

FREEWAY RAMP CONTROL SYSTEM

by

William R. McCasland
Research Engineer

Preliminary Draft

Research Report 24-26 (Final)
Level of Service
Research Study Number 2-8-61-24

Sponsored by
the Texas Highway Department
in cooperation with
U. S. Department of Transportation
Federal Highway Administration
Bureau of Public Roads

August, 1969

TEXAS TRANSPORTATION INSTITUTE
TEXAS A&M UNIVERSITY
College Station, Texas

ABSTRACT

A summary of the development of freeway ramp control systems is presented in this report, which is the last of a series of twenty-six (26) reports on various aspects of freeway operation, traffic surveillance, and control. The theme of this report is that freeway ramp control systems have been adequately tested and evaluated and proven to be effective in improving freeway operations. The design and operation of the control systems will continue to be studied and improved, but considerable benefits can be derived by the freeway motorist from the immediate implementation of these systems on freeways now experiencing traffic congestion.

The suggested warrants, designs, and operational considerations presented in this report are based on three years of operation of the Gulf Freeway Surveillance and Control System and the information derived from conferences with other control projects in the United States. This material and the reports referenced in the pages provide the necessary framework on which a city or state transportation department can plan and develop an urban freeway control system.

TABLE OF CONTENTS

	<u>Page</u>
I. SUMMARY	1
A. The Problem	1
B. Reality of Freeway Control Design	2
C. Recommendation for Implementation	4
II. INTRODUCTION	6
A. Preface	6
B. Control System Theory	7
C. Freeway Ramp Control	8
III. FREEWAY CONTROL SYSTEM WARRANTS	12
A. Justification for Control	12
B. Need for Studies	14
C. Study Techniques	15
IV. DESIGN OF A RAMP CONTROL SYSTEM	25
A. Display	25
B. Surveillance	29
C. Transmission	35
D. Control	41
V. OPERATIONAL CONSIDERATIONS	43
A. Multilevel Approach to Freeway Control	43
B. Selection of Control Program	46
VI. CONCLUSIONS	62
VII. PUBLICATIONS	63

LIST OF FIGURES

	PAGE	
Figure 2.1	Freeway Control System Functional Diagram	8
Figure 4.1	Ramp Control Signal Used on the Gulf Freeway	26
Figure 4.2	Advance Warning Sign Used on the Gulf Freeway	26
Figure 4.3	Third Level Freeway Control With Adaptive Function Components Identified	31
Figure 4.4	Basic FSK Tone Channel	40
Figure 5.1	Decomposition of Freeway Control Function	44
Figure 5.2	Illustration of Gap Acceptance Mode of Ramp Control	48
Figure 5.3	Relationships Among Freeway Demand Rate, Desired Merge Rate and Release Times of Vehicles from the Ramp Signal	50
Figure 5.4	Prototype Merging Controller	53
Figure 5.5	First Generation Merging Controllers	53
Figure 5.6	Schematic of Gulf Freeway Prototype Surveillance and Control System (inbound)	54

I. SUMMARY

A. Problem

Freeway control projects are in operation in Detroit, Chicago, Los Angeles, Atlanta and Houston and other projects are in the planning stages in Seattle and Dallas. The operational systems have resulted in smooth-running freeways that reduce peak-period accidents and delay to traffic in the freeway corridor. The systems have been proven to be highly cost effective.

Why, then, if freeway control is such a good thing, are there so few operational systems? Certainly, there are hundreds of miles of urban freeways that are experiencing extremely poor operation during the peak periods. There are two major reasons why there are not more controlled freeways. First, there is the misconception that the design, implementation and operation of a freeway control system is so complex and expensive that it borders on the impossible. Certainly, the design and operation of a total freeway control system is complex, but not unsolvable. A control system is expensive but cost effective. Lack of understanding must be attributed to the researcher who does not distinguish between the requirements for an operational control system and a research facility. A lack of communication of research findings to the highway practitioner also contributes to the problem.

The second reason is the lack of a specific delegation of authority to actively assume the responsibility for the operation of urban freeways. Most state highway departments have been engrossed in the design, construction and maintenance of highway systems. The operation of urban freeways

has been left to those organizations responsible for enforcement. But the freeway system crosses many jurisdictional boundaries and the coordination of the design and operation of freeway control systems requires supervision by one agency, which in most instances should be the state highway department.

B. Reality of Freeway Control Design

Almost as if by tradition, it has been assumed that a complete description of a freeway's operating characteristics was needed before the control system could be designed and installed. The implication is that all bottlenecks and their capacities must be determined, that trip origins and destinations for the freeway and its ramps must be established -- all before the control system can be designed. In practice, this information has little effect on the design specifications for a freeway control system.

The basic components in a freeway control system are as follows:

Single Ramp Control System

The minimal system for a single freeway entrance ramp consists of one traffic signal, one signal call detector, one merge detector, one freeway conditions detector and one local regulating controller per entrance ramp. The local regulating controller may be used as an interface unit and back-up control system when higher levels of freeway control are added.

Freeway Ramp Control System

For the control of a system of ramps on a freeway, the single ramp

control configuration would be supplemented by the following items: one ramp volume detector, one vehicle-classification detector, and one queue detector per entrance ramp; one detector per exit ramp; one detector per freeway lane between entrance ramps; and preferably a real-time central computer. Control by a central computer would require an adequate data transmission system.

There can be any number of system configurations ranging between a single ramp local control and a full freeway control system, but the system design is affected more by the designated output specifications than by the results of traffic characteristic studies.

Approach

A freeway is similar to any other street and its control is not too different from that of any other street. On a major arterial, intersections become signalized one by one as control is warranted. Initially, the traffic engineer responsible for the operation does not have to be concerned with network theory or control; he merely installs detectors on each intersection approach, the traffic signals, and the controller; he makes initial settings based on geometrics and historical data; and then calibrates the intersection's operation after the system is operational. Eventually, as more intersections are signalized, a complex network problem may evolve that requires a central controller to coordinate the local controllers on the arterial and local streets; but even then the detection system, the signal system, and the local controller requirements do not change. Only the problem of calibration and coordination changes, and if this has been anticipated at the time

of the signalization of the first intersection (first-level control), the network problem is much more manageable.

Freeway control systems can evolve in the same way, entrance ramp by entrance ramp. Eventually, a collection of controlled entrance ramps must be regulated by some higher order of control in the same manner as a collection of intersections. Yet the components employed in the control of the individual entrance ramps are the same ones needed for the control of the entire freeway. The control functions will be built up level by level as more entrance ramps and freeways are brought under control. While this is being accomplished, those responsible for the operation of the freeway facilities will be gaining expertise.

Ideally, new freeways should be designed for control with the detection and transmission systems built-in. In this way, surveillance using these sensors would begin the day the freeway is opened to traffic, and freeway control would be implemented as needed -- certainly before the demand had increased to a point where it exceeded the capacity.

C. Recommendations for Implementation

From the preceding discussion, several conclusions have been drawn in the form of recommendations for the implementation of control systems on urban freeways:

1. Establish freeway operations groups at State Highway District level with the responsibilities of planning, design, operation and maintenance of freeway control system.
2. Establish priorities for freeway control projects in all urban areas.

3. Initiate design procedures for the higher priorities to include:

- (a) Development of plans and specifications
- (b) Conduct of initial traffic studies
- (c) Development of alternative levels of operation (planning for stage implementation)

These activities can be conducted simultaneously.

4. Install and operate the first stage of control.

The procedure beyond step 4 will be one of the objectives of the planning phases of the freeway operations group. The ordering of priorities and the decision of the administrator on expansion or refinement of the system will affect the direction of development.

II. INTRODUCTION

A. Preface

If urban freeways are to operate at acceptable levels of service during periods of peak traffic demand, these facilities must be controlled. At the present time, six freeway surveillance and control projects are in the operational stage or under development. Most notable are those on the Eisenhower Expressway in Chicago, John C. Lodge Freeway in Detroit, and the Gulf Freeway in Houston. Projects in Seattle, Los Angeles and Dallas are under development. As exemplified by these projects, freeway traffic control has evolved into a large complex system with many interfaces and interrelated problem areas.

After more than a decade of promising experimentation with freeway control, in which its feasibility and success have been established, the theory of designing and operating such a control system has not yet reached the stage where a single unified approach has emerged. However, freeway control systems have reached a level of development where implementation is practical and advisable.

This paper reports on the conduct and findings of the Level of Service Project and other related research projects performed on the Gulf Freeway in Houston. Specific recommendations on studies, designs, and operational procedures are presented to assist and encourage the implementation of these research findings on urban freeway systems.

The final report of a project that has been in operation for five years could be an awesome document. Fortunately, the Level of Service Project used the procedure of writing special reports when various phases of the work were completed. There are twenty-five (25) project reports listed in the Appendix.

No attempt will be made to reproduce these reports in this publication. References to the studies and findings will be made to summarize the most significant results of this project, which are the development and implementation of freeway control systems.

In addition to the twenty-five reports evolving from this project, numerous papers, technical memorandums, and reports for other projects on freeway and corridor control systems have been prepared by the staff of Texas Transportation Institute. Discussions and findings of these reports were used in sections of this paper.

B. Control System Theory

A freeway control system is essentially an array of surveillance, communication, and control components designed and connected so as to command or regulate traffic operations. One general scheme for representing freeway control systems is given in Figure 2.1. The blocks C, A, P and D stand for the "controller" (analog and/or digital device), the "actuator" (entrance ramp traffic signals and other traffic control displays), the "process" (freeway traffic operation), and the "detectors" (sensors, surveillance, instrumentation, and measuring subsystems). Two fundamental variables are the system input and system output, denoted by r for reference inputs such as traffic demand, desired speeds, etc. and by c for controlled variables such as volume, density, etc. Since the freeway operations phenomenon must be characterized by a multiplicity of inputs and outputs, r and c are vectors.

In Figure 2.1, d is a disturbance vector which represents unintended inputs to the freeway system which cannot be adjusted such as weather conditions,

accidents, and incidents; m represents information supplied to the controller regarding those disturbance vector components and output vector components which can be instrumented; n represents signals from the controller to the control devices in the actuator subsystem; and u is a manipulatable vector which represents those freeway inputs which can be influenced by control. The vector w represents the broad set of operating specifications, restrictions, and hypotheses pertinent to the control problem. Conceivably, r might be considered as a subset of w except that in some cases it is convenient to distinguish between the two.

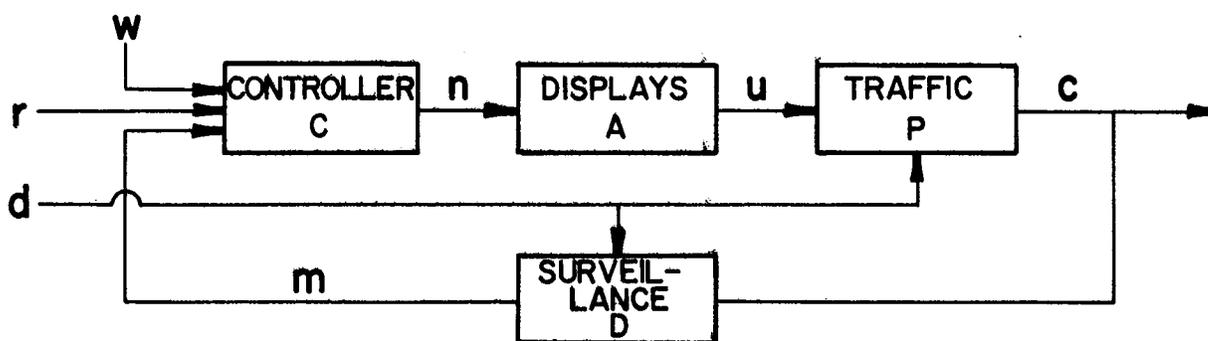


Figure 2.1. Freeway Control System Functional Diagram.

