HIGHWAY RESEARCH RECORD

Number 279

Freeway Traffic Characteristics and Control

19 Reports

Subject Area

53 Traffic Control and Operations
55 Traffic Measurements

HIGHWAY RESEARCH BOARD

DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

Washington, D. C., 1969 Publication 1656

Price: $4.20

Available from

Highway Research Board
National Academy of Sciences
2101 Constitution Avenue
Washington, D.C. 20418

Department of Traffic and Operations

Harold L. Michael, Chairman
Purdue University, Lafayette, Indiana

HIGHWAY RESEARCH BOARD STAFF

E. A. Mueller

COMMITTEE ON HIGHWAY CAPACITY
(As of December 31, 1968)

Carlton C. Robinson, Chairman
Automotive Safety Foundation, Washington, D.C.

Peter A. Mayer, Secretary
Automotive Safety Foundation, Washington, D.C.

W. R. Bellis
Donald S. Berry
Robert C. Blumenthal
Arthur A. Carter, Jr.
Robert R. Coleman
Kenneth W. Crowley
Laurence A. Dondanville
Donald R. Drew
Robert W. Duis
Gordon K. Gravelle
Howard C. Hanna
John E. Hartley
Wolfgang S. Homburger
Jack Hutter
Donald C. Hyde
James H. Kell
Jack E. Leisch
James H. Little
Gary E. Mathias
Adolf D. May, Jr.
Karl Moskowitz
Louis J. Pignataro
Walter S. Rainville, Jr.
E. J. Rockefeller
Gordon A. Shunk
Richard I. Strickland
T. Darcy Sullivan
S. S. Taylor
William P. Walker
Leo G. Wilkie

COMMITTEE ON TRAFFIC CONTROL DEVICES
(As of December 31, 1968)

Robert E. Conner, Chairman
Bureau of Public Roads, Washington, D.C.

Robert E. Titus, Secretary
West Virginia State Road Commission, Charleston

Donald S. Berry
Robert L. Bleyl
James W. Booth
Abner W. Coleman
Robert D. Dier
William H. Dorman
Roy D. Fonda
Joseph V. Galati
Michael J. Gittens
Albert L. Godfrey, Sr.
Alan T. Gonseth
J. T. Hewton
George W. Howie
Rudolph J. Israel
Lester R. Jester
Barnard C. Johnson
Anthony C. Kanz
Leonard I. Lann
Joseph E. Lema
Holden M. LeRoy
Sam E. Luebbert
J. Carl McMonagle
J. P. Mills, Jr.
Zoltan A. Nemeth
Robert J. Nolan
Albert S. Palatnick
A. R. Pepper
Frank G. Schlosser
R. G. St. John
Asriel Taragin
James A. Thompson
Robert Wert
Arthur M. White
Earl C. Williams, Jr.
James K. Williams
Robert M. Williston

COMMITTEE ON THEORY OF TRAFFIC FLOW
(As of December 31, 1968)

Donald G. Capelle, Chairman
Alan M. Voorhees and Associates, Inc., Los Angeles, California

Edmund A. Hodgkins, Secretary
Bureau of Public Roads, Washington, D.C.

Patrick J. Athol
John L. Barker
Martin Beckman
Jack B. Blackburn
Donald E. Cleveland
Lucien Duckstein
Leslie C. Edie
Robert S. Foote
Antranig V. Gafarian
Denos C. Gazis
Daniel L. Gerlough
Frank A. Haight
Walter Helly
Robert Herman
Hyoungkey Hong
James H. Kell
Frederick Lehman
Russell M. Lewis
Mo Chih Li
Adolf D. May, Jr.
Joseph C. Oppenlander
Richard W. Rothery
Charles C. Schimpeler
A. D. St. John
Asriel Taragin
Joseph Treiterer
William P. Walker
Joseph A. Wattleworth
George Weiss

COMMITTEE ON FREEWAY OPERATIONS
(As of December 31, 1968)

Adolf D. May, Jr., Chairman
University of California, Richmond, California

Robert S. Foote, Secretary
The Port of New York Authority, New York, New York

Patrick J. Athol
John H. Auer, Jr.
John L. Barker
Roye G. Burnfield
Donald O. Covault
Donald R. Drew
Conrad L. Dudek
Robert E. Dunn
James L. Foley, Jr.
Paul Fowler
Denos C. Gazis
Edward F. Gervais
John J. Haynes
Gerhart F. King
Alger F. Malo
Stuart F. Millendorf
William J. Miller, Jr.
Donald E. Orne
Jack G. Sample
G. W. Skiles
T. Darcy Sullivan
Asriel Taragin
Joseph Treiterer
James R. Turner
Ronald W. Tweedie
John L. Vardon
Joseph A. Wattleworth
Morton I. Weinberg
James E. Wilson
Richard D. Worrall

COMMITTEE ON CHARACTERISTICS OF TRAFFIC FLOW
(As of December 31, 1968)

Joseph C. Oppenlander, Chairman
Purdue University, Lafayette, Indiana

Patrick J. Athol
Jack B. Blackburn
Martin J. Bouman
Kenneth W. Crowley
Olin K. Dart, Jr.
Robert F. Dawson
H. M. Edwards
John J. Haynes
Clinton L. Heimbach
James H. Kell
Russell M. Lewis
Jack C. Marcellis
Peter A. Mayer
F. William Petring
O. J. Reichelderfer
August J. Saccoccio
Charles C. Schimpeler
Joseph Seifert
William P. Sheldon
William C. Taylor
Kenneth J. Tharp
Robert J. Wheeler

Foreword

The ten papers and nine abridgments that make up the contents of this RECORD are concerned with determining freeway traffic characteristics or providing better traffic control through control devices and techniques. The increasingly difficult problems of building new urban freeways have focused attention on better operation of existing freeways. Highway or transportation department officials will want to be familiar with the research presented in these papers, if only in a cursory way, while traffic or highway engineers concerned with freeways, and researchers similarly inclined, will need to be intimately acquainted with the contents.

The first paper, by three Michigan researchers, is concerned with traffic flow on Detroit's Lodge Freeway. Two distinct modes were found: steady flow, where linear and parabolic relations between density, flow rate, and velocity apply, and oscillatory flow, wherein speed and density appear out of phase. Flow switches from one mode to the other at lane occupancy percentages between 12 and 18 percent. The research presents other interesting characteristics of the two modes of freeway flow.

The second paper, by an Ohio and a Michigan researcher, gives results of a survey concerning supplemental Interstate service signing. The signing increased the Interstate user's awareness of service availability and resulted in better utilization of advance information.

The third paper presents research performed in Detroit on the response to a CB Radio Driver Aid Network. The researchers found that such a system might furnish a feasible interim solution to the problem of roadway surveillance, incident reporting, and action implementation to provide safe and efficient traffic flow.

The next paper, by four Texas experimenters, deals with the definition of the functional and operational requirements for freeway control systems and the actual design and installation of a prototype. The approach taken in the design of the freeway control system prototype is based on the multilevel concept as used in process control. The freeway is viewed as an entity, with the control law split into several levels, lower levels being directed toward recognizing the influence of short-term factors (gap availability for merging ramp vehicles) and higher levels reserved for factors that influence performance on a long-term basis (freeway capacity reductions due to accidents, incidents, or geometric bottlenecks).

In the fifth paper, results of emergency ramp control measures to lessen Los Angeles traffic congestion on holiday weekends are reported. The data collected and observations noted provide information on a unique application of ramp control and an emerging philosophy in freeway and expressway operation.

The sixth paper is concerned with ramp capacity and service volume as related to freeway control. Ramp capacity and service volume are defined in terms of entrance ramp geometry, gap acceptance, and merging controller operational characteristics. An optimal control policy is established for an isolated entrance ramp and a policy for integrated total freeway control policy is presented.

In the next report, Texas researchers using the Gulf Freeway in Houston as a laboratory studied a number of traffic characteristics associated with gap acceptance control. Data collection was performed using manual methods, a closed-circuit television system, and an electronic data-acquisition system including field sensors and a digital computer.

The measurement of commonly sought characteristics of macrosocopic traffic flow by means of automatic sensing equipment using a high-speed digital computer for data collection and analysis is discussed in the next paper. Independent measurements of speed, density, flow, and kinetic energy were obtained from ultrasonic presence detectors and from an optical speed trap.

The last two unabridged papers in this RECORD, also by Texas researchers, are concerned with freeway traffic control. One deals with the control of individual entrance ramps and presents design and installation requirements of a gap acceptance merging control system. The other sets forth the requirements for a moving vehicle merging control system and mathematically models it.

The RECORD concludes with nine abridgments of freeway operations papers that were reviewed by the Board's Freeway Operations Committee. They give short resumes of the significant findings of research in this field.

Contents

Dual Mode Behavior of Freeway Traffic

H. S. MIKA, J. B. KREER, and L. S. YUAN, Michigan State University

Analysis of traffic flow data from the John C. Lodge Freeway in Detroit indicates that flow can be categorized into two distinct modes: (a) steady flow, where the usual linear and parabolic relations between density, flow rate, and speed apply; and (b) an oscillatory mode, in which speed and density exhibit out-of-phase periodicities when plotted as a function of time. The system switches from one mode of flow to the other when operation is near the vertex of the lane occupancy-flow rate parabola. The location of this vertex is between 13.1 and 16.1 percent lane occupancy on the center and median lanes, and between 12.4 and 18.1 percent on the curb lane. The recovery from the oscillatory mode is quite rapid whenever lane occupancy falls below the critical value.

The frequency of the oscillation was found to be ¼ cycle per minute from calculation of the autocorrelation and power density spectrum of the oscillatory waveforms. By cross-correlation of the waveforms measured at pairs of locations, it was determined that the onset of the oscillatory mode propagates upstream at a rate of 16 mph. The data also show that in the steady-flow mode the median and center lanes behave similarly, but the curb lane is significantly different. The curb lane has only about 86 percent of the capacity of the center and median lanes.

•TWO previously unreported phenomena in the traffic flow on an urban freeway were found in an analysis of data collected by an automatic electronic data-collection system located at the National Proving Ground for Freeway Surveillance, Control and Electronic Traffic Aids Project (John C. Lodge Freeway, Detroit). The first phenomenon was the existence of an oscillatory mode of flow, in which speed and density exhibit out-of-phase periodicities when plotted as a function of time. It was found that at a lane occupancy of about 15 percent (45 vehicles per mile) traffic flow switches to this mode. The second phenomenon was a difference in the capacity of the curb lane relative to the center and median lanes.

The onset of the oscillatory behavior was abrupt, remarkably regular, and varied by more than a factor of two in amplitude as a function of location. The frequency of oscillation as determined by a Fourier transformation of the autocorrelation function was about ¼ cycle per minute or a period of 4 minutes. A cross-correlation analysis indicated that a disturbance is propagated upstream with a velocity of about 16 mph. The flow-speed-density relationships were obviously different from those in the steady mode, and it would be difficult to justify the simple linear and parabolic relationships in the high-density region. In all 22 locations studied it was observed that the switch to the oscillatory mode occurred at or near the vertex of the speed-flow rate parabola. In the center lane this corresponded to a lane occupancy range of 13.7 to 16.1 percent. For the curb lane the range was 12.4 to 18.1 percent, and for the median lane the range was 13.1 to 16.0 percent.

Paper sponsored by Committee on Theory of Traffic Flow and presented at the 48th Annual Meeting.

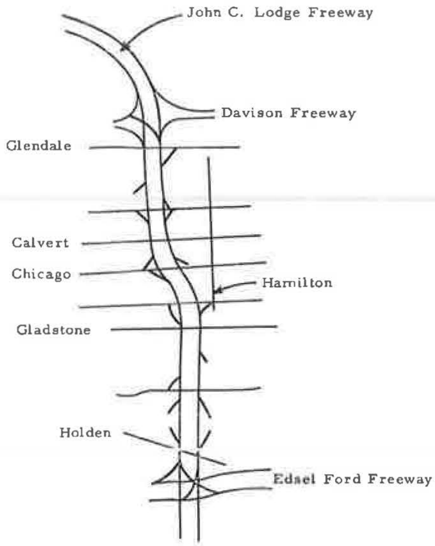


Figure 1. Section of John C. Lodge Freeway studied.

In the steady-flow mode, i.e., below 15 percent occupancy, the flow is characterized by the usual parabolic relationship between average speed and flow rate or density and flow rate, and by a linear relationship between average speed and density. In the steady-flow mode, the performance of the median and center lanes is essentially the same, and efficiency in terms of capacity is essentially the same. In the length of freeway analyzed, this capacity ranged from 1823 vph to 2244 vph. The performance on the curb lane was significantly different from the other lanes, with an average capacity of about 86 percent of the other lanes.

SYSTEM AND INSTRUMENTATION

The NPG System has been described in detail elsewhere (1). Briefly, the data used in this analysis cover 2.38 miles of the John C. Lodge Freeway extending from Glendale as the northern terminus to Holden (Fig. 1). Sonic sensors were placed at 13 locations along this strip as follows: Southbound—5 locations, 3-lane sensing, and 8 locations, center-lane sensing only; Northbound—1 location, 3-lane sensing, and 12 locations, center-lane sensing only. All data from all stations were obtained simultaneously and recorded at 1-minute intervals. Southbound data were recorded from 6:00 a.m. to 9:30 a.m. Northbound data were recorded from 2:00 p.m. to 7:30 p.m.

Although other information was recorded, only the following variables were important to this analysis: time, vehicle count, lane occupancy, and average car speed.

The ultrasonic sensors were mounted above the traffic. The presence of a vehicle below the sensor was determined from the echo receive-time compared to the echo receive-time from unoccupied pavement. This was recorded in vehicles counted during 1 minute. Lane occupancy was measured as the relative amount of time that a vehicle echo was received. Vehicle speed was calculated by dividing 17.5 ft (assumed average length of vehicle) by the time during which an individual echo was received and by averaging all the vehicle speeds during 1 minute.

Data for one day, August 16, 1966, were made available for this study. From study of the log, this appeared to be a typical operating day with no unusual occurrences. Throughout the test the weather was either overcast or clear, and except for a period from 6:00 a.m. to 6:53 a.m. the pavement was dry. (The wet pavement condition was not reflected in any special remarks in the log.) Temperature ranged from 57 F at 6:00 a.m. to 79 F at 2:30 p.m.

DUAL MODE BEHAVIOR

The switching from a steady mode to an oscillatory mode is shown in Figure 2. This figure corresponds to the location described in Figure 11 and shows that flow is reasonably steady until the lane occupancy reaches about 15 percent at 6:38 a.m. Then suddenly lane occupancy and average speed start to oscillate.

This behavior is typical of every location studied. For example, Figure 3 shows the same switching action at the Gladstone location. In this paper the emphasis is on the southbound direction because this was analyzed in greater detail. However, Figure 4 is included to show that the oscillatory mode is observed in the northbound direction during evening rush hours.

It is common practice in traffic engineering to depict behavior by a graph of speed vs density similar to that shown in Figure 5—in this case representing the Calvert

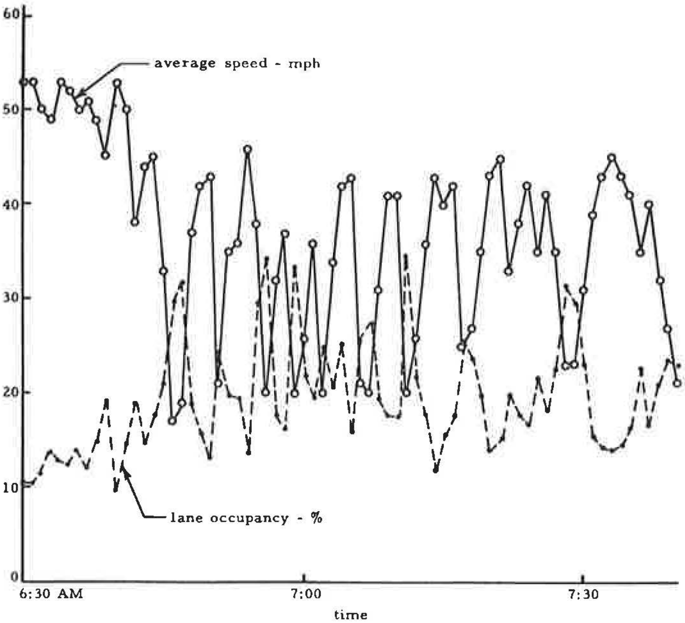


Figure 2. Station 22, Hamilton: Center lane, a.m., southbound; on ramp—50 ft upstream, on ramp—1945 ft downstream.

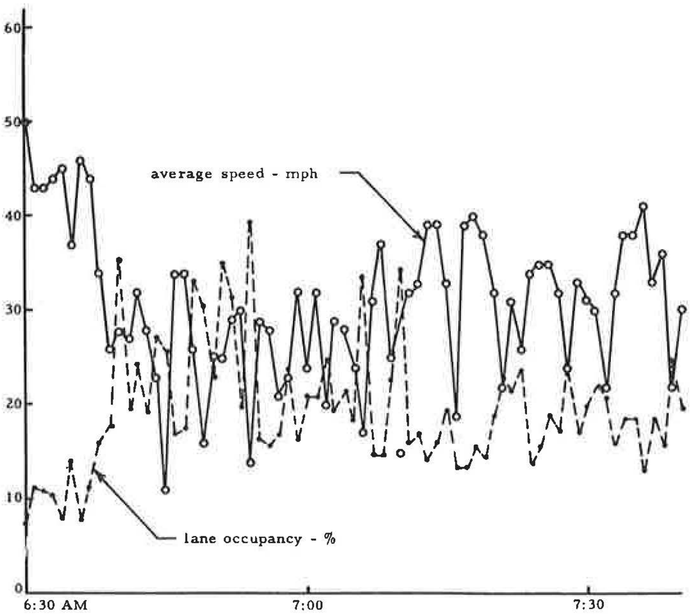


Figure 3. Station 20, Gladstone: Center lane, a.m., southbound; on ramp—25 ft downstream, on on ramp—1970 ft upstream.

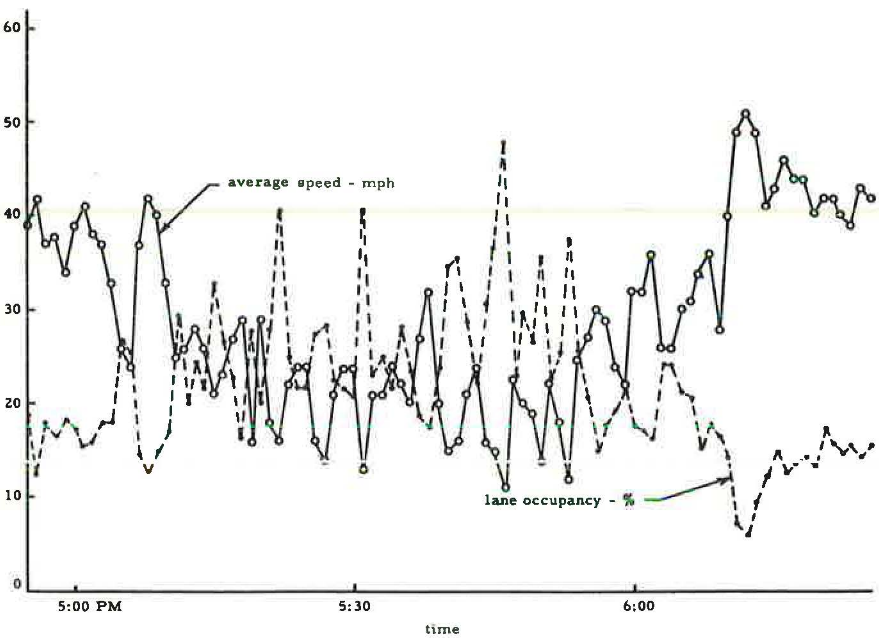


Figure 4. Station 27, Glendale: Center lane, p.m., northbound.

center lane. It would appear reasonable to approximate this relationship by a straight line as shown in the figure. However, much greater insight into the traffic behavior can be obtained by the corresponding dynamic representation shown in Figure 6.

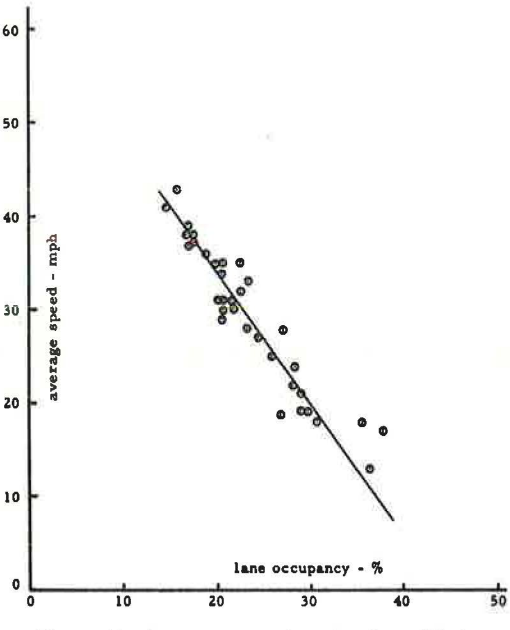


Figure 5. Average speed vs density: High-density region (> 15%), Station 23, Calvert, center lane, a.m., southbound.

Determination of the Period of Oscillation

The question arises as to whether the oscillation is random or has a periodic component. If the oscillation has a periodic component, its frequency may provide some clue of the mechanism in the traffic stream that causes the oscillation.

A very useful technique that mathematically detects a periodicity in an apparently random waveform is the calculation of the autocorrelation function and its associated power density spectrum. This technique, although analytically very powerful, is subject to limitations imposed by record length, sampling time, linearity, and stationarity, and it becomes difficult at times to prove the validity of its application. A full discussion is beyond the scope of this paper. The reader is referred to other sources (2) for details on these matters.

The length of time the traffic flow is in the oscillatory mode imposes a limit on the time record that may be used for analysis. Since the sampling rate is once per minute, only about 36 sample points

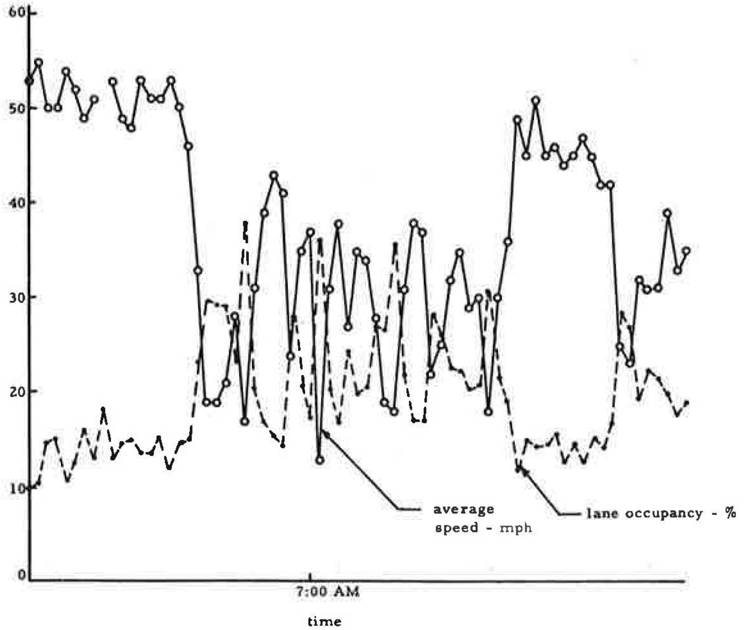


Figure 6. Average speed and percent lane occupancy as function of time: Calvert, a.m., southbound, center lane.

are available for analysis of the oscillatory portion of Figure 6. It would be desirable to have about 250 sample points. To get more sample points, one has two choices: (a) sampling at a higher rate, or (b) obtaining a longer record. Neither of these alternatives was possible in this situation.

The autocorrelation function (more precisely, autocovariance) was computed for the Calvert center-lane speed according to the numerical relation

$$R(\tau) = \frac{1}{N - p} \sum_{n=1}^{N-p} s(t_n)\, s(t_n + \tau)$$

where

N = total sample points,
τ = time shift of a speed time record with respect to itself,
p = sample points lost due to the shift τ, and
$s(t_n)$ = average speed at time t_n, with adjusted mean.

This autocorrelation function is plotted in Figure 7 for a total shift of 20 percent. The power density spectrum was determined by a Fourier transformation of the autocorrelation function. In numerical form

$$S(\omega) = S\left(\frac{n\pi}{m}\right) = \sum_{n=0}^{m\Delta\tau} R(\tau) \cos\left[\left(\frac{n\pi}{m}\right)\tau\right] \Delta\tau$$

τ	$R_{23}(\tau)$
0	60.1
1	9.3
2	-30.1
3	-3.3
4	+20.8
5	+1.1
6	-19.2

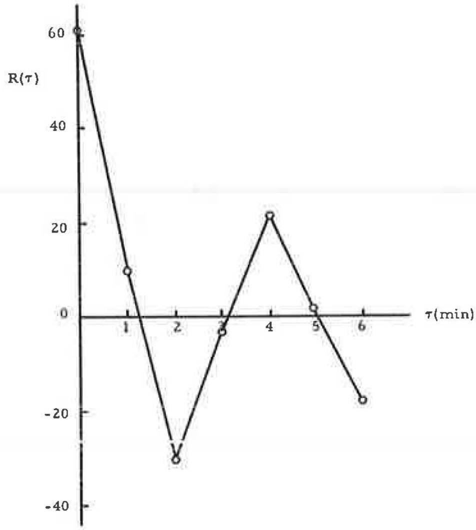


Figure 7. Autocorrelation function: Calvert, a.m., southbound, center lane.

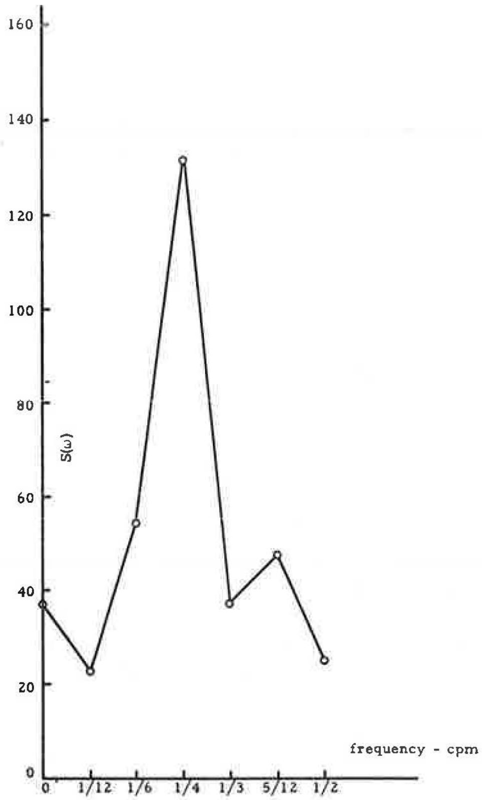


Figure 8. Spectral density: Calvert, a.m., southbound, center lane; folding freq = ½ cpm, "accordion freq" ≈ ¼ cpm.

where

ω = 2π times the frequency, i.e., radians/minute;
$R(\tau)$ = autocorrelation function at shift τ;
$\Delta\tau$ = sampling interval = shift interval = 1 minute; and
m = number of equally spaced $R(\tau)$ intervals.

The result is plotted in Figure 8. No smoothing was incorporated in the computation of the power density spectrum because of the inherent accuracy limitations on the $R(\tau)$ function. Although the amplitude values of the $S(\omega)$ function may be poor estimates of the expected values, it is felt that the relative values indicate the presence of a strong periodic component of the velocity waveform whose frequency is 0.250 cycles per minute.

Propagation Velocity of Mode Switching

Cross-correlation was used to determine the rate at which mode switching propagated in the traffic stream. The cross-correlation function is, in general, subject to the same theoretical considerations as the autocorrelation function. The major difference is that, whereas the autocorrelation function reflects a comparison of a function shifted with respect to itself, the cross-correlation (more pre-

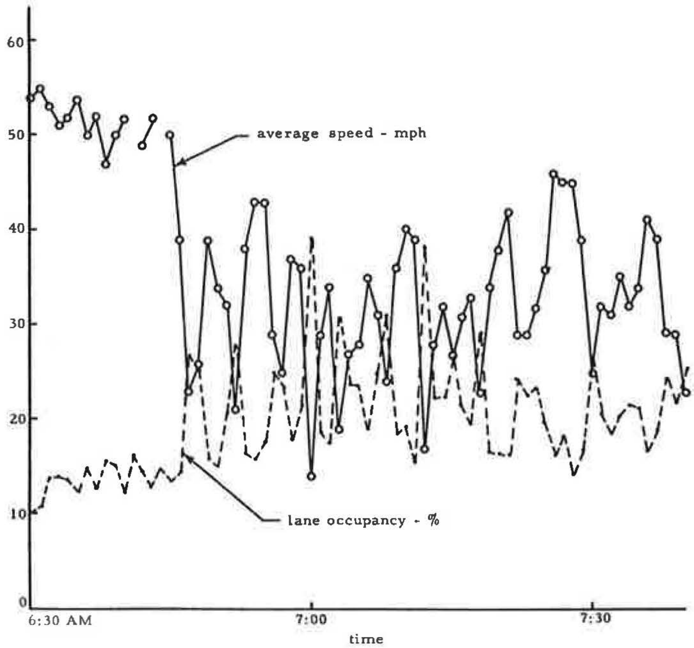


Figure 9. Average speed and percent lane occupancy as function of time: Chicago Blvd., a.m., southbound, center lane.

cisely, covariance) function is a comparison between two different functions shifted with respect to each other.

For example, cross-correlation was applied to the Calvert and Chicago center-lane locations, which are separated by 1435 ft, to see the degree of correlation in traffic behavior between these locations. The two locations were chosen on the basis of geometrical similarity so that the effects of these variables would be a minimum. The average speed and lane occupancy waveforms for the center lane at Chicago Boulevard are shown in Figure 9. A relatively flat cross-correlation function would indicate no connection between behavior at the two locations. On the other hand, a sharp peak would imply a correlation between behavior at two locations but shifted a time τ_m corresponding to the location of the peak in the cross-correlation function. Thus, this would appear to be a good way to trace the propagation of a disturbance, such as a shock wave or switching of mode.

The cross-correlation function was computed from the numerical form

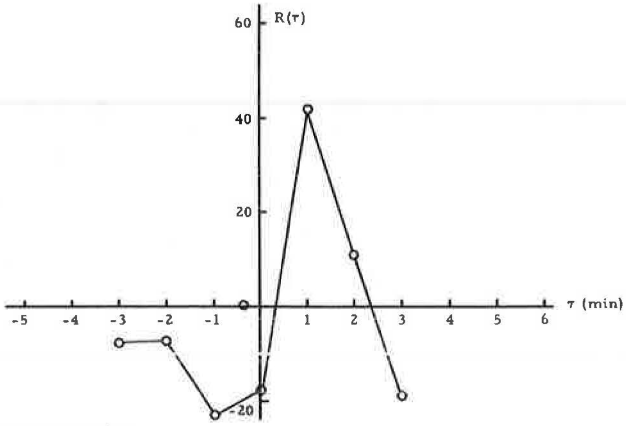
$$R_{34}(\tau) = \frac{1}{N - p} \sum_{n=0}^{N-p} s_3(t_n)\, s_4(t_n + \tau)$$

where

$s_3(t_n)$ = speed at Calvert at time t_n, with adjusted mean;
$s_4(t_n)$ = similar speed at Chicago;
N = total data points common to Calvert and Chicago;
τ = shift in one-minute increments between Calvert and Chicago; and
p = points lost due to shift.

A definite peak in the cross-correlation function was observed at $\tau \approx +1$ minute (Fig. 10). From the separating distance of 1435 ft between the two locations, it was then possible to estimate a propagation of the disturbance upstream at a speed of 16 mph.

τ	$R_{23,\,24}(\tau)$
-3	- 8
-2	- 7
-1	- 23
0	- 18
1	+ 42
2	+ 11
3	- 19



Distance between stations = 1435 ft.
Max correlation = 1 min.
Velocity of propagation (upstream) = 16. 3 mph

Figure 10. Spatial cross-correlation: Chicago-Calvert, a.m., southbound, center lane.

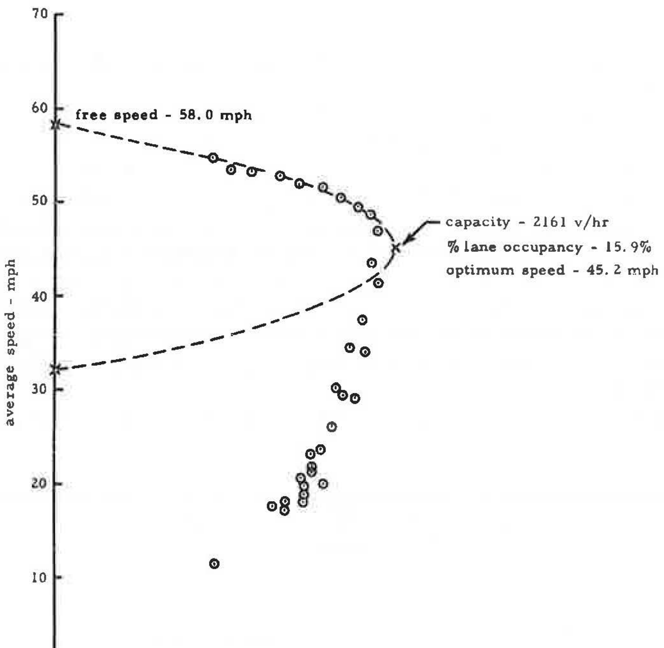


Figure 11. Station 22, Hamilton: Center lane, a.m., southbound; on ramp—50 ft upstream, on ramp—1945 ft downstream.

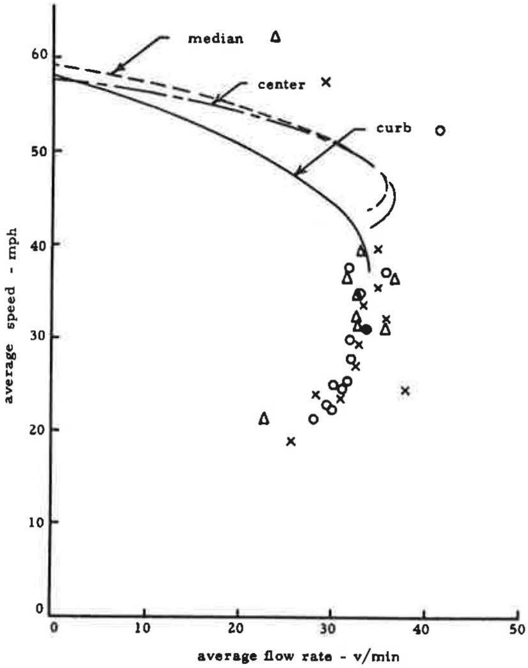


Figure 12. Lane effect: Calvert, a.m., southbound.

