COMPUTER MODEL FOR OPTIMAL FREEWAY ON-RAMP CONTROL

Jin J. Wang and Adolf D. May, University of California, Berkeley

Regulating input volume to a freeway system through ramp metering, or ramp closure, maintains traffic flow at an efficient level and improves overall system performance. This paper describes the development of a computer program, LINCON, that can determine the desired fixed-time metering rates for a group of on-ramps to be controlled. The linear programming technique is used to formulate a decision model that is then integrated with a previously developed deterministic freeway simulation model, FREEQ, to become a ramp-control model, RAMPCON. To take into consideration the effect of traffic diversion under control, the decision model was formulated in such a way that, at each on-ramp, the trips with shorter freeway travel distances could divert proportionally more than the trips with longer freeway travel distances. Two objective functions, maximizing total vehicular input and maximizing total freeway vehicle-miles of travel, are considered. The program user has the option of choosing either objective.

•FREEWAY ON-RAMP CONTROL as a potential tool to improve freeway operations did not receive much attention until the early 1960s when increasing congestion on urban freeways became a serious problem. Basically, the idea is that, by regulating the input of the freeway system through ramp metering (or ramp closure in the extreme case), traffic flow on the system can be kept at a more efficient level and thus overall system performance will be improved. Reduction of accidents also produces smoother flow conditions. Many theoretical investigations and empirical analyses have been reported. Recognizing the need to develop a control strategy for a group of ramps as an interrelated system, more research has been directed to the application of mathematical programming techniques to the problem of ramp control. Studies by Wattleworth (1, 2), Payne (3), and Kreer (4) are examples. This paper describes the development of a computer model that can determine the desired metering rates for a group of on-ramps to be controlled. The set of desired metering rates will be called control strategy throughout this paper.

SYSTEMS ANALYSIS

The system of freeway on-ramp control consists of three basic elements: the roadway, the traffic, and the control. The performance of the elements is dependent on each other. Changes in control would cause traffic patterns to change, which in turn would affect weaving section capacities and travel conditions on the roadway. However, roadway characteristics and traffic demands are basic parameters in determining the control strategy. Recognition of this interdependency is most important in studying freeway on-ramp control.

Publication of this paper sponsored by Committee on Freeway Operations.

Objectives and Measures of Effectiveness

Efficiency and safety are the two most important objectives of freeway operations. Efficiency in terms of total vehicles served can be measured by the total input volume. Efficiency in terms of total travel on the freeway can be measured by total vehicle-miles of travel, which tends to favor ramps having drivers with longer average trip lengths. Better utilization of all sections of freeway can also be expected. If travel distance via arterial route is approximately proportional to travel distance via the freeway route, then this criterion will result in smaller diverted vehicle-miles to the arterial street and thus minimize the adverse effect on the arterial system.

Use of total input as the measure of effectiveness tends to favor ramps with shorter average trip lengths. With a smaller number of trips diverted, the impact on the existing trip pattern is reduced. The choice between these two measures of effectiveness is not obvious. In the case where diverted traffic is dispersed evenly throughout the entire arterial network and does not significantly affect the travel on the arterial streets, the total vehicle-mile criterion tends to be better from the point of view of overall corridor traffic. However, if diverted traffic causes significant delay to arterial traffic, the total input criterion might be better. In developing fixed-time control strategies, we cannot use total travel time as the objective inasmuch as the total number of vehicles entering the freeway is not a constant. If the study includes the arterial street traffic where the total number of vehicles can be considered as a constant, then total travel time can be used as the objective.

A safety objective is difficult to measure and almost impossible to predict. Elimination of freeway congestion improves both efficiency and safety. However, further reduction of traffic might improve safety but reduce efficiency. Thus, in on-ramp control, the safety objective is not taken into account explicitly. The constraint that no congestion is allowed on the freeway may be considered as a limit establishing an acceptable level of safety.

Constraints

Because no congestion is allowed on the freeway, one set of constraints is that input demand for each subsection should be less than or equal to the capacity. A modified constraint would be that the demand should be less than the capacity by at least a certain amount. In operation, two more constraints are often necessary. One is the minimum metering rate, which is necessary to prevent excess violation at the metering signal. Sometimes a higher minimum metering rate is used so that accessibility to an area is not unjustifiably reduced. An area without suitable alternative routes is a typical example. The other constraint is the maximum metering rate. This may be limited by the minimum cycle length of the signal or, in certain conditions, to prevent unacceptably long queues on the ramp.