C. Freeway Ramp Control

The Level of Service Project, initiated in 1963, had as the principal objective the development of a system that would reduce and/or eliminate congestion on an urban freeway during peak traffic flow. From the analysis of traffic operations on a freeway system, a basic conclusion was formed; if traffic demand could be controlled so that it never exceeded the capacity of the roadway anywhere in the system, congestion would be eliminated. Further, it was determined that inputs to the freeway lanes were the most logical points

at which to begin the study of control of freeway demand. The area of influence of the control system includes frontage roads and parking areas, but the first step was taken to control the entrance ramps to the freeway.

Control in the glossary of the traffic and transportation engineer is equivalent to prohibition or regulation of vehicle movements. The first approach, prohibition of movement, was applied to the Gulf Freeway entrance ramps in a control experiment to demonstrate the effectiveness of ramp control in improving freeway operations and to show that available capacity does exist on most alternate routes of the freeway. Since adequate alternate routes may not be available to all urban freeways, the decision to install ramp control systems may require modifications in design, operation and control of the arterial street system.

The closing of entrance ramps is a simple but effective means of controlling demand, but many problems may arise from this approach. The design of the closure may be expensive. If the ramp is to be permanently closed, the cost is low, but if the ramp is to be closed only during predetermined time periods, an automatic or manually operated gate will be required. The complete elimination of ramp demand may in some situations result in inefficient operation of the freeway in terms of traffic throughput. The closure of entrance ramps may create social and economic pressures that cannot be tolerated. For these reasons, regulation of traffic demands was selected as the concept for control of the freeway entrance ramps. Rather than closing entrance ramps, traffic signals were placed on the ramps to regulate, or meter, the traffic flow as it approaches the freeway lanes. The question of how to operate the traffic signal has been and will continue to be the subject of many research and operational

studies. This is not surprising since there are many acceptable modes of operation, and many situations that are not adaptable to all modes. For example, the merging control (gap acceptance) mode cannot be used effectively when an exit ramp is located within 1000 feet upstream of the merge area or when an added lane or auxiliary lane is provided at the merge area. On the other hand, there are some situations for which merging control is the best approach, such as on-ramps with very short acceleration lanes.

The following is a list of some of the ramp control concepts that have been employed:

- (1) Time clock with fixed rates of flow
- (2) Merging control with gap detection and projection
- (3) Capacity-Demand with measured flow rates
- (4) Variable flow rates with measured lane occupancy
- (5) Fixed rates selected by measured overrides, such as queue lengths, freeway speeds and delay at signal.

In (1), the control action depends on prior calibration using historical data. This is not a completely satisfactory approach since the system to be controlled, its environment, and the expected inputs cannot be completely described beforehand. Therefore, for the purpose of classification, this simple open-ended control system can be considered as being of "zero" level because it has no feedback and no memory. Another example of this form of control is complete ramp closure. A discussion of classification by designation of control levels is presented in Chapter V, but basically, the highest levels have more extensive surveillance and control systems.

The use of surveillance -- the continual sensing of freeway traffic in

time and space -- represents one means of replacing simple open-loop controls with a closed-loop freeway control system which has a higher sensitivity to such parameter changes as surges in traffic demand and such environmental disturbances as a stalled vehicle or an incident on the freeway. Numbers (2) through (5) require data from surveillance systems of traffic conditions on the freeway and entrance ramp and are classified as first-level control systems because direct feedback is present.

The closed-circuit television systems on the John C. Lodge Freeway in Detroit and the Gulf Freeway in Houston provide feedbacks for the comparison of actual freeway operation with desired operation so that appropriate control action can be taken. However, one of the recommendations of this research is that television surveillance not be used in the freeway control system except to assist in the design and calibration activities.

Other examples of first-level freeway-oriented control are advisory speed messages and lane closure warnings activated in response to a deterioration in freeway operations.

III. JUSTIFICATION FOR FREEWAY

CONTROL SYSTEMS

A. Objectives of a Freeway Control System

Prior to the decision for installation of traffic control systems, the objectives and expected results should be clearly defined. Freeway control systems do not increase the capacity of the freeway; they regulate traffic and travel patterns so that traffic flow will be at or near capacity for longer periods of time. Results which can be expected from a freeway control system are:

1. reduced traffic congestion
2. reduced accident rates
3. decreased travel time
4. improved quality of travel

In summary, freeway control systems increase the probability that freeways will operate in the manner for which they are designed. Therefore, a comparison of existing traffic conditions to the expected or designed traffic conditions will indicate the need for freeway control systems. For example:

If an urban freeway is frequently experiencing traffic congestion at approximately the same location and the same time of day, it can be assumed that the traffic demand exceeds the capacity of the freeway at one or more locations. Traffic congestion is characterized by speeds below 30 to 35 mph, vehicle densities exceeding 60 to 70 vehicle/mile/lane, and flow rates less than 1800 vph/lane. This demand-capacity relationship can

be obtained in two ways: (1) the demand increases until it exceeds the capacity or (2) the capacity is effectively reduced by traffic operations or environmental conditions. It is important to know which of the two conditions cause the congestion.

If the accident experience on the entrance ramps and freeway lanes is high, it may be due to poor operational characteristics on the ramp and/or the freeway. High accident experience may be difficult to define. In Houston during three years of control of eight ramps, no accidents have been reported or observed on the ramps during the time of control and the freeway accident experience in the controlled section dropped 25 percent.

If an urban freeway is forced to operate with traffic flow patterns for which it was not designed, a freeway ramp control system may be necessary to re-establish the planned flow rates on the entrance and exit ramps. A comparison of design volumes with the existing volumes will indicate the change in traffic flow patterns.

In summary, every urban freeway will need a control system at some time. The factors to be resolved are: (1) the limits of the system, (2) the type of control, (3) the type of surveillance, communications and control design, and (4) the system cost effectiveness.

The opinion has been expressed by some engineers involved in freeway control programs that the best approach is to build the control system into

the new freeways and to initiate control when the roadways are opened to traffic. This is a sound approach but until a greater interest is shown toward providing installations on existing congested freeways, the probability of controls on new freeways is small.

B. Need for Traffic Studies

The previous section has implied that only basic traffic data may be sufficient to establish the need for some type of freeway control system. However, there are several reasons why it is also necessary to conduct special traffic studies of freeway operation during the planning, design and operation of ramp control systems:

1. Justification of control systems - Operational and cost effectiveness evaluations prior to installation may be necessary to supplement the studies made from basic traffic data as suggested above.
2. Establishment of priorities - When more than one freeway is to be controlled, a priority list based on quantitative information is required.
3. Planning and design of freeway control systems - Detailed design of freeway and arterial street network surveillance and control require some detailed information on traffic conditions.
4. Calibration of control systems - Adjustments in local ramp freeway system or arterial street control systems are required

to optimize operation.

5. Evaluation of control systems - Feedback for priority and justification studies and for improving designs are essential.

The studies under 1, 2, and 5 are necessary in administrative decisions and those under 3, 4, and 5 are used for operational and design decisions. It is the conclusion of this report that the design of local control of entrance ramps does not require data from operational studies of the freeway whereas the design of surveillance and control systems do.

C. Traffic Study Techniques

Several studies have been employed to describe freeway operations as they relate to freeway control systems:

1. Input-Output Study

After the boundaries of the control section are determined, coordinated vehicle counts at all inputs and outputs to a freeway section provide information on volumes, speeds, and densities. The upstream boundary should be upstream of all congestion in order that the counts at this location represent the demand on the freeway. The downstream boundary should be downstream of present congestion and possible future congestion (if control measures are successful in increasing the flow out of the presently congested system, congestion may develop at some downstream locations). Data from aerial photography can be useful in the selection of the boundaries of the system of interest.

When the system is defined, it may be necessary to divide it into closed subsystems for the analysis. Manpower requirements normally make this step necessary. Each closed subsystem consists of: a freeway input count, a freeway output count, and a count of each of the intermediate entrance and exit ramps.

The subsystems should not overlap, but together they should include the entire system of interest. The division points between subsystems should be chosen to be a known or suspected bottleneck. In this way the freeway counts can be used in the estimation of the capacity of the bottlenecks.

Data from all count locations are recorded simultaneously at regular intervals, such as five minutes, and in cumulative form. In this way the total number of vehicles entering and leaving the system in the time interval can be determined as can the change in the number of vehicles within the system. The number of vehicles in the subsystem is also known every five minutes (each time the data are recorded) since the net number of vehicles crossing the system's boundary is known.

At freeway count locations a second man is used to obtain speed samples -- one sample per lane per minute. In this way, the quality as well as quantity of flow, is determined and the data are used to aid in the location of critical bottlenecks.

Study Limitations - When properly conducted, a study of this type yields valuable data. There are, however, many factors which can void

part of these data.

It is essential that the watches of the study personnel be synchronized before the study so that all count data are recorded at the same time. This is necessary to assure that at each recording time the number of vehicles in the system is accurate.

The accuracy of the counts is extremely critical in studies of this type. A check on this accuracy is available since the total input in the study period should equal the total output in the period plus vehicles stored.

If a surveillance system of vehicle detectors and a central computer are available, the input-output study can be made automatically. Several methods could be used to begin the analysis; however, a method that does not require visual surveillance or manual input is preferable. One such method calculates speed profile using several speed detectors throughout the section. After the program has run five minutes, the speed profile is used to estimate the travel time between the input and the output detectors. The input detectors are started at time t and the output detectors are started at time t plus the estimated travel time.

Recommendation - If a control system must be justified and evaluated on a quantitative basis, this study should be conducted. A study period of five (5) to ten (10) incident-free days is normally sufficient. References 8 and 15 contain a more detailed description of this study.

2. Entrance Ramp Origin-Destination Study

The distribution of entrance ramp traffic to the downstream exit

points is important in the analyses of traffic assignment and estimations of traffic demand. An origin-destination (O&D) survey can be made by the roadside interview method but is very difficult to conduct during peak hours. A technique that has been used with success is the distribution of questionnaires to all drivers using the entrance ramps in the study area. The forms are returned by mail in envelopes provided with the questionnaire. This method is expensive but not as difficult as the interview method. Results of studies in Houston, Detroit, and Chicago show that 50 percent return, which provides for very reliable information, can be expected.

Another technique that uses a mathematical model for estimating the distribution of entrance ramp traffic has been developed. This method is adequate in most instances for determining the O&D information necessary for ramp control program development.

Recommendation - Unless there is a specific need for the trip origin-destination data, the roadside interview or mailed questionnaire should not be made. There are exceptions such as:

- (1) A combined study with distribution of traffic bulletins
- (2) A critical traffic assignment problem on arterial streets.

Reference 7 presents a detailed description of this study technique.

3. Bottleneck Capacity Studies

Volume counts are made at suspected or known bottleneck locations. The count data are recorded every five minutes. Speed samples of all lanes are obtained at the count station at one minute intervals to insure that downstream restrictions do not influence the volume counts. Studies of this type taken at successive intervals will indicate the critical bottleneck sections. The highest 5 to 15 minute flow rates are used to establish the capacity of the section.

Recommendation - This information is essential in the design of a freeway control system. The counts can be taken as a part of the input-output studies. Data taken on three to five days of consistent operation provide a reliable estimate. Capacity estimates for different environmental conditions, such as rain, fog, and snow will be required in some control systems. References 10 and 13 present detailed descriptions of this study technique.