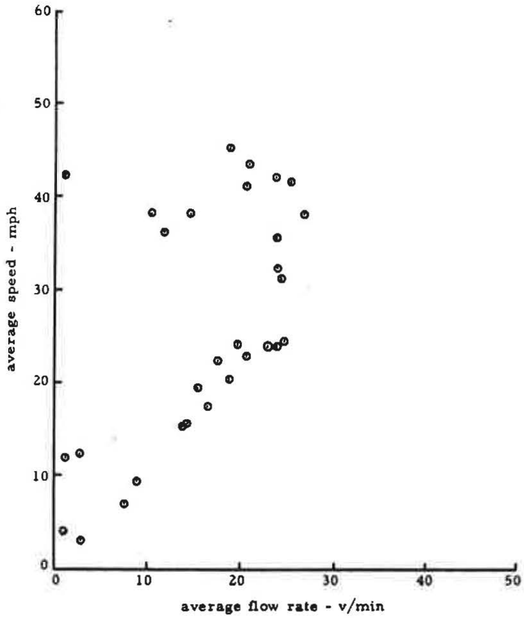


Figure 13. Station 20, Gladstone: Center Lane, a. m., southbound; on ramp—25 ft downstream, on ramp—1970 ft upstream.

This is the same speed as would be predicted for propagation of shock waves using Lighthill and Whitham's theory (3).

STEADY-FLOW MODE CHARACTERISTICS

Figure 11 shows a typical plot of average vehicle speed as a function of average flow rate. It represents the center lane (No. 2) at the Hamilton location in the southbound direction during the 6:00 to 9:30 a. m. period. The average speed in this figure is the average of 1-minute (average) speeds classified in 1 percent lane occupancy intervals. The average flow rate corresponds to the same 1-minute speeds. A least-squares parabolic fit was made using only the data corresponding to traffic flow in the steady flow or non-oscillatory mode as observed from plots of speed, density, and flow as a function of time. In all instances, the parabolic fit was superior to a linear relationship, using the variance σ^2 as a criterion. The capacity, peak occupancy, and optimum speed refer to the values obtained at the vertex of the parabola.

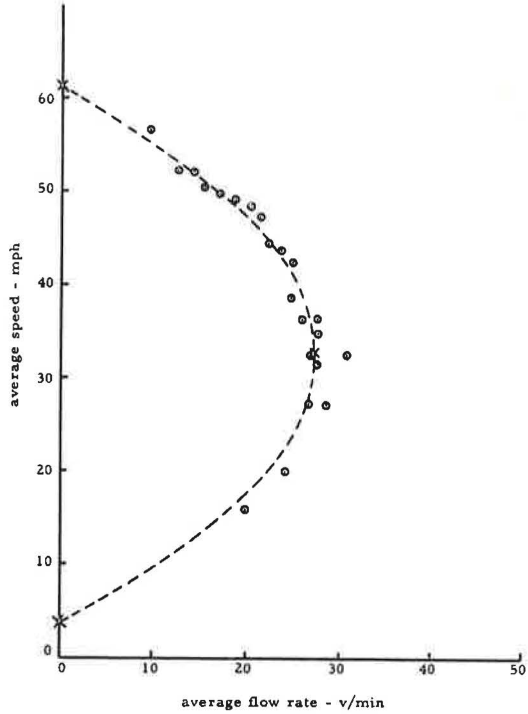


Figure 14. Station 3, Monterey: Curb lane, a. m., southbound; off ramp—40 ft upstream.

TABLE 1

PERFORMANCE AS FUNCTION OF LANE AND LOCATION—SOUTHBOUND

Station	Capacity	Peak Occupancy	Optimum Speed	Free Speed	σ^2	Ramp Location*
Glendale						On: 550 ft US
Curb	1664	0.136	40.46	55.47	0.95	690 ft US
Center	1871	0.145	42.99	58.69	1.68	425 ft DS
Median	1861	0.131	46.97	60.53	4.95	
Monterey						On: 1225 ft US
Curb	1635	0.167	32.47	61.31	0.95	1915 ft DS
Center	1934	0.147	43.75	60.57	8.29	Off: 40 ft US
Median	1948	0.146	44.15	58.92	0.83	
Webb						On: 800 ft DS
Curb	1631	0.124	43.76	60.15	0.84	Off: 1155 ft US
Center	2051	0.152	44.59	61.22	5.29	
Median	2132	0.153	46.15	58.70	5.71	
Calvert						On: 625 ft US
Curb	2029	0.181	37.22	57.94	3.72	Off: 850 ft DS
Center	2191	0.161	45.14	58.32	3.07	
Median	2114	0.152	46.37	59.11	3.40	
Chicago						On: 525 ft DS
Curb	1930	0.170	37.78	56.75	1.15	Off: 700 ft US
Center	2092	0.156	44.49	60.08	2.00	
Median	2244	0.160	46.51	61.20	10.09	

*US: Upstream
DS: Downstream

Typical behavior as a function of lane at the Calvert location is shown in Figure 12. In all three lanes the parabolic fit in the steady flow region was very good—all three fits had $\sigma^2 \approx 3.0$, i.e., a standard deviation of about 1.73 vehicles per minute. Very little difference is indicated between the center and median lanes, but the curb lane is significantly different. Note that the curve in Figure 11 is almost exactly a duplicate of the data in Figure 12 despite the fact that the two locations are separated by 2280 ft.

A comparison of Figure 11 and Figure 13 shows a significant difference in center-lane performance, apparently due to on-ramp effect. Data obtained for Figure 11 were observed 50 ft downstream of an on-ramp, whereas those in Figure 13 were observed 25 ft upstream of an on-ramp. This difference is reflected primarily in the steady-

TABLE 2

PERFORMANCE PROFILE ALONG CENTER (2) LANE—SOUTHBOUND

Station	Capacity	Peak Occupancy	Optimum Speed	Free Speed	σ^2	Lanes
Glendale	1871	0.145	42.99	58.69	1.68	4→3
Monterey	1934	0.147	43.75	60.57	8.29	4→3
Webb	2051	0.152	44.59	61.22	5.29	3
Calvert	2191	0.161	45.14	58.32	3.07	3
Chicago	2092	0.156	44.49	60.08	2.00	3
Hamilton	2161	0.159	45.21	58.02	1.44	4
Clairmount	1931	0.137	46.78	61.13	5.77	4
Gladstone*	1590		38.0			4
Euclid	2086	0.153	46.14	59.84	1.75	4
Seward	1964	0.143	45.66	60.71	2.17	4
Pallister	1923	0.151	42.10	61.47	6.65	4
W. Grand	1823	0.158	38.34	60.10	2.50	3
Holden	INSUFFICIENT DATA					

*No fit possible; assumed peak values from highest flow rate value—26.5 v/min at 38 mph (Fig. 13).

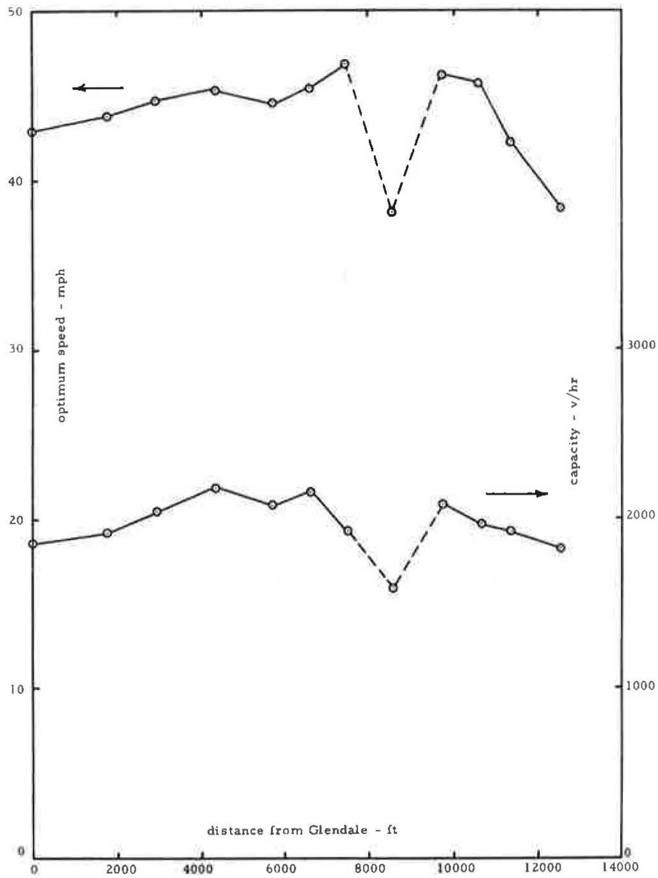


Figure 15. Optimum speed and capacity profile along center lane.

flow mode. No fit was possible to the data shown in Figure 13. This was the only instance where a fit could not be made in the steady-flow mode. Unfortunately, the Gladstone location was the only one in this study that allowed observation of traffic flow immediately upstream of an on-ramp, so that it is difficult to say if it is typical. However, this single observation does indicate that any decision to locate a traffic detector-sensor immediately upstream of a ramp should be weighed very carefully.

No corresponding irregularity was found in the data taken 40 ft upstream of an off-ramp. Figure 14 shows curb-lane data for such a location. If an off-ramp were going to affect the flow, it would be reasonable to expect the irregularities to be most pronounced in the curb lane.

The behavior in the steady-flow mode on a lane-by-lane basis is summarized in Table 1. Note that the peak occupancy (lane occupancy at maximum flow rate) lies in the range of 0.124 to 0.181 for all stations. This represents the vertex of the parabola fitted to the steady-flow mode. It is also near the observed value of occupancy where the system switches into the oscillatory mode. The optimum speed is lowest in the curb lane and highest in the median lane. Capacity is always lowest in the curb lane.

Table 2 includes steady-state behavior for the entire length of the system studied (2.38 miles) along the center lane. Optimum speed and capacity profiles are shown in Figure 15. There is a strong indication that more study should be devoted to determining whether the sharp dip in performance at the Gladstone location is due to inaccurate data or due to some unique disturbance phenomenon at this location (Fig. 13).

CONCLUSIONS

Several highly significant phenomena have been observed in this preliminary study that, if confirmed by further study, would have an important bearing on the development of a good mathematical model for traffic flow:

1. Traffic flow can be divided into two distinct modes: (a) a steady-flow mode below approximately 15 percent lane occupancy, and (b) an oscillatory-flow mode above 15 percent lane occupancy.
2. The oscillatory flow has a periodic component with a frequency of 0.250 cycles per minute for the freeway studied.
3. A disturbance propagation speed was found to be about 16 mph in the case studied.
4. Data collection facilities that include electronic vehicle detectors and a digital computer such as provided at the NPG constitute a powerful research tool for discovering quantitative properties of traffic flow that would go unnoticed by other data collection facilities.

REFERENCES

1. Gervais, E. F. Instrumentation Capabilities and Listing of Reports. Jan. 1966.
2. Blackman, R. B., and Tukey, J. W. The Measurement of Power Spectra. Dover Publications, New York, 1959.
3. Lighthill, M. J., and Whitham, G. B. On Kinematic Waves, II: A Theory of Traffic Flow on Long Crowded Roads. Proc. Royal Society (London), Vol. A229, p. 317-345, 1955.

Service Signing and Motorist's Choice

THOMAS J. FOODY, Ohio Department of Highways, and
WILLIAM C. TAYLOR, Wayne State University

The purpose of this research was to determine the effect of supplemental service information signing on the rural Interstate System in Ohio. The basis for this determination was motorist use measured directly and indirectly. Variations in service facility use patterns were taken as a direct measure of motorist use. Questionnaires, distributed by each service facility, were designed so that motorist use could be inferred indirectly through the measurement of variations in motorist awareness of service availability. The results of this study are based on a series of before-and-after analyses of data collected at five signed interchanges and five unsigned interchanges.

The primary conclusions are that the use of supplemental service signing increases the Interstate traveler's awareness of service availability and that this advance information is utilized. The latter conclusion was substantiated by an overall increase in the service facility use at signed interchanges, an increase in gasoline credit card use at signed interchanges, and an increase in the number of gallons of gasoline in the automobiles of motorists stopping to refill at signed interchanges.

•THE Federal-Aid Highway Act of 1958 provides for a 13-year program of federal-state cooperation to complete the most important highway system in the United States, the National System of Interstate and Defense Highways. This new highway system is being built and operated under a strict set of standards. Sharp curves and steep grades are prohibited wherever possible, as are railroad grade crossings and at-grade intersections.

The system is identified by a special route marker series, and all signing is specifically identified by type, message, and color. Advertising signs are not allowed within a corridor extending 660 ft on either side of the centerline, except at business locations.

Provisions were made to allow the use of standard "Gas-Food-Lodging" signs in advance of interchanges offering these services. However, this method does not supply the Interstate user with any information concerning the particular identity of the services available. Many of today's travelers have definite preferences as to which brand of gas, food, and lodging they use, possibly due to the increasing use of credit cards.

The nation's legislators recognize that today's Interstate user may desire, and more importantly may need, more specific information related to the services available at rural interchanges. The Highway Beautification Act of 1965 contained the following paragraph:

> The House accepted the Senate number and enabling clause and struck out all other language and substituted its own language which did include the amendment by Senator Cooper that service signing be provided on the Interstate System to give the public information as to food, fuel, and lodging, and that Trade names, Trade marks, etc., be used for this purpose.

Paper sponsored by Committee on Traffic Control Devices and presented at the 48th Annual Meeting.

Before the concept of service signing can become a reality, certain questions must be answered. What is the maximum number of services that can be included on each sign? What is the maximum distance from the interchange that a service can be located and still be included in the program? These are just two of the questions that must be answered before a service-signing program can become operative.

However, there is a more fundamental question that highway administrators must consider first: Does the motorist need supplemental service signing? The determination of motorist need is a very difficult and complex task. One approach that would be of value to the highway administrator in this determination is to gage motorist use as an indirect measure of need. The design of this study was directed toward this end—the determination of motorist use.

The purpose of this study was to determine the effect of supplemental service information signing on the rural Interstate System in Ohio. The basis for determination was motorist use measured directly and indirectly. Variations in service facility use patterns were taken as a direct measure of motorist use. Questionnaires, distributed by each service facility, were designed so that motorist use could be inferred indirectly through the measurement of variations in motorist awareness of service availability.

It was also the intention of this research to define the relationship between the effect of service signing and the location of the service facility or its distance from the Interstate. This determination was to aid in the formulation of warrants governing the use of supplemental service signing if this signing was determined beneficial to the Interstate traveler.

A complete inventory of the services available at rural interchanges was taken preceding the selection of test sites. This inventory included all services within 1.5 miles on each side of Ohio's 111 interchanges. The survey indicated that over 90 percent of the available services were located within ½ mile of the Interstate. This fact, coupled with Ohio's relatively flat terrain, meant that the majority of service facilities were visible to the motorist as he made his exit from the Interstate. This compactness and visibility rendered any relationship between service signing and distance from the interchange virtually meaningless. Therefore, the objective was confined to the determination of motorist use.

PROCEDURE

Selection of Test Sites

Ten interchanges were selected as test sites for investigation. Six of these were located successively on I-71 north of Columbus, and the remaining four were located successively on I-75 south of Dayton. Five of the ten interchanges were chosen as study sites: three interchanges on I-71 and two on I-75. These interchanges were signed with the service signing, supplementing the standard "Gas-Food-Lodging" signing. The remaining five interchanges were used as control sites, receiving no supplemental signing.

Every attempt was made to equalize the number of the three types of services available at the study site interchanges and the control site interchanges. An attempt was also made to equalize the distribution of name brand services between the study and control interchanges. Strip maps of I-71 and I-75 showing the service facilities by type and location are included in the Appendix.

Sign Design

Two separate signs were designed for this study, the first to include gas information and the second food and lodging information. The design of these signs was the result of a cooperative effort between the Bureau of Traffic and the HOP Committee (Highway Oil Planners), an ad hoc committee made up of representatives of the Highway Department, Oil Industry, Motel Association, and Restaurant Association, whose function is the resolution of mutual problems created by the overlapping activities of the member groups.

Figure 1. Typical gas service sign installation.

Figure 2. Typical food-lodging service sign installation.

The identification of gasoline service facilities was accomplished by using company emblems only. A maximum of 6 emblem panels per sign were permitted for this study. The overall dimensions of the sign displaying the maximum number of panels were 9 by 14 ft. The use of color was permitted for the emblem design, while the sign background was the standard Interstate blue color (Fig. 1).

Legend and/or emblems were permitted in the identification of food and lodging service facilities. The maximum dimensions of the "Food-Lodging" sign were 14 by 20 ft. This design permitted eight legend panels, a maximum of four for either service. The color design of this sign was white on blue (Fig. 2).

Method of Data Collection

The data collected for analysis were of three main types: (a) manual vehicular movement counts at each interchange; (b) questionnaire distribution at each service facility; and (c) spot speed determination at all study site interchanges.

Data collection for each of the three phases was conducted simultaneously in the before-and-after testing periods. The before testing periods extended from the third week in November 1966 through the second week of February 1967. The after testing period began the second week of March 1967 and ended in the first week of June 1967.

The vehicular movement counts were conducted on one side of an interchange at a time. These counts were taken in one-hour segments varying between 8:00 a.m. and 10:00 p.m., Monday through Friday. The actual number of hours of use counts and the particular hours of the day that they represented were held constant for the before and after testing periods at each interchange. A minimum of 10 hours of counts were collected on each side of an interchange.

Figure 3 shows the types of typical vehicular movements that were recorded. Several distinct movements can be identified as they were utilized to compute the percentages for the service facility use analyses. These percentages are as follows:

1. Ramp use—raw number of vehicles using the exit ramp (movement 1, Fig. 3);
2. Turning percentage—percent of vehicles turning from exit ramp toward the available services on that side of the interchange (% = movement 3 ÷ movement 1);
3. Service facility use percentage—percent of vehicles described (2) that enter any service facility (% = movement 4 ÷ movement 3); and

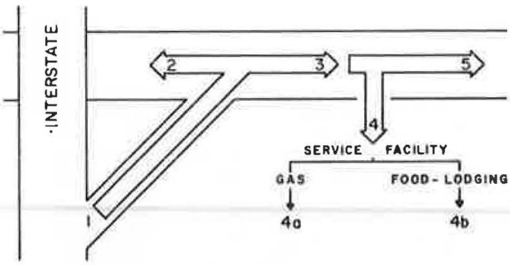


Figure 3. Description of use analyses variables by vehicle movement type.

4. Specific service facility use pattern—percent of vehicles described (2) that enter either a gas service facility or a "Food-Lodging" service facility (% = movement 4a or 4b ÷ movement 3).

The questionnaire distribution was accomplished with the cooperation of the management of each service facility. The distribution and collection of questionnaires was made the responsibility of the employees of each service facility. This involved the distribution of 100 questionnaires by each service facility in both testing periods at all study and control interchanges.

Separate questionnaires (Appendix) were designed for each of the three service facility types, but each delivered the following basic information about the motorist for analysis:

1. Direction of travel prior to entering the service facility to determine exposure to the supplemental service signing;
2. Degree of use of the interchange (frequent, occasional, or first time) to establish interchange familiarity; and
3. Knowledge of the existence of this particular service facility prior to leaving the Interstate.

Only those questionnaires completed by motorists traveling on the Interstate were retained for analysis. In order to eliminate the bias created by interchange familiarity, the sample was further refined to include only those questionnaires completed by motorists using an interchange for the first time. This procedure reduced the questionnaire sample size from 8728 to 1388.

The questionnaires distributed at the gasoline service facilities rendered two additional pieces of information that were used in the analysis. The first was the amount of gasoline in the tank when the motorist stopped to refill. The second concerned the use of a credit card toward this purchase.

Daytime spot speed checks were taken in advance of the supplemental service signing to ascertain the detrimental effect, if any, of the sign design on traffic flow. These checks were made approximately 350 ft in advance of each sign and, as a measure of control, approximately 1½ miles in advance of each interchange having supplemental service signing. This same procedure was followed in the before period, using the proposed location of each sign as the spot speed check site. The control interchanges were not surveyed in this phase of the study.

Description of Analysis

The overall analysis was designed to answer the following specific questions:

1. Were the ramp use patterns altered following the installation of the supplemental service signing?
2. Were the service facility use patterns altered following the installation of the supplemental service signing?
3. What effect did name brand have on these use patterns?
4. Did supplemental service signing provide the motorist using a particular service facility with any additional advance information?
5. Did supplemental service signing influence the motorist's decision of when to leave the Interstate to obtain gasoline?
6. Did supplemental service signing influence credit card use at gasoline service facilities?
7. Did the supplemental service sign design have any effect on the Interstate speeds?

TABLE 1
SUMMARY OF ANALYSIS DESCRIPTION

Data Source	Variable Tested	Statistic Used[a]
Vehicle count	Percent change in exit ramp volume	"t" normal dist.
Vehicle count	Turning percentage	"t" binomial dist.
Vehicle count	Service facility use percentage—total	"t" binomial dist.
Vehicle count	Service facility use percentage—gasoline	"t" binomial dist.
Vehicle count	Service facility use percentage—food-lodging	"t" binomial dist.
Vehicle count	"Better-known vs lesser-known" use—gasoline	Chi-square
Vehicle count	"Better-known vs lesser-known" use—food-lodging	Chi-square
Questionnaire	Service facility awareness percentage—total	"t" binomial dist.
Questionnaire	Service facility awareness percentage—gasoline	"t" binomial dist.
Questionnaire	Service facility awareness percentage—food	"t" binomial dist.
Questionnaire	Service facility awareness percentage—lodging	—
Questionnaire	Credit card use percentage—gasoline	"t" binomial dist.
Questionnaire	Gasoline tank residual	"t" normal dist.
Questionnaire	Credit card use vs supplemental service signing	ANOVA

[a]All tests were of the before-vs-after type, conducted once using study site data and once using control site data.

Table 1 gives a summary of the variables used, the data source for each variable, and the statistical analysis employed with each. Several terms used under "Variable Tested" require some further explanation.

The analysis involving the service facility use percentage was used to determine the effect of supplemental service signing averaged across all gasoline service facilities and all food-lodging facilities. The chi-square analysis involving the better-known vs lesser-known use was employed to determine if this overall effect was related to the presence of brand-name facilities. With respect to food and lodging facilities, better-known facilities were defined as those under chain operation while lesser-known were those operated independently. This determination of better-known and lesser-known for the gasoline service facilities was less obvious. Of the nine different oil companies operating facilities at the ten interchanges included in this project, three operated over half of all the gasoline service facilities in Ohio. It was assumed that these three were better-known than the six remaining oil companies.

The term "awareness percentage" refers to the percent of the Interstate travelers using an interchange for the first time and giving an affirmative response to the "knowledge" question: Did you know this particular service was located here when you pulled off the Interstate? Insufficient sample size prevented any individual statistical analysis of the questionnaires completed at motels.

The term "gasoline tank residual" refers to the number of gallons of gasoline in the automobiles of those Interstate travelers stopping to obtain gasoline and using an interchange for the first time. These data were obtained indirectly from the questionnaire responses. Each interviewee was asked to supply the make and year of the automobile and the number of gallons purchased. This information, used in conjunction with detailed automobile specifications supplied by the Automobile Manufacturers Association, was used to compute the variable of "gasoline tank residual."

The spot speed was conducted by computing the speed variation for each service sign location. The speed variation was defined as the difference in the change in mean speed at the sign location and the corresponding no-sign location. A negative value of the speed variation would indicate that the mean speed did decrease in the approach to a supplemental service sign:

$$\text{Speed Variation} = (\bar{V}_b - \bar{V}_a)_{\text{no sign}} - (\bar{V}_b - \bar{V}_a)_{\text{sign}}$$

TABLE 2

SUMMARY OF ANALYSIS RESULTS

Data Source	Variable Tested	Significance[a]	
		Study	Control
Vehicle count	Percent change in exit ramp volume	No	
Vehicle count	Turning percentage	No	No
Vehicle count	Service facility use percentage—total	Yes (+)	No
Vehicle count	Service facility use percentage—gasoline	Yes (+)	Yes (+)
Vehicle count	Service facility use percentage—food-lodging	No	Yes (-)
Vehicle count	"Better-known vs lesser-known" use—gasoline	No	No
Vehicle count	"Better-known vs lesser-known" use—food-lodging	No	No
Questionnaire	Service facility awareness percentage—total	Yes (+)	Yes (+)
Questionnaire	Service facility awareness percentage—gasoline	Yes (+)	No
Questionnaire	Service facility awareness percentage—food	No	No
Questionnaire	Service facility awareness percentage—lodging	—	—
Questionnaire	Credit card use percentage—gasoline	Yes (+)	No
Questionnaire	Gasoline tank residual	Yes (+)	No
Questionnaire	Credit card use vs supplemental service signing	Yes	—

[a]Level of Significance = 0.05; (+) indicates increase in after period.

Then, this variable of speed variation was plotted against the variable of number of messages per sign type to determine their relationship. Also, the average difference in the before-and-after mean speeds at the signed locations was compared with the average difference at the no-sign location.

RESULTS

Table 2 gives a general summary of the results in terms of statistical significance. Figures 4 through 9 show the results in numerical terms.

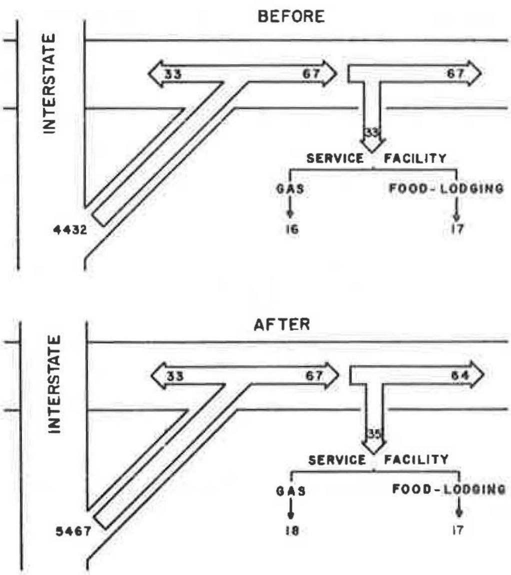


Figure 4. Summary of service facility use data at study interchanges.

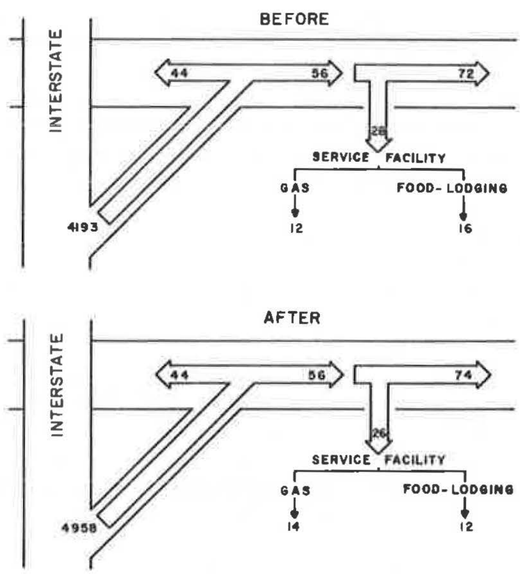


Figure 5. Summary of service facility use data at control interchanges.

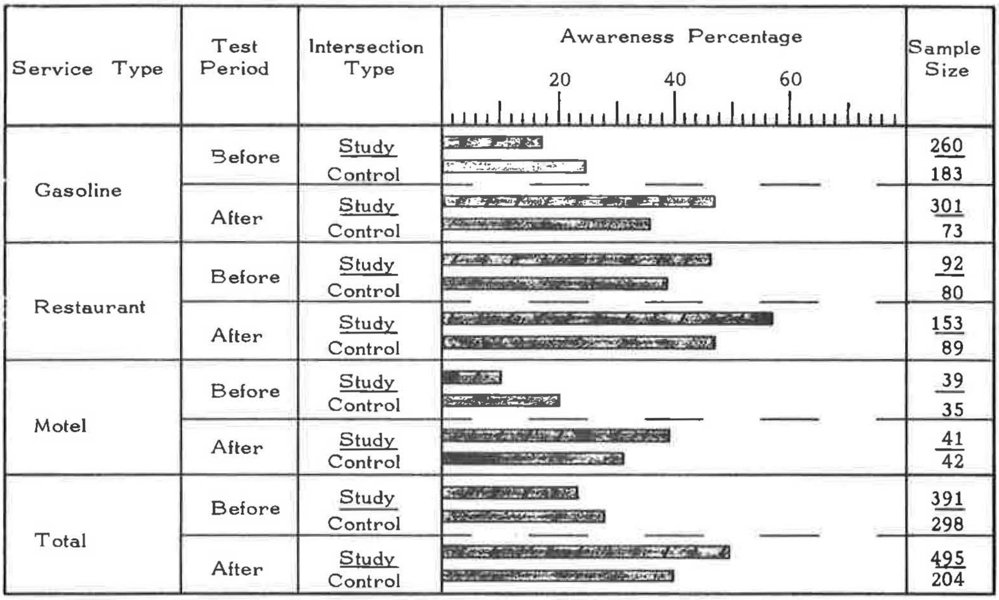


Figure 6. Awareness percentage of first-time interchange users.

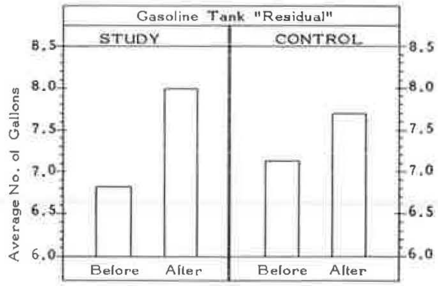
The primary results of the analyses with respect to the variables tested were the following:

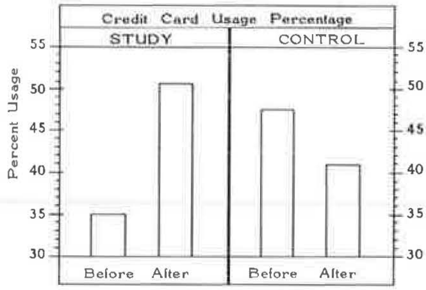
1. The increase in exit ramp volume at signed interchanges (27 percent) was not statistically different from the increase at unsigned interchanges (16 percent).
2. There was an increase in total service facility use at the signed interchanges but not at the unsigned interchanges (Figs. 4 and 5). This increase at signed interchanges was the product of an increase in gasoline service facility use averaged with a "no-change" result at the food-lodging facilities. The "no-change" effect at the unsigned interchanges was the result of a significant increase at the gasoline service facilities compensated by a significant decrease in use at the food-lodging facilities.
3. The variable of name brand (better-known vs lesser-known) was found to have no effect on the service facility use patterns at either the signed or unsigned interchanges (Table 3).
4. The awareness percentage increased significantly at both the signed and unsigned interchanges. Figure 6 indicates that the increase at the signed interchanges was

TABLE 3

RESULTS OF NAME BRAND ANALYSES

Interchange	Service Facility	Brand Name	No. of Facility Users Before	No. of Facility Users After	Chi-Square	Sig.
Signed	Gas	Better-known	303	405	1.32	No
		Lesser-known	159	246		
	Food-Lodging	Better-known	454	549	1.90	No
		Lesser-known	66	101		
Unsigned	Gas	Better-known	199	268	1.18	No
		Lesser-known	75	122		
	Food-Lodging	Better-known	256	228	0.29	No
		Lesser-known	119	97		



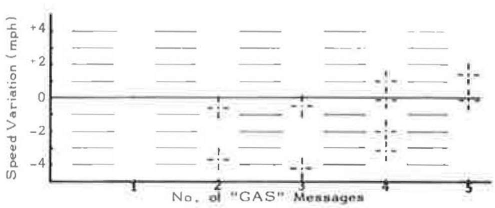


Gasoline Tank "Residual" Analysis					
Signing Condition	Before	After	Difference	"t"	Sig.
Study	6.81	7.99	+1.18	+3.26	Yes
Control	7.14	7.72	+0.58	+0.89	No

Figure 7. Results of gasoline tank residual analysis.

Credit Card Usage Percentage Analysis					
	Before	After	Difference	"t"	Sig.
Study	0.350	0.507	+0.157	+5.02	Yes
Control	0.475	0.411	-0.074	-0.93	No

Figure 8. Results of gasoline credit card use analysis.



Summary of Mean Speed Change Analysis					
Location	N	$\overline{X}$, where $X = (\overline{V}_b - \overline{V}_a)$	$\overline{X}_{ns} - \overline{X}_s$	"t"	Sig.
"Gas"	10	-0.97	-1.18	-1.35	No
"No-Sign"	10	-2.15			
"Food-Lodging"	10	-0.97	-1.18	-1.17	No

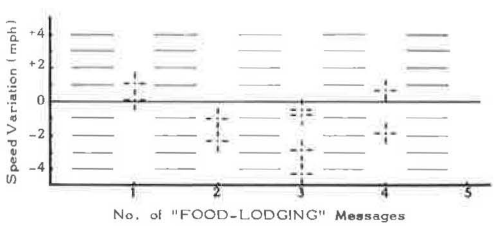


Figure 9. Results of spot speed analyses.

greater than that at the unsigned interchanges in all facility types.

5. There was a significant increase in the average amount of gasoline in the automobiles of those motorists leaving the Interstate to obtain gasoline at unfamiliar but signed interchanges. There was no statistical change in this variable at the unsigned interchanges (Fig. 7).

6. Gasoline credit card use increased at signed interchanges but not at unsigned interchanges (Fig. 8).

7. The mean Interstate spot speeds were not affected by the presence of supplemental service signing (Fig. 9). However, there was an indication of a relationship between number of messages per sign and Interstate speed. This relationship should be considered in the determination of the sign design. Sign design is one phase of the concept of supplemental service signing that merits further research.

The results of the awareness percentage analysis indicate that the Interstate traveler stopping at signed interchanges was provided with more advance information than those choosing to stop at unsigned interchanges. The analysis of the service facility use patterns indicates that this additional information did influence the interchange distribution (result No. 2), but not the name brand distribution (result No. 3).

The results of the restaurant and motel service facility use analyses indicated that it was this group of users that influenced the overall interchange redistribution. The restaurant and motel service facility use at signed interchanges remained constant following the installation of the supplemental service signing, while a significant decrease in use was recorded at the unsigned interchanges. These use results were interpreted as the results of a combination of two factors: first, the presence of supplemental service signing and, second, an outside factor unaccounted for or undefined in this study. The presence of this unaccounted-for factor in the after period resulted in a general decline in use at all restaurant and motel facilities during the data sampling periods. But, the addition of supplemental service signing at selected interchanges within a system of interchanges produced a redistribution of users from the unsigned to the signed interchanges, thus aiding the general decline in use at the unsigned interchanges while offsetting it at the signed interchanges.

