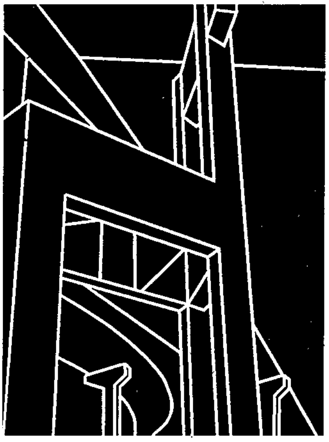
RESEARCH REPORT 1468-1

PB99-145229

INTEGRATED ARTERIAL AND FREEWAY OPERATION CONTROL STRATEGIES FOR IVHS ADVANCED TRAFFIC MANAGEMENT SYSTEMS: RESEARCH REPORT

Hani S. Mahmassani, Didier M. Valdes, Randy B. Machemehl,
John Tassoulas, and James C. Williams



CENTER FOR TRANSPORTATION RESEARCH
BUREAU OF ENGINEERING RESEARCH
THE UNIVERSITY OF TEXAS AT AUSTIN

SEPTEMBER 1998

Technical Report Documentation Page

1. Report No. FHWA/TX-99/1468-1 PB99-145229	3. Recipient's Catalog No.
4. Title and Subtitle INTEGRATED ARTERIAL AND FREEWAY OPERATION CONTROL STRATEGIES FOR IVHS ADVANCED TRAFFIC MANAGEMENT SYSTEMS: RESEARCH REPORT	5. Report Date September 1998 6. Performing Organization Code
7. Author(s) Hani S. Mahmassani, Didier M. Valdes, Randy B. Machemehl, John Tassoulas, and James C. Williams	8. Performing Organization Report No. 1468-1
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3208 Red River, Suite 200 Austin, TX 78705-2650	10. Work Unit No. (TRAIS) 11. Contract or Grant No. 0-1468
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Transfer Section/Construction Division P.O. Box 5080 Austin, TX 78763-5080	13. Type of Report and Period Covered Research Report (8/96 — 9/97) 14. Sponsoring Agency Code

15. Supplementary Notes

Project conducted in cooperation with the Federal Highway Administration.

16. Abstract

The main focus of this study is congestion, primarily that occurring on freeway corridors in metropolitan areas. Lack of coordination in the operation of various components of the system is often a major source of inefficiency, resulting in greater delays to motorists than what might be achievable with the existing physical infrastructure. Inefficiency owing to a lack of coordination may be the result of jurisdictional issues in terms of different entities having operational responsibility for different parts of the system. Typically, the respective control settings for the various subsystems in a freeway corridor are not designed to operate together in an integrated way. The consequences are particularly acute when incidents occur and where there is an attendant loss of capacity, accompanied by possible redistribution of flows; moreover, the control settings along likely diversion paths are not designed to react to accommodate the unfolding situation. The main objective of the study is to improve corridor network management by coordinating the various control elements in a freeway corridor, for both recurrent and nonrecurrent congestion situations.

17. Key Words Congested freeway corridors, arterial and freeway management strategies, integrated control system design methodologies, DYNASMART	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.		
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 136	22. Price

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

INTEGRATED ARTERIAL AND FREEWAY OPERATION CONTROL STRATEGIES FOR IVHS ADVANCED TRAFFIC MANAGEMENT SYSTEMS: RESEARCH REPORT

by

Hani S. Mahmassani
Didier M. Valdes
Randy B. Machemehl
John Tassoulas
James C. Williams

Research Report Number 1468-1

Research Project Number 0-1468

Project title: *Integrated Arterial and Freeway Operation Control Strategies for IVHS Advance Traffic Management Systems*

Conducted for the

TEXAS DEPARTMENT OF TRANSPORTATION

in cooperation with the

U.S. DEPARTMENT OF TRANSPORTATION

Federal Highway Administration

by the

CENTER FOR TRANSPORTATION RESEARCH
Bureau of Engineering Research
THE UNIVERSITY OF TEXAS AT AUSTIN
in cooperation with
THE UNIVERSITY OF TEXAS AT ARLINGTON

September 1998

IMPLEMENTATION RECOMMENDATIONS

The findings of this study can be used by traffic engineers, Texas Department of Transportation and Federal Highway Administration officials, and by cities in Texas to evaluate and reduce urban congestion. Final project implementation recommendations are provided in Project Summary Report 1468-S.

This report was prepared in cooperation with the Texas Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.

DISCLAIMERS

Hani S. Mahmassani, P.E. (Texas No. 57545)
Research Supervisor

ACKNOWLEDGMENTS

This report is based on research funded through a cooperative study between the Center for Transportation Research at The University of Texas at Austin and the Texas Department of Transportation (Research Study No. 0-1468). The authors are grateful to Jon P. Underwood, project director, and Tom Newbern, project coordinator, for their guidance and support. We also acknowledge the contributions made by the individual members of the project monitoring committee, which included C. Molina (TRF), T. Nyguyen (FTW), C. Rawson (TRF), A. Rolla (TRF), and D. Vanover (HOU).

TABLE OF CONTENTS

SUMMARY

Many urban areas suffer from congested freeway corridors. Frequently such congestion is the result of motorists choosing to stay on a freeway to bypass uncoordinated arterial signals at freeway interchanges. Integrated arterial and freeway management strategies can provide effective coordinated incident management to help minimize the effects of urban congestion. The objective of this study is to enable achievement of truly integrated freeway, frontage road, and arterial street corridor traffic control. This study has focused on examining current freeways, frontage roads, and arterial street control systems, including ramp, lane, interchange, and intersection controls, in an effort to determine more effective integration methodologies.

Several strategies were tested using the simulation assignment model integrator DYNASMART. In addition, the TEXAS model and other traffic simulation models were used to perform a more detailed analysis. The results of the experiments suggest that adjusting the control settings of diamond and arterial signals to receive additional demand can have significant implications for the evolution of conditions in the traffic system. New control strategies were also devised and tested, particularly path-based coordination. Some improvement was found with the application of the new control strategy. On the other hand, the application of information and route guidance strategies benefits the whole network to a much higher degree. The experiments performed as part of this study suggest that traffic management objectives can be better achieved by combining traffic control and route guidance in the ATMS/ATIS strategies.

The results of this study are applicable in the design and implementation of integrated freeway corridor traffic control systems. Methodologies and techniques presented in this report will not only reduce freeway corridor congestion costs and incidents, but will also reduce fuel consumption costs, vehicular emissions, and personal time loss.

CHAPTER 1. STUDY OBJECTIVES AND ACHIEVEMENTS

1.1. INTRODUCTION

This study has been prompted by the need to address the problem of congestion, primarily that associated with freeway corridors in metropolitan areas. Lack of coordination in the operation of the various system components is often a major source of inefficiency, resulting in delays to motorists greater than what might be achievable through the existing physical infrastructure. Inefficiency resulting from a lack of coordination may be the result of jurisdictional issues — that is, the fact that different entities have operational responsibility for different parts of the system. However, the present state of the art in traffic systems management and control is still lacking in terms of integrated control strategy designs that explicitly consider arterials, diamond interchange configurations, and freeways as a single interacting system. Typically, the respective control settings for the various subsystems in a freeway corridor are not designed to operate in an integrated way. The consequences are particularly acute in cases where incidents lead to a loss of capacity accompanied by flow redistributions, and where the control settings along likely diversion paths are not designed to react to accommodate the unfolding situation. The main motivation, then, is to improve corridor network management by coordinating the various control elements in a freeway corridor, for both recurrent and nonrecurrent congestion situations.

Representing a primary bottleneck in efforts to coordinate operations — and hence a significant area for improvement — is the diamond intersection, where, when incidents occur, the flow patterns can differ substantially from those for which the control was designed. Further compounding the problem is the fact that such situations do not allow adequate time for extensive analysis; it is necessary, therefore, to have strategies in place for immediate deployment in response to these problems.

Significant opportunities have emerged for improving current operations. Information and communication technologies that can enable control integration have developed rapidly over the past few years; such technologies can now be used not only to detect unfolding conditions in the system (permitting appropriate responses), but also to anticipate how flows may evolve within the corridor. Applications of intelligent transportation system (ITS) technologies to advanced traffic management systems (ATMS) provide capabilities for the integrated management of various roadway functions, as well as real-time responsive traffic control during incidents. These ATMS technologies thus allow greater responsiveness of the control strategies to actual conditions.

ATMS concepts envision a traffic management center (TMC) in charge of monitoring operations and developing and implementing control actions over the traffic network in a metropolitan area. The TMC receives information in real-time regarding prevailing conditions from several sources, including various detection devices (e.g., loop detectors in the pavement, video imaging, and automatic vehicle identification devices) and vehicle probes — vehicles equipped with location and navigation devices and two-way communication capabilities that relay information about prevailing speed and location continuously to the control center. With this information, the TMC can detect incidents and can identify congested locations, bottlenecks, and underutilized capacity. This information, in turn, serves as a basis for undertaking traffic control actions, including responding to incidents, setting traffic signals, generating variable message signs that can inform users and influence their route choices, and providing route guidance instructions to vehicles equipped with two-way communication capabilities and in-vehicle display units (Refs 1, 2, 3). The provision of real-time information to tripmakers is referred to as *advanced traveler information systems* (ATIS). Coordinated freeway corridor network control lies at the confluence of ATMS and ATIS capabilities.

To support the above role of the TMC with regard to ATMS/ATIS functions, several methodological capabilities are required to (1) process the volumes of incoming information, (2) analyze network operations, and (3) determine control actions that optimize network performance. Central among these methodologies are dynamic traffic assignment techniques and several associated support functions that must be integrated and bundled into a dynamic traffic assignment (DTA) system. Two essential capabilities are required of the DTA system (Ref 4). The first capability, referred to as *descriptive*, consists of describing how flow patterns develop spatially and temporally in a traffic network, typically given a set of desired trips between origins and destinations. This descriptive capability allows both *estimation of the current state of the network* (especially when the network is only partially observable) and *prediction of future network states over time*. At its core is the ability to model the outcomes of tripmaker decisions, primarily the decision regarding which route to take between origin and destination, as well as (possibly) the decision when to depart and what mode to use. To the extent that these decisions are predicated on network conditions, which in turn depend on the users' decisions, network states must be determined simultaneously with tripmaker choices, generally in an iterative scheme. Both the estimated state of the network and the predicted future state, in terms of flows, travel times, and other time-varying performance characteristics on the various components of the network, are used in the on-line generation and real-time evaluation of a wide range of ATMS measures and ATIS messages.

The core of the descriptive DTA capability is a traffic simulation model that seeks to capture the dynamics of traffic flow movement in the network. Developed at The University of Texas at Austin, this simulation model, dubbed DYNASMART (Dynamic Network Assignment-Simulation Model for Advanced Road Telematics), presently defines the state of the art in this arena. DYNASMART combines the functions of traffic simulation, which entails moving vehicles in a system, with the path assignment or routing of the vehicles in that system. Unlike conventional traffic simulation packages that assume steady state with fixed turning movements, DYNASMART can capture real-world dynamic traffic patterns. Within DYNASMART, it is possible to assign vehicles to particular paths and then track their movements along these paths. The DTA capability can either specify which path a particular vehicle should follow, or model the path followed by motorists as they distribute themselves through the network. These methodological developments allow the TMC not only to consider the traffic systems per se, but also to influence the distribution of flows over the network, thereby integrating routing with signal and traffic control (Refs 5, 6). DYNASMART is described in more detail in Chapter 3 of this report.

Furthermore, with greater recognition of the need for integrated management, more emphasis is placed on developing and implementing mechanisms that can achieve the kind of multijurisdictional coordination that is required. Such coordination is an essential complement to the deployment of new technologies and associated methodologies.

1.2. STUDY OBJECTIVES

The main objective of this study is to facilitate truly integrated freeway, frontage road, and arterial street corridor traffic control. The following components have been designed to accomplish this overall study objective:

1. Examine current freeway, frontage road, and arterial street control systems, including ramp, lane, interchange, and intersection controls, and identify impediments to, and requirements for, effective integration.
2. Identify specific solutions to integration problems stemming from methodological, hardware, or institutional issues.
3. Develop state-of-the-art control methodology implementing solutions to integration problems, enabling seamless freeway main lane, frontage road interchange, and arterial street system traffic control integration.
4. Develop specific guidelines for implementation of the integrated control strategies in the selected test bed, with such a test serving to demonstrate general applicability and to determine subsequent specific steps required for "live" testing.

This study classifies congestion as either "normal" or "nonrecurrent." These two types of congestion are more precisely categorized as (1) a nominal state (for everyday congested situations) and (2) surges related to demand (e.g., special events) or to supply (e.g., incidents).

1.3. STUDY ACHIEVEMENTS AND REPORT ORGANIZATION

The methodological approach devised to address the main objective included the conceptual integrated control system design, simulation model selection and modification, test bed selection and simulation, and analysis of each specific subsystem in an urban freeway corridor. The integrated control system design was developed in the first stages of this study. It included the specification of system components and their interrelations. Overall, this part of the study provided a guide for the rest of the conceptual and methodological development.

Another stage of the study addressed the selection and modification of the simulation model to support the design and evaluation of integrated operational strategies. Under this activity, the DYNASMART simulation model was selected as "system integrator." Important additions were made to the model in order to simulate all the pertinent interactions occurring among the different subsystems of the urban freeway corridor. The analysis of the operational strategies was conducted in a specific test bed (a portion of the Fort Worth network was selected for this purpose). The necessary data were collected and all the required files were produced in order to simulate the selected test bed.

Several important accomplishments were attained in this study in terms of the development and integration of control strategies. Extensive tests and associated simulations were conducted in order to develop control strategies specifically targeting coordination of diamond intersections, frontage roads, and path-based arterial streets. Among the strategies tested, we selected the path-based coordination using vehicle-actuated traffic signals, given that it provided the best results for the overall network under consideration. However, the most effective control strategies for individual subsystems are also presented in this report. These strategies could be used in conjunction with path-based coordination to generate many possible designs for the control strategy that is ultimately implemented under particular conditions in the actual freeway corridor.

Route guidance strategies, which were also tested using the test bed under congested conditions, demonstrated excellent performance in terms of travel-time savings. This is among the most significant achievements of this study: The findings suggest that, in order to obtain the benefits envisioned under ATMS/ATIS strategies, it is necessary to implement higher-level control measures combined with information strategies. A central focus of this study was to combine both control and route guidance strategies, a combination that yielded benefits ranging between 20% and 40% in terms of savings in travel time for all the vehicles in a congested network. The details of this effort are described in the set of experiments included in Chapter 4.

The remaining portion of this report is organized as follows: Chapter 2 describes the modeling framework developed to study the system control integration for a freeway corridor. Chapter 3 then discusses the details of the models used to simulate every subsystem as well as the model used to simulate the integration strategy for the entire network. Chapter 4 presents the framework application through five sets of simulation experiments that included both route-guidance strategies and control strategies. This chapter also includes a demonstration of the possible benefits of implementing combined control and route-guidance strategies in congested freeway corridors. Concluding comments are presented in Chapter 5.

CHAPTER 2. THE MODELING FRAMEWORK

2.1. INTRODUCTION

The conceptual design of control integration takes into account all the control elements of each type of facility involved in a freeway corridor, including freeway mainlanes, ramps and frontage roads, diamond interchanges, and arterial streets. While the operation of each individual facility is generally well understood in practice, the operation of the system in an integrated fashion still requires additional development, both methodologically and in the field. The diamond interchange, the principal element that joins freeways and arterial streets, represents a special challenge for integration, given that it is the node complex wherein all vehicles come together and where many problems may develop. Indeed, diamond interchanges can be the source of bottlenecks that prevent these various components from working at their optimum level and, moreover, can constrain the overall system's efficiency and throughput.

Understanding these challenges to control element integration has important implications for the study approach and for the tools selected and/or developed to achieve the study objectives. Most traffic analysis tools have been developed for either freeways or arterial streets, and only recently have the tools for integrated networks started coming on line.

One of the key elements that make integration possible is the ability to allocate or redistribute vehicles to the available facilities in a near optimal manner. There are many situations where it might be possible to productively shift some of the traffic from the freeway to the arterials in a way that improves overall traffic flow. In other situations, some turning movements at critical locations could be prohibited temporarily, with vehicles re-accommodated. One such situation is the occurrence of a traffic incident. During an incident, the available capacity can be used more effectively by blocking certain turns while the incident is resolved. In addition, improvements in system performance could be obtained if the system operator (i.e., the traffic management center, or TMC) can direct the drivers to particular paths or facilities through the provision of real-time route guidance information via variable message signs or possibly via onboard ATIS.

There are several tools available for the analysis of freeways and city streets. Among these, CORSIM is the main network traffic tool supported by the Federal Highway Administration (FHWA); however, this program does not consider routes that vehicles may follow and thus does not provide the possibility for redistributing the flow at any major strategic point. Another analysis tool, dubbed DYNASMART, was specifically conceived for the integration of Advanced Traffic Management Systems/Advanced Traveler Information Systems (ATMS/ATIS). Developed at The University of Texas at Austin as part of the FHWA ITS research and development program, DYNASMART applies the necessary routing decisions at the core of the program. It also allows representation of a full corridor network of freeways and intersections. In addition to simulating traffic movement and driver decisions, it allows assignment of vehicles to different paths according to a variety of possible rules and objectives.

On the other hand, a network-level tool like DYNASMART that includes all these features cannot provide the kind of fine-grained local detail desired for specifically analyzing or developing control strategies for diamond intersections (i.e., network-level analysis necessarily requires that details of diamond intersection traffic be simplified). Accordingly, this study uses the TEXAS model, which is the premier microscopic tool available for analyzing isolated intersections and intersection configurations in the form of diamond interchanges. This project thus combines the capabilities of the TEXAS model, which is a responsive tool for representing the diamond intersection geometry, with DYNASMART, which is appropriate for integrating all the various subsystems at the network level.

2.2. CONCEPTUAL INTEGRATED CONTROL SYSTEM DESIGN

Strategic Framework

As shown in Table 2.1, the main subsystems for control integration include freeway mainlanes, ramps and frontage roads, diamond interchanges, and arterial streets. These are the primary controlled facilities involved in the conceptual design of an integrated freeway corridor control strategy. In this portion of the study, each facility is considered a separate subsystem having different types of control. The key parts for integration purposes are also presented. Subsequent chapters will consider the integration of all the subsystems.

Table 2.1: Main elements in control integration

FREEWAY	ARTERIAL STREETS	STREET NETWORK / INTERSECTIONS
Ramp metering	Signal coordination	Signal control (diamond coordination)
Variable message signs	Lane control: reversible, movement assignment (lanes, turns, etc.)	Transit operations / bus preemption, reverse lanes, etc.
Lane control (on/off or closure)	Variable message signs	Lane control
HOV	HOV	Parking control
	Parking control	

Strategies for Intersection Control

Strategies for the control of individual intersections include those described as *pretimed* and those described as *actuated.* The primary interest is in situations where the signal control can be responsive. Some of the strategies available include:

1. Traffic responsive with no prediction (sense and react to the presence of vehicles),
2. Traffic responsive with prediction, (e.g., UTCS-3 strategies [Ref 7]), and

3. “Real-time adaptive” (test generation of intelligent controllers) that combines prediction with sophisticated optimization algorithms (e.g., RT-TRACS).

Strategies for Diamond System Control

As mentioned before, diamond intersections represent the critical connections between system elements. They are potential bottlenecks that can also involve possible jurisdictional and coordination issues as the arterial system connects with the freeway. Signal coordination within the diamond (as well as within the connecting arteries), movement prohibition, and variable message signs are examples of diamond-related control strategies.

Strategies for Arterial Street Control

The PASSER program (Ref 45), used for signal timing along arterial streets, offers coordinated operation through offsets intended for bandwidth maximization. PASSER is a tool for pretimed optimization coordination along an arterial. However, like all pretimed approaches, it does not take into account changes in flow patterns or magnitude as a result of unexpected events. One way to apply this capability in real time is to resolve repeatedly and at regular intervals the program’s output with new input flows from current field measurements. The rolling horizon approach provides a suitable framework for repeated responsive application of this capability.

The integrated management of the arterials within the corridor network calls for modifying the bandwidth in a traffic-responsive manner that depends on the anticipated flows. Flows can be anticipated by explicitly considering a routing element in the strategy, since the flows depend on the routing of motorists through the network — a particularly important element in achieving integrated operation and exploiting its benefits, especially under incident conditions. Consequently, the integrated scheme includes the provision of bandwidth that changes according to the incoming traffic associated with integrated corridor operation. Thus, in addition to responsiveness introduced through quasi-real-time application, the key to integrated operation is to explicitly recognize traffic routing in the corridor.

Strategies for Network Control

In terms of a network of signalized intersections, offsets among clusters of neighboring intersections can be determined, as is proposed under common signal network operation software, such as SCOOT and SCAT (Refs 8, 9). However, a potentially powerful strategy that remains to be further exploited is one that identifies and coordinates major movements in the network. In this sense, the main strategy is the coordination of signals to accommodate major vehicular streams with minimum delay. Implementation of such a strategy would require explicit representation of paths, allowing major movements to be identified and the associated arterial streets to be optimized.

Strategies for Freeway Control

Ramp metering is one of the main tools available for freeway control. The basic strategies for ramp metering include those described as either *pretimed* or *traffic responsive.* In both cases, there have been many recent developments (Ref 36). Some assume a global perspective, in which an optimal entry rate at different ramps is calculated taking the entire freeway into account. Others may be entirely local, where only the immediate section (where the on-ramp is located) is considered. The most interesting ones tend to be hybrid in nature, whereby a global perspective is adopted while a locally determined responsiveness is also incorporated in order to handle deviations from the global plan (Ref 20). It has been found (Ref 18) that strategies that take into account traffic conditions in the local neighborhood tend to work better than the more global ones. ALINEA is a particular technique that has been applied successfully in different freeway environments, mostly in Europe (Ref 17). It was developed originally in Germany, tested extensively in the French freeway system, and modified to make it more of a hybrid type of strategy. It is a feedback-control-type law that senses occupancies and determines metering rates according to the sensed occupancies. It has been modified to consider, first, occupancies upstream and downstream of the on-ramp section, and then to determine the rates accordingly. This modification has yielded good results both in a simulation environment and in the field (Refs 17, 18).

Another freeway control strategy, intended for mainline control, consists of providing speed advisory measures. This strategy has shown promise in several locations in Europe, primarily in Germany and neighboring countries. It is most often applied on the freeway mainlanes to prevent the freeway flow from "breaking down" and sliding into unstable, congested conditions. This is accomplished primarily with advisory dynamic speed limit signs intended to slow drivers and to regularize speed among the drivers, thereby preventing or delaying the breakdown. This strategy appears to be robust vis-`a-vis driver compliance, with the benefits obtained even when only a fraction of the vehicles respond to the messages. While not particularly used in the U.S., it is included in this discussion because such strategies become increasingly possible with the kind of ITS technologies being deployed.

The Use of Variable Message Signs (VMS) as an Integration Strategy

Variable message signs (VMS) could become a key element in efforts to introduce integrated control in a corridor under near-term scenarios of ITS deployment. The strategy here is to view routing as a control method, and the distribution of flows in the network as a mechanism available to traffic managers seeking to optimize the performance of the system. Thus, by integrating the routing and the traffic control, it would be possible to reduce congestion and improve capacity. With the ATMS strategies available today, VMS could be placed at strategic decision points on freeway access roads — that is, where drivers must decide whether to get onto the freeway. More signs could be placed along the arterial streets where additional indications and/or messages could be presented to the drivers. One key aspect of the success of VMS is the ability to update the message rapidly in order to distribute flows. Eventually, when vehicles are equipped to receive onboard information, route guidance information can be distributed more efficiently. Initially, VMS will serve as a primary tool for displaying traffic and route information, and for sending messages jointly at different points to yield the desired spatial distribution of flows over the network.

HOV Integration

High-occupancy vehicles (HOVs) represent one approach to obtaining higher person-carrying capacity from a roadway traffic lane. However, it is not clear how we can influence that choice on a real-time basis as part of an integrated corridor control framework. Inducing greater HOV lane usage is more the target of planning activities aimed at travel demand management. Eventually, this framework might incorporate HOV choice and integration, but it will not be the first element to become operational in this framework.

With regard to HOV operation, the main control decision is whether the HOV status of a lane should remain as such under an incident situation. Another control decision is related to special events on the facilities within a freeway corridor. In these cases, the HOV restrictions could be relaxed. If there are strict HOV requirements, for instance 4+ passengers, we might compromise this level to 3+ or 2+ if we need to accommodate those vehicles.

2.3. FREEWAY RAMP AND MAINLANE CONTROL SYSTEMS

Freeway Ramp Control Operation: General Characteristics

Ramp metering represents the main strategy in freeway control. Its purpose is to maintain uninterrupted, noncongested flow on the freeway, which means that the freeway operates in a stable mode for as long as possible, avoiding or delaying the formation of bottlenecks created by excess demand (with respect to capacity). One way that ramp metering seeks to smooth operations is by preventing the traffic density from entering an unstable mode, thereby maintaining high traffic flow capability at all times.

Procedures developed for freeway traffic control through ramp metering fall into one of two basic schemes: fixed control or traffic responsive. Under fixed control ramp metering, changes in metering rates on the ramps are determined ahead of time according to a predetermined plan. Traffic-responsive metering allows more flexibility, given that the control system is predicated on measurements of prevailing traffic conditions. One form of traffic-responsive metering known as merge control (because it consists of releasing vehicles from the controlled ramps based on the detection of acceptable gaps on the freeway lane adjacent to the entrance ramp), has been largely abandoned because it is not effective as a tool for traffic congestion control. The main form of traffic-responsive metering is not concerned with microscopic details of spacing between consecutive vehicles. Instead, it considers average conditions on the freeway according to the measurements of one or more traffic flow variables, such as occupancy, speed, and/or volume. These variables are considered on a real-time basis and then used to determine the metering rates necessary to control the entrance ramps in the vicinity of the measured section. The discussion of traffic-responsive ramp metering strategies will be focused on those of the second type.

Fixed-Control Ramp Metering

Fixed-control ramp metering is intended for repeatable, steady-state traffic conditions. Formulations have been developed to compute the optimum metering rates that maximize the total output rate for the facility subject to constraints on the capacity of each section, to the input ramp rates not exceeding the available demand for any section, and to non-negativity. Early contributions include Watleworth's use of linear programming techniques to solve the above ramp metering problem, under assumptions of steady state conditions, constant O-D pattern for a single freeway over the time period of analysis, and traffic diverted from one on-ramp not entering onto other on-ramps (Ref 10). The solution consists of the respective ramp metering rates (assumed fixed over the analysis period), including possible closure of certain on-ramps (corresponding to a zero input ramp rate). Some enhancements to Watleworth's formulation include (1) maximum queue constraints in order to limit the number of vehicles that divert, (2) spreading excess demand equally over all ramps, and (3) merging capacity constraints.

Several researchers have extended Watleworth's model by modifying or adding constraints and incorporating new performance measures into the objective function. Some of the modifications include the use of an objective function that maximizes total excess capacity (Ref 10), the use of vehicle miles of travel on the freeway instead of the total output rate (Ref 11), the addition of some characteristics of vehicle diversion (Ref 12), and consideration of surface streets as part of the diversion routes coupled with a user behavior model (Ref 13). All of these studies assumed steady-state conditions. This assumption, while essential for simplifying the models, is not adequate for highly variable peak-period conditions.

Other efforts have attempted to take into account the propagation of traffic by including the time delay between a volume change at a ramp and its subsequent effect at a point downstream (Ref 14). This representation better captures the temporal dimension but increases the number of decision variables dramatically, thereby making the optimization process highly cumbersome. Other efforts have addressed the computation of ramp metering rates for specific situations, like provision of information on a limited basis (Ref 15) and consideration of "balanced" queue lengths at all entrances to maximize the total demand processed (Ref 16).

The advantage of fixed-control ramp metering is that it is relatively simple to implement. On the other hand, the main disadvantage is that if the demand pattern assumed in computing the fixed-time program differs from the actual one, ramp metering is less effective. Since demand patterns could change for a variety of reasons (e.g., day-to-day fluctuations, user response to the effects of the control strategy, or accidents and other unexpected disturbances), fixed-control ramp metering tends to be ineffective under such conditions. Traffic-responsive ramp metering is intended to provide better performance under such conditions.

Traffic-Responsive Ramp Metering

Ramp metering is used as a way of regularizing the flow on the freeway. One of the concerns with some of the steady-state formulations is that they do not account for the dynamics of freeway flow (e.g., ramp metering does not account for the phase transition phenomena, wherein traffic shifts from a stable to an unstable flow).

Traffic-responsive metering seeks to respond to conditions as they are unfolding, instead of following pretimed metering under conditions that may no longer be representative of prevailing actual conditions. Ramp metering rates are therefore set so as to take into consideration the current mainlane volume and ramp conditions.

Methods of traffic-responsive ramp metering fall into the following classes: demand-capacity control, occupancy control, and gap-acceptance control. Demand-capacity control responds to the volume upstream, taking into account the capacity downstream. The on-ramp rate is set so as to prevent capacity from being exceeded downstream, given the upstream volume. The capacity is determined by sensing the upstream volume; the number of vehicles allowed to enter onto the freeway should not cause the downstream section to exceed its nominal capacity. The drawback in this situation is the unclear definition of the downstream capacity value. The local capacity could be obtained based on previous measurements. But there are fluctuations around the average value that are not captured under real-time conditions.

One situation of interest occurs when the remaining capacity is zero. In this case, some vehicles should still be allowed to enter, or else the ramp must be closed to avoid spill-back problems. Therefore, if the residual capacity downstream is zero, the ramp rate is set to its minimum value. One typical suggested value is 3 veh./min-ln. In addition to the upstream volume, consider also the regime under which the facility is operating. If it is already operating in the congested regime, no more vehicles should be allowed, so the operation is set at the minimum rate. In other cases, as many vehicles as can be handled efficiently should be let in. Another important consideration is that measured values approaching capacity normally imply greater fluctuations and less stability of the system. Therefore, instead of setting the capacity value to its maximum, a margin of safety should be allowed by setting the capacity to some reasonable value that provides some protection against falling into the unstable regime. On the other hand, when the on-ramp demand is also high, the entry queues will become very long on the ramp. The trade-off almost always works in favor of not overloading the freeway because once the flow breaks down, then the loss of service rate renders the queues even worse. Therefore, the vehicles should be held back to maintain the service rate in the facility (i.e., the overall throughput) as high as possible.

The prevailing freeway regime can be determined based on occupancy measurements, particularly those obtained from loop detectors. As an example, the threshold used in California is an occupancy of 18%. This means that if the occupancy is greater than 18%, the ramp is set to operate at the minimum rate. This is an empirical rule, based on a locally reported value that therefore needs to be calibrated in each particular environment. Consequently, each area should experiment and decide what value may be used to define the congested regime for which vehicles enter the freeway at the minimum rate.

Rules based on occupancy thresholds correspond to the second strategy, the so-called occupancy control. In this case, the metering rate is determined based on real-time measurements of lane occupancy taken upstream of the entrance ramp. Typically these rates are presented in an occupancy-volume graphic. In this way, the values of the ramp rates in veh./min. are defined according to the measurement of occupancy.

The third technique, gap acceptance control, was an idea tried in the late sixties and early seventies. The main idea of this type of ramp metering is to let vehicles enter only if they can merge with the oncoming traffic without interfering with it. Therefore, it is more an aid to the merging drivers as they enter the stream, rather than a congestion or flow management tool. In addition, if conditions are congested, then there will not be a sufficient number of gaps, and consequently some degree of friction will take place among the vehicles to accommodate high on-ramp demands.

One problem with all of these responsive strategies is that when the conditions are congested, the pressure of letting vehicles in at rates greater than the minimum rate is very high, thereby posing an enforcement problem. Effectiveness is naturally reduced if enforcement is not ideal. Nonetheless, responsive ramp metering has been shown to be effective under many different conditions.

Special Strategy: ALINEA

Papageourgiou and collaborators (1991) have proposed a rather simple linear feedback type of model called ALINEA (Asservissement LINeaire d'Entrée Autoroutiere), or Lineal Control of Freeway Entrance (Ref 17). This is a closed-loop control law, one derived by linearization of the nonlinear freeway traffic model equations around an optimal steady-state condition. The feedback rule takes into account the deviations from nominal conditions to calculate the change in ramp rate. The basic equation is given by:

$$r(k) = r(k-1) + K_R\,[\hat{o} - o_{out}(k)]$$

where the terms are defined below.

The entrance ramp rate *r(k)* for time interval *k* (approximately 1 minute) is calculated based on the on-ramp rate *r(k-1)* of the previous time interval *(k-1)* plus a term that takes into account the effect of variation in occupancy downstream of the entrance. To calculate this effect, the difference between the desired or nominal occupancy, $\hat{o}$, minus the current occupancy o_{out} is obtained based on measurements taken typically 40 meters downstream of the entrance. The difference is multiplied by a parameter K_R that needs to be calibrated. One advantage is that the performance of the rule is insensitive to the value of K_R over a certain range of values, which makes the rule rather stable. However, as the K_R increases, the reaction becomes stronger; in other words, the higher K_R, the greater the reaction to changes or deviations in occupancy from the nominal. In fact, for K_R values that are very large, oscillations and chaotic behavior may take place because the entrance rate moves back and forth rather abruptly and can create unstable conditions that may cause even worse traffic conditions.

The main purpose of this metering rule is to stabilize traffic flow at a high throughput level without underloading the freeway. To maintain its operation at a high traffic level, we need to let vehicles in. Another way to stabilize traffic flow would be by preventing the entrance of vehicles at lower capacity values; however, that would not be desirable because the freeway would not be as efficient as it should be. Therefore, stabilization should be done at a high throughput without overloading the freeway. By fine tuning these parameters, the strategy has been shown to work quite well. Experiments and real-world applications have shown the advantage of using this strategy over some other strategies used for traffic-responsive ramp metering (Ref 18).

Online Control: Integrated Strategies

Certain ramp metering strategies consider both the current section and additional segments. The logic is to avoid overloading not only the current segment, but also downstream sections as well, recognizing both spatial and temporal interactions taking place in the traffic system.

Several approaches to developing strategies for integrated ramp control have been proposed. These typically include a mathematical formulation with a specific objective function to optimize (subject to various constraints on acceptable control values) and to capture traffic flow dynamics (Refs 19, 20). From that formulation, a set of optimal control rules may be derived. Typically, these rules are applied in real time.

The more interesting situations where this approach is likely to be of particular benefit are those in which there is nonrecurrent congestion. For recurrent and predictable conditions, the steady-state problem can be solved, or a time-dependent problem with historical data can be considered. On the other hand, for less predictable nonrecurrent congestion, the dynamic changes are very important and downstream bottlenecks should be considered in a real-time fashion.

Strategies that seek on one hand to meet the anticipated demand while at the same time to be responsive to nonrecurrent conditions tend to have structures comprised of several layers of control. At a minimum, the so-called multilayer or multilevel control logic consists of the following two layers: (1) a higher-level layer, with relatively "long" intervals between updates (somewhere between 10-15 min), which provides the overall control strategy for predicted conditions; and (2) one or more lower-level layers that can update the controls at much shorter intervals and operate on a decentralized basis (e.g., at the level of the individual ramps). Meanwhile, the higher-level layer adopts a centralized control perspective, taking the whole facility and/or system into consideration. The primary function of the lower layer is to detect fluctuations that significantly deviate from the anticipated nominal state, and to respond to them through appropriate, locally determined adjustments in the control. For instance, if a major surge in occupancy is sensed, the lower-level controller could respond by letting fewer vehicles in, even though the higher-level layer may have established a different rate for the predicted conditions. As local fluctuations are detected, the rate is adjusted locally according to the measured state of the system. Additionally, if an incident is detected, the rate may be dropped down to the minimum value or the ramp may be shut-off completely if necessary. In this sense, the lower-level control is reactive in nature and will remain in force until a new "global" plan is generated by the higher levels for updated future demand predictions. Thus, the lower layer deals with fluctuations and responds to them by deviating from the plan established at the higher level.

One of the earlier formulations of this type is credited to Isaksen and Payne, who adapted concepts from control theory and applied them to traffic control (Ref 21). Their approach led to many additional developments in the real-time control of freeways. Several enhancements of that original formulation have been proposed. In this discussion, the particular formulation developed by Papageorgiou (Ref 20) is considered.