Constraints can also be set up so that certain levels of service can be maintained. For example, it may be desirable to keep the speed on the freeway high. Given that there is a direct relationship between speed and volume, this can be accomplished by requiring the demand to be less than the capacity by a certain amount.

General Framework

Figure 1 shows the general framework in developing ramp-control strategy. A freeway model is needed to determine the traffic performance before and after control, and a decision model is needed to select the best control strategy. The control causes some changes in the origin-destination pattern, which affects the traffic performance and may in turn require modification of the control strategy. This study has attempted to take into account the change of O-D pattern under control by some intuitive assumptions. Better understanding of the nature of traffic diversion is needed to improve the formulation.

THE FREEWAY MODEL AND MANUAL PROCEDURE

The Freeway Model: FREEQ

In 1970, a computer simulation model was developed at the Institute of Transportation and Traffic Engineering in Berkeley. This model, FREEQ, simulates freeway traffic performance under given roadway and demand characteristics. Details of this model are described in a previous report (5). Briefly, the model takes the roadway parameters, which include section length, capacity, ramp location, lane configuration, and design speed, and the load parameters, which are in the form of 15-min origin-destination tables, and computes the traffic performance in terms of speed, volume-capacity ratio, density, travel time, vehicle-hours expended, vehicle-miles expended, actual capacity, and queue length.

Manual Procedure

With the aid of the freeway model, a manual procedure can be used to evaluate ramp control strategies. Essentially, the decision model (Fig. 1) can be performed by the analyst manually. Output of the freeway model is studied, and, by following some rules or judgment, the suitable metering rates are determined. These metering rates are then used to modify the load parameters, which, when entered into the freeway model, determine the traffic performance under the proposed control scheme. The scheme may be modified when necessary. An example of this program can be found in an earlier report (6).

APPLICATION OF THE LINEAR PROGRAMMING TECHNIQUE

Application of the linear programming technique to the problem of freeway on-ramp control was first demonstrated by Wattleworth in 1965 (1, 2). Later work was done by Goolsby and McCasland (7), Messer (8), and Brewer et al. (9) in 1969. The greatest advantage of this technique is its extremely short computation time. This is very important when applied in a real-time traffic-responsive control system. The other advantages are its simplicity in formulation and its systematic methodology for finding the optimum solution for a given objective function and a set of constraints.

Basic Assumptions

1. The demand rates are constant for the time slice under consideration,
2. A steady-state condition is assumed,
3. Traffic diverted from one on-ramp will not enter other on-ramps, and
4. Traffic will not divert from one time slice to another time slice.

In the case of fixed-time control, the time slice is usually no less than 15 min. Except for very long sections of freeway, the second assumption is usually not critical. The first assumption affects primarily the main-line input inasmuch as other ramps are regulated by ramp signals and will produce constant demand rates unless the cumulative demand is less than the cumulative output for some period of time. In the case of traffic-responsive control, the time slice is very short, and the first assumption is not critical as far as the freeway is concerned. The effect of the second assumption can be reduced by breaking up the freeway into segments of 2 to 3 miles bounded by key bottlenecks. Each segment will operate almost independently except that the main-line output of the upstream segment is the main-line input of the downstream segment.

The third assumption is quite unrealistic in many cases, particularly when backtracking to enter upstream on-ramps is favorable. There are two kinds of backtracking. The first kind occurs when the original demand pattern is distorted because some traffic uses downstream on-ramps to bypass a bottleneck. When the control is implemented and the traffic in the bottleneck is flowing smoothly, drivers will use upstream on-ramps as they would have in the case of no bottlenecks. This condition can be corrected by adjusting the original demands to the situation with no congestion on the freeway. These adjusted demands are then used in the formulation. The second kind occurs when upstream on-ramps are not controlled and some traffic finds it is faster to backtrack a little even though the travel distance is longer. This effect can be reduced by extending the control farther upstream. Diversion from upstream on-ramps to a controlled downstream on-ramp is possible but unlikely unless queue delay at the downstream on-ramp is much shorter. The forth assumption is not critical for the beginning and ending time slices of the control. Some trips will start just a little bit earlier or later in order to avoid the control. However, control usually begins before congestion actually occurs and ends after the peak period. Thus, diversion in time at both ends of the control period is not critical from the operations point of view. For other time slices, the diversion is caused mainly by queuing on the ramps, which delays the entry of some vehicles at the end of the time slice. The effect depends on the change of destination pattern from one time slice to another. If the destination pattern is relatively stable, the effect will be small.