In reviewing the results of gasoline service facility use analyses for signed and unsigned interchanges (i.e., identical significant increases of 2 percentage points), it must be noted that all gasoline service facilities employed the on-premise, highrise advertising towers. These results plus those pertaining to the service facility awareness, the credit card use, and the gasoline tank residual were interpreted in the following manner. The use pattern was not affected by the supplemental service signing because today's motorists have definite brand preferences. Many of these motorists, in approaching a signed interchange and being given the names of the available services as well as the distance to the next interchange, made the conscious decision not to stop in the hope that their brand would be available at the next interchange. However, the additional decision-making time and information concerning total service availability provided by supplemental service signing did benefit some motorists. This is evidenced by the motorists stopping at unfamiliar but signed interchanges with an increased awareness of service availability, increased gasoline residual, and an increased use of credit cards.

CONCLUSIONS

The following conclusions have been reached regarding the effect of supplemental service signing:

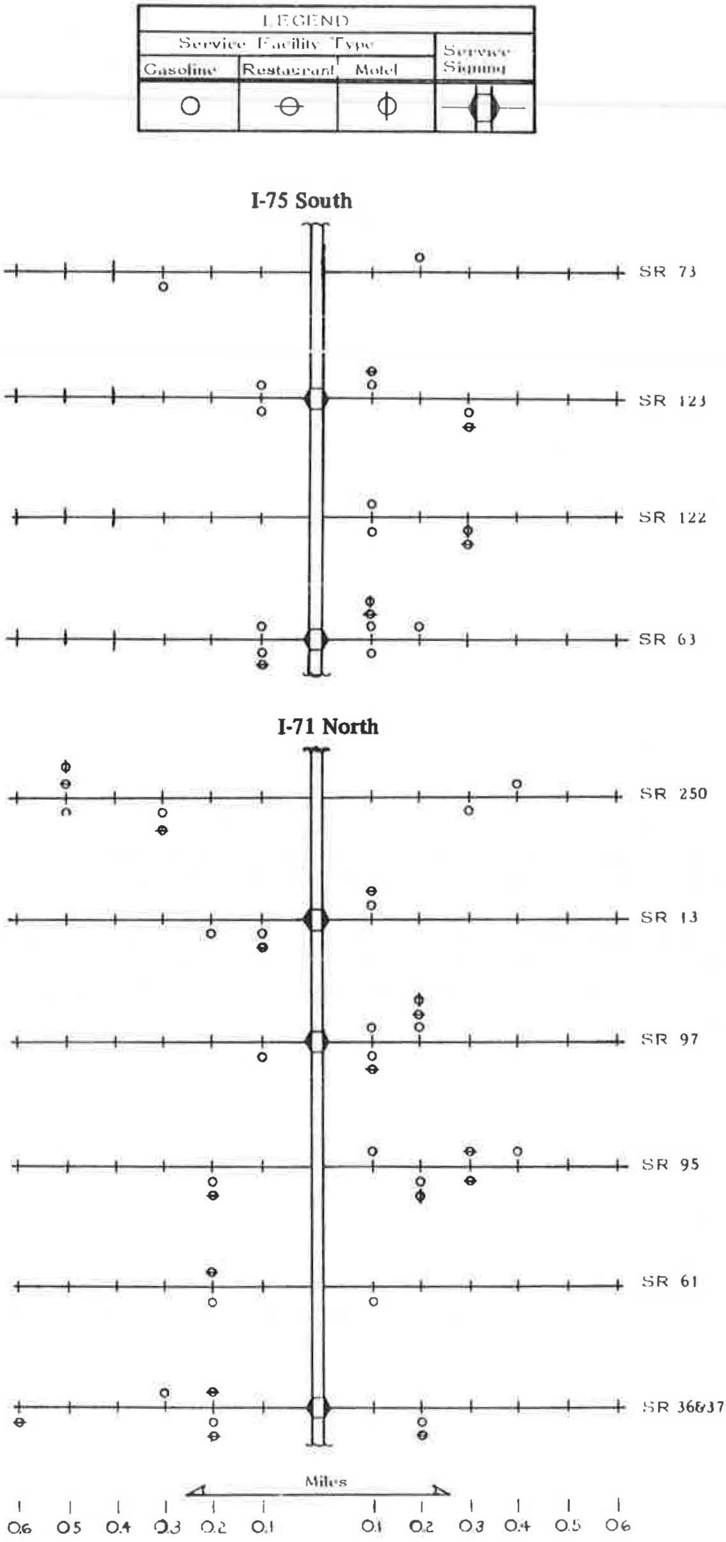
1. The awareness of the Interstate traveler of the existence of specific service facilities is increased by using supplemental service signing.
2. The advance information afforded by supplemental service signing is used by the Interstate traveler as evidenced by: (a) the increase in the service facility use percentage at the signed interchanges; (b) the increase in the amount of gasoline in the tanks of those Interstate travelers stopping at unfamiliar but signed interchanges; and (c) the increase in credit card use at gasoline service facilities at signed interchanges.
3. The overall distribution of service facility users between better-known and lesser-known facilities is not altered by the use of supplemental service signing.
4. Based on the results of the spot speed analysis, the basic sign design used in this study is adequate.

REFERENCES

1. Walker, Helen M., and Lev, Joseph. Statistical Inference. Holt, Rinehart and Winston, New York, p. 157-158, 1953.
2. Li, Jerome C. R. Statistical Inference I. Edward Brothers, Ann Arbor, Michigan, p. 461, 1964.
3. Forbes, T. W., Moskowitz, Karl, and Morgan, Glen. A Comparison of Lower Case and Capital Letters for Highway Signs. HRB Proc., Vol. 30, p. 355-373, 1951.
4. Foody, Thomas J., and Taylor, William C. Service Signing and Motorist's Choice. Report No. 1-141800, Bureau of Traffic, Ohio Department of Highways, 1968.

Appendix

SERVICE FACILITY DISTRIBUTION



TYPICAL QUESTIONNAIRES BY FACILITY TYPE

"GAS"

STATE OF OHIO DEPARTMENT OF HIGHWAYS

BUREAU OF TRAFFIC DRIVER QUESTIONNAIRE

DATE ________ TIME ____ AM ____ PM

1. MAKE OF AUTOMOBILE ________ YEAR ______

2. IS THIS AUTOMOBILE LICENSED IN OHIO Yes__ No__

3. YOU ARE DOING WHICH OF THE FOLLOWING:
   - a. continuing on the interstate ________
   - b. entering the interstate ________
   - c. leaving the interstate ________
   - d. not traveling on the interstate ________

4. HOW OFTEN DO YOU USE THIS INTERCHANGE
   - a. frequently ________
   - b. occasionally ________
   - c. first time ________

5. DID YOU KNOW THAT THIS PARTICULAR SERVICE STATION WAS LOCATED HERE WHEN YOU PULLED OFF THE INTERSTATE Yes____ No____

6. DID YOU USE A CREDIT CARD FOR THIS PURCHASE Yes____ No____

7. HOW MANY GALLONS OF GASOLINE DID YOU PURCHASE ________ GALLONS

"FOOD"

STATE OF OHIO DEPARTMENT OF HIGHWAYS

BUREAU OF TRAFFIC DRIVER QUESTIONNAIRE

DATE ________ TIME ____ AM ____ PM

1. YOU ARE DOING WHICH OF THE FOLLOWING:
   - a. continuing on the interstate ________
   - b. leaving the interstate ________
   - c. entering the interstate ________
   - d. not traveling on the interstate ________

2. HOW OFTEN DO YOU USE THIS INTERCHANGE:
   - a. frequently ________
   - b. occasionally ________
   - c. first time ________

3. WHY DID YOU STOP AT THIS RESTAURANT:
   - a. coffee ______
   - b. snack ______
   - c. meal ______

4. DID YOU KNOW THAT THIS RESTAURANT WAS LOCATED HERE WHEN YOU PULLED OFF THE INTERSTATE ? Yes____ No____

"LODGING"

STATE OF OHIO DEPARTMENT OF HIGHWAYS

BUREAU OF TRAFFIC DRIVER QUESTIONNAIRE

DATE ________ TIME ____ A.M. ____ P.M.

1. YOU ARE DOING WHICH OF THE FOLLOWING:
   - a. continuing on the interstate ________
   - b. leaving the interstate ________
   - c. entering the interstate ________
   - d. not traveling on the interstate ________

2. HOW OFTEN DO YOU USE THIS INTERCHANGE:
   - a. frequently ________
   - b. occasionally ________
   - c. first time ________

3. DID YOU KNOW THAT THIS PARTICULAR MOTEL WAS LOCATED HERE WHEN YOU PULLED OFF THE INTERSTATE Yes____ No____

4. DID YOU HAVE A RESERVATION AT THIS MOTEL Yes____ No____

Response to a CB Radio Driver Aid Network

HERBERT J. BAUER and CLARK E. QUINN,
General Motors Research Laboratories; and
ALGER F. MALO, Detroit Department of Streets and Traffic

The need for a method to provide surveillance of roadways has generated research for many agencies, including the General Motors Research Laboratories (GMR). The Detroit Department of Streets and Traffic is currently operating the GM-sponsored CB Radio Driver Aid Network. This system makes possible the reporting of traffic-related information such as accidents and traffic flow interference. On the basis of earlier demonstration efforts, the system has been expanded to cover the entire Detroit system of surface streets and freeways. The specific functional aspects of the system and a description of its major technical equipment are given.

The paper discusses the results of the earlier operations, presents data on current operational activities, findings of a questionnaire survey, and reactions and conclusions resulting from these efforts. The system is viewed as providing a feasible interim solution to the problem of roadway surveillance, incident reporting, and action implementation to provide safe and efficient traffic flow.

•THE importance of being able to survey dynamics on roadways is a significant problem that has received much attention in the last several years. Specifically, it has been deemed not only desirable but necessary to provide a means of two-way communication with road users. Clearly, communications, taken in the broadest sense, are vital not only for effective and efficient traffic flow, but also for the protection of life and property.

A variety of methods of implementation of surveillance and communications techniques have been suggested. Many of them have been implemented on an experimental basis. These have ranged from simple systems of human observers to highly complicated fixed and mobile systems with varying degrees of complexity and costs. To date, no one system has evolved that meets adequately and simultaneously the technical, economic, and functional requirements of street, highway, and freeway surveillance. The general requirements of such a system have been identified as a need to:

1. Detect situations,
2. Be able to verify the fact that a need for some form of action exists,
3. Be able to establish the criticality of a need,
4. Be able to determine the type of assistance, if any, that is required,
5. Facilitate the dispatch of appropriate aid or initiate necessary control actions, and
6. Acquire data for statistical analysis and research purposes.

The perfect system to satisfy these needs lies in the future. However, it is felt that a meaningful interim solution of the problem for roadway surveillance and two-way communication has been attained by the General Motors-sponsored Citizens Band (CB) Radio

---

Paper sponsored by Committee on Freeway Operations and presented at the 48th Annual Meeting.

Driver Aid Network that has been operated by the Detroit Department of Streets and Traffic since July 1966. This system meets, to a greater or lesser degree, the surveillance-communication requirements stated.

Fundamentally, the system consists of CB radio-equipped vehicles whose drivers report observed incidents to a single base station where an operator is in two-way communication with the incident reporter (CB vehicle operating on CB Channel 9). The base-station operator then transmits action needs via telephone to the appropriate authority or agency that, in turn, dispatches the type of assistance required (Fig. 1).

The fundamental needs of a surveillance-communication system are met by this project in the following ways:

1. The detector is the collection of motorists who are knowledgeable about the system and are equipped to make reports. Clearly, the detection capability of the system is a function of the number of people on the road at any given time with CB units and an interest in reporting their observations, the hours when the base station is monitoring, and the amount of radio traffic and static.

2. The verification function is initially accomplished by the observations of the person reporting, further established by possible additional two-way communication with the base operator, and the judgment of the operator in his decision to report an incident to a given authority. In some instances, the Police Department makes it own additional verification of needs, such as those for a police wrecker to tow away a stalled vehicle. It does this by its "drive-by-and-verify" procedure, in which a patrol car is dispatched to the scene and the officer then makes his own verification and calls his dispatcher.

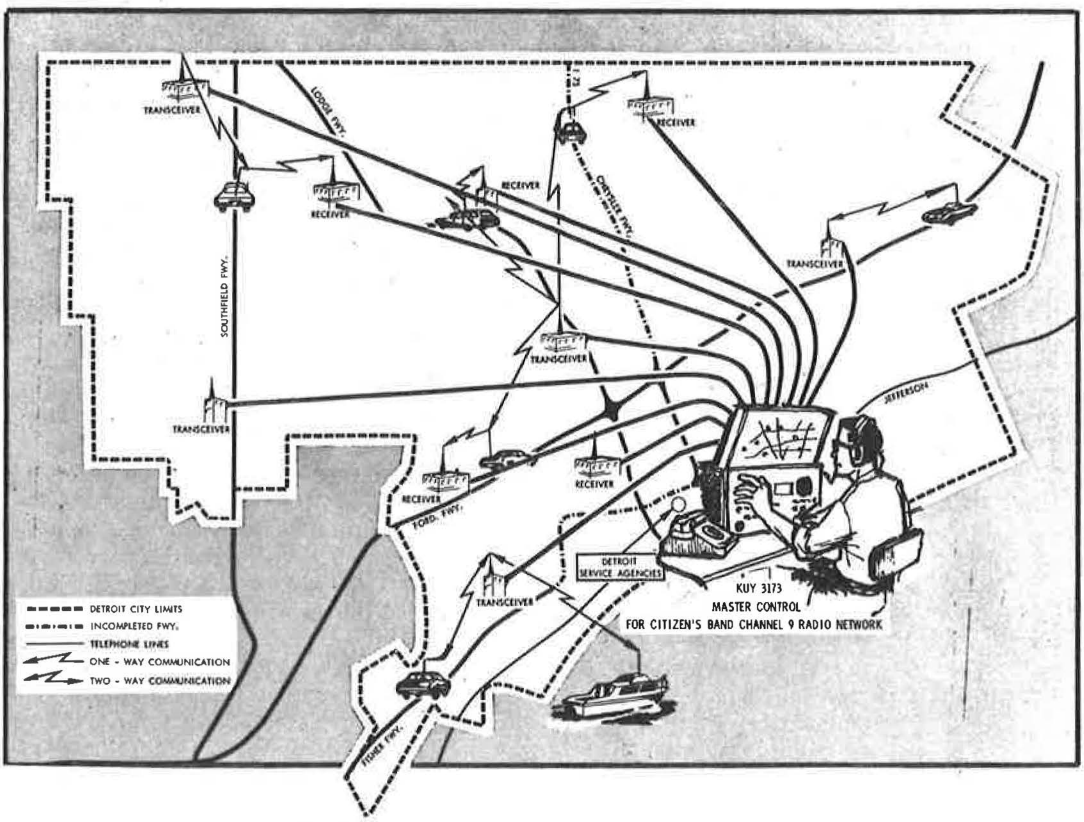


Figure 1. Basic equipment distribution in the project area.

3. Need criticality is established as an integral part of the verification function by the observer, the base-station operator in his inquiries, and, as just stated, in some instances by the responding authority.

4. What has been stated with regard to verification and criticality is applicable in meeting the need for being able to determine the type of assistance or response to be made, and by what authority, agency, or functionary it ought to be made.

5. The dispatch of aid or control action is initiated in the system when the base-station operator makes the telephone call to the appropriate authority or agency. In the case of incidents requiring police response, the base-station operator has at his disposal a direct-line police telephone that is an extension of the police telephone system. Thus, in calling for police response, it is not necessary for the call to be made through the public telephone system. For contact with agencies such as private electric power companies the public telephone system is used.

6. Feedback can be attained, in some cases, by the base-station operator or by further contact with CB mobile units. It can also be obtained by a return call to the authority or agency to whom the report was forwarded. Verification is not a base-station operator need, but a need of the responding action-implementing group, such as a fire department. It is they who must know whether or not a situation has been remedied. This they do in the normal course of their regular routine.

7. Data acquisition is accomplished in this CB radio system by log-keeping. Pertinent data are recorded on cards that are later punched and submitted to computer manipulation.

It is apparent that, while not totally foolproof nor completely satisfactory, the system does meet to a significant degree the fundamental needs of a roadway surveillance system. Some of the shortcomings, such as certain technical imperfections, can be remedied by increasing receiver and transmitter locations, by Federal Communication Commission assignment of a clear channel exclusively for this type of service, and by round-the-clock, seven-days-a-week monitoring. Perhaps the most remarkable attribute of this system is that it does so much toward solving the surveillance problem with relatively unsophisticated equipment.

EARLY DEMONSTRATION SYSTEM

The CB Radio Driver Aid Network concept had its inception as part of the GM Research Laboratories' continuing research effort for safer and more efficient highway travel. Since 1958 highway communications systems projects have been under way (1). In early 1966 the Laboratories proposed a cooperative program with the Detroit Department of Streets and Traffic, the Detroit Police Department, and the National Proving Ground for Freeway Surveillance. The program plan was agreed to and a special Federal Communications Commission (FCC) Class D Citizens Radio Service licence (and later waivers) was obtained. The Citizens Band was used for this project because FCC authorized frequencies are available that may be used with inexpensive, reliable, solid-state transceivers operating directly on power supplied from a 12-volt DC vehicle battery. Channel 9, while in some respects not the most desirable channel, was selected as the operating frequency because the majority of CB'ers are equipped to operate on this channel, and because by original intention this channel was designated as a national "calling channel" when it was established by the FCC. For these reasons Channel 9 was the most logical choice for the demonstration project. On July 26, 1966, CB radio station KUY 3173 and its mobile units began to function. Phase I of the three-phase program was under way.

The original activities were intended to provide coverage only for the John C. Lodge Freeway from Cobo Hall near the Detroit River to Eight Mile Road, a distance of about 12 miles. One transceiver was used. Although transmission was adequate under this arrangement, reception from mobile units was inadequate over approximately 50 percent of the Lodge Freeway due to high RF (radio frequency) interference and the below-ground-level structure of the roadway. Therefore, Phase II was initiated (2). Three additional remote receivers were installed at strategic locations, all feeding via telephone line to the base station. The patterns of coverage, as based on a 2½-mile reception radius, are shown in Figure 2.

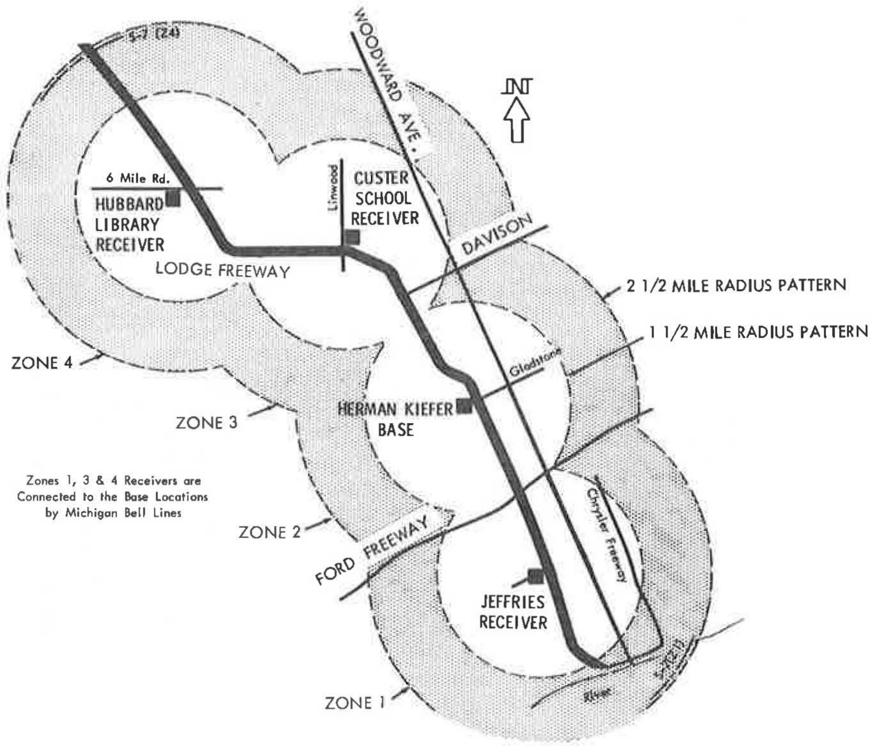


Figure 2. Equipment location and area coverage on Detroit's Lodge Freeway during Phase II operations.

The initial participants of the program, i.e., the "reporters" of incidents, were comprised of approximately 20 City of Detroit employees and about 80 employees of General Motors. Their cars were equipped by GM Research Laboratories with mobile CB transceivers (Fig. 3). Participants were selected primarily on the basis of how frequently they traveled the Lodge Freeway and their willingness to participate in the program. The base station was manned from 6:00 a.m. until 8:00 p.m. 5 days a week.

By coincidence, rather than design or specifically organized publicity, the CB radio community in general soon concluded correctly the purpose and functions of Station KUY 3173 and its mobile units. They spontaneously became an unofficial part of the network and the demonstration program.

PHASES I AND II RESPONSE AND EFFECTIVENESS

The initiation of Phase III was dependent upon the outcome of Phases I and II. A few salient statistical facts are of interest here (3). From August 1 through December 31, 1966, the average number of calls made to the base station was 121 per month. During the same period in 1967 the base was called on the average of 206 times per month. The total number of calls received from the first full operating month of the program through December 1967 was 2,888, or an average of 169.9 calls per month. This volume was not influenced by any public announcement of the system.

It was stated previously that the general CB radio operator community availed itself of the use of the services afforded by the network. Tabulations revealed that during the first 6 months of operation 35.9 percent of the calls to the base were made by persons other than the "official" (GM and City employees) participants in the project.

During April 1968, the Transportation Research Department of the General Motors Research Laboratories sent a two-part questionnaire to the official program participants. One part of the questionnaire required that the respondents identify themselves

Figure 3. Typical CB radio (transceiver) mobile installation.

because it dealt, in part, with technical matters relating to their GM-furnished equipment. Two points gleaned from the identified portion of the questionnaire are worth noting. The 112 participants questioned were asked if they wished to withdraw from the project. Only one individual chose to withdraw! It became apparent from the responses to the questionnaires that many persons made calls—some successfully, others not—from regions well beyond that originally designated as the test area. Their inclination toward wider coverage seemed to be implied.

The second portion of the questionnaire was to be responded to anonymously in order to obtain frank opinions. Virtually everyone suggested that:

1. The monitoring hours to be increased—many wanted 24-hour coverage;
2. Weekends be included in the monitoring schedule; and
3. The geographic area covered by the system be increased.

One section of this part of the questionnaire made provisions for unstructured, voluntary comments. With few exceptions, the comments were very laudatory of the program and/or suggested other functions that the network might assume (Fig. 4).

NETWORK DESCRIPTION

The encouraging results obtained during the first two phases of the Lodge radio reporting project and the enthusiastic support of the Detroit Department of Streets and Traffic and participating drivers led to a proposal by the GM Research Laboratories that the system be expanded to monitor all of Detroit from a master control site with emphasis on the freeways.

This proposal was based on experience with the especially designed control console used for the Lodge network where 4 receivers and one transmitter are monitored and controlled via telephone lines by one operator. Experience with this network had also determined the reception and transmission ranges that could be expected on below-ground-level freeways and in areas of high radio interference. These findings were:

1. Reception range from an average 5-watt mobile on CB Channel 9 to a fixed receiving site of average antenna height of 50 feet, under the stated conditions, is approximately 3 miles.

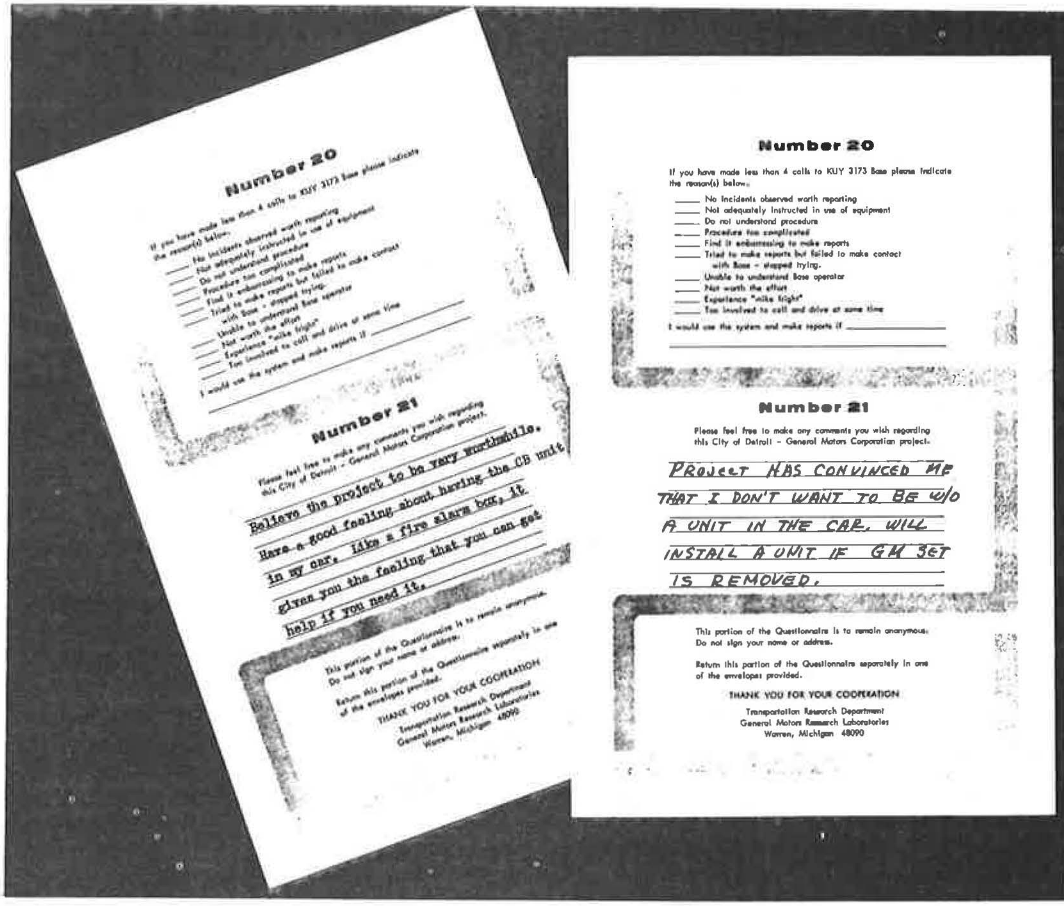


Figure 4. Samples of responses to item 21 of the General Motors questionnaire regarding user opinions of the CB radio network.

2. Transmission range from a fixed location with an average antenna height of 50 feet to a mobile station under the stated conditions is approximately 4 miles.

Based on these criteria and with emphasis on coverage of the Detroit freeways, the city-wide network was laid out on the basis of using city-owned buildings where power and telephone line facilities were available for control and audio transmission purposes. It was determined that in addition to the existing Lodge Freeway system, 6 receiving sites and 4 transmitters would be required. Reception and transmission coverages obtained under the present 10-receiver/5-transmitter arrangement and locations are shown in Figures 5 and 6.

To substantiate the projected network coverage, a radio survey was made from the new sites using city-owned buildings of the Detroit Board of Education and Detroit Fire Department located as near as possible to the proposed locations. The survey indicated that with reasonable freedom from diathermy and CB traffic interference, the projected network would provide satisfactory Detroit freeway coverage as well as covering about 90 percent of the Detroit area surface streets.

Under Part 95 of the Citizens Radio Service Regulations, wire-line control of a transmitter is permitted only if all the equipment is located on the same premises. It was at this point that the GM Communications Section negotiated with the FCC for a special waiver of the rules to permit wire-line control of 5 transmitters at different locations. Control is interlocked so that only one transmitter can be on the air at any given time.

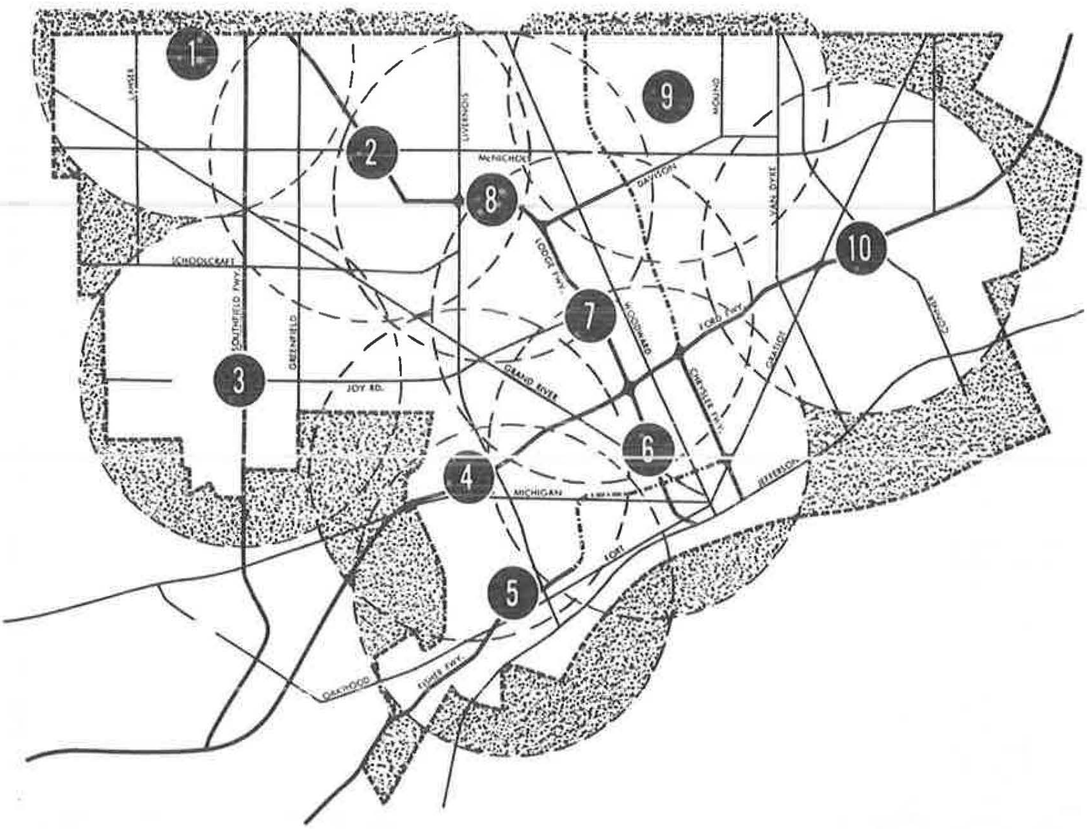


Figure 5. Location and reception zones of the 10 receivers.

With the FCC waiver and remote site approvals in hand, design of the master control console was undertaken and the reworking of commercial transceivers for use in the remote locations and installation of antennas at the sites was begun.

Even to the uninitiated, it is obvious by observing the maps in Figures 5 and 6 that to monitor the 10 receivers located throughout the city and to employ the proper transmitters to answer incoming calls require that the master control console have operator aids. Emphasis had to be placed on simplicity of controls, with the number of operator controls held to a minimum.

EQUIPMENT

Master Control Console

The master control console designed by General Motors Research Laboratories is shown in Figure 7 (4). The unique design provides the operator with aids for orienting himself in relation to the city for visually indicating which remote receiver is picking up a message and to locate the nearest transmitter to employ for answering the caller. Conventional speakers and headphones are provided for hearing the incoming messages.

The upper section of the console is comprised of an illuminated map showing the freeways and principle surface streets in the city. The 10 receiving zones are divided in half by North-South Livernois Avenue. The left group of 5 receivers is heard on the left speaker or headphone and vice versa (Fig. 7). Also shown are the transmitter coverage patterns and pilot lights for the 5 transmitters. This is similar to Figure 6. Selection of any zone receiver for answering a caller illuminates a series of green dots on the map outlining the receiving zone. By observing the map, the operator can determine the closest transmitter to employ for answering the caller.

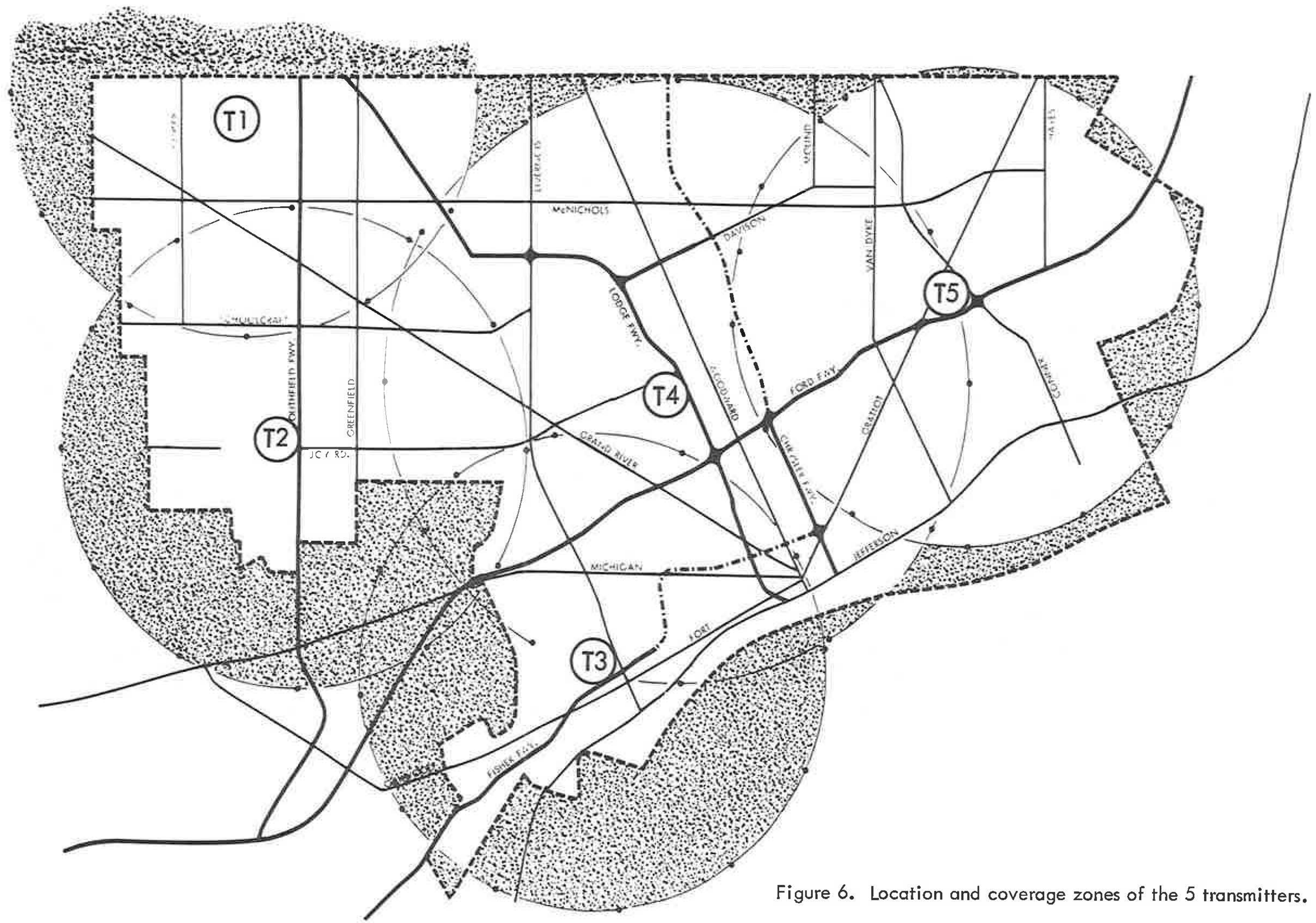


Figure 6. Location and coverage zones of the 5 transmitters.