The multilayer approach to corridor control takes into account the fact that some parameters of the system vary slowly, while others vary more rapidly. For instance, the demands $d_i(t)$ on the facility (through a given ramp i at time t) and the associated origin-destination (O-D) matrix (expressed as a set of parameters α_{ij}, which denotes the fraction of vehicles entering at ramp i that pass through segment j) tend to be slow-varying disturbances in the sense that under normal conditions, they will be fairly predictable (over the short to medium term). In addition to these slow-varying conditions, there will be unexpected disturbances, such as incidents of various types that are essentially unpredictable within a time horizon of 10 to 15 minutes. In this case the system relies on detection schemes to infer the presence of these disturbances through their effects on state variables, such as concentration (or occupancy) or speed. One indication of varying conditions is when the actually measured concentration (or occupancy) in the segment increases above that predicted in the model. The presence of an incident may be detected through the use of a good performance model of the system operating under these slowly varying disturbances, thereby providing a reference level against which to evaluate deviations of the measured values for concentration (or occupancy) in a certain segment. The solution is presented in layers, such that a higher layer is used to optimize for the slowly varying disturbances that have a greater degree of predictability, while a lower layer is implemented to respond quickly to the conditions sensed in the system.

The control scheme proposed by Papageorgiou (Ref 20) is a three-layer approach. Such a scheme is not necessarily optimal under all possible conditions for these problems. In other words, a better solution could in many cases be found *if complete future information* were available. If what has happened is known after the fact (hindsight), then it is possible to optimize control. However, the control must be determined *a priori* — that is, in advance of what actually occurs. Therefore, its performance should be evaluated considering the limited information available and the quality of the prediction procedures. All these schemes and approaches have to be viewed as heuristic approaches: suboptimal but robust. Therefore, the trade off in performance is in terms of the quality of the objective function versus the robustness of the system.

The three layers in Papageorgiou's framework are optimization, direct control, and adaptation layer. The optimization layer solves a simplified (steady-state) optimization problem for the overall process or system. Given predicted demands, this first layer gives an optimal plan, one that includes the nominal or average value for the state variables that are controlled. This is essentially a centralized function operating on a time frame of 10 to 15 minutes. Below it is the direct control layer, which performs a decentralized function that responds to fluctuations and deviations to maintain the system at the desired nominal levels of the state variables. This means that the system is kept in the vicinity of the optimal control point obtained from the initial layer. However, there may be in the system some changes in the process itself in terms of the supply or the demand, with the result being changes in some of the parameters of the relations. It is desirable that the system be able to learn from observation. This is the role of the third layer, called the adaptation layer, which provides a learning function that serves to update the prediction of process conditions, particularly the demand and the capacity. To summarize, such a system has (1) an optimization layer at the central level that communicates with local controllers, (2) a second layer that then provides local direct control, and finally, between these, (3) an adaptation layer that receives information from the local units in order to learn and provide feedback to the optimization layer.

Several additions and other modifications to this basic scheme have been proposed. For example, one concern is what might happen to the system's operation during a breakdown of one of the local controllers or communication links. One way to introduce fault tolerance is to use a cascading design (Ref 19), where the local controllers are interconnected in case of malfunction. Then, when a local controller breaks down, another can assume its functions. This strategy is termed *cascading* because there are multiple levels in the local controllers to allow fault-tolerance redundancy in the system in case it is needed.

The above scheme, developed mostly for freeway facilities, can be readily generalized to encompass freeway corridors as integrated systems, including arterial and surface streets in the same general freeway corridor, or multiple freeway systems, and eventually at the entire network level. At the corridor level, for instance, the logic would be very similar, except that there are many more control variables. Instead of setting only ramp meters and speed limits it is also possible to control the traffic lights and provide coordination among traffic lights, cycle times, phases, phase sequence, and phase duration. Furthermore, it is possible to influence movement through variable message signs or (eventually) via in-vehicle information systems. Using this type of framework, it is possible to manage the whole corridor; but because there are different systems, in addition to the central controller that is providing traffic management functions for the whole system, subcontrollers are typically present to run separate arterials, diamond intersections, the freeway system, etc. It is possible to have several layers below this, until the local controllers are reached. Thus there will be a local controller in the ramp connected with other subsystems but essentially receiving instructions from the integrator on top.

Mainlane Control Techniques

In the case of mainlanes, instead of controlling the entrances to the freeway, the strategies seek to influence traffic already in the freeway's mainlanes. While this approach has not been used much in the U.S., it has gained acceptance in certain European countries, particularly Germany. Specifically, the control seeks to maintain the level of flow and the quality of service above that associated with unstable stop-and-go conditions. Through judicious use and timing of speed regulation using variable message signs, coupled with precise sampling of various traffic parameters (especially speeds), this control approach can sustain traffic flow at levels higher than what would otherwise be possible without the control.

The primary objective of mainlane control is to maintain stable flow and, hence, delay the phase transition from stable to unstable regimes. The unstable regime is typically characterized by stop-and-go traffic or stop-start waves. In the modeling of freeway flow (Ref 22), it has been observed that when traffic approaches the point or threshold at which phase transition occurs (between stable and unstable regimes), the speed distribution widens. The conclusion is that this behavior of the speed distribution provides a good early warning indicator, one that is more effective than concentration or occupancy. Specifically, if we take the standard deviation of the speed over vehicles that we are measuring over a certain period of time, we observe that, as the density increases and approaches the phase transition point, the standard deviation also increases and shoots up asymptotically at a certain value that constitutes the effective cut off between the two regimes. One factor that results in the widening of the speed distribution around the phase transition situation is the coexistence of very fast vehicles and very slow vehicles that are beginning to slow down. The resulting increase in the standard deviation provides a good early predictor of impending phase transition.

The idea is to find some reasonable value of the standard deviation of speed that provides a good warning of the onset of the phase transition, and then attempt to prevent it. One way to lower the value of the standard deviation is to regularize speed by displaying a speed limitation sign. That is the rationale for this type of mainlane control. As the standard deviation approaches a certain threshold (i.e., one that needs to be determined on a facility-specific basis), VMS could be used to indicate a different speed limit in an attempt to homogenize the prevailing speeds and to prevent the occurrence of the critical extreme values. If we can slow down some of the approaching vehicles, the variance would be reduced. In the previously mentioned study, the author, after testing the predictor against many data sets, determined the standard deviation threshold to be approximately 17 km/hr for conditions on German highways.

Fixed Location and Movable Message Sign Systems

Variable message signs, or changeable message signs (VMS or CMS), are used to provide real-time information to motorists (Ref 29). Every corridor having ITS capabilities has a

VMS network installed at strategic corridor locations. These networks serve several purposes: For example, they are used to alert motorists to traffic problems ahead, to indicate the location of an incident, or to offer an advisory message. In some cities where these systems are already deployed, they are used to indicate the availability of additional information sources (e.g., radio stations), where more detailed incident information is provided. These VMS systems are also used to provide information during adverse weather conditions or during construction activities.

When an incident occurs, VMS systems are used to inform drivers of varying traffic conditions. Multiple message signs can indicate alternate routes that will allow drivers to avoid excessive delays. In addition, lane control signals (LCS) can indicate what lanes are affected by an incident and can then guide motorists toward available, unaffected lanes.

Such message sign systems have already been implemented worldwide. In the U.S., the message sign benefits that have been observed in many cities are currently being quantified. Many European cities have also implemented various types of variable message signs, including those relating to multimessage posting, lane closure, and variable speed limits. Studies have confirmed that sign systems significantly reduce primary and secondary accidents and, moreover, can increase throughput during peak periods (Ref 29).

In a typical setting, the control center receives information about the road and traffic conditions through automatic sensing detectors and closed-circuit television cameras (CCTV).. When a special situation is detected (from the information that is processed and evaluated), the VMS is activated and relevant information is displayed to drivers.

Even within a well-established VMS network, it is useful to also have portable variable message signs during special events or to better manage traffic when unpredictable major incidents occur. Typically, portable VMS systems are deployed to inform motorists of special problems on the roadway. The displaying capabilities of portable VMS systems are similar to those of a permanent VMS, with the added advantage of their mobility. Using a portable VMS, additional information can be provided to the users about roadwork or incident-related congestion.

Even though recent developments associated with in-vehicle information systems improve information accessibility to users that can afford to buy the equipment, VMS systems are still important for providing focused information to every motorist using a specific facility.

2.4. FRONTAGE ROAD AND DIAMOND INTERCHANGE CONTROL SYSTEMS

Frontage Road: General Characteristics

As a part of the freeway facility, the frontage road plays an important role in managing traffic. The frontage road facilitates the exchange of freeway and arterial street traffic through on-ramps and off-ramps.

When congestion arises, vehicles may be diverted to the frontage road. In the event of ramp closures, the vehicles approaching an entrance can continue along the frontage road up to the next available entrance. When the volume of traffic to be diverted is high, additional diversion routes may be used, increasing the turning movements at diamond intersections. In all such situations, efforts to manage traffic flow along frontage roads under congested conditions are similar to those used on arterial streets. On-ramp closures typically force drivers to use the frontage road as an alternative route, one that must be used until they reach another point at which they can then enter the freeway to continue a trip (Ref 23).

The diverted traffic needs to be treated in an integrated manner. Depending on the cause of congestion, on-ramp closures may sometimes be planned (as for typical peak period congestion); however, under nonrecurrent congestion, all these measures need to be deployed in accordance with the situation at hand. Consequently, integration between the two facilities (freeway and frontage road) is necessary. The integration mainly consists in changing the control settings of the signal along the frontage road in response to new conditions; at the same time, consideration must be given to the additional intersection traffic arriving from nearby arterial streets and freeway off-ramps.

The main intersection points along the frontage road are special types of intersections called *diamond interchanges.* In order to obtain optimum mobility in a network, special consideration must be accorded to these intersections, which are discussed below.

Diamond Intersections: General Characteristics

Diamond interchanges can assume a variety of geometric configurations that range from the conventional to the specially designed. Conventional diamond interchanges are formed by a one-way diagonal ramp in each quadrant connecting adjacent legs of two grade-separated intersecting roadways. The design is most suitable in suburban and urban locations, where traffic volumes are usually moderate and where right-of-way is restricted. While diamond interchanges may be designed with or without frontage roads, those that include frontage roads are more common in urban areas, particularly in Texas.

Specially designed interchanges may be two-level or three-level configurations. The three-level diamond has a third-level structure formed by four pairs of ramps that handle turning movements and provide uninterrupted traffic flow on intersecting highways. The signalized, conventional diamond interchange characterized by a pair of one-way frontage roads on both sides of the freeway is the simplest and most common type of interchange used in Texas; accordingly, it is the main interchange configuration considered in this research.

Figure 2.1 illustrates the typical geometric arrangement of diamond interchanges. As shown, external approaches refer to lanes carrying traffic movement approaching the upstream, or first-encountered, intersection. Internal approaches refer to traffic movement on lanes between the two intersections of the interchange.

Traffic Patterns: The conventional diamond interchange consists basically of two separate intersections. Yet, because of its close spacing, a diamond operates much differently than two isolated intersections (Ref 24). Many studies have shown that turning movement volumes and percentages at diamond interchanges are often two to four times greater than that associated with isolated signalized intersections (Ref 25). The turning patterns are also likely to change during a typical day. For example, heavy left turns during the morning peak at one of the intersections frequently result in heavy right turns during the afternoon peak at the other intersection. Traffic queuing, another special consideration at diamond interchanges, is the result of the short spacing between intersections. Signal timing plans that use phases that tend to fill the interior lanes are limited by the distance between the intersections that controls the number of cars that can be queued.

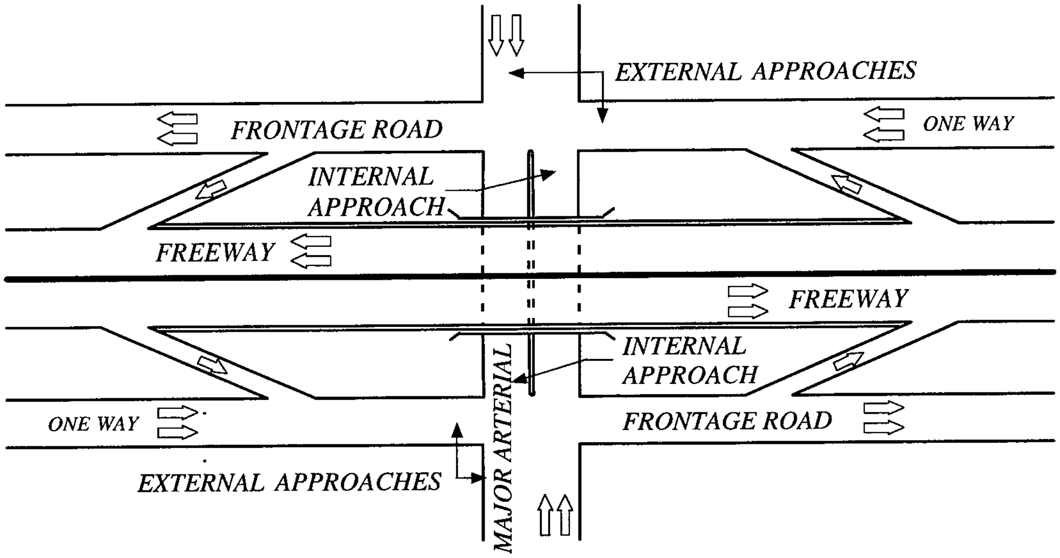


Figure 2.1: Diamond interchange, conventional arrangement

Signal Timing: Determination of the optimal phase timing for a diamond interchange differs little from that of a single signalized intersection. Efficiency requires that diamond interchange cycle lengths be kept low in order to minimize delay, yet there must be some coordination between ramp intersections for high traffic volume movements (Ref 26). Only small overlaps of normally conflicting phases occur and are able to take advantage of the space provided by the internal lanes. An interchange controller may either be pretimed or actuated.

Pretimed controllers are best used where a limited number of traffic patterns are found that repeat on a daily basis. These signals can be easily interconnected with adjacent intersections. The basic phasing can be modified through changes in the split, and offset if both intersections of the interchange have pretimed control. It should be noted that the Texas Department of Transportation no longer purchases pretimed controllers. Instead, it purchases actuated controllers and uses them as pretimed controllers when warranted.

Actuated controllers are appropriate where there are many traffic patterns that vary significantly on a daily basis. Unlike pretimed control, the time of each cycle length depends on short-term traffic volumes, number of signal phases in a cycle, and traffic controller settings of initial green, gap extension, and maximum green time for each phase. These controllers are not easily interconnected with adjacent traffic signals. Consequently, actuated controllers are primarily used at isolated diamond interchanges.

Phasing Sequence: The two intersections of a diamond interchange can either be timed separately to minimize intersection delay (three-phase control) or timed together to maximize interchange progression and, thus, minimize vehicle queuing between the two intersections (four-phase control). Neither method is universally better than the other is and each will result in different optimal cycle lengths. These control strategies may be combined to provide optimum phasing for changing traffic conditions. The Arlington phasing scheme and the Texas Diamond Controller are two examples of such combined strategies.

Three-phase control generally operates best on wide interchanges where a high percentage of the total traffic flow is represented by through movements. This control strategy generally requires the short cycle lengths and wider interior spacing that can permit greater phase flexibility and smoother traffic flow through the interchange.

There are basically two types of three-phase sequencing. The basic sequence begins with both frontage roads receiving green, followed by the two through-movement phases from the arterial street. Alternatively, this sequence may be arranged to accommodate the simultaneous movement of the arterial approaches and two phases for the separate frontage road approaches. With this basic sequence the two simultaneous approach movements receive the same amount of green time. Such phasing is satisfactory when these approach volumes are approximately the same (Ref 27). Phasing splits may be used to make the sequence more responsive to volume variations.

Another common three-phase sequencing technique moves both frontage road approaches during Phase 1 and moves both arterial street approaches during Phase 3. Phase 1 tends to fill the interior of the interchange with vehicles, as left-turning vehicles normally must stop at the downstream intersection to allow for the concurrent servicing of that frontage road approach. Accordingly, Phase 2 is used to flush those vehicles that queue in the internal lanes during the Phase 1 frontage road movements.

The typical three-phase diamond interchange control strategy used in Texas features simultaneous movement of the frontage road approaches and the possibility of separate movements for each direction of the arterial street (Ref 28). Phase 1 is the simultaneous movement of the frontage road approaches, while Phase 2 is the movement of the arterial street approaches. Left-turning movements from the arterial street may be allowed during this phase if traffic conditions are appropriate. Phase 3 consists of the arterial street's protected left-turning movements.

The most frequently used signal phase sequence is probably the four-phase-with-overlaps strategy. This variation of the four-phase control strategy separates the external interchange approaches into four individual phases. The increased number of phases often results in longer cycle lengths and, consequently, increased delays. Operational efficiency is gained, however, by overlapping the two frontage road phases with the two arterial street phases. This allows the frontage road approach phase to run concurrently with the arterial approach phase at the opposite intersection for the short time required for the arterial traffic to cross the interchange interior lanes before joining the frontage road traffic. This is typically implemented using dual-ring controller compatabilities (Ref 29).

In 1988 the City of Arlington, Texas, developed a diamond interchange control strategy that utilizes a phase selection process based on the detection of queue development (Ref 26). In this process, three-phase or four-phase sequences can be selected on a cycle-by-cycle basis, based on prevailing traffic conditions. The strategy also achieves increased operational efficiency by minimizing service to the interior left-turning movements. Protected left turns are provided only when there is substantial demand. Thus, this phasing scheme usually operates at the shortest cycle lengths needed. Yet, the strategy tends to have extremely long cycle lengths during periods of heavy traffic, since the queue detection system may have two or more phases simultaneously extended to the maximum.

The Texas Diamond Controller is a single, software-modified, eight-phase NEMA controller unit equipped with special internal programming to provide a combination of either four-phase or three phase operation (Ref 28). This control strategy maximizes the benefits of both phasing schemes. Variations in traffic demand may yield improved efficiency with one phase sequence during one part of the day, and another phase sequence during another part of the day. The Texas Diamond Controller allows for a change from the three-phase sequence to the four-phase-with-overlaps sequence by time of day.

Details About Phasing Sequence: Much work has been undertaken to determine the most efficient signalization for a common diamond interchange. This study, however, focuses on the signalization that can best benefit the entire freeway network under congested conditions. Accordingly, a pretimed signal control scheme was chosen for its coordination capabilities. The added capacity to freeway corridors achieved through signal coordination among interchanges along a frontage road could compensate for flow problems caused by an incident in the freeway mainlanes.

To achieve signal coordination, all interchanges along a frontage road normally have the same cycle length. The experiments in this study (described in Chapter 4) consider various cycle lengths. The cycle selected is such as allows the most phase flexibility for an interchange without adding excessive delay. To begin with the most common signal phasing, a four-phase plan with equal green time to each approach was assumed for the base case. This is also the phasing used by the network test bed. No offsets were utilized, as the simplest case was desired and since complex control schemes beyond the basic three-phase and four-phase plans have been shown to improve interchange efficiency only in small increments (Ref 30).

The basic three-phase control sequence begins with both frontage roads receiving the same green duration, followed by the cross street through and protected left-turn movements for one and then the other intersection (Figure 2.2). This phasing scheme is satisfactory in cases where the two frontage road volumes are nearly equal; modifications can be made when there is significant asymmetry between the traffic volumes along the frontage roads.

Two additional "three-phase" sequences can be used favoring the frontage road that experiences higher traffic volumes. Green times are then provided to the frontage roads in proportion to their traffic volumes. To favor the downstream frontage road, an additional phase for this frontage road is added to the three-phase scheme. If the basic three-phase scheme consists of phases A-B-C, then this sequence consists of phases A-A1-B-C (Figure 2.2). To favor the upstream frontage road, the technique is a little more complicated. The additional phase for the frontage road, A2, is added after phase A but is then followed with Phase C and then Phase B. Thus, the phase sequencing becomes A-A2-C-B (Figure 2.2). By reversing the order of Phases B and C, the traffic engineer can smooth the traffic flowing through the interchange; short left-turn movements within the interchange are also minimized (Ref 27). Three-phase operation can, however, be sensitive to limited internal approach left-turn storage capacity and to significant left-turning volume.

In four-phase control, the diamond interchange is treated as a single intersection having four approaches. The most common four-phase control uses two overlap phases having a duration equal to the travel time between the left and the right intersections (Figure 2.3). This strategy reduces the number of stops because all external approach traffic volumes progress through the interior of the interchange.

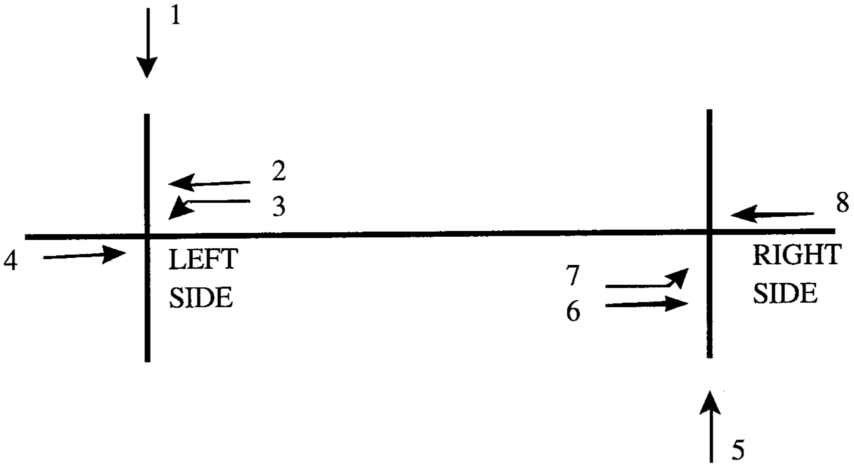
Phase Sequence Strategies for the Fort Worth I-35W Diamond Interchanges Corridor: For the base case, the diamond interchanges are operating under three-phase or four-phase control without overlap, as illustrated in Figures 2.3 and 2.4. For incident scenarios and where these phase sequences are not efficient (where the level of service, or LOS, in terms of delay is not satisfactory), a special three-phase control that prohibits left turns from the internal approaches is provided (Figure 2.4). If this phasing control does not yield acceptable congestion relief, then two phase-control with prohibition of all left turns is provided (Figure 2.4).

Frontage Road Coordination

Traffic Patterns: The frontage road typically serves as the interface between the street network and the freeway. Accordingly, its traffic pattern is a combination of the patterns that characterize associated diamond intersections and on/off ramps. The interactions become more complicated in the presence of freeway congestion, which can force the frontage road to serve as an arterial; under such situations, effective frontage road operation requires coordination to handle the additional users resulting from the congested conditions on the freeway.

2.5 CLOSURE

This chapter has reviewed the various components of a freeway corridor network, as well as the principal forms of traffic control normally applied to these components. In all such efforts, the goal is to provide integrated control for the efficient operation of the entire freeway corridor system. In discussing the role of each component in the system, this chapter also described how these components may be integrated with the other components. To develop integrated control and operation strategies that recognize the role of each component and the available forms of control, a modelling methodology is required to represent the traffic system and to evaluate its performance under different control strategies. This framework is described in the next chapter.



BASIC THREE PHASE	FAVOR LEFT SIDE	FAVOR RIGHT SIDE
↓ 1 ↑ 5	↓ 1 ↑ 5	↓ 1 ↑ 5
	↓ 1 7 ↗ 6 →	← 2 ↙ 3 ↑ 5
4 → 7 ↗ 6 →	4 → 7 ↗ 6 →	← 2 ↙ 3 8 ←
← 2 ↙ 3 8 ←	← 2 ↙ 3 8 ←	4 → 7 ↗ 6 →

Figure 2.2: Three types of three-phase signal timing

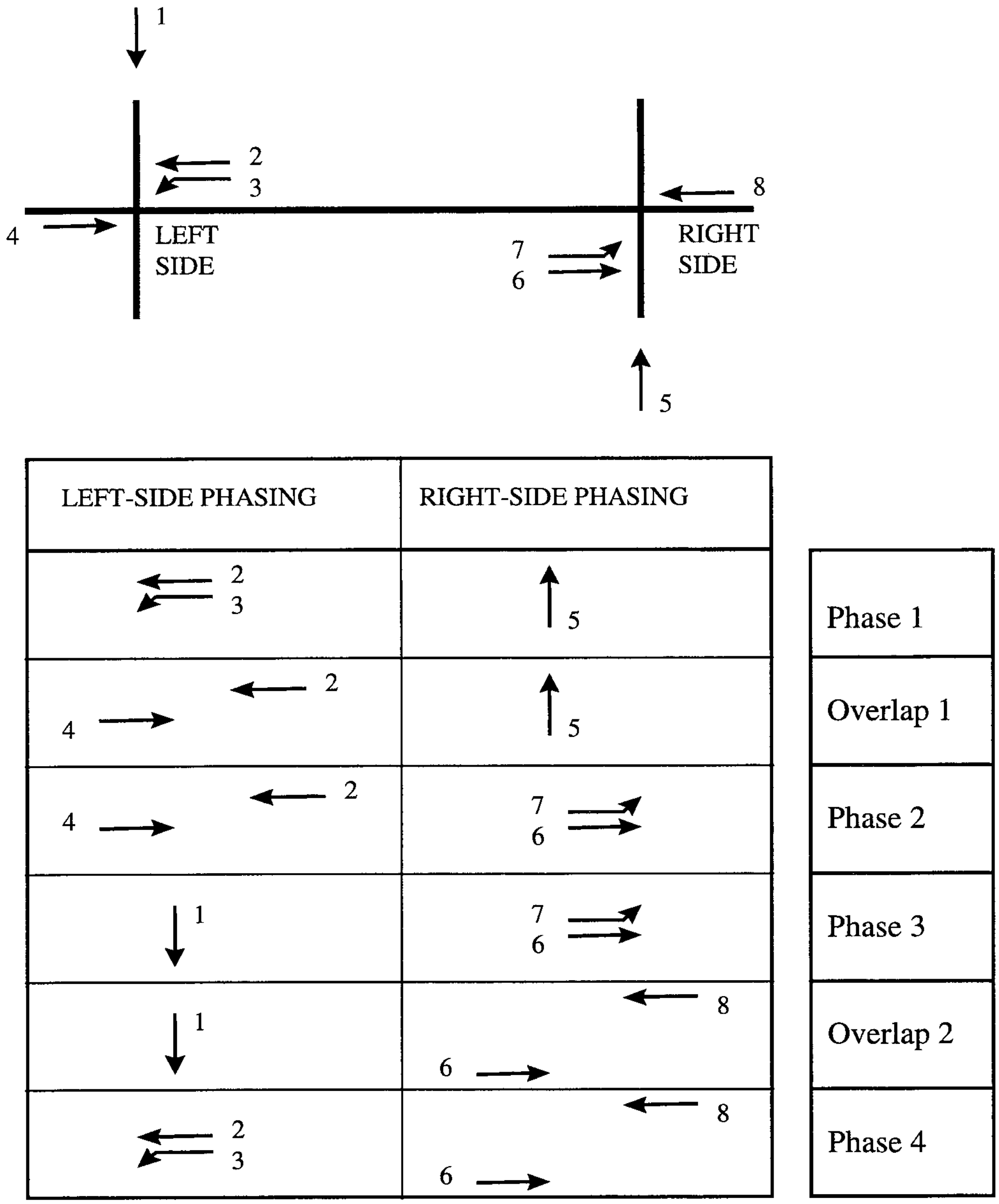
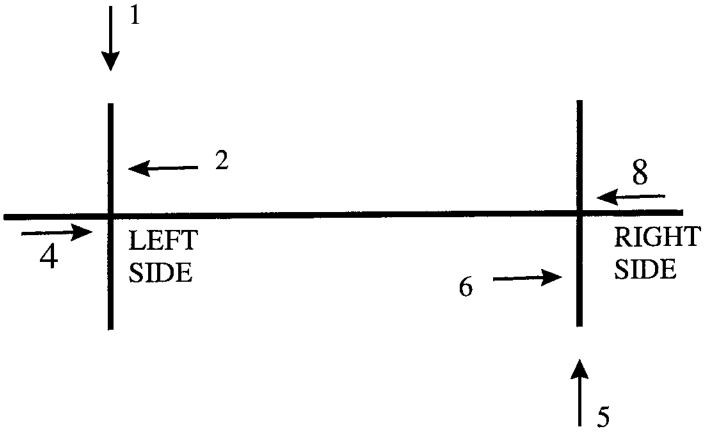


Figure 2.3: Four phases signal timing with overlaps



LEFT-SIDE PHASING	RIGHT-SIDE PHASING	3-phase control Prohibition of all left turns from the internal approaches
4 → ← 2	↑ 5	
4 → ← 2	6 → ← 8	
↓ 1	6 → ← 8	

LEFT-SIDE PHASING	RIGHT-SIDE PHASING	2-phase control Prohibition of all left turns
↓ 1 no left turns	↑ 5 no left turns	
4 → ← 2	6 → ← 8	

Figure 2.4: Three- and two-phase control with prohibition of left-turn movements

CHAPTER 3. METHODOLOGY TO TEST INTEGRATION STRATEGIES

3.1 INTRODUCTION

While field implementation and evaluation represent the ultimate tests of any traffic management strategy, the complexity of some of the interactions taking place among the various components of a freeway corridor network precludes direct field testing prior to careful analysis. Thus, computer simulation provides a most suitable tool for this type of development and analysis, especially in view of the large number of strategies to be tested and the variety of conditions under which the performance of these strategies must be evaluated. Consequently, the principal methodological approach followed in this study has been the use of computer simulation tools to evaluate control strategies developed off line — for both individual system components as well as for overall integrated operation — for on-line implementation and traffic-responsive operation. However, the practical usefulness of the simulation-based evaluation largely depends on the realism of the situations considered in the simulation experiments that form the basis of the evaluation. For this reason, an actual test bed was carefully selected and developed for the purpose of this study. Specifically, a portion of the Fort Worth network was delineated for this purpose, with input provided by TxDOT engineers on the appropriateness of the test bed and its representativeness of situations that are commonly encountered in other metropolitan areas in Texas and elsewhere.

In selecting and developing the simulation methodology for this study, it was realized early on that most traffic simulation and optimization tools used in practice have targeted one or two components of the overall system of interest, e.g., the freeway itself, or intersections, or arterials, but that no tool offered the desired integrated capabilities needed to address the challenges raised by the objectives of this particular study. The only exception is a simulation-assignment tool recently developed at The University of Texas at Austin, called DYNASMART, which is described in greater detail in the next section. DYNASMART was developed primarily for the U.S. Federal Highway Administration's Research and Development Program aimed at providing the next generation of tools and methodologies needed for ATMS/ATIS operation and deployment. As will be described next, this model offers the principal capabilities required to evaluate the kind of integrated control strategies for freeway corridor networks envisioned for this study.

However, any network-level simulation-assignment model must by necessity make certain trade-offs in terms of the level of detail in the representation of the traffic systems (e.g., geometrics and control features), as well as in the movement of vehicular traffic in the network. For this reason, component-specific models that consider all relevant operational details of a particular component of the network were also used in conjunction with DYNASMART to enhance its capabilities with regard to the performance of those components. In particular, given the key role of diamond interchanges, which were described in the previous chapter, a high degree of detail was sought in examining traffic processes taking place at those special intersections. The TEXAS model, which provides a microscopic level of traffic detail and allows a range of diamond-oriented control strategies to be tested, was selected for this purpose. Strategies developed through extensive experimentation using the TEXAS model were subsequently evaluated in an integrated framework using DYNASMART.

Similarly, because both DYNASMART and the TEXAS model are descriptive simulation tools and not optimization tools (i.e., they are useful in evaluating the performance of the network under different control strategies specified by the analyst, but are not designed to solve automatically for the optimum control) other models typically used in practice to optimize the control for specific components were selected. These were used in the study to generate control plans that were subsequently evaluated in an integrated framework by DYNASMART.

In order to perform all the functions required for this study, particularly in terms of simulating different types of control strategies generated by other models for specific components of the network, the DYNASMART model was enhanced and modified specifically for the purpose of evaluating integrated freeway corridor network control strategies. In particular, additional sensitivity for representing specific diamond interchange control options were introduced, along with various modifications intended to provide greater flexibility in representing and evaluating the various control strategies. These are described in the next section, together with the structure and logic of DYNASMART.

The TEXAS model for diamond interchange is then described in Section 3.3, followed in Section 3.4 by the TRANSYT 7-F, PASSER II, and CORSIM models used for signalization frontage and arterial traffic control and evaluation. The selected test bed network is described in Section 3.5, including geometric and operational characteristics and various assumptions made in the process of generating the input data for the analysis. The structure of the simulation experiments is described in Section 3.6, followed by concluding comments in Section 3.7.

3.2. INTEGRATED NETWORK SIMULATION-ASSIGNMENT: DYNASMART

DYNASMART (**DY**namic **N**etwork **A**ssignment **S**imulation **M**odel for **A**dvanced **R**oad **T**elematics) was selected as the primary integrated network modeling tool for this study. In its present form, DYNASMART is primarily a descriptive analysis tool used to evaluate information supply strategies, traffic control measures, and route assignment rules at the network level. The model is designed around a flexible structure that provides (1) sensitivity to a wide range of traffic control measures for both intersections and freeways, (2) capability to model traffic disruptions caused by incidents and other occurrences, and (3) representation of several user classes corresponding to different vehicle performance characteristics (e.g., cars vs. trucks), access to physical facilities (e.g., HOV lanes), different information availability status, and different behavioral rules. DYNASMART is a comprehensive simulation-assignment framework comprising traffic flow models, path processing methodologies, behavioral rules, and information supply strategies (Ref 35). As such, it is the only modeling tool available that allows investigation of traffic networks under a full range of ATMS and ATIS technologies and capabilities.

Several functional requirements must be incorporated into a model in order for it to be effective in evaluating the performance of networks under ATMS/ATIS. These requirements — all satisfied by DYNASMART — include the:

1. ability to model the route choice and diversion behavior of motorists with access to varying types and degrees of information, ranging from on-board ATIS, to VMS, to no information availability;
2. responsiveness to dynamic origin-destination (OD) information, such as might become available to the TMC from historical databases, sensors, probes, and other sources;
3. ability to track the location of drivers, both those that receive ATIS guidance and those that do not;
4. ability to determine the time-dependent impedance (travel time) based on the assignment decisions;
5. ability to model a range of traffic control strategies, such as traffic-responsive signal settings and ramp metering; and the
6. ability to comprehensively model both freeway traffic and surface street traffic in an integrated manner.

DYNASMART provides the above capabilities from the standpoint of the present study. We note that it can simulate several traffic control types, including pretimed and actuated intersection signal control and freeway control measures. Additionally, incidents can be specified by indicating their respective location, severity, and expected duration. As the primary evaluation framework for the integrated control strategies in the kind of freeway corridor networks of interest to this study, it offers the following advantages and unique capabilities:

1. The representation of all link types is standardized and detailed, allowing a comparable (and accurate according to the type of segment in the network) evaluation of all network components.
2. The assignment procedure considers the effects of dynamic traffic demands and controls. It represents traffic conditions at varying congestion levels from undersaturated to oversaturated, considering the build up and dissipation of queues in the network.

3. The model integrates flow modeling and assignment procedures. This feature allows assignment impacts to be properly reflected in the flow model, while correctly reflecting flow characteristics in the assignment procedures.
4. The signal control capabilities considered include any feasible phasing pattern, cycle length, splits, and offsets for pretimed and vehicle-actuated signals. The model also represents stop- and yield-controlled intersections, as well as freeway ramp metering. In all these cases, oversaturation, queue build-up and dissipation, and queue spill-back are considered.
5. The model can account for different vehicle types (e.g., ATIS-equipped and nonequipped) and different levels of penetration of in-vehicle route-guidance systems.
6. It is able to simulate nonequipped vehicles with prespecified paths (selected by the model or exogenously determined based on historical patterns).
7. It is able to apply and maintain the FIFO (first-in, first-out) property in serving queues and moving vehicles under various control strategies.
8. It is able to represent different demand patterns and demand intensities corresponding to different degrees of congestion, including extreme oversaturated conditions associated with major incidents.

DYNASMART was used to simulate the network in the experiments performed in this study. The objectives and description of the experiments are presented in Section 3.6.

