, surveillance should not be limited to these measures only. Additional measures might include travel time, queue length, headway, origin/destination, vehicle classification, etc. These measures have been difficult to obtain in the past but can now be measured because of improvements in technology.

In the past, various surveillance concepts have been investigated in an attempt to meet different functional requirements, and have failed. Some of these concepts failed

Table 3-3. Examples of Performance Measures for a Surveillance System.

Category	Examples
Reliability	• Percent of time that system produces desirable results in various environmental conditions. • Amount of maintenance required.
Accuracy	• Percentage of false alarms generated from an automatic incident detection system. • Difference between a value measured by the detectors in the field (e.g., speed, volume, etc.) and the actual value.
Timeliness	• Average time between when an incident occurs and when it is detected. • Average time between when hazardous weather conditions occur (e.g., ice, fog, flooding, etc.) and when they are detected.

because they were not appropriate; however, others have failed because of a lack of ability to provide the needed data. Therefore, designers/planners should look at the various alternatives for meeting the functional requirements of a traffic management system and not just at the traditional approach. Even though some concepts have failed in the past, they may become viable alternatives due to improvements in technology.

Since the functional requirements of a surveillance system relate to those of the element that the surveillance system is supporting (e.g., ramp metering, incident in the above section on establishing goals and objectives, additional requirements of a surveillance system might include those that are involved in monitoring the performance of a certain system. Table 3-4 shows examples of functional requirements as they relate to the goals and objectives of a surveillance system.

In defining the functional requirements of a system, it is important to determine the needs of all partners involved. The project team should prioritize needs in order to identify those that must be met and identify those needs that may be desirable but not required. For example, a system to detect the location of incidents might be required, but a system to monitor incident clearance might only be desirable.

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

After functional requirements of a surveillance system are defined, the relationships between functions, the data required by each function, and the information produced by each function must be defined. The National ITS Architecture Implementation Strategy defines the various ITS elements in terms of market packages.[5] This concept recognizes that various ITS components must work together to achieve system goals. They are " tailored to fit-separately or in combination-real world transportation problems."[5] The market packages related to surveillance are:

Table 3-4. Examples of Functional Requirements Related to Monitoring System Performance.

Goal	Objective	Functional Requirement
Ensure that proper message is being displayed on variable message sign.	Determine if sign is operational and identify message being displayed.	View operation of variable message sign.
Maximize benefits of a certain control strategy (i.e., ramp metering, lane control, etc.).	Determine the effects of the control strategy on traffic operations.	Monitor traffic operations where control strategy has been implemented.

- **Network Surveillance Marketing Package**. This market package includes the roadside sensors with appropriate control and communication infrastructure to interface with other market packages such as traffic management and traveler information.

- **Probe Surveillance Marketing Package**. This market package is an alternative to traditional network surveillance elements and does not require the extensive distributed roadside infrastructure but does require wireless communication between probe vehicles and other market packages such as freeway control and surface street control.

- **Emissions and Environmental Hazards Marketing Package**. This market package provides emissions and hazards information to the Traffic Information Dissemination market package.

Development of the surveillance system should recognize the national architecture standards and be tailored to fit local issues and requirements.

IDENTIFY AND SCREEN TECHNOLOGIES

After the functional requirements and architecture of the system have been defined, the next step is to identify alternative technologies that can meet the defined requirements. The following steps are involved in identifying and choosing the appropriate technologies for the surveillance system:

- Identify alternative technologies.

- Evaluate alternative technologies.

- Select the appropriate system.

Identify Alternative Technologies

There are a number of alternative technologies available for collecting traffic data. The characteristics, applications, and requirements for various existing traffic detection technologies are discussed in a later section of this module. This review can be used as a starting point for identifying alternative technologies; however, due to continuous technological advancements and system improvements, available systems and their capabilities are constantly changing. Therefore, the current state-of-the-art in surveillance technologies should be identified

in this stage. To identify available technologies, it is important to continually interface with the following groups and individuals during the identification process:[4]

- Manufacturers.
- Suppliers.
- Users.
- Consultants.
- Researchers.
- Other interested individuals.

By keeping in contact with these groups and individuals, the analyst can accomplish the following:

- Keep up with current trends in technology.
- Identify the advantages and disadvantages of available systems.
- Obtain information on system specifications and costs.

Evaluate Alternative Technologies

The first step in evaluating alternative technologies is to identify the selection criteria (see previous section, *Establish Performance Criteria*). The next step involves the following measures:

- Initial screening of all available technologies.
- Detailed evaluation of the remaining alternatives.

During the initial screening process, the following factors should be considered for each type of surveillance technology included in the analysis:

- Location of sensors (i.e., embedded or non-intrusive).
- Installation, operation, and maintenance requirements.
- Reliability.
- Expected life.
- Life-cycle costs.
- Type of communications medium available.
- Requirements for future expansion.

After the initial screening, the detailed evaluation typically includes the following steps:[4]

- Estimating costs and benefits of each alternative.
- Performing comparative analyses.
- Selecting the system offering the greatest potential.

The advantages and disadvantages of the system must be quantified or weighted and evaluated. There are many techniques available to perform an analysis of the costs and benefits for each alternative. **Module 11** contains descriptions of procedures for performing a benefit-cost analysis.

Select Appropriate Technology

Surveillance technologies that are viable alternatives should have benefits that outweigh the costs. The argument can be made that the best system will have the greatest benefit-cost ratio; however, this is

not always the case. For example, a simple, low-cost system with fewer benefits may have the same benefit-cost ratio as a more sophisticated, affordable system with more benefits. Therefore, the analysis should include allowable expenditures for the system and the net benefits of each alternative.[4]

PLAN DEVELOPMENT

After the technologies that will be used in the system have been selected, the next step is to develop a plan for implementation. The Implementation Plan documents the results of the previous steps and identifies how the system will be implemented in the field. The Implementation Plan should also assess the phasing, procurement, staffing, and funding options available for implementing the system. **Module 2** contains a description of elements that should be included in the Implementation Plan.

One of the tasks at this stage is to develop the design plans and specifications. This task involves transforming the needs and goals of the surveillance system into design documents and specifications suitable for competitive bids. The plans include drawings that show the physical layout of the system, and the specifications define the quality and type of workmanship and materials. [4]

Design Plans

The design plans provide information for the contractor and equipment supplier to prepare a project bid and for the construction manager to aid in controlling the construction. Table 3-5 shows typical information that should be provided on the design plans. [4]

Specifications

Specifications contain detailed information on the minimum acceptable standards for the surveillance system's equipment along with procedures on how the equipment should be installed. Detailed specifications are important to ensure that the proper equipment is obtained. Sources for equipment and materials specifications include the following:[4]

- Federal Highway Administration.
- Institute of Transportation Engineers.
- Other State transportation departments.

These specifications may be used as they exist or modified to fit particular needs. Guidelines for specifications for the various components of a surveillance system are provided in table 3-6.[4]

IDENTIFY FUNDING SOURCES

As described in **Module 2**, funding sources should be identified during the development of the Implementation Plan. Funding for surveillance systems may come from both the Federal and State levels. On the Federal level, funding is available through the Intermodal Surface Transportation Efficiency Act (ISTEA) and continuing with the proposed National Economic Crossroads Transportation Efficiency Act (NEXTEA). Under the ISTEA act, capital and operating expenses for traffic management programs are eligible for Federal aid. By including projects in the Transportation Improvement Program (a spending plan required by the Federal government), agencies can receive federal funds. Under Title 23 of the United States Code (as amended by the ISTEA of 1991) the following funding sources can be used to purchase and operate freeway management surveillance systems:

Table 3-5. Typical Information in Design Plan. [4]

Item	Description
Title Sheet	Shows name, location, and scope of project.
Summary of Quantities	Shows material and equipment quantities.
General Notes	Frequently call special attention to critical requirements to ensure they are not overlooked; must avoid conflicts with specifications.
Site Schematics	Scale drawings of each site showing the roadway geometry, location of poles, conduit locations, and any other pertinent site information.
Construction Design Drawings	Shows locations of detectors to be installed and minimum requirements for construction dimensions.
Traffic Control Plan	Indicates handling of traffic during construction.
Details of Barricades and Signing	Illustrates the construction phase of the project.
Diagram of Underground Utilities	Shows details of location.
Standard Plans	Possessed by each agency, generally provide numerous standard drawings of frequently encountered details already approved for use in situations applicable to the project.

Table 3-6. Guidelines for Specifications. [4]

Specification	Guidelines
Detectors	Include physical properties, electrical properties, environmental conditions under which the equipment must operate, controls, and methods of operation.
Computer Software	Provide functional specifications for control software, compilers, assemblers, utilities, and diagnostic programs.
Video Terminals	Specify sample operator screens and controls, screen size, refresh rate, and colors.
CCTV Monitoring	Specify monitors, cameras, and interface protocols.

- Surface Transportation Program (STP).
- National Highway System (NHS).
- Congestion Mitigation and Air Quality Program (CMAQ).

STP funds are available on eligible projects with no time limit. NHS and CMAQ funds have definite time limits, and cannot be used for maintenance.[6]

Although some Federal funds are available for operating costs, Title 23 funds are generally used for system deployment and start-up assistance. Typically, funding for both operations and maintenance costs are provided on the State and local levels through maintenance budgets. Funding operations and maintenance costs for surveillance systems through the maintenance budgets often causes problems because of competition from other traditional maintenance activities. Therefore, it is crucial to identify the funding needs and funding sources for operation and maintenance activities early in the decision process.

Public/private partnership is an approach that many agencies are using to increase the source of funding for traffic management purposes. For example, many cellular telephone companies offer free or reduced rates to agencies for incident detection purposes. This approach allows other partners to get involved in a fair and equitable manner.

IMPLEMENTATION

The implementation process includes the activities involved in installing the components of the surveillance system to meet the established goal and objectives. **Module 2** gives a description of issues to be addressed during the implementation process. Additional issues that should be considered when implementing a surveillance system include the following:

- Phasing of installation.
- Training.

During the installation of a surveillance system, some disruptions to traffic are to be expected. To minimize delay to traffic, construction should take place during off-peak hours. In addition, the maximum number of lanes that can be simultaneously closed should be specified.[4] When possible, phasing of installation for the surveillance system should be coordinated with other freeway construction to minimize the overall delay to traffic.

Another area that must be addressed during the implementation of a system is the development of training programs for operators and maintenance personnel. Training should provide the technical skill necessary to effectively operate and maintain the system. The amount of training required depends upon the qualifications and knowledge of existing personnel and the requirements of the new system. For example, maintenance personnel that are familiar with an inductive loop detection system will require extensive training to maintain a video image processing system. Additional information concerning training requirements is contained in reference 9.

EVALUATION

After a surveillance system has been installed, the system should be evaluated for its effectiveness in meeting the objectives. The objectives of a surveillance system may include any of the following:

- Monitoring traffic operations.

- Detecting incidents.
- Supporting implementation of control strategies.
- Monitoring environmental conditions.

Monitoring Traffic Operations

The goals of monitoring traffic operations include:

- Providing measures of traffic operations (e.g., speed, flow, density, etc.).
- Identifying locations of congestion.
- Indicating the severity of congestion.

Records should be kept to track system operations and note any problems. An overall evaluation of the equipment should include monitoring the following:

- System reliability.
- Ability to provide required data.
- Timeliness of data.
- Accuracy of data.
- Ability to perform under various environmental conditions.
- Operational and maintenance requirements and costs.
- Any other problems with system.

Detecting Incidents

Automatic incident detection systems apply data collected from the field equipment to computer algorithms. The algorithms are designed to detect incidents by identifying discontinuities in traffic operations (e.g., speed, occupancy, etc.) measured upstream and downstream from an incident.

After an incident detection system has been implemented, it must be monitored and evaluated to calibrate the incident detection algorithm. The system is monitored by detecting incidents through alternative means such as the following:

- CCTV.
- Mobile reports (cellular call-ins, call boxes, service patrols, etc.).
- Emergency channels.

The algorithm is calibrated by adjusting the threshold values. The measure of effectiveness typically used for calibration is the false alarm rate. Threshold values are adjusted until an acceptable false alarm rate is achieved. The acceptance level is dependent upon available means of verifying the incident. For example, if an incident can be verified using CCTV, then a higher acceptance level might be tolerated. The threshold values may need to be recalibrated for major operational or geometric changes.[7]

Implementing Control Strategies

There are two types of control systems that may be used for freeways:

- Lane use control (see **Module 4**)
.
- Ramp control (see **Module 5**).

The objectives of these systems are to meter the demand or prohibit a certain movement in order to keep freeway volume below capacity to ensure continuous movement on the freeway. Detectors provide traffic data to these control systems to aid in the

decision making process about the extent to which control should be implemented. The detectors used in a control system should be evaluated based on the following criteria:

- Reliability of detectors.
- Ability to provide needed data.
- Timeliness of data.
- Accuracy of data.

Monitoring Environmental Conditions

Typical environmental conditions on freeways that are monitored include the following:

- Ice.
- Wind.
- Fog.
- Rain.
- Dust.

The goals of monitoring these conditions include the following:

- Warn drivers of dangerous driving conditions.
- Reduce number of incidents.
- Reduce secondary incidents.
- Help maintenance personnel monitor pavement conditions.
- Improve response time.
- Reduce costs associated with monitoring and treating ice on pavements.

Additional conditions that require monitoring in tunnel sections include noxious gases, such as carbon monoxide.

The evaluation of environmental detection systems should include the following determinations:

- Ability to meet established goals.
- System reliability.
- Ability to provide required data.
- Timeliness of data.
- Accuracy of data.
- Operational and maintenance requirements and costs.
- Any other problems associated with system.

3.3 TECHNIQUES AND TECHNOLOGIES

SYSTEM COMPONENTS

The complexity and size of a surveillance system will vary with the size of the freeway system being monitored and with the specific functions that it will perform. For example, when the area being surveyed is relatively small, a section of the traffic engineering offices may be sufficient to house the control center. As the system expands and the responsibilities of the control center increase, a larger center with more sophisticated equipment may be required. Figure 3-2 shows a simple example of the control center components included in a surveillance system.

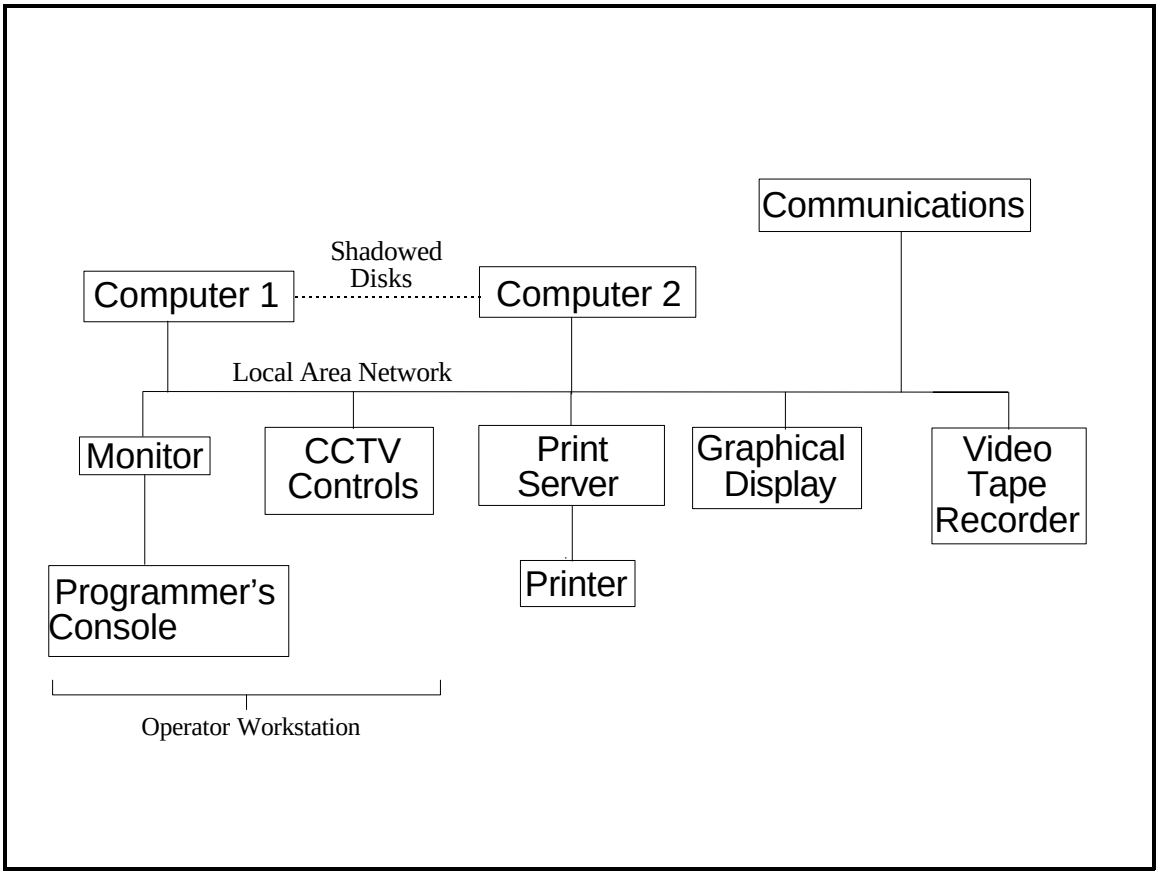


Figure 3-2. Control Center Components Typically Included in a Surveillance System.

The components of a surveillance system include the following:

- Detection methods.
- Hardware.
- Software.
- Communications.

Table 3-7 provides a list of typical system components.

Detection Methods

There are many technologies available for collecting traffic and environmental data for surveillance purposes. These technologies can be divided into the following groups:

- Embedded detectors.
- Non-intrusive detectors.
- Vehicle probes.
- Mobile reports.
- Closed-circuit television cameras.
- Environmental sensors.

Available detection methods are listed in table 3-7. Descriptions of each of these technologies is provided in a later section, entitled *Surveillance Technologies*.

Table 3-7. Typical Components for a Surveillance System.

Component	Alternatives
Detection Methods	Inductive Loop Magnetometer Microwave Radar Infrared Ultrasonic Acoustic Video Image Processing Automatic Vehicle Identification Automatic Vehicle Location Cellular Telephone Probes Cellular Call-Ins Freeway Service Patrol Reports Call Boxes/Emergency Telephones Closed-Circuit Television Environmental Detectors
Hardware	Computers Disk Drives Printers Monitors Controllers Displays Video Tape Recorders
Software	Incident Detection Algorithm Real-Time Expert System Interface Software
Communications	Internal - Local Area Network External - Fiber Optic - Coaxial - Twisted Pair - Microwave - Radio - Cellular Telephone - Citizen Band Radio

Hardware

Computers play significant roles in the operations of a surveillance system. The functions performed by the computers in the control center include the following:[7]

- Reception of data transmission from the field devices.
- Transmission of data from the control room to the field equipment.
- Reception of operator commands from keyboards or control panels.
- Control of graphical displays.
- Data processing to perform such functions as detecting incidents, deriving traffic flow characteristics, and identifying equipment failures.
- Storing data to create a historical database.

Computer 1 in figure 3-2 is an online computer that receives and transmits information to and from the field devices. Computer 2 serves as a backup and performs offline functions such as program development and database updating. The computers are typically microprocessor versions of minicomputers. The disk shadowing feature transfers data from the online computer to backup a system.[4]

The operator workstation typically consists of a personal computer unit. Functions that take place at the workstation include monitoring traffic operations, incident detection, decision making, and implementation of control strategies.

Closed-circuit television (CCTV) can be used for monitoring traffic conditions or detecting and confirming incidents. Controls for operating the CCTV system are also located at the workstation. The operator may have software installed on the personal computer to control the cameras or may have a separate control panel. Camera controls usually include camera selection, pan, tilt, zoom, and focus.

The printers interface with the local area network communication system through the use of print servers. Printers may be used to provide any of the following:[4]

- Hard copy reports of mode status and equipment failure logs.
- Traffic data summaries.
- Database status summaries.
- Special reports requested by the operator.
- Logs of system operations.

The functions of graphical displays are to provide observations of system operations and camera field of views for CCTV. Graphics may be provided on monitors at the workstations or on a large screen graphics display. The large screen displays have generally replaced the wall map displays used in older centers. The display may consist of a projection video display, a large video screen, or an array of smaller video screens. This new system has proven to be much easier to modify when new systems come online than the wall map system.[4]

Another important component of a surveillance system is the video tape recorder. Traffic operations can be recorded for analyzing and comparing conditions over a given period of time or for demonstration purposes. Depending on the size of the system and the recording requirements, one or more recorders can be provided to

interface with any camera in the system.[7] Some traffic management centers choose not to record and archive incidents because of the potential drain on their staff (video tapes requested by claims adjusters, lawyers, private citizens, etc.) and the possiblity of having to appear in court.

Software

Software is an important component of a surveillance system. Without software, the data collected by the surveillance system has little value.

Three general types of software used with surveillance systems include the following:

- Software such as incident detection algorithms, etc. that apply the data supplied by the surveillance system
- Software that supports the operations support software such as real-time knowledge-based systems.
- Software for interfacing and communicating with the field devices.

Incident Detection Algorithms

One of the primary objectives of a surveillance system is to detect non-recurring congestion, such as that caused by an incident. Incidents cause significant amounts of delay for freeway vehicles; however, because they are random, they are impossible to predict and difficult to detect. Incident detection algorithms attempt to automatically detect incidents based on field data received from detection equipment. **Module 7** provides further discussion on incident detection algorithms.

Real-Time Expert Systems

Real-time expert systems are currently being developed for application in traffic operation centers. These expert systems provide decision support for operations personnel so they can more efficiently carry out their functions.[8] The knowledge-based support is based on the knowledge of experts in the field of traffic management. Functions of an expert system include:[8, 9]

- Detecting, verifying, and estimating the severity of incidents.
- Providing advice in areas such as dispatch of traffic and incident management teams.
- Developing and coordinating messages on variable message signs and other methods of information dissemination.
- Controlling traffic with lane control and ramp control systems.

The University of California at Irvine is in the process of developing a prototype real-time knowledge-based expert system called FRED (Freeway Real-Time Expert System Demonstration). The expert system uses advanced processing capabilities to integrate diverse types of traffic surveillance data for freeway monitoring and control purposes. After detecting an incident, the system uses additional information from several sources to formulate appropriate responses. The information includes the following:[10]

- Traffic data from detectors (e.g., speed, volume, occupancy, etc.).
- Information on traffic conditions from CCTV.
- Field reports from police officers and other official personnel.

- Cellular and emergency telephone calls from motorists.

Figure 3-3 illustrates the overall system layout of FRED.[10] In terms of incident detection and verification, this system is consistent with many traffic operation centers around the country. The system uses two methods to detect incidents:

- An incident detection algorithm on a central computer receives 30-second occupancy counts from freeway loop detectors and notifies FRED if an incident is detected.

- Outside reports from various agencies and motorists are received by a communications center that filters the reports and sends the high priority ones to FRED.

The knowledge-based support in FRED comes from the real-time expert system shell, G2. This system has the ability to determine what data is required at any particular stage, and to send for it. Given the incoming data, decisions are made on the basis of a predetermined set of rules. Once an incident has been detected and verified, the system recommends actions to be taken, such as using ramp metering or closing a ramp.[10]

Communications

Communications between the major components in the control center are accomplished through a local area network (LAN). The benefits of a LAN architecture include the following:[4]

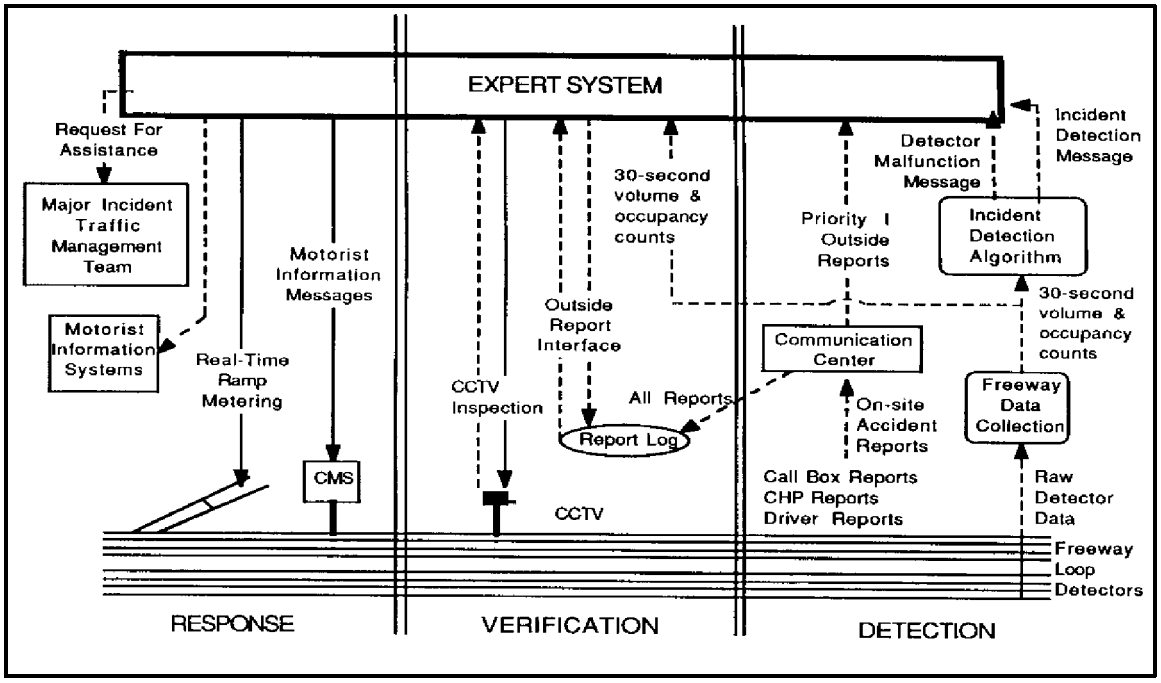


Figure 3-3. External System Overview for Real-Time Expert System. [10]

- Facilitate the addition of components to the network.
- Serve more devices in the operations center.
- Interface with additional field devices and controllers.
- Communicate with additional outside agencies and traffic data users.

Communications between the control center and the field devices occur through an external communication system. There are two basic categories of communication media:

- Physical cable (fiber optic, coaxial, twisted pair).
- Airwave transmissions (microwave, radio, and cellular telephone).

The type of communication medium required depends on the type of detection equipment being used. For example, transmission of video data requires a wide bandwidth that exceeds the capacity of many communication media. Transmission of voice and traffic data, however, requires a much narrower bandwidth which can be met by standard communication media. Table 3-8 lists some available communication media and their capabilities.[3] **Module 9** provides a detailed discussion of communication systems.

SURVEILLANCE TECHNOLOGIES

There are many technologies available for collecting traffic data. Although inductive loop detectors are currently used more than any other method, other technologies are beginning to replace loop detectors in many applications. This section provides descriptions of currently available detectors. The discussion is organized under the following headings:

- Embedded Detectors.
- Non-intrusive Detectors.
- Vehicle Probes.
- Mobile Reports.
- Closed-circuit Television Cameras.
- Environmental Sensors.

Table 3-9 provides a summary of the detectors discussed in this section. The majority of detectors used for traffic data collection purposes are of either the embedded or non-intrusive type. Within these two groups, there are many alternative detectors, and each has advantages and disadvantages.

Table 3-10 provides summaries of the characteristics for embedded and non-intrusive detectors that are either currently available or in the developmental stage. The following discussions will address each of these detectors in more detail. The following areas are addressed for each traffic detector discussed:

- Characteristics.
- Applications.
- Installation requirements.

EMBEDDED DETECTORS

An embedded detector system consists of sensors in or below the surface of the roadway. These detectors are currently the most widely used form of vehicle detection. The main detectors being used include the following:

Table 3-8. Communication Media. [3]

Medium	Type	Data Type	Bandwidth	Distance Capabilities
Fiber Optic	Physical	Video, digital, voice	High	Long
Coaxial	Physical	Video, digital, voice	Moderate	Short
Twisted Pair	Physical	Digital, voice	Low	Long
Microwave	Airwave	Video, digital, voice	High	Line-of-Sight
Radio	Airwave	Digital, voice	Low	Long
Spread Spectrum	Airwave	Digital, voice	Low	Line-of-Sight
Cellular Telephone	Airwave	Digital, voice	Low	Moderate
Citizen Band Radio	Airwave	Digital, voice	Low	Moderate

- Inductive loop detector.
- Magnetometers.

Inductive Loop Detector

Characteristics

The inductive loop detector (ILD) is by far the most common form of detector used for traffic management purposes. Figure 3-4 illustrates the used of loop detectors for freeway surveillance purposes. The principal components of an inductive loop detector include the following:[11]

- One or more turns of insulated wire buried in a narrow, shallow saw-cut in the roadway.
- Lead-in cable that connects the loop to the detector via a roadside pull-out box.
- Detector unit (or detector amplifier) that interprets changes in the electrical properties of the loop when a vehicle passes over it.

The loop system becomes active when the detector unit sends an electrical current through the cable, creating a magnetic field in the loop. When a vehicle passes over the loop, the ferrous material in the vehicle causes a decrease in the inductance of the circuit. This increases the frequency of oscillation that is sensed by the detector unit.

Applications

Loop detectors can measure many traffic parameters including the following:

- Vehicle count.
- Speed.
- Occupancy.
- Presence.
- Vehicle classification.

Loop detectors can operate in either a pulse or presence mode. In the pulse mode, a short signal (typically 0.125 s) is sent from

Table 3-9. Summary of Detectors.

Detector Type	Detector	Data Collection	Information Type
Embedded	Inductive Loop	Site Specific	Data
	Magnetometer	Site Specific	Data
Non-Intrusive	Radar	Site Specific	Data
	Infrared	Site Specific	Data
	Ultrasonic	Site Specific	Data
	Acoustic	Site Specific	Data
	Video Imaging	Site Specific	Video, Data
Vehicle Probes	Automatic Vehicle Identification	Site Specific	Data
	Automatic Vehicle Location	Variable Location	Data
	Cellular Telephone Probes	Variable Location	Data
Mobile Reports	Cellular Reporting	Variable Location	Voice
	Service Patrols	Variable Location	Voice
	Call Boxes	Site Specific	Coded Message
Closed-Circuit Television (CCTV)	Fixed Location Cameras	Variable Location	Video
	Portable Cameras	Variable Location	Video
Environmental Sensors	Freeways	Site Specific	Data
	Tunnels	Site Specific	Data

the loop to the detector unit and can be used to provide volume counts. In the presence mode, the signal that is sent to the detector unit lasts as long as the vehicle is in the detection area. Presence detectors are used to provide volume counts and occupancy measurements. Presence detection is used for most detector applications, and is the preferred mode for most system management purposes.

Table 3-10. Summary of Traffic Detectors.

Detector Type	Detector	Description	Advantages	Disadvantages
Embedded	Inductive Loop	Coil of cable embedded in the pavement surface that creates a magnetic field. Vehicle is detected when this magnetic field is disturbed.	• Flexible design. • Wide range of applications. • Provides basic traffic parameters (e.g., volume, speed, presence, occupancy).	• Installation requires pavement cuts. • Installation and maintenance requires lane closure. • Detectors subject to stresses of traffic.
	Magnetometer	Small cylinders containing sensor coils that operate in a manner similar to inductive loops. Developed as alternative to loop detectors for special situations.	• Can be used in situations where loops are not feasible (e.g., bridge decks). • Less susceptible than loops to stresses of traffic.	• Installation requires pavement cuts. • Installation and maintenance require lane closure. • Small detection zone. • Typically used onlyto provide count and occupancy.
Non-Intrusive	Microwave Radar	Transmits electromagnetic energy toward vehicles on roadway. Traffic parameters are calculated by measuring the return signal frequency from vehicles.	• Generally insensitive to weather conditions. • Provides day and night operation.	• Requires FCC license for operation and maintenance. • May lock on to the strongest signal (e.g., large truck).
	Infrared	Active infrared detectors transmit electromagnetic energy. Passive infrared detectors do not transmit energy but measure the amount of energy that is emitted by objects in the field of view.	• Active detector emits narrow beam allowing for accurate determination of vehicle position. • Provides day and night operation. • Provides most basic traffic parameters. • Passive detectors can be used for strategic loop replacement.	• Operation affected by precipitation (e.g., rain, fog, etc.). • Difficulty in maintaining alignment on vibrating structures.

Table 3-10. Summary of Traffic Detectors (cont.).

Detector Type	Detector	Description	Advantages	Disadvantages
Non-Intrusive (Continued)	Ultrasonic	Transmits sound waves at frequencies between 20 and 200 kHz. Detects vehicle by measuring return waves.	• Provides most basic traffic parameters.	• Environmental conditions (e.g., temperature, humidity, air turbulence, etc.) can affect performance. • Snow covered vehicles are difficult to detect. • High level of special maintenance capability is required.
	Acoustic	Uses microphones along with signal processing technology to listen for sounds associated with vehicles.	• Completely passive. • Generally insensitive to weather conditions. • Provides day and night operation.	• Relatively new technology for traffic surveillance.
	Video Image Processing	Video image processors receive information from video cameras and use algorithms to analyze the video image input.	• Location or addition of detector zones can easily be done. • Provides basic traffic parameters. • Provides wide-area detection.	• Inclement weather, shadows, and poor lighting can affect performance. • May require significant processing power and a wide communication bandwidth.

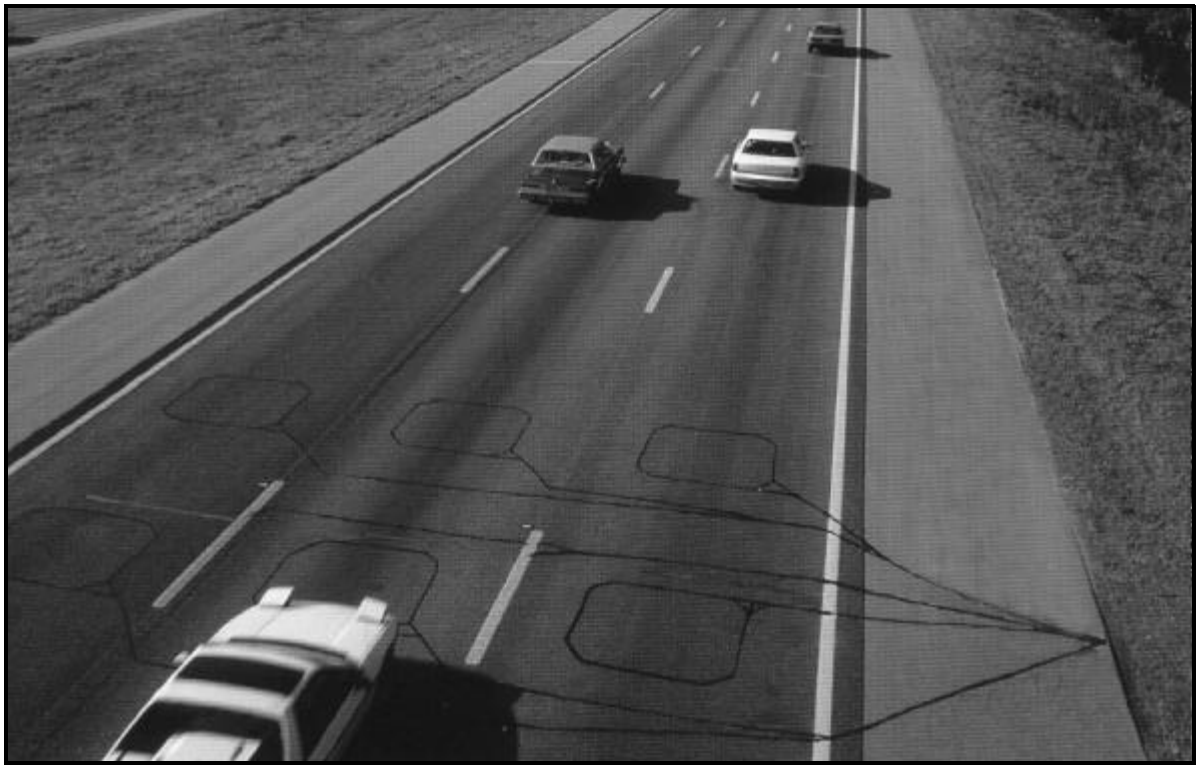


Figure 3-4. Inductive Loop Detectors Used for Freeway Surveillance. [12]

Loops can be used to detect vehicle speeds by placing two loops in pulse mode a short distance apart (see figure 3-5).[4] The distance between the loops divided by the time required for a vehicle to travel between the loops provides the speed of the vehicle.

Installation Requirements

The ILD provides for a wide range of vehicle detection because of the flexibility of its design. Loop configurations are generally grouped into two areas: short loops and long loops. For vehicle detection, the short loop configuration is recommended. The most common loop size for traffic management purposes is 1.8 m by 1.8 m (6 ft by 6 ft). The maximum length of a short loop should be 3.1 m (10 ft).[3] Detailed information concerning the design, installation, and maintenance of ILDs is contained in the *Traffic Detector Handbook*.[11]

The most important process in implementing a loop detection system is the one related to the installation procedures. Installation procedures will have a significant effect on the long term operational effectiveness of the detector system. Improper installation techniques may result in detector failures and require extensive maintenance. Installing the wire loops in the pavement requires the following steps:

- Cutting a slot in the pavement.
- Cleaning and drying the slot.
- Laying in the detector wire.
- Sealing the saw cut.

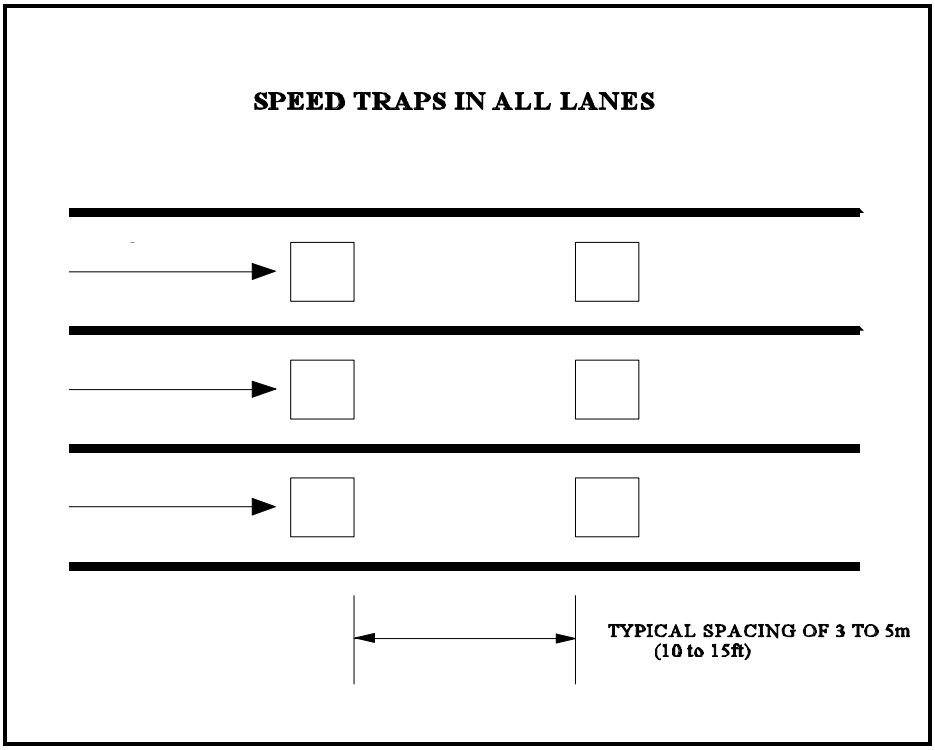


Figure 3-5. Typical Loop Installation for Measuring Speed. [4]

- Connecting the wire loop to the lead-in cable.
- Connecting the lead-in cable to the detector unit.

The main concern with the use of ILDs is reliability. Because the detectors are embedded within the pavement, they are subject to the stress of traffic traveling on them and to pavement deterioration. In northern States, other problems are associated with freeze-thaw conditions.[13]

The primary causes of loop failure include the following:

- Pavement problems (cracking and moving).
- Deterioration of wire insulation.
- Poor sealants or inadequate sealant application.
- Inadequate splices or electrical connections.
- Damage caused by construction activities.
- Improper detector unit tuning.
- Detector unit failure.
- Lightning/electrical surges.

Most of these problems can be traced back to improper installation techniques.

Magnetometer

Characteristics

Magnetometers contain small probes that range in size from 5 cm (2 in) to 11 cm (4.25 in) in diameter. The magnetometer was developed as an alternative for loop detectors in special situations, such as bridge structures. Figure 3-6 illustrates a typical installation for magnetometers. Similar to loop detectors, the components of a magnetometer detection system include the following:[11]

- One or more small probes (sensors) embedded in the pavement.
- Probe cable.
- Lead-in cable that connects the probe to the detector via a roadside pull-out box.
- Detector unit.

Magnetometers measure the density of vertical flux lines of the earth's magnetic field. When a vehicle passes over the probe, the ferrous material in the vehicle increases the density of the flux lines. Magnetometers sense this increase and interpret it as the presence of a vehicle.

Applications

Magnetometers can operate in either presence of pulse modes and are most effective in determining occupancy and volume. This type of system is usually used when the only information required is that a vehicle has arrived at a specific point.[11] Speed can also be measured by placing two sets of probes a known distance apart.

Installation Requirements

Similar to the installation of inductive loops, installing magnetometers requires cutting the pavement; however, because magnetometers are more compact in size, not as much cutting is required. Detailed information concerning the design, installation, and maintenance of magnetometers is contained in the *Traffic Detector Handbook*.[11]

For a magnetometer to detect a vehicle, some part of the vehicle must pass over the probe. Therefore, optimum lateral placement of probes in the pavement is dependent upon the following factors:

- Width of narrowest vehicle to be detected.
- Width of lane.
- Required detection quality.

Since most vehicles are wider than 1.5 m (5 ft), a single probe might not be adequate for lane widths greater than 3 m (10 ft). For lane widths of 3.7 m (12 ft), two probes per lane should be adequate for detecting all four-wheeled vehicles.

Installing the magnetometers requires placing the detectors in the pavement. The installation procedures include the following steps:

- Drill holes in pavement for probes.
- Saw slots in pavement for detector wires.
- Clean and dry saw cuts.

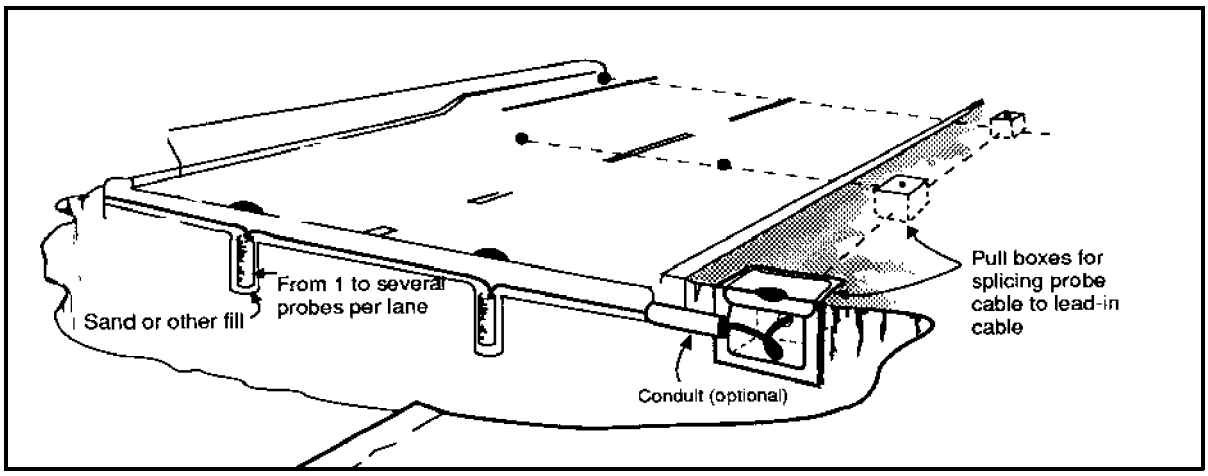


Figure 3-6. Typical Layout for Magnetometer Installation.[4]

- Place each probe in the proper hole.
- Secure probe with housing, sand, or sealant.
- Connect detector wire to lead-in cable.
- Connect lead-in cable to the detector unit.
- Test system.
- Seal drilled holes.

Because the probes are buried several centimeters below the surface of the pavement, magnetometers are less susceptible than ILDs to deteriorating pavement conditions. Therefore, magnetometers are primarily used in northern States that suffer pavement deterioration due to freeze-thaw cycles. In addition, the magnetometer can usually be used on bridge decks where cutting the pavement for loop installation is not an option.

When using magnetometers on bridge structures, the probes should be placed at the maximum distance from steel supports (see figure 3-7). The presence of steel over or under the probes has little effect on performance; however, vertical structural steel members may influence the magnetic field surrounding the probe, and affect performance.

The causes of magnetometer failures are usually either improper installation or maintenance procedures. The major factors that affect the operation of a magnetometer detector system include the following:

- Proper burial depth of probe.
- Stability of the probe in the pavement.
- Moisture penetration into the probe cable.
- Saw slot maintenance.

NON-INTRUSIVE DETECTORS

In response to an increasing demand for alternatives to loop detectors, a broad range of non-intrusive detectors have become

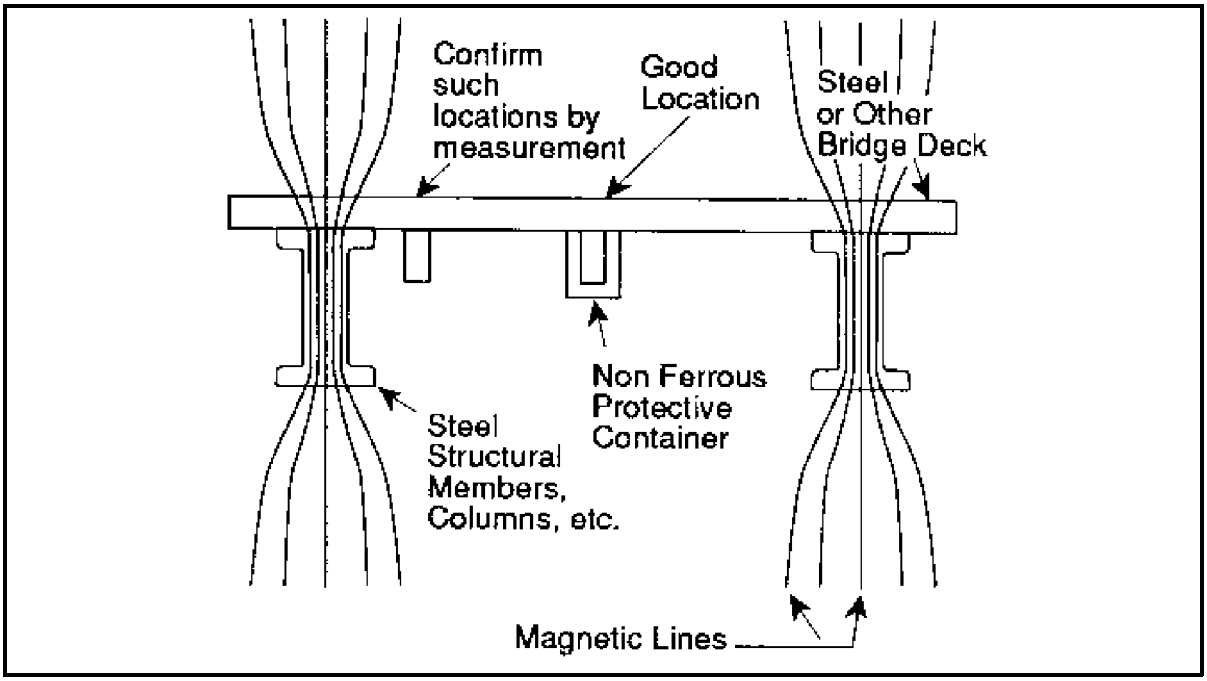


Figure 3-7. Typical Bridge Deck Installation for Magnetometer. [4]

available. [14] Non-intrusive detectors are mounted on a structure above the surface of the pavement. Figure 3-8 shows an example of a non-intrusive detector. Advantages that non-intrusive detectors have over embedded detectors include the following:[15,16]

- Modifications to pavement are not required for installation.
- Detectors can be moved or replaced more easily.
- Lane closure may not be required during installation and maintenance.
- Detectors are not subject to stresses of traffic.
- Detectors can be used during and after any reconstruction or maintenance activities.

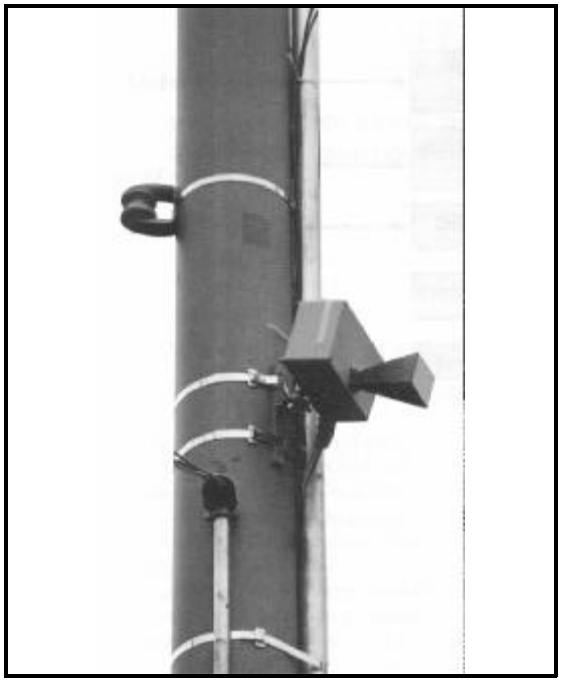


Figure 3-8. Example of Non-Intrusive Detector. [17]

A disadvantage of non-intrusive detectors is that they may produce unreliable results during adverse weather conditions.

Over the past few years, however, improvements in technology have improved the performance and reliability of certain non-intrusive detectors, and costs have become more and more competitive with loops. There are currently many non-intrusive technologies that provide viable alternatives to loop detectors. These alternatives include the following:

- Microwave radar.
- Infrared.
- Ultrasonic.
- Acoustic.
- Video image processing.

Microwave Radar

Characteristics

Microwave radar detectors have been used in both law enforcement and traffic management for some time to monitor vehicle speed. For traffic management, radar sensors are mounted either above or beside the road and direct a beam of microwave energy onto a detection area. Most radar detectors transmit electromagnetic energy at the speed of light at the K-band (24 gHz) or the X-band (10 gHz).[18]

The form of the electromagnetic wave transmitted by the detector determines the type of data that can be obtained by the unit. Three types of radar frequency include the following:[18]

- Continuous wave.
- Frequency modulated continuous wave (FMCW).
- Pulsing waveform.

Continuous Wave. With constant wave radars, electromagnetic energy is transmitted at a constant frequency. This type of radar is the most common form of radar detection. As shown in figure 3-9(a), vehicle speeds are calculated by measuring the Doppler shift that occurs in the return signal frequency. The shift in the return signal is proportional to the speed of the vehicle. Because only moving vehicles cause a frequency shift in the return signal, these detectors cannot detect stationary vehicles, and thus cannot be used as presence detectors.

Frequency Modulated Continuous Wave (FMCW). FMCW radar detectors can be used to measure the presence of vehicles as well as the speed. Figure 3-9(b) illustrates a typical waveform of these types of detectors. Presence detection occurs in the portion of the waveform where the frequency changes with time. The portion of the waveform that is constant is used to measure the speed of vehicles in the same manner as the continuous wave detectors.

Pulsing Waveform. Pulsing waveform is a variation of the FMCW signal. As illustrated in figure 3-9(c), it is generated by using differences in frequency change rates, time duration of each waveform segment, and numbers and sequencing of the waveforms. Detectors that use this type of waveform can measure differences between the range to the road and the vehicle, making it possible to provide not only vehicle counts and presence detection, but also occupancy. Speeds can also be determined by measuring the elapsed travel time between two detectors.

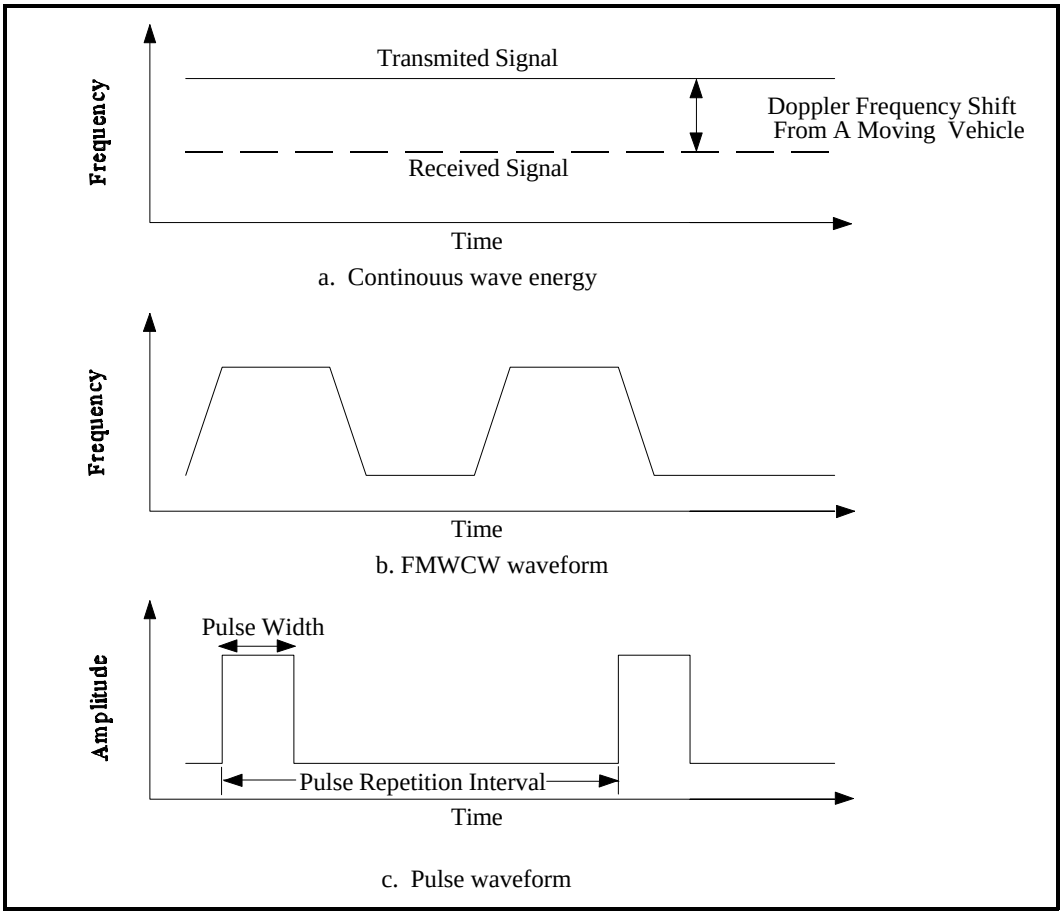


Figure 3-9. Waveforms Used with Microwave Detectors. [18]

Depending on the width of the radar beam, radar detectors may be used to monitor from one to several lanes. Three types of radar detectors exist:[13]

- Wide beam.
- Narrow beam.
- Long-range.

Wide Beam. Wide beam detectors monitor all lanes of traffic in one direction of the freeway (see figure 3-10(a)). They provide information on the overall traffic speeds; however, every vehicle is not detected.

Narrow Beam. Narrow beam detectors monitor only a single lane of traffic in one direction (see figure 3-10(b)). They are typically installed where freeway lanes have varied uses (e.g., through lanes and exit lanes, or through lanes and HOV lanes).

Long-Range. Long-range detectors project the radar beam over greater distances, up to 245 m (800 ft). This type of detector may be installed alongside a freeway with traffic traveling in one direction to measure speeds of freeway traffic traveling in the opposite direction (see figure 3-10(c)).

One advantage of radar detectors is that they can discriminate between approaching and receding traffic. [15] This capability would be

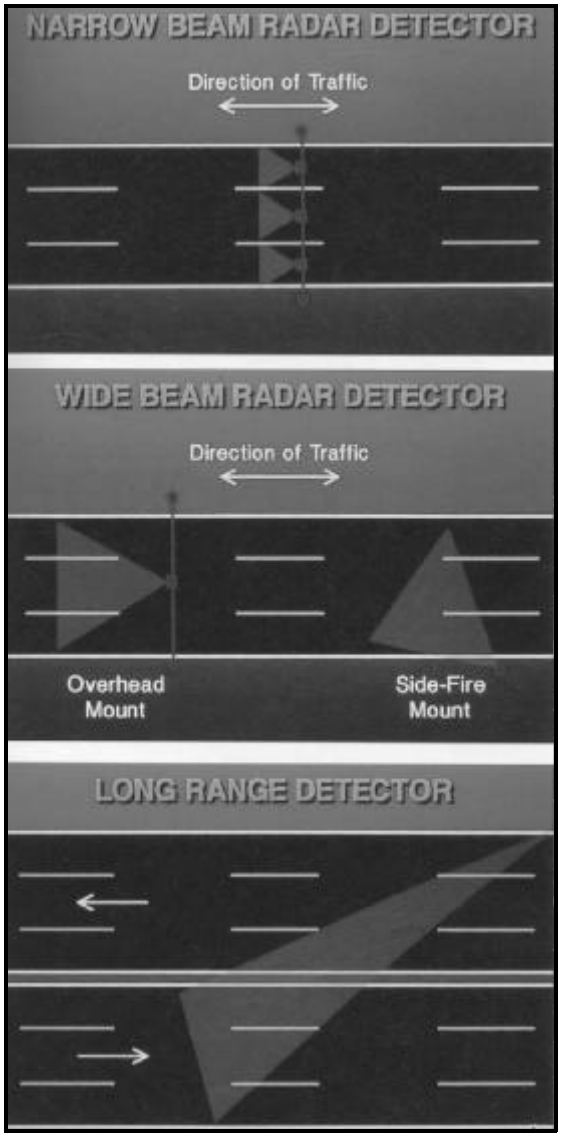


Figure 3-10. Alternative Radar Beam Widths. [13]

useful for monitoring operations on reversible or contraflow lanes.

Applications

The most common applications of microwave radar detectors are to measure vehicle speeds, and in some applications to provide volume counts. As mentioned above, advanced units that are available can be used as presence detectors. In addition, some radar units can measure vehicle profiles to perform classification counts.[18]

An area in which the use of microwave radar detectors is increasing is incident detection. Incident detection algorithms used with radar detectors typically detect incidents by using speed data. This differs from the majority of incident detection algorithms, which detect incidents based on occupancy data. A recent study suggests microwave radar incident detection systems are highly accurate and reliable.[13]

One of the biggest advantages of microwave radar detectors is their ability to perform adequately in all weather conditions. Because the detectors are mounted above the pavement surface, they are not subject to the effects of ice, sand, and salt during the winter months, as are many embedded detectors. Microwave radar detectors can be useful under the following conditions:[13,19,20]

- Rain.
- Fog.
- Snow.
- High winds.
- Day and night.

Installation Requirements

One advantage of using non-intrusive detectors is ease of installation. This is especially true if the detectors can be mounted on existing structures, such as the following:

- Bridges.
- Luminaire poles.

- Sign structures.

Radar detectors can be mounted in either the side-fire position (see figure 3-11) or the overhead position (see figure 3-12). Setup of the detector is simple and involves simply aiming the unit at the traffic flow. Unlike other detection systems, most current radar units do not require setup, configuration, or calibration to provide accurate data.[13]

Maintenance costs of radar detectors have also proven to be relatively low. One study conducted in a northern State showed only a 2 percent failure rate for radar detectors over 2 years. This compares with a 34 percent failure rate of loop detectors over a 5-year period.[13] Costs are also kept lower, because maintenance of the detectors typically does not require lane closure on the freeway.

Infrared

Characteristics

There are two types of infrared detectors:

- Active.
- Passive.

Active. Active infrared detectors (see figure 3-13) operate by directing a narrow beam of energy toward a background, such as the surface of a roadway, at a certain pulse rate. A portion of the beam is directed back to the sensor, and vehicles are detected by changes in the characteristics of the infrared beam. The infrared beam can be transmitted from one side of the road to the other, or from an overhead or roadside position to a device in the pavement surface. Disadvantages of active infrared sensors include the following:[3]

- Difficulties of maintaining alignment on vibrating structures.
- Inconsistent beam patterns caused by changes in infrared energy levels due to passing clouds, shadows, fog, and precipitation.
- Lenses used in some devices may be sensitive to moisture, dust, or other contaminants.
- Inconsistent reliability under high volume conditions.

An advantage of active infrared detectors is the very narrow beam width that they emit. This allows for accurate determination of the spatial position of a vehicle on a road. In addition, the profile of a vehicle can be measured to within a few centimeters for some systems.[22]

Passive. Passive infrared detectors (see figure 3-14) do not transmit energy themselves, but measure the amount of energy that is emitted by objects in the field of view. The amount of energy that is emitted by an object depends on its surface temperature, size, and structure, but not on its color or the surrounding lighting conditions. For years, these detectors have been used as motion detectors; however, detectors are currently available for detecting presence as well.[14] As with active infrared detectors, the ability of passive sensors to detect vehicles can be affected by environmental effects, such as fog and precipitation, that can scatter and emit energy of their own. [18]

Applications

Active. Currently available active infrared detectors can provide speed, count, density, and vehicle classification. Accuracy is based

Figure 3-11. Detector in Side-Fire Position. [21]

Figure 3-12. Detector in Overhead Position. [19]

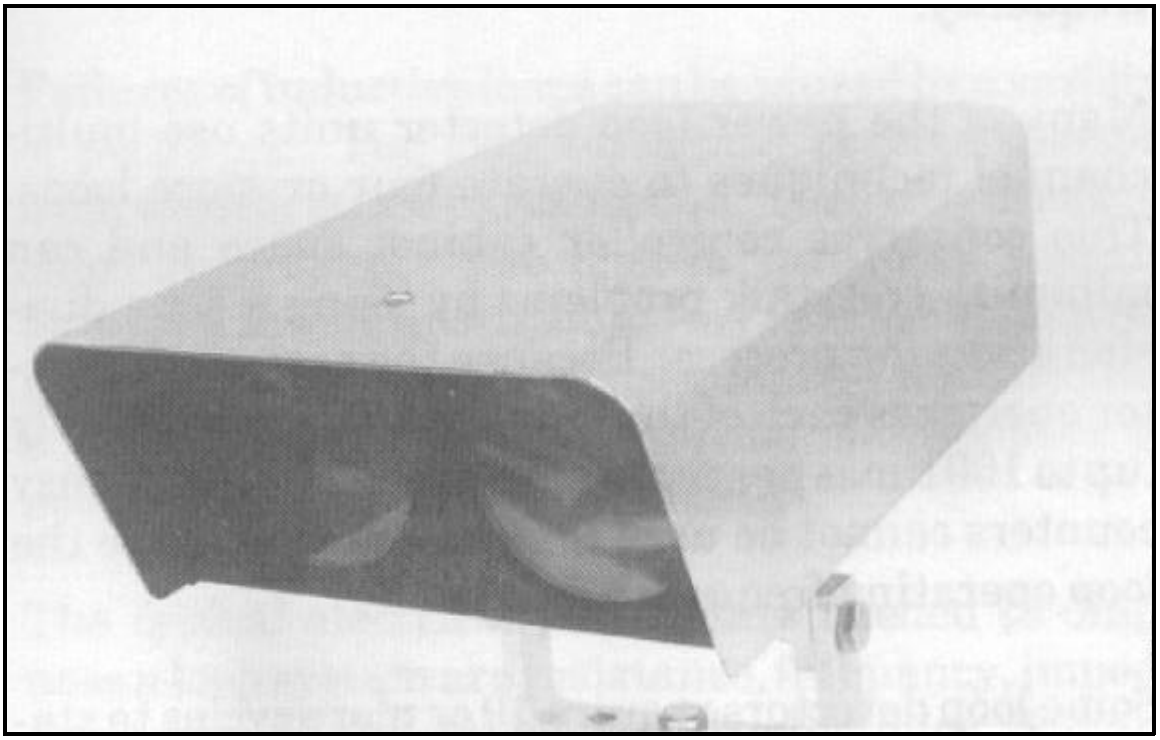


Figure 3-13. Active Infrared Detector. [11]

on the pulse rate, scan rate, and speed of vehicle. Faster pulse rates and scan rates improve the accuracy of the system up to a certain point.[22]

The presence of a moving or stationary vehicle is determined by measuring the round-trip propagation time of an infrared pulse. This time will be shorter when a vehicle is present. Speeds are measured by using two fixed beams, one slightly ahead of the other. By comparing the times at which the front of a vehicle passed through each beam, the speeds can be determined.[22]

One system that is coming onto the market can determine vehicle classification by measuring the profile of a vehicle. A two-dimensional profile of a vehicle passing through the infrared beam is obtained by measuring the distance from the detector to the vehicle (see figure 3-15). If the beam is scanned across the roadway at a fast enough pulse rate, a very accurate profile of a vehicle can be obtained. The vehicle may then be classified by using an algorithm to compare the vehicle's profile against defined profiles for various vehicle classifications. [22]

Passive. Passive infrared detectors can provide volume counts as well as presence detection. Using a multichannel presence detector, passive infrared detectors can be made to mimic the pattern of an inductive presence loop, and can replace malfunctioning loop detectors.[14]

Installation Requirements

For speed measurement and presence detection, infrared detectors may be installed in either the side-fire or overhead position.

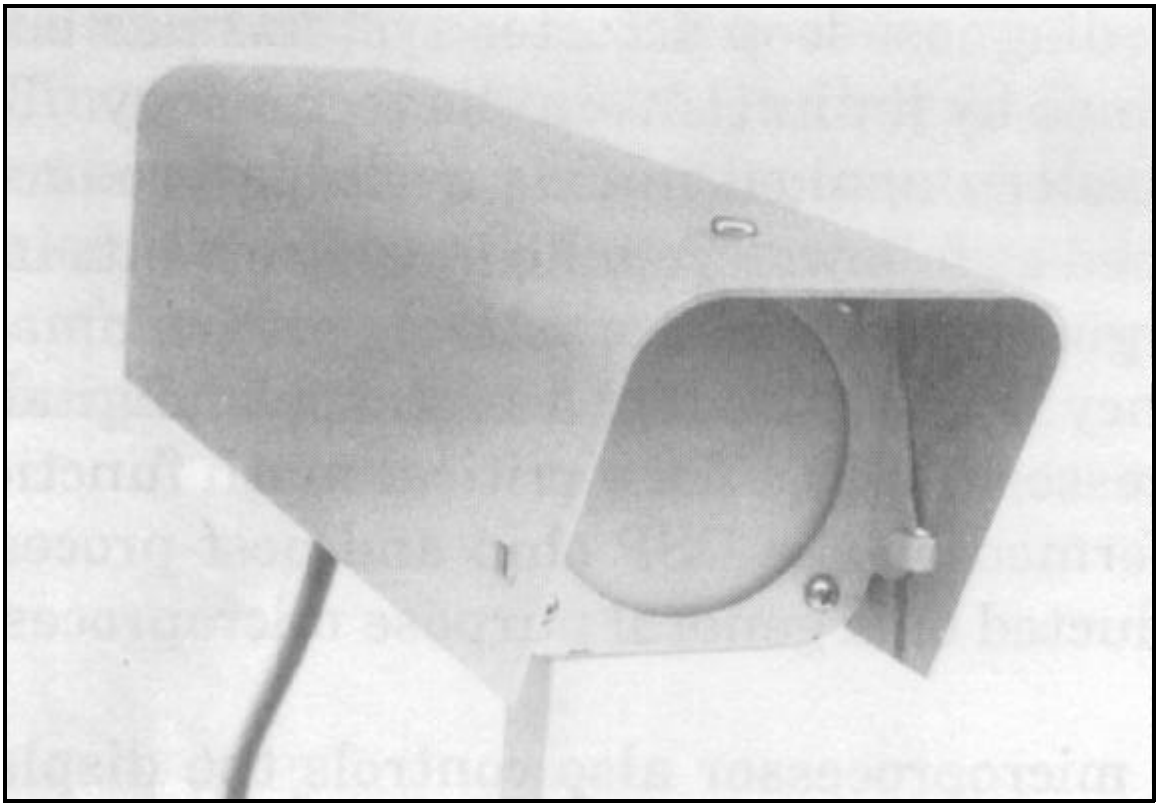


Figure 3-14. Passive Infrared Detector. [11]

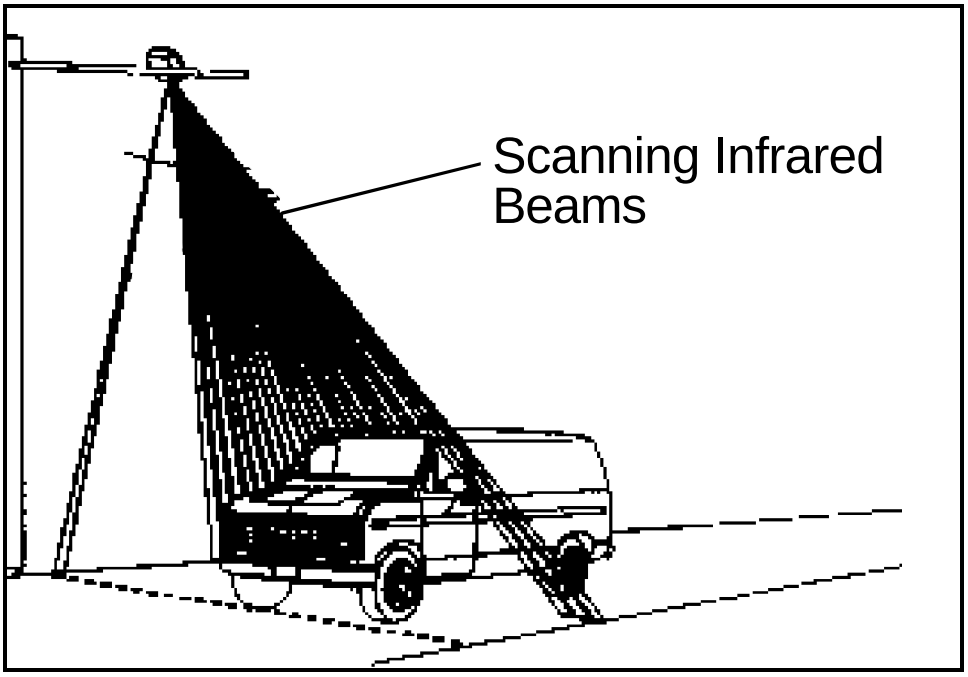


Figure 3-15. Active Infrared Detector Measuring Vehicle Classification. [23]

In order to obtain vehicle classification, the detector must be mounted in the overhead position. Active infrared detectors are suitable for single lane detection, but can be used only at short ranges. Passive infrared detectors offer lane discrimination at long range, but have relatively slow response time. [14]

Some currently available active infrared systems have incorporated microcontrollers that offer continuous built-in testing. For these systems, all adjustments to the system are automatically performed, and no initial calibration is required during installation. In addition, maintenance or design changes may be performed by uploading the system's program code, thereby, eliminating the need to remove the detector from the mount.[22]

Ultrasonic

Characteristics

Ultrasonic detectors use electronic sound wave signals and a receiving unit to detect vehicles traveling in a traffic stream. These detectors operate on the same principle as microwave radar detectors in that both transmit a beam into an area and receive the reflected beam to detect a vehicle. [11] Similar to radar detectors, ultrasonic detectors operate using the following waveforms:[22]

- Continuous.
- Pulse.

For a continuous waveform, a continuous signal is emitted, and vehicle presence is measured using the Doppler principle. This form can be used to detect volume, occupancy, and speed. For pulse waveform, the detector operates by pulse allowing the measurement of classification, as well as volume and presence.

The development of ultrasonic devices for vehicle detection began in the mid-1950s. Michigan, Illinois, New York, and California were among the early users of ultrasonic detectors in the 1960s. The use of these detectors was abandoned for the most part because of the problems experienced. Recently, however, there have been efforts to improve ultrasonic vehicle detectors. Results from these efforts continue to show promise.[11]

Disadvantages of current ultrasonic vehicle detection systems include the following:[3]

- Environmental conditions (such as temperature, humidity, and air turbulence) can affect operations because the detectors use sound waves that propagate through the air.
- Surface of a vehicle may affect the performance of the detector. For example, porous or textured surfaces (e.g. snow on a vehicle) produce weaker reflected sound waves.
- Speed is not measured directly.
- Systems require a high level of special maintenance capability.

Applications

Ultrasonic detectors can be used for both presence and pulse applications. Using signal processing techniques, the reflected waves may be converted into volume, speed, and occupancy measures. In addition, some detectors can be used for presence detection and vehicle classification. Vehicles are classified by comparing the sonic signature created by a vehicle with a set of pre-programmed signatures of vehicles of various classes.[22]

Installation Requirements

Ultrasonic detectors can be mounted in either the side-fire or overhead position. Detectors can provide either single lane or multi-lane coverage. An example of an ultrasonic detector installation is shown in figure 3-16.

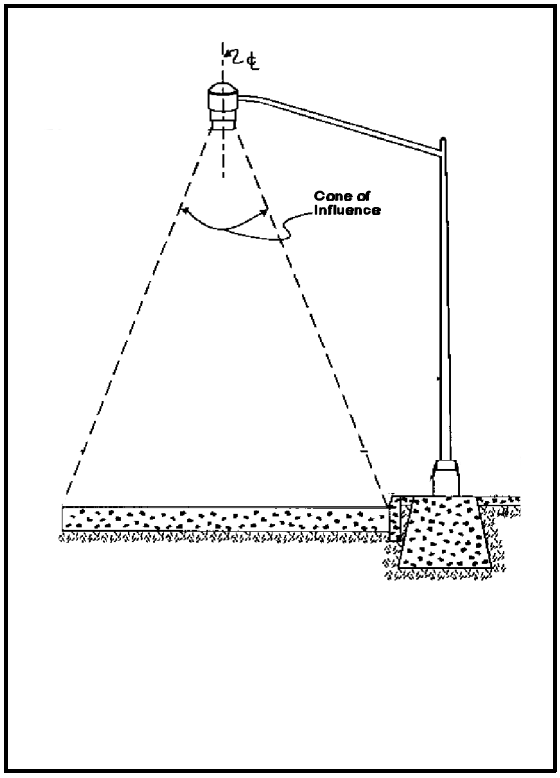


Figure 3-16. Example of Ultrasonic Detector Installation. [11]

Acoustic

Characteristics

Acoustic detectors are completely passive devices. Vehicles are detected by using microphones along with signal processing technology to listen for sounds associated with vehicles. The acoustic technology used by these detectors was originally developed for military defense purposes, such as tracking enemy submarines. The use of this acoustic technology for vehicle detection purposes has emerged within the last few years. The advantages of acoustic detectors are that they work well in all lighting conditions and in wide temperature and humidity extremes. [22]

Applications

Acoustic detectors can be used to measure volume, occupancy, and presence. Vehicle classification can also be determined by comparing the sonic signature of a vehicle to pre-programmed sonic signatures of vehicles for various classes. [22]

Speed can be measured by using an array of microphones as shown in figure 3-17. The configuration shown in this figure consists of two microphones, one mounted above the other. The concept behind this configuration is that the time delay of sound arrival will be different for the upper and lower microphones. This difference in time delay will vary as a vehicle approaches the sensors. For example, when the vehicle is a certain distance from the sensors, the sound arrives at both microphones almost simultaneously. When the vehicle passes under the sensors, the sound reaches the lower microphone first. By plotting time delay versus time, the relationship can be used to estimate speed (see figure 3-18).[24]

Installation Requirements

Acoustic detectors can be mounted either on the side of the roadway or in an overhead position. The detectors can monitor one or many lanes. Multilane detectors cannot distinguish between lanes, and may experience interference among vehicles in different lanes. Single lane detection is accomplished using directional microphones; however, this increases the complexity and cost of the system. [24]

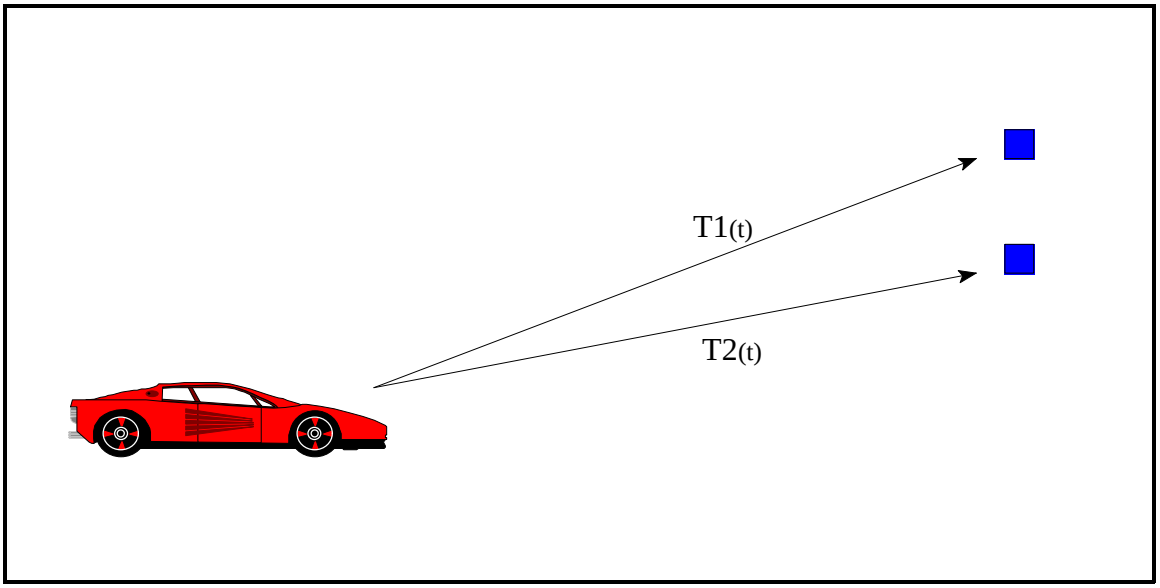


Figure 3-17. Acoustic Detector.

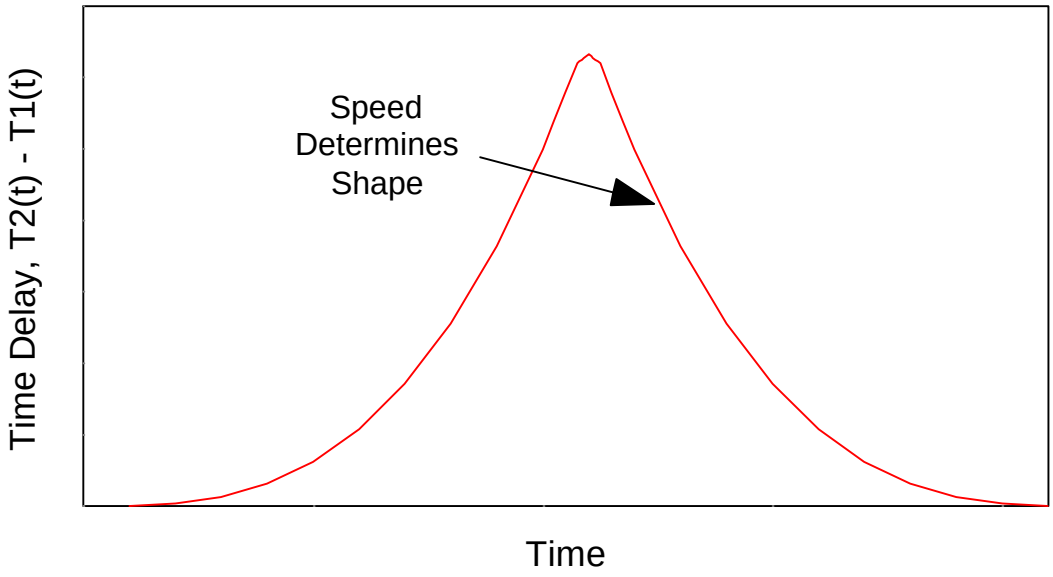


Figure 3-18. Time Delay Versus Time Curve.

Video Image Processing

Characteristics

An emerging technology that appears to be very promising for meeting future data collection and surveillance needs is known as video image processing (VIP). VIP systems detect vehicles by monitoring specific points in the video image of a traffic scene to determine changes between successive frames.[25] The components of a VIP system are shown in figure 3-19. Most systems consist of the following major elements:

- One or more video cameras.
- Microprocessor-based system for processing the video image.
- Software for interpreting the processed images as vehicle detections.

The microprocessor-based system receives video inputs from the cameras. After receiving video input, the processor analyzes the variation of gray levels in a series of pixels from the video image. The processor must filter out gray level variations resulting from weather conditions, shadows, and day- or night-time-related aspects. The resulting image consists of a blank background containing only identified objects such as vehicles, motorcycles, and/or bicycles. By analyzing successive video frames, the system is able to calculate vehicle-related information.[20]

Classes of VIP Systems. VIP systems have evolved through the following three classes:

- Tripline.
- Closed Loop Tracking.
- Data Association Tracking.

Tripline. These first generation systems are the least demanding in terms of computer power and speed. Most of the VIP systems that are commercially available at this time are of this class. Tripline systems operate by allowing the user to define a limited number of detection zones in the field of view of the video camera. When a vehicle enters one of these detection zones, it is identified in a manner analogous to that of inductive loops. In fact, tripline systems are the functional equivalent of inductive loops and are intended to replace inductive loops in areas where a large number of loops are employed. Limitations of tripline systems include the following:

- Difficulty in detection due to the presence of shadows or light changing conditions.
- Problems of occlusion (vehicle hidden by another vehicle or object).

Closed Loop Tracking. These systems are second generation VIP systems. They are an extension of the tripline approach in that detection is performed using the same type of detection zones. Closed loop tracking systems represent the first attempt to perform vehicle tracking. These systems have the same problems as tripline systems with shadows and occlusion. They provide more traffic flow information than tripline systems, but the complexity of both hardware and software subsystems is significantly greater than for tripline systems.

Data Association Tracking. These systems, commonly used in satellite surveillance systems, are third generation VIP systems. A basic requirement for these systems is the capability to identify and track a distinguishable object as it passes through the field of view of the camera. In this mode, the computer identifies vehicles by

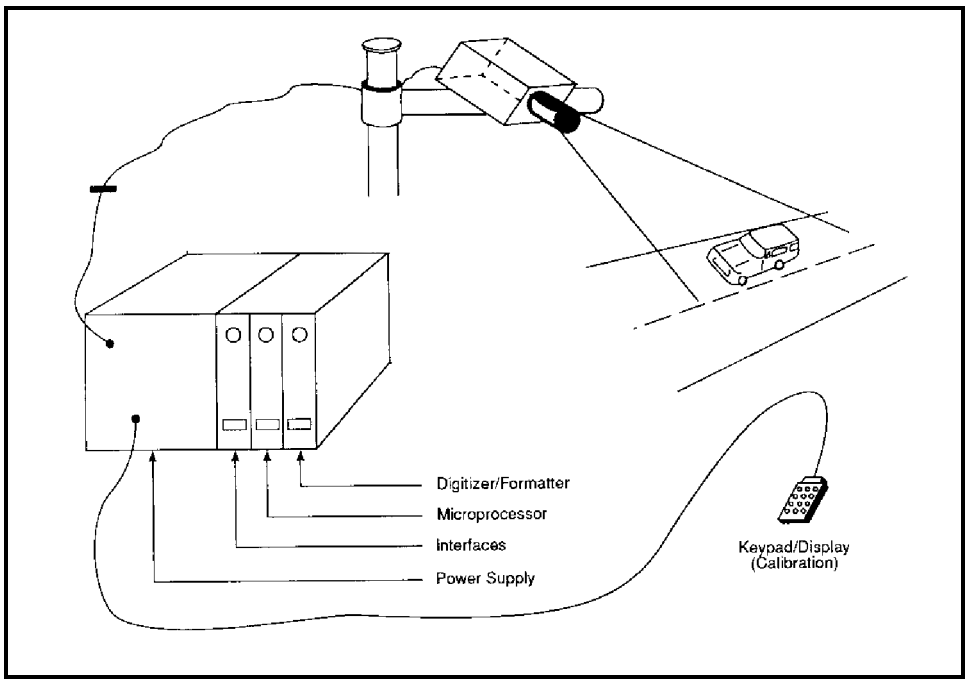


Figure 3-19. Components of Video Image Processing System. [11]

searching for connected areas of pixels that indicate motion when compared with the background information. A series of such vehicle detection is then associated to produce tracking data for each vehicle.

Data association tracking systems require less processing power than closed loop tracking systems, because they do not have to operate at the frame rate of the camera. They also offer better performance with regard to shadows and occlusion. Shadows are addressed using image analysis. Observed differences in the geometry of the image reduce the effects of occlusion. A greater reliance on software sophistication may reduce the hardware costs for these systems. An additional advantage of these systems is that a series of video cameras can be used to cover a wide area, and a vehicle can be "handed off" and tracked from one sensor to another as it passes from one field of view to another.

VIP Detector Performance. Factors influencing detector performance include the following:[20]

- Type of image being processed (i.e., upstream or downstream image).
- Mounting height of the video camera (affects occlusion of vehicles).
- Number of lanes being processed.
- Stability of video camera with respect to wind and vibration.
- Inclement weather, shadows, and poor lighting.

In most instances, the factors that affect VIP detection capability the most are shadows and reduced visibility due to inclement weather and poor lighting on detection capability. As discussed above, a VIP system detects vehicles by analyzing changes in pixels. Therefore, it may be difficult to

differentiate between vehicles and changing light conditions in all situations.[16]

During the implementation of a VIP system at signalized intersections in Oakland County, Michigan, the effects of shadows and lighting were minimized by improving the detection algorithms.[16] This improvement made it possible to distinguish the direction a vehicle enters a detector, thereby, eliminating false detections due to "wrong-way" entries. Typical causes of wrong-way entries include the following:

- Shadows from vehicles in adjacent lanes.
- Light reflections (day or night) from other objects.
- Vehicles that enter the detection zone from the wrong direction.

An evaluation of the system during various environmental conditions revealed both improvements in the detection accuracy and reductions in false detections.[16] Table 3-11 presents the results from that evaluation.

Advantages of VIP Systems. The purpose of the Oakland County project was to implement an Advanced Traffic Management System (ATMS), which included installing AutoScope VIP systems at 90 signalized intersections. After 1.5 years of operation, several benefits of VIP over inductive loop detectors were noted:[16]

- Life cycle cost of video detection is lower than when using conventional loops for a typical situation.
- Visual inspection of detection performance allows for detector size and placement optimization.
- Installation requires less set up time and fewer pieces of equipment.
- Installation and maintenance can be done year-round, and require minimal traffic disruption.
- Relocation or addition of detector regions can be done with software, resulting in no disruption to traffic.

Table 3-11. Results from Oakland County Evaluation of VIP System.

Environmental Conditions	Detection Accuracy (%)	False Detection Rate (%)
High Winds, Shadows	98.1	3.0
Overcast Skies, Snow, Wind	99.6	4.2
Overcast Skies, Wet Road, Wind	96.1	2.0
Day/Night Transition	97.1	7.4
Night	98.1	6.1
Partial Cloudiness, Light Wind	96.2	4.1

- Detection is unaffected by resurfacing or construction projects.

Additional advantages of a VIP system include the following:[16]

- Camera pointing is easily done and verified by a portable video monitor or hand-held viewer.

- Camera location is very flexible, allowing mounting on existing structures.

- Installation of VIP system requires minimal training.

- Detector layout can be done either in the field or in the office.

CCTV Applications. Closed-circuit television (CCTV) cameras can be equipped with VIP capability to give operators the added capability to immediately verify incident warning messages. If an incident has occurred, operators may switch the VIP from the video imagery mode to a standard CCTV mode and monitor nearby incidents via pan/tilt/zoom controls. Once the camera viewing field has been modified, the VIP's preset detection zones are lost and must be reset prior to switching back to the video imagery monitoring mode. In the past, resetting the detection zones has been a cumbersome task; however, several VIP manufacturers have stated that this problem is being solved.[26]

Applications

Currently available VIP systems can detect traffic at a number of locations within the camera's field of view. The detector locations are specified with interactive video graphics and can be easily changed at any time. Using a video monitor, detection lines are placed in the field of view by means of a mouse or keyboard (see figure 3-20). Different detector sizes and locations can be selected. Vehicles are detected when they cross a detection line.[11]

The processing system of a tripline VIP system provides outputs comparable to those provided by loop detectors. These outputs include presence and passage as well as speed. From these measurements, other traffic parameters can be extracted, including volume and occupancy. Similar to loop detectors, speed trap detectors can extract vehicle speed and vehicle classifications as defined by length.[4] An advantage of VIP systems is that a single camera can replace many loops, offering true wide-area detection.[11]

For the more sophisticated systems in which vehicles are tracked, vehicle location and travel time through the detection zone may be obtained, as well as a higher level of accuracy in vehicle speed and vehicle length. Detector lines and speed traps are not needed with these systems. Instead, vehicles are tracked frame-by-frame through the entire field of view. The time and distance traveled between successive video frames is known, thereby providing an accurate measurement of speed.[28]

VIP systems are also very effective in incident detection. For tripline systems, incident detection algorithms similar to those used for loop detectors may be used. A video tracking system can provide traffic information to automatic incident detection systems, and can often implement these algorithms directly with the same hardware platform. Since all vehicles are tracked with this system, parameters pertaining to each vehicle are known.[28]

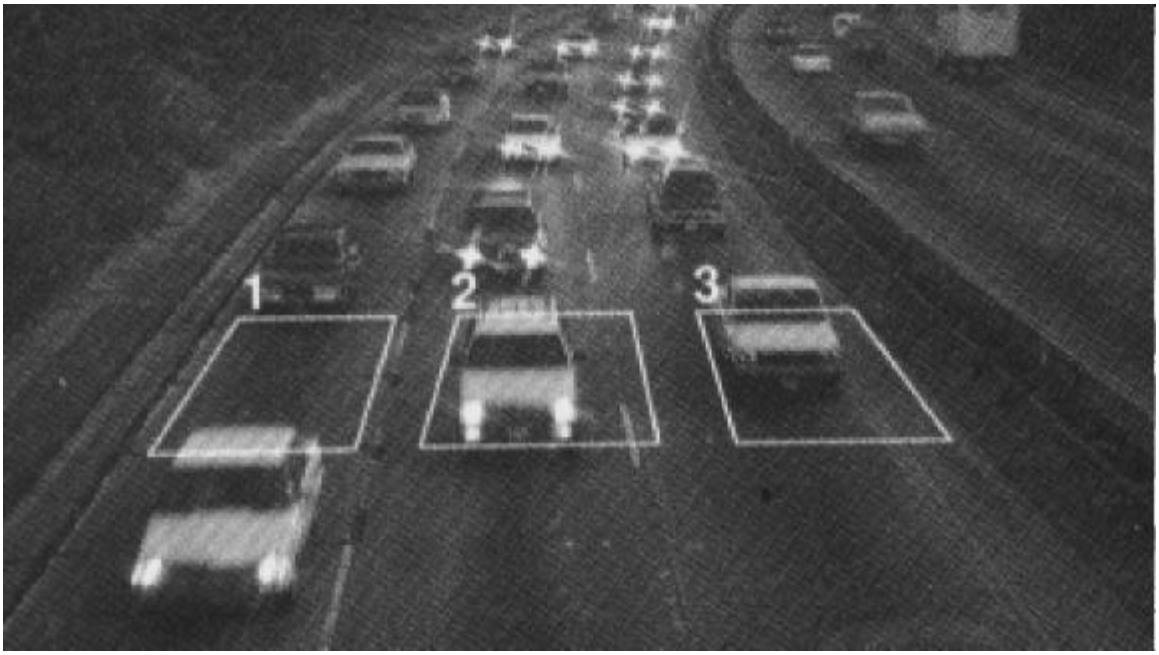


Figure 3-20. Identification of Detection Lines in Field of View.

Installation Requirements

The video detection system includes the following elements:

- Roadside cameras.
- Video processor.
- Appropriate software.

The cameras are mounted above the roadway, typically on existing poles, bridges, or other structures. Occlusion caused by other vehicles or objects can be minimized by selecting proper camera location and defining optimal detector sizing.[28]

Some existing VIP systems offer complete fail-safe operation. For example, each camera is monitored for video signal quality and picture usability. If a camera fails, the affected detection zones will be switched to a user-defined mode, such as "always on." The nature of the problem is also reported to operators in the traffic control center.[28]

COMPARISON OF EMBEDDED AND NON-INTRUSIVE DETECTORS

A summary of the advantages and disadvantages of each of the systems discussed was given toward the beginning of this section (see table 3-10). Table 3-12 lists the functional capabilities of the various systems, as well as estimated installation and detector costs. These estimates will vary by manufacturer and quantity purchased.

VEHICLE PROBES

With the recent advances in computer, communications, and vehicle locating technologies, the vehicle itself can become an important surveillance tool for monitoring traffic conditions in the roadway network. Vehicles, acting as moving sensors (or probes), can provide information about traffic conditions on each link traversed. This information can be transmitted to a central computer system where it can then be merged with information from other sources to provide an accurate representation of actual travel conditions in the transportation

Table 3-12. Characteristics of Traffic Detectors.

Detector	Applications					Lane Coverage per Sensor	Communi-cation Bandwidth	Life	Reliability	Technology	Cost	
	Count	Presence	Speed	Occup-ancy	Classifi-cation						Install	Detector (each)
Inductive Loop	X	X	X	X	X	Size of loop	Low	Moderate	Moderate	Current	Moderate ($1,000)	Low ($500-$800)
Magnetometer	X	X	(1)	X		Single lane	Low	Long	High	Current	Moderate ($1,000)	Low - Moderate ($500 - $1500)
Microwave Radar	X	(2)	X	(2)	(2)	Multiple	Moderate	Long	High	Current	Low ($500)	Low - Moderate ($700 - $3,000)
Infrared	X	X	X	X	(2)	Single (active); Multiple (passive)	Low - Moderate	Long	High	Developing	Low ($500)	Moderate - High ($1,000-$8,000)
Ultrasonic	X	X	X	X	(2)	Single	Low	Moderate	High	Developing	Low ($500)	Low - Moderate ($600 - $1,500)
Acoustic	X	X	X	X	(2)	Multiple	Low - Moderate	N/A	N/A	Developing	Low ($500)	Moderate ($1,500)
Video Image Processing	X	X	X	X	X	Multiple	Moderate - High	Moderate	Moderate	Developing	Low ($500)	Very High ($10,000 - $25,000)

(1) Speed can be calculated by spacing sensors a known distance apart.

system. Advantages of vehicle probes are that the following measurements can typically be obtained:

- Link speeds.
- Link travel times.
- Origin and destination of vehicle traveling through system.

Types of Vehicle Probes

Emerging technologies that utilize vehicles as probes include the following:

- Automatic Vehicle Identification (AVI).
- Automatic Vehicle Locating (AVL).
- Cellular Telephone Probes.

Automatic Vehicle Identification (AVI)

AVI systems permit individual vehicles to be uniquely identified as they pass through a detection area. Although there are several different types of AVI systems, they all operate using the same general principles. A roadside communication unit broadcasts an interrogation signal from its antenna. When an AVI-equipped vehicle comes within range of the antenna, a transponder (or tag) in the vehicle returns that vehicle's identification number to the roadside unit. The information is then transmitted to a central computer where it is processed. In most systems, the transponder and reader/antenna technology are independent of the computer system used to manage and process the vehicle identification information.[3]

Currently, the most common application of AVI technology is for automatically collecting tolls on tollways. In this application, toll charges are electronically deducted from the driver's account when he or she passes through a toll plaza. Because tolls are collected automatically, the vehicle can pass through the toll plaza without stopping.

AVI technology may also be used as a means of automatically collecting travel time information along freeways. In Houston, Texas, AVI systems have been installed to monitor traffic operations on the main lanes and the high-occupancy vehicle (HOV) lanes on three major freeways. Vehicles equipped with transponders are used as probes to collect current travel time information. This information is used to alert freeway operators to potential incidents and congestion on both the main lanes and the HOV facilities. [29]

The original AVI technology, which has been in use for several years, uses a radio frequency signal from the roadside to activate a transponder located in the vehicle. Transponders can be classified according to the type of source required to power the transponder and the degree to which the transponder can be programmed.

Classes of transponders, based on type of power source, include the following:[30]

- Active.
- Passive.
- Semi-Active.

Active. With active transponders, power to the transponder is supplied from either an internal battery or a connection to the vehicle's power supply. The transponder is activated by an interrogation signal from the roadside communication unit. It responds to the signal by broadcasting its own signal (which contains the identification number for the vehicle) from an internal transmitter. This type of transponder generally has a

greater operating range and is more reliable than other types of transponders. The life expectancy of an active transponder is between 7 and 10 years.

Passive. With passive systems, the transponder does not require any internal or external power supply. Instead, the interrogation signal from the antenna is modulated and reflected to the reader. Because the return signal is weaker, passive systems typically produce less lane-to-lane interference than active systems. The weak return signal, however, causes passive transponder systems to have shorter operating distances. Due to the simplicity of their circuitry, passive transponders have a life expectancy of approximately 40 years.

Semi-Active. Semi-active transponders use an operating approach similar to that of passive transponders in that they are activated only after an interrogation signal is received from the reader. Unlike passive transponders, however, semi-active transponders use an internal power to boost the return signal to the reader. This increases the reading distance of the transponder. Like active transponders, a semi-active transponder has a life expectancy of approximately 7 to 10 years.

Classes of transponders, based on the degree to which they can be programmed, include the following:[25]

- Type I.
- Type II.
- Type III.

Type I. Type I transponders are read-only tags that contain fixed data, such as a vehicle identification number. They can initially be programmed either at the manufacturing facility or by the agency issuing the transponder; however, they cannot by reprogrammed without returning the transponder to the manufacturer.

Type II. These transponders have read/write capability. In these transponders, some of the memory contains permanent information (such as vehicle identification number) and cannot be reprogrammed. However, additional memory can be provided and may be reprogrammed or written remotely from the reader. This type of transponder is typically used in toll systems to record time, date, location, and account balance for vehicles.

Type III. These transponders are also known as "smart cards." They have extended memory and are capable of full two-way communication. With this system, vehicles can be warned of incidents, congestion, or adverse weather conditions, enabling drivers to take alternative routes. This type of system requires sophisticated technology for both the roadside and vehicle-based equipment.[4]

Automatic Vehicle Location (AVL)

AVL systems enable the approximate location of a vehicle to be determined and tracked as it traverses the transportation network. These systems have many uses for many different customers, including the following:[31]

- Emergency Services - aid in dispatching emergency vehicles.
- Transit Agencies - track vehicles and provide passengers with arrival time estimations through information displays.
- Delivery Companies - plan the most efficient dispatch of fleet vehicles.

- Private Citizens - allow instant dispatch of tow truck in the event of a vehicle breakdown or to recover a stolen vehicle.

This technology can also be used to determine the severity of congestion or the occurrence of an incident, by obtaining probe reports from vehicles traveling in the network. Software in a control center can automatically monitor travel speeds and transit times of vehicles equipped with AVL technology.

There are numerous techniques and technologies than can be used for locating the vehicle, including the following:[31]

- Dead-Reckoning and Map-Matching.
- Signpost.
- Ground-Based Radio Navigation.
- Loran-C.
- Global Positioning Systems (GPS).
- Differential GPS.

Dead-Reckoning and Map-Matching. Dead-reckoning systems monitor the vehicle's internal compass and odometer and calculate its position by measuring its distance and direction from a known central starting point. Dead-reckoning systems frequently get off track and can be corrected using a technique called map-matching. Map-matching systems store a map of the vehicle's coverage area in a database and assume that when a vehicle changes direction, it must have turned from one road on to another. When a vehicle does make a turn, map-matching systems alter the vehicle's record location to the nearest possible point at which the turn could have taken place. Because of the low degree of positional accuracy of dead-reckoning and map-matching, most AVL systems use more advanced technology options.

Signpost. When vehicles, such as transit buses, regularly travel a fixed route, many fleet operators have found that signpost-based positioning systems offer an alternative to more advanced AVL technologies. Antennas are placed at locations throughout the vehicle's route and record the time when the vehicle passes nearby. A signpost-based AVL system can also be a valuable extension of systems intended for other purposes. Reader antennas that communicate with vehicle tags for electronic toll collection can also track the location of vehicles from one toll booth to another. The Harris County Toll Authority in Houston and the Illinois Toll Authority in Chicago are currently using such systems.

Ground-Based Radio-Navigation. In "terrestrial" or "ground-based" radio-navigation, the AVL vendor sets up several receiving antennas in a metropolitan area. Each appropriately equipped vehicle broadcasts a radio frequency (RF) signal to all nearby receiving antennas. By measuring the time it takes for the signal to travel to the antenna, the distance from the vehicle to the antennas can be determined. If the vehicle's signal was received by three or more antennas, the vehicle's position can be uniquely determined. A disadvantage of radio-navigation is that RF signals have difficulty transmitting through large obstructions, such as mountains, tunnels, parking garages, and metropolitan canyons formed by the large buildings that line many downtown city streets.

Ground-based radio-navigation systems are among the most inexpensive AVL systems for the user. However, since constructing

the necessary infrastructure requires significant financial investment on the part of the AVL vendor, these systems are usually available only in dense metropolitan areas with large market potential.

LORAN-C. LORAN-C is similar to ground-based radio navigation, except that it uses a communication system set up by the U.S. Government. Uncertainty about the government's future plans for this system have lead to a decrease in the number of commercial AVL systems using LORAN-C. In addition, LORAN-C communication signals often experience errors due to atmospheric conditions.

Global Positioning System. Global positioning systems (GPS) use a network of 24 satellites that are continuously orbiting the Earth to locate any object anywhere on the planet. The satellites are available free-of-charge to anyone with a device capable of receiving the satellite signals. The U.S. Department of Defense (DOD) launched the satellites in order to track objects of interest on the ground. The position of the objects is determined measuring how long a radio signal takes to reach the object from multiple satellites. GPS is by far the most accurate global navigation system ever devised, with accuracies in the range of 5 to 30 meters. Similar to radio-navigation, GPS signals have difficulty transmitting through large objects. The signals also have trouble transmitting through opaque objects, such as leaves on trees.

Differential GPS. Differential GPS is a technique used to improve the accuracy of standard GPS. With differential GPS, a receiver placed at a known location calculates the combined error in the satellite range data. By knowing the error, correction factors can be applied to all other receivers in the same locale, virtually eliminating all errors in measurements.

Cellular Telephones

Using radio frequency receivers and triangulation techniques, it is possible to determine a vehicle's location by measuring signals resulting from cellular phone usage within the vehicle. In conjunction with map-matching algorithms, vehicles (cellular probes) can be tracked as they traverse area freeways. The ability to track vehicles via cellular telephones allows vehicle speeds, as well as travel times for various freeway segments to be measured. With so many cellular telephones currently in operation, this system has a potential to provide an inexpensive source of traffic surveillance.[31]

A test project was conducted in the Washington, DC, area to test the feasibility of using cellular probes to obtain real-time travel information.[32] The surveillance system was based on passive geolocations (only vehicles that have initiated a call will be used for the period during which the call is active). This method typically provides a sufficient probe population, since cellular phone usage is directly correlated with traffic congestion.

The cellular surveillance system consists of two basic components:

- Geolocation control system (GCS).
- Traffic information center (TIC).

At the GCS, a direction finding system (consisting of a series of towers) is used to provide latitudes and longitudes of cellular probes. This information is sent to the TIC, where information on traffic conditions is derived from a series of probes. Traffic information is then disseminated to the users (e.g., operators at control center).

When an initiated call is detected by the GCS, the following information is collected and sent to the TIC:

- Time.
- Probe ID number (randomized to ensure privacy).
- Latitude and longitude.
- Information as to whether a key number (such as 911) was called.
- Information on whether the probe was within the boundary of interest.
- Confidence factor associated with the accuracy of the latitude and longitude.

If the message is within bounds, the TIC uses the position to determine which traffic link the probe is traveling on. If a key number is used, the TIC requests additional information from vehicles within a certain distance of the call. Once it has been verified that the requested probes are on the traffic link in question, speeds are calculated. The TIC keeps historical information on speed profiles for various traffic links by time-of-day. Incoming speeds are compared with these historical files. If the difference between current speed and historical speed exceeds a set threshold value, a potential incident flag is set. Data from additional probes is then requested by the GCS. If the threshold value is exceeded a certain number of times, an incident is posted.

MOBILE REPORTS

Potential freeway surveillance techniques that should not be overlooked are those classified as mobile reports. Examples of mobile report methods include the following:

- Cellular telephones.
- Freeway service patrols.
- Call boxes.

Mobile reports are most often used for incident detection. A survey of existing freeway management systems around the United States revealed that a significant portion of incidents are detected either through service patrols or cellular calls.[33] Frequently, mobile reports help detect incidents faster than many methods of automatic incident detection.[34]

Cellular Telephones

This method of quick detection is becoming more of a resource as the number of cellular telephones on the roadways continues to increase. Some traffic management systems have established free cellular call numbers for reporting incidents or requesting aid. These toll free numbers connect the caller directly to the traffic management center or to other agencies responsible for responding to incidents. Numbers that are easy to remember (such as 999 or CALLMAP) should be used to help these systems be effective. It is estimated that half of all incidents are reported via cellular telephones. Other advantages of this system include the following:[3]

- Low start-up costs.
- Two-way communication between caller and response agency.

The effectiveness of this method of incident detection depends on the number of cellular telephones on the roadway and the willingness of drivers to report incidents. To increase effectiveness, a campaign to inform drivers about the cellular call-in system and

the benefits of reporting an incident is recommended.[3]

Freeway Service Patrols

Another effective method for monitoring traffic conditions and detecting incidents is freeway service patrols. A freeway service patrol consists of a team of trained drivers who cover a particular area of freeway, monitoring traffic operations. Many different vehicles are used for freeway service patrols around the United States, including light trucks, mini-vans, and tow trucks. An example of a patrol vehicle and equipment used by the Illinois State Department of Transportation is illustrated in figure 3-21. The most noticeable benefits of service patrols are those involving incident management.[35] Typical objectives of a service patrol include the following:[36]

- Helping stranded motorists.
- Locating incidents, clearing them, and contacting the appropriate agency.
- Providing assistance at major incidents.
- Providing traffic reports to particular agencies.

Freeway service patrols have the benefit of not only being able to detect an incident, but also to assist travelers and remove vehicles upon detecting them. By performing the entire incident management process on their own, service patrols greatly reduce the time to remove an incident which, in turn, minimizes negative effects on traffic operations. The operations of service patrols are covered more extensively in **Module 8**.

Call Boxes/Emergency Telephones

Call boxes or emergency telephones may also be used to detect incidents or to locate motorists in need of help. These devices are located on the side of the freeway (see figure 3-22) and are typically spaced from 0.40 km to 0.80 km (0.25 to 0.50 mi) apart.[3] Motorists can stop and use these devices to report a problem.

Using emergency telephones, motorists can contact a dispatcher to report conditions and request help. With call boxes, motorists simply press certain buttons to request various services (such as police, fire, or ambulance). Because there are no voice transmissions with call boxes, they are relatively inexpensive to install as compared to telephones. However, emergency telephones result in fewer false alarms and greater likelihood that the appropriate services will be provided.[3]

CLOSED-CIRCUIT TELEVISION (CCTV)

Closed circuit television (CCTV) systems have been used for many years to provide visual surveillance of the freeway system.[1] Control centers typically use CCTV systems for the following purposes:

- Detection and verification of incidents.
- Monitoring traffic conditions.
- Monitoring incident clearance.

The discussion in this section will address two types of CCTV systems:

- Fixed location.
- Portable.

Figure 3-21. Illinois DOT's Advanced Technology Quick-Tow Incident Response Vehicle with Equipment.

Fixed Location

For fixed location CCTV systems, video cameras are permanently mounted either on existing structures along the freeway or on specially installed camera poles (see figure 3-23). This type of system consists of various components, including the following:[39]

- Video camera unit.
- Mounting structure (existing or installed).
- Controller cabinet housing the control equipment.
- Communication system connecting camera to control center.
- Video monitors and camera controls located in control center.

CCTV systems allow control room personnel to visually monitor sections of roadway and to react directly to the actual conditions on the roadway. Since operators can lose interest if required to constantly view CCTV monitors, and may fail to notice incidents immediately after they occur, current systems are being designed to automatically position cameras at suspected incident locations (as signaled by incident detection algorithms) and to alert the operator.[26]

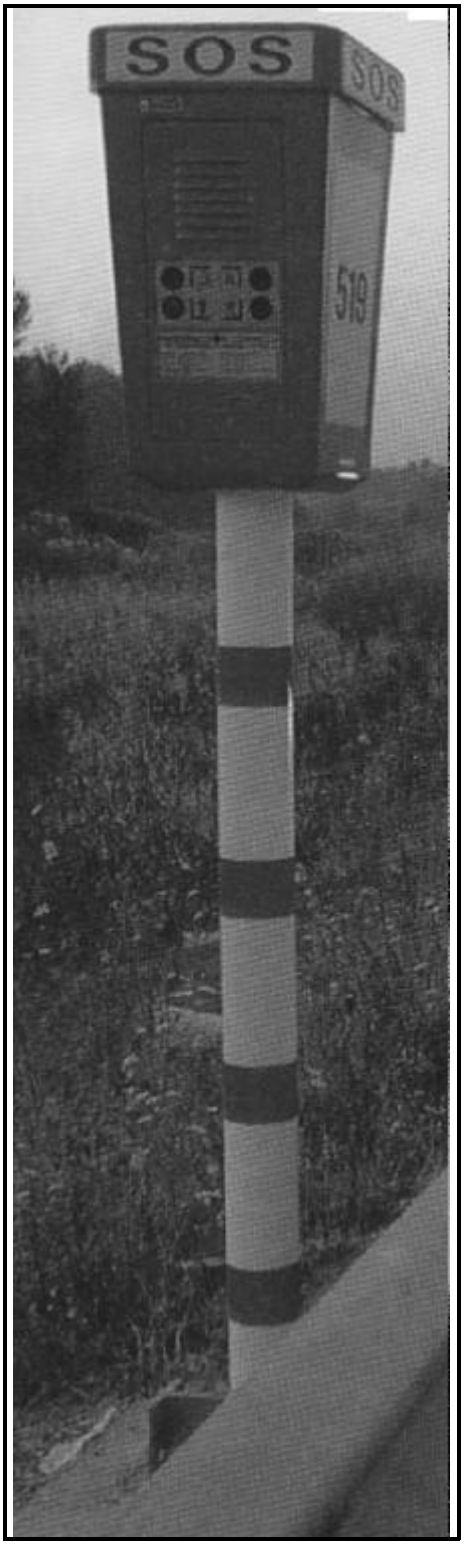


Figure 3-22. Call Box. [38]

Figure 3-23. Fixed-Location Closed-Circuit Television Camera. [16]

Video images from cameras may be transmitted to a control center using one of the following video transmission types:

- Full motion video.
- Compressed video.

Full Motion Video

Using full motion video, real-time video is transmitted to the control center. Real-time video is typically transmitted at a rate of 30 frames per second. This transmission type results in no information loss; however, it requires a wide communication bandwidth, such as that provided by coaxial cable or fiber optic cable. [3]

Compressed Video

When it is not feasible to install the communication medium required for full motion video, compressed video offers an attractive alternative. An advantage of compressed video transmission is that video data can be transmitted over conventional telephone lines and cellular channels. [3]

With compressed video techniques, transmission rates of 8 to 10 frames per second are possible. Because some information is lost between pictures frames, the resulting image appears slightly "jerky". The image, however, is adequate for monitoring freeway operations.[27]

A compressed video system typically includes the following:

- Compression and decompression computer (standard industrial PC) for each camera/monitor link.
- Appropriate software.
- Communications medium (typically a leased ISDN [Integrated Service Digital Network] line).

Standard cameras, monitors, and control hardware can be used, and therefore, can be reused if the communications medium is upgraded to allow for full motion video transmission.[26]

Portable

Portable CCTV systems can serve several purposes including the following:[39]

- Short term traffic monitoring in areas with non-recurring congestion (e.g., work zone, critical incident, etc.).
- Traffic monitoring at special traffic generators (e.g., stadiums, parades, etc.).
- Determination of optimum camera location for fixed location CCTV systems.

Portable CCTV systems are typically mounted in a light truck or van or on a trailer (see figure 3-24). Components of a portable system include the following:[40]

- Camera mounted on a pan-and-tilt unit.
- Telescopic boom.
- Television monitor.
- Video recorder.
- Camera control unit for controlling pan, tilt, and zoom functions.
- Generator for powering equipment.
- Air compressor for operating telescopic boom.

Video transmission can be accomplished using the same techniques available for fixed location CCTV systems. A system used in Dallas, Texas, consists of portable surveillance trailers with the ability to transmit video images over temporary telephone lines.[26] A system in Los Angeles, California uses a video van that is capable of transmitting full motion video back to the control center via wireless communication.[39]

SUMMARY OF TRAFFIC SURVEILLANCE APPLICATIONS

Table 3-13 summarizes the traffic surveillance technologies discussed and particular applications for which they are best suited. The recommendations shown in this table are not meant to imply that other detection methods cannot be used; rather, the table shows which technologies currently work best for specific applications. However, it should be taken into consideration that as technological improvements are made, the applications of certain technologies will be expanded.

Figure 3-24. Portable Closed-Circuit Television System. [17]

The factors used to assign the technologies to certain applications were detector performance and reliability. Cost was not taken into consideration. As technology continues to improve, detectors that are currently not cost feasible will become more competitive.

ENVIRONMENTAL DETECTORS

Environmental detectors on freeways are used to detect adverse weather conditions such as ice. These systems alert motorists to dangerous conditions and in some situations inform maintenance personnel so that evasive actions may be taken. An early warning system for detecting icy conditions on certain freeway sections can reduce the following:[4]

- Response times.
- Staffing levels.
- Anti-icing material.

Other areas where environmental detectors are used are in tunnels. In these areas, special sensors are needed to monitor the level of noxious fumes, such as carbon monoxide. The sensors are typically used to control the ventilation systems in tunnels.

Freeways

There are a number of commercially available systems for monitoring environmental conditions on freeways. These systems can be used to warn drivers

Table 3-13. Traffic Surveillance Applications.

Monitoring Application	Inductive Loop	Magneto-meter	Microwave Radar	Infrared		Ultrasonic	Acoustic	Video Image Processing	CCTV		Vehicle Probes	Mobile Reports
				Active	Passive				Fixed	Portable		
Incident Detection	X		X					X			X	X
Incident Removal									X	X		
Traffic Conditions	X	X	X	X	X	X	X	X	X		X	X
Special Event	X		X	X	X	X	X	X	X	X		
Implement Control Strategies	X	X	X	X	X	X	X	X	X		X	
Maintenance Activity	X	X	X	X	X	X	X	X	X	X	X	X
Vehicle Classification	X			X				X				
Ramp Metering	X							X				

of potentially dangerous driving conditions. In addition, the information can be used by operators in traffic management centers to do the following:[40]

- Control dynamic message signs.
- Decide whether to close roads.
- Integrate weather information into incident detection algorithms.

There are several environmental sensors available today. Categories of environmental sensors include the following:[40]

- Road condition sensors - measure surface temperature, wetness or dryness, presence of snow, and surface moisture conductivity.
- Visibility sensors - detect presence of fog, smog, heavy rain, snow, or sandstorms.
- Thermal mapping - detects presence of ice.

Figure 3-25 shows a roadside environmental monitoring system that is capable of measuring visibility and precipitation.

Several manufacturers also produce complete weather stations. An example of such a system is the SCAN (Surface Condition Analyzer) developed in the United States. SCAN is able to monitor pavement and environmental conditions and provide the following information:[42]

- Pavement temperature.
- Presence of water, ice, snow, or chemicals on pavement.
- Precipitation.

Figure 3-25. Roadside Environmental Monitoring System. [41]

- Air temperature, humidity, and dew point.
- Wind speed and direction.

The system consists of sensors the size of hockey pucks that are embedded in the pavement. Capacitors within the sensors measure electrical charges from water, snow, ice, or chemicals. Thermostats are used to measure surface temperatures. SCAN systems are currently used in 48 States and in 8 countries.[41]

Tunnels

Monitoring systems in tunnels are used to limit the build-up of noxious fumes from vehicle exhausts. Output from the monitors is used to perform the following functions:

- Control tunnel ventilation.
- Warn drivers of poor visibility.
- Set safe speed limits within the tunnel.

Most systems monitor the level of noxious gases and visibility by using infrared absorption techniques. A gas monitoring system typically consists of the following:

- Infrared transmitter and receiver.
- Electronic signal processing unit.

Outputs from the monitors are sent to the system controlling tunnel ventilation. Tunnel ventilation becomes more cost effective when operated in conjunction with a monitoring system. Cost savings arise from reduced frequency of operation and maintenance of the ventilation system.[43] There are currently many systems available for monitoring air quality within a tunnel; however, it is important that a monitoring system be reliable, accurate, and able to operate with minimal maintenance.[44]

3.4 LESSONS LEARNED

OPERATIONS AND MAINTENANCE

Effective operations and maintenance of a traffic management system are required to ensure that the system meets certain objectives, such as the following:

- Continuously accomplishes the goals and objectives for which it was designed.
- Responds to changing technologies and transportation system demands.

The ability of agencies to effectively operate and maintain a freeway management system is vital to the success of the system.[6] This section discusses issues associated with the operations and maintenance of surveillance systems.

Operations

Operations includes those tasks involved in the day-to-day function of a system. The primary responsibility of personnel in charge of operating a surveillance system is to monitor freeway conditions. Freeway monitoring may serve several purposes, including the following:

- Monitoring traffic operations.
- Detecting incidents.
- Monitoring incidence clearance.
- Monitoring environmental conditions.
- Supporting the implementation of control strategies.

Operators may monitor freeway conditions by using any of the following devices:

- Displays or printouts showing conditions either graphically or in a tabular format.
- Closed-circuit television systems.
- Mobile reports.

Additional tasks with which an operator may be involved include the following:[4, 6]

- Controlling and managing equipment and ensuring continuity of operation.

- Monitoring system performance criteria.
- Updating system databases.
- Notifying maintenance personnel of system malfunctions.
- Communicating with transportation and emergency response agencies.
- Documenting daily operational functions and events such as incident occurrences, equipment failures and repairs, and changes in software and hardware components.

Maintenance

Proper design and installation techniques are important to ensure the success of a system. Maintenance, however, is also a critical task that must be considered if the system is going to operate as intended. Maintenance activities can be grouped into three categories:[4]

- Remedial.
- Preventive.
- Modification.

Remedial maintenance is necessitated by malfunctions and equipment failures, and usually demands emergency repair to restore operations. Preventive maintenance includes work done at scheduled intervals to minimize equipment failure. Modification involves replacing equipment and usually results from equipment design flaws or required upgrades to improve system performance. All three maintenance categories should be considered in determining budget requirements.

For a maintenance program to be successful, the following factors should be considered:[4]

- Address maintenance needs (e.g., maintenance requirements, maintenance costs, personnel requirements) during the planning process.
- Ensure that required budget and staffing needs can be met.
- Perform cost tradeoff analyses during the design phase to identify techniques for reducing maintenance requirements.
- Consider that complex surveillance technologies require higher personnel skill levels to maintain.
- Ensure that maintenance personnel receive adequate training.
- Keep accurate records of maintenance activities to aid in predicting future maintenance needs and in analyzing costs.

Operations and Maintenance Costs

The success of a system begins with the type and quality of the equipment being utilized. Budgetary and funding problems that traffic agencies typically face may result in the selection of equipment mainly on the basis of initial start-up cost, rather than on lifetime cost. Consequently, less expensive equipment is purchased during the installation process. However, this is not always the most cost-effective solution. [11] Therefore, operations and maintenance costs should be considered together with start-up costs during the equipment selection process.