4. Moving Vehicle Travel Time Studies

Travel time runs made at frequent intervals during the congestion period provide the means to (1) estimate total travel time, (2) locate critical bottlenecks, (3) evaluate the system in the perspective of the individual driver and (4) provide data for comparison studies of effectiveness of control. The studies may be conducted in an instrumented vehicle which can provide large quantities of data, or in a private vehicle with manual observations. Speed profiles and contours are developed from these studies. Acceleration noise and fuel consumption studies may be made if an instrumented vehicle is available.

Recommendation - Travel time studies should be taken before and after control systems are installed. These studies provide reliable information and are simple and inexpensive.

The instrumented vehicle should be used if a more detailed analysis of the traffic operations is required. References 10, 11 and 13 present detailed descriptions of these study techniques.

5. Estimation of Demand at Freeway Bottlenecks

Demand-capacity analyses are made at each bottleneck. These data estimate how each bottleneck will operate if it were to operate independently of downstream bottlenecks; i.e., if downstream congestion did not back past the bottleneck. From these analyses it is possible to estimate for each bottleneck, (1) the amount by which the demand exceeds the capacity, (2) the time period for which the demand exceeds the capacity, (3) the duration of congestion, and (4) the number of vehicles that would be stored upstream of the bottleneck because the demand exceeds the capacity there. These represent estimates of what would occur at each bottleneck due only to traffic demand upstream of the bottleneck. The demand is determined by combining upstream input volumes with origin-destination data for these same inputs and with the capacities of upstream bottlenecks. The capacity determination was discussed in section 3.

Recommendation - An estimate of demand should be made prior to the installation of a control system. These estimates will be useful in anticipating diversion volumes for which adjustments in signal timing

may be required on alternate routes.

If origin-destination data are not available, estimates of ramp traffic distribution patterns should be used. Reference 8 presents a detailed description of this study technique.

6. Assignment of Traffic

After the ramp demands have been measured and the expected ramp capacities assigned for the control system, an analysis of the diverted traffic can be made. If O&D data are available, the travel patterns at each ramp can be used to reassign the diverted traffic. Otherwise, a subjective analysis is made using arterial street volumes as a guide.

Recommendation - Some consideration, even if it is only a cursory evaluation, should be given to the arterial street system that approaches each entrance ramp. Detailed studies should be made on the more important arterial streets. Reference 7 presents a detailed description of this study technique.

7. Critical Intersection Capacity Studies

Those intersections that will experience a change in traffic demands caused by the diverted traffic from the freeway ramps are analyzed. Delay and capacity studies are conducted to determine if signal timing or geometric changes will be required. Input-output studies as described in paragraph 1 can be made at these intersections.

Recommendation - Critical intersection capacity studies are an

extension of the re-assignment of freeway traffic. It is important to make these investigations at critical intersections. Transferring the freeway problem to the arterial street is not a solution. Reference 9 presents a detailed description of these studies.

8. Arterial Street Moving Vehicle Studies

The same studies described in 4 should be conducted on the alternate routes of the freeway.

Recommendation - Travel time studies should be taken before and after the control systems are installed. These data are for defensive purposes in that they are used to determine if conditions after control have worsened.

9. Cost-Effectiveness Studies

Cost-effectiveness is an analytic study designed to assist a decisionmaker in identifying a preferred system choice from among possible alternatives. In the traffic engineering context, typical analyses might determine what type of traffic controller should be applied to a particular arterial street; or, in light of this analysis, why type of control system should be installed on a freeway. The basic premise in a cost-effectiveness analysis is that there are alternate ways of reaching an objective and each alternative produces a level of effectiveness for a given amount of resources. The analysis is designed to systematically examine and relate cost and effectiveness of alternative ways of accomplishing an objective.

Cost-effectiveness analysis is specifically directed to problems in which the output cannot be totally evaluated in terms of a dollar value but where the resource inputs can be appropriately evaluated at market prices. Road user level-of-service factors such as comfort, convenience, traffic impedances and quality of ride are typical of outputs of a highway system which currently are not priceable.

Essential Elements of a Cost-Effectiveness Analysis^{2,3} - The essential elements of a cost-effectiveness analysis are:

- Objective(s) (functions to be accomplished).
- Alternatives (feasible way or courses of action for attaining the objectives(s)).
- Costs associated with each alternative.
- A set of mathematical or logical relationships among the objectives, alternatives, environment and resources (models).
- Criteria for choosing the preferred system.

Objectives are the aims that need to be accomplished with the alternate systems; the selection of objectives, therefore, is a basic part of defining the problem that the analysis is designed to solve.

Alternatives are the various proposed systems that are capable of attaining the specified objective(s). Each alternative has its own price tag with respect to facilities, time, men, and money.

The costs associated with alternative systems generally fall into four main categories:

- Research, Development, Test and Evaluation Costs (RDT&E)
- Initial Investment Costs
- Annual Operating Costs

- Attrition Costs

Mathematical models are used in the analysis to cope with the variables that are inherent in problems of the future.

Criteria are the tests used to make the selection between alternatives. There are two widely used criteria in studies for selecting the preferred alternative:

- Equal Cost - This criterion is used when there is a fixed budget, and the analysis determines which alternative gives the greatest effectiveness for the same expenditures of resources.
- Equal Effectiveness - This criterion assumes that a specified and measurable effectiveness (capability) is required and the analysis is to determine which alternative achieves this effectiveness at the least cost.

Ideally, a cost-effectiveness analysis should result in a set of curves which illustrates the relationship between absolute levels of effectiveness of each system and their associated increases in cost.

A complete discussion of the cost effectiveness analysis as it relates to the Gulf Freeway Control System is presented in Research Report 24-24.

Recommendation - Cost studies of a ramp control system should be made as part of the preparation of plans and specifications. In the cost-effectiveness studies for justification of control, the alternatives of no control and the first level (simplest) control systems should be compared by relating to the existing conditions of traffic operations. In the cost-effectiveness of the levels of sophistication of control, effectiveness measures must come from data provided by the experience of other comparable systems and from the experience of the design engineer.

IV. DESIGN OF A RAMP CONTROL SYSTEM

A ramp control system consists of four subsystems:

- Display - Traffic signs, signals, and markings.
- Surveillance - Vehicle detection system
- Transmission - Data and control communications.
- Control - Data processing and signal controller.

The design of these four subsystems can be very simple or very complex. The following sections discuss the basic design requirements, the system that operates in Houston, and some of the alternatives that are available to new control projects at the present time.

A. Display

Figures 4.1 and 4.2 illustrate the signing and signal design and location used in Houston. Variations of these designs, used in Chicago, Atlanta, Detroit and Los Angeles, are summarized in Table 4.1. A description of the display as designed for the Gulf Freeway follows:

1. Signal Head - Three (3) indication signal with 8" lenses. -
The same three indication signal that is standard for other traffic control installations is used for ramp control. The amber indication gives greater flexibility because single or multivehicle control can be used with greater safety. A flashing amber signal also can be used to convey a special message. The use of the amber indication does not restrict the normal single vehicle operation.



FIG. 4.1. RAMP CONTROL SIGNAL USED
ON THE GULF FREEWAY.



FIG. 4.2. ADVANCE WARNING SIGN USED
ON THE GULF FREEWAY.

TABLE 4.1

RAMP METERING SIGNAL INVENTORY

State and City where ramp metering signals are installed or planned	California Los Angeles	Illinois Chicago	Georgia Atlanta	Texas Houston	Michigan Detroit
Number of signal lenses for each ramp	2	2	2	1	1
Number of signal lenses for each signal face	2*	2	2	3	3
Color of signal lenses	Red & Green	Red & Green	Red & Yellow	Red, Yellow & Green	Red, Yellow & Green
Longitudinal Location of signal standard	Stop Line	Stop Line	Stop Line	Stop Line	Stop Line
Mounting height of signal faces	4'-6" & 10'-0"*	5'-0"	5'-0"	5'-0"	4'-6"
Period of day that ramps are metered	Peak Periods	Peak Periods	Continuously	Peak Periods	Peak Period
Indication when ramps are not metered	Steady Green	Steady Green	Not Applicable	Off	Off

*At present the ramp metering signals in California are three section heads with red, yellow, and green lenses. The mounting heights vary. Permanent installations being planned will have two signal faces with red and green lenses. The two signal faces will be mounted on one signal standard with the bottom and top signal faces having 4'-6" and 10'-0" mounting heights above the base.

2. Number of Signal Heads - One (1) signal standard to the left of the ramp. - The number of signal heads on a ramp should be determined by the economy, need and desired reliability. Sight distance restrictions and high speed approaches to the ramp may necessitate two or more signal heads; for example, the S. H. 225 entrance to the Gulf Freeway. The effect of a burned out bulb is not as critical for ramp control installations as intersection control, because there are no conflicting movements involved at the signal.

3. Mounting Height - Five-(5)-foot mounting height to the bottom of the signal head. - A mounting height of approximately 5 feet has become standard for all ramp metering systems. This permits the ramp driver to view the signal through the front or side windows.

4. Sign Message at Signal - "STOP HERE ON RED" - The messages on the sign mounted at the signal vary from state to state but all convey the same meaning. Its purpose is to position the lead vehicle over the signal detector and to improve obedience to a signal located in an unfamiliar location.

5. Advance Warning - Warning signs and flashers - When the signals are not in operation for part of the day, the advanced warning devices are necessary. The devices are advisable under any circumstances because of the lack of experience and the unfamiliarity of the drivers with this type of control system. Sign message "RAMP

METERED WHEN FLASHING" is clear and concise.

6. Location of Control Devices - The location of the signs and signals will be dependent on the geometrics of the ramp and the type of control logic to be used. But in general, the preferred locations are:

. Traffic signal -- 200 to 250 feet upstream of the nose of the entrance ramp and on the left side of the ramp. The nose is defined in this statement as that location where a ramp driver can first encroach on the outside lane.

. Advanced warning sign -- 200 to 250 feet in advance of the ramp signal.

The differences in design as noted in Table 4.1, are minor with the exception of the number of signal indications and the type of operation in Atlanta. National standards will probably be formulated in the near future to eliminate or reduce these differences.

Recommendation - Based on the experience in Houston, it is the recommendation of this report that the design, as shown in Figure 4.1 and 4.2, described in this section, and used on the Gulf Freeway be incorporated in new ramp control projects until standards have been developed and incorporated in the manual on Uniform Traffic Control Devices.

B. Surveillance

Electronic surveillance in the form of a vehicle detection system is necessary in all but the simplest form of ramp control. The number, type

and location of the vehicle detectors are dependent on the control system selected. Figure 4.3 shows what is now considered the ultimate design for a detection system. The following sections discuss the design and location of each detector. The justification for these detectors will be discussed in the section on control subsystems.

1. D_q - a queue detector that senses the presence of vehicles waiting to enter the freeway. The area of detection should be 20 feet long and 6 feet wide, and positioned in the center of the lane. In the event a double queue is permitted, the detector should be placed in the predominant lane, which will probably be the left lane.

The detector is placed a minimum of 50 feet downstream of the intersection, ramp or driveway it protects.

2. D_i - an input detector that senses the presence of vehicles at the head of the queue at the signal. It is a demand detector which places the call to the signal controller. The area of detection should be at least 20 feet by 6 feet. If a double queue is permitted, a similar detection area is provided in the second lane.

The detector is placed so that the trailing edge of the detection area is 10 feet upstream of the traffic signal and, for a 20-foot detector, the leading edge is 30 feet upstream of the traffic signal.