In the remainder of this section, the structure of the DYNASMART model is first described, followed, respectively, by the traffic simulation component, the user decisions component, and the path processing component. The model capabilities in terms of incidents and ATMS measures are then presented.

Model Structure

The framework of the DYNASMART simulation-assignment model and its three principal components are illustrated in Figure 3.1 (Ref 32). The simulation model is an extension of a macroparticle simulation model initially developed by Chang, Mahmassani, and Herman as a special-purpose code for experimental studies of commuter behavior dynamics in traffic corridors (Ref 31). DYNASMART integrates traffic flow models, path processing methodologies, behavioral rules, and information supply strategies into a single simulation-assignment framework. The input data include a time-dependent origin-destination matrix (or a schedule of individual departures) and network data. Given the network representation, which includes link characteristics as well as control parameters, the simulation component will take a time-dependent loading pattern and process the movement of vehicles on links and the transfers between links according to specified control parameters. These transfers, which are determined by path processing and path selection rules, require instructions that direct vehicles approaching the downstream node of a link to the desired outgoing link. The user behavior component is the source of these instructions.

Traffic Simulation Component

DYNASMART uses established macroscopic traffic flow models and relationships to model the flow of vehicles through a network. Whereas macroscopic simulation models do not keep track of individual vehicles, DYNASMART moves vehicles individually or in packets, thereby keeping a record of the locations and itineraries of the individual particles. This level of representation has also been referred to as "mesoscopic." Multiple-user classes of different vehicle performance characteristics are modeled as packets consisting of one or more passenger car units; for instance, a bus is represented by a packet with two (or other user-specified values) passenger car units. The traffic simulation consists of two principal modules: link movement and node transfer, as described below.

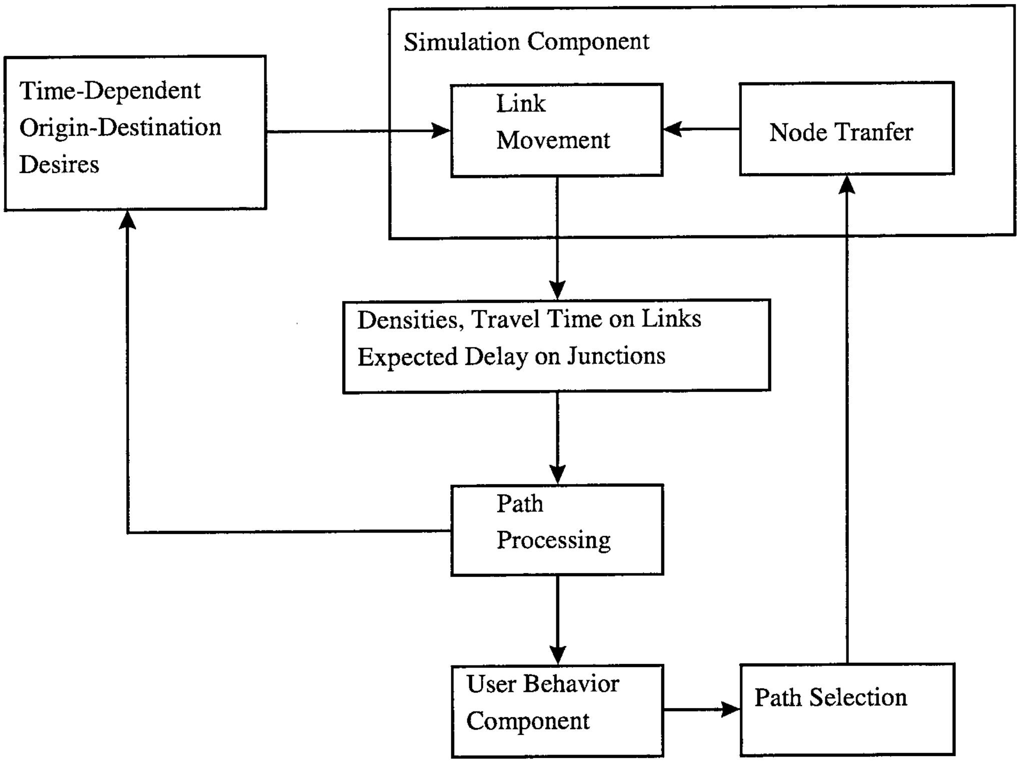


Figure 3.1: Structure of DYNASMART simulation-assignment model (Source: Reference 35)

Link Movement: The link movement is a process for moving vehicles on links during each scanning time interval in the simulation (time step). Note that the network links are subdivided into smaller sections or segments for traffic simulation purposes. The vehicle concentration prevailing on a section over a simulation time step is determined from the solution of the finite difference form of the usual continuity equation, given the concentration and the inflows and outflows over the previous time step (Ref 33). Using the current concentration, the corresponding section's speeds are calculated according to a modified Greenshield speed-density relationship, namely:

$$V_i^t = \left(V_f - V_0\right)\left(1 - \frac{K_i^t}{K_0}\right)^{\alpha} + V_0$$

where

V_i^t , K = mean speed and concentration in section i during the *t-th* time step,

V_f, V_0 = mean free speed and the minimum speed, respectively,

K_0 = jam concentration, and

α = a parameter used to capture the sensitivity of speed to the concentration.

Other traffic stream models may also be incorporated into DYNASMART based on field investigation.

Node Transfer: The node transfer module performs the link-to-link or section-to-section transfer of vehicles at nodes. For interrupted link flow, it appropriately allocates the right-of-way according to the prevailing control strategy. The output of the node transfer includes the number of vehicles that remain in queue and the number added to and subtracted from each link section for each simulation time step. A wide range of traffic control measures for both intersections and freeways is reflected in the outflow and inflow capacity constraints that govern the node transfer (Ref 35).

User Behavior Component with Real-Time Dynamics

One of the principal features of DYNASMART that allows it to interface with activity-based behavioral models is its explicit representation of individual tripmaking decisions, particularly for path selection decisions, both at the trip origin and en-route. Behavioral rules governing route choice decisions are incorporated, including the special case in which drivers are assumed (required) to follow specific route guidance instructions. Experimental evidence presented by Mahmassani and Stephan (Ref 34) suggested that commuter route choice behavior exhibits a boundedly rational character. This means that drivers look for gains only outside a threshold, within which the results are satisfying to them. This can be translated into the following route switching model (Ref 35):

$$\delta_j(k) = \begin{cases} 1 \text{ if } \mathrm{TTC}_j(k) - \mathrm{TTB}_j(k) > \max\left(\eta_j \cdot \mathrm{TTC}_j(k), \tau_j\right) \\ 0 \text{ otherwise} \end{cases}$$

where $\delta_j(k)$ is a binary indicator variable equal to 1 when user j switches from the current path to the best alternate, and 0 if the current path is maintained; $TTC_j(k)$ and $TTB_j(k)$ are the trip times along the current path and along the best path from node k to the destination on current path, respectively; η_j is a relative indifference threshold, and τ_j is an absolute minimum travel time improvement needed for a switch.

The threshold level may reflect perceptual factors, preferential indifference, or persistence and aversion to switching. The quantity η_j governs users' responses to the supplied information and their propensity to switch. The minimum improvement τ_j is currently taken to be identical across users according to user-defined values. Results of laboratory experiments indicate that τ_j is on average equal to 1 minute, while η_j is about 0.2 for typical urban commutes (Ref 46).

Path Processing

The path processing component of DYNASMART determines the route-level attributes (e.g., travel time) for use in the user behavior component, given the link-level attributes obtained from the simulator. For this purpose, a multiple-user class *K*-shortest path algorithm with movement penalties is interfaced with the simulation model to calculate *K* different paths for every origin-destination pair. However, in order to improve the model's computational performance, the *K*-shortest paths are not recalculated every simulation time step, but only at prespecified intervals. In the interim, the travel times on the set of *K* current paths are updated using the prevailing link travel times at each simulation time step (or every few steps to further reduce computational requirements). There are two important ways that this path information is used:

1. *Initial Routes.* At the beginning of trips, nonequipped drivers need to be assigned to specific paths or initial routes. While there is no universally agreed-upon process for assigning initial routes, some researchers have suggested user equilibrium or stochastic user equilibrium assignment for these initial routes. In DYNASMART, initial routes are modeled in an explicit way, allocating drivers to the *K*-shortest paths according to a prespecified rule. Of course, when DYNASMART is used as a simulator in conjunction with an algorithmic search procedure, initial paths may be determined by the search. In practice, such assignments for some vehicles may also be available from historical information based on actual measurements.
2. *Current Path Information.* Current path information forms the basis of driver path choice decisions at every node according to the user behavior component module. In its present version, only current trip times are available to drivers. The current path information is used in equipped vehicles as well as in the variable message signs (VMS) route control module. (The latter is explained further in the VMS sections.) A time-dependent *K*-shortest path routine has also

been developed and could be incorporated within DYNASMART to simulate anticipatory information supply strategies. Such “anticipatory” strategies are now provided with the system optimal, user equilibrium, or multiple user class assignment algorithms. Additional anticipatory strategies with predicted time-dependent trip times can also be easily implemented if a data fusion and prediction function is provided (in a separate module).

Traffic Simulation in DYNASMART

DYNASMART uses macroscopic traffic models to quantify interactions among vehicles and to calculate movements of vehicles along links. However, there are features that need to be included in order to capture traffic complexities and to provide essential capabilities for ATIS/ATMS applications. This section addresses these features in the modeling process.

Traffic Control Elements

DYNASMART provides the ability to explicitly model the array of control elements listed in Table 3.1. The major element for surface streets is signal control, which includes pretimed control and actuated control, as well as signal coordination along arterial streets. Ramp metering and variable message signs (VMS) are the major controls for the freeway system. The geometric configurations and measures of effectiveness that are to be included are also listed in Table 3.1. The following sections address these elements in detail.

Table 3.1: Traffic control strategies in DYNASMART

Surface Street	**Freeway System**
I. Control Types	a. Ramp metering
a. No control	b. Changeable message signs
b. Yield control	
c. Stop signs	
d. Signal control (green, red, amber time, cycle time, offsets, phases)	
Pretimed	
Pretimed coordinated	
Multidial pretimed	
Actuated (full)	
II. Geometric Configurations	a. Number of lanes
a. Saturation flow rate	b. Capacity
b. Number of lanes	c. HOV lanes
c. Number of approaches	
III. Measure of Effectiveness	
a. Average speed	a. Average speed
b. Average travel time	b. Average density
c. Average delay	c. Average ramp queue length

Capacity Control

The node transfer is designed to simulate the input and output flows of vehicles on each approach at intersections operating under a number of control strategies. It calculates the number of vehicles traversing each intersection in the network during each simulation time step, as well as the number of vehicles entering and exiting the network. Several concepts regarding the modeling of vehicle flows in the node transfer (discussed below) include outflow and inflow capacity constraints, equivalent green time for unsignalized intersections, and signalized control.

Outflow Capacity Constraints: The outflow constraints limit the maximum number of vehicles allowed to leave each approach lane at an intersection. These constraints are described in the following equation, which states that the total number of vehicles that enter an intersection (from a given approach) depends on the number of vehicles waiting in the queue at the end of the current simulation interval (time step), ΔT, and the capacity of this approach. The definition of capacity follows the HCM, and consists of the maximum number of vehicles that can be served under prevailing traffic signal operation.

$$VI_i = \min\{VQ_i ; VS_i\}$$

where

VI_i = maximum number of vehicles that can enter the intersection during ΔT,

VQ_i = number of vehicles in queue on link i at the end of ΔT,

VS_i = maximum number of vehicles that can enter the intersection during ΔT, i.e., $G_i S_i$

G_I = remaining effective green time during ΔT,

S_I = saturation flow rate, and

ΔT = the simulation interval.

Inflow Capacity Constraints: The inflow constraints determine the maximum number of vehicles allowed to enter a link. These constraints bound the total number of vehicles from all approaches that can be accepted by the receiving link; they include the maximum number of vehicles from all upstream links wishing to enter link j, the available physical space constraint on the outbound link, and the section capacity constraint of link j.

$$\mathrm{VO_j} = \min\left\{\sum_{k\in U} \mathrm{VI_{kj}}, \mathrm{VE_j}, \mathrm{C_j}\Delta\mathrm{T}\right\}$$

where

VO_j = number of vehicles that can enter link j,

U = set of inbound links into link j (i.e., in the backward star of j),

VI_{kj} = number of vehicles wishing to move from k to j,

VE_j = the available space on link j, and

C_j = approach capacity of link j.

Signal Control

Signal control can be separated into pretimed signal control, pretimed coordinated control, multidial pretimed signal control, and actuated signal control. All such signal controls are modeled explicitly in DYNASMART. Mahmassani et al. have described detailed input data preparation in a technical report (Ref 35).

Equivalent Green Time for Unsignalized Intersections: DYNASMART uses the equivalent green time concept to allocate the right-of-way based on the incoming volume at unsignalized intersections. This can be applied to no control, stop sign, and yield sign control.

$$GE_i = \left(\frac{CVQ_i}{\sum_k CVQ_k} \right) \Delta T$$

where

GE_i = equivalent green time for i-th phase,

CVQ_I = critical vehicle volume in queue of i-th phase, and

ΔT = simulation time step.

Greater detail in modeling unsignalized intersections is not warranted for ATIS/ATMS applications because such intersections tend to be relatively uncongested and tend to serve mostly local traffic needs.

Pretimed Signal and Pretimed Coordinated Control: Input data in this module include phase number, offset, green time, red time, and amber time for every phase. For pretimed signal control, green times are set for every phase according to these data. Since DYNASMART is not intended as an optimizer of signal system control, the model user has to input offsets obtained exogenously from other models to coordinate arterial streets or the network as a whole. Alternatively, optimization modules could be developed for ATMS applications.

Actuated Signal Control: Instead of detecting individual vehicles, DYNASMART uses an appropriate macroscopic method that determines equivalent green times that are updated to reflect prevailing approach volumes. Two alternative methods are provided in the current version of DYNASMART to represent the actuated signal control.

In the first method, green splits are apportioned according to Webster's rule for the measured arrival flow rate. This approach attempts to capture the essential features of actuated signal control: "max out" and "gap out." Max out occurs when the green time for a given phase reaches a preset maximum green time; it is modeled explicitly here. Gap out occurs in the field when a preset time elapses with no detector (generally specified to avoid excessively long delays at conflicting approaches) actuations for the phase in progress, resulting in discontinuation of the green for that phase. In the simulation, because detector actuations are not directly simulated, gap out is emulated under the first method, as the provided green time is intended to serve only vehicles present on a particular approach. The input data set in DYNASMART includes maximum green time, minimum green time, default cycle length, and other signal data (such as phase number). The equation used in calculating green time under the first method for actuated signal control emulation is given below. The concept is to allocate green time depending on incoming volume. If the required green time is larger than the maximum green time or smaller than the minimum green time, the maximum or minimum green time is assigned, respectively.

$$G_i = \left(\frac{CV_i}{\sum_j CV_j} \right) (C - lost\ time)$$

subject to

Min Green = G_i = Max Green,

CV_i = critical volume for phase I, and

C = default cycle length.

If CV_i is less than the maximum number of allowable vehicles, the green time will be reduced accordingly. These calculations are performed at the end of the current cycle. Cycle length will change every cycle. This modeling will be fairly accurate in allocating green time as congestion increases in the network. It will be somewhat less accurate under light traffic conditions; however, the dynamic assignment capabilities in ATMS are of primary concern during congested periods.

In the second method, the green time for a given phase is determined based on the number of vehicles that would have reached the intersection at the end of the current simulation interval. This green is subsequently extended as appropriate for each simulation interval until max out is reached, or terminated if no longer needed, thereby emulating gap out. This second method does not require a default cycle length and may skip a phase altogether if no vehicle demand exists and no minimum green is specified.

Real-Time Signal Control: DYNASMART provides an independent module for real-time signal control that includes an interface to update signal parameters during the simulation. These parameters can be controlled by user-specified rules or be prepared exogenously in advance. The module is intended to assist in testing different real-time control strategies.

Communication Interface between Simulation and Path Processing

In DYNASMART, the path-processing component utilizes the travel time information generated from the simulation. The travel time information for links is separated into two parts: travel time for vehicle movement and queuing time. Traffic on each link segment is modeled as consisting of two parts (shown in Figure 3.2): those vehicles in the upstream (moving part) and those in the downstream (queuing part).

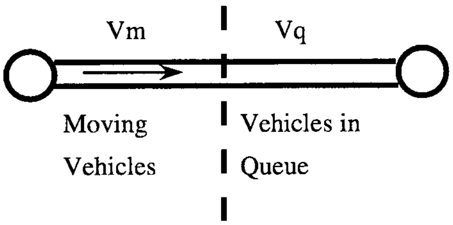


Figure 3.2: Conceptual portions on a link segment

Average Travel Time for Moving Part: The average travel time on link segments for each time step is calculated directly from the traffic stream model used in the simulation. The density and speed are obtained for every simulation interval; travel time is then calculated from available length and associated speed.

Average Queue Delay: The average queue delay is considered to be the time for clearing the queue at the service rate experienced in the recent past over a certain period (say 3 minutes). The queue discharge calculation considers the average outflow rate and the congestion of downstream links.

$$Queue\ Delay = \frac{V_q}{AS_i}$$

where

V_q = number of vehicles in queue (vehicles),

AS_i = average flow rate (vehicles/seconds),

$$AS_i = \frac{\sum_{k=t}^{t+T} f_k}{T}$$

f_k = flow rate at *k*-intervals, and

T = period over which the average queue service rate is calculated.

Incident Modeling

Incidents are modeled in DYNASMART to reflect accidents, lane closures, or other capacity-reducing occurrences. Basically, incidents are modeled completely based on external data and can be specified to occur at any time during the simulation on any link or segment. All incidents cause the reduction of lane capacity. If a whole link or segment is closed, all vehicles (equipped as well as nonequipped) otherwise using the link are diverted to other paths. Some features of incident modeling in DYNASMART include the following:

1. Incidents are specified as reductions of link capacity for a specified time period.
2. All calculations are based on user-specified input information about the incident specifics (location, start time, end time, severity).
3. Complex incidents can be modeled as a series of consecutive incidents.
4. Nonequipped vehicles will be diverted for a street closure only when they reach the upstream node of the blocked link.

Freeway Control

Freeway management techniques can be categorized as *capacity management* and *demand management.* Capacity management, such as ramp control and variable speed control, attempts to maximize throughput and to maintain a certain level of service. Demand management, on the other hand, attempts to reduce the number of vehicles at the peak period. In DYNASMART, two important elements of freeway management are implemented, namely, entrance ramp control and HOV lanes. In addition, variable message signs (including speed control for mainline regulation) may be modeled, though these are not limited to freeway links.

Ramp Control

Ramp control is the most widely used freeway control measure. Its purpose is to limit the number of entering vehicles in order to maintain a satisfactory level of service within a capacity limit. Ramp control includes entrance ramp control as well as exit ramp control. Since exit ramp control is seldom used, it is not explicitly modeled in DYNASMART. However, it could be simulated through other built-in modules, such as lane closure and VMS. According to the *Traffic Control Systems Handbook* (Ref 36), there are five types of entrance ramp control: closure, ramp metering, traffic-responsive metering, gap-acceptance merge control, and integrated ramp control. The first three methods, explicitly modeled in DYNASMART, are explained as follows:

1. *Closure.* For ramp closure, drivers need to select alternate routes to their destination. Since equipped vehicles receive current traffic information, they can respond to ramp closure before they reach the ramp. On the other hand, nonequipped vehicles do not have this advantage, so they will choose another

route after they reach the closed ramp. However, the VMS can be applied on arterial streets as early warning, so nonequipped vehicles can be diverted prior to their arrival. The choice of alternate route for nonequipped vehicles also depends on driver behavior, and requires an observational basis to develop appropriate path selection rules. DYNASMART provides a flexible way to divert the nonequipped vehicles that allows users to define a *k*-th best path number or to randomly choose a path from path files.

2. *Ramp Metering.* Basically, DYNASMART controls vehicle flow under in-flow and out-flow constraints. In ramp metering, a fixed ramp rate or a dynamic ramp rate that determines the maximum number of entering vehicles can be determined in conjunction with the capacity calculations during a specified time period.
3. *Traffic-Responsive Metering.* Traffic-responsive metering is directly controlled by the mainline and ramp traffic conditions during the metering period. Occupancy control and demand control are two widely used methods for traffic-responsive metering. ALINEA (Ref 17), a local feedback control law for on-ramp metering, is implemented in DYNASMART. A typical feedback law is given as follows:

$$r(k) = r(k-1) + K_R\,[\hat{o} - o_{out}(k)]$$

where

K_R = rate adjustment parameter (default value 0.32),

$\hat{O}$ = nominal (target) occupancy (default value 0.2),

$o_{out}(k)$ = detector occupancy at time k, and

$r(k)$ = entrance ramp rate at time k. max: 35 - 25 vehs/min-lane; min: 5 vehs/min-lane.

The given default values of K_R and $\hat{o}$ are from numerical results obtained by Joseph (Ref 47) and are intended for illustrative purposes only.

High Occupancy Vehicle Priority Control

Priority for high occupancy vehicles is to provide preferential treatment through HOV lanes for buses and carpools. The purpose of HOV lanes is to encourage carpools or buses in order to reduce overall vehicle demand. Methods of priority control include separated facilities, reserved lanes, and priority access control. In DYNASMART, HOV lanes are part of a traffic network represented by links and nodes. In order to preclude non-HOVs from using the HOV lanes, the travel times on these links are set to infinity for non-HOVs for the path calculation.

Left-Turn Movement

Left-turn movements, as shown in Figure 3.3, are a critical delay-causing factor in urban networks. However, it is very difficult to model the left-turn movement in a macroscopic simulation model. In this section, the left-turn issue is discussed and the modeling process used in DYNASMART is introduced.

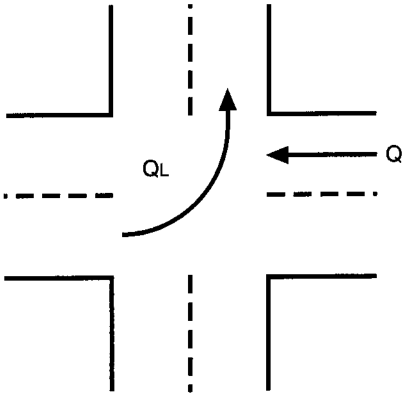


Figure 3.3: Left-turn movement

For left-turn movements without a turning phase, the analytical approach is to calculate the blocked time by opposing vehicle flow at the onset of green and then use gap acceptance models to calculate the actual number of vehicles that can pass the intersection during the residual green interval.

Left-turn capacity is determined by several factors, including opposing volume, number of lanes of the opposing approach, and green time for this phase.

The blocked time from the onset of green that left-turning vehicles cannot use is calculated as follows:

Blocked Time:

$$T_b = \frac{\left(Q/N \right)(L1 + L2 + R)}{S - \left(Q/N \right)}$$

where

Tb = blocked time, time blocked by opposing traffic; clear time for queue,
Q = total opposing flow (vehs/hr),
N = number of opposing lanes,
$L1$ = lost time for opposing traffic,
$L2$ = lost time for start-up,
R = red time, and
S = saturation flow for opposing traffic.

Then, usable time for left-turn vehicles can be calculated as:

$$T_u = G + (T_a - L1) - L2 - T_b$$

where

T_u = usable time of cycle for left-turn (seconds),

G = green time, and

T_a = amber (yellow) time.

The maximum number of possible left-turn vehicles is equal to:

$n = (T_u / h) + 1$, where h : minimum turning headway ($\cong 2.5$ seconds)

Thus, a gap acceptance model (Ref 48) can be used to calculate the left-turn capacity as follows:

$$Q_L = \left(\frac{T_u}{C}\right) Q_{LT}$$

$$Q_{LT} = \frac{Q\, e^{-(Q/3600)T_c}}{1 - e^{-(Q/3600)h}}$$

where

Q_L = left-turn capacity,

Q_{LT} = left-turn saturation flow, veh/hr,

T_c = critical gap, seconds, and

H = turning headway, seconds ($\cong 2.5$ seconds).

The modeling process for the left turn is complex and not easy to combine with any macroscopic simulation. Therefore, a heuristic modeling process is used to capture the effects of left turns in DYNASMART. The process is summarized as follows:

1. Count left-turn vehicles.
2. Calculate maximum flow rate for left turns; this rate can be calculated under different situations:
   a) Protected left-turn phase: saturation flow rate.
   b) Permissive phase: from gap acceptance models or established tables.
3. Calculate an average number of left-turn vehicles and also reduce the saturation flow rate for straight and right-turn approaches.
4. Follow outflow/inflow constraints to transfer vehicles from link to link.

5. Calculate the left-turn delay for the *K*-shortest path calculation.

Left-turn capacity estimation determines the number of left-turn vehicles that can enter the intersection without delays owing to opposing volume. Different approaches have been used in determining the left-turn capacity. For example, a gap acceptance model is applied in TRANSYT 7F for permissive movement (Ref 49). A review of left-turn capacity issues can be found in Lin et al. (Ref 50). DYNASMART adopted the left-turn capacity values from Lin et al. (1984) that derive from simulations using the TEXAS (Ref 51) model. The left-turn capacity is determined by several factors, such as opposing flow, number of opposing lanes, and signal timing. The saturation flow rate for other movements is adjusted according to the 1985 HCM (Ref 37). The left-factor in the adjustment is based on four variables, namely, exclusive or shared lanes, type of phasing, proportion of left-turn vehicles, and opposing volume. The left-turn capacity and adjusted saturation flow rate are used in inflow/outflow capacity constraints.

Multiple-User Classes

DYNASMART allows for different classes of users with different information availability and/or behavioral responses and/or traffic performance characteristics. Vehicle classes can differ by vehicle type, network restrictions, and information availability. Since a variety of attributes are generated for vehicles, vehicles are not identical even within the same class. Currently, seven different classes are modeled in DYNASMART for illustration purposes, and more classes can be included. The seven classes are:

1. nonequipped passenger car,
2. nonequipped truck,
3. nonequipped high occupancy passenger car,
4. equipped passenger car,
5. equipped truck,
6. equipped high occupancy passenger car, and
7. bus.

All the equipped vehicles follow the rules stipulated in the user decisions component. In the current version, the default is the boundedly rational behavior rule discussed earlier, with a relative indifference band and a minimum threshold value. Different vehicle sizes are modeled as packets of different passenger car units specified by the user. The packet size is used in calculating concentration, available capacity, and inflow and outflow constraints. With this ability, DYNASMART can model virtually any network restrictions, such as turning prohibitions, and special facilities, such as bridges. (The HOV concept was described in a previous section. Bus operation is discussed in a later section.)

Variable Message Signs (VMS)

One way to provide dynamic route information to drivers is by means of variable message signs (VMS), where visual word, number, or symbolic displays can be electronically or mechanically varied according to current traffic conditions. VMS displays can address a considerably wide range of traffic management functions; however, drivers are not usually required to follow all messages posted on a VMS. Since the response of drivers to different VMS displays is still in need of further study, the use of the VMS module in DYNASMART should be accompanied by a reasonable assumption regarding driver behavior. The VMS module in DYNASMART includes three parts: speed advisory, route advisory, and route warning messages.

Speed Advisory: Speed advisory is mainly used for mainline control of freeway systems. Experiments with speed advisory changes have been undertaken in several European countries. Through field experiments, it has been reported that reasonable speed limitations imposed during rush hours increase capacity (Ref 18). In DYNASMART, speed advisory applies at VMS locations when the density exceeds a prespecified value. Then, all the vehicles are assigned the advised speed.

Route Advisory: Route advisory may provide an alternative path for vehicles in order to avoid a congested section. In DYNASMART, the user needs to define a *k*-th number of paths (or a fixed path) to be displayed, and all the vehicles will follow the new path to their destinations. Of course, a more comprehensive set of response rules will need to be specified in the user behavior component as results of related targeted research become available.

Route Warning: This form of real-time information instructs drivers to divert in advance of a congested section. In the current implementation, the warning message is generated when the concentration of downstream link reaches the maximum concentration and a given fraction of the vehicles are diverted to other randomly generated routes. The intent is to retain the flexibility to incorporate more complete instructions as ongoing research into ATMS strategies produces testable concepts.

Bus Operations

In DYNASMART, buses are treated as packets having predefined paths; each packet includes two passenger car units. Simulation of bus operations largely depends on related input information, namely:

BUS ID: an identifier for bus
Start Time: the start time of the bus
Average stop (dwell) time
Number of nodes in the route
The sequence of nodes
The activities on links
0: no stop

1: stop at the near side
2: stop at the midblock
3: midblock curb stop (or bus bay)

During the simulation, each bus is treated as a packet of two passenger car units. In the link movement, buses are mixed with other vehicles when calculating the prevailing average speeds and concentration. In the node transfer, capacity with two PCUs is used for transferring a bus from link to link. Loading and unloading of buses will cause the short-term blockage of traffic, and this situation is modeled in DYNASMART according to the locations of bus stops. If the location of the stop is near an intersection, one lane of outflow capacity will be dropped. If the location of a bus stop is in the middle of a block, the short-term blockage will be simulated as a short-term incident. The blockage time is defined as the average dwell time (the user needs to include the average additional time loss resulting from starting). According to the 1994 HCM, where the buses stop in a lane that is not used by moving traffic (a curb parking lane or a bus bay), the time loss to other vehicles is approximately 3 to 4 seconds per bus. The blockage time of midblock curb stop is set at 4 seconds, but can of course be readily changed to reflect actual conditions.

The above modeling of bus operations is not limited to buses, but may also be applied to any other vehicle type having a fixed route and schedule.

Other Considerations

Driver Compliance Factors: Driver compliance factors are modeled as part of the user decision component. Several possible rules can be postulated for this behavioral process — rules that will eventually be developed based on empirical experimental evidence. The ability of DYNASMART to explicitly model multiple user classes on the basis of behavior provides the necessary flexibility to accommodate a wide range of possible compliance rules.

Output Information: For different analyses, three levels of output can be obtained from DYNASMART:

1. Overall system performance (the statistics are also reported for different user classes)
   - average overall travel time
   - average travel (moving) time
   - average entry queue time
   - average stop time
   - average travel distance
   - congestion index
   - simulation summary report

2. Selective information
   - Link

  average speed
  average density
  average end queue
  total number of vehicles passed by
- Vehicle
  behavior attributes
  travel time
  travel distance
  traveled path

3. Detailed information
   vehicle trajectories
   signal timing
   path information
   concentration profiles

Input Data Description

For modeling purposes, the following data should be provided to the program:

1. *Network data*: Zoning, nodes, arcs, and type of arc and number of lanes, saturation flows, maximum speed, and length of each arc in the network.
2. *Movement data*: Indication of connection among arcs and permitted/prohibited turns in the network.
3. *Signal control data*: Signal nodes, type of control, phasing, green time allocation and offset.
4. *Demand data*: Indication of the number of loading intervals and associated time periods, and definition of an O-D matrix for each period in order to generate and load vehicles into the network. There is also a direct way to indicate the demand data if the characteristics of each vehicle including its path are known.
5. *Ramp control data*: Number and location of entrance ramps and ramp metering parameters.
6. *Incident data*: Number of incidents, location in the network, starting time during the simulation, and severity.
7. *Variable message sign data*: Location, type of message, and parameters.
8. *Bus data*: Number of buses considered during the simulation, route and departing time for each bus, and operation parameters.
9. *Scenario data:* Simulation parameters, users classes, equipped vehicles, and user parameters.

Output Description

There is a wealth of output produced by DYNASMART. The following are descriptions of the output files that are used the most to determine the network's performance measures.

Summary of network performance: Extensive information about the network is provided in this file, including basic information about the network, history of the vehicles loaded during the simulation, and the overall statistics report, including trip times, stop times, and trip distance.

Vehicle trajectories: For each one of the vehicles simulated, the following information is provided: path and its characteristic trip time on every link, and stop time.

Volume on links: Number of vehicles on each link at the end of each simulation interval.

Vehicle queue: Number of vehicles in queue at the end of each simulation interval.

Speed: Average speed on each link during each simulation interval.

Concentration: Average concentration on each link during each simulation interval.

Green time: Green time for each approach during each simulation interval.

Real-Time Implementation: DYNASMART-X and the Rolling Horizon Approach

The principal mechanism proposed for implementing dynamic traffic assignment (DTA) capabilities in real time is the rolling horizon (RH) approach, used previously for production-inventory control (Ref 38) and in transportation engineering for on-line demand-responsive traffic signal control (Refs 39, 40). The underlying philosophy behind the RH approach is that current events will not be influenced by events "far" into the future, i.e., that vehicles currently assigned will not be influenced by vehicles assigned "far" into the future, as the currently assigned vehicles will probably be out of the system by that time. The stage length h in Figure 3.4 depicts that length of time (its value in actual problems is network specific). The roll period l represents the short duration into the future for which O-D desires are available with reasonable reliability. To make an assignment of vehicles to various paths for the current period, the estimation and prediction functions of the DTA system need to be exercised to produce O-D desires for the rest of the stage. The O-D desires beyond the stage length h are assumed to be zero. The path assignments in each stage are determined for the entire stage, but implemented only for the roll period.

For a given stage, the problem encountered is analogous to the complete information availability scenario, albeit only for the duration h of the stage. The system is solved for optimality only for this duration, and O-D desires for the roll period are assigned to the paths determined. The time frame is now "rolled" forward by the roll period, and the above process is repeated until the end of the duration of system operation, possibly over an infinite horizon. Hence, a series of optimizations are performed in quasi real time. From a simulation standpoint, it is necessary to ensure proper initial conditions as one advances from one stage to the next.

The selection of the values of l and h in the RH approach depends on a careful consideration of the trade-offs among quality of O-D prediction, computational requirements, rate of change in the traffic system, and solution quality. Simulation experiments have suggested roll periods of the order of 10 to 15 minutes, with stages of about 20 to 30 minutes in duration for DTA implementation (Ref 41).

In addition to the simulation-assignment model and normative route guidance modules used to generate a solution for the next roll period, the support routines generating the O-D predictions and performing consistency checking and resetting must be executing, possibly several times per roll period, in a real-time implementation of the overall DYNASMART DTA system (called DYNASMART-X).

Real-time systems are expected to produce responses within a definite time limit. Designing such systems requires identification of the relevant tasks and events and the associated deadlines. For the DTA system, these events include trip starts, traffic sensor outputs, controller updates, probe data, traffic incidents, scheduled traffic events, and route guidance outputs, among others. These external processes operate on their own time scales; actions carried out in the computer must relate to the time scales of the external processes. Most real-time systems perform a mixture of tasks that can be classified as clock based (periodic), event based (aperiodic), and interactive, with both hard, firm, and soft time constraints, in addition to activities that are not real time.

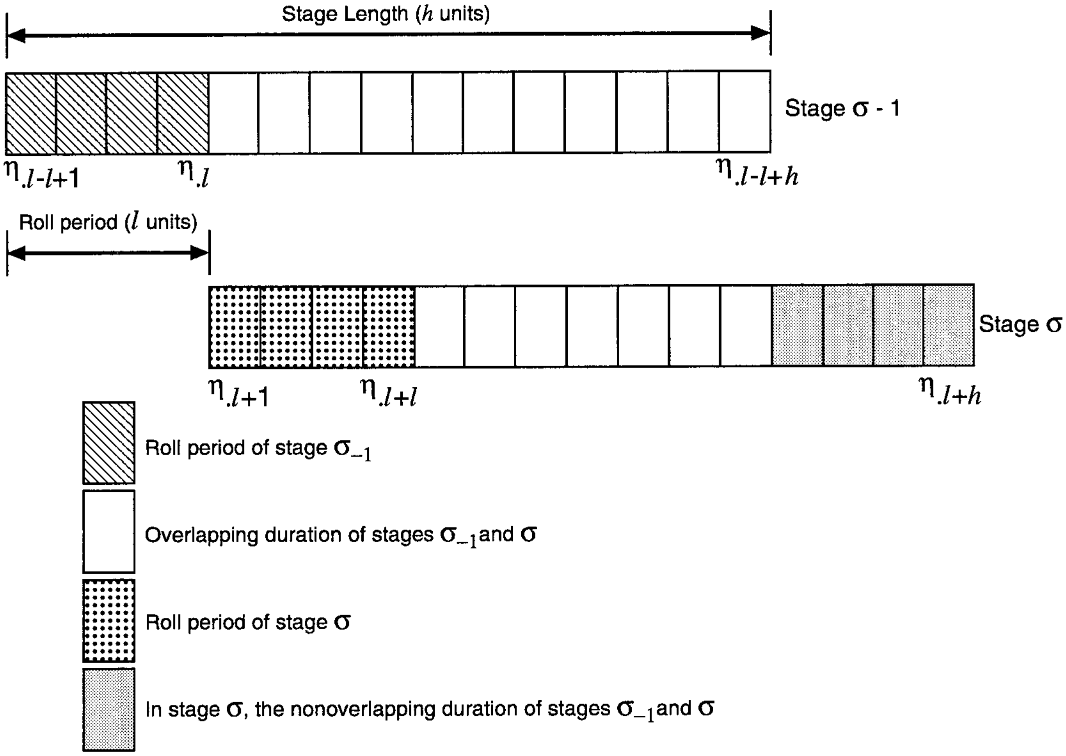


Figure 3.4: Rolling horizon implementation framework

DYNASMART-X can be viewed as an event- and data-driven system. External information about the traffic network is made available on a regular, periodic basis. The DYNASMART-X system operates on these data and provides ATMS/ATIS normative outputs in response, also on a periodic schedule. DYNASMART-X, however, does not poll sensors and does not directly update controller timing; nor does it provide route guidance to individual travelers. Instead, it interfaces to a set of external real-time functions that perform these tasks. Under normal operation, a set sequence of real-time software modules must run each time that DYNASMART-X produces an output (Ref 52).