A Texas Department of Transportation (TxDOT) research project established guidelines for estimating operations and maintenance costs.[6] The cost estimates were based on the following sources:

- TxDOT metropolitan districts with substantial ITS deployment.
- ITE report, *Operation and Maintenance of Electronic Traffic Control Systems*.
- FHWA report, *Cost Estimates and Assumptions for Core Infrastructure*.
- Texas municipalities and transit agencies.
- Technical journals.
- Equipment suppliers.

Table 3-14 shows the operations and maintenance cost estimates for a surveillance system. This table lists the estimated unit operations cost, estimated unit maintenance cost, combined operations and maintenance cost, and the assumptions related to factors included in each of the costs. Using quantities measured in terms of the base units shown in the table, total operations and maintenance costs can be determined. Note that the estimated maintenance costs include maintenance personnel costs, while the estimated operations costs do not included personnel costs.

Most of the operations and maintenance costs in table 3-14 are given as ranges in order to allow for adjustments for the following factors:

- Age and quality of equipment.
- Personnel skill levels.
- System designs.

In many situations, outside financing sources are more readily available for the development and construction of new systems than for operating and maintaining the systems once they are built. Operating and maintenance costs are typically funded through State and local maintenance budgets. These funds are often insufficient because of competition with other maintenance functions. Therefore, it is important that continued funding for operations and maintenance be addressed in the planning stage.

PRIVACY CONCERNS

Most surveillance technologies have little impact on privacy, because there is no need for vehicle identification. Privacy, however, is a critical concern when using technologies that identify vehicles, such as AVI, AVL, and cellular telephone tracking. Privacy is also a concern with the use of closed-circuit television.

A survey of drivers in the United States showed that in 1978, 67 percent of people were concerned about privacy. Results from 1994 showed that the percentage had climbed to 84 percent, revealing that privacy is a major concern for drivers.[45]

Although some people may not mind if their vehicles alone are identified, many fear that surveillance from different sources (e.g., credit card purchases, phone calls, travel habits, etc.) could be combined to build a detailed personal profile of their lifestyles. The ability to compile information about an individual's travel patterns, toll payments, and other activities creates the potential for a database that has not previously existed. Due to advances in computer technology, this kind of information is highly transportable, and opportunities exist for inappropriate access to and use of this personal information.[46]

It must be taken into account that privacy is one of the goals against which any system's effectiveness may be measured. For example, one of the first AVI tolling systems

Table 3-14. Estimated Operating and Maintenance Costs for Surveillance Systems.[6]

Description	Base Unit	Estimated Annual Operations Cost/Unit	Estimated Annual Maintenance Cost/Unit	Estimated Annual Combined O&M Costs/Unit	Cost Assumptions
CCTV	each	$0	$500 - $1300	$500 - $1300	Costs include routine maintenance for CCTV cameras, as well as the camera controls, housing, and/or support pole.
CCTV Cameras w/ Video Image Processing Capability	each	$0	$700 - $1800	$700 - $1800	Costs include routine maintenance for CCTV cameras, as well as camera controls, housing, and/or support pole, and VIP calibrations.
Imbedded Detectors:					
Inductive Loop	per station	$0	$200 - $300	$200 - $300	Costs include contract maintenance/replacement of loops. Costs assume four lanes per station, with two loops per lane. Costs also assume loop failure rates of 4% to 6% per year.
Non-Intrusive Detectors:					
Radar, Ultrasonic, Acoustic, Infrared	per station	$0	$200 - $300	$200 - $300	Costs include routine maintenance of detectors. Costs assume four lanes per station, with one detector per lane.
Video Image Processing	per station	$0	$500	$500	Costs include routine maintenance and calibration of detectors. Costs assume one video detector per station.

Table 3-14. Estimated Operating and Maintenance Costs for Surveillance Systems (cont.).[6]

Description	Base Unit	Estimated Annual Operations Cost/Unit	Estimated Annual Maintenance Cost/Unit	Estimated Annual Combined O&M Costs/Unit	Cost Assumptions
AVI:					
Transponders	each	$0	$10 - $15	$10 - $15	Costs include routine maintenance and replacement of transponders.
Readers	per station	$0	$500	$500	Costs include routine maintenance.
AVL:					
Leased Transceivers/ Antennae	per vehicle	$750 - $1200	$0	$750 - $1200	Operations costs include the lease cost of the transceivers and antennae, and the associated communications cellular air time charges; maintenance is provided by the transceiver/antennae provider.
Owned Transceivers/ Antennae	per vehicle	$250 - $500	$100 - $200	$350 - $700	Operations costs include communications cellular air time charges; maintenance costs include routine maintenance and replacement of transceivers and antennae.

to be implemented was in Hong Kong in 1983. Each month, drivers using the electronic tolling system were sent statements that contained itemized accounts of where they had driven and when. The drivers didn't like the idea of a government agency tracking their daily highway travel; therefore, the system was removed two years after it was introduced, despite the fact that the technology worked.[46] For reasons like these, it is very important to design any surveillance with privacy in mind.

Fortunately, it is relatively easy to design most surveillance systems with privacy in mind.[45] For example, AVI systems can be designed to assign vehicles a random identification number for monitoring purposes, in which case the identification of the driver is never made. This same process can be used with AVL and cellular-based traffic surveillance systems.[32]

In general, consumers may be willing to give up a certain degree of privacy if they feel that the benefits will outweigh the risks (as in the case of credit card use). It is therefore important that the public be informed about the benefits of using certain surveillance technologies. The key to overcoming privacy concerns is to inform the public about what information is being collected and how the data will be used.[47]

SPACING AND PLACEMENT ISSUES

Selecting the location of detectors is an important step in the implementation of a surveillance system. Spacing and placement affect both the cost and performance of the system. The detectors should be placed close enough together to provide adequate surveillance of the system, but as the spacing decreases, the cost of the system increases. In addition, it is desirable to place the detectors in a location that will result in minimum delay to traffic during installation and maintenance procedures. The following discussion presents recommendations for the spacing and placement of mainline traffic detectors and CCTV cameras.

Mainline Traffic Detectors

Spacing

For monitoring traffic conditions and incident detection, detectors should be placed along the freeway and on ramps to detect changes in traffic conditions. Guidelines for detector spacings are as follows:[3,4]

- Typical mainline detector spacings range between 0.40 and 0.80 km (0.25 and 0.5 mi).
- Spacings over 0.8 km (0.5 mi) are inadequate for incident detection purposes.
- Spacings below 0.4 km (0.25 mi) generally produce little or no increase in effectiveness.
- Tunnels or covered roadway sections are exceptions, and may require spacings of less than 0.4 km (0.25 mi).
- Minimally, at least one traffic detector station should be provided between interchanges or freeway ramps, with maximum spacings of 1.6 to 2.4 km (1.0 to 1.5 mi).
- Detectors are also typically placed before and after freeway bottlenecks (e.g., entrance ramps, lane closures, etc.) and metered entrance ramps.

Placement

As discussed earlier, most existing traffic detectors can be classified as either

embedded or non-intrusive. An embedded detector system consists of sensors in or below the surface of the roadway. Non-intrusive detectors are mounted on a structure above the surface of the pavement.

The following issues associated with installing and maintaining embedded detectors:

- Installation typically requires drilling into or cutting the pavement.
- Embedded detectors usually have higher installation costs than non-intrusive detectors.
- Installation may require lane closure, which results in delay to vehicles.
- When feasible, embedded detectors should be installed during the construction or rehabilitation of a freeway to avoid delays to traffic.
- Because modifications to the pavement structure are required for the installation, deterioration of the road surface is escalated; therefore, there are further ongoing maintenance costs associated with replacing damaged detectors or repairing degraded road surfaces.

Non-Intrusive Detectors

Issues associated with installing and maintaining non-intrusive detectors include the following:

- Can typically be installed and maintained without disrupting traffic.
- Installation costs are significantly reduced when an existing mounting structure can be used.
- Ongoing maintenance costs may be minimal.

CCTV Cameras

The primary objective of a CCTV system is to provide visual monitoring of freeway conditions and to verify incidents. The California Department of Transportation, District 7, has developed guidelines for CCTV camera site selection.[39,48,49] The following is a summary of these recommendations.

Site Selection Criteria

Most cameras with a standard 10-to-1 zoom have a viewing range of 0.8 km (0.5 mi) in all directions. Full freeway coverage would therefore require a maximum camera spacing of 1.6 km (1 mi). Full coverage of all freeways, however, is frequently cost prohibitive. Therefore, proper camera site selection becomes a critical factor. Factors that should be considered in the selection of a camera site include:

- Surveillance of congested freeway sections to provide maximum visibility and coverage of weaving and merging areas near ramps and interchanges.
- Ability to provide a clear view of areas with high accident frequencies (based on accident data).
- Ability to provide suitable access for maintenance personnel based on the geometric layout of the area.
- Minimal obstruction to view caused by landscaping, billboards, freeway signs, or topography.
- Ability to verify messages of variable message signs.

- Ability to effectively address power, communication, and underground utility considerations.

- Ability to provide suitable access to the proposed communication system.

Candidate sites for CCTV systems may be prioritized on the basis of accident frequency (high, medium, and low). Threshold values to classify accident frequencies should be based on local conditions. Areas with high accident frequencies will require video surveillance to expedite incident detection and clearance. Areas with medium and low accident frequencies may also require video surveillance. Site selections for these locations will be based on local criteria, such as the importance of a particular freeway section. In addition, video surveillance may be used to monitor ramps and freeway sections near heavy traffic generators (e.g., stadiums) during special events.

Preliminary Site Selection

Preliminary site selection should be based on the criteria discussed above. These sites should be noted on a scaled base map, along with the following characteristics of the freeway:

- Accident frequencies plotted over 0.4 km (0.25 mi) intervals.

- Strategic roadway infrastructure.

- Locations of variable message signs.

- Locations of underground high-risk utility lines.

Cameras should be located in the vicinity of high-accident frequency areas and near variable message signs and other strategic areas. Areas with high-risk utility lines should be avoided if possible. For cameras with standard 10-to-1 zoom and a viewing range of 0.8 km (0.5 mi), video surveillance can be provided for a 1.6 km (1 mi) freeway section (assuming no blockage due to horizontal or vertical curves).

The next step in the preliminary site selection process is to perform a field check. The objectives of the field check are to verify the following:

- Veiling range at “ground eye level.”

- Geometric layout.

- Maintenance access.

- Power and utility concerns.

While in the field, each preliminary site selected should be marked so that the site is easily identifiable for the final site selection process.

Final Site Selection

The purpose of the final task is to perform a video survey to verify camera viewing ranges and select correct mounting heights. For high-speed freeway sections, a mobile CCTV camera system is recommended.

Following are considerations for selecting mounting heights:

- Mounting heights on poles generally vary from 6.0 to 13.7 m (20 to 45 ft).

- Heights below 6.0 m (20 ft) may be susceptible to vandalism.

- Pole heights above 13.7 m (45 ft) may result in maintenance problems and safety hazards.

- Lower heights (between 6.0 and 9.1 m [20 and 30 ft]) are preferred when it is necessary to look beneath overpasses.

- Cameras at heights below 10.7 m (35 ft) can be accessed by a small bucket truck.

- Cameras at heights between 10.7 and 13.7 m (35 and 45 ft) would require a larger hoist truck.

When possible, cameras should be mounted directly on existing structures. This will result in a significant reduction in the initial installation costs.

3.5 EXAMPLES

MIAMI AUTOMATIC VEHICLE LOCATION (AVL) [32]

The City of Miami, Florida, needed an accurate, inexpensive method of collecting average travel speed data in order to calculate roadway level of service as required by the 1985 Florida Growth Management Act. In 1988, the City of Miami staff determined that automatic vehicle location (AVL) technology could be used to measure speeds on specific corridors. (AVL is a means of continuously monitoring the location of vehicles in a road network. Vehicles are equipped with a transponder device that transmits a radio-frequency (RF) signal to a central location at regular intervals.) In 1993, the City of Miami contracted with the Center for Urban Transportation Research (CUTR) to set up a field operational test of AVL to measure vehicle operating speeds on 17 transportation corridors in the city. CUTR recruited AirTouch Teletrac to serve as the vendor for the AVL system. AirTouch Teletrac agreed to provide 25 transponders, one workstation, one copy of its FleetDirector™ software, and training on how to use the software to the City of Miami for 112 days at a nominal cost.

The City of Miami was responsible for recruiting 25 volunteer drivers willing to have their vehicles equipped with transponders. Many of the drivers recruited commuted from the periphery of the city to downtown, and this provided coverage for 5 of the 17 corridors in the peak direction during the peak period. Installation of the transponders for each vehicle took about one hour. A single AVL unit consisted of a control unit (about the size of a video tape cassette) and a “pancake” antenna.

AirTouch Teletrac provided training on the use of the software to CUTR researchers and the City of Miami Planning staff and then provided a workstation, which was a Compaq 386 computer with an internal high-speed modem. The city provided a dedicated phone line for communication between the workstation and AirTouch Teletrac’s operations center in Fort Lauderdale. (AirTouch Teletrac maintains a network of 27 receiving antennas throughout Dade, Broward, and Palm Beach counties in South Florida. RF signals from the receiving antennas are transmitted to AirTouch Teletrac’s operations center in Fort Lauderdale.)

The software was configured to poll each of the 25 vehicles’ positions every 30 seconds when the vehicle ignition was on and every 5 minutes when the vehicle ignition was off. Each time a vehicle’s position was recorded, the following information was obtained: vehicle number, speed, time, date, and location. Data were recorded for the following time periods:

- Weekday mornings, 5 a.m. to 10 a.m.

- Weekday afternoons, 3 p.m. to 8 p.m.

- Saturdays, 10 a.m. to 2 p.m.

Data were gathered for 112 days, and data for over 4400 vehicle trips were obtained. CUTR staff wrote two software programs to analyze vehicle-location data and report average speed. In order to determine the accuracy of the calculations for average trip speeds, CUTR researchers also set up a process to compare observed values with calculated values for the average speed over a vehicle trip. Method 1 yielded a ± 8.3 percent accuracy, and Method 2 yielded a ± 7.1 percent accuracy. Since Method 2 was more accurate, it was used in compiling the output.

CUTR has demonstrated that the AirTouch Teletrac automatic vehicle location system can be used to measure average vehicle operating speeds on Miami's 17 transportation corridors in the peak direction during the peak period, given sufficient electronic storage capacity, appropriate data analysis software, and volunteer drivers. This system could be used anywhere in the six metropolitan areas where AirTouch Teletrac is available. However, in order to conduct this experiment outside of those six areas, an AVL vendor would have to be enlisted.

TRANSCOM ELECTRONIC TOLL AND TRAFFIC MANAGEMENT (ETTM)[50,51]

ETTM, or electronic toll and traffic management, has seen rapid growth. Currently, more than 20 agencies in 8 countries have installed a system and are collecting tolls automatically. Another 23 agencies in 7 countries are in the process of testing, studying, or implementing automatic toll collection. ETTM systems have the potential to be used for other ITS applications that could include congestion pricing, advanced payment systems, traveler information, and traffic management. ETTM requires the installation of readers with the capability of identifying tagged vehicles at periodic intervals along the roadway.

TRANSCOM is an umbrella organization serving the New York/New Jersey metropolitan area in coordinating and disseminating transportation information. In 1991, TRANSCOM retained a team led by PB Farradyne, Inc., to establish the feasibility of using ETTM equipment for traffic surveillance and incident detection applications. The project was called TRANSMIT (TRANSCOM's System for Managing Incidents and Traffic) and was divided into two phases. The first phase concentrated on evaluating the use of ETTM technology for traffic surveillance and on developing a preliminary design of a traffic surveillance system for an ETTM incident detection system. The second phase included the final design and construction of the initial stage of the system as an 28.8-km (18-mi) operational field test along sections of the New York State Thruway and the Garden State Parkway. The primary objective of the TRANSMIT project was to process available traffic flow data for early detection of incidents and abnormal traffic congestion.

An analysis was conducted to determine the reader antenna spacing and toll tag penetration necessary to detect an incident within five minutes with a false alarm rate of two percent or less. The analysis used the following roadway classes: arterial, two-lane highway, and three- to four-lane highway. The results indicated that the arterial roadway class would result in the longest detection time. The analysis concluded that the five-minute detection time could be achieved if the following percentages of equipped vehicles were provided:

- Two-lane highways: 0.9 percent.

- Three- and four-lane highways: 2.1 percent.

- Arterials: 7.6 percent.

An 28.8-km (18-mi) corridor along the New York State Thruway in Westchester and Rockland Counties and the Garden State Parkway in Bergen County were defined for an operational field test. Twenty-two ETTM sites were constructed, and existing overhead sign structures and overpasses were used for the placement of reader antennas. New structures were built where existing structures were not available.

The incident detection algorithm developed by PB Farradyne, Inc., determines the probability of an incident when tagged vehicles detected at an upstream reader are not detected at the downstream reader within the expected arrival time. The later a vehicle is, the higher the probability of an incident.

At the time this project was designed, the technology for Electronic Toll Collection (ETC) in the Greater New York/New Jersey Metropolitan Area was undetermined, so it was decided that the ETTM antennas and reader cabinet equipment would be leased, not purchased. Therefore, if the traffic surveillance concept is found to be successful and is expanded throughout the region, the equipment can be made compatible with the regional ETC equipment from which the tag population originates. The construction cost estimate for the operational field test was $1,195,600, which included equipment leasing costs for a 24-month period.

HOUSTON AUTOMATIC VEHICLE IDENTIFICATION (AVI) [52]

The Texas Department of Transportation awarded the first of three contracts to instrument the metropolitan freeway system in Houston with an automatic vehicle identification (AVI) system to monitor traffic conditions. The AVI system was designed to provide travel time information on about 120 miles of freeways and 100 miles of reversible high-occupancy vehicle (HOV) lanes.

The initial planning began in 1991. The North Houston Corridor, with three parallel radial freeways, was selected as the study site. This corridor had several favorable characteristics: one toll facility with available capacity; one freeway with heavy congestion because of inadequate capacity and reconstruction activities; and one freeway with moderate congestion and an HOV lane. The corridor also served the Houston Intercontinental Airport, the central business district, and other major activity centers. A demonstration project was designed using cellular telephones for the collection of traffic condition information. This project recruited 200 volunteers who commuted to and from work using one of the three freeway routes. The demonstration project was conducted for about 18 months, during which time a more permanent and automated system for traffic monitoring was being developed using AVI technology. The demonstration project was a success in that valid, useful real-time travel time information was collected and made available to the public in a variety of ways.

CONNECTICUT MICROWAVE RADAR [13]

The Connecticut Department of Transportation (ConnDOT) chose radar detection for the recent design of their freeway surveillance system. Connecticut has two separate systems operating within the State. The first system is a federal demonstration project covering approximately 19 km (12 mi) of Interstates

I-84 and I-91 in the Hartford area, and it is the first surveillance system in the country to utilize radar as an incident detection technology. Because of the early success in Hartford, radar detection is also being used in the design of the I-95 Freeway Management System. The I-95 system is installed along the southern coast of Connecticut, covering 90 km (56 mi) from the New York State border east to Branford. Both systems rely on speed data being sent back from the radar detectors to monitor traffic flow condition and to detect incidents.

Three different types of radar detectors are being used on the projects: wide beam detectors, narrow beam detectors, and long range detectors. The Radar detectors for the Hartford Area ATMS (44 individual detectors) have been installed for 2 years. During this time, only one detector has failed, and the others have survived harsh winters with record snowfalls and significantly below-average cold temperatures. In addition to their reliable operation, the radar detectors have proven to be easily installed and setup.

The detectors have been installed for only a few years, but the data looks promising. Testing has been performed for the various detectors and installation configurations to determine the detector accuracy. The difference between vehicle and detected speed ranges from 4 km/h (2.5 mi/h) for the narrow beam detector, to 5.6 ki/h (3.5 mi/h) for the wide beam detector in a side-fire configuration, and up to 10 mi/h for the long range detector. The errors associated with the long range detector can be attributed to the operating principle and placement of the detector. The detector operation is based on the Doppler principle: it is most accurate when it is aimed directly at the traffic flow.

Testing and day-to-day operations have shown that the radar detectors are highly accurate during varying traffic and weather conditions. The degree of reliability and accuracy demonstrated by the detectors has allowed ConnDOT engineers to be confident in the data displayed. With the use of the video cameras that are also part of the system, ConnDOT engineers have been able to detect and verify incidents that would otherwise have been unknown to them, and to initiate a timely response. Lessons learned from the Connecticut system have been applied in the development of systems in Maryland, New Jersey, Massachusetts, and Georgia.

3.6 REFERENCES

1. Sumner, R. et. al. *Freeway Management Handbook, Volume 2: Planning and Design*. Report No. FHWA-SA-88-027. FHWA, U.S. Department of Transportation, Washington, DC, May 1983.

2. *Highway Capacity Manual.* Special Report No. 209. TRB, National Research Council, Washington, DC, 1994.

3. Texas Department of Transportation, Highway Design Division. *Operations and Procedures Manual*, Part IV, Interim Metric Version.

4. Gordon, R.L. et. al. *Traffic Control Systems Handbook*. Report No. FHWA-SA-95-032. FHWA, U.S. Department of Transportation, Washington, DC, February 1996.

5. *ITS Architecture, Implementation Strategy*. Joint Architecture Team: Loral Federal Systems & Rockwell International, Prepared for the Federal Highway Administration, revised January 1997.

6. Daniels, G., Starr, T., and Stockton, W.R. Guidelines for Funding Operations and Maintenance of ITS/ATMS. Report No. FHWA/TX-97/1494-1F. Texas Transportation Institute, College Station, TX, 1996.

7. Sumner, R.L., et. al. *Freeway Management Handbook, Volume 3: Operations and Maintenance.* Report No. FHWA-SA-88-028. FHWA, U.S. Department of Transportation, Washington, DC, May 1983.

8. Vacca, J.R. Creating Intelligent Traffic Control Centres. *Traffic Technology International.* UK and International Press, Autumn 1995.

9. Zhang, H., and Ritchie, S.G. Real-Time Decision-Support System for Freeway Management and Control. *Journal of Computing in Civil Engineering.* American Society of Civil Engineers. Washington, DC, January 1994, pp 35-51.

10. Prosser, N.A., and Ritchie, S.G. Real-Time Knowledge-Based Integration of Freeway Surveillance Data. *Transportation Research Record No 1320.* TRB, National Research Council, Washington, DC, 1991, pp 40-46.

11. *Traffic Detector Handbook*. Second Edition. Institute of Transportation Engineers. Washington, DC, 1991.

12. Peek Traffic, Inc., Automatic Data Recorders (Produce Brochure).

13. Mauritz, M.R., and Decker, H. Radar: Solving the Incident Detection Loophole. *Traffic Technology International*, UK and International Press, Autumn 1995.

14. Steinback, B.G. Loop Substitution for Intersection Control. *Traffic Technology International '96*. UK and International Press, 1996.

15. Bullimore, E.D., and Hutchinson, P.M. Life Without Loops. *Traffic Technology International '96*. UK and International Press, 1996.

16. Michalopoulos, P.G., Jacobson, R.D., Anderson, C.A., and Barbaresso, J.C. Field Deployment of Machine Vision in the Oakland County ATMS/ATIS Project. *Proceedings*, IVHS America 1994 Annual Meeting, Atlanta, GE, April 1994, pp 335-342.

17. Krogmeier, J.V., and Sharmo, D.E. Overhead and Wireless: An ATMS for the Borman Expressway. *Traffic Technology International*. UK and International Press. April/May, 1996.

18. Klein, L.A., MacCaalden, M.S., and Mills, M. Detection Technology in IVHS. *Proceedings*, IVHS America 1993 Annual Meeting, Washington, DC, April 1993, pp 644-653.

19. Manor, D. Multiple Zone Radar Detection by RTMS. *Traffic Technology International '96*. UK and International Press, 1996.

20. Roussel, J.C., Petrucci, J., Lion, D., and Jaouen, R. BEATRICS Radar System for Automatic Incident Detection. *Traffic Technology International '96*. UK and International Press, 1996.

21. EIS Electronic Integrated Systems, Inc. Radar Technology (Produce Brochure).

22. Polk, A.E., Kranig, J.M., and Minge, E.D. Field Test of Non-Intrusive Traffic Detective Technologies. *Proceedings*, 1996 Annual Meeting of ITS America, Houston, TX, April 1996, pp 456-467.

23. Myers, T. Laser Sensors for Traffic Monitoring and Control. *Traffic Technology International '96*. UK and International Press, 1996.

24. Duckworth, G.L., et. al. A Comparative Study of Traffic Monitoring Sensors. *Proceedings*, IVHS America 1994 Annual Meeting, Atlanta, GE, April 1994, pp. 283-293.

25. Michalopoulos, P.G. Incident Detection Through Video Image Processing. *Applications of Advanced Technologies in Transportation Engineering*. American Society of Civil Engineers. New York, NY, 1991.

26. Carvell, J.D., Seymour, E.J., Walters, C.H., and Starr, T.R. *Dallas Area-Wide Intelligent Transportation Plan*. TxDOT Report 591F, Texas Transportation Institute, Dallas, TX, June 1996.

27. Odetics (Advertisement). *Traffic Technology International*. UK and International Press, Winter 1995.

28. Sowell, W.H., and LaBatt, J.S. Video Detection Takes a New Track. *Traffic Technology International.* UK and International Press, Winter 1995.

29. Levine, S. Houston Regional Area System. Presented at the 73rd Annual Meeting of the Transportation Research Board, National Research Council, Washington, DC, 1994.

30. Fenno, D. Electronic Toll Collection System. *Graduate Student Papers on Advanced Surface Transportation Systems*. Texas A&M University, College Station, TX. August 1993.

31. Polk, A.E., and Pietrzyk, M.C. The Miami Method: Using Automatic Vehicle Location (AVL) for Measurement of Roadway Level-of-Service. *Proceedings*, 1995 Annual Meeting of ITS America, Washington, DC, April 1995, pp 343-402.

32. Sumner, R., Smith, R., Kennedy, J., and Robinson, J. Cellular Based Traffic Surveillance: The Washington, DC Area Operational Test. *Proceedings*, IVHS America 1994 Annual Meeting, Atlanta, Georgia, April 1994, pp 794-798.

33. Balke, K., et. al. Freeway Management Systems: A State-of-the-Practice Review. *Task B of Project, Develop a Freeway Management Handbook*. FHWA, U.S. Department of Transportation, Washington, DC, January 1996.

34. Fink, K.L. An Analysis of Incident Management Programs in North America. *Graduate Student Papers on Advanced Surface Transportation Systems*. Texas A&M University, College Station, TX, August 1993.

35. Nowlin, R.L. Institutional Concerns Regarding the Implementation of State Managed Service Patrols." *Graduate Student Papers on Advanced Surface Transportation Systems*. Texas A&M University, College Station, TX, August 1994.

36. Reiss, R., and Dunn, W. *Freeway Incident Management Handbook*. FHWA-SA-91-056. FHWA, U.S Department of Transportation, Washington, DC, July 1991.

37. McDermott, J.M., McLean, C.H., and Smith, A.T. Three Decades of Progress: Freeway Traffic Management in Illinois. *ITE Journal*, Vo 62(3). Institute of Transportation Engineers, Washington, DC, March 1992, pp 27-31.

38. Aselsan (Advertisement). *Traffic Technology International '96*. UK and International Press. 1996.

39. Lee, L.K., Zachariah, M., and Everett, P. CCTV Camera Site Selection: A Field Experience. *Intelligent Transportation: Serving the User Through Deployment.* Proceedings of the 1995 Annual Meeting of ITS America. ITS America, Washington, DC, March 1995.

40. Purdom, N. Whatever the Weather. *Traffic Technology International*. UK and International Press, Autumn 1995.

41. Scientific Technology, Inc. (Advertisement). *Technology International*. UK and International Press, Winter 1995.

42. Hartwell, G. Scanning the Grapevine. *AASHTO Quarterly Magazine*. American Association of State Highways and Transportation Officials, Washington, DC, October 1991.

43. Luchian, S.F., and Krechmer, D. Traffic Control on the Boston Central. *Traffic Technology International*. UK and International Press, April/May 1996.

44. Dingle, M. Limiting the Build-up of Noxious Fumes. *Tunnels and Tunneling*. Morgan-Grampian (Construction Press) LTD, London, United Kingdom, October 1992.

45. Wright, T. Eyes on the Road: Privacy and ITS. *Traffic Technology International*. UK and International Press. Autumn 1995.

46. Dingle, J. IVHS and the Law of Privacy. *Pacific Rim Transtech Conference Proceedings, Volume 1*. American Society of Civil Engineers, New York, NY, 1993, pp 318-325.

47. Costantino, J. Overcoming Institutional Barriers to IVHS. *Transportation Quarterly*. Eno Foundation for Transportation, Lansdowne, VA, Winter 1994.

48. *Guidelines for Camera Placement*. IVHS/TOS Design Guidelines No. 94-01, California Department of Transportation, District 7, Los Angeles, CA, October 1994 (Revised).

49. *Guidelines for CCTV Site Surveying*. California Department of Transportation, District 7, Los Angeles, CA, February 1994.

50. Yermack, L., Gallagher, M., and Marshall, K.R. ETTM - An Early ITS Winner. *ITE Journal*, Vol. 65(12), Institute of Transportation Engineers, Washington, DC, December 1994, pp 42-48.

51. Marshall, K.R.l and Batz, T. The TRANSCOM TRANSMIT Project: An Alternative Approach for Traffic Surveillance and Incident Detection. *Moving Toward Deployment*, Proceedings of the IVHS America 1994 Annual Meeting, ITS America, Washington, DC, 1994, pp 555-563.

52. Levine, S.Z., and McCasland, W.R. Monitoring Freeway Traffic Conditions with Automatic Vehicle Identification Systems, *ITE Journal*, Vol. 64(3). Institute of Transportation Engineers, Washington, DC, March 1994, pp 23-28.

MODULE 4. LANE USE CONTROL

TABLE OF CONTENTS

4.1 INTRODUCTION 4-3
BACKGROUND 4-3
TREATMENTS 4-3
MODULE OBJECTIVES 4-4
MODULE SCOPE 4-4

4.2 DESIGN PROCESS 4-4
PROBLEM IDENTIFICATION 4-4
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 4-5
Potential Partners 4-5
Consensus Building 4-6
ESTABLISH GOALS AND OBJECTIVES FOR LANE USE CONTROL 4-6
ESTABLISH PERFORMANCE CRITERIA 4-6
DEFINE FUNCTIONAL REQUIREMENTS 4-7
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 4-7
IDENTIFY AND SCREEN TECHNOLOGY 4-8
PLAN DEVELOPMENT 4-9
FUNDING SOURCE IDENTIFICATION 4-9
IMPLEMENTATION 4-9
EVALUATION 4-10

4.3 TECHNIQUES AND TECHNOLOGIES 4-10
STATIC SIGNING 4-10
Truck Restrictions 4-10
HOV Restrictions 4-11
INFORMATION DISSEMINATION SYSTEMS 4-11
USE OF NARROW LANES AND SHOULDERS 4-15
MAINLINE METERING 4-15
FREEWAY-TO-FREEWAY RAMP METERING 4-16
CHANNELIZING DEVICES FOR WORK ZONES 4-17
TOLL FACILITIES/CONGESTION PRICING 4-17
AUTOMATED HIGHWAY SYSTEMS 4-19
Concept 4-20
Vehicle Characteristics 4-21
Roadway Infrastructure 4-21
Command and Control 4-21
Entry and Exit Infrastructure 4-21
Goals 4-22

4.4 LESSONS LEARNED 4-22
PLANNING 4-22
Freeway-to-Freeway Ramp Metering 4-22
Toll Facilities 4-23
Congestion Pricing 4-23
Channelizing Devices for Work Zones 4-24
DESIGN/CONSTRUCTION 4-26
LCS and DMS Placement 4-26
LCS Visibility 4-26
LCS Spacing and Mounting Locations 4-26
DMS Placement 4-26
Use of Narrow Lane and Shoulders 4-27

4.5 EXAMPLES 4-27
SAN ANTONIO (TRANSGUIDE) 4-27
Description of System 4-27
Effects of Information 4-28
SR 91, ORANGE COUNTY, CALIFORNIA CONGESTION PRICING 4-29

4.6 REFERENCES AND SUGGESTED READINGS 4-31
REFERENCES 4-31
SUGGESTED READINGS 4-34

MODULE 4. LANE USE CONTROL

Figure 4-1. Proper Lane Use Control Can Facilitate Safe and Efficient Operations.

4.1 INTRODUCTION

BACKGROUND

Another approach to improving the safety and efficiency of freeway operations is the control of traffic on the freeway mainline. This involves the regulation, warning, and guidance of traffic on the freeway mainline so that some or all of the following objectives are addressed:[1]

- To achieve a more uniform and more stable traffic flow as the freeway demand approaches capacity, thereby improving the utilization of the facility and preventing the onset of congestion.
- To reduce the probability of rear-end collisions caused when motorists unexpectedly encounter congested conditions.
- To distribute total delay in a more equitable manner, saving some freeway capacity for downstream segments.
- To increase the efficiency of operation under restricted-capacity conditions caused by incidents or maintenance operations.
- To divert some freeway traffic to alternative routes or to alternative departure times in order to make better use of corridor capacity.

TREATMENTS

Traditionally, freeway lane use control is accomplished through one or more of the following methods:

- Static signing.

- Dynamic message signs.
- Temporary traffic control devices.
- Law enforcement/legal restrictions.
- Economic incentives/disincentives.

These methods are often combined to form an effective subsystem from both a safety and operations perspective. For example, both appropriate static signing and temporary traffic control devices are required in order to temporarily close a freeway travel lane for maintenance or construction activity.[2] In some instances, dynamic message signs may be provided upstream to encourage diversion from the freeway or to further warn of the downstream lane closure. As another example, the restriction of large trucks to certain travel lanes may be accomplished by static signing to notify truck drivers of the restriction, and by police enforcement of those restrictions.

MODULE OBJECTIVES

The objectives of this module are as follows:

- To provide a summary and description of different lane use control philosophies and treatments.
- To provide insight into the issues associated with lane use control subsystems for freeway traffic management.

MODULE SCOPE

This module begins by applying the decision process described in **Module 2** to the lane use control subsystem analysis. The techniques used to control lane use on freeways are then described. Special issues associated with lane use control are addressed following description of the techniques. The module concludes with examples of lane use control subsystems, describing the dynamic information system component of TransGuide in San Antonio, and the use of congestion pricing on the SR 91 Express Lanes in Orange County, California.

4.2 DESIGN PROCESS

PROBLEM IDENTIFICATION

The development of a lane use control subsystem begins by developing a full understanding of the problems or situations that need to be addressed through this subsystem. At the most basic level, the problems treated via lane use control are similar to those specified for other freeway management subsystems. These include excessive peak period vehicle demands that result in congestion, safety problems, excessive vehicle emissions that degrade the air quality of a region, etc. Most generally, the types of problems that need to be treated through lane use control are as follows:

- One or more vehicle lanes must be closed to all traffic for some period of time (due to scheduled work activity or an incident).
- One or more types of vehicles need to be separated from each other in one or more vehicle lanes.
- Speeds and flow rates in a given lane are judged to be too high for safe and efficient operations.

These problems lead directly into the three basic categories or situations that are most commonly treated through some form of lane use control:

- Restricting the use of freeway lanes and shoulders by all vehicles.

- Restricting the use of freeway lanes or shoulders by specific vehicle types.

- Metering how vehicles utilize a freeway lane or shoulder.

The restriction of freeway lane usage by all vehicle types includes such specific actions as temporary lane closures for roadway maintenance and construction activities or incidents when crashes or vehicle stalls block individual travel lanes, and peak-period congestion relief actions where vehicles are allowed to utilize a shoulder as a temporary travel lane during the peak period. Lane use restrictions based on vehicle type covers such topics as high-occupancy vehicle lanes and large truck restrictions. Finally, strategies to meter a freeway lane or a shoulder include freeway-to-freeway connector metering or the reduction of speed limits through dynamic advisory speed signs or speed enforcement.

For freeway lane use control, it is imperative that decision makers have a true understanding of the problem(s) being addressed through lane use control, and also a sense of the degree of urgency or severity of those problems. Not all strategies available for use are equally effective or costly in managing freeway lane usage. For example, both dynamic advisory speed signs and speed enforcement strategies can be viewed as methods of reducing freeway speeds in one or more lanes. However, the effectiveness and costs would be expected to be significantly greater for enforcement than for advisory signing.[3]

IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING

For all aspects of freeway management, information sharing is essential in today's environment. This is particularly true for lane use control strategies. The political implications of these control strategies can be very important, particularly with respect to lane restrictions which require legislation and enforcement to enact. Also, support by enforcement agencies of the decisions made regarding lane use control is essential if those decisions are to achieve their intended purpose.

Potential Partners

The partners involved in lane use control will vary depending on the particular problems being addressed. The partners that are most commonly associated with these types of activities include the following:

- State and local DOTs.

- State and local law enforcement agencies.

- Elected officials.

- Transit agencies.

- Trucking companies.

- Private contractors.

Concern over the interaction between automobiles and trucks on freeway facilities is often a controversial part of freeway management activities. Fortunately, trucking companies have demonstrated a willingness to work with public transportation and enforcement agencies to develop compromises about when and where trucks will travel during peak periods, major freeway construction activities, or special

events.[4] Consequently, they can be an integral partner in decisions regarding lane use control as well.

Consensus Building

After identifying the partners affected by lane use control actions, a consensus about the problems and need for solutions must be developed. Early on, it is critical to establish support for the general concepts about lane use control from elected officials and the general public. This is particularly important when not all motorists are affected equally by the actions that may be implemented (via tolls, lane restrictions for certain vehicles, etc.). Support is also needed from upper management at each of the agencies listed above.

ESTABLISH GOALS AND OBJECTIVES FOR LANE USE CONTROL

Proper decision making about the selection, design, and implementation of lane use control strategies requires goals and specific objectives about what is to be accomplished by using these strategies. General goals and objectives that feed directly into this step may have been established early on in the development of a comprehensive freeway management system . If not, it is imperative that the exact objectives to be accomplished via lane use control be specified. As discussed in **Module 2**, goals are broad statements of intended outcomes, whereas objectives specify exactly what is to be accomplished. Examples of objectives which lane use control may be able to address include the following:

- Increasing the available throughput over a given section of freeway.
- Providing at least a travel time incentive to high-occupancy vehicles.
- Extending pavement life (by restricting trucks from particular lanes).
- Reducing truck-automobile accidents over a given freeway segment.
- Reducing peak-period traffic demand on the freeway.

ESTABLISH PERFORMANCE CRITERIA

Performance criteria must be established for each of the objectives identified in the previous step of the process. These criteria are used to determine whether each of the objectives is met or exceeded. For the most part, performance criteria to be used will be fairly obvious if the objective(s) to be accomplished are specified correctly. For example, an objective to reduce high-occupancy vehicle travel time would utilize a travel-time measure as its performance criterion. However, additional performance criteria may be specified as checks of the primary performance criteria, or to measure other constraints that may exist on freeway lane operations. As another example, performance criteria concerning single-occupant vehicle violation rates might also be used to evaluate the need for a barrier-separated HOV lane (see **Module 6**) versus a concurrent flow HOV lane with no barrier separation. The following list presents some typical performance criteria for lane use control subsystems.

- Level of service on a facility (see **Module 2**).
- Compliance (lane violation) rates.
- Change in mode split, average occupancy rates.
- Accident reductions.

- Results of public and/or special interest opinion polls.

DEFINE FUNCTIONAL REQUIREMENTS

The functional requirements of a lane use control component in a freeway management system define specific actions or activities that are to be performed in order to achieve one or more of the objectives for that component. In theory, the functions should be defined independent of the technology used to implement them. However, some lane use control objectives may be so narrowly defined and governed by standards or policies (temporary freeway lane closures for maintenance, for instance) that this step in and of itself defines the technology to be used.[5] For other objectives, though, the technology required may not be so obvious, and so a definition of functional requirements would be warranted. Figure 4-2 presents an example of possible functional requirements of a lane control subsystem to increase peak period freeway capacity.

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

The functional relationships, data requirements, and information flows show how the lane use control functions will be integrated with each other and with the other freeway management system components (such as surveillance or motorist information dissemination). When defining the functional relationships, data requirements, and information flows as they impact dynamic message signs and surveillance, the National Transportation Communications Interface Protocol (NTCIP) and the National ITS Architecture should be followed.[6] This approach provides the following advantages:

- Automatically determine when freeway lane volumes reach 90 percent of estimated capacity.
- Check to ensure that no stalled vehicles are located on the shoulder in the affected section of freeway.
- Notify motorists at the beginning of the affected section that freeway shoulder can be used as a travel lane.
- Notify motorists at the end of the affected section that they should return to the normal freeway travel lanes.
- Terminate motorist notification of allowability of freeway shoulder usage at the end of the peak period.

Figure 4-2. Example of Possible Functional Requirements of a Lane Use Control Objective.

- It allows components from different vendors to be used interchangeably, increasing competition and reducing costs.

- It eases future upgrades or expansions of the components. The necessary interfaces to other components of the freeway management system (i.e., surveillance) have already been established.

For other lane use control functions, informal functional relationships, data requirements, and information flows may be acceptable, but should be prepared nonetheless. For example, the functional relationships, data requirments, and information flows for managing a temporary total freeway closure over the weekend might include how real-time data concerning traffic conditions, project status, and/or the effects of weather are transferred between the transportation agencies (such as the traffic, maintenance, and public information divisions), enforcement agencies, the media, and the private contractor.

IDENTIFY AND SCREEN TECHNOLOGY

After the functional relationships have been specified for the functions to be accomplished through freeway lane use control, decision makers must then assess available technologies and strategies to determine which are most appropriate to meet the desired objectives of this subsystem. Depending on the types of functions being accomplished, the assessment should include considerations of both the spatial and temporal characteristics of the freeway system and motoring public. For example, spatial considerations address technology/strategy adequacy from the aspect of roadway design and construction, and may require the decision maker to answer questions such as the following:

- Is the pavement on the shoulder adequate to support vehicle travel?

- How will restricting large trucks to a specific lane or lanes affect vehicle merging or diverging maneuvers at entrance and exit ramp locations?

Temporal considerations are also important for assessing technologies for real-time lane use management. For example, before considering truck lane restrictions on a freeway facility, a decision maker may want to consider whether truck-automobile conflicts and crashes are a problem at all times of the day or only during certain periods. Likewise, decisions regarding when and how many lanes will be closed to accomplish a given work activity are a major part of the work zone planning process.[7] For instance, it may be possible to close more travel lanes at night to do the work, but this requires channelizing devices and other traffic control devices that provide higher levels of retroreflectivity (increasing planning and traffic control costs).[7]

Another factor to consider when identifying and assessing alternatives for lane use control are the political sensitivities and ramifications to those alternatives. Those alternatives which would adversely impact some motorists in order to provide an advantage to other motorists have typically been met with resistance. For instance, those HOV projects which have taken a regular-use travel lane and converted it into an HOV lane (i.e., take-a-lane alternatives) have generally been less successful than those which have constructed a new HOV lane while maintaining the same number of regular-use lanes as before (i.e., add-a-lane alternatives).[4] Alternatives that involve direct cash outlays, such as toll facilities or

congestion pricing schemes, have also been met with resistance.[8] These alternatives require more public outreach and political interaction between the partners to ensure that they will be successful if implemented.

PLAN DEVELOPMENT

After the various lane use control alternatives available have been screened and the one(s) most appropriate for use selected, decision makers then prepare detailed plans regarding implementation of the selected alternatives. The actual format and content of the plan varies dramatically by the type of alternative being implemented. The plan for a temporary lane closure for maintenance work may be as simple as a standard traffic control plan (TCP) that has been adopted by the agency for that particular roadway/road work condition. Conversely, the plan to implement a congestion pricing scheme based on real-time traffic conditions and utilizing ITS technologies may need to follow detailed implementation plan requirements as specified by FHWA (see **Module 2**). [9]

Enforcement should be a key element of all lane use control plans that are developed. Plans should address both the management and coordination requirements between the various partners, relative to enforcement. Operational issues such as staff requirements and citation locations also need to be addressed during plan development.

FUNDING SOURCE IDENTIFICATION

Many of the lane use control strategies employed during freeway management activities are supported as part of an operating agency's normal activities, or are addressed during development of other components of the freeway management system (i.e., information dissemination subsystems). Consequently, the primary concern for these strategies is the extent to which introducing the new strategy affects the existing budget for operations and maintenance, and whether this impact can be accommodated through a reallocation of agency funds.

Those strategies which are larger in scope and/or applied across jurisdictional boundaries often require more innovative and complex funding mechanisms. Funding of lane use strategies involving HOV lanes, information dissemination, and/or incident management subsystems are addressed in **Modules 6**, **7**, and **8**, respectively.

Two of the major lane use control options, implementation of tolls and/or congestion pricing strategies, result in revenues that are used to offset the cost of constructing and operating these strategies. Traditionally, toll facilities were converted to "free" roadways once the bonds used to construct the roadway had been paid. Recent changes in legislation, however, now allow agencies to continue toll operations after bond payment, and to use the revenues to fund other traffic management activities.

IMPLEMENTATION

Experiences from past freeway management projects indicate that it is best to implement strategies and techniques incrementally where possible in order to develop operational experience with the strategies, and to demonstrate the advantages of the techniques to elected officials and to the public.[4] This may be true for some of the more "innovative" or controversial lane use control strategies as well. For these situations, consideration should be given to initiating small, demonstration-type projects at a location or over a section of freeway where the benefits are expected to be the greatest. In this way, the partners can

illustrate the benefits of the strategy and generate the support necessary to proceed with more extensive implementation if desired.

EVALUATION

The final step in the decision process for lane use control is to establish the mechanism for evaluation of the strategies once they have been implemented. It is important to monitor the impacts and benefits of new strategies and techniques as they are implemented to determine if they meet the intended objectives and functions for which they were designed. Also, it is important that these data be collected so that they can be collated and disseminated in an ongoing manner to elected officials and the general public. In this way, continued funding for these strategies can be obtained more readily, and expansion of activities to further improve facility operations will be more readily accepted.

4.3 TECHNIQUES AND TECHNOLOGIES

Lane use control can be implemented through the following means:

- Static signing.
- Dynamic information systems.
- Channelizing devices.
- Use of narrow lanes and shoulders.
- Mainline metering.
- Toll facilities/congestion pricing.
- Automated highway systems.

A brief description of how each of these technologies applies to the lane use control concept is provided below.

STATIC SIGNING

Signs are used to provide guidance and warnings as needed to ensure the safe and informed operation of individual elements of the traffic stream. Signs can provide lane use control by restricting certain vehicles from using or not using a particular lane. Examples include the following:

- Restricting trucks to the right most lane(s) or to specific times of day that they can use a facility.
- Establishing lane(s) for high-occupancy vehicle use only.

Discussions of these examples follow.

Static signing is also used for warning purposes in advance of construction and maintenance work zones to inform motorists that one or more travel lanes are closed downstream. These generally supplement the channelizing devices placed at the point of closure to require drivers to vacate the lane at that point (channelizing devices are discussed later in this module).

Truck Restrictions

Several States restrict the lanes in which trucks may operate. The objective in restricting trucks to the right lane or lanes is typically to improve highway operations and reduce accidents. Also, to provide for uniform pavement wear, trucks are sometimes restricted from the right lanes. Lane restrictions through construction zones are used to move the trucks away from workers and from narrower lanes. Table 4-1 is a summary of experience regarding lane restrictions in various states and in research

studies.[5] Several States adopted lane restrictions because trucks were often observed traveling abreast across several lanes, denying passing opportunities for other vehicles.

A survey of State practice in 1986 by the FHWA identified the most common reasons given for using truck lane restrictions:

- To improve operations (fourteen States).
- To reduce accidents (eight States).
- For pavement structural considerations (seven States).
- Because of restrictions in construction zones (five States).

A total of 26 States used lane restrictions, according to survey information. The survey of FHWA field offices also indicated that, in most cases, restrictions have been applied without detailed evaluation plans, including before-and-after studies. Where accident analyses were undertaken, little change in accident experience was noted under any of the restrictions.[10]

Truck restrictions can be implemented in a number of other ways as well. Table 4-2 summarizes the constraints and impacts of different types of restrictions.

HOV Restrictions

Another form of lane use control is when a lane is restricted for use by high-occupancy vehicles only. Several types of static signs have been used to communicate the restriction. (For discussion on high occupancy vehicle lanes, see **Module 6**).

INFORMATION DISSEMINATION SYSTEMS

Information dissemination systems can be employed to advise motorists of freeway conditions so that appropriate actions can be taken to enhance the efficiency and safety of freeway operations. **Module 7** provides a discussion of information dissemination principles and technologies. For lane use control applications, both lane control signals (LCS) and full-matrix dynamic message signs (DMSs) can be used to convey lane use and lane status information to drivers. LCSs are a fixed-grid DMS (refer to **Module 7**) that use both color and symbols to convey information. The Manual on Uniform Traffic Control Devices (MUTCD) defines LCSs as special overhead signals having symbols that are used to indicate whether the use of a specific lane or lanes of a street or highway is permitted or prohibited, or to indicate the impending prohibition of use.[2] In the United States, LCSs have most commonly been used for reversible-lane control. However, they are also appropriate for use on freeways in the following situations:[2]

- On a freeway, where it is desired to keep traffic out of certain lanes at certain hours to facilitate the merging of traffic from an entrance or exit ramp or other freeway.
- On a freeway, near its terminus, to indicate a lane that ends.
- On a freeway or long bridge, to indicate a lane that may be temporarily blocked by an accident, a breakdown, or some other incident.

In addition, at least one operating agency (Virginia) uses LCS to indicate to motorists that a shoulder can be used as a travel lane

Table 4-1. Lane Restrictions Experiences.[5]

Location/Study	Conditions	Results/Comments
Florida I-95, Broward County	Conducted a 6 month, 7 am to 7 pm study in 1988	Public feels safer with lane restrictions for trucks. Overall accidents up 6.3 percent (7 am to 7 pm period); truck accidents down 3.3 percent.
Georgia	Beginning Sept. 1986, trucks were restricted to the right lane(s) except to pass or to make a left-hand exit.	On I-285, trucks were at fault in 72 percent of lane-changing violations. Before the restriction, trucks were observed occupying all lanes thus prohibiting passing.
New Jersey	Turnpike Authority (NJTA) imposed lane restrictions in the 1960s. Restrictions do not allow trucks in the left lane of turnpike roadways that have three or more lanes by direction.	Sources at the NJTA stated that the compliance rate for truck lane restrictions is very high.
Illinois	Began in 1964.	Public feels safer, and better operations.
Maryland Capital Beltway	Believed to have been implemented as a reaction to a major truck accident.	Public feels safer. Effects on safety not well known.
Virginia Capital Beltway	Four studies, one for 24 months, others for 12 months.	Public and political perception: safer highways. Engineering study recommended removal. Accident rate increased 13.8 percent during 2-yr study. Second study also showed increase.
Michigan	Statewide restrictions require trucks to use the right two lanes on roadways that have three or more lanes.	Establishment was thought to be politically motivated No studies available to evaluate the countermeasure.
Garber Study	Simulation based on data from nine sites.	Decreased headways in right lane Slight increase in right lane accidents.
Hanscom Study	Two 3-lane suburban sites, all <100,000 AADT.	Beneficial traffic operations and reduced congestion.

Table 4-2. Summary of Impacts from Truck Restrictions. [11]

Action	Constraints/Limitations	Impacts
Lane Restrictions	• Lane drops at freeway-freeway interchanges limit application. • Could be difficult to enforce. • Could accelerate pavement deteriorations. • Could reduce visibility of overhead signing (if trucks restricted to outside lanes).	• For freeway segments with lane drops, would concentrate lane changes in short section of freeway. • Would increase merging conflicts.
Time-of-Day Restrictions	• Truck traffic peaks do not coincide with typical commuter peaks. • Could be difficult to enforce. • Could be challenged on legal basis.	• Negligible impact on operating speeds. • Could divert trucks to other less congested time periods, or other, lower quality roadways. • Could negatively impact trucks that must travel during restricted periods.
Speed Restrictions	• Differential speed limits for trucks and non-trucks could be difficult to enforce. • Could require extensive enforcement program. • May require use of innovative detection, apprehension, and citation strategies.	• Reduction in speed (differentials) could have positive safety impacts.
Route Restrictions	• Efficient routing plan could not exclude freeways.	• Negligible impacts on safety and operations. • Could have positive impacts if applied to transportation of hazardous materials.
Driver Training/ Certification	• Requires strict application and enforcement of regulations.	• Short-term impacts minimal. • Long term impacts could be significant.
Increased Enforcement of Existing Regulations	• Would require additional enforcement personnel. • Could require incorporation of enforcement requirements in design/re-design of freeways.	• Increased enforcement could lead to increased compliance with traffic laws. However, there is no conclusive proof that increased compliance reduces accidents.

during peak travel periods.[12] Figure 4-3 shows a typical freeway LCS.

The MUTCD gives information on the design, location and operation of lane-use control signals. The meanings of freeway lane-use control signals are as follows:[2]

- A steady DOWNWARD GREEN ARROW means that a driver is permitted to drive in the lane over which the arrow signal is located.

- A steady YELLOW X means that a driver should prepare to vacate, in a safe manner, the lane over which the signal is located because a lane control change is being made, and to avoid occupying that lane when a steady RED X is displayed.

- A steady RED X means that a driver shall not drive in the lane over which the signal is located, and that this indication shall modify accordingly the meaning of all other traffic controls present.

In contrast to LCSs that display a rather small number of colors and symbols, full-matrix DMSs are capable of providing a large number of descriptive messages to indicate that one or more travel lanes are closed downstream, that a freeway shoulder can be used as a temporary travel lane, that an HOV lane is open to mixed-use traffic (as might occur during a major incident), or that trucks normally restricted to a right lane can use other travel lanes as well (as also might occur during a major incident).

Figure 4-3. Typical Freeway Lane Control Signal.

LCSs are smaller than full-matrix DMSs, and so are considerably cheaper to purchase and maintain. This means that LCSs can be installed more frequently along a freeway than can larger and more expensive DMSs. Also, since LCSs use symbols and colors rather than words to convey information, they can be more readily understood by non-English speaking motorists. However, the amount and type of information that can be displayed via LCSs is much more limited than via a typical DMS.[13]

Regardless of the technology used, information dissemination systems alone cannot force vehicles to vacate a lane or utilize a shoulder as a travel lane. They do provide guidance about lane status, and are intended to promote safer operations by warning motorists upstream of an actual lane blockage.

USE OF NARROW LANES AND SHOULDERS

Research has confirmed that shoulders and narrow lanes can be used effectively to increase capacity in congested metropolitan corridors. Based on a recent NCHRP study, it is recommended that these strategies be reserved for use as techniques to improve traffic flow in congested corridors.[14] Widening a corridor over an extended length through the use of these strategies is not recommended. Rather, applications of these strategies should be viewed as a temporary technique for congestion relief. It is recommended that this technique should typically be limited to sections of 1.61 km (1 mi) or less.

Reduction in the travel-lane width to 3.4 m (11 ft) should be the first modification considered. Next, reduction of the left shoulder should be considered before reducing the right shoulder. Research and observations by enforcement personnel indicate that the right shoulder is the preferred refuge area. Also, emergency response is easier to provide if the right shoulder is maintained. If the right shoulder is used and the left shoulder maintained, emergency equipment entering a congested area must work its way across the queue to the left shoulder as opposed to proceeding on the right shoulder. In some cases, the right shoulder or both shoulders have been used. Table 4-3 summarizes the primary advantages and disadvantages of each approach.

Where there is a high percentage of truck traffic (i.e., 5 to 10 percent) during the peak period, use of shoulders and narrow lanes is not recommended. Also, for projects involving possible application of shoulders and narrow lanes, a step-by-step approach (site specific) must be used to ensure an adequate evaluation.

MAINLINE METERING

Mainline metering controls the mainline traffic entering a freeway control section. While the technique can create congestion on the mainline upstream of the control section, it can help maintain uncongested flow on the mainline downstream through the control section. An application of mainline metering has been documented in Japan, where it has been made effective by regulating the number of toll booths open, thereby controlling the traffic entering the freeway. Another application is the San Francisco-Oakland Bay Bridge mainline metering in Oakland, where controls are used after vehicles pass through the toll plaza at a point prior to entering the ridge.[15] Other applications of mainline metering on controlled-access facilities such as terminals, bridges, and toll roads have been found to be effective, but the concept has yet to be applied to a typical metropolitan freeway system.

Table 4-3. Primary Advantages and Disadvantages of Shoulder Removal. [14]

Design Alternative	Advantages	Disadvantages
Use of Left Shoulder	• Left shoulder not used as much for emergency stop/or emergency enforcement • Least expensive if width is available • Trucks often restricted from left lane	• Usually requires restriping • Sight distance problem with some median treatments
Use of Right Shoulder	• Often the easiest to implement	• Right shoulder is preferred area for emergency stops and enforcement • Sight distance changes at merge and diverge areas of ramps
Use of Both Shoulders	• Not recommended • Use ONLY in extreme cases	• Requires restriping • Safety concerns (no refuge) • Enforcement is difficult • Incident response longer • Maintenance more difficult and expensive

Another potentially effective use of mainline metering that has not yet been applied is in a construction zone where traffic demand greatly exceeds the capacity available, and some reserve capacity for vehicles entering at downstream ramps needs to be provided.[1]

Another application of mainline metering involves using DMSs to display reduced advisory speed limits over sections of freeway. This technique is commonly used on portable DMSs upstream of work zone lane closure activities.[16] However, its application to freeway management in the U.S. has been less favorable. Nonetheless, this approach is used with reasonable success outside of the U.S.[17]

FREEWAY-TO-FREEWAY RAMP METERING

The primary benefit of ramp metering is that it reduces congestion and shockwaves on freeways. Metering eliminates the heavy surges or platoons of entering traffic and allows the freeway to carry substantially higher volumes at higher speeds. It also substantially reduces accidents, especially the rear-end type crashes that result from stop-and-go driving. Reduction in accidents and congestion also results in overall reductions in fuel consumption and air pollutant emissions. Freeway-to-freeway ramp metering uses a meter on a ramp from another freeway to control the entrance of vehicles onto the new freeway. Minnesota Department of Transportation (MnDOT) has been successfully metering freeway-to-

freeway ramps for over 20 years and is currently metering 74 ramps. Although there is some concern over metering these high volume ramps, MnDOT believes it is desirable to do so because heavy surges of traffic, if left unmetered, can severely affect the operation of the receiving freeway by causing congestion and accidents.[18]

CHANNELIZING DEVICES FOR WORK ZONES

Traffic management in construction and maintenance work zones is important to the safety of both workers and motorists. Time is required to properly develop and implement the traffic control when lanes must be closed to complete the work. No one sequence of traffic control devices can be designed for all situations. The traffic control plan should be designed to provide the same number of lanes of traffic during construction as before and ideally, be designed for the same free-flow traffic speed that existed before freeway construction began. A work zone cannot always maintain the same number of lanes and, in addition, where motorists drive in a work zone can be influenced by the traffic control devices present.

Static signs are used for directing traffic in advance of and within a work zone. However, variable message signs and arrow panels supplement channelizing devices and provide additional target value and suitable messages that attract motorists' attention as they approach a work zone. Highway Advisory Radio (HAR) can also be used to provide more information, including route detours. Newspaper articles and traffic broadcasts can also be employed to alert motorists to general construction and maintenance work zone locations, but provide limited detailed information about specific lane closures and traffic delays.[1]

Within the work zone itself, traffic must be channeled from the lane(s) being closed to one that is open. A number of different types of channelizing devices can be used, depending on the duration of the closure, traffic speeds, etc. These devices include but are not limited to the following:[2]

- Cones.
- Tubular markers.
- Vertical panels.
- Drums.
- Barricades.

Regardless of the device used, it is critical that proper spacing and length of taper be employed to safely transition motorists from the closed lane(s) to the open lane(s).[2]

Portable barriers may be used to separate traffic from the work area and to protect the construction workers. Special signing and temporary delineation and/or route detours may be needed when these barriers are moved and traffic is shifted. Portable barriers have been used to create HOV lanes in Dallas and Boston. This device could also be used as a lane control technique in a construction area. (See the discussion in **Module 6** for additional information).

TOLL FACILITIES/CONGESTION PRICING

Toll facilities are fully controlled access roadways designed to the same high standards of design as freeways. However, because of both the economic influence they exert on traffic demand for the facility and the traffic metering effect that occurs at the toll collection plazas, toll facilities can also be considered another form of lane use control. Toll facilities form the most direct

user charge for providing revenues based on the costs of travel. Often, they can be implemented more quickly, because the capital funding is available up front and because toll roads often do not have to comply with federal regulations. [19] Also, toll facilities provide adequate funding for ongoing operations and maintenance.

In recent years, legislation has been enacted to facilitate the development of toll facilities through innovative financing. For example, the Federal Highway Act of 1987 provided for eight demonstration projects across the nation, allowing a mixture of toll revenues with State and federal funds on new projects.[20] California enacted legislation that allows the development of joint-power authorities to collect developer fees for transportation projects, and awarded four franchises to private entities, allowing private designing, financing, construction, and operation of toll facilities for 35 years.[20]

Two basic types of collection methods are used: mainline barriers at intermittent spacings, and toll collection at interchanges. Where mainline barriers is used, the location of the barrier should be far enough from exit or entrance ramps to avoid weaving problems. Toll plazas should be located on tangent alignment with decision sight distance provided on the approach. Grades approaching the plaza should not exceed 0.5 percent. The design of the plaza itself should be based on the expected number of peak-period arrivals and the rate at which they can be processed. Sufficient toll collection lanes of great enough length should be provided to minimize the length of queuing at the plaza. Where toll collection occurs at exit and/or entrance ramps, care should be taken in interchange ramp design. Plazas should be placed far enough upstream from the ramp diverge to provide for normal deceleration and braking to a queue.[1]

A new technology experiencing increased use is Automatic Vehicle Identification (AVI) technology for automatic toll collection. The technology is being implemented to increase capacity and decrease delays and congestion at toll plazas. Figure 4-4 illustrates the proposed use of the AVI technology on the toll roads. An AVI system speeds the toll collection process by having a device identify each vehicle as it approaches a toll plaza or passes a check point, and charging the proper toll to the user's account, eliminating the need for vehicles to stop.

Congestion pricing is a transportation management technique similar in concept to toll facilities, but which attempts to spread peak traffic demands to less congested segments of the network and to less congested periods of the day. The application of congestion pricing has been very limited within the U.S. Experience in Hong Kong in the mid-1980s proved the program to be a technological success, although not widely accepted by the public.[22]

Recent advances in technology are making congestion pricing a viable option in the U.S. Automatic fare collection systems can be used for this purpose, including optic scanners and radio frequency technology. The system would automatically charge the user by prepayment, direct billing, or credit card. A billing statement would be received by the user each month. Typically, the user would receive advance notification of billing rates, ideally one month prior to charging. Table 4-4 shows a typical user fee schedule.

Key issues relative to congestion pricing in the U.S. are as follows:[23]

- Technology compatibility — between toll collection systems across the country and across travel modes.

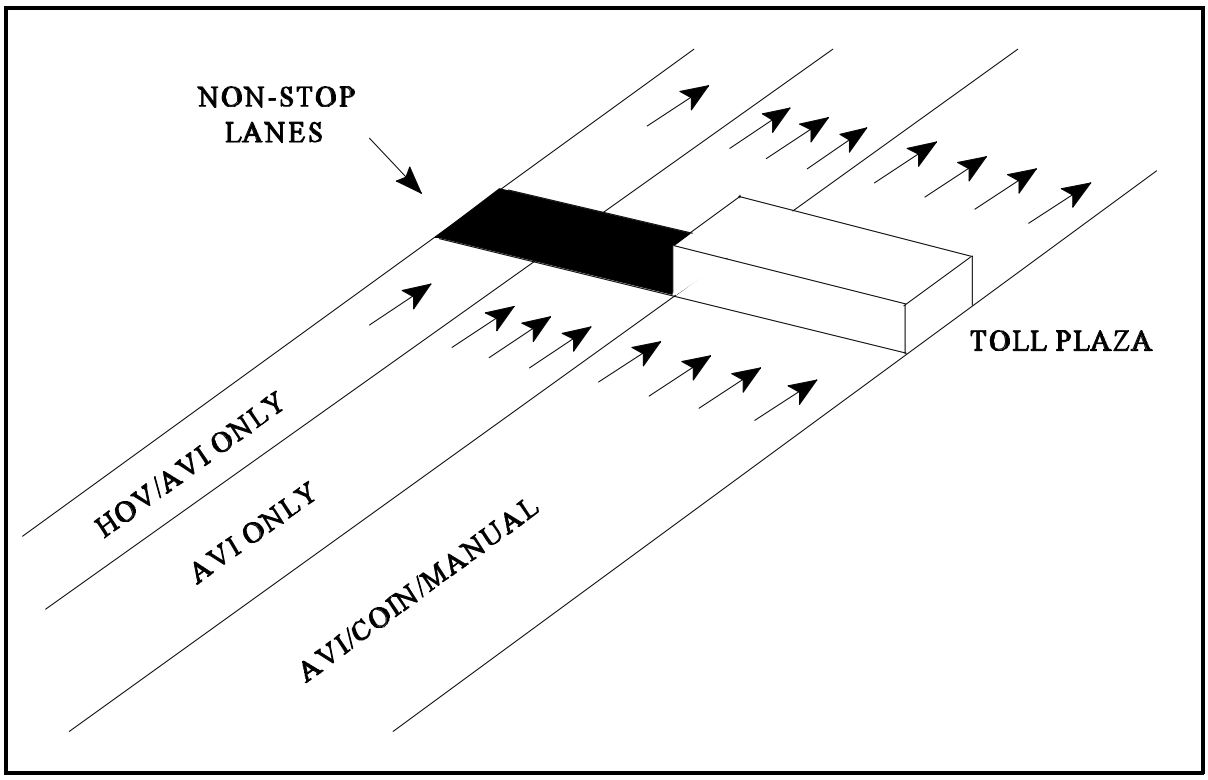


Figure 4-4. Schematic of an AVI Toll Plaza. [21]

- Enforcement — concerns about legislative changes needed to allow vehicle owners to be ticketed, rather than drivers. Also, verification of occupancy requirements (if implemented) needs to be accommodated without unduly affecting the overall operations of the facility.

- Privacy — concerns about the potential tracking of individuals through the AVI technology. One alternative proposed is the use of "Smart" cards which maintain their own internal data about the individual's account.

- Price determination — questions have to be resolved (should charges be based on short-term marginal road user costs of an additional vehicle added to the traffic stream, or on long-term costs of providing an equivalent amount of extra capacity to the facility to accommodate that alternative pricing structures upon traffic demand, air quality, land use, etc. are difficult to estimate at this time.

AUTOMATED HIGHWAY SYSTEMS

Automated Highway Systems (AHS), a component of the Intelligent Transportation Systems (ITS), is seen as the next major performance upgrade of the United States vehicle-highway system. It should be mentioned, however, that AHS is still in its infancy and is many years from being an operational reality. The Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 established a program to determine the feasibility of an AHS. Research has shown that AHS has the potential to double or triple the nation's roadway capacity and reduce the frequency and severity of incidents while providing more stable traffic flow.[24,25] The increased capacity will result from the closer longitudinal and lateral

Table 4-4. Typical User-Fee Schedule.[22]

Route Segment Number	Peak Period (6:01 am - 9:00 am; 3:01 pm - 7:00 pm)				Off-Peak Period (9:01 am - 3:00 pm; 7:01 pm - 6:00 am)			
	1	2	3	4	1	2	3	4
A	$0.09* 45 mph	$0.09 45 mph	$0.09 45 mph	$0.07 50 mph	$0.05 55 mph	$0.05 55 mph	$0.05 55 mph	$0.04 55 mph
B	$0.04 60 mph	$0.04 60 mph	$0.04 60 mph	$0.04 60 mph	$0.02 60 mph	$0.02 60 mph	$0.02 60 mph	$0.02 60 mph
C	$0.06 25 mph	$0.06 25 mph	$0.05 30 mph	$0.05 30 mph	$0.04 35 mph	$0.04 35 mph	$0.02 40 mph	$0.02 40 mph
D	$0.10 40 mph	$0.10 40 mph	$0.08 45 mph	$0.08 45 mph	$0.08 45 mph	$0.08 45 mph	$0.06 50 mph	$0.06 50 mph
E	$0.10 20 mph	$0.10 20 mph	$0.08 25 mph	$0.08 25 mph	$0.08 25 mph	$0.08 25 mph	$0.06 30 mph	$0.06 30 mph
F	$0.05 30 mph	$0.05 30 mph	$0.08 30 mph	$0.08 30 mph	$0.04 35 mph	$0.04 35 mph	$0.06 35 mph	$0.06 35 mph
G	$0.04 45 mph	$0.04 45 mph	$0.07 40 mph	$0.07 40 mph	$0.03 55 mph	$0.03 55 mph	$0.05 55 mph	$0.05 55 mph

* Dollar amount indicates user-fee rate (dollars per mile) for using a specific route. The mile-per-hour figure indicates anticipated travel speed when using a specific route. For example, travel on Route B/Segment 3 during the peak period would cost $0.04 per mile, with an anticipated average travel speed of 60 mph.

spacing that will be allowed due to the full control of the AHS expressway. The Federal Highway Administration (FHWA) has established an AHS program that will be able to carry out the following:

- Identify and analyze alternative AHS concepts.
- Demonstrate the potential feasibility of AHS in 1997.

Concept

AHS refers to the use of modern electronics, sensors, and communications on highway vehicles to provide "fully automated" vehicle operation. The automated system of roadways will run adjacent to the existing infrastructure providing an option to choose the AHS facility or remain on the conventional freeway. As a vehicle approaches and accesses an AHS lane of an AHS expressway, the vehicle's steering, braking, and acceleration are controlled until the vehicle exits the lane, after which the driver once again assumes control of the vehicle. An AHS expressway will have four major components that are fundamentally different from a conventional expressway. These segments are as follows:[26]

- Vehicle characteristics.
- Roadway infrastructure.
- Command and control.
- Entry and exit infrastructure.

The functional characteristics of each of these segments are discussed in the paragraphs below.

Vehicle Characteristics

The function of the AHS vehicle is to carry the driver, passengers, and goods as they are moved through an AHS system. The vehicle must do the following:

- Provide for controlled vehicle movement.
- Interact with the AHS roadway infrastructure component to obtain traction and support for operation while in the system, and to obtain lane boundary indications.
- Interact with the entry and exit infrastructure component to provide smooth and rapid entry and exit to AHS.
- Provide accurate control responses to directions received from the command and control component regarding vehicle braking, steering, throttle, and lights.
- Detect and maintain the status of critical vehicle functions.
- Support access to and from the command and control component.
- Interact with the driver, on a user-friendly basis.

Roadway Infrastructure

The primary function of the roadway infrastructure is to:

- Provide traction and support for vehicle operation, including vehicles operating properly and those that are malfunctioning.
- Enable safe vehicle operation by ensuring vehicle separation in case of severe system malfunction.
- Provide connectivity for entering and exiting vehicles and connectivity to other AHS systems.
- Provide passive or active indication of lane boundaries.
- Provide sensing of environment or obstacles, or both.
- Support/enable command and control and communication access to AHS vehicles and to roadway conditions.
- Support access to roadway by emergency and maintenance vehicles.

Command and Control

The AHS expressway is controlled primarily by the command and control component. The approach that is taken to accomplish the AHS command and control functions will have a strong influence on how the entire system is to be implemented. The five primary functions of the AHS command and control segment are listed below:

- Traffic flow management.
- Intervehicle coordination.
- Incident management.
- Vehicle control.
- Vehicle management.

Entry and Exit Infrastructure

The entry and exit infrastructure refers to the component of the system that transitions to and from the AHS expressway onto or from

the regular freeway lanes. There are several methods for making this transition, but it is hypothesized that these transitions will include separation by vehicle classification and type of in-vehicle AHS technology. The classification separation is made so that lateral space can be optimized.

Goals

In order for an Automated Highway System to be successful, it must provide positive and noticeable benefits. Benefits must be realized in the areas of capacity, safety, energy, level of service, and environmental and community impacts. AHS will provide a more efficient utilization of right-of-way by decreasing lateral and longitudinal spacing of vehicles, thus increasing capacity. Safety improvements will be realized through the implementation of AHS due to the automation of usually manual tasks that create the possibility for driver error. Due to the smooth operation and flow of the vehicles along an AHS expressway, increased energy efficiency will be experienced. By decreasing total trip time, increasing level of comfort, and maintaining flexibility, the level of service of the system can be increased. The AHS system is expected to improve each of these components. Lastly, the AHS will aid realization of environmental goals related to noise and air pollution, community disruption, and user acceptance.[27]

4.4 LESSONS LEARNED

PLANNING

Freeway-to-Freeway Ramp Metering

In considering whether to implement a freeway-to-freeway ramp metering system, several concerns must be examined (e.g., whether or not to provide a dynamic warning device to alert approaching motorists that there is a ramp meter in operation, and possibly a substantial queue downstream). MnDOT, which is currently metering 74 ramps, bases its analysis upon available sight distance, approach speeds, and queue lengths. In general, MnDOT uses ramp meters at most freeway-to-freeway ramps, except loops. If an advance warning device is used, it is a 30 cm (12 in) flashing yellow signal head, accompanied by a standard symbolic “traffic signal ahead” sign, mounted on a 3 m (10 ft) traffic signal pedestal. It is placed in flashing operation when the ramp meter is cycling.[18]

In the operation of the MnDOT freeway-to-freeway ramp meters, metering rates and geometrics are carefully analyzed. Freeway-to-freeway ramps use different timing parameters than other meters. A narrower range of rates is used so that under severe incident conditions ramp flow is reduced by 25 percent, as compared to a 50 percent reduction at other ramps. These parameters help to guard against long queues. When necessary, MnDOT constructs an extended storage lane upstream of the meter to prevent queuing from backing up onto the intersecting freeway. Also, signing is used to encourage motorists to wait in line on shoulders, rather than in the through lanes of intersection freeways.

MnDOT has noted that there are several major concerns at freeway-to-freeway ramps that do not exist at other metered ramps, and they may be the reasons why some traffic management agencies have been reluctant to meter these ramps. The biggest concern is over accidents on freeway-to-freeway ramps, or upstream on the approaching freeway. This concern is related to stopping high speed/high volume ramp traffic, and the resulting queues. Experience over the past 25 years in Minnesota has indicated that there have actually been very few accidents

at these ramps and that there is a net reduction in crashes.

Another concern over freeway-to-freeway ramp meters is that they may result in motorists being metered several times while making one trip. For example, a motorist entering the freeway from a metered ramp at an arterial street may also use several other freeways on a trip, thus encountering several ramp meters. There is also concern over negative reaction from the public when freeway-to- freeway ramp meters are deployed. The faster metering rates used at freeway-to- freeway ramps, however, have resulted in substantially fewer complaints about freeway-to-freeway ramp meters as compared to other ramp meters.

Toll Facilities

Toll road facilities are a means of financing roadway improvements. The intent is to finance roadway construction only if there is sufficient demand willing to pay a premium for services rendered by the facility. Furthermore, the user population must be willing to pay for the opportunity to save time in the system because toll road facilities are generally less congested. Currently, the concept of toll roads is being considered as a realistic procedure to finance the construction of needed facilities in a short time period. There are several advantages and disadvantages involved in the implementation of toll roads as opposed to free roads. Table 4-5 lists some of the advantages and disadvantages of financing a road via user tolls.

Congestion Pricing

There are logistical, institutional, and attitudinal barriers that must be addressed when implementing a congestion pricing system. A feasibility study must therefore be undertaken prior to implementation. The following tasks should be conducted:[22]

- Review of appropriate state-of-the-art technology.
- Attitudinal surveys of the system users.
- Organization of participating agencies and private-sector partners.
- Conceptual design for select U.S. cities.
- Preliminary capital and operational and maintenance cost estimates.
- Development of a conceptual implementation plan (i.e., time frame, lead agency, potential funding sources, institutional requirements).

Congestion pricing has seen very limited use in the United States, though already proven successful in efforts to reduce congestion in Singapore and several Norwegian cities. Implementation of roadway pricing has been discouraged in the U.S. due to technical and political problems. The use of AVI technology overcomes some of the technical barriers such as congestion at toll booths. In order to overcome political barriers, demonstration projects can be performed to introduce the concept. Congestion pricing should not be introduced on a facility that traditionally has no charge. Two recommended facilities that should be used to introduce this concept are as follows:[29]

Table 4-5. Advantages and Disadvantages of Toll Road Financing. [28]

Advantages	Disadvantages
• Timely construction, no delays. • Reduced construction costs. • Earlier realization of benefits, such as lower fuel costs, greater safety, comfort and convenience. • Regular inspections and maintenance required by lenders. • Inclusion of operating costs in financing and income plans; costs include police, emergency services, snow and ice control. • Ability to free tax funds for "free" roads.	• Higher maintenance and operations costs, in most cases. • Infrequent access. • Possible adverse effect on other roads in corridor. • Delays at toll booths. • Added costs of borrowing, collecting tolls.

- Existing toll facilities, where off-peak discounts and peak-hour surcharges can be introduced in order to increase ride-sharing incentives and reduce congestion.

- Completely new facilities.

A pricing schedule should be designed so that initial costs are fairly low. These costs should be maintained for a period of 6 to 12 months to "permit behavior patterns to stabilize," after which the cost should increase. This process should continue until the cost is at the desired level. The use of congestion pricing is expected to improve the traffic flow, ridesharing, and emission-reduction experienced on the roadway. These improvements should be quantified in the demonstration project to prove these benefits.

Table 4-6 presents barriers to the implementation of congestion pricing using AVI technology, and recommended strategies. Three arguments that justify the use of congestion pricing are:[30]

- Economic argument—automobile users should pay the full cost of the congestion their use of the road imposes on the public.

- Congestion is reduced.

- Environmental improvement—less congestion means less air pollution due to slow moving traffic flows.

Channelizing Devices for Work Zones

There is a wide variety of channelizing devices currently available for use in highway work zones. The MUTCD presents basic design standards for these devices and general guidelines for their use; however, it is the highway agency's choice where and when to use particular devices or sets of devices. Typically, work zone channelizing devices are chosen on the basis of one of the following practices:

1. Select the device with the lowest initial cost.

2. Select the device that is normally used by the agency.

Table 4-6. Barriers to Congestion Pricing Implementation. [30]

Barrier	Description	Recommended Strategies
Technical		
AVI technology	• Equipment unreliability.	• More research.
Logistical Problems	• Numerous structures at checkpoints needed. • Equipping vehicles with proper equipment. • Geography of the city.	• Provide financial incentive for equipment installation on vehicles. • Many problems solved with time.
Uniformity	• Different AVI technologies used by each agency.	• Develop standards for AVI systems. • Will require legislation but should not be a problem.
Social and Political		
Congestion Pricing as an Additional Tax	• Public perceives it as another form of taxation instead of a price to use the road.	• Stress that this is a user fee to support roads, not a tax. • Stress that revenues from the system will be linked to further expenditures on the system. • Describe lower rates at non-peak periods as a reward for traveling at that time.
Privacy	• Public does not want their vehicular movements monitored.	• Stress that a congestion pricing payment is no different than an itemized telephone bill. • Information not used for enforcement. • Pictures only taken of license plates. • Not looking at individuals, just tags.
Equity	• Forces poorer drivers off the road--politically undesirable.	• Show that it would not be any more equitable or inequitable than present system.
Business Interests	• May create decreased interest in downtown shopping areas.	• Effects not certain--need demonstration project.

3. Select a device already in stock.

4. Select the "very best" device just in case.

Each of these approaches has drawbacks, and collectively they have resulted in inflated job costs, unnecessarily large inventories, lack of uniformity, and in some cases, improper device use. As an alternative to the typical methods used for selecting channelizing devices for work zones, the value engineering approach (see **Module 11**) can be used. The approach involves 7 steps:[31]

- Determine the intended purpose of the devices.

- Identify available alternative devices.
- Select appropriate measures of device performance.
- Determine the performance of the alternative devices on the basis of selected performance measures.
- Estimate the total cost of each acceptable alternative.
- Calculate the relative value of each acceptable alternative, where value equals performance divided by cost.
- Select the alternative with the greatest value.

The following recommendations should be followed when using this approach:[31]

- Base value engineering study on comprehensive and accurate information.
- Use a team approach—team members are well trained and diverse in experience and technical background.
- Consider value engineering approach most appropriate for central office use—through pooling central office staff and data-gathering resources.

DESIGN/CONSTRUCTION

LCS and DMS Placement

A panel of eight TxDOT managers and engineers with expertise in Lane Control Signal (LCS) design and operation in freeway traffic discussed problems and potential solutions regarding LCS. Their recommendations included the following:[13]

LCS Visibility

- Drive throughs should be performed by TxDOT personnel when a red X is displayed due to its low legibility distance.
- If the department chooses to require double-stroked symbols for LCS displays, a maximum pixel spacing and/or effective stroke-width-to-letter-height should also be specified.
- A regular cleaning and bulb replacement schedule should be implemented.
- Back plates or back panels should be considered for placement behind LCS on overhead sign structures.

LCS Spacing and Mounting Locations

- LCS should be placed every 0.8 to 1.6 km (0.5 to 1.0 mi), but special geometric characteristics and driver decision points should also be considered during this placement.
- Mounting LCS on a cross-street bridge structure rather than on an overhead sign structure is desired.
- Positive guidance principles should be employed when determining the placement of the LCS.

DMS Placement

It is recommended that the following points should be considered when placing dynamic message signs for freeway management:[32]

- Signs should be targeted at an audience that is made up of local drivers generally familiar with the surrounding street system. This allows less detail in the messages presented.

- A 3-line sign with 18 characters per line is recommended.

- A letter height of 46 cm (18 in) is recommended to allow sufficient legibility distance.

- The minimum spacing between a DMS and other guide signs not co-located with the DMS should be 305 m (1000 ft).

- The maximum distance from a detour exit that a DMS should be placed is 1220 m (4000 ft).

- DMS located upstream from an interchange should not be used for messages requiring a detour maneuver beyond the interchange.

- Placing DMS on the far side of a sharp horizontal or vertical curve may limit the driver's ability to read the sign.

- An unrestricted visibility distance of 183m (600 ft) should provide the driver plenty of time to read and comprehend the message.

Use of Narrow Lane and Shoulders

Design guidelines for the implementation of projects involving the use of shoulders and narrow lanes were developed in a recent NCHRP research study based on the experience of agencies that participated in the study. Additional details are provided in the NCHRP report. The guidelines developed in the research represent more of a guide for *applying* design standards. They are intended to supplement, rather than supersede, existing standards. Following is a selection from the guidelines: [14]

- Field observations indicate that operational impacts of reduced shoulder or lane widths are most notable in the transition area. It is recommended that the transition area be located on a tangent, preferably in an area where there are no crossing structures, retaining walls, or other roadside appurtenances.

- On facilities with high truck percentages, it is recommended that trucks be restricted from using a right shoulder lane, which typically does not have adequate pavement structure to support heavy trucks.

- Emergency turnouts and crossovers should be provided along altered sections. It is recommended that enforcement and emergency response personnel be involved in selecting locations. These turnouts should be large enough to accommodate a tractor trailer unit and at least one piece of emergency equipment. The location of crossovers should be considered in conjunction with incident management plans.

4.5 EXAMPLES

SAN ANTONIO (TRANSGUIDE)

Description of System

In late 1995, the Texas Department of Transportation began operation of an advanced traffic management system on approximately 41.6 km (26 mi) of freeway in San Antonio, Texas. The system, labeled TransGuide, includes inductive loop detectors and closed-circuit television cameras for traffic surveillance, and a combination of dynamic message signs and lane control signals to convey freeway traffic conditions and lane status information to motorists on or approaching the freeways.

Other components are being introduced as the technology is developed. The system is also being expanded to include additional freeways in the San Antonio area.[34]

Lane control signals are placed approximately 1.6 km (1 mi) apart between major freeway-to-freeway interchanges. Signal spacings are then reduced to about 0.8 km (0.5 mi) spacing near the interchanges. Dynamic message signs are located upstream of major detour points on the freeway, and on the entrance ramps to the freeway. When an incident or work zone activity requires one or more lanes to be closed anywhere within the limits of TransGuide operations, predefined information dissemination plans are called from the computer database and implemented on the lane control signals and dynamic message signs. The actual plan (termed a scenario) selected is based on the expected duration of the incident, the time of day, and the number of lanes blocked.

A scenario consists of changes in up to three sets of lane control signals upstream of the incident. The two sets of signals closest to the incident and over the lane(s) that are blocked display a red X, and the open lanes display a green downward arrow. At the third signal location (the farthest upstream), signals over the lanes that are blocked display a downward diagonal yellow arrow to the left or the right. These diagonal arrows are not currently specified in the MUTCD as acceptable lane control signal symbols.[5] However, they are utilized extensively in Canada and other countries for this purpose. Also, recent research suggests that diagonal arrows may be better understood as a transition symbol than the yellow X that is currently allowed by the MUTCD for this purpose.[33]

Effects of Information

Recent evaluation of the TransGuide system indicates that drivers are obtaining and responding to the overall dynamic information system. For example, the number of drivers reportedly diverting to alternative routes in response to available information (primarily radio traffic reports before TransGuide, reports plus the dynamic message signs and lane control signals after TransGuide implementation) has increased from 33 to 80 percent.[34] However, the effect of the dynamic information system upon driver lane usage upstream of the incident has not been as well defined to date. Limited data suggests that a small portion of motorists do move from the closed lane to the open lane farther upstream of an incident when the lane control signals indicate a downstream blockage.[35] Survey results also suggest that the signals are seen and correctly interpreted by most drivers.

Operators of the TransGuide system have indicated the importance of providing accurate and useable information on the dynamic message signs in conjunction with the lane control signal displays in order to obtain appropriate driver responses. In one instance, operators closed two lanes where glass had been dropped on the road. They indicated “debris in left two lanes” and activated the appropriate lane control signal sequences, but saw drivers continuing to utilize the left lanes for travel (through the area of glass). They then switched the sign to read “glass in left two lanes.” Nearly all drivers vacated those lanes and remained out of those lanes until the lane control signals indicated that all lanes were open.[35]

SR 91, ORANGE COUNTY, CALIFORNIA CONGESTION PRICING

Recent construction of four express lanes (two each way) in the median right-of-way of State Route 91 in Orange County, California, was a $126 million undertaking that was accomplished through an innovative franchising agreement between the public sector and private enterprise (California Private Transportation Company, CPTC). To pay for construction, CPTC will collect tolls from the facility for the next 35 years, after which it will revert to State operation. The project has received extensive publicity as the first implementation of congestion pricing in the U.S. Users of the facility will be charged a toll ranging from $0.25 to $2.50, depending upon the time of the day and traffic demand levels.

A key component of the SR 91 system is the sole use of electronic toll collection equipment. No toll booths are present along the facility. Instead, only overhead radio receivers, which read small transponders that motorists place in their windshields are used to approve approaching vehicles to enter the express lanes and make appropriate toll charges to the motorist's account. Toll violators may be ticketed by the California Highway Patrol or via the mail and receive a $100 to $300 fine.[36] The revenues from toll collections are used to pay for the costs of enforcement. At the same time, the facility is designed to encourage carpooling. Buses, vehicles with three or more occupants, motorcycles, and vehicles designated as emission-free can use the express lanes for free (although they must still have a transponder in their vehicle).

Because of the system's recent initiation, experiences to date are limited. However, it was reported that within just a few weeks of its opening, the toll lanes and nearby Metrolink trains drew so many commuters that congestion levels on the mixed-use lanes on SR 91 dropped to levels not seen in 15 years.[36] After only a few months of operation, over 50,000 transponders are in use in the SR 91 corridor.

4.6 REFERENCES AND SUGGESTED READINGS

REFERENCES

1. *Traffic Engineering Handbook.* Institute of Transportation Engineers, Washington, DC, 1992.

2. *Manual of Uniform Traffic Control Devices*. FHWA, U.S. Department of Transportation, Washington, DC, 1988.

3. Warren, D.L. Speed Zoning and Control. In *Synthesis of Safety Research Related to Traffic Control and Roadway Elements, Volume 2*. Report No. FHWA-S-82-233. FHWA, U.S. Department of Transportation, Washington, DC, December 1982.

4. Dudek, C.L., and Ullman, G.L.. *Freeway Corridor Management*. NCHRP Synthesis of Highway Practice No. 177. TRB, National Research Council, Washington, DC, March 1992.

5. Middleton, D., Fitzpatrick, K.,Jasek, D., and Woods, D. *Truck Accident Countermeasures on Urban Freeways*. FHWA-RD-92-059. FHWA, U.S. Department of Transportation, Washington, DC, May 1994.

6. *Building the ITI: Putting the National Architecture into Action*. FHWA, U.S. Department of Transportation, Washington, DC, April 1996.

7. Abram, C.M., and Wang, J.J. *Planning and Scheduling Work Zone Traffic Control*. Report No. FHWA-IP-81-6. FHWA, U.S. Department of Transportation, Washington, DC, October 1981.

8. Arrillaga, B. U.S. Experiences with Congestion Pricing. *ITE Journal*, Volume 63, No. 12, December 1993, pp. 39-43.

9. *Implementation Plan Guidance (23 CFR 655.409)*. Federal-Aid Policy Guide, FHWA, U.S. Department of Transportation, Washington, DC, December 7, 1994, Transmittal 12.

10. *Effects of Lane Restrictions for Trucks.* Draft Report. FHWA, U.S. Department of Transportation, Washington, DC, June 1986.

11. Stokes, R., and McCasland, W.R. Truck Operations and Regulations on Urban Freeways in Texas. *ITE Journal*, Vol. 56, No.2, pp 17-21, February 1986.

12. *The Urban Transportation Monitor*. Volume 7, Number 13, July 9, 1993.

13. Ullman, G.L., Tallamraju, S.S., and Trout, N.D. *Visibility and Spacing of Lane Control Signals for Freeway Traffic Management*. Report No. FHWA/TX-95/1498-1. Texas Transportation Institute, College Station, TX, November 1994.

14. Curren, J.E., *Use of Shoulders and Narrow Lanes to Increase Freeway Capacity*. NCHRP Report 369. TRB, National Research Council, Washington, DC, 1995.

15. MacCalden, M. S. A Traffic Management System for the San Francisco-Oakland Bay Bridge. ITE Journal, Vol 54, No. 5, May 1984, pp 46-51.

16. Hanscom, F.R. *Effectiveness of Changeable Message Sign Displays in Advance of High-Speed Freeway Lane Closures.* NCHRP Report No. 235. TRB, National Research Council, Washington, DC, September 1981.

17. Dudek, C. L. *Guidelines on the Use of Changeable Message Signs.* Report No. FHWA-TS-90-043. FHWA, U.S. Department of Transportation, Washington, DC, May 1991.

18. Carlson, G., Lau, R., and Bieniek, J. Freeway to Freeway Ramp Metering Twin Cities Metro Area. *Compendium of Technical Papers*, 66th Annual Meeting of the Institute of Transportation Engineers, Minneapolis, Minnesota, September 1996, pp 131-137.

19. *A Toolbox For Alleviating Traffic Congestion*. Publication No. IR-054A. Institute of Transportation Engineers, Washington, DC, 1989.

20. Hartje, R.L. Toll Roads in California. *ITE Journal*, Vol. 61, No. 6, June 1991, pp 17-21.

21 Innovative Financing Makes California Toll Roads a Reality. *American City & County*, Vol 108, No. 11, October 1993. pp. 50

22. Edelstein, R., and Srkal, M. Congestion Pricing. *ITE* Journal, Vol. 61, No. 2, February 1991, pp 15-18.

23. Decorla-Souza, P. Congestion Pricing: Issues and Opportunities. *ITE Journal*, Vol. 63, No. 12, December 1993, pp 27-32.

24. Rao, B. Investigations into Achievable Capacities and Stream Stability with Coordinated Intelligent Vehicles. Presented at 70th Annual Meeting of the Transportation Research Board, Washington, DC, 1993.

25. Smith, E.R. Anticipating the Far Reaching Impact of the Automated Highway. *Proceedings of the 1996 Annual Meeting of ITS America*, Houston, TX, pp 619-628, 1996.

26. Stevens, W.B. Use of System Characteristics to Define Concepts for Automated Highway Systems. *Transportation Research Record No. 1453*. TRB, National Research Council, Washington, DC, 1995, pp 36-45.

27. Bender, J.G. Overview of Systems Studies of Automated Highway Systems. *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 1, February 1991.

28. Johns, K.B. Toll Roads--What's the Difference? *Compendium of Technical Papers*, 54th Annual Meeting of the Institute of Transportation Engineers, September 1984, pp 124-127.

29. Poole, R.W. Jr. Introducing Congestion Pricing on a New Toll Road. *Transportation Research Record No. 1359*. TRB, National Research Council, Washington, DC, 1992, pp 19-25.

30. Hudders, N.A. Strategies for Implementing Automatic Vehicle Identification for Congestion Pricing. In *Graduate Student Papers on Advanced Traffic Management Systems*. Report No. SWUTC/91/712410-2. Southwest Region University Transportation Center, Texas Transportation Institute, College Station, TX, August 1992, pp D1-D18.

31. Richards, S.H., and Dudek, C.L. Selection of Work Zone Channelizing Devices Using the Value Engineering Approach. *Transportation Research Record No. 1035*. TRB, National Research Council, Washington, DC, 1985, pp 78-82.

32. Dudek, C L. *Changeable Message Signs*. NCHRP Synthesis of Highway Practice 61. TRB, National Research Council, Washington, DC, July 1979.

33. Wohlschlaeger, S.D., Ullman, G.L., and Dudek, C.L. Motorist interpretation of X and Yellow Diagonal Arrow in Freeway Lane Control Signal Array. *Transportation Research Record No. 1495*. TRB, National Research Council, Washington, DC, 1995, pp 9-16.

34. Henk, R.H., Molina, M.E., and Irwin, P.L. Before-and-After Analysis of the San Antonio TransGuide System. *Compendium of Technical Papers*, 66th Annual Meeting of the Institute of Transportation Engineers, Minneapolis, MN, September 1996, pp 518-524.

35. Ullman, G.L., Parma, K.D., Peoples, M.D., and Trout, N.D. *Yellow Transition Lane Control Signal Symbols for Freeway Traffic Management*. Report No. FHWA/TX-97/1498-2 (Draft). Texas Transportation Institute, College Station, TX, September 1996.

36. Finch, G. Congestion Pricing: Reducing Traffic Jams Through Economics. *Public Roads*. FHWA, U.S. Department of Transportation, Washington, DC, Autumn 1996, pp 4-5.

SUGGESTED READINGS

- *A Toolbox for Alleviating Traffic Congestion*. Publication No. IR-054A. Institute of Transportation Engineers, Washington, D.C., 1989.

- *Traffic Engineering Handbook*. Institute of Transportation Engineers, Washington, D.C., 1992.

- Dudek, C.L. and G.L. Ullman. *Freeway Corridor Management*. NCHRP Synthesis of Highway Practice No. 177. TRB, National Research Council, Washington, D.C. March 1992.

- Dudek, C. L. *Guidelines on the Use of Changeable Message Signs.* FHWA-TS-90-043. Federal Highway Administration. Washington, DC. May 1991.

- Dudek, Cl. L., and R. D. Huchingson. *Manual on Real-Time Motorist Information Displays.* Report No. FHWA-IP-86-16. August 1986.

- McConnell, C.H. Compatibility of Automated Highway Systems (AHS) and Freeway Management Systems (FMS). *Compendium of Technical Papers*, 64th Annual Meeting of the Institute of Transportation Engineers, Dallas, Texas, October 1994, pp. 15-19.

MODULE 5. RAMP CONTROL

TABLE OF CONTENTS

5.1 INTRODUCTION 5-5
DEFINITION OF RAMP CONTROL 5-5
APPLICATION OF RAMP CONTROL 5-6
Entrance Ramp Metering 5-6
Entrance Ramp Closure 5-6
Exit Ramp Closure 5-6
Systemwide Ramp Control 5-6
RELATION TO OTHER FREEWAY MANAGEMENT FUNCTIONS 5-6
Surveillance 5-7
Vehicle Detection 5-7
Closed-Circuit Television 5-7
Environmental Sensors 5-7
HOV Treatments 5-7
Information Dissemination 5-7
Communication 5-7
Control Center 5-7
BENEFITS OF RAMP CONTROL 5-8
Improved System Operation 5-8
Improved Safety 5-8
Reduced Vehicle Operating Expense 5-8
Means for Positive Freeway Traffic Control/Management 5-8
Reduction in Vehicle Emissions and Fossil Fuel Consumption 5-9
Coordination With Other Corridor Management Elements 5-9
Promotion of Multimodal Operation 5-9
MODULE OBJECTIVE 5-9
MODULE SCOPE 5-9

5.2 DECISION PROCESS 5-9
PROBLEM IDENTIFICATION 5-10
Level of Service / Capacity Deficiency / Bottlenecks 5-10
Vehicle Crash Experience 5-10
Inventory of Infrastructure 5-11
Freeway System 5-11
Existing Freeway Management Systems 5-11
Existing Ramp Control Systems 5-12
Other Relevant Field Systems 5-12
Inventory of Traffic Characteristics 5-12
Traffic Composition 5-12
Traffic Flow 5-12
Other Traffic Parameters 5-12
Temporal Variations 5-13

Ramp Geometric Limitations 5-13
Cross Streets 5-14
Service Roads 5-14
Summary of Problem Definition 5-14
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 5-14
Relation to Other Agencies 5-14
City/County Traffic Operations Agencies 5-14
Enforcement Agencies 5-14
Emergency Management Agencies 5-15
Public Transportation Agencies 5-15
Relationship to Elected Official / Political Environment 5-15
Importance of Enforcement / Judicial System 5-15
Relationship With Media 5-15
ESTABLISHING GOALS AND OBJECTIVES 5-16
ESTABLISH PERFORMANCE CRITERIA / MEASURES OF EFFECTIVENESS 5-16
DEFINE FUNCTIONAL REQUIREMENTS 5-17
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 5-17

5.3 TECHNIQUES AND TECHNOLOGIES 5-18
ENTRANCE RAMP CONTROL 5-18
Ramp Closure 5-18
Application 5-19
Methods 5-19
Ramp Metering 5-20
Metering Rates 5-20
Congestion. 5-20
Safety. 5-22
Pretimed Metering 5-22
System Components 5-23
System Operation 5-24
Single-entry metering 5-24
Platoon metering 5-24
Tandem Metering. 5-24
Two-abreast Metering 5-25
Traffic-Responsive Metering 5-26
Fundamental Traffic Flow Relationships 5-26
Basic Strategy 5-26
Demand-Capacity Control 5-28
Occupancy Control 5-28
System Components 5-29
System Operation 5-30
Gap-Acceptance Merge Control 5-31
Basic Concepts 5-32
System Components 5-32

System Operation 5-33
System Ramp Control 5-35
System Pretimed Metering 5-35
Example 1 5-36
Example 2 5-38
Example 3 5-38
Practical Considerations 5-40
Example 4 5-41
System Traffic-Responsive Metering 5-43
System Operation 5-43
Metering Rates 5-43
System vs. Independent Ramp Control 5-43
Controller Interconnection 5-43
Incremental Benefits of Various Levels of Control 5-44
Variations in the Ratio of Mainline to Entrance Ramp Demand 5-44
Variations in Overall Traffic Demand Pattern 5-44
Mainline Capacity Reductions 5-44
EXIT RAMP CONTROL 5-45
EMERGING TECHNOLOGIES 5-45

5.4 LESSONS LEARNED 5-46
IMPLEMENTATION 5-47
Public Relations 5-47
Media Relations 5-47
Implementation Strategies 5-48
Implementation Summary 5-48
OPERATIONS AND MAINTENANCE 5-49
DIVERSION OF TRAFFIC 5-50
ENFORCEMENT 5-51
EQUITY 5-51

5.5 EXAMPLES IN RAMP CONTROL 5-52
HISTORY OF RAMP CONTROL 5-52
ENTRANCE RAMP METERING CASE STUDIES 5-52
Portland, Oregon 5-52
Minneapolis/St. Paul, Minnesota 5-53
Seattle, Washington 5-54
Denver, Colorado 5-55
Detroit, Michigan 5-56
Austin, Texas 5-56
Long Island, New York 5-56
San Diego, California 5-57
SUMMARY OF RAMP METERING BENEFITS 5-57

5.6 REFERENCES 5-59

MODULE 5. RAMP CONTROL

Figure 5-1. Ramp Meter with HOV By-Pass in Minneapolis, MN.

5.1 INTRODUCTION

The geometric design of a freeway ramp (width, curvature, vertical alignment, etc.) can have a positive or negative influence on both the operation of the ramp itself and on the freeway at, or upstream of, the merge point. Freeway design standards generally address those considerations. Ramp control, on the other hand, seeks to regulate the flow of vehicles at freeway ramps in order to achieve some operational goal such as balancing demand and capacity or enhancing safety. Other than freeway-to-freeway interchanges, freeway ramps represent the only opportunity for motor vehicles to legally enter or leave a freeway facility and, therefore, the only point at which positive control can be exercised. Freeway ramp control systems have been in operation at various locations throughout the country since the early sixties. It is estimated that ramp control systems are operated in over 20 geographical areas at present, with individual metered ramps numbering over 2300.[1]

Most ramp control systems have been proven to be successful in terms of reduced delay and travel time (and the concomitant reductions in fuel consumption and vehicle pollutants, and in accident reduction.) They are more effective when they are part of an integrated transportation management plan that incorporates other systems as described in other modules of this document. Deployment of ramp control systems has been somewhat limited due to some public resistance to being stopped on a freeway ramp for no readily apparent reason, although the ramp metering rate may reflect a downstream bottleneck such as an incident.

DEFINITION OF RAMP CONTROL

Freeway ramp control is the application of control devices such as traffic signals, signing, and gates to regulate the number of

vehicles entering or leaving the freeway, in order to achieve some operational objective. Typically, the objective will be to balance demand and capacity of the freeway in order to maintain optimum freeway operation and prevent operational breakdowns. Ramp metering may also be applied for safety considerations where certain geometric inadequacies or other constraints exist.

APPLICATION OF RAMP CONTROL

The primary application of ramp control, commonly known as ramp metering, has been on freeway entrance ramps. However, ramp control has been applied in other situations as well.

Entrance Ramp Metering

Metering on entrance ramps involves determination of a metering rate (typically "4 to 15 vehicles per minute" are minimum and maximum rates for single lane metering) according to some criteria such as measured freeway flow rates, speeds, or occupancies upstream and downstream of the entrance ramp. The rates may be fixed (pre-timed) for certain periods, based on historical data, or may be variable minute-by-minute (traffic responsive) based on measured traffic parameters. The entry of vehicles at that rate is regulated by one or more traffic signals beside the ramp at driver's-eye height. Vehicle sensors may be located at points along the ramp to signal the blockage of the merge area or backing of the ramp queue into a cross street.

Entrance Ramp Closure

Typically lower metering rates (say 2 to 4 vehicles per minute) over a sustained period of time are not acceptable to drivers, and they will tend to disregard the signal. In the extreme case where the metering rates must be sustained at lower levels, it may be necessary to physically close the ramp with automatic gates or manually placed barriers. Obviously, this extreme case may cause negative public reactions and should be applied only after considerable planning and a public information program.

Exit Ramp Closure

Metering of exit ramps is obviously not appropriate but closure with automatic gates or manually placed barriers, with adequate freeway warning signs, may be used to accomplish certain operational objectives. For example, if the exit ramp terminus at the cross street has such inadequate capacity that exit ramps queue onto the freeway, the ramp may be closed to encourage drivers to exit upstream or downstream where more capacity is available.

Systemwide Ramp Control

Although individual ramps may be metered or closed for specific reasons, ramp control is most effective when ramps are metered in an integrated system manner. Individual metering rates are determined by conditions over a larger portion of the freeway, not just in the immediate area of the ramp. Although local controllers may suffice in individual ramp metering as described above, system-wide control requires a central or distributed control system master with control algorithms and interconnection by some communications media.

RELATION TO OTHER FREEWAY MANAGEMENT FUNCTIONS

Ramp control is closely related to other infrastructure elements in a freeway management system. The widespread, widely embraced Intelligent Transportation Systems (ITS) movement has further emphasized the benefits of integrated system elements. Other modules in this handbook

describe specific subsystems of a freeway management system. The following paragraphs briefly describe the relationship of those elements to freeway ramp control.

Surveillance

The surveillance subsystem includes various techniques for determination of freeway and ramp operating conditions that may have an influence on metering rates or operational overrides. Specific information on surveillance technology can be found in **Module 3** of this handbook. The paragraphs below provide a description of the types of surveillance used in conjunction with ramp control.

Vehicle Detection

Vehicle sensors located on the freeway can serve multiple purposes if located correctly during the design and construction phase. Detectors located in the freeway lanes generally have the purpose of input to incident detection algorithms and for system operation evaluation. Freeway detectors can also be used as input data in determining metering rates in traffic responsive operations. Counting detectors located on entrance and exit ramps serve as input and output data in defining a closed system operation for estimating average delay in the system.

Closed-Circuit Television

Closed-circuit television (CCTV) are used to detect and verify incidents in the overall surveillance subsystem. Cameras can also be used to fine tune and monitor operation of individual metered ramps, precluding the necessity for on-site field observation.

Environmental Sensors

Due to grades on ramps it is often necessary to adjust ramp metering rates or terminate operation during extreme weather conditions such as icy or extremely wet roadway surfaces. Environmental sensors will give early warning when such conditions exist.

HOV Treatments

Preferential treatment of high-occupancy vehicles at entrance ramps has been used successfully in several locations on entrance ramps. These systems have primarily involved a separate lane to bypass the ramp signal, and single occupant vehicle queue.

Information Dissemination

Notification of travelers of ramp closures can be effected by either pre-trip information dissemination devices such as kiosks, Web site, and Community Access Television (CATV), or by on-road devices such as variable message signs or highway advisory radio. Other operational changes in ramp operations that may be of interest or assistance to travelers may also be communicated.

Communication

Unless the controlled ramps are isolated and operate in a nonsystem mode, the communication subsystem must accommodate for the control, detection, and signal hardware.

Control Center

While ramp control systems generally have the capability to operate in an isolated manner without supervision from a central or distributed master, most are interfaced to a central management system through the communication system.

BENEFITS OF RAMP CONTROL

Positive benefits of ramp control have been documented widely and can be found in the general literature.[1] Benefits have been most commonly reported in typical measurable traffic operations parameters such as reduced delay and travel time, increased throughput and operating speeds, and reduced accident experience. Other benefits less easily quantified may also accrue from ramp control. The case studies in the following section summarize some reported benefits of ramp control. The following paragraphs describe typical benefits, both quantifiable and less easily quantified.

Improved System Operation

Freeway traffic operating characteristics that can be expected to be influenced by ramp control systems are: speed, travel time, and delay. Typically, freeway operation has been described as a series of relationships between volume, speed, and density (or occupancy). The general objective of most freeway management systems is to optimize throughput while maintaining freeway operation in the non-congested area of the curve. By controlling the number of vehicles entering the freeway based on available downstream capacity to accommodate upstream freeway vehicles and entering ramp vehicles, freeway operation is enhanced. In another scenario, the objective may be to maintain some target level of service (as indicated by speed.) Again, by controlling the rate at which vehicles are metered onto the freeway, a target operating condition is maintained. Improvements on the freeway must be weighed against ramp delays and travel times which may be increased for travelers who choose to divert to other facilities.

Improved Safety

Freeway ramp control can effect decreased vehicle crash experience on both the ramp (and merging area) and on the freeway. By breaking up platoons of vehicles, which may enter the ramp from discharge at an adjacent intersection or traffic generator, the incidence of rear end vehicle crashes is decreased in the merging area, where multiple vehicles compete for gaps. Vehicle crashes on the freeway are also reduced as the merge becomes smoother, and freeway drivers in the outside (merging) lane are less likely to have to brake abruptly or make lane-change maneuvers. Finally, in system-wide operation the overall freeway is maintained in a more stable, uniform operational mode and vehicle crashes resulting from stop and go operations are reduced.

Reduced Vehicle Operating Expense

Improved system operation has the direct and quantifiable result of reduced vehicle operating expense. Reductions in the number of stops and speed changes translate to related reductions in vehicle operating expense. The most significant savings are related to the reduction of vehicle crashes.

Means for Positive Freeway Traffic Control/Management

There are few opportunities to actively “control” freeway traffic on a routine basis. Obviously, police officers working freeway incidents control freeway traffic, but not on an everyday basis at the same location. Passive control, such as suggestions or advisories via pre-trip planning information sources or en route signing, may either be followed or ignored. Ramp control offers a means to regulate or control freeway bound vehicles.

Reduction in Vehicle Emissions and Fossil Fuel Consumption

The direct correlation between improved traffic operations and the reduction of fuel consumption and vehicle emissions is well-known. Reductions in delay and numbers of stops, together with the maintaining of more uniform speeds as described previously will, in virtually every situation, result in a similar reduction in fuel consumption and vehicle pollutants. An exception might be where speeds are in higher ranges than is typically experienced during peak periods on metropolitan freeways.

Coordination With Other Corridor Management Elements

Intelligent Transportation Systems (ITS) defines certain core infrastructure elements known collectively as an intelligent transportation infrastructure. The importance of the interrelationship of the various subsystems applies to ramp control as a subsystem of Advanced Transportation Management Systems (ATMS) as well. Examples include the following:

- Ramp metering systems should be coordinated with surface street traffic signals to account for spill back of ramp queues.
- Information on ramp closures may be communicated by off-freeway information devices.
- High Occupancy Vehicle programs may involve special treatment of HOV at entrance ramps.
- Special ramp operating procedures may be instituted during incident conditions.

Promotion of Multimodal Operation

By giving preferential treatment to High Occupancy Vehicles at entrance ramps, the ramp control subsystem can promote travel mode shifts and reduction of single occupancy vehicles.

MODULE OBJECTIVE

The objective of this ramp control module is to provide insights into and guidelines on the issues associated with planning, designing, constructing, operating, and maintaining a ramp control subsystem in a freeway management system. This module also gives guidance to planners, designers, managers, and operators in public relations aspects of freeway ramp control.

MODULE SCOPE

The scope of this ramp control module is intended to include general guidelines as well as serving as a guide to references and other documentation that may be of benefit to planners, designers, and operators of freeway management systems. It is not intended to provide detailed design specifications or other construction documents. Typical plans, specifications, and estimates documents can usually be obtained from agencies already operating ramp control systems.

5.2 DECISION PROCESS

Freeway ramp control is one of the few direct means of controlling access to the freeway main lanes. Indirect control would include such methods as encouraging diversion to other facilities, or mode changes through communications with travelers prior to their trips or en route. However, direct limiting of access through ramp control can be effective and accepted by the driver only

if it is applied in those circumstances where traffic characteristics, demand patterns, and infrastructure are conducive to the technique.

PROBLEM IDENTIFICATION

Traffic engineers and other professionals will no doubt have an intuitive feel for where freeway operational deficiencies exist in a congested freeway environment. However, in order to address potential solutions to alleviate such problems, it will be necessary to quantify deficiencies in both time and space, i.e. during what portions of the day, and at which locations within the freeway system, are such deficiencies present. It is important to document the freeway operations from a traffic characteristics and infrastructure aspect in order to identify and define the problem as well as to provide a basis for measurement of effectiveness and to monitor for future changes. Several techniques may be used to illustrate a systemwide picture of freeway traffic characteristics, design features, capacity deficiencies, vehicle crash experience, and other features of interest, including the following:

- Schematic maps, color coded or otherwise delineated to show various levels of operation, other traffic characteristics, and crash experience at various periods of the day.
- Schematic maps, color coded or otherwise delineated to show various infrastructure characteristics.
- Spreadsheet or other tabular format.
- Descriptive write-up.
- A combination of the above items.

Level of Service / Capacity Deficiency / Bottlenecks

The *Highway Capacity Manual* provides definitive guidance in determining qualitative and quantitative pictures of freeway operations. Capacity deficiencies, or freeway bottlenecks, will be a function of traffic demand and characteristics as well as the geometrics and other design features of the roadway itself. It will be necessary to pinpoint where such deficiencies exist, and the contributing factors. Subsequent sections of this module provide a detailed listing of data required for analysis of capacity and level of service in relation to ramp control, as well as other analyses. **Module 2** provides more guidelines for capacity and level of service analyses.

Capacity and Level of Service (LOS) should be determined for existing traffic characteristics and infrastructure as well as those parameters for future conditions in some horizon or build-out year. Planned additions to the freeway section under consideration, or to alternative routes or modes, may either obviate the need for ramp control or influence its implementation schedule.

Vehicle Crash Experience

The occurrence of vehicle crashes on freeways may be attributed to a variety of factors, some of which may not be correctable by ramp control techniques. Those types of accidents most likely to be alleviated by ramp control include:

- Rearend crashes on freeway main lanes due to over-capacity operation (bottleneck conditions).
- Lane change crashes on freeway lanes due to over-capacity merging conditions.

- Lane change crashes on freeway lanes due to inadequate sight distance or to other geometric deficiencies in the merge area.

- Run-off-the-road crashes caused by drivers avoiding shock waves.

- Rearend crashes on the entrance ramp due to queuing in the merge area.

Crash records may be summarized by section of freeway, location, time of day, and type of crash to determine if ramp control has the potential to reduce collision experience.

Inventory of Infrastructure

Except in the case of an isolated, single entrance ramp location, a ramp control system is generally a subsystem of a comprehensive freeway management system. Much of the infrastructure data required for problem identification will likely be available. The following types of data should be assembled for the freeway system under consideration.

Freeway System

- **Lane Configuration.** Number and types of freeway lanes (through, weaving, acceleration, deceleration) should be determined and tabulated and/or graphically displayed.

- **Ramp Locations.** Entrance and exit ramps should be located, with link distances between ramps determined.

- **Geometrics.** Typical geometrics such as freeway lane and entrance ramp width, vertical and horizontal alignment, ramp length, ramp storage capacity, merging area, and sight distance restrictions should be determined and tabulated. Type of ramp design (loop, linear) should be noted.

- **Frontage Roads.** Presence of frontage roads and their lane configurations should be determined and tabulated.

- **Interface to Crossing Freeways.** Freeways will generally be interfaced or connected via a freeway interchange. The proximity of another freeway's connections to the entrance ramps being considered for ramp control should be noted to determine if any special measures are needed.

- **Interface to Crossing Arterials.** The relationship of entrance ramp metering to an upstream cross street is critical. If not properly considered, queuing from the ramp signal into the cross street can cause concerns to the agency responsible for arterial street operation, as well as public resentment. Type of crossing roadway, traffic control, mix of traffic, ramp storage area, and other factors should be noted for each ramp.

Existing Freeway Management Systems

Normally there will be only one agency responsible for freeway operations in a particular geographic area, but there are some situations where more than one agency may be involved. For example, a dense metropolitan area may extend into two States or a tollway operated by a toll authority may interface to a state-operated freeway. As part of the inventory, the existence of such systems should be confirmed and documented to include the following:

- Participating agencies.

- Type and location of control center facility.

- Type of control system (central, distributed, hybrid, local).
- Surveillance and detection.
- Information dissemination (pre-trip, en route).
- Communication system (medium, leased, or owned).

Existing Ramp Control Systems

In lieu of a full blown freeway management system, some entrance ramps may be metered in an isolated manner with a local, non-system controller. Inventory should include the following:

- Responsible agency.
- Type of controller.
- Surveillance and detection.
- Communication system.

Other Relevant Field Systems

Other relevant field operational systems that may have an effect on freeway operation should also be identified. Such systems would include the following:

- High occupancy vehicle lanes or ramps.
- Incident management teams.
- Accident investigation sites.
- Courtesy and motorist assistance patrols.
- Hazardous material routing and restrictions.

Inventory of Traffic Characteristics

Certain traffic and flow characteristics will influence the potential success and the design of freeway ramp control systems. Typical traffic characteristics are listed below. **Module 2** provides a more detailed description of individual traffic parameters.

Traffic Composition

The composition of the traffic stream on the freeway main lanes and the entrance ramp will influence both the type of control and the design of the system. A determination of the percentage of passenger vehicles, commercial vehicles, and transit vehicles should be made for peak periods.

Traffic Flow

Traffic volumes and traffic flow rates during peak periods will be required for capacity and level of service determinations to define the operating conditions and problem locations that might be addressed by ramp control techniques. Traffic flow data will also be used in determining metering rates and periods of operations. Traffic flow data requirements will include the following:

- Traffic volumes and flow rates, generally by 15-minute periods, on freeway lanes and entrance and exit ramps.
- Distribution of freeway vehicles by lane.
- Traffic volumes and flow rates on adjacent service roads.
- Traffic volumes and flow rates on cross streets served by the freeway ramps.

Other Traffic Parameters

Other typical traffic parameters that may be of value either in defining operating

conditions and problem locations or in developing control strategies include the following:

- **Lane Occupancy.** Defined as the percentage of time a particular sampling "spot" on the freeway is occupied, this parameter may not be economically measured until such time as a surveillance system is in place. Its primary use is in selecting metering rates, although it can identify operational problems if reasonably available. It may be derived from speed and volume data, which may be more readily available prior to system implementation. The reader is referred to **Module 2**, or the *Highway Capacity Manual*, for a discussion of the relationship of lane occupancy to freeway level of service.[2]

- **Traffic Density.** Defined as the number of vehicles per lane per mile, traffic density may be determined with aerial photos or by freeway input/output counts. The reader is referred to **Module 2** or the *Highway Capacity Manual* for a discussion of the relation of traffic density to freeway level of service (LOS).[2]

- **Speed.** Vehicle speeds are another indicator of freeway LOS and may be determined by traditional speed measurement techniques prior to system installation.

- **Vehicle Occupancy.** As opposed to *lane* occupancy, *vehicle* occupancy is generally defined as passengers per vehicle and is usually determined by manual observation. This parameter may be useful in determining the viability of preferential treatment of high occupancy vehicles (HOV) at entrance ramps.

Temporal Variations

As previously mentioned, it is important that traffic operations characteristics be collected and analyzed in incremental time periods so that ramp control operation schedules can be developed optimally. Even though the system may be traffic responsive, it may be advantageous to operate either on a predictable schedule or with limited variations in schedule. Plotting various parameters by time period in 15-minute increments will help predefine those operational periods. Although ramp control is usually associated with peak periods, plotting data over a longer period may indicate other times when ramp metering may be appropriate.

Ramp Geometric Limitations

Inventory of infrastructure elements and field observations will provide information to evaluate the physical viability of individual ramps to support ramp metering. The following physical factors should be considered:

- **Ramp Storage.** How many vehicles can reasonably be stored or queued on the ramp upstream of the metering signal without interfering with cross street traffic?

- **Ramp Width.** Is there adequate width for side-by-side metering and/or preferential HOV bypass lanes?

- **Grade**. Are ramp grades restrictive during adverse weather or for certain types of heavy vehicles?

- **Merge Area.** Does the present design facilitate a smooth merge?

Cross Streets

Limited vehicle storage for queuing at ramp signals may adversely affect operation of an upstream cross street. Presence of such conditions should be noted so that they can be considered during design of control strategies.

Service Roads

As with cross streets, limited vehicle storage for queuing at ramp signals may adversely affect operation. The type of cross street (major arterial, collector, etc.), traffic demand, presence of signals, and their operation must be considered.

Summary of Problem Definition

Traffic characteristics and demand, as well as geometric factors, are important in evaluating existing and future conditions and the potential applicability of freeway ramp control. While not all data items listed above may be available to the designers and planners, it is important to collect and assemble as much relevant data as feasible for the analysis. Many of the data items noted above may also be used during the design of the system and development of the control strategies and software.

IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING

Freeway ramp control is the primary method of managing demand once drivers have committed to use the freeway for their trip. It has been proven to be an effective means of balancing capacity and demand and reducing delay and vehicle crashes. It can also be one of the most controversial traffic control techniques. Delay at ramp signals or closing of a ramp may be considered too drastic by some drivers, and even an infringement on their rights. Such delays will be offset by overall system improvement, but this is not always apparent to the driver. Without ramp control, drivers may experience even more delay on the freeway than they would have experienced at the signal. Again, this may not be readily discernable to the driver.

In most instances, a State Transportation Department or Toll Authority will have the responsibility for operation, but not necessarily enforcement, of a ramp control system. It is important for the agency responsible for operation of the ramp control system to identify and establish relationships and communications with all agencies that may have a role in operation and enforcement so that they may be brought into the planning and design process. It is also important that the benefits of ramp control, which are realistic and measurable, be fully explained and that it not be oversold as adding capacity (such as adding a lane). It should be characterized as a means to make maximum use of available capacity by managing capacity and demand.

Relation to Other Agencies

City/County Traffic Operations Agencies

Because of the close relationship and interface between surface street traffic operation and signalization and access to and from freeway ramps, it is important to involve those agencies and build a consensus for the system at all levels, from the agency head to the operations engineers to the control system operators. To the extent possible, system goals and objectives should be developed mutually.

Enforcement Agencies

Depending on State and local ordinances or interagency agreements, State, local, or transit police may be responsible for

enforcement of ramp control devices. Compliance with ramp control signals is essential if the system is to operate efficiently. Enforcement agencies must be brought into the process early and must understand the goals and objectives of the system and the operating philosophies. The signals must be enforced, but over-enforcement can have a detrimental effect on driver attitude and, in fact, cause deterioration of operation as drivers are stopped on the freeway shoulder. Compliance with the signals must be established early and monitored to ensure that an acceptable level is maintained. A program of public information and police support is essential.

Emergency Management Agencies

Fire, police, medical, hazardous materials, motorist assistance patrols, and other agencies responsible for emergency management on the freeway system should be aware of the proposed system and be fully informed as to its operation and benefits. Any special support required of the particular agency should be solicited.

Public Transportation Agencies

Public transit agencies that access freeways via metered ramps**,** or that exit on ramps which may be closed during certain periods of the day, should be also be brought into the planning and design process at an early stage. This is particularly important where preferential treatment of high-occupancy vehicles such as buses is being considered.

Relationship to Elected Official / Political Environment

Although a support base and consensus may be built at the staff and agency level, it is important to build support with elected officials as well. As stated above, benefits of ramp control are real and measurable in the overall system, but may not be apparent to the individual driver who experiences delay at an entrance ramp or must reroute due to a ramp closure. Citizen (voter) complaints can have an adverse effect on the success of ramp control projects. System planners, designers, and operators must help those in office understand the goals, objectives, and operating characteristics of the system prior to system turn-on.

Importance of Enforcement / Judicial System

The importance of enforcement of ramp control has been previously stated. Accordingly, enforcement must be supported by the judicial system. A standard ramp traffic signal that meets the requirements of the *Manual On Uniform Traffic Control Devices* (MUTCD) is a legally enforceable device.[3] However, because ramp control systems are not as familiar as intersection signals, certain judges may be inclined to dismiss related citations. It is important to ensure that the proper laws and ordinances are in place and that judges to whom appeals of citations may be taken are informed of the system goals, objectives, and operating characteristics prior to system turn-on.

Relationship With Media

Local news media, both print and electronic, can have a profound effect on the success of ramp control systems. It is important that a media relations plan be developed to help ensure that positive support is secured. Methods for disseminating information about ramp control system include brochures, town meetings, and handouts.

As stated previously, it is important that the benefits of ramp control, which are realistic and measurable, be fully explained and that they not be oversold as adding capacity (as

in the case of adding a lane). It should be characterized as a means to make maximum use of available capacity by managing capacity and demand.

ESTABLISHING GOALS AND OBJECTIVES

Module 2 describes the process of establishing system goals and objectives. Goals and objectives of the ramp control system should complement and not conflict with overall system goals. In the rare case of stand-alone ramp control system, the goals and objectives may differ from those in an integrated system.

Typical overall system goals and objectives and how they may be supported by a ramp control system are listed below.

- **Reduced Accident Experience.** Maintaining smoother freeway flow by metering and improving merge conditions on the ramp.

- **Maintaining Acceptable Freeway Level of Service.** Metering on entrance ramps to maximize freeway flow rates within acceptable ranges.

- **Balancing Demand/Capacity in Freeway Corridor.** Metering on entrance ramps to encourage drivers to shift to other ramps or facilities with available capacity, or to change trip time.

- **Reduction of Single-Occupancy Vehicles.** Preferential treatment of car pools on entrance ramps.

- **Reduced Vehicle Delay.** Metering on entrance ramps to limit freeway flow rates within acceptable ranges.

- **Incident Management**. Closing ramps upstream of a freeway incident and increasing metering rates downstream.
- **Promotion of Multimodal Operation.** Preferential treatment of buses on entrance ramps.

- **Reduced Noise.** Smoother Traffic Flow Reduces Engine Revving.

- **Reduced Vehicle Operating Costs.** A result of smoother traffic flow and reduced stops.

- **Reduced Fuel Consumption**. A result of smoother traffic flow and reduced stops.

- **Reduced Vehicle Emissions.** A result of smoother traffic flow and reduced stops.

ESTABLISH PERFORMANCE CRITERIA / MEASURES OF EFFECTIVENESS

Performance criteria express broad goals in tangible or measurable terms. Better operation is obviously a goal to be strived for, but is difficult to measure and may have different meanings for different people. With the exception of the first goal (balancing capacity and demand), the goals and objectives listed above are tangible and measurable in readily understandable terms both before and after system turn-on. Level of service can be calculated, vehicle crash rates can be tabulated from law enforcement data bases, speed and delay studies can determine operating conditions that can be used to calculate delay, fuel consumption, and vehicle emissions. Transit records are available to establish changes in bus patronage, and field studies can measure vehicle occupancy.

DEFINE FUNCTIONAL REQUIREMENTS

Functional requirements for the ramp control subsystem are fairly straightforward and are summarized below:

- **Displays.** Signals on the ramp for vehicle drivers and advance warning signs.
- **Local Controller.** Device to receive and store vehicle detector information and operate signals according to internal logic or according to a central supervisory system.
- **Vehicle Detectors.** Devices to measure conditions on the freeway and ramp.
- **Control Logic.** Programs residing in the local controller for non-system operation, or at a central system processor for system operation.
- **Communications.** Leased or owned communication link between field location and central management site for data and control command transmission.
- **Central Control System.** Computer, peripherals, and operator interface devices.

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

The June 1996 ITS Architecture Executive Summary states (italics indicate adaptation to ramp control systems):

> The National ITS Architecture provides a common structure for the design of intelligent transportation systems. It is not a system design concept. What it does is define the framework around which multiple design approaches may be developed, each one meeting the needs of the user, while maintaining the benefits of a common architecture. The architecture design defines functions (*e.g., collect data from freeway and ramp detectors; and operate and monitor ramp meter signals*) that must be performed to implement a given user service, the physical entities or subsystems where these functions reside (*e.g., detectors on the freeway and ramp, signals on ramps, and local controller near the ramp*), the interfaces/information flows between the physical subsystems and the communication requirements for the information flows (*e.g., signal wirelines from the detector to the local controller and from the controller to the ramp signal; two-way wideband communication between the field controller and the central management site.)* In addition, it identifies and specifies the requirements for the standards needed to support national and regional interoperability.

In all likelihood, the functional relationships, data requirements, and information flows for a ramp control system will be dictated by the design of the broader freeway management system. However, in the case of an isolated ramp control system, the architecture will be more in the realm of typical signal design at an arterial street intersection. In any event, an open architecture (one that can be interfaced with in the future) should be employed.

5.3 TECHNIQUES AND TECHNOLOGIES

The great majority of improvements and innovations in freeway traffic management have been in the area of computing and communications technology capability. Computers are faster, have more memory and storage capability, and are more user friendly, and virtually every person involved in freeway management has ready access to a personnel computer. Development of improved communications technology has paralleled development of the more capable computers. Broad band fiber optic cable, which accommodates both high speed digital data and video, has become the standard in most freeway management systems, rather than twisted-pair and coaxial cable for hub-to-hub transmission. Wireless technology (such as cellular, microwave, packet radio, and other media) has provided a means for quick implementation until the more capital intensive construction of fiber can be funded. Many systems operate with a hybrid communication system that combines multiple media including leased telephone lines and fiber cable. The freeway management techniques and strategies documented in the 1983 *Freeway Management Handbook* were much the same as those documented in the 1985 *Traffic Control Systems Handbook*, although newer technologies were described.[5,6]

The 1996 *Traffic Control Systems Handbook* documented further developments in computing and communications hardware which had application in freeway management.[7] Other modules of this handbook specifically address hardware and software that have application in the freeway management arena.

The basic freeway ramp control techniques have not changed appreciably. Field displays and control strategies such as pretimed metering, traffic responsive metering, and system metering algorithms are still valid but with the increased computing and data transmission techniques, those algorithms can operate faster and virtually in real-time leading to more efficient control and evaluation. The techniques described below have been drawn from previous handbooks and updated as necessary to reflect changing techniques.

ENTRANCE RAMP CONTROL

Ramp Closure

Entrance ramp closure is a seldom-used technique except on a short term basis, and is included here for information purposes, and should not be considered comparable to other ramp control techniques. The closure of an entrance ramp during peak traffic conditions is the simplest and most positive form of entrance ramp control. It is also the most restrictive. Therefore, it is usually the least popular and it is also subject to considerable public opposition. However, it has been used successfully as part of a system in a number of cities in the United States and Japan (e.g., Houston, Los Angeles, San Antonio, and Fort Worth, and Osaka and Tokyo, Japan). Closure has also been effectively used in single spot improvements at entrance ramp applications, such as on freeways in Beaumont and Corpus Christi.[6] Closure may be the appropriate measure where an entrance ramp introduces serious weaving problems. Although this type of entrance ramp control can provide the same operational benefits to freeway traffic as the other types, it lacks flexibility. Consequently, if applied inappropriately, it can result in underutilizing freeway capacity, with the consequent overloading of alternate routes.

Application

Because of its limitations, entrance ramp closure should not be considered except under the following circumstances:

- Adequate storage is not available at the entrance ramp to prevent queues of vehicles waiting to enter the freeway from interfering with surface street traffic. The closure of the entrance ramp would eliminate the storage problem.

- Traffic demand on the freeway immediately upstream of the entrance ramp is at capacity, and an alternate route with adequate capacity is available. The closure of the entrance ramp would prevent demand from exceeding capacity on the freeway section immediately downstream from it, and it would divert the traffic demand at the ramp to an alternate route. Even if the upstream traffic demand is less than downstream capacity, the rate at which traffic could be allowed to enter the freeway might be so low that it would not be possible to control the entrance of ramp traffic without a large number of violations. In this case, it would be more practical to close the ramp in order to prevent congestion on the freeway.

With regard to the second circumstance, it should be noted that the required demand-capacity relationship could occur because of nonrecurrent congestion as well as because of recurrent conditions. Therefore, closure might be used as a response to incidents on the freeway, as is done in Japan.[8,9]

Ramps may be closed on a temporary basis, on a scheduled basis, or permanently.

- **Temporary Closures.** Entrance ramps may be closed temporarily in response to maintenance or construction activities either on the freeway or the adjacent frontage road or surface street. It is not uncommon for a ramp to be closed by police during management of a downstream incident.

- **Variable Schedule.** Because of extreme recurring downstream capacity deficiencies, ramps may be closed during certain peak periods and open at off-peak times.

- **Permanent Closure.** A ramp may be closed on a permanent basis due to changes in the freeway systems or demand patterns. Concrete barriers or other physical constraints are recommended.

Methods

Methods of entrance ramp closure that have been used in current systems include the following:

- Manually placed barriers such as cross bucks, barrels, or cones.

- Automated barriers such as those used at railroad crossings.

- Signing.

Experience in Detroit and Los Angeles has indicated that signs alone cannot effect a positive entrance ramp closure.[10,11] Automated barriers enable an entrance to be closed and opened automatically, which tends to increase the flexibility of closure as a means of control. Since manual placement of barriers is labor intensive, this approach is best suited for short-term or trial control projects.

Ramp Metering

Metering is a method of regulating traffic flow. When applied as a form of entrance ramp control, metering is used to limit the rate at which traffic can enter a freeway. Maximum practical single lane rate is generally at 900 vph, with practical minimum of 240 vph. When the metering rate is not directly influenced by mainline traffic conditions, the control is referred to as "pretimed metering." This does not, however, necessarily imply the absence of vehicle detectors.

Metering Rates

The calculation of metering rates depends on the purpose for which the metering is being used. Normally, metering is used either to eliminate congestion on the freeway or to improve the safety of the merging operation as follows:

Congestion. If the metering system is intended to eliminate or reduce congestion, demand must be kept at less than capacity. Therefore, the calculation of the metering rate at a ramp would be based on the relationship between upstream demand, downstream capacity, and the volume of traffic desiring to enter the freeway at the ramp. Downstream capacity may be determined by the merging capacity at the ramp or by the capacity of the freeway section downstream. Of course, if the sum of upstream demand and ramp demand is less than or equal to downstream capacity, metering is not needed to prevent congestion. On the other hand, if the upstream freeway demand alone is greater than downstream capacity, metering at the ramp would not eliminate congestion. Otherwise, the desired metering rate is set equal to the difference between upstream demand and downstream capacity (assuming upstream demand is less than downstream capacity and the metering rate does not create excessive queuing).

For example, in the situation shown in figure 5-2, the upstream demand is 5,100 vph, the downstream capacity is 5,400 vph, and the ramp demand is 500 vph. Since the total demand (5,600 vph) is greater than the downstream capacity, ramp metering might be a feasible solution. Therefore, if a metering rate equal to the difference between upstream demand and downstream capacity (300 vph) were used, the freeway would be able to accommodate the upstream demand and maintain uncongested flow while also handling 300 vph of the ramp demand.

However, the ultimate test of the feasibility of ramp metering at a rate of 300 vph would involve consideration of the following questions:

- Is adequate additional capacity available in the corridor for the 200 vph that are likely to be diverted? And, if so, is it likely that the 200 vph would utilize that extra corridor capacity? If not, capacity would have to be added to the corridor and/or made more attractive for this number of vehicles per hour to be diverted. Otherwise, ramp metering would solve only the problem on the freeway.

- Is adequate storage available at the ramp to accommodate the queue of vehicles that would have to wait at the ramp before entering the freeway? If adequate storage could not be provided at the ramp, alternatives to be considered would be closure of the ramp, or metering at other ramps upstream to reduce upstream demand, which would in turn permit a higher metering rate and require less storage at the ramp.

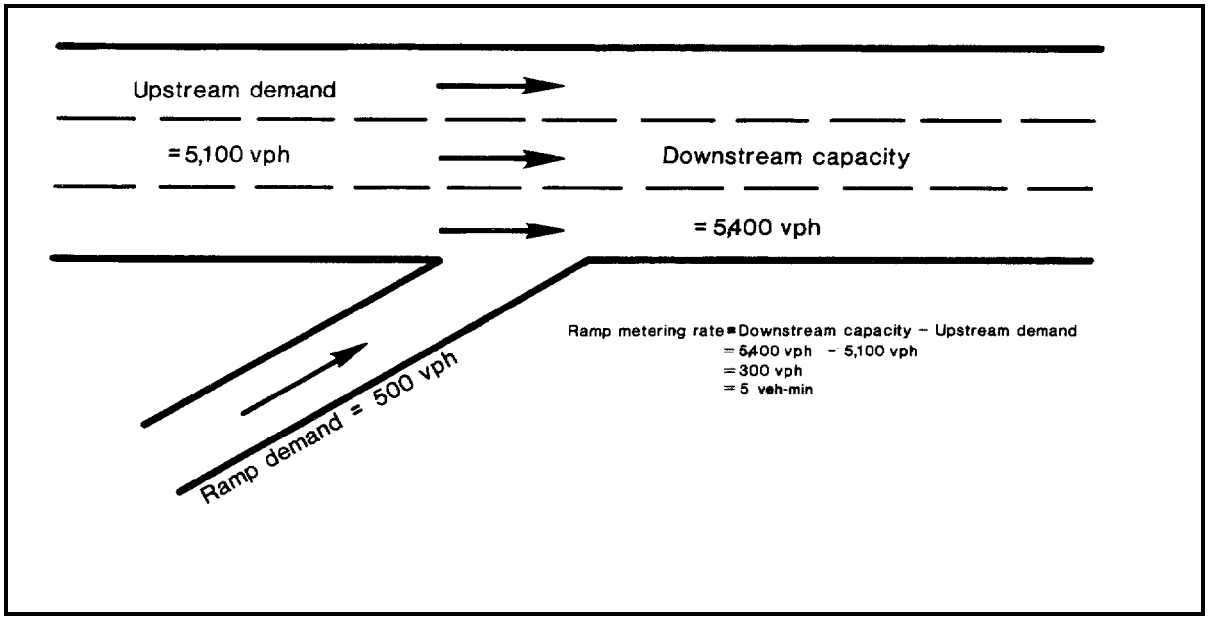


Figure 5-2. Example of Pretimed Entrance Ramp Metering Rate Calculation.

- Is the specified metering rate (300 vph) too restrictive? If so, consideration should be given to closing or metering other ramps upstream to reduce upstream demand, which would permit a higher metering rate at the ramp. However, metering other ramps upstream might lead to the underutilization of the freeway.

Signal Timing. Given that a metering rate has been set, specific signal timing parameters must be determined. (See figure 5-3 for general detector positioning.)

- **Signal Cycle.** Cycle is the inverse of the metering rate or forced headway between released vehicles. For example, a 10-second metering rate results in a 6-second cycle or headway between released vehicles.

- **Minimum Green.** The green interval is just long enough to allow one vehicle to cross the stop line at the signal, usually 0.5 to 1.3 seconds. Some systems use the checkout detector (a pulse detector) to signal the controller to terminate green.

- **Clearance Interval.** If a yellow clearance interval is used, it is typically 0.7 to 1.0 seconds, making the total green plus yellow 1.2 to 2.3 seconds. If no yellow clearance is used, the 1.0 second clearance is added to the minimum green to ensure safe clearance.

- **Red Interval.** The red interval is, then, the difference between the total cycle length and the green plus yellow or the green only interval.

- **Queue Detector.** If the queue defector (a presence loop) is occupied more than some maximum length of time (say 2.0 seconds) indicating an excessive queue, the controller may increase the metering rate in order to reduce or clear the queue.

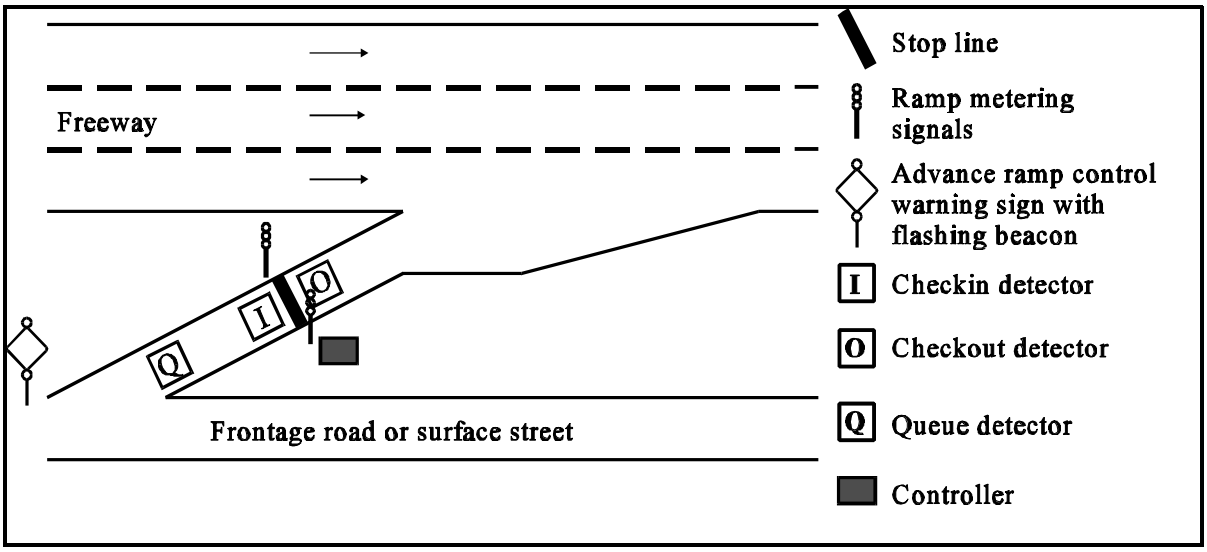


Figure 5-3. Pretimed Ramp Metering Layout.

- **Merge Detector.** If the merge detector (a presence loop) is occupied more than some maximum length of time (say 3.0 seconds), indicating the merging area is blocked, the controller may hold the ramp signal in red to avoid stacking on the ramp.

The settings given above are typical but should be fine tuned in the field to account for unique geometrics, grades, driver characteristics, vehicle mix, and other factors.

Safety. If metering is to be used primarily as a means of improving the safety of the merging operation, then the metering rate is simply set at a maximum consistent with merging conditions at the particular ramp. The primary safety problem of the merging operation is incidence of rearend and lane-change collisions caused by platoons of vehicles on the ramp competing for gaps in the freeway traffic stream. Therefore, metering is used to break up these platoons and to enforce single-vehicle entry. For this to happen, the metering rate selected must ensure that each vehicle has time to merge before the following vehicle approaches the merge area. The time it takes a vehicle to merge depends on the following factors:

- Distance the vehicle is stopped from the freeway.
- Geometrics of the ramp (grade, sight distance, and length of the acceleration lane).
- Type of vehicle.
- Availability of acceptable gaps in the freeway traffic stream.

If the average time to merge is 6 seconds, the metering rate will be 600 vph or 10 vpm.

Pretimed Metering

Pretimed metering refers to a fixed metering rate that is not influenced by current mainline traffic conditions. The rate will normally be set on the basis of historical data. However, pretimed metering does not necessarily imply the absence of detectors.

System Components

Pretimed metering is the simplest form of entrance ramp metering. Typical components are shown in figure 5-3 and are described below.

- **Ramp metering signal.** Usually a standard 3-section (red-yellow-green), or 2-section (red-green) signal display that controls the ramp traffic.

- **Local controller.** Frequently a standard pretimed or Type 170 controller with capability to vary metering rates by time of day or to accept detector inputs. However, national cooperative efforts among industry, operating agencies, and FHWA have developed a design for an open architecture protocol for local controller. This controller, more properly termed a "field processor,"acts as a communication processor with extended capability for other freeway management functions such as control of variable message signs, detector processing, and closed-circuit television cameras.

- **Advance ramp control warning sign with flashing beacon.** A sign which indicates to traffic approaching the ramp that it is being metered. In California, a blank-out type "METER ON" sign is used at many installations in lieu of the flashing beacon.

- **Vehicle Detectors.** There are four types of detectors that are generally used with this type of ramp metering strategy:

  - Check in (demand) detectors.

  - Checkout (passage) detectors.

  - Queue detectors.

  - Merge detectors.

In some applications of pretimed metering a check-in detector is placed on the approach to the ramp metering signal so that the signal will remain red until a vehicle is detected at the stop line, as shown in figure 5-3.[12] When a vehicle is detected by the check-in detector, the ramp metering signal will change to green, provided the minimum red time has elapsed. With this type of operation, it is desirable to have a minimum metering rate (e.g., 3 vpm) at which the signal is set in case there is no detector actuation, because of possible detector failure or because of vehicles stopping too far back from the stop line to actuate the detector. In some cases, two detectors are used to provide redundancy to reduce the impact of detector failures.

In some systems, a checkout detector has been used to ensure single-vehicle entry. When a vehicle is permitted to pass the ramp metering signal, it is detected by the checkout detector, which is installed just beyond the stop line (usually about half a car length past it). The green interval is then terminated as soon as the vehicle is sensed by the checkout detector. In this way, the length of the green interval is made sufficient for the passage of only one vehicle.

In some pretimed metering systems, a queue detector is used to detect backing of ramp traffic into frontage roads or surface streets. The queue detector is placed at a strategic point on the ramp, or on the frontage road, in advance of the ramp metering signal. When a queue is sensed by a vehicle occupying the loop for a selectable period of time, indicating that the queue of vehicles waiting at the ramp metering signal is sufficient to interfere with traffic on the frontage road or surface streets, a higher metering rate may used to shorten the queue length. This can be self-defeating, however,

since shorter queues often attract higher demands.

A merge detector is a device that senses the presence of vehicles in the primary merging area of the ramp and freeway mainlanes. When the merge detector senses that a vehicle has stopped, blocking the merge area, the signal may be held in red for some preset maximum time in order not to clog the area and to reduce the possibility of a rear end collision. On a well designed entrance ramp with adequate acceleration and merging distance, a merge detector is not necessary or practical.

Placement of these auxiliary detectors is discussed in more detail in the subsequent section on traffic responsive metering. Figure 5-4 shows ramp metering signals and advance warning signs that have been used. Also, for a discussion of standards for various system components, refer to the recommended practice for freeway entrance ramp displays prepared by the Institute of Transportation Engineers (ITE).[13]

System Operation

In the operation of a pretimed metering system, the ramp signal operates with a constant cycle in accordance with a metering rate prescribed for the particular control period. However, timing the red, yellow, and green intervals of the cycle (many systems use ramp signals that have only red and green intervals) depends on whether the type of metering used is single-entry metering or platoon metering.

Single-entry metering. In the case of single-entry metering, the ramp metering signal is timed to permit only one vehicle to enter the freeway per green interval. Therefore, the green-plus-yellow (or just green if yellow is not used) interval is just long enough (usually about 1.5 to 2 seconds) to allow one vehicle to proceed past the signal. The red interval varies with the number of vehicles being metered. For instance, if a metering rate of 600 vph or 10 vpm were to be used, and the green-plus-yellow interval were 2 seconds, a red interval of 4 seconds would be used. If the metering rate were 300 vph, or 5 vpm, and the green-plus-yellow interval were 2 seconds, a red interval of 10 seconds would be used.

Platoon metering. When metering rates greater than 900 vph are required, platoon metering, which permits the release of 2 or more vehicles per cycle, may be used to achieve such high metering rates. For pretimed platoon metering, the cycle length is determined on the basis of the desired metering rate and the average number of vehicles to be released per cycle. For example, in the case of a metering rate of 1,080 vph, or 18 vpm, and a release of 2 vehicles per cycle, 9 cycles per minute would be required. Therefore, the cycle length would be 6.67 seconds. Similarly, if a release of 3 vpc were used instead, the cycle length would be 10 seconds. However, the timing of the cycle intervals (i.e., green, yellow, red) would depend on the form of platoon metering used, tandem or 2-abreast.

Tandem Metering. In the case of tandem metering, the vehicles are released one after another. Therefore, the green-plus yellow time is made long enough to permit the clearance of the desired number of vehicles per cycle. A yellow interval should be used to minimize the rearend collision potential. Thus, for the 7-second cycle with 2-vehicle platoons, a 4-second green-plus-yellow and a 3-second red might be used. And for a 12-second cycle with 3-vehicle platoons, a 9-second green-plus-yellow and a 3-second red might be used. Experience indicates that 2-vehicle platoons can be handled satisfactorily and that 3-vehicle platoons are a practical maximum. In either case, a

Figure 5-4. Typical Field Displays for Ramp Meter Installations.

maximum metering rate of 1,100 vph can be expected.[14]

Two-abreast Metering. With two-abreast metering, two vehicles are released side by side per cycle. This form of metering requires two parallel lanes on the entrance ramp plus a sufficient distance beyond the ramp metering signal for the two vehicles to achieve a tandem configuration before merging with freeway traffic. The more common practice in two-lane situations is to alternate the release—one from the left lane followed by one from the right. The timing of the cycle intervals for multiple-lane metering is similar to that for single-entry metering in that the green-plus-yellow interval is just long enough (usually about 3 seconds) to allow one vehicle in each lane to proceed past the ramp metering signal. The remainder of the cycle is red. With alternate release metering, maximum metering rates of about 1,700 vph may be achieved.

Compared to single-entry metering, platoon metering is a more complex operation and may cause some drive confusion which may lead to disruptions of ramp flow. Therefore, single-entry metering should always be given first consideration, and platoon metering should not be used unless it is necessary to achieve higher metering rates. However, platoon metering has been successfully used in several locations and drivers can adapt with proper design and pre-operation publicity.

It has been shown that entrance ramp control can be extremely cost effective.[1] Experience has indicated that the biggest net

gain in benefits is realized in going from no control to pretimed metering. Pretimed metering offers both advantages and disadvantages. The most important advantages are that it gives the driver a dependable situation to which he can readily adjust, and that it tends to be associated with lower costs. The major disadvantage is that the system can neither respond automatically to significant changes in demand, nor adjust to unusual traffic conditions resulting from incidents. Because of this inability to automatically respond to changes in traffic conditions and the relative difficulty of dissipating resultant congestion, pretimed metering rates have usually been set so that operation will be at volumes slightly below capacity at the desired LOS.

Traffic-Responsive Metering

In contrast to pretimed metering control, traffic-responsive metering is directly influenced by the mainline and ramp traffic conditions during the metering period. Metering rates are selected on the basis of real-time measurements of traffic variables indicating the current relation between upstream demand and downstream capacity.

Fundamental Traffic Flow Relationships

In order to determine or predict demand-capacity conditions on the basis of real-time measurements of traffic variables, a description or model of traffic is necessary. Most frequently used as indicators of operating conditions for traffic-responsive metering are functional relationships between flow rate, q; space-mean speed, u; and density, k.[15]

A generalized relationship between each of the variables is depicted in figure 5-5 and can be summarized as follows:

- At zero density, or when no vehicles are on the roadway, the flow rate is zero. and traffic is permitted to travel at its free speed, u_f.

- As density increases to a value, k_m, the flow rate increases to a maximum value, q_m, which is the capacity of the roadway, and speed decreases to a value, u_m.

- As density increases from a value, k_m, to a maximum value, k_j (jam density), the flow rate decreases to zero because the roadway is blocked with too many vehicles for traffic to move.

The values of q_m, u_f, u_m, k_m, and k_j—and the shapes of the curves—depend on several factors including geometrics of the roadway, composition of traffic, and weather conditions. Therefore, these values may be different for different sections of the roadway, and each section may have more than one set of these values. Although these are theoretical relationships based on the assumption of uniform traffic flow, the trends expressed by these relationships do exist.[16]

Basic Strategy

As explained earlier, congestion occurs whenever demand exceeds capacity. Therefore, as indicated in figure 5-5, the values of q_m, u_m, and k_m define boundaries between congested flow and uncongested flow. The purpose of metering is to prevent or reduce congestion, or in other words, to keep the values of the fundamental traffic flow variables at levels that define points on the uncongested-flow portions of the traffic flow curves. Thus, the basic strategy of traffic responsive metering is as follows:

- Obtain real-time measurements of traffic variables on the freeway.

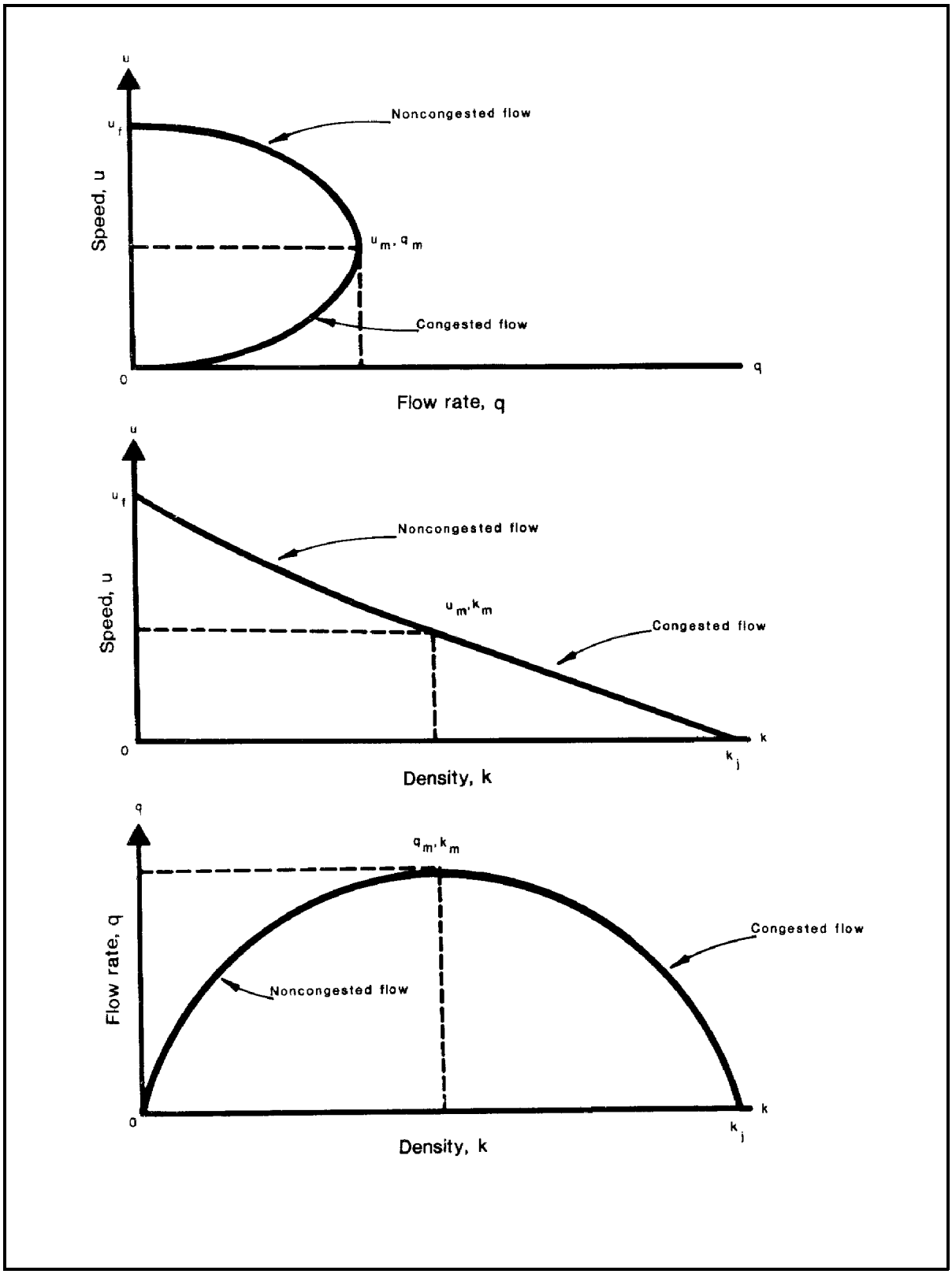


Figure 5-5. Fundamental Flow Rate-Speed-Density Relationships.

- On the basis of these measurements, determine where on the fundamental traffic flow curves the freeway section is operating with respect to capacity.

- Determine the maximum ramp metering rate at which vehicles can be permitted to enter the freeway.

A refinement that is often made to this strategy is to select the traffic flow curves on the basis of traffic composition and weather conditions.

Lane occupancy (a surrogate measure for density) and flow rate (volume) are the two traffic variables generally used to describe freeway traffic conditions for traffic responsive metering. These are the control parameters usually used, because they can be measured in real time using vehicle detectors.

Several variations on the basic strategy of traffic-responsive metering utilize different combinations of traffic variables. Although most are reported as having positive effects on freeway operations, none has been generally accepted as being superior to the others. In fact, new strategies are still being formulated to find better modes of control. However, the principal traffic-responsive strategies remain demand-capacity control and occupancy control.

Demand-Capacity Control

Demand-capacity control features the selection of metering rates on the basis of a real-time comparison of upstream volume and downstream capacity. The upstream volume is measured in real time and compared with either a preset value of downstream capacity determined from historical data or a real-time value computed from downstream volume measurements. To be most effective, the downstream capacity used should account for the effects on capacity of weather conditions, traffic composition, and incidents.

The difference between the upstream volume and the downstream capacity is then determined and used as the allowable entrance ramp volume. This ramp volume is expressed as a metering rate to be used during the next control interval (usually 1 min). If the upstream volume is greater than the downstream capacity, a minimum metering rate is used (e.g., 3-4 vpm). Theoretically, if the upstream volume were greater than the downstream capacity, a zero metering rate, or ramp closure, should be used in order to prevent congestion. It has generally been found that metering rates lower than 3 vpm are not effective, because vehicles waiting at the ramp will judge the ramp metering signal to be malfunctioning and will proceed through on red.

Downstream capacity may also be measured directly from freeway detector(s) to reflect for variations in traffic composition, weather, or other limiting factors which would not be accounted for in a fixed value of capacity.

Since a low upstream volume could occur in congested as well as uncongested flow, volume alone does not indicate degree of congestion. Therefore, an occupancy measurement also is usually made to determine whether uncongested or congested flow prevails. If the occupancy measurement is above a preset value (e.g., 18 percent, as used in Los Angeles).[17] which is determined from historical data, congested flow will be assumed to exist and a minimum metering rate used.

Occupancy Control

Occupancy control utilizes real-time occupancy measurements generally taken

upstream of the entrance ramp. One of a number of predetermined metering rates is selected for the next control interval (usually 1 min) on the basis of occupancy measurements taken during the current control interval. For a given entrance ramp, the metering rate to be used for a particular value of occupancy would be based on a plot of historical volume-occupancy data collected at each measurement location. An example of a typical plot from Chicago is shown in figure 5-6.[18] From such a plot, an approximate relationship between volume and occupancy at capacity is determined. For each level of occupancy measured, a metering rate can be determined that corresponds to the difference between the predetermined estimate of capacity and the real-time estimate of volume. If the measured occupancy is greater than, or equal to, the preset capacity occupancy, a minimum metering rate will be selected instead of a zero rate or ramp closure. This choice would be based on effective and practical entrance ramp control considerations, as explained earlier for demand-capacity control. Table 5-1 shows a recommended range of metering rates based on measured occupancy.[19]

System Components

A traffic-responsive metering system contains the same components as described for pretimed metering. These include ramp metering signal(s), local controller, advance warning sign with flashing beacon, and detectors. The local controller unit for traffic-responsive metering requires additional logic over and above that required for pretimed metering in order to monitor traffic variable measurements, select or calculate metering rates, and respond to override-type conditions such as excessive queues. Queue, check in, checkout, and merge detectors are normally also included in traffic-responsive metering systems.

Some traffic-responsive metering systems have also included detectors used to determine traffic composition and weather conditions.[9, 17] Input from these detectors enables the system to account for the effects of these factors on traffic flow.

Table 5-1. Local Actuated Metering Rates as Function of Occupancy. [19]

Occupancy (%)	Metering Rate(Vehicles/ Minute)
≤ 10	12
11-16	10
17-22	8
23-28	6
29-34	4
> 34	3

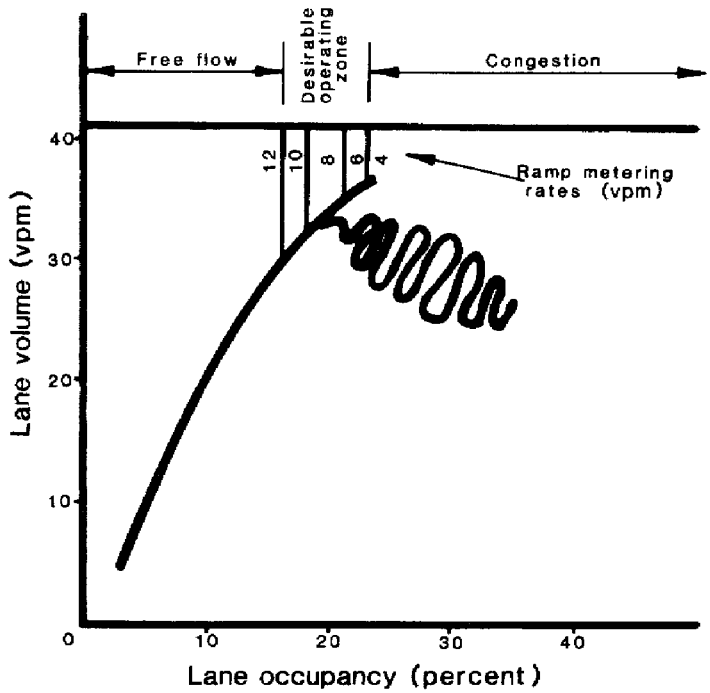


Figure 5-6. Typical Volume-Capacity Plot Related to Ramp Metering Rate. [18]

The typical location of these components on a ramp is shown in figure 5-7. For a discussion of standards for various components, the reader is referred to the publication on recommended practice for freeway entrance ramp control displays prepared by the Institute of Transportation Engineers (ITE).[13]

System Operation

The operation of a traffic-responsive metering system is similar to that of a pretimed metering system, except in regard to the following:

- **Metering Rate Selection.** Single-entry metering is normally used to time the red-yellow-green (or red-green) intervals for a given metering rate. However, if high metering rates (e.g., higher than 13 vpm), are required, platoon metering might be used. Single-entry metering and platoon metering should not both be used at any one ramp.

  In a traffic-responsive metering system, the selection is based on real-time measurements of traffic variables which describe traffic flow conditions on the freeway. The control interval, which is the time period during which a selected metering rate remains in effect, is much shorter for a traffic-responsive metering

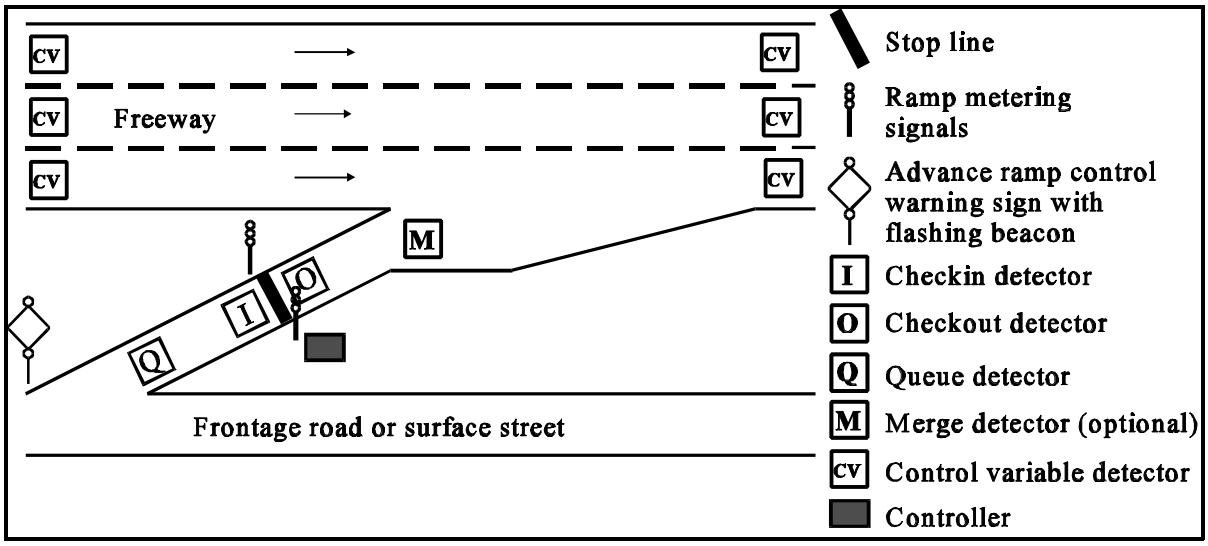


Figure 5-7. Traffic Responsive Ramp Metering Layout.

system (e.g., 1 min) than for a pretimed metering system (e.g., 30 min., 1 hr., or the entire peak period).

- **Override Features.** Override features of a traffic responsive system adjust metering rates in accordance with certain operational considerations as follows:

- **Continued Actuation of the Queue Detector.** Indicates that the queue of vehicles waiting at the ramp metering signal is approaching the frontage road or surface street and is likely to interfere with traffic on either or both. Therefore, a higher metering rate may be used to shorten the queue length.

- **Actuation of the Merge Detector.** Indicates that a vehicle is still in the merge area. Therefore, in the case of single-entry metering, subsequent green intervals are preempted until the vehicle merges.

- **No Actuation of the Checkout (Passages Detector After a Green Interval).** Indicates that a vehicle has missed the green signal. Therefore, the ramp metering signal is returned to or left in green.

- **Continued Actuation of the Queue Detector With No Actuation of the Check in (Demand) Detector.** Such a condition indicates that a vehicle on the ramp has stopped short of the check in detector. Therefore, the ramp metering signal is turned to green to allow this vehicle to proceed.

Gap-Acceptance Merge Control

Gap-acceptance merge control has been implemented and tested, but is little used, if at all, today. The concept of matching a merging vehicle to a specific freeway gap is attractive, but many variables can cause it to fail. Certain elements, such as slow vehicle detection, may have application in other types of ramp control operation. Gap-acceptance merge control might have application where geometries are substandard and the primary concern is safety.

The merge-control concept of entrance ramp metering is intended to enable a maximum

number of entrance ramp vehicles to merge safely without causing significant disruptions in freeway traffic. The concept involves maximum utilization of gaps in the traffic stream of the freeway lane into which ramp vehicles are to merge. It may or may not involve the calculation of ramp metering rates in accordance with the demand-capacity constraint. The problem is mainly one of inserting entrance ramp vehicles into freeway gaps. However, a provisional metering rate based on system calculations may be established. If a gap is found in a "window", say 3 seconds either side of the calculated release point, it is considered to have satisfied the metering rate, and a vehicle is released.[20] Gap acceptance metering has not been widely used, but may be warranted where geometrics are substandard or the safety of the merging operation can be improved.

Basic Concepts

The concepts of gap acceptance at freeway entrance ramps are important in describing the interaction of the freeway and ramp traffic. It is assumed that the ramp driver measures each gap in the adjacent freeway lane and compares it with an acceptable gap which he/she judges as large enough for a safe merge.

The minimum acceptable gap is dependent on several factors, such as the following:[20]

- Entrance ramp and freeway geometrics.
- Vehicle performance characteristics.
- Driver behavior.
- Traffic conditions.
- Weather conditions.

Merge-control systems are designed to improve the merging operation at the entrance ramp by providing the driver with the information needed to coordinate in time and space entry onto the freeway. These systems operate according to the following basic guideline procedures:

- Detection of an acceptable gap on the freeway into which a ramp vehicle could merge.
- Projection of the arrival of the acceptable gap at the merging point of the entrance ramp.
- Release of the ramp vehicle in sufficient time for it to accelerate and merge into the moving gap.
- If a gap is not detected within some maximum time, say 60 seconds, the vehicle is released.

System Components

Gap-acceptance merge-control systems use many of the same components as those described for pretimed metering, which include ramp metering signals, local controller, advance warning sign with flashing beacon, and detectors. A mainlane gap/speed detector is located in the shoulder lane of the freeway upstream of the ramp merge to provide data from which the controller unit can determine presence and speed of available gaps in which to insert merging ramp traffic. Queue, check in, checkout, and merge detectors are normally included in gap-acceptance merge-control systems.

Another override-feature component that might be added to the system is a slow-vehicle detector, which senses the presence of a slow-moving vehicle on the entrance ramp between the ramp metering signal and

the merge detector. A schematic layout for gap-acceptance operation combined with traffic responsive operation as implemented in the Dallas Corridor Study is presented in figure 5-8.[22] Also, for a discussion of standards for various system components, the reader is referred to the publication on recommended practice for freeway entrance ramp control displays prepared by the Institute of Transportation Engineers (ITE).[13]

System Operation

A gap-acceptance merge-control system does not normally operate in accordance with a constant metering rate for a specified control interval as do pretimed and traffic-responsive metering systems. Instead, it operates in response to the availability of acceptable gaps in the lane of the freeway into which ramp vehicles are to merge.

Usually, the system is designed to operate in a single-entry metering mode, with the ramp metering signal resting on red when no vehicles are waiting on the ramp. Experience on the Gulf Freeway in Houston has indicated that it is usually not desirable to operate the ramp metering signal in either of the following two ways:[24]

- If it gives a green indication at the proper time, whether or not there is a vehicle waiting.

- If it normally rests on green when there are no vehicles waiting.

Procedures can be summarized as follows for the nominal operation of a gap-acceptance merge-control system, with single-entry metering and the ramp metering signal resting on red:

- A vehicle stops at the ramp metering signal and actuates the check in detector.

- The controller begins to measure gaps and vehicle speeds which are sensed by the gap/speed detector that is located upstream from the ramp in the lane of the freeway into which ramp vehicles are to merge.

- The controller compares each measured gap to a preset minimum gap size to determine whether or not it is an acceptable gap.

- If the system includes demand-capacity features as described above, the controller determines if the gap falls within a "window" and adjusts the release time accordingly.

- If a gap is not acceptable, the controller considers the next gap. If it is acceptable, the controller computes the time at which the vehicle at the ramp metering signal should be released in order to arrive at the merge point at the same time as does the acceptable gap. This calculation involves the following factors:

  - Speed of the traffic flow measured in the lane of the freeway into which ramp vehicles are to merge.

  - Distance of the gap/speed detector location from the merge point.

  - Predetermined ramp travel time of a vehicle stopped at the ramp metering signal to the merge point.

- At the proper instant, the controller causes the ramp metering signal to change to green.

- The ramp metering signal remains on green for a fixed interval long enough to release a single vehicle. Then, it changes to yellow for a short fixed interval before

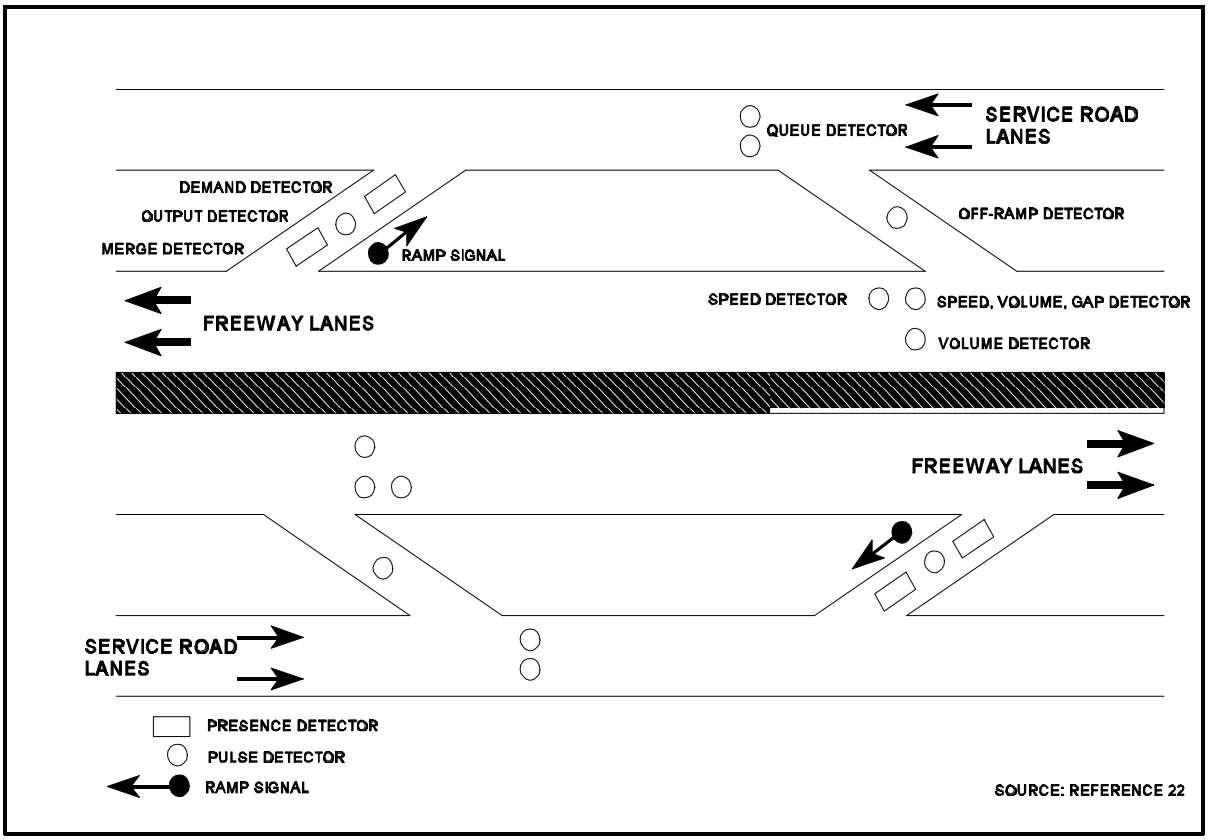


Figure 5-8. Gap Acceptance / Traffic Responsive Ramp Metering Layout.

it changes to red. (Where permitted by State law, the yellow interval may be omitted.) The green plus yellow (or green only) interval is usually about 3 seconds long. It is necessary that the ramp metering signal remain on red long enough to give the next vehicle in line time to pull up to the signal. Thus, the minimum time for a full green-yellow-red (or green-red) cycle should be 4 to 5 seconds.

The operation of the override features of a gap-acceptance merge-control system is essentially the same as for a traffic-responsive metering system. However, a gap-acceptance merge-control system may have the following additional override features:

- **Low-speed, Fixed-rate Metering.** When congested flow develops on the freeway, small space headways between successive vehicles constitute large time headways because of the low speeds. For example, if traffic on the freeway should come to a complete stop, the measured time headways will be infinitely large. Thus, unless an appropriate override were provided, the controller would release a number of entrance ramp vehicles to enter the freeway during the congested flow, a response which would be contrary to the objective of improving freeway operations. Therefore, if the speed of the freeway traffic drops below a preset value (e.g., 25 mi/h), ramp vehicles are metered at a minimum fixed rate {usually 3 to 4 vpm).

- **Slow-vehicle, Red-interval Extension.** At entrance ramps where there are relatively high percentages of trucks and buses, it might be desirable to make

special allowances for their performance characteristics. Accordingly, a slow-vehicle detector might be provided to measure the travel time of vehicles from the ramp metering signal to their location. If the travel time is greater than a preset value, the ramp metering signal is held on red until the vehicle has cleared the merge detector or until the merge detector is actuated.

Benefits to be realized from a gap-acceptance merge-control system are similar to those realized from traffic-responsive metering system. A study conducted by the Texas Transportation Institute, which compares a gap-acceptance merge-control system with a demand-capacity-control, traffic-responsive metering system, has reported the following results:[23]

- Gap-acceptance merge-control resulted in a higher percentage of ramp-metering-signal violations by ramp vehicles, which was probably due to its irregular pattern of operation and longer queue delays (metering rates ranged from 1 veh/4 seconds to 1 veh/25 seconds).

- Gap-acceptance merge-control resulted in lower travel times from the ramp metering signal to the merge area, which indicates a smoother merging operation.

- Demand-capacity control resulted in higher metering rates and higher peak-hour entrance ramp volume.

In general, for entrance ramps that have well-designed geometrics, a gap-acceptance merge control is less cost-effective than either pretimed or traffic-responsive metering systems. However, gap-acceptance control might be warranted at locations where the geometrics are substandard and the primary concern is to improve the safety of the merging operation.

System Ramp Control

System ramp control refers to the application of ramp control to a series of entrance ramps where a single ramp meter cannot address the excess freeway demand. The primary objective of system ramp control is to prevent or reduce the occurrence of congestion on the freeway. Therefore, the control of each ramp in the control system is based on the demand-capacity considerations for the whole system rather than on the demand-capacity constraint at each individual ramp. This concept does not necessarily imply the use of large computer control systems, since small subsystems may be coordinated by the use of mutual coordination of adjacent ramp meter controllers.

If congestion is to be prevented or reduced on the freeway system, the concept of system ramp control must be used in the design of a system of controls for a section of freeway with more than one entrance ramp. It may be applied in the following types of systems:

- System pretimed metering (including ramp closure).

- Traffic-responsive metering.

- Gap-acceptance merge control.

A discussion of system ramp control applied to each of these systems follows.

System Pretimed Metering

System pretimed metering refers to the application of pretimed metering to a series of entrance ramps. The metering rate for each of these ramps is determined in accordance with demand-capacity constraints at the other ramps as well as its own local demand-capacity constraint.

Determining these metering rates, which are computed from historical data pertaining to each control interval, requires the following information:

- Mainline and entrance ramp demands.
- Freeway capacities immediately downstream of each entrance ramp.
- Description of the traffic pattern within the freeway section to be controlled.

This information provides the basis for establishing the demand-capacity constraints of the entrance ramps and their interdependencies.

Fundamental metering rate calculations—given the required data, the fundamental procedure for computing metering rates involves five steps:

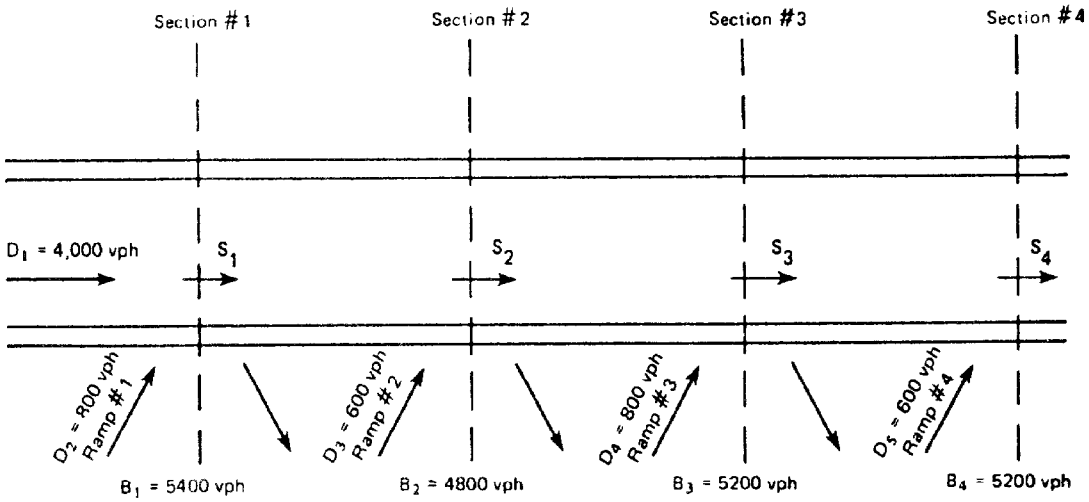
1. Start with the entrance ramp that is farthest upstream.
2. Determine the total demand (upstream mainline demand plus ramp demand) for the freeway section immediately downstream of the ramp.
3. Compare the total demand to the capacity of the downstream section, and proceed as follows:
   a. If the total demand is less than the capacity, metering is not required at this ramp by this demand-capacity constraint. Therefore, skip step 4 and go immediately to step 5.
   b. If the total demand is greater than the capacity, metering is required at this ramp by the demand-capacity constraint. Therefore, proceed to step 5.
4. Compare the upstream mainline demand to the capacity of the downstream section and proceed as follows:
   a. If the upstream mainline demand is less than the capacity, then the allowable entrance ramp volume (or metering rate) is set equal to the difference between the capacity and the upstream mainline demand.
   b. If the upstream mainline demand is greater than or equal to the capacity, then the allowable entrance ramp volume is zero, and the ramp must be closed. If the upstream mainline demand is greater than the capacity, the volumes permitted to enter at ramps upstream must be reduced accordingly. The total reduction in the allowable entrance ramp volumes upstream is equal to the difference between the upstream mainline demand and the capacity, adjusted to account for that portion of the traffic entering upstream that exits before it reaches the downstream entrance ramp being closed.
5. Select the next entrance ramp downstream and go back to step 2.

This procedure is illustrated by the following examples.

Example 1 [5,6]

In the example case shown in figure 5-9, pretimed metering rates are calculated for an integrated, pretimed control system comprised of four entrance ramps. In reviewing this example, the following points should be noted:

- Since only entrance ramp control is being considered and not mainline control, the allowable mainline volume at



X_i = allowable volume at input i

D_i = demand at input i

B_j = capacity at section j

A_{ij} = decimal fraction of vehicles entering at input i which pass through section j

S_j = demand at section j

A_{ij} Values

i \ j	1	2	3	4
1	1.00	0.95	0.90	0.85
2	1.00	0.75	0.70	0.60
3	–	1.00	0.90	0.85
4	–	–	1.00	0.90
5	–	–	–	1.00

Compute X_i's starting at section #1

- Set $X_1 = D_1 = 4000$ vph
- $S_1 = A_{11}X_1 + A_{21}D_2 = (1.00)(4000)+(1.00)(800) = 4800 \text{ vph} < B_1 = 5400 \text{ vph}; \therefore X_2 = 800 \text{ vph}$
- $S_2 = A_{12}X_1 + A_{22}X_2 + A_{32}D_3 = (0.95)(4000)+(0.75)(800)+(1.00)(600) = 5000 \text{ vph} > B_2 = 4800 \text{ vph}; \therefore X_3 = 400 \text{ vph}$
- $S_3 = A_{13}X_1 + A_{23}X_2 + A_{33}X_3 + A_{43}D_4 = (0.90)(4000)+(0.70)(800)+(0.90)(400) + (1.00)(800) = 5320 \text{ vph} > B_3 = 5200 \text{ vph}; \therefore X_4 = 680 \text{ vph}$
- $S_4 = A_{14}X_1 + A_{24}X_2 + A_{34}X_3 + A_{44}X_4 + A_{54}D_5 = (0.85)(4000) + (0.60)(800) + (0.85)(400) + (0.90)(680) + (1.00)(600) = 5432 \text{ vph} > B_4 = 5200 \text{ vph}; \therefore X_5 = 368 \text{ vph}$

Conclusion:

- Ramp #1: No control needed.
- Ramp #2 : Meter at a rate of 400 vph.
- Ramp #3: Meter at a rate of 680 vph.
- Ramp #4: Meter at a rate of 368 vph.

Figure 5-9. Integrated Entrance Ramp Control: Example No. 1 Calculation of Pretimed Metering Rates. [5,6]

Section 1, Xj is set equal to the mainline demand Dj.

- With the notation given in figure 5-9, the demand, S_j, at a section, j, is computed by the following equation:

$$S_j = (\sum_{i=1}^{j} A_{ij}X_i) + A_{j+1}D_{j+1}$$

where:

X_i = Allowable volume at input i

D_j = Demand at input i

A_{ij} = Decimal fraction of vehicles entering at input i which pass through Section j

S_j = Demand at Section j

As it happens, the metering rate computed for each entrance ramp in this particular example is determined solely by the demand-capacity constraint at the section immediately downstream and is not influenced by the demand-capacity constraints at other ramps.

Example 2[5,6]

The data given in the example shown in figure 5-10 are the same as those given in the previous example. except that the mainline demand, D_1, is 4,600 vph instead of 4,000 vph. In this case, the metering rates at Ramps 2, 3, and 4 are determined solely by their respective downstream demand-capacity constraints, as was the case in the previous example. However, the metering rate at Ramp 1, rather than being determined by the demand-capacity constraint at Section 1, is established in accordance with the demand-capacity constraint at Ramp 2, as is described below.

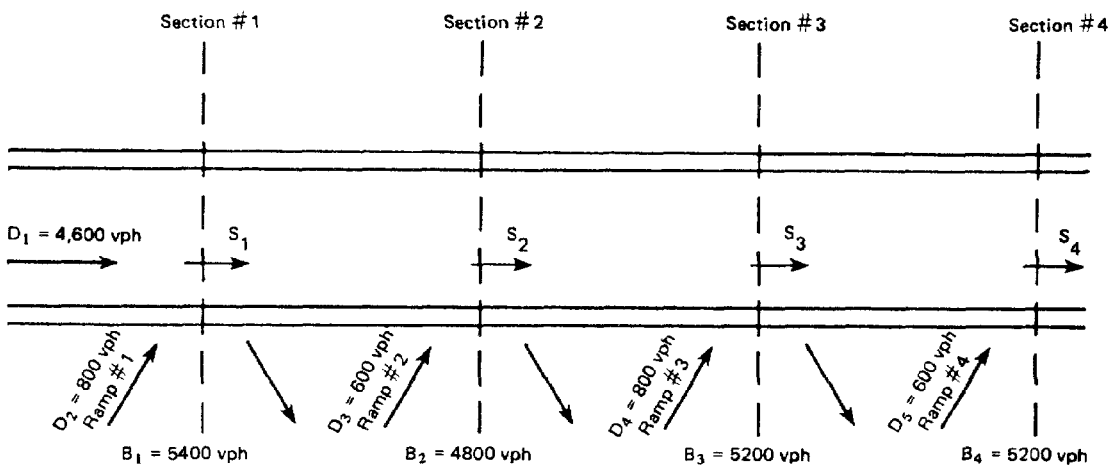
The demand, S_2, at Section 2 is 5,570 vph, which is 770 vph greater than the capacity, B_2, at Section 2 (4,800 vph). If Ramp 2 is closed, the demand at Section 2 is reduced to 4,970 vph, a volume which also exceeds the capacity, B_2. Therefore, it is necessary to reduce the allowable volume, X_2, entering at Ramp 1 (input 2). The allowable volume, X_2, must be reduced enough to reduce the demand, S_2, by 170 vph. The amount of the reduction is equal to the 170 vph divided by the decimal fraction, A_{22}, of the vehicles entering at Ramp 1 and passing through Section 2 (170 vph/0.75 - 227 vph). Therefore, the allowable volume, X_2, at Ramp 1 would be 573 vph instead of 800 vph.

In this procedure, excess demand, S_j - B_j, at any section, j, is removed by reducing the allowable volume on the entrance ramp immediately upstream. If' instead, the allowable volumes on entrance ramps farther upstream were reduced, a large number of vehicles would have to be removed from these ramps in order to reduce the demand, S_j, sufficiently at any section, j. This is necessary because some of the vehicles that enter at these ramps will exit the freeway before they reach Section j.

Example 3

Again, in the situation presented in table 5-2, allowable ramp volumes would be calculated as follows. If the excess demand, 1200 vph, at Section 2 were to be removed by reducing the allowable volume, X_2, at Ramp 1, the volume at Ramp 1 would have to be reduced by 267 vph. The allowable entrance ramp volumes are summarized accordingly in table 5-2.

The total input of 2,172 vph, however, is less than that of 2,248 vph, the volume obtained if Ramp 2 is metered as in Example 1. Thus, the fundamental approach



X_i = allowable volume at input i

D_i = demand at input i

B_j = capacity at section j

A_{ij} = decimal fraction of vehicles entering at input i which pass through section j

S_j = demand at section j

A_{ij} Values

i \ j	1	2	3	4
1	1.00	0.95	0.90	0.85
2	1.00	0.75	0.70	0.60
3	–	1.00	0.90	0.85
4	–	–	1.00	0.90
5	–	–	–	1.00

Compute X_i's starting at section #1

- Set $X_1 = D_1 = 4600$ vph
- $S_1 = A_{11}X_1 + A_{21}D_2 = (1.00)(4600)+(1.00)(800) = 5400$ vph $= B_1 = 5400$ vph; $\therefore X_2 = 800$ vph
- $S_2 = A_{12}X_1 + A_{22}X_2 + A_{32}D_3 = (0.95)(4600)+(0.75)(800)+(1.00)(600) = 5570$ vph $> B_2 = 4800$ vph;

  $\therefore X_3 = 0$ and volume entering upstream must be reduced by 170 vph;
  $\therefore X_2 = 800 - 170/A_{22} = 800 - 170/0.75 = 573$ vph
- $S_3 = A_{13}X_1 + A_{23}X_2 + A_{33}X_3 + A_{43}D_4 = (0.90)(4600)+(0.70)(573)+(0.90)(0)$
  $+ (1.00)(800) = 5341$ vph $> B_3 = 5200$ vph; $\therefore X_4 = 659$ vph
- $S_4 = A_{14}X_1 + A_{24}X_2 + A_{34}X_3 + A_{44}X_4 + A_{54}D_5 = (0.85)(4600) + (0.60)(573) + (0.85)(0) + (0.90)(659)$
  $+ (1.00)(600) = 5447$ vph $> B_4 = 5200$ vph; $\therefore X_5 = 353$ vph

Conclusion:

- Ramp #1: Meter at a rate of 573 vph.
- Ramp #2: Close.
- Ramp #3: Meter at a rate of 659 vph.
- Ramp #4: Meter at a rate of 353 vph.

Figure 5-10. Integrated Entrance Ramp Control: Example No. 2 Calculation of Pretimed Metering Rates. [5,6]

Table 5-2. Allowable Entrance Ramp Volumes for Example 3.

Ramp No.	Volume (vph)
1	533
2	600
3	687
4	352
Total Input	2172

described will result in the optimal utilization of the freeway. It maximizes the sum of the allowable entrance ramp volumes, a procedure which corresponds to maximizing system output for steady-state, uncongested flow conditions.[26] It also maximizes the total travel in the system.[27]

Linear Programming Formulation—The fundamental procedure described in Examples 1 and 2 can be formulated as a linear programming model.[26] This model may be used to compute optimal allowable entrance ramp volumes. In terms of the notation defined in figures 5-8 and 5-9, the linear programming model would be as follows:

- Maximize ΣX_j, where n is the number of inputs
- Subject to the following constraints:
  - Demand capacity:

$$\sum_{i}^{n} A_{ij} X_i \quad B_j;\ j = 1,.....n-1$$

  - At Section 1, allowable mainline volume ≤ mainline demand:

$$X_1 = D_1$$

  - Allowable entrance ramp volume ≥ entrance ramp demand:

$$X_1 \leq D_1;\ i=2,.....,n$$

  - Allowable entrance ramp volume equals minimum allowable ramp volume:

$$X_1 \geq \min x_i \geq 0;\ i = 2,.....n$$

The use of the linear programming model yields allowable entrance ramp volumes identical to those obtained by using the fundamental procedure described above.

Practical Considerations

The allowable entrance ramp volumes (or metering rates) calculated for an integrated ramp control system should be evaluated with respect to the following practical considerations:[6]

- Metering rates of less than 180 to 240 vph (3 to 4 vpm) are not feasible because drivers required to wait longer than 15 to 20 seconds at a ramp metering signal often believe that the signal is not working correctly. They will, therefore, proceed on a red indication by the signal. Thus, if a metering rate of less than 180 to 240 vph is calculated, consideration should be given either to closing the ramp or to metering it at a higher rate.

- Practical maximum metering rates are about 900 vph for single-entry metering and approximately 1,100 vph for platoon metering. Therefore, for a metering rate greater than the maximum for the metering type to be used, the setting should be less than or equal to the practical maximum rate, and the metering rates at the other entrance ramps should be adjusted accordingly.

- Metering rates at each entrance ramp should be evaluated with regard to available storage at the ramp and potential resulting congestion on the adjoining surface street system. If the storage is not sufficient, it may be necessary either to close the ramp or to increase the metering rate.

- Metering rates equal to zero indicate that an entrance ramp closure is necessary. However, the closure of a particular entrance ramp may not be acceptable. Therefore, it may be necessary to increase a zero metering rate to some minimum acceptable rate.

- The procedure described for computing metering rates gives preference to traffic entering the system near the upstream end. Consequently, metering rates at entrance ramps downstream may be too restrictive to be acceptable to the motoring public. Therefore, it may be necessary to increase the metering rates computed for some of the downstream entrance ramps, and thus to reduce accordingly the metering rates for some of the upstream entrance ramps.

If any of the computed metering rates were to be altered because of one or more of the practical considerations mentioned above, the metering rates at the other entrance ramps would have to be adjusted accordingly to ensure both an optimal utilization of the freeway and an uncongested flow.

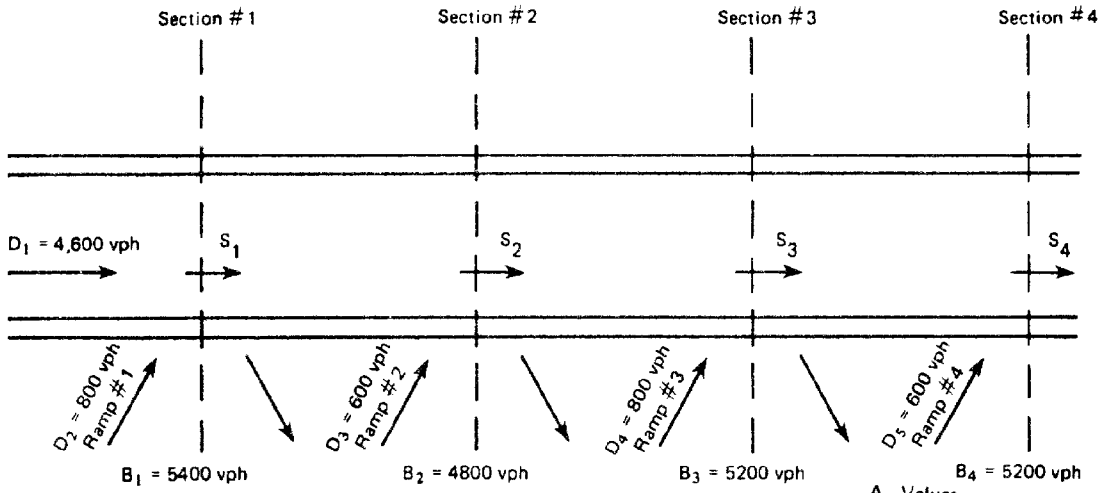
Example 4 [5,6]

If it were necessary to maintain a metering rate of at least 240 vph at Ramp 2 in the example presented in figure 5-10, it would be necessary to follow the adjustment procedure for the metering rates at the other entrance ramps (as shown in figure 5-11). The allowable volume. X_2 (573 vph), at Ramp 1 would have to be reduced by 320 vph in order to allow 240 vph to enter at Ramp 2 and still satisfy the demand-capacity constraint at Section 2. This reduction also decreases the mainline demand at Sections 3 and 4. Thus, the allowable volumes at Ramps 3 and 4 are increased to maximize the utilization of the freeway at these sections.

It is usually difficult to obtain reliable estimates of the A_{ij} values, because these vary with time and generally exhibit a high variance. Also, the O/D type studies used to collect these data are expensive and do not provide real-time data.

As indicated in the 1996 *Traffic Control System Handbook*, it may be unfeasible to reduce ramp volumes sufficiently to effect changes on freeway main lanes because of circumstances such as the following:[7]

- Minimum metering rate constraints.



X_j = allowable volume at input i

D_i = demand at input i

B_j = capacity at section j

A_{ij} = decimal fraction of vehicles entering at input i which pass through section j

S_j = demand at section j

A_{ij} Values

i \ j	1	2	3	4
1	1.00	0.95	0.90	0.85
2	1.00	0.75	0.70	0.60
3	–	1.00	0.90	0.85
4	–	–	1.00	0.90
5	–	–	–	1.00

Compute X_j's starting at section #1

- Set $X_1 = D_1 = 4600$ vph
- $S_1 = A_{11}X_1 + A_{21}D_2 = (1.00)(4600)+(1.00)(800) = 5400$ vph $= B_1 = 5400$ vph; $\therefore X_2 = 800$ vph
- $S_2 = A_{12}X_1 + A_{22}X_2 + A_{32}D_3 = (0.95)(4600)+(0.75)(800)+(1.00)(600) = 5570$ vph $> B_2 = 4800$ vph;

  Since X_3 must equal at least 240 vph; $X_3 = 240$ vph and volume entering upstream must be reduced by 410 vph;
  $\therefore X_2 = 800 - 410/A_{22} = 800 - 410/0.75 = 253$ vph
- $S_3 = A_{13}X_1 + A_{23}X_2 + A_{33}X_3 + A_{43}D_4 = (0.90)(4600)+(0.70)(253)+(0.90)(240) + (1.00)(800) = 5333$ vph $> B_3 = 5200$ vph; $\therefore X_4 = 667$ vph
- $S_4 = A_{14}X_1 + A_{24}X_2 + A_{34}X_3 + A_{44}X_4 + A_{54}D_5 = (0.85)(4600) + (0.60)(253) + (0.85)(240) + (0.90)(667) + (1.00)(600) = 5466$ vph $> B_4 = 5200$ vph; $\therefore X_5 = 334$ vph

Conclusion:

- Ramp #1: Meter at a rate of 253 vph.
- Ramp #2: Meter at a rate of 240 vph.
- Ramp #3: Meter at a rate of 667 vph.
- Ramp #4: Meter at a rate of 334 vph.

Figure 5-11. Integrated Entrance Ramp Control: Example No. 4 Calculation of Pretimed Metering Rates. [5,6]

- Lack of vehicle queuing storage.
- Too large a capacity deficiency.

The reader is referred to the handbook for a detailed example demonstrating the interaction among ramp metering requirements, diversion impacts, and ramp storage requirements. A detailed procedure can be found in reference 19.

Systemwide ramp metering strategies provide the opportunity to distribute vehicle demands over a larger number of ramps.

System Traffic-Responsive Metering

System traffic-responsive metering is the application of traffic responsive metering to a series of entrance ramps where the metering rates at each ramp are selected in accordance with both system and local demand-capacity constraints.

System Operation

During each control interval, real-time measurements are taken of traffic variables (usually volume, occupancy, and/or speed). The data are used to define the demand-capacity conditions at each ramp. Then, on the basis of these measurements, both an independent and an integrated metering rate are calculated for each entrance ramp. Of these two metering rates, the one that is the more restrictive is selected to be used during the next successive control interval.

Metering Rates

The methods used to calculate independent and integrated traffic-responsive metering rates are basically the same as those used to compute independent and integrated pretimed metering rates. Instead of calculating metering rates in real time, a set is precomputed for the range of demand-capacity conditions expected from which the metering rates are then selected in real time. The linear programming model is often used to calculate predetermined sets of integrated, traffic-responsive reentering rates. Also, the metering rates are usually subject to the merge-detector, queue-detector, and maximum-red-time overrides used in traffic-responsive metering

System vs. Independent Ramp Control

Comparisons of system and independent entrance ramp control indicate that increased benefits are realized with system ramp control.[26,27] Improvements occur in terms of the following:

- Lower travel time.
- Higher total travel.
- Fewer crashes.

In traffic-responsive metering, the greater system flexibility provided by system ramp control enables an optimal system response to individual variations in traffic demands and capacities resulting from incidents on the freeways.

Controller Interconnection

A significant feature of system ramp control is the interconnection among local ramp controllers, which permits conditions at one location to affect the metering rate imposed at one or more other locations. Real-time metering plans are computed and updated by a central master computer which issues metering rates to the respective local ramp controllers on the basis of freeway traffic information obtained from vehicle detectors throughout the system.

Although the decision-making capabilities are centralized within the central computer

system, the processing of control intelligence may be distributed among the individual entrance ramps. For economic (and possibly reliability) reasons, there is a trend toward decentralized decision-making, distributed computation, and hierarchical control.[28]

RAMBO (Ramp Adaptive Metering Bottleneck Optimization) is a suite of programs developed for the Texas Department of Transportation by the Texas Transportation Institute.[29] RAMBO I is a software tool designed to assist in developing ramp metering plans using the TxDOT ramp meter specification, while operating either in the isolated mode or in local control. The program provides Transition Point Patterns for each metering level and evaluates traffic operations. RAMBO II likewise develops and evaluates ramp metering plans based on forecasted traffic conditions along an extended section of freeway containing up to 12 metered entrance ramps and 12 exit ramps operating either in the system mode, or in a hierarchically distributed system having real-time local control with systems-based metering objectives. The program was implemented in Houston, TX, in 1996.

The total software package can perform capacity analysis of the freeway system, assess projected metering operation, and assist in developing optimal ramp metering plans for either local ramp metering operations (using RAMBO I) or system ramp metering operations (using RAMBO II). RAMBO II can translate system-based results into local metering control parameters that can be downloaded into the local ramp meters if some minor modifications are made to the current ramp meter specifications. The programs include extensive interactive graphic screens.

Incremental Benefits of Various Levels of Control

As discussed earlier, the benefits offered by pretimed metering (including ramp closure) versus no access control include increased mainline speeds (reduced travel time), higher service volumes, less delay, safer merging operations, and reduced user costs. Beyond pretimed metering, the incremental benefits gained from traffic-responsive metering (local or systemwide) depend on the factors discussed below.[21]

Variations in the Ratio of Mainline to Entrance Ramp Demand

As mainline demand approaches capacity, the permissible metering rates become more and more constrained. On the other hand, as the mainline demand decreases. more traffic can be allowed onto the freeway from entrance ramps, and ramp metering control can exert greater impact on the quality of freeway flow, thus producing greater benefits.

Variations in Overall Traffic Demand Pattern

Traffic demand on the freeway and entrance ramps exhibits two types of variations: (1) shift in demand level, and (2) short-term fluctuations. The larger the magnitude of these types of demand variation, the higher the potential for benefits from traffic-responsive metering.

Mainline Capacity Reductions

Reductions in mainline freeway capacity result from accidents, traffic incidents, and adverse weather conditions. As the frequency and impact of these capacity-reducing factors increase, more need develops for traffic-responsive metering to cope with the variations in available

capacity. To determine the appropriate level of ramp metering control for a given freeway, the incremental benefits produced by local and systemwide traffic-responsive metering (relative to a base of pretimed metering) must be estimated. Computer simulation can be effectively used in evaluating control benefits. In addition, the incremental system costs (installation, operation, maintenance) and the incremental user costs (travel time, vehicle operating costs, accidents, air pollution emission) must be estimated. Incremental benefits and incremental costs can then be used to conduct a benefit/cost or utility/cost analysis to decide upon the most desirable type of ramp metering.

The growth in traffic demand over the lifetime of a ramp metering project may reduce the incremental benefits of a traffic-responsive type of ramp metering control (local or systemwide). As traffic demand grows substantially over the lifetime of the project, the controllability index of the freeway decreases. Since the benefits are nonlinearly related to the controllability, it is possible that the benefits could decrease faster than the growth rate in demand. In planning ramp metering installations, the engineer should be aware of this effect. It is recommended that the analysis be repeated for as many years as are in the expected project-life duration.

The incremental benefits analysis is but one component of a system selection process which, in turn, is a component of a freeway traffic management decision process. The major components of this decision process include the following:

- Developing a basic analysis of freeway operations.
- Making a detailed analysis of freeway operations and determining improvement alternatives.
- Examining the feasibility of ramp control as an improvement alternative.
- Analyzing the site conditions and selecting the control level.

A detailed discussion of the incremental benefits of different types of ramp metering control is provided in NCHRP Report 232.[19]

EXIT RAMP CONTROL

Exit ramp control is seldom used as a means of freeway traffic control because the opportunities for its effective application are limited. In many situations, the use of exit ramp control may actually be contrary to the objective of safe and efficient freeway operations. Also, it should only be used where destinations can easily be reached by using alternate exits.

Exit ramp closure can be used effectively to reduce safety hazards and congestion caused by excessive weaving between closely spaced ramps and long queues on exit ramps. Also, exit ramp closure can be used at a lane drop location by closing downstream exit ramps in order to encourage more traffic to leave the freeway at the exit ramps before the lane drop and thus decrease the demand on the freeway section beyond the lane drop. However, as in the case of entrance ramp closure, exit ramp closure might not be acceptable because of the increased travel it creates for some motorists.

EMERGING TECHNOLOGIES

As mentioned in the opening paragraph of this section, most of the advances and

emerging technologies in freeway management systems are in the computing hardware and communications technologies. While development of those fields will continue to enhance the ramp control process, there will be emergence or at least advancement, from preliminary stages of freeway ramp control systems. Such advancements include:

- **System Operation.** As freeway systems expand with more communications links, and detector data become available, there will be increased system operation of entrance ramp meters, with metering rates being determined on a system or subsystem basis.

- **Integrated Systems.** Earlier freeway management systems generally operated independently of the operation of surface street signal systems. In lieu of actually being integrated through hardware and communications links, traffic state and local system managers sometimes communicated informally to bridge the two systems. Future systems will likely be fully integrated, with data exchange and control decisions being made automatically with real-time data.

- **Information to Motorists.** Advanced information systems being installed or planned as part of the ITS deployment and expansion will assist motorists in selecting or bypassing entrance ramps where metering rates may be restrictive. Such diversion can be considered in integrated freeway and surface street systems.

- **Advanced Control Algorithms.** The National ITS Architecture Implementation Strategy provides an evaluation of ITS Technology Areas as to their maturity (mature, immature, mixed) to assess their potential deployment horizon.[2] While most of the elements are hardware oriented,"processing technology and advanced algorithms that enable advanced vehicle and traffic control application" are designated as mixed, meaning that there is opportunity for emerging applications.

- **Advanced Ramp Metering Concepts.** Because queues become critical under heavy ramp demands conditions, improved queue management algorithms based on multipoint detection are under development. Also, traffic responsive activation of ramp control will likely be used to manage traffic during off-peak or weekend incident conditions.[7]

5.4 LESSONS LEARNED

Although ramp control systems have been in operation in various metropolitan areas throughout the country for over a quarter of a century, they are still sometimes viewed as a "new or radical" approach to traffic control and management. Intersection traffic signals, on the other hand, are accepted by most drivers as necessary and, in fact, their installation is often requested by citizens. The two systems essentially perform the same function: **Facilitate use of available capacity between conflicting vehicular movements on the basis of demand levels and safety considerations with traffic signals.** However, the ramp signal may be viewed negatively by drivers, because freeways have been traditionally designed for unrestricted flow. In reality, the flow is often restricted by recurring and non-recurring congestion that may have a greater effect than that of the meter signal, which may encourage the driver to divert. For these reasons, there are certain "lessons learned" associated with ramp control which

may not be a factor in other traffic and freeway control elements.

IMPLEMENTATION

Public Relations

Ramp metering systems can be successful only if they receive public support from political leaders, enforcement agencies, and the motoring public. To gain this support in advance of implementation, a comprehensive public relations and information program should begin well in advance. To the public, ramp meters are often seen as a constraint on a roadway normally associated with a high degree of freedom. Although definite benefits may be achieved by metering and have been demonstrated statistically, the benefits may not be recognized by individual motorists. A 3-minute wait at an entrance ramp, however, is easily recognized. A proactive public relations program should be an integral part of every metering project.[1]

It is important not to oversell the benefits of ramp metering. It is not a substitute for a new freeway lane. The benefits are measurable systemwide, but may not be readily discernable to the individual driver at the ramp signal. Successful public relations campaigns will explain the difficulties of mitigating freeway congestion problems and the cost effectiveness of management techniques such as ramp metering.[1] The campaigns should also provide realistic expectations of the system's benefits, and show how taxpayers will experience improved freeway conditions. The most common method of disseminating ramp metering information is through brochures or media advertisements on television and radio. Some examples of public relations brochures are shown in reference 1. In Minneapolis and Los Angeles, the "public" has actually requested additional metered ramps. This public input has become one of the factors in evaluating and selecting new metered locations.

Public relations aspects of the ramp control system should begin well in advance of turn-on. In Seattle, the Washington State DOT (WSDOT) has developed a methodical approach to implementing ramp metering. [30] Their process describes what needs to be accomplished starting five years prior to ramp metering all the way up to one week before, and continuing through six months after start-up. The procedure includes public input, the design process, and the public relations focus. In Tacoma, Washington, the WSDOT went beyond the typical public relations campaign of brochures and media advertisements. WSDOT has incorporated a ramp metering lesson into both public and private driver education school curricula. The lesson, which lasts about 30 minutes, helps students to understand what ramp meters are and what they mean to the driver. The information packet for this lesson includes a lesson plan, information sheets, brochures, key chains, and a well-developed 12 minute video entitled "Ramp Meters: Signals for Safety."

A promotional videotape from the FHWA entitled "Ramp Metering: Signal for Success" is another example of how the merits of ramp metering can be presented to the public.[1] This 17-minute videotape, which is intended for citizens and public officials, explains the principles and benefits of ramp metering. It addresses key issues such as safety, efficiency, equity, and public relations. Copies are available through the FHWA or the Institute of Transportation Engineers (ITE).

Media Relations

The print and electronic media can be great allies or great deterrents to the success of ramp control systems. When the Dallas

Corridor Study metering system was implemented in 1974, a radio reporter in the control center (with CCTV and other displays) reported that the system was working great, while a television reporter interviewing the 20th vehicle in a ramp queue proclaimed the system a failure.[31] The system perspective (which was understood by the reporter in the control center) must be stressed. As with the general public, the media must be informed as to system goals and expectations, schedules, operations, and results. It is also important to maintain communication with the media after system turn-on. Beat reporters are often reassigned, and the new reporter may need to be briefed before an uniformed, negative story is written.

Implementation Strategies

Scheduling of ramp control turn-on should be carefully considered. Incremental implementation of individual sections should be considered, rather than a total system launch. In particular, locations that have the best alternate routes and the highest probability of disruption of traffic flow should be considered first. Ramps should be operated with metering rates that cause little disruption. As drivers become familiar with and accustomed to the system, metering rates can be tightened and other locations implemented.

An interesting approach has recently been employed in Houston. Some of the pioneering efforts in ramp control took place in the mid-sixties.[12] However, due to reconstruction of freeways, ramp metering had not been in operation for some time. When ramp metering was recently reimplemented, a conservative philosophy was developed. The Implementation philosophy was as follows: [32]

> *... drivers and their views are important and a very high priority. No ramp delays (for a while at least) will be more than 2 minutes, and this must be verified. When queues or delays get too long, the signals are shut off until the queues clear, no matter what happens to the freeway. For the first three months, metering during the peak of the rush hour was sometimes terminated. No written complaints were received. However, continuous quality improvement for the freeway traffic flow is stressed. Freeway drivers have called by cell phone and by Internet asking TranStar (the freeway management center) for "more" ramp metering. Now, the simple explanation for this is that we have "teased" the freeway traffic into this position. But we have not followed any ramp control strategy mentioned in the traditional freeway ramp control manuals. The traditional demand/capacity methods are for marginally overloaded well-disciplined systems, and that goal of demand/capacity control is only a faint vision in Houston at the moment. We are simply pushing back up the q/k curve toward capacity from stop-and-go conditions, and not from the other side.*

Implementation Summary

The successful implementation of a freeway ramp control system is dependent on many factors outside of the hardware, software, and control algorithms. The implementation plan must include involvement, education, and support by the public, media, and political leaders. Additionally, the strategy with which individual ramps and subsystems

are "turned on" must be carefully considered, planned, and executed.

OPERATIONS AND MAINTENANCE

Operations and maintenance considerations are not unlike those for other freeway control subsystems or for other traffic signal control systems. While the strategies may differ, there is still a necessity for operating agencies to commit the funds for personnel to operate, maintain, evaluate, and update the control system.

- **Personnel.** Adequate personnel for system operation and maintenance are essential if systems are going to succeed and continue to succeed. While improved hardware and software capabilities have allowed many tasks to be automated in system operation, personnel must be assigned to ensure continued efficient operation.

- **Training.** Training for system operations and maintenance is usually provided by the systems contractor. Continuing training programs will be essential as new personnel are assigned and as hardware and software upgrades are implemented.

- **Documentation.** Initial documentation for system operation and maintenance should (**must**) be provided by the systems contractor. Operations and maintenance personnel must also ensure that documentation is updated as system changes or hardware upgrades are made. Detailed logs should be kept for such changes. Modern systems often incorporate automated logging capability to facilitate the task and ensure that records are consistent.

- **Evaluation.** Although effectiveness of ramp control techniques has been well documented in the literature, it is usually necessary to perform "before and after" studies to document results of each system. It will also be important to continue to sample system operation with the same type data used in the initial evaluation to detect changes in system operation performance.

- **Updating Initial Strategies.** Based on continued system monitoring, as mentioned above, changes in individual control parameters or control strategies may be warranted. These may require minor changes to the data base or more significant changes to the control programs. Changes in the roadway system, both freeway and surface streets, must also be monitored and considered.

- **Incorporating New Strategies.** As ramp control systems continue to grow and mature, new ramp control algorithms will likely also be developed and tested. Continued communications among system operators and participation in professional organizations such as the Transportation Research Board (Freeway Operations Committee), Institute of Transportation Engineers, and Intelligent Transportation Society of America will be beneficial in becoming aware of such strategies.

- **Hardware and Software Maintenance.** Hardware maintenance may be performed either by the agency or by contract, or by a combination of the two. The responsible agency will likely maintain standard traffic control equipment and communications cables. Computer and communications hardware will usually be maintained by contract. Software data bases will normally be maintained by the responsible agency,

while applications and system software will be maintained by contract. Whatever the method, agency or contract, maintenance responsibilities should be clearly defined and understood in advance of system implementation. Sufficient funding must continue to be committed for hardware and software maintenance.

DIVERSION OF TRAFFIC

A major issue that is raised in connection with metering is the potential diversion of freeway trips to adjacent surface streets to avoid queues at the meters. Extensive evaluations of existing metering systems show that adjustments in traffic patterns, after metering is implemented, take many forms.[1] However, it is possible to predict the likely impacts of metering before it is installed. Factors that enter into the analysis include trip length, queue length, entry delay, and especially the availability of alternate routes. The impact of attractive and efficient alternate routes can be a key factor in the effectiveness of a ramp metering system.[33] The probable new traffic patterns, including diversion, can then either be accommodated in the design and operation of the system, or become part of a decision that metering is not feasible.

Metering may, in fact, divert some short trips from the freeway. In concept, freeways are not intended to serve very short trips, and diverting some trips may even be desirable if there are alternate routes that are under-utilized. Diverting traffic from high volume, substandard, or other problem ramps to more desirable entry points should be an objective of metering where it is feasible. Such an action does require a thorough analysis of the alternate routes and the impacts of diversion on those routes, and improvements on the alternate routes when and where they are needed.

In Portland, city officials were very concerned about entrance metering creating problems on parallel streets. Before the meters on I-5 were installed, the city and State agreed that if volumes on adjacent streets increased by more than 25 percent during the first year of operation, the State would either abandon the project or adjust the meters to reduce the diversion below the 25 percent level. Following meter installation, the increase in local street volume was not substantial. Evaluations of the impact of metering on adjacent streets have been conducted in Los Angeles, Denver, Seattle, Detroit, and other cities. Significant diversion from the freeway to surface streets did not occur in any of these locations. Formal and informal agreements are common between State and local jurisdictions in connection with metering projects, and close advance coordination between jurisdictions is highly recommended.[1]

In some cases, there may not be feasible alternates routes, due to barriers such as rivers, railroads, or other major highways. Metering still can and does operate effectively where diversion is not an objective of the system. The systems in Denver, Northern Virginia, and Chicago, for example, operate under a so-called non-diversionary strategy. In these systems, metering is sometimes terminated at least until the queue dissipates. (See discussion of Houston ramp metering above). Significant benefits in freeway flow and accident reduction still result from nondiversionary metering. The onset of mainline congestion consistently begins later in the peak period and ends earlier. On many days, the mainline does not break down at all. Accidents and accident rates are also reduced. For example, in Denver it was observed that many drivers entered the freeway earlier in the morning. Peaks or spikes in volumes were thus leveled out over a longer period of

time resulting in better utilization of freeway capacity.[34]

ENFORCEMENT

The effectiveness of ramp metering, like that of any other traffic regulation, is largely dependent on voluntary driver compliance. As part of the public information effort, it should be made clear that ramp meters are traffic control devices that must be obeyed.[1] The laws and penalties should be clearly explained. In cities where the advance publicity has been positive and plentiful, violation rates has been lower. Again, as with any other regulation, enforcement is needed. Cooperation with police agencies is essential. Effective enforcement requires good enforcement access, a safe area for citing violators, adequate staff, support by the courts, and good signs and signals that are enforceable. Enforcement needs must be considered and accommodated **early** in the project development and design stages. Enforcement personnel should also be included **early** on in the planning and design of ramp metering projects. Compliance is critical to the success of a ramp metering system. Compliance rates, have generally been good in most areas across the country. However, violations are contagious and can multiply quickly. The result can be an extremely ineffective ramp metering system.

EQUITY

The complaint that ramp metering favors longer trips at the expense of shorter trips can be a controversial issue.[1] Close-in residents argue they are deprived of immediate access to the freeway, while suburban commuters can enter beyond the metered zone and receive all the benefits without the ramp delays.

Again there are strategies that have been employed to mitigate the equity issue. Initial metering in Detroit operated only in the outbound direction to minimize the city-suburb equity problem. Once the effectiveness of the metering was established, the system was expanded with less objection. This strategy was used in Atlanta where northbound I-75, leaving the city during the evening peak, will be the first section metered. [33] In Seattle, the system was designed to allow more restrictive metering rates farther away from downtown. With the long trip length, motorists originating from the suburbs have the most to gain from improved freeway conditions. The minor additional delay experienced at the meters is more than offset by the reduced mainline travel times. In Milwaukee, where the question of equity has been a limiting factor in the expansion of metering, it is now proposed to expand the system by metering each ramp that contributes traffic to congested freeway segments. Metering rates will be designed to be comparable for all ramps. For example, if it is determined a 10 percent reduction in demand is needed on the freeway segment, metering rates will be established to reduce all ramp volumes by 10 percent. In addition, each ramp metering rate will be adjusted to the extent possible in order to ensure average motorist delays are about equal for outlying ramps and for closer in ramps.[35] In Dallas, there was concern that suburbs were being favored over areas closer to the central business district. Ramp counts and license plate studies revealed that approximately as many vehicles were exiting the freeway before they reached downtown as were entering downstream of the adjacent suburbs, so equity was achieved.[36]

Even if only a few drivers experience increased travel times, there may still be objections simply because some have to wait at the ramps and others do not. A reasonable analogy can be made between a metered freeway and a signalized arterial. Vehicles entering an arterial from a minor

street must generally wait at a traffic signal while traffic already on the arterial is given priority. In both cases, the freeway and the arterial, the entering vehicles experience some delay in order to serve the higher volume facility. [1]

5.5 EXAMPLES IN RAMP CONTROL

There is extensive documentation of ramp control systems in the literature, much of which are cited in the reference lists in this handbook. An excellent summary of ramp control status, Ramp Metering Status in North America, was published by FHWA in 1995.[1] The history and case studies cited below were adapted from that report.

HISTORY OF RAMP CONTROL

The first metered ramp, as we know it today, was installed in Chicago on the Eisenhower Expressway in 1963. This first application, however, was preceded by successful tests of the effectiveness of metering traffic entering New York tunnels, and by ramp closure studies in Detroit. In Los Angeles, ramp metering began in 1968. That system has been expanded continually until there are now over 800 ramp meters in operation in L.A. County—the largest system in North America. Currently ramp meters are in operation in 23 metropolitan areas in North America These metering systems vary from a fixed time operation at a single ramp to computerized control of every ramp along many kilometers of a freeway.

Many reports have been written that document the potential successes and benefits of ramp metering. However, the true measure is in the continued growth of ramp metering installations. Since 1989, the number of operating meters in North America has increased from about 1600 to more than 2300, an increase of about 45 percent. Additionally, many existing systems are proposing expansions and/or upgrades. On the planning side, new ramp metering is being considered in numerous other cities as part of ITS early deployment plans or feasibility studies. By the year 2000, at least 33 cities in the United States and Canada will have functioning ramp meters. This will be 11 more systems than existed in 1989.

ENTRANCE RAMP METERING CASE STUDIES

The abbreviated case studies presented here are just a few examples of effective ramp metering operations. The benefit statistics presented are not consistent from city to city as there is no uniform evaluation criteria. Additionally, the measures of effectiveness (MOEs) vary depending on the objectives of the system. Further, complicating the matter, many ramp metering installations are implemented at the same time as other freeway improvements such as increased capacity, high-occupancy vehicle (HOV) lanes, surveillance systems, traffic information systems, and incident management programs. In these cases, it is not always possible to evaluate the individual components of the larger projects. The conditions of the evaluations of these case studies are noted for each discussion.

Portland, Oregon

The first ramp meters in the Pacific Northwest were installed along a 10 kilometer section of I-5 in Portland in January 1981. The meters are operated by the Oregon Department of Transportation. I-5 is the major north/south link, and is an important commuter route through the metropolitan area. This initial system consisted of 16 metered ramps between downtown Portland and the Washington state line. Nine of the meters operated in the

northbound direction during the p.m. peak, and seven controlled southbound entrances during the a.m. peak. The meters operate in a fixed time mode. There are currently 58 ramp meters operating on 5 different freeways.

Prior to metering, it was common along this section of 1-5 for platoons of vehicles to merge onto the freeway and aggravate the already congested traffic. The northbound PM peak hour average speed was 26 ki/h. Fourteen months after installation, the average speed for the same time period was 66 ki/h. Travel time was reduced from 23 minutes (but highly variable) to about 9 minutes. Premetered conditions in the southbound a.m. peak were much less severe, hence the improvements were smaller. Average speeds increased from 64 to 69 kph, resulting in only slight reductions in southbound travel times.

Additional benefits that were evaluated for the p.m. peak period included fuel savings and a before-and-after accident study. It was estimated that fuel consumption, including the additional consumption caused by ramp delay, was reduced by 2040 liters of gasoline per weekday. There was also a reduction in rearend and sideswipe accidents. Overall, there was a 43 percent reduction in peak period traffic accidents. [37]

Minneapolis/St. Paul, Minnesota

The Twin Cities Metropolitan Area Freeway Management System is composed of several systems and subsystems that have been implemented over a 25-year period by the Minnesota Department of Transportation. The first two fixed time meters were installed in 1970 on southbound I-35E north of downtown St. Paul. In November 1971, these were upgraded to operate on a local traffic responsive basis and 4 additional meters were activated. This 8-kilometer section of I-35E has been evaluated periodically since the meters were installed. The most recent study shows, that after 14 years of operation, average peak hour speeds remain 16 percent higher, from 60 to 69 ki/h, than before metering. At the same time, peak period volumes increased 25 percent due to increased demand. The average number of peak period accidents decreased 24 percent, and the peak period accident rate decreased by 38 percent.

In 1974, a freeway management project was activated on a 27-km section of I-35W from downtown Minneapolis to the southern suburbs. In addition to 39 ramp meters, the system included 16 closed-circuit television (CCTV) cameras, 5 dynamic message signs (DMS), a 2-km zone of highway advisory radio (HAR), 380 vehicle detectors, and a computer control monitor located at the MnDOT Traffic Management Center in Minneapolis. This project also included extensive “freeway flyer” (express bus) service, and 11 ramp meter bypass ramps for HOV's. An evaluation of this project after 10 years of operation shows that average peak period freeway speeds increased from 55 to 74 ki/h, or 35 percent. Over the same 10-year span, peak period volumes increased 32 percent, the average number of peak period accidents declined 27 percent, and the peak period accident rate declined 38 percent. Over one million dollars a year in road user benefits are attributed to reduced accidents and congestion. This system also has positive environmental impacts. Peak period air pollutant emissions, which include carbon monoxide, hydrocarbons, and nitrogen oxides, were reduced by just under 2 million kilograms per year. [38]

Over 300 additional ramp meters have been implemented from 1988 to 1995, and there are currently 400 meters in operation. Further projects are now in the design and construction phases. The plans are to

complete the ramp metering system which will cover the entire Twin Cities freeway network over the next five years.[39] The success of the Twin Cities system has shown that the staged implementation of a comprehensive freeway management system on a segment-by-segment, freeway-by-freeway basis, over a long period of time, is an effective way of implementing an area-wide program.

Seattle, Washington

In September 1981, the Washington State Department of Transportation (WSDOT) implemented metering on I-5 north of the Seattle Central Business District. Initially the system, which is named FLOW (not an acronym), included 17 southbound ramps that were metered during the a.m. peak, and 5 northbound ramps that were metered during the p.m. peak. Currently, the ramp metering system includes 54 meters on I-5, I-90, and SR 520. These meters are all operated under centralized computer control. Future expansion plans include additional ramp meters on SR 520 east of Lake Washington, on all of I-405, and on I-5 south of Seattle.

One evaluation of the initial 22 meter system showed that between 1981 and 1987, mainline volumes during the peak traffic periods increased 86 percent northbound and 62 percent southbound. Before the installation of metering, the travel time on a specific 11-km course was measured at 22 minutes. In 1987, the travel time for the same course was measured at 11.5 minutes. Over the same 6-year time period, the accident rate decreased by 39 percent.[40]

A somewhat unique application of metering was implemented in Seattle on SR-520 in 1986. While diversion caused by metering is often controversial, one of the objectives of metering SR-520 was to reduce commuter diversion through a residential neighborhood. The meters were installed on the two eastbound ramps on SR-520 between I-5 and Lake Washington. One of these ramps, the Lake Washington Boulevard on-ramp, is the last entry onto SR-520 before the Evergreen Point Floating Bridge. Because there were no bottlenecks downstream of this ramp, traffic would normally flow freely on the bridge and beyond. Motorists, especially commuters from downtown Seattle, were using residential streets to reach the Lake Washington Boulevard on-ramp to avoid congestion on SR-520. This on-ramp, however, was a major contributor to congestion on SR-520 because of the high entering volumes. By metering the ramp, it was anticipated that traffic diverting through the adjacent neighborhood from downtown would be discouraged by the delay caused by the meter. Motorists would instead use the Montlake Boulevard on-ramp, which was also metered at the same time. A HOV bypass lane was also installed at the Montlake Boulevard on-ramp. Two other objectives of this project were to improve flow on SR-520 and to encourage increased transit use and carpooling.

An evaluation of this two-ramp meter "system" after four months of operation showed there was a 6.5 percent increase in mainline peak period volume, a 43 percent decrease in the volume on the Lake Washington Boulevard on-ramp, an 18 percent increase in the volume on the Montlake Boulevard on-ramp, and a 44 percent increase in HOVs using the Montlake Boulevard on-ramp.[41] Another indication of the effectiveness of the combination of the HOV bypass and the improved SR-520 flow is a decrease of 3 minutes in METRO (King County Department of Metropolitan Services) transit travel times for buses traveling from downtown to the east, and a 4-minute

decrease for buses traveling from University District to the east. The reliability of the bus travel times also improved, and METRO adjusted the schedules for these routes accordingly.

In 1993, the WSDOT implemented weekend ramp metering for the first time. Three ramps north of Seattle on southbound I-5 have been metered several hours due to heavy weekend volumes. Because of this success, in March of 1995, weekend metering was expanded to include four additional southbound ramps.

In April of 1995, WSDOT began operating seven southbound I-5 meters during the evening commute. This is WSDOT's first implementation of metering both directions of a corridor during the same peak period. The motivation behind this operational change is that the traditional reverse commute direction has become increasingly congested. Prior to this change, metering along this section had operated southbound (inbound toward Seattle) during the morning commute and northbound (outbound) during the evening commute.

Denver, Colorado

The Colorado Department of Transportation activated a pilot project to demonstrate the effectiveness of ramp metering on a section of northbound I-25 in March 1981. The initial system consisted of five local traffic-responsive metered ramps operated during the a.m. peak on a 4.7-km section of I-25 south of the city. Periodic after-evaluations revealed significant benefits. An 18-month after study showed that average peak period driving speed increased 57 percent and average travel times decreased 37 percent. In addition, incidence of rearend and side-swipe accidents declined 5 percent due to the elimination of stop-and-go conditions.

The success of the pilot project led to expansion of the system. In 1984, a central computer was installed and a System Coordination Plan was implemented that permits central monitoring and control of all meters. Since 1984, additional ramp meters have been added, until reaching the current number of 28. In late 1988 and early 1989, a comprehensive evaluation of the original metered section was conducted. A number of changes occurred between 1981 and 1989, the most significant of which was the completion of a new freeway, C-470, which permitted more direct access to I-25 from the southwest area and generated higher demand for I-25. Volumes during the 2-hour a.m. peak period increased from 6200 vph in 1981 to 7350 vph in 1989 (on 3 lanes). Speeds measured in late 1988 decreased from the original evaluation, but remained higher than the speeds before metering was implemented: 69 ki/h before, 85 ki/h after, in 1981, and 80 ki/h in late 1988. The frequency of accidents during the a.m. peak period did not increase between the time of original evaluation and 1989. As a result, the accident rate decreased significantly because of the increased volumes. Rearend and sideswipe type accidents decreased by 50 percent during metered periods.

An interesting unplanned "evaluation" of the system occurred in the Spring of 1987. To accommodate daylight savings time, all of the individual ramp controllers were adjusted one hour ahead. Unfortunately, the central computer clock was overlooked. The central computer overrode the local controllers, and metering began an hour late. Traffic was the worst it had been in years. However, this oversight did have a bright side for the Department of Transportation. Since this incident, the media has been even more supportive of ramp metering than before. [34]

In 1988, the Colorado Department of Transportation conducted a study to evaluate different levels of ramp metering control. The study compared ramp meters operating in local traffic-responsive mode versus meters operating under centralized computer control. The results showed that if local traffic-responsive metering could maintain freeway speeds above 90 ki/h, centralized control offered little or no additional benefit. However, if local traffic-responsive metering was unable to maintain speeds near the posted speed limit of 90 ki/h, centralized control was very effective. Data showed speeds increased 35.5 percent, from 50 to 68 ki/h, and vehicle hours of travel were reduced by 13.1 percent.[42] This evaluation shows the importance of implementing operating strategies that correspond to the needs of the freeway network.

Detroit, Michigan

Ramp metering is an important aspect of the Michigan Department of Transportation's (MDOT) Surveillance Control and Driver Information (SCANDI) System in Detroit. The SCANDI metering operation began in November 1982 with six ramps on the eastbound Ford Freeway (I-94). Nineteen more ramps were added on I-94 in January 1984 and three more in November 1985. An evaluation performed by Michigan State University for MDOT determined that ramp metering increased speeds on I-94 by about 8 percent. At the same time, the typical peak hour volume on the three eastbound lanes increased to 6400 vehicles per hour from an average of 5600 VPH before metering. In addition, the total number of accidents was reduced nearly 50 percent, and injury accidents came down 71 percent. The evaluation done by Michigan State also showed that significant additional benefits could be achieved by metering the three freeway-to-freeway connectors on this section of I-94.[44]

Austin, Texas

In the late 1970s, in Austin, the Texas Department of Transportation implemented traffic responsive meters at 3 ramps along a 4.2 km segment of northbound I-35 for operation during the a.m. peak period. This section of freeway had two bottleneck locations that were reducing the quality of travel. One was a reduction from 3 to 2 lanes and the other was a high volume entrance ramp just downstream of a lane drop. Metering resulted in an increased vehicle throughout of 7.9 percent and an increase in average peak period mainline speeds of 60 percent through the section. The meters were removed after the reconstruction of I-35 eliminated the lane drop in this section.[44] This situation shows the versatility of ramp metering in that it can also be used effectively as a temporary solution.

Long Island, New York

At the other end of the spectrum from Austin is the INFORM (Information For Motorists) project on Long Island. The INFORM project covers a 64-km long by 8-km wide corridor at the center of which is the Long Island Expressway (LIE). Also included in the system is an east-west parkway, an east-west arterial and several crossing arterials and parkways, a total of 207 kilometers of roadways. System elements include 70 metered ramps on the LIE and the Northern State/Grand Central Parkway.

In 1989, an analysis of the initial metered segment was conducted after 2 months of operation. For the peak period, the study showed a 20 percent decrease in mainline travel time (from 26 to 21 minutes) and a 16

percent increase in average speed (from 47 to 56 ki/h). Motorists entering at metered ramps also experienced an overall travel time reduction of 13.1 percent and an increase in average speed from 37 to 45 ki/h. The MOEs for this project include vehicle emissions. For this initial segment, the analysis indicates there was a 6.7 percent reduction in fuel consumption, a 17.4 percent reduction in carbon monoxide emissions, a 13.1 percent reduction in hydrocarbons, and a 2.4 percent increase in nitrous oxide emissions. The last is associated with the higher speeds. Initial observations of the effect of metering the 4-lane parkway on the INFORM project indicates the benefits may be even greater than those achieved on wider freeways. Intuitively this makes sense, because the impact of an unrestricted merge on only two lanes (in one direction) can be severe. [45]

A more extensive evaluation of the INFORM project was completed in 1991. Data from this study showed much more conservative results. It is believed that this study is more representative of the true traffic conditions. The main reason for this is related to the "queuing off" (shut-down of the meter due to excessive queuing) of the ramp meters. The original study did not include areas where metering was usually shut off due to heavy ramp volumes, while the later study accounted for all ramps. This evaluation showed that while throughput had increased only about 2 percent, the average mainline speeds had increased from 64 to 71 ki/h, or about 9 percent. However, for two separate bottleneck locations, data showed increases of 53 to 84 and 53 to 89 ki/h, or gains of about 36 and 40 percent respectively. This evaluation also included calculation of a "congestion index." This index is the proportion of detector zones for which speeds were less than 48 ki/h (30 mi/h). While no benefit was shown in the evening peak period, the morning peak period showed an improvement of 25 percent in the congestion index. The accident frequency rate also showed encouraging improvement, with a 15 percent reduction as compared to the control section. [46]

San Diego, California

In San Diego, ramp metering was initiated in 1968. That system, installed and operated by the California Department of Transportation (Caltrans), now includes 134 metered ramps on 110 plus kilometers of freeway. No detailed evaluations of metering have been conducted on the San Diego system since the early installations, but sustained volumes of 2200 vph to 2400 vph, and occasionally even higher, are common on San Diego metered freeways. A noteworthy aspect of the program is the metering of eight freeway-to-freeway connector ramps. Metering freeway-to-freeway connectors requires careful attention to storage space, advanced warning, and sight distance. If conditions allow, freeway connector metering can be just as safe and effective as other ramp metering. [47]

SUMMARY OF RAMP METERING BENEFITS

Metering entrance ramps can significantly improve mainline traffic flow. These case study evaluations, as well as others, show that metering consistently increases travel speeds and improves travel time reliability, both of which are measures of reduced stop-and-go, erratic flow. It should be emphasized that these benefits occurred even though, in most instances, mainline volumes had significantly increased. Metering helps smooth out peak demands that would otherwise cause the mainline flow to breakdown. A strong case can be made from the data reported that metering actually increases the throughput of a freeway. The

data from Minneapolis, San Diego, Seattle, Detroit and Denver shows mainline volumes well in excess of 2100 vph per lane on metered sections, and sustained volumes in the range of 5 percent to 6 percent greater than for pre-metered conditions. Improved traffic flow, particularly the reduction in stop-and-go conditions, also reduces certain vehicle emissions. This has been shown in both the INFORM project and in the Twin Cities Freeway Management System.

The other direct benefit, but one that has not been fully quantified, is the reduction in accidents attributed to metering. The Dallas corridor provided a unique opportunity to compare vehicle crash experience in a ramp metering system.[48] Evaluation studies showed significant improvements in system operating characteristics as compared to the "before" conditions. However, during the first year of operation, metering was exercised only in the peak direction of flow. During that year, crashes in the metered direction decreased by 24 percent as compared to the previous year, while crashes in the non-metered direction increased by 12 percent. During the first three years of metering, total weekday (24 hour/day) crashes increased by 8 percent while accidents during ramp metering decreased by 18 percent. The other case studies presented in this report consistently show a reduction in crash rates of 24 to 50 percent. Minnesota Department of Transportation estimates over 1000 vehicle crashes are prevented each year on Minneapolis/St. Paul metropolitan area freeway due to ramp metering.[39] However, the benefits derived from accident reduction go well beyond the direct costs related to medical expenses and vehicle damage. To illustrate, assume an incident blocks one lane of three at the beginning of the peak period on a freeway with a 2-hour peak demand of 6000 vph. Studies show that an accident blocking one of three lanes reduces capacity by 50 percent. A 20-minute blockage would cause 2100 vehicle-hours of delay and a queue over 3 kilometers long, and take 2 1/2 hours to return to normal, assuming there were no secondary accidents or incidents. Clearly the safety aspects of metering are a major benefit.

5.6 REFERENCES

1. Piotrowicz, G., and Robinson. J.R. *Ramp Metering Status in North America*. Report DOT-T-95-17. FHWA, U.S. Department of Transportation, Washington, DC, 1995.

2. *Highway Capacity Manual*. Highway Research Board Special Report No. 209, Third Edition, TRB, National Research Council, Washington, DC, 1994.

3. *Manual on Uniform Traffic Control Devices*. FHWA, U.S. Department of Transportation, Washington, DC, 1988.

4. *ITS Architecture Summary*. Joint Architecture Team, FHWA, U.S. Department of Transportation, Washington, DC, 1996.

5. Sumner R., et. al. *Freeway Management Handbook*, FHWA, U.S. Department of Transportation, Washington, DC, 1983. (Four volumes.)

6. Wilshire, R., et al. *Traffic Control Systems Handbook - Revised Edition - 1985*, FHWA-1P-85-11, FHWA U.S. Department of Transportation, Washington, DC, 1985.

7. Gordon, R., et. al. *Traffic Control System Handbook - Revised Edition - 1996*, FHWA-SA-95-032, FHWA, U.S. Department of Transportation, Washington, DC, 1996.

8. *Hanshin Expressway Traffic Control System*. Osaka Management Division, Hanshin Expressway Public Corporation, Osaka, Japan, March 1973.

9. *Traffic Control of Metropolitan Expressway*, Metropolitan Expressway Public Corporation, Tokyo, Japan, May 1973.

10. Wattleworth, J.A., Courage, K.G., and Carvell, J.D. *An Evaluation of Two Types of Freeway Control Systems*. Report 488-6, Texas Transportation Institute, Texas A&M University System, College Station, TX, 1968.

11. Newman, L., Dunnet, A., and Meis, G. Freeway Ramp Control - *What It Can and Cannot Do*. Freeway Operation Department, District 7, California Division of Highways, February 1969.

12. McCasland, W.R. *Freeway Ramp Control System*. Report 24-26F, Texas Transportation Institute, Texas A&M University System, College Station, TX, 1969.

13. Institute of Transportation Engineers, Technical Committee 4M-11. "Displays for Metered Freeway Entrance Ramps." *ITE Journal*, Vol. 54, No. 4, Washington, DC, April 1984, pp 14-18.

14. Masher, D. P., et al. "Guidelines for Design and Operation of Ramp Control Systems." *NCHRP Project 3-22 Final Report,* December 1975. (Loan copy available on request to NCHRP.)

15. Courage, K. "Freeway Surveillance and Control: A State of the Art Summary." Paper presented at the Computer Traffic Control Workshop. University of Minnesota, March 20-22, 1974.

16. Drake, J.S., Schofer, J.L., and May, Jr., A.D. "A Statistical Analysis of Speed Density Hypotheses." *Highway Research Record 154*, HRB, National Research Council, Washington, DC, 1967, pp 53-87.

17. Breiman, L. "A Case Study - LAAFSCP." Paper presented at the Urban Freeway Surveillance and Control Short Course. Technology Service Corporation, Bethesda, Maryland, March 1974.

18. McDermott, J.M., Kolenko, J.J., and Wojcik, R.J. Chicago Area Expressway Surveillance and Control: Final Report. *Expressway Surveillance Report No. 27*. Illinois Department of Transportation, Springfield, IL, March 1979.

19. Blumentritt, C. W., et al. "Guidelines for Selection of Ramp Control Systems." *NCHRP Report 232,* Washington, DC, May 1981.

20. Carvell, J.D. *Freeway Operations Manual. Dallas Corridor Study*. Research Report 953-11, Texas Transportation Institute, Texas A&M University System, College Station, TX, 1974.

21. Drew, D.R., LaMotte, L.R., Wattleworth, J.A., and Buhr, J.H. "Gap Acceptance in the Freeway Merging Process." *Highway Research Record 279*, HRB, National Research Council, Washington, DC, 1969, pp 137-149.

22. Buhr, J.H., McCasland, W.R., Carvell, J.D., and Drew, D.R. "Design of Freeway Entrance Ramp Merging Control Systems." *Highway Research Record 279*, HRB, National Research Council, Washington, DC, 1969. pp 137-149.

23. Drew, D., McCasland, W.R., Pinnell, C., and Wattleworth, J.A. *The Development of An Automatic Freeway Merging Control System.* Research Report 24-19. Texas Transportation Institute, Texas A&M University System, College Station, TX, 1966.

24. Wattleworth, J.A., and Berry, D.S. "Peak-Period Control of a Freeway System - Some Theoretical Investigations." *Highway Research Record 89*, HRB, National Research Council, Washington, DC, 1965, pp 15-20.