Previous Formulation

The basic linear programming model as formulated by Wattleworth is

$$\text{Maximize} \sum_{i=1}^{n} X_i \tag{1}$$

subject to

$$\sum_{i=1}^{n} A_{ik} X_i \leq B_k \tag{2}$$

for k = 1, 2, ..., m,

$$X_i \leq D_i \tag{3}$$

for i = 1, ..., n, and

$$X_i \geq 0 \tag{4}$$

for i = 1, ..., n, where

X_i = desired input rate from on-ramp i (the metering rate),
n = number of on-ramps,
m = number of subsections,
A_{ik} = fraction of traffic from on-ramp i going through subsection k,
B_k = capacity of subsection k, and
D_i = demand rate on on-ramp i.

Equation 1 simply states that the objective of the control is to maximize total input rate from all ramps. Equation 2 is the capacity constraint (total demand for any subsection should not exceed its capacity). Equation 3 states that the input rate for any on-ramp cannot be more than the demand, and Eq. 4 states that the input rates cannot be negative. This formulation can be solved efficiently by the standard simplex method. Four possible modifications were suggested by Wattleworth:

1. B_k may represent the service volume of a desired level of service to be maintained for the freeway traffic.
2. A constraint may be added that will limit the number of vehicles diverted. Expressed mathematically,

$$D_i - X_i \leq Q_i \tag{5}$$

where Q_i is the maximum number of vehicles diverted for ramp i.

3. A different type of constraint on the number of vehicles diverted may be in the form of

$$D_i - X_i = D_{i+1} - X_{i+1} \tag{6}$$

for all i. This constraint will spread all access demand equally over all on-ramps.

4. A merging capacity constraint in the form of

$$P_a \sum_{j=a+1}^{n} A_{jk}X_j + X_a \leq L_a \tag{7}$$

for a = 1, . . . , n - 1 may be added. P_a is the fraction of trips upstream of ramp a, which is in lane 1. L_a is the merging capacity.

Other modifications that have been suggested are listed as follows:

5. In the discussion of the paper presented by Wattleworth, Foote (2) suggested use of the objective in minimizing total excess capacity. Mathematically,

$$\text{Minimize} \sum_{k=1}^{m} \left(B_k - \sum_{i=1}^{n} A_{ik}X_i\right) \tag{8}$$

6. In discussing the same paper, May suggested the use of the objective of maximizing total vehicle-miles of travel.

$$\text{Maximize} \sum_{i=1}^{n} X_i l_i \tag{9}$$

where l_i is the average trip length of all traffic from on-ramp i.

Two additional modifications may be considered.

7. Equation 5 may be expressed as

$$X_i \geq D_i - Q_i = N_i \tag{10}$$

where N_i is the minimum metering rate for ramp i.

8. A maximum metering rate constraint in the form of

$$X_i \leq M_i \tag{11}$$

can be added where M_i is the maximum metering rate.

Alternative Formulation

Wattleworth's formulation implicitly assumes that, for each on-ramp, the destination pattern before and after control is the same. This is reflected by the parameter A_{ik}, which is computed from the destination pattern before the control. In operation, a queue will form on the ramp and cause a certain amount of delay. Some traffic will find it better to use alternative routes. Logically, traffic with better alternative routes or, in terms of travel time, a shorter travel time for the alternative route than for the freeway route will more likely divert first. The exact pattern of diversion is undoubtedly stochastic in nature and depends on the actual origin and destination of each trip and on driver characteristics. As an approximation, it is assumed that trips with shorter freeway trip lengths will divert proportionally no less than trips with longer freeway trip length. This can be expressed as

$$P(i,j) \leq P(i,j+1) \tag{12}$$

where P(i, j) is the percentage of the original demand between on-ramp i and off-ramp j that still uses the freeway after control. Equation 12 does not guarantee that all short trips will divert first; it only states that short trips could potentially divert proportionally more than long trips. This may appear to be an inconsistent assumption inasmuch as no definite diversion rule is established. However, because of the stochastic nature of the true diversion pattern, we feel that this assumption will produce results closer to reality than the assumption that all short trips will divert first.

The following notations are used in the alternative formulation.

TRIP(i, j) = original demand rate from on-ramp i to off-ramp j,
P(i, j) = percentage of original demand from on-ramp i to off-ramp j that is not diverted,
B_k = capacity (or service volume) of subsection k, and
L(i, j) = freeway travel distance between on-ramp i and off-ramp j.