However, a key motivation for intelligent transportation systems and real-time control is that operation will not always follow normal predictable conditions. Exception handling is therefore a major consideration for DYNASMART-X. When external events require recalibration of the model parameters to achieve greater consistency between internal values and externally observed quantities, deadlines must still be met, and the system may enter a degraded mode of operation where the ATMS/ATIS outputs can continue to be provided, but with a reduced degree of systemwide optimality. This type of strategy is common in complex real-time systems.

From a software design standpoint, it is desirable to place activities with different types of time constraints (including non-real-time portion) into separate modules, and then explore the options for decreasing real-time software cycle execution time by increased concurrency and multiprocessing. This is of particular concern for this application, which involves large-scale networks and elaborate simulation-based algorithms. The principal mechanism for achieving smaller cycle times is to utilize a distributed processing implementation that significantly increases the concurrency of module execution and also exploits the inherent parallelism in the algorithms.

Special DTA Simulation Capability: Decentralized DTA Using Local Rules

The decentralized real-time traffic assignment framework is an alternative control approach that stands in contrast to centralized control architecture. The decentralized control envisions a set of local controllers scattered or distributed in the network, where every controller can extract only limited "row" information (speed, travel time, concentration, etc.) from network detectors; it utilizes this information using local control "rule" to guide the within-territory vehicles to their respective destinations (Ref 42).

The local rules represent the logic of the distributed system responsible for the assignment. Local control units communicate with equipped vehicles under their rule only (in territory). Territory size is primarily governed by the processing capabilities (e.g., memory resources) of the control units. Local control rules use available partial information to evaluate alternative subpaths emanating from the decision node *(i)* towards the destination *(j)*, and assign vehicles at node *(i)* among the links immediately downstream. The spatial extent of the local area ruled by controller *(i)* is highlighted as the shaded circle in Figure 3.5 (Ref 42). Only current traffic measurements (e.g., travel time, concentrations, etc.) are extracted by detectors installed in the local area. For clarity and uniformity in presentation, the local area is represented by the set of links (and nodes) with depth less than or equal to a prespecified knowledge level (denoted by *K*). This level is basically governed by the available technology of the control units (computational facilities and resources), the level of available investment, and the desired accuracy.

Decisions are reached by control units after considering the relative merit or disutility of alternative subpath options, as reflected by local and nonlocal state variables. For any origin-destination pair *(i,j)*, there exists a finite number of subpaths that can be evaluated and ranked individually on the basis of the expected disutility imposed on the system as a result of assignment a vehicle to the subpath.

Three rules are defined in connection with the noncooperative structure as follows: Rule 1: Only travel time estimates are used to evaluate the subpath and all-or-nothing assignment is performed; Rule 2: Similar criteria, but different assignment procedures that split vehicles among a set of subpaths; Rule 3: Generalized evaluation function of multiple-state variables and a splitting assignment procedure; Rule 4: Defined for cooperative structure. This decentralized approach to route diversion and guidance has been shown to be very responsive under unexpected incident conditions, and will be illustrated in the numerical experiment.

Model Limitations

DYNASMART is able to track the movement and location of individual vehicles, as already mentioned. However, it does not consider microscopic maneuvers such as car following and overtaking. These are of concern when examining the detailed performance aspects at a problem location such as a diamond interchange — hence the use of the TEXAS model in a complementary role to DYNASMART for this study. DYNASMART provides a number of network performance measures, as well as detailed output information on a variety of performance measures that can be aggregated according to the needs of the specific study. This requires customized postprocessing of the output files in order to aggregate the performance measures needed for a particular type of investigation. Special capabilities to produce corridor-level and component-level measures of effectiveness have been built into DYNASMART specifically for the needs of the present study.

3.3 DIAMOND INTERSECTIONS: TEXAS MODEL

TEXAS Simulation Model

The TEXAS Model for Intersection Traffic is a powerful computer simulation tool that allows the user to evaluate in detail the complex interaction among individual driver-vehicle units as they operate in a defined intersection environment under a specific type of traffic control (Refs 53, 51). In its current version (Version 3.0), the model has the ability to handle not only vehicular traffic at a single intersection, but also traffic moving through the closely spaced, at-grade intersections of a diamond interchange. Traffic control simulated by the TEXAS model includes unsigned, yield sign, stop sign, and signalized control (pretimed or actuated).

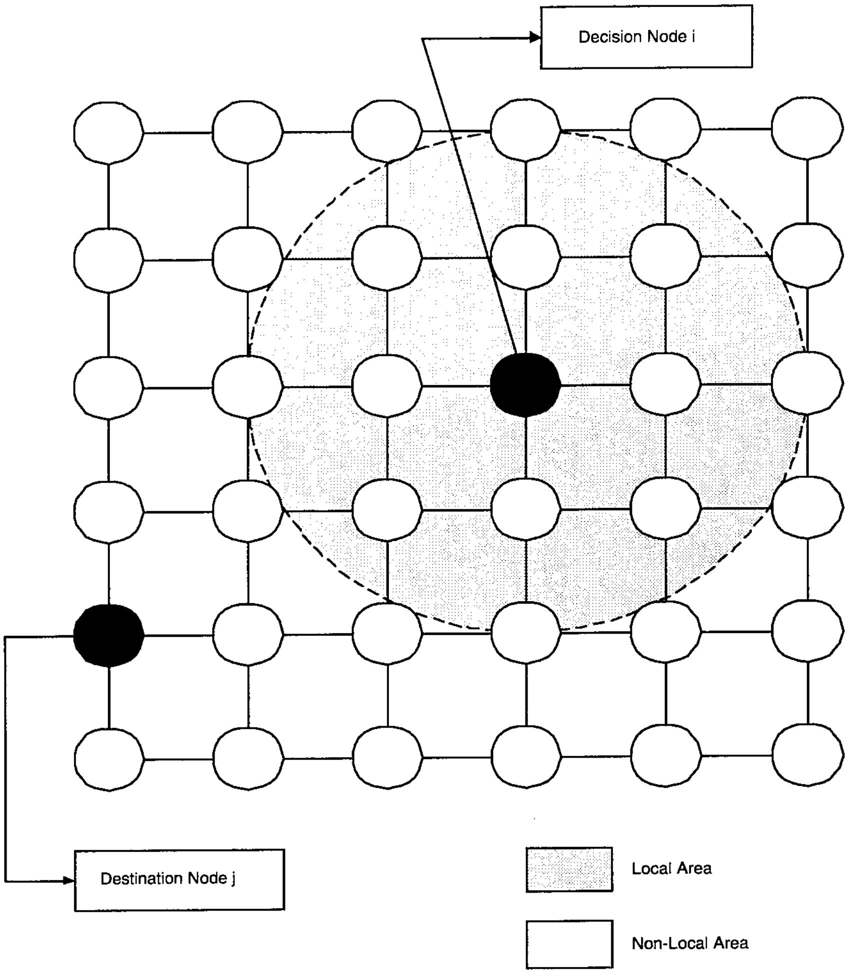


Figure 3.5: Local and nonlocal area in simple network

Model Structure

The TEXAS Model for Intersection Traffic includes four data processors: GEOPRO (Geometry), DVPRO (Driver-Vehicle), SIMPRO (Simulation), and DISPRO (Display) for describing, respectively, the geometric configurations, the stochastically arriving traffic, the behavior of traffic in response to the applicable traffic controls, and the animated graphics of the traffic.

GEOPRO develops a geometric definition of the interchange as specified by the user. DVPRO assembles the arriving traffic by utilizing assigned characteristics for each driver and vehicle class and generating attributes for each individual driver-vehicle unit. Each unit is identified by inputs for driver class, vehicle class, desired speed, desired outbound intersection leg, and lateral inbound lane position. SIMPRO processes the traffic behavior of each unit in response to the momentary surrounding conditions of traffic control device indications, adjacent traffic, and geometric features that might be applicable. DISPRO develops the animated graphics of the simulated traffic.

Geometric Terminology

The TEXAS model utilizes a specific numbering system to describe the various geometric characteristics of a diamond interchange, such as the intersection legs, lanes, and approaches. This system must be used consistently when inputting data or when interpreting statistical output files. This report also refers to the same terminology when identifying lanes or approaches. The numbering system associated with the geometry of diamond interchanges as used by the TEXAS model is described in Figure 3.6. The numbering of the lanes and legs are used mainly by the user only to input data. However, in order to correctly read the statistics files and interpret the output in this report, it should be noted that approaches 3 and 8 refer to the inbound frontage road approaches of the interchange. The inbound arterial street approaches are shown as approaches 2 and 7.

Output

The output from diamond interchange simulation regarding the performance of each driver-vehicle unit and the assignment of green time to each phase and phase combination are gathered during simulation and presented in summary form at the end of each run. Delay statistics that are collected include the total delay, queue delay, and the stopped delay incurred by all processed vehicles. Each delay type is summarized by movement and by approach. Total delay is the difference between travel time for a vehicle through the system and the time it would have taken the vehicle at its desired speed. Stopped delay is the time spent by a vehicle that has a velocity less than 3 feet/second. Delay statistics show the overall influence of the intersection environment on traffic passing through the intersection. Queue length statistics include average queue length and maximum queue length. Statistics are also collected and summarized with respect to vehicular speed, acceleration, travel time, and the number of vehicles processed.

Additional output from the TEXAS model used by the animated graphics processor includes the instantaneous speed, location, and time relationship for every simulated vehicle. Intersection geometry is extracted from the input files and displayed on the screen. The position of each simulated vehicle is represented on the screen by an outline of the vehicle, scaled to size and color coded according to performance capability, with respect to time. With this animated graphics display the user can study overall traffic performance of an interchange and examine in great detail the behavior of an individual vehicle in the traffic stream.

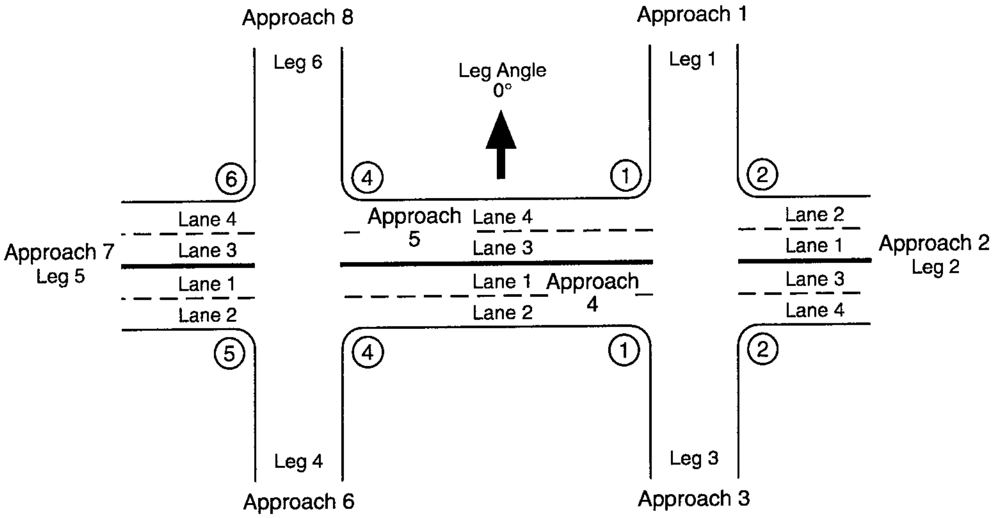


Figure 3.6: Diamond interchange nomenclature and terminology

Model Limitations

The TEXAS model provides an effective tool for analyzing traffic behavior at an isolated intersection or interchange that would otherwise require time-consuming, conventional field study techniques. Yet users should be aware of the limitations associated with this simulation tool; such an awareness can ensure more effective application of the model and better interpretation of the output statistics.

The TEXAS model is intended primarily for use under average steady-state traffic conditions; it is not designed to simulate dynamic traffic control strategies, which limits its ability for the portions of the research where dynamics are of the essence (such as incident conditions). Before an incident occurs, there is a base case traffic volume scenario for a diamond interchange. When a freeway incident occurs, traffic will begin to divert to the frontage road and through the diamond interchange system. If the interchange traffic control is not changed instantaneously to accommodate this sudden increase in volume, queues will develop. These queues, present at the beginning of the traffic control strategy implementation, cannot be simulated by the TEXAS model. The model begins each simulation without any existing queues. Therefore, this study must assume that the change of traffic control at the diamond is quick, if not instantaneous, so that no queues have developed. Another model will need to be used to simulate the development of the queues during any delay of a control strategy change and the rate of dissipation of the queues when a strategy is implemented.

Several other limitations of the TEXAS model result from the underlying logic embedded in the simulation. The traffic simulation attempts to replicate the actions of actual drivers. Every simulated driver is able to logically predict his/her own as well as other vehicles' positions and velocity for each time interval. The driver-vehicle units are also programmed to obey the following traffic laws and regulations:

- Drivers will not jump the green signal.
- Drivers will not block other vehicles that already have the right-of-way.

In the real world, such traffic rules may not necessarily be respected at all times by every driver. Also, many decisions made by some real-world drivers will differ from those programmed into the model.

3.4 SIGNALIZED FRONTAGE ROADS AND ARTERIALS

FRONTAGE ROAD: TRANSYT-7F

Simulation Using TRANSYT-7F: TRANSYT-7F is a version of a network signal timing optimization program that was first developed in 1967 by Robertson (Ref 43). It represents travel with a macroscopic time-based simulation model that tracks platoons of vehicles in small time increments and describes platoon size changes using a platoon dispersion algorithm. Although the program can be successfully used for single intersections, its macro-simulation nature is designed to enable efficient network signal timing optimization.

Principles of TRANSYT-7F: Four major and rather restrictive basic assumptions are made by TRANSYT-7F :

- All major network intersections are signalized.
- The proportion of left- and right-turn movements is the same for every signal cycle.
- The cycle length for all network signals is the same or one-half this value.
- A constant entry rate is specified for each location where traffic enters the network.

Network Representation: The program has two main elements: a traffic model and an optimization routine. The traffic model is totally deterministic and calculates, for any given signal settings, the expected behavior of the vehicles as they pass through the network. The network is represented on a link-by-link basis. Each intersection is represented by a node and each approach by a link (Figure 3.7). The user defines the order in which the links of the network are numbered. TRANSYT-7F evaluates traffic behavior link by link from upstream to downstream. Computed traffic patterns of upstream links are used to describe traffic patterns for the next links. Another important concept is the dispersion of moving platoons. This dispersion process clearly represents real-life driver behavior as drivers attempt to attain unique desired speeds. The performance index (PI) that is discussed in the next paragraph measures the overall impedance to traffic.

The optimization routine gives the optimum cycle length and signal offsets minimizing the performance index. First, a simulation-based evaluation of a user-specified range of cycle lengths is made in terms of the performance index and employing a user-specified cycle length increment. The program suggests the optimum cycle length in terms of the performance index and identifies saturated links.

Performance Index (PI): The performance index is a weighted linear sum of delay and stops and, optionally, a queue size penalty or existing operating costs. It is defined as follows:

$$PI = \sum_{i=1}^{n} \{[w_{di} d_i + k w_{si} s_i] + u_i [w_{di} - d_{i-1} + k w_{si-1} s_{i-1}] + B_i [w_q (q_i - c_i)^2]\}$$

where

w_{di} = link specific weighting factor for delay for link *i*,

w_{si} = link specific weighting factor for stops for link *i*,

w_{si-1} = link specific weighting factor for stops for link *i*-1,

w_q = a network wide "penalty,"

d_i = delay on link *i* in veh-hr,

d_{i-1} = delay on link *i*-1 in veh-hr,

k = a coefficient specified by the user expressing the importance of stops relative to delay,

s_i = stops in link *i* in stops/sec,

s_{i-1} = stops in link *i*-1 in stops/sec,

B_i = 1, if link-to-link weighting has been established, 0, elsewhere,

q_i = computed maximum queue "capacity" for link *i*, and

c_i = maximum queue "capacity" for link *i*.

The optimization procedure used is referred to as a "hill climbing" technique. It is an iterative, gradient search technique that requires one macrosimulation run for each evaluated timing condition.

Platoon Dispersion: Drivers with unique desired travel speeds determine the operation of individual vehicles in the network. Owing to this speed heterogeneity the compact platoon pattern that leaves every signalized intersection will be modified as the platoon traverses a downstream link.

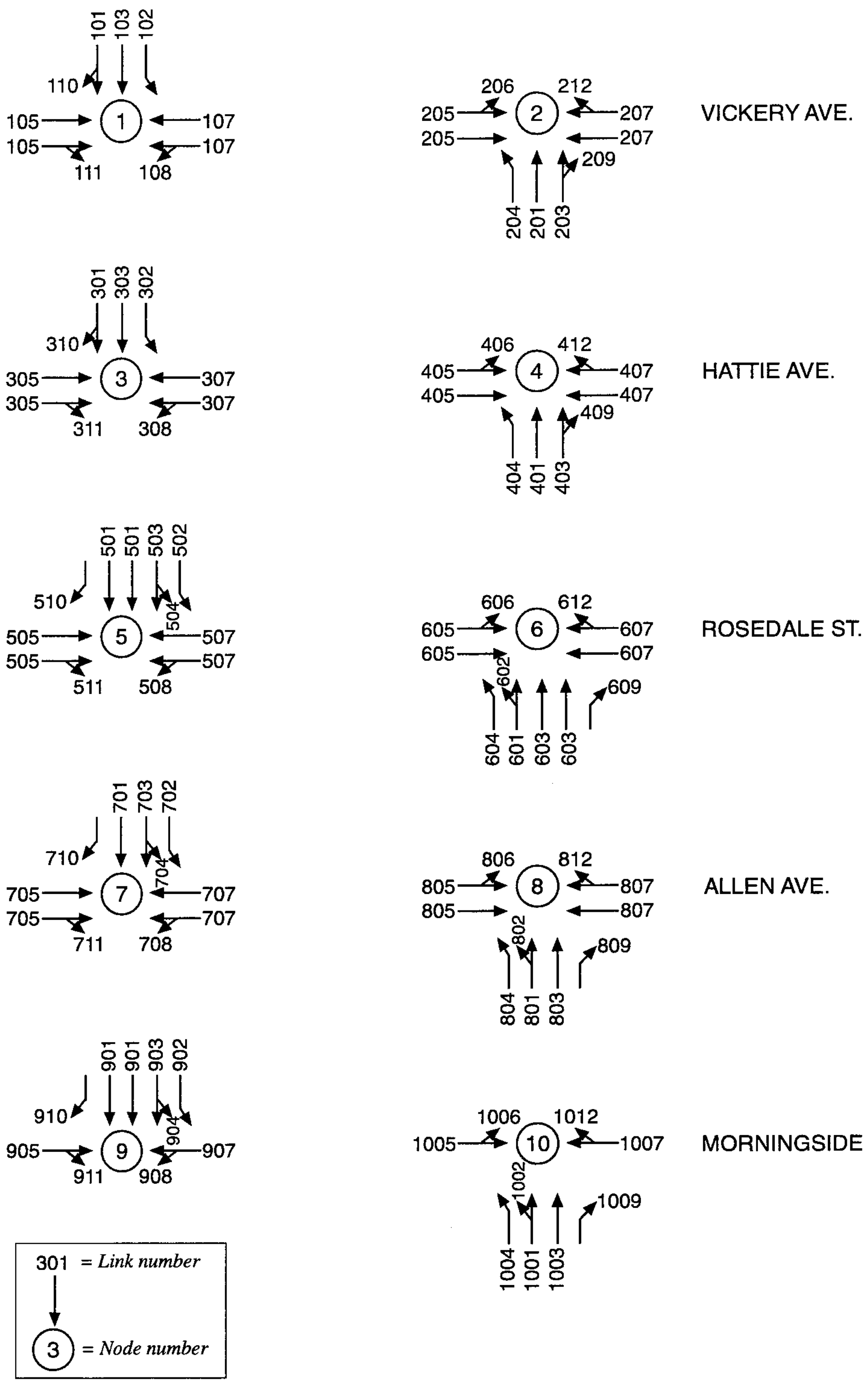


Figure 3.7: Diamond interchanges representation with links and nodes

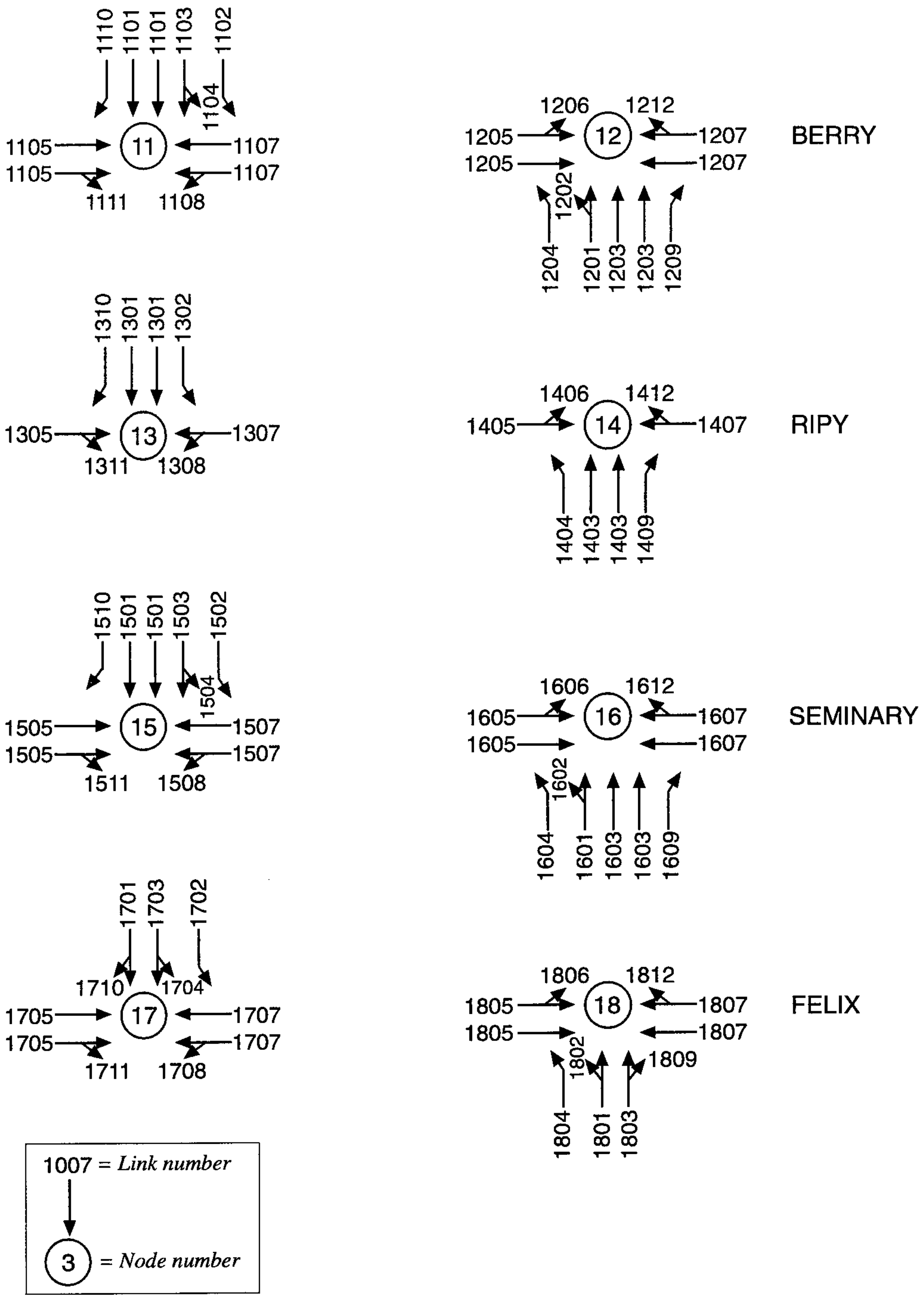


Figure 3.8: Diamond interchanges representation with links and nodes

Road Research Laboratory researchers at four sites in West London observed over 700 platoons. The method that TRANSYT-7F uses to predict dispersal of an average platoon is based on these observations. A smoothing factor F is used to fit the actual and the calculated platoon shapes. This smoothing factor is related to T (where T is 0.8 times the link cruise travel time) by the expression:

$$F = \frac{1}{1+\alpha T}$$

where α is an empirically derived constant.

British experience suggests a value of α of approximately 0.4; however, researchers have found that for the U.S. urban environment, $\alpha = 0.35$ is more adequate. The α value may vary as the user considers factors like grades, degree of curvature, parking, and opposing traffic flow. Robertson in 1969 (Ref 43) clearly outlined the importance of the platoon dispersion patterns used by the program.

The inclusion of a model of traffic dispersion means that the overall solution automatically takes into account the importance of having good progression on short links.

Measures of Effectiveness: As discussed previously, the program represents travel using vehicle platoons where movements are simulated using small time increments and a platoon dispersion algorithm. Among the measures of effectiveness used are average and total delay in seconds, numbers of stops, and fuel consumption.

Average Delay: In traffic studies, average delay is one of the most important measures of effectiveness. It can be considered a surrogate for costs of excess fuel consumption and time loss. A primary task of this research study is to minimize average delay through an efficient signal timing plan.

The well-known Webster delay model (Figure 3.9) estimates average total delay per vehicle on an intersection approach. On the other hand, the Highway Capacity Manual (HCM) (Ref 37) estimates only stopped delay per vehicle for the subject lane group. However, empirical evidence suggests that, on the average, approach delay is about 30% higher than stopped delay. Thus, one would expect that the HCM delay model will yield lower delay estimates in comparison with the Webster model.

TRANSYT-7F uses a method that is similar to Webster's model, expressed as:

$$D = \frac{C(1-\lambda)^2}{2(1-\lambda x)} + \frac{x^2}{2Q(1-x)} - 0.65\left(\frac{C}{Q^2}\right)^{1/3} x^{(2+5\lambda)}$$

where

D = average approach delay in seconds,

C = cycle length,

λ = portion of cycle that is effective green for this approach,

Q = traffic volume in veh/sec,

x = degree of saturation (Q / S), and

S = saturation flow in veh/sec.

Since Webster's original delay model yields reasonable estimates for degrees of saturation up to 95%, but is not applicable when saturation flow reaches 100%, the TRANSYT-7F delay model uses a modified version of the Webster model that gives better results where flows reach saturation levels. The relationship between Webster's original and the TRANSYT-7F model for delay versus degree of saturation is shown in Figure 3.9.

The program's delay estimates have been shown to properly match field measurements when the degree of saturation is lower than 100%. In this study the delay estimates seem to be reasonable for degrees of saturation up to 105%; however such estimates must be used very carefully. When the degree of saturation exceeds 105% the delay estimates are descriptive of extreme congestion and delay magnitude is near observation duration.

The HCM defines the level of service (LOS) in terms of the stopped delay per vehicle (column 2, Table 3.2). Average total delay can be related to stopped delay as follows:

$D = 1.3 * d$, where D is the average total delay in sec/veh and d is the average stopped delay in sec/veh.

The strategies and the signal timing provided for the diamond interchanges network have a goal of providing LOS E or better. Thus, all the average stopped vehicle delays must not exceed 60.0 seconds.

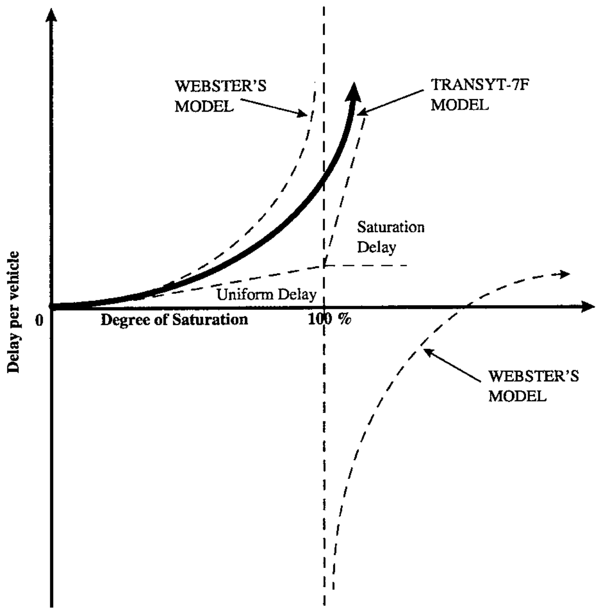


Figure 3.9: Comparison of TRANSYT-7F and Webster's delay models (Source: Ref 54)

Table 3.2: Level of service criteria for delay (Source: Ref 25)

Level of Service (LOS)	Stopped delay (sec/veh)	Average Delay (sec/veh)
A	less than 5.0	less than 6.5
B	5.1 to 15.0	6.6 to 19.5
C	15.1 to 25.0	19.6 to 32.5
D	25.1 to 40.0	32.6 to 52.0
E	40.1 to 60.0	52.1 to 78.0
F	greater than 60.0	greater than 78.0

Number of Stops: Average delay is related to the number of vehicle stops, which is also computed in the model. However, a filtering algorithm is used to separate momentary stops (delayed vehicles) from longer-term or effective stops. TRANSYT-7F assumes that delayed vehicles are also stopped. The percentage of stopped vehicles is related to the duration of delay (Figure 3.7). When the degree of saturation is close to 100% the stops estimate is also close to 100%.

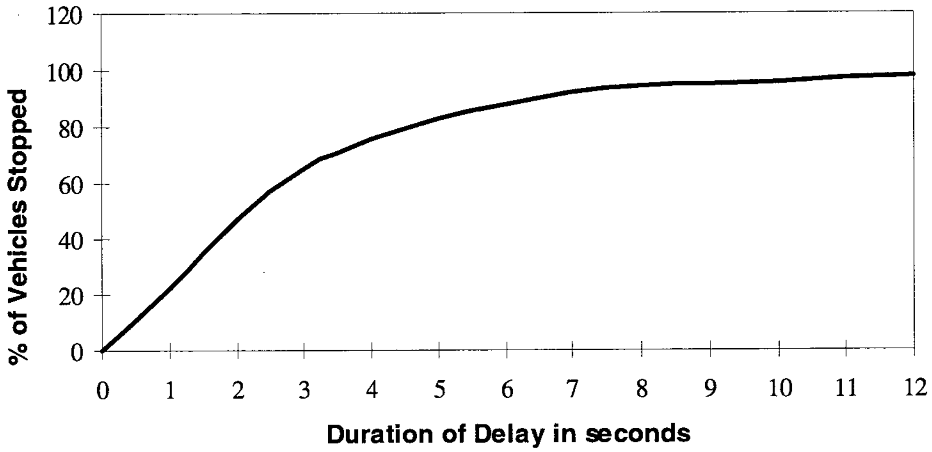


Figure 3.10: Relation of number of stops with the duration of the delay

Fuel Consumption: Vehicle stops and total delay play a significant role in determining fuel consumption. The program's fuel consumption model, based on experimental studies, is the following:

$$F = k_1 TT + k_2 D + k_3 S$$

where

F = fuel consumption per hour,

TT = total travel time per hour,

D = total delay,

S = total stops, and

k_1, k_2 = coefficients of regression equation calibrated to the 1977 vehicle fleet mix.

Since the *k's* represent the 1977 type vehicles, the fuel consumption estimates are probably overestimated.

Output: The output data provided by TRANSYT-7F contain a brief input data report with the number of links, input traffic volumes, and the saturation flow specified by the user. The measures of effectiveness, including average and total delay and the number of stops, are printed for each link, subtotaled by intersection and for the whole network. The timing settings (pretimed) for each intersection and a cycle evaluation summary for the specified cycle lengths are also printed. Time-space diagrams, platoon progression diagrams, and flow rate histograms are also available.

Summary: TRANSYT-7F can be used to evaluate several four-, three-, or two-phase control schemes for the study network's diamond interchanges. The program to obtain the optimum cycle length seeks to minimize a weighted linear factor (performance index) that combines delay and stops. Among the measures of effectiveness used are average and total delays in seconds, numbers of stops, and fuel consumption. The network simulation results for the base case are presented in Chapter 4.

ARTERIAL STREETS: PASSER II, CORSIM

CORSIM: Several computer simulation models are available, but very few can integrate the operations of street networks and freeways, or the flow of traffic on and between them. While CORSIM was intended to achieve this, it succeeds only partially.

CORSIM (CORridor SIMulation), a microscopic simulation model developed by the Federal Highway Administration, is a component of the TRAF family of traffic models. The TRAF software system includes NETSIM, for the simulation of signalized arterial networks, and FRESIM, for the simulation of freeways. CORSIM combines both models by the inclusion of interface nodes. In CORSIM, a vehicle enters the network on an arterial or freeway segment through an entry node. If a vehicle then wishes to cross from one subnetwork (NETSIM or FRESIM) to the other, it will do so by passing through an interface node.

Specifically, interface nodes represent a point on a ramp that connects a frontage road to a freeway lane, which is where the vehicle crosses from the street network to the freeway (or vice versa).

It is important to note that CORSIM became available in 1995 and was, at the time of this study, still in beta testing. Therefore, the user experienced some difficulties, but the program performed well overall.

Input Data Description: To input the network data into CORSIM, the first step is to build a network of links and nodes to describe the geometrics of the network. The links represent the connecting roadways and the nodes represent either an intersection or a change in the geometry of the roadway. An upstream and a downstream node identify each link.

The characteristic data of the links and nodes necessary to be input into CORSIM include:

1. Geometric data: Number of lanes of the roadway, their length, channelization, direction, left-turn and right-turn bays, grade, and pavement condition.
2. Operational data: Type of control at each intersection, volumes, desired free-flow speed, turn movements, and transportation mode.

These characteristics may be input into a text file — the input file — composed of a series of cards. (Note: Lines of input data are often referred to as cards, reflecting the technology of the seventies, when many of these programs were originally developed. Actual cards, of course, are rarely used today.) Each card describes a set of relevant characteristics for a specific link or node in the network. For example, one card is used to input the entry volumes, where the upstream and downstream node numbers of the entry link and the respective volume are written. Another describes the type of control at each node. Some cards are required to build the NETSIM subnetwork and others to build the FRESIM subnetwork, while still others are required for both and others are optional, depending on the user's needs.

The program reads each card and takes the relevant information to build the network by recognizing that certain characteristics in the specific card it is reading should be under a certain column. For example, in card type 11, for the input of the NETSIM subnetwork geometric data, columns 1–4 are reserved for the upstream node. When it reads the card, the program will interpret the number found under columns 1–4 as the upstream node of a particular link.

CORSIM contains a preprocessor that meticulously checks the input file for errors before performing a run. These subroutines flag any apparent coding error made by the user.

ITRAF: The Federal Highway Administration has designed a Windows-driven input module for any of the TRAF components called ITRAF. This software provides a user-friendly alternative to building an input file on a text editor, where much attention must be paid to placing the data in the correct columns and where there is no graphical interface for the network being built to guide the user. With ITRAF, a user can graphically build the network and, with the aid of a mouse, click on all the links and nodes to input the necessary attributes. The program is designed to disallow as input any characteristics that are considered fatal errors by the CORSIM preprocessor.