25. May, A.D., Jr. "Peak-Period Analysis and Control of a Freeway System." *Highway Research Record 157,* HRB, National Research Council, Washington, DC, 1967, pp 15-20.

26. Courage, K.F., and Levin, M. *A Freeway Corridor Surveillance Information and Control System*. Research Report 438-8, Texas Transportation Institute, Texas A&M University System, College Station, TX, December 1968.

27. Messer, C. J. "Development and Evaluation of a Multi-Level Freeway Control System for the Gulf Freeway." Paper presented to the Freeway Operations Committee, 50th Annual Meeting of the Highway Research Board, Washington, DC, 1971.

28. Looze, D.P., et al. "On Decentralized Estimation and Control with Application to Freeway Ramp Metering." *IEEE Transactions on Automatic Control*, Vol. AC-23, No. 2, April 1978, pp. 268-275.

29. Messer, C.J., et al. *Ramp Adaptive Metering Bottleneck Optimization (RAMBO) User's Manual*. Research Report 1232-31, Texas Transportation Institute, College Station, November 1994.

30. *TSMC SC&DI Operations/Implementations Plan*. Washington State DOT, October 1994.

31. Carvell, J.D. *Dallas Corridor Study-Final Report*, Report 953-29, Texas Transportation Institute, College Station, TX, August , 1976.

32. Personal Correspondence with Dr. Carroll Messer, Research Engineer, Texas Transportation Institute, College Station, TX, November 1996.

33. Yagar, S., "Predicting Impacts of Freeway Ramp Metering on Local Street Flows and Queues*", ITE 1989 Compendium of Technical Papers*, September 1989.

34. Corcoran, L.J., and Hickman, G.A. "Freeway Metering Effects in Denver." *ITE 1989 Compendium of Papers*, September 1989.

35. *Recommended Ramp Metering Strategy for the Atlanta Regional Advanced Transportation Management System*, NET/TRW for the Georgia Department of Transportation, September, 1993.

36. *A Freeway Management System Plan for the Milwaukee Area.* Southeastern Wisconsin Regional planning Commission, November 1988.

37. I-5 North Freeway Ramp Metering, Portland, Oregon: Project Development *Operation*, Oregon Department of Transportation-Metropolitan Branch, June 1982.

38. *Freeway Traffic Management Program: Status Report*, Minnesota Department of Transportation, January 1989.

39. *Overview of the Minnesota DOT Ramp Metering Program*, Minnesota DOT, 1995.

40. Henry, K.C., and Mehyar, O. *Six-Year FLOW Evaluation*, Washington State Department of Transportation, District 1, January 1989.

41. *Evaluation of the Eastbound SR-520 Traffic Management System*, Washington State Department of Transportation, no date.

42. Lipp, L.E., Corcoran, L.J., and Hickman, G.A. "The Benefits of Central Computer Control for the Denver Ramp Metering System." *TRB Annual Meeting*, January 1991.

43. Maleck, T.L., Kostyniuk, L.P., Taylor, W.C., and Hamad, A.I. *An Evaluation of the Detroit Freeway Operations (SCANDI) Project*, Michigan State University, June 1988.

44. Marsden, B. *Ramp Meters and Travel Quality in Austin, Texas, Texas DH&PT*, May 1981.

45. *Before and After Study, Ramp Metering, Eastbound Long Island Expressway, Suffolk County*, Technical Memo., James H. Kell & Assoc., April 28, 1989.

46. *INFORM Evaluation, Volume 1: Technical Report* , FHWA-RD-91-075, January 1992.

47. Harvey, S. "Ramp Meters-What They Can and Cannot Do." *District 6 Newsletter*, July-August 1985.

48. *Evaluation of Ramp Control Operation: Dallas Corridor Study, Report 836-1*, Texas Transportation Institute, College Station, TX, August 1972.

MODULE 6. HOV TREATMENTS

TABLE OF CONTENTS

6.1 INTRODUCTION 6-5
TREATMENTS 6-6
MODULE OBJECTIVES 6-6
MODULE SCOPE 6-7

6.2 DESIGN PROCESS 6-7
IDENTIFY PROBLEMS/NEEDS 6-7
IDENTIFICATION OF PARTNERS 6-8
CONSENSUS BUILDING 6-10
ESTABLISH GOALS AND OBJECTIVES 6-10
ESTABLISH PERFORMANCE CRITERIA / MOES 6-10
DEFINE FUNCTIONAL REQUIREMENTS 6-11
IDENTIFY AND SCREEN TECHNOLOGY 6-11
System Planning 6-13
IMPLEMENTATION 6-15
EVALUATION 6-16

6.3 TECHNIQUES AND TECHNOLOGIES 6-18
HOV FACILITIES 6-18
Operational Considerations 6-18
HOV Roadway Operations 6-20
Operating Efficiency 6-20
Considerations for 2+ Versus 3+ Occupancy Requirement 6-20
Hours of Operations 6-22
Design Elements 6-22
Design Speed 6-22
Cross Section Width 6-22
Vertical Clearance 6-23
Signs and Markings 6-23
Signs 6-23
Markings 6-23
Enforcement 6-25
Goals and Objectives 6-25
Methods 6-27
Enforcement Area Design 6-28
ACCESS 6-29
Ramp Connections 6-29
Design 6-29

PRIORITY ACCESS 6-34
Ramp Metering 6-34
Bypass Lanes 6-34
Signs 6-35
Markings 6-35
PARK-AND-RIDE FACILITIES 6-35
Location of Facilities 6-36
Establishment of Site Selection Criteria 6-36
Identification of Alternative Sites 6-38
Evaluation of Alternatives 6-38
Determining Size of Facility 6-38
Typical Designs 6-39
Internal Lot Layout 6-39
Functional Area Design. 6-39
Internal Circulation. 6-40
Amenities. 6-40
Pavement and Drainage. 6-40
Landscaping. 6-41
Lighting. 6-41
Security. 6-41
Environmental Considerations. 6-41
Roadway Interface 6-41
General Access and Egress Considerations 6-42
Access Points 6-42
Access Roadways 6-42
Traffic Control Devices and Traffic Signals 6-42
Handicapped Considerations 6-42
Parking 6-42
Capacity 6-42
Location 6-42
Design 6-43
Signing and Marking 6-43
Bus Loading Area 6-43
PARKING INCENTIVES 6-44
Preferential Parking 6-44
Free/Low-Cost Parking 6-44
Differential Parking Rates 6-44
Benefits of Parking Incentives 6-44
Barriers to Implementing Parking Incentives 6-45
EMERGING TECHNOLOGIES 6-46
Enforcement 6-46
Movable Barrier 6-46

6.4 LESSONS LEARNED 6-48
PLANNING 6-48
Public and Political Support and Consensus Building 6-50
Public Involvement 6-51
System Approach to HOV Planning 6-52
Possible Pitfalls in Implementation 6-52
Development Concerns for Park-and-Ride Facilities 6-53
Program Planning 6-53
Design Requirements 6-55
Operation and Maintenance 6-55
Program Marketing 6-56
DESIGN/CONSTRUCTION 6-57
Tradeoffs in Retrofitting Cross Sections 6-57
Impacts of Ingress/Egress on Adjacent Freeway and Arterial System 6-57
Signing Guidelines 6-59
Pavement Marking Guidelines 6-60
OPERATIONS 6-61
Surveillance, Communication, and Control 6-61
Incident Response 6-62
Enforcement Strategies 6-62
Other Users 6-63
MAINTENANCE 6-64
CONCURRENT FLOW LANES 6-65

6.5 EXAMPLES OF HOV TREATMENTS 6-66
HOUSTON 6-66
SEATTLE 6-71

6.6 REFERENCES AND SUGGESTED READINGS 6-75
REFERENCES 6-75
SUGGESTED READING 6-77

MODULE 6. HOV TREATMENTS

Figure 6-1. HOV Lane in San Diego, CA.

6.1 INTRODUCTION

Preferential treatments for high occupancy vehicles (HOV) have proven to be one of the most flexible, cost effective alternatives for increasing the person-moving capacity of congested metropolitan transportation systems. The concept emphasizes *person movement* rather than traditional *vehicle movement*. It offers multi-person vehicles the opportunity to travel in reserved lanes that allow higher operating speeds and more reliable travel times. HOV treatments are not appropriate in every situation, but their applications are growing as more and more metropolitan areas face the challenge of improving mobility, particularly during peak travel periods.

Constraints on metropolitan mobility have been identified as among the most serious transportation issues affecting the economic and social vitality of the nation. Many regions of the country, recognizing this critical importance of continued mobility, have established the following areawide goals:

- Reduce vehicle-miles of travel.
- Conserve energy.
- Reduce air pollution.

HOV facilities can be a major component in regional efforts to improve the operational efficiency of a freeway by:

- Increasing the people-moving capacity of the facility.
- Offering high-speed travel to a larger number of people (to decrease the average travel time and make it more predictable).

- Providing an incentive for people to share rides (to increase the number of persons carried per vehicle).

The objective of HOV lanes is, of course, to move more people in fewer vehicles than conventional highway lanes. A dramatic example of the person-moving capability of an HOV lane is on 4 km (2.5 mi.) of I-495 in New Jersey, between the New Jersey Turnpike and the Lincoln Tunnel entering New York City. Here a contraflow lane, a lane in the off-peak direction, separated from the adjacent lane with plastic posts during the morning peak, carries 35,000 people in 750 buses during the peak hour, and more than 70,000 people in 1650 buses in the 3-hour peak period. It would require 15 conventional freeway lanes to move the same number of people in the peak hour. Certainly, no other city has the employment density or transit use of the New York/New Jersey Metro area, but HOV facilities throughout the country are accomplishing the same objective by providing a reduced travel time and travel time reliability incentive for drivers to change mode to HOV. Existing freeway HOV lanes carry the person-equivalent of two to five conventional lanes, and do so at a higher level of service than the conventional lanes. If an HOV lane becomes congested, the occupancy requirement can be raised, or other adjustments can be made to accommodate additional vehicles and thus additional people.[1]

TREATMENTS

Priority treatments may be either dedicated mainline lanes or isolated improvements that bypass a bottleneck. These treatments include the following:

- Exclusive HOV Facility - Separate Right-of-Way.
- Exclusive HOV Facility - Freeway Right-of-Way.
- Concurrent Flow Lane.
- Contraflow Lane.
- Queue Bypass HOV Facility.

These treatments are best served when implemented with a number of other transportation demand management measures to encourage use such as:

- Rideshare matching services.
- Guaranteed ride home.
- Employer incentives through preferential parking.
- Parking cost incentives.
- Supporting facilities that help to collect and distribute passengers.
- Public information.
- Education programs.

MODULE OBJECTIVES

The objectives of this module are to accomplish the following:

- Present the process for developing HOV treatments within a freeway system.
- Provide a summary and description of different HOV treatments in use, including both HOV lanes and supporting facilities such as park-and-ride and preferential parking programs.

- Discuss planning, design, construction, operating, and maintenance issues associated with the different HOV treatments.

MODULE SCOPE

In view of the impressive HOV success stories of Houston, Seattle, and other cities (see **section 6.5**), a high interest exists in implementing HOV treatments within freeway systems. This module presents the basics for HOV treatments associated with freeways. Treatments on arterial roadways can also support the freeway system, however those treatments are not discussed here. Information on the different freeway treatments and their supporting parking-related facilities are presented along with discussion on issues associated with those facilities. An emphasis of this module is on the need for cooperation between the partners and the presentation of issues that need to be considered during the planning process.

6.2 DESIGN PROCESS

IDENTIFY PROBLEMS/NEEDS

High Occupancy Vehicle facilities are viable components of freeway systems. The needs that can be addressed by HOV facilities include the following:

- Increasing person-carrying capacity of a highway corridor.
- Reducing total travel time.
- Reducing or defer the need to increase highway vehicle-carrying capacity.
- Improving efficiency and economy of public transit operations.
- Reducing fuel consumption.
- Improving air quality.
- Inducing mode shift.

Factors that influence consideration of HOV facilities include:

- Congestion.
- Travel time savings.
- Person throughput.
- Vehicle throughput.
- Local agency support.
- Enforceability.
- Physical roadway characteristics.
- Support facilities.
- Environmental mitigation.
- Compatibility with other modes.

The two criteria that most commonly appear to influence HOV viability are *congestion* and *travel time savings*. Without existing or forecasted congestion, the HOV alternative offers no substantial benefits for single-occupant drivers to switch to carpool, vanpool, or bus. Although the definition of "congestion" varies from one locale to another, a good measure of congestion is average freeway speeds of 48 km/h (30 mi/h) or less during the peak hour, or 56 km/h (35 mi/h) or less during the peak period. In some instances, an HOV alternative has been considered for a congested freeway that could operate relatively smoothly with an added mixed-flow lane, but for which future congestion is predicted. Travel time savings has become one of the most reliable

predictors of HOV viability, and it must potentially exist to encourage mode shifts. For most treatments, a projected 5-minute or more savings per trip is generally recognized as a prerequisite. Time savings of less than 5 minutes may still justify consideration of queue bypass where a modest investment benefits many drivers.[3]

An assessment of HOV projects in six U.S. cities—Houston, Minneapolis-St. Paul, Pittsburgh, Seattle, Washington, DC, and Orange County, California—found the following common elements in the decision-making process: [4]

- *Corridor and areawide characteristics*. An awareness of the need to address increasing traffic congestion problems in the corridor had developed.
- *Lack of a fixed-guide way transit plan for the corridor*. No decision had been made on the development of a fixed-guide way transit system in the corridor where the HOV facility was ultimately developed.
- *Planned or scheduled highway improvements*. HOV projects were considered and implemented as part of an extensive program of highway improvements. This coordination helped maximize available resources and minimize impacts on implementation.
- *Project champion or champions*. Individuals in positions of authority in highway and transit agencies supported the HOV concept and promoted it throughout the project development process.
- *Legislative direction and policy support*. Legislative or agency policies and directives played an important role in the decision-making process in some HOV projects.

To assist with defining the problems that exist within the freeway system so that appropriate decisions can be made, an inventory of physical and organizational components is needed. Table 6-1 lists some of the specific types of information needed.

IDENTIFICATION OF PARTNERS

The key partners in the development of an HOV facility component of a freeway management system include the following:

- State Department of Transportation.
- Transit Authority (or Authorities).
- Federal Highway Administration.
- Metropolitan Planning Organizations.
- Enforcement Agencies.
- County or City Departments of Transportation.

In addition, the following can be key in developing a successful plan, rather than one that is challenged during the development and/or construction of a facility:

- Elected Officials.
- Media representatives.
- Citizens.
- Representatives from private businesses.
  - Major traffic generators.
  - Businesses with ride-sharing programs.

Table 6-1. Information Needs for Successful HOV Operations.

Person-Moving Capacity of the Roadway Facility
Potential for Support Facilities Such as Park-and-Ride Lots
Origins and Destinations
Activity Centers
Point of Origin
Average Trip Length
Trip Times
Traffic Operation on Freeway Mainlane
Average Delay
Peak Period Volumes
Average Travel Speeds
Travel Time
Bottlenecks
Location
Duration
Causes
Vehicle Occupancy
Percentage of Peak Period 2+ and 3+ Vehicles in the Traffic Stream
Future Demand
Growth Factor of Corridor
Growth Factor of Parallel Routes
Transit and Rideshare Patronage
Types of Operations That Could Benefit from HOV
Existing and Planned Transit Services
General Design Limitations of Existing Freeway Facility
Safety and Accident Data
Violation and Enforcement Data
Information on the Perception of Users, Non-Users, and the General Public

In some cases, for example the development of a park-and-ride facility has included other non-traditional partners. Other important partners that can be involved in the decision-making process include the following:

- Owners/operators of major traffic generators.
- Churches whose parking lots could be used for weekday commuters.

Successful HOV systems requires good coordination between respective governmental agencies. Two examples of areas where effective coordination has occurred include Houston and Seattle.[5] Agencies primarily involved in the implementation of HOV facilities in Houston included:

- Texas Department of Transportation.
- Metropolitan Transit Authority.
- Federal Highway Administration.

Agencies involved in Seattle included:

- Washington State Department of Transportation.

- Puget Sound Council of Governments.
- Seattle Metro Transit.
- Pierce County Transit.
- Municipality of Metropolitan Seattle (Metro).
- King, Pierce, and Snohomish Counties.
- City of Bellevue.
- Federal Highway Administration.

Additional information on the Houston and Seattle systems is provided later in this module.

CONSENSUS BUILDING

Having an overall consensus and developing good working relationships within and between agencies are important in implementing an HOV facility system. During initial meetings, identifying and understanding the differences in operational philosophies and priorities of the different partners will assist in the process. For example, major business associations may be more concerned with the location of ramps to and from an HOV lane, and whether those ramps will go into a parking garage, while a transit agency's priority may be to have the ramps exit to a downtown street.

The importance of public support needs to be recognized early in the process. Extensive public relations and media campaigns may be needed to show the public the benefits of the HOV facilities. An important lesson that has been learned from project failures to date is that the public must be involved and must be able to understand and appreciate the role that HOV systems can serve. Public participation will need to be focused and pursued at various levels.

A multi-agency review group (also called an advisory committee or steering committee) should be formed. Their role includes technical and policy guidance, concurrence powers at major decision points, coordination and liaison with others in the respective agencies, and outreach to greater public participation efforts as needed. Apart from the multiagency review group, focus groups composed of local civic associations, special interest groups, politicians, the media, and others may be necessary to link public participation with the process.[5]

ESTABLISH GOALS AND OBJECTIVES

After the relevant partners are identified, the goals and specific objectives that are to be addressed through this component of the freeway system are developed. Goals are broad statements of the intent of the system or of one of its components, whereas objectives are specific statements about what the system or component of that system will attempt to accomplish. A given goal may have more than one objective specified to reach that goal. Table 6-2 lists examples of goals and objectives that the team might develop for HOV facilities.

ESTABLISH PERFORMANCE CRITERIA / MOES

Performance criteria and measures-of-effectiveness need to be identified in order to assess the extent to which HOV treatments are meeting goals and objectives. Table 6-3 lists suggested objectives and measures of effectiveness.

Table 6-2. Examples of Goals and Objectives for HOV Treatments.

Category	Examples
Goals	• Produce a Better Operating Freeway • Increase the Person-Moving Efficiency of the Roadway Facility • Encourage Mode Change
Objectives	• Reduce User Travel Time • Reduce or Defer the Need to Increase Highway Vehicle Capacity • Improve Efficiency and Economy of Public Transit Operations • Reduce Fuel Consumption • Improve Air Quality by Reducing Air Pollution in a Corridor

Evaluations are necessary to ensure that the improvements are providing the desired benefits and that the expenditure of public funds is justified. Interest exists in the results of these evaluations among many groups, for example, transportation professionals, elected officials, and the general public. While there appears to be general agreement among transportation professionals that HOV facilities should be evaluated, a consensus does not exist regarding the most appropriate measures to use, the performance thresholds the projects should meet to be considered effective, or the preferred data collection techniques. The challenge to transportation professionals is to provide accurate and objective evaluations of HOV facilities that focus on key criteria and that can be easily understood by the different partners. Analyses and findings contained in several reports provide some guidance.[6-8]

DEFINE FUNCTIONAL REQUIREMENTS

The functional requirements of HOV treatments define specific actions or activities that are to be performed in order to achieve one or more of the objectives. The functions should be defined independent of the technology to be employed in the system, so that the focus is on *what* the system is designed to do rather than on *how* the system will be doing it. Table 6-4 presents examples of functional requirements for various objectives.

Example treatments include the following:

- HOV Lanes.
  - Facilities on separate rights-of-way.
  - Long-distance HOV lanes serving buses, vanpools, and carpools, and located within or adjacent to the freeway right-of-way.
- Priority Access at Ramps.
- Terminal Facilities.
- Preferential Parking.

IDENTIFY AND SCREEN TECHNOLOGY

It is appropriate to assess the actual technologies available to meet the functional requirements. The assessment can begin with the discussion of available treatments contained later in this module. Because of the high interest in HOV facilities, the Transportation Research Board, among others, frequently plans meetings to discuss

Table 6-3. Suggested Objectives and Measures of Effectiveness for HOV Facilities. [6]

Objective	Measures of Effectiveness
Improve the capability of a congested freeway corridor to move more people by increasing the number of persons per vehicle	• Actual and percent increase in the person movement efficiency • Actual and percent increase in average vehicle occupancy rate • Actual and percent increase in carpools and vanpools • Actual and percent increase in bus riders
Increase the operating efficiency of bus service in the freeway corridor	• Improvement in vehicle productivity (operating cost per vehicle-mile, operating cost per passenger, operating cost per passenger mile) • Improved bus schedule adherence (on-time performance) • Improved bus safety (accident rates)
Provide travel time savings and a more reliable trip time to HOVs utilizing the HOV facility	• The peak-period, peak-direction travel time in the HOV lane(s) should be less than the travel time in adjacent freeway lanes • Increase in travel time reliability for vehicles using HOV lane(s)
Favorable impacts on air quality and energy consumption	• Reduction in emissions • Reduction in total fuel consumption • Reduction in the growth of vehicle miles of travel (VMT) and vehicle hours of travel
Increase the per lane efficiency of the total freeway facility	• Improvement in the peak-hour per lane efficiency of the total facility
Not unduly impact the operation of the freeway mainlanes	• The level of service in the freeway mainlanes should not decline
Be safe and not unduly impact the safety of the freeway general purpose mainlanes	• Number and severity of accidents for HOV and freeway lanes • Accident rate per million vehicle miles of travel • Accident rate per million passenger miles of travel
Have public support	• Support for the facility among users, non-users, general public, and policy makers • Violation rates (percent of vehicles not meeting the occupancy requirement)
Be a cost-effective transportation improvement	• Benefit-cost ratio

Table 6-4. Examples of Functional Requirements for HOV Treatments.

Objective	Functions
Improve the capability of a congested freeway corridor to move more people by increasing the number of persons per vehicle	To support the HOV facilities, assist in the development of ridesharing programs
Provide travel time savings and a more reliable trip time to HOVs utilizing the HOV facility	Design facility so that incidents are appropriately handled (e.g., quick identification and removal)
Have public support	Involve public at appropriate stages during the process. Develop and execute a public relations campaign

the benefits, challenges, and other aspects of HOV treatments. Publications from these sources can provide additional information on treatments.

Factors that should be considered at this stage include cost (construction, operation, maintenance, replacement, etc.), operations and maintenance requirements, and personnel, equipment, and facility needs. The process for identifying and screening different technologies for inclusion in a freeway management system is often iterative, because there are multiple ways that different technologies can be combined to achieve an objective. For example, decisions made about an HOV treatment can influence decisions made about surveillance techniques.

System Planning

In order for HOV systems and facilities to be properly integrated within the freeway system, system planning needs to occur at all levels, including strategic planning, long-range system planning, short-range planning, and service or operations planning. At the strategic planning level, freeway and transit agencies need to determine their roles, missions, and types of HOV services they want to provide in a metropolitan area. (This type of activity is discussed in detail in **Module 2**.) Through the long-range planning process, agencies can ensure that HOV facilities and services are incorporated into the future design of freeway systems and that funding for capital-intensive facilities are programmed into area transportation improvement plans. The short-range planning process can be used to assess administrative, funding, and service changes that need to occur usually within a 5-year period. Service or operations planning is an ongoing activity—often on a route or corridor basis— and is intended to identify improvements to improve service efficiency and effectiveness of HOV facilities.

Figure 6-2 illustrates a system planning methodology that can be used to identify HOV alternatives designed to service peak-hour person demand at the lowest total cost to the public, while at the same time, providing system continuity.[8] The public costs have been identified as travel delay, construction and right-of-way, and operation costs of the facilities. The methodology also recognizes that some

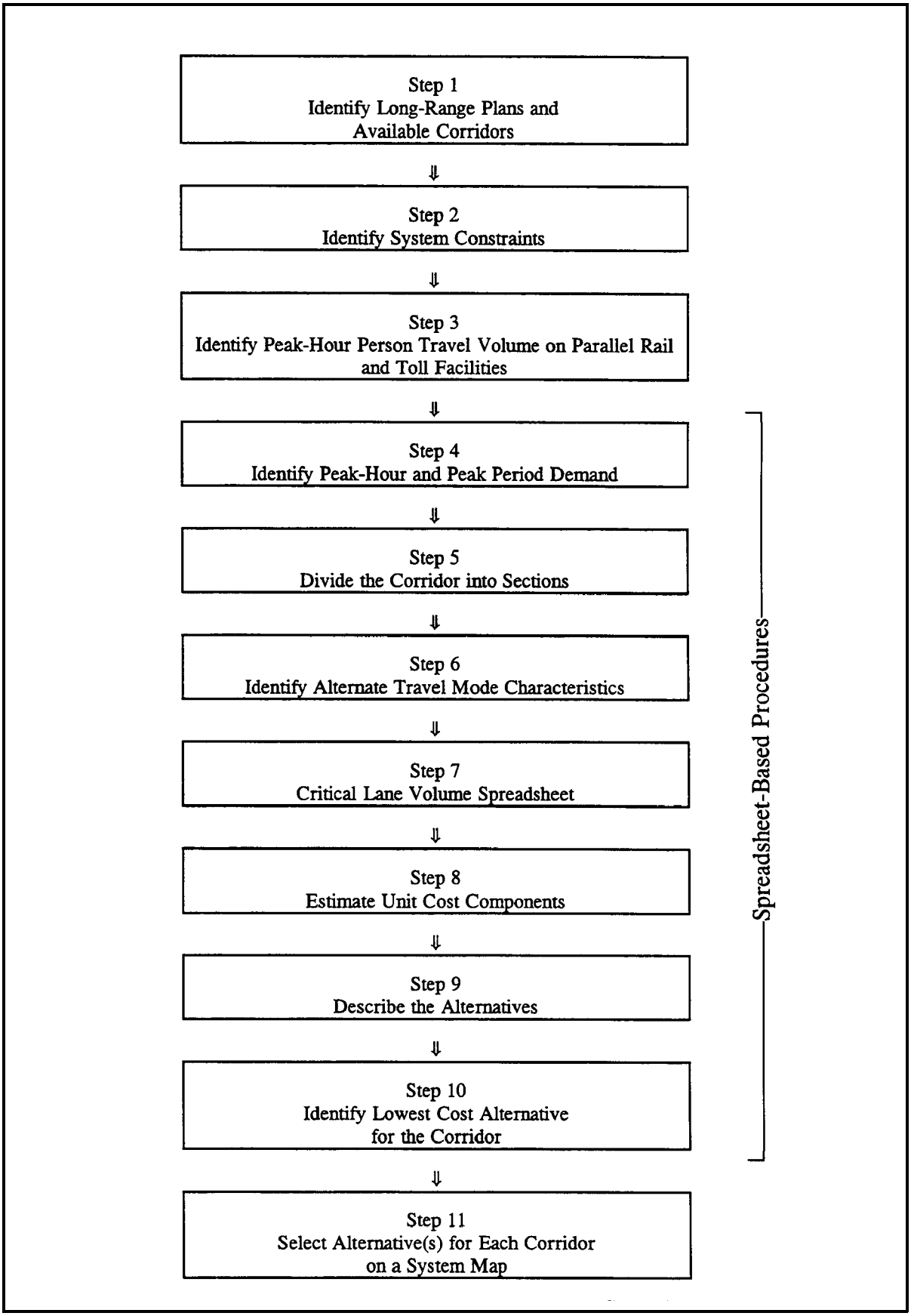


Figure 6-2. Multimodal System Planning Technique—Steps in the Process.[8]

motorists will change their mode of travel when given the opportunity to avoid congestion, resulting in more transit and carpool use. The methodology can be summarized in the following steps, which begin with corridor and system concerns, become more specific, and conclude with a system-level assessment of the results:

- Identify constraints in the corridor and the system.
- Estimate demand.
- Test alternatives for corridors.
- Examine results of individual corridors for system consistency and adjust improvements.

The methodology includes consideration of facilities such as parallel urban rail transit and toll highways, high-occupancy vehicle lanes, express freeway lanes, and general freeway lanes. The process can also incorporate operational and travel demand management improvements.

The multimodal system planning process includes a spreadsheet-based analytical procedure and several considerations before and after the spreadsheet operation. The technique generally follows an inside-outside geographic arrangement. System constraints (usually most frequent near downtowns) are initially identified. Individual corridors are analyzed and the alternatives optimized, possibly for several factors. The corridors are brought together on a system map, and the match points harmonized.

The system planning methodology provides information to quantify decisions regarding the most efficient expenditure of transportation funds for a multimodal system. It includes operational experience in the framework and provides a balance in difficult concepts such as congestion level and mode shift to ridesharing alternatives. It can optimize transportation systems based on the lowest cost to the public and also optimize within agency construction and operation budgets. The spreadsheet-based analytical program is open to user assumptions, and all default values are supported by documentation.

In practice, the outcomes of the corridor analyses vary depending on travel demand. Low demand corridors (less than 150,000 daily trips) where little or no construction is warranted are usually optimized with freeway general purpose lanes. Moderate demand corridors (up to 200,000 daily trips) may require some limited access express lanes, but the lowest public cost is usually achieved with only general purpose improvements. High demand corridors (in excess of 200,000 daily trips) are usually most efficient with a combination of HOV lane and freeway improvements.

IMPLEMENTATION

A natural result of the system engineering approach is the implementation of the plan. Issues that could be of concern with the implementation of an HOV treatment include the following:[5]

- **Scheduling.** If the HOV project is a retrofit, implementation scheduling can be complicated by a variety of unknowns, generally related to the policies and procedures of the various agencies involved. If the project requires daily monitoring (e.g., contraflow or reversible-flow operations), an additional 1 to 3 weeks following construction completion should be included for pre-operation testing.

- **Design/Construction Packaging.** Proper packaging of the project design can affect quality and cost. There are various reasons to segregate a project's elements into multiple design and construction packages. If the package is too large, competitiveness is reduced in the construction bidding process. If the improvement represents substantially different construction trades (i.e., roadway work versus electronic surveillance), it may also make sense to segregate the improvements to achieve the best responsiveness and quality of work. Conversely, it makes sense in projects involving HOV and adjacent freeway improvements, to combine common elements of work into the same construction packages. This simplifies construction management on the job site.

- **HOV Operation During Construction**. One of the most effective methods of cultivating an early market for an HOV project is to start offering preferential treatment during the construction phase. Additionally, this approach can be a cornerstone of the traffic management plan aimed at preserving corridor flow during construction activities. These benefits often outweigh the complications this approach creates for contractors and throughout the construction period.

- **Pre-Operation Testing**. Some pre-operation testing is desirable for any HOV facility. This period allows police to refine enforcement strategies and the operations team to make minor adjustments in the facility design prior to operation. For reversible-flow or contraflow projects, pre-operation testing is essential. At least 1 to 3 weeks is needed to check out any of the automated features that will be changed on a daily basis and acquaint bus operators, deployment staff, police, and others with how to handle daily operation, maintenance, and emergencies.

- **Operation.** Opening of a project should be preceded by significant public awareness efforts. Target users should be provided information on how to take advantage of the project, including maps, rules and regulations for use, and instructions on how to react in an emergency.

Prior to opening the facility to traffic, a plan must be devised for enforcing the restrictions on the facility. The plan should include how often, how long, and where enforcement activities should be performed.

EVALUATION

The final step in the design process is to evaluate the effectiveness of the HOV treatment. This should not be considered a one-time activity, but should be part of a periodic review of the effectiveness of the component and of the overall system. In addition to providing information to the sponsoring agencies on the effectiveness of the treatment(s), the information would be helpful in communicating the effectiveness of the project to the public and enhancing a general understanding of the role that the HOV project has performed.

Table 6-5 provides an overview of suggested objectives, data collection efforts, and corresponding measures of effectiveness for evaluating HOV facilities. The source of the information in this table contains an approach for conducting evaluations of freeway HOV facilities, including identification of appropriate evaluation objectives, corresponding measures of

Table 6-5. Suggested Objectives, Data Collection Efforts, and Measures of Effectiveness for Evaluating HOV Facilities. [6]

Objective	Veh. & Occup. Counts		Travel Time Runs		Surveys (1)			Corresponding Measure of Effectiveness (MOEs) (3)
	Freeway (2)	HOV Lane	Freeway (2)	HOV Lane	Freeway	HOV Lane	Other	
Increase occupancy	✶	✶			✶✶	✶✶	✶✶ (4)	Percent increase in peak-hour, peak-direction person volume; increase in average vehicle occupancy; and modal shift
Cost effective	✶✶	✶	✶	✶				Benefit-to-cost ratio
Travel time savings			✶	✶	✶✶	✶✶	✶✶(5)	Amount of travel time saving by HOV users; reliability of trip time for HOV users
Public Support		✶✶			✶	✶	✶✶(6)	Percent of users, non-users, and general public who approve of HOV facility; violation rates
Energy and air	✶	✶	✶	✶	✶✶	✶✶	✶✶(7)	Reduction in CO, HC, and NO emissions; reduction in energy consumption and noise level
Freeway operations	✶		✶		✶✶			Increase in peak-hour per lane efficiency of freeway
Safety	✶✶	✶✶					✶(8)	Number and severity of accidents; accident rate per million vehicle miles of travel and per million passenger miles of travel

✶ Indicates the top priority data collection efforts needed to evaluate the objectives.
✶✶ Indicates data collection efforts which ideally should be conducted, but are not absolutely necessary to evaluate the objectives.
1 Involves periodic use of surveys of HOV users (bus riders, carpoolers, and vanpoolers), non-HOV users in the general traffic lanes, and in some cases, the general public.
2 It is strongly suggested that this data be collected for both the freeway lanes adjacent to the HOV facility and the control freeway.
3 Some, but not necessarily all, of the suggested MOEs associated with gauging the attainment of the objectives are shown.
4 Vehicle and occupancy counts on alternate arterial routes to identify any changes in throughput for the corridor, counts at park-and-ride lots, and vehicle and occupancy counts on a "control" freeway.
5 Monitoring bus on-time performance and schedule adherence before-and-after implementation of the HOV lane(s).
6 Identifying violation rates for the HOV lane (i.e., those vehicles not meeting the minimum occupancy requirement). Monitoring complaints, media, and policy actions.
7 Monitoring air quality and noise levels along the corridor.
8 Identifying freeway accident rates and types before-and-after implementation of the HOV lane(s), as well as obtaining accident rates on the HOV facility.

effectiveness, data collection methodologies, and an evaluation process.[6]

6.3 TECHNIQUES AND TECHNOLOGIES

HOV FACILITIES

There are essentially four different types of high-occupancy vehicle (HOV) facilities used on freeways.

- ***Exclusive HOV Facility - Separate Right-of-Way.*** A roadway or lane(s) developed in a separate and distinct right-of-way and designated for the exclusive use of HOVs (see figure 6-3).

- ***Exclusive HOV Facility - Freeway Right-of-Way.*** Roadways or lanes built within the freeway right-of-way which are physically separated from the other freeway lanes but reserved for exclusive use by HOVs, at least during portions of the day (see figure 6-4).

- ***Concurrent Flow Lane***. A freeway lane in the peak direction of flow (normally the inside lane) that is not physically separated from the other freeway lanes but is designated for use by HOVs at least for a portion of the day (see figure 6-5).

- ***Contraflow Lane***. A freeway lane in the off-peak direction of flow (normally adjacent to the median) that is designated for use by HOVs traveling in the direction of peak flow for at least a portion of the day. Normally, the contraflow lane is "separated" from the off-peak (or opposite) flow by insertable cones, pylons (see figure 6-6), or movable concrete barriers.

Operational Considerations

The importance of incorporating operational considerations into both the planning and design process for HOV facilities cannot be overstated. The operation of an HOV facility is critical and should be considered when making planning, design, and implementation decisions. Also, consideration must be given to a range of needs involving support services and facilities, such as park-and-ride lots, bus service planning, marketing rideshare matching, and enforcement. Following is a summary of several operational considerations as discussed in the AASHTO *Guide*.[10]

Figure 6-3. Examples of Exclusive HOV Facility - Separate Right-of-Way. [9]

Figure 6-4. Examples of Exclusive HOV Facility - Freeway Right-of-Way. [9]

Figure 6-5. Examples of Concurrent Flow Lane. [9]

Figure 6-6. Examples of Contraflow Lanes. [9]

HOV Roadway Operations

The operation of separated HOV roadways may be reversible or two-way. The facility can be restricted to HOVs during peak periods only or throughout the day. The latter is less expensive to sign and operate. Limiting access to a reversible HOV facility is crucial if the facility is to be operated in a safe and efficient manner. A system of gates should be considered at each end to prevent wrong-way traffic from entering the facility. In addition to these features, this type of facility should also have a system of variable message signs (VMS) which inform commuters as to the operational status of the facility (open or closed).

Operating Efficiency

The reduction of user travel time is the commonly used measure for assessing the benefits of HOV facility operation. A vehicle breakdown on an HOV facility can be anticipated to occur approximately every 64,000 vehicle-kilometers (40,000 vehicle-miles) traveled. A disabled vehicle will cause a decrease in traveling speeds or, in extreme cases, a total blockage of the facility. Figure 6-7 illustrates the consequences of a 15-minute total blockage on an HOV facility on a separated roadway having a demand of 6,000 persons per hour and a capacity of 12,000 persons per hour. As shown, the 15-minute blockage causes 22,500 person-minutes of delay, and it will take 30 minutes from the time of blockage until the queue totally dissipates.

Considerations for 2+ Versus 3+ Occupancy Requirement

HOV facilities should be implemented in a way that balances the flexibility of HOV growth and the public perception as to the use of a facility. An initial minimum vehicle occupancy requirement must be selected to optimize the efficiency of the facility. The selection must allow for growth in traffic volumes as more commuters choose to switch to carpooling arrangements and take advantage of the travel time and fuel savings. Title 23 United States Code 102(A) allows State departments of transportation to establish the minimum occupancy requirements for vehicles operating on HOV lanes; except that no fewer than two occupants per vehicle may be required and that motorcycles and bicycles shall not be considered single occupant vehicles.

Retaining the potential to carry more people over time offers important operational flexibility. At the same time, though, public perception of the adequacy of HOV lane usage must also be addressed. Peak hour

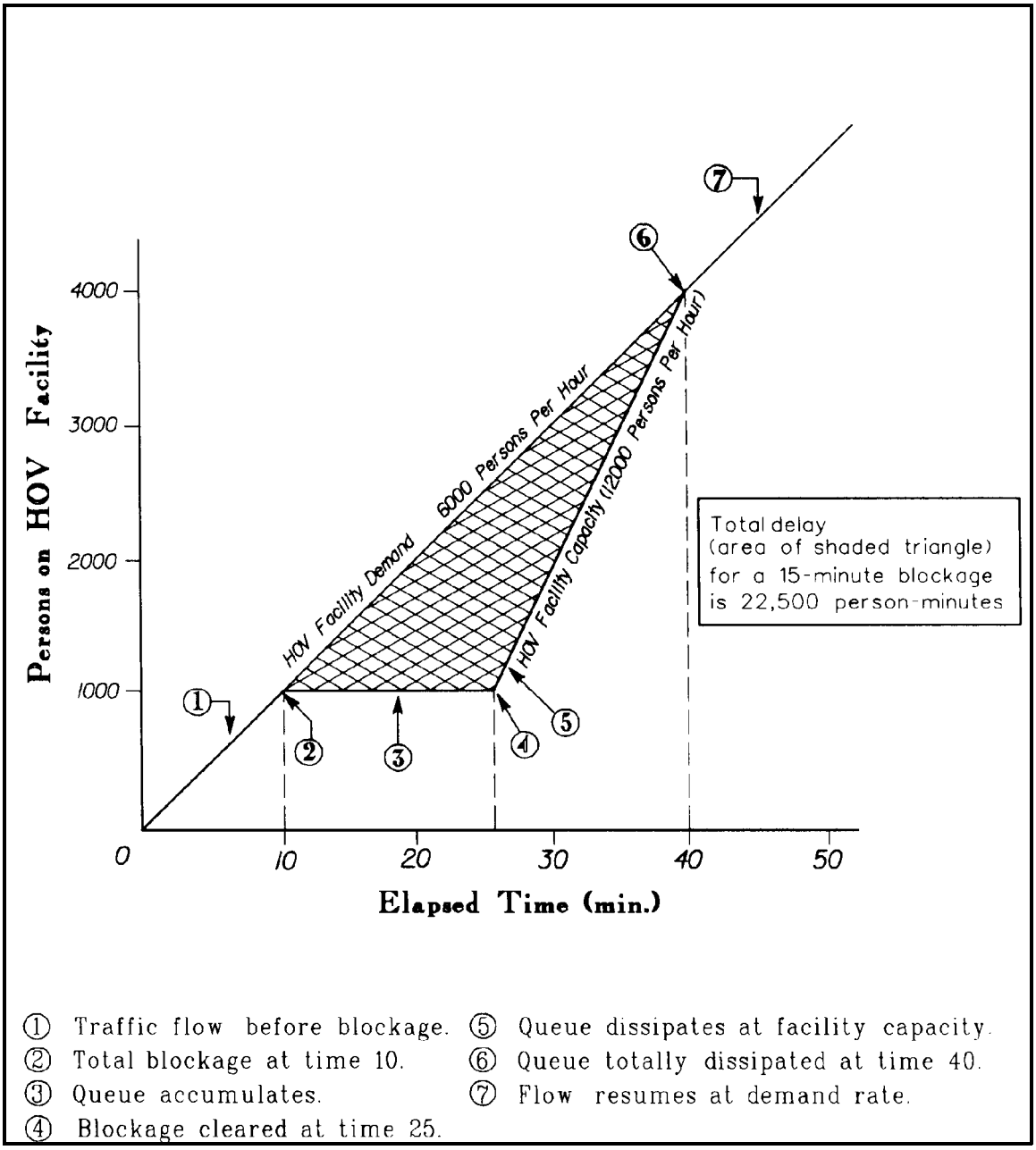


Figure 6-7. Example of Delay on an HOV Facility Due to a Total Blockage Incident. [10]

HOV traffic volumes need to be high enough to help mitigate public concerns over underutilization of HOV facilities. The positive aspect of 2+ eligibility is that a staged resource of commitment to ridesharing is being established. Less work is involved in forming a 2+ carpool versus a 3+ carpool, and the base volume to draw from is considerably greater. There may be less eventual resistance to adding a third passenger than to forming an initial 3+ carpool.

Subsequent changes in occupancy requirements need to be weighed with projected future demand. To go from 2+ to 3+ occupancy could reduce vehicular demand by as much as 75 to 85 percent.

This could be severe if only a 10 to 20 percent reduction in demand is necessary for the near future. A new HOV 3+ lane typically may carry only a few hundred peak-hour vehicles, while an adjacent freeway lane is carrying 1500 to 2000 peak-hour vehicles. Even though the HOV lane may be carrying more peak-hour person trips than an adjacent freeway lane, the traveling public may perceive the lane to be underutilized. Consideration should be given to changing from 2+ to 3+ occupancy when the level of service of the HOV lane is approaching Level of Service D.

In other words, the flexibility to change occupancy requirements is a strong operational management tool. However, changing these requirements frequently or varying the requirement by time of day may create enforcement problems and public resistance.

Hours of Operations

An HOV facility may be operated only during peak periods only or for 24 hours. A number of factors, including geometric design, volumes of HOV and mixed-flow traffic, and hours of congestion will influence HOV operating hours.

Twenty-four hour HOV use of priority facilities is sometimes preferred, because violations tend to be lower and there is less motorist confusion. Also, 24-hour use may provide a greater overall incentive for the formation of new carpools. Some HOV facilities, such as reversible lanes, may not be conducive to 24-hour operation. The hours of operation for reversible facilities must allow time for a variety of necessary functions, such as clearing the lane, moving gates, and changing signing.

Part-time use of a shoulder as an HOV facility should be implemented only after careful consideration of operational and safety problems. The shoulder HOV facility differs from a part-time HOV lane that reverts to mixed-flow use during off-peak periods. The shoulder facility requires special delineation and signing, and involves separate enforcement problems for both peak and off-peak periods. Motorists may tend to use the shoulder as a freeway lane during off-peak hours when it should be used as a shoulder.

Design Elements

Design Speed

A purpose of HOV facilities is to provide a travel time savings for HOVs. Therefore, it is preferable to use a design speed for the HOV facility that is comparable to the adjoining freeway. AASHTO freeway standards for design speed should be used to provide for a high level of service. The design of the facility should consider the specific physical dimensions and operating characteristics of the vehicle types that are expected to be well represented in the vehicle mix. For example, the difference in braking and acceleration characteristics may suggest a different roadway geometry than just using passenger cars.

Cross Section Width

Following is a summary of dimensions used in the design of transit ways.[11] The preferred cross section for single lane at-grade, one-way transitways is a 3.6-m (12-ft) travel mainlane and 1.5-m (5-ft) clearance offsets (shoulders), while the usual cross section is 3.6-m (12-ft) travel mainlane and 1.143-m (3.75-ft) clearance offsets (shoulders). A total minimum width of 6.1 m (20 ft) is recommended for single-lane facilities, as this will allow a disabled vehicle to be passed. Reduced cross sections do not

provide the passing capability, and therefore are recommended only for short sections, approximately 610 m (2000 ft) or less, that involve physical constrictions.[12] For multiple lane at-grade, two-way facilities, the travel mainlanes are 3.6 m (12 ft), the center shoulder separations are 3.0 m (10 ft), and the clearance offsets are 0.6 m (2 ft).

HOV lanes should have a minimum pavement width of 3.6 m (12 ft). The pavement should be widened through horizontal curves to account for the offtracking of buses. A minimum width of 4.0 m (13 ft) is recommended for HOV ramps. The typical cross slope for a transitway is two percent, the same cross slope found on most freeway mainlanes.

To allow water to drain, a minimum grade of 0.35 percent should be provided. A maximum grade of 6 percent is recommended to prevent buses from slowing down on the HOV lane. The desirable superelevation is 0.04 to 0.06 for speeds of 64 to 80 km/h (40 to 50 mph) and 0.06 to 0.08 for speeds of 80 to 97 km/h (50 to 60 mph). Vertical clearances are 4.42 m (14.5 ft) minimum and 5.03 m (16.5 ft) desirable, while lateral clearances are 0.6 m (2 ft) minimum and 2.4 m (8 ft) desirable.

Vertical Clearance

Vertical clearance to structures passing over the HOV facility should desirably be the same as for the adjacent freeway at 5.03 m (16.5 ft). While this is more than sufficient allowance for the maximum height of a standard transit bus at 3.429 m (11.25 ft) (double-deck bus is 4.32 m [14.2 ft]), it does allow for the possibility of emergency use or for future use of other types of vehicles, including large commercial trucks. In situations of restricted vertical clearance, a reduced (usual) clearance of 4.42 m (14.5 ft) is generally acceptable. This includes some allowance in vehicle operation and future pavement resurfacing.[4]

Signs and Markings

Signs and markings should conform to the *Manual on Uniform Traffic Control Devices (MUTCD)* to the fullest extent possible.[13] Preferential lane markings should be used to indicate that the lanes are restricted, with supplemental signs or signals conveying the specific restrictions. At the entrance to reversible facilities, particular attention must be paid to the control devices. In addition to static signs, variable message signs may be necessary. These should be supplemented with gates or barriers to further prevent entry by vehicles going in the wrong direction or to allow only authorized vehicles by special designation to enter the facility.

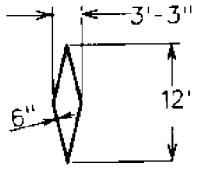
Signs. Regulatory signing for HOV lanes should follow the standard regulatory signing principles (e.g., black legend on white background, rectangular shape, and reflectorized or illuminated). The diamond symbol (white on black background) should be incorporated into the sign format. Guide signing may be necessary on HOV lanes that have designated ingress, ingress/egress, or egress points to inform the motorist of the appropriate HOV lane entry or exit point to use. The guide signing should follow the standard guide signing of the *MUTCD*. The diamond symbol should also be incorporated into the sign format, preferably in the upper left corner of the sign.

Figure 6-8 gives sample signing and pavement markings for HOV lanes used by Caltrans, along with some general applications criteria.

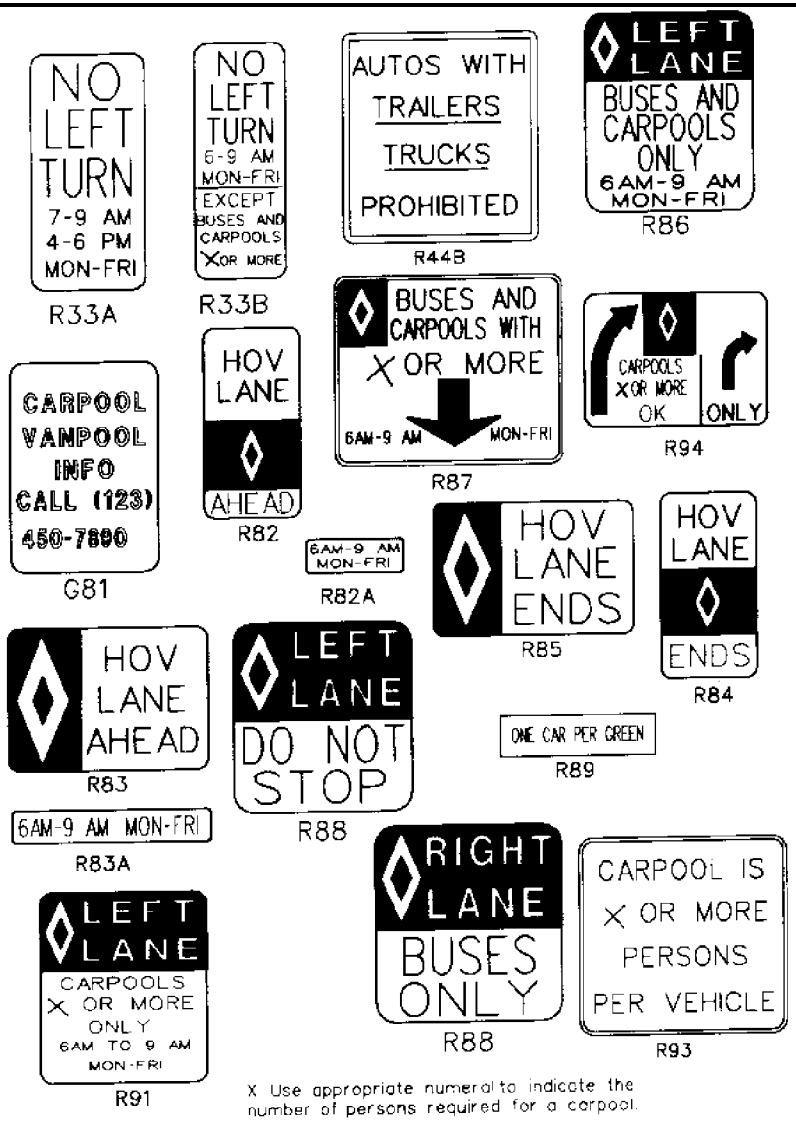
Markings. When a lane is assigned full- or part-time HOV use, HOV lane markings are

HOV PAVEMENT MARKING AND SIGNING NOTES

1. Two advance carpool signs, R94, should be installed on local street when striped for mandatory right turn.
2. Two or more carpool signs, R33B, should be installed on local streets (with concurrence of local agency) wherever left turns are restricted to carpools during peak hours.
3. Two carpool signs, R91, should be installed on extra long entrance ramps with two Detail 'A' pavement markings.
4. Four or more carpool signs, R91, should be installed on extra long ramps with two or more Detail 'A' pavement markings.
5. Spacing of Detail 'A' may be as close as 80 feet on city streets, while a spacing of 1/4 mile may be appropriate for a freeway.
6. Signing and pavement marking details and notes shown on this sheet should be used wherever feasible. Exceptions and proposed experimentation with other signs and markings should be reviewed in advance by the Operations Unit.
7. The bus-carpool lane on the freeway entrance ramp may be either on the left or right side.
8. Additional required signs and markings are shown in the Caltrans Traffic Manual, Section 4-05, Ramp Terminal Signing.
9. A number may be placed in the center of the Detail 'A' pavement symbol to indicate the number of persons required for a carpool.
10. Posted time is not required for ramp bypass lane metered by traffic signals.
11. The pavement legend "CARPOOL LANE" may be used to supplement the Detail 'A' installations on new projects. The word message should then be allowed to wear out.
12. For 24 hour HOV operations, time restrictions are deleted.



DETAIL 'A'
(See Note 9)



X Use appropriate numeral to indicate the number of persons required for a carpool.

Figure 6-8. Sample Signing and Pavement Markings for HOV Facilities.[10]

necessary. The HOV lane marking should be an elongated diamond placed along the longitudinal center of each restricted lane (see figure 6-8). The marking is intended to convey that a restriction on the use of the lane exists, and it is supplemental to the signs or signals conveying the specific restriction. Signs or signals should be used with the HOV lane markings.[10]

The frequency with which the marking is placed is a matter for engineering judgement based on prevailing speed, distance between interchanges and other considerations necessary to adequately communicate with the driver. A spacing of 0.40 km (0.25 mi) would be appropriate for most freeway situations, except on crest vertical curves where a shorter spacing should be used. Initially, the markings "CAR" "POOL" "LANE" may be painted between the diamond symbols on new projects to supplement, but not substitute for, the HOV lane markings. The word markings should then be allowed to wear out once drivers become familiar with the facility.[10]

The striping pattern for the lane line between the HOV lane and the adjacent mixed-flow lane should be in accordance with the *MUTCD*.[13] Typical HOV lane striping and pavement marking schemes are shown in figure 6-9. This figure also provides examples of signing appropriate for concurrent HOV lanes.

Enforcement

Goals and Objectives

An objective of enforcement by police officers on HOV facilities is to maintain the operational integrity and safety of the facility for those high occupancy vehicles designated or authorized to use it. In this regard, detection and apprehension of violators, and effective prosecution of violators, are essential. Therefore, law enforcement personnel with full capability to issue citations must be employed on HOV facilities. In addition, police officers help ensure the safe and efficient operation of the facility. Depending on the type of facility and priority users, the potential safety and operational problems caused by vehicle breakdowns, wrong way movements and/or other vehicles' encroachments into the HOV facility may have an adverse impact on operations and must be a concern of the enforcement authority.

Table 6-6 summarizes selected goals and objectives of enforcement personnel, as well as strategies for implementation and measures of effectiveness.

A 1988 Texas Transportation Institute study of the enforcement procedures for HOV lanes determined the following key concepts related to effective HOV enforcement:[14]

- The level of enforcement needed is dependent upon facility type. In general, concurrent flow facilities require more enforcement than do separated roadway and contraflow facilities.

- To be effective, an officer must have a safe and convenient place to issue citations or warnings. The enforcement activity should be in view of HOV users so that they can see when the lane restrictions are being enforced; however, it should not interfere with traffic on the HOV and mixed-flow lanes.

- To preclude high violation rates, a highly visible enforcement presence has to be maintained at a level where potential violators and legitimate users believe that violators have little chance to use the lane without getting caught.

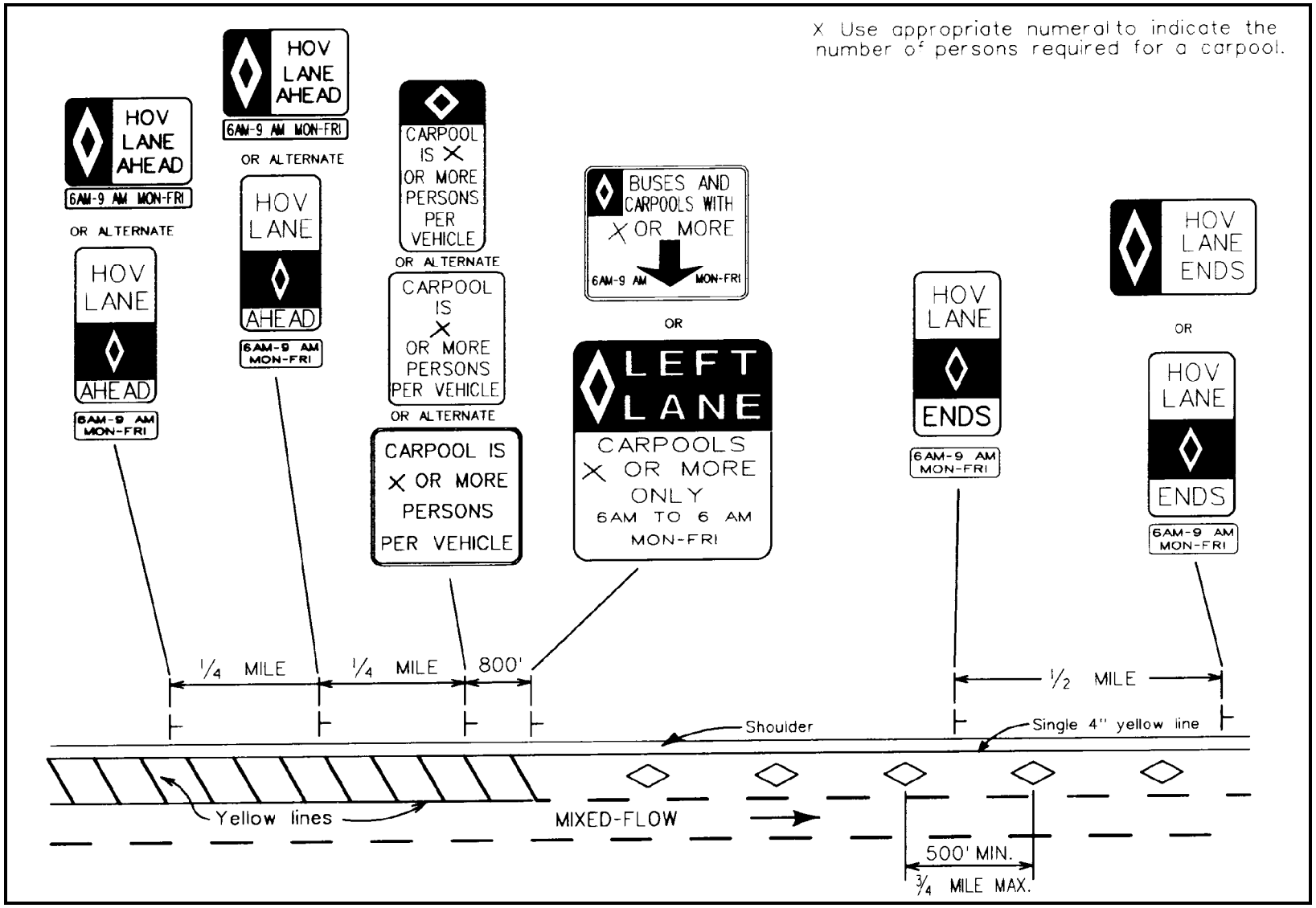


Figure 6-9. Typical Contiguous Concurrent HOV Lane w/o Buffer Signing and Pavement Markings.[10]

Table 6-6. Goals and Objectives of Enforcement Personnel.

Goal	Objectives	Strategies	Measures of Effectiveness
Maintain operational integrity	Help minimize delay Minimize violation rates	Strict enforcement of occupancy requirements Clear communication of nature of facility High visibility of enforcement officers Swift, safe removal of violators	Violations Violation rates Travel times
Maintain safe operation	Minimize accidents Help minimize incident response and clearance times	Strict enforcement of operating rules Clear communication of nature of facility Swift, safe removal of violators	Accidents Accident rates Incident response and clearance times

- On limited access facilities, diverting potential violators before they can traverse some part of an HOV lane can be safer and more efficient than apprehending them after the fact. Whenever possible, enforcement areas should incorporate this concept.

Methods

Where enforcement is difficult to accomplish, or perceived as being unsafe, police may avoid apprehending violators, resulting in increasing numbers of illegal vehicles using the lane. Where enforcement has been a problem, 60 percent or more of the vehicles that used the lanes were violators. Experience suggests that steady doses of routine enforcement, combined with moderate application of special enforcement, can generally keep violation rates on exclusive HOV facilities in the 5 to 10 percent range. Heavy, consistent doses of special enforcement would be necessary to have violation rates below 5 percent. There are locations where no amount of enforcement can bring violation rates to an acceptable level.[10]

In some metropolitan areas, programs have been initiated where motorists can call in to report HOV facility violators. Appropriate literature is sent to frequent violators, and enforcement personnel can make a point of watching for these vehicles in the HOV lane. These "so called" HERO programs have been helpful in reducing violation rates. Also, a system of video cameras combined with officer observation may be considered. Another factor that will have a positive

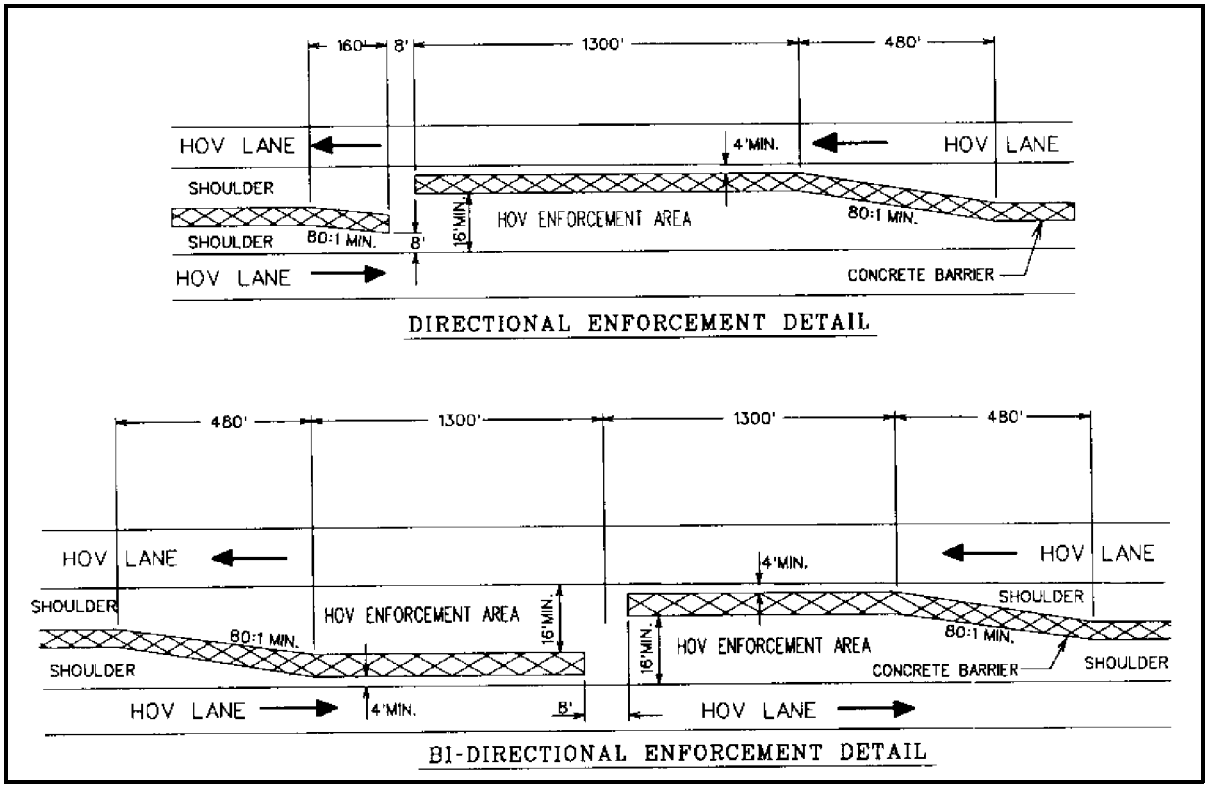


Figure 6-10. Median Enforcement Area for Median 6.7 to 8.8 m (22 to 29 ft) Wide.[10]

impact on the violation rate is the cost of the fine for a violation. Fines exceeding $250 for first offenders have been used, significantly lowering the violation rate.

Enforcement Area Design

When totally new concurrent HOV facilities are to be built and there is adequate width for a median 9 m (30 ft) or wider, consideration should be given to providing a 4.3-m (14-ft), paved median shoulder in both directions as a continuous enforcement area. When facilities are operating at high speeds (81 km/h (50 mi/h) or above), enforcement officers are uncomfortable having to stop and approach another stopped vehicle in less width than this.[10]

On other facilities where the median width is less than 9.2 km/h (30 ft) but equal to or greater than 6.7 m (22 ft) wide, it is possible to accommodate enforcement pockets by narrowing the median shoulder on alternating sides of the center barrier, which permits on-line enforcement. Figure 6-10 shows possible designs for enforcement areas. Closely spaced raised pavement markers could be used to provide added delineation between the enforcement area and the HOV lane in order to improve safety of the enforcement area. California, for example, uses raised pavement markers spaced out at 1.8 m (6 ft) along the edgeline to provide an audible warning to those entering the enforcement area.[10]

Where the median width is less than 6.7 m (22 ft) and enforcement pockets are required, it will be necessary to shift, in one or both directions of travel, away from the centerline in order to obtain the minimum

width to accommodate a minimum 4.3-m (14-ft) wide pocket, 0.6-m (2-ft) barrier, and 0.9-m (3-ft) offset to the barrier from the opposing HOV lane. Tapers for lane shifts should be 115:1 or greater. Conceptual design is shown on figure 6-11.

Pockets should be no more than 3.2 to 4.8 km (2 to 3 mi) apart, and should be located between interchanges to avoid interference from interchange structures and also to minimize both the distraction to other drivers and the lane changes taking place at any one time. Pockets should be located to provide minimum sight distance requirements.[10]

ACCESS

Ramp Connections

Types of ramp connections to HOV facilities include the following:

- ***Slip Ramps*** - The at-grade slip ramps are the easiest and least expensive to build. An opening large enough for normal merge/diverge maneuvers is placed in the barrier. This type of ramp is usually from a park-and-ride lot to the frontage road, the freeway, or the HOV lane (see figure 6-12). Elevated slip ramps are also used in some locations (see figure 6-13).

- ***Flyover Ramps*** - The second type of park-and-ride connection is the flyover ramp (see figure 6-14). This ramp resembles an elevated freeway ramp, except that flyover ramps can be either one- or two-way. The elevated flyover ramp is directly connected from the park-and-ride lot to the HOV lane. Because the flyover ramp does not create a tee intersection, fewer conflict points exist.

- ***Grade Separated Interchanges*** - Grade separated interchanges are more expensive than at-grade ramps, but they provide greater flexibility and movement. They are basically aerial tee intersections that can be either one or two directional (see figure 6-15). Acceleration and deceleration lanes should be included. Grade separated interchanges allow vehicles to travel directly from the park-and-ride facility to the HOV lane.

Design

The design of access connections depends on the decision of how to interface high occupancy vehicles with general purpose vehicles on and off the freeway.[10] Several options are available. One option is to connect the HOV facility directly to the freeway with the use of elevated flyover ramps or at-grade slip ramps. Another option is to link the HOV facility directly to the frontage road or surface street system. Of course, transition treatments must be provided in some manner at both ends of the roadway.

If feasible, the terminal connections to HOV facilities from the adjacent freeway mainline should be made with flyover ramps at both terminal end connections. This allows buses and other vehicles using the HOV facilities to exit and enter the freeway mainline on the right instead of having to enter the inner high speed lanes. This eliminates the high-speed lane merge, which is inherently more difficult to execute, especially for HOV traffic such as buses and vans. Depending on the interchange spacing, it could also eliminate the need for the HOVs to make several rapid lane changes in order to access the HOV lane or exit the freeway.

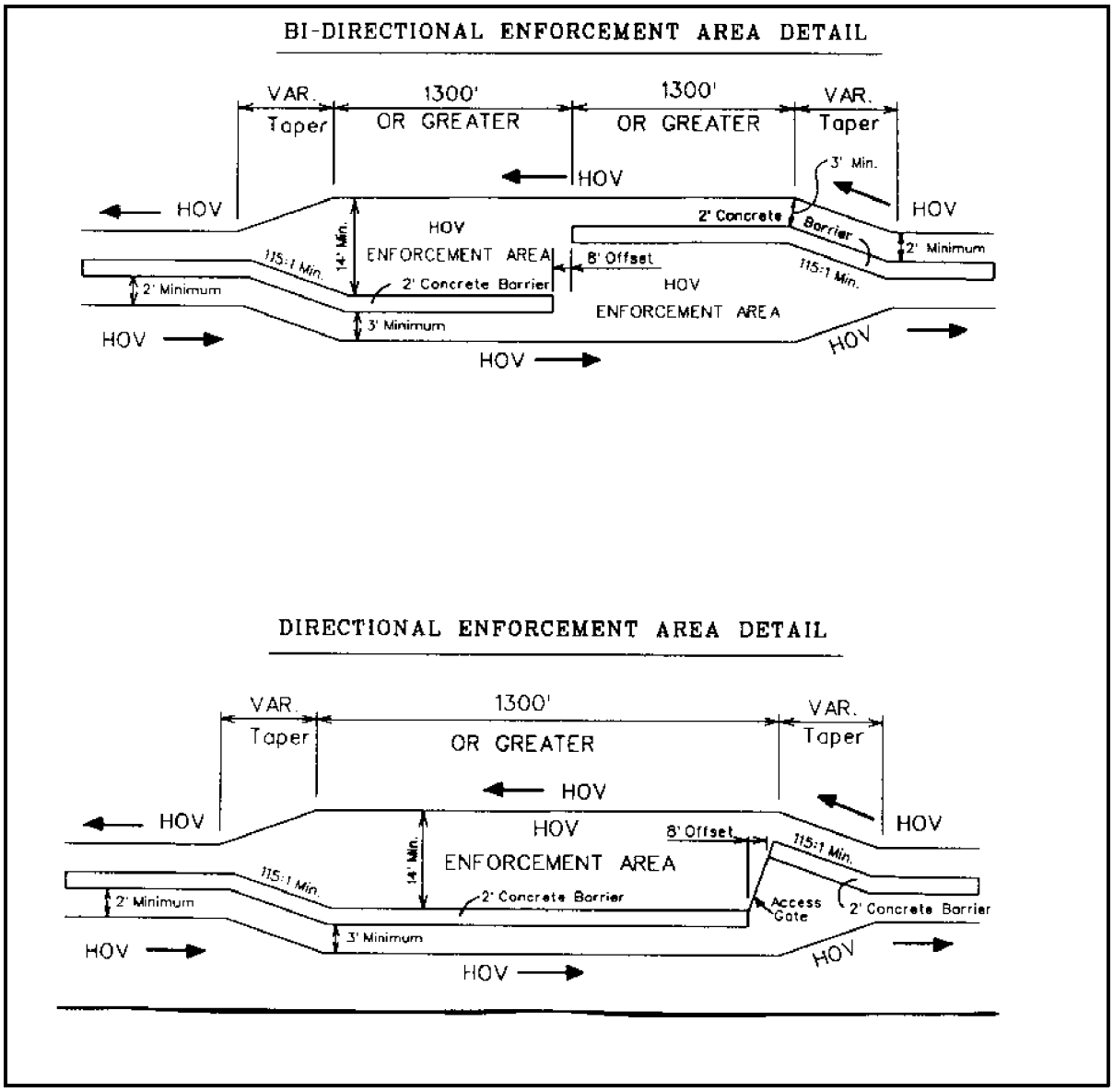


Figure 6-11. Median Enforcement Area for Median Less than 6.7 m (22 ft) Wide.[10]

Figure 6-12. Slip Ramp. [9]

Figure 6-13. Elevated Slip Ramp. [9]

Figure 6-14. Flyover Ramp.[9]

Figure 6-15. Grade Separated Interchange.[9]

Figure 6-16 shows typical flyover ramps connecting the HOV lanes to the righthand lanes of the freeway mainline. Where limited right-of-way and/or high costs prohibit the use of elevated flyover ramps, at-grade slip ramps can be used. At-grade slip ramps are also appropriate where the HOV facilities are reversible. This type of slip ramps terminal is particularly applicable to temporary or phased HOV lane implementation. Figure 6-17 illustrates the general concept of how an HOV facility with slip ramps can be incorporated into the freeway median. Since the HOV lanes shown are reversible, signing and/or barriers will be required to eliminate the wrong-way entry or exit. If traffic patterns warrant, separated HOV facilities should tie in to the existing street system within the central business district.

Direct ramps from a median HOV facility may be expensive and result in operational problems. However, they are preferable to merging HOV traffic with other freeway traffic in advance of the central business district, provided conditions permit. This concept is shown in figure 6-18 with connections into surface streets. It may also be desirable to connect the HOV facility to the through lanes of the freeway to provide for the through movement motorist. From outer areas, connections may be provided into freeway frontage roads for either collection or distribution of high-occupancy vehicles. Figure 6-19 illustrates the use of HOV flyover ramps to the freeway frontage roads. Where the HOV lanes are reversible, signing and/or barriers will be required to eliminate the wrong-way entry.

Intermediate connections to the HOV facility allow access on and off the facility to the through lanes of the freeway at critical locations, transit transfer centers, park-and-ride lots, and park-and-pool areas. These connections may be made at-grade with intermediate slip ramp openings or by grade separated interchanges.

All terminal and intermediate access connections should have high design standards. Tapers on entrance and exit ramps should be designed the same as for other freeway ramps, except that special consideration should be given to the acceleration and deceleration characteristics of loaded buses. This is especially critical where ramp grades are significant. Very long, gradual tapers should be avoided on exit ramps, as traffic may inadvertently follow the taper, assuming it is the main roadway. Ramps that connect to adjacent facilities or to cross streets should be designed to the same standards as comparable facilities that connect freeways to cross roads.

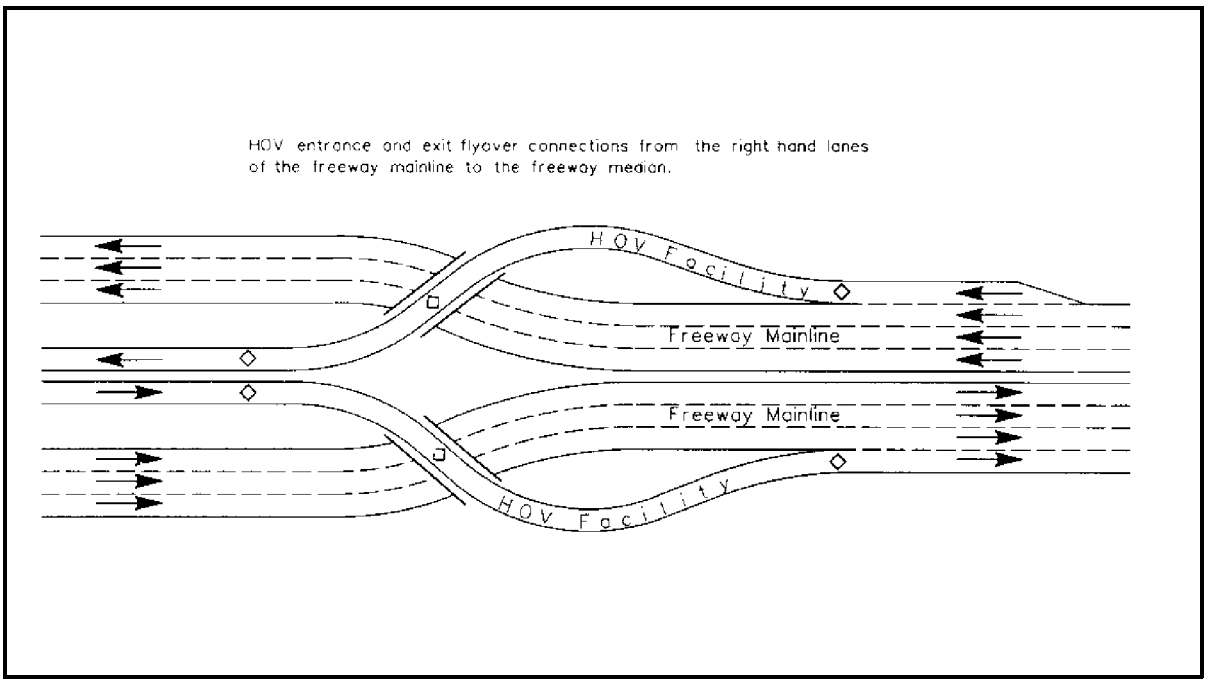


Figure 6-16. Terminal Flyover Connection to Freeway Mainline.[10]

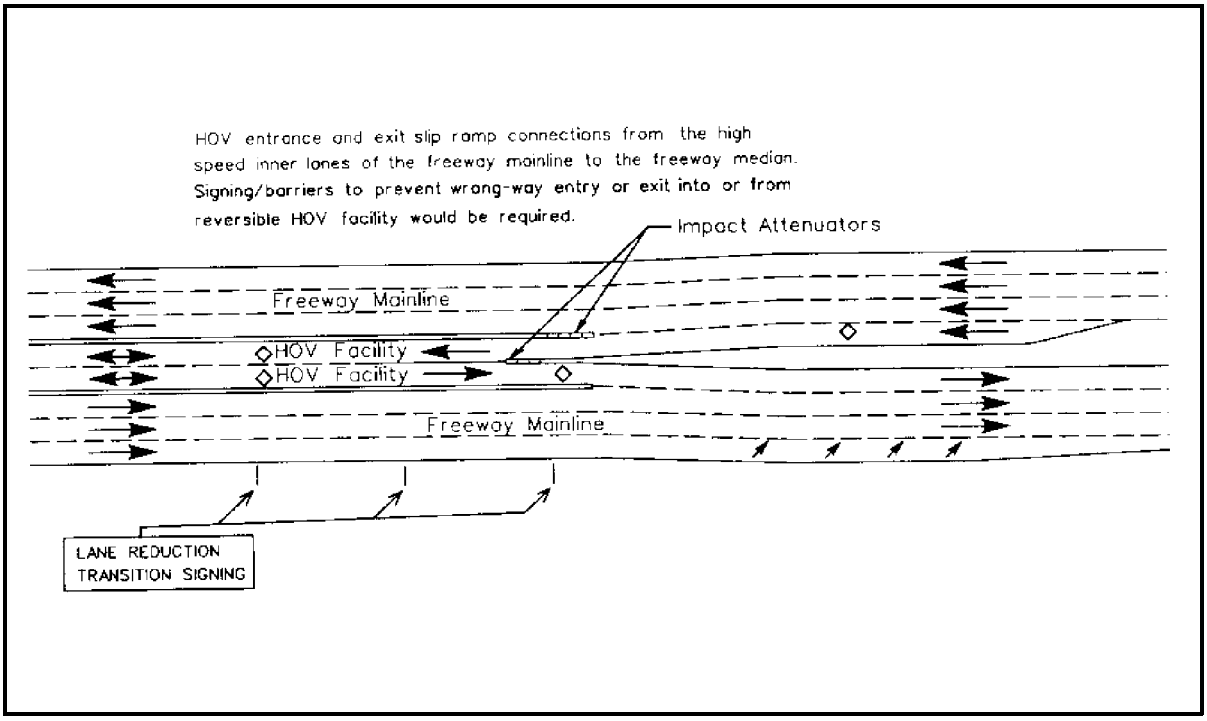


Figure 6-17. Terminal Slip Ramp Connection.[10]

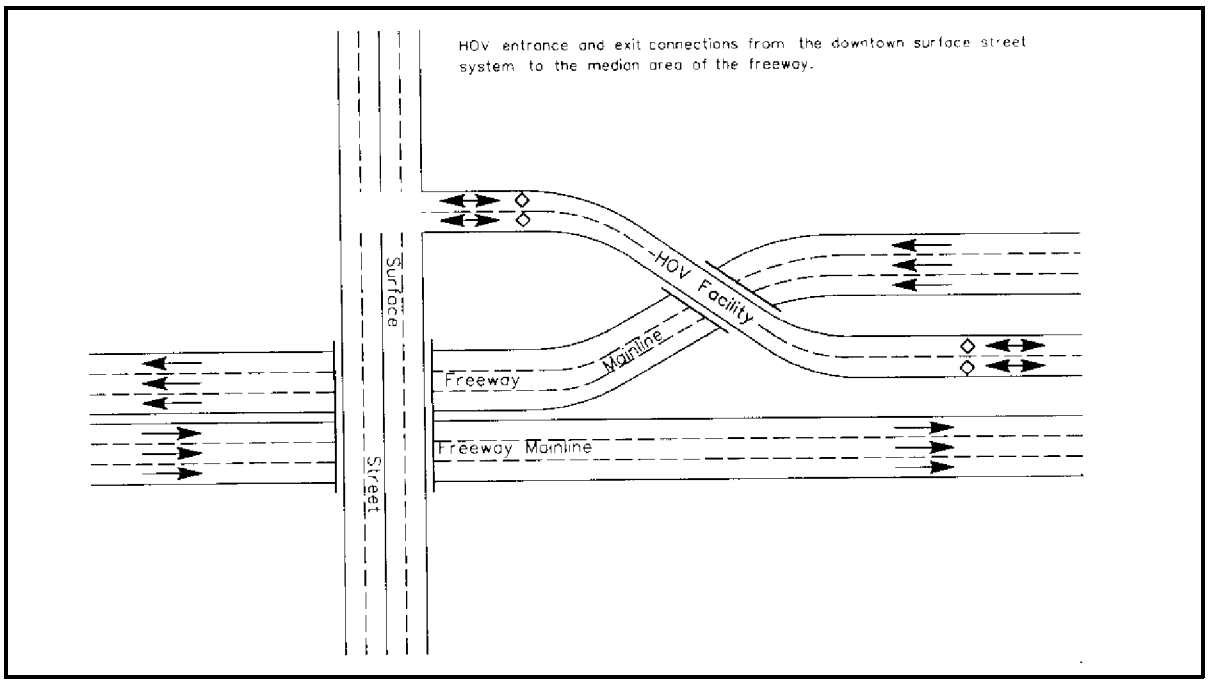


Figure 6-18. Terminal Flyover Connection to Surface Street System.[10]

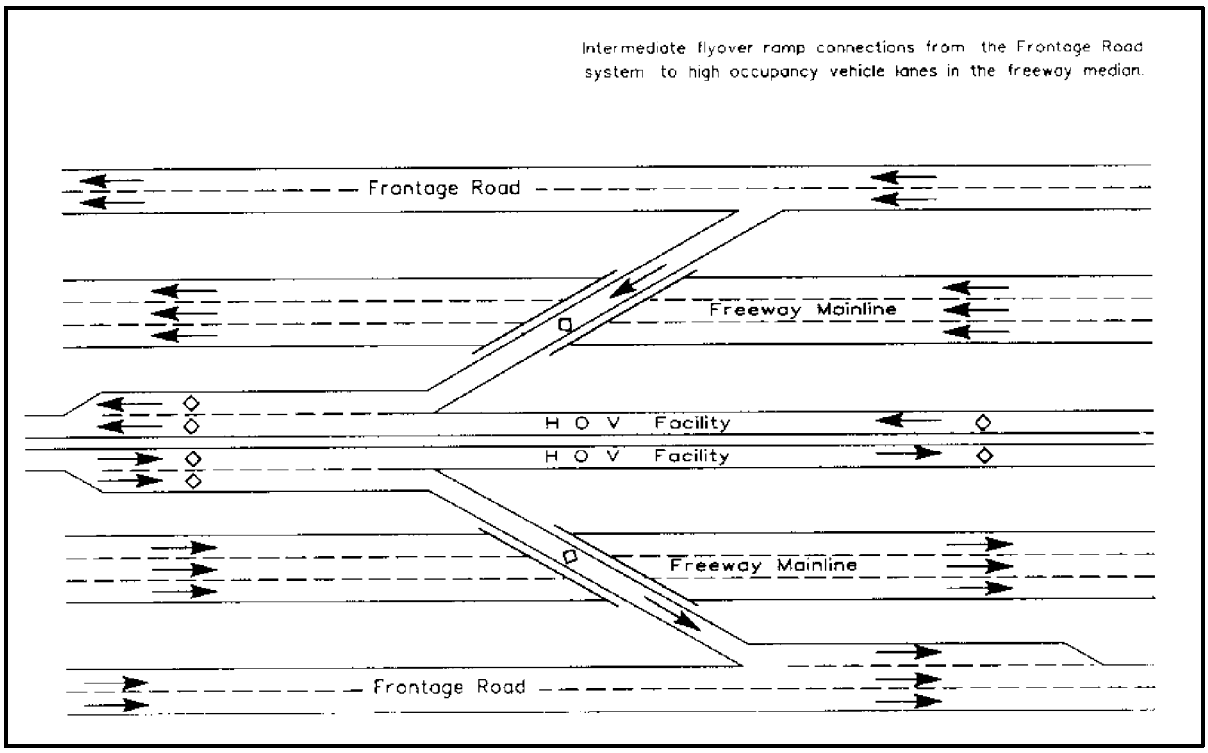


Figure 6-19. Intermediate Flyover to Frontage Road.[10]

Ramps are generally designed for speeds between 0 and 64 km/h (0 and 40 mi/h). Acceleration and deceleration lanes should be used for all types of HOV ramps. A minimum width of 4.0 m 13 ft) is recommended for HOV ramps. The minimum grade is the same as on the HOV lane, 0.35 percent; the maximum grade, however, is 8 percent.[11] Additional information is provided in the latest edition of the *Green Book*. [15]

PRIORITY ACCESS

Ramp Metering

Ramp meters are installed on freeway entrance ramps to do the following:

- Maintain uninterrupted, noncongested flow.
- Increase safety and efficiency by spreading out platoons of entering vehicles.
- Reduce congestion at the mainline ramp merge point.
- Limit amount of traffic entering the freeway so that demand for any section does not exceed capacity.

Ramp metering accomplishes these objectives by means of traffic signals on the entrance ramps that control the rate at which vehicles enter the freeway to maintain a balanced demand-capacity relationship. Usually, the rate of entry to the freeway is adjusted in response to the level of congestion or capacity available on the freeway. As freeway congestion increases, the rate at which vehicles are allowed to enter the freeway decreases. Also, excess freeway vehicular demand is encouraged to shift to alternative routes, to less congested time periods, or to HOV modes of transportation.

Ramp metering provides an opportunity to give priority treatment to, and encourage the use of, HOVs. This priority treatment can be in the form of a bypass of the meter or a preferential metering rate as compared with that of the general purpose ramp-metered lane. These treatments can be used in conjunction with mainline HOV lanes or where a freeway management system maintains an acceptable level of service and no HOV lanes are provided.

These bypass lanes can be restricted to buses only or can be made available to all HOVs. The decision on eligibility for use of the bypass should depend on the goals of the community, the number and types of vehicles that will use the bypass, and geometric conditions at the site affecting enforcement and operation of the ramp.

Bypass Lanes

The configuration of an HOV bypass lane can be that of a two-lane entrance ramp which tapers to a single-lane ramp prior to the merge with the freeway mainline, or a separate HOV bypass lane which merges with the entrance ramp downstream of the ramp meter. Figure 6-20 shows an example of the two-lane entrance ramp bypass.

The design of the ramp meter bypass should be determined by the conditions at each location. Bypasses should be 3.7 m (12 ft) wide with full ramp shoulders where possible, and should extend 91 m (300 ft) beyond the metering signal to permit HOVs to merge with normal ramp traffic. The ramp bypass traffic should merge first with regular ramp traffic, and then with freeway traffic.[10]

Figure 6-20. Two-Lane Entrance Ramp with HOV Bypass. [9]

Signs

For the two-lane entrance where one lane is designed as an HOV bypass lane, signs and pavement markings need to indicate the use of the right or left lane as the bypass lane. Such regulatory signs typically bear the legend "RIGHT/LEFT LANE BUSES AND CARPOOLS ONLY," and have the diamond symbol on the sign. The diamond symbol is also used as a pavement marking in the bypass lane. If a separate HOV bypass lane is constructed, the entrance to the bypass lane often has a sign installed reading "BUSES AND CARPOOLS ONLY," and the white-on-black diamond symbol.

Markings

Pavement markings have three major functions in a ramp metering HOV bypass installation:

- Indicate to drivers where to stop.
- Guide vehicles to form a single lane if the ramp begins as two lanes and tapers to one.
- Indicate the HOV bypass lane.

Ramp meter bypasses may be separated from the mixed-use lane by solid white pavement markings 20 cm (8 in) wide or wider. Diamond symbols should be used at 23 to 31 m (75 to 100 ft) intervals. Word pavement markings such as "CARPOOL ONLY", or "BUS ONLY" can also be used as appropriate and should be supplemented with signs.[10]

PARK-AND-RIDE FACILITIES

Park-and-ride facilities are an integral part of a multimodal transportation system. The purpose of these facilities is to provide a location for individuals to transfer from a low-occupancy mode of travel to a high-occupancy mode of travel. Park-and-ride facilities are typically provided in conjunction with transit services.

Most of the HOV facilities in operation around the country are connected either directly or indirectly to park-and-ride lots. For HOV lanes, transfers at park-and-ride lots are usually made from a vehicle to a bus; however, transfers may also occur from single-occupancy vehicles to carpools or vanpools. Examples of park-and-ride facilities used with HOV lanes are shown in figure 6-21.

There are many benefits associated with the proper use of park-and-ride facilities, including the following: [16]

- Encouragement of use of high-occupancy travel to maximize the efficiency of the transportation system.
- Improvement in efficiency of transit system by providing high-density areas for transfers and by increasing ridership.
- Assistance with congestion management, through a reduction in the number of single-occupancy vehicles on the freeway.

Figure 6-21. Park-and-Ride Lots with HOV Lane. [9]

- Reduction in energy consumption and air pollution.

The following sections provide discussions on locating and designing park-and-ride facilities. The information in these sections was gathered from several sources, and the reader is encouraged to refer to these sources for further information.[16,17,18]

Location of Facilities

Choosing the optimum location for a park-and-ride facility involves investigating several factors. Factors that should be considered when selecting an appropriate location for HOV-related park-and-ride facilities include the following: [16]

- Locate facilities in advance of areas experiencing major traffic congestion.
- Locate facilities so that drivers have a direct route to the lot.
- Orient facilities to ensure good accessibility and visibility.
- Locate facilities at appropriate distances apart.
- Locate at transit stations.

Encourage cooperation among agencies involved in developing and operating the facilities.

The four general steps in choosing an optimum location for a park-and-ride facility include:

- Establishing site selection criteria.
- Identifying alternative sites.
- Evaluating alternative sites.
- Determining size of park-and-ride facility.

Establishment of Site Selection Criteria

During the site selection process, site selection criteria must be established. The criteria should be used both in the initial screening of potential sites and in the evaluation process. The American Association of State Highways and Public Transportation Officials (AASHTO) publication, *Guide for the Design of Park-and-Ride Facilities*, provides the following descriptions for the criteria to be considered:[17]

- **Facility Development Policy.** The park-and-ride facility development program may make use of publicly

owned property, excess right of way, or property used with the permission of private owners, that may be used for other parking purposes. The policy establishes the guidelines for the relative importance of each type of facility within the program.

- **Site Availability.** The first step in the site selection process is the development of an inventory of potential sites.

- **Site Accessibility.** A site must be easily accessible to commuters and transit vehicles when transit service is anticipated. Park-and-ride facilities that are inconvenient for potential users will not, in most cases, be well utilized and so should be avoided.

- **Site Visibility.** Potential sites should be visible from their access roads. Visibility of park-and-ride facilities contribute to the recognition by passing motorists of their availability. Visibility of a park-and-ride site is a deterrent to vandalism and an asset to the security of vehicles.

- **Projected Demand.** Demand for park-and-ride space is based on analysis of individual travel corridors. The indicated demand provides a guideline for the number of potential spaces and estimated lot size that must be identified within each corridor.

- **Transit Service Availability.** Potential sites for park-and-ride facilities should be located along existing or potential transit routes. The potential for transit service should be considered even at those lots which are intended for carpool and van-pool users. If transit service is to be provided by the local transit agency, coordination is a very important element.

- **Accessibility to HOV Facilities.** Sites located adjacent to HOV lanes, HOV priority ramps or other priority facilities provide benefits to park-and-ride users. Coordination of the location of park-and-ride facility sites with HOV facility development can increase the usefulness, demand, and efficiency of both facilities.

- **Development and Operating Costs.** Since most park-and-ride lots do not collect fees, sites that can be developed economically are desirable. Potential development costs include related non-site costs such as added traffic signalization or intersection improvement costs. Shared use of a private facility, church parking lot, mall, etc., can be a factor in development and operating costs.

- **Available User Benefits.** Sites that provide users with the benefits of both travel time and cost savings are preferable to those that provide only a transfer opportunity.

- **Staged Construction Potential.** In many cases, it may be desirable to test demand analysis accuracy before committing funds for larger lots or more permanent construction. In these cases, it is desirable to have a site that can be developed in stages to reduce resource commitments until space requirements are verified. Staged construction potential is also desirable when projected land use development is expected to generate additional park-and-ride demand.

- **Development of Sites in Environmentally Sensitive Areas.** Special attention must be given to consideration of the placement of park-

and-rides in areas determined to be especially environmentally sensitive.

Identification of Alternative Sites

Alternative sites for park-and-ride facilities can be identified using any one or a combination of the following techniques:[17,18]

- **Office Survey.** The goal is to identify those properties which would most readily be developed for parking and which have suitable access.
- **Field Observations.** Field reconnaissance along travel corridors is the preferred method of potential site identification.
- **Aerial Photography.** Aerial photographic mapping provides an ideal means of locating vacant land and private parking areas that can be easily related to travel patterns and corridor routes.
- **Local Contacts.** District highway engineers, local government officials, law enforcement agencies, transit representatives and other local groups can be used as sources to identify sites that are suitable for park-and-ride use.
- **Inventory Records.** Potential park-and-ride sites should be documented on an inventory record form.

Evaluation of Alternatives

Once alternative sites have been selected for the location of a park-and-ride facility, the next step is to evaluate the alternatives and choose the most appropriate location. Many transit agencies use some type of form or check list to evaluate alternative locations. The AASHTO *Guide* presents an evaluation form that includes criteria for establishing a priority rating system.[17] The primary categories of factors are classified below:

- Location Criteria.
- Site Considerations.
- Economic Considerations.
- Potential User Costs and Time.

The criteria used in the AASHTO approach will vary depending upon an agency's preference and needs.

Agencies may establish their own evaluation technique based on local practice and available resources; however, the procedure used should provide a comparative analysis of each site's potential that results in the most economical and efficient use of park-and-ride facility resources. During the site selection process, local officials should be given an opportunity to comment on sites proposed in their areas.[17]

Determining Size of Facility

Once an ideal location for the park-and-ride facility has been identified, the next step is to estimate the parking area required to serve the estimated demand. Factors affecting the required size of the lot include the following:[16,17,18]

- **Traffic Demand.** The major factor affecting the required size of the parking facility is the estimated average daily demand. It is recommended that the lot be designed to accommodate at least 10 percent more vehicles than the estimated average daily demand in order to ensure that adequate parking spaces are available on a typical day.[16,18]

- **Maximum Walking Distance.** Recommended maximum walking distances range from 120 to 300 m (400 to 1000 ft); however, maximum walking distance should not exceed 195 m (650 ft) whenever possible. Walking distance can be minimized by moving the transit station to a central location. Walking distances longer than 300 m (1000 ft) may require consideration of additional transit stations. [18]

- **Bus Service.** The frequency of buses using the park-and-ride facility affects the demand that can be accommodated. For example, as the headways between buses decrease, the number of passengers that can be serviced increases. The increase in transit capacity may increase demand, which will directly affect the required number of parking spaces.[18]

- **Rideshare Use.** Park-and-ride facilities used strictly for rideshare use (e.g., carpool and vanpool) will not have the same restrictions as those facilities accommodating bus service. For example, maximum walking distance will not be as big a factor because drivers will typically have a prearranged meeting location, rather than walking to a central waiting area.[16]

- **Access.** Inadequate capacity on nearby roadways and intersections may severely restrict the volume of traffic that can enter or leave the lot during a given time.[18]

- **Land Availability.** The size and shape of a lot may be restricted by land availability and/or site development costs.

Typical Designs

After an appropriate location is selected for the park-and-ride facility, the next step is to design the facility. There are a number of existing design standards that can be used as guidelines during the design process. Some of the better known documents are the previously mentioned FHWA and AASHTO reports.[17, 18]

Internal Lot Layout

Park-and-ride facilities that accommodate bus services encompass design elements for both parking lots and transit stations. Elements that need to be addressed in the design process are discussed below:[16]

Functional Area Design. The functional design of a park-and-ride facility should meet the requirements of user groups with different access modes. Examples of access modes that may need to be accommodated include the following:

- Long-term parking.
- Drop-off (or kiss-and-ride) areas.
- Handicapped parking.
- Bicycle racks.
- Pedestrian walkways.

In addition, facilities that allow parking for carpooling or vanpooling will need to provide space accordingly. An example of a typical layout for a park-and-ride facility is shown in figure 6-22.

Figure 6-22. Example of Park-and-Ride Layout. [9]

Internal Circulation. Park-and-ride lots should be designed to provide internal circulation to meet all modes of transportation and to minimize pedestrian/vehicle conflicts. Internal circulation should provide for safe and efficient movement of vehicles, buses, motorcycles, bicycles, and pedestrians.

Amenities. Amenities are elements that are considered to enhance the comfort and convenience of using a park-and-ride facility. The number and type of amenities provided will depend on a number of factors, such as: type of facility, anticipated use, local policies, and available funding. Typical amenities that may be considered for use at various park-and-ride facilities include the following:

- Public telephones.
- Trash receptacles.
- Newspaper vending machines.
- Other vending services.
- Transit information displays.
- Transit shelters.
- Seating/benches.
- Bicycle storage facilities.

Pavement and Drainage. Providing adequate pavement design and ensuring proper drainage are other objectives that must be addressed in the design stage. The pavement design will depend on the functional area that it will serve. Typical sources of design guidelines include the following:

- AASHTO standards.
- Local and state pavement specifications.
- Agency guidelines.

Landscaping. Using landscaping for park-and-ride facilities will enhance the appearance of the facility, which will in turn improve public acceptance and enhance the feeling of security. Landscaping should be designed to be compatible with the type of facility and the surrounding area. It should not interfere with sight distance, safe operation, or access to the facility. Guidance for landscaping is available in several publications, including publications by AASHTO and Federal Transit Administration (formerly UMTA).[19,20]

Lighting. Adequate lighting should be provided at park-and-ride facilities to promote safety and security. Factors that must be considered when designing for lighting include:

- Type of lighting.
- Mounting height.
- Spacing of luminaries.
- Intensity of lighting.

Recommendations for designing lighting at park-and-ride facilities are provided in the AASHTO *Guidelines*.[17]

Security. An important factor in ensuring the success of a park-and-ride facility is providing adequate security. Both personal safety and automobile security are important to users of the facility. Measures that are used to enhance safety and security include:

- Lighting.
- Fencing and gates.
- Security monitoring booths.
- Cameras and surveillance equipment.
- Signing.
- Ensuring adequate visibility from all parts of the facility.
- Quick removal of graffiti.

Environmental Considerations. Any effects that a park-and-ride facility will have on the environment should be considered in the design process. For example, effects on air quality can be addressed by minimizing the number of idling buses and vehicles. Factors besides air quality that should be considered include the following:

- Groundwater runoff and water quality.
- Noise impacts.
- Visual and traffic impacts.

Roadway Interface

Providing adequate access between a roadway and a park-and-ride facility is important to ensure efficient operation of the facility and to minimize the effects on surrounding roadways. Both the FHWA *Guidelines* [18] and the AASHTO *Guidelines*[17] provide design recommendations for interfacing the park-and-ride facility with the surrounding roadway. Design issues that are associated with roadway interface are discussed below.

General Access and Egress Considerations. Providing adequate access and egress for all modes of transportation should be considered when designing a park-and-ride facility. The location of entrances and exits can be controlled by a number of factors, including the following:

- Topography.
- Position relative to adjacent roadways.
- Types of local roadways.
- Types of traffic control.
- Connecting service provided by transit.

Access Points. Another factor that should be considered when determining the location of access points for park-and-ride lots is the effect on the surrounding transportation system. To determine the optimum location of access points, a traffic impact study should be conducted. The study should involve determining the current operating conditions on surrounding roadways and estimating the effects of different access points on existing traffic operations. Factors that may be investigated in the analysis include the following:

- Existing roadway capacity.
- Current traffic volumes.
- Normal projected growth.
- Projected growth due to park-and-ride lot.
- Impact of commercial development that may occur due to location of lot.

Access Roadways. On the basis of results from the traffic impact study, the capacity of the existing roadways should be analyzed to determine if they can handle the additional demand due to the park-and-ride facility. The analysis should investigate both current operating conditions and future (projected) conditions. The outcome of the analysis can be used to identify any needed improvements to the surrounding roadways.

Traffic Control Devices and Traffic Signals. The purpose of traffic control devices is to improve the safety and efficiency of traffic operations by providing guidance to drivers through the use of signs, signals, and roadway markings. The need for new traffic control devices, or for modifications to existing traffic control devices, should be identified in the traffic impact study. Guidance on the use of traffic control devices is found in the *MUTCD*.[13]

Handicapped Considerations

For park-and-ride facilities, handicapped considerations must be addressed during the design process for both the parking and bus loading area.

Parking

The AASHTO *Guidelines* contain standards for providing handicapped parking at park-and-ride facilities.[17] Parking-related factors that must be considered for handicapped provisions include capacity, location, design, and signing and marking.

Capacity. Standards for the minimum number of handicapped parking spaces given in the *Uniform Federal Accessibility Standards* are shown in table 6-7.

Location. Handicapped facilities are to be designed in accordance with State or local codes and should be provided at the nearest possible location to the bus loading zone.

Table 6-7. Accessability Standards. [21]

Total Parking Spaces	Required Minimum Number of Handicapped Accessible Spaces
1 to 25	1
26 to 50	2
51 to 75	3
76 to 100	4
101 to 150	5
151 to 200	6
201 to 300	7
301 to 400	8
401 to 500	9
501 to 1000	2%
Over 1000	20 plus 1 for each 100 over 1000

The facilities should be in accordance with the following considerations:

- Preferably no access roads or bus lanes should be crossed by handicapped patrons en route to the bus loading zone.
- The handicapped patron must never be forced to travel behind parked cars (in their circulation path).
- To facilitate the movement of physically handicapped patrons, wheelchair ramps must be provided.

Design. The designer should consider grading the handicapped parking stall so that the head of the stall is at the same level as the sidewalk and the rear of the stall is at the appropriate aisle grade. This will reduce the need for depressed curbs, keep the handicapped patron from crossing any type of storm drainage, and simplify construction of the handicapped parking area.

Signing and Marking. Appropriate signing and/or pavement markings should indicate the restricted use of the space for handicapped persons. Curbs to and from the bus loading area should be depressed for wheelchairs, as dictated by local standards.

Bus Loading Area

The Americans with Disabilities Act (ADA) of 1990 establishes guidelines for providing reasonable access to and use of buildings, facilities, and transportation. All new or renovated transit facilities must comply with the accessibility regulations in the ADA guidelines. Rules applying to transportation and transportation facilities are contained in *Transportation for People with Disabilities*.[22] The *Americans with Disabilities Act: Accessibility Guidelines for Buildings and Facilities, Transportation Facilities, and Transportation Vehicles* publication, hereafter referred to as the *Guidelines*, contains guidelines for meeting

ADA requirements for transit facilities.[23] Section 10 is titled "Transportation Facilities" and contains guidelines related to transit facilities. It also references portions of section 4 in the *Guidelines* that relate to transportation. A TCRP (Transit Cooperative Research Program) report discusses information in the *Guidelines*

relevant to the design of bus stops.[24] Another TCRP report presents guidelines on the design and location of bus stops, and includes a summary of ADA issues.[25]

PARKING INCENTIVES

Implementing special parking policies can provide incentives for commuters to use high-occupancy modes of travel. For instance, adopting preferential parking for carpools and vanpools can increase the attractiveness of ridesharing. Examples of parking programs to induce commuters to use high occupancy vehicles include the following:[26]

- Preferential parking.
- Free or low-cost parking.
- Differential parking rates.

Preferential Parking

Preferential parking involves providing special parking spaces at the final destination for high-occupancy vehicles. The advantage of this strategy is that the utilization of existing parking facilities is increased without having to add additional parking spaces. Preferential parking strategies include the following:[26]

- Guaranteeing spaces for carpools and vanpools where parking is limited.
- Assigning closest and most convenient spaces to carpools and vanpools.
- Assigning specific garage spaces to carpools and vanpools.

Preferential parking strategies should also be considered as a part of new development. Using these strategies can reduce the parking requirements by designating parking spaces for high-occupancy vehicles during the planning process. This approach allows developers flexibility in meeting minimum parking requirements, while increasing ridesharing. [26]

Free/Low-Cost Parking

Free or low-cost parking is another strategy that can be implemented to encourage ridesharing. Using this strategy, carpools are allowed to park in designated areas for no cost or for a small monthly fee. Some of the parking programs that have been established around the United States allow carpools to purchase a permit for a small monthly fee. The permit allows these vehicles to park in designated spaces or to park at metered parking spaces without paying the standard fee. [26]

Differential Parking Rates

Another parking program to increase ridesharing involves charging commuters a flexible parking rate based on the number of occupants in a vehicle. The cost of parking decreases as the vehicle occupancy increases. An alternative is to maintain a constant parking rate and provide employers subsidies that vary according to vehicle occupancy. An example of this strategy that has been used by some private firms is shown in table 6-8.[26]

Benefits of Parking Incentives

Parking incentives have been shown to have a positive impact on the number of commuters that choose ridesharing alternatives. A study conducted in Washington, DC showed a 20 to 40 percent increase in commuters who chose to carpool when parking incentive programs were adopted. [27] A company in Boston, Massachusetts implemented a parking program in which the daily parking fee was

Table 6-8. Example of Differential Parking Rates. [22]

Vehicle Occupancy	Percentage of Parking Paid for by Employer
Single Occupant	0 %
Two-Person Carpool	50 %
Three-Person Carpool	100 %
Vanpool	100 %

dropped for vehicles with three or more occupants. The company reported that at least 34 percent of their employees shifted to carpools when this strategy was adopted.[26]

Most parking incentive programs are partially or wholly subsidized by employers for a number of reasons, including the following:[26]

- A sense of responsibility to promote energy conservation.
- A desire to reduce the costs of maintaining current parking arrangements.
- The need to eliminate new parking construction.
- Improvement in employee morale when employers take an active role in reducing transportation costs for employees.
- Possible favorable media coverage for participating employers when they are identified as contributing to energy conservation, air pollution control, and reduced traffic congestion.

Barriers to Implementing Parking Incentives

A critical barrier that typically impedes implementation of preferential parking strategies is public and business opposition. Opposition from these groups occurs for the following reasons:[26]

- Merchants, employers, and employees traditionally resist changes in the status quo.
- Merchants are leery of any changes in municipal parking supplies that they feel would reduce their competitive position with other retailers who offer free customer parking.
- Employers and employees often resist changes because parking is part of an overall benefits package.
- Some executives may fear that they will lose their parking privileges by adopting this type of strategy.

There have been some factors, however, that have influenced officials in some cities to call for policies that would enhance the coordination between traffic demand management and parking incentives. Factors to be considered include the following:

- Auto ownership continues to increase. Ownership has increased at a rate of 2.5 times the population growth rate over the past few years.

- Structure parking is becoming prohibitively expensive.

- Metropolitan freeway construction is extremely expensive and is widely perceived to be environmentally disruptive.

EMERGING TECHNOLOGIES

Enforcement

The Texas Transportation Institute (TTI), in a research study sponsored by the Dallas Area Rapid Transit (DART), Federal Highway Administration (FHWA), Federal Transit Administration (FTA), and Texas Department of Transportation (TxDOT), is investigating the use of advanced technologies in HOV lane enforcement. The study is aimed at identifying technologies to improve the safety and cost-effectiveness of HOV lane enforcement. The study's objective is to identify the most promising technology and perform a 6-month operational test on the East R. L. Thornton (I-30) Contraflow HOV lane in Dallas, Texas. The results of the operational test will help DART improve enforcement procedures in HOV lanes in Dallas.

The enforcement system will perform the following basic functions: [28]

- Collect and transmit video images of vehicle license plates and vehicle compartments for all HOV lane users to a remote computer workstation.

- Perform automatic license plate character recognition on the license plate video image.

- Synchronize captured video images of vehicle occupants with license plate numbers.

- Search a license plate database containing vehicle occupancy histories and, based upon failure to meet set criteria, display the vehicle license plate number and vehicle compartment images on a computer monitor for review and enforcement purposes.

After the equipment is procured and installed, an operational test will follow to evaluate and refine the enforcement procedures. The operational test is expected to be concluded by August 1997.

Movable Barrier

The movable concrete barrier system consists of a barrier transfer vehicle and a movable concrete barrier (see figure 6-23). The main advantage of the barrier is its ability to provide a solid, physical separation between opposing flows of vehicles. The movable barrier consists of 0.9 m (3 ft) concrete segments jointed together by pins. A specially designed conveyor system on the self-propelled transport and transfer machine (TTM) is used to shift the barrier laterally across the roadway (see figure 6-24). The distance of the shift can vary from 1.2 to 6.7 m (4 to 22 ft). The T-shaped barrier top is engaged by conveyor wheels on the TTM. The barrier is then lifted several inches off the ground, moved sidewinder fashion through an elongated S-curve, and repositioned to form a new lane. Barriers can be moved at up to 8 km/h (5 mi/h), depending on the circumstances.[29]

The use of the movable barrier to create the HOV lane on the East R. L. Thornton (RLT) Freeway in Dallas began in September 1991. The movable barrier shifts the barrier approximately 6.7 m (22 ft) laterally to

Figure 6-23. Movable Barrier. [9]

Figure 6-24. Movable Barrier Used in Dallas. [9]

create an extra travel lane for the peak direction of flow. The implementation of this HOV lane was accomplished by narrowing freeway lane widths to 3.4 m (11 ft) and reducing the inside shoulder of the freeway in some locations (see figure 6-25). The contraflow lane is 8.4 km (5.2 mi) long. The cost to construct this contraflow lane (in 1995 dollars) was $15.4 million. In December 1995, the East RLT HOV lane served 13,572 daily person trips.[26]

The second contraflow HOV lane to use a movable barrier opened in November 1995 on the I-93 Southeast Expressway in the Boston metropolitan area. Like the project in Dallas, the I-93 contraflow HOV lane is created by moving a concrete barrier from the median across one traffic lane to create an additional HOV lane in the peak direction of travel. After six months of operation, the HOV lane, or *zipper* lane as local commuters call it, has met the operational goals for the initial phase of the project. In 1996, the lane carried over 14,400 people each weekday. Concerns expressed about operation during winter weather were dispelled when the lane was opened 99 percent of scheduled times during the snowiest winter in Boston's history (250 plus cm (100 plus in).[29]

6.4 LESSONS LEARNED

PLANNING

A planning process that yields information on the best possible HOV candidates should be undertaken to ensure an efficient and effective HOV facility. This can aid in determining the type of HOV facility needed and the proper vehicle eligibility requirements. Operational considerations such as operating periods, eligibility, active traffic management, incident response, and enforcement should be addressed during the planning process. When operational issues are not considered, the end result can be an HOV facility that is mismanaged, vulnerable to long periods of nonuse when incidents occur, and difficult to enforce.[5]

Fuhs developed a generic process for evaluating HOV alternatives. [5] The four stages are as follows:

- **Stage 1: Conceptual Viability.** The first of the four stages of HOV development is an evaluation of HOV conceptual viability, accomplished by applying a list of criteria to the candidate corridor. If the corridor fails to meet these criteria, it should be excluded from further consideration, or efforts should be undertaken to better understand what may be needed for the criteria to be met.

- **Stage 2: Alternative Development.** The second stage involves a more qualitative and quantitative assessment of the merits of the candidate corridor and an identification of the specific types of HOV designs and operational elements that are applicable and feasible. There is less likelihood that a candidate corridor will be excluded at this stage; however, a number of design and operation alternatives may be created at this stage that warrant further study. If this is evident, follow-on studies may occur here to better assess the relative merits of each.

- **Stage 3: Development of Recommended Alternative(s).** The third stage involves development of one or more recommended alternative(s). This stage focuses on specific design and operation issues that deserve attention. Activities include the following:

    - Selected elements of preliminary engineering.

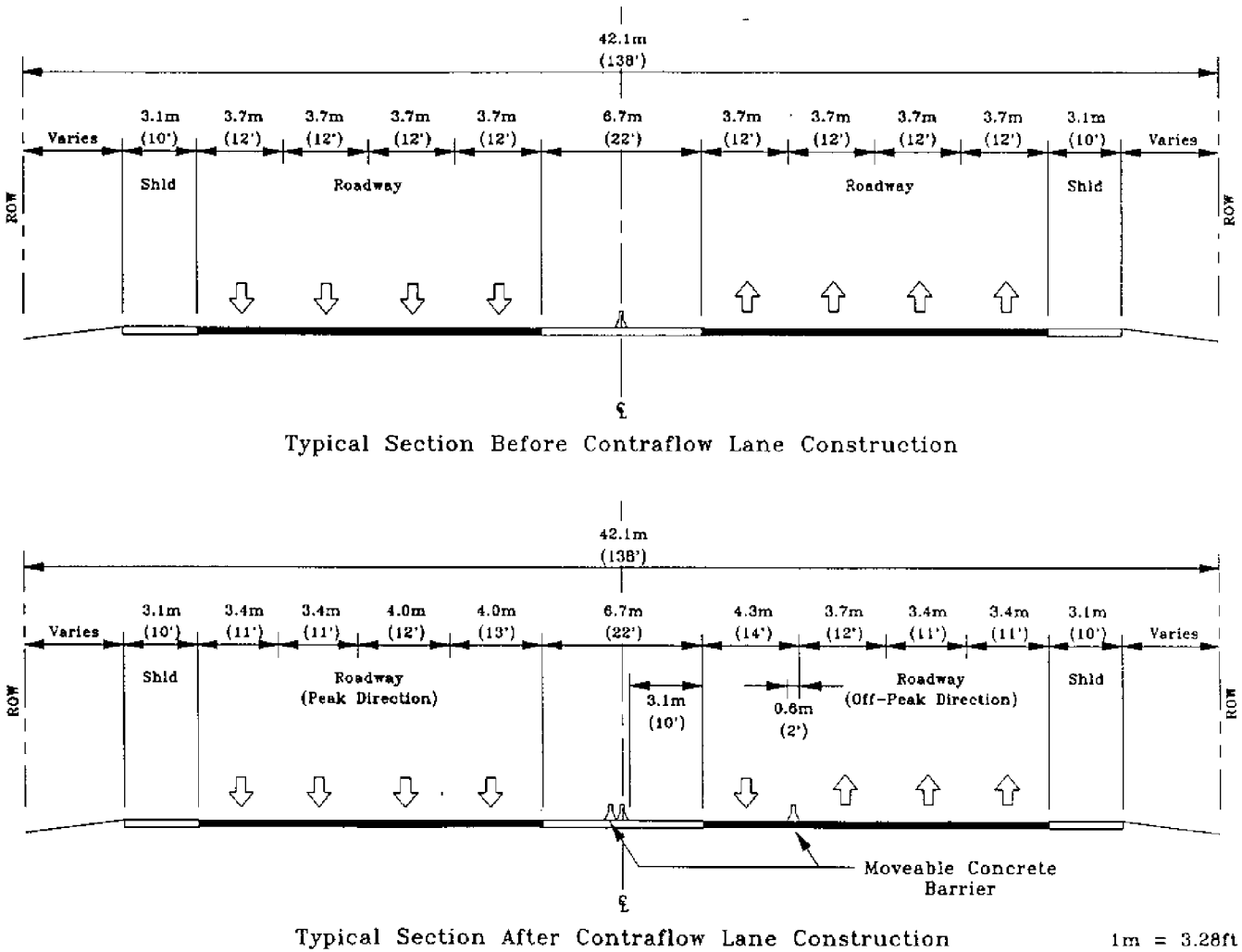


Figure 6-25. Typical Sections, Before and After East RLT Contraflow Lane Construction.[26]

- Checks for constructibility and implementation staging.
- Support facilities (park-and-ride and/or park-and-pool lots, and on- and off-line bus facilities, etc.).
- Identification of supporting programs like ridesharing.
- Implementation phasing.
- Cost estimating and cash flow analyses.
- Other corridor-specific issues.

Products from this third stage include an operation plan and a geometric design plan.

- **Stage 4: Plan Adoption.** The fourth and final stage is formal adoption of the local preferred HOV plan. The results of the HOV planning study are officially sanctioned by respective boards, commissions, or other official bodies, and an implementation process is defined. The adopted HOV plan may be a stand-alone approach or may be one of a number of other capacity alternatives (highway or transit) that are subjected to alternative and environmental analysis before being eligible for design and construction.

Public and Political Support and Consensus Building

HOV concepts are often not understood by the public, politicians, or media. Within the scope of the planning study, a multiagency review group (also called an advisory committee or steering committee) should be formed. Their role includes the following:

- Technical and policy guidance
- Concurrence powers at major decision points.
- Coordination and liaison with others in the respective agencies.
- Outreach to greater public participation efforts as needed.

HOV projects are particularly in need of this type of oversight, where an ongoing commitment is needed among a number of agencies to effectively plan, implement, operate, enforce, maintain, and modify the HOV project as necessary to ensure continued viability. The composition of the review group may vary during the planning and implementation process, but at a minimum, representatives from affected local agencies, police, the state DOT, and the FHWA should usually be involved. If a project is adopted for implementation, this group should continue to function through the design phase, and some semblance of the group should be named to provide operation and maintenance overview of the HOV facility once it is opened.[5]

Apart from the multiagency review group, focus groups composed of local civic associations, special interest groups, politicians, the media, and others may be necessary to bring public participation into the planning process. This process should be highly structured to keep issues focused. Appropriate in many areas will be the need for some public awareness outreach to market the HOV concept. There are four steps to an effective public education program:[5]

- Information should be disseminated three to four months prior to the opening of the project to prepare the public for the new facility and its regulations.

- During the first few weeks of project operation, there should be an intense public relations effort until users are familiar with the HOV facility and its operation.

- As soon as data on project use are available, they should be presented to the public to disseminate information about the positive impacts from the project.

- The public education program should be continued, as needed, to reinforce an understanding of concept objectives and the role of enforcement.

An important lesson that has been learned from project failures to date is that the public must be involved and must be able to understand and appreciate the role that HOV systems can serve. To this end, any HOV planning process will necessarily need to be an information, awareness, and education process as well.

Public Involvement

Traditional methods of public notification and community relations are important to gaining public acceptance of changes to transportation services. The public interest, however, has progressed to a desire to become directly involved in the initial and ongoing decisions about service and design. Public involvement is not public relations. Public relations strategies attempt to position an issue, project, or plan in the best possible light for decision-makers and the public. Frequently the objective is to gain public acceptance of a prior decision. Public involvement, on the other hand, involves the public in the shaping of policy and project decisions that affect a community. Public involvement is best initiated before project alternatives are developed or decisions made. Public involvement can be loosely divided into four components—data gathering, public and private communications, promotion, and evaluation.[2]

- Data Gathering.
  - Literature Search.
  - Surveys.
  - Focus Groups.
  - Focused Interviews.
- Public and Private Communications.
  - Kick-Off Briefing.
  - Community/Jurisdiction/Elected Official Briefings.
  - Public Meetings, Hearings, Open Houses, Workshops.
- Promotion (Heightening General and Targeted Awareness).
  - Media Relations.
  - Newsletters.
  - Speakers Bureau/Public Forums.
  - Advertising and Public Service Announcements.
  - Special Events and Ceremonies.
  - Customer Service Representatives and Other Staff.
- Evaluation.
  - Survey.
  - Focus Groups.

- Focused Interviews.
- Public/Private Partnerships.
- Sponsorships: Corporate & Media.

System Approach to HOV Planning

Most HOV projects have been implemented on a corridor basis. As such, *corridor planning* has been pursued with focused evaluation on the specific needs and travel characteristics of the corridor in mind. As additional corridors have been addressed in a region, there has been a recognition of the need for *regionwide system planning*, in which a broader perspective of HOV applications for a metropolitan area is studied. There have been few cases where regional system planning occurred prior to the study of a particular corridor. Pursuing a regionwide approach assumes a much higher level of commitment to the HOV concept as a valid regionwide option.