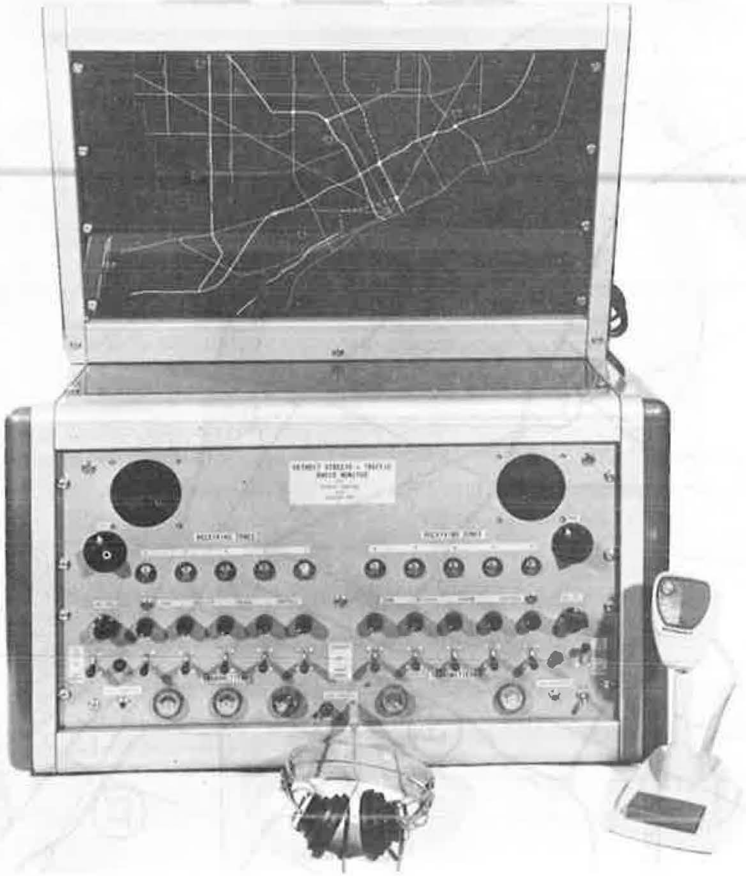


Figure 7. Base station operator's console for control of the 10 receivers and 5 transmitters.

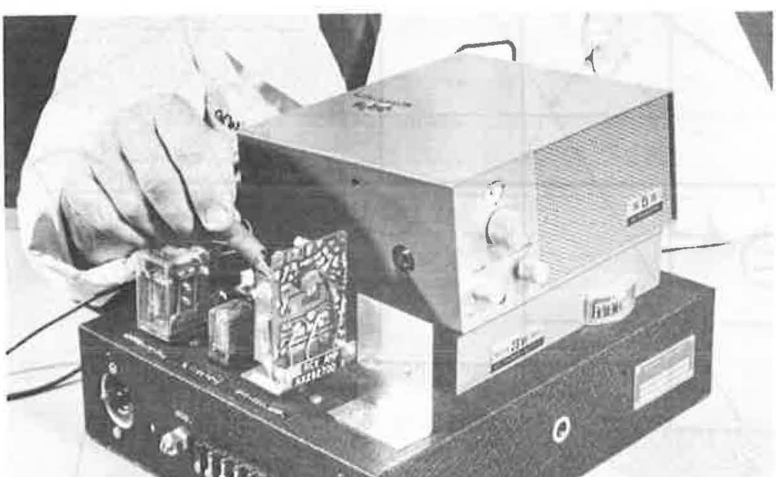


Figure 8. Remote fixed equipment transceiver assembly (cover partially removed) showing relays for wire-line control of power and mode.

The lower section of the console contains the solid-state amplifiers and power supplies. The functional controls are mounted on the front panel of this chassis. Outboard controls on each side of the panel select the mode and adjust volume and tone for each group of 5 receivers feeding individual speakers or phone units. A neon lamp at the top of each zone control group flashes when a voice signal or other modulation is received. An individual zone selector switch mutes all other zones, selects the zone mode, and illuminates the green dots outlining the zone on the map. The bottom row of 5 pushbuttons is used one at a time to activate the remote transmitters. Connections for the telephone lines to the remote units are made to the rear of the chassis.

Remote Receivers and Transceivers

Commercially available transceivers were reworked and mounted in fireproof housings along with control relays and matching transformers for remote telephone line control (Fig. 8).

Interference

Monitoring a network of 10 receivers on Channel 9 is complicated not only by the continuously increasing CB radio traffic, but also by the large number of diathermy machines in the Detroit area operating at frequencies encroaching upon the citizens band. The machines are by far the worst source of interference because they not only have a strong carrier that varies in frequency to swamp nearby receivers, but also they are modulated by harmonics of the powerline frequency and remain turned on for 20- to 30-minute periods.

Figure 9. KUY 3173 base-station operator's area in the offices of the Detroit Department of Streets and Traffic.

OPERATIONAL CONSIDERATIONS

Operating experience during Phases I and II suggested that a more isolated and permanent location be provided for the base station. A separate and exclusive radio room has been provided in the Detroit Department of Streets and Traffic (Fig. 9). It is apparent that the utility of the CB Radio Driver Aid Network can be enhanced by extension of the monitoring hours and the number of days monitored. It is anticipated that, ultimately, such an expansion of operations will be implemented. As of November 1968, plans were under way to utilize the CB network as the communications backbone of a Department of Transportation project for providing emergency medical aid in roadway accidents. Monitoring hours of KUY 3173 will be extended in hours and days of coverage at least for this program.

Another conclusion based on past experience is a realization of the need for operator training. A specific training program has been initiated through the mutual efforts of General Motors Research Laboratories and the Detroit Civil Service Training Division. The course serves as a refresher for current operators, provides basic initial instruction to new employees, facilitates training "standby operators" from the Department of Streets and Traffic, serves to insure a maximum utilization of equipment potential, and provides a uniform, legal procedure for all operators.

It was also decided on the basis of Phase I and II operations to computerize the data recording (log-keeping) system. The system provides greater accuracy in data collection and facilitates data retrieval and manipulation, thus enhancing the routine use of collected information, as well as making research in depth more readily possible (5).

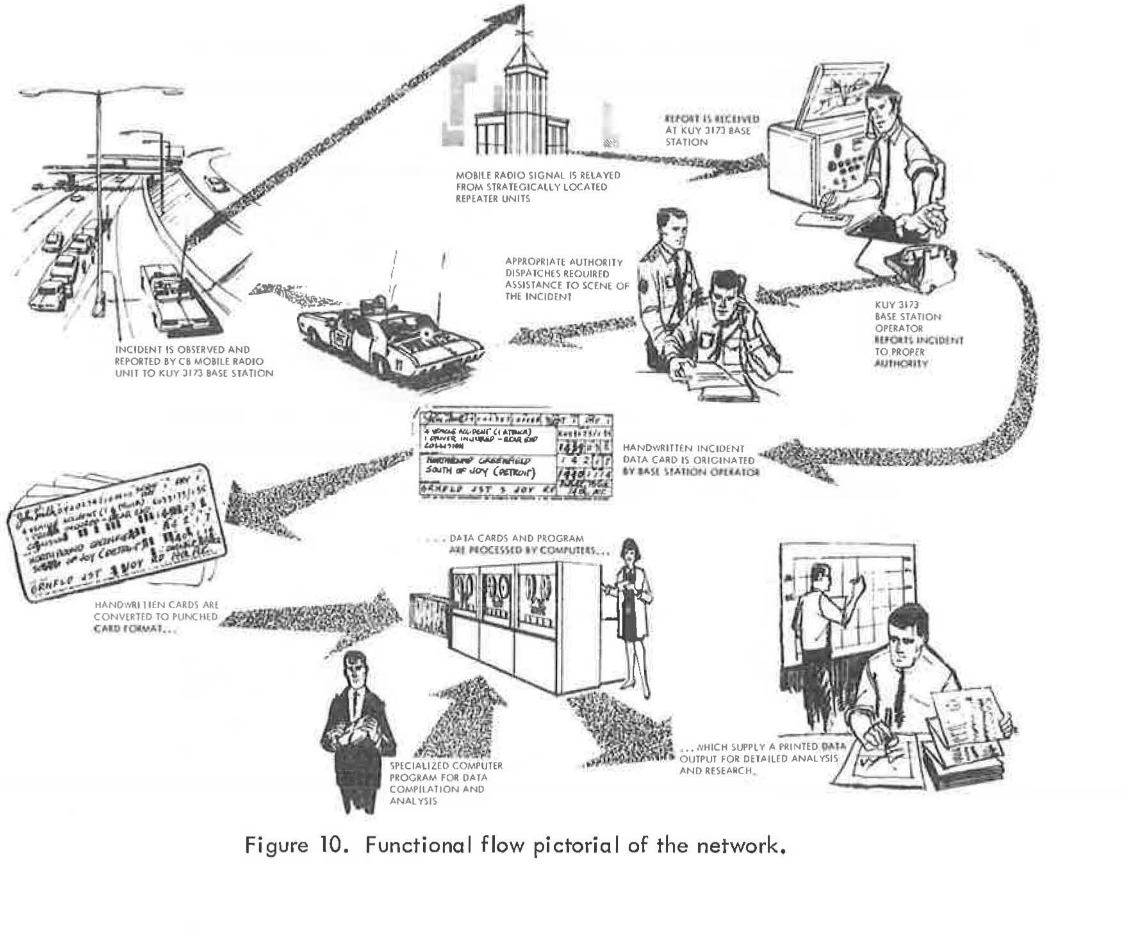


Figure 10. Functional flow pictorial of the network.

DETAIL OF CURRENT SYSTEM FUNCTION

The overall procedures and processes (Fig. 10) of the expanded, city-wide operation begin with the observation by a CB-equipped motorist of an incident that he believes merits reporting because of its influence on traffic flow or the safety of persons and/or property. The motorist calls KUY 3173. The call is received by the nearest receiver and is relayed via telephone line to the operator at the base station. Essentially, 5 basic incidents are of importance:

1. Accidents,
2. Vehicle-caused traffic flow interferences,
3. Non-vehicle-caused traffic flow interferences,
4. Hazardous roadway conditions, and
5. Public equipment or utilities failures.

Once the input information is received by the base operator, he makes a preliminary judgment about whether or not a city authority or particular department should be advised of the incident. A fundamental CB project-authority liaison exists with the Detroit Police Department, Fire Department, various maintenance departments, and public utilities companies, and all of these in Detroit's surrounding communities.

All the transactions and exchanges with the KUY 3173 base station are recorded on cards and, occasionally, on audio tape (Fig. 11). Included in the classes of information retained are:

1. Incident serial number,
2. Date,
3. Operator identification,
4. Weather and roadway conditions,
5. Type of roadway involved,
6. Incident location,
7. Nature of the incident,
8. Actions taken by the base operator, and
9. Technical and equipment-related information.

Some of the data are prerecorded while some information is added after the communications have terminated.

A great potential exists for statistical analysis and comparison. At given time periods, the handwritten cards are keypunched for computer use. A basic program exists that decodes the various entries and provides a computer-originated printed output. Many variations of the program are possible. Many combinations and varieties of data analysis may be undertaken for traffic engineering purposes, special events analysis, and research purposes.

OPERATING EXPERIENCE

Earlier, reference was made to the increasing number of calls made to the KUY 3173 base station. Figure 12 illustrates the growth. As network coverage expanded, more and more of the public CB community came to use the facilities of KUY 3173, and, we believe, to depend on it for reporting emergency traffic incidents and conditions. An unusual "high" occurred in June 1967, when there were several severe rainstorms in the Detroit area. In July 1967, the base station was inoperative for several days as a consequence of a local civil disturbance. Figure 12 reflects these facts.

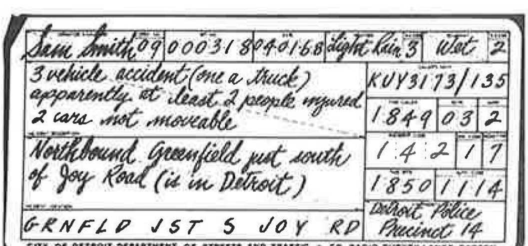


Figure 11. Computer card log record of KUY 3173 base station communication exchange.

As the demonstration program moved from Phase I to Phase II—from a single transceiver to 4 receivers and one trans-

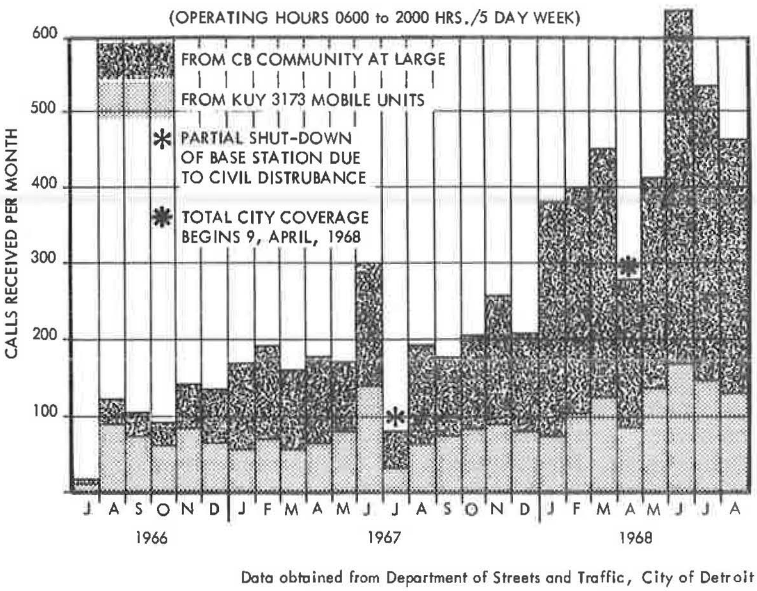


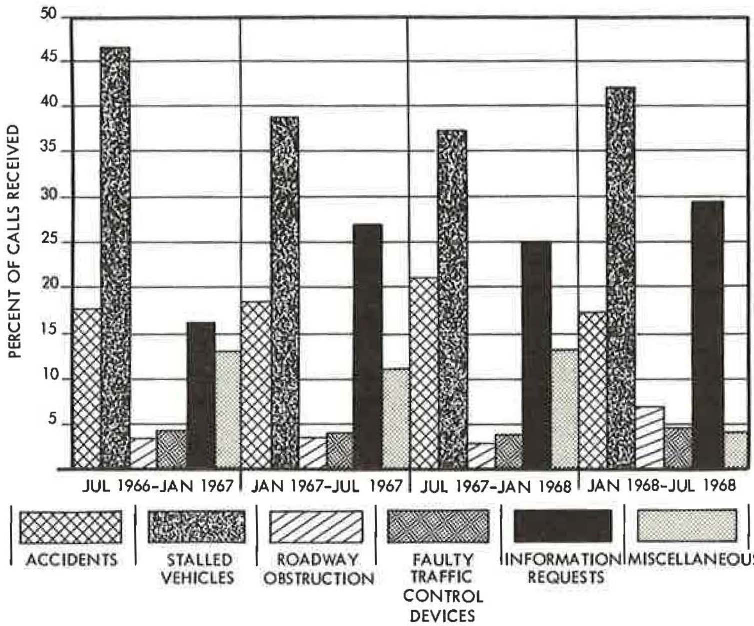
Figure 12. Total monthly calls received by KUY 3173 base station for period from June 1966 through August 1968.

mitter—the average number of calls per month increased from 142 to 255. Phase I ended April 2, 1967. Phase II ended April 8, 1968. Phase III—10 receivers and 5 transmitters providing city-wide coverage—went into operation April 9, 1968.

The bars in Figure 12 are a composite of two sources of calls: those from KUY 3173 mobile units, and those from non-KUY 3173 (other) CB operators. Throughout all phases of the demonstration program, the number of KUY 3173 mobile units has remained fairly stable and the traffic corridors and times of day traveled by these units has remained essentially the same. As would be expected, the number of calls from these units each month has remained fairly constant. By contrast, as other CB'ers became aware of the KUY 3173 operation and functional purpose, the percentage of calls from the general CB community changed from about 35 percent during initial operations to about 65 percent during Phase II. Presently, this source of calls is about 75 percent of all calls received.

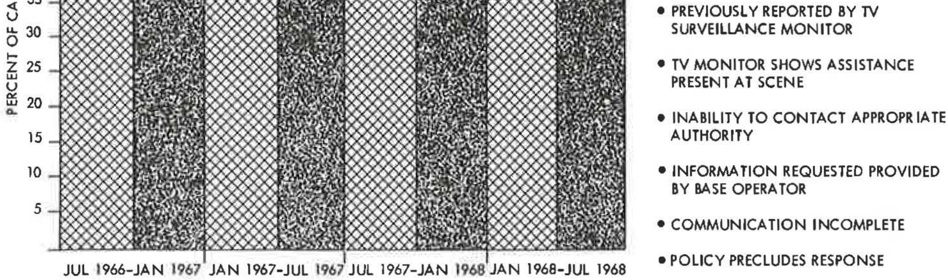
Although the present network covers the entire city, most of the calls received still involve incidents on the freeways. Figure 13 shows that most of the calls made to the KUY 3173 base involve stalled vehicles. On a freeway where traffic density is much greater than on surface streets, and where any interruption of traffic flow creates greater problems than on surface streets, stalled vehicles are a very serious and aggravating problem. Accident reports vary between about 15 percent and 20 percent of the total calls logged. Requests for information are primarily related to traffic conditions on the freeways. The incidents of such requests vary considerably in a given month, depending primarily on the severity of weather conditions during the period.

Not all of the calls received at the KUY 3173 base station are reported to some action-implementing authority (Fig. 14). The high incidence of requests for information has been cited. These do not require authority action, but are merely a relay of information already in the possession of the base-station operator. Obviously, a given incident, e.g., freeway accident, may be reported by several observers. Such redundant reports are not forwarded. In some cases the base-station operator must make a judgment regarding further transmittal to an authority and will, on occasion, decide that no authority action is required. He is equipped to make these judgments on the basis of the operator training program and his immediate experience.



Data obtained from Department of Streets and Traffic, City of Detroit

Figure 13. Distribution by type of incident of calls received by KUY 3173 base station for period from July 1966 through July 1968.



Data obtained from Department of Streets & Traffic, City of Detroit

Figure 14. Operator disposition of calls received by KUY 3173 base station for period from July 1966 through July 1968.

FOLLOW-UP STUDY PLANS

The present record system of the KUY 3173 operations can provide action data through the point of the reporting of an incident to a given authority for action. (The authority or agency to whom the report is made is recorded.) However, the record system does not indicate the nature of the action taken by an authority. Presently, it is not possible to obtain accurate follow-up information. However, it is intended that, to a certain extent, follow-up information will be available through the coordination between the Driver Aid Network functions and another Detroit project.

In July 1968, Detroit received a Department of Transportation grant to develop and evaluate an emergency response system for the care and transport of the injured in traffic accidents. An important part of this study is the evaluation of the various accident reporting systems. One of the major systems to be evaluated will be the CB Radio Driver Aid Network (KUY 3173). The investigators propose to develop models of various systems, structuring each component of the communications from the detection of the accident through the actions of the related authority. To obtain the necessary evaluative information, the investigators will study the records of the various dispatching and responding agencies to relate their actions to incidents reported to them by the KUY 3173 base-station operator.

To insure that the maximum benefits are being derived from this CB radio reporting system, the emergency medical project will evaluate the significance of various operating schedules, including hours and days of coverage not presently provided. Based on the results of this experience, schedules can be developed that should provide for the most efficient use of manpower and equipment.

In summary, incidents occur, they are reported by mobile CB radio to the base station where the information is screened, and, if necessary, the related authority or department is informed. The latter then makes the appropriate response. All the transactions and associated technical information are recorded on cards. These are keypunched at a later time and processed as desired through a card sorter or computer, thus making printed data on activities available.

CONCLUSIONS

The results obtained with this radio reporting program indicate a desire on the part of the motorist to have two-way communication with an authorized base station where he can report traffic problems and malfunctions as well as request aid from police or other services.

It is felt that this experimental program has shown a need for clear highway communications channels that are out of the band covered by general CB radio usage and diathermy interference and are restricted for uses such as those illustrated by the General Motors-City of Detroit CB Radio Driver Aid Network.

REFERENCES

1. Hanysz, E. A., et al. DAIR—A New Concept in Highway Communications for Added Safety and Driving Convenience. Presented at IEEE Vehicular Communication Conference, Montreal, Dec. 1, 1966.
2. Quinn, C. E., Spreitzer, W. M., and Gibson, A. C. Implementation of a Citizens Band Radio Communication System for the John C. Lodge Freeway. IEEE Automotive Conference, Detroit, Sept. 22, 1967.
3. Bauer, H. J., and Quinn, C. E. The Detroit Citizens Band Radio Driver Aid Network. Proc., Automotive Safety Seminar, General Motors Corp., Paper No. 27, July 12, 1968.
4. Quinn, C. E. Instruction Manual for the Detroit Citizens Band Radio Network. General Motors Research Laboratories Publication GMR-798, Warren, Michigan, Aug. 22, 1968.
5. Bauer, H. J. CB Radio Station KUY 3173 Base Log Keeping Instruction. General Motors Research Laboratories Publication GMR-779, Warren, Michigan, July 7, 1968.

Discussion

RICHARD A. PERRY, Systems Engineer, Texas Instruments Inc.—It is becoming increasingly evident that some form of two-way communication with road-user vehicles is essential to public safety and to the expeditious movement of traffic. The work described in this paper is significant in that it provides some of the first quantitative data about an operational system.

Two items brought out by the project are of special significance. First, the system was configured of relatively low-cost, unsophisticated equipment. This is of prime importance in gaining public acceptance of such a system. Even the base-station costs and expenses should not significantly affect the local tax rates. Second, the spontaneous participation of other radio operators points up the unfulfilled need for such a communication capability.

The second principal point of the report is well taken: a separate clear channel is needed for this particular service. The future will also require other additional channels for local road and traffic condition broadcasts and for general service location inquiries such as the location of motels and service stations.

It might be suggested that while vehicle location and route guidance systems could better operate through buried loops or other vehicle detectors, the present emergency vehicle dispatching could well continue to operate on the public safety radio channels. The result would be the three major communication functions: (a) emergency vehicle dispatch, (b) private vehicle communication, and (c) vehicle location and guidance, performed on a non-interfering basis by proper selection of new radio channels and utilization of the existing systems. This work adds significantly to the background of knowledge necessary to allow synthesis of a practical vehicle communication system.

CLARK E. QUINN, Closure—Although the Detroit CB Radio Driver Aid Network is accomplishing its intended purpose, i.e., communication with the driver (in January 1969, KUY 3173 handled 1,197 calls), the authors wish to add a word of caution regarding the establishment of a permanent system for public use where no secrecy of communication exists to protect the person in trouble on the highway.

It may be of interest to know that channels for two-way communication and audio signing have been under investigation by the G. M. Research Laboratories for over 10 years. Our early audio signing research indicated a need for at least two one-way channels to cover local and regional conditions with two extra channels for two-way radio traffic and miscellaneous information resulting in a total of 4 channels. As pointed out by Mr. Perry, these highway department services are now entirely feasible using low-cost HF equipment. Original requests for frequency allocations for highway communication were made by General Motors and the Automobile Manufacturers Association and to our knowledge are still awaiting action by the FCC. Frequencies have been allocated in the UHF band, but low-cost equipment with adequate range is unavailable for their implementation.

The authors feel that communication with the vehicle operator will follow the pattern of a simple two-way voice system at first, followed by audio signing, updating of the two-way system by some means of coding the requests or reports, and last by a driver aid routing system similar to the goal of the GMR DAIR System.

Multilevel Approach to the Design of a Freeway Control System

DONALD R. DREW, KENNETH A. BREWER, JOHANN H. BUHR, and
ROBERT H. WHITSON, Texas Transportation Institute, Texas A&M University

This report deals with the definition of the functional and operational requirements for freeway control systems and the actual design and installation of a prototype. The specifications proposed for designing the prototype control system are (a) the optimal use of acceptable freeway gaps by merging ramp vehicles and (b) the prevention of congestion. The underlying philosophy 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 near entrance ramps. The theory behind this is based on utilizing gap availability and gap acceptance models. Another theory suggests that the prevention of congestion ultimately results in moving more traffic faster. Theoretically, congestion is prevented if demand never exceeds some service volume.

The approach taken in the design of the freeway control system prototype is based on the multilevel concept. The freeway is viewed as a single entity with the control law being split into several degrees of sophistication or levels, with the lower levels directed toward recognizing the influence of short-term factors (gap availability for merging ramp vehicles) and the higher levels reserved for factors that influence performance on a long-term basis (freeway capacity reductions due to accidents, incidents, or geometric bottlenecks).

•IF URBAN freeways are to operate at acceptable levels of service during peak traffic periods, the facilities must be controlled. Although projects to control urban freeways have been established in Detroit, Chicago, Houston, Los Angeles, and Seattle, the theory of designing a control system has not reached the stage where a single unified approach has emerged. Accordingly, the Texas Transportation Institute was awarded a research contract by the U. S. Bureau of Public Roads to develop system design specifications for freeway control by computer control functions integrating local merging control parameters with freeway control parameters. This report concerns functional and operational requirements for such a system and the design and installation of a prototype freeway control system.

The control action in early freeway control systems depended on prior calibration, using historical data. This was not a completely satisfactory approach because the system to be controlled, its environment, and the expected inputs could not be completely described beforehand. Therefore, for classification purposes, this single open-ended control system is of "zero" level because it has no feedback and no memory. Examples of this form of control are complete ramp closure and fixed ramp metering.

Paper sponsored by Committee on Freeway Operations and presented at the 48th Annual Meeting.

The use of surveillance—the continual viewing of freeway traffic in time and space—represents one means of replacing crude open-loop controls with a closed-loop control system (one with a higher sensitivity to such parameter changes as surges in traffic demand and such environmental disturbances as a stalled vehicle or incident on the freeway). Certainly, the closed-circuit television systems on the John C. Lodge Freeway in Detroit and the Gulf Freeway in Houston provide feedback—the comparison of actual freeway operation with the desired operation so that appropriate control action can be taken. Thus, this system may be termed first-level because direct feedback is present for control. Traffic-responsive ramp metering, advisory changeable speed message, and lane closure in response to a deterioration in freeway operations are examples of first-level freeway surveillance-oriented control. Although this type of control system can react to some stimulus related to congestion, it cannot make a conditional choice of action and, therefore, may be described as tactical rather than strategic in nature.

During the last decade, various complex military systems incorporating radar, computers, communication networks, and weapons systems have evolved. Rational methods to control system constituents, which may be spread over a whole continent, and to arrive at a smoothly working entity have been developed. The extension of engineering and scientific methods for designing these systems to achieve proper balance, performance, and economy is called systems engineering. Applying systems engineering to the rational improvement of the undesirable characteristics exhibited by given processes or operation through the addition of instrumentation, controllers, or computers gives automatic control the character of a science and appears to hold great promise as an approach to freeway control.

FREEWAY CONTROL MODEL

Control System Theory

A freeway control system is simply the array of surveillance, communication, and control components designed and connected so as to command or regulate traffic operations. Figure 1 shows one general scheme for representing freeway control systems. 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 "plant" or "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 and desired speeds) and c (for controlled variables such as volume and density). Because the freeway operations phenomenon must be characterized by a multiplicity of inputs and outputs, r and c are vectors.

In Figure 1, d is a disturbance vector that represents unintended inputs to the freeway system that cannot be adjusted, such as environmental factors, weather conditions, accidents, and incidents; m represents signals supplied to the controller regarding those disturbance vector components and output vector components that are instrumentable; n stands for signals from the controller to those control devices in the actuator sub-

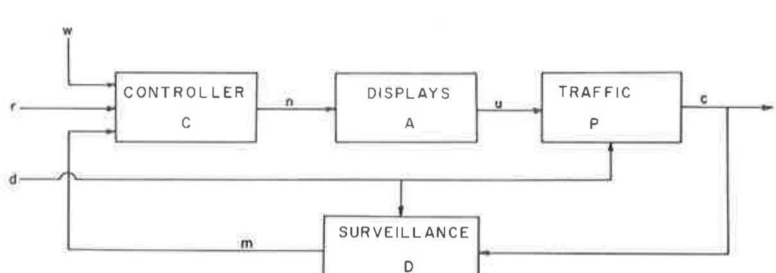


Figure 1. First-level freeway control model.

system, and u is a manipulatable vector that represents those freeway inputs that 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, as shall be seen.

Transfer functions—mathematical descriptions of the ratio of the output of a component to its input—may be written. It is apparent that

$$u = n = f(r, m) = f(r, c, d) \tag{1}$$

In the matter of controlling the system, Eq. 1 may be termed the control law. Ideally a control law is determined so that some performance functional, p, is maximized over a fixed period of time (1). Expressed mathematically, we are trying to accomplish

$$\text{Max. } p = f(u) \tag{2}$$

subject to $c = f(u, d)$ and $f(c, u, r) \geq 0$.

There are several practical problems associated with this development and with system control in general. First, the system vectors and equations developed are rarely known exactly or completely. Second, the mathematical techniques for optimization implied in Eq. 2 are not sufficiently developed to handle the realities and complexities commonly encountered in engineering processes and operations. Third, even where these difficulties can be resolved, the computational problems associated with a single, central controller may make implementation of such a system for on-line control impractical.

There are, moreover, additional problems peculiar to freeway system control. First, there is still a lack of a comprehensive understanding of the traffic stream's behavior. A second area of difficulty regards a suitable analytic description of the system's performance criterion. Ideally, such level of service factors as safety, economy, comfort, and convenience should be included in a system performance criterion.

Failure to understand precisely the complex interactions occurring in freeway system operations is no excuse for proceeding in a haphazard way. The best approach to overcoming these problems lies in separating and resolving the total freeway system control problem into its constituent parts or elements. Two aspects of this approach are (a) the decomposition of the freeway system, and (b) the decomposition of the freeway control function. Both aspects have the objective of reducing the total complex problem into subproblems. The final requirement is a coordination of the subproblem solutions so as to avoid sub-optimization and to achieve the best overall performance of the entire system (2).

Decomposition of Freeway System

After the overall freeway control area has been defined, the system is divided into closed subsystems for analysis. Division points between subsystems should be based on some criteria consistent with the freeway control philosophy. Thus, the division points between subsystems are the freeway "bottlenecks," if the "demand-capacity" philosophy of freeway control is to be employed. If, however, control is to be based on the availability of acceptable gaps in freeway merging areas, each subsystem should be defined so as to contain an entrance ramp, its merging area, and upstream area of influence. For a freeway with substandard entrance ramp geometrics in which the merging areas are in fact the bottlenecks, the subsystem configuration under both control philosophies might very well be the same.

Decomposition of Freeway Control Function

The approach first explained implies that a relatively complex system such as a freeway can be reticulated into a number of independent subsystems, each of which has its own local control law. Another approach, described as the decomposition of the control function, applies a relatively comprehensive control system to the operation of one of these independent subsystems. The freeway subsystem is viewed as a single entity with the control law being split into several levels or degrees of sophistication.

The general question of control of interacting processes has recently been considered through this new viewpoint as an integral part of a theory of "multilevel systems" (3, 4, 5). The concept is relatively new, with its principal value being the provision of a rational means for developing control configurations for extremely complex systems. The multilevel approach is directed toward establishing a hierarchy of control that results not only in an efficient system, but one that can be implemented in stages. The control hierarchy is established so that lower levels—the zero-level and first-level systems discussed previously—are directed to recognizing the influence of short-term factors, whereas higher levels are reserved for factors that influence performance on a long-term basis. There is also a certain ordering of hierarchical levels based on the degree of complexity of computation, the degree of uncertainty, and the required speed of reaction to a change in operating conditions. The central idea is to share the effort of solution among two or more levels, each of which communicates both with the level directly above and that directly below. Generally, the (n + 1)st level influences or even directs the decisions of level n.

Figure 2 shows the conceptual form of a four-level control configuration. 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 layer are identified in the following paragraphs. Most of the variables have been established in previous sections, with superscripts used to denote the respective levels (Fig. 2).

The Regulating Function—This controller accomplished what might be called the basic "subgoal" of the control system. Although the goal of the control system is to provide the best possible level of service on the freeway, its components, and its interfaces, various subgoals have been advanced on which the regulating control subsystem may be based. Implicit is the assumption that optimizing the subgoals will optimize the primary performance criterion. The optimal use of available gaps in the freeway merging process is such a subgoal, and it is accomplished by the regulating controller (Fig. 2). 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).

The Optimizing Function—The object of this controller (Fig. 2) is the determination of optimum operating conditions based on the appropriate performance criterion and

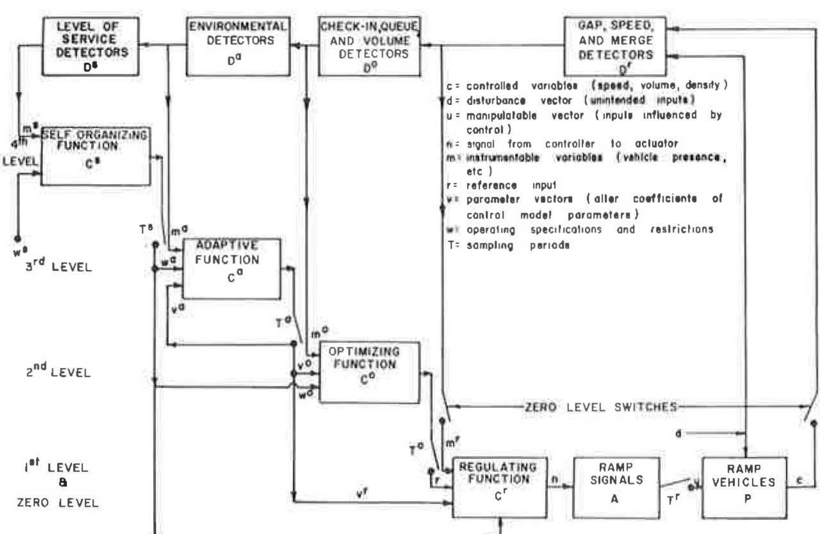


Figure 2. Decomposition of freeway control function.

mathematical model of the process. For example, if the 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. The form of this optimizing function for a freeway control model will be discussed later.