The linear programming formulation is as follows:

$$\text{Maximize} \sum_{i=1}^{n} \sum_{j=1}^{m} P(i,j)\ \text{TRIP}(i,j) \tag{13}$$

where n is the number of on-ramps and m is the number of off-ramps, subject to

$$P(i,j) \leq 1 \tag{14}$$

for all (i, j),

$$P(i,j) - P(i,j+1) \leq 0 \tag{15}$$

for all (i, j),

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \delta \times P(i,j) \times \text{TRIP}(i,j) \leq B_k \tag{16}$$

and

$$P(i,j) \geq 0 \tag{17}$$

In Eq. 16, $\delta = 1$ if i is upstream of subsection k and j is downstream of subsection k; otherwise, $\delta = 0$. In this formulation, P(i, j) is the decision variable and TRIP(i, j), B_k, and δ are constant. Equation 13 is replaced by

$$\text{Maximize} \sum_{i=1}^{n} \sum_{j=1}^{m} P(i,j)\text{TRIP}(i,j)\ L(i,j) \tag{18}$$

when the selected objective is to maximize vehicle-miles of freeway travel.

In the remainder of this report, the formulation by Wattleworth will be referred to as the formulation with proportional diversion, and the alternative formulation de-

Figure 1. General framework in developing ramp-control strategies.

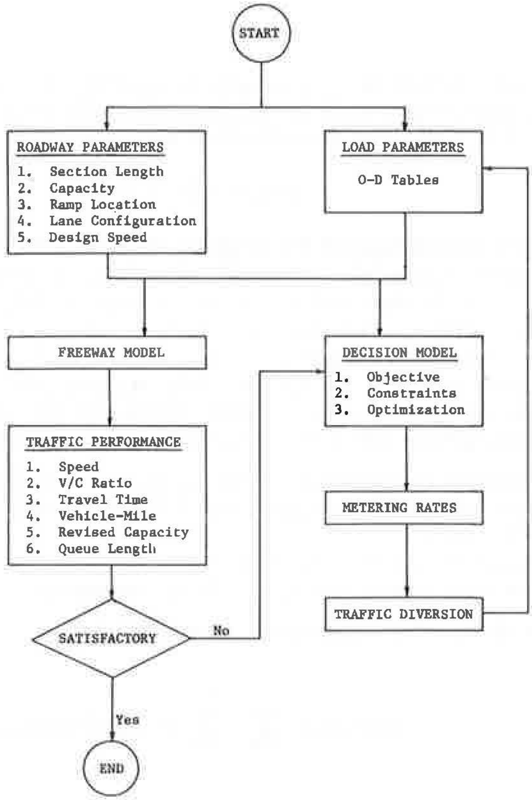
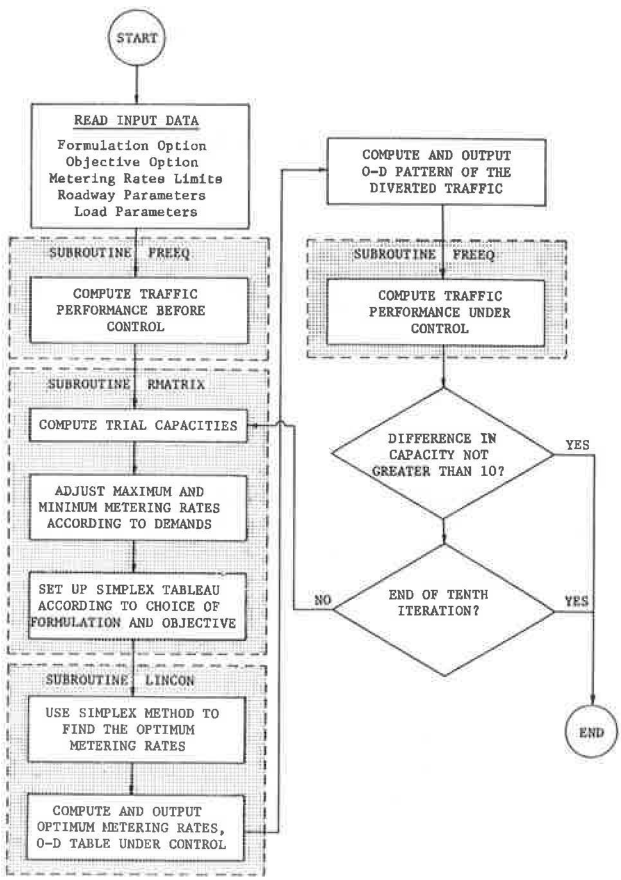


Figure 2. Program RAMPCON.



veloped in this section will be referred to as the formulation with short-trip diversion.