Regionwide planning recognizes the relationships among system elements, particularly in staging system implementation to maximize early benefits. Regionwide issues may influence operation polices in which the same hours of operation or rules regarding eligibility are uniformly established for all projects in a given area. This policy of operation consistency has been adopted for Orange County, California, while the Seattle and Houston areas have policies allowing operation policies to be tailored to the unique attributes of each corridor. Both policy approaches can be locally popular and successful.[5]

Possible Pitfalls in Implementation

The following are typical issues to be aware of and to avoid in the implementation process:[5]

- **Policies and Procedures***:* Involved agencies are often unaware of the strings attached to funding, review, and approval steps, and the procedures required for project implementation. Involved agencies should become familiar early on with the rules and policies associated with funding, programming, and implementing projects. This is particularly important when more than one agency is actively involved in the process. FTA and FHWA guidelines are not always in agreement. State policies and procedures may not be familiar to local agencies assisting the State in various activities.

- **Involvement**: In trying to accelerate the effort, some of the agencies that should be involved are overlooked.

  Keep all appropriate agencies involved in decisions being made. This will keep the project properly focused and on schedule.

- **Commitment**: One agency pursues the project without additional external, and sometimes internal, support.

  A dedicated commitment is needed from all the respective parties before proceeding with a project. This provides a broader base of support for sustaining the project.

- **Taking Unnecessary Risks**: In an attempt to achieve faster results, authorization is given to proceed on future activities before all reviews and approvals have been acquired from prior actions.

  Risk management is involved in any implementation process. If such risk is

to be taken, it should be based on a general consensus from the project team with an understanding of the possible downside impacts.

- **Changing Course**: A prior approval or commitment is rescinded, for any of a number of reasons.

  Inevitably, as any project progresses into greater and greater detail of development, there can be justification to re-explore a prior matter that was thought resolved. In such instances, the project manager can be effective in bringing together the affected parties to reach a renewed understanding and, if necessary, to change the course or scope of the project. Recurring problems of this sort may require resolution at higher levels in the organization, providing another reason why a top-down commitment to a project is essential before undertaking implementation.

- **Scope, Budget, and Schedule**: An early effort should be undertaken to understand any change in scope, schedule, and budget and to obtain a resolution that rebalances these variables. Performance and delivery expectations should be based on a realistic understanding of the project requirements. The risks of fast-tracking a schedule, for example, should be defined and understood. Again, a single project manager can be the focal point for this understanding.

Tables 6-9 to 6-10 list characteristics of successful HOV facility implementations and potential pitfalls, respectively.

Development Concerns for Park-and-Ride Facilities

There are many factors that must be considered when implementing a park-and-ride program. These factors are discussed below.

Program Planning

Program planning involves developing the operating plan that defines specific program requirements. These planning requirements are discussed below:

- **Site Development Policies.** Once it has been determined that park-and-ride facilities will accompany specific HOV lanes, the first step is to establish site development policies. These policies will provide guidance in developing the facilities. Typical policies that need to be addressed include the following:

  - Stage construction guidelines.
  - Maximum or minimum size lots.
  - Extent of amenities (e.g., telephones, lighting, etc.) to be provided.
  - Transit service guidelines.
  - Joint accommodation of carpools and vanpools with transit users.

- **Inventory of Potential Sites.** The next step is to identify potential sites. This can be accomplished by a number of means, including the following:

Table 6-9. Successful Characteristics of HOV Facility Implementation. [2]

- ☞ Clear set of objectives and measures of success
- ☞ Development of the HOV lane as an additional lane
- ☞ Existing congestion in corridor (able to save 1/minute/mile & 5 to 8 minutes total)
- ☞ Projections for continued increase in demand
- ☞ Enforcement commitment/collaborative working relationships with enforcement agencies/courts along the corridor
- ☞ Reason to believe you can get support from both agencies and public
- ☞ Reason to believe you can provide a lane that can be safely operated and enforced
- ♥ Policies and programs supporting transit use
- ♥ Rideshare program in corridor
- ♥ Successful HOV facilities already in operation in same corridor or adjacent corridors
- ♥ High existing volume of 2+ HOV's (700 or more vehicles per hour)
- ♥ Traffic system management program already in place along the corridor
- ♥ High level of convenient transit service along the corridor (local/express/Park & Ride routes)
- ♥ Commute trip reduction legislation
- ♥ Existing communication network with employers along the corridor
- ♥ Collaborative working relationships with environmental agencies/groups along the corridor
- ♥ Collaborative working relationships with neighborhood/community groups along the corridor
- ♥ Collaborative working relationships with local jurisdictions/transit agencies/DOT along the corridor
- ♥ Commitment to evaluation to accurately show benefits/disbenefits
- ♥ Origin-destination pattern that can benefit from the HOV lane

Legend
☞= Essential Characteristic
♥= Desirable Characteristic

  - Aerial photographs.
  - Excess property lists.
  - Field reconnaissance.

- **Facility Implementation Priority Rankings.** Once potential sites are identified, a list of the sites is prepared for use in defining the short-term and long-term programs. The next step is to establish a methodology to allocate available funding and to address the most critical needs on a priority basis.

- **Short-Term and Long-Term Site Development Programs.** Short-term and long-term site development programs are established to provide for planning, design, construction, funding, scheduling, and resource allocation.

- **Transit System Interface.** The location of park-and-ride lots should be coordinated with existing and proposed transit routes to maximize the benefit of these facilities.

Table 6-10. Potential Pitfalls to Successful HOV Facility Implementation. [2]

- Converting existing general purpose lane to HOV lane, may result in negative impacts (increased accidents, increased travel time, etc.) in general purpose lanes or protest by the public
- High accident rate along the corridor that will not be improved by HOV lane
- Little support from enforcement authorities (State Police/Patrol, municipal judges/magistrates)
- Low existing volume of HOV
- Poor working relationships with local media
- Poor working relationships with neighborhood/community groups along the corridor
- Poor working relationships with elected officials (especially critical during election years)

➢ Changing lane designation from general purpose to HOV during lane construction (example: begin construction as general purpose, change designation during construction phase to HOV)
➢ Low level of transit service
➢ Lack of transit funding
➢ No/low level of support facilities
➢ No incident management program in place
➢ No/inadequate ride matching services
➢ Poor pavement maintenance of existing facility that would not be helped by HOV lane construction

Legend
•= Potential Fatal Flaw
➢= Possible Problem

- **Roadway Interface.** The development of park-and-ride lots should also be coordinated with HOV lanes to provide for proper design of access routes and entrances. The objective should be to provide access to HOV lanes in a way that will maximize travel time savings for the users of the facility.

Design Requirements

Both transit planners and traffic engineers should be involved during the design process to ensure proper operating conditions for buses and cars. Standard design manuals for park-and-ride facilities are available and should be used to control design quality.[16,17,18] A following section titled *Typical Designs* provides a summary of some existing design guidelines.

Operation and Maintenance

It is very important that operation and maintenance issues be addressed during the development process. For example, it is important to ensure that adequate funding is available for the operation and maintenance of a facility. Also, the agencies and groups responsible for the operation and maintenance of the facility should be identified early in the development process:

- **Operation.** The successful operation of park-and-ride facilities will require close cooperation among local transit agencies, the community, and the State DOT. The operation of a park-and-ride facility involves a number of elements, including the following:[18]
  - Parking fee structures.
  - Frequency of transit service.
  - Fares for transit service.
  - Carpool use.
  - Providing security.
- **Maintenance.** Providing proper maintenance of a facility also has a significant impact on the success of a park-and-ride program. Ensuring that the facility is clean, attractive, and well maintained will have a positive impact onusers. In developing a comprehensive maintenance program, several factors must be considered:[16]
  - Periodic inspection.
  - Pavement repair.
  - Shelter or station repair.
  - Traffic control devices (e.g., signs and pavement markings).
  - Lighting.
  - Mowing.
  - Sweeping and cleaning.
  - Trash removal.
  - Landscaping.
  - Site furnishings.
  - Snow and ice maintenance.
  - Security/gates.

Program Marketing

The objective of a marketing program is to inform motorists about the park-and-ride facility and available transit services. It is important to develop a marketing program during the implementation of a new park-and-ride facility; however, ongoing marketing efforts are also important for the continued promotion of the facility. Two important needs that should be addressed when developing a marketing campaign are as follows:[16]

- Identify target audience.
- Determine effective mechanism for communicating desired information.

Examples of actual marketing techniques include the following:[16]

- Direct mail.
- Radio.
- Television.
- Newspaper.
- Billboards.
- Roadside signs.
- Lot location maps, transit maps, and transit schedules.
- Employer-focused efforts.

For further information on the marketing of HOV systems, the reader is referred to the

High Occupancy Vehicle (HOV) Lane Marketing Manual.[30]

DESIGN/CONSTRUCTION

The development of HOV facilities within freeway corridors is typically a retrofit process. Not all desirable design standards are likely to be obtainable. Some compromise in design requirements will probably be needed, and the elements of compromise will differ from one location to another. Tradeoffs need to be considered on a case-by-case basis, according to an agreed-upon understanding of what design treatments are acceptable to the appropriate reviewing agencies. Therefore, HOV practitioners recognize that the typical recommended designs provided in any of the currently available design resources must be considered as *guidelines* and not *standards*. The guidelines are of two types, *desirable* and *reduced*.

- *Desirable* design guidelines are those that should be considered in the concept development of any new or substantially reconstructed freeway corridor or HOV guideway in separate right-of-way. Desirable guidelines are generally recognized as reflecting a permanent facility treatment with the same commensurate design as that which would be applied to a new freeway with full design standards. HOV projects constructed with these dimensions have had significant costs and have been accompanied by substantial freeway reconstruction as well.

- *Reduced* design guidelines are those that may be considered for retrofit projects on a case-by-case basis, where space or cost considerations might otherwise make application of the HOV treatment environmentally infeasible or cost prohibitive. Reduced designs often reflect restriping or otherwise modifying the existing freeway to achieve an HOV operation that may be termed "interim," with a longer-range intent for meeting desirable guidelines. The reduced dimensions have also been used in short sections of projects otherwise meeting desirable standards, where significant cost would have been incurred to provide the desirable widths.

Tradeoffs in Retrofitting Cross Sections

The process of retrofitting HOV facilities onto existing freeways requires design compromises. Often, bridge columns, narrow structures, limited right-of-way, and other isolated constraints are encountered which make adherence to desirable or even reduced criteria difficult without some commensurate reduction in the overall cross section of the freeway. Design compromises should be thoroughly investigated before an HOV facility is implemented. The key concerns relevant to this issue appear to be cost effectiveness of the remedy and length of the design exception required. If reductions in typical guideway standards are necessitated, general practice indicates that the sequence of tradeoffs presented in table 6-11 should be considered when making these determinations.[5]

Impacts of Ingress/Egress on Adjacent Freeway and Arterial System

HOV ingress/egress designs have evolved through experience, and much has been learned from various applications. Following are a few general guidelines:[5]

- Where possible, the same geometric criteria should be applied as would be used for a freeway ramp, including locally recognized entrance and exit design standards.

Table 6-11. Tradeoffs in Reducing Cross Section Elements.[5]

Compromise Sequence	Cross Section Element		
	For Reversible-Flow HOV Facilities	**For Two-Way Barrier-Separated HOV Facilities**	**For Buffer-Separated HOV Facilities**
First	Reduce single-lane HOV envelope to no less than 6 m (20 ft), or two-lane envelope to no less than 8.5 m (28 ft).	Reduce left HOV lane lateral clearance to no less than 0.6 m (2 ft).	Reduce left HOV lane lateral clearance to no less than 0.6 m (2 ft).
Second	Reduce freeway left lateral clearance to no less than 0.6 m (2 ft).	Reduce right HOV lane lateral clearance to no less than 2.4 m (8 ft).	Reduce freeway right lateral clearance (shoulder) from 3.1 m (10 ft) to no less than 2.4 m (8 ft).
Third	Reduce freeway right lateral clearance (shoulder) from 3.1 m (10 ft) to no less than 2.4 m (8 ft).	Reduce freeway left lateral clearance to no less than 0.6 m (2 ft).	Reduce buffer separation to no less than 0.3 m (1 ft).
Fourth	Reduce HOV lane width to no less than 3.4 m (11 ft) (some agencies prefer reversing fourth and fifth steps when buses are projected to use the HOV facility).	Reduce freeway right lateral clearance (shoulder) from 3.1 m (10 ft) to no less than 2.4 m (8 ft).	Reduce HOV lane width to no less than 3.4 m (11 ft) (some agencies prefer reversing fourth and fifth steps when buses are projected to use the HOV facility).
Fifth	Reduce selected mixed-flow lane widths to no less than 3.4 m (11 ft) (leave at least one 3.7-m [12-ft] outside lane for trucks).	Reduce HOV lane width to no less than 3.4 m (11 ft) (some agencies prefer reversing fifth and sixth steps when buses are projected to use the HOV facility).	Reduce selected mixed-flow lane widths to no less than 3.4 m (11 ft) (leave at least one 3.7-m [12-ft] outside lane for trucks).
Sixth	Reduce freeway right lateral clearance shoulder from 2.4 m (8 ft) to no less than 1.2 m (4 ft).	Reduce selected mixed-flow lane widths to no less than 3.4 m (11 ft). Leave at least one 3.7-m (12-ft) outside lane for trucks.	Reduce freeway right lateral clearance shoulder from 2.4 m (8 ft) to no less than 1.2 m (4 ft).
Seventh	Convert barrier shape at columns to a vertical face.	Reduce freeway right lateral clearance shoulder from 2.4 m (8 ft) to no less than 1.2 m (4 ft).	Transition barrier shape at columns to a vertical face or remove buffer separation between HOV and mixed-flow lanes.
Eighth		Convert barrier shape at columns to a vertical face.	

Adapted from Caltrans *HOV Guidelines*, as reported in the reference.

- The typical volumes using HOV access ramps are less than for commensurate freeway ramps. HOV volumes accessing a local street seldom overwhelm a signalized intersection. Average peak-hour ramp demand at an intersecting street probably will not exceed 300 to 500 vehicles per hour.

- Sight distance is particularly critical due to the proximity of barriers to ramp lane alignments. Lateral clearances are often no further than 0.6 m (2 ft) from the edge of the travel lane to barrier. Where practical, removal of barrier-mounted glare screens or slight adjustments in striping alignment may be necessary within the ramp envelope to accommodate the proper design speed.

- The location of ingress/egress facilities is influenced by a number of factors. For example, direct access ramps to/from local streets should be made with candidate streets that currently do not have freeway access, so as to better distribute demand and prevent overloading existing intersections. For at-grade access to the adjacent freeway lanes, designated outlets should be strategically positioned so as to minimize erratic weaving to reach nearby freeway exits.

- Left- or right-hand exits from a one-lane HOV facility are equally valid and equally safe. The standard "right hand only" rule for entrance and exit ramps should not apply to HOV facilities.

- Where possible, entering and exiting movements should be shared on the same structure or space envelope between retaining walls to reduce cost.

- Adequate advance signing should be provided, and pavement markings should emphasize the mainline (possibly through use of skip stripe markings across the diverging exit ramp).

- Safety lighting should be applied for all ingress/egress locations, using the same warrants as are applied for metropolitan freeway entrance and exit ramps.

- Where possible, provision for entrance ramp metering and/or enforcement should be considered (these are project-specific considerations based on a number of local issues and input from enforcement agents).

Signing Guidelines

Signing guidelines for HOV facilities should be no different than for any other high-speed highway application. Regional consistency should be considered when arraying periodic information regarding rules, policies, and operating procedures. The pitfalls commonly evidenced on projects include the following: [5]

- Lack of adherence to MUTCD color standards.

- Lack of diamond symbol exhibited on signs.

- Confusing regulatory sign information.

- Sign lettering too small to be read at the posted speed.

- HOV guide signs that can be read and misconstrued by mixed-flow drivers.

- Signs that have not been adequately upgraded to reflect the most current operating rules.

The following are general recommendations for handling these shortcomings and applying the MUTCD chapters:[5]

- The standard sign for HOV regulatory and guide signing is black-on-white, with a white diamond symbol on black background in the upper left corner. The diamond symbol should appear on all signs related to HOV user and nonuser communication (regulations and guidance).

- The size of the sign should be commensurate with the design speed, using the same relationship in letter size as is required for any other highway signing.

- Signs should be considered on the approaches to all HOV ingress and egress locations and at the ingress/egress locations themselves. It may be necessary to mount such signs over the lane to provide proper sizing. Guide signing for HOV users is necessary on HOV facilities that have designated ingress/egress locations.

- Regulatory signs related to lane restrictions should be mounted over the lane at regular intervals.

- Regulatory signs located in advance of a preferential lane should be mounted on the appropriate side of the approach roadway.

- If dynamic signing is used, the diamond symbol should be applied on or above the variable sign in the upper left corner.

- Regulatory signing should periodically clarify user eligibility, hours of operation, directionality (if appropriate) and other restrictions on the operation along the length of the facility. Regulatory information should be repeated along nonbarrier-separated facilities at 0.8 to 1.6 km (0.5 to 1.0 mi) intervals. Regulatory information for barrier-separated facilities can be displayed at entrances and exits only.

- Carpools and vanpools should be defined in terms of "PERSON" numbers. (Sample message: Buses and 2+ *Person* HOVs only.)

- Where lateral clearance is limited, as for a ground-mounted sign affixed to the median barrier, the sign panel may have to be skewed as much as 30 degrees to allow for an increase in panel length.

- All signing should be reevaluated whenever operation policies change. Signs should be changed in accordance with the revised policies and not be cluttered with amended information that confuses or contradicts the primary message or information being communicated.

- Where guide signing for similar destinations can potentially be seen by both HOV users and nonusers, the HOV message should be reinforced with the diamond symbol. All guide signs in the area of confusion should be mounted over the appropriate roadways.

Pavement Marking Guidelines

There is generally more diversity in pavement markings, and this is due to wide variation in locally accepted standards and uncertainty over proper pavement marking treatments. Following are general recommendations for interpreting and applying guidelines in the MUTCD: [5]

- Reserve the use of the diamond symbol on signs and pavement markings for

HOV treatments. Although the MUTCD recommends this symbol for bicycle lanes, and for commercial or other purposes not related to HOV, practitioners now concur with restricting the diamond symbol to HOV applications. The placement of the diamond on the pavement should be frequent enough to communicate the restrictive nature of the lane. Along higher-speed freeway facilities, the interval should be from 153 to 305 m (500 to 1,000 ft).

- The colors of pavement markings should, as a general rule, be in compliance with the MUTCD. White is usually applied to communicate a concurrent-flow condition; yellow is applied to communicate a potential for adjacent oncoming vehicles. Solid lines denote that crossing of the marking is not permitted. There are local exceptions.

- For a barrier-separated two-way configuration, a solid white stripe should delineate a right shoulder and a solid white or yellow stripe delineate the left edge of the travel lane. For buffer-separated lanes, the buffer should always be delineated with solid stripes, preferably white, if continuous access is not allowed. Some buffer-separated projects use a combination of white and yellow. Nonseparated lanes that allow continuous ingress/egress and become mixed-flow lanes during at least part of the day should apply standard freeway lane delineation. Contraflow lanes should apply a separation stripe that is either a single or double yellow dashed stripe . Queue bypasses should apply the same pavement markings as are routinely applied along ramps. The stop bars for meters should be highly visible and quite large (typically 20 to 30 cm [8 to 12 in] wide).

- Use of dashed (skip-stripe) markings across an exit ramp should be applied where there is possible confusion over the orientation of the mainline HOV facility. This is more likely where HOV ramps enter and exit from the left.

- Cross hatching of buffers of gore areas helps direct traffic, but is difficult to maintain. Some projects include cross hatching initially; however, they are not maintained. Drivers' experience and knowledge of the HOV facility provide guidance after the markings fade.

- Avoid use of raised reflectorized delineators of buttons in conjunction with pavement markings wherever motorcycles are allowed as eligible users. Consult with local police if motorcycle enforcement patrols are being considered.

OPERATIONS

Surveillance, Communication, and Control

An HOV facility operating without a surveillance, communication, and control (SC&C) system must rely on area traffic bulletins, police reports, or bus operator communications to locate and remove incidents. This approach can be acceptable and has proven adequate in most HOV applications. With the SC&C system, loop detectors and CCTV provide operators with information on the real-time operating status. An incident can be located in seconds. It is this reduction of time in locating a disabled vehicle, and its subsequent prompt removal, that provide travel time savings to the HOV facility users in the event of an incident. And, like any

freeway operation, the greater the traffic volumes, the greater the delay that is associated with the increasing number of incidents.[10]

For the HOV facility to maintain a higher quality of service at all times, electronic surveillance and communication tools can help operators respond to an incident that disrupts flow. Applying technology to both the HOV facility and the mixed-flow facility in a common corridor may bring incremental benefits and economies of scale. Either consideration or expanded use of surveillance and communication techniques is prudent in the design phase, even if the only steps taken are to allow for easier adoption of this technology at a later date. Typical treatments in the roadway design that complement future installation include the following:[5]

- Inclusion of conduits adjacent to or under the outside shoulders (normally two, 7.6-cm (3-in) conduits are adequate). Placement of conduits inside barrier alignments is not recommended due to the difficulty in accessing the conduits.

- Inclusion of conduit valve openings within any longitudinal box beam structures and abutments.

- Placement of additional induction loops in Portland cement concrete pavement at 0.4 to 0.8 km (¼ to ½ mi) intervals. Up to three loops in a series should be applied, even if all are not ultimately used.

Incident Response

Once an incident is detected, a key to minimizing delay to HOVs is the speed with which the incident is cleared. Effective incident response must include service equipment and facilities that, upon detection and location of an incident, allow for the rapid removal of that incident.[10]

One important consideration in incident management on HOV facilities is the cooperation of the agencies responsible for providing the needed response. Often more than one department of an agency, or more than one agency, is involved. Since the priorities within each agency may be different, it is sometimes difficult to achieve the full cooperation of all parties. Matters involving multiple jurisdictions can also complicate the management process. To overcome these differences, it may be necessary to create an incident management team composed of representatives of the major operating agencies and governmental entities. At a minimum, the incident management team should coordinate incident response with existing groups or freeway incident management personnel, if any.[10]

An additional resource in incident response is the media. The dissemination of real-time traffic information by a central control facility to radio and/or television stations allows all citizens, especially commuters, to receive up-to-the-minute traffic information.[10]

Enforcement Strategies

Enforcement strategies are influenced by the compliance goal for a facility, accepted local practices of the enforcement agencies, and available staffing and resources. An acceptable violation rate varies from one type of priority treatment to another. In general, violation rates should be capable of being managed to no more than 10 to 20 percent of the observed traffic stream in the HOV facility.[5] Four primary enforcement strategies are being used on HOV facilities:[31]

- **Routine enforcement** activities are those activities which are randomly conducted in concert with normal police monitoring duties. Routine enforcement could be appropriate if the violation rate experienced is considered acceptable by the project management.

- **Special enforcement** is characterized by continuing, systematic staff allocations and enforcement tactics specifically dedicated to enforce HOV violation rules. It is used when routine enforcement cannot effectively address HOV needs without sacrificing other enforcement duties.

- **Selective enforcement** combines both routine and special approaches. Although most enforcement is routine, special tactics are applied periodically to specific problem areas where violations have been observed.

- **Self enforcement** means having motorists and HOV users police themselves, identifying noncompliance with the HOV facility and taking voluntary actions to report violators.

Other Users

Other users that are eligible (or considered for eligibility) for some HOV facilities include the following:[5]

- **Motorcycles.** Inclusion of motorcycles has been promoted by some as an effective and safe way of segregating these vehicles from the mixed-flow traffic stream. The 1982 Surface Transportation Act passed by Congress leaves the decision, in effect, to local authorities. The act specifically includes motorcycles as eligible to use HOV facilities constructed with federal aid funds, unless the responsible operating agency(s) take measures to have motorcycles prohibited for safety reasons. There does not appear to be any nationwide evidence to substantiate reports of positive or negative impacts associated with motorcycles. Inclusion of motorcycles as eligible users seems to be a local issue that has been effectively addressed at the local/state level.

- **Commercial vehicles.** The role of commercial trucks in moving goods to market has expanded significantly. Reducing commercial vehicle transportation costs can mean a higher domestic standard of living and a more competitive position for the U.S. in the world marketplace. For these reasons it is becoming increasingly viable to think of trucks as priority vehicles. A policy that includes large commercial vehicles (e.g., any vehicle with more than two axles) in HOV facilities has been proposed over and over, particularly by the general commuting public and politicians, as a means of remedying conflicts in the mixed-flow traffic stream. With greater focus being given to restricting trucking operations on metropolitan freeways or during peak hours, the possibility of opening HOV facilities in the off-peak periods could have merit. However, the differing origins and destinations of commuters and commercial trucks cannot be easily accommodated. HOV facilities often contain design elements that are restricted within the geometric limitation of the surrounding freeway, making accommodation for trucking movements difficult. Weaving movements of trucks can adversely affect mixed-flow operation where median HOV ingress/egress is primarily via these adjacent lanes. There are also some questions about the safety of such a

vehicle mix, particularly the effect on sight distances and operating speeds.

Even so, Environmental Protection Agency (EPA) regulations at 40 CFR 88.313-93, which are based on section 246(h) of the Clean Air Act, added by the 1990 Clean Air Act Amendments, allows fleet vehicles from centrally fueled fleets of 10 or more that have been certified and labeled as inherently low emission vehicles (ILEV) to use HOV lanes in certain ozone or carbon monoxide nonattainment areas, regardless of the number of occupants, unless they would "create a clear and direct safety hazard."[32] If ILEVs will be allowed to use HOV lanes, the lanes need to be designed to limit the negative aspects of mixing commercial vehicles and HOVs discussed above.

- **Deadheading Vehicles.** The term "deadheading" refers to transit operators' operating empty buses on HOV facilities for the return trip to their routes. By allowing transit operators to deadhead on HOV facilities, transit services are afforded greater operation efficiency and more visibility on HOV facilities. However, it is not desirable to issue blanket approval of transit operator deadheading, because there are instances where the appearance of empty vehicles on HOV facilities may create a public perception problem. Also, there could be numerous private transit operators in a metropolitan area who may wish to deadhead on HOV facilities, possibly causing unnecessary congestion on the facilities.

- **Emergency Vehicles.** Emergency vehicles are generally allowed on all HOV projects. In practice, however, few emergency vehicles partake of the HOV lanes, because the ingress and egress locations are not generally oriented to the nearest emergency facility (e.g., fire station or hospital). Inclusion of emergency vehicles does not affect the design envelope that is applied, especially if all classifications of buses are accommodated. Any emergency vehicle that uses an HOV facility should be properly identified and involved in a legitimate mission. Use of the lane by off-duty enforcement or unmarked vehicles deteriorates public respect for the operation and its policies and should be discouraged.

MAINTENANCE

Maintenance needs, like incident handling, tend to be the same as for any other metropolitan roadway. HOV lanes can be designed to reduce certain maintenance needs. Following are a few recommendations:[5]

- Drainage inlets should not be placed under barriers. Inlets may be cast into barriers, but slots should be wide enough to be maintained. No inlets should be located in a lane of travel.

- Access to overhead signs should be available. Closed-circuit television, variable message signs and lighting need special attention to ensure that they can be accessed for maintenance without blocking a travel lane.

- Snow removal may require special consideration. This need can dictate the width of shoulders for temporary storage of plowed snow, or intermediate openings in barriers for access by snow removal equipment. HOV lanes should be afforded a higher priority for snow removal than other portions of the roadway.

- Litter removal from the median carries a higher priority on most HOV treatments, particularly on treatments that do not provide adequate lateral clearances with barriers. Even small objects that are blown against a barrier can affect HOV operation. Barrier-separated facilities tend to collect debris blown in from mixed-flow lanes. Some projects have resorted to sweeping or driving the HOV facility in advance of operating periods to remove major obstacles.

- Transit support facilities, most notably bus transfer stations, can be subject to extensive vandalism, and require an inordinate amount of maintenance. Special attention to material selections that are in keeping with anticipated use (staffed versus unstaffed facilities) are recommended.

- The typical one-lane HOV roadway is subject to extensive wear along a specific wheel path. It is not uncommon for ruts in asphaltic concrete surfaces, or tears in the seams between differing pavement materials, to create ongoing maintenance problems. Every effort should be made during the design process to develop a consistent pavement section with sufficient quality for longevity.

- Barriers between parallel directions of travel (HOV and mixed-flow) should be only loosely fixed to the pavement, and barriers preferably precast. This design treatment allows sufficient flexibility in the future to move barriers when the inevitable need arises to handle traffic during pavement resurfacing and rehabilitation activities.

- Pavement markings and HOV signing should be designed with longevity in mind, and properly maintained to minimize confusion and ensure proper operation.

- The design of the HOV facility should not attempt to address every detail. To optimize operation, there will be a need for maintenance activities to address minor modifications to signing, pavement markings and other elements of the project that require some refinement as operating experience is acquired. Capital or maintenance funds should be budgeted for these inevitable needs.

CONCURRENT FLOW LANES

The following issues should be considered when planning and designing concurrent flow HOV lanes:

- Phased construction is sometimes used to build concurrent flow HOV lanes. Often, the finished portions are opened to regular traffic. Converting these lanes to HOV operation can cause opposition and resentment. System planners and operators should either keep the lane free of traffic until it is completely finished or impose HOV restrictions on any opened partial section. In both cases, marketing and public relations campaigns may be needed to explain how the HOV lanes will be used during construction.

- Careful consideration should be given to how vehicles merge into and out of concurrent flow HOV lanes. Designers should be careful not to create a design where unsafe driving conditions are generated because of the speed differential between HOV and regular traffic lanes.

- Concurrent HOV lanes can be used by aggressive drivers to pass slow moving

vehicles, and to weave in and out of regular flow traffic. Regular enforcement activities may be needed to curb this activity.

6.5 EXAMPLES OF HOV TREATMENTS

Two case studies of HOV Systems are offered. These studies provide an overview of the HOV facilities in Houston, Texas, and Seattle, Washington.

HOUSTON

Houston, Texas hosts the most extensive network of barrier-separated HOV facilities in the United States.[33] Ultimately, the Houston HOV system will incorporate 166 km (103 mi) of HOV lanes.[34] The Houston HOV system has enjoyed a great deal of success largely due to the coordinated support offered by the Metropolitan Transit Authority of Harris County (METRO) and the Texas Department of Transportation (TxDOT).[33]

The HOV facilities in Houston were created in response to a surge in traffic congestion that began in the 1970s. The City of Houston and TxDOT sought to improve transit operations and reduce traffic demands by increasing vehicle occupancy from an average occupancy of 1.2 persons per vehicle through HOV priority treatments. Houston's HOV experience began with the development and operation of a 14.4 km (9 mi) contraflow lane on the North Freeway (I-45). The contraflow lane hosted only authorized buses and vans and, when opened for use in 1979, was extremely successful. The contraflow lane was operated only for 2.5 hours during each peak period, but was credited with moving over 8000 persons during those times. As a result of the success of the contraflow lane, Houston began development of a large-scale HOV system concept.[33]

Having committed to the development of HOV facilities in Houston, the City of Houston and TxDOT laid plans for the development of a 166 km (103 mi) HOV lane system as illustrated in figure 6-26. While two-direction HOV facilities are being developed, the typical HOV lane in Houston currently provides for reversible operations, and is separated from general-purpose freeway lanes by concrete median barriers (see figure 6-27). The HOV lanes are typically located in the freeway median and measure approximately 6 m (20 ft) in width. Due to geometric constraints, freeway mainlanes and inside shoulder widths were narrowed in some locations to accommodate HOV lane construction.[33]

Access to the Houston HOV facilities is provided by several means and is typically provided at 5 to 8 km (3 to 5 mi) intervals. Most access to the HOV lanes is provided by grade-separated interchanges. At these interchanges, the HOV lane is elevated in the freeway median, and grade-separated ramps are used to convey traffic over the freeway lanes to select local streets, transit facilities, and park-and-ride lots (see figure 6-28). The grade-separated ramps were chosen for their ability to eliminate interference with freeway mainlane operations, enhance safety, reduce travel time, and provide locations for enforcement personnel to operate. The construction costs of the grade-separated interchanges are estimated to range between 2 and 7 million dollars each. In addition to the grade-separated ramps, some access to the median HOV lanes is provided through the use of slip ramps, as shown in figure 6-29.[34]

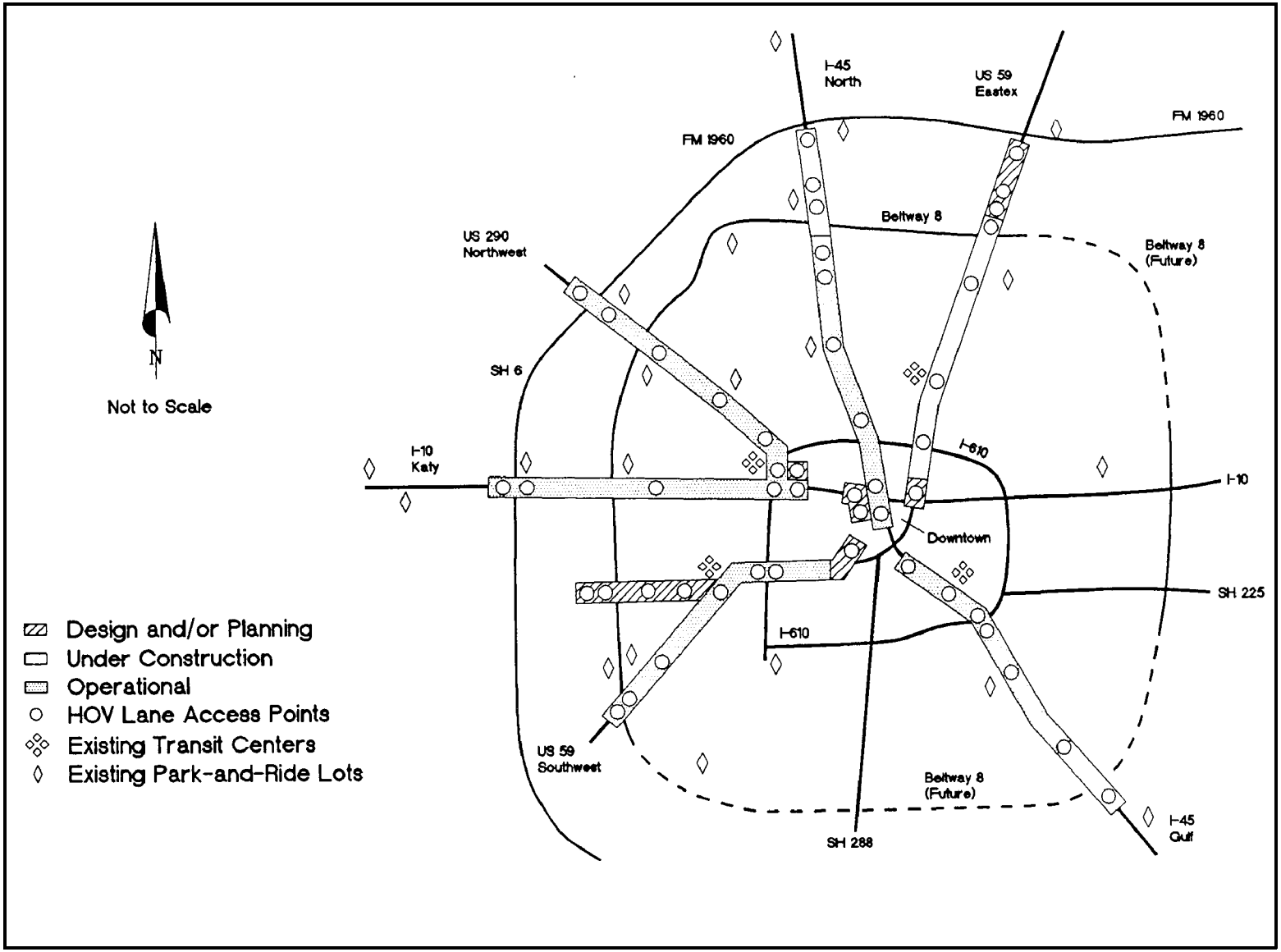


Figure 6-26. Houston HOV System.[34]

Figure 6-27. Typical Houston HOV Lane.[9]

Figure 6-28. Houston Grade Separated HOV Lane Access Ramps.[9]

The occupancy requirements of the Houston HOV lanes, as well as the vehicles that are allowed on them, continue to evolve. Changes have been made to both the occupancy and vehicle requirements to manage traffic volumes on the HOV lanes (see figure 6-30). The changes were made to ensure the desired high speeds and reliable trip times that contribute to the success of HOV lanes.[33]

As of December 1995, the system was hosting in excess of 77,000 person-trips per day.[34] Peak-hour vehicle occupancy has increased by approximately 20 percent on freeways with HOV lanes as compared to their pre-HOV conditions. Freeways without HOV lanes witnessed a decrease in occupancy during the same periods. Given the quasi-independent status of the HOV lanes in relation to the freeway mainlanes, operations on the freeways and other parallel routes have not been significantly affected by the HOV lanes. No significant changes in freeway and parallel route congestion have occurred. This finding is attributed to the HOV's ability to reduce the growth rate in congestion, as opposed to reducing the existing levels of congestion.[33]

Figure 6-29. Houston HOV Slip Ramp.[9]

Figure 6-30. Example of Dynamic Message Sign Showing New Occupancy Requirement.[9]

Contributing to the success of the HOV lanes has been the development of several strategically located park-and-ride lots, transit centers, carpool lots, and downtown bus lanes (see figure 6-31). Park-and-ride lots themselves have experienced a tremendous growth in demand due to use of the HOV lanes. In two of Houston's HOV lane corridors, use of the park-and-ride facilities increased 200 percent or more from pre-HOV levels. Similar corridors not equipped with the HOV lanes experienced

only small changes during the same time period. [34]

Travel time savings offered by the HOV lanes have also served as a lucrative means of attracting new users. Because the Houston system is still relatively new, it is not believed to have reached its full potential yet. Nevertheless, time savings range from five minutes on the Gulf Freeway HOV lane to 18 minutes on the Katy Freeway HOV lane. The current 102 km (63.6 mi) system is estimated to result in 41 minutes of time savings, or one minute per mile. Recent survey evidence indicates that the time savings perceived by users of the HOV lanes are much higher and it is expected that future benefits will be even greater as congestion on freeway mainlanes increases.[34]

On the basis of survey evidence, the typical users of the HOV lanes are characterized as young, educated, white-collar commuters. Furthermore, over 90 percent of the transit bus riders using HOV lanes have an automobile at their disposal, indicating a significant expansion of the transit market. Principal reasons cited for using the HOV lanes include the following:[28]

- Freeway too congested (22%).
- Saves time (18%).
- Time to relax (16%).
- Reliable trip time (15%)

Sixty-five percent of the total person trips on the HOV lanes are currently served by vans and carpools. The remaining 35 percent are served by buses. [34]

The latest regional plan for the HOV system estimated that the entire Houston HOV system will have cost approximately $900 million, or $5.5 million per km ($8.8 million per mi), when the system is complete. These costs, estimated in 1995 dollars, include the HOV lanes, HOV lane access and egress ramps, all park-and-ride lots, park-and-pool lots and bus transfer centers, and the HOV surveillance, communication, and control systems. The exact capital cost of the HOV lanes will be difficult to determine because the HOV lanes have usually been constructed in conjunction with freeway renovation projects. [34]

Figure 6-31. Examples of Park-and-Ride Lot and Bus Station.

SEATTLE

Seattle's HOV system is sponsored mainly by the Washington State Department of Transportation (WSDOT) and is eventually expected to include the HOV lanes depicted in figure 6-32. The Seattle HOV system traces its origins to several failed heavy rail and freeway expansion planning projects that were initiated during the 1960s. From those early defeats arose an HOV system that is now considered to represent a long-term component of Seattle's transportation system that doesn't compete with other transit guideway technologies under development. [3]

The first HOV project in Seattle arose as a result of an express bus service demonstration project sponsored by the Urban Mass Transit Association (UMTA) and known as the Blue Streak. The Blue Streak Project involved the operation of eight different transit routes using reversible express median lanes along I-5 in conjunction with downtown ramps. Owing to the success of the Blue Streak program, funding was allocated to further develop and evaluate the Blue Streak program and led to additional HOV operations. In 1972, the Puget Sound Governmental Conference and the regional transit agency published their first transit plan that addressed exclusive/preferential lanes. Further expansion of HOV use came with projects involving I-90 during the late 1970s. During the 1980s, an underground bus tunnel through the downtown area was completed, the I-5 north reversible lanes were extended, concurrent-flow HOV treatments were used on I-5, I-405, and SR 67, and the SR 520 HOV operation witnessed the expansion of its operational hours and eligibility requirements to meet increased demand. [3] Figures 6-33 to 6-35 show typical Seattle HOV facilities.

As each of the HOV projects progressed, the planning process that supported their development evolved. Interagency panel reviews and a task force representing all affected agencies was formed to provide guidance in HOV planning and operations. [3] Public information measures have been used, including telephone and mail-out surveys, newsletters, media events, and other means. In 1992, WSDOT initiated a major policy shift with one of its reports that stated, "When new capacity options are proposed, one of the alternatives to be considered shall be the conversion of a general purpose lane to an HOV lane." In the past, WSDOT had created HOV lanes through construction of new lanes. The first application of this new concept involved the creation of a HOV lane on a 7.5 mile segment of I-90 which had been experiencing lane balance problems. After receiving FHWA approval in the spring of 1993, a combination of lane conversion and new lane construction was undertaken to create new eastbound and westbound HOV lanes. Other sites were later placed under consideration for similar lane conversions.[36]

In most cases, WSDOT serves as the sponsoring agency for HOV lane projects. Funding has been secured from a variety of sources, including a special high-capacity funding package based on regional sales tax revenue.

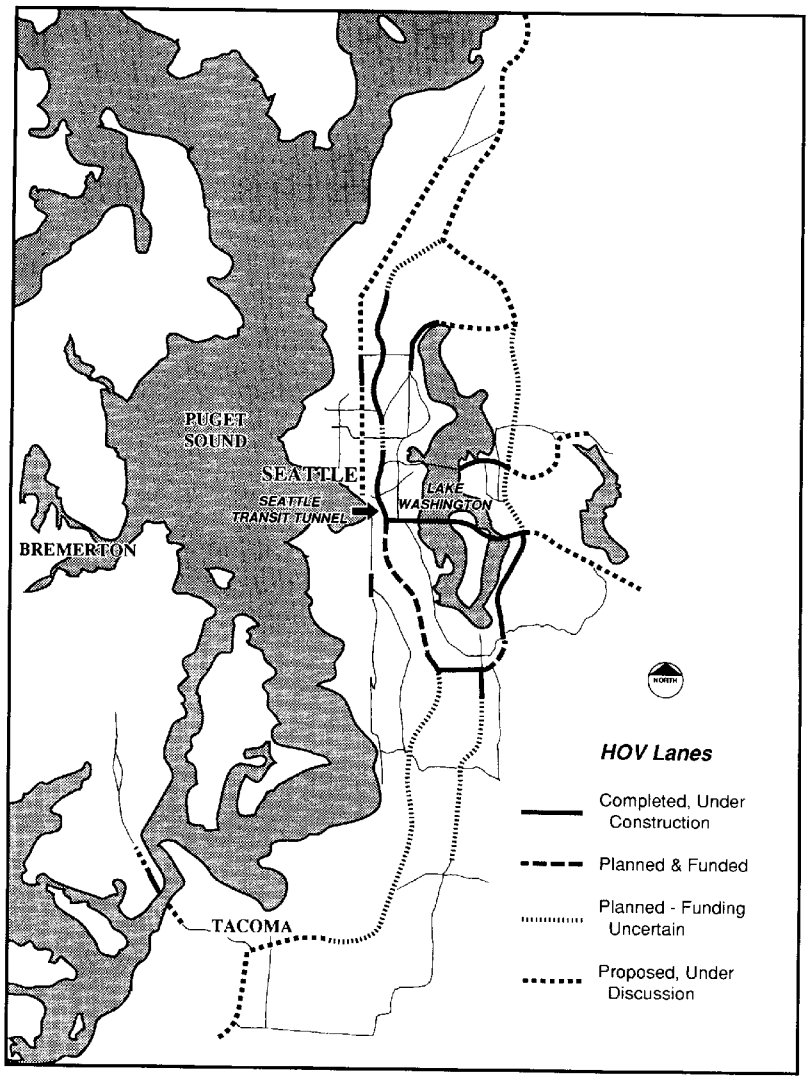


Figure 6-32. Seattle Area HOV System. [35]

Figure 6-33. I-5 Reversible-Flow HOV Lanes. [9]

Figure 6-34. I-5, Seattle Concurrent Flow HOV Lanes (Median Location). [9]

Figure 6-35. I-405, Seattle Concurrent Flow HOV Lanes (Shoulder Location).[10]

6.6 REFERENCES AND SUGGESTED READINGS

REFERENCES

1. Larson, T.D. "High Occupancy Vehicle (HOV) Facilities" Memorandum to Regional Federal Highway Administrators, October 4, 1990.

2. *Public Involvement and Promotion for the Transportation Professional.* A Handbook Developed for the Federal Highway Administration Seminar on *Planning, Operation, and Design of High Occupancy Vehicle Facilities.* Crystal City, VA, 1993.

3. Fuhs, C.A. *Preferential Lane Treatments for High-Occupancy Vehicles*. NCHRP Synthesis of Highway Practice 185, TRB, National Research Council, Washington, DC, 1993.

4. *Guidelines for High Occupancy Vehicle Lanes, A Recommended Practice.* Publication No. RP-017. Institute of Transportation Engineers, Washington, DC, 1986.

5. Fuhs, C.A. *High-Occupancy Vehicle Facilities: A Planning, Design, and Operation Manual.* Parsons Brinckerhoff Quade & Douglas, Inc., 1990.

6. Turnbull, K. F., Henk, R. H., and Christiansen, D.L. *Suggested Procedures for Evaluating the Effectiveness of Freeway HOV Facilities.* UMTZ/TX-91/925. Texas Transportation Institute, Texas A&M University System, College Station, TX, 1990.

7. Speilberg, F., et al., *Evaluation of Freeway High Occupancy Vehicle Lanes and Ramp Metering.* DOT-OST-78-050. U.S. Department of Transportation, Washington, DC, 1980.

8. Walters, C.H., Lomax, T.J., Skowronek, D.A., Middleton, M.D., and Oey, H.S. *Multimodal System Planning Technique*: *An Analytical Approach to Peak Period Operation.* Draft Report FHWA/TX-95/1451-1F. Texas Transportation Institute, Texas A&M University, College Station, TX, 1995.

9. *FHWA/PB HOV Interactive1.0*. FHWA, U.S. Department of Transportation in Association with Parsons Brinckerhoff, 1996.

10. *Guide for the Design of High Occupancy Vehicle Facilities.* American Association of State Highway and Transportation Officials. 1992.

11. *A Short Course on Freeway Design and Operations.* Prepared by the Texas Transportation Institute in cooperation with the Texas Engineering Extension Service. Texas A&M University System. Revised July 1994.

12. *Planning, Operation, and Design of High-Occupancy Vehicle Facilities Seminar Notes*. Prepared for the Federal Highway Administration. Presented in Atlanta, Georgia on August 10-11, 1993.

13. *Manual on Uniform Traffic Control Devices*. FHWA, U.S. Department of Transportation, Washington, DC, 1988.

14. *High Occupancy Vehicle Lanes Enforcement Survey*. Prepared for Metropolitan Transit Authority of Harris County. Texas Transportation Institute, Texas A&M University System, College Station, TX, 1988.

15. *A Policy on Geometric Design of Highways and Streets*. American Association of State Highway and Transportation Officials, Washington, DC, 1994.

16. Turnbull, K.F. *Effective Use of Park-and-Ride Facilities*. Synthesis of Highway Practice 213, National Cooperative Highway Research Program, Transportation Research Board, Washington, DC, 1995.

17. *Guide for the Design of Park-and-Ride Facilities.* American Association of State Highway and Transportation Officials, Washington, DC, 1992.

18. Bowler, C.E., Noel, E.C., Peterson, R., and Christiansen, D. *Park-and-Ride Facilities - Guidelines for Planning, Design, and Operation.* FHWA, U.S. Department of Transportation, Washington, DC, 1986.

19. *A Guide for Transportation Landscape and Environmental Design*. American Association of State Highway and Transportation Officials, Washington, DC, 1991.

20. *Transit Planting: A Manual*. UMTA, U.S. Department of Transportation, Washington, DC, 1986.

21. *Uniform Federal Accessibility Standards.* General Services Administration, TRB, National Research Council, Washington, DC, 1984.

22. *Transportation for People with Disabilities.* U.S. Department of Transportation, 49 CFR 37, Docket 47483, Notice 91-14, Final Rule, Effective October 7, 1991.

23. *Americans with Disabilities Act: Accessibility Guidelines for Buildings and Facilities, Transportation Facilities, and Transportation Vehicles*, U.S. Architectural and Transportation Barriers Compliance Board, Washington, DC, 1994.

24. Fitzpatrick, K., Hall, K., Perkinson, D., Nowlin, R. L., and Koppa, R. *Location and Design of Bus Stops, Final Report*. TCRP Report. Washington, DC, July 1996.

25. Fitzpatrick, K., Hall, K., Perkinson, D., Nowlin, R. L., and Koppa, R. *Guidelines for the Location and Design of Bus Stops*. TCRP Report 19. Washington, DC, November 1996.

26. *The Coordination of Parking with Public Transportation and Ridesharing*. Prepared for the Urban Consortium for Technology Initiatives. Report Number DOT-I-82-29. Public Technology, Inc. Washington, DC, 1982.

27. Hirst, E. "Transportation Energy Conservation Policies." *Science.* American Association of Advancement of Sciences, Washington, DC, April 1976.

28. Turner, S. "Video-Based HOV Enforcement System To Be Tested," *Urban Transportation Monitor*. April 26, 1996, pp 3.

29. "Second Contraflow HOV Lane to Use Moveable Barrier Opens in Boston" *HOV System Notes*. No. 8. TRB, National Research Council, Washington, DC, July 1996.

30. *High Occupancy Vehicle (HOV) Lane Marketing Manual.* Report Number DOT-T-95-04. Prepared by Systan, Inc., Los Altos, CA. for the Federal Highway Administration, September 1994.

31. Miller, C. et al. *Enforcement Requirements for High-Occupancy Vehicle Facilities*. FHWA-RD-79-15. FHWA, U.S. Department of Transportation, Washington, DC, 1978.

32. Memorandum on "Exemption from HOV Lane Restrictions for Inherently Low Emission Clean-Fuel Fleet Vehicles" from Associate Administrator for Safety and System Applications, FHWA, U.S. Department of Transportation, October 23, 1996.

33. The High-Occupancy Vehicle Facility System, Houston, Texas. Prepared by the Texas Transportation Institute, Texas A&M University System, College Station, TX, 1991.

34. Henk, R.H., Morris, D.E., and Christiansen, D.L. *An Evaluation of High-Occupancy Vehicle Lanes in Texas, 1995*. Draft Report FHWA/TX-97/1353-4. Texas Transportation Institute, Texas A&M University System, College Station, TX, 1996.

35. Christiansen, D.L., and Morris, D.E. *An Evaluation of the Houston High-Occupancy Vehicle Lane System*, Report DOT-T-92-15, FHWA, U.S. Department of Transportation, Washington, DC, 1991.

36. Berg, D.B., Jacobson, L.N., and Jacobson, E.L. "Converting a General Purpose Lane to an HOV Lane: The Interstate 90 Project", in *ITE 1994 Compendium of Technical Papers*, Institute of Transportation Engineers, Washington, D.C., 1994, pp 74-78.

SUGGESTED READING

- Fuhs, C. *Preferential Lane Treatments for High-Occupancy Vehicles, Synthesis of Highway Practice 185*. TRB, National Research Council, Washington, DC, 1993.

- *Guide for the Design of High Occupancy Vehicle Facilities.* American Association of State Highway Officials, Washington, DC, 1992.

- Notes from the FHWA *Planning, Operation, and Design of High-Occupancy Vehicle Facilities* Seminar in Atlanta, GE, August 10-11, 1993.

- "Effective Use of Park-and-Ride Facilities." *Synthesis of Highway Practice 213*. National Cooperative Highway Research Program, TRB, Washington, DC, 1995.

- *Park-and-Ride Facilities - Guidelines for Planning, Design and Operation*. Federal Highway Administration, January 1986.

- *Guide for the Design of Park-and-Ride Facilities*. American Association of State Highway and Transportation Officials, 1992.

- *I-394 HOV Lane Case Study - Final Report.* Minnesota Department of Transportation, 1995.

MODULE 7. INFORMATION DISSEMINATION

TABLE OF CONTENTS

7.1 INTRODUCTION 7-3
MODULE OBJECTIVES 7-4
MODULE SCOPE 7-4

7.2 DESIGN PROCESS 7-4
IDENTIFY NEEDS 7-4
IDENTIFY INFORMATION DISSEMINATION PARTNERS 7-6
BUILD CONSENSUS AMONG PARTNERS 7-6
Identify Differences in Operating Philosophies Among Partners 7-6
Establish Common Ground Between Partners 7-7
ESTABLISH GOALS AND OBJECTIVES 7-7
ESTABLISH PERFORMANCE CRITERIA AND MEASURES 7-8
Information Credibility 7-8
Market Penetration 7-8
Traveler Response 7-9
FUNCTIONAL REQUIREMENTS 7-10
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 7-11
IDENTIFY AND SCREEN TECHNOLOGIES 7-11
DEVELOP IMPLEMENTATION PLAN 7-12
IDENTIFY FUNDING SOURCES 7-13
IMPLEMENT 7-14
EVALUATE 7-14

7.3 TECHNIQUES AND TECHNOLOGIES 7-15
ON-ROADWAY INFORMATION TECHNOLOGIES 7-15
Types of DMS 7-16
Light-Reflecting DMSs 7-17
Fold-out. 7-18
Scroll (or Tape). 7-18
Rotating Drum. 7-19
Reflective Disk Matrix. 7-19
Light-Emitting DMSs 7-19
Neon or Blank-Out Signs. 7-19
Lamp (Incandescent Bulb) Matrix Signs. 7-19
Fixed-Grid or Shuttered Matrix Fiber-optic Signs. 7-19
Fixed-Grid or Matrix Light-Emitting Diode (LED) Signs. 7-20
Hybrid DMSs 7-20
Reflective Disks with Fiberoptics/LEDs. 7-20
Static/DMS. 7-20
Advantages and Disadvantages of the DMS Technologies 7-20

IN-VEHICLE INFORMATION 7-21
Auditory In-Vehicle Information Technologies 7-21
Highway Advisory Radio 7-21
Automated HAR (AHAR) 7-24
Cellular Telephone "Hotlines" 7-25
Commercial Radio 7-26
Citizen-Band Radio 7-26
Visual In-vehicle Information Technologies 7-26
Video Display Terminals 7-26
Head-Up Displays 7-27
OFF-ROADWAY INFORMATION 7-27
Television 7-28
Telephones 7-28
Pagers 7-28
Personal Data Assistants (PDAs) 7-29
Computers (Internet) 7-29
Kiosks 7-29

7.4 LESSONS LEARNED 7-30
NTCIP STANDARDS 7-30
INFORMATION MESSAGE DESIGN 7-30
ENVIRONMENTAL ISSUES CONCERNING DMSs 7-31
Sun Position 7-31
Weathering 7-31
"AT-REST" DISPLAY CONDITIONS 7-31

7.5 EXAMPLE OF AN INFORMATION DISSEMINATION SYSTEM: TRANSCOM 7-31
ORGANIZATION 7-31
OPERATIONS 7-32
FUTURE ACTIVITIES 7-33

7.6 REFERENCES 7-35

MODULE 7. INFORMATION DISSEMINATION

Figure 7-1. Dynamic Message Signing on New Jersey Turnpike.

7.1 INTRODUCTION

It is well known that the key to successful driving task performance is efficient information gathering and processing.[1] Likewise, properly communicating with motorists is critical to successful freeway traffic management and operations. Motorists rely on a wide variety of information to properly accomplish the control, guidance, and navigational aspects of the driving task. The roadway alignment and general terrain itself provides a great deal of this information through visual "cues;" sources such as pavement markings and regulatory, warning, and guide signs also contribute greatly to the overall information system. However, in an effective freeway management system, dynamic methods of conveying information to motorists or travelers are often needed to better operate and control the system.

Freeway management systems rely on various information dissemination components to apprise all types of travelers (motorists, transit users, commercial vehicle operators, etc.) of current and anticipated travel conditions so that informed mode, departure time, and route choice decisions can be made.

Information dissemination is also managed in order to improve travel conditions in the

corridor by influencing traveler behavior (by recommending diversion routes around an incident, for example). This information can be disseminated from a variety of sources (State departments of transportation, transit agencies, private-sector information service providers, etc.) using a variety of methods (dynamic message signs, commercial radio traffic reports, traffic information kiosks, etc.).

MODULE OBJECTIVES

The objectives of this module are threefold:

- To describe how to define, establish, and coordinate effective traveler information components in a freeway management system.
- To identify existing and emerging technologies available to facilitate information dissemination to travelers.
- To illustrate how information dissemination components can integrate with each other and with other components of a freeway management system.

MODULE SCOPE

This module addresses both traditional and emerging information dissemination processes and technologies for freeway management systems. Whereas a number of technologies (i.e., dynamic message signs, highway advisory radio) have been around for a number of years, there are new opportunities today. In the past, information was distributed in response to an incident. Today, the focus is to provide a continuous flow of information to travelers, businesses, and commercial carriers in order to make their trip travel time more predictable. The focus of this module is to emphasize the need for integration among all components and technologies being utilized for information dissemination within the integrated regional transportation management system.

7.2 DESIGN PROCESS

Decisions about when, where, and how to disseminate travel-related information to the public have become much more complex in recent years, due to improved traffic/ weather/transit surveillance capabilities and expanded information dissemination options. As with the other components that can be included in a freeway management system, the decisions necessary to develop and integrate information dissemination into the system can be best accomplished by following the basic decision process described in Module 2. Specific application of this process to the task of incorporating information dissemination into a freeway management system is discussed in the following sections.

IDENTIFY NEEDS

The first step in the decision process is to identify the need to be addressed through information dissemination, or stated another way, the information needs that exist in the freeway corridor. Two basic categories of information dissemination exist:

- Pre-trip planning.
- En route guidance and information.

Table 7-1 presents examples of some of the specific types of need and/or information needs in these categories. Certain information may be needed both pre-trip and en route, whereas other information may be needed for either one or the other. To the extent possible, these information needs should be further defined by the following:

Table 7-1. Information Need Categories.

Category	Examples of Information Need
pre-trip planning	• current/anticipated traffic conditions - speeds - incident locations - other congestion locations - upcoming road closures • weather effects - pavement conditions - road closures • route guidance - around incidents - to special events • transit information - bus schedules and status - transfer locations - rideshare matching (preplanning and real-time matching)
en route guidance and information	• current traffic conditions - speeds - incident locations - other congestion locations • weather effects - pavement conditions - road closures • route guidance - around incidents - to special events • lane/shoulder/ramp use status

- Audience.
- Location.
- Time-of-day.

These characteristics affect how well information can be received by the users, and what types of responses can be expected from the users who have that information available.

In addition to the needs to be addressed through information dissemination efforts, it is also important early on to identify those factors or issues that will influence the ability of agencies and/or the private sector to provide that information. Tort liability concerns can be one such factor. Some agencies may avoid providing current pavement condition information, for example, for fear of establishing a precedent that may be used against them if they fail to warn of that pavement condition (or a similar one) in the future. Preestablished agreements between public agencies and private sector companies regarding access to

agency data may also influence the direction of future dissemination efforts.

The problem identification step of the decision process also includes an inventory of existing information sources, including media reports and private sector initiatives. The inventory can include such data items as target audiences, accessibility, frequency of reports, and information accuracy. If possible, assessments of user satisfaction with the information should also be obtained.

IDENTIFY INFORMATION DISSEMINATION PARTNERS

The key partners in the development of the information dissemination component of a freeway management system include the following:

- Traditional State and local public sector agencies (transportation and public works, transit, toll authorities, law enforcement).
- Commercial media.
- Private sector traffic reporting services (distributing through commercial media venues or through direct subscription to motorists).
- Local fleet operators (delivery services, taxis, etc.).

In some cases, the owners/operators of major traffic generators (malls, tourist attractions, annual special event promotions, etc.) could also be important partners to include in the decision-making process.

BUILD CONSENSUS AMONG PARTNERS

Identify Differences in Operating Philosophies Among Partners

Developing a consensus regarding the importance of information dissemination and its role as part of an overall freeway management system is the next step in the decision process. It is important that project partners take ownership of the effort up front, or it may never work. To make this happen, it is important to identify and understand the differences in operational philosophies of the different partners involved.

For example, highway and transit agencies may view the presence of some degree of recurrent congestion on the freeway in very different ways if an HOV lane is included within the freeway right-of-way. Whereas diverting freeway traffic to arterial streets might be a primary goal of the highway agency, the transit agency may desire to promote bus utilization and the HOV lane as a means of reducing that congestion. Partners need to communicate these concerns to each other so that a consensus about goals can be reached.

The partners must also be aware of the differences in operational philosophies between public agencies and private sector entities. Whereas equity concerns typically drive public agency decisions regarding the dissemination of travel-related information, marketing opportunities and profit motives will generally dictate private sector interests, perceptions, and decisions. A clear understanding of the differences among these philosophies is required.

Establish Common Ground Between Partners

Once operational and philosophical differences among the various partners have been identified, the next step is to establish areas of common concern or priorities between the partners with respect to travel-related information dissemination. Common ground must be established both at the upper management and political level, and at the day-to-day operations level. Information dissemination requires upper management and political support of a common vision of information availability in order to ensure continued funding support. Meanwhile, consensus among partners at the operations level is needed to promote true operations integration among the various information dissemination components utilized within the corridor.

ESTABLISH GOALS AND OBJECTIVES

Once a consensus has been established among the partners involved with information dissemination efforts, it is necessary to define the goals and specific objectives that are going to be addressed through this component of the freeway management system. As discussed in Module 2, goals are broad statements of the intent of the system or one of its components, whereas objectives are specific statements about what the system or component of that system will attempt to accomplish. A given goal may have more than one objective specified to reach that goal. Table 7-2 presents examples of goals and objectives an agency might have for the information dissemination component of its freeway management system.

Table 7-2. Examples of Goals and Objectives for Information Dissemination.

Category	Examples
Goals	• Reduce motorist demands upstream of a freeway incident • Reduce motorist errors in locating unfamiliar destinations • Reduce transit user uncertainty about bus arrivals
Objectives	• Warn motorists of adverse weather conditions • Notify motorists of downstream incidents • Advise motorists when to seek alternative routes • Provide motorists with origin-to-destination route guidance assistance • Inform motorists at park-and-ride lots when the next bus will arrive

In general terms, each of the objectives developed for information dissemination should be specific enough to answer the following questions: [1]

- Who is being communicated with?
- What responses are desired or anticipated?
- Where will the responses take place?

ESTABLISH PERFORMANCE CRITERIA AND MEASURES

In order to assess the extent to which information dissemination efforts within a freeway management system are meeting goals and objectives, a set of performance criteria and measures-of-effectiveness pertaining to these efforts must be identified. Relative to information dissemination, performance criteria have three different dimensions that are of interest:

- Information credibility.
- Market penetration.
- Traveler response.

These three dimensions are interrelated. The following sections provide additional details and examples of criteria for each of the above dimensions.

Information Credibility

An information dissemination tool must be credible to travelers if it is to be utilized and have an impact upon traffic operations. The following criteria define how credibility is established: [2]

- The information must be accurate.
- The information must be timely.
- The information must be relevant to its intended audience.

While these are generally accepted concepts, it is sometimes difficult to identify and obtain objective and quantifiable measures with which to judge them. This task is further complicated by the fact that the measures themselves may depend on the specific message or unit of information that a partner is trying to convey. Examples of performance measures that could be used to evaluate the credibility of information being disseminated to motorists are provided in Table 7-3. Local concerns and capabilities will dictate which performance measures are most appropriate for evaluating information credibility in a given locale.

Market Penetration

Market penetration refers to the percentage of the potential audience reached by the information dissemination efforts. Performance criteria regarding market penetration may be appropriate for evaluating certain system goals and technologies, particularly those emerging as part of Advanced Traveler Information Systems (ATIS). It is expected that some technologies, such as in-vehicle dynamic route guidance, will require only limited market penetration in order to achieve operational benefits. Other technologies, such as information kiosks in major traffic generators, may require agencies to strive for as great a market penetration as possible in order to distribute the information to a wider audience and possibly attract private sector advertising and sponsorship.

Table 7-3. Examples of Performance Measures for Information Dissemination.

Category	Examples
Information accuracy	• Difference in the number of incidents in the system and number of incidents reported • Difference between reported expected arrival times of buses and the times the buses actually arrive at a transit station • Number of complaints received from the public about inaccurate information (by device and type of information)
Information timeliness	• Average delay time between when an incident is verified and when information about the incident is disseminated to travelers
Information relevance	• Number of travelers who access a given information component or unit

Traveler Response

Ultimately, the purpose of providing information to travelers is to effect some-change in traveler behavior that will cause an improvement in safety or operations. Thus, performance measures are also needed to determine the extent to which information dissemination accomplishes this purpose. Changes in traveler mode, departure time, and route (if appropriate) are appropriate for evaluating the effectiveness of real-time travel-related information. However, it may be very difficult and expensive to obtain actual data for these measures. Traveler opinions about the effectiveness of the information being provided can be another important eval-uation measure. Reductions in travel time, turning and route choice errors, or similar measures may also be useful to evaluate certain types of information. [3]

It is important to recognize that because of the complex travel patterns of travelers at any point in the roadway, it may not be possible to adequately measure the overall effects of many types of information or dissemination modes upon traffic volumes, speeds, or delays. The day-to-day variances in travel patterns themselves may mask the effects of any information disseminated during a specific event such as an incident, particularly if the information is intended for a very specific audience (such as vehicles within a freeway traffic stream destined for a specific downstream exit).

Consequently, it may sometimes be necessary to include performance measures that evaluate the effect of information dissemination at an individual traveler's level. For example, trip diaries that identify a specific driver's travels on any given day may need to be compared to determine whether the presence of information had any influence on driving behavior for that select group with drivers on that particular day. This approach can be very costly and manpower intensive, however.

FUNCTIONAL REQUIREMENTS

The functional requirements of information dissemination components in a freeway management system define specific actions or activities that are to be performed in order to achieve one or more of the objectives. Initially, the functions should be defined without considering the dissemination technology or system architecture that will be employed. Functions simply specify what information will be disseminated and possibly when and where it is to be presented, not how this will be done. Table 7-4 presents some examples of functional requirements for various information dissemination objectives.

The ITS National Architecture should serve as the basic building block of the functional requirements definition process for information dissemination.[4] The functions described in the National Architecture must then be detailed to match the needs and desires of the local agencies.

Details regarding who should receive information, as well as when and where that should occur, all become a part of the functional definitions. The intent of this step in the design process is not only to specify functions independent of the technology that could be used to achieve those functions, but also to highlight what and why other components of the freeway management system must link with the information dissemination components.

Table 7-4. Examples of Functional Requirements for Information Dissemination Objectives.

Examples of Objectives	Examples of Possible Functions
• Warn motorists of adverse weather conditions	• Notify freeway motorists of downstream roadway flooding whenever more than 100 mm of water collects in an underpass section • Notify motorists approaching an entrance ramp whenever the freeway is closed because of adverse weather conditions such as high winds or ice
• Notify motorists of downstream incidents	• Within “X” minutes of its occurrence, notify freeway motorists of any incident that occurs in the travel lanes within the next exit or decision point
• Advise motorists when to seek alternative routes	• If the road is closed due to an accident, recommend that freeway motorists traveling to the CBD exit at an upstream connector to a parallel toll road whenever travel times to the CBD on the freeway are more than “X” minutes longer than on the toll road

For example, an information dissemination function that gives route specific travel times requires that travel time monitoring be a function in the freeway management system. Likewise, a function of notifying motorists about downstream lane-blocking incidents within a set time also affects the design of the freeway surveillance component of the freeway management system (what type of vehicle detection technology can be used, how close together detectors must be placed, etc.). As a final example, notifying motorists that lanes are closed downstream for maintenance work implies that a mechanism for determining when and where such closures occur is available via direct communication between the maintenance and operations divisions of the transportation agency, observation of the closure via closed-circuit television, or other methods.

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

The purpose of defining functional relationships, data requirements, and information flows is to establish an understanding of how the various information dissemination functions that are to be accomplished will be integrated with each other and with the other components in the freeway management system. In this step, the relationships between information dissemination functions are further refined to incorporate local issues, concerns, and capabilities.

The relationships between the desired information dissemination functions and necessary data flows to those functions can be particularly important when designing a freeway management system, especially if several partners are involved in the information dissemination process. Locations of potential information dissemination conflicts can be identified beforehand (a highway and a transit agency may want to display information at the same roadway location, for example). Perhaps more commonly, locations and situations where information sharing between partners can occur may be identified, and result in a more efficient system design.

IDENTIFY AND SCREEN TECHNOLOGIES

Once the system requirements for information dissemination have been developed, it is then appropriate to assess the actual technologies available to meet the functional and system architectural requirements that have been developed for meeting the goals and objectives of information dissemination. Technologies that are available for information dissemination can be grouped into one of three categories:

- Those located on the roadway where the information transfer to the traveler occurs at a specific point or within a very small segment of roadway (i.e., using dynamic message signs or highway advisory radios.)

- Those located within the vehicle where the information transfer is not constrained to a point or a small segment of roadway (e.g., using radio, cellular telephone, or in-vehicle navigation devices).

- Those located off of or away from the roadway altogether (e.g., using television, computer linkages, or kiosks in major traffic generators).

A review of the technologies in each of these three categories is provided later in this module. This review can serve as the

starting point for the technology screening process. However, technology in the information dissemination realm continues to change rapidly and should be reassessed each time this point is reached in the decision process.

Initial screening can be used to determine which of the three basic categories can be used to effectively accomplish each of the specified information dissemination functions. More than one of the categories may be useable. Then, within each category, a more detailed screening and evaluation process can help select among the various alternatives of that technology. Screening should include such considerations as the following:[5]

- Legibility (for visual information components).
- Reliability.
- Costs to the user and public agency (these should be life-cycle costs, discussed in greater detail in Module 11).
- Potential market size.

If, for some reason, the screening process does not identify technologies that can adequately perform the functions that were established, then the analyst must step backwards and reassess the objectives and/or functions of the information dissemination component of the freeway management system. Objectives and associated functions may have to be scaled back or otherwise modified.

DEVELOP IMPLEMENTATION PLAN

The next step in the process is the development of an implementation plan. Implementation plan development is governed by federal guidelines, and is required for all new or expanded traffic control system projects that utilize federal funds.[6] Actually, an implementation plan is a good idea for all traffic control system projects, regardless of whether or not they utilize federal funds.

Module 2 described the general guidelines regarding overall plan development for a freeway management system. In this section, special issues of plan development are discussed that relate specifically to the information dissemination components of a freeway management system.

Because information dissemination must be coordinated among the various partners, a portion of the plan should be devoted to a discussion of the management structure and the agreements that will be utilized to achieve this coordination. Items to be included in this portion of the plan are as follows:

- Names of contact person(s) for each partner.
- Protocols and methods to be utilized to coordinate.
- Definition of each partner's responsibilities regarding information exchange.
- Estimates of each partner's financial contribution to the effort (cash, in-kind exchange of equipment or services, etc.).
- Letters or memoranda of agreement regarding the desire to coordinate among partners.

Additional details may need to be included in the letter of agreement if some of the partners will jointly operate some of the

information dissemination technologies to be implemented. For example, a transportation agency may wish to allow law enforcement personnel to access and utilize their dynamic message signs or highway advisory radio equipment for managing major incidents during late-night hours when the transportation agency does not have someone on duty.

A detailed implementation plan is also needed in order to properly coordinate information dissemination within the freeway management system. Whereas the letters of agreement and discussion of the management structure lay out broad administrative boundaries and consensus to coordinate, the implementation plan spells out in detail how this will occur.

An implementation plan may include specific preplanned actions that will be taken in response to certain common or expected situations (i.e., which dynamic message signs will be activated and what will be displayed on them in response to an incident that occurs on a specific section of freeway), as well as more general rules that will be used to decide what information to present for unusual, unexpected situations. For example, the TransGuide Traffic Management System in San Antonio, Texas has over 60,000 preplanned "scenarios" coded into a computerized database. Depending on the time of day, location of an incident, and the number of lanes that are blocked, a specific scenario is selected that defines which dynamic message signs and lane control signals are to be activated and the messages that are to be displayed on each (i.e., "LEFT LANE CLOSED AHEAD"). Most of the time, system operators simply activate one of these preplanned actions. Occasionally, a freeway problem develops for which preplanned actions have not been developed. In these cases, system operators use basic information dissemination rules for the system they have learned to decide what messages need to be displayed, on which signs they should be displayed, etc.

IDENTIFY FUNDING SOURCES

Identification of funding source alternatives actually occurs as part of the plan development process. As with the other components of freeway management systems, the partners involved in information dissemination can share costs through some combination of the following:

- Individual agency hardware purchases and operation/maintenance which is coordinated with another agency's efforts in the region.
- In-kind contribution of equipment and services to a pooled service.
- Establishment of user fee schedules (direct or indirect).

Several different cooperative institutional arrangements have been successfully utilized to distribute information dissemination costs in a rational and equitable manner to public and private partners. For the TRANSCOM organization in New York/New Jersey, cash and in-kind contributions from each member agency fund the operations of the organization.[9] Each partner has an official assigned to the TRANSCOM executive committee, which works together to establish the actual fee structure, management policies, and other matters of concern to the agencies as a group.

In Boston, the SmarTraveler cellular call-in system is sponsored in large part by FHWA and the Massachusetts Highway Department, but is managed by a private-sector limited partnership, SmartRoute.[10] Several other private-sector companies (cellular

telephone companies, a paging company, and broadcast media) purchase some of the information being managed and disseminated, which has helped offset some of the costs to the public agencies. [11]

In San Antonio, Texas, both public and private-sector partners have entered into a Cooperative Industry Product Agreement to facilitate the development and implementation of the software that will decode the video and data transmissions from a low-power television station that the Texas Department of Transportation has purchased and will operate. [12] This has allowed the various potential beneficiaries of the system (i.e., media stations, fleet operators, etc.) to contribute to software development for a relatively small price (and risk). Furthermore, it provides a mechanism to allow future partners to enter into the agreement in a fair and equitable manner. [12]

IMPLEMENT

Implementation of information dissemination components in the freeway management system occurs as a natural result of following the systems engineering process. Module 2 described, in general terms, the implementation activities that should occur at this point in the process. In addition to the issues brought out in that module, other concerns that pertain directly to information dissemination include:

- Implementation schedule.
- User training.

Information dissemination is the main link between the freeway management system and the motoring public. Experience indicates that the implementation of information dissemination devices should occur only after the surveillance and control infrastructure has been installed. Otherwise, the credibility of the overall system with the traveler can be severely degraded by having the devices in place and not using them or by having the devices display wrong information. One of the problems reported by officials of the INFORM project in New York was that the DMSs for the project were installed early on in the contract and then sat there unused for several years as the agencies built the rest of the infrastructure. Since it was not immediately apparent to motorists why the devices could not be used to communicate about downstream roadway and traffic conditions, public opinion about the project suffered.

Another important facet of information dissemination implementation is the development of training necessary to assist travelers in correctly interpreting and responding to some of the information dissemination technologies. For example, the Texas Department of Transportation in Fort Worth displays the meaning of lane control signals mounted over each travel lane periodically throughout the freeway network. As a result, motorist comprehension of the available symbols for the signals are somewhat higher in Fort Worth than in other locations in Texas. [13] Other possible training alternatives include public service announcements or mailing inserts in local utility bills.

EVALUATE

The final step in the design process is to evaluate the effectiveness of the information dissemination component of the freeway management system. However, this evaluation should not be considered a one-time activity, but should be part of a periodic review of the effectiveness of the information dissemination components and of the overall system. As discussed previously, it is rather difficult to measure the impact of information dissemination

upon overall traffic measures, because of the stochastic nature of travel demands and behavior in a given freeway corridor, as well as the events which cause disruptions to these demands (incidents, special events, etc.).

Care should be taken not to overestimate the benefits achieved by the implementation of information dissemination components in a freeway management system. Specifically, it is important to recognize that travel patterns in a freeway corridor are quite dynamic, and that some drivers will divert naturally when they encounter freeway congestion regardless of whether or not they receive information beforehand about that congestion. Therefore, information dissemination should be considered as having an *incremental* effect upon traffic conditions by modifying where and how some travelers respond to congestion (where travel routes are changed, how many motorists change departure times, how many changed modes, etc.).

7.3 TECHNIQUES AND TECHNOLOGIES

Information dissemination components of a freeway management system can range from a single device owned and operated by one agency, to an integrated collection of devices and mechanisms under the control of several agencies and several private sector entities. In this handbook, a basic distinction is made between kinds of information depending on which of three main locations it comes from:

- On-roadway information.
- In-vehicle information.
- Off-roadway information (typically at origin of a trip).

Specific technologies available in each of these categories are discussed below.

ON-ROADWAY INFORMATION TECHNOLOGIES

One of the most fundamental technologies available for disseminating traffic-related information from the roadside is that of dynamic message signs (DMS). DMSs are sometimes referred to as changeable message signs (CMS) or variable message signs (VMS). DMSs allow operating agencies to visually disseminate travel information to motorists on a near real-time basis. DMSs use words, numbers, or symbols to convey information. They are extremely flexible and powerful traffic management tools in freeway operations.

DMSs can be either portable or fixed, and can be operated either on a fixed-time basis with on-site control or interconnected with a traffic control center to provide remote control.[14] DMSs can be used to perform the following functions:

- Inform motorists of varying traffic, roadway, and environmental conditions.
- Provide specific information relative to the location and delays associated with incidents.
- Advise motorists on detour routes because of construction or roadway closure.
- Suggest alternate routes to avoid freeway congestion.
- Reassure drivers on unfamiliar alternate routes.
- Redirect diverted drivers back to freeways.

Table 7-5 lists the applications for which DMSs can be used.[14]

Types of DMS

DMSs can be conveniently classified into three categories, namely:

- Light-reflecting.
- Light-emitting.
- Hybrid.

Light-reflecting signs reflect light from some external source such as the sun, headlights, or overhead lighting. Figure 7-2 shows examples of different types of light-reflecting DMSs. In comparison, light-emitting DMSs generate their own light on or behind the viewing surface. Examples of different types of light-emitting DMSs are shown in figure 7-3.

Some manufacturers have combined two DMS technologies (e.g., reflective disk and fiber-optic) to produce hybrid displays that

Table 7-5. Applications of DMSs. [14]

Category	Applications
Traffic management and diversion	• Freeway traffic advisory and incident • Freeway-to-freeway diversion • Special events • Adverse road and weather conditions • Speed advisory
Warning of adverse conditions	• Adverse weather and environmental conditions (fog, smog, snow, rain, dust, wind, etc.) • Adverse road conditions (ice, snow, slippery pavement, high water, etc.) • Low bridge clearance
Control at crossings	• Bridge control • Tunnel control • Mountain pass control • Weigh station control • Toll station control
Control during construction and maintenance operations	• Advisory of upcoming construction/maintenance • Speed advisory • Path control
Special-use lane and roadway control	• Reversible lanes • Exclusive lanes • Contraflow lanes • Restricted roadways • Temporary freeway shoulder use control

Figure 7-2. Examples of Light-Reflecting DMS Technologies.

exhibit the qualities of both. Some agencies have also combined DMSs with static displays to form what can also be considered to be hybrid displays.

Several references are available that describe the operational and performance characteristics of various DMSs (see references 1, 14, 15, 16) The following paragraphs, taken from those references, briefly describe the different types of DMSs in each category.

Light-Reflecting DMSs

The basic categories of light-reflecting DMSs are as follows:

- Fold-out.
- Scroll.
- Rotating drum.
- Reflective disk matrix.

Basic operations and characteristics of each are described below.

Figure 7-3. Examples of Light-Emitting DMS Technologies.

Fold-out. A fold-out sign is a conventional highway sign with a hinged viewing face. This type of sign can either display two messages (one with the hinged face closed, one with it open) or show a message only when the hinged face is open (no message is displayed when the hinged face is closed). Typically, signing materials that conform directly to the *Manual on Uniform Traffic Control Devices* (MUTCD) are used to make each message. For freeway applications, they are most often used to indicate icy bridge or other hazardous roadway conditions, or to indicate whether truck weigh stations are open or closed.

Scroll (or Tape). A scroll sign uses a tape or film upon which desired messages are imprinted. The tape is then rolled one way or the other (or up or down) to place a desired message within a display window. As with fold-out signs, scroll signs can use colors and other message characteristics that conform with the MUTCD. Typically, up to 8 to 12 messages can be stored on the tape. The inclusion of more than 12 messages can

create long change-times if one or more of the messages is at the opposite end of the tape.

Rotating Drum. Rotating drum signs are made up of one to four multifaced drums, each containing two to six messages. Each face of each drum portrays one line of a fixed message, and pivots about its axis in order to display the appropriate face for a given message. Colors and lettering characteristics that conform to the MUTCD can also be used on rotating drum signs.

Reflective Disk Matrix. These types of DMSs were very popular for freeway management purposes in the 1970s due to their low energy requirements relative to light-emitting DMS technologies. The viewing face of a reflective disk matrix DMS is comprised of an array of permanently magnetized, pivoted indicators that are flat matte black on one side and reflective yellow or a similar color on the other. The indicators may be square, rectangular, or circular in shape. An electrical current activated when a given pixel is to be turned flips the indicator from a black matte finish to the reflective side. The sign itself can exist of a series of arrays of these indicators (e.g., 5 pixels by 7 pixels, or similar grid for each letter position), or a continuous grid or matrix over the entire sign face. The latter provides more flexibility to operators, but at a higher purchase price.

Light-Emitting DMSs

The following types of light-emitting DMS technologies are described in this section:

- Neon or blank-out.
- Lamp (incandescent mulb) matrix.
- Fixed-grid or shuttered matrix fiber-optic.
- Fixed-grid or matrix light-emitting diode (LED).

Neon or Blank-Out Signs. Neon signs use neon tubing to form characters and messages that are to be displayed. Two basic sign designs are possible:

- Separate each message on the sign face.
- Stack the neon tubing for each message one over another.

The stacked design has reported drawbacks, in that emitted light will be diffused as it passes through the overlayed neon tubing, reducing its legibility. Conversely, the separate message design will require a fairly large sign face in order to display even a moderate number of messages.

Lamp (Incandescent Bulb) Matrix Signs. A lamp matrix DMS relies on an array of incandescent bulbs affixed to a dark background to create the characters needed for a given message. The DMS can be a continuous array of bulbs, or a series of array modules, each of which can display a single character in the overall message. This design provides considerable message flexibility, limited only by the size of the sign (typically 4 lines of text, each comprised of 12 to 20 characters).

Fixed-Grid or Shuttered Matrix Fiber-optic Signs. Fiberoptic DMSs funnel light energy from a light source through fiber bundles to the sign face. For a fixed-grid fiberoptic sign, the ends of these bundles (pixels) are positioned on the sign face to create a given message. Sets of bundles and a separate light source are then used for each message portrayed on the sign. Conversely,

the shuttered matrix fiberoptic DMS positions the ends of the fiber bundles on the sign face in an array similar to that used for other matrix signs. The light sources for all of the fiber bundles remain on constantly, and shutters at the ends of the bundles open and close to create the characters needed to display each message. As with the lamp matrix DMS, the shuttered fiberoptic signs can be either a continuous array of pixels or a series of array modules, each of which portrays a single character. Because of its flexibility, this tends to be the more common type of fiberoptic DMS utilized for freeway traffic management purposes.

Fixed-Grid or Matrix Light-Emitting Diode (LED) Signs. Light-emitting diodes (LEDs) are semiconductors that glow when voltage is applied. Recent advances in LED technology have allowed DMS manufacturers to utilize this technology to construct both fixed-grid and matrix style DMSs. Typically, several individual LEDs are "clustered" together in order to create each pixel. Early versions of the LED DMSs experienced difficulties in maintaining acceptable performance levels over time. However, LED manufacturers have strived to alleviate many of these types of problems in the next generation of DMSs that have been developed. [16]

Hybrid DMSs

Two of the most common types of hybrid DMSs are the following:

- Reflective disk with fiberoptics/LEDs.
- Static sign/DMS combination.

Reflective Disks with Fiberoptics/LEDs. To combat the times of the day and environmental conditions when reflective disk DMSs are not very visible, some manufacturers have combined fiberoptic or LED technologies into these types of signs to give them the visual boost they need. During the times of the day that the sun shines upon the sign and provides adequate visibility, the light sources are turned off by the sign. In addition to providing the additional visual boost during nighttime conditions, this technology also eliminates the need for external lighting on the sign. This eliminates the potential for glare off of the sign face as well. Some agencies are retrofitting reflective disk DMSs with hybrid reflective disk/fiberoptic technology.[15]

Static/DMS. The other major type of hybrid DMS currently in use employs components of both static signing and one of the DMS technologies into a single sign. This can reduce costs in situations when part of a message will be constant and another part of the message will be changed.[1] Also, the static components can be constructed to conform to MUTCD principles. However, this approach does limit the use of the sign to the specific situations addressed by the static component of the message.

Advantages and Disadvantages of the DMS Technologies

The selection of appropriate DMS technology is a complex task requiring analysis of trade-offs between display capability to fulfill a specific need and display cost (including operating and maintenance considerations). Further complicating the selection process is the large number of signing techniques available, each possessing quite different design and operating features. Tables 7-6 through 7-8 present a summary of the advantages and disadvantages that have been reported for each type of DMS technology (see references 1, 14, 15, 16).

Table 7-6. Advantages and Disadvantages of Light-Reflecting DMSs.[Adapted from 1,2,14-16]

Type	Advantages	Disadvantages
Foldout	Simple operation Can conform to MUTCD regulatory and warning signing principles	Limited capacity (1 or 2 messages) Higher potential for environmental and mechanical failures (i.e., panels jamming)
Scroll	Simple operation Can conform to MUTCD regulatory and warning signing principles	Limited capacity (8 to 12 messages) Time to change messages may be significant
Rotating Drum	Simple operation Can conform to MUTCD regulatory and warning signing principles	Limited character size Limited message capabilities (but more than fold-out or scroll)
Disk Matrix	Total message display flexibility Wider angle of legibility than fiberoptic or LED DMSs (see Table 7-7) Low power consumption relative to light-emitting DMSs	Visibility typically lower than similar size light-emitting matrix DMSs Disks sometimes stick or fail prematurely due to excessive dirt or moisture Illumination is required at night, sometimes causing glare or blurring problems

Too often agencies will purchase DMSs before signing objectives and messages are determined. Often, this causes disappointment in the DMSs when these agencies cannot display the desired messages, or when the signs provide lower than expected target value and legibility for the environmental conditions present at the site.[2] Conversely, agencies may end up purchasing a more expensive DMS with capabilities that exceed their actual needs.

IN-VEHICLE INFORMATION

Another realm of information system technologies available are those located within the vehicle itself. In-vehicle information can be disseminated to the motorist by audio or visual means. The following sections describe the basic types of in-vehicle information technologies.

Auditory In-Vehicle Information Technologies

Highway Advisory Radio

Although not as widely used as dynamic message signs, Highway Advisory Radio (HAR) is another means of providing highway users with information in their vehicles. Traditionally, information is relayed to highway users through the AM radio receiver in their vehicles. Upstream of the

Table 7-7. Advantages and Disadvantages of Light-Emitting DMSs.[Adapted from 1,2,14-16]

Type	Advantages	Disadvantages
Neon (blank out)	Simple operation	Limited message capacity Requires a fairly large sign for even moderate number of messages Current designs to not allow for nighttime dimming
Lamp (Incandescent) Matrix	Simple, proven operation High target value Message flexibility	Continuous energy supply required to display message High operation and maintenance costs (energy, bulb replacement)
Fixed Grid Fiberoptic	Low power usage Can display symbols and messages	Narrow cone of vision Limited message capacity
Shuttered Matrix Fiberoptic	Simple operation Message flexibility Low failure rates reported	Narrow cone of vision Mechanical component (shutters) increases potential maintenance costs
Fixed Grid LED	Low power usage Solid state devices (no bulbs) Reported reliability of LED lamps is very high	Narrow cone of vision Limited message capability Intensity adversely affected by high temperatures Long-term performance not known Super-bright LED lamps must be used for adequate daytime visibility
Matrix LED	Low power usage Solid state devices (no bulbs) Reported reliability of LED lamps is very high	Narrow cone of vision Intensity adversely affected by high temperatures Long-term performance not known Super-bright LED lamps must be used for adequate daytime visibility

Table 7-8. Advantages and Disadvantages of Hybrid DMSs.[Adapted from 1,2 14-16]

Type	Advantages	Disadvantages
Disk Matrix with Fiberoptics or LEDs	Some energy conservation (light-emitting technology is turned off in daytime conditions) Eliminates need for nighttime illumination of sign	Use of LED technology limited to date (long-term performance not known) More complex operation (maintenance may be more costly)
Combined Static/DMS	Portions of sign can conform to MUTCD regulatory and warning signing principles Some cost savings is possible relative to full-matrix DMSs	Static portions of sign limits message flexibility Light-emitting DMS technology may wash out static portions of sign if too bright

HAR signal, users are instructed to tune their vehicle radios to a specific frequency via roadside or overhead signs. Usually, the information is relayed to the users by a pre-recorded message, although live messages can also be broadcast.

Message transmissions can be controlled either on-site or from a remote location through telephone or radio interconnects. Most HAR systems operate at the 530 or 1610 kHz frequency level; however, any available frequency can be used as long as a low enough power level is used. A license from the Federal Communication Commission is required to operate a HAR system at high power levels (10 watts or greater).

HAR is used in many different applications. With respect to freeway management, uses of HAR are similar to those of DMSs, and include the following:

- Warning of roadway incidents or congestion.
- Warning of adverse environmental conditions (fog, ice, etc.).
- Notification of highway construction or maintenance.
- Alternate route information.
- Advisories within and regarding transportation terminals (airports, special events, and other major traffic generators, etc.).

One of the major advantages of using HAR is that information is received through a different sensory channel (audio), which reduces information overload received visually while driving.[17] Also, longer and more complex messages can be provided to motorists than can be accomplished using DMSs. However, traditional HAR systems have their drawbacks. They require motorist action (i.e., tuning radio to appropriate frequency) in order to receive information. Also, if a motorist misses one part of an HAR message, he or she must listen to the entire message again in order to obtain the missed part.

There are two types of HAR systems: vertical “whip” antenna systems and induction cable antenna systems. Vertical “whip” antenna systems use individual

antennas or a series of antennas electronically connected together to transmit information. The signal radiates from the antenna in all directions providing a circular area of transmission. Vertical antenna systems have the following advantages:

- They are small.
- They are easy to install.
- They can be placed within several hundred feet of the roadway.
- They are less costly to purchase and install than induction cable systems.

Vertical antenna systems also have the following disadvantages:

- They are subject to damage by weather, accidents, and vandalism.
- They often require special equipment to ensure that the signal is stable, reliable, and easily tuneable.
- Motorists can lose the HAR signal when traveling through tunnels, canyons, etc.

Induction cable antenna systems use a cable installed either under the pavement or adjacent to the roadway. This type of antenna design produces a strong but highly localized signal within a short lateral distance 30 to 45 m (100 to 150 ft) from the cable. Induction cable systems have the following advantages:

- The signal is strong enough to provide full coverage of a multilane facility without causing interference to other HAR systems.
- Messages can be individualized by direction of travel.
- Interference with other radio systems in the area is minimized.

Disadvantages of induction cable HAR systems include the following:

- Induction cable systems are more expensive to purchase and maintain since the cable must extend the full length of the desired coverage area.
- Once installed, induction cable HAR systems cannot be transported from one location to another.

Automated HAR (AHAR)

One of the disadvantages of HAR systems is that they require advance signing to notify motorists of the availability of information, and an action by the motorist to activate and receive that information. Automated highway advisory radio (AHAR) systems have been proposed that require no actions by the motorist to obtain radio information. These systems emit a special electronic notification signal upstream of an HAR message, which interrupts the specially-designed radio/cassette/compact disc player to broadcast the message to the motorist.

FHWA sponsored the prototype development and pilot testing of an AHAR system in the early 1980s. [18,19] However, the difficulties associated with producing and marketing a receiver within a price range acceptable to motorists hindered further AHAR implementation. However, recent emphasis on ITS has sparked renewed interest in AHAR systems. Michigan is experimenting with alternative AHAR designs at this time. [20] Meanwhile, Europe has been experimenting with a version of these types of systems (termed the Radio Data System [RDS] Traffic Management Channel [TMC]) for the past few years. This system relies on a silent data channel

broadcast via FM from existing radio stations. With an appropriate radio receiver, a motorist is then able to receive information from the closest or strongest FM radio signal at that location. [15]

Cellular Telephone "Hotlines"

An in-vehicle communication technology that has seen dramatic growth in the past few years is cellular telephones, which gives the motorist the ability to call special "hotline" systems for traffic information from within their vehicle. Originally, these systems allowed motorists and transit users to call for information to assist in pre-trip decisions from their homes. Information can now be accessed en route via cellular telephone, and decisions can be made whether to alter travel routes. The creation of call-in systems has been a popular traffic impact mitigation strategy for many major urban freeway reconstruction projects in recent years. [21]

One of the more successful examples of this technology is the two SmarTraveler systems now in place in Boston and Cincinnati. [10] Both systems employ special telephone numbers to keep the service free to cellular telephone users. Touch-tone menus allow callers to receive route-specific traffic information (delays, construction activity, recommended alternative routes), bus and train scheduling, and special event transportation information.

This type of in-vehicle communication has the advantage over HAR of giving the motorist some control over the type and amount of information he/she wants to obtain through the touch-tone menus. In addition, it is also possible to generate two-way communication between the motorist and the information source.

Many metropolitan areas have established cellular "hotlines" for motorists to call in and report accident information to the highway agency. Examples include #77 and *SP. The public agency must negotiate with the cellular companies to provide the toll free calls.

Recommendations for establishing cellular telephone "hot-line" systems include the following:

- The call must be toll-free to users.
- The telephone number must be easy to remember and dial.
- The information must be concise.
- If a menu system is used, a long and tedious menu selection process should be avoided.
- A sufficient number of telephone lines should be provided to prevent the majority of users from receiving a busy signal.
- If a system is going to be used to gather information from users, there must be a method of ensuring the accuracy of the incoming information.
- "Official" use of tipster information should include procedures for verifying that information.
- If incident information is to be received, a human operator is recommended so that secondary questions can be asked to clarify confusing or unclear reports.

As with HAR systems, this technology also requires action by the motorist to access information. There are also significant operating costs associated with this technology, as any calls made using cellular

telephones must be paid for by either the motorist, or a public agency, or else absorbed by the corporation providing cellular telephone communication capabilities in the region. Finally, there is some concern that cellular telephone usage while driving may degrade motorist attention and operating capabilities. Manufacturers have developed "hands-free" telephones that allow motorists to listen and talk without holding the telephone receiver.

Commercial Radio

The public has learned to depend upon the media to provide them with "almost" real-time traffic information. Commercial radio has proven to be a good means of providing travelers with traffic information both in and out of their vehicles. Traffic and roadway condition reports have become standard programming items on many commercial radio stations. Commercial radio has the best potential of reaching the greatest number of commuters, since most of them have radios in the vehicles they drive to and from work.

The primary disadvantage of using commercial radio relates to the accuracy of the information. Traffic reports often are transmitted only when normal scheduling permits. This may cause considerable time delays between when an incident occurs and when it is reported by the media. Often, many incidents go unreported or are cleared by the time they are reported on the radio and television. The accuracy of the information provided by commercial radio is a function of the time between the broadcaster's last communication with the incident reporting source and the number of incidents that have occurred and/or have been cleared during that time.

Some transportation agencies have made substantial efforts to improve coordination and cooperation between themselves and the media traffic reporters. For example, some agencies allow private traffic reporting agencies to place personnel in the traffic management center to obtain information on traffic conditions and expected agency responses in an accurate and timely manner.

Citizen-Band Radio

Even though it was once considered an excellent means of providing motorists with two-way communications from their vehicle, Citizen-Band (CB) radio has declined in popularity in recent years. However, there are still a significant number of vehicles, particularly commercial vehicles and trucks, equipped with CB radios. In the past, CB radios have been used primarily in motorist-aid systems. A disabled or passing traveler broadcasts a request for assistance on channel 9. The channel is monitored 24 hours a day, 7 days a week by a police or volunteer organization, which dispatches aid to the stranded traveler. The primary advantage of a CB radio system is that it permits two-way communication between the traveler and the response agency. The effective range of many CB radios is approximately 32 kms (20 mi), depending upon geographic conditions.

Visual In-vehicle Information Technologies

Video Display Terminals

One of the newer technologies for communicating with motorists in their vehicles is through a video display terminal (VDT) mounted in the dashboard. This is primarily a private sector industry, which has not been used widely for information distribution. These systems can be used to provide motorists with route guidance and navigational information in one of two different formats. One approach is to

present the driver navigation and route guidance information in the form of maps or equivalent displays. With these systems, a global picture of the traffic network can be provided. Recommended routes can be highlighted on the video map display as well. In another approach, simple symbolic signals (e.g., arrows, text instructions, or a combination of both) guide the driver along a recommended route. Some prototype systems use a variety of displays depending upon whether or not the vehicle is in motion, the functions selected, and level of informational and navigational displays available.

In-vehicle VDTs offer a number of advantages over available technologies in providing information to motorists while driving. These include the following:

- Travel information is more readily accessible to the driver (providing continuous access to current position, routing, and navigational information).

- Computer-generated navigational maps and displays are logical extensions of traditional forms of providing drivers with route guidance and navigation information.

Information can be displayed in text, graphics, or both and tailored to the needs and desires of each motorist.

There are also limitations to in-vehicle VDTs. These include the following:

- Drivers have to take their eyes off the roadway in order to receive the information.

- In-vehicle VDTs present the driver with complex maps and diagrams that may create a potential to overload the driver with too much information.

- VDTs may also add to the visual clutter already inside the vehicle.

Head-Up Displays

As technology continues to improve, the Head-Up Display (HUD) has become another alternative to in-vehicle VDTs for presenting visual navigational and route guidance information to motorists. Although originally developed for the aviation industry, several automobile manufacturers are beginning to develop HUDs for presenting vehicle status and navigational information to drivers.

A wide variety of options for displaying information may be available using HUDs. Through both icons and alpha-numeric text, navigation and route guidance information may be projected directly into the driver's field of view. This is expected to reduce the need for visual scanning between two information sources (the inside instrument panel and the outside environment), and the associated visual accommodation time.

However, as might be expected, numerous concerns exist regarding the safety and applicability of using HUDs in the driving environment.[23] Currently, most HUDs under development and implementation provide drivers with only relatively simple information, such as speed indications.

OFF-ROADWAY INFORMATION

The final type of information systems discussed in this module are those that drivers access off-roadway, at their homes, offices, shopping malls, etc. A major thrust of ITS development and implementation focuses on putting more information into the hands of motorists and passengers so as to improve departure time, mode choice, and route choice decisions. This emphasis has resulted in increased utilization of a number

of traditional information technologies, and has yielded a number of new off-roadway information sources and several new dissemination technologies. Off-roadway information dissemination technologies include the following:

- Television.
- Telephones.
- Pagers.
- Personal Data Assistants (PDAs).
- Computers.
- Kiosks.

Television

Television (together with radio) was one of the first off-roadway motorist information technologies available to motorists. Even today, commercial television stations in most major cities provide traffic reports to viewers as part of their morning programs to indicate incident locations, traffic signal malfunctions, and other traffic "hotspots." These locations are usually shown on some type of wall-mounted or computer-generated map display.

Public access television is a means of overcoming many of the disadvantages associated with privately owned media stations. Many city governments are responsible for franchising cable television service within the corporate limits of the city. As part of awarding the franchise to a company, many city governments offer their own programming on one or more of these dedicated channels. Public access channels can be used by traffic management agencies to broadcast continuous traffic information during peak hours. Either "crawl" messages across the bottom of the screen or map displays accompanied by voice messages can be used to provide users with information. Traffic reports can also be provided by interrupting normal programming. The primary disad-vantage of using public access television is that the information is available only to cable subscribers. Travelers living outside the service area or not subscribing to the particular cable company do not have access to the information.

One freeway traffic management center in the U.S. has developed the capabilities to provide video and digital data directly to television signals. [(12)] They broadcast live video, a computerized map display, and current control and management actions being implemented (which DMSs have been activated). These signals are scrambled so that only those with the appropriate software and hardware can properly receive the signal. This is believed to reduce the need for censoring the images that are being presented, since they are not available to the public at large.

Telephones

As discussed earlier, telephone "hotlines" can allow both freeway and transit users to obtain pre-trip information via their telephones at home or work, or in their car if they are willing to pay for cellular airtime charges to call the hotline. Information can be provided using recorded messages, synthesized voice messages, or human operators. A touch-tone menu system can be used to tailor information to a specific route or travel corridor. Telephone calls are also a means of developing real-time rideshare matching systems that are being explored in several U.S. locations.

Pagers

A more recent technology now used to help disseminate travel information to users away

from the vehicle are personal pagers. In Bellevue, Washington (east of Seattle), alphanumeric pagers provide their owners with hourly information regarding available carpools looking for riders as well as current traffic conditions on the Puget Sound-area freeway system.[24] The Genesis project of the ITS Guidestar program in Minneapolis, Minnesota and TRANSCOM in New Jersey also utilize alphanumeric pager technology to disseminate travel conditions in real-time to its owners. A key constraint of pager technology is that broadcast messages are limited to a small number of characters. Consequently, efforts are underway to determine which information is most relevant to pager owners and how to format that information succinctly, but accurately.[25]

Personal Data Assistants (PDAs)

Personal Data Assistants (PDAs) are the next higher level of sophistication in off-roadway information dissemination technology. PDAs are computer products that have enough power to support applications such as time management and handwriting recognition. By adding radio frequency (RF) communications technology, PDAs allow users to interact directly with travel information systems. This interaction allows users to obtain route planning assistance, traffic information broadcasts, and other pertinent information. Through keypad entry, the user can log on to the information system, request pertinent information, and then log off. PDAs offer the user increased communication and information transmission/receiving power over alphanumeric pagers. However, they are more complex to use and expensive to purchase. [25]

Computers (Internet)

Personal computers, both desktop and laptop/notebook, can also be used to disseminate information to travelers. As with PDAs, two-way communications between the computer user and the travel information center can be established by telephone modem or an RF communication link. This two-way flow of information allows the user to request information specifically tailored to his or her needs (route planning, traffic information on specific roadway links, transit bus schedule and status, etc.).

Kiosks

The last off-roadway information dissemination device to be discussed is the information kiosk. Kiosks are video monitors mounted on a stand-alone cabinet, in a wall, or on a counter-top. They may have input devices such as keypads, track-balls, or touch-screen displays. Kiosks can serve as an important point of access to travel information networks in a variety of public locations, including the following:

- Hotels.
- Restaurants.
- Airports.
- Gas stations (truck stops).
- Retail establishments (shopping malls).

Kiosks are one type of information dissemination technology where private/public partnerships are particularly relevant. For instance, some agencies are looking to the private sector (through advertising revenue) to help defray the operating and maintenance costs of these systems.[26, 27] Proper kiosk placement is also a critical issue. Many kiosks are not well utilized because they are not located where they can draw attention and reach their intended market.[27]

7.4 LESSONS LEARNED

NTCIP STANDARDS

Equipment compatibility, especially with respect to communications, has long been a problem for many transportation agencies operating various components of freeway management systems. Incompatibilities between DMS and HAR hardware from different vendors have required the agency to have separate systems to communicate with each type of DMS or HAR. This has made it difficult to operate the equipment efficiently and in an integrated manner.

In response to these and other problems, a National Transportation Communications Interface Protocol (NTCIP) is under development to reduce the difficulties associated with developing an integrated freeway management system. These protocols establish standard communications linkages between various types of hardware and software used for traffic management purposes. This ultimately reduces the need for an agency to select and stay with a particular vendor or model of equipment.

In the information dissemination arena, standards have been proposed, or are under development, for DMS and for HAR. Standards for other equipment are likely to emerge as the need arises. Partners are strongly encouraged to utilize NTCIP standards where possible.

INFORMATION MESSAGE DESIGN

Although traditionally a concern associated with public agency information dissemination mechanisms such as DMS and HAR, the proper design and packaging of the information being disseminated is critical to successfully meeting the goals and objectives for all components of an information dissemination subsystem. Evidence already exists on some of the newer ATIS technologies that improper message designs have an adverse impact upon customer (i.e., motorist) interpretation of the worth and usefulness of these technologies.[29] Without proper attention to packaging, the intended audience may not be able to adequately perceive, interpret, and respond to the information. The following factors are critical to proper message design:

- Message load.
- Message length.
- Message format.

Message load refers to the quantity of information that is included in a given message. Message load is an important consideration for DMS operations, because the driver must attend to many different tasks at the same time, while trying to read and comprehend a message. If the message presents too much information, message overload may occur. Some transportation agencies have reported traffic slowing down and other adverse operational problems in the vicinity of DMSs when messages are too long.[1]

Message length refers to the number of words or characters that are displayed on one or more lines of a sign or sequence. Obviously, message length and load are related. However, message length and how the message is formatted, or "chunked," both affect message load.

Message formatting is a term that refers to the arrangement of message elements (words or lines) on an information dissemination device (e.g., a sign) to form a message. Use of abbreviations also relates to message formatting. Placement of message elements on the wrong line or in the wrong sequence

will result in confusion and increased message reading and comprehension times.

The completed and detailed set of guidelines regarding message design from which these paragraphs were drawn are available elsewhere. [2]

ENVIRONMENTAL ISSUES CONCERNING DMSs

Because they are located outdoors next to the freeway, the effectiveness of DMSs as an information source is also affected by environmental factors. Two factors that can be significant are sun position and weathering effects.

Sun Position

Both sunlight falling directly upon the face of a DMS and sunlight located directly behind the DMS adversely affect a driver's ability to see the message. Sign positions with the rear surfaces oriented west-southwest and east-northeast should therefore be avoided if possible. The west-southwest orientation is more problematic, however, because the setting sun tends to be more intense than the rising sun.[1]

Weathering

Because of their proximity to freeway travel lanes, road grime can be a problem for certain DMSs, and significantly reduce the distances at which they are legible. Consequently, DMS manufacturers often place clear Lexan covers over the DMS face to protect the sign from road grime and other environmental effects. One study suggests that legibility distances of a fixed-grid fiberoptic DMS (a freeway lane control signal without a Lexan cover) may be reduced by as much as 25 percent after 18 months in the field. [29] It appears that these losses in legibility can be minimized to a large extent by a regular sign maintenance (pixel lens cleaning, light bulb replacement) program.

"AT-REST" DISPLAY CONDITIONS

Another DMS-specific issue is what to do with the signs when there is nothing special to report to travelers. There are generally two schools of thought regarding when information should be provided in dissemination devices.[2] One perspective is that some type of information such as current time and temperature, an agency slogan, etc., should be provided at all times. Proponents argue that this indicates to motorists that the devices are working. The other perspective is to display information only when there is something new to tell travelers. Proponents of this school of thought argue that, according to human factors principles, trivial or non-traffic specific information should be avoided so that the devices do not lose credibility with the travelers.

7.5 EXAMPLE OF AN INFORMATION DISSEMINATION SYSTEM: TRANSCOM

ORGANIZATION

TRANSCOM (Transportation Operations Coordinating Committee) is a consortium of 14 transportation and public safety agencies in the New York, New Jersey, and Connecticut area whose goal is to provide a cooperative, coordinated approach to regional traffic management. TRANSCOM is funded, staffed, and governed by its member agencies. It has an Operations Information Center (OIC) that is staffed 24 hours per day, 7 days per week. TRANSCOM shares incident information via

alphanumeric pager, phone, and fax to more than 200 highway and transit facilities, police agencies, and radio traffic services. It also serves as a forum for incident and special event management planning, construction coordination, and for shared testing and implementation of regional traffic and transportation management technologies.

The planning process for TRANSCOM began at the end of 1984. TRANSCOM began operations in 1986. Originally, the Port Authority of New York envisioned TRANSCOM as a coordinating mechanism to facilitate traffic flow and operations at the Hudson River crossings. However, TRANSCOM's focus quickly expanded to all traffic facilities in the region.

TRANSCOM is managed by a steering committee made up of senior and mid-level managers from each agency. This committee directs the technical and operational focus of TRANSCOM. Meanwhile, an executive committee consisting of the chief executive officer from each agency decides major policies pertaining to TRANSCOM.

Although TRANSCOM is a public sector information dissemination example, it is considering moving towards a public/private partnership. Public/private partnerships are a much more appropriate example of what is likely to occur in the future.

OPERATIONS

The TRANSCOM OIC receives information regarding incidents and transit facilities from all over the metropolitan area. The operations personnel in the OIC are employees of one of the member agencies and generally have some background in dispatch, operations, or media reporting. The TRANSCOM partnership has allowed several projects to be developed and implemented to facilitate information dissemination and incident response in the region. These include the following:

- Regionwide initiatives for coordinated deployment and operation of VMSs, HAR, and enhanced traffic monitoring via CCTV.

- An enhanced traffic advisory/diversion system at the intersection of the New Jersey Turnpike and Garden State Parkway, which will focus on alternative routing for New Jersey transit buses.

- Expansion of traffic monitoring along the I-287 Tappan Zee Bridge corridor.
- Initiation of the TRANSMIT (TRANSCOM's System for Monitoring Incidents and Traffic) Operational Test. This project uses vehicles with transponders on a highway system equipped with readers/antennas.

TRANSCOM receives information regarding incidents and transit facility conditions from all over the area. The OIC is normally staffed with two operators who receive, collate, and disseminate transportation-related information. The information is relayed to the agencies affected by the incident via alphanumeric pagers, facsimile machines, and voice communications. This information is then disseminated by the various agencies according to their capabilities. In some instances, TRANSCOM has been granted authority to operate the information dissemination technologies for certain agencies. [9]

TRANSCOM operations are fully funded by its member agencies. This funding consists of both monetary contributions and in-kind services such as providing operators to staff the OIC.

FUTURE ACTIVITIES

The executive and steering committees are looking at a number of issues relating to future TRANSCOM operations. TRANSCOM will continue to pursue research and testing of advanced technologies to help the region more effectively manage travel in the congested northeast corridor. The TRANSMIT system is an example of this commitment to future operations. [30] TRANSCOM is participating with other agencies in the development of a computerized Information Exchange Network (IEN) to automate the information retrieval and dissemination process.

7.6 REFERENCES

1. Dudek, C.L. *Guidelines on the Use of Changeable Message Signs*. Report No. FHWA-TS-90-043. FHWA, U.S. Department of Transportation, Washington, DC, May 1991.

2. Dudek, C.L. and Huchingson, R.D. *Manual on Real-Time Motorist Information Displays*. Report No. FHWA-IP-86-016. FHWA, U.S. Department of Transportation, Washington, DC, August 1986.

3. Inman, V., Sanchez, R., Porter, C., and Bernstein, L. *TravTek Evaluation Yoked Driver Study*. Report No. FHWA-RD-94-139. FHWA, U.S. Department of Transportation, Washington, DC, October 1995.

4. *Building the ITI: Putting the National Architecture into Action.* Report No. FHWA-JPO-96-011. FHWA, U.S. Department of Transportation, Washington, DC, April 1996.

5. Vorce, K.E. and Plass, M. A Practical Approach to Changeable Message Sign System Design and Deployment: The I-595 Experience in South Florida. *Proceedings*, Third ITS World Conference, Orlando, FL, October 1996.

6. Implementation Plan Guidance (23 CFR 655 409). Federal-Aid Policy Guide. FHWA, U.S. Department of Transportation, Washington, DC, December 1994.

7. McGowan, P.F., Olson, M.D., and Irwin, P.L. Texas Department of Transportation Advanced Traffic Management System, San Antonio, Texas. *Compendium of Technical Papers*, 64th Annual Meeting of the Institute of Transportation Engineers, Dallas, TX, October 1994, pp 1-9.

8. Roper, D.H., Murphy, R.J., and Zimowski, R.F. Alternative Route Planning: Successful Incident Traffic Management. Special Report No. 153, TRB, National Research Council, Washington, DC, 1975, pp 141-144.

9. Wilson, F.J. The TRANSCOM Coalition: Multi-Jurisdictional Issues in ITS. *ITS Quarterly*, Vol. IV, No. 2, Spring 1996, pp 83-95.

10. Bregman, S., Koses, D.G., and Wilson, A.P. Targeting the Car Phone Market: Findings from a Survey of Callers Using Cellular Telephones to Access the *SmarTraveler* Advanced Traveler Information System. *Proceedings*, Third ITS World Conference, Orlando, FL, October 1996.

11. Dial and Drive -- or Don't. *Inside IVHS*, Vol. 5, No. 3, January 30, 1995.

12. Dellenback, S.W., and McGowan, P.F. Advanced Traveler Information Systems Using Low Power Television Stations. *Proceedings*, 1996 Annual Meeting of ITS America, Houston, TX, April 1996, pp 82-86.

13. Ullman, G.L., Wohlschlaeger, S.D., Dudek, C.L., and Wiles, P.B. *Driver Interpretations of Existing and Potential Lane Control Signal Symbols for Freeway Traffic Management*. Report No. FHWA/TX-93/1298-1, Texas Transportation Institute, College Station, TX, November 1993.

14. Dudek, C.L. *Changeable Message Signs*. NCHRP Synthesis of Highway Practice No. 61. TRB, National Research Council, Washington, DC, 1979.

15. Gordon, R.L., Reiss, R.A., Haenel, H., Case, E. R., French, R.L., Mohaddes, A., and Wolcott, R. *Traffic Control Systems Handbook*. Report No. FHWA-SA-95-032. FHWA, U.S. Department of Transportation, Washington, DC, February 1996.

16. Lavigne, R.C. *Assessment of Changeable Message Sign Technology*. Report No. FHWA/RD-87/025. FHWA, U.S. Department of Transportation, Washington, DC, December 1986.

17. Walker, J., Alicandri, E., Sedney, C., and Roberts, K. *In-vehicle Navigation Devices: Effects on the Safety of Driver Performance*. Report No. FHWA-RD-90-053. FHWA, U.S. Department of Transportation, Washington, DC, May 1990.

18. *Automatic Audio Signing, Volume I: Executive Summary*. Report No. FHWA/RD-82/012. FHWA, U.S. Department of Transportation, Washington, DC, November 1981.

19. Turnage, H.C., Hawthorne, R.C., and Birdseye, M.C.L. *Automatic Audio Signing, Volume II: Prototype Development and Prototype Pilot Testing Demonstration Program. Report No. FHWA/RD-84/039*. FHWA, U.S. Department of Transportation, Washington, DC, June 1984.

20. Rajendra, K., and Maki, R.E. The IVHS Strategy in Michigan. *Proceedings of the IVHS America Annual Meeting*, Atlanta, GE, April 1993, pp 402-406.

21. Von Tomkewitsch, R. Dynamic Route Guidance and Interface Transport Management with ALI-SCOUT. *IEEE Transactions on Vehicle Technology*, Vol. 40, No. 1, February 1991.

22. Dingus, T.A., Carpenter, J.T., et al. Human Factors Engineering the TravTek Driver Interface. Vehicle Navigation and Information Systems Conference Proceedings, Report No. SAE/P-91/253, Society of Automotive Engineers, Inc., Warrendale, PA, October 1991, pp 749-756.

23. Kiefer, R.J. A Review of Driver Performance with Head-Up Displays. *Proceedings*, Third ITS World Conference, Orlando, FL, October 1996.

24. Haselkorn, M., Spyridakis, J., Gable, B., and Michalak, S. Bellevue Smart Traveler: An Integrated Phone and Pager System for Downtown Dynamic Ride Sharing. *Proceedings of the IVHS America Annual Meeting*, Atlanta, GE, April 1993, pp 126-131.

25. Starr, R.A., and Wetherby, B.C. Traffic Information on Alphanumeric Pagers: Evaluation Results from the Genesis Project. *Proceedings*, Third ITS World Conference, Orlando, FL, October 1996.

26. Schroeder, J.L., and Green, J. The Emergence of Smart Traveller Kiosk and the User Interface Requirments for their Successful Deployment. *Proceedings of the IVHS America Annual Meeting*, Atlanta, GE, April 1993, pp 831-837.

27. Pohlman, J.M., and Long, T.I. Advanced Traveler Information Project Travelink. *Proceedings*, Third ITS World Conference, Orlando, FL, October 1996.

28. Guiliano, G., Hall, R., and Golob, J.M. *Los Angeles Smart Traveler Field Operational Test Evaluation*. Report No. UCB-ITS-RRR-95-41. California PATH Program, Institute of Transportation Studies, University of California-Berkeley, Berkeley, CA, December 1995.

29. Ullman, G.L., Tallamraju, S.S., and Trout, N.D. Visibility and Spacing of Lane Control Signals for Freeway Traffic Management. Report No. FHWA/TX-95/1498-1. Texas Transportation Institute, College Station, TX, November 1994.

30. Tarnoff, P.J., and Batz, T. The TRANSCOM TRANSMIT Project: A Unique Institutional Approach to a Unique Project. *Proceedings of the IVHS America Annual Meeting*, Atlanta, GE, April 1993, pp 233-237.

MODULE 8. INCIDENT MANAGEMENT

TABLE OF CONTENTS

8.1 INTRODUCTION 8-3
BACKGROUND 8-3
MODULE OBJECTIVES 8-4
MODULE SCOPE 8-4

8.2 DECISION PROCESS 8-4
IDENTIFY NEEDS 8-5
IDENTIFY INCIDENT MANAGEMENT PARTNERS 8-7
BUILD CONSENSUS AMONG PARTNERS 8-8
ESTABLISH INCIDENT MANAGEMENT GOALS AND OBJECTIVES 8-9
ESTABLISH PERFORMANCE CRITERIA 8-9
DEFINE FUNCTIONAL REQUIREMENTS 8-10
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS 8-10
IDENTIFY AND SCREEN TECHNOLOGY 8-11
Incident Detection and Verification 8-11
Incident Response and Clearance 8-12
DEVELOP PLANS 8-13
IDENTIFY FUNDING SOURCES 8-14
IMPLEMENT 8-15
EVALUATE 8-15

8.3 TECHNIQUES AND TECHNOLOGIES 8-15
INCIDENT DETECTION AND VERIFICATION 8-15
Non-Automated Detection Techniques 8-15
Automated Surveillance Techniques 8-16
Technologies 8-16
Detection Algorithms 8-19
INCIDENT RESPONSE AND CLEARANCE 8-21
On-Site Response Techniques 8-21
Equipment 8-21
Identification Arm Bands and Vests. 8-22
Incident Response and Hazardous Materials Manuals. 8-22
Total Stations Surveying Equipment. 8-22
Inflatable Air Bag Systems. 8-25
Equipment Storage Sites. 8-25
Tow Truck/Removal Crane Contracts. 8-25
Management Strategies/Operations 8-26
Emergency Response Vehicle Parking Coordination. 8-26
On-Site Traffic Control. 8-26
Incident Response Teams. 8-27

Vehicle Removal Laws. 8-28
Techniques to Improve Emergency Vehicle Access and Traffic Flow 8-28
Equipment 8-28
Barrier Openings. 8-28
Barrier Gates. 8-29
Emergency Ramps. 8-29
Management Strategies/Operations 8-29
Police Escorts and Wrong-Way Entrance of Non-Emergency Response Vehicles. 8-29
Shoulder Utilization. 8-30
Contraflow Diversion. 8-30
Accident Investigation Sites 8-31
Equipment 8-31
Management Strategies/Operations 8-31
Freeway Service Patrols 8-33
Equipment 8-33
Management Strategies/Operations 8-33
Electronic Traffic Management and Control Techniques 8-34
Equipment 8-34
Ramp Metering. 8-34
Information Dissemination Components. 8-34
Management Strategies/Operations 8-34
Alternative Route Diversion Techniques 8-34
Equipment 8-35
Management Strategies/Operations 8-37
Work Zone and Special Event Traffic Management 8-37
Equipment 8-37
Management Strategies/Operations 8-37

8.4 LESSONS LEARNED 8-37
WATER ACCESS 8-37
SERVICE PATROL CHARGES FOR ASSISTANCE 8-39
TRAINING 8-39

8.5 EXAMPLES OF FREEWAY SERVICE PATROLS 8-39
CHICAGO EMERGENCY TRAFFIC PATROL 8-39
MINNEAPOLIS HIGHWAY HELPER PROGRAM 8-40
LOS ANGELES FREEWAY SERVICE PATROL 8-40
HOUSTON MOTORIST ASSISTANCE PATROL 8-41
SAMARITANIA, INC. 8-41

8.6 REFERENCES 8-43

MODULE 8. INCIDENT MANAGEMENT

Figure 8-1. Example of an Incident Response Site.