The Adaptive Function—While the two lowest levels of the control hierarchy are developed on mathematical models approximating the real system, the adaptive controller's function (Fig. 2) is to compensate for the errors introduced by the models by adjusting the parameter values, v^o and v^r. A parameter vector v^a is supplied to the adaptive controller so that, in effect, it 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.

The Self-Organizing Function—This controller (6) 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 humans in the system as represented by w^s. The worth or decision vectors generated by the self-organizing function act to control the overall system to achieve the best total system performance. These decisions are based on the accumulated experience and understanding of the system, and are subject to the specifications, goals, and constraints embodied in the worth vector w^s. These decisions w^r, w^o, and w^a alter the control laws that are applicable at each level in the hierarchy.

In conclusion, it is apparent that the higher up the hierarchy one goes, the less rapidly do the environmental and operational conditions pertinent to a given level vary. For this reason, the outputs of the higher levels are considered to be discrete. This is represented in Figure 2 by the samples ("switches" in the logic circuits), which operate with different periods T, in which

$$T^s > T^a \gg T^o \gg T^r \tag{3}$$

Control Synthesis

Decomposition is but a means to an end. It remains to put the submodels together. The final requirement is coordination of the subproblem solutions to ensure their compatibility. An example of some of the difficulties involved may be seen in Figure 2 in that the constraint that $c_i = r_{i+1}$ must be recognized. One can envision the problem arising from a controller in subsystem i attempting to minimize c_i at the same time a similar controller in subsystem i + 1 expects an unlimited r_{i+1} to achieve its optimum performance.

To match the conditions between subsystems to guarantee compatibility, some higher level controllers in the hierarchy must be specified to assess information supplied by the lower level controllers and to provide additional worth or decision vectors to the lower level controllers in order to resolve any conflicting requirements and to introduce environmental or policy conditions not directly accessible by the lower level controllers themselves.

In the freeway control model, it is proposed that this task be resolved at the third (adaptive) level, added to its primary control function. In principle, an n level controller could supervise any number of upstream n - 1 level controllers under its influence. This at least suggests the utilization of an n + 1 level control function to help manage the coordination task of the nth, giving rise to a pyramid of control functions.

FREEWAY CONTROL PROTOTYPE

The process of designing a freeway control system may be described by assuming that the design passes through well-defined phases in chronological order, realizing, however, that a phase is often not recognizable until it has passed. The phases are initiation, organization, preliminary design, principal design, prototype installation, testing, training, and evaluation. This section deals principally with the modeling and design of the prototype based on a multilevel approach.

Prototype Specification

Even in the design of a prototype, the designer must have some specification in mind. That is, he must know what the prototype is supposed to do and establish the means for deciding how well the prototype does it. Thus, the specification is composed of four parts: (a) a description of the location (the plant) and the environment, (b) the description of the inputs (traffic demand, speeds, and other variables), (c) specification of outputs, and (d) measures of effectiveness.

The Plant—The Gulf Freeway in Houston was selected as the proving grounds for this research prototype. Operation on this facility is typical of many urban freeways that have been suffering severe congestion and high accident rates. The Gulf Freeway has three 12-ft lanes in each direction, separated by a 4-ft concrete median with a 6-in. barrier-type curb. The section in the study area extends about 6 miles from the Reveille Interchange to the downtown distribution interchange at Dowling Street. Between these interchanges are eight diamond interchanges containing eight entrance ramps. Frontage roads are one-way and continuous except at three railroad crossings. The through lanes of the freeway overpass the intersecting streets at the interchanges, with the effect of this grade line being a tendency to produce bottlenecks at the overpass as well as limiting the sight distance for the entrance ramp merging maneuver.

Description of Inputs—One difficulty in specifying inputs is in separating those factors that are to be treated as inputs and those that are to be treated as actions of the environment on the system to be controlled. This distinction is important for two reasons: (a) an input is included in functional models, whereas an environmental influence is only important because it modifies the properties of elements; and (b) in using the multilevel approach, inputs are acted upon by the regulating and optimizing controllers, whereas the response to environmental factors is usually handled at the adaptive control level.

In the design specifications for a freeway control system, inputs should be limited to such common traffic variables as volume, speed, density, and gaps, leaving such unintended inputs as bad weather, darkness (during the peak hours in winter), accidents, and incidents to be treated as environmental factors. In the past, comprehensive descriptions of the freeway control system inputs have been made. The relevancy of these descriptions will be discussed later.

Specification of Outputs—One does not just install a control system and then "see what can be done with it." There is often a problem, however, in finding a suitable analytic description of what is desired. Although it is generally accepted that the function of the freeway is to present an environment that permits a driver to operate his vehicle economically, safely, and with a minimum amount of anxiety, it is easier to give a qualitative description than to find an analytic one.

The output specification proposed for the design of the Gulf Freeway prototype consists of two objectives: (a) the optimal use of acceptable freeway gaps by merging ramp vehicles, and (b) the prevention of congestion. The underlying philosophy of the first specification is that minimizing intervehicular 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. The theory behind this first specification is based on the utilization of gap availability and gap acceptance models (7, 8). Behind the second specification is the idea that the prevention of congestion ultimately results in moving more traffic faster. Theoretically, congestion is prevented if demand never exceeds some service volume.

Measures of Effectiveness—The fourth, and in many ways the most difficult, component of a system specifications is a set of measures by which the success of a system can be evaluated. Various figures of merit have been proposed to evaluate freeway operations. To some extent they can be categorized according to whether they are macroscopic or microscopic in nature, rational or empirical, designer- or user-oriented, and according to their sensitivity and capability of automatic measurement.

Because of the complexity of the freeway phenomenon and the relevancy of most common measures of effectiveness to the two output specifications discussed, not one, but several measures were employed in the design and evaluation of the Gulf Freeway control system prototype. The actual utilization of the figures of merit in the test and evaluation of the control system prototype are discussed by Whitson et al (9).

System Design Phase

We have identified the output specification of the prototype control system for the Gulf Freeway as consisting of (a) merging control—making optimal use of gaps in merging areas—and (b) freeway control—preventing a breakdown of operation on the freeway between merging areas. A multilevel control model consisting of a regulating, an optimizing, an adaptive, and a self-organizing function has been developed to provide the theoretical basis for implementing the prototype control system and thus fulfilling the stated objectives of the output specification. Roughly, the regulating and optimizing (first and second) levels are used to accomplish the merging control specification, and the adaptive and self-organizing (third and fourth) levels are employed to fulfill the freeway control specifications.

First-Level Control

To meet the merging control specifications one must be able (a) to detect acceptable gaps on the outside freeway lane, (b) to predict when these acceptable gaps will reach the merging point, and (c) to arrange for the speed adjustment of the merging ramp vehicle so it hits the gap at the merge point. A sensor is required to measure the time interval between successive freeway vehicles (gap detection) and vehicular speed (gap projection). A standard traffic signal installation on the ramp offers a conventional means of communicating with the ramp driver to standardize the required speed adjustment in the merging maneuver. The effect is to stop all ramp vehicles at some point on the ramp far enough from the merge point to allow these vehicles to accelerate to the speed of the freeway traffic system. A call for the green signal is made when the projected gap reaches the position in time (designated the decision point) at which the travel time of the gap to the merge area is the same as the travel time of the ramp vehicle from the signal to the merge area. If the gap is equal to or greater than the designated acceptable gap size for more than one vehicle, the controller holds the green signal until the gap passes the decision point. A loop detector is placed in the pavement of the ramp just upstream from the merge area. All vehicles entering the freeway from

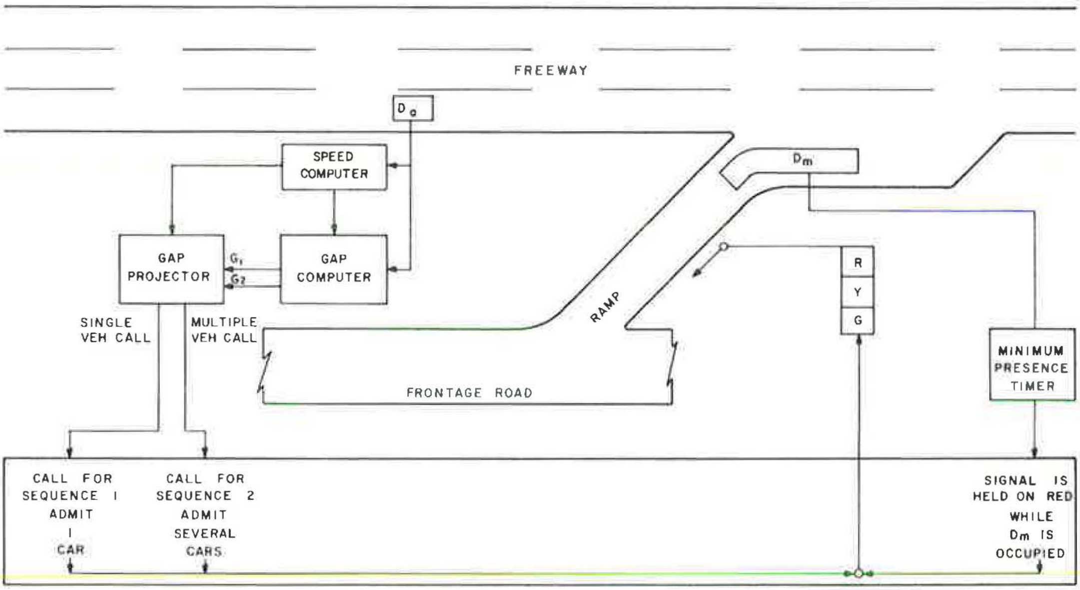


Figure 3. First-level freeway control showing regulating function components.

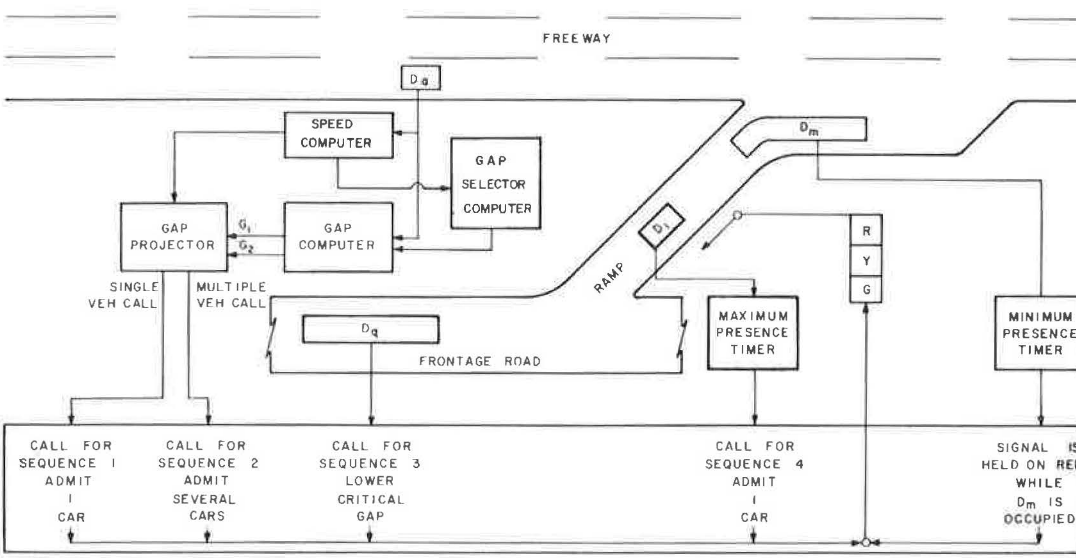


Figure 4. Second-level freeway control with optimizing function components identified.

the ramp will actuate the detector. If a vehicle stops on the ramp in this area, blocking the entrance to the freeway, the detector will time out and the signal controller will hold on red until the detector is cleared.

Second-Level Control

Figures 3 and 4 show the functions and components of the first-level and second-level control systems. Figure 3 shows that the control of the ramp signal is accomplished at the first level basically through the detection and projection of gaps. As a provision for keeping the ramp area from the ramp signal to the freeway clear, the presence of a vehicle in the merge area precludes a green signal indication, thus preventing a queue from forming at the merge point and reducing driver anxiety and the potentiality of accidents.

Whereas the first-level regulating controller compares measured gaps in the outside lane of the freeway with an arbitrary gap setting and then meters ramp vehicles into acceptable gaps, there is no assurance that the resulting operation has been optimized with respect to any criteria. The objective of the second-level optimizing controller is to adjust the gap settings on the first-level regulating controller in response to the outside lane freeway operation (volume and speed) so as to maximize the ramp service volume. This is accomplished by the gap selector computer component shown in the second-level freeway control block diagram in Figure 4 according to a family of curves plotted in Figure 5 developed from a mathematical model described by Brewer et al (10). Reference to high-, intermediate-, and low-type operation is based on the criteria of relative speed of merging vehicles with respect to the freeway traffic stream under stable flow conditions.

In the control of an entrance ramp with an hourly demand of q_r, if the freeway outside lane volume q_f can be sensed, the service gap setting T_s can be readily determined. However, this assumes that the freeway is operating under stable and free-flow conditions, and although the control system is designed to maintain such flow conditions, there will be occasions in practice when the freeway reverts to unstable and forced-flow conditions. Any control system should, of course, be versatile enough to accommodate such occurrences. Based on the volume sensed, the control system is unable to distinguish between free flow and forced flow unless speed is also sensed. Therefore, it has been found most practical to sense speed only because this is also required for project-

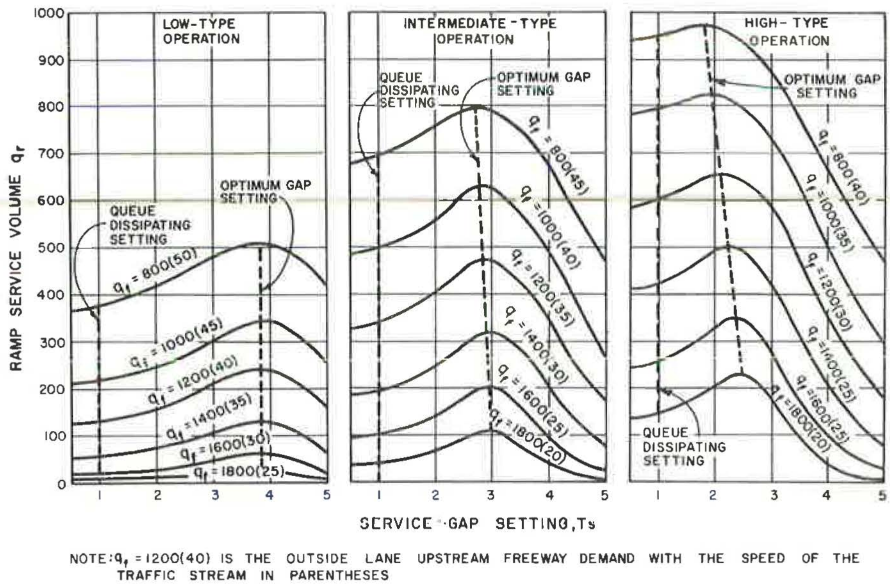


Figure 5. Service volume control curves.

ing gaps. To convert the volumes in Figure 5 to speeds, use was made of the speed-flow relationship measured on the Gulf Freeway. These appear in the parentheses on the curves of Figure 5. For unstable and forced-flow conditions, gap selection may be described by

$$T_s = T_{min} + (k/v); \; v \leq 25 \text{ mph} \tag{4}$$

where T_s is the service gap for freeway speed v, T_{min} is the minimum service gap at high freeway speed, and k is a constant assumed to equal 15.

Two additional functions of the second-level controller are depicted by sequences 2 and 3 in Figure 4. The ramp may be so located that an excessively long queue at the ramp signal will back into an intersection of the frontage road with a cross-street, thus adversely influencing an adjacent traffic system. To minimize this interference with off-freeway traffic systems, it is necessary to detect such an occurrence with a suitably placed presence detector D_q, which, if continuously occupied for longer than a certain period, will cause vehicles to be metered at a faster rate by reducing the service gap to a minimum gap setting (Fig. 5).

Another loop detector D_i is placed in the vicinity of the ramp signal. If a vehicle is delayed at the signal for longer than a certain period, chances are that the driver will assume the signal to be out of order and proceed past the signal anyway, thus violating the control. This period varies among drivers, of course, but is considerably shorter than at a traffic signal on a regular surface street intersection probably because of the somewhat unconventional location of the signal and the absence of any immediate danger in violating the signal. The violating driver is then more often than not forced to stop in the merge area. It is therefore necessary to have a maximum red phase, insuring that the signal will turn green at least once every so often. In practice, on the Gulf Freeway a 20-second maximum waiting time is used.

Third-Level Control

In designing a freeway control system, the automatic detection and location of a reduction in capacity must be given high priority. This reduction may result from either a bottleneck caused by a deficiency in design or from an incident blocking one or more lanes.

In the vicinity of each entrance ramp, six detectors were installed to monitor the accumulation of traffic in an entrance ramp subsystem. Three detectors, located from 1000 to 1500 ft upstream from the entrance ramp nose, are used to determine freeway demand. The speeds at the location of these detectors are also monitored. Three more detectors are located from 500 to 1000 ft downstream from the ramp nose (past the end of the acceleration lane) and are used primarily to detect reduced capacity operation. Low speeds in this area indicate congestion from a downstream bottleneck, with volume counts at this location used to estimate the capacity of the critical bottleneck.

A downstream capacity reduction, other than a geometric bottleneck, may be caused by either the prevailing ambient conditions or by an incident on the freeway. When this effect is detected and the degree of capacity reduction is measured, using the third-level adaptive function components (Fig. 6), these new parameters must be fed to the optimizing controller, which in turn will modify the critical gap setting on the regulating controller. It should be pointed out that a capacity reduction due to the ambient conditions is predictable, whereas an incident can only be ascertained after the fact. The ambiance components are envisioned ultimately as containing a light-meter, thermometer, and rain gage so as to describe driving conditions as evidenced by visibility and the condition of the pavement.

The adaptive controller's function is to adjust the parameters of the lower controllers to compensate for deviations from the assumptions inherent in the mathematical models governing the lower control functions. Another way of looking at this is that the third-level, or adaptive, controller handles the unexpected inputs, often referred to as environmental factors. Ambient conditions and operational incidents represent two such environmental factors; trucks on the entrance ramp may be interpreted as a third.

The two detectors D_i and D_t will be used to classify vehicles. Classification, although normally thought of as distinguishing between trucks and normal passenger vehicles, is not that simple. The significant difference between the two classes of vehicles as inputs to a control system is not their size, shape, weight, etc., but rather their accelerating characteristics. A fast-accelerating, empty truck may well be placed in the same category as an ordinary passenger vehicle. On the other hand, a slow-accelerating passenger vehicle may have the same effect as a truck. Consequently, the purpose of the classification really is to distinguish between fast- and slow-accelerating vehicles and, once slow-accelerating vehicles have been detected, revise the operation of the metering equipment accordingly. Use of the two detectors should also provide an estimate of the vehicle's length.

Each of the controllers at each entrance ramp should be operated in a manner that results in optimal operation of that particular system. Until now, no stipulation has been made regarding controller configuration (central or local) or type of hardware (all three levels of computer controllers described could be either analog or digital or a combination of both). In the discussions that follow, a central digital computer is envisioned in the role of either a third- and fourth-level controller or both, monitoring the first- and second-level control functions for each entrance ramp subsystem as well as guaranteeing the compatibility between subsystems.

The controlled portion of the Gulf Freeway is divided into 6 closed subsystems that are being monitored by an IBM 1800 digital computer. Since the shock wave resulting from a capacity reduction travels back through the system at a speed of from 15 to 25 mph, it is important to devise some means of early detection so that the entrance ramp metering rates may be adjusted. Using the system of upstream and downstream detectors (D_{1V} - D_{6V}) shown in Figure 6, a point of reduced capacity can be pinpointed to within a few hundred feet.

A capacity reduction can be caused by three different situations. The most common and frequent stoppage is an incident such as a stalled vehicle or accident, either of which could affect the freeway up to several minutes. Adverse ambient conditions and geometric bottleneck represent two other capacity mitigating factors that can produce shock waves. The effects on freeway operation can be just as dramatic. Whereas capacity reductions produced by any of the three causes can be detected, it is apparent that the effects of environmental factors and geometric bottlenecks can be anticipated.

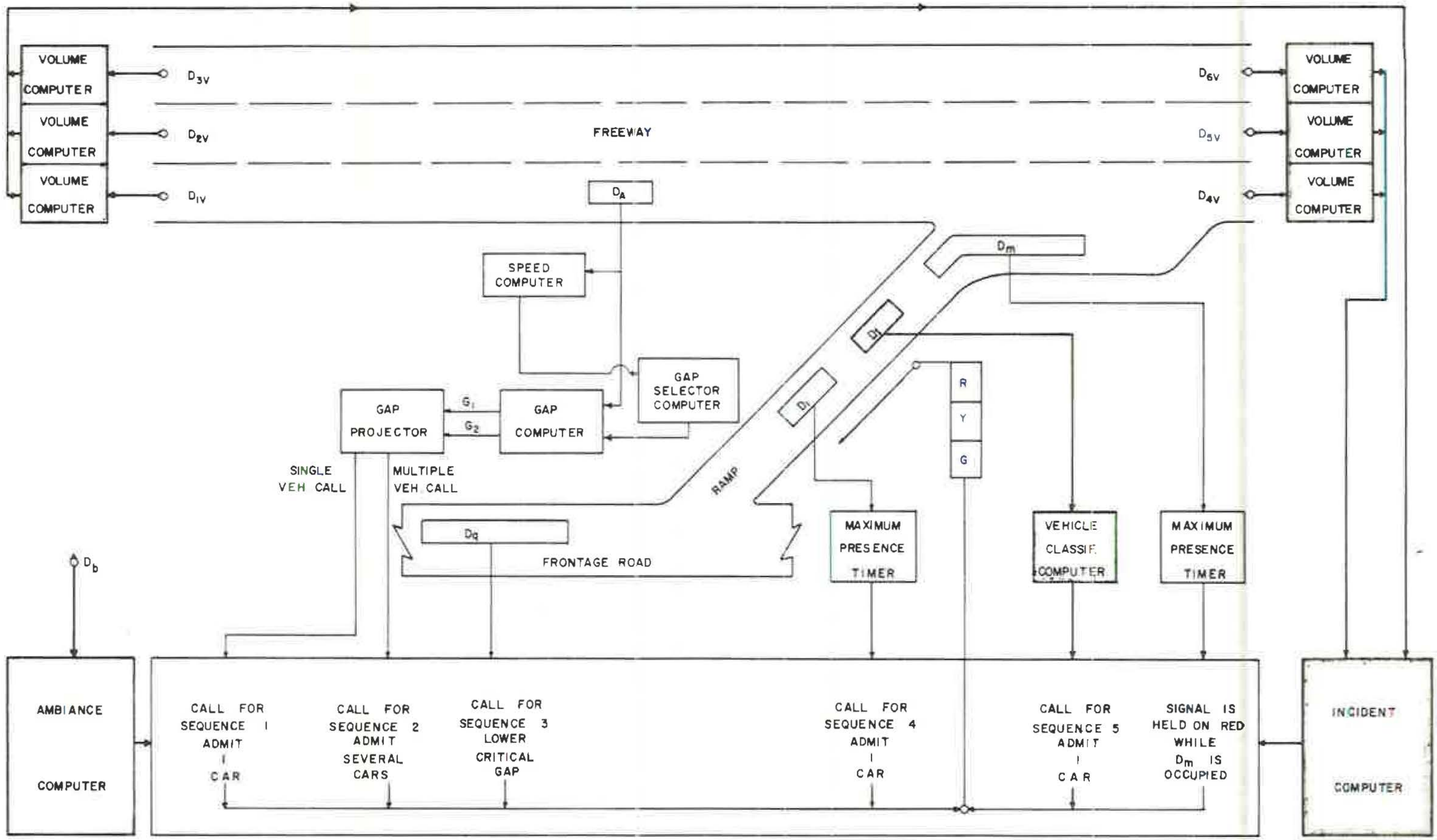


Figure 6. Third-level freeway control with adaptive function components identified.

Fourth-Level Control

In automatic control technology, a number of expressions have been coined to designate the various control systems, such as regulating, optimizing, adaptive, and self-organizing. The fundamental property of a self-organizing or learning control system, as it is often called, is its ability to perform better as time progresses. Using the notation of Figure 2, learning might be implemented as follows: Suppose the optimum performance with respect to a given output specification is accomplished for the parameter settings v_i^j (j = r, o, a) when the input is r_i. Corresponding to a given input r_1, for example, the optimizing system and adaptive systems previously discussed would ultimately settle on the vectors v_1^r, v_1^o, and v_1^a, with the search procedure carried out by C^a always being the same. However, in the fourth-level self-organizing control system, the results of previous computations are stored in memory, which makes it unnecessary that the same lengthy process of attaining the optimum settings v_1^r, v_1^o, and v_1^a be repeated each time the command input r_1 is observed by D^a (Fig. 2). The memory of this simple self-organizing level would ultimately consist of a table such that for each r_1 there would be a corresponding value of v_1^r, v_1^o, and v_1^a. Let us see how this concept may be used in the optimization of a freeway control system.

The fourth-level computer can be programmed to automatically update the parameters used in the three lower control levels. Capacity reductions offer an example of its application to third-level control. The capacities of geometric bottlenecks, an icy pavement, a section of freeway being paint-striped by a maintenance crew, etc., can be "learned". The curves in Figure 5 afford an example of the utility of the self-organizing controller to second-level control. Since the classification of the merging operation at a given ramp as to high, intermediate, and low is based on relative speed, one function of the fourth-level computer is to measure these relative speeds to be sure that the assumption of a given ramp's classification is, in fact, correct.

Complex control algorithms also may be devised. One such approach utilizes a linear programming model (11). Briefly, this model maximizes the output of the freeway system subject to constraints that keep the demand less than the capacity or some specified service volume through each subsystem and that maintain the feasibility of the solution. Additional constraints applicable to the freeway control system proposed in this paper are the control of ramp queues and maintaining the sum of the merging volume in the outside freeway lane and the entrance ramp less than or equal to a specified merging service volume. The translation of these constraints into local controller gap-settings for the Gulf Freeway control system is treated in another project report (10). One special subroutine in this algorithm particularly dependent on this fourth-level concept is the procedure for learning freeway trip origin and destinations as inputs into the linear programming model.

System Hardware

The hardware required to implement a multilevel freeway control system can be categorized into six basic subsystems: sensors, controllers, traffic signals, transmission subsystem, digital computer, and displays. Sensors (devices embedded in or placed above the roadway to detect vehicles) may be of one of the following types: pressure sensitive, inductive loop, ultrasonic, radar, or magnetic. Controllers transform the computer commands into controls for the signals on the entrance ramps. The traffic signals simply present the traffic control indications to the ramp drivers. The transmission, or communication, subsystem provides a means of transferring information from the sensors and controllers to the computer, and transferring command information from the computer to the ramp controllers (if local controllers are used). The digital computer accumulates the incoming data, performs analyses, makes decisions, and sends commands up and down the four-level hierarchy. Displays are incorporated so that the operator and other observers can monitor the status of the computer, the individual ramps, and the overall freeway traffic operation.

Figure 7 shows the Gulf Freeway Surveillance and Control Center, located on the frontage road south of the Wayside Interchange. Analog controllers, built to the first-

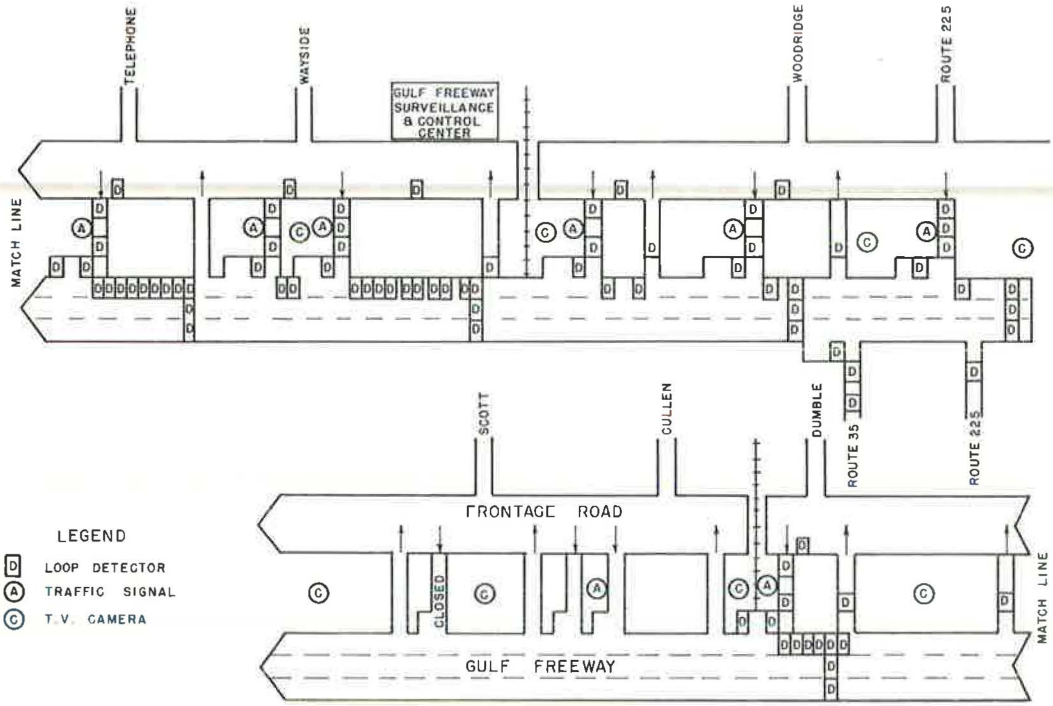


Figure 7. Gulf Freeway prototype surveillance and control system (inbound).

and second-level functional requirements described, have been installed for the control of eight inbound ramps. The new analog computer equipment is shown in Figure 8. In addition, an IBM 1800 digital computer has been installed in the control center (Fig. 9). This computer can be used to either perform the third- and fourth-level functions in conjunction with the local analog controllers, or to perform all four levels as a central computer controller. With the analog and digital computers now installed and operating in the Surevillance and Control Center, a wide range of control measures can be effected—from the simplest to the most sophisticated. This redundancy is not recommended for operational projects, but in a research project this flexibility is needed to compare various control system configurations, to establish control warrants, and to perform cost-effectiveness analyses.

As the name implies, the Center contains surveillance equipment as well as control components. Figure 7 shows the location of a portion of the camera stations for the closed circuit television surveillance system. Closed circuit television is not a part of a freeway control system in the context of the multilevel approach advanced in this paper, it is merely a useful display device to complement the control system. Strategically placed cameras give the observer a view of critical merging areas and bottlenecks for observing the results of decisions changing system operational characteristics. There are other display devices in the control center, such as the digital computer peripheral equipment (plotter, keyboard, etc.), meters in conjunction with the analog controllers, and various event recorders.

PERSPECTIVES

The Enigma of Freeway Control

Freeway traffic control projects are being conducted on sections of the John C. Lodge Freeway in Detroit, the Eisenhower and Dan Ryan Expressways in Chicago, and the Gulf

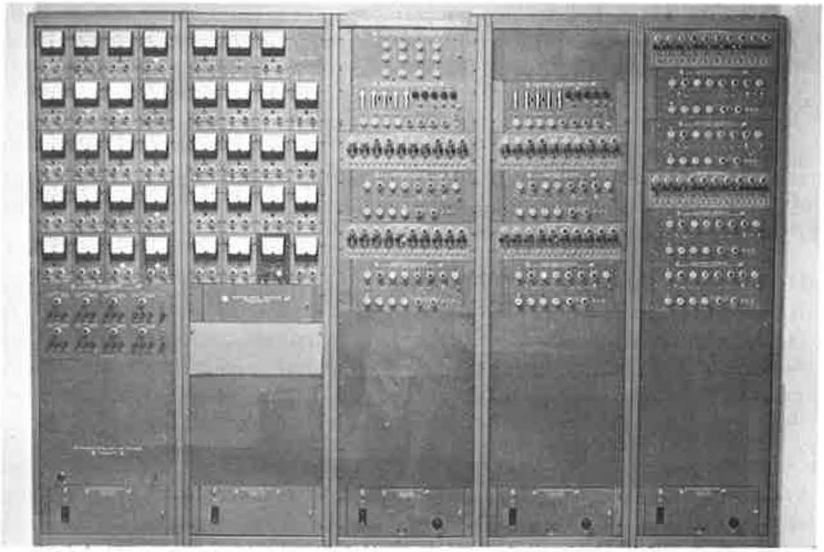


Figure 8. Functional analog computer-controller equipment in Gulf Freeway surveillance and control office.

Freeway in Houston. The results are not only smooth-running freeways and substantial reduction in peak-period accidents, but the aggregate delay to all traffic in the respective corridors has been reduced. Freeway control furnishes the highway administrators with the answer to the creeping innuendo that "freeways won't work."

Figure 9. IBM 1800 digital computer in Gulf Freeway surveillance and control office.

One wonders why, then, if freeway control is such a good thing, there are less than a half-dozen freeways in the country being controlled. Certainly, there are scores of urban freeways throughout the United States experiencing extremely poor operation during the peak hours.

The reason there are not more controlled freeways is the misconception that the design, implementation, and operation of a freeway control system is so complex as to border on the impossible. For example, in a report called "Analysis and Projection of Research on Traffic Surveillance, Communication, and Control" prepared for the Highway Research Board (12), Moskowitz writes:

> Highway administrators will never be able to buy a [freeway control] system, install it, and let it run itself. They will have to face up to the fact that there is such a thing as Operation, and create an organization to operate completed highways. It appears that the most practical way of propagating freeway operational techniques would be to create teams of experts who could act as resident engineers during the formative stages of operating departments.

One can only speculate as to whether or not the same fatalistic prediction was made regarding process control, i. e., that industrialists "will never be able to buy a control system, install it, and let it run itself." Thousands of steel mills, hydroelectric plants, refineries, and chemical plants would never have been automatically controlled if they had required "teams of experts who could act as resident engineers during the formative stages of the operating departments."

The Reality of Freeway Control

Certainly the design and operation of a freeway control system is a challenge; it is not, however, a mystery. Well-meaning researchers have contributed to the confusion surrounding freeway control by not distinguishing between the requirements for an operational control system and a research facility. Almost 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 thus capacities must be determined, that trip origins and destinations for the freeway and its ramps must be known, and that gap availability and gap acceptance characteristics must be established—all before the control system can be designed. The procedure has been to (a) perform manual system input-output studies requiring as many as 30 people, (b) use time-lapse aerial photography, (c) conduct questionnaire studies to obtain trip information, and (d) employ moving-vehicle study techniques. As it turns out, this part of the design specification for the actual design of a freeway control system is unnecessary.