3. D_t - a vehicle classification (truck) detector that senses the presence of vehicles on the entrance ramp downstream of the traffic signal. This detector serves several purposes: (1) count-

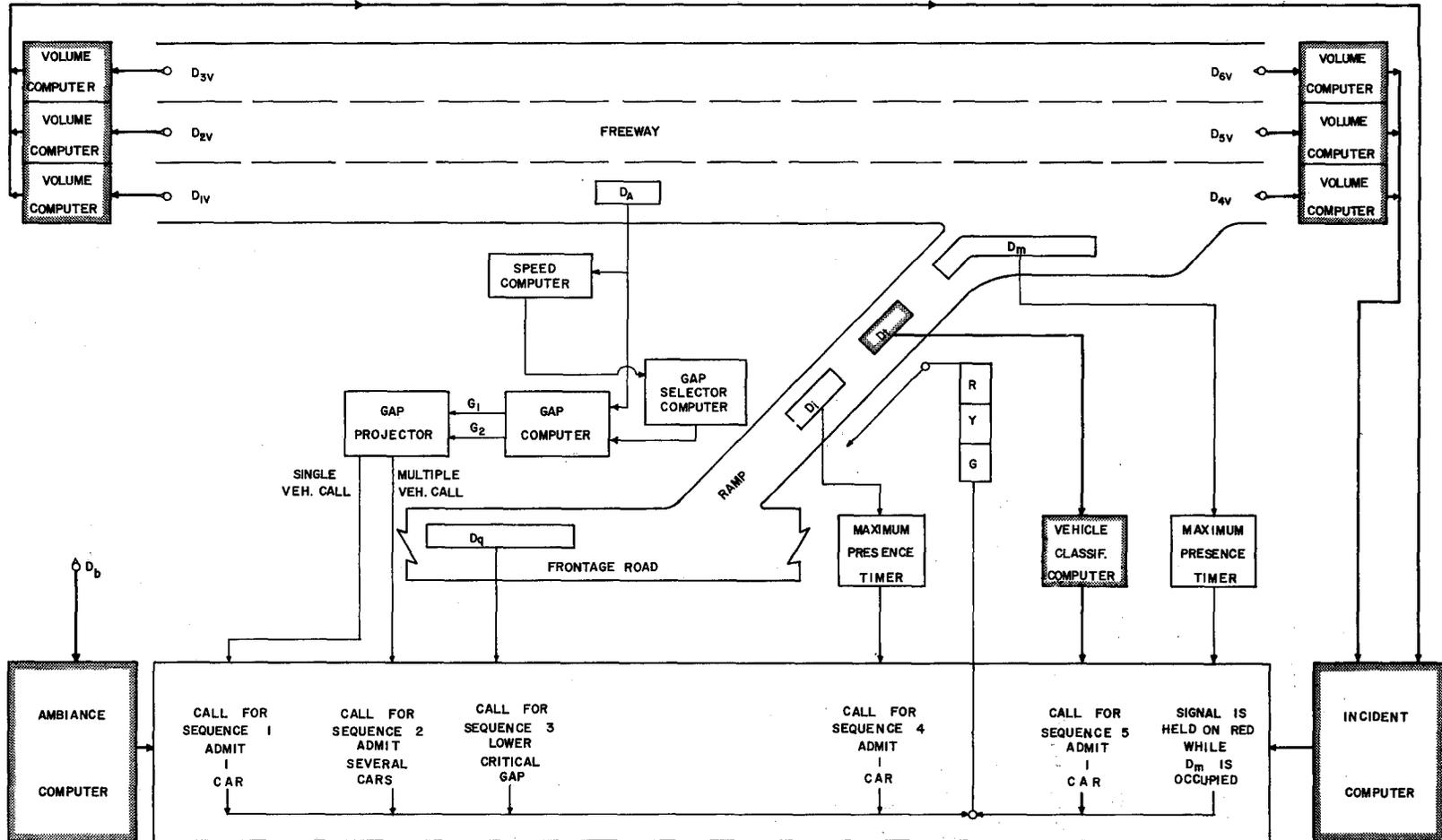


FIGURE 4.3. Third level freeway control with adaptive function components identified.

ing detector for measuring ramp volumes, (2) travel time detector for measuring the acceleration characteristics of vehicles leaving the ramp signal, (3) vehicle classification detector for distinguishing between trucks and passenger vehicles. The area of detection depends on the type of detector used and the purposes of the detection. If, for example, only (1) and (2) are required, an area of detection from two to six feet square is sufficient. (It may be necessary to combine the vehicle classification function (3) with the D_1 detector to alter the control logic.)

The detector is located 50 to 100 feet downstream of the traffic signal in the center of the ramp.

4. D_m - a merge detector that senses the presence of vehicles stopped at the end of the entrance ramp and beginning of the acceleration lane. The area of detection should be 40 to 50 feet long and 6 to 8 feet wide.

The detection area should be so positioned that a vehicle which would block the ramp near the ramp nose would be detected.

5. D_a - an approach detector that senses the presence of vehicles in the outside freeway lane. The approach detector is used to measure traffic speed and vehicle spacings in time. Speed can be calculated from the time an average length vehicle occupies the area of detection. However, since time measurements based on the speed will be used in the control, it is recommended that a speed trap, formed by two detectors 20 feet apart, be used.

The D_a is located upstream of the entrance ramp nose at a distance equal to or greater than d ; where

$$d = S_{\max} (TT_R + G_{\max})$$

and

S_{\max} = Highest speed in the outside lane during time of control (approx. 50 MPH)

TT_R = Ramp travel time from the ramp signal to the merge area (approx. 10 sec.)

G_{\max} = Maximum gap setting in the controller (approx. 5 sec.)

The detection area should be positioned so that only traffic in the outside lane is detected.

6. $D_{1v}, D_{2v}, D_{3v}, (D_{4v}, D_{5v}, D_{6v})$ - volume detectors that sense the presence of vehicles in the freeway lanes. The detectors serve several purposes: (1) counting detectors for measuring traffic flow by lane, (2) presence detectors for measuring average traffic speed by lane, (3) presence detectors for measuring percent occupancy by lane.

The detection areas are positioned in the center of each lane so that individual counts by lane can be made accurately. The location with respect to the entrance ramps will depend on other factors in the design of the total system. Usually the detection stations are located at the adjacent exit or entrance ramp.

7. D_6 - an ambience detector that senses changes in atmospheric conditions that affect traffic flow, such as rain, snow, ice or fog.

The detector is generally located at bottleneck locations that will first be affected by change in conditions. The exact type and locations of detectors will depend on the function served.

It has been the experience of all freeway control projects that detectors which measure presence provide more usable data than other types. For example, if the primary function of a detector is counting, the measurement of presence for average speed or percent occupancy can be a byproduct that is incorporated into a control program.

Three types of detectors have been used in Houston: induction loops, ultrasonic and magnetometers. Each type performed reasonably well, and the advantages and disadvantages were noted:

- . Loop detectors are more easily applicable to large detection areas, and the area of detection is more precise than the sonic detector. Some problems in tuning were noted in Houston to obtain the accuracy needed to measure speeds between two detectors. This problem is probably not confined to loop detectors.

- . Ultrasonic detectors require a mount above and/or to the side of the roadway. The sidefire position was not acceptable for the measurement of presence because of the inconsistency of the target provided by the side of the vehicle.

- . Magnetometers are more expensive but require less installation time and cost. Although these units have not been used extensively, they should be given consideration for application to freeway surveillance systems. One notable advantage is the long cable runs

that can be used between the detector head and the amplifier.

Recommendation - The loop detector used for most detection functions by the existing freeway surveillance and control projects should be the first consideration for design. However, magnetometers have special features that are attractive in the design of a freeway system.

C. Transmission

The transmission system is defined as the equipment necessary to transmit relay contact closures from the detector amplifiers to the ramp controllers and to transmit the response from the controllers to the displays. There are several feasible transmission systems that can be considered. The decision on which system to use will have to be based on economic and technical evaluations of each installation. One of the first determinations will have to be the type of system control with respect to location; local or centralized.

A local control system collects and processes the data and determines the control logic at a location near the control display. Because of the short distances between the detectors, controllers and displays, direct wire connections are the most feasible type of transmission system.

A centralized control system is an extension of local control. Data is transmitted to a remote controller, the control logic is applied, and the control functions are returned to the local controller. Studies of different types of transmission systems become more important as the distances and quantity of data become greater. A general discussion of equipment and factors influencing system design will follow.

1. Types of Carriers

a. Direct Cable

An extension of the hard wire system of the local control system to remote control centers has proven to be an effective means of data transmission. The Gulf Freeway Project utilizes this system over cables owned by the Texas Highway Department. Many other control systems of this type utilize the telephone network for the interconnect from the local control system to the remote center. This system of direct wire connection should be the first alternative in an evaluation process.

b. RF Systems

The transmission of data by radio is feasible. However, the lack of frequencies may eliminate this alternative in the urban areas. There are many other disadvantages that remove these systems from serious consideration: cost, complexity, maintenance, reliability, extraneous interference, and bandwidth limitations.

c. Microwave

Microwave systems are widely used for data transmission but the characteristics of traffic surveillance and control systems are not readily adaptable to make this type of data system cost effective. The data points are too widely spaced along the freeway. If, however, there were central receiving points from which all data were transmitted to a remote computer center, microwave should be considered.

Good resolution of data is available and reliability and freedom from interference is adequate. Specialized maintenance is required and

the licensing of transmitters is a difficult and lengthy process.

2. Multiplexing Techniques

Telemetry transmission systems can utilize the principle of multiplexing, which means dividing a communication medium into pieces or slots, each capable of carrying information from a separate input. Multiplexing may take place by time or frequency division.

a. Frequency Division Multiplexing

Frequency division is used in the parallel mode of transmission, with each channel assigned a specific frequency band. (By using the principle of frequency division, it is possible to have a leased or privately owned voice grade line provide a number of low speed lines.) Each low speed line is allocated its own individual frequency. This allows independent, simultaneous transmission on each frequency band.

b. Time Multiplexing

Time multiplexing is the dividing of the medium into discrete time slots, each of which is capable of carrying information from a different input.

Several lines supply simultaneous and independent signals to a scanning device which then assembles a composite signal on a single line. The advantage of such an arrangement is that a single line can service multiple lines terminating at the scanning device.

The character assembled from the low speed lines is a composite made up of one bit from each of the automatically scanned lines. This composite character is then transmitted to the remote end and auto-

matically demultiplexed. This technique is sometimes referred to as concentrating by time division multiplexing.

Time division multiplexing by method of addressing remote units has some disadvantages: (1) a remote unit cannot economically be installed for a cabinet containing relatively few inputs and outputs, (2) the interactive role of the computer in the communication system creates an overhead problem, (3) the coding of bit trains for maximum protection against errors causes reduction of the effective data transmissions rate, (4) turn-around time for query and receipt of data are high and, (5) interface modems (modulator/demodulators) and party line interface equipment add to system costs.

3. Audio Tone Equipment

If the carrier frequencies are in the audio range for a frequency division multiplexing system, then the transmitting and receiving apparatus is called audio tone equipment. This parallel mode of transmissions is advantageous to the extent that functional events are not time constrained and are transmitted independently of each other. The parallel mode implies that all functions may occur and be transmitted simultaneously and within a reasonable response period. This simultaneous occurrence is, of course, subject to the limits of the logic design, as for instance, the inability of the system to display red, amber and green lights concurrently at a given signal.

A "tone channel" consists of a tone transmitter and a tone receiver. The tone transmitter is simply an audio signal generator whose frequency is precisely controlled and which is arranged for on-off frequency-shift keying.

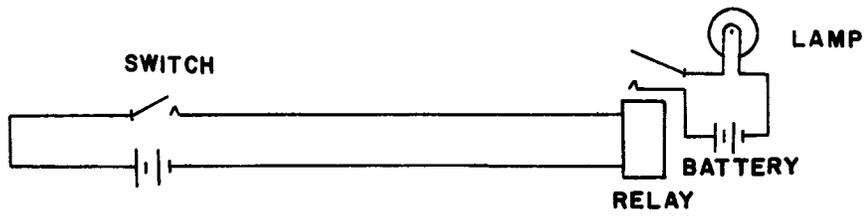
The tone receiver is similar to a radio receiver except that it is tuned to intercept a specific audio frequency signal. Instead of driving a loud-speaker, the tone signal is converted into a DC signal and used to control a relay or electronic switch.