8.1 INTRODUCTION

BACKGROUND

Congestion occurs when the amount of traffic wishing to use a facility (demand) exceeds the traffic-carrying capabilities of the facility (capacity). There are two types of congestion: recurring and nonrecurring. Recurring congestion occurs when normal demand exceeds the physical capacity of the freeway. This type of congestion typically occurs during high volume periods (such as the a.m. and p.m. peak periods) and is predictable in terms of its location, duration, time, and effect. Through experience, most motorists have learned to deal with the effects of recurring congestion by planning for it in their daily routines.

Nonrecurring congestion, on the other hand, is a result of either 1) a temporary reduction in capacity of the freeway (such as an accident or a work zone lane closure) or 2) a temporary excess of demand (caused by a special event or other similar activity). Generally, the activities that cause nonrecurring congestion can be either unpredictable (as in a stalled vehicle) or planned (such as a construction activity). The primary factor distinguishing nonrecurring and recurring congestion is that nonrecurring congestion is unexpected by motorists. As a result, nonrecurring congestion can be a considerable safety hazard and cause excessive delays to uninformed motorists.

Most of the activities that cause nonrecurring congestion can be classified as incidents. Incidents are responsible for a significant proportion of the delays and costs to the motoring public. For example, Caltrans estimates that over 50 percent of all delays experienced by motorists on their freeway system are caused by incidents. [(1)] In Texas, freeway incidents in 1990 were the source of over 440,000 hours of delay, costing the motoring public approximately $2.2 billion. [(2)] By 2005, the impacts of incidents in terms of hours of delay, wasted fuel consumption, and excess road user costs are expected to increase 5 fold over levels experienced just 10 years ago. [(3)]

One method for combating nonrecurrent congestion problems is to implement an incident management program or component of a freeway management system. Incident management refers to a coordinated and planned approach for restoring traffic to its normal operations as quickly as possible after an incident has occurred. It involves a *systematic* use of human and mechanical processes for the following purposes:

- Detecting and verifying the incident.
- Assessing its magnitude and identifying what is needed to restore the facility to normal operation.
- Providing the appropriate response in the form of control, information, and aid.

MODULE OBJECTIVES

The goal of this module is to provide guidance on how to develop or enhance incident management as part of a freeway management system. The specific objectives are as follows:

- To illustrate the decision-making process associated with developing improved incident management capabilities as part of a freeway management system.
- To describe and provide current summary data on the various techniques and devices available to facilitate freeway incident detection, verification, response, clearance, and traffic management.
- To identify and discuss unique issues and lessons learned pertaining to the planning, design, operation, and maintenance of an incident management component of a freeway management system.

MODULE SCOPE

This module is intended to be an overview of the incident management development process. The reader should note that a number of other documents dealing with incident management are available, and should be reviewed for additional information (See references 4,5,6,7.) The focus of this module is to present the decision-making process to assist in identifying, selecting, and implementing an integrated package of incident management actions that are most appropriate for the region of interest.

8.2 DECISION PROCESS

Incident management is a coordinated and planned approach for responding to incidents when they occur on the freeway systems. It involves the systematic use of human and mechanical processes for detecting, responding to, and clearing incidents. In this section, the systems engineering process that serves as the common thread throughout this handbook is applied to the development of this incident management module. It is important to note that incident management

is a separate metropolitan component of an integrated regional transportation system. It is very often implemented with freeway management systems.

IDENTIFY NEEDS

Before an incident management component can be properly designed, it is important to have an understanding of the characteristics and impacts of incidents on freeway traffic in the region. Non-recurrent congestion is caused by one of three events:

- Roadway incidents (crashes, stalls, spilled truck loads, work zones, etc.).
- Special events.
- Regional transportation emergencies (flooding, hurricanes, chemical plant emergencies, etc.).

All of these events can reduce the amount of roadway space available to carry traffic. Furthermore, the last two categories can also increase traffic demands on a roadway segment, causing additional congestion.

To date, data regarding the frequency of incidents remains somewhat limited. Table 8-1 presents a summary of some of the different freeway incident frequencies that have been reported, and the units of measure utilized to describe these frequencies.

Rates of between 12 and 125 incidents per million-vehicle-kilometers (MVK) have been reported in the literature over several years.[1,7] Other studies have reported freeway incident frequencies as a daily or hourly rate per lane-kilometer.[8] Generally, these rates represent all reported incidents, from very minor vehicle stalls lasting only a few moments to major incidents involving hazardous materials or fatalities that take several hours to clear. Rates of more severe incidents that typically involve multiple agency response are significantly lower. For example, data from Houston indicated that a major lane-blocking freeway incident lasting more than 45 minutes occurred once every 137 MVK, or a rate of 0.73 per 100 MVK.[9] Undoubtedly, use of local data is the preferable method of assessing the severity of the incident problem at a location or in a given region.

Table 8-1. Summary of Reported Freeway Incident Rates.

Unit of Measure	Value
Incidents per million-vehicle-kilometers [1]	12-125
Incidents per million-vehicle-kilometers[7]	65
Lane-blocking incidents per hour per lane-kilometer[8]	0.006-0.023
Lane-blocking incidents lasting more than 45 minutes per 100 million-vehicle-kilometers[9]	0.68

Incident severity is also described in terms of its capacity-reducing effect on the roadway. Generally speaking, an unplanned incident or work zone activity will reduce freeway capacity by an amount greater than the reduction in roadway space due to that incident. For example, an incident that blocks one lane of a three-lane freeway section reduces available roadway space by 33 percent, yet reduces the capacity of the freeway by about 50 percent.[1] This additional capacity loss is due to the turbulence caused in the traffic stream as drivers attempt to move from the closed lanes into open lanes, and from driver "rubbernecking."

Table 8-2 summarizes the effect of stalls and accidents on freeway capacity.[1] Data regarding roadway capacities for work zone lane closures are shown in table 8-3.[11] Note the similarity in the estimated percentage reductions in capacity for one and two-lane blockages on freeway sections with three travel lanes per direction.

Generally speaking, frequencies of special events and major transportation emergencies are much smaller than those referred to above. However, when they occur, they can have a severe impact upon mobility and traffic conditions in an entire region. Fortunately, they are usually confined to a specific location and future time period, so that more preparation time can be spent analyzing and implementing various alternatives to help combat any adverse impacts expected.

It is important not only to attempt to quantify the overall magnitude of the incident problem that must be addressed, but also to identify in as much detail as possible the types of incidents that are a problem and the magnitude of their impacts. The degree to which specific problem areas can be identified in this step of the process determines how effectively possible mitigation alternatives are identified and implemented.

Table 8-2. Typical Capacity Reductions During Incident Conditions.[1]

Type of Incident	Normal Number of Lanes	Number Lanes Blocked	Capacity Reduction (%)
Accident on shoulder	3	0	26
Vehicle stall	3	1	48
Non-injury accident	3	1	50
Accident	3	2	79

Table 8-3. Typical Capacities During Work Zone Activities.[(Adapted from 11)]

Normal Number of Lanes	Number of Lanes Closed	Average Capacity (vph)	Reduction in Roadway Space (%)	Capacity Reduction (%)[a]
2	1	1340	50	63
3	1	2980	33	45
3	2	1170	67	81
4	1	4560	25	37
4	2	2960	50	59
5	3	2740	60	70

[a] Percent reduction based on an assumed normal vehicle capacity of 1800 vehicles/hour/lane
vph = vehicles per hour

As a final note, it is important to limit problem identification at this point in the process to the symptoms that need to be addressed, and not try to attach responsibility (or blame) for that symptom to an individual or organization. Part of the difficulty in establishing effective incident management programs is developing the lines of communication, cooperation, and trust within and between the various agencies that need to be involved. These cannot be established effectively in an environment where affected parties are pointing fingers at each other.

IDENTIFY INCIDENT MANAGEMENT PARTNERS

A key to effective freeway incident management lies in the ability of multiple agency partners to coordinate and cooperate before, during, and after an incident. Generally speaking, some or all of the following agencies will have a vested interest in improving freeway incident management capabilities and procedures within a region:[(12)]

- Elected officials.
- State and city/county DOTs.
  - traffic operations.
  - traffic management.
  - maintenance.
  - public relations.
- Transit operators.
- State, city/county, and transit law enforcement agencies.
- Fire and emergency medical services.
- Hazardous materials (Hazmat) contractors.
- Other emergency agencies (office of emergency management, etc.).
- Environmental protection agencies.

- Towing services.
- Corporate service patrol providers.
- Regional authorities (metropolitan planning organizations, council of governments, etc.).
- Media representatives.
- Special event promoters.

Depending on the number of agencies potentially involved in the incident management process, it may be advantageous and necessary to establish smaller subgroups of partners who must interact and cooperate for specific types of incidents.[13] For example, fire and emergency medical services, environmental protection agencies, and offices of emergency management typically have little involvement in the minor accidents and stalls that are of concern to police, DOTs, transit agencies, towing services, etc. Of course, partners such as police and State DOTs will be involved in nearly all subgroups.

BUILD CONSENSUS AMONG PARTNERS

One of the most critical activities that must occur early on in the incident management development process is consensus building among the partners. A consensus is needed on both the importance of optimizing the incident management process and the importance of adopting a "team" approach to addressing the need for coordinated incident management in a region.

Consensus-building does not just "happen." It must be fostered and developed over time. One way to facilitate consensus-building is for each partner to develop a true understanding of the goals, responsibilities, and capabilities that the other partners have within the incident management process.[9] Also, a critical review should be made not only of what the roles and responsibilities of each partner currently are, but also of what they could be (or what the partner wants them to be).[5] Finally, legal ramifications pertaining to each partner's involvement in incident management must be assessed. Formal agreements of understanding and cooperation may need to be written and signed by upper management of each partner to further promote a unified team approach.

Traffic management teams (TMTs), established in numerous regions nationwide, serve as an excellent beginning of a cooperative environment in which to improve incident management. Basic guidelines for successful TMTs are shown below:[1]

- Regular meetings (monthly or bimonthly).
- Attendance by the same personnel at each meeting.
- Attendance by personnel with authority to commit the partner's resources.
- Informal interaction.
- Preparation of agenda for each meeting.
- Focus on reaching verbal consensus on issues being discussed.

It is important that the organization responsible for organizing and running the TMT meeting make an active, consistent effort to identify agenda items and tasks to be undertaken by the team. This is essential to keeping other partners interested and involved. The Palm Beach and Broward TMTs in Florida have hired a consultant to run their meetings. The consultant is seen as having a more organizationally neutral

viewpoint on agenda items, and has more time to keep the various partners interested and involved.

ESTABLISH INCIDENT MANAGEMENT GOALS AND OBJECTIVES

Once a consensus to cooperate and coordinate incident management activities has been reached among the various partners, the goals and objectives that the partners wish to accomplish relative to improved incident management must be defined. As stated in **Module 2**, goals are broad statements about the intent of the system or one of its components, whereas objectives are specific statements about what the system or component of that system will attempt to accomplish. A given goal may have more than one specified objective. Furthermore, the emphasis at this point in the process is on identifying what the subsystem or component will accomplish, not what technology will be employed. Table 8-4 presents examples of some possible goals and objectives relative to incident management.

ESTABLISH PERFORMANCE CRITERIA

Performance criteria are used to determine whether the objectives established in the previous step are being achieved. For incident management, these criteria can be both quantitative and/or qualitative. Examples of some types of performance criteria for incident management are shown in table 8-5.

Table 8-4. Examples of Goals and Objectives of Incident Management.

Category	Examples
Incident management goals	• Reduce the impact of incidents upon peak-period traffic • Reduce the potential for injury to motorists stranded by disabled vehicles • Reduce the freeway congestion that develops at exits to a regular special event
Incident management objectives	• Detect all lane-blocking peak-period incidents within 2 minutes of occurrence • Provide first response to an incident within 5 minutes of occurrence • Reduce the time to clear an incident by 15 minutes • Reduce freeway traffic volume approaching a peak period lane-blocking incident by 10 percent • Detect 75 percent of all disabled vehicles within 20 minutes after they have stopped on the shoulder • Divert 25 percent of traffic that normally uses a given ramp to access a special event facility to other exits

Table 8-5. Examples of Incident Management Performance Criteria.

Objective	Examples of Performance Criteria
Detect all major incidents within 2 minutes of occurrence	• Average detection time • Percent of incidents detected within 2 minutes of occurrence
Provide first response to an incident within 5 minutes of verification	• Average response time • Percent of incidents responded to within 5 minutes of verification
Reduce the time to clear an incident from the freeway by 15 minutes	• Change in average clearance time due to improvements in incident management procedures

DEFINE FUNCTIONAL REQUIREMENTS

Once the goals and objectives for incident management have been identified, the functions (specific features or activities) that are required to achieve the objectives need to be defined. As with the objectives, the functions should be described independent of the technologies that could be used to accomplish the objectives (focusing on what should be done, not how it should be done).[14] Often, these functional requirements can be outlined in a hierarchial order. Figure 8-2 provides an example of possible functional requirements for the incident management component of a freeway management system. Note that each of the functional elements defines an action that is to be performed, and is independent of the technology that could be used to accomplish that action.

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

The purpose of defining the functional relationships, information flows, and data requirements is to establish an understanding of how the various incident management functions that are to be accomplished will be integrated with each other and among the other components in the freeway management system. Much of these relationships, flows, and requirements have been prepared through the National Architecture development effort for ITS. The National ITS Architecture for incident management can be adapted to incorporate local issues, concerns, and capabilities.

Within this step in the decision process, local and regional details that further define when and how incident management functions will be accomplished are established. For example, it may be necessary to subdivide the freeway or corridor into smaller segments because of jurisdictional differences, and adjust the design of the system slightly for each segment, depending on the partners involved, chain-of-command for contacting each partner, etc. Special legislation may need to be established to allow certain incident management tasks (such as service patrol operations) to be privatized.[5] These and other institutional issues identified during problem

Identify incident
- Identify incident location
- Identify incident characteristics (vehicle types, injuries, etc.)
- Identify incident impacts upon traffic flow

Formulate response actions
- Identify necessary emergency vehicle response
- Select incident information for dissemination to travelers
  - Notify transit riders at a park-and-ride lot that their bus may be delayed
  - Notify traffic information services of incident location
- Identify traffic control strategies
  - Increase green time on parallel arterial nearest to the freeway
  - Initiate ramp metering upstream of the incident

Initiate and monitor response
- Initiate coordinated response by the appropriate agencies
  - Implement emergency vehicle response
  - Provide incident information to travelers
  - Implement traffic control strategies
- Monitor response
  - Arrival of emergency vehicles
  - Implementation of traffic control
  - Clearance of incident
  - Clearance of congestion

Figure 8-2. Examples of Incident Management Functions.

identification and goal/objective definition are resolved during this step of the process.

IDENTIFY AND SCREEN TECHNOLOGY

After the various institutional issues relative to incident management have been addressed, the decision process then reaches the point where it is necessary to identify and screen technologies available to achieve the functional requirements and incident management architecture defined for the freeway management system. A review of the various techniques available to facilitate incident management is provided later in this module. This section provides a review of the major issues pertaining to each of the basic phases of incident management (detection and verification, response, and clearance) as they relate to freeway management systems.

Incident Detection and Verification

Incident detection and verification requires the following:[1]

- A means of sensing that an incident has occurred.
- A means of verifying the incident's existence and location.
- A focal point for the fusion of data from multiple detection sources.
- Communications links between detectors and receivers of incident data.

- A means of displaying and recording incident information.

Many different technologies exist which can be used for incident detection and verification. Incident detection in many locations involves a combination of several different detection technologies, and appears to work quite well. Overall, detection and verification technology screening requires a consideration of the following issues:

- Detection speed.
- Accuracy.
- Costs.
- Maintainability.
- Personnel requirements.
- Usefulness of data for other freeway management purposes.
- Speed with which the technology can be implemented and benefits begin accruing.

It is the responsibility of the various partners involved in incident management to decide upon the relative importance of each of these items. The last bulleted item is particularly relevant to all phases (detection, verification, response, and clearance) of incident management. Several agencies have strongly recommended initially implementing low-cost technologies that can provide immediate and demonstrable benefits to the general public, and then building upon those successes to further enhance and improve incident management capabilities in an incremental fashion.[5,12] For instance, cellular phone calls from motorists and CB radio reports from truck drivers have become very important components of incident management efforts in many metropolitan areas, and are often the fastest means available for detecting incidents.[15] Methods of incorporating these detection sources into incident management efforts early on can yield significant benefits almost immediately.

Details of various incident detection and verification technologies are addressed in the Techniques section later in this module, and also in **Module 3**.

Incident Response and Clearance

Technology screening relative to incident response and clearance will depend on the specific types of problems identified initially in the process. Certain types of problems, such as lengthy clearance delays for large truck incidents, can be reduced by making sure that specialized response equipment (heavy-duty wreckers, inflatable air bag systems to upright overturned trucks, etc.) are available from private-sector services. However, specialized hardware or software to assist in incident response and clearance is only one part of the overall technology screening process. Other focus areas to be screened include:

- Minor geometric modifications to enhance response and clearance capabilities (i.e., installing or moving median barrier gates, constructing staging areas for incident management, etc.).
- Institutional arrangements to facilitate cooperation and coordination of incident response and clearance activities among the partners.
- Legislation supporting vehicle removal policies or other clearance activities.

As with incident detection and verification, the screening process for incident response

and clearance should include consideration of the following issues:

- Costs.
- Maintainability.
- Personnel requirements.
- How quickly the technology can be implemented and benefits begin accruing.

Also, it is important to recognize that the technology screening process for incident management cannot occur in isolation from other ongoing or planned freeway management system activities. Implementation actions related to ramp control, information dissemination, HOV treatments, and lane use control all affect and tie into incident management efforts and initiatives as well.

Additional details regarding incident response and clearance are discussed later in the Techniques section of this module.

DEVELOP PLANS

This step in the decision process involves the development of a plan to implement the technologies that have been determined to best meet the goals and objectives of incident management and that are the most feasible for the unique geographic, environmental, and institutional characteristics of the region. As discussed in detail in **Module 2**, this plan documents the following features of the incident management system:[16]

- Needed legislation.
- System design (architecture, technologies utilized, etc.).
- Procurement methods.
- Construction management procedures (if construction is involved).
- Start-up plan.
- Operations and maintenance plan.
- Institutional arrangements.
- Required personnel and budget resources.

In conjunction with the implementation plan (or possibly as a separate development effort), an incident response plan can also be an important activity to facilitate improved incident management within the region. This plan is a detailed document which specifies the following:

- Key partner roles.
- Communication links.
- Detailed traffic management procedures.
- Resources (and their locations).
- Logistics.

Table 8-6 summarizes the information provided for each of the items listed above. Commonly, a detailed response plan recognizes that traffic demands vary over time and location within the region, and so the management procedures required to effectively accommodate an incident will depend upon where and when it occurs as well as its severity. Consequently, some agencies have found it convenient to develop varying levels of incident management implementation. These levels may range from minor on-site management activities during low-volume conditions to a full-scale integrated major response effort involving

Table 8-6. Information in an Incident Response Plan. [1]

Section	Function
Participating Agencies	• Lists all participating partners and telephone numbers of the incident management coordinator for each • Lists other agencies that may participate (i.e., resource locations)
Summary	• Describes major plan elements
Levels of Implementation	• Describes the series or levels of incident management intensity and conditions under which each level is to be invoked
Traffic Management	• Describes communications procedures for each roadway section, agencies to be informed, diversionary routes, traffic control locations, local and regional signing
Resources and Responsibilities	• Lists each partner, contact numbers, key personnel, responsibilities, and other pertinent information
Media Contacts	• Lists each media contact, contact numbers, key personnel, responsibilities, and other pertinent information
Team Coordinators	• Lists all involved partner coordinators and telephone numbers

diversion strategies, information dissemination efforts, and other techniques.[1]

IDENTIFY FUNDING SOURCES

Because incident management is the coordination of multiple entities and techniques to achieve an overall goal (reduced disruptions to travel due to incidents), many different funding sources can play a part in the development and operation of an incident management subsystem. Typically, infrastructure-based investments (automated surveillance, roadway information dissemination tools, traffic signal timing improvements, accident investigation sites) have typically come from the public agencies (State, local) that have direct jurisdiction over the facilities on which they are located.

Initially, courtesy service patrols were funded primarily by DOTs, with the agency benefiting both from positive public relations and from the improvements in traffic operations.[12] In recent years, though, several regions have successfully implemented patrols in which different funding mechanisms have been used.[17] These mechanisms include the following:

- Additional gasoline taxes.
- Additional sales taxes.
- Department of motor vehicle fees.
- Federal construction funds.
- Bridge or highway tolls.
- Congestion Management and Air Quality (CMAQ) funds (for start-up and initial operating costs).
- National Highway System funds.

- Surface Transportation Program funds.
- Private funds.

IMPLEMENT

Experiences from several past incident management implementation efforts indicate that partners should utilize a building-block approach, initiating low-cost components first to demonstrate to the public and to elected officials the benefits of incident management activities.[12] It is then easier to "sell" other components (such as electronic surveillance) that are more capital intensive.

In general, the initiation of service patrols should be one of the first incident management activities that partners consider for their region. Benefit-cost ratios between 2.3:1 and 36:1 have been reported.[17] Again, partners should start small, and gradually build upon this service as they learn more about its operation and develop public and elected official support for it over time.[12]

EVALUATE

Evaluation is a critical component of the incident management decision process. Evaluation helps to define the benefits of incident management in order to maintain and improve funding levels. Also, evaluation is necessary to assess the extent to which the goals and objectives established for incident management are being met. Finally, evaluation is important in identifying new, unforseen difficulties that arise in the incident management process, as well as in identifying possible solutions for dealing with those difficulties.

Methods of evaluation of an incident management program include the following:[18, 19]

- Quantifying the impact upon traffic flow (delays, fuel consumption, emissions, secondary accidents, etc.).
- Critiquing the program through periodic traffic management team reviews or special debriefings held after major incidents.
- Staging mock drills or exercises of a major incident as a training activity and process review.
- Holding post-incident debriefings to review effectiveness of incident management activities and identify areas of improvement.

8.3 TECHNIQUES AND TECHNOLOGIES

INCIDENT DETECTION AND VERIFICATION

Rapid detection is a critical element in the incident management process. The sooner an incident can be detected, the quicker a response to clear the incident can be initiated. Technologies available for detecting incidents range from low-cost non-automated methods to sophisticated automated surveillance techniques requiring extensive public agency investments. It should be noted that emerging Intelligent Transportation Systems (ITS) technologies offer promise for dramatically improving detection capabilities and reliability. The various technologies used by transportation agencies to detect incidents are discussed in the following sections.

Non-Automated Detection Techniques

Whereas most thoughts of incident detection for a freeway management system focus on the various automated technologies that are

available, experience suggests that non-automated detection methods serve a valuable and often primary role in incident detection efforts in many instances. These methods typically utilize motorist call-in technology or manual surveillance methods to achieve incident detection. Non-automated detection methods include the following:

- Cellular telephone calls to 911 or incident reporting hotline.
- Dedicated freeway service patrols.
- Peak-period motorcycle patrols.
- Citizen-band radio monitoring.
- Motorist call boxes.
- Aircraft patrols.
- Fixed observers/volunteers.
- Closed-circuit television.
- Fleet operators (taxis, transit, delivery drivers).

Table 8-7 summarizes the advantages and disadvantages of each of these non-automated detection technologies. It should be noted that the growth of cellular telephone popularity has resulted in that becoming the most important detection technology in most metropolitan areas.[15]

Freeway service patrols are able to begin incident response and clearance activities immediately, which increases their attractiveness from an overall incident management perspective. They are also a valuable public relations tool for the sponsoring agency. [12] Conversely, the use of fixed observers is most applicable to short-term needs (during special events, major freeway construction activities, etc.).

Closed-circuit television can serve as a manual means of incident detection by having system operators continually watch the monitors. However, this technology is more commonly used as a verification tool for incidents detected via cellular telephone reports or automated detection algorithms (discussed in the next section).

Automated Surveillance Techniques

Technologies

Most freeway management systems include some form of automated surveillance which can be used to detect freeway incidents. Types of automated surveillance are listed below:

- Inductive loop.
- Magnetometer.
- Microwave/radar.
- Laser.
- Infrared.
- Ultrasonic.
- Acoustic.
- Machine vision.
- Vehicle probes (automatic vehicle identification, automatic vehicle location).

Module 3 presents additional details about these detectors. The first two technologies, inductive loops and magnetometers, are placed within the pavement. The others are non-intrusive surveillance technologies,

Table 8-7. Comparison of Non-Automated Incident Detection Technologies.[1, 5]

Technology	Definition	Advantages	Disadvantages
Cellular telephone calls to 911 or incident reporting hotline	Motorists use their cellular phones or call from a roadside telephone to report incident.	Often fastest detection method available.	Dependent upon motorist input. Verification needed. May need additional staff to handle calls.
Freeway service patrols	Special vehicles circulate to provide breakdown assistance.	Serves detection, verification, and response functions.	Congestion reduces circulation frequency. Labor-intensive.
Peak-period motorcycle patrols	Motorcycle police officers patrol freeway segments.	Serves detection and verification functions. Already in place as part of regular police functions. Can travel through stopped traffic to get to incident.	Congestion reduces circulation frequency. Labor-intensive.
Citizen-band radio monitoring	Can establish a special frequency for incident reporting.	Inexpensive. Generally can be monitored by existing staff.	Detection dependent upon number of trucks/CB owners on facility. CB owners may need to be trained to use.
Motorist call boxes	Devices located on side of road which motorists can use to notify authorities.	Incident reporting can occur 24 hours/day. Citizen acceptance is high.	Start up costs are high. Requires motorists to walk to activate. Potential for vandalism. May require additional staff to handle calls.

Table 8-7. Comparison of Non-Automated Incident Detection Technologies.[1,5] (Cont'd.).

Technology	Definition	Advantages	Disadvantages
Aircraft patrols	Use of airplanes or helicopters to locate incidents.	Can be used to detect and verify incidents. Can cover a wide region, and see things from a larger perspective (easier to assess impacts of incident).	Costly, significant delays may occur between passes over a given segment. May not be useable during severe weather, fog, etc. Some areas may have airspace restrictions.
Fixed observers/volunteers	Observers are positioned on towers or buildings to watch traffic and report incidents.	Flexible. Useful as an interim measure such as for special events or during roadway construction.	Labor-intensive. Not practical during severe weather conditions.
Fleet operators (taxis, transit, delivery drivers)	Drivers call in incidents they encounter during their normal travels.	Large number of observers can be recruited. Little or no cost to public agency.	Accuracy, reliability cannot be controlled. Limited by fleet size.
Closed-circuit television (CCTV)	Cameras located in traffic management center are continuously monitored by system operators.	Provides detection and verification functions together.	Can cause operator boredom problems. Effectiveness dependent upon camera placement.

located above the travel lanes or off to the side. The first eight technologies typically measure speed, volume, and/or occupancy. Conversely, vehicle probe detection systems monitor vehicle position and elapsed travel time.

Although not listed as an automated detection technology, closed-circuit television is commonly used with automated detection systems to verify that an incident is truly present at a location and to begin evaluating and anticipating appropriate response measures.

Detection Algorithms

The effectiveness of automatic detection technologies depends in part on the type of algorithm used to analyze the detector data. Three parameters are generally used to monitor the performance of an incident detection algorithm: detection rate, detection time, and false alarm rate.[20] Detection rate is defined as the percentage of the total number of capacity-reducing incidents that are detected by the computer algorithm. Detection time is defined as the time between when an incident occurs and when it is detected by the algorithm. The false alarm rate is generally used to provide an indication of how many times the algorithm incorrectly indicates that an incident condition exists when, in fact, no incident is present.

As shown in figure 8-3, there is a general relationship that exists between detection rate, false alarm rate, and detection time. With most incident detection algorithms, the false alarm rate increases as the detection rate increases. Also, the false alarm rate increases as the detection time decreases. This is because as the sensitivity of the algorithm is adjusted to detect less severe incidents more quickly, minor fluctuations in traffic can trigger the algorithm to signal that an incident is present, when in fact no incident exists on the freeway.

Agencies must decide for themselves what is an acceptable balance between detection sensitivity and false alarm rate for their detection system. False alarms can be tolerated in order to achieve a higher detection sensitivity, so long as they are not too frequent and liable to be ignored by the system operator. Evaluation of false alarm rates should generally be based on the their frequency over a given time period (i.e., hourly or peak period). An algorithm may be reported as having a very small false alarm rate percentagewise (based on the total number of detection "checks" it performs), but yield a fairly high number of false alarms because the total number of such "checks" made during a given time period is so high.

The four general categories of incident detection algorithms that rely on volume, speed, and or occupancy data include the following:[21]

- Comparative.
- Statistical.
- Time-series/smoothing.
- Modeling.

Comparative (or pattern recognition) algorithms compare traffic parameters at a single detector station or between two detector stations against thresholds that define when incident conditions are likely. Statistical algorithms use statistical techniques to determine whether observed detector data differ statistically from historical or defined conditions. Time series and smoothing algorithms compare short-term predictions of traffic conditions to measured traffic conditions. Modeling algorithms use standard traffic flow theory

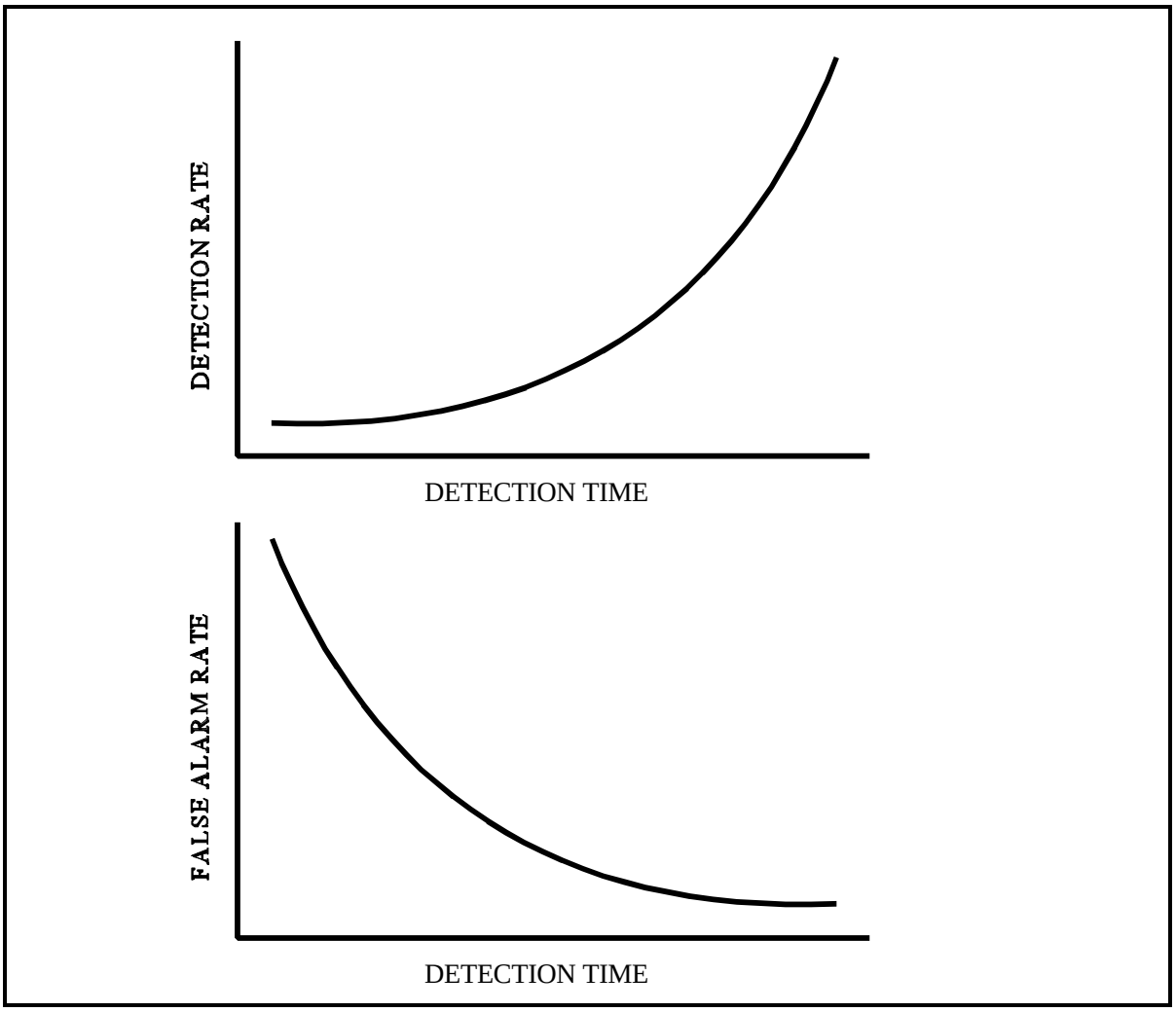


Figure 8-3. Relationship between Detection Rate, False Alarm Rate and Detection Time.[20]

to model expected traffic conditions on the basis of current traffic measurements. Table 8-8 lists the algorithms available in each category.

A recent review of the algorithms indicated that two of the Modified California Algorithms (#7 and #8) and the McMaster Algorithm rated the highest on the basis of reported performance, operational experiences, and model complexity. When calibrated properly, these algorithms can be expected to achieve detect 70 to 85 percent of all incidents, while incorrectly triggering a false detection alarm about 1 percent of the time or less.[22] Conceptually, these algorithms were also judged to be easy to understand and implement from an operator's perspective.

Detectors that monitor traffic parameters other than volume, speed, and occupancy (i.e., travel times, individual vehicle movements, queue lengths, etc.) have only recently been introduced. Consequently, only limited experimental data is available on the feasibility of these technologies for detecting incidents.[23, 24] More work in this area will be needed before AVI/AVL or

Table 8-8. List of Available Incident Detection Algorithms.[22]

Algorithm Type	Algorithms	
Comparative	California All Purpose Algorithm Pattern Recognition (PATREG)	Modified California (10 different algorithms)
Statistical	Standard Normal Deviate (SND)	Bayesian
Smoothing/Filtering	Exponential Smoothing	Low-Pass Filtering
Traffic Model	Dynamic Model	McMaster

other non-traditional technologies will be viable for incident detection purposes.

INCIDENT RESPONSE AND CLEARANCE

Response is defined as the activation, coordination, and management of appropriate personnel, equipment, communication links, and information media as soon as there is reasonable certainty that an incident is present. Steps in the response process include:

- Verifying the existence and location of the incident.
- Assessing the incident to determine the type of response needed to clear it.
- Initiating the appropriate response.
- Removing the incident.

A quick and timely response by the necessary resources to clear the incident can significantly reduce its duration. The following sections provide a summary of the technologies available to facilitate incident response and clearance, and thereby reduce its total duration and impact upon travel in the region. Each category is discussed in terms of equipment, management strategies, and operations involved.

On-Site Response Techniques

One of the most critical facets of incident response is the utilization of equipment and management strategies on-site. The effectiveness of incident response is a function of both the use of appropriate techniques for the situation and how well these techniques are managed by the personnel present. A well managed response that utilizes minimal available resources can still operate more effectively than a poorly managed response that has all of the latest and greatest equipment. In the following sections, equipment and management strategies/operations relative to on-site incident response are summarized.

Equipment

Several types of equipment are available to help with on-site response. These range from items that improve the management of resources, to special-use equipment items that help reduce the response and clearance times of certain types of incidents. Some of the more effective items are discussed below.

Identification Arm Bands and Vests. These can be used to quickly differentiate respondents from members of the public or media who may also be at the incident site. Specially-designated colors and/or patterns can be provided for each of the agencies. This can reduce confusion, and help in controlling who is allowed within the incident site.[5]

Incident Response and Hazardous Materials Manuals. Having adequate documentation detailing how to handle an incident situation can significantly reduce response times. Generally, a transportation agency needs to develop two types of manuals: a response manual and a hazardous materials manual.

- *Response Manual* - The response manual outlines how to respond to specific incident situations and who should be contacted.[6] It should list all of the resources available for responding to an incident regionwide, including both public agency and private sector sources. This list should include the locations and possible operators of large tow trucks, special incident handling equipment, and equipment suitable for handling hazardous materials.

  A key to the effectiveness of these lists is a commitment to keeping them current, updating them as personnel and other resources change within the region. Agencies may wish to consider keeping the manual on-line in a computer database to help keep it current. The manual may also include other important information such as maps, diagrams of selected interchanges, milepost identifiers, utility locations, and sensitive or hazardous off-freeway facilities. It should be organized to facilitate immediate access to specific information or telephone numbers. The type of information generally included in a response manual is shown in table 8-9.

- *Hazardous Material Manual* - Although the primary rule for responding to hazardous materials spills is "leave it to the expert," it is necessary to provide those agencies who are likely to respond first to an incident scene with some basic guidelines for when they suspect that hazardous materials may be involved.[6] These guidelines should be provided in the form of a reference manual. Two hazardous materials manuals are generally developed: one for use by field personnel (e.g., police and DOT personnel) on the scene, and one for use by response dispatchers. The field manual should include guidelines on identifying types of hazardous materials and how to stabilize them at the incident scene. In addition to describing who should be contacted in case of a hazardous materials spill, the dispatcher's manual should also include more detailed information that can be used to consult and advise the field personnel as to any situation that may occur before a hazardous materials response team arrives on the scene. Table 8-10 shows potential subject matter to be included in a hazardous materials manual.

Total Stations Surveying Equipment. This equipment utilizes infrared surveying technology to measure distances critical to an accident investigation. This technique is currently being used by law enforcement personnel in several States as an aid in investigating major accidents.[25, 26]

Table 8-9. Example of Resource Material Included in Response Manual. [6]

Police State City Park Tollway County (including sheriffs) Military	**Fire/Rescue** State City Airport County Industrial Military
News Media Radio stations Newspapers Television stations Traffic reporting services (Metro, Shadow, Smartroutes, etc.)	**Local and State Agencies** Health Pollution control Agriculture Air control
Highway Department Engineering Maintenance Cleanup Traffic Management Center Tollway or turnpike authority Traffic Management Team	**Emergency Medical Services** Coroner Red Cross Funeral homes Helicopters Special medical vehicles Ambulance Hospital emergency rooms Rescue squads-extrication
Special Vehicle and Equipment Cranes Oversize wreckers Tanker trucks Trucking companies Local transit service Livestock trailers Earthmoving equipment	**Towing and Road Service** Auto clubs Franchised tow truck operators Private-gas stations, garages, junkyards Public-police, hwy authority, service patrol
Special Hazard Teams Chemical Electrical Mechanical Biological Radioactive Ordinance disposal	**Utilities** Telephone Electric Gas Water Sewer Cable
Federal Agencies Department of Energy Energy Resources Development Administration Federal Aviation Administration Department of Defense U.S. Public Health Service Defense Civil Preparedness Agency Office of Emergency Transportation Environment Protection Agency Department of Agriculture Postal Service Federal Emergency Management Agency	**Other** National Guard and Reserve Accident investigation teams Vehicle rental companies Institutions Humane society Game warden Military personnel Railroads Weather bureau Pipeline companies Water authorities Scuba divers Transportation services

Table 8-10. Example of Material to Be Included in a Hazardous Materials Manual.[6]

- **INTRODUCTION**
- **RESPONSE AT SITE**
  - **Information to be gathered**
  - **Specifics of spill**
    - Liquid/gaseous
    - Description of leak
      - Rate of flow/quantity spilled
      - Odor
      - Color
      - Density
    - Type of container
      - Box, box trailer
      - Tanker type
    - Precise labels from truck
      - UN numbers
      - Company name
  - **Drainage systems in area**
    - Ditches
    - Bodies of water
  - **Weather conditions**
  - **Traffic flow**
    - Number of lanes open/blocked
  - **Communications**
    - Communications with central command post
      - Fire
      - Other police
      - Ambulance
      - Environmental protection
      - Other
  - **Securing the scene**
    - Establish field command post
    - Cordon off area
      - Green zone
      - Yellow zone
      - Hot zone
    - Types of vehicles to position in each zone
    - Implement traffic diversion plans
- **RESPONSE AT CENTRAL COMMAND POST**
  - **Query field personnel to obtain all relevant information on spill**
  - **Notify other agencies**
    - Environmental Protection Agency
    - Local Boards of Health
  - **Notify local contractors if required for clean-up**
  - **Notify personnel on scene as to protection required**
  - **Notify media**
  - **Utilize available literature and guides**
    - US DOT Guidebook Chemtrec Center, Washington, D.C.
  - **Notify shippers**
- **APPENDICES**
  - Drills and Training
  - State Regional Coordinators
  - County Offices - Boards of Health
  - Traffic Control Guidelines
  - Blank Forms for Environmental Protection Agencies
  - Radiation Accidents
  - References to Laws and Regulations
  - List of References

The equipment is used to quickly obtain measurements needed to reconstruct the accident, and can significantly reduce the time required to collect the necessary data to reconstruct the accident. It also allows more accident-related information to be collected with greater accuracy. A supporting computer system is required to achieve the maximum benefits. In addition, training and continued use is required to ensure that the officers are using the equipment properly. For example, the Maryland State Police have three "crash" teams which operate total stations equipment and are proficient in its use. They are on call 24 hours a day and report to an incident within 30 minutes.

Inflatable Air Bag Systems. These can be used to right overturned heavy vehicles.[5] The system consists of several heavy rubber inflatable cylinders of various heights. These bags are placed under the overturned vehicle at strategic locations, and inflated to right the vehicle. This system can be used in almost any location to right an overturned vehicle. Wreckers are still required to assist in the process, and, depending upon the size and shape of the vehicle, some vehicles can puncture the bags. Most large-scale towing and recovery specialists in large metropolitan areas have this type of equipment available. Appropriate contracts may need to be established between public agencies and these specialists to ensure that the equipment is available for use in an incident.

This type of system is ideal for righting vehicles with fragile loads or tankers where other means of righting the vehicle may rupture or damage the cargo. It is also ideal for working in constrained areas, such as tunnels, bridges, and overpasses, where larger towing or response vehicles may have difficulty maneuvering. Figure 8-4 illustrates the use of an air bag system.

Equipment Storage Sites. These sites are key locations, normally suffering from high incident rates, where incident response equipment is stored (particularly traffic control devices). To be effective, they need to be easily accessible and used by all response agencies. Also, agreements must be established as to who will stock and maintain the equipment and materials contained at the storage sites.[5]

Tow Truck/Removal Crane Contracts. These contracts can achieve minimum response times in specific sections of freeway. Tow trucks can be summoned to an incident location either by using a rotation list (in which each wrecker service in a specific section of freeway is called in order) or by securing a bid contract for service for a specific section of freeway. The use of tow truck contracts or agreements may require local agencies to adopt ordinances that ensure that existing wrecker services are not adversely affected. It is also essential to have a local ordinance providing the police with the authority to establish wrecker contracts, to lay the ground rules for using such contracts, and to determine penalties for non-compliance by wreckers with the ordinance. Some requirements that are commonly used in wrecker contracts include the following:[1]

- Minimums on equipment, storage space, insurance, and licensing.
- Specifications calling for the availability of heavy duty towing and recovery equipment.
- Twenty-four hour availability of wrecker service.
- Specified minimum response times (30 minutes is commonly used).

Figure 8-4. Air Bag Use to Overturn Tractor Trailer.

Management Strategies/Operations

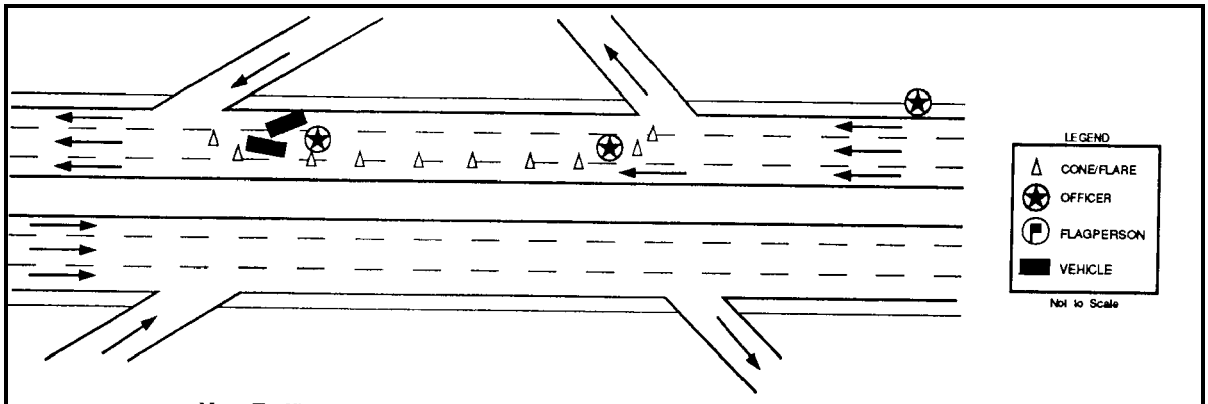
In addition to special equipment to facilitate on-site response, a number of management strategies can be developed and invoked, as necessary, to assist in the response process, as summarized in the following sections.

Emergency Response Vehicle Parking Coordination. This coordination may be enacted to help facilitate an "order" as to the number of response vehicles parked in travel lanes at an incident site and how other response vehicles will be parked on the shoulder or similar off-roadway location in the most efficient manner possible. This "order" may differ depending on the type of incident.[5] However, the overall goal is to minimize the amount of maneuvering required by the different response vehicles, to ensure that no more space is taken up by the incident scene than is necessary, and that lanes can be reopened with minimal movement of response vehicles or disruption of incident clearance activities.

On-Site Traffic Control. This control is required to facilitate the orderly movement of traffic past the incident site or off of the freeway (in the case of a total freeway closure).[27] Channelization of traffic can be accomplished with flares or cones, depending on the anticipated duration of the incident. If the incident is anticipated to last several hours or days, a more elaborate traffic control plan, similar in content to a work zone lane closure set-up, should be employed. Figures 8-5 and 8-6 illustrate

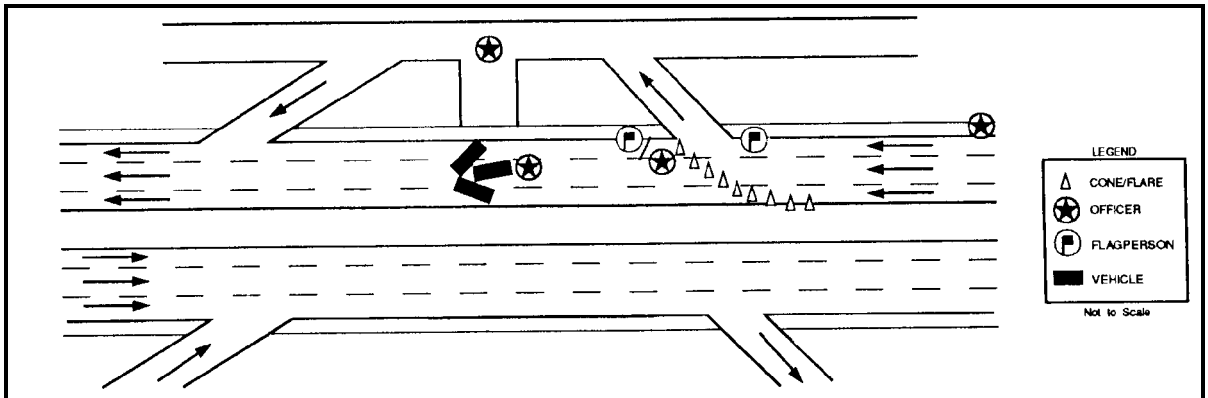
typical set-ups for traffic flow past an incident site and diverted completely from the freeway, respectively.

Incident Response Teams. These teams are often needed to clear major incidents. By assembling a major incident response team prior to an actual incident, a faster and more coordinated response can be provided.[1] Most major incident response teams are composed of individuals from law enforcement, traffic engineering, maintenance, and fire and emergency services. These individuals should be of sufficient rank to make decisions about committing the resources of their agencies without further approval from their superiors.



Note: Traffic control plan depicted is for illustration purposes only under emergency conditions.

Figure 8-5. Typical Set-up for Traffic Flow Past a Freeway Incident. [27]



Note: Traffic control plan depicted is for illustration purposes only under emergency conditions.

Figure 8-6. Example of a Typical Total Freeway Closure at an Incident. [27]

At a major incident, the team is responsible for assessing the situation and coordinating the implementation of a prepared response within each member's own agency. The team can also make on-site adjustments to the plan to conform to the specifics of the situation. Examples of the types of decisions that will be required by team members include:[1]

- Where traffic will be diverted.
- How and when the wreckage will be cleared.
- How and when repairs to the roadway will be made.
- When the roadway can be partially reopened.

The most important aspect of a major incident response team is that the same individuals report to the incident scene each time, and are familiar with the personnel, authority, and resources of each of the responding agencies. The major function of the team is to handle all logistical problems in support of the incident commander, and to smoothly transfer control of the scene from one agency to another as the incident progresses. Also, the team can be the single source for information to the media, so that consistent information is given out. One means of effecting ongoing improvement in incident response team operations is to hold post-incident debriefings to evaluate procedures and identify areas of improvement.

Vehicle Removal Laws. These laws are legislative or administrative policies that promote the fast removal of disabled, abandoned, or damaged vehicles that constitute a hazard to other motorists. Local laws or ordinances can be enacted that require motorists to move their vehicles (if driveable) off the freeway immediately after an incident. These laws can also define the maximum time limit for leaving a vehicle unattended in the right-of-way, and can establish procedures for removing vehicles by local authorities with push bumpers or tow trucks. To be effective, the enactment of these laws must be followed up with an extensive publicity campaign informing motorists of what to do in case of an incident. Local agencies must also translate the new laws or policies into effective operating procedures and must provide a mechanism for enforcing the new law.

Techniques to Improve Emergency Vehicle Access and Traffic Flow

The next category of techniques to be discussed are those related to improving vehicle access to, from, and around an incident. Whereas the controlled-access and barrier-protected designs of freeways promote the highest degree of safety possible, they are problematic when it comes to trying to get emergency vehicles to an incident site when traffic has queued behind the incident. Equipment and management strategies relative to vehicle access are discussed below.

Equipment

Barrier Openings. These openings can be designed into freeway sections with inadequate access for emergency vehicles, thereby reducing the time it takes for those vehicles to reach an incident location. However, adequate crash cushion protection must be provided for the exposed end of the barriers. Also, an improperly designed opening can be a temptation to freeway motorists to use as a turn-around if they miss an intended exit.

Barrier Gates. These gates have been used in some jurisdictions to combat unauthorized vehicle use. The gates are kept closed and locked until access is needed by an authorized vehicle. One of the major concerns with previous gate assemblies, however, has been with their crash worthiness. Houston has recently installed new remote-controlled barrier gate systems that the developers claim will meet NCHRP Report 350 standards for longitudinal barrier crash worthiness.[28] Figure 8-7 shows this barrier gate.

Emergency Ramps. As agencies reconstruct their aging freeways, entrance and exit ramps are often removed or relocated to improve freeway operations. Unfortunately, this can also reduce the ability of emergency vehicles to reach an incident scene. In some instances, emergency ramps are employed to maintain emergency freeway access. These ramps are closed to regular traffic, but can be utilized by authorized vehicles through a manual or electronic gate system.

Management Strategies/Operations

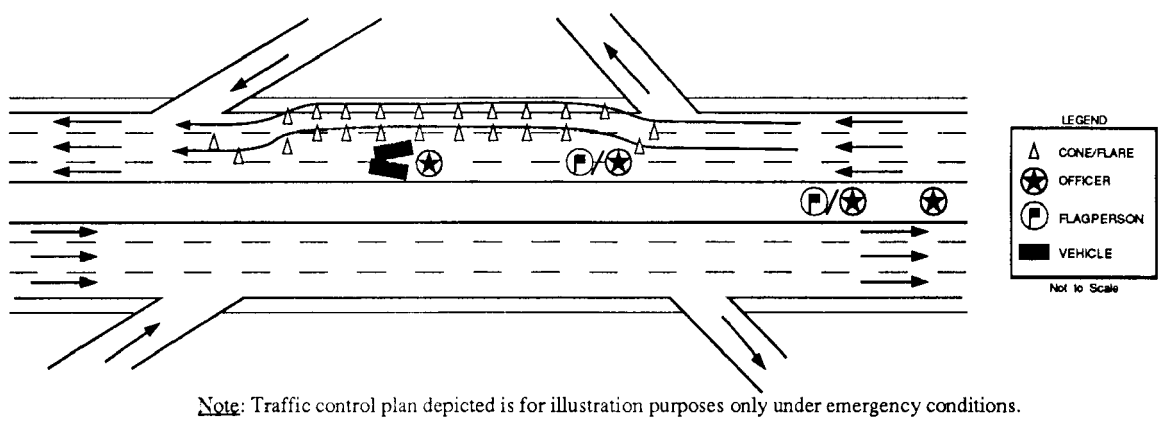
Police Escorts and Wrong-Way Entrance of Non-Emergency Response Vehicles. This strategy can yield significant incident response benefits. Since both towing services and most DOT equipment utilize yellow flashing lights, it is difficult for them to reach an incident scene. Providing for police escorts of these vehicles, or allowing them to enter the freeway in the wrong direction (i.e., from downstream of the incident) can reduce access times of these

Figure 8-7. Example of Crash Worthy Barrier Gate.

vehicles. Of course, this strategy must be planned beforehand and implemented through ongoing training efforts.

Shoulder Utilization. This management technique has been successfully used to increase traffic flow capacity around an incident site. This is normally implemented best by uniformed officers positioned upstream of the incident site. For this technique to be effective, a minimum of 3 m of clearance is needed from the outer edge of the paved shoulder to the incident or any emergency vehicles that are needed at the response scene. Figure 8-8 illustrates the set-up of a shoulder utilization technique at an incident site.[28]

Contraflow Diversion. This technique can sometimes be used at an incident site that will close the entire freeway for several hours. This involves getting traffic across the roadway median (or across the concrete barrier via a barrier opening or gate as described earlier), utilizing a travel lane from the opposing traffic direction to bypass the incident, and returning the traffic back to its original side of the freeway. It also requires close coordination between law enforcement and transportation personnel, and is generally warranted for only the most severe, special cases. Figure 8-9 illustrates a contraflow diversion set-up at a freeway incident site.



Note: Traffic control plan depicted is for illustration purposes only under emergency conditions.

Figure 8-8. Example Set-Up for Shoulder Utilization at a Freeway Incident Site. [27]

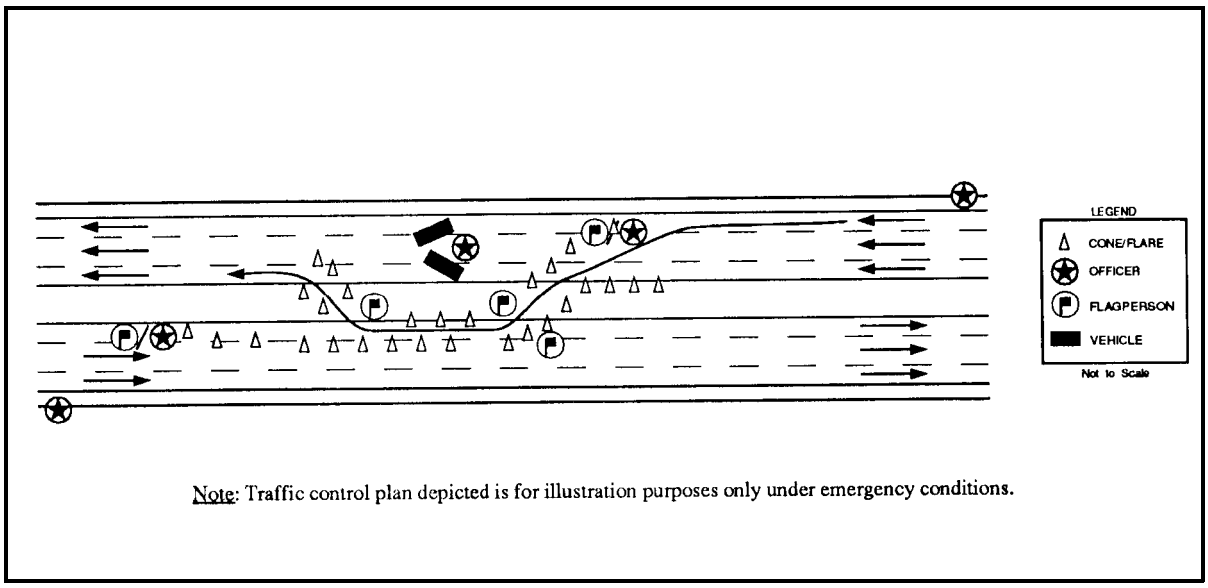


Figure 8-9. Example of a Contraflow Diversion Set-Up at a Freeway Incident. [27]

Accident Investigation Sites

Accident investigation sites are special designated and signed areas off the freeway or roadway where drivers of damaged vehicles can exchange information, and police and motorists can complete necessary accident report forms. These sites can also serve as an incident relocation point by agency-operated clearance equipment and as a media assembly location. In order to reduce rubbernecking, accident investigation sites are generally located so that motorists involved in an accident, the investigating police officer, and the tow truck operators are out of view from the freeway. The benefits of an accident investigation site include reduced motorist delays, reduced vehicle operating costs, reduced secondary accidents, reduced pedestrian exposure, and more efficient use of a public agency's personnel. Unfortunately, such sites have not been well received by the public or law enforcement in some instances, due to security concerns, poor access, or a general lack of awareness about their existence and location.

Equipment

Typical locations of accident investigation sites include under a freeway overpass, on a side street or parallel frontage road, or in a shopping center parking lot out of view of the freeway. An accident investigation site should have space for parking a minimum of five vehicles. This equates to a minimum size of 92 square m. Also, there should be a minimum of 31 m, longitudinally, to pull into and out of curb parking accident investigation sites.[29] Figure 8-10 illustrates examples of accident investigation site design.

Management Strategies/Operations

The criteria for locating an accident investigation site includes the following: [29]

- Easy access to and from the freeway.
- Sufficient overhead lighting and other provisions to ensure personal safety.

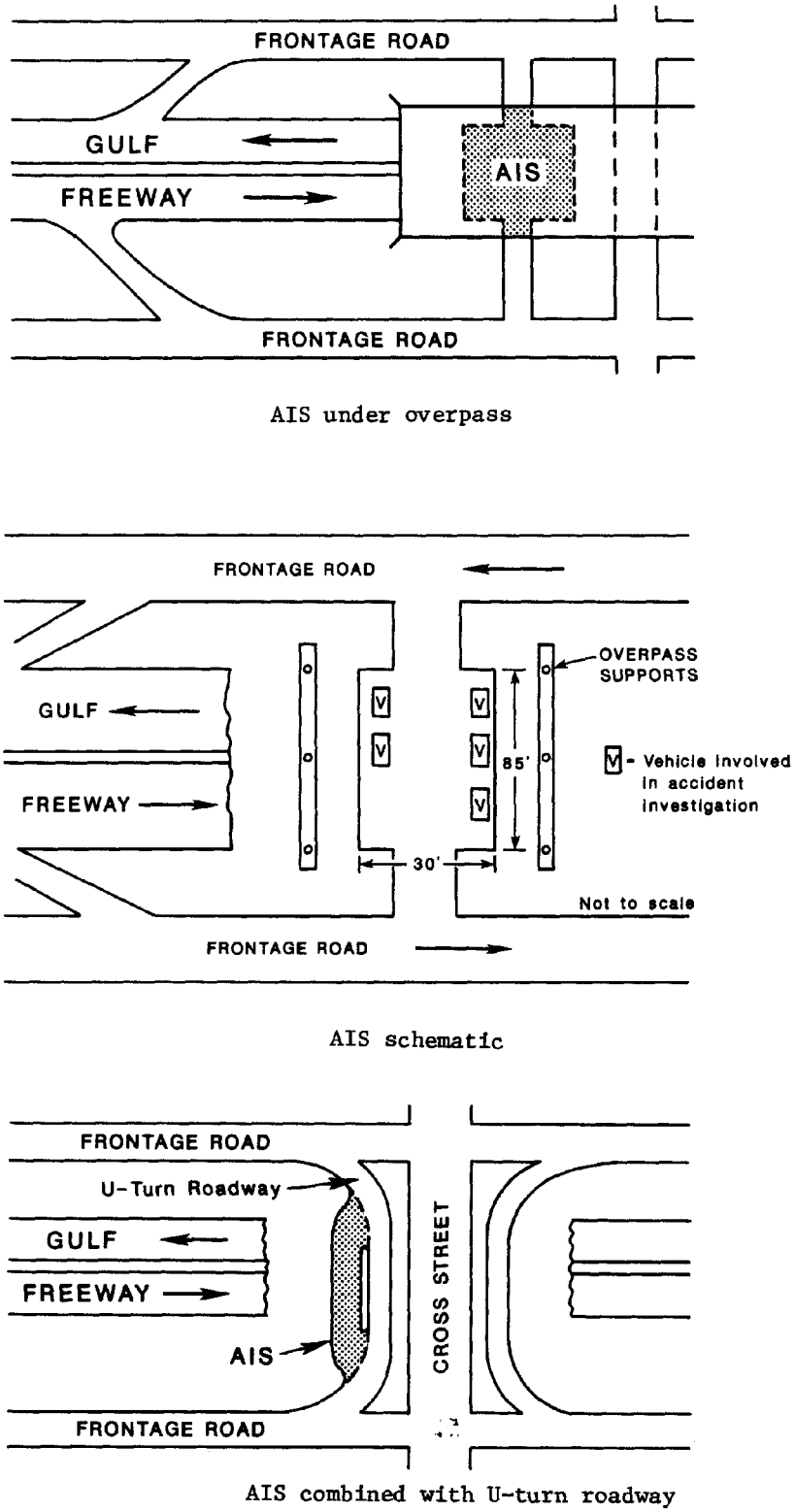


Figure 8-10. Examples of Accident Investigation Site Locations.[29]

- Concealed from the freeway.
- Well delineated and signed.
- Located near a high accident location.
- Little or no construction required.
- Access to a public telephone.

Also, the provision of accident investigation sites must be supported by proper vehicle removal legislation and by law enforcement personnel in order to be effective.

Freeway Service Patrols

Freeway service patrols are a major tool for combating the effects of incidents in metropolitan areas. The primary objectives of service patrols are to locate incidents, reduce the risks to motorists and patrol personnel around the incident, and reduce incident duration so that full capacity can be restored to the freeway as quickly as possible.

Equipment

Typical service patrols rely on light-duty trucks or vans, equipped with a wide assortment of supplies that could be needed to assist a stranded motorist. Some agencies, such as the Illinois DOT in Chicago, use medium-duty wreckers in their service patrol fleet to allow them to quickly relocate automobiles ans small trucks from the freeway lanes. Some of the more common supplies include the following:

- Gasoline.
- Water.
- Tools for minor automotive repair.
- Jumper cables.
- Flares.
- First-aid kit.
- Warning lights.
- Push bumpers.

Access to more extensive and specialized equipment may need to be worked out through private-sector agreements or by other methods. This equipment includes such things as heavy-duty tow trucks, removal cranes, sand spreaders, and rescue and extricator trucks. For large incident management operation in a major metropolitan area, this equipment will need to be accessible 24 hours per day, 7 days per week. [30]

Management Strategies/Operations

Freeway service patrols can be operated from stationary points to deal with spot locations (i.e., bridges, tunnels, construction zones, etc.), dispatched on a call-in basis, or circulated throughout a coverage area (the preferred method for dealing with larger freeway sections).[12] They can be organized and funded by several different organizations:[17]

- Enforcement agencies.
- Transportation agencies.
- Private organizations.
- Multijurisdictional cooperative arrangements.
- Private sector contracts with public agencies.

Electronic Traffic Management and Control Techniques

Equipment

The various traffic management and control components of a freeway management system are intended to assist in incident management activities. These components help to warn motorists approaching the incident about downstream traffic conditions, advise about reduced advisory speeds, reduce approaching traffic demands, and adjust control settings on other roadways that are accommodating increased traffic volumes due to diversion from the freeway. The management and control components that can be used for this purpose include ramp metering and information dissemination.

Ramp Metering. Ramp metering can be implemented upstream of incidents to reduce traffic demands entering the freeway. This technique requires metering of all entrance ramps along a section of freeway (since incident locations vary) and real-time control of these ramp meters.[31] Ramp metering typically operates during peak traffic periods, and may be adjusted slightly if an incident occurs during these periods. However, implementation of ramp metering explicitly for incident management during off-peak periods (when ramp metering is not normally operating) is less common. Additional information regarding ramp metering can be found in **Module 5**.

Information Dissemination Components. These components are typically activated to warn motorists upstream of an incident which travel lanes are closed and to encourage motorists to leave the freeway early or not enter the freeway at all. Notification can occur via any of the different information dissemination technologies (discussed in **Module 7**) present within the region, but must be coordinated and managed so that the information remains current (particularly with respect to location, expected duration, and its impact upon traffic conditions).

Management Strategies/Operations

Although the exact location, severity, and other characteristics of any given incident cannot be predicted in advance, some level of planning is appropriate for the management and operation of both ramp metering and motorist information subsystems for incident response. Table 8-11 summarizes some advance planning steps to facilitate the use of certain information dissemination technologies for incident response.

Alternative Route Diversion Techniques

Alternative route planning is a systematic process that involves examining where and how much traffic should be diverted whenever an incident or other blockage occurs on any section of freeway at any time of the day. In effect, alternative route contingency plans are developed for various levels of freeway incidents anywhere in the system.[12] Coordination and cooperation with local agencies during alternative route plan development is essential for these routes to be safely and effectively implemented when needed.

Alternative route planning involves determining not only where and how much traffic should be diverted, but also when diverting traffic would produce positive benefits. Since diverting traffic to alternate routes is often politically sensitive, how long a freeway is to remain closed before an official detour route is established is often a policy decision. For example, some areas

Table 8-11. Advance Information Dissemination Planning for Incident Response.[Adapted from 1]

Technique	Planning Required
Variable Message Signs	• If messages are not computer selected, establish responsibilities for message selection and display. • Establish criteria for message selection/display. • If portable, establish dispatch procedures.
Highway Advisory Radio	• If messages are not computer selected, establish responsibilities for message composition and recording. • Establish criteria for message selection/activation. • If portable, establish dispatch procedures.
Private-Sector Information Service Providers (ISPs)	• Information transfer protocols and agreements need to be established between ISPs and agencies providing data.
Commercial Radio/Television	• Set up means for transfers of information to radio/television stations. • Meet with station managers to establish guidelines for reporting frequency. • If information not transmitted by computer, set up procedures and responsibilities for message transmissions.
Print Media	• Define formats for press releases. • Set up procedures and responsibilities for issuing press releases.
Telephone Hotlines	• Set up procedures and responsibilities for message updating and recording.

divert traffic only when an incident is likely to last more than one hour.

Equipment

The information needed on any alternative route plan may differ from jurisdiction to jurisdiction. Most commonly included are specifications about the equipment and manpower that will be needed to implement the specific plan, such as police officer control locations, barricades, signing, etc.[12]

Initially, alternative route plans were prepared on hardcopy printouts that were distributed to law enforcement personnel and to the incident response team. An example of an alternative route plan is provided in figure 8-11. Recently, however, these plans have began to be converted to computerized format on Geographic Information System (GIS) or other platforms to assist in organization and retrieval. Methodologies are also being developed to computerize

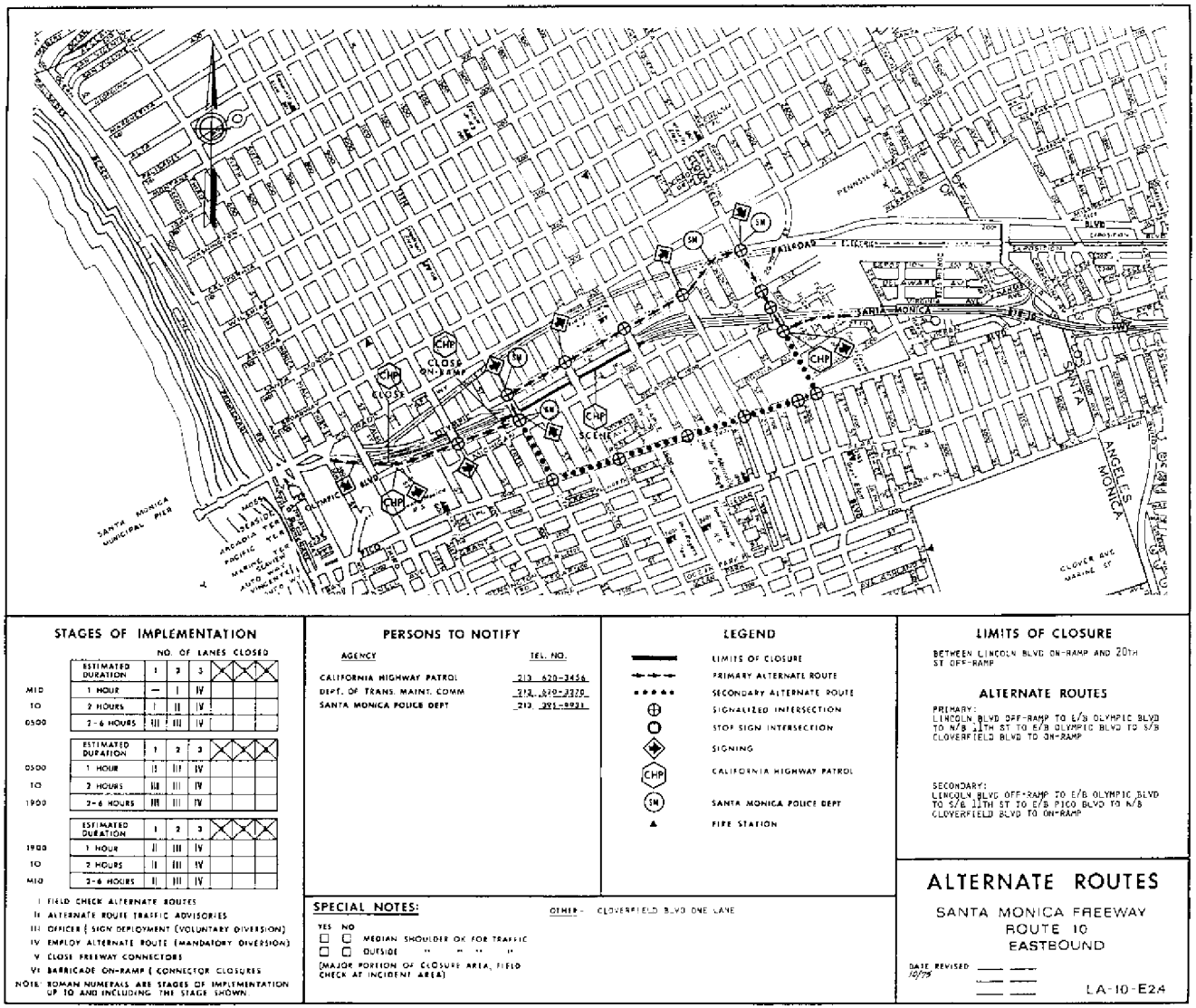


Figure 8-11. Alternative Route Plan.[12]

some of the decision-making processes involved in alternative route selection. [32]

Management Strategies/Operations

Planning of alternative routes must be done by a team consisting of State and local transportation and enforcement agencies, as a minimum.[12] Not all arterials near a freeway may be desirable alternative routes. Features that make an arterial undesirable as an alternative route include the following:

- Schools.
- Hospitals.
- "Sensitive" neighborhoods.

Work Zone and Special Event Traffic Management

Work zones and special events constitute a special type of incident in that the location and duration are usually known in advance. This allows agencies to analyze beforehand the potential impacts of a work zone or special event, and if necessary, to develop a customized package of transportation system improvements to mitigate those impacts.

Equipment

General guidelines for traffic control are provided in the *Manual on Uniform Traffic Control Devices*.[33] In addition, various state and local agencies adopt their own version of these guidelines. These guidelines cover appropriate advance signing, barricades, channelizing devices (where lanes are to be closed) and other approved devices. Many of these devices are appropriate for site control at and near special events as well.

Management Strategies/Operations

The *MUTCD* also provides guidance on appropriate management strategies for various types of work zone lane closures, indicating such things as minimum cone taper lengths, locations of advance signing, and proper layout of channelization devices. These guidelines are dependent upon the duration of the scheduled work activity, the time of day, the roadway type, and the operating conditions.[32]

Major freeway reconstruction activities and special events place tremendous demands upon the freeway and surrounding surface street system to accommodate unusual traffic patterns. The planning horizon available for these types of activities allows for a number of demand management strategies to be implemented within the freeway corridor, if appropriate. Table 8-12 summarizes some of the impact mitigation strategies that have been implemented at a sample of major reconstruction projects and special events nationwide.[12]

8.4 LESSONS LEARNED

WATER ACCESS

Not all incident response and management activities must involve a high degree of technical complexity in order to provide a significant benefit to the public and to incident response agencies. Often, simple low-tech solutions can offer substantial benefits to a response agency or to the public.

In San Antonio, for example, officials have commented about problems encountered when trying to combat vehicle fires on elevated freeway sections. Simple standpipe assemblies that provide a fire department with access to a fire hydrant from an

Table 8-12. Traffic Management Actions During Major Freeway Construction and Special Events. [12]

Types of Actions	Actions Implemented
Actions to Improve Alternative Routes	Traffic signal timing adjustments Traffic signal equipment improvements Left-turn restrictions at critical locations Parking restrictions Police control of critical intersections Reversible lanes Implementation of alternating, one-way pairs Intersection widening and channelization Resurfacing and other pavement repairs Signing and lighting improvements Use of real-time information systems to encourage diversion Constructing temporary pedestrian overpasses at vehicle/pedestrian conflict points
Actions to Improve HOV and Transit Utilization	New or expanded commuter rail service Expanded rapid transit service New or expanded bus service Implementation of HOV only ramps and lanes New or expanded park-and-ride lots New or expanded ridesharing programs Eliminating or moving single occupant vehicle parking locations Increasing parking fees for single occupant vehicles
Actions to Improve Public Understanding, Cooperation, and Acceptance	Traditional P-R tools (press conferences, new releases) Special publications Toll-free hotlines Highway advisory radio systems Special freeway signing Employment of an ombudsman Rescheduling event start and end times to avoid peak periods Developing agreements with trucking agencies to avoid peak period deliveries

elevated freeway section can greatly facilitate their response capabilities. In other jurisdictions nationwide, the DOTs and fire departments have worked together to mark the location of hydrants, and to provide access doors through noise barrier walls to connect to hydrants on the other side.

SERVICE PATROL CHARGES FOR ASSISTANCE

A question often arises relative to courtesy patrol operations about whether or not to charge motorists for "service supplies" (typically gasoline) that are provided. For years, Chicago requested that motorists pay for gasoline provided by its Emergency Traffic Patrol; however, auditors recently found it more cost effective for them to give the fuel away than to try and recoup these expenses.[30] Likewise, a recent analysis of the Los Angeles service patrol system concluded that charging motorists for gasoline by that agency would cost more to administrate than would be recouped through revenues.[17] The differences in these two operational philosophies could be due to the different administrative structures utilized by each for managing service patrol operations (see the Examples section at the end of this module for a description of various service patrol operations).

TRAINING

A component of incident management that is sometimes overlooked is the need for proper training and retraining. This is particularly important for dealing with larger-scale incidents that involve multiple agencies. Because these types of incidents may occur very infrequently (hopefully), it is difficult to test out the processes and procedures that have been developed for these situations beforehand. In addition to working out the "kinks" of coordination, training also helps to demonstrate to incident response personnel the need for, and benefits of, some of the procedures and protocols that have been established. Methods of training for such large events include the following:[19]

- Workshops.
- Conferences.
- Instructional videos.
- Mock disaster exercises.
- Post-incident debriefings.

It appears that the liability an agency bears for failing to properly train its personnel to react to normal day-to-day situations is slowly being extended in the courts to cases involving more dynamic emergency situations.[34] In the future, if it can be proven that adequate training to handle emergency situations could have prevented injuries or damages, an agency may be forced to assume at least some liability for failing to provide that training.

Interagency training is also important. Fire departments, for instance, can provide training for police and DOT personnel in hazardous materials identification and response.

8.5 EXAMPLES OF FREEWAY SERVICE PATROLS

CHICAGO EMERGENCY TRAFFIC PATROL

The Chicago Emergency Traffic Patrol (ETP), also known as the "Minutemen" patrol, began operations in 1961 as an Illinois DOT effort. This program is an example of a completely State-DOT-operated project. The ETP consists of 58 drivers, 35 patrol vehicles, 11 light-duty four-wheel drive vehicles, three heavy-duty tow rigs, a crash crane, a tractor-retriever, a sand spreader, heavy rescue and extricator truck, and a hazmat response trailer.[17]

The ETP patrols 125 centerline kilometers of freeway 24 hours per day, 365 days per year.

Annually, the patrol assists at over 100,000 incidents.[30] The patrol drivers are trained to handle most incidents. They also provide fuel, water, and minor repairs, and can tow or push vehicles to the shoulder or frontage road.

Overall, the Chicago freeway management system (including the ETP and other incident management activities) have been estimated to provide a benefit-cost ratio of 17:1.[30] The ETP alone is estimated to provide an annual 9.5 million hour reduction in delay, for a $95 million delay savings per year.[35] The ETP costs approximately $5.5 million to operate annually, equivalent to about $55 per assist. Funding for the ETP comes from State taxes on motor fuel. Additional funding is being sought to equip ETP vehicles with automatic vehicle locating (AVL) systems.

MINNEAPOLIS HIGHWAY HELPER PROGRAM

The Minnesota Highway Helper program began in December 1987. The program initiated out of the Minnesota Department of Transportation's District Maintenance office, and was operated only during peak periods. In March 1993, the traffic management center in Minneapolis took over management of the program and increased operations to daytime off-peak hours as well.[17]

As of 1994, the Highway Helper program patrolled 109 centerline kilometers (68 centerline miles) of freeway. Seven pick-up trucks logged approximately 125,000 kilometers (78,000 miles) during the year, assisting 13,000 motorists.[17] According to program logbooks, about 34 percent of stranded motorists were assisted within 5 minutes of disablement. Meanwhile, 26 percent of motorists were assisted within 5 to 10 minutes of disablement, and 20 percent were assisted within 10 to 20 minutes.[17]

Annual operating costs for the program are approximately $550,000.[17] A 1994 evaluation of the program yielded a benefit/cost ratio of 2.3 to 1.[37]

LOS ANGELES FREEWAY SERVICE PATROL

Los Angeles is home to the largest freeway service patrol in the U.S. The patrol was initiated in 1991, and consists of 144 tow trucks patrolling 610 centerline kilometers of freeway. The Los Angeles program contracts out the actual patrol activities to private wrecker services. Currently, 20 different towing companies are participating in the program.[17] The program is funded and administered through a cooperative effort between the following agencies:

- California Department of Transportation (jointly responsible for the overall supervision of the program).
- California Highway Patrol (jointly responsible for overall supervision and for dispatching patrols).
- Metropolitan Transit Authority (responsible for the contractual arrangements with the private towing companies).

Patrol routes and vehicle frequency were designed so that a patrol vehicle passes each point on the route every 15 minutes, which was recently verified by an evaluation of average response times.[17] Also, patrol vehicles are outfitted with automatic vehicle locating (AVL) systems that allow dispatchers to know the exact location of patrol vehicles and whether or not the unit is available to respond to a call.

On average, the Los Angeles service patrol program assists at over 750 incidents per day, equaling almost 250,000 assists per

year. Funding for the Los Angeles service patrol program comes from state funds and a 0.5 percent local sales tax. Costs of the program are estimated to be between $15 million and $20 million, yielding a per assist cost of approximately $80.[35]

HOUSTON MOTORIST ASSISTANCE PATROL

The service patrol in the Houston region is known as the Motorist Assistance Patrol (MAP). This program began operations in its current form in 1989, and is an example of a cooperative venture between several public agencies and private-sector companies. Partners in the Houston MAP include the following:[36]

- Texas Department of Transportation (supervising dispatch and operations, providing space and personnel in the interim control center, partially supporting MAP personnel salaries, supporting necessary equipment purchases).
- Metropolitan Transit Authority (supervising operations, partially supporting MAP personnel salaries, supporting necessary equipment purchases).
- Harris County Sheriff's Department (providing deputies to operate the MAP vans, maintaining and operating the MAP vans).
- Houston Automobile Dealers Association (providing MAP vans).
- Houston Cellular (providing cellular phones, air time, and a toll-free number for motorists [CALLMAP]).

The program operates between 6 a.m. and 10 p.m. weekdays, patrolling 240 centerline kilometers of freeway. In 1993, Houston MAP assisted at over 25,000 incidents.[36] It is estimated that the MAP program saves 2,376,000 hours of delay on Houston freeways annually. The current funding for Houston MAP is $1.4 million per year, yielding a per assist cost of about $56.[17]

SAMARITANIA, INC.

The Samaritan program, operated by Samaritania, Inc., is an example of a privately financed service patrol. Samaritania, Inc., based in Franklin, Massachusetts, has established service patrols in the following locations:[12, 17]

- Albany.
- Boston.
- Cincinnati.
- Hartford.
- Indianapolis.
- Philadelphia.
- Providence.
- Washington, DC
- White Plains.
- Worchester.

Samaritania, Inc. seeks out large metropolitan areas, typically with populations greater than 100,000, that have significant incident traffic problems on area freeways. Potential corporate sponsors with annual receipts of over $200 million are contacted about sponsoring a service patrol. Samaritania, Inc. demonstrates the marketing and public relations benefits of such sponsorship. Once a sponsor is

obtained, their logo is prominently displayed on the side of the service patrol van.

Local personnel in each area operate the patrols, and are trained in basic emergency medical procedures, traffic pattern analysis and reporting, and emergency vehicle repairs. If the motorist needs more assistance than the patrol driver can provide, the patrol driver will contact the appropriate organizations for the motorist.

Samaritania reports that 75 to 90 percent of all disabled vehicles encountered on patrol are returned to the freeway at no cost to the motorist.[17] Some of the patrols also provide traffic information to local radio and television stations for dissemination to motorists.

8.6 REFERENCES

1. Reiss, R.A., and Dunn, W.M. *Freeway Incident Management Handbook*. Report No. FHWA-SA-91-056. FHWA, U.S. Department of Transportation, Washington, DC, July 1991.

2. Schrank, D.L.,Turner, S.M., and Lomax, T.J. *Estimates of Urban Roadway Congestion -- 1990.* Report FHWA/TX-90/1131-5. Texas Transportation Institute, College Station, TX, March 1993.

3. Robinson, J. What's New in Freeway Incident Detection and Response. *Technical Papers from ITE Specialty Conferences 1990, 1989, and 1988.* Institute of Transportation Engineers, Washington, DC, pp 244-257.

4. Roper, D.H. *Freeway Incident Management*. NCHRP Synthesis of Highway Practice No. 156. TRB, National Research Council, Washington, DC, December 1990.

5. Mannering, F., Hallenbeck, M., and Koehne, J. *Framework For Developing Incident Management Systems*. Report No. WA-RD 224.1 Washington State Transportation Center, University of Washington, Seattle, WA, August 1991.

6. Sumner, R. et al. *Freeway Management Handbook*. Report No. FHWA-SA-88-026. FHWA, U.S. Department of Transportation, Washington, DC, May 1983.

7. Skabardonis, A., Petty, K., Noeimi, H., Rydzewski, D., and Varaiya, P.P. I-880 Field Experiment: Data-Base Development and Incident Delay Estimation Procedures. In *Transportation Research Record 1554.* TRB, National Research Council, Washington, DC, 1996, pp 204-212.

8. Goolsby, M.E. Influence of Incidents on Freeway Quality of Service. In *Highway Research Record 349.* TRB, National Research Council, Washington, DC, 1971, pp 41-46.

9. Ogden, M.A., Mounce, J.M., Middleton, D.R., and Ullman, G.L. *Traffic Management in Response to Major Freeway Incidents, Volume I.* Report No. FHWA/TX-94/1345-2F. Texas Transportation Institute, College Station, TX, August 1994.

10. Gordon, R.L., Reiss, R.A., Haenel, H., Case, E.R., French, R.L., Mohaddes, A,. and Wolcott, R. *Traffic Control Systems Handbook*. Report No. FHWA-SA-95-032. FHWA, U.S. Department of Transportation, Washington, DC, February 1996.

11. *Highway Capacity Manual*. Special Report 209 (3rd Edition). TRB, National Research Council, Washington, DC, 1994.

12. Dudek, C.L. and Ullman, G.L. *Freeway Corridor Management*. NCHRP Synthesis of Highway Practice No. 177. TRB, National Research Council, Washington, DC, March 1992.

13. Cyra, S.J. and Stilp, K.A. *Developing the Southeastern Wisconsin Incident Management Program: An Application of the ITS Planning Process*. Presented at the 3rd ITS World Congress, Orlando, FL, October 1996.

14. *Building the ITI: Putting the National Architecture into Action*. Report No. FHWA-JPO-96-011. FHWA, U.S. Department of Transportation, Washington, DC, April 1996.

15. McLean, C.H. Cellular Phones -- A Key Traffic Management Component. *Compendium of Technical Papers*, Institute of Transportation Engineers, Washington, DC, 1991.

16. Implementation Plan Guidance (23 CFR 655.409). Federal-Aid Policy Guide, FHWA, U.S. Department of Transportation, Washington, DC, December 7, 1994, Transmittal 12.

17. Fenno, D., Ogden, M., Morris, M., Thompson, B., and Benz, R. Guidelines for Establishing Freeway Service Patrols. *Compendium of Technical Papers*, 66th Annual Meeting of the Institute of Transportation Engineers, Minneapolis, MN, September 1996, pp 54-61.

18. Petty, K.F., A. Skabardonis, and P.P. Varaiya. Methodology for Estimating the Impacts of Incident Management Measures. Presented at the 3rd ITS World Congress, Orlando, FL, October 1996.

19. Ullman, G.L. and Trout, N.D. Traffic Management for Major Emergencies. *Compendium of Technical Papers*, Institute of Transportation Engineers, Washington, DC, 1994, pp 293-297.

20. Wilshire, R., Black, R., Grochoske, R. and Higginbotham, J. Traffic Control Systems Handbook. Report No. FHWA-IP-85-11. FHWA, U.S. Department of Transportation, Washington, DC, April 1985.

21. Payne, H.J., Helfenbein, E.D., and Knobel, H.C. Development and Testing of Incident Detection Algorithms, Volume 2: Research Methodology and Detailed Results. Report No. FHWA-RD-76-20. FHWA, U.S. Department of Transportation, Washington, DC, April 1976.

22. Balke, K.N. An Evaluation of Existing Incident Detection Algorithms. Report No. FHWA/TX-93/1232-20. Texas Transportation Institute, College Station, TX, November 1993.

23. Balke, K.N., G.L. Ullman, W.R. McCasland, C.E. Mountain, and C.L. Dudek. *Benefits of Real-Time Travel Information In Houston, Texas*. Report No. SWUTC/95/60010-1. Southwest Region University Transportation Center, Texas Transportation Institute, College Station, TX, January 1995.

24. Parkany, E. and Bernstein, D. Design of Incident Detection Algorithms Using Vehicle-to-Roadside Communication Sensors. In *Transportation Research Record 1494*. TRB, National Research Council, Washington, DC, 1995, pp 67-74.

25. Jacobson, L.N., Legg, B., and O'Brien, A.J. Incident Management Using Total Stations. In *Transportation Research Record 1376*. TRB, National Research Council, Washington, DC, 1992, pp 64-70.

26. Agent, K.R., Deacon, J.A., Pigman, J.G., and Stamatiadis, N. Evaluation of Advanced Surveying Technology for Accident Investigation. In *Transportation Research Record 1485*. TRB, National Research Council, Washington, DC, 1995, pp.124-133.

27. Mounce, J.M., Stokes, R.W., Brackett, R.Q., and McCasland, W.R. *Guidelines for Utilization of Police Officers in Traffic Control and Enforcement on Urban Freeways*. Report No. FHWA/TX-86/75+410-7F. Texas Transportation Institute, College Station, TX, September 1986.

28. Promotional literature for BarrierGate™ system. Energy Absorption Systems, Inc., Chicago, IL. 1993.

29. Dudek, C.L., McCasland, W.R., and Burns, E.N. *Promotional Issues Related to Off-Site Accident Investigation*. Report No. FHWA/RD-87/036. FHWA, U.S. Department of Transportation, Washington, DC, July 1987.

30. McDermott, J.M., Mclean,C.H., and Smith, A.T. Three Decades of Progress: Freeway Traffic Management in Illinois. *ITE Journal*, Vol. 62, No. 3, March 1992, pp 27-31.

31. Ritchie, S.G., and Prosser, N.A. Real-Time Expert System Approach to Freeway Incident Management. In *Transportation Research Record 1320*. TRB, National Research Council, Washington, DC, 1991, pp 7-16.

32. Albert, S., Siegfried, R., and Levine, S. Application of Geographic Information Systems to Houston Freeway Incident Management Program. *Compendium of Technical Papers*, Institute of Transportation Engineers, Washington, DC, 1990, pp 131-132.

33. Part VI, Standards and Guides for Traffic Controls for Street and Highway Construction, Maintenance, Utility, and Incident Management Operations. *Manual on Uniform Traffic Control Devices*. FHWA, U.S. Department of Transportation, Washington, DC, September 1993.

34. Knight, B.L. General Liability Issues for Local Government in Hurricanes. Presented at the National Hurricane Conference, Houston, TX, April 1990.

35. Nowlin, R.E. Institutional Concerns Regarding the Implementation of State Managed Service Patrols. In *Graduate Student Papers on Advanced Surface Transportation Systems*. Report No. SWUTC/94/72194-2. Southwest Region University Transportation Center, Texas Transportation Institute, College Station, TX, August 1994.

36. Wohlschlaeger, S.D. and K.N. Balke. Incident Response and Clearance in the State of Texas: Case Studies of Four Motorist Assistance Patrols. Report No. FHWA/TX-92/1232-15. Texas Transportation Institute, College Station, TX, November 1992.

37. *Highway Helper Summary Report, Twin Cities Metro Area.* Report No. TMC 07450-0394. Minnesota Department of Transportation, Minneapolis, MN. March 1994.

MODULE 9. COMMUNICATIONS

TABLE OF CONTENTS

9.1 INTRODUCTION 9-3
MODULE OBJECTIVES 9-4
MODULE SCOPE 9-4

9.2 DECISION PROCESS 9-4
PROBLEM IDENTIFICATION 9-5
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 9-8
ESTABLISH SYSTEM GOALS AND OBJECTIVES 9-9
ESTABLISH PERFORMANCE CRITERIA 9-9
FUNCTIONAL REQUIREMENTS 9-10
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS,
AND INFORMATION FLOWS 9-12
Logical 9-12
Physical 9-12
Central 9-14
Distributed 9-16
Trunking 9-16
Backbone 9-16
Multimedia 9-18
IDENTIFY AND SCREEN TECHNOLOGIES 9-18
IMPLEMENTATION 9-23
National ITS Standards 9-23
Installation 9-26

9.3 TECHNIQUES AND TECHNOLOGIES 9-26
LAND-LINES 9-26
Twisted-Wire Pairs 9-27
Coaxial Cable 9-31
Fiber Optic 9-35
WIRELESS 9-39
Area Radio Network 9-39
Microwave 9-42
Spread Spectrum Radio 9-43
Cellular Radio 9-45
Packet Radio 9-48
Satellite Transmission 9-50

9.4 LESSONS LEARNED 9-51
DESIGN PRINCIPLES 9-51
OWNED VERSUS LEASED COMMUNICATIONS 9-53

NTCIP 9-56
SHARED RESOURCE PROJECTS 9-59
TRANSIENT PROTECTION 9-63
COMMUNICATIONS WITHIN THE CONTROL CENTER 9-64

9.5 REFERENCES 9-65

MODULE 9. COMMUNICATIONS

Figure 9-1. Fiber Optic Cable Installation in Fort Worth, TX.

9.1 INTRODUCTION

The communications subsystem is one of the most critical and expensive elements of a freeway management system. The function of the communications system is to link the devices in the field with the operating personnel in the control center. It is the backbone of the entire freeway management system because it allows the following types of information to be transferred from the field components and the control center: [1]

- Commands to the various field components.
- Data from the system detectors and sensors.
- Status checks of field equipment to detect malfunctions.

Typically, the communications system accounts for approximately 25 percent of the total capital costs associated with a freeway management system. Failure to plan, design, and properly maintain the communications system guarantees the entire freeway management system will have problems in achieving its operational goals and objectives. Poor planning and installation of the communication system can also result in high operating and maintenance costs. To avoid problems with the communications system, system planners, designers, and operators must be familiar with existing and emerging communication technologies and system architectures.

Three general types of communications systems are used in a freeway management system:

- A communications system allowing the transfer of surveillance data and control commands between field devices and the control center.

- A communications system allowing the computers within the control center to transfer data and displays between each other.

- A communications system allowing the operators in the control center to exchange data and information with other operators and field personnel outside the control center.

This module focuses on the first type of communications system (i.e., the communications system needed to transfer surveillance information and data, and control commands between the field devices and the control center). For information on the other types of communications systems, the reader is referred to Module 10.

MODULE OBJECTIVES

The objectives of this module are as follows:

- To outline the process that can be followed in planning and designing a new freeway management system or updating or modifying an existing freeway management system.

- To highlight the different types of communication technologies that are commonly used in a freeway management system.

- To provide insight into the issues associated with planning, designing, constructing, operating, and maintaining a communications system in a freeway management system.

MODULE SCOPE

The focus of this module is on the decision process for planning, designing, and selecting a communications system for a freeway management system. It is intended to help system designers and traffic engineers make informed decisions about communications systems. It is not intended to provide detailed information about the operation of different communications systems or technologies. For more detailed information about the different communications technologies, the reader should consult the *Communications Handbook for Traffic Control Devices* and the *Traffic Control Systems Handbook*.[1,2]

9.2 DECISION PROCESS

As with the planning and design of the other functional elements of a freeway management system, a systems engineering approach can also be applied to the planning, design, and use of a communications system

for a freeway management system. An overview of the systems engineering approach is presented in **Module 2**. The information in this section highlights some important issues that need to be discussed, evaluated, and decided upon as part of the systems engineering process. The user is encouraged to consult other reference materials, such as the *Communications Handbook for Traffic Control Systems* for a detailed discussion of the technical aspects associated with planning and designing a communications system for traffic control applications. [1]

PROBLEM IDENTIFICATION

The first step in planning and designing a communications system for a freeway management system is to identify the physical, institutional, and other factors that might affect the design of the communications system. System designers and planners need to collect information on the following attributes and characteristics about each of the other elements and subsystems that will be supported by the communications system. The types of information that should be gathered include the following:

- The type of data that will be transmitted by each field device.
- The content and format of the data transmitted.
- How much processing of data will occur in the field devices and where this processing will occur.
- The total number of field devices and where they will be situated.
- The type and presence of existing communications technologies near the freeway management system.
- The relative importance of the information and how a loss of communication from each device would affect the overall operation of the system.

Table 9-1 summarizes some physical, institutional, and other issues that might affect the planning and design of the communications system for a freeway management system.

One question that must be addressed in planning and designing a communications system is the type of data that will be transmitted over the communications medium: voice, data, video, or a combination of all three. A communications medium is the pathway over which a signal is transmitted (i.e., twisted wire pair, coaxial cable, fiber, radio, etc.). "Voice" is the chosen term for audio signals, since the impetus for modern communications systems was the need to transmit vocal messages over the telephone. While some freeway management applications, such as Highway Advisory Radio, still depend on voice communications, data is the most general type of information transmitted in a communications system in a freeway management system. Digital data is the most common form of transmitted information in today's communications systems. With digital data, information is represented by a combination of bits, where a bit can assume one of only two states: "0" or "1." Digital data can be physically transmitted as voltage differences between pairs of wires, light intensities in an optical fiber, frequencies in a transmission line, or phases of a radio wave. Video is animated imagery and represents the most demanding form of data transmission.

The requirements for sending video signals differ significantly from data communications requirements. Systems with

Table 9-1. Resources and Constraints Considerations in Communication Systems. [3]

PHYSICAL FACTORS

- Number and location of field cabinets to be served, and location of control center
- Location and type of existing communication facilities
  - Cable
  - Conduit
  - Pole lines
- Nature of terrain to be trenched and backfilled (conduit installation)
  - Roadway
  - Sidewalk
  - Structure
  - Railroad
  - Soil
- Nature of terrain to be spanned (aerial installation)
  - Waterway
  - Railroad
  - Elevated roadways
- Location of utility equipment and underground structures that may interfere with installation
- Air-path propagation characteristics
  - Trees
  - Hills
  - Buildings
- Climactic conditions affecting communications
  - Temperature extremes
  - Moisture
  - Lightning
  - Ice storms
- Planned or current construction activities
  - New conduit installed
  - Existing conduit removed/relocated

INSTITUTIONAL ISSUES

- Rights-of-way
- Franchise agreements between utility companies and government
  - Right of the agency to use utility conduits and pole lines
  - Responsibility of clearing ducts and utility adjustments
- Franchise agreements with CATV for government use
- Telephone company tariffs and policies
- Other agreements (formal and informal)
- National and local codes (National Electrical Code)
- Federal Communications Commission (FCC) rules and regulations
- Restrictions on work procedures and traffic maintenance
- Rules regarding different types of conduit, overhead cabling, conduit installation, junction boxes, antenna structures, etc.

OTHER

- Personnel and skill levels for communications maintenance
- Other maintenance resources (budget, contract, etc.)
- Vandalism threat
- Presence of contractors in area with skill/experience in installation of communication networks

video components generally require a greater bandwidth, and if transmitted digitally, higher signal rates. Bandwidth is the range of signal frequencies that a communications medium (or channel) will carry without excessive loss of signal strength. The bandwidth requirements for video communications vary depending on the type of transmission mode. Generally, three modes are available for transmitting video images back to a control center:

- Full motion analog video.
- Freeze frame/slow scan.
- Compressed video.

Each of these techniques for transmitting videos images to a control center is discussed in more detail in the **Techniques and Technologies** section below.

Another critical piece of information that needs to be determined before a communications system can be designed is the type of data that will be transmitted. Examples of the type of data carried by the communications system for freeway management purposes include the following:

- Volume, speed, and occupancy from field detectors.
- Alphanumeric messages for DMS displays.
- Codes to select/implement stored messages or control strategies.
- Device status and malfunction reports.
- Weather/environmental sensor data.
- Video control.

Another issue that affects the design of the communications system is how much processing of the data will occur in the field. Generally, those systems where most of the processing occurs at a central location have greater communications needs than those in which the processing of the data is distributed between the field devices and the control center. Generally, those systems where most of the data processing occurs at a central location require shorter polling cycles (the time required to communicate with all the field devices on a communications line once). Those systems that distribute the data processing generally have long polling cycles and more robust field storage capabilities. The advantages and disadvantages of having longer polling cycles are listed in table 9-2.

The total number of field devices, and their location also influences the overall design of the communications system. Generally, the more devices that need to transfer data and information within a freeway management system, the greater the communications requirements. Also, the type of communications architecture can change dramatically if the devices are dispersed over a wide geographic region. For example, using a wireless communications architecture may be more economical where few field devices are widely dispersed throughout an area. If, on the other hand, all the devices are found in close proximity, a wire-type of architecture may be more suitable.

Another important factor and/or constraint that may influence the architecture design and medium selection is the presence of existing communications systems. Often, in-place, usable cable, conduit, and field equipment can be used in a new design. Since the conduit/cable is often the major cost in a system, a plan to use existing equipment can become the least expensive communications alternative.

Table 9-2. Advantages and Disadvantages of Longer Polling Periods. [1]

Advantages	Disadvantages
• Usually results in lower data rate requirement for each field unit to perform the same functions. • Enables more field units to be serviced by a communication channel of given bandwidth or data rate. • Enables higher resolution surveillance data to be communicated. • Avoids use of data overflow schemes.	• Requires more extensive processing of data for the field devices. • Requires large field data base. • Delays transmission of data.

The relative importance of the information (and the impact the loss of the information has on the operation of the system) can have a significant impact on the type of transmission medium and architecture used in a system. Through policy, some jurisdictions may place certain constraints on the types of communications media or architectures possible. Examples of these constraints include the following: [1]

- Rejection of radio-based technologies.
- Preference for owned communications media over leased media.
- Avoidance of communication designs containing points of single failure that disrupt communication to many devices.

Maintenance of the system is also an important issue that needs to be considered in the initial planning and design of a communications system. The increasing sophistication of many newer communications media may require a level of maintenance experience and capability beyond what many local agencies may be willing to supply. Maintenance issues that need to be considered during the initial planning and design phases include the following:

- Who will maintain the communications (i.e., in-house forces, contract forces, or a combination of both)?
- How will maintenance of the communications system be funded?
- Can a simpler technology that can be maintained using in-house forces be used and still provide the same functionality as a more sophisticated technology?
- What kind of leased-line service can be used to reduce the need for in-house maintenance forces?

IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING

A final issue is what other organizations/ agencies will want, or if they would be willing to share resources in constructing and maintaining a communications system. Traditionally, transportation agencies have wanted to own and operate their own communications systems for traffic management purposes; however, the cost of

installing and maintaining these systems can be substantial. As a result, many agencies are looking for methods to share the cost of installing, operating, and maintaining their communications systems. Some agencies are discovering that leasing their communication needs is the most cost-effective method. System planners and designers should attempt to identify other organizations and agencies that also need to install communications systems as potential cost sharing partners. However, along with sharing resources among agencies and organizations comes the need for building coalitions and consensus on the design of the system. Potential agencies and organizations that should be considered in the initial planning and design of a communications system include the following: [3]

- Other TMCs in the area.
- Media.
- Emergency service providers.
- Private communications providers.

The maintenance responsibilities of these agencies should also be defined during the initial planning and design of the system.

ESTABLISH SYSTEM GOALS AND OBJECTIVES

The next step is to establish goals and objectives for the system. The goals and objectives of a communications system are influenced by the goals and objectives of the other functions and systems included in the freeway management system. For example, an objective of an incident management subsystem might be to detect all incidents within two minutes of their occurrence. The goal of the communications system should be to make sure the control center has all the data it needs to detect incidents on the freeway two minutes after they occur. In this example, a communications media and architecture that results in a five-minute polling cycle of all the field detectors does not allow the incident management system to achieve its stated objective.

Agencies should develop both technical and nontechnical goals and objectives for evaluating the performance of communication alternatives. Technical goals and objectives relate to criteria that establish the effectiveness and efficiency of the communications system. Examples of nontechnical goals and objectives include the following: [3]

- The ability to provide and/or the potential for providing intra- and inter-agency data sharing.
- The potential for expanding the communication system to permit future growth and adding functions to the freeway management system.
- How much redundancy or reliability is built into the system.
- The life-span of the system (i.e., whether the system is permanent or temporary), and the potential for installing the system in phases.

ESTABLISH PERFORMANCE CRITERIA

Establishing performance criteria and measures of effectiveness is extremely important in the initial planning stages of a communications system. The criteria can be used not only to evaluate different design alternatives, but also to measure the reliability and expandability of the system. Agencies need to decide up front what represents a “good” level of operation for the communications system.

Most agencies use system reliability as the primary measure of performance for their communications system. System reliability is measured in two ways:

- Transmission errors.
- System uptime.

Transmission errors are primarily caused by noise in the communications system. Noise is any unwanted signal or disturbance of a signal that interferes with or distorts the original communication signal. Noise causes the receiver to produce incorrect outputs and errors in the bit stream that transmits the information. Noise can be caused by many factors including the following: [1]

- Temperature extremes.
- Natural radio or other electronic signals (such as lightning or cosmic/solar bursts).
- Human-made electrical signals (such as motors, car ignitions, power lines, etc.).
- Signals from another communication channel (i.e., crosstalk).

Transmission errors are measured in Bit Error Rate (BER). BER is the ratio of incorrectly transmitted bits to correctly transmitted bits. Values of about 10^{-6} or better for end-to-end communications represent an acceptable BER for most computer and traffic control communications systems. [1]

Most systems have processes to detect errors in communication signals. Table 9-3 describes the error detection techniques commonly used in traffic control systems. When an error in a message is detected, most systems either ignore the message until the next polling cycle, or request that the device retransmit the message. In either case, transmission errors reduce the effective throughput of the communications channel.

System uptime is also another common measure of performance of a communications system. System uptime represents the portion of the normal operating time of the system during which a link or the entire communications system is functioning properly. System uptime can be used to identify problem locations that might need special communication considerations (e.g., a link susceptible to outages due to construction activities or environmental conditions).

It should be noted that bandwidth is NOT a measure of system performance. Bandwidth is the range of signal frequencies that a medium or channel will respond to, or carry without excessive attenuation. [1] It is a measure of the characteristics of a system and does not tell anything about the performance of the system.

FUNCTIONAL REQUIREMENTS

The purpose of a communications system is to transfer information and data from one freeway management function to another with no loss in accuracy. Therefore, the functional requirements of the communications system need to describe how it can help the other elements of the system achieve their goals and objectives. For example, a functional requirement of a communications system might be to ensure that the volume, occupancy, and speed data are transported to the incident detection algorithm every 20 seconds. Likewise, if an objective of an incident management system is to ensure that appropriate aid is dispatched to an incident scene in response to any detected incident, a functional requirement of the communications system might be to have video images transmitted

Table 9-3. Commonly Used Error Detection Techniques. [1]

Techniques	Description
Parity (also known as vertical parity)	An additional bit is added to each data byte or character. The sum of the 1s in the byte and the additional bit must be an odd or even number as specified. This technique detects an odd number of bit errors in the byte.
Longitudinal Redundancy Check	An additional byte is provided after an entire message or portion of a message (block). A bit in the new byte is computed from the corresponding bit in each data byte in a way similar to the parity check. An odd number of bit errors is again detected. When used in conjunction with parity this is a powerful technique.
Checksum	An additional byte or character is added at the end of the message or block. An algorithm is used which computes the checksum byte as a function of the message bytes. The receiving station performs a similar computation and determines whether the checksum byte is consistent with the received data.
Cyclic Redundancy Code (CRC)	An additional two or more bytes are added to the message or block. Algorithms are used to compute these bytes which provide protection, particularly against bursts of errors.
Repeat Transmission	The entire message is repeated. At the receiving station the messages are compared and an error is detected if they are not identical.

directly to emergency service providers and police dispatchers.

The communications functional requirements need to be developed for every element in the system. The information that should be specified in the functional requirements for a communications system includes the following:

- The type of data required by each element in the system (i.e., voice, data, video, or a combination of all three of these).
- The need for one-way versus two-way communication between the element and the control center.
- The type of messages being transmitted (i.e., traffic flow data, text messages, video images, device commands, etc.).
- The frequency at which information is being transmitted or required (i.e., continuously, once every second, once every 20 seconds, once every minute, etc.).
- The desired level of control (e.g., monitoring traffic conditions vs. full operational control through field devices).
- The importance of the information (i.e., critical, important, non-critical).

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

Once the functional requirements of the communications system have been identified, the functional relationships, data requirements, and information flows also need to be identified. Generally, these relationships and requirements can be grouped in two categories: logical and physical.

Logical

The logical relationships and requirements of a communications system defines what information flows from one element to the next in the system. Generally, they can be represented by a series of data flow diagrams and process specifications that illustrate what type of information is used by each element in the freeway management system. Identifying the logical requirements and relationships requires that all the functional requirements of the system be broken down into smaller and smaller subfunctions until each process in the system can be specified. Process specifications describe the following in detail:

- The type of information needed to execute a process.
- How the information is processed or manipulated.
- What information or action is produced by the process.

The data flows is constructed by combining the data flows of the various elements (e.g., the incident management subsystem, the ramp control subsystem, the information dissemination subsystem, etc.) that are being installed in the freeway management system. The data and information flowing out of the elements are what needs to be transported by the communications system. Figure 9-2 illustrates, at a high level, a logical architecture for a typical freeway management system.

Once the logical relationships have been mapped out, the transmission mode (i.e., the direction of flow over a communications channel) can readily be determined. The following are the transmission modes commonly used in communications systems:

- Simplex.
- Half Duplex.
- Full Duplex.

Table 9-4 summarizes the characteristics and applications of each of these transmission modes. Most systems use either a half duplex or a full duplex transmission mode so that the status of the field devices can be checked and transmission errors corrected.

Physical

Once the logical relationships of the communications system has been identified, the physical design can be established. The physical design highlights how information flows from one element of the system to the next. It forms the basis for selecting the transmission media used to connect the different elements in the system. Factors affecting the selection of the physical architecture include the following:

- The number and location of the field devices.
- Distance between field devices and control center.

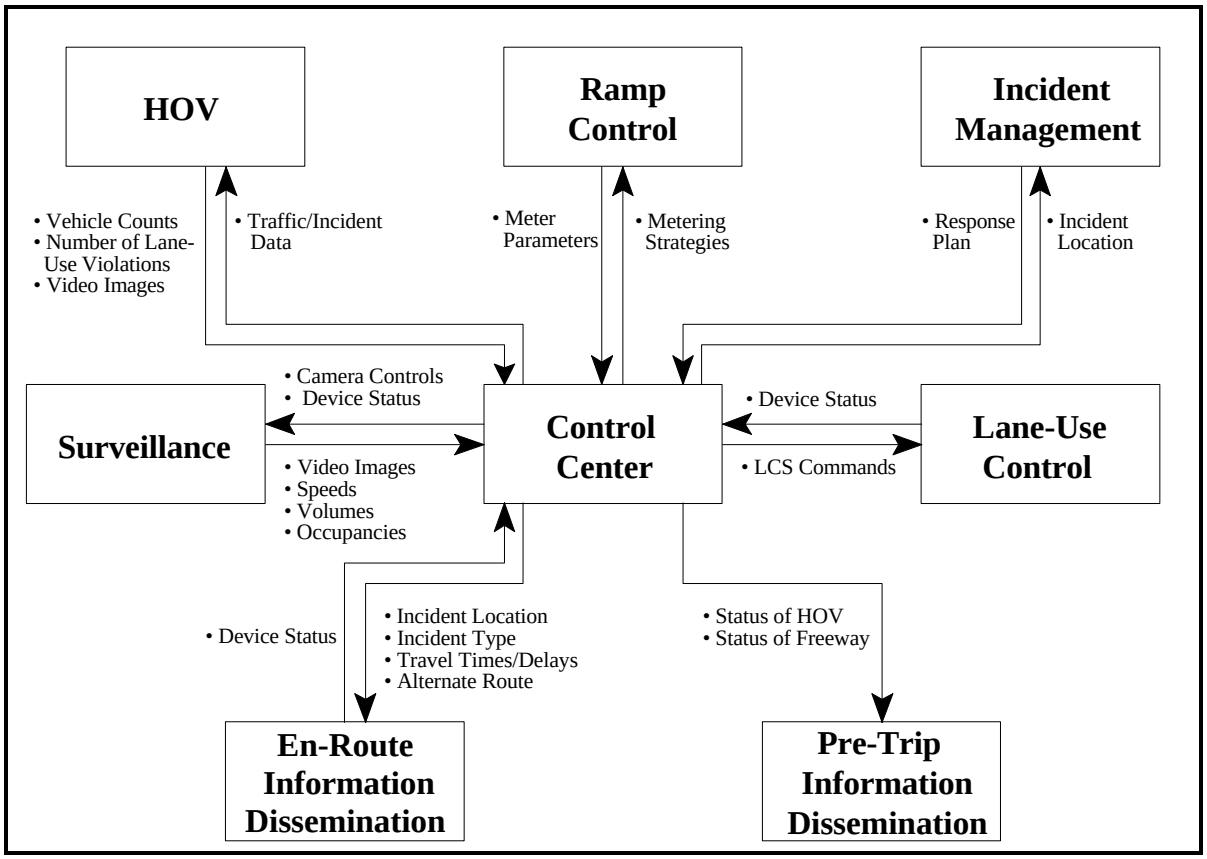


Figure 9-2. High Level Logical Architecture for a Communications System.

- Type and frequency of data being transmitted.
- Availability of right-of-way for placing communication media.
- Proximity of existing communications systems, including private communications systems.
- Presence of geologic or manufactured features that might prevent wireline connections.

The type and importance of the data have an impact on the design of the system and the type of error correction technique used. For example, with a centralized system using once-per-second communications, a single undetected error in detector data represents only one second's worth of sampling time. To discard this information would not be critical, since new information will be provided during the next polling cycle. If, on the other hand, the error occurred in a command to operate a dynamic message sign or other traffic control device, it might cause an unsafe condition on the freeway, or reduce credibility of the system. A distributed system having longer periods between updates (typically one minute for detector data, and possibly hours for controller information) requires more elaborate error correction techniques to avoid losing information caused by communication errors.

The physical architecture of most communications systems used in traffic management applications fall into the following five major categories: [1]

Table 9-4. Characteristics of Transmission Modes. [1]

Mode	Data Flow Direction	Characteristics	Comments
Simplex	Data flow in one direction only.	• Does not provide verification that data were received and acted upon. • Does not provide answer-back, status reporting, or validity checking.	• Commercial radio and television are examples. • Traffic control systems which provide no return information to a master controller or traffic operations center use this mode.
Half Duplex (HDX)	Data flow in either direction, but only in one direction at a time.	• Requires modem at each end of the line. • Requires control capability to assure proper operation. • Uses latency time or turnaround time (the time period required to turn the line around) for the process in which the direction of data transmission is reversed, which can be time consuming.	• In a copper wire transmission medium, HDX requires two wires but may be used with four wires (four wires provide improved interference characteristics).
Full Duplex (FDX)	Data flow possible in both directions at the same time.	• Acts like two simplex channels in opposite directions. • Permits independent, two-way, simultaneous data transmission • May raise cost of channel. • Reduces the one-way capacity if frequency multiplexing is used on a single channel.	• In a copper wire transmission system, some FDX modems require four wires while others require only two wires. In the latter case, the modem divides the channel into two subchannels to achieve simultaneous bidirectional service.

- Central.
- Distributed.
- Trunked.
- Backbone.
- Multimedia.

Characteristics of these communications architectures are summarized in table 9-5.

Central

A *Central* communications architecture is one that possesses only one level of communication before the signal reaches the field controller. In other words, all of the field devices link directly to the control center. Because data flows directly from the field devices to the control center, it is not processed along the way. As a result, only one data rate exists between the field controller and the control center. Because the field devices and the control center are linked directly, central communications architectures require only one

Table 9-5. Common Applications of Communications Architectures. [1]

Traffic System Communications Architecture	Common Application	Examples
Central	Communication requirements limited to a small number of field controller and video channels at each field location.	• Traffic signal systems controlled by computer at traffic operations center. • Small or medium sized freeway surveillance systems with limited video.
Distributed	Traffic control system computations performed at locations other than traffic operations center and field controllers.	• Closed loop traffic signal control systems.
Trunking	Achieves economies by concentrating data onto high speed channels for long runs to traffic operations center.	• Large freeway surveillance systems with long runs to traffic operation center.
Multimedia Channel	Geometrics and/or economics render single medium impractical.	• Signal systems and freeway surveillance systems with no right-of-way connection to traffic operations center requiring leased media to access control center. • Change from land line to wireless medium to cross a physical obstacle.
Backbone and Distribution System	Very heavy communication requirements (usually including video) that make the use of high speed channels economical for the longer transmission links.	• Large area-wide freeway surveillance systems and corridors.

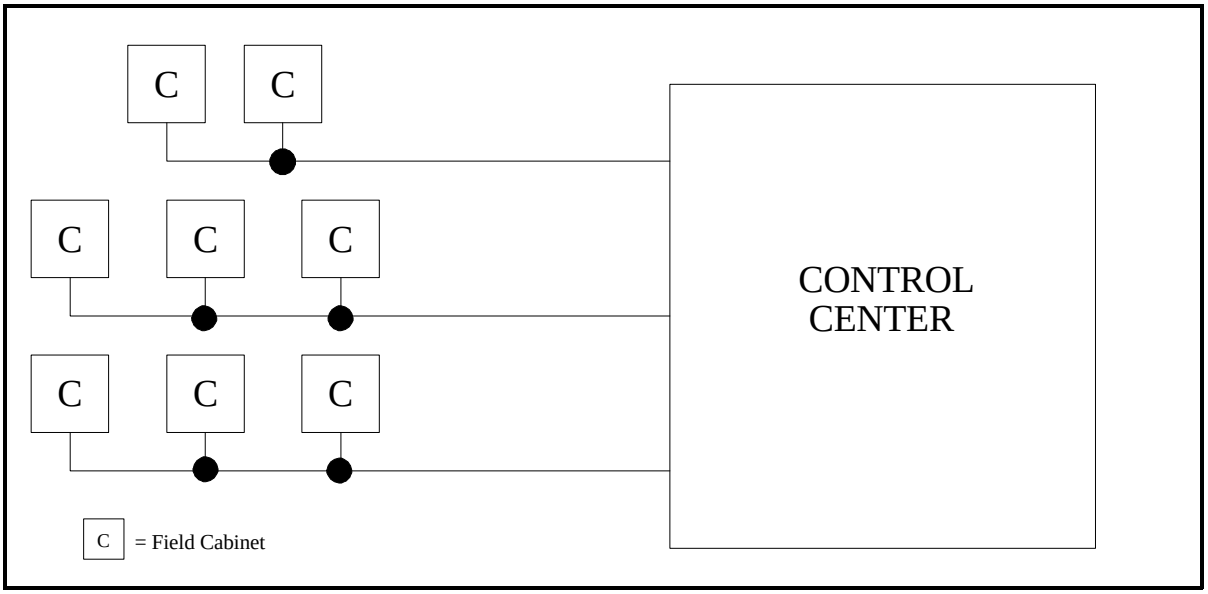


Figure 9-3. Illustration of a Central Communications Architecture. [1]

communication protocol. The primary advantage of this type of communications architecture is that a direct link is provided between each individual field device and the control center; therefore, if a link fails, communications are cut only to that one device. Figure 9-3 illustrates the concept of a central communications architecture.

Distributed

A *Distributed* architecture is generally used as follows:

- Where multiple levels of computations occur between the field devices and the control center.
- Where changes in the data rate occur between the field devices and the control center.

This type of architecture commonly uses field master controllers to collect data from many of the local control units, processes the data, and then transmits the processed information back to the control center (see figure 9-4).

Trunking

A special type of distributed architecture is a *Trunking* architecture. A communications system is trunked if the following holds true: [1]

- The communications system collects information from, and distributes it to several field controllers.
- At some field location, the data rate or bandwidth of the communication channel increases to require fewer channels for communicating with the control center.

Figure 9-5 shows an example of a trunked type of communications architecture.