Basically, the components in the freeway control system consist of one traffic signal per entrance ramp, one merge detector, one gap-speed detector on the outside lane of the freeway, and a regulating controller (Fig. 3). These are the minimum requirements for controlling a single entrance ramp. For controlling the entire freeway, one check-in detector, one vehicle-classification detector, and one queue detector can be added per entrance ramp; one detector per exit ramp; one detector per freeway lane between entrance ramps; and preferably a real-time central computer controller (Fig. 6). This is what is needed and all the studies and measurements explained in the previous paragraph will not change these needs once the output specification proposed in this report, the optimal use of gaps and the prevention of congestion, is chosen. If, however, characteristics are desired, they may be obtained by using the same detection system rather than the manual or aerial techniques.

The purpose of this report is to provide some rationale for the design of a freeway control system based on the multilevel systems approach used in process control. A freeway is just another 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 the control is warranted. In the beginning the traffic engineer responsible for the operation does not have to know anything about network theory or control; he merely installs a detector on each intersection approach, the traffic signals, and the controller, and worries about calibration of the intersection's operation after the system is working. Eventually, as more intersections are signalized, it may evolve into a complex network problem necessitating a central computer controller to coordinate the local controllers on the arterial and neighboring 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 becomes much more tractable.

Many freeway control systems will 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 just as in the case of 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 and 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. Although the multilevel systems approach has its basis in the theory of controlled processes, it is compatible with the practicalities of the stage implementation of freeway traffic control.

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 built up to a point where it exceeded the capacity. After more than a decade of experimentation with surveillance and control, progress in the implementation of freeway control systems on existing facilities is still lagging because of the lack of a rational, unified approach to the design of such systems. It is believed that one answer lies in the application of the multilevel control theory advanced in this report.

REFERENCES

1. Brosilow, C. B., Lasdon, L. S., and Pearson, J. D. Feasible Optimization Methods for Interconnected Systems. Jour. Automatic Control Preprints, p. 79-84, 1965.
2. Lasdon, L. S., and Schoeffler, J. D. Decentralized Plant Control. ISA Trans., Vol. 2, p. 175-183, April 1966.
3. Sanders, J. L. Multilevel Control. Trans. IEEE, Applications and Industry, p. 473-479, Nov. 1964.
4. Lefkowitz, I. Multilevel Approach Applied to Control System Design. Jour. Automatic Control Prymois, p. 100-109, 1965.
5. Rekoff, M. G. Multilevel Systems. 1966, unpublished.
6. Mesarovic, M. D. Self Organizing Control Systems, Descrete Adaptive Processes. Trans. IEEE, p. 265-269, Sept. 1964.
7. Drew, D. R., McCasland, W. R., Pinnell, C., and Wattleworth, J. A. The Development of an Automatic Freeway Merging Control System. Texas Transportation Institute Research Rept. 24-19, 1966.
8. Drew, D. R. Traffic Flow Theory and Control. McGraw-Hill, New York, 1968.
9. Whitson, R. H., Buhr, J. H., Drew, D. R., and McCasland, W. R. Real Time Evaluation of Freeway Quality of Traffic Service. Texas Transportation Institute Research Rept. RF 504-4, 1968.
10. Brewer, K. A., Buhr, J. H., Drew, D. R., and Messer, C. J. Ramp Capacity and Service Volume as Related to Freeway Control. Paper presented at 48th Annual Meeting and included in this Record.
11. Wattleworth, J. A. Peak Period Analysis and Control of a Freeway System. Texas Transportation Institute Research Rept. 24-15, 1965.
12. Moskowitz, Karl. Analysis and Projection of Research on Traffic Surveillance, Communication, and Control. NCHRP Rept. 87, unpublished.

Traffic Control on a Two-Lane, High-Speed, High-Volume Freeway Entrance

HARVEY HAACK, Chicago Area Transportation Study, and
LARRY MADSEN, LEONARD NEWMAN, and WILLIAM E. SCHAEFER, California Division of Highways

This report covers efforts to solve a serious traffic problem resulting when holiday-weekend travelers returned to the Los Angeles area. Intolerable congestion on Memorial Day 1966 led to emergency measures on July 4th. A form of ramp control provided partial relief. On subsequent holidays, more sophisticated plans were attempted with a variety of results. The data collected and the observations noted provide a wealth of information on a unique application of ramp control and an emerging philosophy in freeway and expressway operation.

•ON holiday weekends in 1966, traffic returning to Los Angeles from resort and recreation areas north of the city experienced severe congestion at the merge between the Golden State Freeway (Interstate 5) and Sierra Highway (California 14). The interchange between these two routes is located in Saugus Pass, which slices between the Santa Susana and the San Gabriel Mountains just inside the north city limits.

This report covers traffic control techniques used to relieve the southbound merge between Sierra Highway and the Golden State Freeway prior to interim widening, which was completed in the summer of 1967. Practical benefits were sought through recently developed methods for improving freeway operation by controlling entrance ramps. Basic philosophy was derived from on-going research in Chicago and Houston. The results of that work, however, had to be adapted to the unique physical and operating characteristics of the Golden State Freeway-Sierra Highway interchange.

Plans were developed through the cooperative efforts of the City of Los Angeles Traffic and Police Departments, the Los Angeles County Sheriff's Office, the California Highway Patrol, and the State Division of Highways. The nature of the project provided opportunities both to improve traffic and to evaluate techniques of ramp control. The specific objectives were the following:

1. To reduce congestion and delay to motorists using Sierra Highway and the Golden State Freeway during inbound peaks on holiday weekends,
2. To evaluate ramp closing and ramp metering techniques, and
3. To demonstrate ramp control as a tool for improving freeway operation.

DESCRIPTION OF CONTROL PLANS

This section provides some background on the street and highway system controlled and gives a summary of each of the three plans used.

Background

Figure 1 shows the affected street and highway system. Control operations took place on the southbound connector from Sierra Highway to the Golden State Freeway.

Paper sponsored by Committee on Freeway Operations and presented at the 48th Annual Meeting.

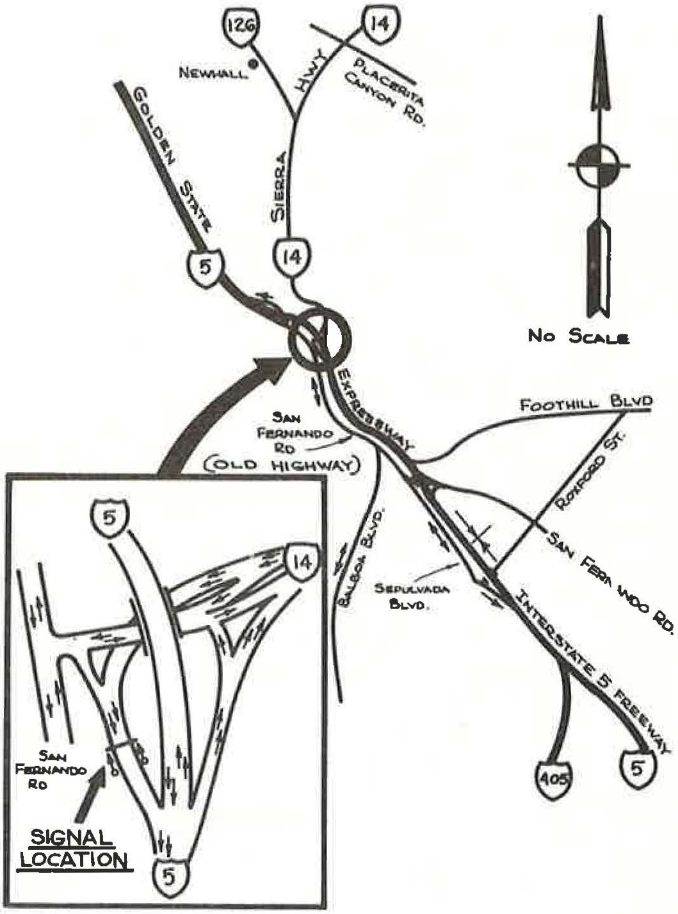


Figure 1. Area map.

The Golden State Freeway, an Interstate route, had been enlarged to eight lanes both upstream and downstream of the existing interchange. Complexity of the future interchange, however, had delayed construction here and only two lanes served southbound traffic in 1966. A frontage road, consisting of the old highway (San Fernando Road) and Sepulveda Boulevard to the south, provided a bypass around this two-lane section.

Figure 2 shows an aerial view of the interchange. A three-level structure carries the Golden State Freeway over the top, Sierra Highway at grade, and the southbound connector underneath. The critical merge occurs where the two-lane connector from southbound Sierra Highway enters the Golden State Freeway's two southbound lanes. The most severe congestion occurred on three-day holiday weekends, such as Memorial Day, July 4th, and Labor Day in 1966. On occasion the Golden State Freeway backed up 10 miles and Sierra Highway as much as 5 miles. No serious congestion developed during normal day-to-day operation, however.

Beginning July 4, 1966, traffic control was used to relieve excessively heavy demand. On this day the connector was closed for several hours and Sierra Highway traffic routed over the old highway. On Labor Day, 1966, Sierra Highway traffic was metered onto the Golden State Freeway by means of manual control. No problems were anticipated for Thanksgiving weekend, and no plans were made for control. The heavy congestion that did develop, however, led to installation of temporary traffic signals on the connector road. Arrangements were made to operate these signals on Christmas Monday in 1966, on New Year's Monday, the first Sunday of trout season, Memorial Day, and July 4th in 1967.

Figure 2. The Golden State Freeway and Sierra Highway interchange.

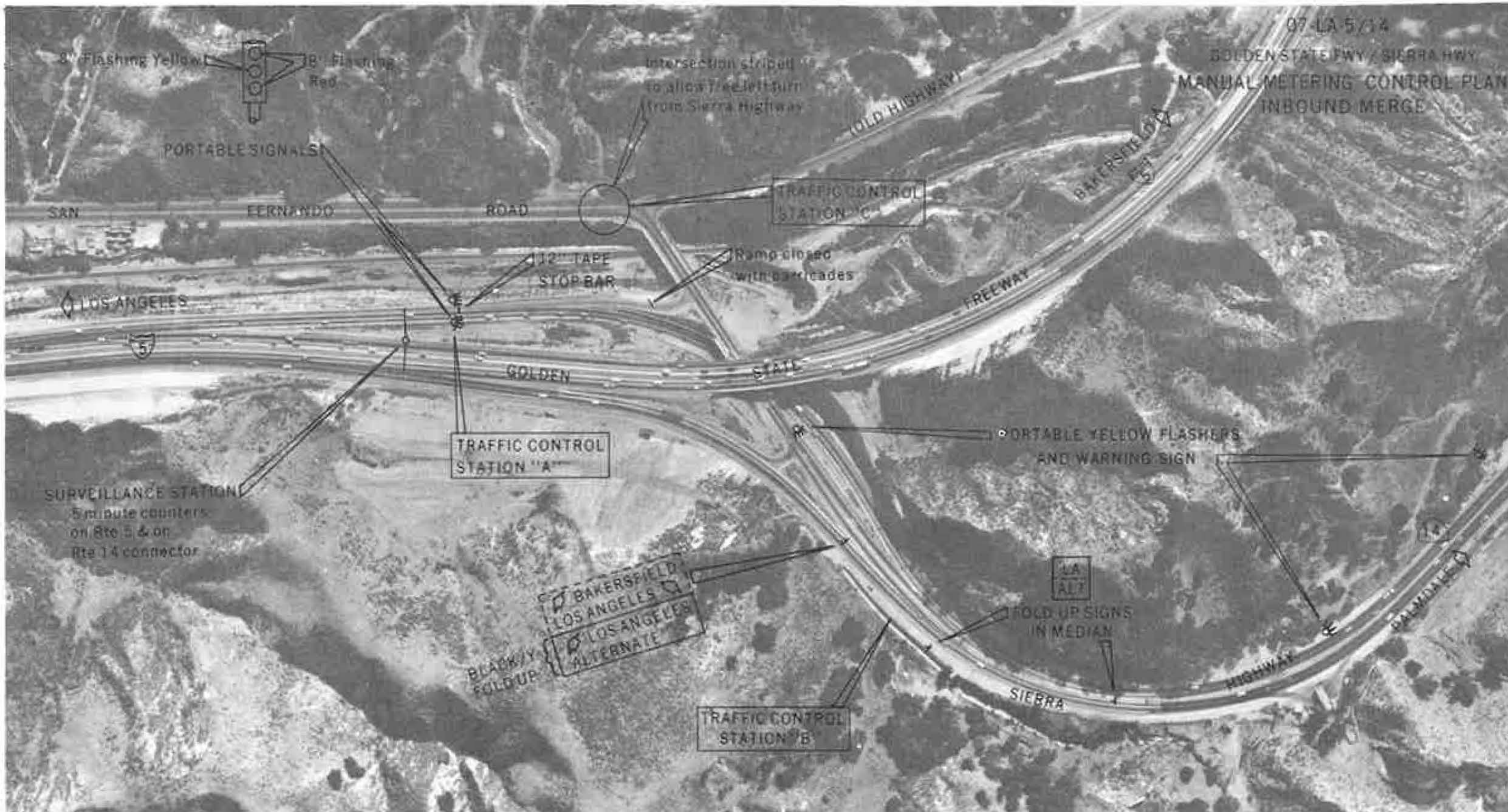


Figure 3. Plan for manually metering Sierra Highway traffic.

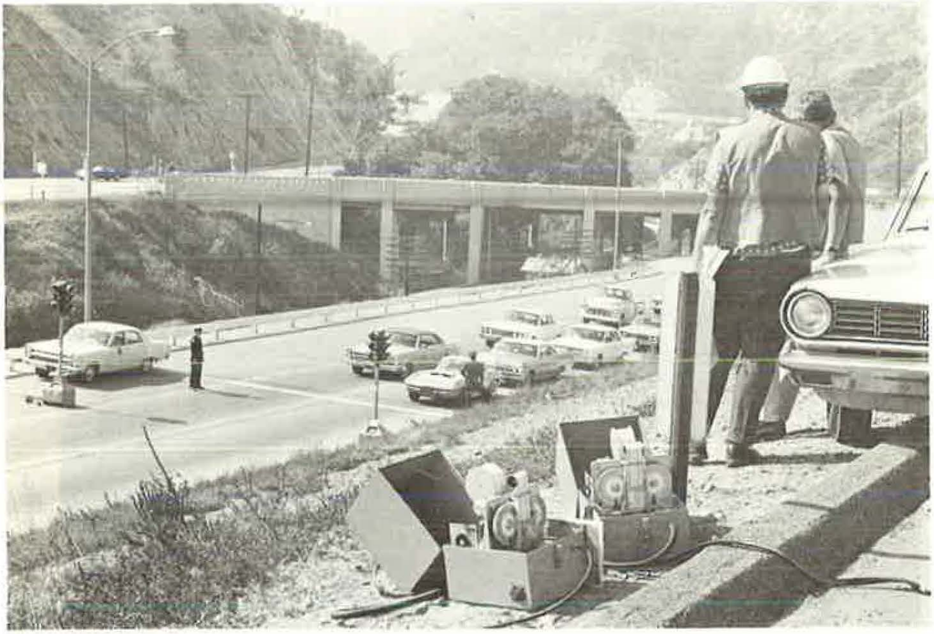


Figure 4. Manual metering control operation.

An interim widening project was completed before Labor Day in 1967. This project provided three lanes southbound instead of the original two. The increased capacity proved adequate to handle peak rates of 4600 vph on Labor Day, 1967. Traffic control was no longer required.

Control Plans

The first control plan was more or less an emergency measure (undertaken without adequate advance traffic data) to avoid the "intolerable" situation that had occurred on Memorial Day weekend in 1966. Other plans evolved as more was learned about traffic operation over the system. The basic idea was to relieve the overloaded Golden State Freeway by forcing all of Sierra Highway traffic onto the old highway. Portable barricades were used to close off the connector. On this attempt, however, the old highway could not handle all diverted traffic. This led to the manual metering plan.

Figures 3 and 4 show the manual metering plan. The main objective here was to balance flow between the freeway and the old highway. Police officers stationed on the connector regulated traffic entering the Golden State Freeway. Officers released vehicles in platoons from two lanes every 30 seconds. Initial platoon size was determined from observations made on other occasions. During operation, however, an observer located on the shoulder of the freeway watched traffic flow in the merge and downstream. He instructed officers on when to adjust platoon sizes. The method proved quite sensitive to fluctuations in upstream demand, which turned out to be unexpectedly severe. Originally, 5-minute machine counts were to provide surveillance data. The 5-minute lag proved too great, however.

Minor control was provided as needed at two other critical locations. In all, six police officers were required to operate the plan. Because police officers are in particularly high demand on holiday weekends, to reduce the number required, a plan calling for traffic signal control was developed.

Operation of this plan was essentially the same as with manual control. In this case, however, traffic signals replaced police officers on the connector. The object was to develop a system that could be put into effect by a single officer on routine

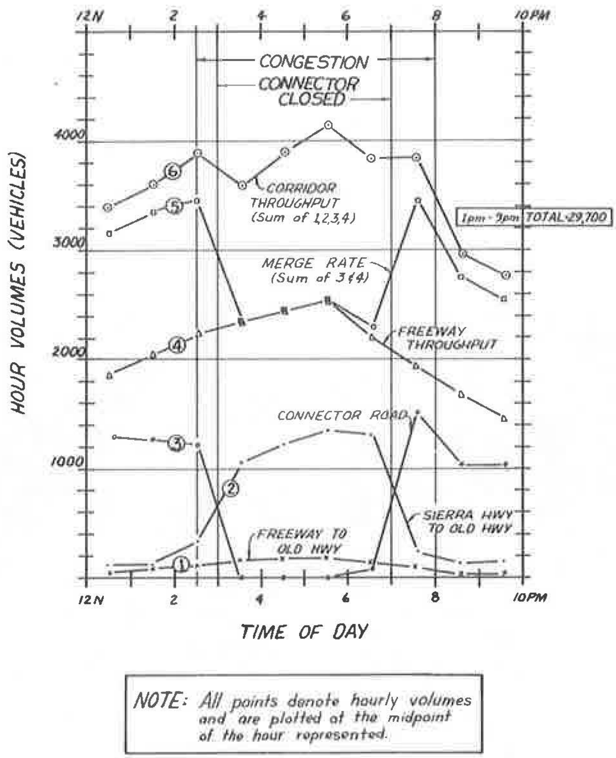


Figure 5. Hourly traffic counts—July 4, 1966.

patrol. Also, since this was the first time entrance control had been used in Los Angeles, experience with signalized control was desired.

OPERATION AND RESULTS

Data were collected on eight weekends. For the sake of brevity, however, only four weekends will be covered here. In general, discussion is organized according to the type of control used. Comparisons of delays are made, but it is emphasized that these are extremely rough estimates. The fact that traffic demands varied so much during the different modes of operation made direct comparisons difficult.

Closed Connector

On July 4, 1966, as a result of "intolerable" congestion on Memorial Day, emergency traffic control measures were taken. Control consisted of closing the connector from Sierra Highway to the southbound Golden State Freeway. The connector was closed for roughly 4 hours from about 3 to 7 p. m. During this time Sierra Highway traffic was diverted to the old highway.

Figure 5 illustrates traffic volumes at various points on the system. The sudden drop in connector volumes at 3 p. m. (curve 3 in Fig. 5) shows when the connector was closed. Reopening is indicated by the jump in connector volumes at 7 p. m. Curve 2 shows the greatly increased flow to the old highway during the period when the connector was closed. The major restriction to this flow was the left turn from Sierra Highway to southbound on the old highway. On this particular occasion, maximum rates of less than 1400 vph were accommodated. This was several hundred vehicles per hour less than demand. Since the old highway provides the only alternate route, excess demand was forced to store upstream on Sierra Highway. Figure 6a shows the density in the queue that formed. At times this queue extended back 5 miles.

Curve 5 in Figure 5 represents merge rates. Ordinarily this would be the sum of freeway throughput (curve 4) and traffic on the connector (curve 3). While the connector was closed, however, merge volumes and freeway throughput were the same. That connector traffic made up a large portion of the merge volume is demonstrated by the large drop in merge rates at 3 p. m. when the connector was closed.

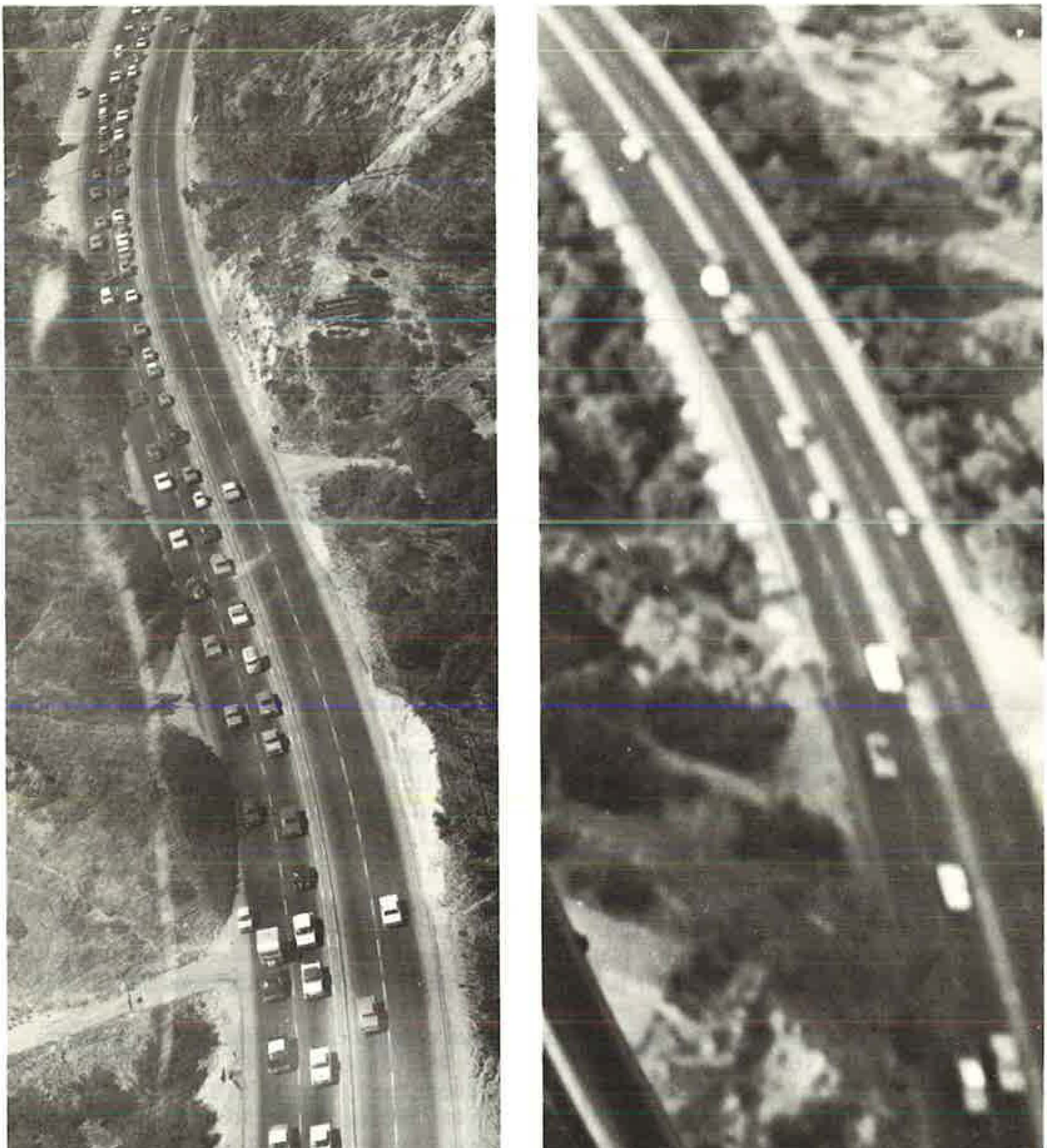


Figure 6. Improved flow: (left) Queue of vehicles backed up on Sierra Highway while the connector was closed (average speed 6 mph); (right) flow during manual control (average speed 40 mph).

The drop in merge rates at 3 p. m. and the subsequent jump at 7 p. m. point out the under-utilization of the freeway during control. Immediately after the connector was opened, merge rates, because of a backlog of demand on Sierra Highway, jumped to 3500 vph. In other words, freeway capacity was at least 1000 vph greater than the average rate of 2400 vph recorded on the freeway while the connector was closed.

During closure, the freeway was able to handle its full demand. In fact, the freeway experienced no congestion at all. Total corridor throughput (curve 6), however, reached a maximum of only 4200 vph. On Sierra Highway, maximum rates totaled about 1400 vph (curve 2 plus curve 3). This is less than the volume that normally would be handled without control.

For the period from 1 to 9 p. m., the total volume was 18,700 vehicles on the freeway approach and 11,000 vehicles on the Sierra Highway approach. The sum of the two (29,700 vehicles) represents full corridor demand because all congestion occurred within this 8-hour period and flow was good at both 1 and 9 p. m. (This was also true for the other days described.)

Total delay on this July 4th was estimated to be 150,000 vehicle-minutes (based on floating car runs, lengths of queue, and duration of congestion). Approximately 7,500 vehicles on Sierra Highway were delayed. This averages to about 20 minutes per vehicle; maximum delays reached 35 to 40 minutes.

Based on subsequent observations on days with and without control, it appears that less total delay would have occurred if the connector had not been closed. Estimates indicate that there would have been about 25,000 vehicle-minutes of delay, and that most of this would have accrued to freeway traffic (i.e., on the high-speed approach). Sierra Highway would have suffered little delay, and congestion on the freeway would have lasted from about 3 to 7 p. m.

In spite of increased overall delay, the fact that the high-speed approach was kept free-flowing is important. It is highly probable that because of this accidents were prevented.

Manual Metering Control

On Labor Day, 1966, control was accomplished by manually metering connector traffic from Sierra Highway to the southbound freeway. Congestion started to develop shortly after 2:30 p. m. Control was begun immediately and continued until demand subsided around 6:30 p. m.; 30-second metering rates were varied constantly between 1200 and 2200 vph. Even so, average flow on the connector over the controlled period remained relatively constant between 1400 and 1500 vph (Fig. 7, curve 3).

The queue, planned for on the connector, developed immediately after control started. The tail of this queue fluctuated between 800 and 1200 ft upstream of the control point. Figure 6b shows the free flow of traffic on upstream Sierra Highway under metering operation. Compare this with the dense queue in Figure 6a—a picture of the same location taken on July 4th.

Signs advising that the old highway provided an alternate to the freeway were displayed. During surges of high demand on Sierra Highway, many drivers diverted voluntarily. This was due, apparently, to the alternate route signs and the lengthening queue on the connector. An officer, provided to direct traffic to the old high-

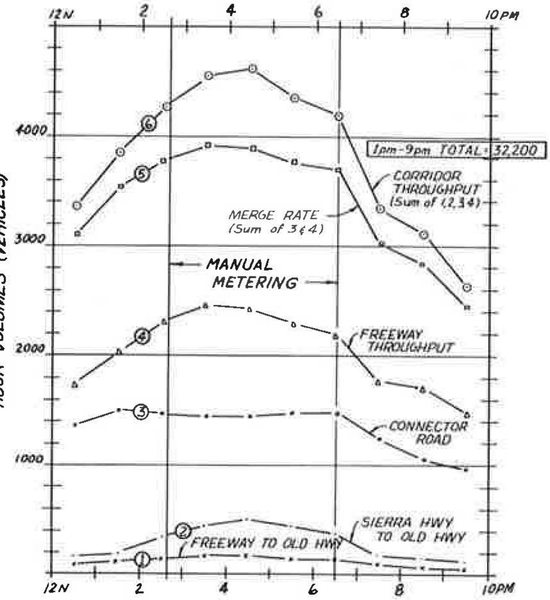


Figure 7. Hourly traffic counts—Labor Day, 1966.

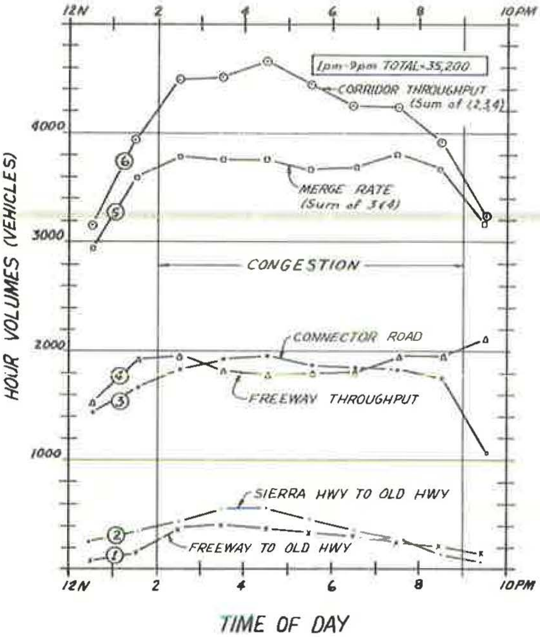


Figure 8. Hourly traffic counts—Thanksgiving Sunday, 1966.

way when the queue extended upstream of the diversion point, was seldom needed. Diversion rates (including those drivers that wanted to use the alternate route) ranged between 400 and 600 vph during the control period (curve 2 in Fig. 7). Thus, Sierra Highway approach volumes averaging around 2000 vph were handled (curve 2 added to curve 3).

The rate of freeway throughput reached a maximum of 2500 vph during control (curve 4). This demand was handled with only minimum reductions in freeway speeds. Fluctuations in demand (5-minute rates ranged from 1600 to 3100 vph) did result in occasional shock waves, but overall freeway delay was negligible.

Total corridor throughput reached 4600 vph (curve 6). This was accomplished with no queues on the freeway and only short queues on the connector. Even so, additional capacity was available on the alternate. Had connector queues extended upstream on Sierra Highway, more vehicles could have been diverted to the old highway and corridor throughput would have been even greater.

Volume for the 8-hr period from 1 to 9 p. m. was 18,400 vehicles on the freeway and 13,800 vehicles on Sierra Highway—a total of 32,200 vehicles. This was 2600 vehicles, or about 9 percent, more than on July 4th. Yet, even with this greater volume, there were insignificant delays to freeway traffic and only about 9000 vehicle-minutes of total delay to Sierra Highway. Maximum delay to individual vehicles on Sierra Highway was less than 2 minutes.

Had there been no control, estimated delays would have been about 170,000 vehicle-minutes. Most of this delay would have been suffered by freeway traffic. Delays to Sierra Highway traffic would have been about the same as with control.

No Control, Heavy Demand

No plans were made to control traffic on Thanksgiving weekend other than to expose signs advising of the alternate route provided by the old highway. Observers had claimed that congestion occurred only during summer months. This proved incorrect. Actually, congestion was greater on Thanksgiving weekend than on any other weekend observed.

Freeway throughput reached an early peak of nearly 2000 vph between 2 and 3 p. m. (Fig. 8, curve 4). Congestion on the freeway began about 2 p. m. A queue extending back 2 miles was observed by 3 p. m., and by 6 p. m. this queue had grown to 10 miles. Evidently, freeway traffic was overpowered by the pressure of increasing demand from Sierra Highway. Merge geometrics give equal advantage to both approaches. In other words, when both approaches were full, each provided roughly half the merge volume. This is borne out in Figure 8 (curves 3 and 4). As a result, freeway throughput (curve 4) dropped slightly during peak Sierra Highway demand. As Sierra Highway demand fell, after 8 p. m., freeway throughput increased, dissipating the 10-mile queue that had formed. It was after 9 p. m., however, before all congestion cleared.

At 3 p. m. only slight congestion existed on the connector (curve 3). The maximum queue observed during the entire afternoon extended back less than ¼ mile. It appeared that many drivers voluntarily diverted to the old highway when it was obvious that there was congestion on the freeway. Voluntary diversion from Sierra Highway reached a high of 600 vph (curve 2).

Voluntary diversion from the freeway also took place. Usually about 200 vph leave the freeway on a normal holiday weekend. However, volumes of 400 vph were recorded on Thanksgiving weekend (curve 1). Unfortunately, not enough freeway users took the old highway early in the peak, and even those that exited to the old highway were trapped in the 10-mile queue. The maximum volume using the old highway reached roughly 1000 vph—600 vph less than volumes recorded on previous occasions when diversion was encouraged by police officers. In other words, the old highway was under-utilized.

Merge rates remained steady at 3700 to 3800 vph throughout the entire 7-hour period of congestion (curve 5). Total corridor throughput reached a peak of 4700 vph from 4 to 5 p. m. and fell gradually as fewer vehicles diverted to the old highway (curve 6).

Volume for the period from 1 to 9 p. m. was 17,400 vehicles on the freeway and 17,800 vehicles on the Sierra Highway—a total of 35,200 vehicles. Estimated congestion totaled some 270,000 vehicle-minutes with individual delays of 30 minutes. Most of this delay was suffered by freeway traffic.

The total corridor volume from 1 to 9 p. m. on Thanksgiving Sunday was 3000 vehicles higher than the comparable volume on Labor Day when control virtually eliminated delay. It appears, however, that had control (manual or signal) been imposed, most of the delay on Thanksgiving Sunday could also have been eliminated. Although corridor ouput was high without control, the alternate route still had additional capacity. Control, of course, would have encouraged greater use of this capacity.

The fact that total corridor throughput was as great as it was can be attributed to the high rate of voluntary diversion to the old highway. This, in turn, can be attributed to (a) driver experience on previous holiday weekends, (b) publicity of previously congested conditions and of the available alternate, (c) signing of the alternate, and (d) perhaps most importantly, the congestion itself.

Traffic Signal Metering Control

Christmas Monday was the first attempt toward automating metering control (Fig. 9). A multi-cycle, variable phase traffic signal had been installed. Cycle selection and

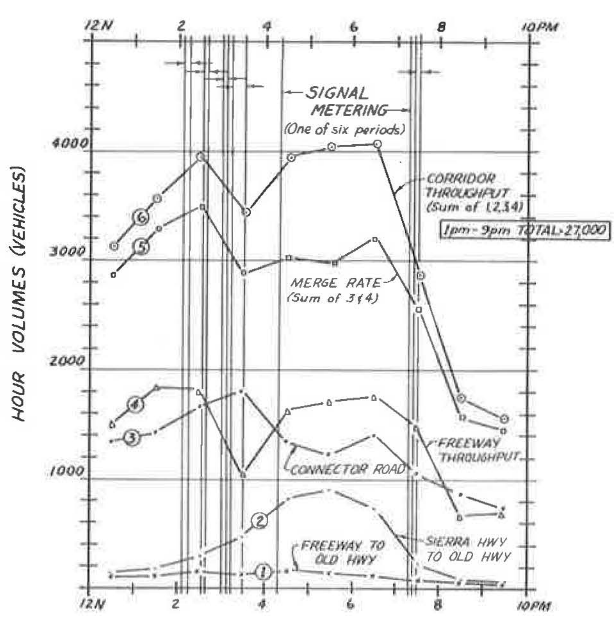


Figure 9. Hourly traffic counts—Christmas Monday, 1966.

phase adjustments were still made manually, however. Surveillance was still accomplished by visual observation of downstream flow supplemented by short manual counts.