In an FSK (frequency shift keyed) tone channel, the frequency of the tone transmitter is shifted by switch contacts (as, for example, the vehicle detector amplifier relay) and the receiver responds to a shift from one specific frequency to another specific frequency by energizing a relay or delivering a DC voltage. This represents the interface to the traffic control computer. Figure 4.4 illustrates basic tone channel operation.

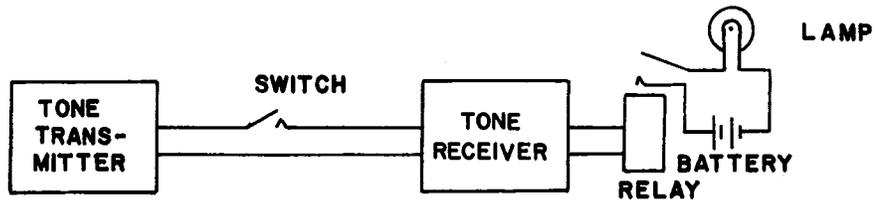
Tone signals can be transmitted over a leased telephone circuit, coaxial cable or almost any pair of wires used exclusively for the tone transmission; or over a radio circuit. Up to 30 low-speed or up to 18 75-band tone channels can be transmitted over what is known as a 300-3000 cps voice circuit. Tone signals can be transmitted simultaneously over a circuit carrying power or other intelligence.

For systems which have a large number of centralized inputs, it can become uneconomical to utilize large numbers of discrete tone channels for transmitting binary information. With careful analysis of data resolution requirements, it is possible to justify a scanner monitoring system to continuously and sequentially sample the condition of remote relays or voltages.

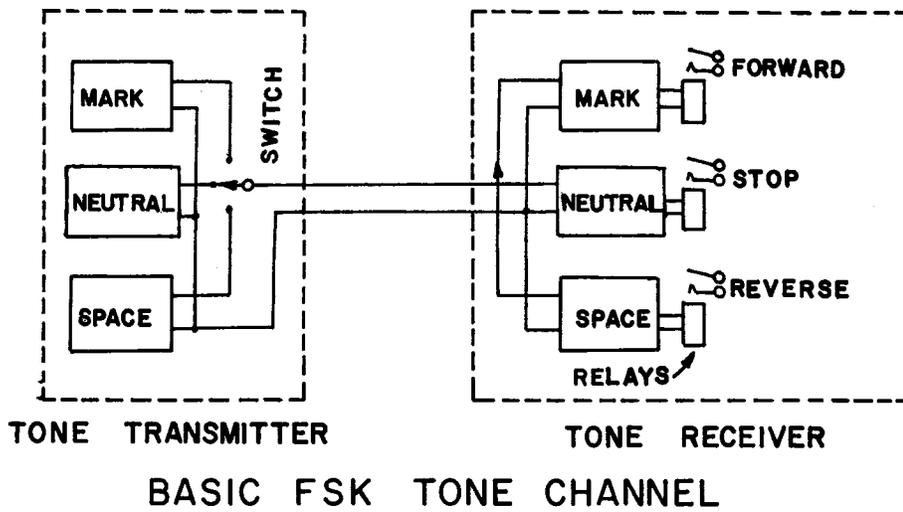
Recommendation - It is not possible to give a general recommendation on the type of transmission system to use. Each project will require a separate analysis to determine the appropriate candidate systems, and an



BASIC D.C. CONTROL CIRCUIT



BASIC ON-OFF TONE CHANNEL



BASIC FSK TONE CHANNEL

FIGURE 4.4.

economic evaluation to determine the final design. However, the audio tone telemetry equipment should receive serious consideration in the design of a freeway traffic control system. Multiplexing is almost mandatory for signal control in outlying areas, and it is also important to consider the flexibility of the telephone system in augmenting control functions as required. Detailed discussions of the systems mentioned above are contained in References 52 and 53.

D. Control Subsystem

The design of a freeway ramp control system is affected by the location of the controller and the detection system required to perform the selected control functions. If a local control system is proposed the transmission system will be greatly simplified. On the other hand, a central control system will require an extensive system to transmit the data and the control functions.

The detection system is a direct function of the control system. For example, if the gap acceptance operation is to be employed, gap and speed detectors are required. A more complete discussion of the control programs and their required detection system is presented in the next chapter.

1. Design Considerations

The design of the signal controller is difficult because there are several different types of control units that can meet freeway control requirements and perform other tasks as well. The multiple use of the control unit requires that both technical and administrative operations be considered in the design. However, the control requirements should determine the type of controller. Conversely, the specifications of the controller should not limit the control subsystem. Some of the decisions that have to be made in the design of the controller are:

- a. The degree of control - The multilevel approach to freeway control is discussed in Chapter V.
- b. The control program - The philosophy of control as discussed in Chapter V.
- c. The application of the controller to other functions - If a digital computer is selected for control, other functions, such as data logging, data analyses, administrative functions are usually considered because of cost-effectiveness considerations.

An example of the analyses necessary in the design of the controller is presented in Reference 24. Other studies that are concerned with the hardware aspects of the control systems are presented in References 25, 39, 41 and 42.

2. Types of Control Systems

- a. Analog Satellite System - A local control system that uses analog computers to control each entrance ramp independently of adjacent controllers.
- b. Digital Satellite System - A local control system that uses digital computers to control each entrance ramp independently of adjacent controllers.
- c. Central Digital Computer System - A central system that coordinates the operation of all signals in the system.

Recommendation - After the control functions have been established, a set of candidate controller systems should be established. An economic evaluation of the systems will be necessary to determine the final design.

V. OPERATIONAL CONSIDERATIONS

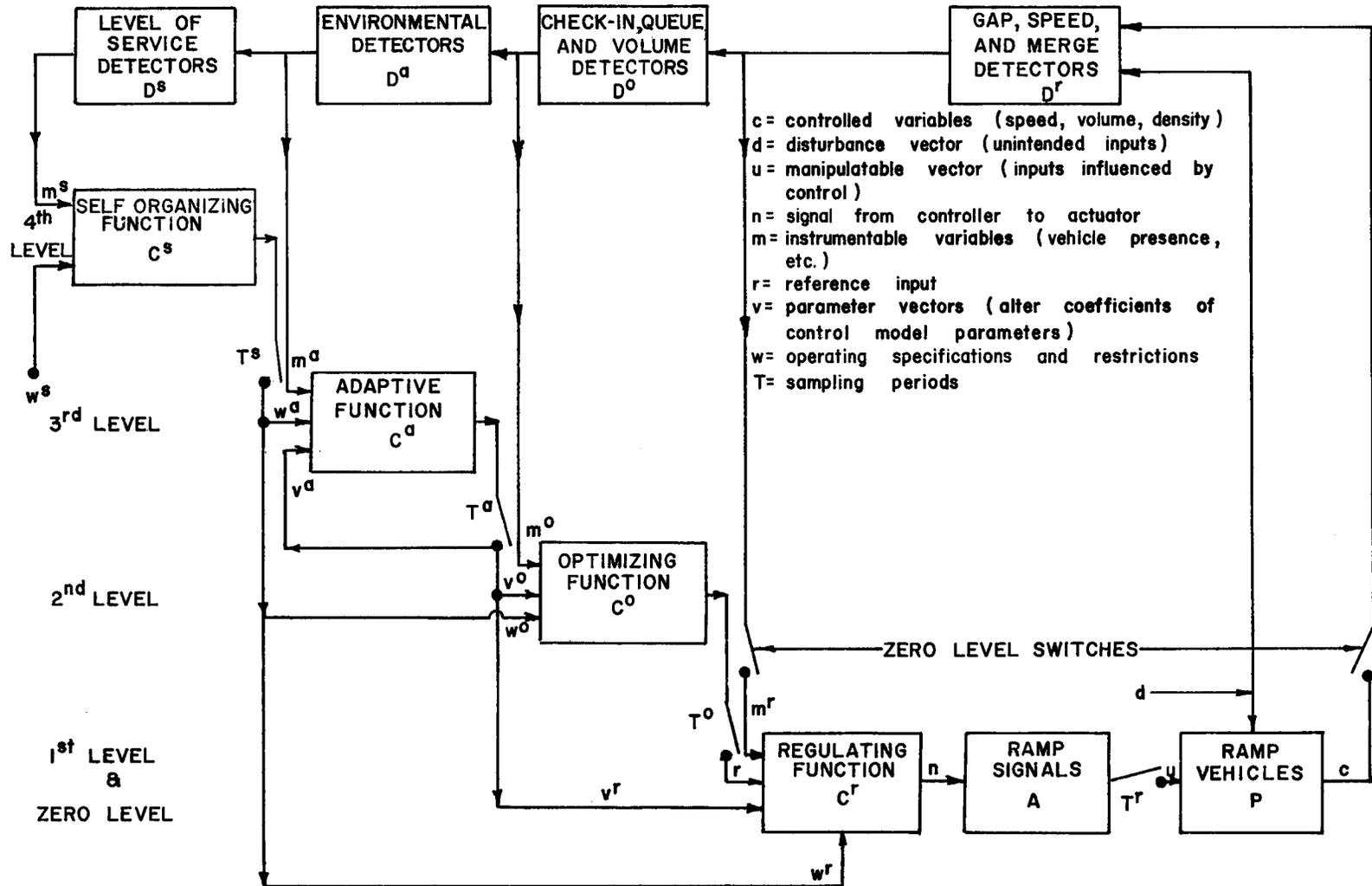
A. Multilevel Approach to Freeway Control (Reference 33)

Multilevel systems is a relatively new concept for providing a rational means to develop control configurations for extremely complex systems. The multilevel approach is directed towards establishing a hierarchy of control that results in an efficient system and one that can be implemented in stages. The lower levels of the hierarchy are directed to recognizing the influence of short-term factors, whereas the higher levels are reserved for factors that influence performance on a long-term basis. There is also a certain hierarchy based on the complexity of computation, the degree of uncertainty and of unscheduled events, and the required speed of reaction to a change in operating conditions.

Figure 5.1 illustrates the conceptual form of a four-level control configuration developed for merging control systems. These levels are, in ascending order of sophistication: the regulating, the optimizing, the adaptive, and the self-organizing control functions. The basic control activities associated with each level are identified in the following paragraphs.

1. The Regulating Function

A controller design to operate at this level accomplishes what might be called the basic subgoal of the control system. Although the goal of the control system is to optimize the level of service on the freeway, various subgoals have been advanced upon which the regulating control subsystem may be based. The assumption that optimizing the subgoals will optimize the primary performance criterion is implied. The



77

FIGURE 5.1. DECOMPOSITION OF FREEWAY CONTROL FUNCTION.

optimal use of available gaps in the freeway merging process is such a subgoal, and it is accomplished by the regulating controller in the block diagram in Figure 5.1. This controller translates the directions of the higher level controllers into direct actions on the operation (the timely release of ramp vehicles by the ramp signal).

2. The Optimizing Function

The object of this controller (Figure 5.1) is the determination of optimum, operating conditions based on the appropriate performance criterion and mathematical model of the process. For example, if the acceptable gap setting on the regulating controller is too high, many marginal gaps are left unfilled; on the other hand, if the setting is too low, many metered vehicles will reject the gaps and be forced to stop in the merging area where their presence, as detected by a loop detector, preempts metering. Obviously, the optimum gap setting is somewhere between "too high" and "too low".