When the file is saved, ITRAF stores it as a DOS text input file, which can therefore be read directly by CORSIM or by the TRAF software being implemented. ITRAF can also interact with any text file created for TRAF software even if it was not originally created in ITRAF.

Since ITRAF is still a prototype, it is subject to occasional fatal coding errors, so some debugging will still be required. Nevertheless, this software can still effectively be used to create a great portion of the input file, followed by the use of a text editor to clear any remaining coding errors.

Time in CORSIM: Time in CORSIM is defined by periods, intervals, and steps. A time period describes the amount of simulation time in which the input characteristics drawn by the user is valid. If the user wishes to simulate one certain input stream for 1800 seconds and then add an HOV lane to the network and simulate it for 1200 seconds, for example, then two time periods should be specified: the first for 1800 seconds and the second for 1200 seconds. CORSIM permits the user to simulate up to 19 time periods with different durations in one run.

Each time period is divided into a sequence of time intervals. In each time interval each of the two subnetworks (NETSIM and FRESIM) is called out of central memory and run once. The manual recommends setting the time interval to the most frequent signal cycle length in the network. In addition, the time period specified must be an integer multiple of the time interval to be used. The FRESIM model partitions each time interval into time steps and is input in tenths of a second. The time interval must be a multiple integer of the time step used.

CORSIM Simulation: When the input file has been debugged of all apparent coding errors, a simulation run begins. The first step in the simulation process is the initialization. During the initialization, vehicles enter and exit the network until steady state is established on both subnetworks or until the initialization time specified by the user in the input file is exhausted. Subsequently, the simulation for the first time period begins, running until the specified simulation time period is exhausted. If a second time period is specified, the same process is repeated with the corresponding input.

Note that when defining incidents to be simulated, a study performed in North Carolina using TRAF-FRESIM reported that the incident definition feature does not wait until the initialization time is completed to start the incident duration. The user must therefore exercise care in defining an incident as having sufficient duration time in the respective input card to account for initialization time when simulating an incident lasting the entire time period (as in a work zone lane blockage).

CORSIM Outputs: A great variety of information is generated during and after a successful simulation run. This includes the input file, embedded data, vehicles missing destinations and the time of occurrence, spillbacks with their starting and ending times, and a wealth of NETSIM- and FRESIM-specific movement and measure of effectiveness tables. Much can be learned from this output information. Problematic areas can be pinpointed to provide some solution, such as signals with poor timing that produce spillbacks, and weaving areas having poor levels of service.

TRAPHIX, a graphical output module, has also been developed by the Federal Highway Administration. TRAPHIX is a Windows-driven software program that can produce animation of a successful simulation run. It is helpful in graphically viewing and understanding the behavior of the traffic and the problems it encounters in a flow. It takes a considerable amount of time to convert the output file to a TRAPHIX animation file, in relation to a CORSIM simulation run.

Model Limitations: The limits in the size of the network and other input variables in a single CORSIM run are shown in Table 3.3.

Table 3.3: CORSIM limits for input variables

SIZE LIMITATIONS OF NETWORK CHARACTERISTICS

Characteristic	NETSIM	FRESIM
Nodes	250	120
Links	500	200
Vehicles	10,000	10,000
Buses	256	256
Bus Stations	99	N/A
Bus Routes	100	100
Actuated Controllers	100	N/A
Detectors	300	200
Detector Data Stations	N/A	50
Events	200	N/A
Incidents	N/A	20
Disjointed Freeway Segments	N/A	10
Entries per Segment	N/A	35
Through Lanes	7	5
Auxiliary Lanes	N/A	6

PASSER II-90: To input the signal timing on all signalized intersections in CORSIM, optimization of these signals was preferred to provide better flow through the network. PASSER II was designed for this purpose. PASSER stands for *Progression Analysis Signal Systems Evaluation Routine.* This signal timing optimization software developed by the Texas Transportation Institute (TTI) belongs to a family of signal optimization programs comprised of PASSERs II, III, and IV. PASSER II optimizes signal timing on an entire arterial street by maximizing bandwidth efficiency. It computes the cycle lengths and splits for street intersections using Webster's equation:

$$C_o = \frac{1.5\,n\,l + 5}{1 + \sum_{i=1}^{n} y_i}$$

where C_o is the cycle length that results in minimum delay, n the number of phases, l the lost time per phase, and y_i the flow ratio in phase i. PASSER II cannot optimize a closed network of signals.

Because PASSER II is not a microscopic simulation program, it does not therefore track vehicles. It uses an algorithm to calculate the probability of queue clearance and the percentage of green time falling within the green band in each direction. It then selects the best performing bandwidth from the possible phasing patterns being tested. The bandwidth efficiency is measured as the sum of both through bands divided by twice the cycle length (Ref 44).

Data Input: PASSER II includes an input/output module, PASSETUP, which provides menus that take the user through the steps necessary to run PASSER II. The input menus allow the user to easily input the necessary data for analysis. The user must provide a minimum and maximum cycle length, permitted phasing options, left-turn phase protection type, hourly volumes, presence or absence of left-turn bays, and distances between intersections. The program has features to assist the user in the calculation of saturation flow and pedestrian crossing time. It uses the NEMA (National Electronics Manufacturers Association) movement numbering convention for identification of movements. The program creates a data file of the input characteristics, which it uses to execute its functions.

Output: After a successful run, an output file is created. The selected optimal phase pattern sequence is displayed using NEMA conventions. The intersection performance is measured in terms of volume/capacity ratios (v/c) and average delay. Average fuel consumption and total number of stops along the arterial are also displayed.

A LEART animation output module that animates the signal timing and vehicle movement along the arterial accompanies PASSER II. It should be kept in mind that since PASSER II does not track vehicles, the LEART animation is more of a progression display than a serious engineering tool.

Model Limitation: A total of twenty intersections can be analyzed in one data file, with two- to six-phase sequences being studied. The LEART animation module displays a maximum of eight intersections.

PASSER III: To optimize signal timing at diamond interchanges, PASSER III was used. This software, also developed by TTI, is capable of calculating signal timing plans for isolated interchanges or for connecting a series of interchanges along frontage roads. When calculating signal timing for several interchanges, PASSER III solves for bandwidth efficiency in a manner similar to that used by PASSER II.

Input Requirements: As in PASSER II, data are input following a series of menus. The data required to analyze interchanges in PASSER III are distances between signals, progression speed, traffic volumes, number of lanes and left-turn bays, interior travel times, and available storage.

3.5. TEST BED SELECTED FOR INTEGRATED SYSTEM DEVELOPMENT

Generalities

To articulate our conceptual framework, use the tools that have been identified, and effectively put together the strategies for integration, it is convenient to work in a specific context. Therefore, in the first stages of this study, a test bed was selected. The test bed offers the opportunity to apply all the control elements in the developed framework, and to test directly the integrated control strategies developed. Therefore, a test bed was selected after consultation with the project director at TxDOT. The test bed is located in the Fort Worth area on I-35W between I-20 and I-30. This location has all the features required for the study.

Test Bed Location and Description

Network Configuration and Characteristics: The study network shown in Figure 3.11 is subdivided in 13 zones and contains 178 nodes and 441 links. The network representation corresponds to one sector of I-35W between I-20 and I-30 in the city of Forth Worth, Texas. The freeway is represented in the middle of the network; the street network, including parallel and crossing arterial streets and local streets, is shown on both sides of the freeway. Zone centroids, shown as filled circles, define the destination of the trips; the origins are distributed all over the network, using as generation location the links on each zone.

Freeway nodes are connected to the street network through entrance and exit ramps. The links corresponding to the freeway, frontage road, and ramps are represented as directed arcs in one direction, while the rest of the links in the network are actually two directed arcs, one per direction.

The network accommodates 16,741 vehicles during a 35-minute period, representing a peak-period situation with high congestion. The first 5 minutes are considered a warming period to allow the network to be reasonably occupied, followed by a 30-minute generation of traffic, for which the performance statistics are accumulated. In terms of control characteristics, sixty-one intersections have actuated signal controls, and thirty-two intersections have stop signs. The remaining nodes have no signal control.

3.6. SIMULATION EXPERIMENTS DESCRIPTION

Situation of Interest

The general integration problem can be described as follows: Assume that users want to reach their respective destinations using the minimum travel time. Thus, they will need to know what the travel time would be using different possible paths in order to find the minimum one. After all the users in the network have followed a similar procedure, we can determine the volume of users at each controlled intersection, arterial street, or entrance ramp in the network. In this way, we can determine the optimal settings for each node in the network so as to optimize a certain performance measure. With those settings in mind, users can again calculate the travel time using different paths. Travel time along different paths will vary because of the new signal settings; the path taken can also change, changing in turn the traffic volumes. This circle of interdependencies is represented graphically in Figure 3.12.

In addition to the situation described, we need to account for time-dependent processes. While the dynamics of the system introduce additional complications, they also allow us to break the circle of interdependencies with time. Therefore, we can consider two complementary strategies: a reactive strategy applied to the neighborhood of the incident link that should start as soon as the incident is detected, and a proactive strategy that includes modifications in the paths of vehicles already in the network using information provided through VMS and other sources of information, as well as modifications in the paths of the vehicles ready to enter the network using all the information capabilities in order to improve the general network performance.

In suggesting a solution methodology for the integration problem, first a simulation of different cases is presented, then the main variables are determined, and finally a solution procedure is suggested. A set of simulation programs has been chosen for this activity, according to the desired characteristics in each case.

Set of Experiments Designed for Strategy Testing

Several groups of experiments were performed in order to analyze the integration of different subsystems to handle congested conditions. The next chapter is devoted to the description and analysis of all the experiment sets designed to test the different strategies considered in this study. The first set of experiments is devoted to the analysis of different information strategies, including descriptive information and normative information with route guidance and different compliance levels. The second group of experiments is focused on control strategies to handle frontage road and diamond intersections that are the main bottleneck for control integration. The third group of experiments explores different incident conditions to establish some typical cases that are modeled in the fourth set of experiments to test control strategies. Finally, the fifth set of experiments combines control and route guidance strategies, presenting the possible benefits that may be achieved using these kind of strategies.

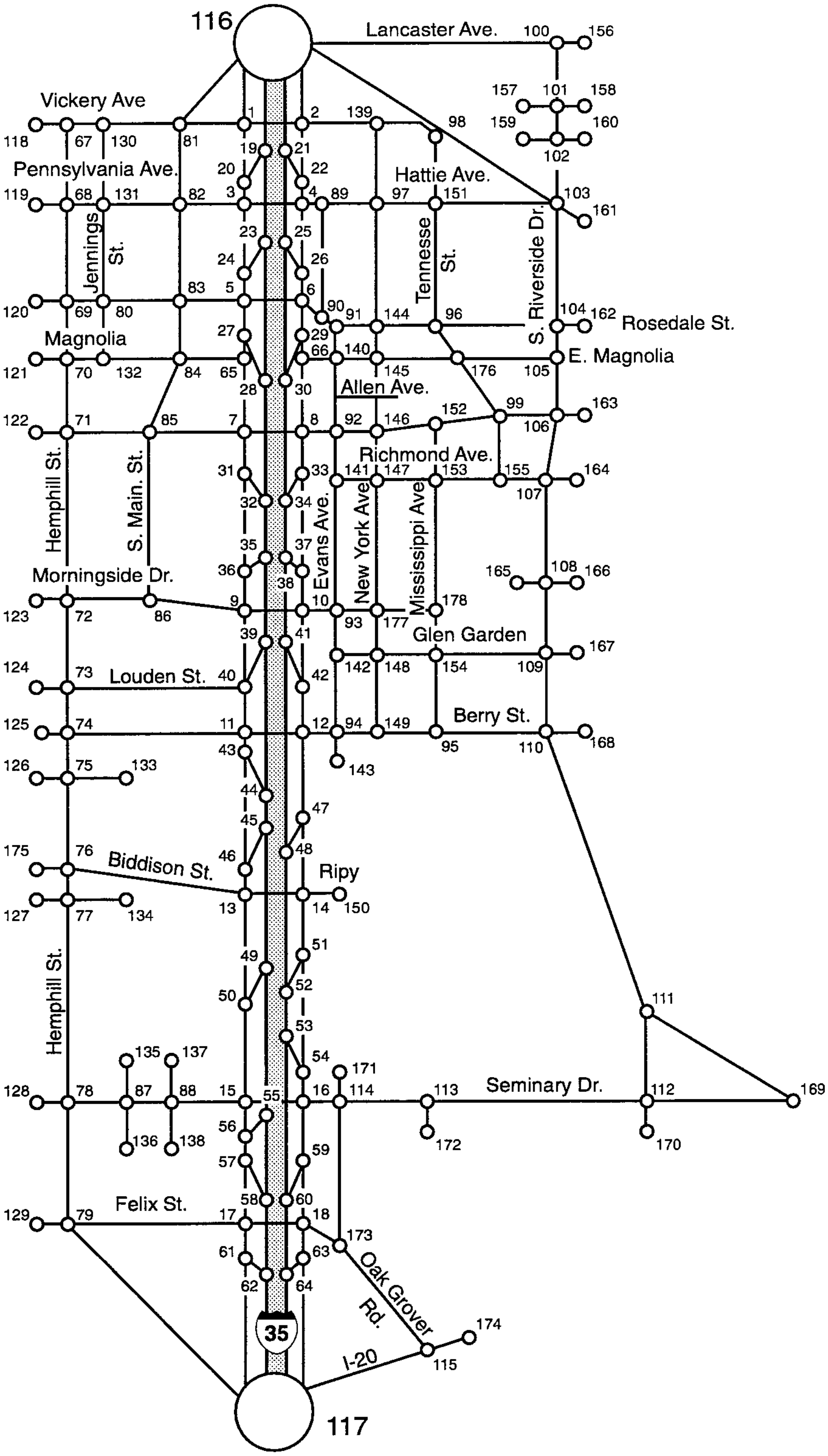


Figure 3.11: Network representation

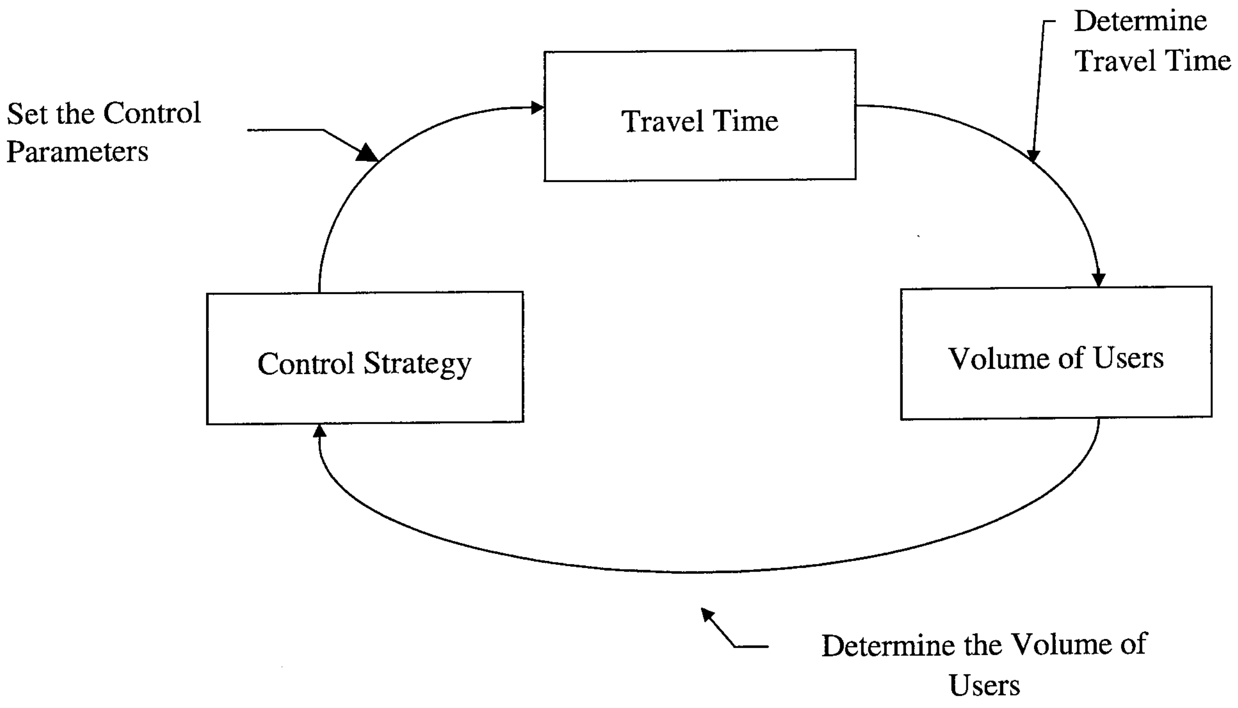


Figure 3.12: Interdependencies among travel time, volume of users, and control settings

CHAPTER 4. FRAMEWORK APPLICATION: SIMULATION EXPERIMENTS

As noted in the previous chapter, five sets of experiments have been designed and conducted in order to demonstrate various integrated control strategies in a corridor network. The Fort Worth test bed described in the previous chapter provides the context within which these strategies are tested. The results are nonetheless general in nature, given the design of the experiments and the representative character of the Forth Worth test bed.

4.1. THE FIRST EXPERIMENTAL SET

Experiment Description

The first set of experiments is intended to show the benefits that may be obtained by using information strategies, including route guidance, to manage incident situations. In this case, an incident is simulated on the freeway link 48-41 on the Fort Worth network (Fig 3.11 in previous chapter). The starting time is considered at minute 10 of the simulation and the ending time is at minute 20, i.e., for a duration of 10 minutes of the traffic interruption. The severity of this incident is set to 0.8, meaning that the blocking effect removes 80% of the capacity normally available on that link.

The strategies for route diversion (Dynamic Traffic Assignment, DTA) are as follows:

Descriptive Information (DES-DTA): In this case, information on prevailing trip times and associated current best paths is available and is provided to the users. The market penetration can vary and consequently different levels of information are modeled to consider this situation. To determine the advantages of this strategy, the base case is set to 0% level of information. To contrast with the base case, the levels of descriptive information considered are 25% and 75%.

Normative with Multiple-User Classes (NMUC-DTA): In this case, four different classes of users are considered. First, some users follow the same path all the time. Those are considered vehicles with prespecified paths. Second, some users receive information (route guidance) from the central controller and follow the given path that was found when solving for the optimum of the system. These are vehicles with SO (system optimum) paths. Third, some users receive information and follow the path that is the best for them individually. These are vehicles with UE (user equilibrium) paths. Fourth, some users receive information and process it, but switch routes only if the improvement in travel time exceeds a certain threshold or minimum level. These are referred to as "boundedly rational" users.

Normative with 100% Optimal Route Guidance or System Optimum (NSO-DTA): This case is used as a benchmark against which to compare the others. The SO solution assumes that all vehicles are guided along paths so as to minimize the total travel time in the network. A SO solution, by definition, is the best that one could possibly achieve in terms of overall performance. In this context, the vehicles are routed along the least marginal cost path to their destinations, paths that impose the least penalty on the system. This situation gives us an upper bound on the benefits attainable with real-time traffic information.

The experimental setup includes both situations without incident and with incident. The situations considered are presented in Table 4.1

Table 4.1: Experimental scenarios

Route-Diversion Strategy	**Characteristics**
	Scenarios without incident
DES-DTA 0%	Base Case: 0% information
NMUC-DTA	Four different user classes (25% of each class in the network)
NSO-DTA	100% route guidance
	Scenarios with incident
DES-DTA 0%	0% information
DES-DTA 25%	25% information
DES-DTA 75%	75% information
NMUC-DTA	Four different user classes (25% of each class in the network)
NSO-DTA	100% route guidance

Each particular scenario was simulated using DYNASMART, with the results then summarized. The performance measure used for analysis is the average travel time in the network.

Figure 4.1 presents the results for the situation involving descriptive dynamic traffic assignment only. The first group corresponds to the base case. This no-incident, no-information case constitutes the base case that is set for comparison purposes. The second group of values corresponds to the case with incident where 0% of the vehicles receive en-route information but where pretrip information is available to all users (e.g., through radio reports or cable TV channels). This means that vehicles departing after the incident has occurred are aware of the evolution at the incident location. The third group corresponds to the case with incident and 25% of the vehicles receive en-route information. The three columns indicate the performance of each group of vehicles: first those with information, then those without information, and, finally, the total combined for all the vehicles in the experiment. The fourth group corresponds to the case with incident, with 75% of the vehicles provided with en-route information. Also presented is the performance of each group, namely, those users that receive en-route information ("With Info" in the figure legend) and those that do not ("W/O Info" in the legend).

Figure 4.2 presents the results for the incident case with different information scenarios considering the descriptive and normative cases, including the multiple user classes. The first group corresponds to the descriptive information case, and shows cases with 0%, 25%, and 75% en-route information, respectively. The second group corresponds to the multiple-users classes (MUC), with 25% of each class including prespecified, SO, UE, and boundedly rational. The third group corresponds to the lowest possible travel time under the incident scenario. It corresponds to the case of 100% vehicles guided to follow SO paths, thus obtaining the optimum for the system under the incident and the (theoretical) benchmark against which to gauge the other strategies.

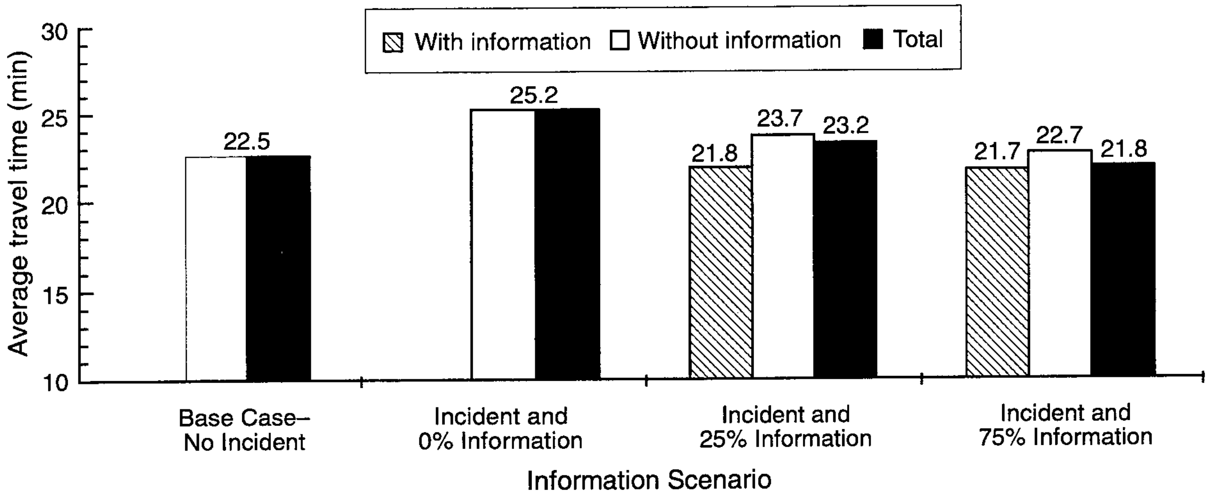


Figure 4.1: Travel time of the descriptive information scenarios

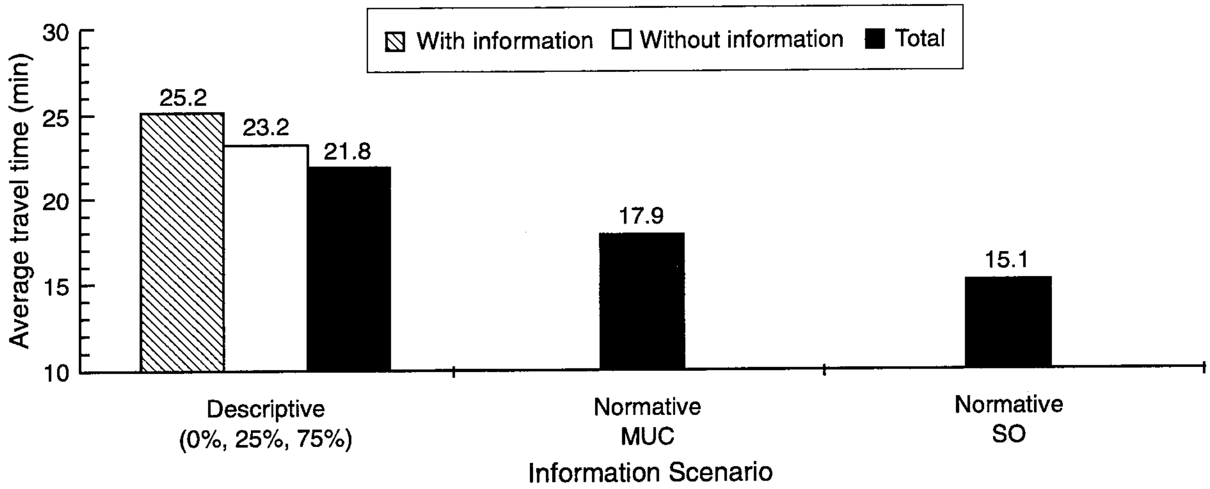


Figure 4.2: Travel time for different scenarios — Incident case

Figure 4.3 presents the results for the no-incident case with three different information scenarios: descriptive, multiple-user classes (MUC), and 100% SO route guidance. Since there is no incident in this case, these values are used for comparison purposes with respect to the incident case.

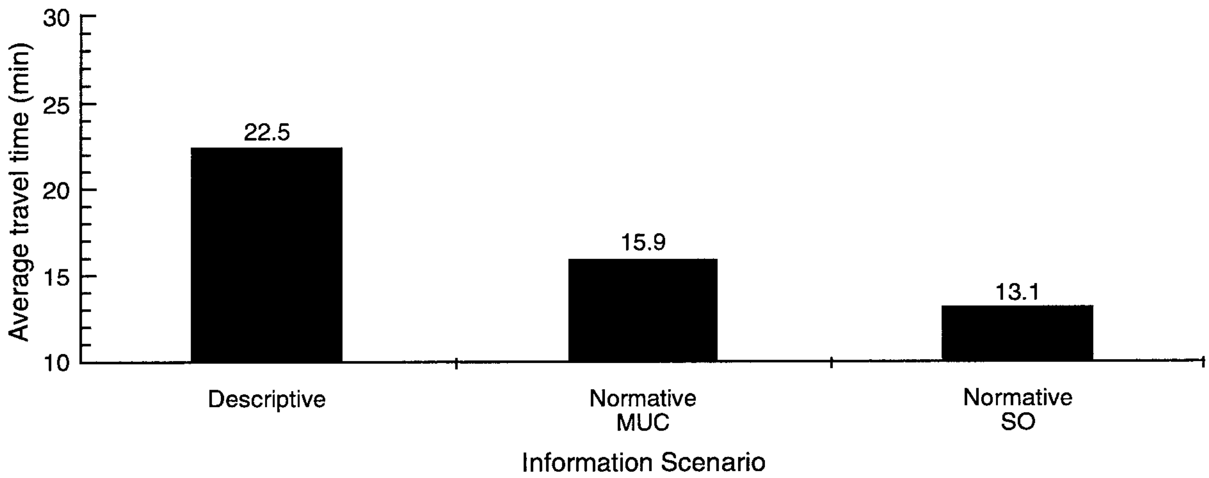


Figure 4.3: Travel time for different scenarios — No-incident case

Analysis

Several interesting insights may be drawn from the values presented in Figures 4.1 through 4.3. The analysis presented here considers not only the values in each figure, but also comparisons between values presented in the different figures. In all the cases, the performance measure used is the average travel time for the vehicles in the system. In applicable cases, separate values are reported for vehicles with information (i.e., equipped) and without information. In addition, the average travel time for all the vehicles in the system is considered in all cases.

Comparing the values presented for the descriptive information scenarios, including the base case, the incident cases are definitely worse than the base case, as expected. On the other hand, in case of incident, when 75% of the users receive descriptive information, the performance of the system is even slightly better than the base case. When only 25% of the users receive descriptive information, the performance of the whole system is not better than that under the base case, but there is improvement with respect to the case of no information, and the vehicles with information do slightly better than the vehicles in the base case (no incident). Of course, site-specific analysis is needed to determine what percentage of vehicles receiving descriptive information works best under particular conditions. In any case, when more vehicles have information, the users without information experience reduced average travel time.

In Figure 4.3, travel time under the normative SO (100%) case indicates how effective the system would be if all the users followed the route guidance supplied by a central controller that optimizes the travel time for the whole system. One interesting feature is that the normative MUC (multiple-user classes) scenario performs very well, close to the 100% SO benchmark, even though only 25% of each class is present in this case. This particular result has been mentioned in previous studies (Ref 55), but this is the first instance in which real networks confirm previous results obtained using hypothetical networks.

Comparing the results achieved under Normative 100% SO (NSO) route guidance information with those obtained under the DES 0% (Descriptive 0% information) information case reveals that the improvement in travel time possible through optimal route guidance is about 40% for both incident and nonincident cases. These results are suggestive of the benefits that might be achieved through optimal route guidance in the network.

Comparing performance under the MUC (multiple-user classes with 25% of each class) case with the DES 0% case, the improvement in travel time is about 30% for both incident and nonincident cases. These results also suggest that under any condition, most of the benefits of optimal route guidance can be achieved even though only a fraction of the vehicles may follow the routes provided through route guidance.

Table 4.2 presents a summary and the relative performance of the strategies considered with the incident case. (The same data are displayed graphically in Figure 4.2.) These results suggest that there is a 40% improvement in travel time under the route guidance strategy with 100%compliance. On the other hand, even when only a fraction of the users in the network are guided, an almost 30% improvement in travel time is achieved in this network.

Table 4.2 Relative performance of normative MUC vs. descriptive information scenarios — Incident case

Route Diversion Strategy	Characteristics	Average Travel Time (min)	Improvement in Percentage
DES-DTA 0%	0% information	25.2	Base
DES-DTA 25%	25% information	23.2	8.0 %
DES-DTA 75%	75% information	21.8	13.5 %
NMUC-DTA	Four different user classes (25% each)	17.9	28.8 %
NSO-DTA	100% route guidance	15.1	40.0 %

4.2 SECOND SET OF EXPERIMENTS: DIAMOND INTERSECTIONS AND FRONTAGE ROAD

This second set of experiments presents various incident scenarios. Freeway incidents can create massive congestion and traveler-time losses. Incident management techniques probably represent one of the most potentially beneficial traffic management concepts. Freeway incident management should include traffic diversion to parallel routes; where available, frontage roads represent a very desirable, convenient parallel path around an incident-blocked freeway section. The first set of experiments illustrated the potential benefits that could accrue through diversion achieved by the provision of real-time information. However, diverted traffic on frontage roads must cross arterial streets at signalized diamond interchanges that have, traditionally, not been tuned to accept diverted freeway flows. Simple modification of diamond control schemes to handle the local demand overflow can create massive network-level congestion problems unless frontage roads and crossing and parallel arterial streets are treated as integrated network elements.

Several freeway incident scenarios were implemented to evaluate network signal timing and to study the adaptability of diamond interchange traffic operations. In all incident scenarios the general traffic policy with regard to the I-35W corridor was to serve the maximum possible portion of traffic via the frontage road in order to minimize traffic congestion, air pollution, and noise on the urban network. The primary criterion for evaluating proposed signal schemes was average delay, although other measures of effectiveness were examined.

Incidents Affecting the Main Three Diamond Interchanges

Five different incident scenarios were examined, each involving one of the three major diamond interchanges of the test network. Traffic diversion to the frontage road was through existing exit ramps and diversion back to the freeway through existing interchange entry ramps. The incident location and the entry and exit ramps are presented in Figure 4.4.

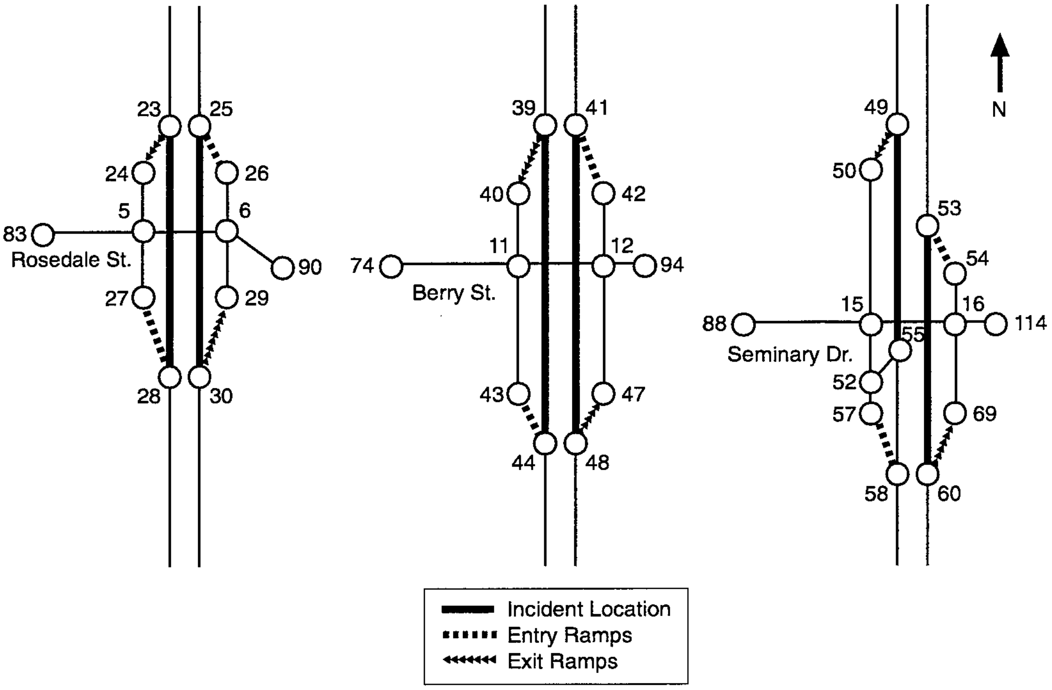


Figure 4.4: Incident locations and entry and exit ramps

The incident scenarios tested during this initial set of experiments were:

Scenario 1: Simulate a minor freeway incident by reducing the capacity by 25% northbound, blocking one lane of four (link 30-25 for Rosedale, link 48-41 for Berry, and link 60-53 for Seminary).

Scenario 2: Simulate a major freeway incident by reducing the capacity by 50% northbound, blocking two of four lanes (same links as in Scenario 1).

Scenario 3: Reduce freeway capacity by 25% southbound (minor incident), blocking one of four lanes (link 23-28 for Rosedale, link 39-44 for Berry, and link 49-55 for Seminary).

Scenario 4: Reduce freeway capacity by 50% southbound (major incident), blocking two of four lanes (same links as in Scenario 3).

Scenario 5: Simulate a major incident involving both freeway directions, reducing capacity by 50% southbound and northbound, blocking two of four lanes in each direction (same links as in Scenarios 1 and 3).

The diamond interchange phasing schemes considered were:

- Four phase without overlaps
- Three phase without overlaps
- Three-phase control that favors the affected side of the incident intersection (the right side when the incident is northbound and the left side when the incident is southbound). This phasing scheme was not implemented when there were incidents in both directions (Scenario 5).
- Three-phase control with prohibition of all left turns from the internal approaches. In this case the left internal traffic flows were added to the straight traffic flows.
- Two-phase control with prohibition of all left turns. Again, all left traffic flows were added to straight traffic flows.

Rosedale Diamond Interchange

Under incident Scenario 1 (minor incident blocking one northbound lane), the best timing scheme was the three-phase control favoring the right side of the diamond interchange (see Table 4.3a) and a 70-second cycle. Phase times and movements are tabulated in Table 4.3b. The four- and three-phase controls were found inadequate insofar as they yielded unacceptable average delays for the left internal westbound and for the external eastbound movements, respectively.

For incident Scenario 2 (major incident blocking two lanes), the optimum phasing scheme was the three-phase control with prohibition of internal left turns and a 70-second cycle (see Table 4.4a). Phase times and movements are presented in Table 4.4b. Straight and left northbound movements experienced higher delays (52.4 seconds), however this phasing sequence was superior to the four-phase control that yielded delays greater than 60 seconds for both the internal and external westbound traffic. On the other hand, the other three phasing controls gave estimates over 60 seconds for interchange average delay.