Plans were made to operate the signal on four occasions other than Christmas Monday; these were New Year's Monday, the first Sunday of trout season, Memorial Day, and July 4th, 1967. Demand sufficient to require control only occurred on Christmas Monday and the first Sunday of trout season. Only Christmas Monday is discussed here.

Traffic conditions on Christmas Monday were not normal by any means—even for a three-day holiday weekend. Two minor, but disruptive, accidents on the upstream freeway played havoc with freeway demand.

Initiating control involved some hazard because approach speeds on the connector were relatively high. Because of the hazard, start of control was delayed until there was reasonable assurance that continuous control would be required during the remainder of the peak. The resultant criterion was formation of a stable queue (rather than just shock waves) on the freeway at the merge. Had the accident on the freeway been anticipated, signal control could have been operated to prevent queues from forming on the freeway. In fact, to lessen the chance of recurrence of such an incident on subsequent occasions, impending shock waves rather than standing queues became the criterion. This, of course, sacrificed some capacity.

Wide fluctuations in freeway demand on Christmas Monday—almost instantaneous drops and jumps ranging from 2400 vph to less than 1200 vph—led to severe, rapid variation in metering rates. For short periods (2 to 3 minutes) rates on the connector were restricted to as low as 600 vph. This was to allow dissipation of the queues that had formed on the freeway because of accidents upstream.

Five different times during lulls in freeway flow connector traffic was released from direct control. At first this was accomplished by switching the signal to steady green. This, however, seemed to encourage connector traffic to higher speeds, around 50 mph. At these speeds, regaining control was hazardous. Further operation during lulls was on flashing yellow.

Flexibility of signal control was taxed by wide fluctuations in demand. Controller limitations required minimum phase lengths (12 seconds for green-yellow and 9 seconds for red) that delimited the range of settings provided by each of the four cycle lengths. These limits were more restrictive for shorter cycles. For instance, the range of metering rates achieved on the 30-second cycle was 1100 to 2200 vph, while the range on the 60-second cycle was 600 to 2200 vph.

Although the longer cycles permitted greater flexibility, compliance (stopping for the red signal) was not as good. For example, compliance was good on the 30- and 40-second cycles throughout their entire range of settings. On the 50- and 60-second cycles, however, the longer periods of green required to maintain metering rates equivalent to the maximum of shorter cycles allowed vehicles to gain higher speeds. A greater reluctance to comply was observed. In fact, had the longer green phases been sustained, it appeared that loss of metering control might have resulted. Violations were also frequent when a queue did not exist on the connector. It seemed that, at free-flow speeds, drivers simply did not have time to comprehend the strange situation of having to stop for a signal on a freeway-to-freeway connector.

The range of metering rates also affected the time required to negotiate a given length of the connector queue. On Labor Day the maximum time to travel the length of the connector was about 2 minutes. On Christmas Monday, however, a much denser and slower moving queue developed. This queue required a maximum of 5 minutes to travel the length of the connector. During heaviest demand the queue extended upstream of the exit to the alternate route.

Also of interest were the platoons released from the signal. Their size varied in proportion to the length of the green-yellow phase. Between the minimum and maximum phase lengths of 12 and 51 seconds, the platoon size ranged from 9 to 36 vehicles. Manual metering on Labor Day resulted in fairly well-spaced platoons. Those released from the signal, however, were much more closely spaced. In each case, platoons with more than 10 vehicles tended to disrupt freeway traffic. However, groups of up to 20 vehicles were able to merge with relative ease. Platoons of greater than 20 vehicles were impractical because of problems in merging, compliance, and control.

All in all, Christmas Monday provided a true test of the flexibility of traffic signal control. Volume for the 8-hour period from 1 to 9 p. m. was 13,100 vehicles on the freeway and 13,900 vehicles on Sierra Highway—a total of 27,000 vehicles. This total was less than the total on July 4th. Even so, Sierra Highway demands were greater on this occasion.

Good estimates of total delay are not available for Christmas Monday. Generally, though, the freeway was kept free-flowing. On Sierra Highway, however, metering rates were kept quite low at times (for reasons noted) and diversion to the old highway, while high, was not as efficient as it could have been. For roughly an hour this resulted in significant delays (up to 10 minutes) to Sierra Highway traffic.

Another factor contributing to Christmas Monday problems was an unusually low merge capacity. Actually, merge capacity is limited by what the downstream section can absorb. On this day the downstream section was not working well at all. In spite of the various problems, it appeared that control was successful and that the relatively large diversion to the old highway allowed better overall operation and reduced travel time. Without this diversion, the large queues on the freeway (caused by accidents) could not have been dissipated and extremely large delays would have occurred.

Increased Capacity Conditions

The interim widening project was completed just prior to the 1967 Labor Day weekend. Operation was observed, and the additional capacity eliminated all problems. Continuous counts were not made, but short (3-minute) counts made during the peak indicated the merge rate varied between 4000 and 4600 vph. Capacity of the merge and downstream section is now about 5400 vph. Use of the Alternate route varied between 300 and 400 vph. Total corridor demands were about 4600 to 4700 vph, which are the same as those occurring during the 1966 Labor Day weekend.

On Thanksgiving weekend (1967) demands were extremely high, with corridor outputs (based on short counts) reaching 5100 to 5300 vph, including approximately 4800 vph on the freeway. Traffic demands on the freeway approach were so high that there was congestion on the two-lane section upstream of the exit to the old highway. Even so, the corridor was able to handle, reasonably well, all traffic that could reach the section. Control would not have improved operation.

SUMMARY AND CONCLUSIONS

Higher than normal traffic volumes occurred at the merge between Sierra Highway and the southbound Golden State Freeway on eight occasions between July 4, 1966, and July 4, 1967. During that time, two basic approaches were used to regulate merging traffic. The first approach was to completely close off the entrance to the freeway; the second approach concentrated on regulating, or metering, entrance traffic.

Steps toward automated surveillance and metering control have not been taken because an interim widening project now provides more than enough capacity at the merge. Consequently, conclusions drawn here are based on metering rates determined by direct visual surveillance and accomplished by manual regulation of traffic or by traffic signal control, whichever applied.

In comparing the various control plans it was found that closing the connector resulted in under-use of the freeway at the expense of Sierra Highway, which backed up 5 miles. It took as much as 40 minutes to traverse this queue, and total delay was about 150,000 vehicle-minutes.

Getting full use of freeway capacity required that a certain amount of traffic be allowed onto the freeway via the connector. Direct control by manual metering resulted in an average merge rate of from 3600 to 3900 vph during the metering period; 5-minute rates on the uncontrolled leg ranged from 1600 to 3100 vph.

The most unique feature of the plan was the magnitude of connector rates allowed onto the freeway; 30-second rates were varied between 1200 and 2200 vph. Achieving these rates required utilization of both connector lanes. It was also necessary to release vehicles in platoons, compared with the one-at-a-time operation usually used. Metered platoons contained 10 to 18 vehicles, depending on available freeway capacity.

Manual metering control resulted in negligible delay to freeway traffic. Individual Sierra Highway delays were less than 2 minutes, vs the 40-minute delays on July 4th. Total delay was about 9000 vehicle-minutes. It only required a diversion of 400 to 600 vph to accomplish this.

A standard traffic signal provided metering control that was less flexible but comparable to manual metering control; 30-, 40-, 50-, and 60-second cycles with 12-second minimum green-yellow and 9-second minimum red phases were used. Cycle-to-cycle rates for signal metering were also varied from 1200 to 2200 vph according to available capacity.

Generally, drivers obeyed signals on 30- and 40-second cycles. It appeared, however, that higher speeds generated from relatively long green phases associated with some 50- and 60-second cycles made compliance more difficult. This difficulty was evidenced in that a number of drivers simply ignored the signal.

Analysis and evaluation of data collected on the eight occasions of heavy congestion provide results from which the following conclusions are drawn.

1. Marginal nature of ramp control: (a) Under normal conditions, diverting just a few vehicles (often as few as 100 to 200 vph) accomplishes great improvements in freeway operation; (b) heavy surges in traffic—following removal of accident vehicles from the traveled way, for example—may overwhelm metering operation.

2. Ramp closing: Closing a high-volume entrance may result in a partially empty freeway. This method should only be considered when the ramp volume is less than the excess demand downstream from the entrance, and/or the alternate routes have adequate capacity to absorb diverted traffic.

3. Variable control: (a) Variable control allows adjustment of ramp volumes to fill excess freeway capacity; (b) variable control can be sufficiently sensitive to smooth normal fluctuations in freeway flow; (c) variable control may be accomplished either manually or by using traffic signals.

4. Manual control: (a) Manual control allows precise adjustment of flow rates; (b) manual control spaces vehicles more evenly in merging platoons, reducing merge friction.

5. Signal control: (a) Cycles of 30 to 40 seconds result in good signal compliance, tolerable platoon size (9 to 20 vehicles), and less flexible metering rates (1100 to 1800 vph) than longer cycles; (b) cycles of 50 to 60 seconds result in signals being ignored (because of the higher speeds generated), intolerable platoon size (up to 36 vehicles), and more flexible metering rates (600 to 2200 vph) than shorter cycles; (c) traffic approaching the control signal should be slowed to less than 40 mph (by advance warning lights for example) prior to and during control; (d) a variable cycle having a fixed green-yellow phase of 12 seconds, and a variable red phase of 48 seconds maximum and 8 seconds minimum is recommended—this would allow metering rates ranging from 600 to 1800 vph while retaining the optimum 10-vehicle platoon for two-lane entrances.

6. Surveillance: Speeds and densities in the merge and immediately downstream provide the best indication of when to adjust metering rates.

7. Bypass routes: (a) If bypass capacity is not adequate to handle diverted traffic, bigger problems may be created than the one solved; (b) a number of drivers bypass congestion of their own accord when they know of alternate routes; (c) ideally, drivers should be persuaded to take an alternate route before congestion develops (however, this is against human nature).

8. Merge gaps: Sufficient gaps exist in freeway flow at 45 mph to allow platoons of 10 or fewer vehicles to merge with little friction. Merge design must be good, however.

9. Safety: (a) Metering improves safety by either stabilizing or eliminating queues; (b) metering high-speed, high-volume traffic is potentially hazardous and requires thoroughly indoctrinated personnel to supervise and to perform metering activities.

10. Costs and benefits: (a) Controlling ramps is relatively inexpensive, especially when compared with the cost of increasing capacity by widening; (b) typically, on this project, time savings exceeded 150,000 vehicle-minutes per day of control.

In general, merge control when properly applied provides an inexpensive and effective means of increasing the amount and quality of traffic flow. Furthermore, merge control reduces overall delay by taking advantage of all available corridor capacity. Much research remains to be done toward application, but as knowledge continues to grow and more advanced techniques are developed, it would seem that merge control will eventually lead to fuller and more efficient utilization of the existing street and freeway system.

ACKNOWLEDGMENTS

Our appreciation is extended to the Los Angeles City Traffic Department and the Assistant City Traffic Engineer, A. L. Hutchison; the Traffic Division of the Los Angeles City Police Department, formerly under Deputy Chief H. W. Sullivan, and now under Deputy Chief Roger Murdock; the Los Angeles County Sheriff's Office; and the California Highway Patrol. Special thanks are given to Jim Hardy and Lieutenant Jim Lane, coordinators, respectively, for the City Traffic Department and the City Police. Most especially, we acknowledge the contribution of Jack Eckhardt, former head of our own Freeway Operation Department.

This report is condensed to meet publication requirements. A more detailed report may be obtained from the Freeway Operation Department, California Division of Highways, 120 S. Spring Street, Los Angeles, California 90012.

Ramp Capacity and Service Volume As Related to Freeway Control

KENNETH A. BREWER, JOHANN H. BUHR, DONALD R. DREW, and CARROLL J. MESSER, Texas Transportation Institute, Texas A&M University

The problem of establishing a policy for freeway control can be structured into an ordered set of control problems within a hierarchy of levels of control. At each level associated with the hierarchy of the problem structure, an interrelation exists among freeway-capacity, freeway-demand, ramp-capacity, ramp-demand, outside-freeway-lane-flow, and ramp-merging characteristics and merging-control-system operation. The extent to which an optimum control policy can be established at each level of control is directly related to the extent to which the pertinent traffic and control system variables can be interrelated.

In this report, ramp capacity and service volume are defined in terms of entrance-ramp geometry and gap-acceptance and merging-controller-operational characteristics. Subsequently, an optimal control policy is established for an isolated entrance ramp and the integrated total freeway control policy modeled.

•DURING the last decade freeway control has evolved from a research experiment to an operational reality. Freeway control systems have been installed and are operational in Chicago, Detroit, and Houston (6, 7, 8). Others are being initiated in Seattle and Dallas. The various forms of control have included freeway main-lane control (9), ramp metering (7), and ramp merging control (8).

In order to optimally control a freeway system, the capacity and flow characteristics of the freeway and the ramps must be known and understood. The capacity of freeway sections has been studied (1), and its effect on freeway control philosophy is well documented (10). The capacity of an uncontrolled ramp is not as well defined nor as precisely quantified, even though much work (1, 2, 3, 4, 5) has been done to relate the outside freeway lane flow to ramp service volumes. Optimizing freeway control system operation requires that the control of an individual ramp be understood in the context of overall freeway system control. Therefore, the capacity and service volume relationships for controlled merging conditions must be determined in terms of variables that functionally define freeway control. These variables include speed and flow on the outside freeway lane, total freeway flow, freeway capacity, ramp capacity, ramp demand, efficiency of the merge, and extent of queuing on the ramp. Furthermore, these variables can be grouped according to those that can be controlled (e.g., ramp flow), those evaluated within a control system (e.g., ramp geometry and vehicular characteristics), and those used to effect the desired control action (e.g., controller settings).

This paper is directed toward determining the relationships among these variable groups, their influence on controlled ramp capacity, and their effect on freeway operation.

Paper sponsored by Committee on Highway Capacity and presented at the 48th Annual Meeting.

CONTROLLED RAMP SERVICE VOLUME MODEL

General Merging Control System Configuration and Operation

An entrance ramp with a "stopped vehicle gap acceptance" mode of control searches for a gap, t, in the outside freeway lane such that $t \geq T_S$, where T_S is the service gap setting on the controller. A selected gap is projected in time to a decision point so that the vehicle waiting at the ramp signal may be released into the selected gap (8, 13).

When a queue of vehicles is always waiting at the ramp signal, the merging control system output is primarily a function of the efficiency of the system components and freeway traffic conditions. In the event that a ramp driver rejects the selected gap, he will suffer some delay while searching for a gap in the traffic stream that is acceptable to him. If during such time the driver stops in the merging area, a presence detector will hold the merging controller in the inoperative state until the merging area is clear. Any additional vehicles that are released at the ramp signal and that are behind the driver who stopped in the merging area (due to the travel time from the ramp signal to the merge area) will also miss their assigned gaps. When a vehicle stops in the merging area, the controller will be held in the inoperative state until all vehicles between the signal and the merge area have departed. Then, the normal gap selection-gap projection process will continue.

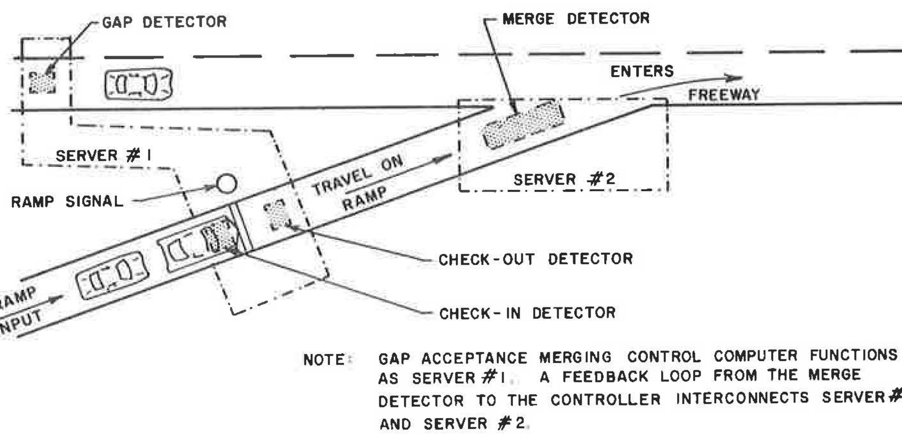
As each vehicle arrives at the head of the queue, it will stop over a "check-in" detector in front of the signal. From the time the ramp signal displays a green indication until a "check-out" detector just beyond the signal is actuated, the controller will dwell in the inoperative state. This "server dwell time" also corresponds to the time required for the next vehicle in the queue to move up over the check-in detector.

A controlled ramp as just described can be considered similar to a two-server-in-series queuing facility. Figure 1 shows a typical controlled ramp facility with each server function physically identified. In order to maximize the output per unit time of such a ramp, the average service time in the total service facility must be minimized.

"Controlled ramp capacity" will be a function of several variables, including (a) availability of gaps in outside freeway lane, (b) size of service gap setting, (c) geometry of ramp, and (d) driver-vehicle characteristics. Control of the ramp is basically a problem of detecting and projecting freeway gaps for a ramp vehicle to utilize. Functional relation of this problem to ramp capacity will be in terms of these variables.

Development of Control System Total Service Time

If time headways in the outside freeway lane are distributed according to an Erlang distribution with a probability density function, f(t), then



NOTE: GAP ACCEPTANCE MERGING CONTROL COMPUTER FUNCTIONS AS SERVER #1. A FEEDBACK LOOP FROM THE MERGE DETECTOR TO THE CONTROLLER INTERCONNECTS SERVER #1 AND SERVER #2.

Figure 1. Ramp control system functional components.

$$f(t) = \frac{\left(aq_f\right)^a t^{a-1} e^{-aq_f t}}{(a-1)!} \tag{1}$$

where t is time headway, q_f is average outside lane flow rate, and a is an Erlang constant.

The expected delay to a ramp vehicle searching for a gap, t, in a stream of gaps has been developed mathematically when the driver has a constant threshold gap (17). For a controlled merge, an equivalent search is conducted by the control computer. Therefore, the average waiting time, T_c, for a vehicle at the ramp signal while the merging controller is searching for a gap (16) is expressed by

$$T_c = \frac{e^{aq_f T_s} - \sum_{i=0}^{a} \frac{\left(aq_f T_s\right)^i}{i!}}{q_f \sum_{i=0}^{a-1} \frac{\left(aq_f T_s\right)^i}{i!}} \tag{2}$$

where T_s is the gap set on the merging controller. Then, the average service time in server No. 1 (Fig. 1), S_1, is given by

$$S_1 = R + T_c \tag{3}$$

where R is the average server dwell time between the release of a ramp vehicle at the signal and the start of the next gap search.

In order to consider the service time in server No. 2, the ramp driver's probability of accepting the assigned gap must be evaluated. Research into the gap-acceptance phenomenon has developed the concept of a gap-acceptance probability, P(t), such that a ramp driver has a probability P(t) of accepting a gap, t, in the merge area (17, 18, 19, 20). The gap-acceptance function is

$$P(t) = \int_0^t g(t)\,dt \tag{4}$$

where g(t) defines the functional change in the probability of accepting a gap as gap sizes change.

Two distributions that have been found to be appropriate for g(t) are the long-normal distribution (17) and the Erlang distribution (21). Since an Erlang distribution is a reasonable description of gap acceptance (Fig. 2), the gap-acceptance function used in this development will be a cumulative Erlang form. Then P(t) becomes

$$P(t) = 1 - e^{-k\mu t} \sum_{i=0}^{k-1} \frac{(k\mu t)^i}{i!} \tag{5}$$

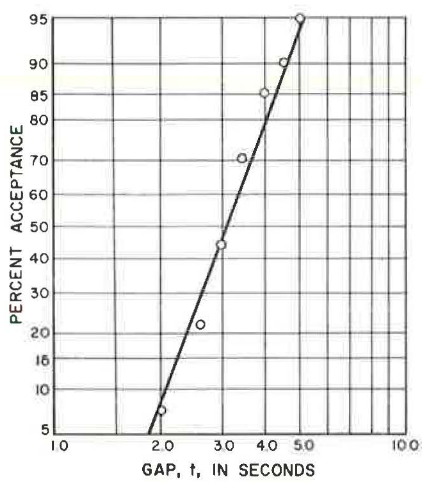


Figure 2. Dumble gap acceptance.

where k is an Erlang constant and μ is the inverse of the mean accepted gap for moving merges. It follows then that the probability of accepting a gap, t, given that $t \geq T_S$, which will be denoted by P(a), is expressed by

$$P(a) = \frac{\int_{T_s}^{\infty} P(t)\, f(t)\, dt}{\int_{T_s}^{\infty} f(t)\, dt} \tag{6}$$

Substituting Eqs. 1 and 5 into Eq. 6 gives

$$P(a) = 1 - \frac{\int_{T_s}^{\infty} \left[e^{-k\mu t} \sum_{i=0}^{k-1} \frac{(k\mu t)^i}{i!} \right] \frac{\left(aq_f\right)^a}{(a-1)!} t^{a-1} e^{-aq_f t}\, dt}{e^{-aq_f T_s} \sum_{i=0}^{a-1} \frac{\left(aq_f T_s\right)^i}{i!}} \tag{7}$$

The integral in the numerator of the fractional portion of Eq. 7 can be expanded into a series of integrals:

$$\frac{\left(aq_f\right)^a}{(a-1)!} \frac{(k\mu)^n}{n!} \int_{T_s}^{\infty} e^{-\left(k\mu + aq_f\right) t}\, t^{a+n-1}\, dt, \quad (n = 0, 1, 2, \ldots, k-1)$$

By letting $b = k\mu + aq_f$, this expression can be manipulated into a series of Erlang integrals with the general term being

$$\frac{\left(aq_f\right)^a (a-1+i)!\,(k\mu)^i}{b^{a+i}\,(a-1)!\,i!} \int_{T_s}^{\infty} \frac{e^{-bt}\, b^{a+1}\, t^{a-1+i}}{(a-1+i)!}\, dt, \quad (i = 0, 1, 2, \ldots, k-1)$$

Integration of each integral results in a series of cumulative Poisson series that has a general term of the form

$$\frac{\left(aq_f\right)^a}{b^{a+i}} \frac{(a-1+i)!}{(a-1)!} \frac{(k\mu)^i}{i!} \left[e^{-bT_s} \sum_{j=0}^{a-1+j} \frac{\left(bT_s\right)^j}{j!} \right], \quad (i = 0, 1, 2, \ldots, k-1)$$

Substituting this expression into Eq. 6 gives

$$P(a) = 1 - \frac{\dfrac{\left(aq_f\right)^a e^{-bT_s}}{b^a} \displaystyle\sum_{j=0}^{k-1} \left[\frac{(k\mu)^j}{b^j} \frac{(a+j-1)!}{(a-1)!\,j!} \sum_{i=0}^{a+j-1} \frac{\left(bT_s\right)^i}{i!} \right]}{e^{-aq_fT_s} \displaystyle\sum_{i=0}^{a-1} \frac{\left(aq_fT_s\right)^i}{i!}} \tag{8}$$

Let p(n) be the probability of n consecutive drivers accepting an assigned gap before a driver rejects his gap. Then, enumerating the possible combinations gives

$$p(0) = 1 - P(a)$$

$$p(1) = P(a) \cdot [1 - P(a)]$$

$$p(2) = [P(a)]^2 \cdot [1 - P(a)]$$

$$\vdots$$

$$p(n) = [P(a)]^n \cdot [1 - P(a)]$$

The expected number of consecutive drivers accepting their assigned gap before one driver rejects his gap is

$$E(n) = \sum_{n=0}^{\infty} n \cdot p(n) = \frac{P(a)}{1 - P(a)} \tag{9}$$

The expected number of vehicles, E(q), on the ramp when a driver rejects his assigned gap is related to the average travel time on the ramp and the service time at server No. 1 by

$$E(q) = 1 + \frac{T_r}{S_1} \tag{10}$$

where T_r = average travel time on the ramp.

The average delay suffered in the merging area by drivers who reject or miss their assigned gap can be estimated by assuming each driver looks for a gap larger than a critical gap, T_m. The average delay for this merging problem (17) is

$$T_d = \frac{e^{aq_f T_m} - \sum_{i=0}^{a} \frac{\left(aq_f T_m\right)^i}{i!}}{q_f \sum_{i=0}^{a-1} \frac{\left(aq_f T_m\right)^i}{i!}} \tag{11}$$

which is the same form as Eq. 2. The average total delay, E(d), to all vehicles on the ramp when the first driver of a group rejects his gap is

$$\begin{aligned} E(d) &= T_d[1 + 2 + 3 + \ldots + E(q)] \\ &= \frac{E(q)\,[E(q) + 1]}{2}\, T_d \end{aligned} \tag{12}$$

The average service time in server No. 2 (Fig. 1), S_2, is the average total delay suffered by vehicles delayed when a ramp driver rejects his assigned gap divided by the sum of the average number of vehicles delayed and the average number of vehicles between such delayed groups. The expression for S_2 is

$$S_2 = \frac{E(d)}{E(q) + E(n)} \tag{13}$$

The expected total service time in the merging control system is the sum of the average service times in each server. The ramp service volume under merging control, q_r, is the inverse of the expected total service time so that

$$q_r = \frac{1}{S_1 + S_2} \tag{14}$$

Effects of Ramp Geometry and Ramp Operation

Institution of merging control on a freeway entrance ramp should reduce the importance and influence of geometric characteristics since a driver approaches a merging area that is free of ramp congestion. There are, however, basic considerations of ramp geometry and operational characteristics that will influence the controlled merge operation. Previous research has been reported dealing with the effect of ramp vehicle speed relative to freeway speed, the effect of angle of convergence of the ramp with the freeway, and the effect of acceleration lane length on uncontrolled gap acceptance (5, 17, 22). Much of the restrictive influence of these variables has been removed by the merging controller as it selects a gap for the driver. However, relative speed at the ramp nose, sight distance as the driver approaches the merging area, angel of convergence, acceleration lane length, and ramp grade do appear to influence controlled merging. In the absence of any quantifiable evidence of the effect of geometrics on the operation of a controlled ramp, and in view of past findings for uncontrolled ramps (17), the type of ramp operation anticipated under controlled conditions may best be classified on the basis of expected relative speed, angle of convergence, and acceleration lane length (Table 1).

TABLE 1

TYPE OF RAMP OPERATION

Relative Speed = Average Freeway Speed - Average Ramp Speed at Nose									Angle of Convergence (deg)
< 5 mph			5 mph–20 mph			> 20 mph			
Acceleration Lane Length (ft)			Acceleration Lane Length (ft)			Acceleration Lane Length (ft)			
0-300	300-600	> 600	0-300	300-600	> 600	0-300	300-600	> 600	
I	H	H	I	H	H	L	I	H	0-6
I	H	H	L	I	H	L	L	I	6-12
L	I	I	L	L	I	L	L	I	>12

H = high type, I = intermediate type, L = low type.

Gap-Acceptance Parameters

The classification of high-, intermediate-, and low-type operations can be assumed to correspond to gap-acceptance characteristics for moving vehicle merges described by Erlang distributions (Eq. 5) with the following Erlang constant, k, and mean accepted gap, $1/\mu$:

1. High type: $k = 2$, $1/\mu = 2.4$ sec;
2. Intermediate type: $k = 6$, $1/\mu = 3.0$ sec; and
3. Low type: $k = 10$, $1/\mu = 4.0$ sec.

Examples of each are shown in Figure 3.

Gap acceptance regressions for fast- and slow-moving vehicles provided a basis for selecting values of T_m, the critical gap for slow and stopped merges, associated with each classification of ramp operation (17). For the relative speed categories used in Table 1, the corresponding representative critical gaps for drivers who had rejected a gap were estimated to be

1. High Type: $T_m = 3.0$ sec;
2. Intermediate type: $T_m = 3.5$ sec; and
3. Low type: $T_m = 4.0$ sec.

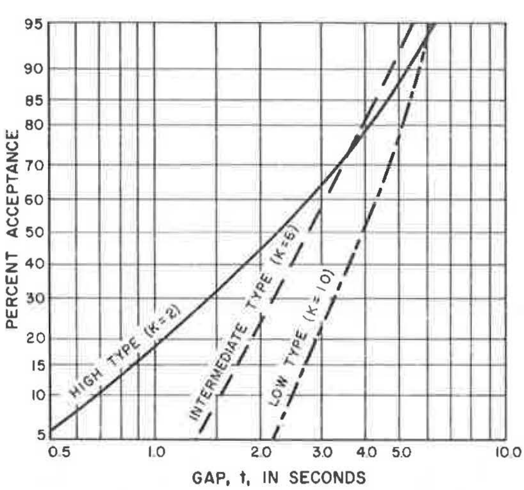


Figure 3. Gap acceptance characteristic for operation types.

Gap Distribution Parameters

In addition to the gap acceptance parameters, the outside lane freeway flow enters into the expression for ramp service volume. So that a specific Erlang distribution may be selected to describe a particular flow level, the Erlang constant, a, must be evaluated. Drew et al (5) found an empirical relation between

TABLE 2

ERLANG a AND OUTSIDE FREEWAY LANE FLOW

Erlang a:	2	3	3	4	5	6
Flow (vph):	800	1000	1200	1400	1600	1800

TABLE 3

SERVER DWELL TIME IN CONTROLLED MERGING AT DUMBLE RAMP

Sample (no.)	No. of Vehicles	Mean (sec)	Variance (sec^2)
1	35	2.73	3.39
2	74	2.72	1.02
3	65	2.83	3.99
4	25	2.25	0.26

the freeway lane flow rate, q_f, and the Erlang constant, a, from which the values in Table 2 were used to evaluate Eq. 14.

Server Dwell Time

A range of likely values from 1.5 to 4.0 sec for server dwell time could be established from previous research (23, 24). But measurements of server dwell time as given in Table 3 indicate that R = 2.6 sec is a good estimate for high- and intermediate-type operation. A value of 3.0 sec for R would appear to be a reasonable estimate for low-type ramps since the initial ramp speeds would likely be lower than for high- and intermediate-type ramps.

Ramp Travel Time

Ramp travel time in the development of Eq. 14 is the time to travel from the checkout detector to the merge area. Travel times, as reported in a companion report (15), can be adjusted for the server dwell time resulting in the values for each type of operation (Table 4). The same travel times are valid for ramp lengths different from those indicated because the ramp grade, sight distance, and ramp traffic composition obviously alter the value of travel time, but these variables are all reflected in the relative speed parameter.

Service Volume Control Charts

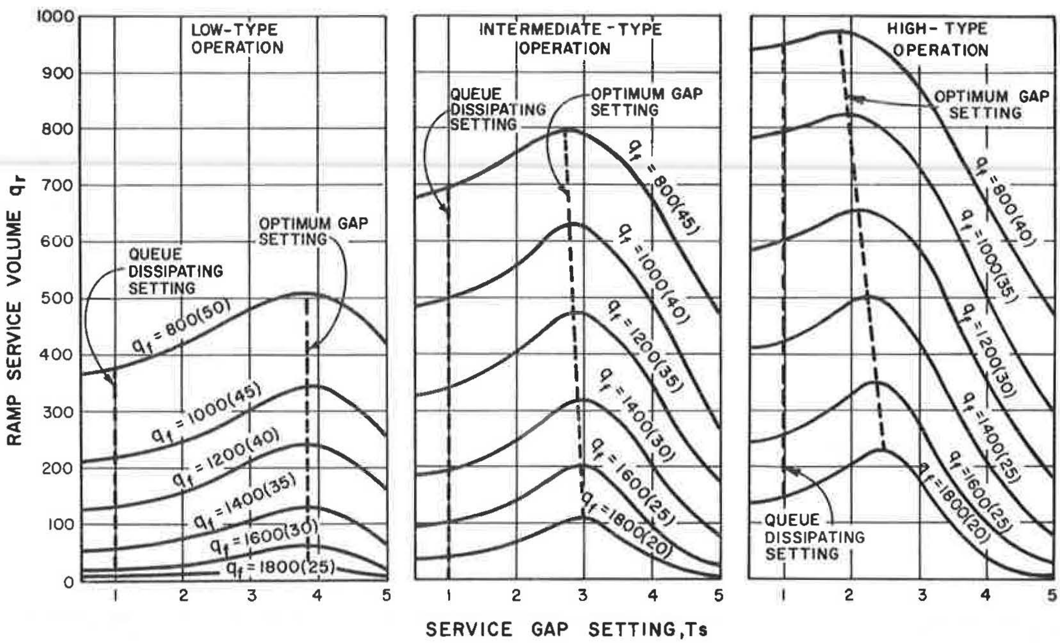
The family of curves in Figure 4 shows the ramp service volumes that can be attained with a certain service gap setting when the outside freeway lane flow is at a given level, as predicted by Eq. 14 under the assumptions stated. Each type of operation has an optimum controller setting (service gap) for a given level of freeway flow. As the controller setting is decreased from the optimum, vehicles are released more rapidly, but more drivers reject their assigned gaps, thus decreasing the service volumes. As the controller setting is increased above the optimum, the service time at the signal becomes excessively long, again decreasing the ramp service volumes. The optimum setting varies quite slowly as the freeway flow level varies.

When the congestion created by the queue of vehicles waiting at the ramp signal begins to reduce the capacity of frontage roads or arterial street intersections, it may be necessary to temporarily store queued vehicles on the ramp between the ramp signal and the merge area. Even under such conditions, it is still desirable to release the vehicles at the signal for an assigned gap. Thus, a queue-dissipating setting is indicated that reduces the overall efficiency of the merging control system operation, but, since the service gap is small, many gaps in the traffic stream larger than it are identified, thereby increasing the rate of release of vehicles at the ramp signal.

TABLE 4

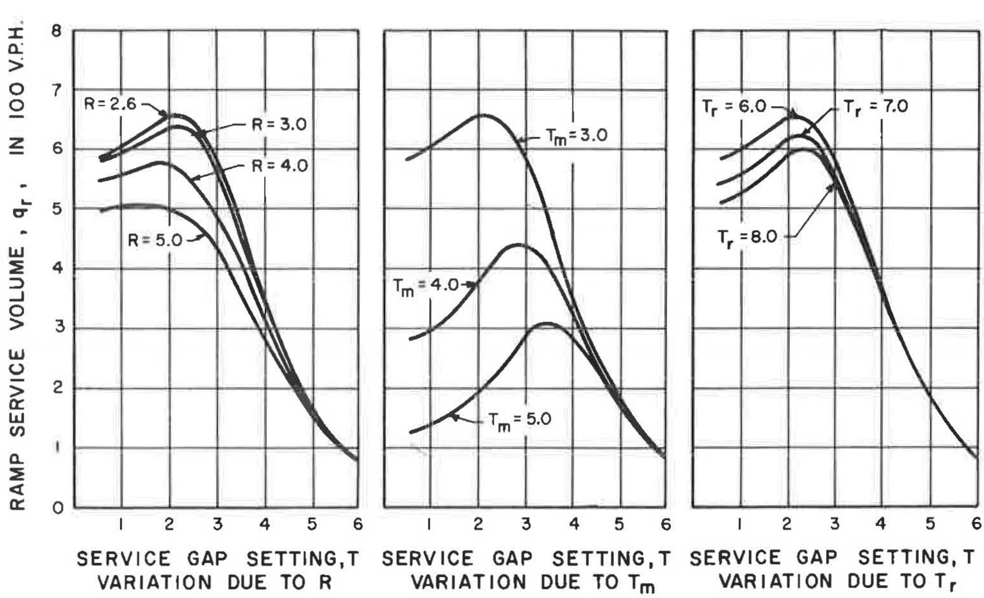
RAMP TRAVEL TIME CHARACTERISTICS

Type Operation	Travel Time (sec)	Distance From Signal to Merging Area for Level Ramps (ft)
High	6.0	150
Intermediate	7.0	200
Low	8.0	250



NOTE: q_f = 1200(40) IS THE OUTSIDE LANE UPSTREAM FREEWAY DEMAND WITH THE SPEED OF THE TRAFFIC STREAM IN PARENTHESES

Figure 4. Service volume control curves.