3. The Adaptive Function

While the two lowest levels of the control hierarchy were developed through mathematical models approximating the real system, the adaptive function in Figure 5.1 serves to compensate for the errors introduced by the models by adjusting the parameter values, v^o and v^r . Note that a parameter vector v^a is supplied to the adaptive controller so that, in effect, the controller can "see" what it has been doing. The parameter vectors v^r , v^o , and v^a , in effect, alter the coefficients of the control laws that are applicable at the lower control levels, but do not change

the laws themselves.

4. The Self-Organizing Function

This controller determines what the "worth" or "decision" vectors w^r , w^o , and w^a should be on the basis of those measurable freeway characteristics m^s and the intervention of human factors into the system as represented by w^s . The worth or decision vectors generated by the self-organizing function act to control the overall system in achieving the best total system performance. These decisions are based on the accumulated experience and understanding of the system.

A similar treatment of the decomposition of the freeway control function can be applied to other ramp control theories that have sub-goals different from the one just described.

B. Selection of Control Program

Freeway ramp traffic is controlled by the operation of traffic signals on the entrance ramps. (Figure 4.1) The basis on which the decision is made to change the signal indications will depend on several factors: (1) the philosophy of the control program, (2) the design of the hardware system, (3) the condition of traffic operations and, (4) the jurisdictional and administrative program of control. As these factors are discussed in the following sections, reference is made to Figure 4.3 and the discussions on surveillance and displays in Chapter 4.

1. Philosophy of Control

There are two general categories of control: pretimed control and traffic responsive control. Under traffic responsive systems, there are numerous variations of control programs utilizing various combinations of

traffic parameters.

a. Pretimed Control

From historical records, acceptable demand levels can be determined for each input to the freeway. A controller, either local or centralized, can be programmed to produce the ramp flow rates determined by the acceptable demand pattern. Real time information of traffic conditions on the freeway lanes is not provided to the controller.

The system consists of a display and control unit, which may be started by a time clock or detection device. If the system has a detector at the signal (D_1), it is traffic actuated and the signal will cycle only on a call from a ramp vehicle.

b. Traffic Responsive Control (Reference 19)

(1) Gap Acceptance (Merging) Control - The control of the ramp signal is based on the detection and projection of acceptable gaps from the freeway approach detector (D_a). An acceptable gap is one which exceeds the control parameter (critical gap) established by traffic conditions on the freeway. The projection rate is established by speed measurements made at D_a . The concept, illustrated in Figure 5.2, is to match the travel time of the ramp vehicle with the travel time of the acceptable gap to the merge area.

(2) Demand-Capacity Control - The control of the ramp signal is based on the demand rate as measured by $D_{1v} + D_{2v} + D_{3v}$ and the

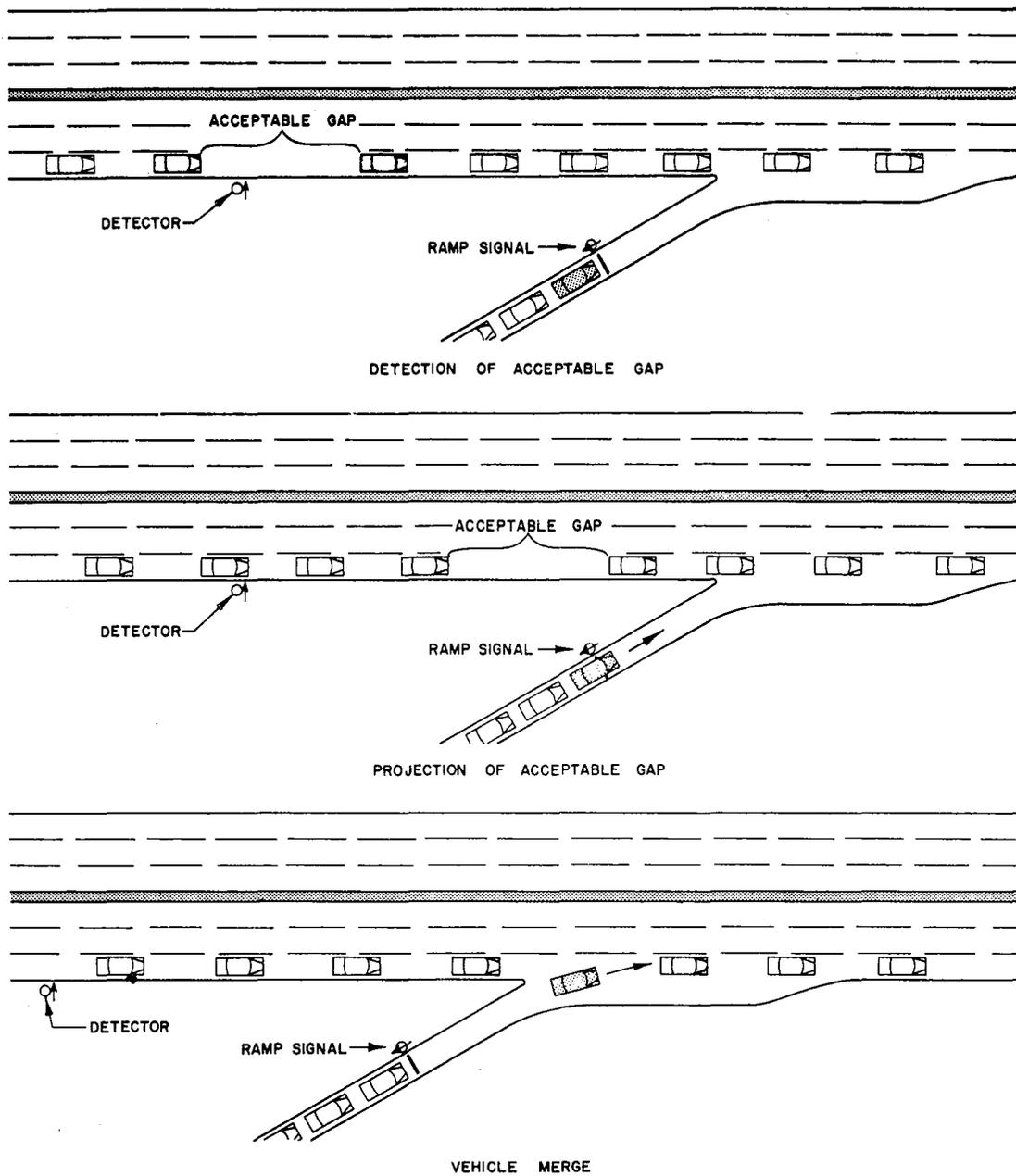


ILLUSTRATION OF GAP ACCEPTANCE MODE OF RAMP CONTROL

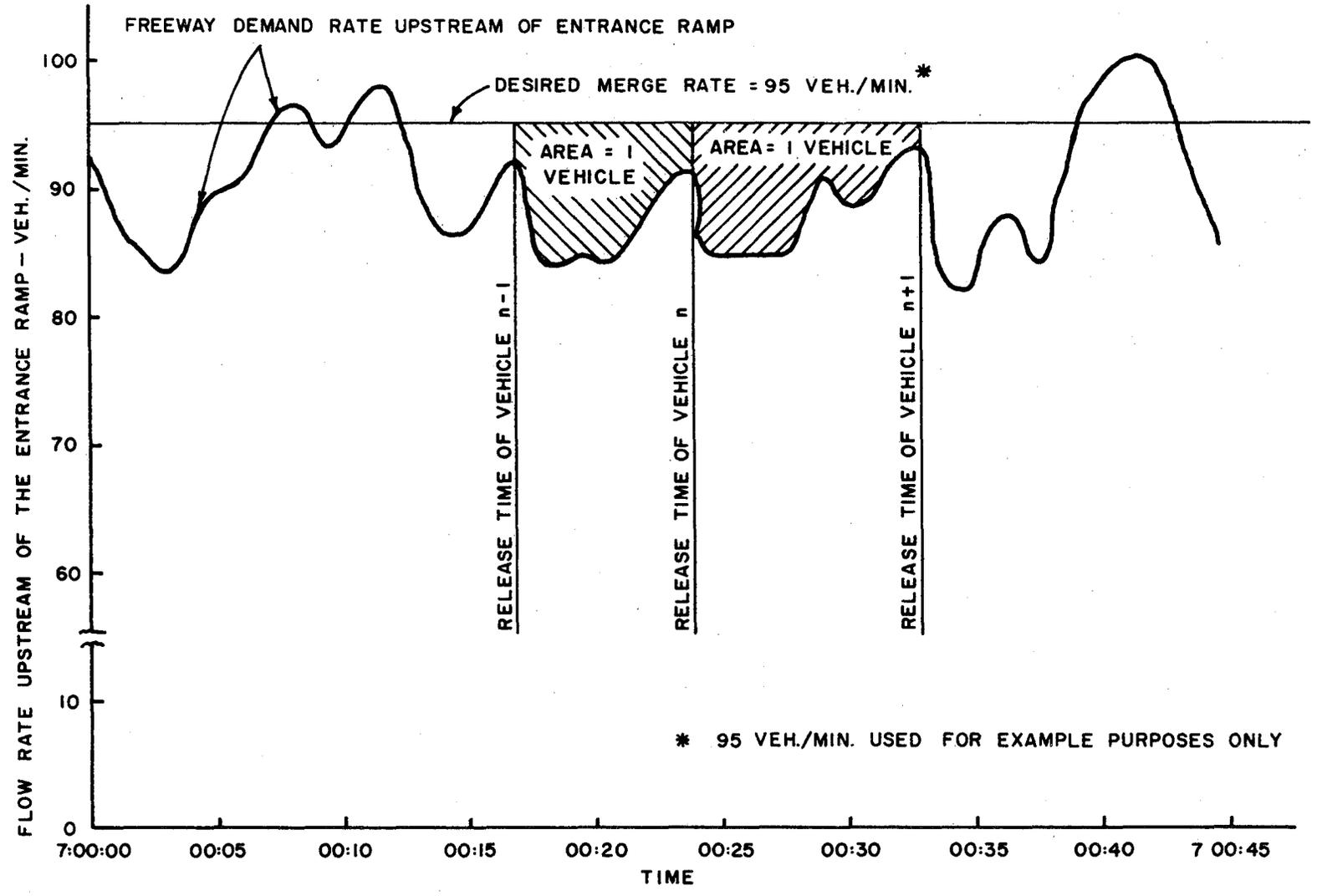
FIGURE 5.2.

capacity of the merge area, as measured by $D_{4v} + D_{5v} + D_{6v}$, or as determined by historical data. Figure 5.3 illustrates the theory of the technique. The difference between capacity and demand is integrated on a real time basis until one unit of capacity is available. When this occurs, a ramp vehicle is released and the controller begins integrating again from zero. Variations of the integration feature can be made to determine ramp flow rates for specific intervals of time (1 minute, for example).

Another variation of the demand-capacity control is the use of one lane instead of the entire freeway. D_{1v} and D_{4v} would be used in this design. This mode of operation is a combination of the two systems described above; the difference being that a specific gap is not detected, but simply space in the merging lane.

(3) Percent Occupancy Control - The control of the ramp signal is based on the percent occupancy measured upstream of the merge area, usually in the second lane (D_{2v}). Percent occupancy is that percent of time that vehicles occupy the detection area. Relationships between occupancy rates and entrance ramp flow rates are established for the control program.

(4) Combinations - A control system may have two or more traffic responsive systems. One that has been used in Houston with success is the gap acceptance/demand-capacity combination. The ramp signal is operated so that ramp vehicles will merge with projected gaps, the acceptable gap size being selected from demand-capacity relationships. This has the advantage of providing both



RELATIONSHIPS AMONG FREEWAY DEMAND RATE, DESIRED MERGE RATE AND RELEASE TIMES OF VEHICLES FROM THE RAMP SIGNAL

FIGURE 5.3.

merging control and system control.