Under Scenarios 3 and 4 (blocking one or two southbound freeway lanes, respectively), the best scheme was again the three-phase control with prohibition of internal left turns (see Tables 4.5a and 4.6a) and a 60- or 65-second cycle. Phase times and movements are tabulated in Tables 4.5b and 4.6b. The three-phase control favoring the left side of the interchange could also be considered a possible solution, though the eastbound approach, which experiences a delay of 60 seconds, could be rendered inadequate if traffic demand increases even slightly.

Incident Scenario 5 (two freeway lanes blocked in each direction), created the heaviest interchange demand, since traffic flows increased in both directions. All diamond interchange management control steps described, through two-phase control, were required to achieve acceptable delay estimates at all internal and external approaches (see Table 4.7a). The optimum cycle length increased to 110 seconds. Table 4.7b gives phase times and movements for this scheme.

Berry Diamond Interchange

The results obtained for all five incident scenarios indicated that the four-phase system was capable of handling Berry diamond interchange traffic demand. Owing to the network configuration and to entry/exit-ramp locations, traffic demand did not exceed four-phase capabilities. A tabulation of traffic flows and average delay estimates (see Table 4.8a) showed that the eastbound external approach yielded delays close to 55 seconds under Scenarios 4 and 5. This value, considering the extremely heavy volumes generated by the incidents, was considered reasonable. Although the cycle length was long for incident Scenarios 4 and 5 (compared with that of the first three incident scenarios), it represented an 18% decrease over that obtained in the Rosedale incident Scenario 5 study. The best cycle length varied for each incident scenario from 65 to 90 seconds. Phase times and movements are tabulated in Table 4.8b.

Seminary Diamond Interchange

For all five incident scenarios, as in the Berry interchange case, the four-phase scheme was found to be sufficient for the Seminary case study. Table 4.9a indicates that for all cases and for all approaches, the average delay estimates were less than 45 seconds. The best cycle length varied from 65 seconds for incident Scenarios 1 and 3 to 100 seconds for incident Scenario 5. The approaches that experienced highest delays were eastbound and westbound approaches. That occurred because TRANSYT-7F gave more green time to northbound and southbound approaches, where the demand for the straight movement increased (see Table 4.9b for phase times and movements).

Table 4.3a: Rosedale diamond interchange: Incident northbound, one lane blocked. average delays in seconds for different phasing schemes

Direction	Movement	Flow (veh/hr)	4-phase control	3-phase control	3-phase control favor right side	3-phase control no left turns	2-phase control
Southbound	Left	368	27.9	34.1	46.5		
	Straight	643	27.9	34.1	46.5		
	Right	10	19.7	22.9	24.4		
Northbound	Left	600	15.2	28.9	31.8		
	Straight	1160	15.2	28.9	31.8		
	Right	275	12.8	21.2	22.4		
Eastbound	Straight	948	14.4	**>60**	38.6		
	Right	56	14.4	**>60**	38.6		
Westbound	Straight	658	24.7	37	36.9		
	Right	54	24.7	37	36.9		
Internal Eastbound	Left	322	29.6	10.1	25.5		
	Straight	994	11.4	12	29.4		
Internal Westbound	Left	941	**>60**	17.6	9.9		
	Straight	317	7.1	41.7	14.1		
Best Cycle Length (sec)			60	70	70		
Interchange Average Delay			29.35	32.11	29.85		

Table 4.3b: Rosedale diamond interchange: Incident northbound, one lane blocked. phase times and movements for proposed scheme (three-phase control, favoring right side)

Intersection	Phase	Time (sec)	Links Moving
Left	1	15	501 502 503 504 510
	1'	24	507 508
	2	7	507 508
	3	24	505 511
Right	1	17	601 602 603 609 604
	1'	7	601 602 603 609 604
	2	19	607 612
	3	27	605 606

Table 4.4a: Rosedale diamond interchange: Incident northbound, two lanes blocked. average delays in seconds for different phasing schemes

Direction	Movement	Flow (veh/hr)	4-phase control	3-phase control	3-phase control favor right side	3-phase control no left turns	2-phase control
Southbound	Left	368	37.4	38.8	**>60**	33.8	
	Straight	643	37.4	38.8	**>60**	33.8	
	Right	10	21.4	23.9	7.8	22.9	
Northbound	Left	800	55	**>60**	**>60**	52.4	
	Straight	2760	55	**>60**	**>60**	52.4	
	Right	475	9.8	11.2	11.2	10.8	
Eastbound	Straight	948	13.5	**>60**	**>60**	4.4	
	Right	56	13.5	**>60**	40.2	4.4	
Westbound	Straight	658	**>60**	**>60**	**>60**	19.8	
	Right	54	**>60**	**>60**	**>60**	19.8	
Internal Eastbound	Left	322	11	**>60**	**>60**	no turn	
	Straight	994	15.2	**>60**	**>60**	55.7	
Internal Westbound	Left	941	**>60**	6.9	**>60**	no turn	
	Straight	517	1.1	16.8	**>60**	4.1	
Best Cycle Length (sec)			60	70	70	70	
Interchange Average Delay			40.38	**>60**	**>60**	25.3	

Table 4.4b: Rosedale diamond interchange: Incident northbound, two lanes blocked. phase times and movements for proposed scheme (three-phase control, no internal left turns)

Intersection	Phase	Time (sec)	Links Moving
Left	1	46	505 511 507
	2	7	505 511 507
	3	17	501 502 503 504 510
Right	1	42	601 602 603 609 604
	2	7	605 607 612
	3	21	605 607 612

Table 4.5a: Rosedale diamond interchange: Incident southbound, one lane blocked. average delays in seconds for different phasing schemes

Direction	Movement	Flow (veh/hr)	4-phase control	3-phase control	3-phase control favor left side	3-phase control no left turns	2-phase control
Southbound	Left	418	23.6	39.2	41.3	23.6	
	Straight	1043	23.6	39.2	41.3	23.6	
	Right	60	16.2	24.1	23	16.2	
Northbound	Left	550	19.2	26.1	27.9	23.6	
	Straight	760	19.2	26.1	27.9	23.6	
	Right	225	17	21.2	24.1	18	
Eastbound	Straight	948	17.5	**>60**	*60*	7.6	
	Right	56	17.5	**>60**	*60*	7.6	
Westbound	Straight	658	16.7	52.2	33.6	7.8	
	Right	54	16.7	52.2	33.6	7.8	
Internal Eastbound	Left	322	16.9	13.9	6.8	no turn	
	Straight	1044	8.3	16.4	9.9	11.6	
Internal Westbound	Left	941	**>60**	16.9	27.5	no turn	
	Straight	267	8.6	45.5	52.9	7.6	
Best Cycle Length (sec)			60	80	75	60	
Interchange Average Delay			38.51	44.77	33.06	14.43	

Table 4.5b: Rosedale diamond interchange: Incident southbound, one lane blocked. phase times and movements for proposed scheme (three-phase control, no internal left turns)

Intersection	Phase	Splits (sec)	Links Moving
Left	1	7	505 511 507
	2	32	505 511 507
	3	21	501 502 503 504 510
Right	1	23	601 602 603 609 604
	2	7	605 607 612
	3	30	605 607 612

Table 4.6a: Rosedale diamond interchange: Incident southbound, two lane blocked. average delays in seconds for different phasing schemes

Direction	Movement	Flow (veh/hr)	4-phase control	3-phase control	3-phase control favor left side	3-phase control no left turns	2-phase control
Southbound	Left	1198	**>60**	**>60**	44.7	20.8	
	Straight	2643	**>60**	**>60**	46.7	20.8	
	Right	260	12.4	12.3	9.2	8.1	
Northbound	Left	550	22.9	32.8	**>60**	19.6	
	Straight	760	22.9	32.8	**>60**	19.6	
	Right	225	20.2	21.7	**>60**	16.2	
Eastbound	Straight	948	**>60**	**>60**	**>60**	22.5	
	Right	56	**>60**	**>60**	**>60**	22.5	
Westbound	Straight	658	15.5	**>60**	24.4	9.9	
	Right	54	15.5	**>60**	24.4	9.9	
Internal Eastbound	Left	322	3.6	14.1	8.2	no turn	
	Straight	1244	5.4	15.1	6.6	11.1	
Internal Westbound	Left	941	**>60**	14	**>60**	no turn	
	Straight	267	2.4	15.6	**>60**	35.7	

Best Cycle Length (sec)	70	65	65	65	
Interchange Average Delay	59.92	**>60**	**>60**	18.74	

Table 4.6b: Rosedale diamond interchange: Incident southbound, two lanes blocked. phase times and movements for proposed scheme (three-phase control, no internal left turns)

Intersection	Phase	Splits (sec)	Links Moving
Left	1	19	505 511 507
	2	7	505 511 507
	3	39	501 502 503 504 510
Right	1	28	601 602 603 609 604
	2	7	605 607 612
	3	30	605 607 612

Table 4.7a: Rosedale diamond interchange: Incidents northbound and southbound, two lanes blocked each direction. Average delays in seconds for different phasing schemes

Direction	Movement	Flow (veh/hr)	4-phase control	3-phase control	3-phase control no left turns	2-phase control
Southbound	Left	1198	**>60**	44.9	30.9	no turn
	Straight	2643	**>60**	46.9	30.9	14
	Right	260	11.1	10.2	9.2	12
Northbound	Left	800	**>60**	**>60**	**>60**	no turn
	Straight	2760	**>60**	**>60**	**>60**	34.8
	Right	475	11.3	12.3	15.9	8.7
Eastbound	Straight	948	**>60**	**>60**	18.8	32.2
	Right	56	**>60**	**>60**	18.8	32.2
Westbound	Straight	658	**>60**	**>60**	11.8	39.3
	Right	54	**>60**	**>60**	11.8	39.3
Internal Eastbound	Left	322	3.6	**>60**	no turn	no turn
	Straight	1244	16.3	**>60**	19.6	34.8
Internal Westbound	Left	941	>60	**>60**	no turn	no turn
	Straight	517	7.1	**>60**	**>60**	1.1
Best Cycle Length (sec)			60	75	60	110
Interchange Average Delay			**>60**	**>60**	53.48	27.14

Table 4.7b: Rosedale diamond interchange: Incidents northbound and southbound, two lanes blocked each direction. Phase times and movements for proposed scheme (two-phase control, no left turns)

Intersection	Phase	Splits (sec)	Links Moving
Left	1	66	501 510
	2	44	505 507 511
Right	1	77	603 609
	2	33	605 607 612

Table 4.8a: Berry diamond interchange: Different incidents scenarios. Flow (veh/hr) and average delay (sec/veh) for the four-phase control

Direction	Movement	Flow (veh/hr)				
		Incident 1	Incident 2	Incident 3	Incident 4	Incident 5
Southbound	Left	319	319	369	569	569
	Straight	429	429	829	2429	2429
	Right	906	906	956	1156	1156
Northbound	Left	171	371	121	121	371
	Straight	858	2458	458	458	2458
	Right	209	409	159	159	409
Eastbound	Straight	420	420	420	420	420
	Right	69	69	69	69	69
Westbound	Straight	584	584	584	584	584
	Right	114	114	114	114	114
Internal Eastbound	Left	420	420	420	420	420
	Straight	319	319	369	569	569
Internal Westbound	Left	77	77	77	77	77
	Straight	678	878	628	628	878

Direction	Movement	Average Delay (sec/veh)				
		Incident 1	Incident 2	Incident 3	Incident 4	Incident 5
Southbound	Left	7.7	7.7	8.8	13.8	14.1
	Straight	7.6	7.6	8.8	13.8	14.1
	Right	23.4	20.5	25.7	38	42.4
Northbound	Left	19.6	29.4	29.2	30.5	23
	Straight	19.6	29.4	29.2	30.5	23
	Right	17.6	11.8	29.2	30.5	12
Eastbound	Straight	36.2	37.1	37	53.8	55.8
	Right	36.2	37.1	37	53.8	55.8
Westbound	Straight	16.7	35.4	8.9	10.7	37.5
	Right	16.7	35.4	8.9	10.7	37.5
Internal Eastbound	Left	14.4	31	1.9	13.6	50.9
	Straight	0.5	7.3	1	0.3	12.5
Internal Westbound	Left	9.8	2.4	22.5	27.3	5
	Straight	13	14.2	22.6	29.3	27.2

Best Cycle Length (sec)	65	70	70	90	85
Interchange Average Delay	17	45.62	17.28	20.96	23.96

Table 4.8b: Berry diamond interchange: Phase times and movements for different incidents (four-phase control)

Intersection	Phase	Movements	Times (sec)				
			Inc. 1	Inc. 2	Inc. 3	Inc. 4	Inc. 5
Left	1	1107 1108	6	6	6	6	6
	2	1107 1105 1111 -1108	14	15	15	17	16
	3	1101 1103 1104 1102 1110	39	43	43	61	57
	4	1107 1108	6	6	6	6	6
Right	1	1201 1202 1203 1204 1209	25	39	16	24	50
	2	1205 1206	6	6	6	6	6
	3	1205 1206	6	6	6	6	6
	4	1205 1207 1212 -1206	28	19	42	54	23

Table 4.9a: Seminary diamond interchange: Different incidents scenarios. Flow (veh/hr) and average delay (sec/veh) for the four-phase control

Direction	Movement	Flow (veh/hr)				
		Incident 1	Incident 2	Incident 3	Incident 4	Incident 5
Southbound	Left	103	103	153	353	353
	Straight	215	215	615	2215	2215
	Right	480	480	530	730	730
Northbound	Left	70	270	20	20	270
	Straight	870	2470	470	470	2470
	Right	503	703	453	453	703
Eastbound	Straight	287	287	287	287	287
	Right	10	10	10	10	10
Westbound	Straight	79	79	79	79	79
	Right	512	512	512	512	512
Internal Eastbound	Left	135	135	135	135	135
	Straight	255	255	305	505	505
Internal Westbound	Left	79	79	79	79	79
	Straight	70	270	20	20	270

Table 4.9a (continued): Seminary diamond interchange: Different incidents scenarios. Flow (veh/hr) and average delay (sec/veh) for the four-phase control

Direction	Movement	Average Delay (sec/veh)				
		Incident 1	Incident 2	Incident 3	Incident 4	Incident 5
Southbound	Left	6.4	6.2	6.5	9.6	10.4
	Straight	6.4	6.2	6.5	9.6	10.4
	Right	9.5	8.7	8.7	9.2	9.9
Northbound	Left	10.9	19.6	12.1	17.4	19.5
	Straight	10.9	19.6	12.1	17.4	19.5
	Right	13.4	13.9	16.2	23.1	15.4
Eastbound	Straight	28.1	31.4	34.5	44.5	44.5
	Right	28.1	31.4	34.5	44.5	44.5
Westbound	Straight	26.7	44.9	21.5	19.7	40.1
	Right	26.7	44.9	21.5	19.7	40.1
Internal Eastbound	Left	0.3	0.5	0.2	0.6	1.1
	Straight	4.5	6.2	4.4	7.2	17.8
Internal Westbound	Left	15.6	18.9	15.3	37.4	45.9
	Straight	6	8.6	6	8.2	20.4

Best Cycle Length (sec)	65	70	65	85	100
Interchange Average Delay	13.68	17.68	13.16	13.58	16.81

Table 4.9b: Seminary diamond interchange: Phase times and movements for different incidents (four-phase control)

Intersection	Phase	Movements	Times (sec)				
			Inc. 1	Inc. 2	Inc. 3	Inc. 4	Inc. 5
Left	1	1507 1508	6	6	6	6	6
	2	1507 1505 1511 -1508	13	13	11	13	17
	3	1501 1503 1504 1502 1510	40	45	42	60	71
	4	1507 1508	6	6	6	6	6
Right	1	1601 1602 1603 1604 1609	35	42	31	37	63
	2	1605 1606	6	6	6	6	6
	3	1605 1606	6	6	6	6	6
	4	1605 1607 1612 -1606	18	16	22	36	25

Incident Northbound on Link 37-34

Link 37-34 northbound I-35W, between Berry and Rosedale diamond interchanges (see test bed representation), was the most saturated link in the network, having a maximum concentration of 7,000 vehicles during the peak hour. An incident at this section represents a worst-case situation. The incident duration was 15 minutes, starting at the 20th and ending at the 45th minute of the DYNASMART simulation. Three freeway lanes of the four northbound lanes were blocked. Links 52-51 and 48-47 (exit ramps) were used to divert traffic to the frontage road (see test bed). The principal goals of the DYNASMART investigation were to evaluate the traffic flow distribution over the entire network and the corresponding impact of variable message sign (VMS) control strategies during the incident, and to examine network traffic diversion impacts. After diversion, vehicles could re-enter the freeway through link 26-25, north of Rosedale interchange. The four diamond interchanges that were most affected included Berry, Morningside, Allen, and Rosedale (see Table 4.10). TRANSYT-7F was used to investigate the optimum cycle length and phasing scheme for the four interchanges. A common cycle length was assumed, but the phasing sequence that yielded the smallest delay was chosen.

First, base case four-phase control with a 70-second cycle was implemented at all interchanges. For Allen and Morningside, delay values were acceptable, but for Rosedale the left internal westbound movement, and for Berry the right external southbound, northbound, and left internal eastbound movements experienced high average delays (see Table 4.11). Since the incident was northbound, three-phase control favoring the left side of the interchange was applied. A tabulation of average delays (see Table 4.12) showed that the eastbound and westbound approaches experienced high delay values at both interchanges. Then, three-phase control with prohibition of internal left turns was applied; with an optimum cycle length of 60 seconds, all approaches yielded acceptable delays (see Table 4.13).

Optimal Phasing Plan and Cycle Length for Incident Scenarios

For incident Scenarios 1 through 5, the cycle length for the incident-affected interchange given in Table 4.14a is most critical. To maintain frontage road progression, all diamond interchanges must operate under the same critical cycle length. Phasing schemes can differ; that is, four-phase control can be used for all diamond interchanges, with the affected one modified to suit critical demands, as shown in Table 4.14a.

Table 4.10: Incident on link 37-34. Traffic flows in veh/hr for Rosedale, Allen, Morningside, and Berry diamond interchanges

Direction	Movement	Flow (veh/hr)			
		Rosedale	Allen	Morningside	Berry
Southbound	Left	160	414	181	587
	Straight	667	176	449	184
	Right	57	183	50	1200
Northbound	Left	603	113	100	165
	Straight	1431	209	533	633
	Right	236	437	22	734
Eastbound	Straight	636	646	379	1444
	Right	50	50	50	46
Westbound	Straight	1401	567	584	623
	Right	59	50	97	82
Internal Eastbound	Left	119	390	193	720
	Straight	677	670	367	1311
Internal Westbound	Left	1186	403	374	102
	Straight	818	301	310	610

Table 4.11: Incident on link 37-34. Average delays in sec/veh for Rosedale, Allen, Morningside, and Berry diamond interchanges for Scenario 1

Direction	Movement	Average Delay (sec/veh)			
		Rosedale	Allen	Morningside	Berry
		4-phase	4-phase	4-phase	4-phase
Southbound	Left	36.7	20.7	26.7	33.8
	Straight	36.7	19.4	26.7	33.8
	Right	25.3	19.3	22.8	**>60**
Northbound	Left	48.2	15.3	31.4	16.3
	Straight	48.2	15.3	31.4	16.3
	Right	20.1	21.3	22.1	**>60**
Eastbound	Straight	9.5	16.1	12.3	16.8
	Right	9.5	16.1	12.3	16.8
Westbound	Straight	41.2	18.4	16.8	19.5
	Right	41.2	18.4	16.8	19.5
Internal Eastbound	Left	20.5	30.5	22.7	**>60**
	Straight	12.7	11.4	4.6	12
Internal Westbound	Left	**>60**	36	26.8	31
	Straight	3	7.8	4.7	4.3

Table 4.12: Incident on link 37-34. Average delays in sec/veh for Rosedale, Allen, Morningside, and Berry diamond interchanges for Scenario 2

Direction	Movement	Average Delay (sec/veh)			
		Rosedale	Allen	Morningside	Berry
		3-phase favor left side	4-phase	4-phase	3-phase favor left side
Southbound	Left	**>60**	20.9	32.1	15.7
	Straight	**>60**	19.7	32.1	15.7
	Right	27.9	19.5	25.4	21.4
Northbound	Left	>60	13.2	33.8	57.2
	Straight	>60	13.2	33.8	57.2
	Right	21.5	18.1	22.1	**>60**
Eastbound	Straight	**>60**	14.9	10.5	**>60**
	Right	**>60**	14.9	10.5	**>60**
Westbound	Straight	**>60**	20.5	17.1	**>60**
	Right	**>60**	20.5	17.1	**>60**
Internal Eastbound	Left	54.5	15.5	17.9	56.1
	Straight	52.5	6.9	3.8	59
Internal Westbound	Left	17.2	12.5	18.3	9.4
	Straight	17.3	5.5	2.1	9.5

Table 4.13: Incident on link 37-34. Average delays in sec/veh for Rosedale, Allen, Morningside, and Berry diamond interchanges for Scenario 3

Direction	Movement	Average Delay (sec/veh)			
		Rosedale	Allen	Morningside	Berry
		3-phase no left turns	4-phase	4-phase	3-phase no left turns
Southbound	Left	34.5	19.6	25.5	14.8
	Straight	34.5	18.3	25.5	14.8
	Right	22.3	18.1	21.1	21.1
Northbound	Left	22	12.6	30.7	21.9
	Straight	22	12.6	30.7	38.1
	Right	13.8	17.3	19.8	38.1
Eastbound	Straight	3.5	14.8	11.5	16.6
	Right	3.5	14.8	11.5	16.6
Westbound	Straight	16.1	19.9	18.1	5.7
	Right	16.1	19.9	18.1	5.7
Internal Eastbound	Left	no turn	11.3	15.9	no turn
	Straight	16.1	5.2	6.5	10
Internal Westbound	Left	no turn	7.5	19.7	no turn
	Straight	5.3	3.1	2.7	7.9

Table 4.14a: Optimal phasing plan and cycle length for incident Scenarios 1-5 for the three major diamond interchanges

	Critical Interchanges		
Incident *	Rosedale	Berry	Seminary
1. Freeway capacity	70 sec	65 sec	65 sec
reduced by 25%	3-phase scheme	4-phase scheme	4-phase scheme
northbound	favor right side		
2. Freeway capacity	70 sec	70 sec	70 sec
reduced by 50%	3-phase scheme	4-phase scheme	4-phase scheme
northbound	no left turns		
3. Freeway capacity	60 sec	70 sec	65 sec
reduced by 25%	3-phase scheme	4-phase scheme	4-phase scheme
southbound	no left turns		
4. Freeway capacity	65 sec	90 sec	85 sec
reduced by 50%	3-phase scheme	4-phase scheme	4-phase scheme
southbound	no left turns		
5. Freeway capacity	110 sec	85 sec	100 sec
reduced by 50%	2-phase scheme	4-phase scheme	4-phase scheme
in both directions			
** Incidents and phase scheme steps are described in detail in paragraph 4.2*			

For the incident case of link 37-34, all diamond interchanges can operate under a 60-second cycle length. The three-phase scheme with prohibition of left turns from the internal approaches is the optimum phasing plan for the Rosedale and Berry diamond interchanges, although the four-phase scheme is the optimal phasing plan for all other interchanges (see Table 4.14b).

Table 4.14b: Optimal phasing plan and cycle length for incident on link 37-34 for the involved diamond interchanges

Incident	Rosedale	Allen	Morningside	Berry
Freeway capacity reduced by 75% on northbound link 37-34	60 sec 3-phase scheme no left turns	60 sec 4-phase scheme	60 sec 4-phase scheme	60 sec 3-phase scheme no left turns

4.3. THIRD SET OF EXPERIMENTS

Experiment Description

This third set of experiments is designed to investigate the sensitivity of the system to different incident characteristics, in order to select cases that represent different levels of incident impact in the system. These will be used in developing and evaluating integrated arterial and freeway control strategies to improve system performance under congested conditions.

Nonrecurrent congestion caused by incidents has special relevance to the objectives of this study. In particular, incidents on the freeway system are of primary concern to such integrated strategies for the following reasons. First, the location on the freeway allows us to analyze the benefits of integration to address problems in the facility that carries the highest flows in the system, thereby affecting a considerable portion of the trips. Second, once a freeway segment has seen a reduction in its capacity, many vehicles will accumulate upstream of the incident link, generating large delays to many users. Third, the situation provides an opportunity to explore diversion routes that could be improved to handle the additional demand imposed by the incident. Finally, the source of congestion is such that the increase in traffic flow can be localized and certain traffic control measures can be applied to the affected areas.

The main characteristics of an incident for modeling purposes are its location, severity, and duration. In these experiments, the incident occurs during the peak period, in which about 17,000 vehicles enter the network in a 35-minute period. The incident data considered for this set of experiments are presented in Table 4.15 (the link numbers are given with reference to Fig 3.11).

Table 4.15: Experimental setup

Location (2 locations)	Severity (4 levels)	Duration (4 levels)
Link 48-41	0.25	5 min - (25 - 30)
	0.50	15 min - (20 - 35)
Link 37-34	0.75	30 min - (20 - 50)
	0.99	60 min - (20 - 80)

The following subsystems are considered for analysis purposes: (1) freeway northbound, (2) freeway southbound, (3) frontage road, (4) diamond intersections, and (5) arterial streets, where details for specific arterial streets are also provided.

The main performance measure considered is the stopped time that users experience on the different subsystems. This time can also be viewed as the delay experienced by each vehicle at each node. To compute the average stopped time, all stopped time incurred by the vehicles using a certain subsystem is considered. (This information is readily available in one of the simulation output files.) Note that in some cases a link might belong to two different subsystems (as is the case with diamond intersections and arterial streets); however, the stopped time is considered for each subsystem separately. As a result, the sum of the stopped time over all the subsystems may not be the same as the total stopped time for the whole network.

In this set of experiments, the impact of the incident is evaluated within a situation where the network users follow exactly the same paths that they would have used in the absence of the incident. This provides a controlled and reproducible situation that does not benefit from the congestion relief effect of any countermeasure. For simulation purposes, the vehicles are generated using the vehicle characteristics and path files (generated by DYNASMART) from the no-incident base case. Thus, during the simulation, the vehicles follow the very same paths that they took when there was no incident in the network; as such, they experience the congestion caused by the incident. This can be considered as a practical worst case situation for the freeway incident, since no diversion is allowed for the vehicles already captured upstream of the incident location.

Table 4.16 presents the results of interest obtained from the experiments of Table 4.15. The facility that is most affected by the incident condition is the freeway northbound. Therefore, the stopped time experienced on this facility provides better differentiation among the cases tested.

Figure 4.5 presents part of the network with the indication of the incident links considered in the experiments. The main difference between the two locations is the affected area when diversion is allowed. In the case of link 48-41, there is a clear diversion path to avoid the incident location, whereas in the case of link 37-34, there are multiple possibilities and more complications in finding diversion paths to improve the situation. According to the stopped time values, the more interesting cases are related to the incident on the link 37-34 because it represents a complicated situation. Nevertheless, there is also the need to explore some cases with the incident located on link 48-41, which, while less complicated than the other link, also generates interesting solutions from the integration stand point.

Table 4.16: Stopped time freeway northbound for each incident situation

Link	Duration	Severity	Stopped time Freeway NB
48 - 41	5	0.25	2646
	15	0.25	2996
	30	0.25	2738
	60	0.25	2733
	5	0.50	4654
	15	0.50	4945
	30	0.50	5002
	60	0.50	5001
	5	0.75	2791
	15	0.75	8878
	30	0.75	10780
	60	0.75	20230
	5	0.99	7488
	15	0.99	15600
	30	0.99	33896
	60	0.99	86704
37 - 34	5	0.25	2427
	15	0.25	2080
	30	0.25	2205
	60	0.25	2205
	5	0.50	5157
	15	0.50	5463
	30	0.50	6826
	60	0.50	6928
	5	0.75	8247
	15	0.75	16935
	30	0.75	23916
	60	0.75	48086
	5	0.99	10078
	15	0.99	20343
	30	0.99	43537
	60	0.99	N/R *

Note: N/R* indicates congestion over feasible limits for the simulation

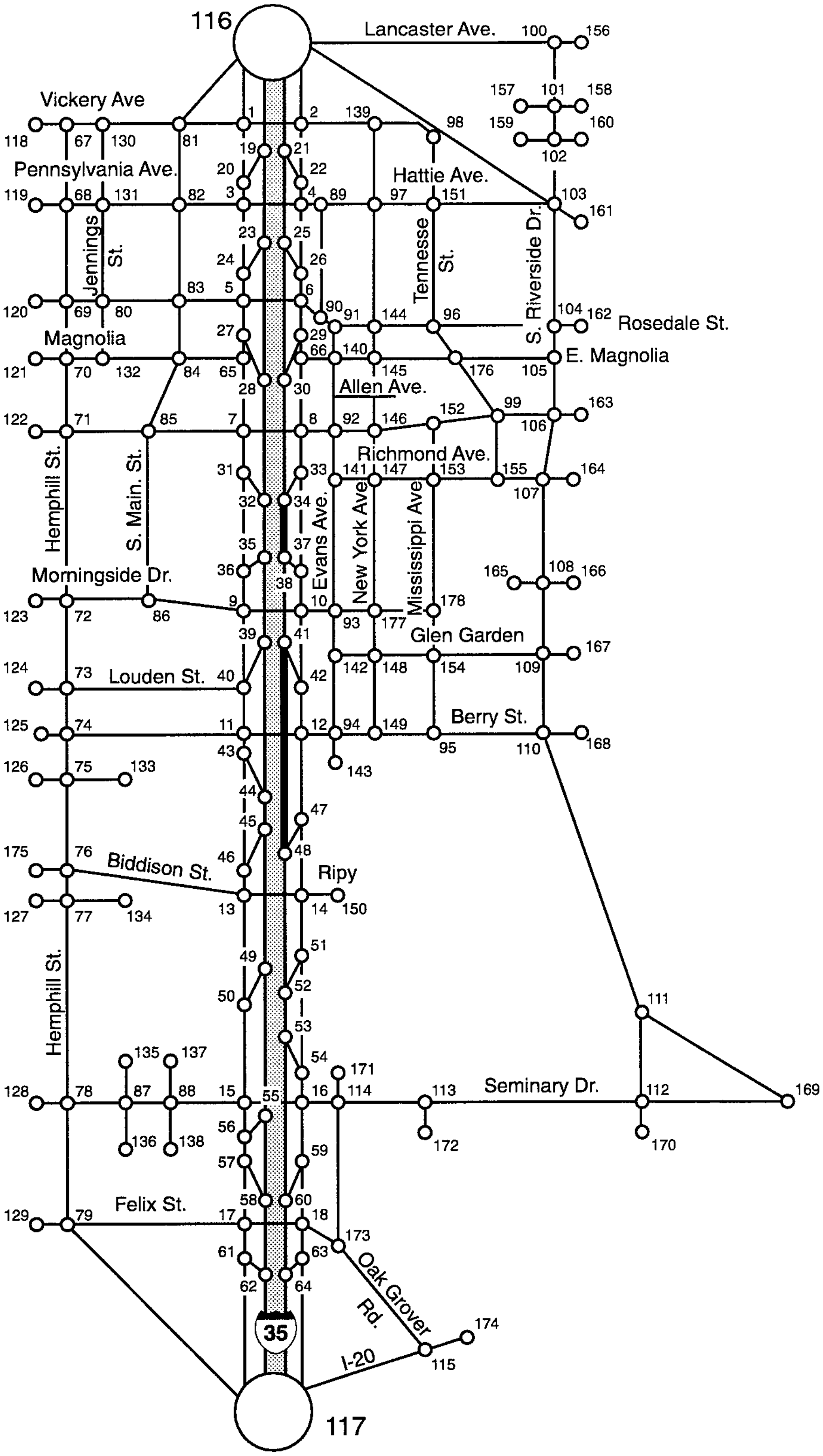


Figure 4.5: Incident location for Experiment 3

Selection of Incident Scenarios

The main interest is to determine the cases that will be used to perform more detailed experiments so as to reflect different levels of congestion, ranging from low to severe. As such, seven cases were selected; these are listed in Table 4.17, from low to high congestion in the freeway northbound. Figure 4.6 depicts the information presented in Table 4.16, but sorted in ascending order of the stopped time.

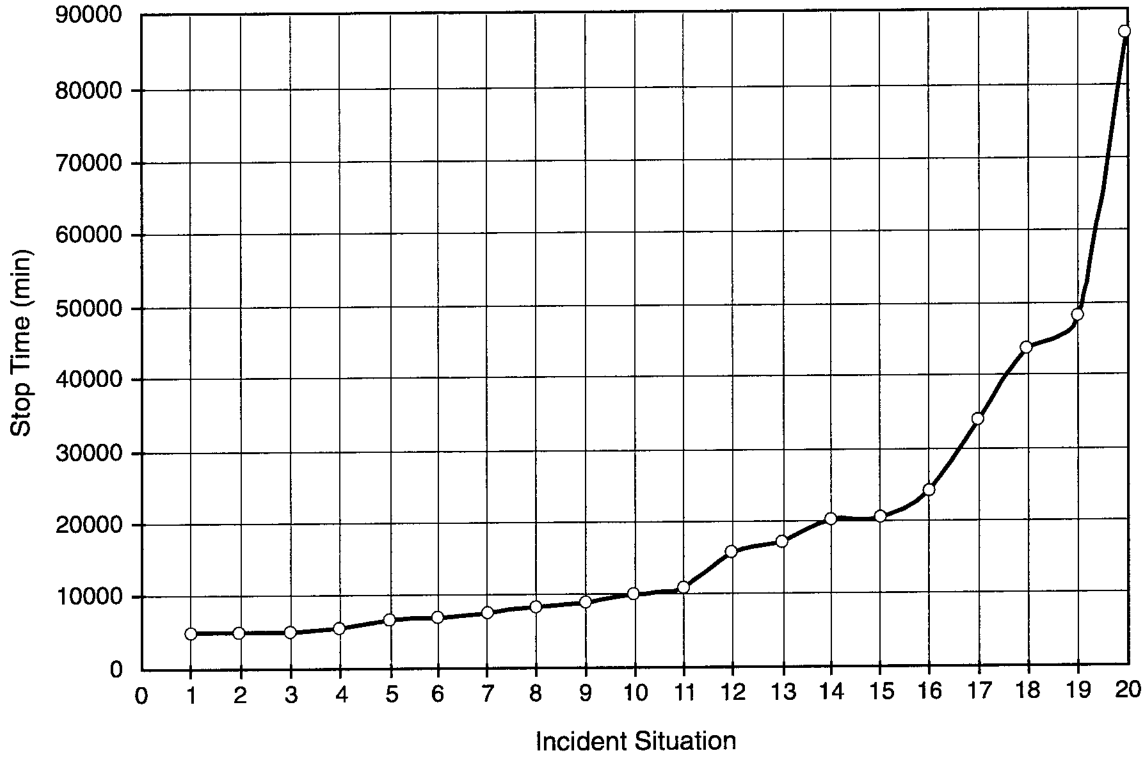


Figure 4.6: Stopped time on the freeway northbound for different incident situations

Table 4.17: Selected incident scenarios

Case	Link	Duration	Severity	Stopped time Freeway NB
1	37 - 34	15	0.5	5463
2	48 - 41	30	0.75	10780
3	37 - 34	15	0.75	16935
4	37 - 34	30	0.75	23916
5	48 - 41	30	0.99	33896
6	37 - 34	30	0.99	43537
7	37 - 34	60	0.75	48086

Experiments to Investigate VMS Effectiveness

The following situations are investigated in order to examine the effect of information provision through variable message signs (VMS) under incident conditions.

First, consider the base case. In this case, users are assumed to have access to only pretrip information to help them select the path that they will follow. The network is loaded through the automated generation process used in the simulator. This is a critical case because the vehicle and path files generated in this initial run are considered to be the same for several of the subsequent runs in order to allow direct comparability. These files (vehicle characteristics and path files) represent the routes that the users will typically take if no additional information is provided. Therefore, they represent "current" conditions without ATMS measures.

Second, consider the case where there is no diversion once en-route, but where users have access to pretrip information at all times during the incident. This case is similar to the base case, with the addition of the incident. Users have only pretrip information on prevailing conditions to help in selecting the path that they will follow. They select the best path available to them at the starting location. When the incident occurs, users that have not yet initiated their trips (i.e., vehicles that have not yet been generated) will know about the congestion caused by the incident, and some may therefore follow paths that differ from the "usual" ones (without the incident).