The Computer Program: LINCON

Based on the formulations previously presented, a computer program (LINCON) was prepared. The program includes the four basic modifications by Wattleworth, modification items 6, 7, and 8, and the alternative formulation. Thus, the user has the flexibility of choosing either type of formulation depending on the applicability of the diversion assumption. The user can also choose either total vehicle input or total vehicle-miles as the objective function.

The number of equations in the linear programming formulation is reduced by not entering the maximum and minimum metering rates as constraints. In the program, the maximum metering rates are first adjusted so that they are not greater than the demands. Then, if the demand of a ramp is greater than the maximum metering rate, D_i in Eq. 3 is replaced by the maximum metering rate, or TRIP(i, j) in Eqs. 13 and 16 is modified such that the number of total trips from an on-ramp is equal to the maximum metering rate. Trips with shorter freeway trip lengths are diverted first. The minimum metering rates are also modified so that they are not greater than the demands. Then, these trips are loaded to the freeway as if they are through traffic. If the diversion assumption of the alternative formulation is adopted, long trips are selected first for freeway use. Parameter values for the simplex tableau are automatically computed in the subroutine RMATRX by using the input data.

INTEGRATION OF FREEWAY MODEL AND LINEAR PROGRAMMING

The linear programming model, LINCON, is now used as the decision model and, in combination with the FREEQ model, results in a completely automatic computer model, RAMPCON. The need to integrate these two models has been shown previously. Figure 1 can be considered as the generalized flow chart of this model. A more detailed flow chart is shown in Figure 2.

Iteration Procedure

From the linear programming formulations, it can be seen that it is the change in B_k value that will cause the modification of the metering rates. The modified metering rates may in turn change the value of B_k caused by change in weaving effect. Thus, an iteration procedure is required to obtain the equilibrium solution. This program handles only the case that B_k is the capacity. Minor modification is required if B_k is to be the service volume of a specified level of service.

The value of B_k is affected by the load parameters because of the weaving effect. The step-by-step procedure is as follows:

1. Input roadway parameters and O-D table into the FREEQ model to obtain the traffic performance under no control conditions.
2. Let C_k be the capacity of section K if there is no weaving effect and WE_k be the weaving effect. Compute a trial value of B_k by using Eq. 19.

$$B_k = C_k - \frac{WE_k}{2} \tag{19}$$

3. With the computed B_k value and other necessary data, use LINCON program to compute the optimum metering rates.
4. Revise the O-D tables based on the metering rates, and rerun FREEQ.
5. Compare the previously computed B_k value with the true capacity, which is equal to $C_k - WE_k$. If the difference is less than 10 for all subsections, terminate the procedure; otherwise, continue to next step.
6. Revise the B_k value by using Eq. 20.

$$B_k = (B_k + C_k - WE_k)/2 \tag{20}$$

Table 1. Summary of test roadway data.

Subsection No.	No. of Lanes	Capacity	Length (ft)	Truck Factor	Ramp	Subsection Description
1	3	5,728	1,660	0.970	O	Central off to Central on
2	3	5,806	1,890	0.970	OD	Central on to Carlson off
3	3	5,520	2,310	0.940		Carlson off to Carlson on
4	3	5,950	1,460	0.980	OD	Carlson on to Potrero off
5	3	5,806	3,800	0.970		Potrero off to Cutting on
6	3	5,880	1,100	0.980	O	Cutting on to grade change point
7	3	5,950	660	0.980	D	Grade change point to Macdonald off
8	3	5,950	1,480	0.980	D	Macdonald off to San Pablo off
9	3	5,728	1,480	0.970		San Pablo off to San Pablo on
10	4	6,850	800	0.980	OD	San Pablo on to Solano off
11	3	5,800	4,690	0.970	D	Solano off to San Pablo Dam off
12	3	5,806	2,190	0.970		Dam Road off to Dam Road on
13	3	5,800	2,320	0.970	OD	Dam Road on to Road 20 off
14	3	5,049	830	0.850		Road 20 off to grade change point
15	3	4,746	1,180	0.793		Grade change point to Road 20 on
16	3	4,700	2,560	0.780	OD	Road 20 on to Mainline Destination

Note: Ramp limit = 1,500 vph.

Table 2. Origin-destination data for test system (15-min volume).

Origin Down	Destination Cross 1	2	3	4	5	6	7	8
1	61	106	56	102	34	121	66	798
2		3	6	9	4	19	7	39
3		7	3		5	14	9	44
4					34	108	40	153
5					14	48	29	152
6							11	55
7								0

Table 3. Summary of test system results.