When merging control breaks down, or when unusual demand patterns develop, the system reverts to demand-capacity control. This is necessary since very large time gaps are generated in stop and go traffic and under very light traffic, total freeway demand, rather than only the outside lane, should determine the metering rates.

2. Discussion of Alternative Systems of Control

a. Pretimed Control

This system relies on historical data and predictive traffic patterns. This type of system does offer benefits over no control situations. However, the costs of the additional equipment to make the system traffic responsive are not large.

Therefore, it is recommended that pretimed systems should be used only as the first stage implementation of traffic responsive systems.

b. Traffic Responsive Systems

From the many studies of freeway ramp control, two distinct control objectives have developed: first, to prevent the breakdown of operations and the formation of congestion on the freeway between merging areas; and second, to make optimal use of gaps in the merging area. The first objective is common to all operating control systems except one (Atlanta), and is often stated in more specific terms, such as: to increase level of service, to reduce total delay, or to increase

throughput. The second objective, however, is only applicable to those systems that are capable of detecting and projecting acceptable gaps. The question to be discussed is: Should all new freeway ramp control projects adopt both objectives in the design of the systems?

(1) Gap Acceptance Control - Considerable work has been devoted to this type of control. The Level of Service Project was the first to use this approach with a prototype controller in 1966. (Figure 5.4) Dr. D. R. Drew developed the theoretical model and functional specifications for this project. Two additional gap projection controllers and six capacity demand controllers were added to the project in 1967 when all eight ramps were installed with automatic controls. (Figures 5.5 and 5.6) Finally, three more gap projection controllers were installed in Houston as a part of another research project. There is also another project conducted by Raytheon Corporation that uses the merging control concept.

The rationale for merging control is that minimizing inter-vehicular interference at entrance ramps reduces the probability of rear-end collisions in merging areas due to false starts, reduces the tension on a merging driver, and prevents shock waves from developing on the freeway in the vicinity of entrance ramps.

(a) Disadvantages of a Gap Acceptance Control System - Regardless of the conceptual appeal of gap acceptance operations, there are some disadvantages that must be noted:

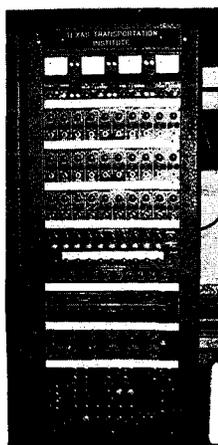


FIGURE 5.4. PROTOTYPE MERGING CONTROLLER.

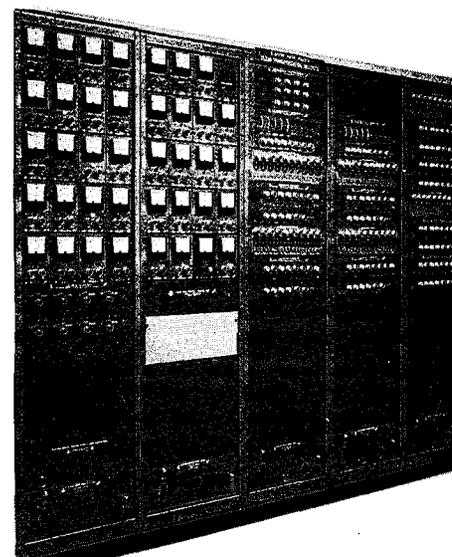


FIGURE 5.5. FIRST GENERATION MERGING CONTROLLERS.

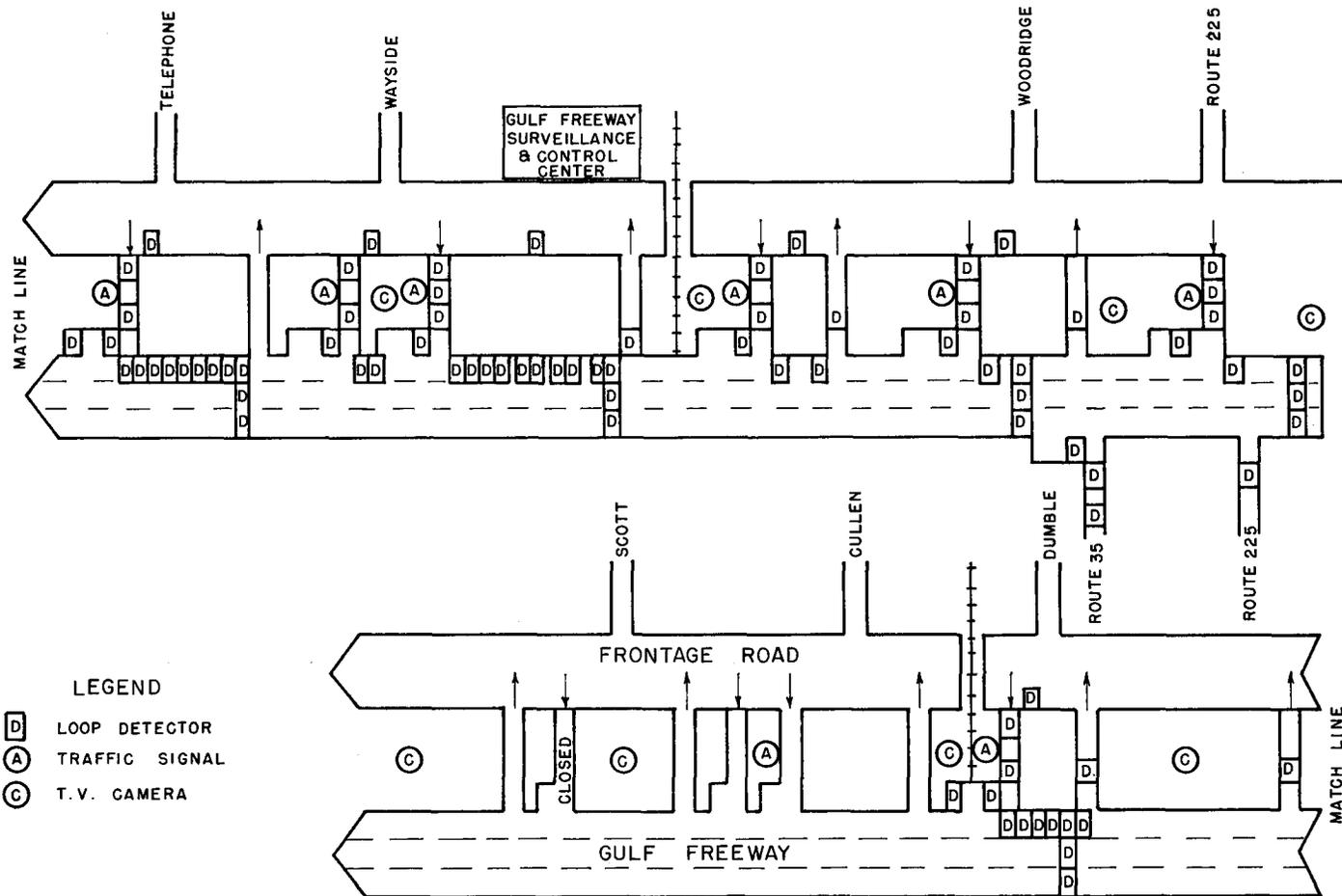


FIGURE 5.6. SCHEMATIC OF GULF FREEWAY PROTOTYPE SURVEILLANCE AND CONTROL SYSTEM (INBOUND).

1. An accurate measure of speed is required at each ramp.
2. The time intervals between the ramp vehicles may be irregular. This type of operation apparently increases traffic violations of the ramp signal.
3. Control of input flow rates to prevent the total freeway demand from exceeding capacity is difficult because of the irregular pattern of gaps in the outside lane.
4. Instability of traffic flow, variations in ramp travel times and lane changing upstream of the merge area reduce the probability of successful merges (moving into the specific gap projected). Studies have shown that 60 to 75 percent of the gaps projected are filled with the ramp vehicles.

(b) Application of Gap Acceptance Control System - Regardless of all the disadvantages just noted, the gap acceptance technique is the best system of control for some situations, i.e.:

1. Older freeways that have very poor merging operations, caused by limited sight distances, inadequate acceleration lanes, grades, etc.
2. Non peak-periods when traffic demand is low and speeds are high.
3. Near peak-periods as traffic demand increases and freeway speeds are uniformly high.

4. Entrance ramp demands are low.
5. Entrance ramp preceded by an exit ramp with high traffic demands, i.e., a merging location with many acceptable gaps.

(2) Demand-Capacity Control - All control systems, regardless of the specific parameters that are used, are based on the premise that the traffic demand should not exceed the capacity of the freeway. Therefore, the most direct method of control is the continuous measurement of the capacity and demands and the absolute control of the ramp flow rates. A prototype controller was developed in 1966 for the Telephone Entrance Ramp on the Gulf Freeway. (Figure 5.4)

(a) Disadvantages of a Demand-Capacity Control

1. Measurement of all traffic demands and flow rates at bottleneck locations to estimate the capacity requires a large number of detectors.
2. Speeds must be measured in addition to flow rates.
3. Demand is generally measured over the entire roadway. Unless individual lane flows are considered in the control program, the capacity of the merging lane can be exceeded, causing some breakdown in operation.

(b) Application of Demand-Capacity Control - The surveillance system required for this type of control is extensive. Therefore, more information on traffic condition is available than from other control systems. This information is useful in other freeway systems:

1. Disabled vehicle detection systems
2. Driver communications systems
3. Freeway maintenance protection systems
4. Data logging systems

If the freeway control is to be a subsystem of an urban freeway and arterial street control network, an extensive surveillance system is required.

(3) Percent Occupancy Control - The system that uses percent occupancy to adjust ramp flow rates measures traffic conditions with one detector at one location. Combining the measures of speed and flow into one parameter often masks the problem for several minutes. Therefore, the sensitivity of the measurement is a factor in its selection as a control parameter.

(a) Disadvantages of Percent Occupancy Control

1. The system using this parameter only samples the upstream traffic conditions. If all lanes were detected, the flow parameter would be used.
2. Data logging and systems control are more difficult with the limited detection system.

(b) Application of Percent Occupancy Control - If a limited detection system is specified, then percent occupancy provides an acceptable measure of traffic operations. Other locations or conditions which are favorable to this control program are:

1. Isolated locations that have good merging facilities.
2. Backup system for a demand-capacity detection system.

3. First stage of a more comprehensive surveillance and control system.

(4) Combinations - From the previous discussion it can be seen that combinations of traffic parameters for the basis of control offer the greatest flexibility and utility.

(a) Gap Acceptance Demand-Capacity Occupancy-This combination of control theories is recommended for urban freeways that have inadequate merging facilities. If the entrance ramps are well designed with long acceleration lanes, gap acceptance control is not a critical control feature.

(b) Gap Acceptance Percent Occupancy - This combination of control theories is recommended for urban freeways that have inadequate merging facilities and limited surveillance systems.

(c) Demand-Capacity-Percent Occupancy - This combination of control theories is recommended for urban freeways that have good merging facilities and extensive surveillance systems.

3. Operational Considerations

In addition to the basic control parameter, several other factors affect the operation of the signal.

a. Merge Detector, D_m

Occupancy of the D_m detector will preempt the control and hold the signal in red. This detector is very important for ramps with poor geometrics, such as restricted sight distances, high angles of entry and short acceleration lanes. Observations of merging operations on a ramp before control is installed will indicate the need for the merge override

feature.

b. Queue Detector, D_q

Occupancy of the D_q detector will preempt the control of the signal, provided D_m is not in effect. The d_q function increases the flow rate of the ramp to reduce the queue length. The rate of flow called for by the D_q is dependent on the necessity for queue control.