Third, consider the worst situation for the freeway with incident. In this case, the tripmakers are assumed to follow their usual paths (i.e., without incident) even though the incident is in progress. For simulation purposes, the vehicles are generated using the vehicle characteristics and path files from the base case. Therefore, the users follow the very same paths as if there were no incident in the network, though they experience the congestion caused by the incident. As noted, this can be considered the practical worst-case situation for the freeway incident, since no diversion is allowed for the vehicles already captured upstream of the incident location.

Finally, consider the case of an incident on the freeway, but with en-route diversion via VMS. In this case, a fraction of the users may be diverted as they approach the incident-caused congestion location. For simulation purposes, the vehicles are generated using the vehicle characteristics and path files from the base case. Thus, the users will follow the same paths, but when they arrive to the links that precede the incident location, a variable message sign will alert them to the incident situation, giving them the opportunity to divert (as determined by the decision rules internal to the DYNASMART model).

Traffic Characteristics under Incident and VMS Cases

The traffic volumes around the incident link were analyzed for the above four cases considered in the experimental setup. Three links were selected to illustrate the temporal pattern of the volume with and without VMS. The selected links are: link (48,47), which is the exit ramp immediately downstream of the VMS; link (47,12), which lies along the frontage road between the downstream off-ramp and the diamond intersection; and link (12,42), which is also along the frontage road, downstream of the diamond intersection. The volumes (in vehicles per hour) on those links are depicted at 5-minute intervals over the simulation period in Figures 4.7, 4.8, 4.9, and 4.10 for each of the four cases considered, respectively.

The volumes shown in Figures 4.7 and 4.9, which correspond to the base case and the worst case, respectively, are very similar. In both cases, no diversion is allowed, hence providing a good representation of the traffic flow patterns that would develop in the freeway corridor when an incident occurs and users remain on the freeway until the incident is cleared.

When vehicles have pretrip information (case 2 in Figure 4.8), the volumes vary substantially relative to the base case. The volumes in the three links considered increase, because some users who now know in advance about the incident select the path formed by these links.

Figure 4.10 shows the volumes when no pretrip information is provided, but a VMS is located along the freeway upstream of the incident link. In this case, more vehicles are diverted using the off-ramp. While some of these vehicles continue traveling along the frontage road (link [12,42]), not all do because the frontage road becomes congested as well.

Discussion of Each Case

Case One — Base Case: The first column in Table 4.18 summarizes the total stopped time for each of the different subsystems considered in this analysis. The first column in Table 4.19 compares the stopped time in the base case to the stopped time in the worst case in terms of the percentage difference (from worst case).

The base case represents a congested (peak) period, but with no incident. The summary output file of DYNASMART gives several performance measures to characterize the network conditions considered for these experiments. The stopped time in all the subsystems is relatively high. The average stopped time (over all vehicles) is 12.45 minutes, and represents more than 50% of the average total trip time (20.21 min.).

The percentage deviation of the stopped time relative to the worst case shows the extent of degradation of network performance caused by the incident on the freeway. It is interesting that some of these values are negative, reflecting the fact that the performance of that specific subsystem under the "worst case" situation is actually better than that under "normal" conditions. This is an illustration of network effects, whereby vehicles trapped by the incident may actually decrease the load over other subsystems (for example, freeway southbound) and, as a result, perform better than they would under the base case.

The variation of the queue length over time upstream of freeway northbound (link [37, 34]) in the base case is shown in Figure 4.11. This is the queue upstream of the incident link. As seen in this figure, there are some vehicles in queue even in the no-incident base case, reflecting the degree of congestion considered in these experiments.

Case Two — Pretrip Information Case: The second columns in Tables 4.18 and 4.19 present information comparable to that discussed in the base case, but for Case two, where only pretrip information is provided to users and the incident occurs.

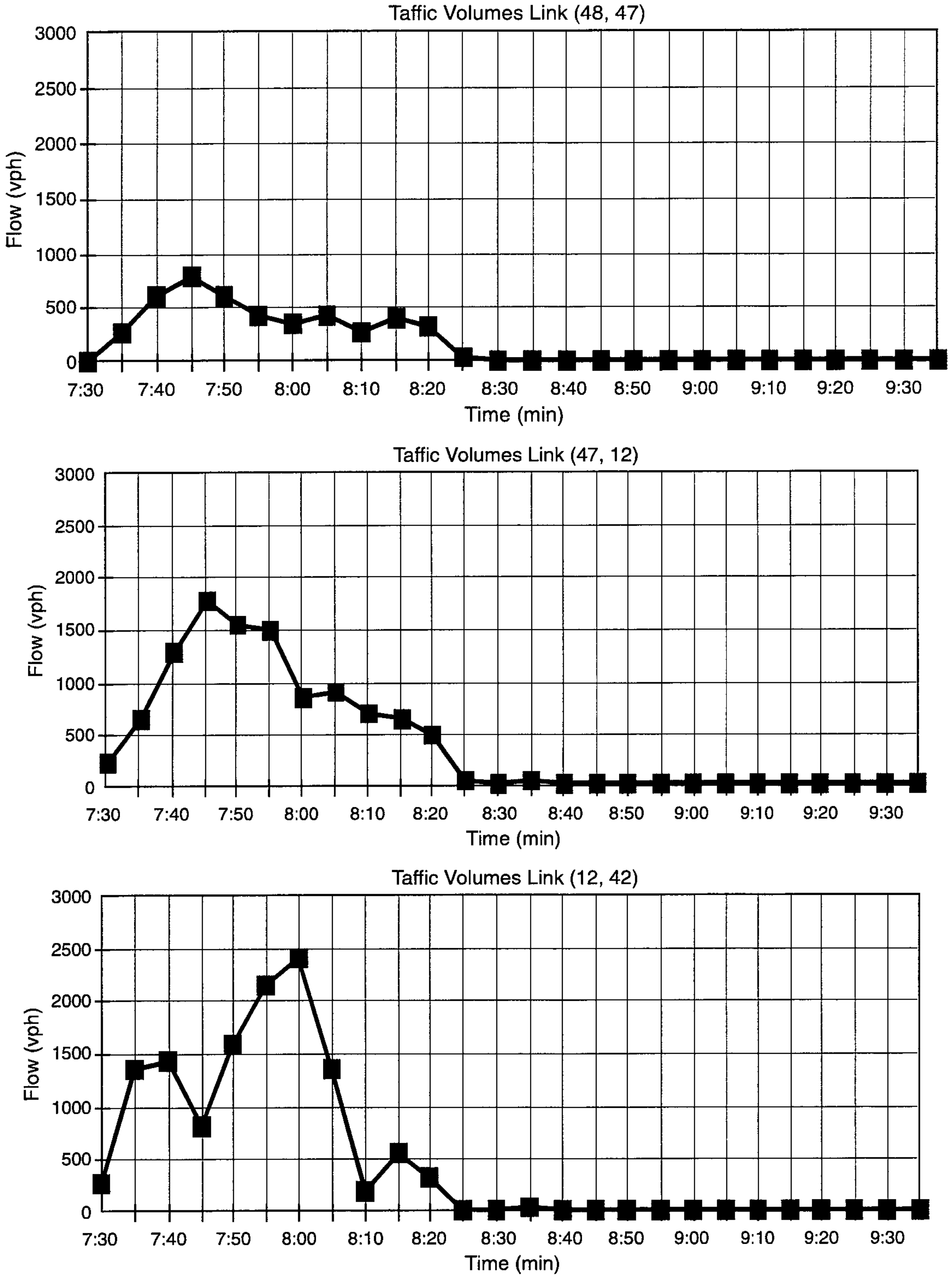
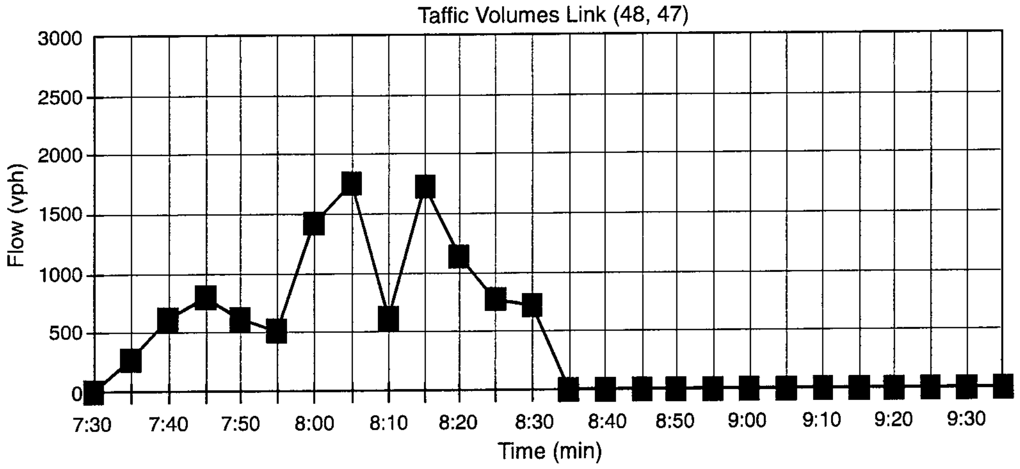
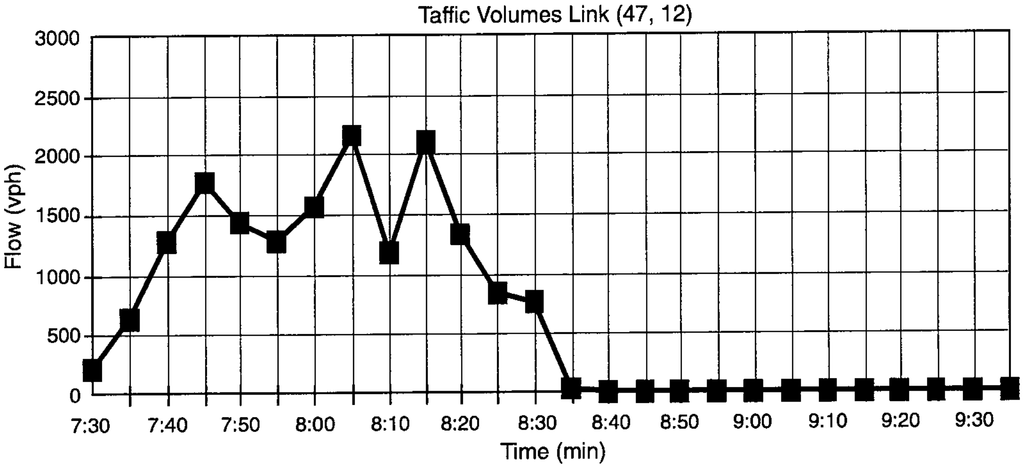


Figure 4.7. Base case — Traffic volumes for selected links





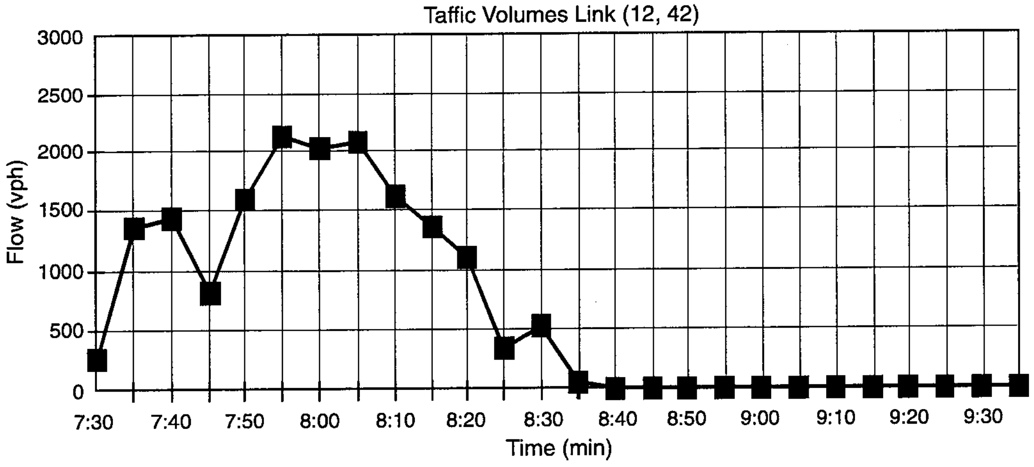
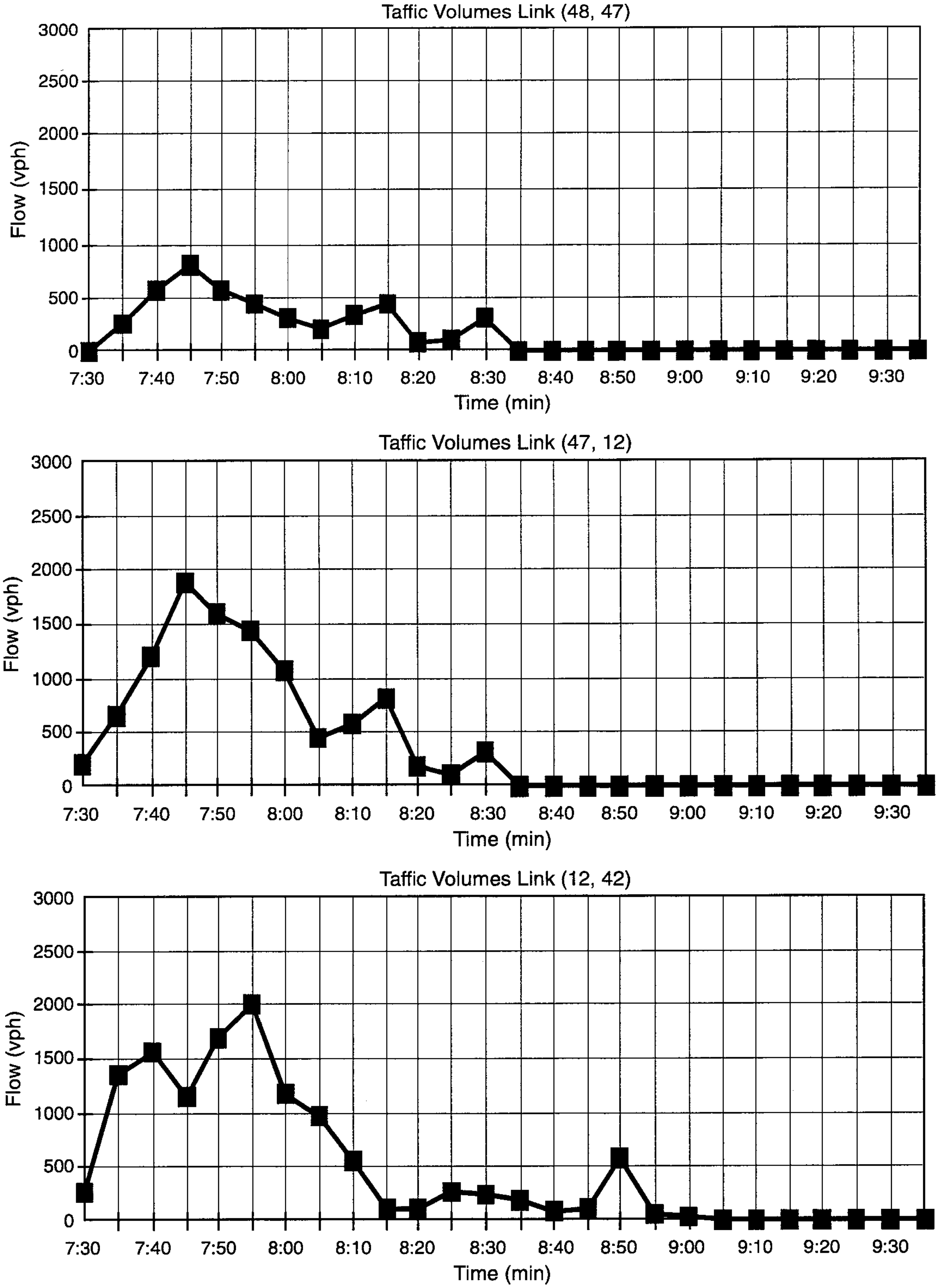


Figure 4.8: Pretrip information case — Traffic volumes for selected links



Taffic Volumes Link (48, 47)

Taffic Volumes Link (47, 12)

Taffic Volumes Link (12, 42)

Figure 4.9: Worst case — Traffic volumes for selected links

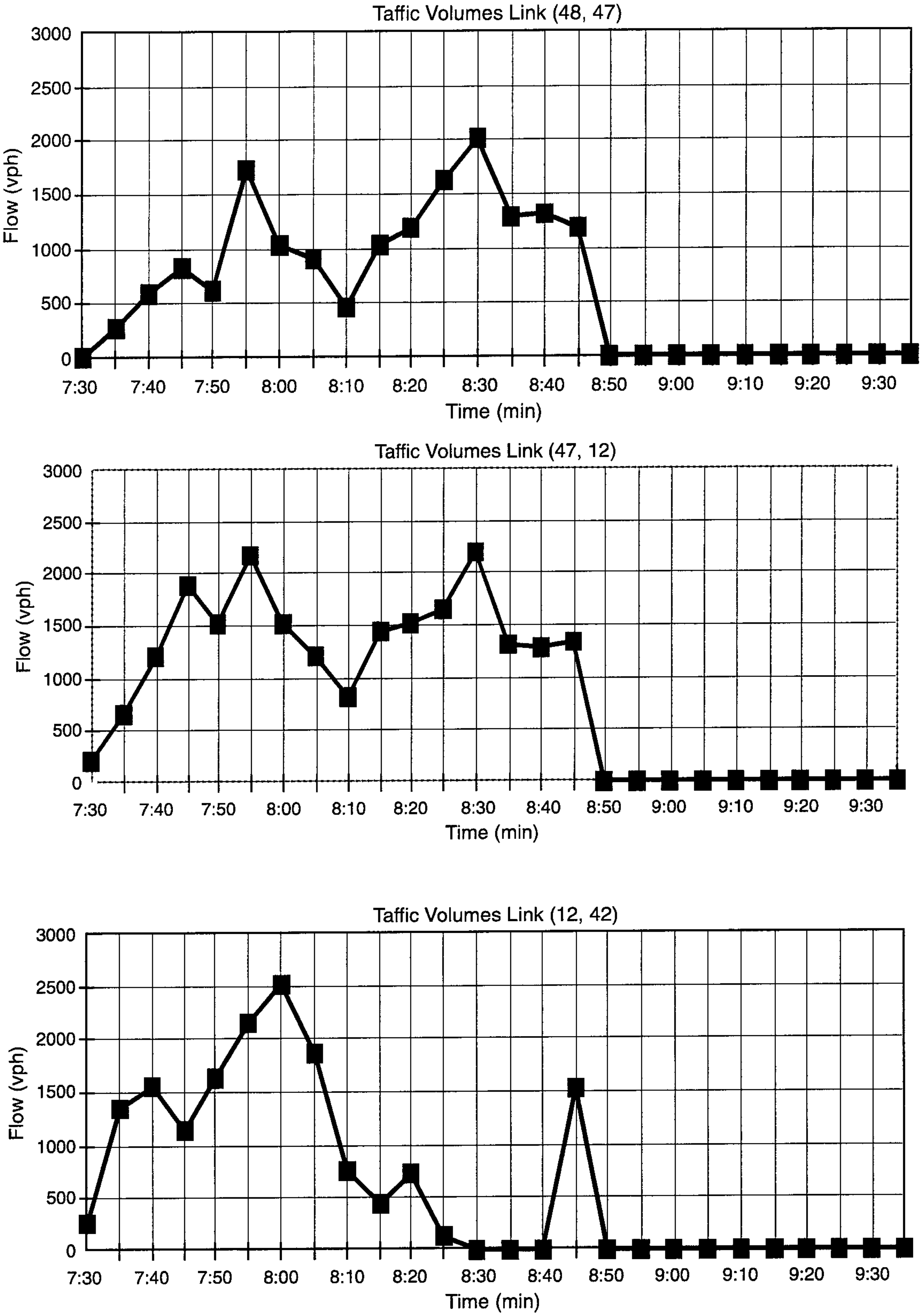


Figure 4.10: VMS — Traffic volumes for selected links

Table 4.18: Total stopped time disaggregated in different subsystems

	TOTAL STOPPED TIME			
Facility	**Base Case**	**Pre-trip inf. Case**	**Worst Case**	**VMS**
Freeway sb	4034	5934	869	823
Freeway nb	2926	19837	48086	30579
Frontage sb	26819	25855	28214	23482
Frontage nb	23614	29887	36637	34268
Diamonds	44154	49956	54527	49896
Riverside Dr	21329	22222	20943	21873
Hemphill St	26968	26571	26909	28043
Rosedale St	23916	25416	25652	24023
Berry St	20987	22244	29769	24663
Seminary Dr	24498	22969	28967	28855
TOT ARTERIAL St	105103	106186	119674	113558
Other links	48364	42055	35245	39920
NETWORK TOTAL	188179	207522	234053	215373

In this case, the percentage deviation of the total stopped time relative to the worst case suggests that when users have access to pretrip information while the incident is in progress, the system in general performs better than when users do not know about the incident in advance. Similar to the base case, some subsystems, like the freeway southbound, actually perform better in the "worst case," for the same reason discussed previously.

Figure 4.12 depicts the variation in queue length on the freeway link upstream of the incident. There is a long queue owing to the incident, which reaches approximately 400 vehicles over four lanes (about 100 veh/lane). Even though this is a long queue, it is shorter than the queue that is formed in the worst case or when using only VMS as information source.

Table 4.19: Comparison of all the cases relative to the worst case – reduction (or increase, if difference is negative) in the total stopped time as a percentage of the worst case stopped time

PERCENTAGE DEVIATION IN TOTAL STOPPED TIME			
Facility	**Base Case**	**Pre-trip inf. Case**	**VMS**
Freeway sb	-364.2	-583.0	5.3
Freeway nb	93.9	58.7	36.4
Frontage sb	4.9	8.4	16.8
Frontage nb	35.5	18.4	6.5
Diamonds	19.0	8.4	8.5
Riverside Dr	-1.8	-6.1	-4.4
Hemphill St	-0.2	1.3	-4.2
Rosedale St	6.8	0.9	6.4
Berry St	29.5	25.3	17.2
Seminary Dr	15.4	20.7	0.4
TOT ARTERIAL St	12.2	11.3	5.1
Other links	-37.2	-19.3	-13.3
NETWORK TOTAL	19.6	11.3	8.0

Case Three — Worst Case: The third column in Table 4.18 corresponds to the worst case condition. As explained, this case is taken as reference against which other cases are compared in Table 4.19.

Comparing the values of the stopped time for this case to the values in columns one and two of Table 4.18, the following observations can be made:

- The total stopped time in the network is greater than that for any other case.
- Different subsystems experience different levels of delay in this congested situation. The freeway NB is the most affected subsystem, while some arterial streets experience delay reductions. The delay on local streets is also reduced. This case clearly illustrates that if vehicles are trapped by the incident, then some relief may be experienced in the congestion on some arterial and local streets.

However, the overall delay for the entire network nonetheless increases about 20% with respect to the base case situation.

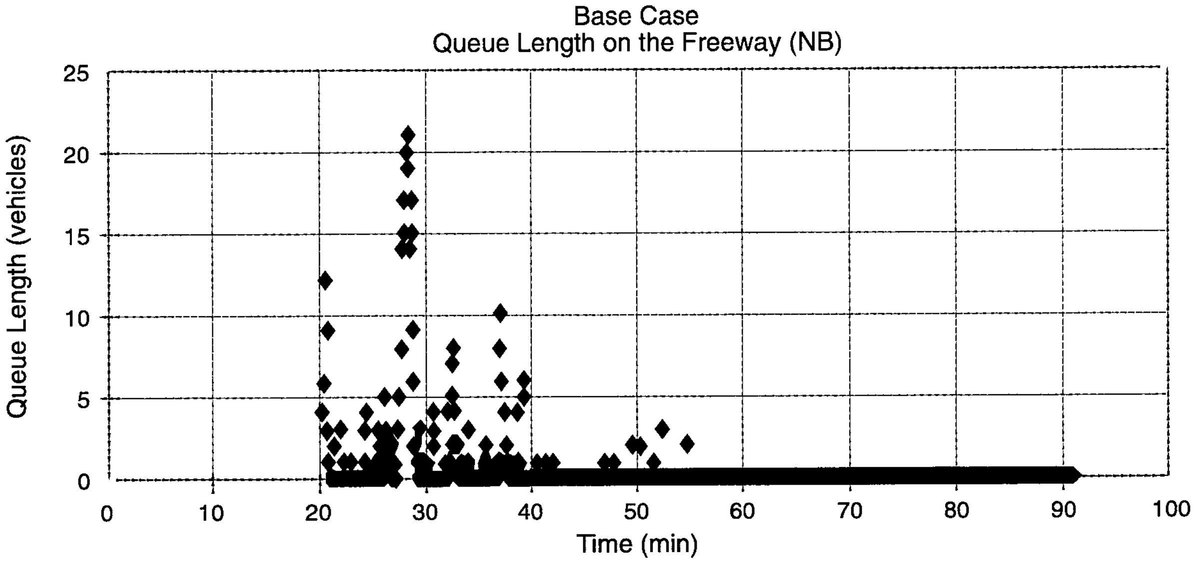


Figure 4.11: Queue length on the freeway (NB) — Base case

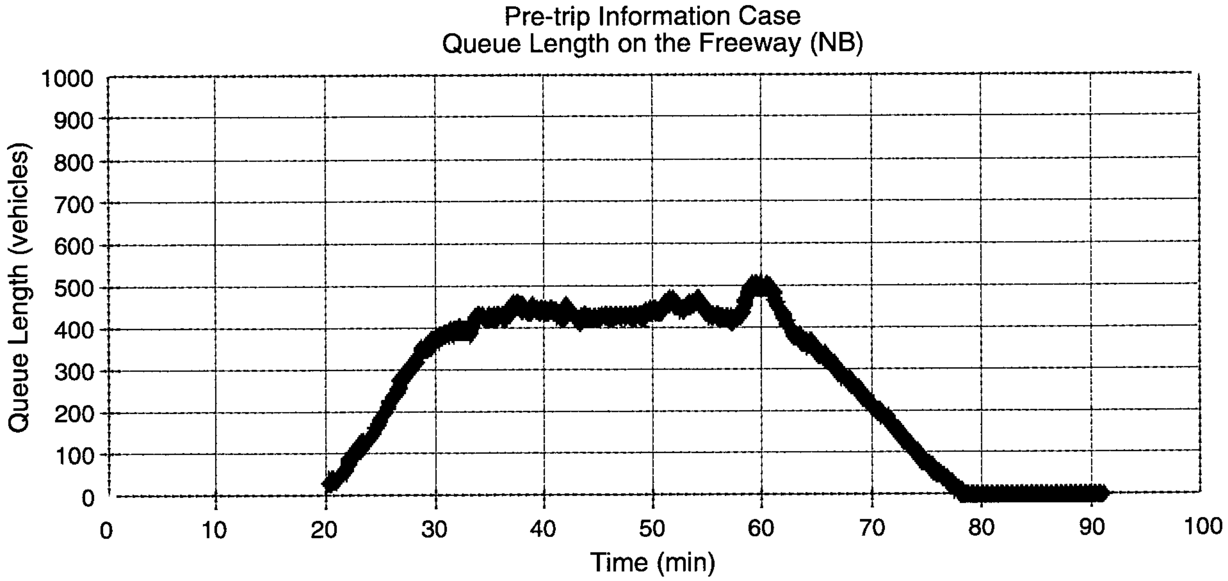


Figure 4.12: Queue length on the freeway (NB) — Pretrip information case

This analysis further illustrates that defining the objective function to be minimized, and more generally the objective of the traffic control measures, should be considered very carefully to account for such network effects. An incident could be beneficial in a local sense for some subsystems in the network. This further highlights the importance of coordinated operation across the various subsystems, and the need for integrated corridor or network-level strategies to optimize overall system performance.

Figure 4.13 shows the variation in queue length upstream of the incident link for the worst-case situation. This figure shows that the queue length reaches a peak of about 950 vehicles in four lanes (about 250 veh./lane). This is definitely a freeway in very bad condition. This happens when the vehicles are trapped in the freeway and no one switches to a different path, admittedly an extreme situation.

Case Four — Information Supplied through VMS: The fourth column in Table 4.18 and the third column in Table 4.19 present the corresponding summary measures for the fourth case.

In this case, even though the vehicles can be diverted, the decision rule used in the simulation only diverts them if the alternative path is better than the current path. Since the rest of the network is already congested, not all the vehicles that can divert actually do.

According to the values presented in Tables 4.18 and 4.19, the diversion using the information given by the VMS improves the situation with respect to the worst case. The overall situation is improved by about 8%, whereas the situation in the affected freeway northbound is improved by about 36%. The improvement is clearly visible in Figure 4.14, which shows the variation in queue length. The queue increases because of the incident but it does not reach the same peak as shown in the worst case. With VMS the maximum number of vehicles in queue is about 700 veh, or about 175 veh./lane. While not a particularly good condition, it is better than the worst case and could be further improved with pretrip information.

4.4 FOURTH SET OF EXPERIMENTS

Experiment Description

This fourth set of experiments is designed to test path-based coordination strategies to be used in case of incident situation or severe congestion. Path-based coordination is a perfect example of an integrated corridor network strategy in that it seeks to optimize system performance by recognizing the particular traffic patterns that develop as a result of diversion from the freeway. It is first proposed and tested in the present study.

This set of experiments expands the previous set by incorporating changes in control to coordinate traffic signals along those paths used when the vehicles divert from the congested freeway. The signal phases along those paths are coordinated using different procedures, including those presented in previous chapters of this report. The results and analysis of those tests are presented below. The general experimental scheme consists of the following cases:

1. Base case, with no incident.

2. Incident with the assumption that the vehicles follow the paths determined in the base case. This is the worst-case situation.
3. Incident situation with initial paths similar to those in the base case, but with VMS to influence the paths of vehicles headed toward the incident. In addition, the integrated path-based coordination strategy is applied.

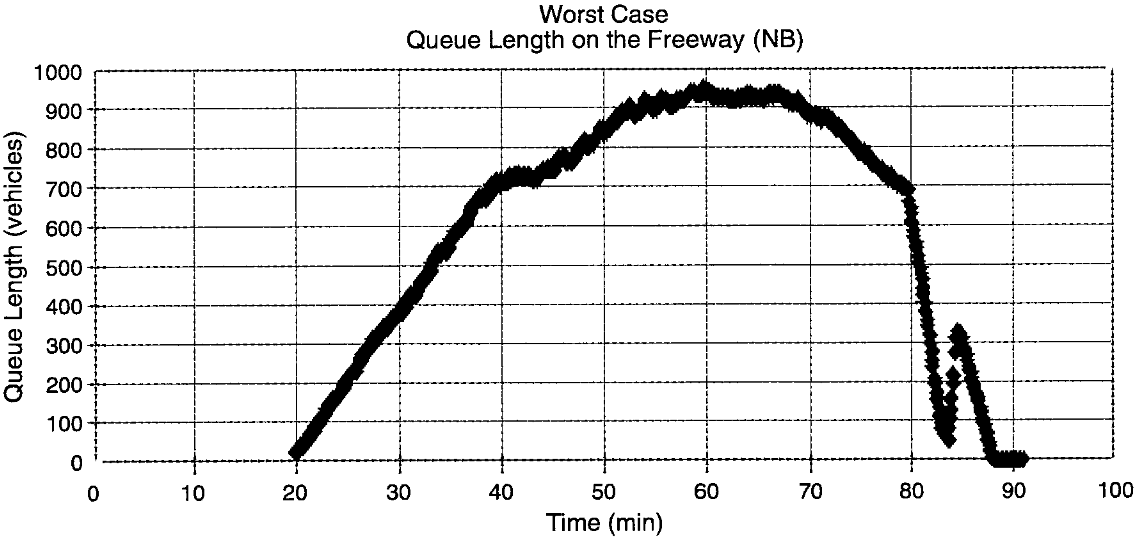


Figure 4.13: Queue length on the freeway (NB) — Worst case

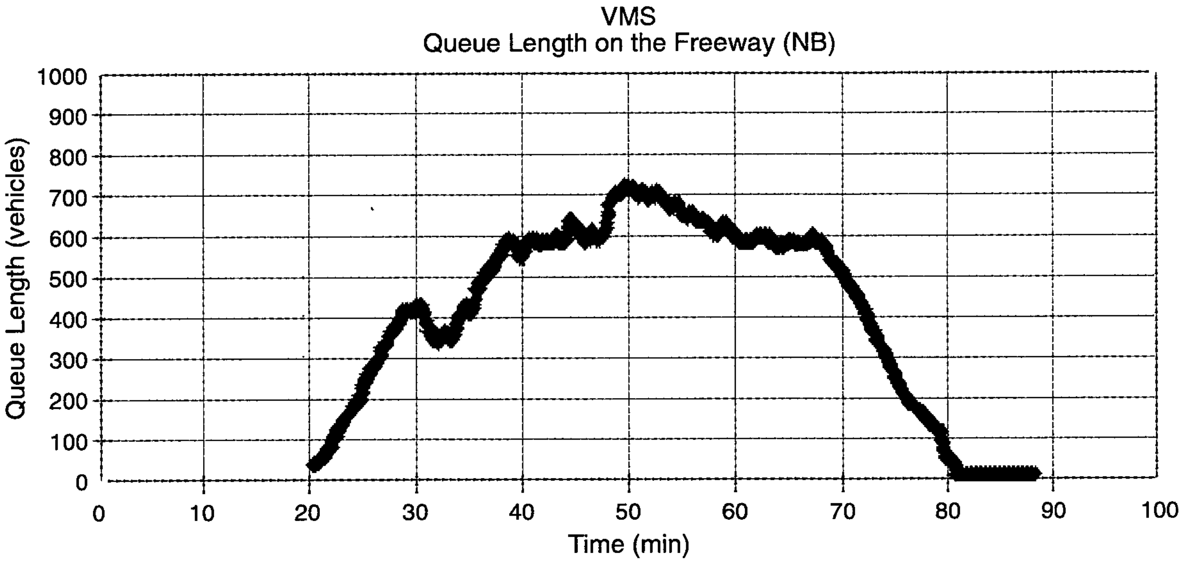


Figure 4.14: Queue length on the freeway (NB) — VMS

The incident situation considered for modeling purposes has the following characteristics: The incident is located on link (37,34), with a severity of 75% and a duration of 60 minutes. The incident occurs during a peak period in which about 17,000 vehicles enter the network in a 35-minute period. The characteristics of this incident will be used also in the final set of experiments. The main reason to consider this incident location is that it allows coordination among a greater number of intersections — i.e., greater than that for the incident located on link (48,41).

In this set of experiments there is an additional VMS located on link (53, 52). With the additional location, more users are informed about the incident situation. Consequently, more vehicles divert to the parallel arterial system. This situation is illustrated by comparing the queue on the freeway caused by the incident when the new VMS is used and the previous situation presented in experiment set No. 3. The number of vehicles in queue decreases with two activated VMS's, thereby increasing the number of vehicles using the arterial (diagonal) surface street system that will be coordinated.

Path-Based Coordination

Coordination of signals to provide minimum disruption to traffic flowing along successive intersections is an essential element of traffic management in networks where the major traffic streams map directly onto major arteries, along which vehicles follow primarily straight stretches of roadway. In the case of diversion from a main freeway, the vehicles may follow different paths, which may have some portions along straight movements but which will also include turning movements. The main purpose of path-based coordination is to provide a progression scheme along the paths of major traffic streams where those paths consist of a combination of straight portions of arterial streets and turning movements.

For this set of experiments, the performance measure used is the average travel time of vehicles traveling over the whole network.

The basic procedure for path-based coordination is the following: First, the most dominant paths followed by the diverted vehicles are determined. After that, we determine the movements that need to be considered, along with the respective phases in which these movements have the green along the selected paths. With these elements, the control settings are changed according to the desired control strategy.

Three basic control strategies could be used. The first is to switch the signals to fixed control (F-C) and calculate the offsets along the paths to be coordinated. The second consists of leaving the signals as vehicle actuated (V-A) and modifying the maximum green. The third is the combination of progression and vehicle actuation. The first two strategies are tested in this set of experiments, with the general results presented below.