Formulation		Proportional Diversion		Short-Trip Diversion	
Item	No Control	Maximize Vehicle-Miles	Maximize Input	Maximize Vehicle-Miles	Maximize Input
Total input rate	8,320	7,753	7,753	7,655	7,691
Total output rate	7,087	7,753	7,753	7,655	7,691
Total vehicle-mile rate	6,699	7,703	7,703	7,751	7,735
Computer time (sec)		17	17	34	40
No. of iterations		7	7	6	7
Key bottleneck subsections	6	4, 6, 11	4, 6, 11	4, 6, 16	4, 11, 16

Table 4. Optimum metering rates on test system.

On-Ramp No.	Original Demand Rate (vph)	Optimum Metering Rate (vph)			
1	348	348	348	348	348
2	328	328	328	328	328
3	1,340	512	512	512	403
4	972	925	925	827	972
5	264	264	264	264	264
6	0	0	0	0	0

The B_k value on the right side of the equation is the value used in the previous run of LINCON.

7. Return to step 3.

Test System and Results

The northbound Eastshore Freeway (located in the San Francisco Bay area) from the Central Street off-ramp to the main-line section north of the Road 20 on-ramp was selected as the test system. A previous study (6) has proposed the implementation of ramp control on this section. The roadway data are given in Table 1, and the traffic data in the format of a 15-min origin-destination table are given in Table 2. This O-D table is the projected 1972 demand (assuming BART in operation) and for the time period from 4:30 to 4:45 p.m. A minimum metering rate of 240 vph is selected for all ramps. A maximum metering rate of 1,080 vph is selected for the San Pablo on-ramp and 800 vph for all other ramps.

Tables 3 and 4 give the results of the test system. Results of both formulations and both objective functions show improvements in total output and total vehicle-miles of freeway travel for the controlled system as compared with the uncontrolled system. Although total input for the uncontrolled system is higher than that for the controlled system, there is severe congestion on the freeway. Under the formulation with proprotional diversion, the results for both objective functions are identical. This may not be the case for other systems. The differences among the four conditions do not appear to be significant. This is probably because subsection 6 is the major bottleneck that effectively reduces the demands of the downstream subsections to less than or only slightly above the capacity.

In summary, the computer model developed by combining a freeway model and the linear programming technique can determine the desired metering rates, taking into consideration the changes of freeway subsection capacities when the demand pattern is altered by the metering rates. The freeway submodel also provides data on freeway performance before and after control for better analysis of the control effect. The model was developed basically for the analysis of fixed-time ramp control. Application to the traffic-responsive control is possible when supplemented by a real-time demand forecast and distribution model.

REFERENCES

1. Wattleworth, J. A., and Berry, D. S. Peak-Period Control of a Freeway System—Some Theoretical Investigations. Highway Research Record 89, 1965, pp. 1-25.
2. Wattleworth, J. A. Peak-Period Analysis and Control of a Freeway System. Highway Research Record 157, 1967, pp. 1-21.
3. Payne, H. J., and Isaksen, L. A Systematic Approach to Freeway Traffic Control and Surveillance. To be published.
4. Kreer, J. B., Koenig, H. E., Goodnuff, J. L., Mika, H. S., Panyan, W. D., and Yuan, L. S. Coordinated Control of Urban Traffic Networks. Div. of Engineering Research, Michigan State Univ., technical rept., Dec. 1968.
5. Makigami, Y., and Woodie, W. L. Freeway Travel Time Evaluation Technique. Highway Research Record 321, 1970, pp. 33-45.
6. Aidoo, J. F., Goedhart, R. W., and May, A. D. Bay Area Freeway Operations Study—Final Report: Analytic Techniques for Evaluating Freeway Improvements on the Eastshore Freeway (I-80) Northbound. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, June 1971.
7. Goolsby, M. E., and McCasland, W. R. Freeway Operations on the Gulf Freeway Ramp Control System. Texas Transportation Institute, Texas A&M Univ., Research Rept. 24-25, 1969.
8. Messer, C. J. A Design and Synthesis of a Multilevel Freeway Control System and a Study of Its Associated Operational Control Plan. Texas A&M Univ., dissertation, Aug. 1969.
9. Brewer, K. A., Buhr, J. H., Drew, D. R., and Messer, C. J. Ramp Capacity and Service Volume as Related to Freeway Control. Highway Research Record 279, 1969, pp. 70-86.