Backbone

Another form of distributed communications architecture using multiple data rates is a *Backbone* type of architecture. This type of architecture is characterized by a high data rate backbone connected to a series of field communications hubs or nodes. Equipment at each hub transforms these high data rate

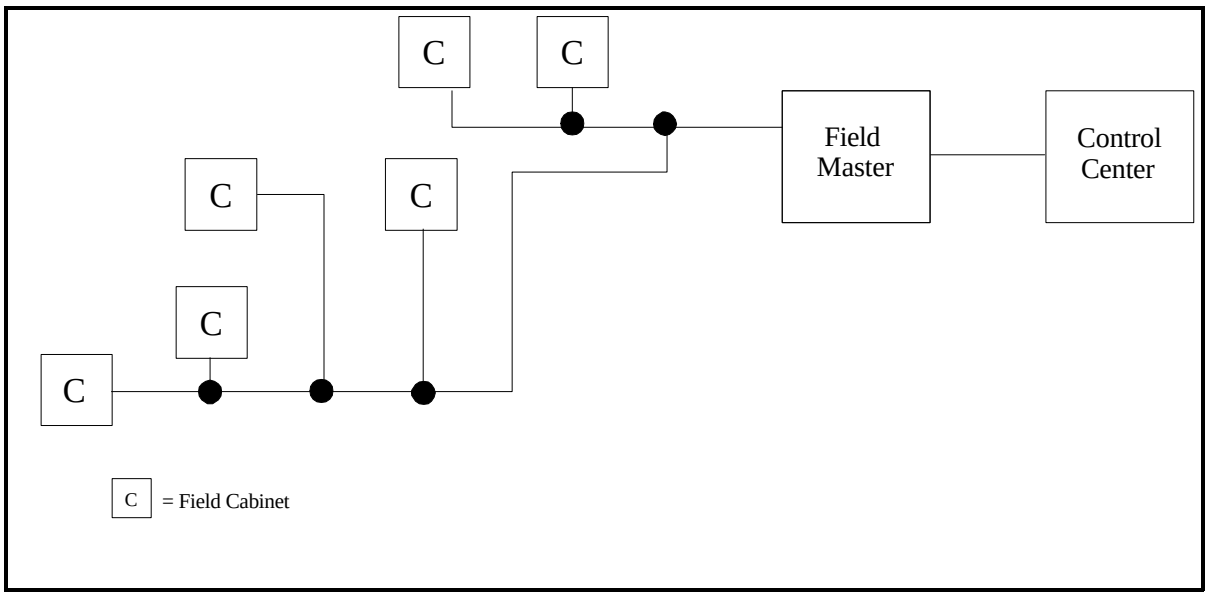


Figure 9-4. Illustration of a Distributed Communications Architecture. [1]

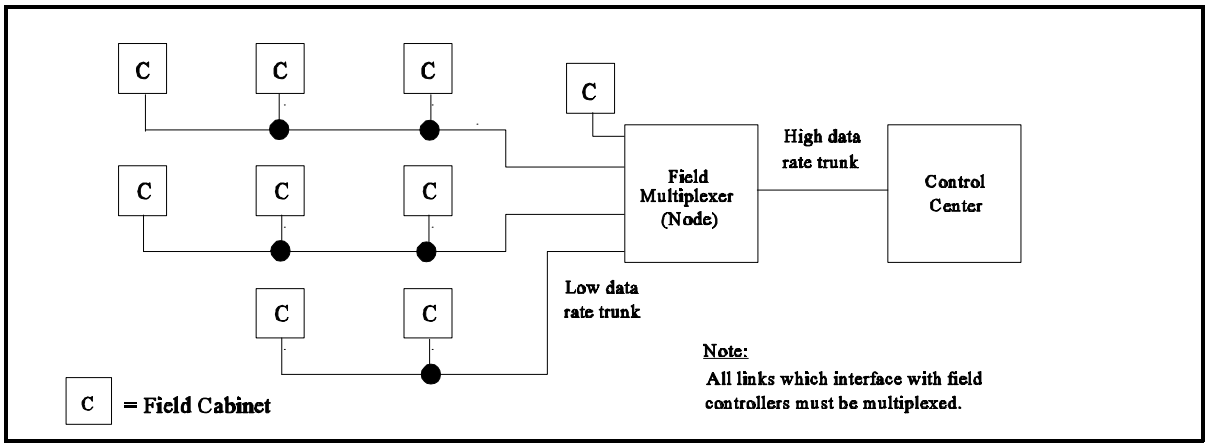


Figure 9-5. Illustration of a Trunked Communications Architecture [1]

channels into many separate low data rate distribution channels. Figure 9-6 shows an illustration of a backbone type of architecture.

Although it can be used with other high bandwidth communication media, the backbone type of architecture is commonly used with fiber optic networks. The four common types of topologies (i.e., network configurations) used with backbone architectures for fiber optic systems include the following:

- Unprotected Ring — Each node (i.e., a communications hub or field device) is connected to two others by unidirectional transmission links, creating a "closed" loop.

- Protected Ring — Two rings are used instead of one, thereby providing two unidirectional transmission paths that may run in opposition directions. Redundant opposite direction paths allow each node to communicate with every other node, even if the communication media is cut.

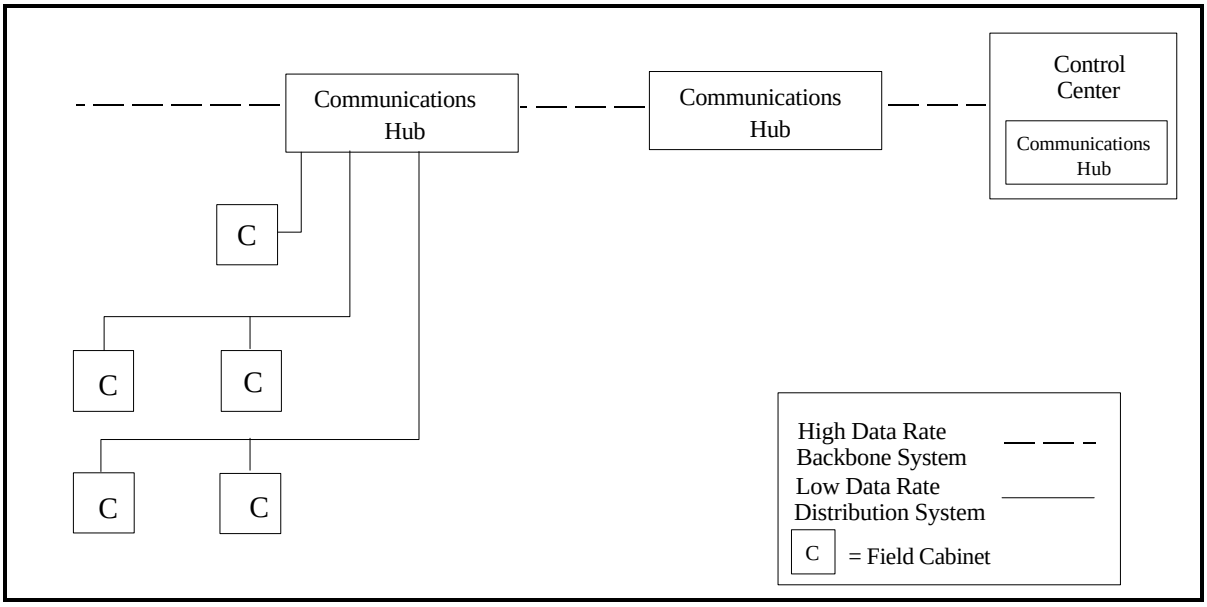


Figure 9-6. Illustration of a Backbone Communications Architecture. [1]

- Linear Drop — Nodes are connected in a string or chain, with transmission data being "dropped" at a designated node.

- Star — Communication links emanate from a source node (e.g., traffic control center) to multiple secondary nodes (e.g., a communication hub or field device).

Figure 9-7 illustrates each of these backbone architecture topologies.

Multimedia

Although not a communications architecture per se, planners and designers need to be aware of the impact changing communication media has on the architectural design of the communications system. A *multimedia* communication channel occurs where more than one medium is used to transmit data and commands to and from the control center and the field devices without altering the data rate and transmission protocol. This type of communication channel is generally used to breach natural or manufactured impediments (i.e., mountains, valleys, etc.) to provide wireline type of communication. Figure 9-8 illustrates the general concept of a multimedia communication link.

IDENTIFY AND SCREEN TECHNOLOGIES

After the functional requirements and system architecture have been established for a communications system, the next step is to identify and screen alternative communications technologies. The *Communications Handbook for Traffic Control Systems* describes a process for identifying, screening, and selecting the communications technologies in a freeway management system.[1] The steps in this process include the following:

- Identify generic or typical links consistently used through the communications architecture. Figure 9-9 illustrates some generic links that occur

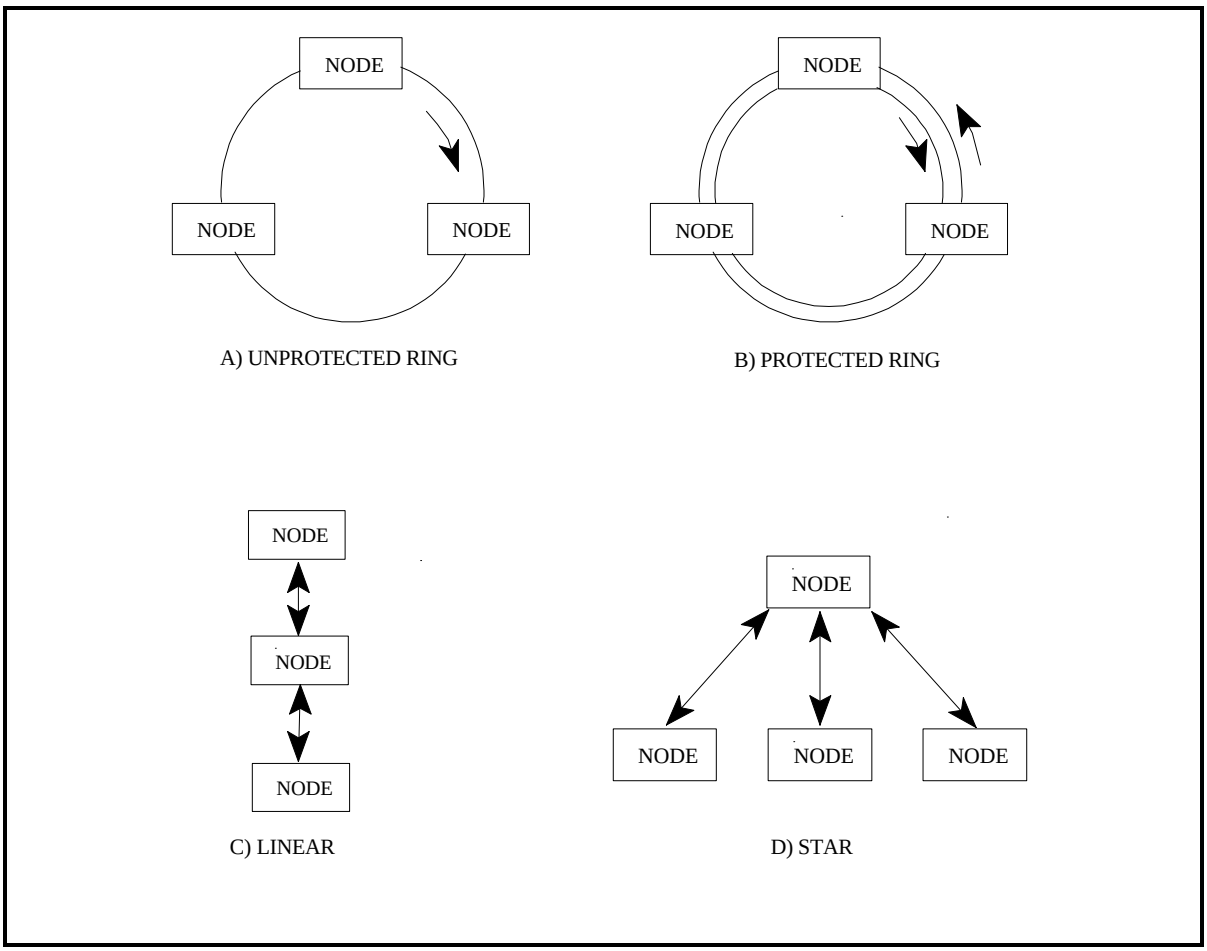


Figure 9-7. Traffic Control Communications Network Topologies. [1]

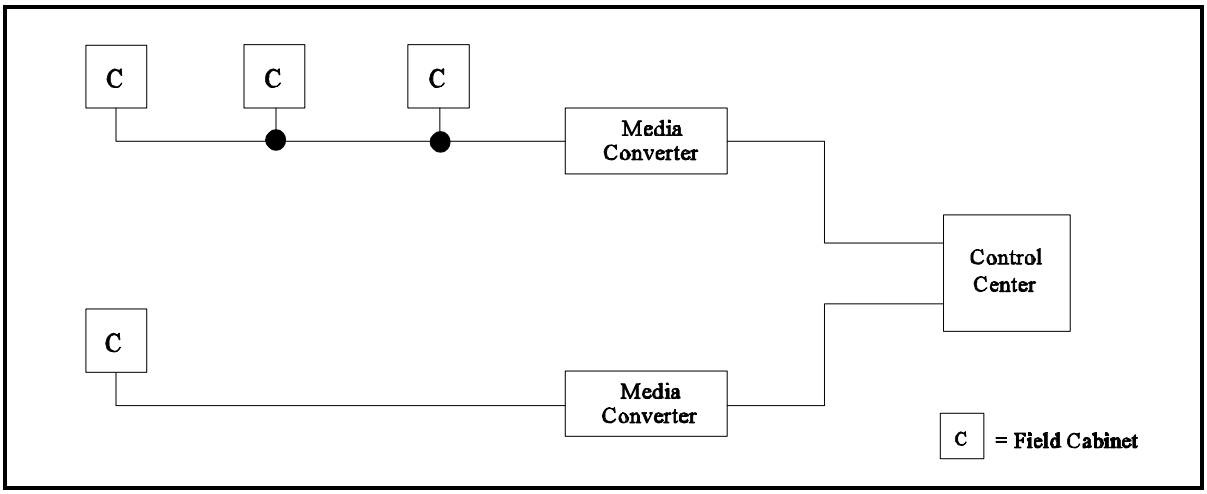


Figure 9-8. Illustration of a Multimedia Communications Channel. [1]

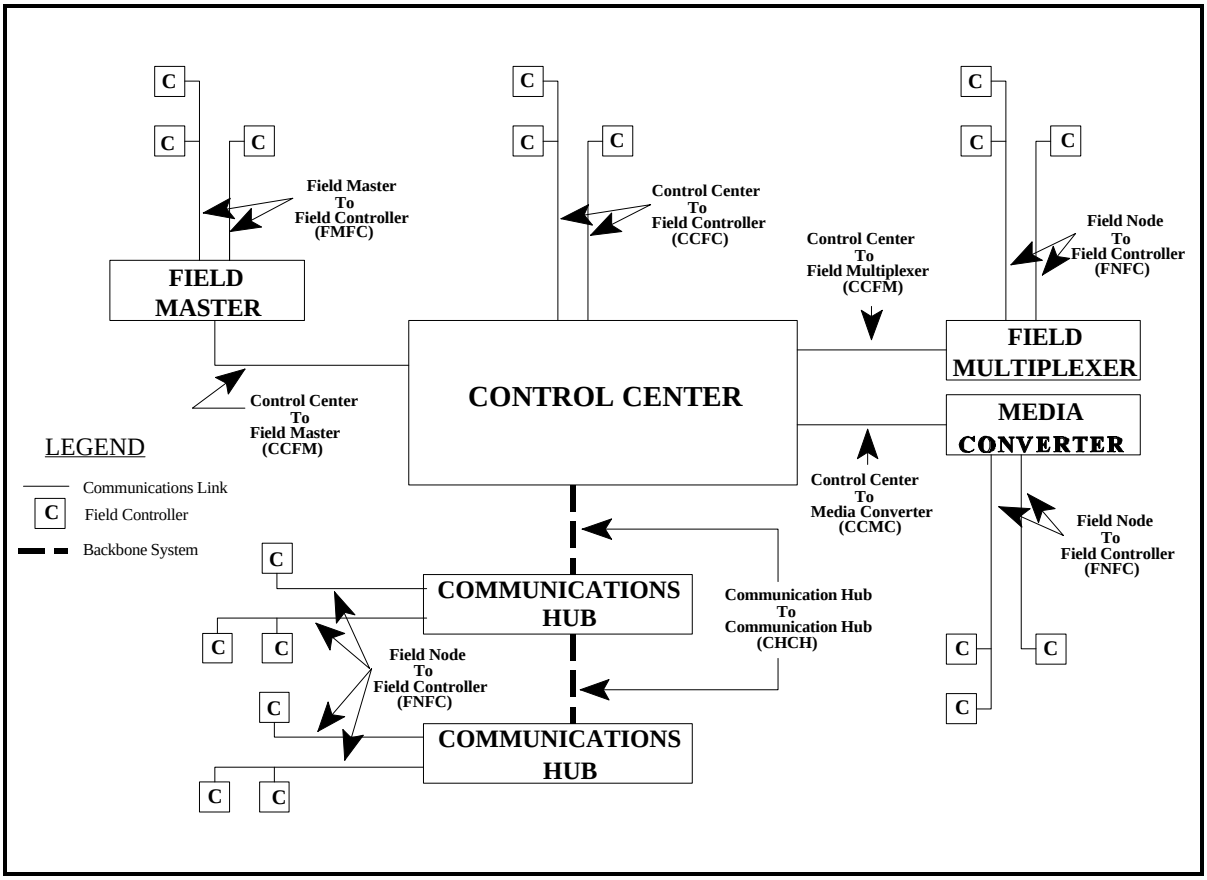


Figure 9-9. Illustration of Generic Communication Links. (1)

in communications systems, while table 9-6 summarizes the characteristics of these links.

- Using tables 9-7 and 9-8, identify candidate media for each generic communications link.

- Conduct a preliminary screening of the media to eliminate any that are impractical or unavailable in an area. The preliminary screening should also eliminate those media and technologies that are not compatible with institutional goals and policies.

- Estimate the data rate requirements for the communications channel serving each field device.

- Eliminate those media and technologies that cannot meet the data rate requirements of the field devices.

- Assess the potential for satisfying the communications requirements via leased communications systems.

- Assess the remaining technologies to decide the following:

  – Whether or not they satisfy the functional requirements of the system.

  – The interface requirements of the field controllers and other field equipment in relation to the communications media.

Table 9-6. Attributes of Generic Links in a Communications System. [1]

Type of Generic Communication Link	Attributes
Control Center to Field Controller (CCFC)	The type of connection between a control center and field controller where no computation or change in data rate occurs.
Control Center to Field Master (CCFM)	The type of connection between a control center and a field master that provides supervisory control to a group of local controllers. The master controller generally does not provide direct control over the site, but rather processes data and adds information and/or control commands.
Field Master to Field Controller (FMFC)	The type of connection between a field master and local field controller. The local field controller provides direct control and collects information from the specific devices at a site. The master controller processes data from a group of field controllers.
Control Center to Field Multiplexer (CCFX)	The field multiplexer site provides higher data rates to the control center link than does the field controller. While the data rates differ, the multiplexer performs no processing related to traffic system functional requirements.
Control Center to Media Converter (CCMC)	The type of communication that occurs between a control center and a device that converts the communication channel to a different type of media (e.g., from twisted pair to fiber optic).
Communication Hub to Communication Hub (CHCH)	This type of connection occurs between two communication hubs. This type of connection typically represents a high data rate backbone link.
Field Node to Field Controller (FNFC)	The field controller connects to a field node for either of the following purposes: • To provide a higher order of multiplexing by means of trunking or backbone systems between the control center and field node (CCFX, CHCH) by means of a field multiplexer. • To use a medium more suitable for communication to the control center while retaining essentially the same communication channel capacity (CCMC). The field node in this case consists essentially of back-to-back modems to service each medium.

Table 9-7. Relationship of Communication Technology to Generic Communication Link for Data Transmission. (1)

Communication Technology	Generic Communication Link						
	CCFC	CCFM	FMFC	CCFX*	FNFC	CCMC	CHCH**
Twisted wire pair voice grade channels	✓	✓	✓	✓	✓	✓	
Leased voice grade channels	✓	✓	✓			✓	
Switched voice grade channels		✓					
Fiber optics channels	✓	✓	✓	✓	✓	✓	✓
CATV channels	✓					✓	
Leased digital channels	✓			✓		✓	
Area Radio Networks (owned)	✓	✓	✓		✓	✓	
Terrestrial Microwave	✓	✓	✓	✓	✓	✓	✓
Spread spectrum radio	✓		✓		✓		

* Trunking Link **Backbone Link

CCFC = Control Center to Field Controller
CCFM = Control Center to Field Master
FMFC= Field Master to Field Controller
CCFX = Control Center to Field Multiplexer
CCMC = Control Center to Media Converter
CHCH = Communication Hub to Communication Hub
FNFC = Field

  – The types of controls needed to provide adequate environmental protection.

- Examine the feasibility of providing a backbone or trunking architecture for the field devices.

- Isolate those links where geometric or physical situations require a change in the communications media.

- Identify multiplexing strategies for each communications media.

- Develop life-cycle cost estimates for the alternative communications technologies for each system link. Cost estimates should include not only capital cost but also operating and maintenance costs.

- Using the estimate of the link costs, estimate the total communications systems costs.

- Assess the impact of the non-cost related factors on the total system costs. Examples of the non-cost related factors include the following:

Table 9-8. Relationship of Communication Technology to Generic Communication Link for Video Transmission. [1]

Communication Technology	Generic Communication Link				
	CCFC	CCFX*	FNFC	CCMC	CHCH**
Twisted wire pair voice grade channels					
Leased voice grade channels					
Switched voice grade channels					
Fiber optics channels	✓	✓	✓	✓	✓
CATV channels	✓	✓	✓	✓	✓
Leased digital channels	✓	✓	✓	✓	
Area Radio Networks (owned)					
Terrestrial Microwave	✓	✓	✓	✓	✓
Spread spectrum radio	✓		✓		

* Trunking Link ** Backbone Link

CCFC = Control Center to Field Controller
CCFM = Control Center to Field Master
FMFC= Field Master to Field Controller
CCFX = Control Center to Field Multiplexer
CCMC = Control Center to Media Converter
CHCH = Communication Hub to Communication Hub
FNFC = Field

- The risk of the leased communication costs escalating in the future.
- The differences in service reliability between owned lines, leased services, and radio communications.
- The ease of maintaining each communication alternative.

- Select the "best" communications system alternative.
- Review assumptions and iterate the process if necessary.

The *Communications Handbook for Traffic Control Systems* or similar references should be consulted for a detailed explanation of the analyses involved in each step, and example applications of the evaluation process.[1]

IMPLEMENTATION

National ITS Standards

In the past, some agencies have experienced problems maintaining and expanding their systems because they used communications equipment that did not conform to standard equipment interfaces and protocols. One method of ensuring maintainability and expandability is to specify communications equipment and system designs that use common standards. System planners and designers must recognize the importance of standards during the design process. The designer must prepare specifications in sufficient detail to ensure the desired level of

standardization of field equipment functions and communication interfaces and protocols. Standards are available for the following elements in a communications system: [1]

- Serial data interfaces.
- Modems.
- Voiceband channels.
- Digital signals.
- Fiber optics.
- Integrated services digital networks.
- Compressed video.
- Local area networks.

The reader should consult the *Communications Handbook for Traffic Control Systems* for more details about the specific standards that relate to these elements.[1]

The advantages and disadvantages of incorporating standards into the design of a communications system are summarized in table 9-9.

Beginning in 1995, a group composed of infrastructure operators, manufacturers, system integrators, and representatives from the FHWA began developing a national communications standard. Communications standards were needed to ensure interoperability and interconnectivity of traffic control and ITS devices such as variable message sign control, camera control, vehicle classification, and general purpose data collection and device control.[4] This standard is known as the National Transportation Communications for ITS Protocol (NTCIP). It is designed to provide a communications interface that allows hardware and software products from different manufacturers to be connected in a traffic management system. The potential benefits of adopting the NTCIP standards for use in designing freeway management systems include the following: [5]

- Reduction in warehousing requirements and costs.
- Reduction in training needs for personnel.
- Improvement of staffing effectiveness if same personnel are used predominately for repairs and operations of unique equipment.
- Reduction in product costs of some devices if a more competitive procurement environment yields cost reductions.
- Mitigation of procurement issues associated with system expansion and spares.
- Reduction in downtime.
- Enhancement of interjurisdicational coordination and integration.
- Shared use of the communications network, providing the opportunity to share the costs of communications with other agencies (or at least avoid the cost of dedicated parallel communications systems).

For more information about NTCIP, the reader should consult the **Lessons Learned** section at the end of this module.

In addition, the National Architecture effort also spurred development of standards for other freeway applications, such as the following:

Table 9-9. Advantages and Disadvantages of Using Communications Standards. [1]

Advantages	Disadvantages
• Ensures the availability and compatibility of equipment spares for repairs and future expansion. • Allows system to be expanded to locations not contiguous to the existing system and accessible by land line. • Ensures the compatibility of test equipment. • Supported by a wide body of literature describing the functions of ports and modems designed to common communications standards.	• Some existing system architectures are incompatible with accepted standards. • May result in higher system costs since equipment with less than optimum data rate may be required to conform to standards. • May require higher level of maintenance personnel.

- Transit Communications Interface Protocols (TCIP).
- Dedicated Short Range Communication (DSRC).
- Location Referencing.
- ITS Data Bus.
- FM ATIS Broadcasting.
- Navigation Messages.
- MayDay Reporting.
- Navigation Human Interfaces.
- ATIS Message Set

The technical edition of the *ITS Deployment Guidance for Freeway Management Systems* lists possible standards for the following different applications within Freeway management systems:

- GIS Applications
- DMS Applications
- CCTV Applications
- ETTM Applications
- ATIS Applications
- Center-to-Center Communications
- Center-to-Roadside Communications
- Roadside-to-Vehicle Communications
- Vehicle-to-Vehicle and In-Vehicle Communication
- Center-to-Remote Access Communication
- Center-to-Vehicle Communication

If circumstances dictate the use of nonstandard equipment, agencies should procure sufficient spare devices for use well into the future. Agencies should also consider acquiring an inventory of specialized parts (e.g., custom integrated

circuitry) that may become difficult to purchase in the future. [1]

Installation

The use of proper installation techniques and procedures is critical to installing a successful and functioning communications system. Failures and unreliable performance in a communications system commonly occur because proper installation procedures were not followed. Examples of improper installation techniques associated with wireline communications systems that can affect overall performance include the following: [1]

- Improperly installing fiber optic cable and coaxial cable connections.
- Exceeding the maximum pulling tension on the cable.
- Using a tighter radius than recommended by the manufacturer when bending a cable.
- Improperly splicing cables.

Installation and construction inspection personnel should receive appropriate related training before starting installation.

Three basic methods of installing cable are commonly used to connect devices in a freeway management system:

- Cable in Conduit.
- Direct Burial.
- Aerial Mounting.

Table 9-10 summarizes the advantages and disadvantages of each of these installation options.

9.3 TECHNIQUES AND TECHNOLOGIES

This section discusses the transmission media that are commonly used in communications systems for traffic management applications. This section also discusses some options available for transmitting video images to a control center. This section is intended only to highlight several major characteristics and issues associated with these media. For more detailed information on any of these media, the reader should consult the *Communications Handbook for Traffic Control Systems* or similar references.[1]

Communication media can be divided into the following two categories:

- Land-Line.
- Wireless.

The media contained in each of these categories are discussed below.

LAND-LINES

Land-lines (i.e., wire cable), whether leased or owned, are by far the most prevalent form of traffic control system communication media.[1] The three primary land-line transmission media commonly used in freeway management systems include the following:

- Twisted Wire Pair.
- Coaxial Cable.
- Fiber Optic.

Table 9-11 provides an overview of the land-line communication technologies that will be

Table 9-10. Advantages and Disadvantages of Installation Methods for Land-Line Communications Media.

Installation Method	Advantages	Disadvantages
Cable in Conduit	• Most secure because cable is placed in conduit that is buried approximately 1 m underground. • Requires the least maintenance if adequately protected by markouts and tight permitting process. • Access to cable provided through junction boxes.	• Frequent junction boxes required, especially when bending cable. • Can be damaged by excavation or installation of sign posts, guard rail posts, etc. • Relatively more expensive than other installation methods.
Direct Burial	• Cable, which is protected by extra insulating jacket, is buried directly in the ground. • Eliminates need for conduit and junction boxes. • Easier to install than conduit and aerial methods. • Suitable for use where roadside development is limited.	• Greater susceptibility to damage by excavation. • More difficult to maintain and repair. • Requires good control of excavations through permitting process.
Aerial Mounting	• Cable hung from existing utility poles in or near right-of-way. • Least expensive installation method (often ¼ the cost of installing cable in conduit or direct burial). • Not susceptible to damage by excavation. • Relatively easy to repair.	• Normally requires yearly rental fee on utility poles. • May also be assessed fee for relocating existing utilities on poles. • Because sun can deteriorate cable, it must be replaced approximately every 15 years. • Susceptible to damage from tree limbs, etc. during wind storms.

discussed and some of their notable features. Each of these media is discussed below.

Twisted-Wire Pairs

Twisted-wire pairs are the most prevalent type of communications media used in traffic control applications. A twisted-pair cable consists of sets of two wires wrapped around each other. The twisting of each pair reduces interference from external sources because the pairs of conductors carrying the signal are always immediately next to each other in the cable. Therefore, the induced signal from the interfering source will affect each conductor of the pair similarly. Figure 9-10 provides an illustration of one conductor in a twisted-wire pair cable.

Since the receiver is measuring only the difference in voltage between the conductors, the fact that both conductors have the same induced voltage will not be noticed. The design of the cable also reduces crosstalk between lines, because the same current is flowing in opposite directions in each conductor (see figure 9-11). Therefore, the electric field radiates from the two conductors in opposite directions (radially around each wire). This results in the two fields canceling each other out. [(1)]

Table 9-11. Summary of Land-Line Technologies and Their Features. [2,6]

Features	Twisted-Wire Pairs	Coaxial Cable	Fiber Optics
Transmission Media	Copper Wires	Center conductor is copper clad aluminum Outer conductor uses aluminum	Glass or plastic fibers
Transmission Range	14 to 24 km (8.7 to 14.9 miles) with repeaters	Commercial subscriber network repeaters at 0.5 km (0.31 mi); 1 km (0.62 mi) or more on dedicated systems; maximum of approximately 60 repeaters	Rarely a limitation when drop/insert units used at communications hubs or drop points
Principal Multiplexing/ Modulation Technique Used	Time Division Multiplex (FSK)	Frequency Division Multiplexing to divide channel bandwidths; Time Division Multiplexing to communicate data	Time Division Multiplex (FSK)
Carrier Frequency Band	300 to 3000 Hz	5 MHz to 350 MHz	850 to 1,550 nanometers
Bandwidth/Channel Bandwidth	Will exceed 2.7 Hz for most systems	6 MHz/channel	Various
Data Rates per Channel	1,200 to 3,100 bps Higher rates possible with different modulation techniques	Up to 7.5 Mbps based on channel subdivision	Up to 2.4 Gbs
Government Regulation of Channel or Service	None	May require licensing from local and state authorities, FCC provides legislation	None
Types of Information Supported	Data, voice, slow scan TV	Data, voice, video	Data, voice, analog TV, Codec
Owned or Leased	Owned	Either	Owned

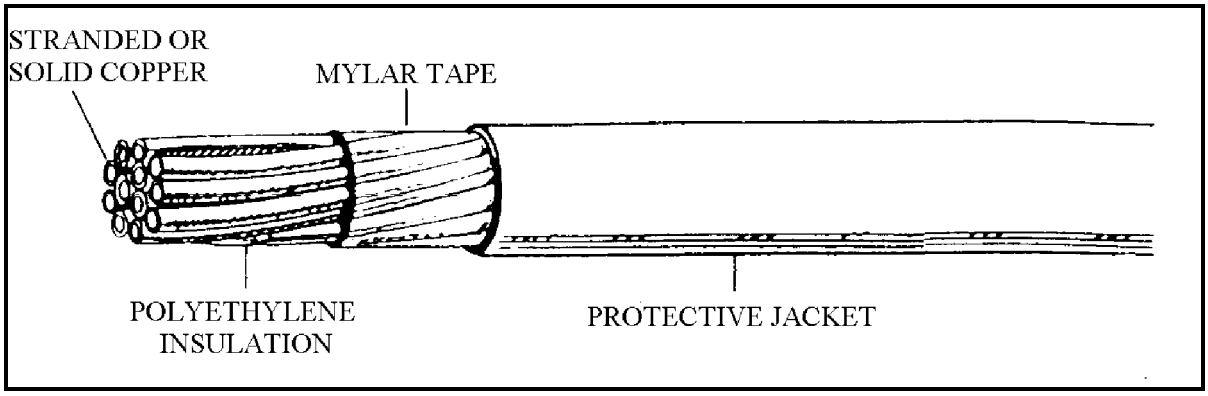


Figure 9-10. Illustration of a Twisted-Wire Pair Cable.

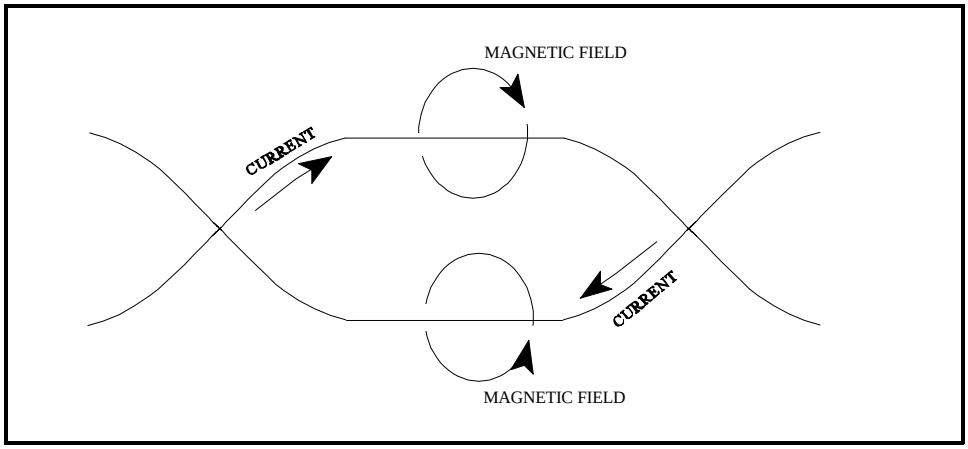


Figure 9-11. Crosstalk Reduction by a Twisted-Wire Pair. [1]

Most twisted-wire pair cabling used in traffic control systems is usually of the voice-grade type. This means the usual bandwidth (the range of signal frequencies a medium or channel will respond to, or carry without excessive loss in signal strength) ranges from 300 Hz to 3000 Hz, the audible frequency range of the human voice. Twisted-wire pair cabling is commonly used in voice telephone communications. A voice-grade cable will accommodate a data transmission rate of 1,200 bits per second (bps), which is adequate for transferring small amounts of data (i.e., loop detector data). When higher data transmission rates are required (for trunking applications), the twisted-pair cable must be conditioned by adding electronic equipment, such as loading coils, to improve the transmission characteristics of the line.

Twisted-wire pair technology consists of two insulated copper wires wrapped around each other and used to convey signals. Twisted wire is manufactured in standard numbers of pairs (6, 12, 18, 25, 50, 75, 100, 150, 200, 300, 400, 600, and 1200) in wire gauges 19, 22, 24, and 26 American Wire Gauge (AWG). [2] The maximum number of twisted pairs of wires manufactured in a single cable is 2600. [7]

The advantages and disadvantages of twisted-wire technology are summarized in table 9-12.

Twisted-wire pair technology is an inexpensive communications medium whose acquisition cost is lower than that of both coaxial and fiber-optic technologies; however, long term operating costs are dependent on installation considerations and are higher when life-cycle costs are considered. [6] The extent of such costs will vary greatly with the different installation and maintenance procedures required to support the installation method used and the terrain covered. For example, aerial installations can be damaged or links severed by sources ranging from falling trees to storm related damage. By comparison, buried cable affords protection from those threats, but may be subject to construction-related severance and will reflect higher installation costs. [1]

Twisted-wire pairs can be stranded together to form color coded binder groups that can then be wrapped around a common axis to increase the number of conductors in a given cable. The amount of twist in a given binder group is used to control how much crosstalk occurs.[1] A cable consisting of multiple

Table 9-12. Advantages and Disadvantages of Twisted-Wire Pairs. [1,6,7]

Advantages	Disadvantages
• Represents a low cost form of transmission. • Easy to splice. • Requires no special interface equipment. • Electrical characteristics are very favorable to basic analog transmission.	• Data cable splicing is not recommended. • There is a bandwidth limitation in that twisted-wire pair tends to attenuate high-frequency electrical signals, thereby limiting the ability to transmit digital information at high data rates. • Bandwidth limitation prevents transmission of live television images, though recent developments permit transmission of slow-scan television. (Prototype equipment is available for transmission of full-motion television over twisted-pair copper wire.) • Low security.

twisted-wire pairs can transmit several channels of data. In a full-duplex network, every two pairs can provide one channel.

The electrical characteristics of twisted-pair copper strongly support basic analog transmissions; however, its capability to transmit digital information at high data rates is limited by attenuation of high-frequency electrical signals.[7] The attenuation of high-frequency signals increases with increased cable length and decreased diameter. [1]
Noise can be a source of problems with some twisted-pair cable communications systems. Most modems used with twisted-pair systems, however, can filter out signals, many sources of noise below 750 Hz, but in noisy high frequency environments, other types of communications media may be more appropriate. Sources of potential noise in a twisted-pair communications system include lightning, high voltage power lines, and large electric motors. [1,6]

The use of equipment from multiple vendors can increase the difficulty in separating a given signal and transmission noise. [6] Noise may also rise if the twisting of the wire pair does not completely cancel the induced electrical fields. Moisture in twisted-pair cable networks can also lead to increased crosstalk and cable noise by introducing signal reflections (cable splices are the primary source of moisture entry). [1]

Several features of twisted wire pair communications should be considered during the design process. Fundamental concepts that should be considered by system designers include the items summarized in table 9-13. Besides those items listed in table 9-13, other items that should be considered during the design and installation processes including the following: [1,2]

- Maximum allowable transmitter power.
- Signal level expected at the receiver.

- Limitations on the number of drops on a multipoint line.
- Wire gauge.
- Environmental effects, such as moisture entry, noise, and transients.
- Aspects involving the number of repeaters, special functions of the system, and testing.
- Susceptibility to inadvertent electrical surges resulting from maintenance activities (consideration should be given to the provision of transient protection for sensitive electrical devices connected to the cable).
- Installation should not exceed the maximum pulling strength of the cable. Installation specifications should require either hand-pulling or tension monitoring using a strain gauge.
- Consideration of the use of existing facilities should take into account the limitations of the system, potential to expand, and physical plant responsibilities.

Finally, it should be noted that this transmission medium offers little security. It may be easily tapped and the electromagnetic field it radiates can be read by sensitive electronic devices. [6]

Coaxial Cable

Coaxial cable technology can transmit either data or video via several communications channels. [2] Its name is derived from its characteristic shape, essentially a set of two concentric circles. As shown in figure 9-12, the cable consists of an unbalanced pair made up of an inner conductor within an outer conductor held in a concentric configuration by a dielectric material that separates the two conductors. [7] Physical structure characteristics typical of coaxial cable include the following: [1]

- A center conductor composed of copper clad aluminum.
- An outer conductor composed of aluminum as a braided metal fabric, corrugated semi-rigid metal, or a rigid metal tube.
- A dielectric composed of solid polyethylene or polyvinyl chloride, foam, Spirafil, or inert gas (when gas is used, the center conductor is kept in place by spacers or disks).
- An outer jacket composed of low density, high molecular weight polyethylene.
- Optional armor for use in direct burial in gopher areas.
- Cables that vary greatly in size and construction. Typical traffic control applications commonly use a 19.05 mm semi-rigid coax for trunk lines, with smaller diameters used for connections between the trunk and field drops.

Table 9-14 summarizes the major advantages and disadvantages of coaxial communication technology.

A coaxial cable system uses frequency-division multiplexing (FDM) and time-division multiplexing (TDM) to fit all traffic control signals on a single conductor. FDM is used to subdivide the cable bandwidth into appropriate channels for data, video, and voice transmission, and TDM is then used to communicate the data.[1]

Table 9-13. Design and Installation Considerations Relating to Twisted-Wire Pairs. [1,2]

Consideration	Design Factors/Options
Number of wire pairs required for startup and expansion	• Total number of field drops • Maximum number of drops allowed on each channel • Network configuration
Cable routing and installation techniques	• Field drop locations • Existing communication facilities • Cable termination requirements • Cost
Installation methods	• Underground, in conduit • Underground, by direct burial • Aerial, using existing/new utility poles • Support of cable or conduit by bridges, overpasses, and other structures • Need for submarine cable
Cable size	• Minimum bending radii • Cable weight
Cable size affect on conduit design	• Available conduit and pole spacing • Size of manholes and junction boxes • Required bend of conduit at entry to cabinet/junction box • Aerial cable weight
Attenuation levels	• Signal frequency • Conductor size • Cable length • Number of splices and connections
Susceptibility to electrical transients from natural phenomena such as lightning	• Storm frequency • Terrain • Type of cable
Moisture prevention	• All splices on telephone terminal blocks should be inside weatherproof cabinets, or in waterproof splice enclosures for aerial installations • Installation must be performed such that the cable does not drag on the pavement or rub against jagged conduit ends causing cuts or abrasions • Be aware of cable jacketing imperfections • Seal cable ends during pulling operations to prevent moisture entry
Thorough testing after installation including at least:	• End-to-end continuity for each pair • The insulation resistance for each conductor (to ground and to the paired-conductor) • Attenuation

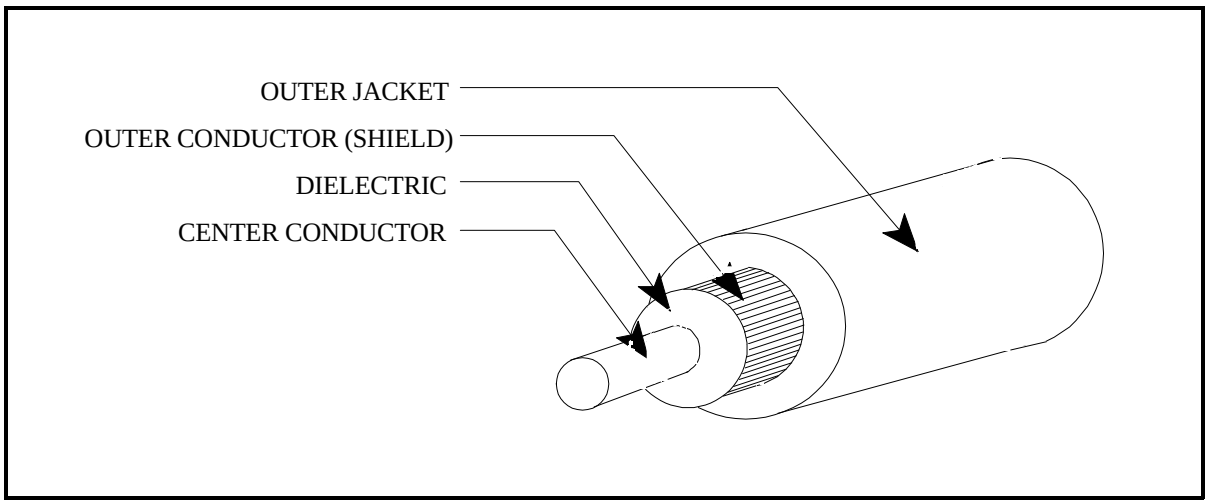


Figure 9-12. Illustration of Coaxial Cable Design.

Table 9-14. Advantages and Disadvantages of Coaxial Cable. [1,6,7]

Advantages	Disadvantages
• Because of its physical structure, coaxial cable is more immune to electromagnetic interference and has a much higher bandwidth than twisted-pair cable. • Minimal signal losses. • Low signal leakage. • Higher bandwidth allows for transmission of video signals (cable television using coaxial cable can transmit as many as 75 independent video signals) and for the transmission of digital data at very high rates. • Bandwidth of coaxial cable permits theoretical transmission rates as high as 700 million bits per second (Mb/s). These rates are more favorable than typical twisted-pair rates that are limited to less than 24 Mb/s for short distances.	• Splice connections are susceptible to noise and transient problems. • Cannot be spliced together by manual strip and twist method. Inherent nature of cable, and the importance of conductor alignment, make the coaxial cable much more difficult to splice. • Lower communication reliability than fiber optic. • Higher maintenance and adjustment effort required compared to fiber optic. • Low security. • Cannot be conventionally “tapped.” Requires termination.

The typical coaxial communications system can be used for two-way communications by dividing the frequency spectrum into two parts. The upstream channels (generally inside the control center) use one range of frequencies to transmit data, while the downstream channels transmit in the opposite direction (from the field to the operations center). Filters and amplifiers are used at the repeater locations to carry the signals simultaneously in both directions. [2]

The coaxial network branches out from the operations center via a series of splitters, directional couplers, taps, and repeaters. Splitters are used to divide and combine signals within the tree-like network depending on their destination. Directional couplers are used to combine signals onto cables by direction. Taps are then used to divert signals to and from the trunk line at field drops. Repeater amplifiers are supplied to boost the signal along the system based on transmission frequencies, cable size, number of connectors, and design tolerances of the system. [1]

Coaxial cable's physical structure is less affected by electromagnetic interference than twisted wire pairs. [7] Repeater amplifiers have low noise levels and can deliver strong signals over a wide range of output levels; however, coaxial cable does experience moderate signal attenuation losses. Coaxial systems amplify the signal at each repeater, instead of regenerating it. As a result, background noise is also amplified at each repeater. This phenomenon limits the number of repeaters that can be placed in series to approximately 60 (this limit has not been reported to have affected traffic control systems using this technology). [1]

The attenuation in a given coaxial cable varies as a function of frequency, temperature, and cable size. Attenuation levels roughly double each time the bandpass frequency quadruples. Temperature variations can also degrade performance. For example, every 10-degree increase in temperature above 68 degrees results in approximately a 1 percent increase on attenuation.[6] Amplifiers are designed to compensate for the effects of temperature, though technicians must occasionally retune them. Unfortunately, the modems that are commercially supplied for coaxial systems often have temperature settings that fall below NEMA specifications. Use of these modems can result in interference with adjacent channels and the disruption of system communications if their temperature limits are exceeded. [1]

While only one cable may be required in a coaxial system, the design of a coaxial network is more complex than that of a twisted pair network. The higher frequencies and lower tolerances associated with noise and interference among coaxial cables translate to both higher skill level and test equipment requirements to support a coaxial network. [1] As indicated in describing the previously cited potential modem problems, it is essential that the designer have a thorough understanding of the equipment required by these systems.

Similar to twisted wire pairs, coaxial cable may be installed underground in a conduit, by direct burial, or overhead on utility poles. Generally, coaxial network connections can be spliced using press-fit connectors that clamp over the ends of the wires and are then screwed together. Many alternative connector and splicing arrangements can be employed for inserting tabs at midpoints in the cable and for ending cable runs. Given these factors, a design must consider the following: [1]

- The placement of amplifiers, couplers, splitters, and pilot generators.

- Provision of power for the amplifiers (repeaters).
- Compensation for temperature extremes and rate of change.
- Minimization of electromagnetic interference (EMI) and radio-frequency interference (RFI).
- Allocation of frequencies to the channels.

Contingent on the method of installation, the installer must ensure the following: [1]

- All conduit has gentle bends between pull boxes. (Severe bending of the conduit may cause the outer conductor to collapse and/or the inner conductor to permanently shift.)
- The maximum pulling tension specified by the manufacturer is not exceeded.
- All splices and connections are made in waterproof cabinets to reduce exposure to moisture and noise ingress.

During the installation process, care should be taken to ensure good quality control. Adequate construction supervision and thorough testing of the entire installed network are needed to maintain good quality control. Testing should consist of an end-to-end examination conducted under various weather conditions that exemplify operations at different temperatures.

Fiber Optic

Fiber optic communication provides a high volume, cost-effective means of transmitting either data or video via several communications channels with immunity to electrical interference. [1] Its name is derived from the medium's use of an optical fiber to transmit light by means of internal reflection off the surrounding surface cladding. Essentially, light impulses are coded and transmitted into a glass fiber structure. The fiber itself confines and guides the beam of light between origin and destination points. Upon reaching its destination, the light signal is detected, converted to electrical pulses, and decoded to an appropriate output. [7] The advantages and disadvantages of fiber optic communications media are summarized in table 9-15.

Fiber optic technologies are more expensive to acquire than comparable metallic transmission media; however, fiber optic cable has a reduced life-cycle cost when compared with copper transmission media, and is increasingly being used to replace coaxial cable systems. [7] Because of the high cost of fiber optic systems as compared with copper transmission, they are typically used for trunking applications in which large amounts of information must be transmitted over long distances or for high-speed local area networks that are to be distributed over a campus area. [2]

Fiber optic communications rely on light to act as a signal carrier within an optical fiber composed of a core region, cladding, and a coating. To transmit video, voice, and other data, the information is converted to a coded pulse of light, and introduced into the optical fiber, and is transmitted by internal reflection of the light wave within the fiber. Depending on the entry angle of the light into the fiber, the light will travel different distances. [2] Light waves that enter the core at an angle greater than the critical angle reflect and propagate within the core. Those light waves that enter at less than the critical angle refract into the cladding and attenuate. [1] As the light travels through the fiber, its bandwidth changes due to the spreading of the light pulses. As the pulses spread, they overlap and may interfere with

Table 9-15. Advantages and Disadvantages of Fiber Optic Communications.[1,7]

Advantages	Disadvantages
• A pair of light tubes can support many more circuits than a metallic path. • Immunicty from electromagnetic interference (EMI) and radio-frequency interference (RFI). • High integrity for data transmission. • Emits no radiation and it is difficult to tap a fiber tuve without detection of the resulting signal loss, thus represents a highly secure means of communication. • Use of small cable diameters and low-weight cable. • Small bending size, small bending radius, and light weight. • Safety in hazardous environments. • Extremely flexible - can be installed to support a low-capacity (low-bit-rate) system and, as the system's requirements expand, can use broadband capabilities of optical fibers and convert to a high-capacity (high-bit-rate) system simply by changing the terminal electronics.	• Designing a fiber optic network tends to require substantial engineering effort due to complexity of networks, light distribution characteristics and medium, and other factors. • Splicing tends to require elaborate equipment and expertise.

each other.[2] This dispersion effect limits the transmission speed and, if the pulse spreads so that the last portions of one pulse arrive after the first portions of the following pulse, may diminish the signal to a point that the receiver can no longer distinguish individual pulses.[1]

Several supplementary equipment items are needed with fiber optic communications systems. These items include modems, multiplexors, and repeaters. Fiber optic modems transmit data over a dedicated pair of fibers (e.g., between a controller and a traffic operations center) by converting an electrical signal to an optic signal and vice versa. This technology allows for full duplex communications over one or two fibers. Conversely, multiplexors transmit voice, data, and/or video over a fiber backbone network.[1]

Multiplexors combine many low- to medium-speed digital data channels into high-speed channels. In fiber optic systems, two types of multiplexors are used: electrical and optical. As implied by the name, electrical multiplexors combine signals electronically, and result in a combined electrical signal that drives the fiber optic transmitter. Single fiber optic channels can be used to carry multiple signals, using wavelength-division multiplexing. This technique is useful for long distance

backbone and trunk lines, but is not cost-effective for local distribution uses. [1]

After entering a fiber, a pulse is carried by the optical fibers to the receiving station, where it is either received or regenerated. At the final receiving station, the light pulses are converted to electric signals, decoded, and converted to the original form of information. To correct pulse dispersion, fiber optic repeaters convert the optical signal back to its original form and then back to an optical signal.[1]

In systems with multiple drops placed in series, common practice uses drop/insert units, which perform the following functions:[1]

- Interception of the optical signal.
- Conversion of the signal to an electrical form for use by the device.
- Injection of a response if necessary.
- Modulation of the electrical signal back to an optical signal.
- Retransmission of the optical signal.

In a daisy-chain multiple drop configuration, the failure of any drop/insert unit results in the loss of all the remaining units beyond that point. This situation can be avoided by using units that contain a passive transfer feature. [1] The passive transfer feature uses a separate battery back-up for each unit which provides continued communication to the downstream units in case of an electrical power failure in the controller. [2]

Fiber optics can also be used to transmit video signals in one of three manners: [2]

- Baseband — transmits unmodulated video signals.
- Modulated — extends the transmission range, using amplitude modulation or frequency modulation.
- Multiplexed — enables transmission of several video signals on the same fiber.

Because of its dielectric nature, fiber optic communication medium is unaffected by electrical signals. As a result, many problems such as radiative interference, ground loops, and lightning-induced damage (when fiber optics have been installed in a cable without metal) can be avoided with fiber optic technology. This property makes fiber optics an ideal communication medium in a noisy electrical environment.

Fiber optics is, however, affected by dispersion and attenuation. The negative impacts of these two factors increase with the number of modes, and limit the length of a given fiber optic link. To limit the effects of dispersion, graded-index fiber can be used. Graded-index multimode fiber design limits the dispersion created by multiple modes by permitting light waves traveling different length paths to travel at similar axial speeds. Attenuation is caused by the following factors: [1]

- Absorption — where propagating light interacts with impurities in the silica glass.
- Scattering — where geometric imperfections in the fiber cause light to be redirected out of the fiber.
- Microbends — which may occur during fiber or cable manufacture.
- Macrobends — which can result from improper installation practice (e.g., exceeding the minimum bending radius).

The extent of the effects of dispersion and attenuation vary depending on the baud rate transmitted. Typical multimode fibers have a bandwidth of 400 MHz-km at a wavelength of 1300 nm. Such a fiber can sustain a baud rate of 400 MHz over a distance of 1 km (0.6 mi) but only 100 MHz over a distance of 4 km (2.5 mi). Traffic control systems that use multimode fiber at the low baud rates typical of distribution channels for small signal control systems will typically be limited by attenuation considerations. [1]

As with the other land-line transmission media, fiber optic cables may be installed underground in a conduit, by direct burial, or overhead on utility poles. Recent traffic control system designs have specified conduit installation of the fiber cables. [1] The proper design of a fiber optic communication system requires a knowledge of the transmission characteristics of the optical sources, fibers, and other devices (such as connectors, couplers, and splices, etc.). Transmission criteria that affect the choice of fiber type used in a system include the following: [7]

- Signal attenuation.
- Information transmission capacity (bandwidth).
- Source coupling.
- Interconnection efficiency.

Interconnection efficiency is usually measured by signal attenuation, the loss of signal strength within the fiber, plus the loss of signal strength in the splices and connectors. [7]

Fiber optics can be installed using single or multimode technology. Single mode fiber boasts higher capacities than multimode fiber and is preferred for long-distance communications. [7] As a result, single mode fiber is recommended over multimode fiber for any communications link that serves as, or could evolve into, part of a backbone network; [1] however, multimode fiber optic systems are sufficient for most transportation related applications. [7]

Splices or connectors are commonly interconnection devices used to install fiber optic systems. To reduce attenuation, interconnection must result in a highly accurate alignment of the fibers so that light can be successfully transmitted between them. Two tools exist for interconnection purposes: splices and connectors. Splices are used to join lengths of fibers together permanently, while connectors are temporary devices that can be disconnected and reconnected as needed. Designs must compensate for the losses introduced by splices and connectors used for interconnection purposes. They also need to provide a closure system that offers environmental protection and mechanical support. [1]

Field splicing of fiber optics is grouped into two methods: fusion and mechanical. Fusion splicing yields lower losses but also requires more time to complete. [2] For temporary connections of individual fibers in controlled environments, connectors and distribution systems are used together. Inexpensive interconnection cables (patch panels) can be used to end the distribution cable. Within these cabinets, small fiber counts of 12 to 48 fibers can be spliced, ended, and stored. Larger distribution frame shelves can be used to end as many as 144 fibers at major hubs. For traffic applications, fibers can be pulled into a breakout box and cable assembly with factory-installed jumper cables on one end and distribution cables on the other. Up to 24 fibers can then be connected from the equipment cabinet to a pull box. [1]

WIRELESS

Several forms of wireless communications technologies commonly used in freeway management systems include the following:

- Areawide radio networks.
- Terrestrial microwave links.
- Spread spectrum radio.
- Cellular radio.
- Packet radio.
- Satellite transmission.

Table 9-16 summarizes the characteristics and features of the wireless communications technology used in freeway management systems. The first three of these communication media are generally owned by the public agency while the last three can generally be leased from commercial providers.

In general, owned communication technologies offer several advantages and disadvantages as compared with commercially leased wireless communication technologies, as summarized in table 9-17.

Area Radio Network

Area radio networks derive their name from their ability to broadcast signals to an area as opposed to a specific location. [(2)] Table 9-18 summarizes the advantages and disadvantages of area radio networks applied to traffic control systems. [(1)]

Area Radio Networks operate in the 150 MHz to 960 MHz bands; the 450 to 470 MHz and 928 to 960 MHz bands are the most commonly used. [(4)] Bandwidth channels in the 25 KHz range are often used for data transfer and can support a signal rate of 9.6K bits/second; however, they will not support video transmission. [(1)] Depending on antenna location, the signals generally radiate uniformly in all directions and rely on the scattering and reflection properties of the signal to propagate throughout the area. During propagation, signals may “bend” slightly over changes in the ground surface, reflect off buildings or other obstacles, or penetrate into buildings.[(1)] Because of potential signal degradation, repeaters are sometimes used. [(2)]

In most applications, radio transmitters use the same transmit and receive channels. Voice communications use receivers all tuned to the same broadcast channel, and users respond only to the messages directed to them. This feature represents a potential limitation in data transfer applications and requires that protocols be established to permit multiple transfers of data on one communication channel. The protocols available vary and include both polling techniques (each data source receives exclusive use of the transmitted channel for a short period) and multiple access contention techniques (compensation is provided for collisions caused by simultaneous transmissions). [(1)]

While the scattering and reflection of radio signals allow the signal to propagate into built up areas, they also reduce the signal strength. Terrain barriers and weather factors can also interfere with the performance of area radio networks.

Operation of area radio networks requires an FCC license. The design of radio communication systems is complex and often requires special expertise. It is recommended that, if a preliminary study shows the usefulness of a radio network, an engineer with the necessary background be retained to evaluate system needs further.

Table 9-16. Wireless Communication Technologies and Their Associated Properties.

Technology	Principal Multiplexing/ Modulation Technique Used	Carrier Frequency Band	Bandwidth/ Channel Bandwidth	Data Rates per Channel	Transmission Range or Repeater Spacing	Government Regulation of Channel or Service	Types of Information Supported	Owned or Service	Constraints on Use
Area Radio Network	Time Division Multiplex, Modulation technique varies	15-174 MHz 405-430 MHz 450-470 MHz 928-960 MHz	25 KHz channels	9.6 Kbps	Several kilometers	FCC licensing of channels for each network	Data	Owned	Channel availability, line of sight in 900 MHz band, multipath sensitivity, geometries
Microwave	Time Division Multiplex, Modulation technique varies	928 MHz to 40 GHz	Various	Up to 7.5 Mbps depending on channel allocation	Varies, may extend several kilometers	FCC licensing of channels except for channels in 31 GHz band for each installation	Data, voice, analog TV, Codec	Owned	Channel availability, line of sight availability, multipath sensitivity, geometries, weather
Spread Spectrum Ratio	Time Division Multiplex, Modulation technique varies	902 to 928 MHz	Various	200 Kbps (typical)	0.8 km (0.5 mi) to several km	No license in the 902-928 MHz band for the network	Data, Codec	Owned	Line of sight, geometries, protocol compatibility
Cellular Radio	Because of narrow channel width, usually only simple polling or other multiple access techniques	825-845 MHz for mobiles, 870-890 MHz for base stations, 20 MHz in reserve	30 KHz per channel	1.2 to 14.4 kbits/sec, depending on modem type used	3.2 - 16 km (2-10 mi) per sector or “cell”	FCC has divided U.S. into 734 service areas, two cellular operators licensed/area	Voice, data	Service	Transmission cost, data transfer limitations
Packet Radio	Multiple access protocol	800 and 900 MHz bands	12.5 and 25 KHz channel widths	4.8 to 8 kbits/sec; 19.2 kbits/sec in major cites	16 - 32 km (10-20 mi) per base station		Data	Service	Service limited to major cities, time delay in delivery
Satellite Transmission	Frequency Shift Keying, Minimum Shift Keying, Phase Shift Keying	C Band (4-6 GHz) Ku Band (2-4 GHz)	Various	Various, depending on system complexity	Essentially unlimited, based on coverage area	FCC allocates frequencies for fixed satellite communication	Data	Service	Availability in some geographic areas, cost

Table 9-17. Advantages and Disadvantages of Owned Radio Communications. [1]

Advantages	Disadvantages
• No need for physical medium since signal propagates through the atmosphere. • No cost of major land-line installations and maintenance. • Used to span natural barriers or provide communications link between points where rights-of-way are not available. • Flexible implementation. • Commercial off-the-shelf equipment available. • Used in a number of traffic control systems.	• Relatively complex design (compared to land-line communication systems) since the local operating environment (e.g., terrain, potential sources of interference, available frequencies, etc.) must be investigated and taken into account for the design process. (However, a variety of theoretical models can be used to predict radio wave propagation for a given set of conditions.) • Limited choices of operating frequencies based on regulatory issues. • Path line of sight constraints (e.g., in the microwave region of 900 MHz and above, line of sight to the receiving antenna(s) generally required). Propagation relationships govern the actual obstacle and adjacent structure clearance required. • Fading considerations. • Turnaround time considerations. • Limited bandwidth. • Requires external antennas and cable. • May require repeaters. • Specialized maintenance.

System designers should consider the following items when installing an area wide radio network: [1]

- Maximum distance to be covered.
- Antenna requirements, including line-of-sight requirements.
- Transmitter power requirements.
- Operating frequency characteristics.
- Receiver sensitivity.
- Line losses at the antenna.
- Required polling intervals.
- Radio transmission considerations such as bending, clearance of obstacles, propagation over bodies of water; and the effects of weather.

Table 9-18. Advantages and Disadvantages of Area Radio Networks. [1]

Advantages	Disadvantages
• Can operate traffic controllers or other traffic control devices. • Can provide voice communications to highway maintenance vehicles. • Can propagate into built up areas and buildings. • Can support 9600 baud data rate. • Can prove cost effective depending on application.	• Terrain may limit range. • Limited channel availability in metropolitan areas. • Requires antenna at each controller site. • Turnaround time excessive for some applications. • Service reliability may limit use of some applications.

Microwave

Microwave systems convey point-to-point messages at very high frequencies that allow for reuse at small distances. [7] The advantages and disadvantages of terrestrial microwave links are summarized in table 9-19.

Microwave systems are expensive due to the infrastructure costs required to interconnect communications. [6] These infrastructure requirements are directly attributable to the operating frequency in use. Higher frequencies require smaller antennas that in turn have fewer extensive infrastructure needs; however, the larger antennas typically required by 2 GHz and 6 GHz systems cost more to purchase, are more difficult to install, and require stronger support structures due to their higher wind loads. [1]

In terrestrial microwave technology, microwave signals are radiated through the atmosphere along a line-of-sight path between highly directional microwave frequency transmitting and receiving antennas. Use of microwave radio allows transmission in both directions simultaneously. [1]

To convey signals, microwave systems use both analog and digital transmission techniques. The analog systems typically use frequency modulation techniques to manipulate the available bandwidth into channels for multiple voice, data and/or video communications. Conversely, digital microwave systems require the use of other modulation techniques including Amplitude Key Shifting, Frequency Shift Keying, and Phase Shift Keying. [1] These techniques do not afford secure transmissions. [7]

Microwave-based “line-of-sight” systems rely on radio waves that travel in paths approximating a straight line. Because of this routing technique, microwave communications are limited by both the effects of the curvature of the earth and the local topography. Interference problems can be overcome to some extent through modification of antenna heights and relative positions. How much compensation can be applied to sources of interference depends on the nature of the interference. For example, as the length of microwave paths increases, the effects of fading also increase due to factors such as atmospheric conditions, ground reflections, and water reflections along the propagation path.

Table 9-19. Advantages and Disadvantages of Terrestrial Microwave Links. [1]

Advantages	Disadvantages
• Useful as a point-to-point trunk. • Can transmit data and a limited number of full motion video channels. • Can control groups of traffic control devices. • Can use both analog and digital transmission.	• Require line of sight path. • In most cases, require FCC license. • Channel availability limited. • May have little choice in operating frequency. • Possible interference due to rain, snow, and atmospheric conditions. • May require antenna tower. • Available bandwidth usually limited.

Rainfall, heavy fog, and other atmospheric factors can reduce the power of the signal below a usable level. This type of fading affects a wide band of frequencies and may last several hours. [7]

Other types of interference vary by season. For example, multipath fading occurs when there is no wind to break up the atmospheric layers. Such conditions are typical of early mornings and nights during the summer season. This type of fading is typically experienced over a limited frequency range and recurs frequently, but lasts only a few seconds. It can be reduced by either "frequency diversity" or "space diversity" techniques, depending on the severity of the problem. [7]

Several aspects of microwave communications should be considered during the design process. The fundamental concepts include the items summarized in table 9-20.

Spread Spectrum Radio

Spread spectrum radio systems operate by transmitting a signal bandwidth over a wide range of the frequency spectrum. To receive such a broadband transmission, the signal is compressed to the original frequency range at the receiver. The major advantages and disadvantages of spread spectrum radio are summarized in table 9-21.

The infrastructure equipment necessary to operate an unlicensed spread spectrum communication system is readily available. Local area traffic control uses will require equipment levels very similar to land-line systems. In contrast, vehicle-to-vehicle communications have no infrastructure requirements. Because of the low-power requirements and general availability of required system components, this communications medium is considered a low-cost technology. [6]

Spread spectrum radio relies on low-power transmitters to provide short-range communications by spreading a signal's transmission spectrum, an application of code division multiplexing. Such systems operate by transmitting a signal bandwidth over a wide range of frequencies and then compressing the signal to the original frequency at the receiver. Each network in an area is operated using a different code,

Table 9-20. Microwave Communication System Design and Installation Considerations. [2,3]

Consideration	Design Factors/Options
System selection factors	• Spectral efficiency requirements. • Performance requirements. • Equipment complexity economics.
Transmission beam is narrowly focused	• Requires precise location of the transmitting and receiving devices. • Antenna towers must be rigid enough to withstand high winds without excessive antenna deflection.
Frequency availability	• FCC allocates a portion of the radio frequency bands of the electromagnetic spectrum for government and public service use. • Most civil government and industrial microwave radio systems operate in the 6, 11, 18, and 23 GHz bands. • Some bands require site licensing, others do not. • Bands above 6 GHz are the easiest to obtain and least costly, but are subject to power fading capabilities.
Repeater spacing	• Geography of a given radio route. • Terminal equipment technology. • Can use point-to-point communications in a daisy chain fashion; path availability and network geometry must be carefully considered. • Transmitter power permitted by the FCC. • Typical repeater spacings in the 6 GHz spectrum are 32 to 40 km (20 to 25 mi), and 1.6 to 4.8 km (1 to 3 mi) in the 23 GHz spectrum; longer spacings are possible when the effects of fading are expected to be minimal.

thereby permitting different networks in the same area to use the same frequency band. [6]

Pseudorandom noise direct sequences (PN) and frequency hopping techniques (FH) are used to provide the necessary expansion of the signal. [6] The PN spreading technique relies on a wide baseband signal to modulate the original baseband signal while the wideband signal's amplitude is continually changed between two states. In contrast, the FH technique uses very short transmissions carried on different frequencies ranging in number from a few hundred to the thousands. [1]

Spread spectrum technology was originally developed by the military to resist enemy radio interception and jamming during World War II. Spread spectrum broadcast techniques allow receivers to decode spread spectrum signals even if background noise levels exceed the signal level and they can resist the effects of multipath interference.

Table 9-21. Advantages and Disadvantages of Spread Spectrum Radio. [1]

Advantages	Disadvantages
• Very flexible installation. • Does not require cable installation and maintenance. • Does not require FCC channel use approval in 902-928 MHz band. • Works extremely well in a high-noise environment. • Currently in use for many industrial process control applications. • Uses low transmitter power. • Can be used in a mixed system of wired or radio interconnected controllers. • No land-line interconnect required. • Relatively low equipment cost. • Potential for broad range of traffic control system applications.	• New technology for traffic control and surveillance applications. • Uncertain range 0.48 to 9.7 km (0.3 to 6 mi); function of area topography. • Higher bandwidth than radio fixed frequency transceivers. • Requires external antenna and cable. • Requires more sophisticated equipment and specialized technicians. • Unprotected channel space.

This technology is considered to have a security advantage over similar wireless transmission techniques; intentional interference with spread spectrum communications is extremely difficult unless the transmission technique is known. [6]

Severe electrical storms have the potential to cause sporadic interference, but the effects are generally short-term. [6] Line-of-sight obstructions can affect transmissions, though obstacles in the 902-928 MHz band generally can be avoided through bending. Though spread spectrum radio itself is noise resistant, it can cause significant interference with other communications systems using the same broadcast frequencies. [1]

Spread spectrum radio currently permits digital communications between equipment in the 200 to 300 Kbps range. This allows uses including communications between a field master and local controllers in closed loop systems, communications between field points and field controllers, and transmission of low-end compressed video. [1] Vehicle-to-vehicle communications can also be supported with this technology. [6]

Fundamental design and installation considerations that should be kept in mind when using Spread Spectrum Radio communications have been summarized in table 9-22.

Cellular Radio

Cellular radio is a new data transmission medium that has grown rapidly since its introduction in 1983. The Federal Communications Commission has designated 666 channels in the 800 MHz band for cellular service purposes. Cellular communication systems are small radio sectors, or “cells” that provide

Table 9-22. Spread Spectrum Radio Design and Installation Considerations. [1,6]

Consideration	Design Factors/Options
System selection and implementation factors	• Maximum peak output power. • Maximum effective radiated power. • Maximum antenna height. • Minimum processing gain. • Favor line-of-sight alignments.
Frequency availability	• Must ensure compliance with FCC regulations. • Potential interference with/from other licensed services.
Operating range	• Varies by application but typically ranges from 0.8 km (0.5 mi) to several kilometers. • Generally used for short-range communications with multiple users. • Repeaters may be required.
Potential for expansion	• Number of users in code division multiplexing systems can be increased by increasing the code length. • Limitations to number of users based on noise levels at receiver. • In FH systems, increasing bandwidth will increase system capacity.

communications coverage in a series of small, slightly overlapping areas. [7] The cells are sized to reflect the density of users in a given area, and typically cover between 0.8 and 8 km (0.5 to 5 mi) each. [6] An antenna is placed in each cell so that, within a given system of cells, antennas relay signals through the system via a series of communications with their closest counterparts. The frequencies used for transmission in one set of cells can be reused in another set of cells by keeping the two signals far enough apart to avoid interference. [7]

With mobile cellular units, the mobile unit establishes a communications link with the closest antenna, and commences communications through that link. The cellular service provider then monitors the strength of the signal link via a mobile telephone switching office (MTSO). The original link antenna will typically be used until the signal power diminishes due to increasing distance and/or interference between the mobile unit and the antenna.

Once the MTSO determines that the signal power has diminished to an established level, the link will shift to the next closest antenna, if that level can better support the link. The communications link then transfers to the antenna closest to the mobile unit. This process of transferring and switching frequency assignments typically requires only milliseconds to complete and often remains undetected during voice communications. [7] In some instances, this transfer process may

result in a "click" that users can hear. In other instances, the system may drop a call if no other antennas are available to support the link.

The advantages and disadvantages of cellular radio are summarized in table 9-23.

The cost for cellular "airtime" varies widely, depending on the provider's service plan and the area of concern. Additional variability in operating costs arises in transmitting data by cellular means and is dependent on the type of modem used. [1] Cellular service can be expensive, depending on usage. Basic service charges apply no matter whether a cellular telephone user originates a call or receives an incoming call. The specification of additional service features beyond the basic operating programs will result in additional fees being assessed.

The costs associated with the use of cellular technology for communications between permanent installations is much higher than those of other communications media. However, it should be noted that cellular radio technology can significantly improve the efficiency of personnel who spend substantial periods traveling in vehicles, and it can be the most cost-effective alternative for temporary communication purposes. [7]

Early cellular communications were handled through analog technology. Currently, conversion programs are underway to transfer operations to digital operating systems to achieve enhanced system quality and capacity. Additional details are offered in reference 8.

In general, noise and interference levels associated with cellular communications typically increase as the distance between the communications source and the nearest receiving tower increase. One major shortfall of cellular technology rests in the fact that it was not originally designed to support data transfer. The switching process historically has caused interruptions that require retransmission of the entire message. Problems with data transmission become especially apparent when operating in the overlap area between two cells. Within this area of overlap, cellular communications may begin "hopping" between the frequencies of the two cells, potentially resulting in data link severance. [8] Consequently, cellular providers often suggest that cellular users do not transmit data via cellular networks while they are in motion. [6] Voice communications are less affected by this form of interference; the audible effects often remain unnoticed. Noise and interference can be reduced through selection of appropriate modems and the use of digital transmission technology. An in-depth analysis of the cellular technologies available is contained in reference 8.

Cellular technology is not currently appropriate for continuous communication service, due to the service costs involved; however, applications that require communications at infrequent intervals or to some remote locations may prove ideal for cellular use. Areas that do not have owned land-line service, those where provision of land-line service is cost prohibitive, and locations of temporary installations may be candidates for communications systems that use cellular technology. [1] Potential traffic control applications include the following: [6]

- Communications with hazardous-materials teams.
- Control of remote equipment.
- Communications with traffic signal/ ramp meter repair crews.

Table 9-23. Advantages and Disadvantages of Cellular Radio. [1]

Advantages	Disadvantages
• May prove cost effective for infrequent communications. • Eliminates need to connect to a telephone company service point or provide owned land-line. • Effective for controlling portable variable message signs. • May prove effective for temporary installations. • Network covers 93% of U.S. population.	• "Airtime" cost excessive for continuous communication service. • Actual data throughput reduced due to protocol overhead. • Remote areas may not have service.

- Communications with motorist call boxes.
- Transmission of dispatch data.
- Control of portable variable message signs/highway advisory radio units.
- Communications with management and other staff members.

During the design and installation process, consideration should be given to the following attributes associated with cellular telephone technologies:

- Local service concerns such as coverage and cost.
- Available data transmission options.
- Compatibility with other equipment (especially important for data transfer).
- Security implications of this communications medium.

A recent report noted that some features of cellular technology as used in portable CMS and HAR applications can be exploited. To prevent the exploitation of portable cellular units, call-back diagnostic features that require the cellular phone on field units to be continuously transmitting should not be used. Continuous monitoring of the telephone bill for each unit (often automatically done by cellular phone companies) is strongly recommended. Further details are offered in reference 9.

Packet Radio

Packet radio services were designed for the transmission of data via wireless means. The distinction between packet radio and the evolving cellular packet data technology is becoming increasingly blurred, except that packet radio does not support voice communications. Table 9-24 summarizes the advantages and disadvantages of packet radio service.

Packet radio operates in the 800 and 900 MHz bands and relies on a system of base stations placed between user terminals to

Table 9-24. Advantages and Disadvantages of Packet Radio Service. [1]

Advantages	Disadvantages
• Designed for data transmission. • Cost effective for short messages. • Can eliminate need for leased or owned land-line. • At least two major providers.	• Not cost effective for continuous communications. • Not cost effective for lengthy file transfers. • Service currently limited to major cities. • Time delay in delivering packet.

convey communications from origin to destination. Radio base stations are positioned on towers or other large structures so that they provide coverage areas of 16 to 32 km (10 to 20 mi). The base stations are connected to backbone communications networks through which they can direct packets of information over one or more frequency pairs. [1]

This technology permits remote terminals to use the same frequency pair for transmitting and receiving data. Multiple access protocol algorithms are employed to reduce the probability of collisions between transmissions from remote terminals. [1]

Packet radio technology is expensive and is not cost effective for file transfers or continuous communication purposes. Specific costs will depend on the service provider and will likely include an initial subscription fee, monthly service fees, and modem and terminal charges. [1]

Noise and interference levels will vary depending on the access protocol used and the provider's quality of service. The packet radio system itself monitors noise and/or interference problems. For example, the ALOHA algorithm transmits data and then waits for the destination base station to confirm that it has received the signal. If the transmission is not acknowledged, the signal transmission is repeated and confirmation is again sought. [1]

The destination of communications can also affect performance issues. Some packet radio service providers offer service adequate for mobile purposes, while their service inside buildings has been characterized as inferior. [8]

The best source of design and installation considerations can be found in project specifications and information available through local packet radio providers. At a minimum, the following factors should be considered in evaluating the feasibility of using packet radio in a freeway management application: [1]

- Number of channels available.
- Channel width.
- Radio frequency protocol.
- Coverage area.
- Raw data transfer rates.
- Modem and terminal requirements.

Satellite Transmission

Satellite communications make use of space-based equipment to relay signals transmitted from earth. Unlike other means of communication, because of the relay equipment's location in space, satellite transmission technology is less dependent on the relative location of transmitters and receivers. [7] Table 9-25 summarizes the basic advantages and disadvantages of satellite communications.

The cost of making frequent satellite transmissions is generally considered excessive in comparison with that of other technologies. As with the other technologies examined, advances in the satellite industry will continue to enhance the economics of communications via satellite transmission. [1]

Satellites in orbit are used to convey signals between earth stations. To transmit a signal, an earth station first receives the data from its source (such as a telephone network) and converts it to a form suitable for radio-frequency transmission. The signal is then amplified to power levels of between one watt and several kilowatts, and is sent to an antenna where it is directed toward the satellite. After being received by the satellite, this signal is then shifted to a down-link frequency band, amplified, and reradiated toward earth. Antennas at various earth stations within the intended coverage area can then receive the transmission from the satellite and, after converting the signal to a usable form, transfer it to other communication media for distribution. [7]

Table 9-25. Advantages and Disadvantages of Satellite Communications. [1,7]

Advantages	Disadvantages
• Cost of circuits independent of their length. • Cost effective for long-haul circuits. • Downlink signals can be received over a wide area. • Cost effective for point-to-multipoint distribution applications (e.g., cable TV). • Uplink signals can originate over a wide area. • Flexibility for "quick setup" or mobile applications. • Provides option to links that cannot be achieved by earth-based communication media.	• Not proven cost effective for local communications. • Limited number of service providers. • Channel leasing costs subject to increases. • Most transportation agencies have no need for the frequent long distance transmission capabilities of satellite technology.

Satellites receive signals transmitted from earth (up-link) on frequencies in one band and then retransmit them back to earth (down-link) on frequencies in a different band. Currently, the C band (5.925-6.425 GHz uplink and 3.7004.200 GHz downlink), and the Ku band (14.0-14.5 GHz uplink and 11.7-12.2 GHz downlink) are designated for use by the FCC. [1]

Satellite access is dominated by two types of networks: preassigned access and demand assigned access. Preassigned access networks, which are best suited to heavy volume trunk service, assign fixed subsets of telephone channels to various destination points. This type of hierarchy is desirable if large subsets are used and load variations among destinations are small. In contrast, demand assigned access is more appropriate for use in communication systems where the demand volumes are lower. This approach makes uses of a pooled collection of channels assigned and used on an as needed basis. [7]

Satellite systems are generally considered to support all-weather operations; however, very severe weather may affect earth-based equipment and interfere with operations. Potential interference problems also exist between some mobile satellite systems and use of the Global Positioning System, and should be investigated for a given application. [6]

Long-haul trucking companies are currently using mobile satellite technology for dispatching purposes. Using inexpensive transmission and reception equipment, messages can be relayed between individual trucks and their dispatching centers, and the truck's location can be monitored. Such technology can also be used for dispatching transportation maintenance crews and emergency equipment from freeway management centers. Unfortunately, these technologies are still being refined and their usefulness in metropolitan areas has been limited due to interference from buildings, trees, and power lines. [7]

Several aspects of satellite communications should be considered during the design process. Table 9-26 illustrates some fundamental concepts that should be considered.

9.4 LESSONS LEARNED

DESIGN PRINCIPLES

Four fundamental principles that should be followed in the design of a communications system for freeway management applications include the following:

1. **Design for "Understandability"** — The communications system should be designed so that its functions, specifications, and processes are easily understood by those who will be responsible for operating and maintaining the system. The underlying principle for achieving understandability is to "Keep It Simple." Simpler designs are easier to operate and maintain. Another general rule-of-thumb is to stay in the mainstream of communications design and not pioneer anything unless it is clearly on the horizon for common communications carriers.

2. **Design for Redundancy** — Many operational problems experienced by existing freeway management systems are caused by communication failures. Designing a communications system where individual field devices or communication hubs can be reached even when a link has been disrupted can help maintain system performance and credibility. Redundancy can be provided

Table 9-26. Satellite Communication System Design and Installation Considerations. [1,7]

Consideration	Design Factors/Options
System selection and implementation factors including data rate and type of satellite	• Long distance communications more cost effective than short distance • High data rates demand strong signal levels at satellite. Requires either: - Low earth orbits; - Higher antenna gain at satellite; or - Higher radiated power transmission level from earth. • Low earth orbits require more satellites to cover a given area continuously • Geosynchronous satellites have lower data transmission rates • Geosynchronous orbits have potential access problems due to local geography • Short transmissions efficiency can be enhanced by using data networks that are better suited than standard gateway connections
Operating frequency	• Availability of desired frequencies • Frequency characteristics vary with propagation effects and equipment performance • FCC regulations govern
Potential for expansion	• Can be achieved by adding satellites • Feeder link traffic limited by FCC transmitter power restrictions

in a system by using dual ring and self-healing types of system architectures. Examples of functional specifications that can improve the reliability of the communication system include the following:

- Provide a redundant communication backbone.
- Limit the number of controllers on one communication channel.
- Provide adequate power margins to account for component degradation without adversely affecting the performance of the system.
- Provide adequate environmental controls (i.e., ventilation, heating, humidity control, pest control, etc.) in field cabinets.
- Provide adequate lightning protection.
- Provide adequate power surge protection.

3. **Design for Maintainability** — An agency can take several steps during the initial planning and design of a communication system that will enhance an agency's ability to maintain the system, including the following:

   - Critically assess the agency's capability to provide timely maintenance of communication equipment and lines. Either a less complex technology or leased communication channels may be preferred if adequate support cannot be provided.
   - Establish a maintenance plan well ahead of the installation of the system. In preparing construction bidding documents for installing the system, specify the type and level of

training that will be provided to the personnel responsible for maintaining the communications system. The training should be specified to occur during the installation of the system.

- Ensure that adequate documentation on the design and operations of the communication system is provided. The documentation should include all communication line connection and controller and field device channel communication assignments.

- Obtain the appropriate quantity and type of communication testing equipment.

- Establish a mechanism to coordinate the use of the right-of-way by utility companies so that construction and maintenance activities do not present a threat to the communication cable and conduit.

4. **Design for Expandability** — Ability to expand the system to meet future growth is a critical feature that can only be incorporated at the initial design phase. Several items that system planners and designers can use to ensure expandability include the following:

   - Use standard design, equipment, and protocols (such as NTCIP).

   - Provide spare communications channel capacity.

OWNED VERSUS LEASED COMMUNICATIONS

Using a leased communications network is an alternative to owning the infrastructure of a communications system. Several forms of leased communications systems are commonly used in freeway management systems, including services available through Local Exchange Carriers (LEC's), cable TV providers, and metropolitan area network providers. [(1)] Leased telephone services fall into three categories: terminal equipment, switching equipment, and transmission facilities. These categories encompass such leased services as office communications service, analog and digital point-to-point services, foreign exchange service, wide area network service, broadband video transmission service, switched digital data service, and channel bank termination service within central offices. [(7)] Several aspects of leased operations should be considered when evaluating their potential use.

Usually, the transmission characteristics and performance of telephone company channels are equivalent to those of a twisted-pair cable network. Because the lessor is paying for a dedicated communications system, they do not have to be concerned with the engineering aspects of the facility. For example, the distance of the transmission line, amplification needs, and other issues should all meet preestablished specifications and tolerances, thereby releasing the lessor of responsibility to maintain the required infrastructure. [(7)] Table 9-27 compares user-owned and leased wire-pair cable. These findings can be easily extrapolated to other means of communication.

The use of leased telephone company channels is most appropriate in areas where a high volume of voice or data communication is required between separate facilities. They may be used in areas where it is not cost-effective to locate other forms of communication technologies (due to physical or political constraints), or in temporary applications. [(7)]

Leased services may also encompass the use of coaxial cable communication systems through cooperation with community

Table 9-27. Comparison of User-Owned and Leased Wire-pair Cable.

Feature	User-Owned	Leased
Capital Cost	High - especially when installed underground in new conduit.	Reasonable - involves connection from access points to cabinets and from central facility to manhole(s) or other cable access.
Ongoing (Maintenance) Cost	Reasonable - owner has full maintenance responsibility.	High - lease costs almost always tied to general service rate increases. Agency must maintain connection described above.
Response to Maintenance Requirements	May be better or worse than leased line service depending on jurisdiction's capability.	In some cases can be specified.
Maintenance Responsibility	Always borne by the owner.	Sometimes difficult to determine whether problems are the responsibility of the lessor or lessee.
Control of Communication Network	Excellent - owner is user and can fully control.	Less control - communication traffic requirements and limitations on leased channel may impede use.
Design of Communication Network	Completely flexible since designer/owner starts from scratch.	Constrained by telephone company cable network. In most cases it is generally compatible with signal system requirements.
System Expansion	Cost determined by location of expansion.	Accessibility to expansion site usually available.
Reliability	Excellent - with proper maintenance procedures.	Good to poor - depends upon lease contract terms and level of owner's maintenance function.

antenna television (CATV) providers. Franchise agreements sometimes require CATV suppliers to provide such services.[(2)] Several features of CATV result in both advantages and disadvantages, as summarized in table 9-28.

After deciding on the requirements of the leased system, various configurations can be explored to evaluate system cost. Typical cost evaluations will include providers' fees and charges, costs associated with constructing any connections between the provider's termination point and the lessor's equipment, and any special hardware costs

Table 9-28. Advantages and Disadvantages of CATV.

Advantages	Disadvantages
• Single 6 MHz channel adequate for data transmission. • Network already in place. • Design efforts and initial installation cost much lower than a dedicated coax system. • Franchise agreement may provide for government use of CATV cable and bandwidth at reduced rates or free, reducing recurring costs. • Second separate coax institutional network (I-net) may exist for the express purpose of providing bi-directional services to commercial subscribers. • I-nets generally provide good levels of service to subscribers.	• Most CATV networks designed and installed with emphasis on downstream transmission of video signals to cable subscribers. • Video channels take up most available bandwidth. • Bandwidth available to traffic control may be very narrow, ranging from a single 6 MHz channel to 4 or 5 channels. • Single 6 MHz channel does not support full motion video transmission in addition to data communications. • Frequencies of available channels often least desirable in susceptibility to noise and interference. • Quality of video signal required for CATV considerably less than required for data. • CATV subscribers sometimes concentrated in residential areas. • Service to CBD and industrial areas sparse or nonexistent. • Area of coverage and network layout may not coincide with control equipment locations. • Traffic control system may have to compete with other public I-net users for more desirable channels.

required for interfacing with the provider's equipment. [(1)]

To illustrate the items that must be evaluated, consider leased telephone lines. These lines typically have a low initial installation cost as compared to an owned network of similar size, especially if underground installation is required. However, the monthly lease cost may be much higher than the recurring costs of an owned network. Furthermore, because telephone company rates and charges are established by tariff and are subject to approval by State regulatory agencies, future costs cannot be guaranteed by the telephone company.

The leasing tariffs and the method by which they are calculated vary from State to State. Some tariffs are based on the "drop-to-central office mileage" (the distance from one drop [traffic signal controllers, a ramp meter, etc.] to the telephone company central office). Other tariffs use "drop-to-central office fixed," in which a fixed charge per drop is assessed, whatever the distance. Additional charges are usually assessed when communications between central offices are required in the path from field to central equipment. These charges may be the most significant element of the total cost. Alternatives to leasing communication lines for great distances through multiple central offices should be explored. [3]

The transmission technique in use depends on the communication medium that is being leased. Local telephone companies (often called local exchange carriers, or LECs) offer both voice grade service and digital private line data services. Such systems support communications between control centers and field devices or maintenance shops through synchronous transmission speeds of 2.4, 4.8, 9.6, or 56 Kbps. LECs also provide coded video transmission capabilities via higher capacity digital transmission channels such as T1. [2] Most digital transmission facilities provided by telephone companies are based on the use of digitized voice channels. These providers typically do not use the same conversion process as modems to transform voice information to digital transmissions. Reference 2 provides additional details concerning this subject. Other communications vendors offer leased transmission techniques including high capacity DS1 and DS3 circuits, voice grade, digital data service, and fractional T1 circuits. [2]

In certain circumstances, leasing fiber optic networks should also be considered. Leasing of fiber optic technology would be most appropriate for traffic control applications that require high-capacity links.[2]

Interference encountered in leased systems varies depending on the communications medium used and the service provider. Specifications should be made in the lease contract about how much noise and interference are permissible.

The limit of the potential traffic control applications for leased technologies is dependent on the technology that is being used. Leased twisted-pair telephone lines are frequently used for traffic surveillance and control system applications such as the interconnection between a central computer and remote field equipment (e.g., traffic signal controllers, variable-message signs, freeway ramp meters, traffic counting equipment, etc.). [7] Other potential applications information can be found listed under each of the technology sections presented in this module.

Leased communications require agreement on a contract between the lessor and the service provider. The development and use of a leased communication system require close cooperation with the service provider at all levels of personnel including engineering, sales, and data transmission representatives. [7] At a minimum, the concerns summarized in table 9-29 should be addressed.

NTCIP

The National Transportation Communications for ITS Protocol (NTCIP) is a standard communication protocol for transmitting data and messages between electronic devices used in ITS, freeway management systems, and arterial traffic control systems. It uses the OSI (Open

Table 9-29. Leased Communications Service Design and Installation Considerations.

Consideration	Design Factors/Options
Desired system features	• Data characteristics including desired transfer rates and bandwidth required. • Acceptable level of noise and interference. • Operational aspects of the system (e.g., once-per-second/dial-up communications, full/half duplex, etc.). • Number of field drops and their locations (e.g., controller/ ramp meter cabinets).
Characteristics of the leased channels and equipment restrictions	• Type and number of channels to be provided (insure both up and downstream channels are provided for CATV applications). • Data transfer rates. • Transmission characteristics and quality. • Frequency allocations. • Any system limitations.
Availability of service	• Time required to initiate service. • Facilities the provider will use to supply service to field locations. • Future expansion capability. • Location of the nearest access point to the provider's transmission line from each of the lessor's field locations.
Understanding of contract requirements	• Rules regarding leased circuit terminations at the central facility, including any equipment which the jurisdiction must furnish. • Division of work between the provider and user in connecting to field locations (i.e., is complete service connection provided or will lessor need to make connections to the provider's equipment). • Respective maintenance responsibilities of the provider and the jurisdiction. • Maintenance response priorities. • Testing criteria and requirements prior to acceptance. • Rates and charges; both one-time charges and monthly fees. • Pending rate increase requests, if any, and an estimate of expected increases in leasing rates. • Expiration/renewal terms.

Systems Interconnection) Reference Model for providing communications links between devices. [10,11] The OSI Reference Model is based on a concept developed by the International Standards Organization (ISO). The model is called the ISO OSI Reference Model because it deals with connecting systems that are "open" for communications with other systems. The OSI model consists of seven hierarchical layers. Each layer performs a related subset of the functions required to talk with another system. Each

layer relies on the next lower layer to perform other functions and to conceal the details of those functions. It in turn provides "services" to the next higher layer. These layers are defined in such a manner so that changes in one layer do not require changes in the other layer. Table 9-30 identifies each layer and gives a brief description of the functions associated with each layer.

NTCIP is a common set of codes, procedures, and relative timing relationships (i.e., communications protocols) that follow the OSI model to allow different types of field devices to communicate with one another without the need of a physical conversion box or reliance on a single vendor to supply all the equipment in the system. [11,12] Rather than defining a unique format for each combination of data items that need to be communicated, the NTCIP defines a structure that uses a modular approach for delivering generalized messages. By defining a series of message pieces that can be collected into groups of data items, system software designers can create virtually any useful message structure without requiring new programming in the field device. NTCIP promotes two important features that could not previously be provided in traffic control systems:

- *Interoperability* — the ability to connect devices of different types to the same communications medium.

- *Interchangeability*—the ability to exchange devices of the same type, but from different vendors, without any lossof functionality or modification to the rest of the system.

To date, four profiles have been developed for different traffic management applications.[10,12] A profile defines the standard communication protocols that are appropriate for the different types of traffic management applications. The Class A profile allows messages to be routed from one device to the next. The Class B profile allows devices to talk with one another over a low speed, low capacity communications system where routing is not required. Class B can be used where legacy communications exist, typically twisted-pair copper wire. The Class C profile supports file transfer and sequencing and routing, while the Class E profile provides routing, sequencing and file transfer capabilities over a point-to-point communications link.

Table 9-31 shows the specifications applicable in each layer of the OSI for the different classes. Additional NTCIP profiles will be developed to satisfy various system architectures and communications media as the need arises.

Besides defining how communications between devices will occur, the NTCIP also defines what information is contained in a data stream (i.e., object definitions). The NTCIP uses a hierarchical method of addressing where data elements are stored in a database, similar to that used widely on the Internet. [12] The hierarchical method allows data to be grouped into logically related data elements. Each of these logical groupings can be expanded over time as new elements are added to the group, without running out of addresses in the group or limiting other groups to fewer addresses. Currently, object definitions have been developed for traffic signals, but object definitions for other traffic management elements (i.e., variable message signs, CCTV camera controls, ramp meters, vehicle detection/count stations, highway advisory radio, etc.) are under development.[13]

More information about NTCIP can be obtained through the Internet at http://www-atms.volpe.dot.gov/ntcip/.

Table 9-30. Layers of Open Systems Interconnection Reference Model. [10]

OSI Layer	Description
Layer 7: Application	Defines procedures for file transfers, access methods, and management of messages.
Layer 6: Presentation	Defines the syntax and semantics of transmitted information. It provides services such as encryption, text compression, and reformatting.
Layer 5: Session	Defines the procedure for different communications equipment to establish dialogues. Provides the control structure for communication between applications. It establishes, manages, and terminates connections (sessions) between cooperating applications.
Layer 4: Transport	Defines the organization of data passing to and from the lower layers. Involves breaking longer messages into packets for transmission.
Layer 3: Network	Defines how packets of data are routed from source to destination. Provides upper layers with independence from the data transmission and switching technologies used to connect systems. It is responsible for establishing, maintaining, and terminating connections.
Layer 2: Data Link	Defines methods for ensuring data integrity such as error correction. It sends blocks of data (frames) with the necessary synchronization, error control, and flow control.
Layer 1: Physical	Defines the mechanical, electrical, and procedural characteristics used to establish, maintain, and deactivate the physical link.

SHARED RESOURCE PROJECTS

The need for increased communications capabilities has led many agencies to consider shared resource projects. Shared resource projects refer to those projects where the private sector would use public highway rights-of-way, previously viewed as entirely within the public domain, to install telecommunications hardware (principally fiber-optic lines, but also including cellular towers). In return, the private sector would "compensate" the public sector, either through barter or cash. In barter or in-kind arrangements, private parties install the communications system. In return for providing telecommunications capacity and/or services to the public agency, they receive access to the right-of-way for installing their own capacity. In cash arrangements, private parties install the telecommunications system, receiving access to the right-of-way in return for monetary compensation to the public agency. Hybrids of the barter and cash alternatives can also be created in which in-kind (communications capacity) and monetary compensations are combined as consideration for private access to right-of-way for private sector use. [14]

At the outset, potential shared resource partners must address the threshold legal and

Table 9-31. Specified Protocols for NTCIP Profile Classes. [10]

OSI Layer	Profiles			
	Class A	Class B	Class C	Class E
Application	STMF	STMF	Telnet FTP SNMP	Telnet FTP SNMP
Presentation	Null	Null	Null	Null
Session	Null	Null	Null	Null
Transport	UDP	Null	TCP	TCP
Network	IP	Null	IP	IP
Data Link	PMPP	PMPP	PMPP	PPP
Physical	EIA 232E FSK	EIA 232E FSK	EIA 232E FSK	EIA 232E FSK

STMF = Simple Transportation Management Framework
SNMP = Simple Network Management Protocol
FTP = File Transfer Protocol
UDP = User Datagram Protocol
TCP = Transmission Control Protocol
IP = Internet Protocol
PP = Point-to-Point Protocol
PMPP = Point-to-Multipoint Protocol
EIA 232E = standard modem interface
FSK = Frequency shift keying

political issues that, if unresolved, can preclude shared resource arrangements. Some of these issues are summarized in table 9-32. In some situations, new statutes or regulations will be required to permit private sector access to right-of-way or to conduits in the right-of-way. In other cases, careful contractual arrangements can ensure effective private sector longitudinal access without violating legal or regulatory constraints; for example, using leasing rather than easements to convey rights. [14]

Most other issues can be addressed without resorting to legislative or regulatory changes. For example, the issue of bond tax exemption can probably be addressed through attention to the ways that bond issues are structured and bond proceeds are used. In some cases, the number of options available is limited by regulation, resulting in little (or even no) choice but to pursue legislative changes (e.g., where cash compensation is precluded by statute). Even under these circumstances, shared resource projects can be effective without legislative initiatives so long as the potential partners are willing to accept an option that is within currently accepted boundaries. [14]

Often, the choice made among ways of addressing these issues is based on preferences and an evaluation of the pros and cons of each option. The choices between exclusivity and multiple partners, allocation of infrastructure relocation and liability for repairs, and selection of barter over cash compensation are often based on preference rather than necessity. [14]

Table 9-32. Issues Associated with Shared Resource Project Development. [14]

THRESHOLD LEGAL AND POLITICAL ISSUES	
Public sector authority to receive and/or earmark compensation	The public sector may be precluded from receiving cash payments, but may still be free to engage in barter arrangements, particularly if they are structured as procurements. In general, state departments of transportation (DOTs) have less flexibility; municipalities and authorities such as turnpike and transit agencies have greater flexibility in dealing with cash flows.
Authority to use public right-of-way for telecommunications	Shared resource arrangements may be precluded if state law mandates free access for utilities or if pubic agencies are not allowed to discriminate among utilities (e.g., permit access for telecommunications but disallow access for gas and sewerage).
Authority to participate in public-private partnerships	Because shared resource arrangements are a form of public-private partnering, legal authority to enter into such agreements is a basic requirement. In some cases, "implied authority" is not considered sufficient and specific legislation or "express authority" must be passed.
Political opposition from private sector competitors	Shared resource arrangements may trigger political opposition, though not necessarily prohibition, from private sector companies resisting the establishment of bypass networks that they perceive as competing with the services they offer. Opposition may be slight when the bypass system is limited to transportation needs, but it is likely to be stronger if the system supplies a greater range of public sector communications needs.
Inter-agency and political coordination	In addition to investing effort in coordination among agencies in the same political jurisdiction, the lead pubic agency may also have to orchestrate agreements between geographically proximate political jurisdictions to ensure continuity for fiber for their private partner(s).
Lack of private sector interest in shared resources	At its core, shared resource arrangements depend on private sector interest in expanding telecommunications infrastructure. Reluctance to enter into partnerships with pubic agencies for access to right-of-way may stem from insufficient market demand for increased communications capacity, cost factors such as more stringent installation specifications along roadway right-of-way, and administration or managerial burden of compliance.
FINANCIAL ISSUES	
Valuation of public resources	Before entering into shared resource agreements, the public sector needs to have some idea of the value of the assets it brings to the partnership; that is, continuous or sporadic access to its right-of-way for placement of private (communications) infrastructure.
Tax implications of shared resource projects	Partnerships between public and private entities may pose unique tax issues, particularly bond eligibility for tax-exempt status when proceeds may benefit profit-making private organizations.
Valuation of private resources	Valuation of the private resources provided in barter arrangements helps the public sector determine whether it is receiving a fair market "price" for its resource.
Public sector support costs	Although shared resource arrangements provide cash revenue or telecommunications infrastructure without public sector cash outlays, such compensation is not without cost since the pubic sector must use agency labor hours for administration, coordination, and oversight.
PROJECT STRUCTURE ISSUES	
Exclusivity	Shared resource arrangements may limit access to pubic right-of-way to a single private sector partner in any specific segment, that is, grant exclusivity. From the public sector point of view, exclusive arrangements have both advantages (administrative ease) and disadvantages (potential constraints of competition among service providers, lower total compensation received by the public sector).
Form of real property right	Shared resource arrangements can be structured in any of several legal formats (easement, lease, franchise, license) with variations in the property rights conveyed. Moreover, the property right may involve access to the right-of-way itself for privately owned infrastructure, or be limited to access to (or use of) publicly owned infrastructure.
Type of compensation	Compensation to the public sector may be in the form of goods (in-kind), cash, or combinations of both. Moreover, in-kind compensation can include not only basic fiber-optic cable but also equipment to "light"* fiber, maintenance, and even operation and upgrading.
Geographic scope	Projects can be extensive in scope, covering long segments of roadway, or more focused on specific areas. The option that is best in any individual context depends on other factors, such as considerations of administrative burden, service interests of potential bidders, and private willingness to install infrastructure in an area larger than their primary area of interest.

*"Light" fiber refers to fiber optics supported by equipment for transmission and receipt of communications signals; "dark" fiber refers to physical fibers devoid of supporting telecommunications equipment.

Table 9-32. Issues Associated with Shared Resource Project Development (Cont'd). [14]

CONTRACT ISSUES	
Relocation	Allocation of responsibility for infrastructure relocation in case of roadway improvements affects private partner willingness to pay for right-of-way insofar as it carries a financial responsibility as well.
Liability	Similarly, allocation of legal liability among partners affects the financial risks assumed by each one. Liability includes responsibility for system repair, consequential damages (economic repercussions), and tort actions.
Procurement issues	Shared resource arrangements face many of the same issues as other procurements regarding selection and screening of private vendors or partners.
System modification	Shared resource arrangements may or may not include explicit provisions for system modification; that is, technological upgrading to keep abreast of technical improvements and expansion of capacity to meet subsequent needs.
Intellectual property	Intellectual property involves intangible components (e.g., software programs) of the operating system that might not be available to the public sector partner when the partnership is dissolved after the lease period unless specifically addressed in the contract.
Social-political issues	Social-political issues involve equity among political jurisdictions or population segments within the right-of-way owner's domain. More specifically, two issues may affect how shared resource arrangements are structured: most-favored community issues — comparable compensation for all communities engaging in shared resource arrangements, and geographic and social equity — equitable access to and benefit from shared resource arrangements.

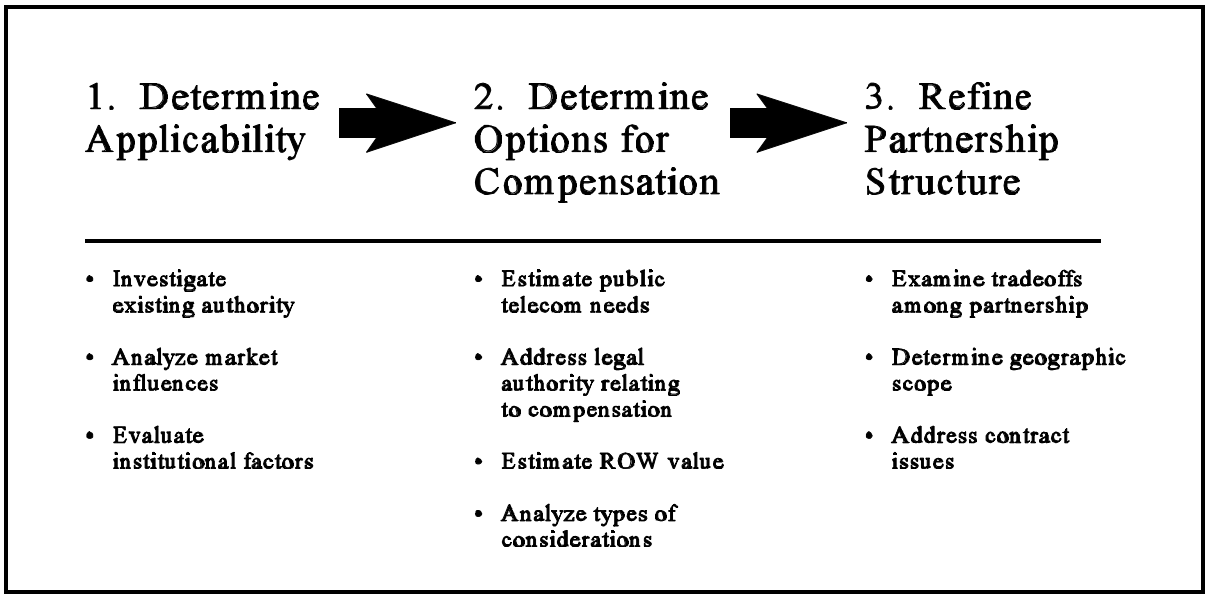


Figure 9-13. Key Decisions and Supporting Information in Evaluating the Need for a Shared Resource Project. [14]

Although shared resource projects have been heralded as an innovative approach to satisfying public sector needs, they are only one of several ways to provide for these communications needs. They are hardly the universal solution. Before initiating a shared resource project, public agencies must evaluate their telecommunications needs (including private sector-supplied services), and then evaluate the appropriateness of

each alternative in light of specified needs. Figure 9-13 graphically summarizes the basic stages in developing a shared resource project. [14]

In addition, the public sector needs to realize that shared resource partnering is market-driven. Because it is market-driven, an upper limit exists to how much compensation will be paid. Furthermore, private agencies will not wait extended periods for access rights to a public agency's right-of-way. Therefore, public agencies must be ready to act when the opportunities present themselves.

The highway right-of-way has no "inherent" value. Concerning telecommunications access, value is derived from the potential revenue telecommunications companies can gain by using the right-of-way as tempered by the cost of other rights-of-way that might be available at that time.[14]

Similarly, market conditions dictate response time for prospective partnering. As market forces and technology change, demand for access to highway rights-of-way may also change. In fact, timing can be a critical factor in the choice among options in structuring a shared resource project. Because the window of opportunity is often narrow and because private sector partners can have access to non-highway right-of-way for infrastructure, public agencies interested in effecting shared resource partnerships must address the associated issues in a timely fashion; otherwise, a public agency may have to wait until market expansion or industry restructuring generates new demand for right-of-way use.[14]

For more information about the issues associated with shared resource projects, the user should consult *Shared Resources: Sharing Right-of-Way for Tele-communications.* [14]

TRANSIENT PROTECTION

Electronic traffic control equipment is subjected to a wide range of electromagnetic threats, including lightning, electrostatic discharge, internally and externally generated inductive switching transients, and radiated electromagnetic interference (EMI) from radio, TV, radar, and mobile communication transmitters. Lines providing electrical power and cable interconnecting equipment to sensors, communications systems, or peripheral hardware provide a direct path for the conduction of disruptive and damaging electrical transients and other EMI into traffic control equipment. In general, traffic control and communications systems can be protected against transients through the following mechanisms: [15]

- Proper grounding of power supply and communications systems.
- Adequate shielding of equipment and interface cables.
- Proper bonding and corrison control.
- Use of terminal protection techniques such as amplitude limiters and filters.
- Good interface designs that use balanced line inputs and fiber optics.

The most severe electromagnetic threat to traffic control equipment is lightning. To determine the level and extent of protection required against lightning-induced transients, the traffic control professional must know the lightning threat levels at traffic controller inputs, the rates-of-occurrence of lightning transients, and the effects of the lightning discharges on electronics within traffic control equipment. The lightning levels will vary depending on the type of interface, the length of the interconnecting lines or cables, and whether the lines are shielded or not.

NCHRP Report 317 Transient Protection, Grounding, and Shielding of Electronic Traffic Control Equipment describes techniques for protecting against transient and electromagnetic interference.[15]

COMMUNICATIONS WITHIN THE CONTROL CENTER

Providing adequate communications systems within the control center is critical to the success of the control center. In general, the types of communications systems implemented in a control center need to allow operators to communicate with the following personnel:

- Other control center operators.
- Supervisory personnel.
- Maintenance dispatchers/forces.
- Computer system operators/ maintainers.
- Incident management personnel.
- Public relations personnel.
- Media.

In general, the types of communication systems that need to be included in the design of the control center include the following:

- Local area computer networks where information and data can be transferred from one computer terminal to another.
- Voice communications systems

Module 10 provides more information about these types of communications systems.

9.5 REFERENCES

1. Gordon, R.L., Reiss, R.A., Dunn, W.M. and Morehead, D.R. *Communications Handbook for Traffic Control Systems*. Report FHWA-SA-93-052. FHWA, U.S. Department of Transportation, Washington, DC, 1993.

2. Gordon, R.L., et. al. *Traffic Control Systems Handbook*. Report FHWA-SA-95-032. FHWA, U.S. Department of Transportation, Washington, DC, 1996.

3. Neudorff, L.G. and Terry, D.C. *Communications in Traffic Control Systems: Final Report*. Report FHWA-RD-88-012. FHWA, U.S. Department of Transportation, Washington, DC, 1988.

4. *Building the ITI: Putting the National Architecture in Action.* FHWA, U.S. Department of Transportation, Washington, DC, 1996.

5. Seymour, E.J. *Standards — NTCIP*. White Paper from the National Conference on Operating and Maintaining Advance Traffic Management Systems (ATMS) Centers. Institute of Transportation Engineers, Washington, DC, 1996.

6. Task C Report. *Assessment of Candidate Communications Systems and Technologies For Use With Intelligent Transportation Systems*. Prepared for FHWA Turner Fairbank Research Center by ARNIC, 1995.

7. Tarnoff, P.J., and Pugh, T. *Transportation Telecommunications.* NCHRP Synthesis of Highway Practice 165, TRB, National Research Council, Washington, DC, 1990.

8. Elliot, S.D., and Dailey, D.J. *Wireless Communications for Intelligent Transportation Systems*. Artech House, Inc., Norwood, MA, 1995.

9. Brehmer, C.L. Addressing Security Issues in ITS/ATMS, *Compendium: Graduate Student Papers On Advanced Surface Transportation Systems*, Texas Transportation Institute, Report Number SWUTC/96/72840-00003-1, August 1996.

10. Seymour, E.J. *Standards — NTCIP*. White Paper for the National Conference on Operating and Maintaining Advance Traffic Management Systems (ATMS) Centers. Institute of Transportation Engineers, Washington, DC, 1996.

11. The National Transportation Communications/ITS Protocol: An Introduction. In *ITE Journal*, Volume 65, Number 12, Institute of Transportation Engineers, Washington, DC, 1995.

12. The NTCIP - Protocol Details. In *ITE Journal*, Volume 66, Number 1. Institute of Transportation Engineers, Washington, DC, 1996.

13. NTCIP — How It Can Be Used. In *ITE Journal*, Volume 66, Number 4. Institute of Transportation Engineers, Washington, DC, 1996.

14. Jakubiak, S. and Relin, A. *Shared Resources: Sharing Right-of-Way for Telecommunications.* FHWA, U.S. Department of Transportation, 1996.

15. Denny, H.W, and Rohrbaugh, J.P. *Transient Protection, Grounding, and Shielding of Electronic Traffic Control Equipment.* NCHRP Report 317, TRB, National Research Council, Washington, DC, 1989.

MODULE 10. CONTROL CENTER

TABLE OF CONTENTS

10.1 INTRODUCTION 10-3
MODULE OBJECTIVE 10-5
MODULE SCOPE 10-5

10.2 DECISION PROCESS 10-5
PROBLEM IDENTIFICATION 10-5
IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING 10-7
ESTABLISH GOALS AND OBJECTIVES 10-7
DEFINE FUNCTIONAL REQUIREMENTS 10-8
DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS,
AND INFORMATION FLOWS 10-11
Function Allocation 10-11
Task Analysis 10-13
Data Flows 10-15
IDENTIFY AND SCREEN TECHNOLOGIES 10-15
Physical Environment 10-15
Workstations 10-17
Controls and Displays 10-17
User Interfaces 10-18
EVALUATION 10-19

10.3 TECHNIQUES AND TECHNOLOGIES 10-20
PHYSICAL DESIGN 10-20
Lighting 10-24
Acoustics 10-25
Heating and Cooling 10-25
INFORMATION DISPLAYS 10-26
Static Wall Maps 10-26
Projection Television Systems 10-26
Video Walls 10-26
Design Guidelines and Recommendations 10-28
OPERATOR WORKSTATION DESIGN 10-28
COMMUNICATIONS SYSTEMS 10-29
Local Area Networks 10-30
Components 10-30
Planning for an LAN 10-31
Voice Communications 10-32

STAFFING 10-34
Qualifications 10-36
Training 10-36
Staffing Sources — Agency versus Outsourcing 10-38
SECURITY 10-38
AUTOMATION 10-39
JOINT OPERATIONS 10-40

10.4 LESSONS LEARNED 10-42
SYSTEM CAPABILITIES 10-42
OPERATIONS MANUAL 10-42
EFFECTIVE USE OF SYSTEM OPERATORS 10-42
SYSTEM MAINTENANCE 10-43

10.5 REFERENCES 10-45

MODULE 10. CONTROL CENTER

Figure 10-1. The Control Center of the TranStar Freeway Management System in Houston, TX.

10.1 INTRODUCTION

The control center is the hub or nerve center of a freeway management system. It is where data about the freeway system is collected and processed, and fused with other operational and control data to produce "information." The information is then used by system operators to monitor the operations of the freeway and initiate control strategies to affect changes in the operations of the freeway system. It is also where agencies can coordinate their responses to traffic situations and incidents. Furthermore, the control center is also the focal point for communicating traffic flow information to the media and the motoring public. Because of its critical role in the successful operation of a freeway management system, it is essential that the control center be designed to allow operators to control and manage the functional elements of the freeway management system.

All too often, system designers have viewed technology and/or automation as the design goal for the control center. Unfortunately, this view is not supported by experience, and can cause operational problems when it is allowed to bring undue influence on the design of the control center. In one operational control center, for example, expert design resulted in an automation

scheme in which the support system would determine the "best" pattern of variable message signs to present motorists with warnings of wrong-way drivers in the travel lanes. Under the design principle, the operator would review the automatic decisions and could change them after they were initiated, if necessary. After the system became operational, however, it was found that the surveillance system could not detect the difference between roadway maintenance equipment working in a closed lane and wrong-way drivers in the travel lane. The error resulted in the automatic closure of a large area of a freeway. This problem arose from a faulty *allocation of functions* to a higher level of automation than was feasible, a common design flaw. Its primary cause was that the design emphasis was on technology rather than on the performance of the total system. [1]

One way to avoid such design problems is to begin with a user-centered approach to designing a control center. [1] A user-centered approach applies system engineering and human factors principles to developing a system design that focuses on what the system is supposed to accomplish rather than on technology. The distinguishing features of a user-centered design philosophy as compared with other approaches include the following: [1]

- **The design focus is on the operator, not the designer.** In the user-centered design, the user (i.e., the operator) is viewed as a critical system component. The characteristics, capabilities, and limitations of the user need to be defined and considered during the requirements analysis and design of the system. Ideally, the user should be involved in the design process at its earliest stages and this involvement should continue throughout the design and testing phases.

- **The selection and acquisition of system components are based on validated functional requirements.** Many control centers have been designed adhering to a *technology-centered design approach* where the focus in designing the system has been on hardware and software. The user-centered design focuses on *what* needs to be accomplished by the system, not *how*. Therefore, system components are selected on the basis of an analysis of the functional requirements of the system. A *validated requirement* is a functional capability (e.g., identify sensor locations at which occupancy exceeds 30 percent during a 2-minute period) that has been formally stated and objectively evaluated.

- **The process is iterative.** Systems are best developed through an *iterative* process, in which a design is tested and validated in a series of stages. This is particularly important in control center design, where the multiple iterations can uncover problems — and opportunities — that are not apparent until they are viewed in the total system.

- **The process extends throughout the life cycle of the control center.** The fact that a new control center has been built and put into operation does not suggest that it is "complete." As the managers and operators of the control center become familiar with the system, they will make recommendations for adding many excellent design features and procedural changes to improve their abilities to control traffic on the freeway.

A user-centered design approach applies system engineering principles to develop a system that views the user (i.e., the operator) as critical to the overall success of the system. Figure 10-2 illustrates how the user-centered design process can be used to identify the elements of a control center. The first step in the design process is a mission analysis, which uses the goals and objectives of the system to define what it is the control center is supposed to do. After the mission analysis, a functional analysis identifies which functions will be performed by various elements (both human and technological) in the control center. Once the functional analysis has been completed, the tasks needed to accomplish each function can be analyzed. The task analysis involves breaking tasks down into their individual elements to include a detailed description of both manual and mental activities that the operators must do, and their duration, frequency, and complexity. The task analysis leads directly to defining operational characteristics and needs for the control center.

MODULE OBJECTIVE

The goal of this module is to identify some major factors and issues that should be considered in the planning and design of a control center. The specific objectives of this module are to identify the some major issues that must be considered designing, operating, and maintaining a control center for a freeway management system.