Different procedures can be used to determine the offsets in a coordinated path; in this set of experiments, a procedure based on simulated travel time is used. The average travel time on the links along the path under consideration was calculated to determine the offset between coordinated phases. Thus, an iterative procedure is required to simulate the situation. The first run is used to determine the most frequently used paths and the link travel times according to the scenario to be tested. With that result, the required offsets are calculated. A second run is then performed to obtain the performance measures.

Different procedures were tested to determine the common cycle length and the offsets to provide coordination along arterial streets in order to improve the situation caused by the incident and the vehicles diverting from the freeway. Among the procedures tested that did not improve the situation, the following can be mentioned: common cycle length according to the timing provided by TRANSYT 7F, as discussed in the second set of experiments described earlier in the chapter, and timing and offset calculations using PASSER II. Both procedures provided minor benefits to the signals that were coordinated but did not improve the overall situation. The main reasons for this appear to be the changes in type of control from V-A to F-C, and the fact that vehicles are not following straight paths that may be readily coordinated using the above-mentioned procedures. The first situation may occur because, in order to coordinate along an arterial street, it is necessary to change the type of control of those signals from vehicle actuated to fixed control. Since in this network the volume variability owing to the incident situation is quite high, and turning movements constitute an important part of the paths, the benefits of coordination are eliminated by using fixed control signals even with progression for the straight direction. The second situation may occur because, in cases of vehicle diversion from the freeway, the most frequently used paths are not straight paths but include turning movements that need to be coordinated as well.

Considering the variability in traffic flow, experiments were conducted using different parameters with the vehicle-actuated (V-A) signals. The purpose of these experiments was to determine the effect that changes in certain parameters of the timing plan may have over the whole network. The parameters that can be changed are the minimum green and maximum green for any given phase. Sensitivity analysis was performed with respect to the maximum green. The minimum green is usually set according to the experience gained in the control of each intersection. Therefore, a simple sensitivity analysis was conducted to observe the network effects of changes in maximum green and, hence, to changes in the maximum cycle length.

For this sensitivity analysis, three cases were simulated. The first corresponds to the base case, referred to as "normal case." In the second case, 10 seconds were added to the maximum green of each phase that had less than 50 seconds of maximum green. Third, 10 seconds were subtracted from to the maximum green of each phase having more than 10 seconds of maximum green. The average trip time in the network for each simulation is presented in Table 4.20

Table 4.20: Sensitivity analysis to changes in maximum green

Simulation	Average Travel Time (min.)
Adding 10 sec.	28.3
Normal Base Case	24.12
Subtracting 10 sec.	34.25

The results presented in Table 4.20 suggest that the control parameters used to simulate the current situation are quite good. Therefore, in order to generate strategies to improve the system performance under congested conditions, it is necessary to modify the current control to explicitly consider those changes in demand caused by the diversion of vehicles coming from the freeway onto the arterial/surface streets.

On the other hand, previous results suggest that changing from V-A to F-C, even while providing progression along straight arterial streets, will not generally improve the situation caused by vehicles diverting from the freeway to the arterial system. Actually, the V-A appear to respond reasonably well to the changes in traffic flow, accommodating the traffic coming from the freeway (owing to the diversion caused by the VMS); it does this by allocating the maximum green time to the corresponding phase and allocating less than the maximum to the other approaches. This is a simple form of self-adjusting control that provides a degree of integration; it also suggests better strategies for system integration.

The integration strategy tested in this and the following set of experiments takes elements of both V-A signals and path-based progression. The main idea is to determine the most frequently used paths when the vehicles divert from the freeway, and coordinate along those paths by increasing the maximum green for the phases where the movements of interest occur.

To test this strategy, experiments were conducted according to the following scheme: First, an initial run was made considering the incident, original paths, and VMS strategy. The results of this initial run were processed to determine the most frequently used paths considering the diversion of vehicles through VMS information. Finally, the maximum green time for the phases affected by the detected movements were increased (thereby allowing for increased throughput) when needed for the phases along the coordinated paths.

In summary, different control strategies were tested. They include the use of off-line programs to calculate coordination along arterial streets, changes from V-A to F-C signals in order to establish a common cycle length for coordination purposes, path-based coordination using the same philosophy of fixed control signal, and, finally, path-based coordination with V-A traffic signals. Among those strategies tested, the path based with V-A signals was the only one that produced improvements over the entire network.

A complete set of experiments was conducted to determine the effectiveness of this control strategy — when used in combination with different route guidance strategies as well. The results of those experiments are presented and analyzed in the next set of experiments.

4.5. FIFTH SET OF EXPERIMENTS

Experiment Description

This fifth set of experiments is designed to test a combined strategy that includes both route guidance and path-based coordination.

In this case, the simulated incident is located on the test bed link (37, 34). The incident starts at minute 20 and ends at minute 80, which means a traffic interruption of 60 minutes. The severity of this incident is set to 0.75, meaning that the blocking effect takes out 75% of the available capacity on that link.

The strategies for information and route guidance (dynamic traffic assignment, or DTA) include those applied in the first set of experiments in this study, augmented by a decentralized approach-to-route guidance proposed by Hawas and Mahmassani (Ref 56). This means that in these experiments the full toolkit of DTA approaches is used. A brief description of each of the DTA strategies follows.

Descriptive Information (DES-DTA): In this case, there are two sources of information. First, all vehicles are assumed to have access to pretrip information. This means that in order to determine the initial path (path when the vehicle is generated), the users rely on knowledge of prevailing conditions in the network. Second, a fraction of the users are assumed to have in-vehicle information and can therefore change their respective paths according to the evolution of the traffic in the network. The percentages of market penetration of in-vehicle information considered for this set of experiments are 25%, 50%, and 75% (in addition to the base case of 0%).

Normative with Multiple-User Classes (NMUC-DTA): Four different classes of users are considered. The first class includes vehicles with prespecified paths, representing those users that do not change paths. The second class includes vehicles with route guidance capabilities that follow the paths given by a central controller after solving for the system optimum. These are vehicles with SO paths. The third class includes the users that have information and follow the paths that minimize their own travel time. These are vehicles with the UE (user equilibrium) paths. Finally, the fourth class includes the boundedly rational users that receive information and process it, but who switch only if the improvement in travel time is significant for them.

Normative with 100% Optimal Route Guidance or System Optimum (NSO-DTA): By definition, the SO solution is the best that can be achieved in the system. As such, it is taken as the benchmark against which to compare the other strategies.

Decentralized Route Guidance (DEC-DTA): In this case, the intelligence for route guidance is distributed and information is provided to the users through local controllers. Each controller has local information about the traffic conditions in the surrounding area. The knowledge level (K) considered determines the size of this surrounding area. Many initial tests were conducted with different knowledge levels. A knowledge level of K=7 was selected because it offers the best results for this network (Refs 42, 56).

The experimental setup considers an incident situation under two different control settings. The corresponding situation without incident was presented in the first set of experiments. The situations considered in terms of route guidance are summarized in Table 4.21. In addition, for each one of these situations, two control strategies are tested. The first control strategy is the one used for the base case. It consists of vehicle-actuated traffic signals with maximum green times that provides a maximum cycle length of 120 seconds. The second control setting is based on the paths followed by the vehicles diverted from the freeway because of the incident. In this case, coordination is provided among the phases that serve the movements along the most frequently used paths. The traffic signals are kept as vehicle actuated, but the maximum green time is modified to allow greater throughput for those movements involved in the diversion paths.

Table 4.21: Route guidance scenarios for experiment No. 5

Route-Diversion Strategy	Characteristics
DES-DTA 0%	0% on-route information
DES-DTA 25%	25% on-route information
DES-DTA 50%	50% on-route information
DES-DTA 75%	75% on-route information
NMUC-DTA	Four different user classes (25% of each class in the network)
NSO-DTA	100% route guidance with the solution that is optimal for the system
DEC-DTA	Decentralized route guidance with local rules

Each particular scenario was simulated using DYNASMART. All performance measures for the entire network, as well as for individual subsystems, were obtained. Of those, the primary performance measure selected for comparison is the average travel time in the network, given that it captures the total savings that can be obtained using ATMS/ATIS strategies.

Table 4.22 presents the numerical results of each one of the experiments for old and new control, as well as the improvement obtained by changing the traffic signal setting for each of the route guidance strategies. This information is presented for all the route diversion strategies considered, which included descriptive, normative, and decentralized strategies. In all cases, there is a positive improvement in travel time using the new control with path-based coordination. In some cases, the improvement appears to be rather small, such as that under the normative SO strategy; yet in others, such as that for the normative MUC, there is meaningful improvement using the new control strategy. However, it should be noted that the paths used for path-based coordination were not recomputed following the SO traffic pattern. In other words, the overall system is not truly fully optimized in this regard.

Another important aspect of these experiments that was initially highlighted in the first experiment is the magnitude of the benefits obtained through route guidance for congested situations. Table 4.23 presents the improvement in average travel time (ATT) generated by each route guidance strategy for each one of the control settings used in these experiments. Using any of the control strategies mentioned, there is an improvement of between 20% and 40% (over the base case of pretrip information only), depending on the route guidance strategy used.

Table 4.22: Relative performance of control setting strategies

Route-Diversion Strategy	Characteristics	Average Travel Time (min.)		Improvement in Percentage
		Old Control	New Control	
DES-DTA 0%	0% on-route info.	24.27	23.98	1.2 %
DES-DTA 25%	25% on-route info.	17.67	17.12	3.1 %
DES-DTA 50%	50% on-route info.	17.92	17.23	3.8 %
DES-DTA 75%	75% on-route info.	18.56	18.01	3.0 %
NMUC-DTA	Four different user classes (25% each)	17.42	16.64	4.5 %
NSO-DTA	100% route guidance	14.69	14.66	0.2 %
DEC-DTA	Decentralized local rules (K=7)	16.28	16.01	1.7 %

Table 4.23: Relative performance of route guidance strategies under different control strategies

Route-Diversion Strategy	Characteristics	Old Control Settings		New Control Settings	
		ATT (min.)	Improv. %	ATT (min.)	Improv. %
DES-DTA 0%	0% on-route info.	24.27	base	23.98	base
DES-DTA 25%	25% on-route info.	17.67	27.2 %	17.12	28.6 %
DES-DTA 50%	50% on-route info.	17.92	26.2 %	17.23	28.1 %
DES-DTA 75%	75% on-route info.	18.56	23.5 %	18.01	24.9 %
NMUC-DTA	Four different user classes (25% each)	17.42	28.2 %	16.64	30.6 %
NSO-DTA	100% route guidance	14.69	39.5 %	14.66	38.9 %
DEC-DTA	Decentralized local rules (K=7)	16.28	32.9 %	16.01	33.2 %

Figures 4.15 and 4.16 present the results for the situations with descriptive dynamic traffic assignment for the old and new control strategies, respectively. In each figure, the first cluster corresponds to the incident case with 0% of the vehicles receiving en-route information and with all vehicles having access to pretrip information. This means that vehicles departing after the incident occurs have information about prevailing trip times at the time of departure. The second cluster shown in the figures corresponds to the case where 25% of the vehicles

receive en-route descriptive information. The three columns (bars) correspond, respectively, to the performance of (1) the vehicles without information, (2) those with information, (3) the combined total for all vehicles in the system. The third cluster in these figures corresponds to the case in which 50% receive en-route descriptive information, while the last cluster corresponds to the case in which 75% receive en-route descriptive information.

Figure 4.17 presents the results for the route guidance strategies tested using both the old and new control strategies. In this figure, the first column (bar) of each cluster corresponds to the case associated with the old control, while the second column corresponds to the new control strategy. This illustrates that the new control using the path-based strategy actually performs better than the previous one for the whole network with any one of the route guidance strategies tested. This figure also allows us to see that the large improvements are obtained mainly with route guidance strategies. The improvements obtained with the new control settings using path-based control are relatively small in comparison.

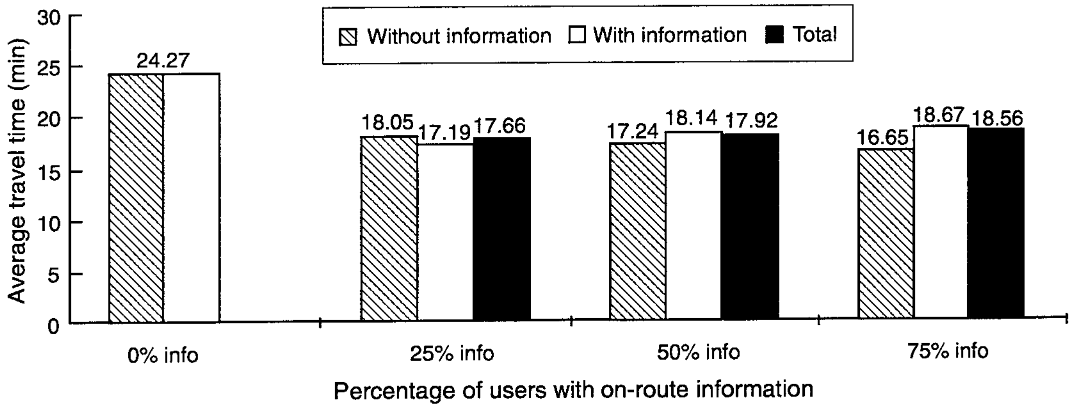


Figure 4.15: Travel time of the descriptive information scenarios — Old control strategy

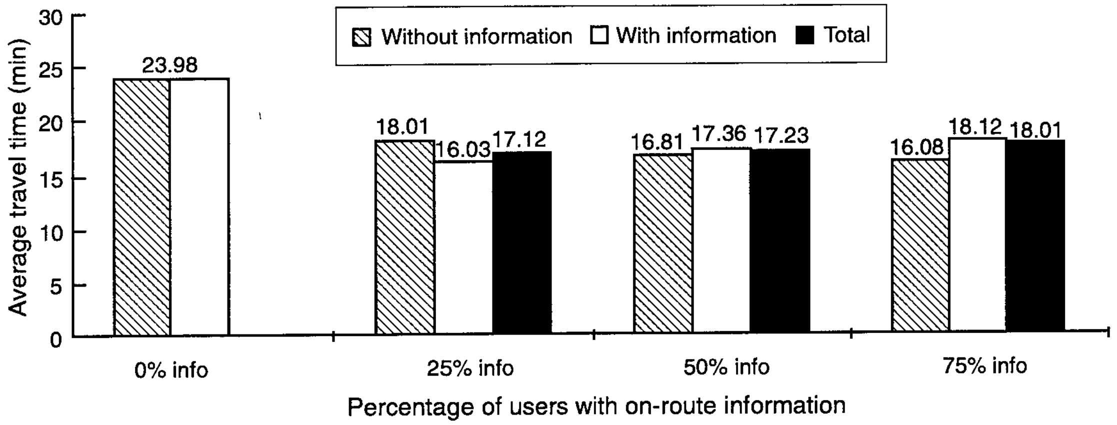


Figure 4.16: Travel time of the descriptive information scenarios — New control strategy

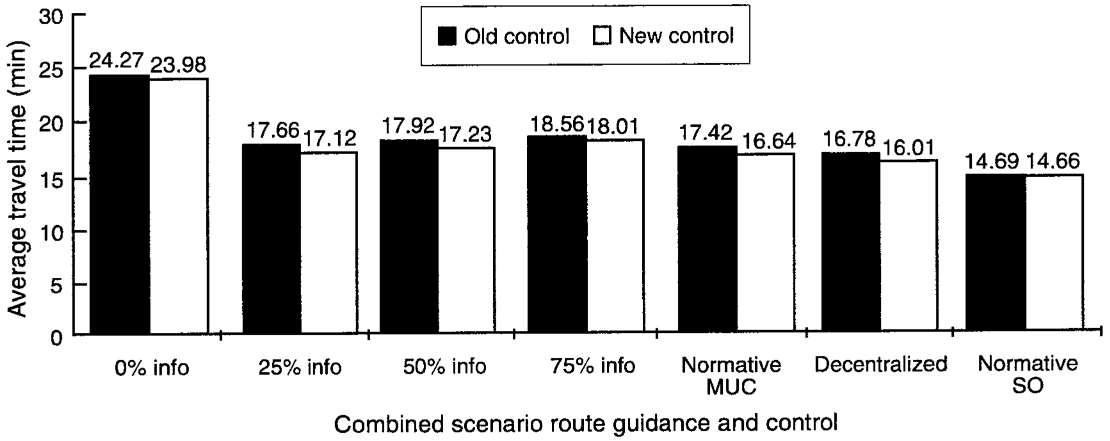


Figure 4.17: Old and new control strategies for different route guidance strategies compared

Analysis

The analysis presented here complements the analysis presented in conjunction with the first set of experiments. In all cases, the performance measure considered is the average travel time for all vehicles in the system. Many interesting insights can be drawn from the tables and figures presented for this set.

The main difference with respect to the first set of experiments is that here a full case is presented interrelating route guidance with a new control setting. In addition, the decentralized route guidance strategy is also considered in this set.

Figures 4.15 and 4.16 illustrate a situation that has been observed previously in various experiments. Owing to the characteristics of the incident considered in these experiments, there are many opportunities for the vehicles with information to select different routes. Because there is no guidance under the descriptive information strategies, the performance of vehicles with information may not necessarily be better than that of the vehicles without information when the fraction of those with information exceeds a certain level (Ref 32). Of the cases shown, only that with 25% of the vehicles receiving en-route information exhibits better performance, on average, than those vehicles without information. In those cases in which 50% and 75% of the vehicles have en-route information, the latter vehicles do worse than the vehicles without information. This suggests that provision of information may not be sufficient by itself: In order to obtain the desired improvements in network performance, there is a need for an integrated strategy that includes information supply and route guidance.

Large improvements can be observed in Figure 4.17 and in Table 4.23. These large improvements are obtained using normative route guidance strategies. At the same time, the improvements obtained with the descriptive strategies are also quite good. Similar to previous experiments, the normative SO case is taken as the benchmark, providing as it does an upper bound of the attainable improvement if all users followed the paths provided under this guidance strategy. On the other hand, the normative MUC (multiple-user classes) scenario performs very well too, even though only 25% of the users are actually following the SO route guidance. It is indeed a remarkable result that the benefits from information strategies, and more generally from ITS, can be attained with only a fraction of the users actually complying with or following the provided advisory information.

Another interesting observation is related to the performance of the decentralized strategy. This case is especially important because the decentralized strategy uses local controllers distributed across the network, just as the signal control strategy uses V-A traffic signals. Therefore, the combination of these two strategies provides a way of using local controllers to perform both traffic control and route guidance functions according to local rules, yielding very good results if the knowledge level and all the required parameters are properly calibrated.

Both the new control settings and the route guidance strategies represent practical benefits to the network that can be obtained by applying the results presented here to the real network. The results presented in Table 4.22 suggest that the improvements obtained through the new control settings vary between 0.2% and 4.5%. The results presented in Table 4.23 reflect the fact that improvements can vary between 23% and 40%. Thus, large benefits could be expected from the implementation of route guidance, while changes in control settings to accommodate changes in demand patterns can be expected to result in comparably more modest (but nonetheless meaningful) improvements.

CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS

5.1. CONCLUSIONS

Traffic congestion in urban freeway corridors places significant demands on our ability to develop and implement integrated control strategies that seek to improve traffic performance. Adjusting the control settings of diamond and arterial signals to receive additional demand has significant implications for the evolution of conditions in the traffic system. Route guidance strategies could be a major source of benefits to the users of urban freeway corridors, especially under nonrecurrent congestion. However, the combination of both control and route guidance strategies has not been explored until recently. This study is among the first attempts to understand the real needs for integrated control and route guidance strategies to produce the improvements expected of ATMS/ATIS technologies.

One significant achievement of the study was the modification made to the computer simulation model DYNASMART: The modified model not only can analyze various interactions of interest, but can also provide possible strategies for handling congested situations in urban freeway corridors, including freeways, frontage roads, diamond interchanges, arterial streets, and local streets. This model is currently the only tool available for studying the dynamics of system performance under both control and route guidance strategies. It can explore such dynamic traffic assignment strategies as (1) descriptive strategies with different levels of information, and (2) normative strategies considering system optimum, user equilibrium, and the more realistic (and general) combination of multiple user classes. It is also possible to test normative route guidance strategies with decentralized intelligence handled by local controllers using local rules that have shown excellent performance and considerable promise in experiments to date.

The development of a detailed database required to simulate an actual network (developed to represent the test bed in this study) is another important achievement of the study. Valuable experience gained from previous studies was used in this study to construct the data set for the Fort Worth test bed. A unique feature of the data set is the level of detail using links of all the subsystems found in this type of corridor. It is important to note the cooperation provided by the Fort Worth District (TxDOT) and the City of Fort Worth's Traffic Division in our efforts to obtain the necessary data. Considerable effort was then required to reduce these data and to ensure their reliability for the purpose of this analysis. A side benefit of this effort was the development of a graphical user interface, including a network and input data editor (dubbed DYNAFACE), to facilitate this process in future applications.

In terms of control strategies, several important accomplishments can be noted. Many tests and simulations were conducted in order to develop control strategies capable of handling diamond intersections, frontage road coordination, and path-based arterial streets coordination.

The TEXAS MODEL was found to be an effective simulation tool for the assessment of isolated diamond interchange operations under various geometries, traffic, and control conditions. The traffic control strategies outlined for diamond intersections effectively reduce delays at congested interchanges where the congestion is caused by freeway incidents. Left-turn prohibitions and two-phase control implementation show reduction in congestion at the intersection level. The reallocation of green time to relatively congested approaches was suggested to improve the situation at the network level.

The TRANSYT-7F program was implemented to handle the next integration level that considers not only the diamond interchange but also the frontage road. Common cycle lengths, but different phasing schemes, were suggested for sequentially located frontage road diamond interchanges. Although these results were interesting, the model's ability to simulate diamond interchanges in a realistic way was found questionable. This conclusion was very important when the strategies were tested with the system integrator model (DYNASMART). The performance obtained using those strategies suggested for handling the situation with diamond interchanges and frontage road did not produce the anticipated benefits over the whole network.

New strategies were also devised and tested, particularly path-based coordination. One of the primary problems found with previous strategies was the need to change from vehicle-actuated signals to fixed-control signals. Even though this is necessary to coordinate along straight paths on arterial streets, it is not always beneficial for the whole network. The final control strategy that produced good results for the whole network consisted of determining the most frequently used paths and, based on that, modifying the parameters of the vehicle-actuated signals along those paths to increase the throughput along the most frequently used path. This strategy improved the average travel time for all the vehicles in the network.

Experiments conducted to test different route guidance strategies produced very good results for the test bed network. With these tools in hand, the final set of experiments was designed to test jointly control and route guidance strategies. Some improvement was found with the application of the new control strategy. This improvement varied between 0.2% and 4.5%, depending on the route guidance strategy used. If no route guidance strategy is used, the experiments suggest that an improvement of about 1.2% could be attained. On the other hand, applying route guidance strategies (providing both information and route guidance) increases the benefits over the whole network. The experiments performed suggest that benefits between 20% and 40% can be obtained, depending on which route guidance and information strategies are used.

The experiments performed as part of this study suggest that, in order to obtain the benefits required by ATMS/ATIS strategies, control and route guidance strategies must be combined. Properly designed route guidance alone can produce meaningful benefits, while control strategies alone produce only marginal benefits. In some cases, control strategies produce benefits at the specific site that is improved, but the effect is not distributed over the entire network. In order to obtain meaningful systemwide benefits, integrated or combined strategies are required. It was noted in connection with these experiments that there exists a considerable difference between the "benchmark" for the system (as indicated by the experiments with 100% of vehicles following system optimal guidance) and the performance simulated under most other situations. These differences suggest meaningful opportunities for improvement in current system operations through properly designed information-based strategies and real-time control.

5.2. RECOMMENDATIONS FOR FURTHER STUDY

This study represents an important step toward developing effective congestion management strategies. The experiments performed have provided insight into the critical factors that must be considered in the implementation of such strategies in actual freeway corridors. The experiments using different route guidance and control strategies provide an organizing structure for continued efforts in this regard. For instance, the experiments have shown the importance of combining strategies that include not only control but also (especially) route guidance to obtain the desired benefits of ATMS/ATIS technologies. However, implementation of such strategies requires additional effort in order to determine how best to provide information and route guidance through all the means that are currently (or soon will be) available. In addition, the implementation of such strategies should be monitored in order to obtain the necessary data to determine if the benefits suggested by the simulation experiments are really achieved in an actual network.

Recently, significant advances have been made in the implementation of information strategies, especially in the context of operational tests; however, information dissemination is still slow. Many questions remain to be answered before these strategies can be applied to actual systems. Because most such questions deal with the implementation itself, an actual laboratory is required to continue testing the implementation of strategies such as suggested in this study. Now that an actual test bed has been successfully set up and simulated, there is an opportunity to implement these strategies on the same network so as to observe actual system performance and actual benefits; in this way, necessary strategy adjustments can be made to ensure a real ATMS/ATIS success.

In terms of control strategies, it would be desirable to continue testing in both the simulated and the actual test bed the kinds of signal control strategies that are coming from the RT-TRAC (Real-Time Traffic Adaptive Control) research effort. The incorporation of new traffic control strategies considering the real-time evolution of the traffic in conjunction with some of the route guidance strategies (e.g., the decentralized local rules) may contribute further to improving network performance. The current test bed represents an important opportunity for testing and developing advanced traffic management methods (as well as the necessary methodological support basis) — *if* it is coupled with data collection and evaluation using the simulation models. The simulation may be used to guide the design and operation of a demonstration project on this test bed, whereas the actual application could provide not only the data required to support model development, but also practical insights for effective development.

REFERENCES

1. Catling, I. (ed.) (1994). *Advanced Technology for Road Transport: IVHS and ATT*, Artech House, Boston.
2. Whelan, R. (1995). *Smart Highways, Smart Cars*, Artech House, Boston.
3. Branscomb, L. M., and Keller, J. H. (1996). *Converging Infrastructures: Intelligent Transportation and the National Information Infrastructure*, MIT Press, Cambridge, MA.
4. Hawas, Y., and Mahmassani, H. S. (1996). *Conceptual Design for Real-Time Traffic Assignment System,* Technical Report ST067-85-TASK C, Center for Transportation Research, The University of Texas at Austin.
5. Jayakrishnan, R., Mahmassani, H. S., and Hu, T. Y. (1994). "An Evaluation Tool for Advanced Traffic Information and Management Systems in Urban Networks," *Transportation Research C*, Vol. 2C, No. 3, pp. 129-147.
6. Mahmassani, H. S., and Jayakrishnan, R. (1990). "Dynamic Simulation-Assignment Methodology to Evaluate In-vehicle Information Strategies in Urban Traffic Networks," *Proceedings of Winter Simulation Conference WSC'90*, New Orleans, LA, pp. 763-769.
7. Tarnoff, P., and Gartner, N. (1993). "Real-Time, Traffic-Adaptive Signal Control," in Yagar and Santiago, *Large Urban Systems*, Proceedings of the ATMS Conference, St. Petersburg, Fl., FHWA, pp. 157-168.
8. Hunt, P. B., et al. (1981). *SCOOT: A Traffic Responsive Method of Coordinating Signals*, Transport and Road Research Laboratory Report, LR-1014, Crowthorne, Berkshire, England.
9. Lowrie, P. R. (1992). *SCATS: Sydney Coordinated Adaptive Traffic System, A Traffic Responsive Method of Controlling Urban Traffic*, Roads and Traffic Authority, Sydney, NSW, Australia.
10. Wattleworth, J. A. (1967). "Peak-Period Analysis and Control of a Freeway System," *Highway Research Record No. 157*, pp. 1-11.
11. May, A. D. (1965). "Gap Availability Studies," *Highway Research Record No. 72*, pp. 101-136.
12. Wang, J. J., and May, A. D. (1973). "Computer Model for Optimal Freeway On-ramp Control," *Highway Research Record No. 469*, pp. 16-25.
13. Payne, Harold, Thompson, J., and Warren, A. (1974). "Allocation of Freeway Ramp Metering Volumes To Optimize Corridor Performance," *IEEE Transactions on Automatic Control*, June.
14. Papageorgiou, M. (1980). "A New Approach to Time-of-Day Control Based on a Dynamic Freeway Traffic Model," *Transportation Research-B*, vol. 14B. pp. 49-360.

15. Pretty, R. L. (1972). "Control of a Freeway System by Means of Ramp Metering and Information Signs," *Highway Research Record No. 388*, pp. 62-72.

16. Yuan, L. S., and Kreer, J. B. (1971). "Adjustment of Freeway Ramp Metering Rates To Balance Entrance Ramp Queues," *Transportation Research* Vol. 5, pp. 127-133.

17. Papageorgiou, M., Hadj-Salem, H., and Middelham, F. (1991). "ALINEA a Local Feedback Control Law for On-Ramp Metering," *Transportation Research Record 1320*, pp. 58-64.

18. Papageorgiou, M., Hadj-Salem, H., and Middelham, F. (1997). "ALINEA Local Ramp Metering: Summary of Field Results," *Transportation Research Record 1603*, pp. 90-98.

19. Goldstein, N. B., and Kumar, K. S. P. (1982). "A Decentralized Control Strategy for Freeway Regulation," *Transportation Research B*, 16B(4), pp. 279-290.

20. Papageorgiou, M. (1983). "A Hierarchical Control System for Freeway Traffic," *Transportation Research B*, 17B(3), pp. 251-262.

21. Isaksen, L., and Payne, H. J. (1973). "Suboptimal Control of Linear Systems by Augmentation with the Application to Freeway Traffic Regulation," *IEEE Transactions on Automatic Control,* AC-18(3), pp. 210-219.

22. Kühne, R. (1991). "Freeway Control Using a Dynamic Traffic Flow Model and Vehicle Reidentification Techniques," *Transportation Research Record 1320*, pp. 251-259.

23. Pinnel, C., and Tutt, P. R. (1963). *Evaluation of Frontage Roads as an Element of Urban Freeway Design*, Departmental Research Report No. 62-3, Texas Highway Department, January 1963.

24. Messer, C. J., and Chang, M. S. (1984). "Guidelines for Installing Traffic Signals at Diamond Interchanges," *Texas Transportation Institute, Research Report 344-1*, Texas A&M University System.

25. Fambro, D., and Urbanik, T. (1992). "Overview of Signalized Interchange Operations," *Conference Session Papers: 1992 TRB Annual Meting*, pp. 16-23.

26. Messer, C. J., and Chaudhary, N. A. (1989). "Evaluation of Arlington, Texas, Diamond Interchange Strategy," *Texas Transportation Institute, Research Report 970-1F*, Texas A&M University System.

27. Messer, C. J., Whiston, R. H., and Carvell, J. D. (1974). "A Real-Time Frontage Road Progression Analysis and Control Strategy," *Transportation Research Record 503*, TRB, National Research Council, Washington, D.C., pp. 1-12, 1974.

28. Herrick, G. C., and Messer, C. J. (1992). "Strategies for Improving Traffic Operations at Oversaturated Interchanges," *Texas Transportation Institute, Report FHWQ/TX/1148-4f*, Texas A&M University System.

29. Institute of Transportation Engineers (1985). *Traffic Control Systems Handbook*, Washington, D.C.

30. Messer, C. J., Fambro, D. B., and Richards, S. H. (1977). "Optimization of Pretimed Signalized Diamond Interchanges," *Transportation Research Record 644*, Transportation Research Board, National Research Council, Washington, D.C., pp. 78-84.

31. Chang, G. L., Mahmassani, H. S., and Herman, R. (1985). "A Macroparticle Traffic Simulation Model to Investigate Peak-Period Commuter Decision Dynamics," *Transportation Research Record 1005*, pp. 107-120.

32. Mahmassani, H. S., and Jayakrishnan, R. (1991). "System Performance and User Response under Real-Time Information in a Congested Traffic Corridor," *Transportation Research,* Vol. 25A, No. 5, pp. 293-307.

33. Mahmassani, H. S., and Jayakrishnan, R. (1988). "Dynamic Analysis of Lane Closure Strategies," *ASCE Journal of Transportation Engineering*, Vol. 114, No. 4, pp. 476-496.

34. Mahmassani, H. S., and Stephan, D. G. (1988). "Experimental Investigation of Route and Departure Time Dynamics of Urban Commuters," *Transportation Research Record 1203*, pp. 69-84.

35. Mahmassani, H. S., Hu, T., Peeta, S., and Ziliaskopoulos, A. (1993). "Development and Testing of Dynamic Traffic Assignment and Simulation Procedures for ATIS/ATMS Applications," *Technical Report DTFH61-90-R-0074-FG*, Center for Transportation Research, The University of Texas at Austin, June 1993.

36. Gordon, R. L., et al. (1996). *Traffic Control Systems Handbook.* Federal Highway Administration, Report No. FHWA-SA-95-032, Washington, D.C.

37. Transportation Research Board (1985). *Highway Capacity Manual.* Transportation Research Board, Special Report 209, Washington, D.C.

38. Wagner, H. M. (1977). *Principles of Operations Research*, Prentice Hall, Englewood Cliffs, NJ.

39. Gartner, N. H. (1982). "OPAC: A Demand-Responsive Strategy for Traffic Signal Control," *Transportation Research Record 906*, pp. 75-81.

40. Gartner, N. H. (1983). "Simulation Study of OPAC: A Demand Responsive Strategy for Traffic Signal Control," *Transportation and Traffic Theory*, eds., Gartner, N. H., and Wilson, N.H.M., Elsevier Science Publishing Company.

41. Peeta, S., and Mahmassani, H. S. (1995). "System Optimal and User Equilibrium Time-Dependent Traffic Assignment in Congested Networks," special edition of *Annals of Operations Research on Transportation.*

42. Hawas, Y. E. (1996). *A Decentralized Architecture and Local Search Procedures for Real Time Route Guidance in Congested Vehicular Traffic Networks*, doctoral. dissertation, The University of Texas at Austin.

43. Robertson, D. (1969). "TRANSYT Method for Area Traffic Control," *Traffic Engineering and Control*, pp. 276-281.

44. Fambro, D., Sunkari, S., and Sangineni, S. (1993) "Implementation Guidelines for Retiming Arterial Networks," *Research Report 1164-4*, study 2-18-89-1164, Texas Transportation Institute, The Texas A&M University System.

45. Chang, E. C., and Messer, C. J. (1991). *Arterial Signal Timing Optimization Using PASSER II-90 – Program User Manual,* Report No. FHWA/TX-90/467-2F.

46. Liu, Y. S. (1997). *Day-to-Day Dynamics of Commuter Behavior under Real-Time Traffic Information*, doctoral dissertation, The University of Texas at Austin.

47. Joseph, T. C. (1996). *Freeway Corridor Control in a Day-to-Day Dynamic Framework*, doctoral dissertation, The University of Texas at Austin.

48. Drew, D. R. (1968). *Traffic Flow Theory and Control*, McGraw-Hill Book Company, New York.

49. Wallace, C. E., White, F. J., and Wilbur, A. D. (1991). "A Permitted-Movement Model for TRANSYT-7F," *Transportation Research Record 1112*, pp. 45-51.

50. Lin, H. J., Machemehl, R. B., Lee, C. E., and Herman, R. (1984). *Guidelines for Use of Left-Turn Lanes and Signal Phases*, Research Report 258-1, Center for Transportation Research, The University of Texas at Austin.

51. Lee, C. E., Rioux, T. W., Savur, V. S., and Copeland, C. R. (1977). *The TEXAS Model for Intersection Traffic Development*, Research Report No. 184-1, Center for Transportation Research, The University of Texas at Austin.

52. Mahmassani, H. S., et al. (1997). *Development of DYNASMART-X Software for Real-Time Dynamic Traffic Assignment*, Technical Report ST067-85-TASK E, Center for Transportation Research, The University of Texas at Austin.

53. Lee, C. E., Machemehl, R. B., and Sanders, W. M. (1989). *Texas Model Version 3.0 (Diamond Interchanges),* Research Report No. 443-1F, Center for Transportation Research, The University of Texas at Austin.

54. Department of Transportation, Federal Highway Administration. (1988). *TRANSYT-7F User's Manual*, Release 6, Version 3.0, October 1988.

55. Hu, T.Y. (1995). *Dynamic Analysis of Network Flows under Advanced Information and Control Systems*, doctoral dissertation, The University of Texas at Austin.

56. Hawas, Y. E., and Mahmassani, H. S. (1996). "Comparative Analysis of Robustness of Centralized and Distributed Network Route Control Systems in Incident Situations," *Transportation Research Board 1537*, pp. 83–90.