The rate increase can be accomplished by preset rates of flow or by reduction of acceptable gap size for the merging control operation.

c. Vehicle Classification Detector, D_t

The utility of this detector as a control parameter has not been established at this time. For a vehicle classification detector to have a meaningful input to the control program, the acceleration characteristics of the vehicle must be known prior to release at the signal. Vehicle size, such as height or length can be measured at D_i but will be of little use. The measure of time from the initiation of queue indication, to the actuation of D_i may give enough data to project the total travel time of the ramp vehicle to the merge area.

The measure of time from initiation of green signal to D_t may be used to activate the merge control operation. Also, the D_t detector as a measure of flow rate is incorporated in the capacity-demand calculations of a freeway system.

d. Freeway Speed Measurement

Any of the detectors located on the main lanes of the freeway can be used to measure speed of traffic flow. Speed reduction is usually the

first indicator of freeway congestion and is used to initiate control restrictions.

e. Systems Measurements

The speed measurement mentioned above when applied to only the nearest entrance ramp is a local control function. If the speed data can be transmitted to several entrance ramp controllers, systems control is then the result. Other measurements, such as volume and percent occupancy, can be handled in the same manner so that conditions several thousand feet from a ramp will be taken into account in the control of the signal.

f. Traffic Signal Phasing

It is the recommendation of this report that a standard three indication signal be used on the ramp control display. The amber indication gives the flexibility of operation that is necessary to optimize freeway operations over a wide range of conditions.

g. Single Vehicle Control

During peak period operation, conditions generally warrant single vehicle control. Merging operations are very difficult, available gaps small, and demand is high. A pretimed signal cycle of $1\frac{1}{2}$ seconds green, 2 seconds amber, and $\frac{1}{2}$ second minimum red, provide excellent control of single vehicle entry. If faster rates of flow are required, a cycle length of 3 seconds can be used, where the phases are 1 second green, $1\frac{1}{2}$ seconds amber, and $\frac{1}{2}$ second red.

h. Multivehicle Control

There are circumstances for which the release of more than one vehicle per signal cycle is desired. The amber indication is of particular importance for this operation.

(1) Added Lane at Merge Area - When a lane is added to the freeway at an entrance ramp, there is usually no need to control the ramp because of local traffic conditions. However, in a systems concept of control, the ramp flow rates will have to be regulated. A multivehicle metering system is acceptable with cycle lengths in the order of one minute with phases of 30 seconds green, 3 seconds amber and 27 seconds minimum red, for example.

(2) Merge Control-Large Gaps - There are locations in a freeway system where large gaps in the right lane will occur, even during peak periods. A merging controller, may release two or more vehicles by regulating the green phase, or with a double queue at the signal, both lead vehicles can be released on the same green phase. The two vehicles then form a single file as they proceed down the ramp.

(3) Ramp Control Under Capacity Reduction Situations - A major change in capacity due to accidents, stalls, maintenance, construction, etc., can effect major changes in traffic flow patterns. Under these conditions, it will be necessary to greatly increase flow rates on some ramps. The multivehicle metering rates with long green phases will accomplish this objective. This approach is considered better than to eliminate control during these "emergency" situations.

VI. CONCLUSIONS

The Level of Service Project has produced many results and conclusions in the twenty five reports, but basic to all the work is the one that states: Freeway ramp control is an effective method of reducing traffic congestion on urban freeways during peak traffic. Freeway ramp control systems are economically and operationally efficient.

Sufficient experimentation and evaluation have been completed and documented to provide operational organizations with the reference material necessary to plan, design and operate freeway ramp control systems. Persons assigned the responsibility for the development of these systems have the opportunity to visit several operational projects in the nation.

A freeway ramp control system is the base of which more comprehensive traffic surveillance, communications and control systems can be developed in the future to coordinate the traffic flow on freeway and arterial street networks in the urban area.

VII. REFERENCE TO
TEXAS TRANSPORTATION INSTITUTE
PUBLICATIONS

1. Report 24-1, "Theoretical Approaches to the Study and Control of Freeway Congestion" by Donald R. Drew.
2. Report 24-2, "Optimum Distribution of Traffic Over a Capacitated Street Network" by Charles Pinnell.
3. Report 24-3, "Freeway Level of Service as Influenced by Volume Capacity Characteristics" by Donald R. Drew and Charles J. Keese.
4. Report 24-4, "Deterministic Aspects of Freeway Operations and Control" by Donald R. Drew.
5. Report 24-5, "Stochastic Considerations in Freeway Operations and Control" by Donald R. Drew.
6. Report 24-6, "Some Considerations of Vehicular Density on Urban Freeways" by John J. Haynes.
7. Report 24-7, "Traffic Characteristics of the Freeway Interchange Traffic of the Inbound Gulf Freeway" by William R. McCasland.
8. Report 24-8, "System Demand-Capacity Analysis on the Inbound Gulf Freeway" by Joseph A. Wattleworth.
9. Report 24-9, "Capacity - Demand Analysis of the Wayside Interchange on the Gulf Freeway" by William R. McCasland.
10. Report 24-10, "Inbound Gulf Freeway Ramp Control Study I" by Charles Pinnell, Donald R. Drew, William R. McCasland, and Joseph A. Wattleworth.
11. Report 24-11, "Investigation of an Internal Energy Model for Evaluating Freeway Level of Service" by Donald R. Drew and Conrad L. Dudek.
12. Report 24-12, "Gap Acceptance Characteristics for Ramp Freeway Surveillance and Control" by Donald R. Drew.
13. Report 24-13, "Inbound Gulf Freeway Ramp Control Study II by Charles Pinnell, Donald R. Drew, William R. McCasland, and Joseph A. Wattleworth.

14. Report 24-14, "Outbound Operation of the Gulf Freeway" by William R. McCasland.
15. Report 24-15, "Peak-Period Analysis and Control of a Freeway System" by Joseph A. Wattleworth.
16. Report 24-16, "An Investigation of the Feasibility of Improving Freeway Operation by Staggering Working Hours" by Gary Lee Santerre.
17. Report 24-17, "Freeway Surveillance and Control Project-Gulf Freeway - Houston, Texas" by Donald R. Drew.
18. Report 24-18, "Considerations for the Installation of U-Turns at Freeway Interchanges" by Sammy L. Wilson.
19. Report 24-19 "The Development of an Automatic Freeway Merging Control System" by Donald R. Drew, William R. McCasland, Charles Pinnell and Joseph A. Wattleworth.
20. Report 24-20, "An Investigation on the Effect of Freeway Grades on Truck Operation" by Jack B. Humphrey.
21. Report 24-21, "Freeway Ramp Control Reduces Frequency of Rear-end Accidents" by Milton L. Radke.
22. Report 24-22, "Evaluation of the Benefits of Traffic Surveillance and Control on the Gulf Freeway" by W. R. McFarland, W. G. Adkins, and W. R. McCasland.
23. Report 24-23, "Effect of Rain on Freeway Capacity" by Roy L. Jones and Merrell E. Goolsby.
24. Report 24-24, "Cost Effectiveness Analysis of Freeway Ramp Control" by Conrad L. Dudek and William R. McCasland.
25. Report 24-25, "Freeway Operations on the Gulf Freeway Ramp Control System" by Merrell E. Goolsby and William R. McCasland.
26. Report 430-1, "A Nationwide Study of Freeway Merging Operations" by J. H. Buhr, D. R. Drew, J. A. Wattleworth, and T. G. Williams.
27. Report 430-2, "Gap Acceptance in the Freeway Merging Process" by D. R. Drew, L. R. Lamotte, J. H. Buhr, and J. A. Wattleworth.
28. Report 430-3, "Operational Effects of Some Entrance Ramp Geometrics on Freeway Merging" by J. A. Wattleworth, J. H. Buhr, D. R. Drew, and F. Gerig.

29. Report 430-4, "The Determination of Merging Capacity and Its Application to Freeway Design and Control" by D. R. Drew, J. H. Buhr, and R. H. Whitson.
30. Report 430-5, "Traffic Interaction in the Freeway Merging Process" by J. H. Buhr.
31. Report 430-6, "Digital Simulation of Freeway Merging Operation" by D. R. Drew, T. C. Meserole, and J. H. Buhr.
32. Report 430-7, "Annotated Bibliography on Gap Acceptance and Its Applications" by Fred S. White.
33. Final Report, "Gap Acceptance and Traffic Interaction in the Freeway Merging Process" by D. R. Drew, J. A. Wattleworth, J. H. Buhr, and T. G. Williams.
34. Report 504-1, "Multilevel Approach to the Design of a Freeway Control System" by D. R. Drew, K. A. Brewer, J. H. Buhr, and R. H. Whitson.
35. Report 504-2, "Traffic Characteristics for Implementation and Calibration of Freeway Merging Control Systems" by J. H. Buhr, R. H. Whitson, K. A. Brewer, and D. R. Drew.
36. Report 504-3, "Design of Freeway Entrance Ramp Merging Control Systems" by J. H. Buhr, W. R. McCasland, J. D. Carvell, and D. R. Drew.
37. Report 504-4, "Real Time Evaluation of Freeway Quality of Traffic Service" by R. H. Whitson, J. H. Buhr, D. R. Drew, and W. R. McCasland.
38. Report 504-5, "Ramp Capacity and Service Volume as Related to Freeway Control" by K. A. Brewer, J. H. Buhr, D. R. Drew, and C. J. Messer.
39. Report 504-6, "Cost-Effectiveness Analysis for Freeway Control" by C. L. Dudek, D. R. Drew, J. H. Buhr, and R. H. Whitson.
40. Report 504-7, "Some Design Considerations of Digital Computer Control Systems Applied to Freeway Traffic Operations" by J. H. Buhr, D. R. Drew, and J. N. Gay.
41. Report 504-8, "A Moving Vehicle Merging Control System" by J. H. Buhr, M. L. Radke, B. M. Kirk, and D. R. Drew.

42. Report 504-9, "Design Considerations of the Central Digital Computer in Freeway Surveillance and Control Systems" by J. H. Buhr, D. R. Drew, J. N. Gay, and R. H. Whitson.
43. Final Report, "Gap Acceptance and Traffic Interaction in the Freeway Merging Process" by the Division of Designing Engineers - Texas A&M University.
44. Report 488-1, "Interim Report on the Evaluation Phase of National Cooperative Highway Research Program Project 20-3" by J. A. Wattleworth, C. E. Wallace, M. Levin, and J. G. Sample, 1967.
45. Report 488-2, "Evaluation of the Operational Effects of an 'On-Freeway' Control System" by J. A. Wattleworth, and C. E. Wallace, 1967.
46. Report 488-3, "Development and Evaluation of a Ramp Metering System on the Lodge Freeway" by J. A. Wattleworth, C. E. Wallace, and M. Levin, 1967.
47. Report 488-4, "Some Traffic System Analysis Techniques" by J. A. Wattleworth and C. E. Wallace, 1967.
48. Report 488-5, "Optimizing Freeway Corridor Operations Through Traffic Surveillance, Communications and Control" by J. A. Wattleworth, Claude Archambault, C. E. Wallace, and J. D. Garvell.
49. Report 488-6, "An Evaluation of Two Types of Freeway Control Systems" by J. A. Wattleworth, K. G. Courage, and J. D. Garvell.
50. Report 488-7, "Some Electronic Measurements of Macroscopic Traffic Characteristics on a Multi-Lane Freeway" by Kenneth G. Courage.
51. Report 488-8, "A Freeway Corridor Surveillance, Information and Control System" by K. G. Courage and Moshe Levin.
52. Report 590-1, "Analysis of the Communications System Requirements for the Dallas North Central Expressway Control" by Charles W. Blumentritt.
53. Report 590-2, "Analysis of the Communication System Requirements for Pilot Control Study of Signalized Intersections" by Charles W. Blumentritt.