MODULE SCOPE

The purpose of this module is to help system planners and designers identify some major issues and factors affecting the design and operations of a control center for a freeway management system. It focuses on how the characteristics, capabilities, and limitations of the operators should be accounted for in the design of the control center. This module is not intended to be a detailed discussion of the human factors requirements affecting the design of the control center. For detailed information about the human factors requirement, the user should consult the *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition) or other references. [1,2,3]

10.2 DECISION PROCESS

PROBLEM IDENTIFICATION

The first step in the design process for a control center is to identify its purpose or role in the overall freeway management system. The functions of a control center can vary from location to location, depending upon the local operating goals and philosophies for the freeway management system. For example, some agencies may be more interested in disseminating information about traffic and travel conditions to travelers, while other agencies may want to influence traffic behavior through control mechanisms, or to manage incidents by closely interacting with other agencies. The design of the control center needs to support the desired operating philosophy. An agency whose primary goal is information dissemination may want to design a control center that allows easy access to the information in the system by users outside the control center. An agency that wants close interaction with other operating agencies (fire, police, emergency services, etc.) may want to provide a physical location in the control center for those agencies to dispatch their emergency responses. Common roles/functions of a control center in freeway management systems include the following:

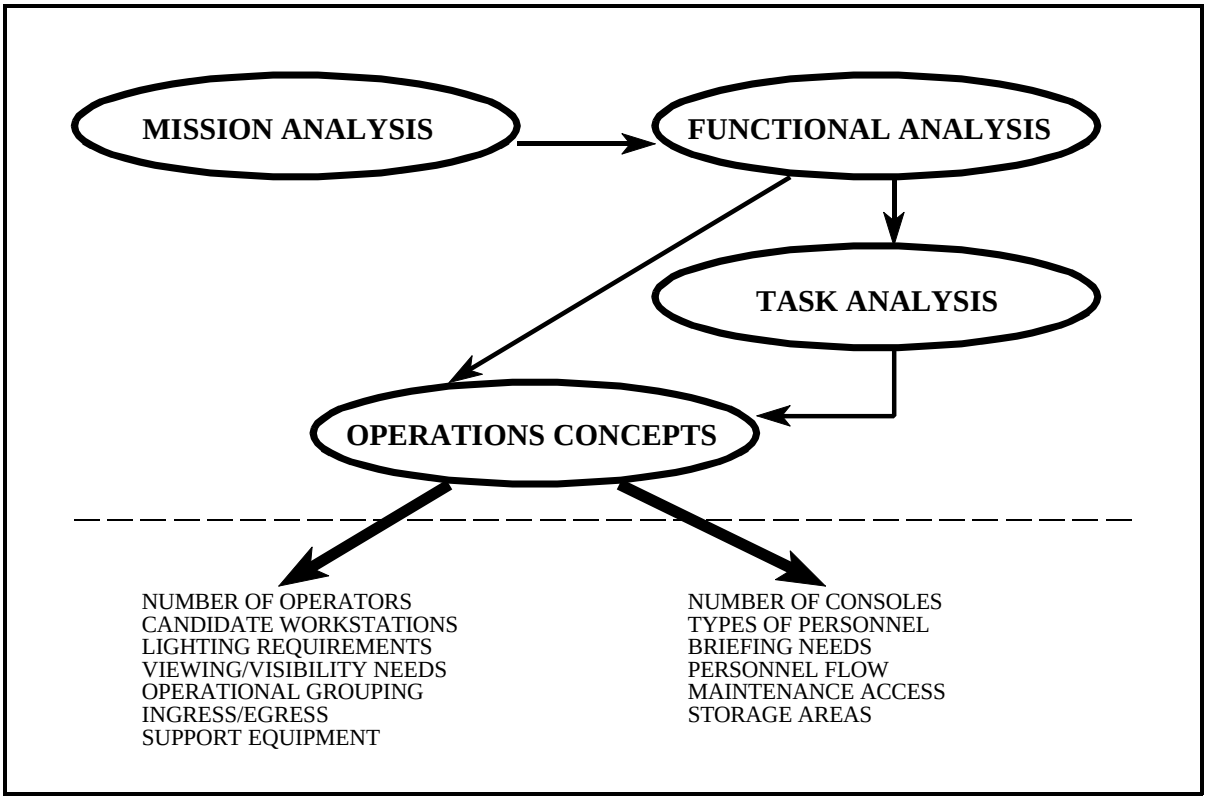


Figure 10-2. A User-Centered Design Approach. [1]

- A location for coordinating and implementing freeway management strategies and controls.
- A dispatching center for incident management and maintenance forces.
- A location for doing maintenance and repairs of malfunctioning or damaged field equipment.
- A central location for distributing freeway traffic and travel information to travelers, elected officials, and the media.
- A command post for coordinating the response to major emergencies.

Designing and constructing a control center can be costly (often millions of dollars). In the past, some agencies have found that their control center facilities have been inadequate to satisfy future growth. Planning accordingly for system expansion is important for agencies. Systems can expand their operations in many ways, including the following:

- Increase the number of freeways and/or roadway facilities covered.
- Add new freeway management functions to the existing functions of the system.
- Permit joint or cooperative operations of several agencies from one location.
- Serve as command post for major emergencies.

With these potentials in mind, agencies can plan for future expansion in the control center by providing the following:

- Adequate space in the operations room to install additional operator consoles/ workstations.

- Sufficient space and capacity to install additional computers and peripherals.

- Spare or expandable communications capabilities.

- Additional office space for operations and maintenance personnel from different operating agencies.

IDENTIFICATION OF PARTNERS AND CONSENSUS BUILDING

Regardless of whether the control center will be operated by a single agency or by several agencies in a joint operation, agencies need to identify potential partners, and build consensus and support within these organizations. A big question that must be addressed in the initial planning stages is who needs to be in the control center. Depending upon local needs and operating philosophy, operating personnel can come from the following agencies:

- State transportation agencies.

- Local transportation agencies.

- Police and emergency service providers.

- Local transit authorities.

- Private media and traffic reporting services.

Joint operations can be structured administratively to occur in different ways, such that varying levels of functional and management control are centralized or individual control is maintained. Joint operation can be structured through the following: [(4)]

- Sharing physical resources that are common to each agency's operation, but operating each system or agency component individually. This could occur through use of a common communications system (e.g., the TranStar System in Houston, TX).

- Operating individual or multiple systems under one designated management structure where operational control is centralized. This could occur by time of day where peak periods are under central control and off-peak is under local jurisdictional or functional agency control. Typically, the participating parties establish operating guidelines that are carried out by an individual agency or group, with the goal being to establish coordinated ongoing operations.

- Delegating day-to-day operations to another agency or group (including a private entity). This type of operation could entail turning the operations and maintenance of individual devices over to another agency under a defined set of conditions (e.g., Transcom operation of certain DMSs and HARs in the region surrounding New York City).

The ability to engage in joint operations is not an easy objective and usually occurs from ongoing relationship building. While a variety of strategies can be undertaken to foster cooperative joint operations, no single technique or action is always appropriate. Each community must assess its unique situation and develop a specific strategic plan that conforms to that situation. [(4)]

ESTABLISH GOALS AND OBJECTIVES

Before beginning on the actual design of the control center, participating agencies need to establish common goals and objectives for the control center. These common goals and objectives are used to develop principles for designing the functional and physical

components of the control center. Examples of the types of issues that agencies need to address in establishing goals and objectives for a control center include the following:

- The functions to be carried out in the control center (e.g., traffic management, dispatching, incident management, etc.).
- The degree to which various elements in the system can be accessed and controlled by potential users of the system.
- The level of automation and the role of the operator in the system.
- The level and types of access that users outside the system (e.g., media, administrative, traffic reporting services, etc.) have to the information (i.e., databases, video, incident reports, etc.) produced by the system.
- The amount and type of customized equipment and software that will be permitted in the design of the control center.
- The level and manner in which existing and planned elements of the system will be integrated.
- The types of operator interfaces (i.e., displays, data entry screens and devices, etc.) that will be provided.
- The hours and operating conditions under which the control center will be staffed.

DEFINE FUNCTIONAL REQUIREMENTS

After establishing design goals and objectives, the functional requirements of the control center can be defined. Functions are broad statements that describe every operation and activity required to meet the defined objectives. Table 10-1 lists some possible generic functions of freeway management control centers. Note that these functions describe what it is the system does, and do not define whether activities are done by humans, by automated equipment, or by a human using a computer.

A *Mission Analysis* is an exercise that may be useful to agencies in identifying the functions of a control center. A mission analysis is used by system planners and designers to identify the operational capabilities required in the control center. Traditionally, two methods are used to conduct a mission analysis: a mission profile and scenario development. A mission profile is a detailed description of normal system operations that occur during a given system activity or over a given interval of time. It consists of listing all the activities to be done by various elements in the total system—operators, supervisors, automated subsystems, sensors, etc. The list also includes any activities done simultaneously (e.g., automated tasks done by system hardware, operator assessments, operator decisions). Activities are described at a high level and no attempt is made to define the roles of the operators or automated system in doing them. This technique provides an organized, high-level framework of system requirements that will support subsequent, detailed design analysis. [(1)]

With the scenario development technique, descriptions of specific scenarios—nonroutine but typical situations that would burden (challenge) the capabilities of the control center—are sometimes useful in providing an understanding of the control center's functions. The scenarios should describe what actions and information are needed to manage traffic during different

Table 10-1. List of Possible Generic Functions Performed in Freeway Management Control Center. [1]

Input	Throughput	Output	Support
• Detect vehicle locations • Detect vehicle speeds • Detect vehicle types • Sense roadway surface conditions • Receive BIT reports • Receive ad hoc component status reports • Sense visibility conditions • Verify incident data • Monitor incident clearance • Receive traffic volume reports • Receive probe vehicle reports • Receive ad hoc travel time reports • Receive ad hoc roadway condition reports • Receive O-D Data • Receive commercial rail traffic data • Receive ad hoc commercial rail traffic reports • Receive weather service data • Receive ad hoc weather reports	• Assess current load • Anticipate near-term traffic conditions • Select best traffic control option • Determine need for ITMS support • Track special vehicles • Predict traffic conditions given options • Determine remedial maintenance needs • Assess predicted traffic conditions given options • Assess traffic management effectiveness • Determine software upgrade needs • Determine hardware upgrade • Determine personnel upgrade needs • Determine preventative maintenance needs • Determine source of anomalies • Identify anomalies in traffic patterns	• Control railroad crossings • Post route advisories on information outlets • Provide route advisories to other users • Post speed advisories on information outlets • Provide speed advisories to other users • Post travel advisories on information outlets • Provide travel advisories to other users • Post mode selection advisories on information outlets • Provide mode selection advisories to other users • Transmit electronic maintenance requests • Issue special maintenance requests • Issue upgrade requests • Transmit electronic incident services requests	• Store electronic network data • Retrieve electronic network data • Store electronic incident data • Store hardcopy of incident reports • Retrieve electronic incident data • Retrieve hardcopy of incident reports • Perform database management • Provide traffic management training • Provide maintainer training • Provide incident management training • Provide special events training • Develop strategic traffic management plans • Develop special event traffic management plans

Table 10-1. List of Possible Generic Functions Performed in Freeway Management Control Center (Cont.). [1]

Input	Throughput	Output	Support
• Receive interagency incident data • Receive ad hoc incident response reports • Receive interagency emergency response data • Receive ad hoc emergency response data • Receive interagency data from alternate transportation modes • Receive ad hoc reports from alternate transportation modes • Receive interagency special events reports • Receive ad hoc special event reports • Receive public comments • Receive requests for public relations activities • Receive requests for historical data • Receive requests for simulation studies	• Predict multimodal demand given options • Determine ITMS responsibilities • Determine need for incident services • Determine appropriate ITMS responses • Assess multimodal demand and capacity • Identify demand regulation options • Assess predicted multimodal demand • Monitor compliance with other advisories • Monitor general compliance with advisories • Assess survey data • Assess ad hoc public comments • Plan public confidence enhancements	• Issue special incident services requests • Issue requests for information • Issue requests for on-site traffic control • Transmit electronic incident reports • Issue special incident reports • Transmit electronic incident management reports • Issue special incident management reports • Provide historical traffic data • Provide simulation reports and recommendations • Provide public relations information	• Develop traffic management contingency plans • Receive directives • Develop policy • Specify procedures • Implement policy and procedures • Perform fiscal planning • Perform budget tracking • Perform evaluations • Perform personnel selection • Maintain personnel records • Maintain communications with incident responders • Coordinate multi-agency incident response • Coordinate multi-agency response to other emergencies • Coordinate multi-agency transportation planning

operating situations such as freeway incidents, major traffic stressors (e.g., large athletic events, inclement weather), or strategic planning episodes.

DEFINE FUNCTIONAL RELATIONSHIPS, DATA REQUIREMENTS, AND INFORMATION FLOWS

In defining the functional relationships, data requirements, and information flows for a control center, the following three tasks must be done:

- Allocating control center functions the operators, computer/machine components in the system, or a combination of both.
- Analyzing the tasks required to complete each function.
- Establishing how data flow from one function to the next.

Each of these tasks is discussed below.

Function Allocation

In the design of the control center, function allocation involves assigning system functions to machine components, human operators, or a combination of human and machine components. Using criteria similar to those shown in table 10-2, each function (or process) is assigned to a human or machine component. Properly allocating functions is critical to ensuring that operators in the control center perform tasks that are within their capabilities and do not become overloaded. The *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition) presents techniques for allocating functions to human operators and machine components. The reader is referred to this document for details on the allocation procedures.[1]

The allocation of functions in the control center is usually the first point in the design process at which critical decisions must be made about the role of the operator. It is also a point at which mistakes, if not identified and corrected in design iterations, can cause serious design deficiency. One common misconception that occurs in allocating functions is that designers presume that a single set of functions should be handled solely by an operator and another set should be handled solely by machines. In fact, many critical functions in the control center can best be handled by the integrated efforts of humans and machines. Identifying these partly automated tasks requires detailed study and analysis. Failing to identify them properly and assign proper interface strategies causes serious operational problems. General guidelines and design considerations for allocating functions in a control center include the following:[1]

- If environmental constraints limit human performance, the function should be allocated to a machine.
- Events that cannot easily be perceived by humans (such as changing levels of traffic moving past a point on a freeway) should be allocated to machines.
- When a function requires a response that is beyond the speed or accuracy of human capabilities, it should be allocated to a machine.
- If the speed and volume of information derived or needed by a function is beyond the capabilities of a human, it should be allocated to a machine.

Table 10-2. Criteria for Assigning Functions to Humans and Machines. [1]

Humans Excel in ...	Machines Excel in ...
Detection of certain forms of very low energy levels	Monitoring (both men and machines)
Sensitivity to an extremely wide variety of stimuli	Performing routine, repetitive, or very precise operations
Perceiving patterns and making generalizations about them	Responding very quickly to control signals
Ability to store large amounts of information for long periods — and recalling relevant facts at appropriate moments	Storing and recalling large amounts of information in short time periods
Ability to exercise judgement where events cannot be completely defined	Performing complex and rapid computation with high accuracy
Improving and adopting flexible procedures	Sensitivity to stimuli beyond the range of human sensitivity (infrared, radio waves, etc.)
Ability to react to unexpected low-probability events	Doing many different things at one time
Applying originality in closing problems (i.e., alternative solutions)	Exerting large amounts of force smoothly and precisely
Ability to profit from experience and alter course of action	Insensitivity to extraneous factors
Ability to perform fine manipulation, especially where misalignment appears unexpectedly	Ability to repeat operations very rapidly, continuously, and precisely the same way over a long period
Ability to continue to perform when overloaded	Operating in environments that are hostile to man or beyond human tolerance
Ability to reason inductively	Deductive processes

- If the information produced by a function is beyond the memory capabilities of a human, it should be allocated to a machine.

- If a function is to be performed continuously, it should be allocated to a machine.

- If the interruption of, and response to, unusual or unexpected events are

required, the function should be allocated to a human.

Following is a list of additional guidelines that can be used to help in allocating functions in a control center:[1]

- Allocate functions so that they make the best use of human abilities.
- Avoid decisions based solely on the ease or difficulty of automation; consider how allocating different functions between humans and machines affects total system performance.
- Avoid allocating functions in such a way that both humans and machines are forced to work at their peak limits all or most of the time.
- Allocate functions to humans so that they can recognize or feel that they are making an important and meaningful contribution to the performance of the system.
- Allocate functions between humans and machines so that a natural flow and processing of information can occur.
- Assign tasks that require extremely precise manipulations, continuous and repetitive tasks, or lengthy and laborious calculations to a machine.
- Design human/system interfaces on the presumption that the human might at some point have to take control of the system.
- Use hardware and software to aid the operator; do not use the operator to complement a predetermined hardware/software design.

Task Analysis

After functions have been allocated to human operators and system components, the next step in designing a control center is to identify the tasks that make up system functions. Each function includes one or more tasks. A task is an independent action, carried out either by an operator or by a machine, that results in an identifiable outcome. Tasks can frequently be decomposed into discrete subtasks that represent activities that are distinct enough to be analyzed separately, but are clearly contributing to the completion of an identified task.[1]

Once the tasks have been identified, they can be grouped to form operational flow or process diagrams. The operational flow diagrams allow designers to identify the actions, information requirements, processes, and decisions that need to be made to accomplish a function. Figure 10-3 illustrates an operational flow diagram for dealing with traffic incidents. Operational flow diagrams are useful tools in the design process, because they allow designers to identify the following elements easily:

- The types of data and communications requirements in the control center.
- The decision-support aids needed to complete operational tasks.
- The data-storage requirements of various processes and tasks in the control center.
- The types of outputs and decisions produced by each task.

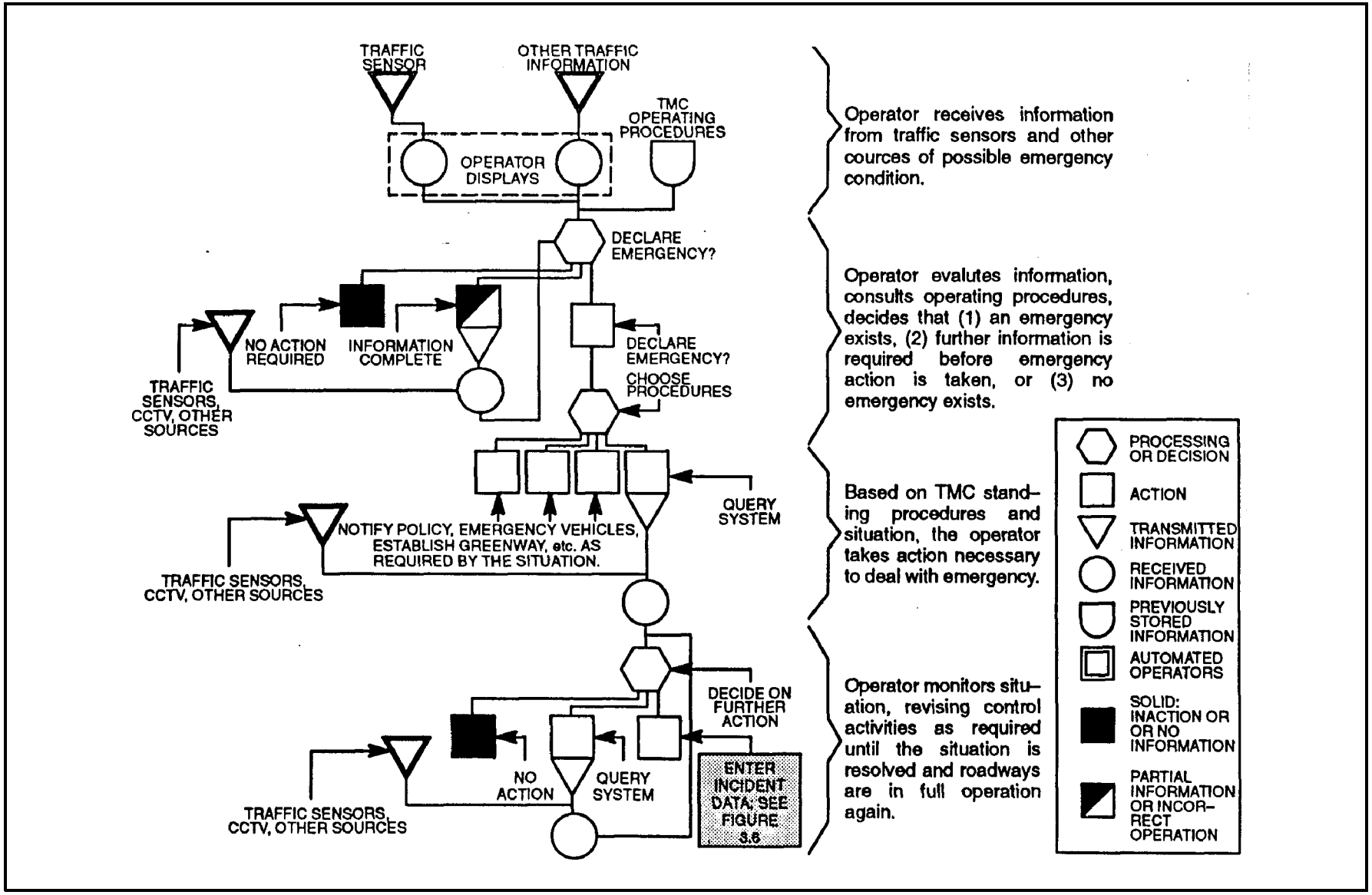


Figure 10-3. Example of Operational Flow Diagram for Managing an Incident. [1]

Data Flows

Establishing data flows is a critical step in designing a control center. Data flows describe the type and frequency of data needed to execute each function of the control center. This step in the design process is important because it allows system designers to assess the communications requirements of each component in the control center. Establishing the data flows also helps to identify the structure of the data streams needed to operate each function of the system.

One way to depict data flow is through data flow diagrams. With data flow diagrams, large circles are used to represent sources and destinations of data. The sources/ destinations can be either subsystems or functions within subsystems. Lines connecting data sources and destinations are used to represent the type of data that flows between two elements of the system. Each data type is labeled so that designers know what information is flowing between components. An example of a typical data flow is provided in figure 10-4. Data flow diagrams need to be prepared for each level of design detail and for each subsystem within the control center.

IDENTIFY AND SCREEN TECHNOLOGIES

Only after the functions of the control center have been identified and the data flow requirements have been assessed should designers begin designing the physical layout and the support elements and related technologies of the control center. In planning the control center, system designers need to be concerned with the following elements:

- The physical environment in which the operators and equipment will function.
- The design and operations of the operators' workstations.
- The design of the controls and displays that the operators will use to operate the system.
- The design of the interfaces through which the operators will be presented with information and initiate control decisions.

The following sections relating to these issues were extracted from the *Human Factors Handbook for Traffic Management Center Design* (First Edition).[1]

Physical Environment

The control center's physical environment consists of design elements that allow the system — both human and machine components — to function effectively. The following lists some physical elements that designers must consider:

- Atmospheric (heating, ventilation, and air conditioning).
- Visual (primary and supplementary lighting).
- Acoustic (background noise and interior acoustical properties allowing operators to communicate).
- Physical design of the work space (access, dimensions, and fixtures).

The requirements for some of these elements are mandated by public law (e.g., access for the disabled). The design of other features of the control center should be based on established design practices (e.g., lighting

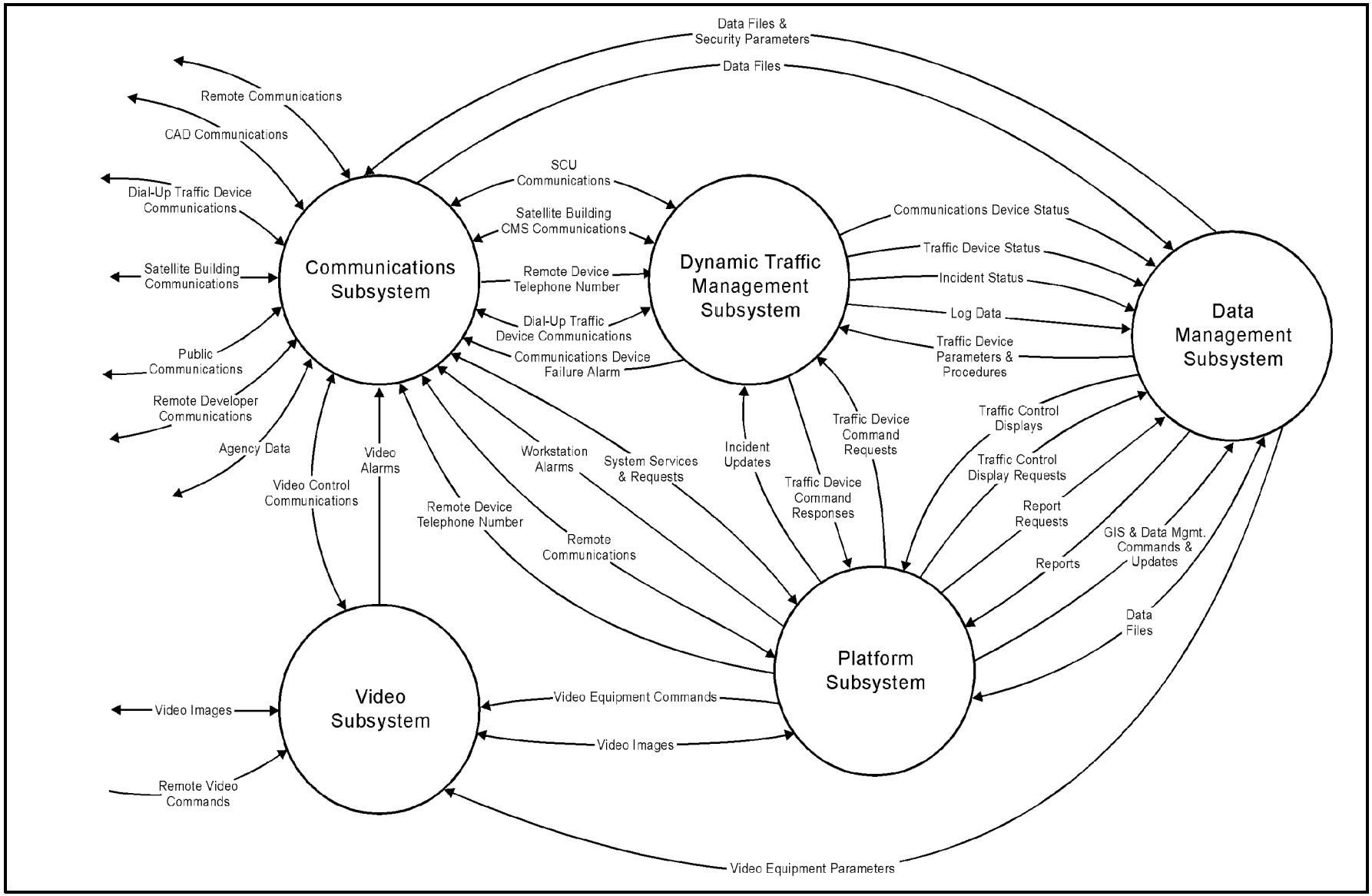


Figure 10-4. Data Flow Diagram of the TranStar System in Houston, TX. [5]

standards for designated work areas). The *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition) provides guidelines and requirements that can be used to design the physical environment of a control center.[1]

Workstations

At the same time the design team is considering how the control center work space will be used, it should begin designing and specifying the workstations that will be used in the control center. A typical operator workstation is shown in figure 10-5.

When workstations are to be configured from commercial off-the-shelf components, this task will require catalog searches and negotiations with vendors. Sometimes—when, for example, the control room is too small to hold separate workstations—customized fixtures designed specifically for the control center under development may be required. The layout of workstations and other furniture and fixtures should be specified as part of the overall work space design. In designing workstations, comfort and suitability should be considered as separate issues, both of which are supported with many experience-based guidelines. The *Human Factors Handbook for Traffic Management Center Design* (First Edition) summarizes many of these guidelines. [1]

Poorly designed workstations and supplemental furnishings (e.g., chairs) can cause discomfort and perhaps occupational injuries (e.g., back strain, cervical stress disorder, carpal tunnel syndrome, and repetitive stress disorders). Poor workstation design can also limit productivity. A good design will contribute to productivity and employee health and morale. Designing the workplace to accommodate the characteristics and capabilities of human operators is sometimes called *ergonomics*. The strong movement toward ergonomic suitability has created many sources of information that support proper workstation design; however, an experienced ergonomist or human factors engineer should be consulted before a large investment is made in workstations.[1]

Controls and Displays

Integral to any control center workstation are the means by which operators enter and retrieve information. Controls allow the operators to guide certain traffic parameters (e.g., traffic flow) within the limits of the center's mission. Displays provide information that operators need to monitor the status of the system and make control decisions. Several methods used to display information to operators in a control center including the following:

- Static display boards.
- Video/CRT monitors.
- Projection television systems.
- Video walls.

Advantages and disadvantages of each of these display devices are discussed in the section 10.3 of this module.

Poorly designed controls and displays can affect the operations of the control center and the operators themselves. Inadequate controls and displays can cause cognitive information processing deficiencies, faulty situation assessments and decisions, inaccurate data and command entry, occupational stress, and a general loss of operating efficiency.

Figure 10-5. Typical Operator Workstation in the Houston TranStar Freeway Management Control Center.

Activities associated with the design and specification of controls and displays are actually subprocesses of the workstation design effort. Controls and displays are considered separately because of the established body of knowledge and standards associated with each of them.

Control center controls and displays should be designed according to (1) the type and quantity of information that they must process and (2) the capacities and demands of the operators who will use the information they provide.[1] Results of a task analysis will suggest information requirements. In turn, these requirements will drive control and display designs.

Once display types have been determined, the guidelines and standards presented in the *Human Factors Handbook of Advanced Traffic Management Center Design* (First Edition) can be used to specify display quality.[1] Emphasis should be placed on the development of display performance standards. Brightness, contrast, and resolution of visual displays, for example, should satisfy operators' requirements for clarity and visual comfort. Designers should be aware, however, that designing controls and displays is an iterative process. A given display, although acceptable when viewed on its own merits, may represent an inappropriate design solution when it is integrated into the control center.

User Interfaces

If the smooth execution of system functions is desired, interfaces between operators and computers must be designed properly. The guidelines and standards for user-computer

interface design are relatively new, and less stable than designers typically prefer. The status of these standards and guidelines will present the most significant challenge to the control center configuration effort.[1]

Many troublesome human factors problems existing in operational control centers are derived from poor user interface design. One common problem is the use of command line interfaces that require operators to memorize commands that have little, if any, inherent meaning. Another common interface problem can be attributed to "error traps" — procedures that carry an unacceptable risk of user errors. Some of these errors may have negligible effects, while others may be devastating to the performance of the system. The steps involved in designing acceptable user interfaces include the following:[1]

- Develop an interface concept. This concept should be based upon a knowledge of the types of interfaces available, and their strengths and weaknesses. The types of tasks to be performed through user interfaces should be considered and should be what drives the design process, rather than technology.

- Develop criteria to govern the design and evaluation of the interface. These criteria include design requirements, dialog strategies, and training requirements. The criteria should also be verified by experienced operators, traffic engineers, and others who have experience in interface design.

- Develop a "look and feel" of the user interface and conduct initial testing of the interface, using simulations of various levels of realism. (The level of realism associated with a given simulation is also called *fidelity*.) This type of simulation activity provides an excellent opportunity to incorporate *rapid prototyping*. Rapidprototyping involves using special software tools to develop and evaluate interface prototypes. Rapid prototyping techniques offer speed, flexibility, and realism to the design effort.

- Develop user interface for full implementation.

- Conduct user acceptance testing of the interface. User expectations and perceptions do not always agree with those of the developers. Rigorous evaluation procedures must be employed during acceptance testing. Preliminary tests of training plans, user aides, and user documentation should also occur.

- Incorporate the user interface into the control center design. Design consistency and suitability should be assessed. Final development of a training plan, a training implementation plan, and user documentation should also be completed.

EVALUATION

As different elements of the control center are installed, they need to be evaluated to ensure that they meet generally accepted principles of human factors engineering. Human factors testing and evaluation should not be a one-shot, pass-or-fail activity conducted near the end of the design or implementation phase. Instead, human factors testing and evaluation should occur throughout the design and implementation process. The most important principle to follow in testing and evaluating the human factors elements of the design is to *test early and often*.[1]

Testing should be conducted by someone who has experience in conducting human factors evaluations. Through human factors evaluations, system operators can be assured of the following:[1]

- The design of the control center will meet accepted standards of human factors engineering and will be tailored to the specific requirements of the system.
- Mistakes and problems in the design of the control center will be detected and corrected early, without compromising the budget or schedule.
- Operators will be more easily trained.

Human factors testing and evaluation should be both formative and summative. Formative evaluations result in an iterative process where various levels of design are evaluated and modifications to the design are made on the basis of the results of the testing. Summative evaluations produce a formal pass/fail judgement of design features. The *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition) provides more detailed information on how to conduct these evaluations.[1]

In conducting human factors evaluations of control center design, obtaining opinions from the following individuals is recommended:[1]

- The control center operators themselves.
- The operators of control centers in other cities.
- The managers and supervisors of the control center.
- Representatives from other agencies who will be operating in the control center or have experience with control centers (i.e., fire/police dispatchers, air traffic controllers, etc.)
- Independent human factors specialists.

10.3 TECHNIQUES AND TECHNOLOGIES

PHYSICAL DESIGN

The size of control centers varies considerably, depending on the design objectives and functions performed by the system. A functional analysis is needed as part of designing the control center. Factors that affect the design of the system include the following:

- The hours and days of operation.
- The types of information displays that will be used in the center.
- The types of monitoring and control strategies in the system.
- The desirability of interagency staffing.
- The need for media and public access to the control center.
- Whether the control center will function as a communications or dispatching center.
- Whether the control center will also function as an emergency operations center.

The size of traffic control centers varies widely throughout the United States. Some control centers, such as TransGuide in San

Antonio, Texas, and Transtar in Houston, Texas, occupy as much as 4,800 m^2 (52,000 ft^2); however, many more successful control centers occupy less space. For example, the control center in Minneapolis, Minnesota occupies only approximately 950 m^2 (10,000 ft^2). Generally, most traffic control centers are between 1,600 m^2 (18,000 ft^2) to 2,300 m^2 (25,000 ft^2).

A typical traffic control center has the following work areas within the building: [6]

- Operations room.
- Computer and peripheral area.
- Communications area.
- A gallery for viewing and training.
- A conference room.
- System support offices.
- A break room.
- A media room.

An example of a typical floor plan for a control center is shown in figure 10-6.

The operations room is the hub of the control center. It houses the operators' workstations and display boards/video terminals used to display information about conditions in the network to the operators. The size of the operations room is dictated by the number of operator workstations that will ultimately be in the control center, and the type and location of the information displays.

The computer/peripheral area houses the computers needed to run the freeway operations center. The size of the computer room depends upon system design decisions regarding how much computer capability will be based at the control center as compared with that installed at hub locations in the field. Often these rooms require strict environmental controls (i.e., heating and air-conditioning). Many locations also have strict access control for this area. Figure 10-7 is an example of a typical computer/ peripheral area in a freeway management control center.

The communications area can be a stand-alone area or can be combined with the computer/peripheral area. The communications area is the terminus for data entering and exiting the control center from the field. It also serves as the distribution point for transmitting data to the rest of the system (i.e., the computer and video control systems).

Many locations like to provide a gallery or viewing area of the operations room. From this area, tours of the control center can be conducted without interrupting the operations in the control center. This area can also be used as a training center where trainees can monitor the operators without disturbing them.

Many control centers also have an area set aside for the media (see figure 10-8). Some control centers use the gallary area to service this function, but others have set aside a separate area as a media area. From this area, the media can watch what is going on in the operations center without interrupting the operations. Often, these areas have several telephones and video terminals that are available for use by the media during incidents or emergencies.

One unique feature of the Houston TranStar System is a media cabinet located in the foyer of the building. As shown in figure 10-9, the cabinet contains power, video, and audio jack that can be used by the media

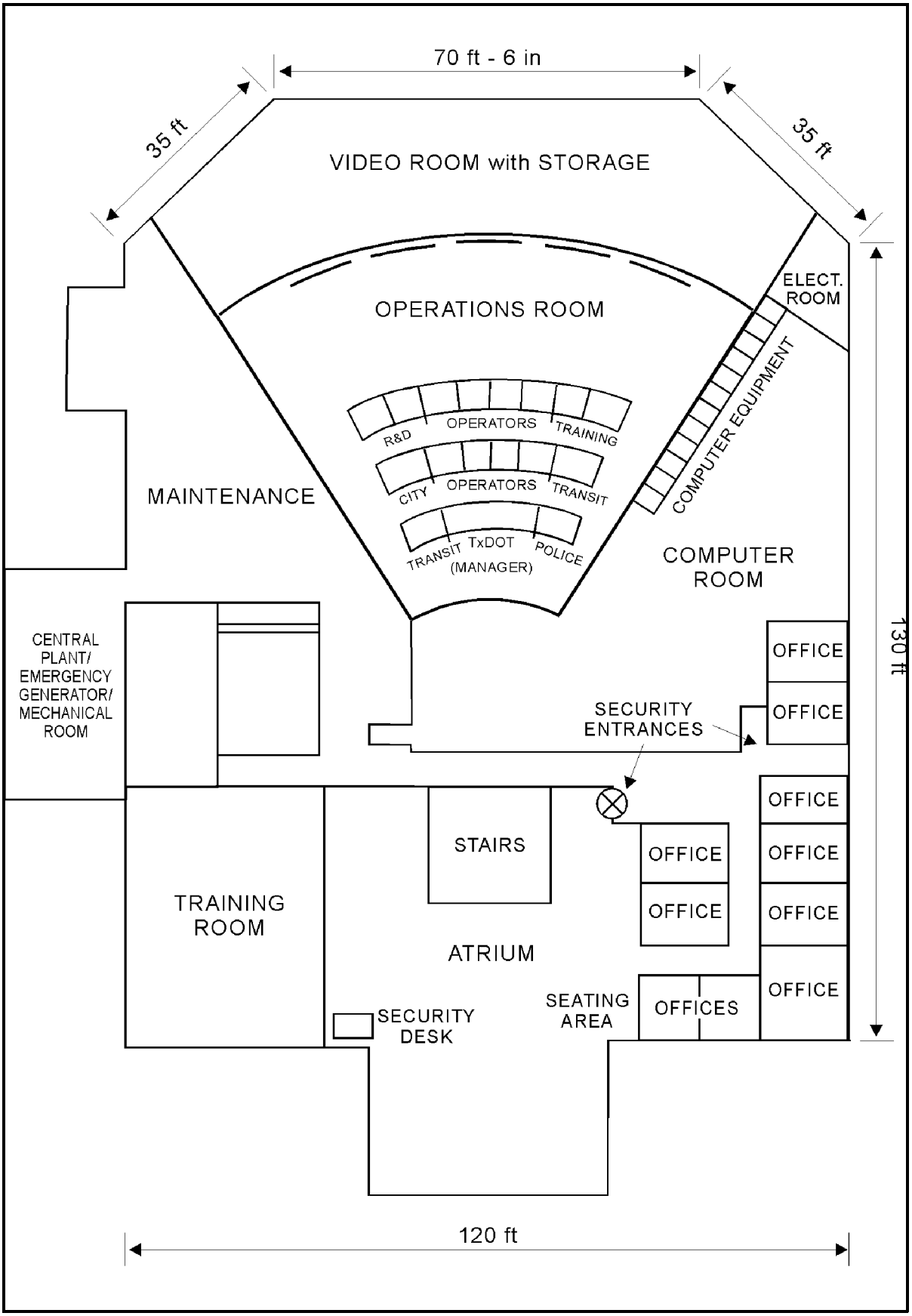


Figure 10-6. Floor Plan of the TransGuide Control Center in San Antonio, TX.

Figure 10-7. Example of a Computer/Peripheral Area in a Control Center.

Figure 10-8. Example of Media / War Room in a Freeway Management Center.

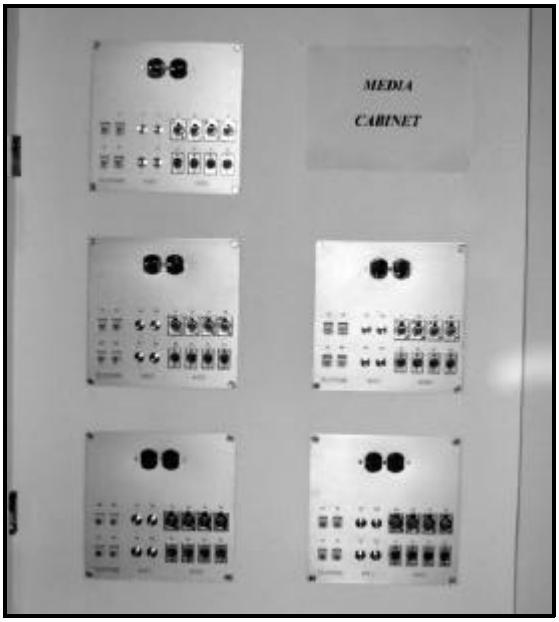


Figure 10-9. Media Cabinet in Houston TranStar Center.

when conducting interviews or reporting on traffic situations. Cables run from the cabinet in the foyer to a cabinet located in the parking lot, where similar connectors are provided for the crews to tie into their respective vans.

Besides these common areas, office space should be provided to those working in the control center. The number of office spaces, again, depends upon what functions will be carried out in the control center. Generally, office space is required for the operations supervisors, the system support staff, the maintenance supervisors, and the engineering and administrative staffs.

Some agencies have included space for their system maintenance personnel at their control centers. If maintenance operations will be based at the control center as well, the space provided at the control center is based on the number of electrical maintenance personnel that will be stationed there. In addition, if the control center is also expected to serve as an Emergency Operation Center (EOC), then additional space is needed for such items as dormitories, a large kitchen, and showers. Some locales may have mandated structural requirements for an EOC that are more stringent than for regular commercial office space.

Lighting

Lighting is an important consideration in the design of a control center. Viewing a large situation map with LED signals or a bank of CCTV monitors, for example, is not compatible with high levels of general illumination. On the other hand, many operators' tasks cannot be performed in low levels of illumination. The lighting scheme and choice of luminaries must be viewed as an integrated whole, and not designed piecemeal. [(1)]

The greatest challenge in designing the lighting in the control center revolves around the need to provide dim general illumination and higher levels of local illumination. Usually, a low level of illumination is provided in the operations room because of the nature of the tasks and displays used. Most centers provide supplemental lighting at each operator console. Some general recommendations related to lighting in control centers are as follows:[(1)]

- Indirect lighting should be employed to provide overall illumination for the control room.
- Canister fixtures should be employed to provide supplementary illumination over work areas.
- Adjustable fixtures should be used in work areas where more intense illumination is required.

For more detailed recommendations on lighting issues, the reader is referred to the

Human Factors Handbook for Advanced Traffic Management Center Design (First Edition).[1]

Acoustics

Communications between operators can be critical, especially in emergencies. Common sources of noise in a control center include the following:

- Alarms.
- Radio/telephone communications.
- Operator conversations.
- Data processing equipment.

Overall, the noise level in a control center should not be high enough to interfere with normal speech between operators. The objective in designing for noise is to balance the different sound sources so that local speech is unaffected, but is sufficient to mask intrusive noise from adjacent spaces. Some general recommendations for reducing the impact of noise in the control center include the following:[1]

- Identify possible noise sources (including machines, telephones frequently in use, loudspeakers, and radios with speakers) during the design phase and eliminate them.
- If noise sources cannot be eliminated, consider strategies for reducing noise level, including textured or sound-deadening wall and ceiling materials.
- Consider placing noisy functions that are not tied to normal operator activity in a separate room or in an area enclosed by acoustic partitions.

The *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition) should be consulted for more detailed recommendations and specifications for accounting for noise in control centers.[1]

Heating and Cooling

Operator comfort and performance can be affected by temperature and air quality. In designing a control center, designers must be concerned about two issues related to the thermal environment of the control center: the general heating, ventilating, and air conditioning (HVAC) standard and the effects of local thermal conditions related to special equipment such as computers and video display units. Sometimes, special rooms in the control center, such as the computer/peripheral room or the communications area, may require separate heating and cooling standards. The *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition) summarizes the heating and cooling standards applicable to control centers. General standards that should be considered in the design of the heating and cooling system for control centers include the following:[1]

- Actual ventilation should be ensured by introducing fresh air into any personnel enclosure.
- Within permanent structures, effective temperature shall be maintained at not less than 18°C (65°F), and not greater than 29.5°C (85°F).
- Approximately 45 percent relative humidity should be provided at 21°C (70°F). The humidity should decrease with rising temperatures, but should remain above 15 percent to prevent irritation and drying of body tissues.

- The temperature of the air at floor level and at head level should not differ by more than 5.5°C (10°F).

- The exhausting of air from instrumentation (video display units, system units, etc.) should be so accomplished as to avoid discomfort to users and others close to the equipment.

- Units should be designed so that forced-air exhausts are not directed toward the operating position, or toward other workers in their work positions.

- External surfaces that can be touched during operation shall have a surface temperature that does not exceed 50° C (122°F). Surfaces intended to be touched during normal operation should not exceed 35°C (95°F).

- Heat build-up from equipment under the work surface (around the operator's knees and legs) greater than 3°C (5.5°F) above ambient should be avoided.

INFORMATION DISPLAYS

A dominant feature in nearly every control center is a large information display board or screen. These devices are used to provide a broad overview of the status of the system to both operators and visitors in the control center. Three different kinds of information displays are generally used in control centers: static wall maps, projection television screens, and video walls. [7]

Static Wall Maps

Static wall maps were used extensively in systems designed in the 1960s and 1970s. These types of displays provide the geographic layout of the roadway system and show the location of the system's resources, such as detector stations, signals, signs, and cameras. Tiny light bulbs or light-emitting diodes (LEDs) connected to sensors in the field are used to display information about the status of the system. In some freeway management centers, different colored lights are used to depict various traffic flow conditions (e.g., green for free flow, yellow for moderately congested, red for congestions, and flashing red for incident location). Since these displays are generally large painted maps, making modifications to these displays as the infrastructure changes is difficult and costly. Figure 10-10 shows a typical static wall map used in a freeway management center.

Projection Television Systems

Projection television systems have also been used to display traffic information in control centers. The primary advantage of this type of system over static wall maps is that they are extremely flexible as to what can be displayed to the operator. Besides showing status information on computer-generated maps, operators can also display live video from CCTV cameras and television images on the screen. Because the images are usually generated by computer, the operator can zoom to various levels of detail on the display. The primary disadvantage of projection television, however, is that the resolution can sometimes make the image become blurry. Furthermore, because of its sensitive optics, it frequently requires realignment and adjustments. Figure 10-11 shows an example of a control center with a projection television map display.

Video Walls

Video walls are being used almost exclusively in new freeway management control centers. A video wall is a matrix of television monitors used as a single display. Each individual monitor can be used to

Figure 10-10. Example of Static Wall Map.

Figure 10-11. Example of a Projection Television Display.

display a single image or can be used to compose part of a larger display. By using a video wall, the operators in the control center have the flexibility of customizing the presentation of the information as conditions warrant.

Design Guidelines and Recommendations

The following guidelines and recommendations are provided to help in designing visual displays in control centers:[1]

- Avoid too much detail on large maps or status boards.
- Limit the number of colors for maps, target symbols, alphanumeric headings, etc.
- If moving objects are displayed on a map, keep the number at a minimum and display only those that move slowly.
- If the display area must have a low ambient illumination in order for individual operator displays to be used effectively, use white or luminescent markings against a dark background for maps as status boards to help operators maintain dark adaptation.
- If color-coded object information is to be used on a large map display, use only a neutral color, such as gray, for the map background; this allows the color targets to have maximum effect (contrast).
- When front projectors are to be used to project information on a large map display, ensure that the projectors are positioned so that they are not readily visible (i.e., causing obstructions and glare) to the operating personnel.
- Determine and provide the properly sized alphanumeric characters and/or symbols on the large-screen displays for the maximum viewing distances at which each set of characters and symbols must be read.
- Consider using unambiguous coding techniques to help operators in discriminating between old and new data.
- Orient maps with north to the top.
- Color codes should agree with commonly accepted practices:
  - *Flashing Red*: Emergency.
  - *Red*: Alert.
  - *Yellow*: Caution, recheck.
  - *Green*: Safe, go ahead.
  - *White*: Used when no right or wrong entry or condition exists.
  - *Blue*: Can be used to denote advisory, but use should be avoided.

OPERATOR WORKSTATION DESIGN

Considerable changes in workstation design have occurred as freeway management systems have evolved over the last three decades. In the past, most control centers were designed with multiple workstations, each with a specific function. For example, one workstation controlled the video surveillance system, another controlled the variable message signs, another controlled the ramp metering system, etc. This type of design required the operators to move between a series of workstations to implement a control strategy in a specific situation. Today, most control consoles are being designed so that a single operator can operate all of the subsystems from a single

workstation. This generally results in better operational control over the system; however, in some situations, two or more operators may be competing for control of the same camera or monitor. Methods of setting priorities whereby multiple functions can be performed at a single workstation need to be established in a control center.

The design of individual workstations in a control center varies depending upon the functions to be performed by the operators. All workstations should be designed to anthropological standards. The placement of video display monitors and input devices (e.g., keyboards, mouse, trackballs, switches, etc.) should also conform to recognized standards and guidelines. Many centers have poor ergonomic designs that contribute to operator stress and discomfort. Common ergonomic design problems include the following:[7]

- Monitors placed above operators' seated line of sight.
- Operators having to look over consoles, monitors, and other equipment to view monitors on the wall.
- Inadequate or improper labeling of control features.
- Console heights not adjustable to extremely short or extremely tall operators.

Glare is one of the most common problems with video display units incorporated into workstations. Glare is generally caused by either lighting sources at the workstation or at other nearby video display units. Glare problems can be eliminated by the following means:[7]

- Provide glare shields between video display units and light sources.
- Cover light sources, including windows.
- Use task lighting.

The *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition) provides detailed standards and recommendations that should be followed in the design of operator workstations.[1]

COMMUNICATIONS SYSTEMS

In general, system designers need to be concerned with three types of communications systems when designing a freeway management system:

- The type of communications system that links the field devices with the control center and permits the transfer of data and commands. This type of system generally requires a high speed, high capacity type of transmission medium.
- The type of communications system that links the computer systems inside the control center that are responsible for processing information and commands, generating displays and reports, and interfacing with the control center operators. Most freeway management systems in operation today use local area networks (LANs) to connect their computer and display equipment in the control center.
- The type of communications system that permits operators in the control center to converse with other personnel (and also the public) by voice.

Because of the costs and criticality of the first type of communications system, it is covered in **Module 9**.

Local Area Networks

The term Local Area Network (LAN) is commonly used to describe the type of communications system that links the digital computers internal to the control center. By definition, an LAN is any telecommunications system that serves a limited geographic area (typically a single building or campus). The term "network" refers to the fact that multiple users are interconnected.[8]

No single LAN design is ideal for use in a control center. The design of the LAN needs to support the functions and the types of data exchanges in the control center. The simplest type of LAN permits the exchange of information between computers and computer-like devices (such as word processors, operator workstations, database managers, etc.). More complex forms of LANs are required to support the transmission of video and audio information besides data.

Components

While the complexity of the LAN varies depending on the type of data being transmitted, every LAN has the following basic components:[8]

- User workstations.
- Supporting processing equipment.
- Peripheral equipment.

The user workstations are the most visible component of the LAN. In control center applications, workstations may consist of conventional PC computers, intelligent workstations, or terminals. All these devices have the following common functions:[8]

- Display output to users, usually using a device such as a video display monitor.
- Receive data from users entered from a keyboard, mouse, or other input device.
- Interface with the communications medium, a function that includes formatting and transmitting data at the appropriate time and rates.
- Processes data, depending upon the LAN design, the capabilities of the workstation, and the type of software used.

Supporting processors, known as servers, are connected to the LAN to execute functions that cannot be provided by individual workstations. Generally, servers are used for the following applications:[8]

- Execute large computer programs that exceed workstation capacities.
- Provide centralized data storage and retrieval functions (file servers).
- Interface with external communication facilities (communications servers).
- Interface with peripheral devices.

The capacity of the supporting processors can vary from that of a PC to that of a large mainframe computer. The number and type of supporting processors depend on several factors including the following:[8]

- The design of the LAN.
- The number of users sharing the supporting processors.
- The functions being performed by each processor.

In most control centers, a variety of peripheral devices are supported on the LAN, including the following:[8]

- Printers.
- Plotters.
- Telex.
- Fax machines.
- Modems.
- Optical character readers.
- Scanners.

In some systems, the LAN is also used to support video and voice transmissions via connections to television equipment and the telephone system.[8]

In larger LANs, a special processor, known as the network control station, is often used to monitor the communications traffic on the LAN continuously, and to accumulate statistics on workstation usage, transmission quality, and network configuration.[8]

Planning for an LAN

Quite often, LANs are selected more on the basis of vendor recommendations than on technical or application requirements; however, when planning an LAN system, the first consideration should be the selection of the type of system or topology that is most appropriate for the application requirements. Generally, three types of topology (i.e., physical shape) are commonly used in designing LANs: a star topology, a bus topology, and a ring topology. Figure 10-12 illustrates each of these topologies. Advantages and disadvantages of each topology are summarized in table 10-3.

For simple system requirements, such as a few PCs and a file server, an Ethernet type of system may be quite appropriate. However, for large numbers of PCs with large data-transfer volumes, a token ring system is more likely to be the best selection. A rule-of-thumb is to use a baseband LAN for small office systems requirements and a token ring for large or host computer-oriented (mini/mainframe) systems.

Connections between LAN nodes can be established over twisted-pair cable, thin coaxial cable, standard coaxial cable, or optical fiber. Each of these alternatives has increasing capacities. Twisted-pair (telephone) cable is smaller and cheaper than other communications media, but is subject to electrical and radio interference. Coaxial can carry higher frequencies and data rates than twisted-pair but is more difficult to manipulate physically. Optical fiber is popular because of its small cable diameter, protection against electromagnetic or radio interference, low attenuation, and large bandwidth, but is generally more expensive to install than the other media. Factors that influence the selection and implementation of communications media used in the LAN include the following:[8]

- The amount of data being transferred between devices (i.e., capacity levels).
- The potential for external interference.
- The physical distance between devices.
- The potential for future expansion.

The ability to diagnose and predict its own failure is an important factor in designing an LAN. Simple layouts and well-established procedures should facilitate quick repair.[8] In addition, system planners should anticipate expansion needs in the design of the system. The internal design of the

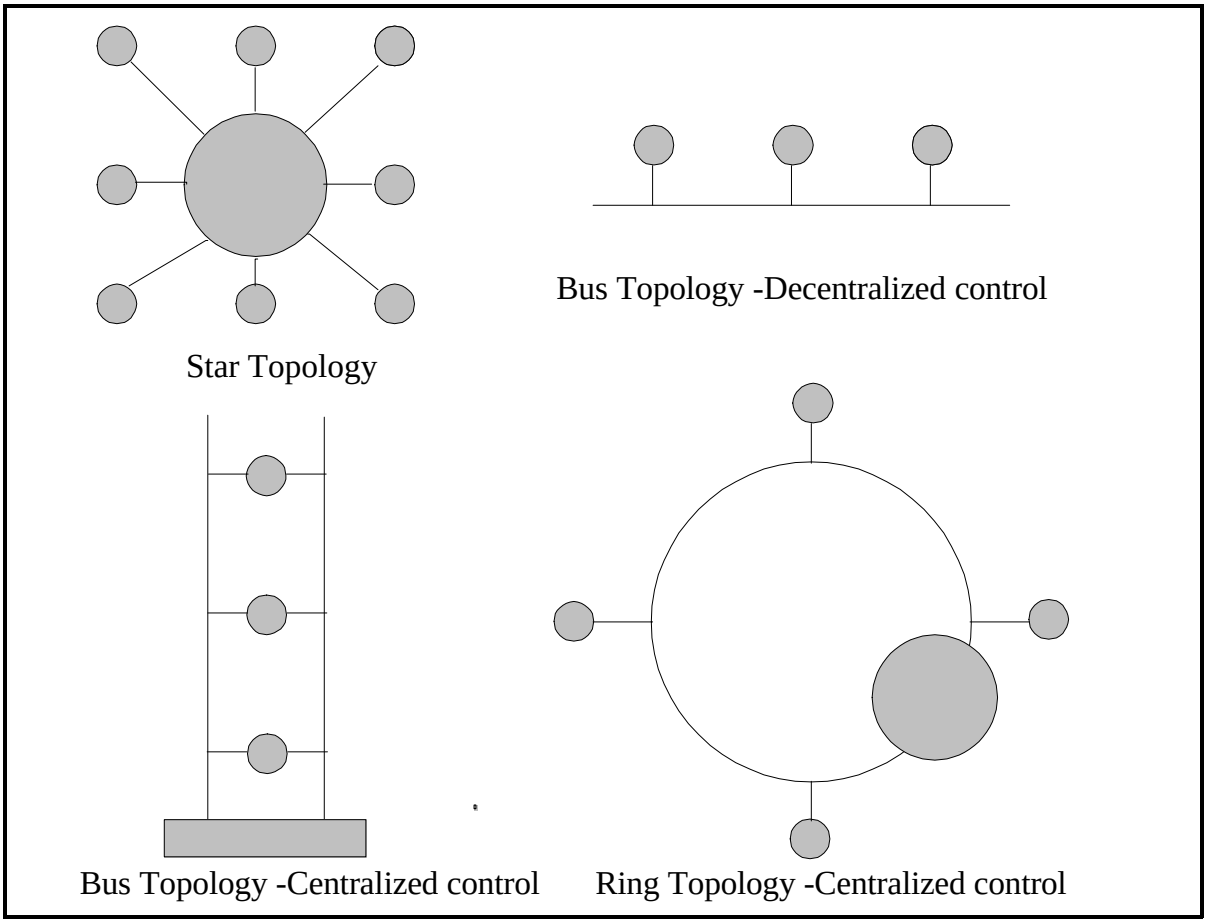


Figure 10-12. Examples of Popular LAN Topologies. [8]

building should incorporate adequate routing ducts and conduits to hold all future growth needs. In addition, access rooms should be provided to ease the job of adding and maintaining an LAN.[8]

Voice Communications

Systems within the control center are also needed that allow operators to talk with individuals outside the control center. The types of voice communication systems needed in a control center depend primarily upon the functions to be performed by the system, existing communications systems, local availability, and agency preference. Common types of voice communications systems included in the control centers include the following:

- Intercom.
- Direct line telephone connects.
- Switching telephone systems.
- Radio.
- Cellular telephones.

Figure 10-13 shows some of the voice communications devices used by operators at a control center.

Voice communications are commonly being used in control centers for the following purposes:

Table 10-3. Advantages and Disadvantages of Common LAN Topologies

Topology Type	Attributes	Advantages	Disadvantages
Star	• Communication protocols. generally controlled by central processor. • Polling techniques used to initiate data transfers. • Used by popular computer operating systems (e.g., UNIX and OS/2).	• Permits purchase of lower-cost user workstations.	• All workstations disabled in event of central computer failure. • System expansion costly.
Bus	• Topology on with both the Ethernet and token bus protocols based. • Individual workstations initiate data transfers.	• Network growth can be readily accommodated.	• Each processor must provide the capacity required to execute the user's application software. • Generally results in more expensive workstations.
Ring	• Each node receives "token" transfer data. • A token is a digital code that allows workstation access to the network. • Each node transfers data in turn.	• "Tokens" can be passed in either direction around the ring.	• Limited exclusively to token passing techniques.

Figure 10-13. Typical Communications Devices Used by Operators in a Control Center.

- Communications with incident response teams (including fire, police, emergency medical service providers, hazardous material teams, etc.).
- Control of remote and portable equipment (e.g., highway advisory radios, and variable message signs).
- Communications with operations and maintenance field personnel.
- Communications with motorist call boxes.
- Transmission of dispatch information, data, and calls.
- Communications with management and administrative staff, both internally and with outside agencies.

STAFFING

No rigid rules exist for determining the number of operators in a control center. The number of operators in a control center primarily depends on the functions that are being performed, the number of facilities covered, and the operating objectives and philosophy of the center. Table 10-4 shows the ultimate staffing requirement of a traffic operations center in Orange County. The primary function of this traffic control center is to detect and clear incidents from the freeway at all hours of the day. It also supports those individuals involved in the continued planning and operation of the system, and houses the staff responsible for developing operational strategies for special events, providing lane closure recommendations for construction contracts, and responding to major traffic incidents. It also houses maintenance dispatching

Table 10-4. Staffing Plan for Orange County Traffic Operations Center. [9]

Function	Number	Hours	Grade
CALTRANS			
TOC System Manager	1	0700-1600 M-F	Senior Engineer
Assistant Manager	1	1000-1900 M-F	Associate Engineer
Receptionist/Clerical	1	0800-1700 M-F	Receptionist
Area Traffic Engineer	5	0700-1600 M-F	Assistant Engineer
Area Traffic Technician	10	0700-1600 M-F	Engineering Technician
TOC System Operators	9	3 Shifts, 7 Days/Week	Engineering Technician
Software Support Programmer	1	0700-1600 M-F	Assistant Engineer
Electronic Maintenance Specialist	1	0700-1600 M-F	Electrical Leadworker
Maintenance Dispatch Supervisor	1	0700-1600 M-F	Maintenance Supervisor
Maintenance Dispatcher	6	3 Shifts, 7 Days/Week	Maintenance Man II
Assistant Maintenance Dispatcher	2	2 Shifts, M-F	Student Assistant
CALIFORNIA HIGHWAY PATROL			
System Coordination Officer	1	0700-1600 M-F	Lieutenant
Duty Officer	6	3 Shifts, 7 Days/Week	Traffic Officer
Media Information Officer	2	2 Shifts, M-F	Media Information Officer

functions for all of the Orange County Caltrans District offices.[9]

Generally, the types of personnel needed in the control center of a freeway management system include the following:[10]

- Electronic/maintenance technicians.
- Control center technicians.
- Communications specialists/operators.
- Software programmers.
- Computer system operators.
- Operations supervisor or director.

Generally, those centers whose functions include significant interaction with police officers in the field should consider including a police liaison officer on their staffs. A police liaison officer fosters stronger interagency cooperation and can aid in

dispatching appropriate police responses to incident scenes.

Qualifications

Overall, the level of qualifications for control center operators should be directly related to the design of, and functional allocation within, the center. If the operators' functions are repetitive, predictable, and non-critical, operators with low qualifications can be used; however, if unique problems are frequent, rapid reactions are required, and/or criticality is high, operators need to have higher levels of training and expertise. As a rule, operators must possess good verbal skills, a degree of computer literacy, and good reasoning skills. Some agencies also require operators to have a good working knowledge of the freeway system and some dispatching experience.[7]

The qualifications for control center operators also vary from center to center. For example, some locations employ part-time students as operators. These students are under the supervision of a manager or a senior operator. Some control centers use operators taken from the agencies' technical staff of traffic engineers or computer scientists. Often, these operators have other assigned technical duties, or may be given special projects to work on along with their duties in the control center. A few centers require their operators to have college engineering or technical degrees. Operators should possess the following characteristics no matter how much education and experience they have:[10]

- Ability to make good judgment calls quickly and accurately.
- Ability to handle stressful situations.
- Good oral and written communication skills.
- Good understanding of the transportation language.
- Good technical skills and good interpersonal skills in dealing with nontechnical personnel.
- Ability and willingness to follow standard operating procedures.
- Initiative to learn more about the operations of the other functions in the control center.
- Willingness to provide suggestions/ideas to promote a more efficient operation.
- Computer literacy.

Training

The training of operators is critical to the success of a freeway management system. Operators need to be provided with three levels of training: technical, operational, and managerial. Both formal and informal training is needed before the system becomes installed, as the system goes on-line, and after the system has been operating for some time.[10]

How much training is required by operators in a control center depends on the functions to be performed and the level of technical competence of the operations staff. Technical training is required in a variety of areas associated with system operations and maintenance, including diagnostic procedures for all hardware as well as for new upgrade procedures.

The basic level of training needed by system operators includes the following:[7]

- General principles, operating philosophies, and concepts of freeway management.
- An overview of the system, including the system schematics, field subsystems, communications, central subsystems, and proposed or planned system functions to be added.
- Operation and interpretation of system software and displays.
- Basic radio and communications codes and procedures.
- Standard operating procedures.
- Communicating with other agencies such as fire, police, etc.

At most control centers, new operators generally receive one-on-one, on-the-job training with an existing operator.

Besides basic training, operators need to be provided with continuous advanced training. This advanced training can be adaptable to specific issues and needs of the operations and maintenance staff. Areas in which advanced training may be provided include the following:[7,10]

- Emergency response procedures.
- Hazardous material spills procedures.
- Major accident and disaster clearance procedures.
- Roadside fire response.
- Multijurisdictional extended pursuits.
- Release of information to the media.

Methods of providing advanced training include the following:[7]

- Lectures or short-courses.
- Videotape.
- Simulated events or table top exercises.
- Computer simulators.
- Site visits to other control centers.

Sufficient training should also be provided when a new element or application is added to an existing system. This is usually accomplished through contract specifications with training line items.[10] In preparing contract specifications for training, the following should be specified:[11]

- The maximum number of persons to attend each formal training session.
- The minimum number of days for each training program. (Defining what a day is may also be important.)
- Who will develop and supply all the necessary manuals, displays, class notes, visual aids, and other instructional materials for the training program.
- Outlines of lectures and demonstrations, and samples of all training materials. These materials should be submitted to the agency for review some specified time before their proposed use. Agency approval should be specified before the training courses can be scheduled.
- Where training is to be conducted (e.g., at a local site designated by the agency or at the contractor's facility). If the training is to be provided at the contractor's facility, the specifications also need to define who is responsible

for paying the transportation and subsistence costs of the agency personnel.

- All training should be conducted during normal business hours of the agency and the training site.

Staffing Sources — Agency versus Outsourcing

Being able to attract and maintain qualified control center personnel can be difficult for many agencies. Agencies have two basic options for staffing their control center: with personnel from with the agency, or by outsourcing. A pure agency staff has the advantage that managers and team leaders have a single personnel management system to deal with, and team cohesiveness is easier to establish and maintain. Outsourcing involves hiring private personnel to perform some or all of the functions in the control center. The primary advantage of outsourcing is the immediate availability of highly-qualified personnel. Seasonal and special events can be adequately staffed with limited training and start-up time. In addition, private employment agencies are not subject to the civil service regulations and policies that public agencies must follow. Private agencies can be more competitive and can pay the market salary rates that make it easier to attract and retain qualified people. To some degree, many agencies already employ outsourcing techniques to perform some functions in the control center (such as computer maintenance, software development, etc.). Those types of positions that can potentially be outsourced include the following:[12]

- Field and electronic technicians with communications or electronics experience.
- Team managers or leaders from a variety of backgrounds.
- Control center and communications technicians with knowledge, experience and training from other private sector hardware and software systems.
- System programmers and computer systems analysts with knowledge and familiarity of control algorithms.

SECURITY

Security surrounding control centers largely depends on the nature of the center and its objectives. Because of the vast differences in the purposes and capabilities of these centers, the appropriate levels of security to protect them vary widely, as do the perceptions of security risk. Many original freeway management control centers around the country were developed strictly as a traffic control measure and, as such, their operators did not see any particular threats to the facility. This stemmed from a perception that there would be little intrinsic value in attempting to attack or otherwise break into such a facility. In contrast, many facilities currently being brought on-line incorporate police, transit, emergency management, and traffic operations. These facilities are perceived to represent a much more likely target for computer hackers, theft of data, and even potential terrorist activities.[13]

Several generic security measures can be taken to limit physical access to facilities.[14,15] A recent article in an architectural magazine notes that “effective security is an interplay of three elements: natural and architectural barriers, including anything from landscaping strategies that discourage access, to the number, location, size, and type of doors and windows; human security, including the protection provided

by guards and other personnel; and electronic security, provided by any one of the array of systems now available."[15]

Obviously, location of the facility will play a central role in determining what security measures are appropriate. Here again, the needs of staff should be considered in selecting what countermeasures are employed.[14] Communication systems, power supplies, access points, physical integrity of the building, and several other issues are all directly affected by security considerations. In addition, what countermeasures can be used is affected by building codes regarding access and egress during emergencies such as fires. Yet another layer of regulatory codes is associated with the Americans with Disabilities Act, which can affect aspects ranging from physical security barriers to systems that must accommodate both the blind and deaf.[15] Security systems also must be designed so that they are not too obtrusive, intrusive, or otherwise intimidating to employees.

Computers in the control center are also a source of security concern for many operating agencies. According to one source, the majority of all computer security losses have been attributed to errors or omissions. Major sources of other computer security losses include dishonest and disgruntled employees, and external threats such as disasters. Only a small percentage of security losses were credited to outside sources, such as hackers.[16] The point was thus made that "the two most heavily publicized types of security problems, hackers and viruses, are among the least serious threats to most systems." The threat from viruses (on a percentage basis) was also apparently small. As much as 80 percent of all computer-related damage was estimated to be caused internally. Nevertheless, because of the publicity surrounding hackers and viruses, an undue amount of attention is often focused on protecting against external threats.[16]

The extent of the risk to computer systems will vary greatly from agency to agency. The determination of how much security is necessary reverts to the need for risk assessment. Multiple layers of firewalls and other security measures may be warranted in some systems. Several publications are available that offer further detailed exploration of computer security measures that can be implemented.[16,18,19,20]

AUTOMATION

The acceptable level of automation varies from site to site. In some locations, the long-term goal for their systems is to be fully automated while, at other locations, the goal is for the operator to continue to be a critical element in the operation of the system. Each philosophy has significant implications with regard to the overall design of the control center.

The role of the operator in a system can be defined in terms of whether a human or a machine makes the decisions (i.e., closes the loop) in a task or process. The role of the operator in the decision-making process can be placed into four categories:[1]

- Direct performer.
- Manual controller.
- Supervisory controller.
- Executive controller.

Under direct performer control, the human operator performs all the functions of the system. In manual control, the machine components are heavily involved in the decision-making process as sensors and

effectors, but the actual loop-closing aspect of the function is solely the responsibility of the human. In supervisory control, a machine component is allowed to close the loop under supervision of a human operator who may intervene and adjust or override the machine's decision. With executive control, the machine is totally responsible for performing all functions of the system—the operator is only there to keep the machine operating.

A continuum of operator roles exists that defines how much automation is needed to accomplish a function. At one end of the continuum, a function is allocated solely to the human, and at the other end, solely to the machine. In between, performance of the function is shared by human and machine components. As shown in figure 10-14, the continuum can be divided into four major regions; each region defines a generic operator role in relation to automation. Because it is a continuum, how much automation occurs varies within each region.

How much automation is acceptable in a control center varies from center to center, depending upon the experience of the operators, the operational goals and philosophy of the agency, and the sophistication of the system. In many locations, routine functions in the control center can be automated. For example, different control centers use different methods for generating messages for display on variable message signs, including the following:

- Messages are entered manually.
- Messages are entered manually with computer assent.
- The operator chooses from selection of canned messages.
- The computer determines a response plan with the assistance of an operator.
- The computer determines and carries out a response plan.

With the current state of technology, it is recommended the operator always remain "in the loop." Automation systems should be designed to help and support the operators. Operators should be fully aware of what information the system software has, what the automated systems is doing, why it is doing it, and what it is going to do next. At a minimum, the operator should be able to switch off the automated system to prevent future problems or to correct improper decisions.

If automated systems are to be used, the level of automation should be gradually increased throughout implementation. Initially, an operator should be present in the control room to review and approve all automated system actions. A detailed log of automation failures should be kept. As failures are eliminated, the role of the operator in the decision-making process can be reduced.

JOINT OPERATIONS

Joint operation can be structured through the following: [4]

- Sharing physical resources that are common to each agency's operation, but operating each system or agency component individually.
- Operating individual or multiple systems under one designated management structure where operational control is centralized.

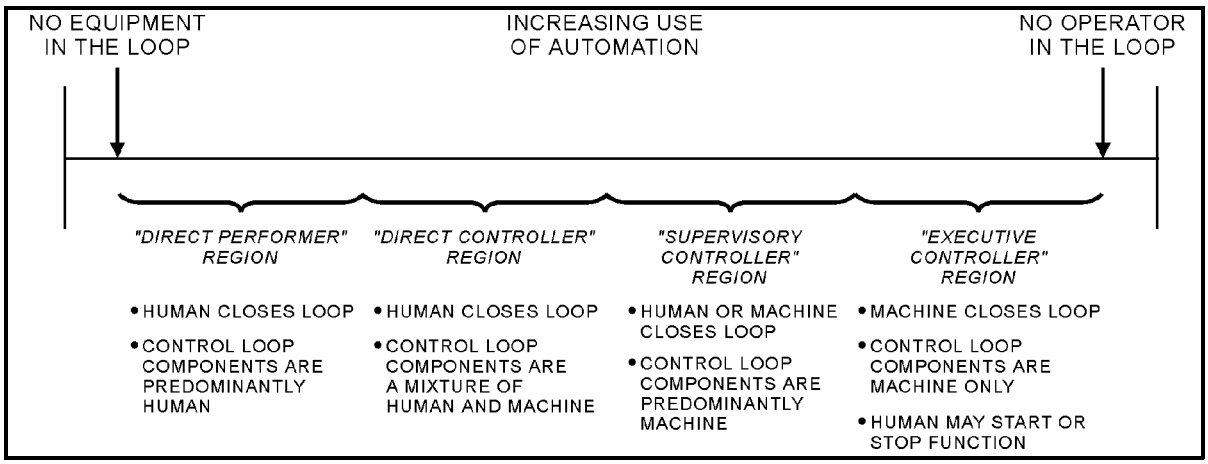


Figure 10-14. Continuum of Operator Roles in Relation to Automation.[7]

- Delegating day-to-day operations to another agency or group (including a private entity).

The ability to engage in joint operations is not an easy accomplishment and usually occurs because of ongoing relationship building. A variety of strategies can be undertaken to foster cooperative joint operations. No one individual technique or action is appropriate for all areas; instead, each community must assess its own unique situation. Strategies that can be employed to foster cooperative and joint operations in a control center include the following: [4]

- Ensure that each agency is represented in the initial stages of the design and program development of the control center.

- Emphasize how projects and programs can affect/address needs and problems of each agency throughout the development process. Look for ways to widen the focus of the initial goals of the system to help other agencies improve their operations.

- Approach joint operations with an open attitude about how overall results can be enhanced by sharing resources.

- To facilitate sharing and build trust among agencies, start joint operations with a relatively small and noncritical task. Building confidence and trust around these smaller elements will facilitate the accomplishment of larger tasks in the future.

- Develop an open-ended and flexible system architecture such that new systems and changes in hardware and operating procedures can be accommodated easily.

- Add functions and responsibilities under joint operations at a manageable rate.

- Develop standard operating procedures for how the devices in the system can be used by each agency in the control room. These operating procedures should be scenario-based and describe the roles and responsibilities of each agency in the scenario.

- Cross-train the staff from each agency so that they can do the jobs of the other agencies' staffs, and so that each operator has an understanding of the roles, responsibilities and limitations of each agency in the control center and

can serve as a backup or substitute in crises.

- Provide a mechanism for positively, reviewing and debriefing each others' operations with the idea of improving the overall operations and capabilities of the system.

10.4 LESSONS LEARNED

SYSTEM CAPABILITIES

It is important to consider system balance when deploying a freeway management system. System balance is achieved when all of the elements of the system are operating at their optimum, designed operating levels. A system becomes unbalanced when one of its elements is forced to operate beyond its capabilities. System operators should not be expected to do more to manage traffic than what the system was designed to do. Information should not be distributed and strategies should not be employed that are beyond the capabilities of the system to monitor and carry out. Erroneous information, and inappropriate and/or inadequate strategies erode public and political support for the system.

OPERATIONS MANUAL

Even though Federally funded projects require an Implementation Plan prior to actual system implementation, preparing an Implementation plan is a good professional practice, regardless of the funding source of the project. The plan defines system objectives, deployment elements to achieve those objectives, a general operations plan of how the elements will work together, a staffing plan for operations and maintenance of the system, and a commitment of resources for operation and maintenance of the system. The Operations Manual will define in more detail the actual day-to-day operations in the control center. It is essential that the manual be prepared in advance of system turn-on and used in the training of operators. In addition to the control console operation (the physical operation and adjustment of hardware and software), the Operations Manual must provide consistent and appropriate responses to similar traffic situations. Consistency can be provided by developing response plans for different levels of incidents, congestion, and traffic situations. The response plans should indicate what actions should be taken by elements of the system for a given traffic situation. Response plans should be well documented and placed in an Operations Manual. The Operations Manual also identifies who is responsible for carrying out specific actions. Figure 10-15 outlines the basic areas and topics to be included in the Operations Manual.

EFFECTIVE USE OF SYSTEM OPERATORS

Operators can be valuable in assessing the quality of operations, and in identifying potential operational and design changes in a system. The system should be designed to support operators and their task of maintaining traffic flow. Because they deal with traffic situations on a daily basis, they are often in a position to know what works, what does not work, and what needs to be done to improve the system. They can provide valuable insight into not only what additional hardware and functions are needed at the center, but also ways to improve the effectiveness and efficiency of the overall operation of the system.

In existing practice, the skills level of the individuals staffing and operating freeway management centers varies widely from system to system. Some agencies use

General Information About the TMC
- Mission and functions of the TMC
- Relationship of the TMC to other transportation agencies
- Organizational relationship of the TMC within the agency responsible for administration
- Diagrams depicting the physical layout of the TMC

Policies and Procedures on Internal Operations and Maintenance of the TMC
- General information on items such as address, main telephone number, fax number, e-mail, etc.
- Hours of operations
- Contact procedures in the case of emergencies involving the TMC
- Telephone procedures
- Personnel
  - Organizational chart
  - Description of duties for each position
  - Training (required classes and training manuals)
  - Rules of conduct
- Equipment
  - Authorized use
  - Maintenance
- Facility management
  - Security procedures and authorized access
  - Backup power
  - Custodial services
- Software
  - Backup procedures
- Media
  - Access to center
  - Media guidelines
- General public, transportation professionals, and VIP's
  - Access to center
  - Conduct of tours
- Disaster recovery plan

Policies and Procedures on Traffic Management
- Incidents (accidents, stalled vehicles, spilled loads, etc.)
  - Identification (vehicle detection, 911 calls, traffic reporters, etc.)
  - Verification (CCTV cameras, police, or DOT personnel)
  - Response plans (ramp metering, traffic signals, VMS, HAR)
  - Documentation of incidents
- Congestion
  - Monitoring
  - Response plans
- Planned events (including roadway closures and maintenance)
  - Obtaining necessary information on planned events
  - Response plans
- Field equipment malfunctions
  - Dispatching of repair crews
  - Documentation
- Interjurisdicational coordination
  - Other TMCs
  - Transit agencies
  - Other agencies (highway patrol, police, fire, etc.)
- Information distribution
  - Media
  - Value-added packages of transportation information
  - General public

Figure 10-15. Recommended Topics to be Covered in Operations Manual. [21]

students as operators, others use maintenance personnel, and still others use engineers. The key to operation of a successful freeway management center is to have highly qualified and motivated individuals who are both knowledgeable about traffic operations and technically competent.

SYSTEM MAINTENANCE

System maintenance is often overlooked by agencies. The performance of any operational traffic control system, whether a traffic signal system or a freeway management system, depends on the commitment of the operating agency to provide effective maintenance for the system. Agencies should not underestimate the budgetary or staffing requirements to properly maintain the control center and the

system as a whole. Agencies must recognize that it takes highly skilled individuals to maintain today's complex systems. Public agencies should be aware of the costs associated with attracting, training, and retaining individuals with the requisite skills.

All personnel needed to operate and maintain the system should be placed under the management of the same individual in the agency's administrative hierarchy. The individual having the administrative responsibility for the control center should report to the same individual who has overall authority for traffic operations and maintenance functions of the agency. This individual may be the Head of Traffic Operations in larger agencies, the Head of Traffic Operations and Maintenance in medium-sized agencies, and the Head of Transportation in smaller agencies. This administrative structure will aid in accomplishing the following:

- Achieving a balance between funds to operate and maintain the system.
- Effectively addressing conflicts between operations and maintenance personnel.
- Developing a high level of technical expertise in operating and maintenance personnel.
- Facilitating interaction and opening lines of communications between operations and maintenance personnel.
- Facilitating interagency and intraagency communication and sharing of data and information.

10.5 REFERENCES

1. *Human Factors Handbook for Advanced Traffic Management Center Design* (First Edition). Prepared for FHWA, U.S. Department of Transportation by the Georgia Institute of Technology, Georgia Tech Research Institute, Atlanta, GE, October 1995.

2. Huchinson, R.D. *New Horizons for Human Factors in Design*. McGraw-Hill Book Company, New York, NY, 1981.

3. Wickens, C.D. *Engineering Psychology and Human Performance*. Harper Collins Publishers, New York, NY, 1992.

4. Wiersig, D.W. White paper on *Joint Operations*. Prepared for the National Conference on Operating and Maintaining Advance Traffic Management Systems (ATMS) Centers. Institute of Transportation Engineers, Washington, DC, 1996.

5. *System Design Specifications for the Regional Integrated Transportation Management System*. Prepared for the Greater Houston Transportation and Emergency Management Center by LORAL Space Information Systems, Transportation Systems Programs, Houston, TX, 1995.

6. Kay, J. *Planning a Traffic Operations Center*. FHWA Freeway Surveillance and Control Course. California Polytechnic State University, San Luis Obispo, CA, 1993.

7. Kelly, M.J., Gerth, J.M., and Whaley, C.J. *Comparable Systems Analysis: Design and Operation of Advanced Control Centers*. Report No. FHWA-RD-94-147. FHWA, U.S. Department of Transportation, 1995.

8. Tarnoff, P.J, and Pugh, T. *Transportation Telecommunications*. NCHRP Synthesis of Highway Practice 165, NCHRP, Transportation Research Board, Washington, DC, 1990.

9. *Orange County SOC Study Site Selection*. Prepared by JHK & Associates for District 12 of the California Department of Transportation, 1993.

10. Mona, J.A. White Paper on *Training*. Prepared for the National Conference on Operating and Maintaining Advance Traffic Management Systems (ATMS) Centers. Institute of Transportation Engineers, Washington, DC, 1996.

11. Neudorff, L.G. *Guidelines for Successful Traffic Control Systems: Final Report*. Report No. FHWA-RD-88-014. FHWA, U.S. Department of Transportation, Washington, DC, 1988.

12. McMichael, L.N. White Paper on *Staffing*. Prepared for the National Conference on Operating and Maintaining Advance Traffic Management Systems (ATMS) Centers. Institute of Transportation Engineers, Washington, DC, 1996.

13. Brehmer, C.L. *Addressing Security Issues in ITS/ATMS*, Compendium: Graduate Student Papers On Advanced Surface Transportation Systems, Texas Transportation Institute, Report Number SWUTC/96/72840-00003-1, August 1996.

14. Cooper, W., and DeGrazio, R. Building Security: An Architect's Guide. *Progressive Architecture*, Vol. 76, No. 3, March 1995, pp. 78-81.

15. Aarons, S. Access Control Systems. *The Architect's Journal*. Vol. 201, No. 15, April 13, 1995, pp. 35-36.

16. *Securing Client/Server Computer Networks*. McGraw-Hill Series on Computer Communications. McGraw-Hill, New York, 1996.

17. Waller, D. Onward Cyber Soldiers. *TIME Magazine*. Vol. 146, No. 8, August 21, 1995, pp. 38-44.

18. Olnes, J. Development of Security Policies. *Computers & Security*, Vol. 8, No. 8, 1994, pp. 628-636.

19. Cohen, F.B. *Protection and Security on The Information Superhighway*. John Wiley and Sons, Inc., New York, 1995.

20. Sydmonds, I. Security in Distributed and Client/Server Systems-A Management View. *Computers & Security*, Vol. 13, No. 6, 1994, pp. 473-480.

21. Rowe, E. White Paper on *Administration*. Prepared for the National Conference on Operating and Maintaining Advance Traffic Management Systems (ATMS) Centers. Institute of Transportation Engineers, Washington, DC, 1996.

MODULE 11. ECONOMIC ANALYSIS

TABLE OF CONTENTS

11.1 INTRODUCTION 11-3
PURPOSE AND SCOPE 11-3
SYSTEM GOALS AND OBJECTIVES 11-4
INTERRELATIONSHIP OF SYSTEM ELEMENTS 11-4
STRUCTURE OF MODULE 11-4

11.2 ESTIMATING COSTS AND BENEFITS 11-4
SYSTEM COSTS 11-4
Types of Costs 11-4
Sources of Cost Information 11-5
SYSTEM BENEFITS 11-5
Types of Benefits 11-5
Sources of Benefit Information 11-6
IMPORTANCE OF LIFE CYCLE COST CONSIDERATIONS 11-6

11.3 SYSTEM EVALUATION 11-7
DEFINITION OF SYSTEM ALTERNATIVES/COMPONENTS 11-7
ASSIGNING BENEFITS TO ALTERNATIVES 11-8
ESTIMATING LIFE-CYCLE COSTS 11-9
ANALYSIS TECHNIQUES 11-9
Net Present Worth 11-9
Description 11-9
Example 11-9
Other Considerations 11-9
Benefit-Cost and Incremental Benefit-Cost Analysis 11-11
Description 11-11
Example 11-11
Other Considerations 11-12
Utility-Cost Analysis 11-13
Description 11-13
Example 11-13
Other Considerations 11-14
Value Engineering 11-14
Description 11-14
Example 11-14
Other Considerations 11-16

Sensitivity Analysis 11-17
Description 11-17
Example 11-17
Other Considerations 11-18

11. 4 REFERENCES 11-19
REFERENCES 11-19
SUGGESTED READING 11-20

MODULE 11. ECONOMIC ANALYSIS

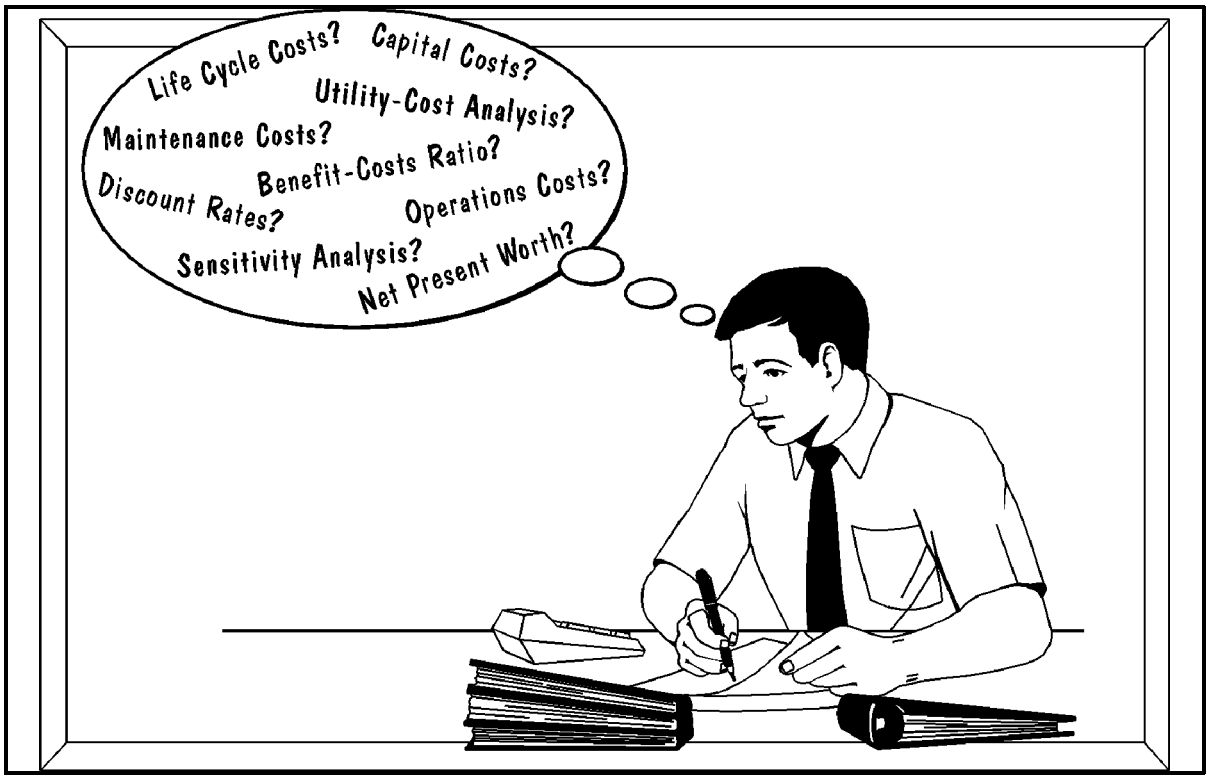


Figure 11-1. An Economic Analysis Has Many Components.

11.1 INTRODUCTION

Since freeway management systems are designed, constructed, and operated and maintained with public funding, it is critical that economic analyses are conducted to ensure that public funds are spent prudently. In addition to being used to determine which alternative system offers the most potential, economic analyses serve to justify the cost-effectiveness of system installations to elected officials who oversee public funding, as well as to the public whom these elected officials serve. If funding for new freeway management systems, or funding for operating and maintaining existing systems is to continue, it is critical that elected officials and the public be made aware of the benefits of the freeway management system.

PURPOSE AND SCOPE

This module serves to give guidance to planners and designers responsible for the economic justification of freeway management systems. Planners and designers must be familiar with the costs and benefits expected from freeway management systems in order to justify the installation and continued operation of these systems. This module provides typical capital costs associated with the design and construction of freeway management systems, as well as typical continuing costs associated with their operation and maintenance. Also provided in this module are typical quantifiable and nonquantifiable benefits that can be expected from the implementation of freeway management systems.

SYSTEM GOALS AND OBJECTIVES

The primary function of freeway management systems is the real-time management of recurrent and nonrecurrent congestion. It is the goal of planners and designers to provide efficient, cost-effective freeway management systems that meet defined system goals. A successful freeway management system meets or exceeds defined system goals, thus producing benefits such as delay reductions and increased safety to freeway users that outweigh the system's initial capital costs and its associated lifetime operating and maintenance costs.

INTERRELATIONSHIP OF SYSTEM ELEMENTS

Costs associated with freeway management systems, whether capital costs or operation and maintenance costs, are readily available and easily measured in monetary terms. However, benefits accrued from various freeway management system elements are sometimes more difficult to quantify, due to the interrelationships that exist among the various system elements. For example, dynamic message signs (**Module 7**) and ramp meters (**Module 5**) both contribute to reducing freeway congestion. The benefits of each of these elements could be measured (via travel time studies, vehicle counts, etc.) if they were implemented alone within the freeway section. However, the benefits of stand alone systems are not necessarily additive. Rather, the effects of these components interact with each other. In the above example, the provision of real-time information via dynamic message signs might cause some drivers normally intending to enter the freeway to utilize another route to their destination. This would reduce the traffic demands at the entrance ramps, and influence the magnitude of benefits that would be achieved with a ramp metering system that was also implemented in the freeway corridor.

STRUCTURE OF MODULE

The next section of this module, Estimating Costs and Benefits, describes the types of costs and benefits normally associated with freeway management systems. The section that follows, System Evaluation, describes the different economic analysis techniques that are available to evaluate and justify freeway management system expenditures, and to compare various system alternatives. Using the cost and benefit estimating procedures described subsequently, designers and planners can utilize their cost and benefit data to conduct economic analyses.

11.2 ESTIMATING COSTS AND BENEFITS

SYSTEM COSTS

Types of Costs

Costs associated with freeway management systems can be classified as follows:

- Capital costs.
- Continuing costs.

Capital costs include all costs associated with the design and construction of freeway management systems (or component thereof). Items classified as capital costs include:

- Design costs.
- Right-of-way costs.
- Equipment costs.

- Construction costs.
- Software development costs (including system integration).

Continuing costs are those associated with ongoing operations of the freeway management system. These costs include the following:

- Equipment and infrastructure maintenance costs.
- Equipment replacement costs.
- Staffing costs to operate the system (operations personnel, clerical personnel, public information personnel, etc.).
- Utilities costs.
- Leasing costs (communications, control center space, etc.).

Sources of Cost Information

Each freeway management system is (presumably) a compilation of components and techniques designed to meet specific goals and objectives of the region. Technology used for freeway management is constantly being improved. In addition, the unique characteristics of each system (the components employed, the method of their integration, etc.) heavily influence the costs of previous systems. Consequently, "typical" costs associated with freeway management systems are generally not available.

Those involved in the planning, design, and evaluation of such systems are well-advised to communicate directly with vendors of the various system components to obtain the most recent cost estimates. As an alternative, personnel in other locations who have recently implemented freeway management systems with components of interest can serve as a source of cost information. However, the reader must be cautioned that these cost experiences can be heavily influenced by site-specific factors, and so may have little (if any) relevance to the freeway management system project of interest.

SYSTEM BENEFITS

Types of Benefits

An economic analysis requires the measurement of the benefit of a new or improved system relative to the existing system. To do this, the analysis compares existing conditions with those anticipated from the improvements. It must always be remembered that a freeway management system can produce a number of benefits, some of which can be quantified (e.g., the reduction in total system delay) and some of which cannot (e.g., improvement in driver perception of the transportation agencies in the region). Furthermore, while some of the quantifiable benefits can be converted to monetary value (e.g., a reduction in fuel consumption and motorist delay), other benefits do not easily lend themselves to monetary conversion (e.g., the reduction in vehicle emissions).

Traditionally, benefits analyses for traffic operations projects such as freeway management systems have typically focused on the reduction in road user costs, which can be categorized as follows:[1]

- Reductions in motorist travel time.
- Reductions in vehicle operating costs.
- Reductions in accident costs.

Vehicle operating costs typically are broken down as follows: [2]

- Fuel consumption.
- Lubricants.
- Vehicle maintenance (labor and parts).
- Vehicle depreciation.
- Interest on loans.
- Other wages.
- Overhead.

In many instances, analysis of alternative accident costs is problematic because of a lack of available data. Consequently, an evaluation may be limited to travel time and vehicle operating costs only. [3, 4]

Although not easily quantifiable in terms of dollar benefits, the effect of a traffic control system on vehicle emissions is typically a key issue. Improved traffic control systems often provide the potential for emission reductions. Fortunately, most traffic simulation models available for evaluating traffic control systems include vehicle emissions estimates as part of the measures-of-effectiveness outputs. Because of the complexities and variabilities associated with vehicle emissions, these values should be used primarily for order-of-magnitude comparisons between the various alternatives.

Sources of Benefit Information

As with cost estimates, estimates of benefits of a freeway management system cannot simply be obtained from the literature. Existing traffic conditions before system implementation, the existence and stability of working relationships between agencies, the specific combination of subsystems incorporated into the overall freeway management system all contribute to its overall success and impact in the region. Experiences with past projects indicate that benefits of freeway management may be quite substantial. [5,6] However, the analyst should focus his or her attention on identifying recent projects that bear some resemblance to his or her own situation.

Generally speaking, traffic simulation analyses provide an objective mechanism to help assess the traffic impacts of certain combinations of freeway management subsystems (traffic simulation models are discussed briefly in **Module 2**). Unfortunately, one of the major limitations of most of these models is the limited representation of how drivers actually respond to the introduction of these subsystems with respect to their route, departure time, and mode choice decisions. These limitations should be taken into consideration when interpreting the outputs of simulation. Again, sensitivity analyses are important to developing an understanding of how reasonable the simulation estimates are, and how much confidence the analyst should place in them accordingly.

IMPORTANCE OF LIFE CYCLE COST CONSIDERATIONS

Frequently, in efforts to secure capital funding for freeway management systems, the continuing costs associated with operating and maintaining these systems are not given proper consideration. Continuing costs are just as important as, if not more important than, capital costs. Adequate funding for operations and maintenance, including funding to replace system components when their useful lives have expired, is essential for successful freeway management.

When conducting economic analyses for freeway management systems, a system's continuing costs must be accounted for

along with its capital costs. Along this note, capital cost savings resulting from low bid procurement procedures can often result in greater continuing costs down the road. Capital cost savings associated with installing equipment that is inferior in terms of constructability and design features may ultimately increase the system's costs over its useful life, due to a reduced capability to handle future technology advances, more frequent replacement of equipment, etc.

Another reason for using life-cycle costs is to account for the fact that different system components often have quite different useful service lives. For example, Table 11-1 presents estimated service lives of some common freeway management system components. As the table illustrates, component service lives can range from 5 years or less to as much as 20 years. [8]

11.3 SYSTEM EVALUATION

The evaluation of the economic viability of the alternatives selected for consideration provides an objective basis for deciding which alternative, if any, should be funded. It also provides information that can be used to gain political and public support of the alternative determined to be the most beneficial (indicating the expected reduction in motorist costs or burdens for the money that will be invested). The alternatives evaluated represent a trade-off between various combinations of system components or subsystems, and can become quite complex.

A number of different considerations are required in any economic analysis of a freeway management system. The analysis itself can involve one or more of the analysis tools available. Some of the major considerations and analysis tools available are discussed in the following sections.

DEFINITION OF SYSTEM ALTERNATIVES/COMPONENTS

The decision process described in **Module 2** and emphasized in each of the other modules in this handbook is intended to provide an objective, systematic method of determining appropriate analysis alternatives. Once again, it is worth reiterating that the analyst and/or designer must base alternatives on the goals, objectives, and functions that the freeway management system is intended to achieve. The development of alternatives for analysis then evolves naturally from these intended functions. One of the options that should be evaluated is the "do nothing" alternative. This is an important benchmark to be used for evaluating the benefits of investments in freeway management alternatives.

As indicated earlier in this module regarding system benefits, a key consideration during this alternative development phase is to fully recognize the synergies that can develop from implementing certain combinations of components or subsystems of a freeway management system. For example, the implementation of closed-circuit television may not only assist in the detection and verification of an incident, but also prove useful to agencies in verifying whether a real-time traffic message is properly displayed on a nearby changeable message sign. At the same time, it is important to realistically assess how certain components or subsystems will actually perform, given the presence of other components in the system. For instance, it may be inappropriate to consider a series of inductive loop detectors installed over a section of freeway as detecting X number of incidents per day (estimated from previous incident experiences elsewhere) when a toll-free telephone hotline has also been established for cellular telephone users in the region to call in and report incidents.

Table 11-1. Selected Freeway Management System Component Service Lives [8]

Freeway Management System Component	Estimated Service Life (yrs)
Surveillance:	
Inductive Loop Detectors	5
Automatic Vehicle Identification	10
Automatic Vehicle Location	10
Video Imaging/Closed Circuit Television	10
Vehicle Emissions	10
Call Boxes	10
Toll Facilities:	
Plazas	20
Readers	10
Ramp Control:	5
Ramp Meters	
Information Dissemination:	
Dynamic Message Signs	20
Highway Advisory Radio	20
Information Kiosks	7
Communications	
Wireline	20
Transportation Management Center	
Computer Hardware	5
Computer Software and Integration	5-20

Experience suggests that the vast majority of incidents will likely be detected from the hotline, rather than from the loop detectors.[7] Additional information concerning incident detection technologies is presented in **Module 8**.

ASSIGNING BENEFITS TO ALTERNATIVES

Also, as stated earlier in this module, it is sometimes very difficult to properly assess the benefits associated with the implementation of a freeway management system, when the system is comprised of several different components which interact and affect traffic in some unknown manner. The possible interrelationship between dynamic message signs and ramp metering subsystems was offered as one example. On the other hand, truck lane restrictions and a ramp metering subsystem could generate possibly adverse interactions (e.g., by having all trucks in the lane where ramp vehicles were attempting to merge).

Unfortunately, no hard and fast rules are available to determine how the introduction or elimination of specific components or subsystems affects the impact of other components in a freeway management system. The analyst must rely on judgement and experience to generate a best-guess in these instances. The uncertainty inherent in

the analysis of benefits is why sensitivity analyses are so important to include as part of the analysis plan. (Sensitivity analysis is discussed briefly later in this module.)

ESTIMATING LIFE-CYCLE COSTS

When estimating the costs of a freeway management system alternative, it is critical to include not only the capital costs associated with designing, purchasing, and/or constructing the system alternative, but also the ongoing operating and maintenance costs that will be required to keep the alternative operational until it reaches the end of its useful life cycle. These life-cycle costs are the most appropriate way to evaluate freeway management system components.

Data from vendors or other operating agencies should be consulted when attempting to identify the various costs associated with a particular freeway management system alternative. In the absence of available data, some analysts have taken a percentage of the capital costs of a component as a measure of the operations and maintenance costs. The recent *ITS Architecture Cost Analysis*, for example, estimates operations and maintenance costs for many of the ITS technologies to be implemented in the future as between 2 and 10 percent of expected capital costs of those technologies. [8]

ANALYSIS TECHNIQUES

Net Present Worth

Description

Computation of an alternative's net present worth involves a conversion of all costs and benefits of an alternative that are incurred at the alternative's initiation and throughout its useful life (life-cycle) to an equivalent current value.[9] The current value of the equivalent costs is subtracted from the current value of the equivalent benefits of the alternative. If the benefits exceed the costs, the alternative can be justified economically. Furthermore, comparisons among alternatives are straightforward; the alternative that provides the greatest additional benefits over costs (sometimes referred to as "excess benefits") is said to have the greatest net present worth.

Example

An analysis of alternative communication systems for a traffic control system in a central business district provides an excellent example of the application of the net present worth method to alternatives evaluation.[10] Alternatives considered included installed twisted-wire pair (TWP), lines leased through the telephone company (TELCO), or installed fiberoptic lines. Table 11-2 summarizes the estimated capital and maintenance costs for each alternative over its 15-year lifespan. As the table illustrates, the TWP alternative provides the lowest net present worth of costs to the agency ($1,925,764).

Other Considerations

This analysis approach is perhaps the simplest to explain and the most understandable to the general public. Unfortunately, the net present worth method does not necessarily convey the relative economic merits of various alternatives as compared to the risk associated with the investment in that alternative. Most freeway management systems must compete for limited funds with other types of transportation improvement projects. Two alternatives may be capable of generating the same amount of excess benefits over their costs. However, if one alternative requires only a fraction of the cost of the other

Table 11-2. Example of a Net Present Worth Evaluation (Adapted from10)

Alternative		Capital Costs($)	Annual Maintenance Costs($)	Maintenance Net Present Worth ($)[a]	Total Net Present Worth ($)
Twisted-Wire Pair (TWP)					
Conduit:	15,600@$50	780,000	39,000	367,501	
RCU:	196@$1,200	235,200	23,520	221,631	
Modems:	33@$180	5,940	594	5,597	
Cable:	76,000@$1	76,000	3,800	35,808	
	1,600@$1.50	2,400	120	1,131	
	2,000@$2	5,000	250	2,356	
Prep:	62,400@$3	187,200	0	0	
TOTAL NPW		**1,291,740**		**634,024**	**1,925,764**
Leased Telephone Lines (TELCO)					
Conduit:	3,920@$50	196,000	9,800	92,346	
RCU:	196@$1,200	235,200	23,520	221,631	
Modems:	33@$180	5,940	594	5,596	
Cable:	3,920@$1.80	7,056	353	3,326	
	3,920@$3.30	12,936	647	6,096	
Other:	196@$143	28,028	0	0	
Annual Lease Charges:		0	138,336	1,303,553	
TOTAL NPW		**485,160**		**1,632,548**	**2,177,708**
Fiberoptic Lines					
Conduit:	15,600@$50	780,000	39,000	367,501	
RCU:	196@$1,500	294,000	29,400	277,039	
Modems:	33@$415	13,695	1,370	12,910	
Cable:	78,000@$3	234,000	11,700	1010,250	
Prep:	62,400@$3	187,200	0	0	
TOTAL NPW		**1,508,895**		**767,700**	**2,276,595**

[a] Annual Maintenance Costs multiplied by present worth factor (7%, 15 years) of 9.4231
RCU = Remote Communications Unit

alternative, the partners receive a better return on the investment by choosing the lower-cost alternative (and are risking a lower amount of capital in so doing). Consequently, a benefit-cost analysis, as described in the next section, provides a more accurate picture of the relationship between the potential benefits and costs of freeway management system alternatives.

Benefit-Cost and Incremental Benefit-Cost Analysis

Description

The Benefit-Cost (B/C) analysis technique is perhaps the most widely accepted methodology for evaluating transportation improvement alternatives. The B/C ratio is simply the equivalent benefit of an alternative divided by the equivalent cost of that alternative:

$$B/C = \frac{\text{(benefits of alternative i)}}{\text{(costs of alternative i)}}$$

Benefit-cost comparisons are possible when the benefits of an improvement can be assigned a monetary value. If the benefits of an alternative exceed its costs, the improvement is economically justifiable. Furthermore, the ratio of each alternative provides a convenient basis for comparison, providing a measure of the dollars of expected benefit of an alternative for each dollar spent on that alternative.

If system alternatives being analyzed build upon each other in terms of the costs, quantities, complexities, etc. of components that meet the system goals and objectives, it may be more appropriate to consider an incremental benefit-cost analysis. For this approach, the benefits and costs considered for each alternative are not the totals, but rather the additional benefits achieved and costs incurred over the next expensive (and presumably effective) alternative. This analysis considers, in effect, whether an investment necessary to achieve the next incremental step in the system can be justified in terms of the incremental benefits that would be achieved.

Example

Table 11-3 presents the results of a benefit-cost sketch-planning analysis of alternatives to address mobility and congestion problems in Salt Lake City, Utah.[11] Analysts identified five alternative improvement categories:

- Highway capacity expansion (HWY) to add freeway lanes.
- Transportation system management (TSM) actions (enhanced bus service in the corridor).
- HOV lane addition.
- Light Rail Transit (LRT) addition to the corridor.
- Travel Demand Management (TDM) action to collect tolls within the corridor.

Capital, operation, and maintenance costs for both the highway and transit agencies were calculated for each alternative, based on available literature. Benefits estimated for each alternative consisted of estimated travel time savings, vehicle operating cost reductions, reduced vehicle emissions, accident cost reductions and other external benefits, and revenues generated through additional parking fees. Based on the assumptions made, some of the alternatives were estimated to experience negative benefits (sometimes referred to as dis-benefits) in one or more of these categories.

Table 11-3. Example of a Benefit-Cost Analysis [11]

	Benefits and Costs (in millions of dollars)				
	HWY	TSM	HOV	LRT	TDM
Benefits					
User	22.5	21.8	24.1	20.6	-16.6
Revenues	0.0	0.0	0.0	2.5	33.1
External	-21.0	2.0	4.3	5.7	22.7
Emissions	-0.9	0.6	2.3	1.9	5.0
TOTAL	0.6	24.4	30.6	30.7	44.2
Costs	4.7	5.9	8.	48.1	14.7
B/C Ratio	0.12	4.14	3.45	0.64	3.01

As table 11-3 illustrates, the TSM alternative was estimated to provide the highest B/C ratio (4.14). This indicates an expected $4.14 benefit to be received for every $1 spent by the public agency on the alternative. The HOV and TDM alternatives also provided a reasonable B/C ratio, returning $3.45 and $3.01, respectively, for every $1 invested on that alternative.

The HWY and LRT alternatives, however, resulted in B/C ratios less than 1. This indicates that these alternatives are not attractive investments, providing only $0.12 and $0.64 benefits per $1 invested. It should be noted that significant external disbenefits (expected accident increases due to increased vehicle demand on the facility) caused the relatively poor result for the HWY alternative. For the LRT alternative, the high capital cost required was the main factor causing the low B/C ratio.

As an example of the incremental B/C analysis, suppose that the TSM alternatives described above had initially consisted of potential bus service improvements on four different routes in the corridor. Estimated benefits and costs for each route might have been as shown in table 11-4. Examining the economic viability of each route separately clearly shows that the expenditures on routes 1 through 3 would be justified (B/Cs for each greater than 1), whereas those for route 4 would not (B/C less than 1). Furthermore, if the analyst had only considered the four routes together without analyzing them incrementally (using the numbers in the last column of table 11-4), he or she would have incorrectly concluded that the TSM improvements should be made to all four routes.

Other Considerations

The benefit-cost (or incremental benefit-cost) analysis methodology provides an objective means of comparing the quantifiable and monetarily-based benefits of an alternative to the costs of that alternative. Unfortunately, not all quantifiable benefits are easily converted to a monetary value (the reduction in vehicle emissions, for example). Likewise, some benefits are not easily quantified (e.g., the improvement in public perception of a transportation agency due to improving the availability of traffic information in the freeway corridor).

Table 11-4. Example of an Incremental Benefit-Cost Analysis

	TSM Components (Bus Service Improvements)				
	Route 1	Route 2	Route 3	Route 4	Total
Benefits	7.1	10.0	7.3	2.0	26.4
Costs	1.0	2.0	2.9	4.0	9.9
B/C Ratio	7.1	5.0	2.5	0.5	2.7

Because of this, alternative analyses are often needed to help assess which alternative systems or subsystems meet their objectives in the most economical manner. The utility cost analysis, described below, is one such analysis approach.

Utility-Cost Analysis

Description

Although a benefit-cost (or incremental benefit-cost) analysis is a direct method of determining whether a freeway management system alternative is economically viable, such an analysis can be performed only if the benefits to be accrued can be estimated in monetary terms. For many goals and objectives of freeway management, this is not possible. In these cases, a utility-cost analysis approach is commonly utilized. The term cost-effectiveness is sometimes used interchangeably with the term utility-cost analysis.[12]

In a utility-cost analysis, utility measures of performance goals or objectives are created to estimate system benefits. Typically, a project team or expert panel subjectively rates (from 0 to 10 or on a similar scale) how well an alternative is expected to achieve each of the objective or performance criteria. Weighting factors (summing to unity) are also estimated for each of the objective or performance criteria, and multiplied by the rating given to that objective/criterion. These "utilities" of each of the objective/criteria are then summed to determine the total system utility. Dividing the system utility by total system cost represents the utility-cost factor for a particular system. The basic steps in a utility-cost analysis are as follows: [4, 9]

- Define goals and subgoals (done as part of the decision process).
- Weigh each goal.
- Weigh each subgoal.
- Rate the utility of each goal/subgoal.
- Compute utility-cost ratio.

Example

Figure 11-2 illustrates a computational procedure presented elsewhere, that rates each alternative of a traffic control project against the specified requirements for that project. [13] The ability of each alternative system to meet each requirement is rated 0 to 10. Zero indicates that the system does not satisfy the requirement at all, while 10 indicates total satisfaction. This internal rating scale measures how well an alternative system satisfies a requirement. Multiplying

the individual rating by a relative weight creates the utility value for that requirement. The sum of individual utility values gives an overall utility rating for each alternative.

Other Considerations

It is tempting to use the utility cost ratio only to compare alternative system designs, selecting the one with the greatest ratio. However, this conclusion may prove faulty because a simple, inexpensive system with low utility and low cost may have the same ratio as a sophisticated, but more expensive system with high utility and satisfying all defined requirements. The analyst must also assess whether the low-cost option satisfies all of the system objectives and functions. Conversely, the analyst must determine if funds are available to implement the more expensive alternative. If the alternative exceeds available funds, one or more of the components may need to be removed or modified, leading to a new rating of expected performance and a new utility cost ratio.

Value Engineering

Description

Value engineering is an organized effort directed at analyzing the function of an item with the purpose of achieving the required function at the lowest overall cost.[14] The relationship between value and function is expressed as follows:[15]

Value = (Functional Performance)/Cost

From this equation, it is evident that value is increased in one of two ways:

- Value is increased by reducing costs, if performance is maintained.
- Value is increased by increasing performance (but only if increased performance is needed and the user is willing to pay for it).

A critical step in any value engineering is to identify and assess the function(s) of the item or service under study. These functions are categorized as basic or secondary. Basic functions are those which are absolutely essential in order for the item or service to perform its purpose. Secondary functions are those related to esteem, appearance, or convenience. Finally, the basic functions are organized into a logical, hierarchial sequence. This makes it possible to identify the principle function(s). It is these principle function(s) which are used to judge value.[16]

Example

An evaluation of candidate channelizing devices to be used during a work zone lane closure on a section of freeway are presented in tables 11-5 and 11-6. Six alternative channelizing devices are considered:

- Type I Barricades.
- Type II Barricades.
- Vertical Panels.
- Cones.
- Tubes.
- Drums.

Key measures of functional performance to be included in the analysis are mean detection of an array of each type of device at the closure, and the mean distance at which drivers change lanes upstream of the closure when each device is present. Table 11-5 summarizes these distances.[16] Table

Requirement	Relative Importance of Each Requirement Sum = 100	Ability of Alternative Systems to Meet Each Requirement Rating (0 to 10)					
		Alternative System 1		Alternative System 2		Alternative System 3	
1. Intensity of CCTV coverage	25	4	100	7	175	10	250
2. Spacing of mainline monitoring station	25	2	50	6	150	8	200
3.	8	5	40	4	32	7	56
4.	4	10	40	6	24	6	24
5.	16	10	160	7	112	7	112
	etc.						
	100		390		493		642

Relative weight scale
How important is this requirement in this jurisdiction? The total of these values must equal 100.

Internal rating scale
How well does this alternative system satisfy this requirement? The range of this value is between 0 and 10 where 0 means the system cannot accomplish this function, and 10 means the requirement is totally satisfied.

Utility value
Obtained by multiplying the internal rating scale by the relative weight of each requirement. These values are summed to determine the total utility of each alternative.

Figure 11-2. Example of Utility Cost Analysis. [13]

11-6 then provides the costs and resulting value estimates for each device.

Engineering judgement must be used to some degree in value engineering analyses. As can be seen in table 11-6, no one device provides the best value for both functional performance measures in both day and night conditions. In this analysis, vertical panels and drums were said to be "good values" (entailing lower costs for unit functional performance) for combined day and night use at freeway work zones. They were at or near the bottom of cost per unit of detection distance or lane change distance for both day and night conditions.[16]

Other Considerations

A project team or expert panel approach is used in this analysis process, just as for a utility-cost evaluation. The principal difference between value engineering and utility cost evaluation is in how item performance is accounted for in the analysis. Whereas the utility cost approach assigns a subjective measure of utility to otherwise nonquantifiable performance measures, the value engineering approach depends on the ability of the analyst (or project team) to define a quantifiable measure of performance for the primary function(s) of the alternative being evaluated.

Table 11-5. Mean Detection and Lane Change Distances for Channelizing Devices. [16]

Device Type	Mean Array Detection Distance (ft)		Mean Lane Change Distance Before Taper (ft)	
	Day	Night	Day	Night
Type I Barricades	4,250	3,150	640	660
Type II Barricades	4,100	2,800	400	810
Vertical Panels	4,400	3,300	370	500
Cones	4,400	1,450	460	250
Tubes	3,200	1,900	620	350
Drums	4,200	3,000	540	560

Table 11-6. Example of Alternative Channelizing Device Values. [16]

Device Type	Device Array Cost ($)	Cost Per 100 ft Array Detection Distance ($)		Cost Per 100 ft Mean Lane Change Distance ($)	
		Day	Night	Day	Night
Type I Barricades	40	0.94	1.27	6.25	6.06
Type II Barricades	45	1.10	1.61	11.25	5.56
Vertical Panels	22	0.50	0.67	5.95	4.40
Cones	18	0.41	1.24	3.91	7.20
Tubes	22	0.69	1.16	3.55	6.29
Drums	25	0.60	0.83	4.63	4.46

Sensitivity Analysis

Description

Uncertainty is a part of most economic analyses of freeway management system alternatives. Quickly changing technology results in cost estimates during construction and operations that may differ from those obtained during the data collection phase of the analysis. Estimates of benefits are commonly based on simulation analyses or upon documented experiences of other agencies that have implemented similar alternatives. The extent to which these experiences will be realized in the project of interest are generally not known until after implementation.

It is always tempting to avoid an explicit evaluation of uncertainty by basing the analysis on a "conservative" estimate of the uncertain variables. Unfortunately, the analyst may not always know whether a conservative estimate—in the sense that it is supposed to lead to an underestimate of net benefits—requires a value that is higher or lower than the most likely estimate.[12] The simplest way of allowing for uncertainty is by means of a sensitivity analysis: systematically varying the value of key assumptions and parameters to objectively evaluate their importance. A sensitivity analysis is appropriate for any type of economic analysis (net present worth, benefit-cost, utility-cost, or value engineering).

Example

The example presented earlier for the benefit-cost analysis provides an excellent example of the importance of conducting sensitivity analyses in economic evaluations. The results presented in table 11-3 indicated that bus scheduling improvements (labeled as the TSM category) provided the best benefit-cost ratio, followed closely by the addition of an HOV lane and transportation demand management (TDM) alternatives. Building additional freeway capacity (HWY alternative) and a light rail transit (LRT alternative) were computed to have benefit-cost ratios less than 1. However, these results assumed that a significant increase in vehicle demand would be generated by the addition of freeway capacity, and that this would increase accident rates and costs significantly. Likewise, TDM actions were assumed to result in significant decreases in

travel demand, significantly reducing travel demands.[11]

Analysts tested the sensitivity of the results to this assumption by computing a benefit-cost ratio for each alternative without assuming significant changes in demand. In other words, they evaluated the effects of these alternatives upon existing travel demand patterns. The results, shown in table 11-7, indicate extreme sensitivity to the assumption about changes in future demand. Whereas the HWY alternative was not viable (benefit-cost ratio less than 1) when significant accident cost increases were assumed to occur, the alternative becomes the most attractive alternative (benefit-cost ratio of 4.60) when accidents are not assumed to increase significantly over the existing levels.[11]

Other Considerations

The range of variation of key assumptions and parameters in a sensitivity analysis can either be a specified amount, say 10 percent above and below the central estimate, or can be related to the inherent uncertainty associated with the central estimate itself. In the latter case, the ranges would represent best-guess estimates of possible "high" and "low" values of the parameters, not necessarily uniformly distanced from the most likely value.

A properly conducted sensitivity analysis quickly shows which estimates are most important and gives the engineer a general idea of the following: [2]

- Those aspects which justify further work to narrow the range of uncertainty.
- The qualitative uncertainty associated with the scheme as a whole.

When a quantitative evaluation of uncertainty about benefits associated with freeway management is required, traffic simulation may be useful (although it is possible to evaluate uncertainty using mathematical expectations, this approach is usually too complicated for all but the simplest applications). The analyst can vary each of the parameters systematically, and determine the influence of that parameter upon the overall estimates of performance obtained with simulation.

Table 11-7. Example of Sensitivity Analysis of Benefit-Cost Ratios. [Adapted from 11]

	Benefit-Cost Ratio				
	HWY	TSM	HOV	LRT	TDM
Benefit-Cost Ratio (with accident increases)	0.12	4.14	3.45	0.64	3.01
Benefit-Cost Ratio (without accident increases)	4.60	4.42	4.25	0.74	3.94

11. 4 REFERENCES

REFERENCES

1. *A Manual on User Benefit Analysis of Highway and Bus Transit Improvements*. American Association of State Highway and Transportation Officials, Washington, DC, 1978.

2. Heggie, I.G. and Thomas, S. Economic Considerations. In *Transportation and Traffic Engineering Handbook.* W.S. Homburger, Editor. Second Edition. Institute of Transportation Engineers. Prentice-Hall, Englewood Cliffs, NJ, 1982.

3. Memmott, J.L. and Dudek, C.L. *A Model to Calculate the Road User Costs at Work Zones*. Report No. FHWA/TX-83/20+292-1. Texas Transportation Institute, College Station, TX, September 1982.

4. Blumentritt, C.W., Pinnell, C., McCasland, W.R., Ross, D.W., and Glazer, J. *Guidelines for Selection of Ramp Control Systems*. NCHRP Report No. 232. TRB, National Research Council, Washington, DC, May 1981.

5. Intelligent Transportation Infrastructure Benefits: Expected and Experienced. Report No. FHWA-JPO-96-008. FHWA, U.S. Department of Transportation, Washington, DC, January 1996.

6. McDermott, J.M., McLean, C.H., and Smith, A.T. Three Decades of Progress: Freeway Traffic Management in Illinois. *ITE Journal*, Vol., 62, No. 3, March 1992, pp 27-31.

7. McLean, C.H. Cellular Phones—A Key Traffic Management Component. *Compendium of Technical Papers*, Institute of Transportation Engineers, Washington, DC, 1991.

8. *ITS Architecture Cost Analysis*. FHWA, U.S. Department of Transportation, Washington, DC, June 1996.

9. Collier, C.A., and Ledbetter, W.B. *Engineering Cost Analysis*. Harper & Row, New York, NY, 1982.

10. Gordon, R.L., Reiss, R.A., Dunn,W.M., and Morehead. D.R. *Communications Handbook for Traffic Control Systems*. Report No. FHWA-SA-93-052. FHWA, U.S. Department of Transportation, Washington, DC, April 1993.

11. Decorla-Souza, P., Cohen, H., and Bhatt, K. Using Benefit-Cost Analysis to Evaluate Across Modes and Demand Management Strategies. *Compendium of Technical Papers*, Institute of Transportation Engineers 66th Annual Meeting, Minneapolis, MN, September 1996, pp 439-445.

12. Gordon, R.L., et. al. *Traffic Control Systems Handbook.* Report No. FHWA-SA-95-032. FHWA, U.S. Department of Transportation, Washington, DC, 1996.

13. Wilshire, R., et al. *Traffic Control Systems Handbook Revised Edition - 1985,* FHWA-IP-8511, FHWA, U.S. Department of Transportation, Washington, DC, April 1985.

14. Miles, L.D. *Techniques for Value Analysis and Engineering*. McGraw-Hill, Inc., New York, NY, 1972.

15. *Value Engineering Conference: Summary Report*. Report No. FHWA-TX-80-246. FHWA, U.S. Department of Transportation, Washington, DC, October 1980.

16. Richards, S.H., and Dudek, C.L. *Selection of Work Zone Channelizing Devices Using the Value Engineering Approach*. Report No. FHWA/TX-85/06+292-3. Texas Transportation Institute, College Station, TX, January 1984.

SUGGESTED READING

Cost and Effectiveness of Transportation Control Measures (TCMs): A Review and Analysis of the Literature. Prepared by Apogee Research, Inc. for the National Association of Regional Councils, January 1994.