NOTE: INTERMEDIATE TYPE RAMP AT A FREEWAY FLOW OF 1200 V.P.H. (OUTSIDE LANE)

Figure 5. Parameter sensitivity.

Parameter Sensitivity

In order to indicate how sensitive the ramp service volume is relative to variations in the server dwell time, the critical gap for stopped vehicles, and the ramp travel time, control curves were developed holding all variables constant except R, T_m, and T_r. The results (Fig. 5) indicate that ramp service volume is relatively insensitive to server dwell time, R, and to ramp travel time, T_r. Variations in the critical accepted gap for slow and stopped merge vehicles, T_m, do produce wide ranges in the maximum ramp service volume (Fig. 5). However, the value of T_m for any specific ramp is a constant that is not related to merging control. When necessary, the curves of Figure 4 can be adjusted for the measured T_m characteristic of a particular location (17).

Ramp Capacity as a Function of Service Volume

The capacity of a controlled ramp is the maximum expected service volume for a given level of flow on the outside freeway lane. For a specific ramp, the geometry and the gap-acceptance parameters will be constant. Thus, ramp capacity will only be a function of the outside freeway lane flow and the service gap setting. Ramp capacity, q_{rc}, under merging control may be defined as

$$q_{rc} = q_r \text{ at } \frac{\partial q_r}{\partial T_s} = 0, \text{ at } q_f = \text{constant} \tag{15}$$

for a specific freeway flow level, q_f. Since evaluating the derivative in Eq. 15 is quite tedious, the approximate ramp capacity as a function of outside freeway lane flow can be taken from Figure 4. The ramp service volume for a given freeway flow level at the optimum service gap setting is the ramp capacity for that flow level. Figure 6 shows the variation of ramp capacity as freeway flow varies for each type of ramp operation.

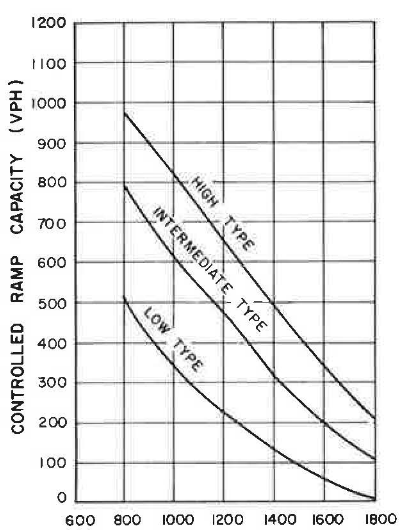


Figure 6. Ramp capacity under merging control.

VERIFICATION OF SERVICE VOLUME RELATION

Study Site

The Dumble Street entrance ramp to the northbound Gulf Freeway in Houston was selected for the data collection site to verify the mathematical model for ramp service volume. Figure 7 shows the site plan and relative location of the detection equipment. This particular site was selected for the following reasons:

1. A gap acceptance merging controller was available;
2. The ramp grade and the freeway grade in the vicinity of the ramp were nearly level;
3. Unrestricted sight distance was present at the location; and
4. Freeway flow in the vicinity of the ramp was usually free-flowing.

Data Collection and Data Analysis

Detection and control-function hardware components of the merging control system

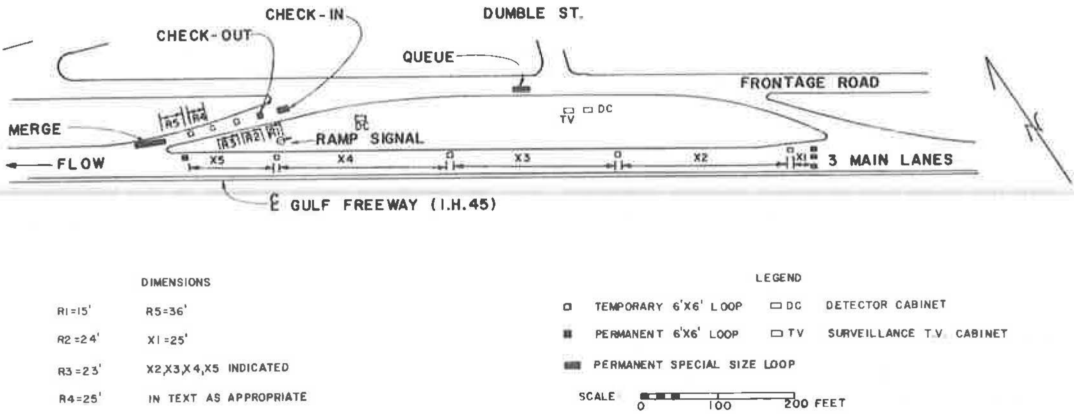


Figure 7. Layout of site.

were monitored during the morning peak-period control. The permanent control system inputs were supplemented with observers and temporary loop detectors placed on both the ramp and the outside freeway lane. All of the data were recorded on a 20-pen graphic recorder at the field site. The graphic record was run through a digital X-Y coordinate analyzer to produce punched card data records for analysis by digital computer.

Parameter Measurement Results

Ramp travel times from the check-out detector to the merge detector were measured, and the results are given in Table 5. The results for the four data sets are quite consistent. The Dumble Ramp is 135 ft long from the check-out detector to the merge detector, and is on a level grade with an acceleration lane of 350 ft and an angle of convergence of 12 deg. The ramp-freeway speed characteristics are given in Table 6. On the basis of the relative speed and the given geometric characteristics, Table 1 indicates an anticipated "intermediate-type" controlled merge operation.

Figure 8 shows the observed and predicted service volume rates for two service gap settings tested at the Dumble Street ramp (T_S = 2.5 sec and T_S = 3.0 sec). Each curve in Figure 8 represents the expected service volume for this intermediate-type ramp as the merging volume (combined freeway flow and ramp flow) varies. The majority of the observations are close to the expected curves, which tends to substantiate the model.

LEVELS OF CONTROL AND RAMP CAPACITY

Control of a freeway can be decomposed into a set of component subproblems with each subproblem responding to a particular level of control within a hierarchial structure of control functions (12). Within each level of control, ramp and freeway capacity assume varying degrees of influence on the establishment of an optimum control policy.

TABLE 5

RAMP TRAVEL TIME

Data Set	Number of Vehicles	Mean Time
1	35	5.6
2	74	5.5
3	65	5.6
4	25	5.5

First Level

The lowest level of control deals only with entrance ramps and considers each ramp to be independent of the total freeway or surface street operation. Merging control at this level seeks to put ramp vehicles into gaps in the outside lane of the freeway. Whenever a gap in the traffic stream arrives that is greater than or

TABLE 6

RAMP-FREEWAY SPEED IN MERGE AREA

Data Set	Average Speed at Ramp Nose (mph)	Average Freeway Speed in Merge Area (mph)	Average Outside Freeway Lane Flow (vph)	Relative Speed (mph)
1	20.0	28.4	1180	8.4
2	19.2	37.4	1210	18.2
3	18.0	38.0	1220	20.0
4	19.0	39.0	1185	21.0

equal to a service gap setting on the controller, a vehicle is released at the ramp signal (Fig. 1). If the ramp driver accepts his assigned gap, the controller continues to operate. If the ramp driver rejects his assigned gap, thereby blocking the merging area, the controller dwells until the merging area clears. Since no consideration is made of freeway capacity, the optimal control policy at this level is to put as many ramp vehicles into gaps as is possible. This means the ramp should be controlled to operate at its "controlled ramp capacity" (at the optimum setting of Fig. 4).

Second Level Control

Within the second level of the problem structure, merging control is expanded to include not just the ramp, but also variations in the speed and volume of the outside freeway lane flow and control alternatives for excessive queuing at the ramp signal. Operation of the merging controller is now functionally related to ramp service volume rather than capacity. The controller seeks to optimize the ramp service volume as the outside lane freeway operation (speed and volume) varies and the ramp signal queue fluctuates. Thus the controller setting, T_S, will move from one optimum to another as the outside freeway lane flow rate varies (Fig. 4), and as the ramp signal queue becomes excessively long the controller setting will drop to the queue-dissipating setting (Fig. 4).

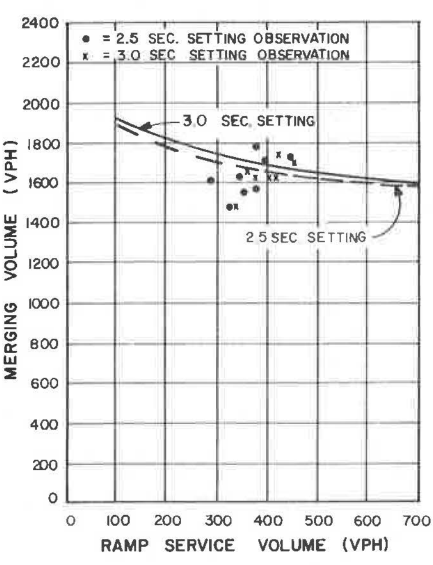


Figure 8. Observed and predicted service volume.

Third Level Control

An entrance ramp cannot often be controlled to operate at its maximum capacity without considering its effect on freeway operation in the vicinity of the ramp. Similarly, merging control will be more effective when it is responsive to capacity reductions on the freeway and to other deviations from established values of control model parameters. In the third level of merging control, controller operation is a function of ramp service volume as related to freeway capacity (Fig. 6).

Fourth Level Control

When the control of a freeway spans several ramps so that control of individual ramps becomes a problem of interrelating ramp and freeway operation over a freeway system, another level of control evolves. An overall system control concept establishes control criteria for each

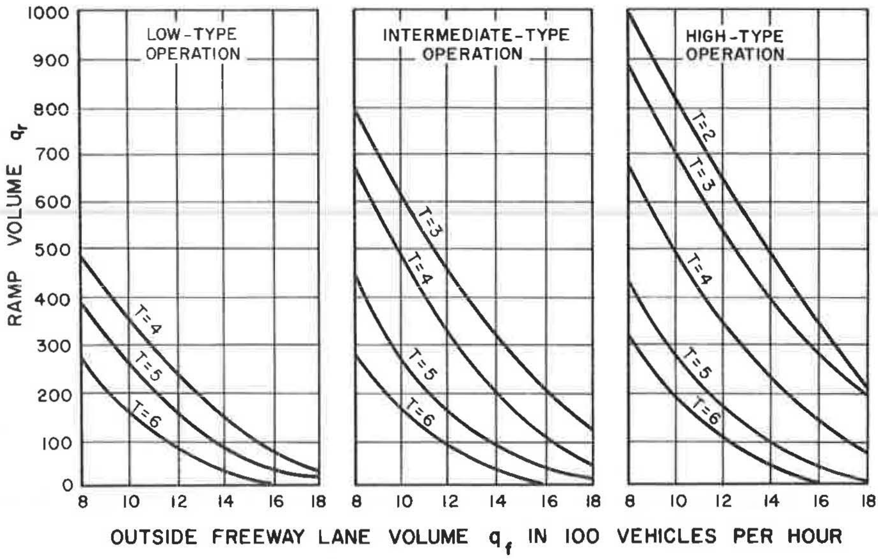


Figure 9. Curves illustrating the theory of reduced capacity operation.

ramp based on each ramp's controlled capacity, each ramp's service volume, freeway capacity constraints, freeway demand, and each ramp's demand. A linear programming model has been developed for long-time measurement periods (10). A mathematical programming model is being developed for short measurement periods (25).

APPLICATION TO FREEWAY CONTROL

Initiation of Control on an Isolated Entrance Ramp

An initial freeway control installation may consist of only a few merging controllers just as an initial traffic signal network installation may have signals at only a few intersections. Individual entrance ramps with merging operation problems may be identified and placed under control much as isolated problem intersections in a street network would be signalized. An isolated entrance ramp could be controlled with either first-level control or first- and second-level control.

If a ramp were to be controlled by a first level of control sophistication, a single service gap setting would be needed to achieve the most nearly optimum operation for expected outside freeway flow conditions during control periods. From Figure 4, the controller service gap setting would be about 2.4, 2.8, or 3.6 sec, depending on whether the ramp was considered a high-, intermediate-, or low-type ramp. This control policy would dictate operating the ramp at or near capacity.

When both first and second levels of control are installed at an isolated entrance ramp, the merging controller becomes responsive to variations in outside lane freeway flow and to the extent of queuing at the ramp signal. Figure 4 shows how the merging controller would vary its operation to achieve such responsiveness. As changes are detected in the speed and flow on the outside freeway lane, the controller service gap setting moves along the optimum setting line to maintain capacity operation (providing that demand exceeds controlled capacity) until the ramp signal queue becomes excessively long. The merging controller service gap setting is then adjusted to the queue-dissipating setting, which is shown to be 1.0 sec in Figure 4. Figure 4 shows that the ramp service volume will decrease but the congestion will be moved from the ramp signal to the ramp merging area. Such a trade-off would be appropriate when a ramp signal queue backs up into a frontage-road intersection or reduces frontage-road capacity at a critical location. Control function and hardware component interrelations to accomplish this operation are described in a companion report (12).

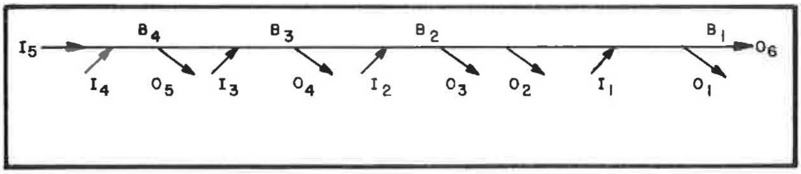


Figure 10. Schematic of freeway system control.

Integrated Multilevel Freeway Control

With a freeway control system capable of exercising various levels of control, each ramp will be controlled as previously described for an isolated ramp under the first and second levels of control (12), but each ramp will also be constrained to broader system-control optimization at higher levels of control.

If a traffic incident, a permanent geometric constriction, or an ambient condition reduces the capacity of the freeway downstream from the controlled ramp, it may be necessary to operate the merging controller so that the ramp service volume is less than optimum. This means that the third level of control will establish a service gap setting larger than the optimum shown in Figure 4. The theory of reduced capacity operation can be illustrated with Figure 9. The basic steps of controller operation are

1. Measure volume at downstream block (bottleneck capacity, Q);
2. Measure upstream freeway demand, q_f;
3. Enter abscissa at q_f and ordinate at $(Q - q_f)$; and
4. T = controller gap setting.

For example, if the ramp is an intermediate-type ramp and the bottleneck capacity of the outside lane is 1500 vph, then, with an outside freeway lane flow of 1200 vph, the allowable ramp volume is 300 vph. This locates a point below the curve for T = 4 so that the service gap setting must be about 3.5 sec or larger.

For a freeway control system to utilize the fourth level of control and achieve system optimization, the control policy must incorporate both the macroscopic (10) and the microscopic (11) approaches inherent in the various lower levels of control into a total system control policy. Messer has proposed a control model that has this capability (25).

The approach basically uses linear programming techniques to determine the optimum input (entrance) ramp flow rates on a system basis. Once the allowable input flow rates and the active input constraints are known, the appropriate service gap settings can be determined. The resulting service gap settings are then used as control inputs to the merging controllers for the particular ramps. This yields the desired flow rate (service volume) and a solution to the merging control problem. To illustrate the approach further, a freeway control example will be evaluated.

Consider the freeway schematic shown in Figure 10 as being the entire freeway system under control. Merging control is in effect at all input ramps with a central computer monitoring and controlling the total system operation.

The linear programming model used to determine the freeway inputs is

$$\text{Maximize} \sum_{j=1}^{5} X_j$$

$$\text{subject to} \sum_{j=1}^{5} A_{jk} X_j \leq B_k \qquad k = 1, \ldots, 4$$

$$\text{and} \qquad b_m P_a \sum_{j=a+1}^{5} A_{ja} X_j + X_a \le C_m \qquad a = 1, \ldots, 4$$

$$\text{and} \qquad X_j \le D_j \qquad j = 1, \ldots, 4$$

$$X_5 = D_5$$

$$\text{and} \qquad X_j \ge 0 \qquad j = 1, \ldots, 5$$

X_j is the input volume to the freeway system; A_{jk} is the decimal fraction of vehicles entering at input j that pass through section k; B_k is the bottleneck capacity (in this example) of freeway section k; A_{ja} is the decimal fraction of vehicles entering at input j that pass by ramp a. The percent of the total freeway volume upstream of input a, which is in the lane adjacent to entrance ramp a (25 percent was used in all cases in this example (5)) is P_a; D_j is the present demand at input j. To determine the b_m and C_m values for a particular ramp input a, it is necessary to first classify ramp a according to type—high, intermediate, low, (m). Then the b_m and C_m values are the slope and the y intercept, respectively, of the linear estimate of the ramp flow vs freeway flow capacity curves in Figure 6.

The required inputs to the model are given in Table 7, and Table 8 gives the optimum flow rates and the appropriate service gap settings for the merging controllers. These service gap settings will always be the optimum setting for the particular freeway flow unless the active constraint for that input ramp is the "bottleneck" constraint. If the "bottleneck" constraint is the active constraint, then the "throttled" gap setting as determined from the appropriate operating curve in Figure 4 is used to compute the required service gap.

TABLE 7

REQUIRED INPUT TO L-P

CLASSIFICATION OF INPUT RAMPS

INPUT RAMP	1	2	3	4
TYPE	INTERMED.	LOW	INTERMED.	INTERMED.

RAMP FREEWAY LANE 1 CAPACITY EQUATIONS

HIGH	$q_r + .770\ q_f$	≤ 1590
INTERMEDIATE	$q_r + .695\ q_f$	≤ 1320
LOW	$q_r + .494\ q_f$	≤ 840

PRESENT DEMAND

INPUT	1	2	3	4	5
DEMAND	300	80	500	400	5400

INPUT		1		2		3		4		5	
FRACTION DESTINED TO X	x	k	a	k	a	k	a	k	a	k	a
	1	1		.86	.76	.52	.60	.48	.59	.79	.82
	2			1.		.93	.93	.91	.91	.90	.90
	3					1.		1.	1.	.94	.94
	4							1.		1.	1.

TABLE 8

RESULTS OF L-P MODEL

INPUT RAMP	1	2	3	4
OPTIMUM FLOW (VPH)	380	61	233	380
BOTTLENECK ACTIVE ?	NO	YES	YES	NO
SERVICE GAP	2.86	5.47	3.92	2.89

REFERENCES

1. Highway Capacity Manual—1965. HRB Spec. Rept. 87, 1965.
2. Hess, J. W. Ramp-Freeway Terminal Operation as Related to Freeway Lane Volume Distribution and Adjacent Ramp Influence. Highway Research Record 99, p. 81-116, 1965.
3. Moskowitz, K., and Newman, L. Notes on Freeway Capacity. Highway Research Record 27, p. 44-68, 1963.
4. Major, N. G., and Buckley, D. J. Entry to a Traffic Stream. Australian Road Research Board Proc., Vol. 1, Part 1, p. 206-228, 1962.
5. Drew, D. R., Buhr, J. H., and Whitson, R. H. Determination of Merging Capacity and Its Applications to Freeway Design and Control. Highway Research Record 244, p. 47-68, 1968.
6. May, A. D. Experimentation With Manual and Automatic Ramp Control. Highway Research Record 59, p. 9-38, 1964.
7. Gervais, E. F. Optimization of Freeway Traffic by Ramp Control. Highway Research Record 59, p. 104-118, 1964.
8. Drew, D. R., McCasland, W. R., Pinnell, C., and Wattleworth, J. A. The Development of an Automatic Freeway Merging Control System. Texas Transportation Institute, Research Rept. 24-19, 1966.
9. Wattleworth, J. A., and Wallace, C. E. Evaluation of the Operational Effects of an "On-Freeway" Control System. Texas Transportation Institute, Research Rept. 438-2, 1967.
10. Wattleworth, J. A., and Berry, D. S. Peak-Period Control of a Freeway System—Some Theoretical Investigations. Highway Research Record 89, p. 1-25, 1965.
11. Drew, D. R. Theoretical Approaches to the Study and Control of Freeway Congestion. Texas Transportation Institute, Research Rept. 24-1, 1964.
12. Drew, D. R., Brewer, K. A., Buhr, J. H., and Whitson, R. H. Multilevel Approach to the Design of a Freeway Control System. Paper presented at 48th Annual Meeting and included in this Record.
13. McCasland, W. R., Buhr, J. H., Carvell, J. D., and Drew, D. R. Analog Merging Controllers for Freeway Entrance Ramps. Texas Transportation Institute, Research Rept. 504-3, 1968.
14. Whitson, R. H., Buhr, J. H., Drew, D. R., and McCasland, W. R. Real-Time Evaluation of Freeway Quality of Traffic Service. Texas Transportation Institute, Research Rept. 504-4, 1968.
15. Buhr, J. H., Whitson, R. H., Brewer, K. A., and Drew, D. R. Traffic Characteristics for Implementation and Calibration of Freeway Merging Control Systems. Paper presented at 48th Annual Meeting and included in this Record.
16. Brewer, K. A. Description and Evaluation of Controlled Freeway Merging. Unpublished PhD dissertation, Texas A&M University, 1968.
17. Drew, D. R., Lamotte, L. R., Buhr, J. H., and Wattleworth, J. A. Gap Acceptance in the Freeway Merging Process. Highway Research Record 208, p. 1-36, 1967.
18. Raff, M. S., and Hart, J. W. A Volume Warrant for Urban Stop Signs. Eno Foundation for Highway Traffic Control, Saugatuck, Conn., 1950.
19. Bissell, H. H. Traffic Gap Acceptance From a Stop Sign. Unpublished Graduate Research Report, ITTE Univ. of California, Berkeley.

20. Drew, D. R. Gap Acceptance Characteristics for Ramp-Freeway Surveillance and Control. Highway Research Record 157, p. 108-143, 1967.
21. Blunden, W. R., Clissold, C. M., and Fisher, R. B. Distribution of Acceptance of Gaps for Crossing and Turning Maneuvers. Australian Road Research Board Proc., Vol. 1, Part 1, p. 188-205, 1962.
22. Worrall, R. D., Coutts, D. W., Echterhoff-Hammershmid, H., and Berry, D. S. Merging Behavior at Freeway Entrance Ramps: Some Elementary Empirical Considerations. Highway Research Record 157, p. 77-107, 1967.
23. Greenshields, B. D., Shapiro, D., and Ericksen, E. L. Traffic Performance at Urban Street Intersections. Yale Bureau of Highway Traffic, New Haven, Conn., 1947.
24. Capelle, D. G., and Pinnell, C. Capacity Study of Signalized Diamond Interchanges. HRB Bull. 291, p. 1-25, 1961.
25. Messer, C. J. Alternatives and Consequences of Real-Time Freeway Control. PhD dissertation (in progress), Texas A&M University, 1968.

Traffic Characteristics for Implementation and Calibration of Freeway Merging Control Systems

JOHANN H. BUHR, ROBERT H. WHITSON, KENNETH A. BREWER, and DONALD R. DREW, Texas Transportation Institute, Texas A&M University

Using the Gulf Freeway in Houston as the laboratory, a number of traffic characteristics associated with gap acceptance control were studied. Data collection was performed using manual methods, a closed-circuit television system, and an electronic data acquisition system including field sensors and a digital computer.

The ramp traffic characteristics presented include the vehicle performance characteristics of speed, acceleration, and travel time, the relationships among these variables, and the effects of vehicle type and ramp grade on these variables. The behavior of ramp vehicles at the control signal was also studied to determine the starting-delay characteristics and stopping position of vehicles at the ramp signal. A frequency distribution of each variable is presented and the application of the results in the implementation and calibration of merging control systems is discussed.

Several freeway traffic characteristics were also studied. These were concerned mainly with traffic stability upstream from an entrance ramp and the effect of this stability on the accuracy with which the trajectory of a ramp vehicle can be matched with a freeway gap. Also presented is the relationship between total travel and total travel time, two measures of effectiveness often used in the evaluation of freeway control systems. The possible application of this relationship in the development of warrants for freeway control is indicated.

•IN March of 1966, the first prototype automatic traffic-responsive merging controller was installed on the Telephone Road inbound entrance ramp of the Gulf Freeway in Houston. Since that time, considerable experience has been gained with the operation of this type of instrument and its effect on traffic behavior. This has led to the development of first-generation and second-generation controllers of an increasingly complex nature. The rapidity of these developments has caused the accompanying expansion of knowledge to be limited to relatively few researchers and practicing traffic engineers. Much of the new technology is of a rather subjective nature, based on the day-to-day observation of the operation of these controllers over a long period of time. Documentation of findings, both speculative and factual, has lagged far behind. Many of the more subjective aspects of the controlled operation are difficult or impossible to measure or otherwise pin down factually, so that researchers are understandably loath to record their views on these matters. However, there are a number of readily measurable traffic characteristics that affect the design and operation of a control system and that are affected by the control system. The objective of this paper is to document these characteristics. It is an outgrowth of a research project

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 48th Annual Meeting.

sponsored by the U. S. Bureau of Public Roads to develop functional and operational requirements of merging control systems and system design specifications for freeway control by centralized digital computer.

RAMP TRAFFIC CHARACTERISTICS

Ramp-Vehicle Performance Characteristics

In the peak-period control of a freeway system, it is normally necessary to bring all ramp vehicles to a stop before releasing them for a gap. This is done not only because it may be necessary to wait for an acceptable gap, but also because the travel time of the ramp vehicle can be more accurately predicted, thus making it possible to fit the vehicle into a specific gap. It is therefore evident that ramp vehicle performance characteristics, in terms of speed and acceleration, are of primary interest.

Theory—The average driver's normal acceleration can reasonably be assumed to have a linear relationship with the speed of his vehicle as represented by the differential equation

$$\frac{du}{dt} = a - bu \tag{1}$$

where u is the speed of the vehicle, t is time, and a and b are constants (1).

If a ramp vehicle, after being stopped by the ramp signal, accelerates normally so that its acceleration decreases as its speed increases, as given by Eq. 1, then after time t its speed is given by

$$u = \frac{a}{b}\left(1 - e^{-bt}\right) \tag{2}$$

and it has traveled x feet down the ramp as given by

$$x = \frac{a}{b}t - \frac{a}{b^2}\left(1 - e^{-bt}\right) \tag{3}$$

Equation 3 gives the time-space relationship of a vehicle accelerating normally from a stopped position.

However, if it is assumed that a ramp driver utilizes uniform or constant acceleration of c ft/sec^2 instead of a non-uniform acceleration decreasing with his speed, then, after time t, his speed will be given by

$$u = ct \tag{4}$$

and he would have covered x feet, where

$$x = \tfrac{1}{2}ct^2 \tag{5}$$

In on-ramp control, the range of interest in the time-space relationship is determined by the location of the ramp signal with respect to the merge point. In gap acceptance control, this distance would normally be between 100 and 200 feet. Over this distance, the time-space trajectories given by Eqs. 3 and 5 are quite similar if the constants are selected carefully, as shown in Figure 1. Because the equations for uniform acceleration are much simpler than those for non-uniform acceleration, it is, in some instances, advantageous to assume uniform acceleration without sacrificing realism. The two theories will, however, indicate considerably different speeds at

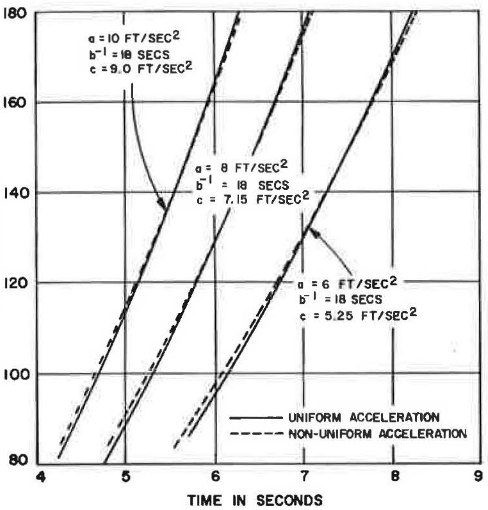


Figure 1. Comparison of time-space traces given by uniform and non-uniform acceleration.

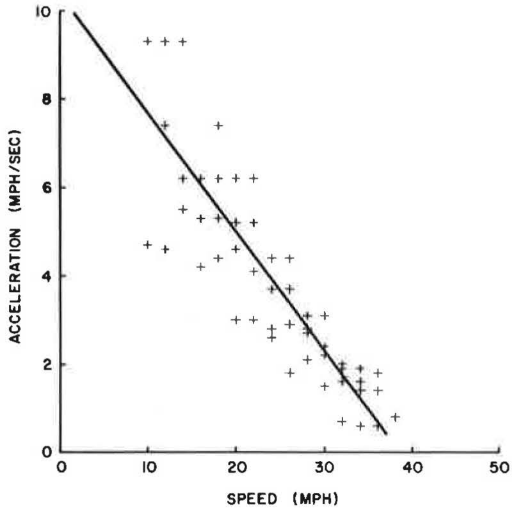


Figure 2. Acceleration vs speed of passenger vehicle during normal acceleration from a standing start on an entrance ramp.

various points on the ramp. The researcher should, therefore, be mindful of the effects of substituting one theory for another.

Study Procedure—In order to evaluate the constants that define the speed-acceleration characteristics of ramp vehicles, several runs were made on both controlled and uncontrolled ramps using a 1966 Plymouth Station Wagon equipped with a V8 engine and automatic transmission and instrumented with a recording speedometer. Several drivers were used and only smooth merges were included in the analysis. On the few uncontrolled ramps studied, the vehicle was brought to a stop on the ramp before accelerating on to the merge area. From the recorded time-speed trace, the acceleration was calculated over every 2-mph increase in speed. The accelerations were then plotted against the speeds, and straight lines fitted through the points.

Results—The results of these experiments showed quite similar speed-acceleration characteristics for different drivers and different ramps, except for some isolated runs. Figure 2 shows the results of four runs made by the same driver on the Telephone inbound entrance ramp (zero grade) on the Gulf Freeway (correlation coefficient = 0.85 with Eq. 1). It is representative of all the runs recorded. The values of the parameters a and b^{-1} yielded by these experiments are approximately 10 mph/sec (15 ft/sec^2) and 4 seconds, respectively.

These results are quite different from those found by other experimenters whose studies, performed under different conditions, yielded values of a ranging between 4 and 7 mph/sec with b^{-1} ranging between 20 and 15 seconds (2, 3). These differences are illustrated in Figure 3.

The Effect of Grades—To evaluate the effect of grades on the speed-acceleration characteristics of passenger vehicles on an entrance ramp, studies as described above were also performed on the Griggs inbound entrance ramp on the Gulf Freeway (3 percent upgrade) and on the Newcastle inbound entrance ramp (5 percent upgrade) on the Southwest Freeway in Houston. The results of four test runs on Griggs and three test runs on Newcastle are shown in Figures 4 and 5, respectively, with correlation coefficients of 0.84 and 0.87. From the figures it would appear that on the steeper ramp the vehicle generally utilized a slightly greater acceleration than on the less steep ramp. To compare these characteristics with those of a level ramp, the results obtained on the Telephone (zero grade), the Griggs (3 percent upgrade), and the Newcastle (5 percent upgrade) ramps are superimposed in Figure 6. These results indicate that vehicles apparently utilize greater acceleration on steeper ramps.

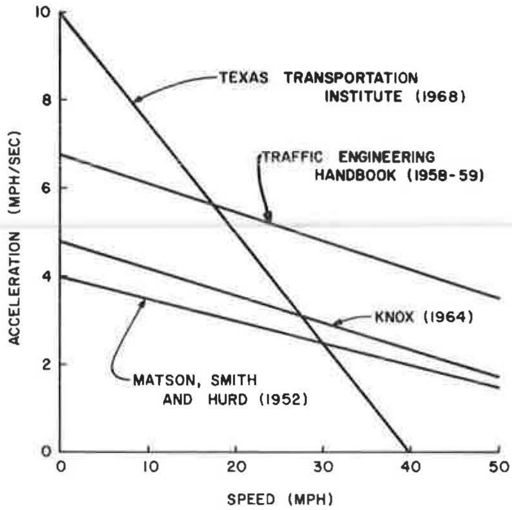


Figure 3. Comparison of speed-acceleration characteristics found by different experimenters.

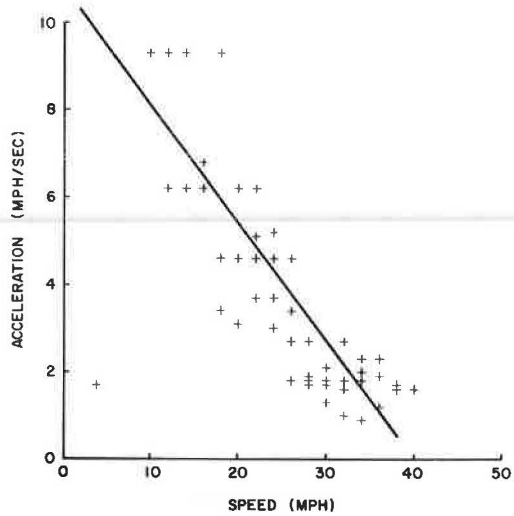


Figure 4. Speed-acceleration characteristics on the Griggs Road entrance ramp (3 percent upgrade).

Ramp Travel Times, Speeds, and Accelerations

The travel time of a ramp vehicle on a controlled ramp is considered here to be that time which expires between the instant the ramp signal turns green and the instant the vehicle actuates a merge detector located at the physical nose of the ramp. This important variable affects the controller settings and places certain limitations on the detector layout of a control system, and its variation affects the number of vehicles that will actually "hit" their assigned gaps (4).

Data Collection—The data required to study the travel-time characteristics of vehicles on a controlled ramp were measured and recorded by the digital computer being used to control the ramps. The data were collected only for those merging vehicles that had a clear and unoccupied ramp ahead as they were released at the ramp signal. Travel times were sampled on each of six ramps during two morning peak hours.

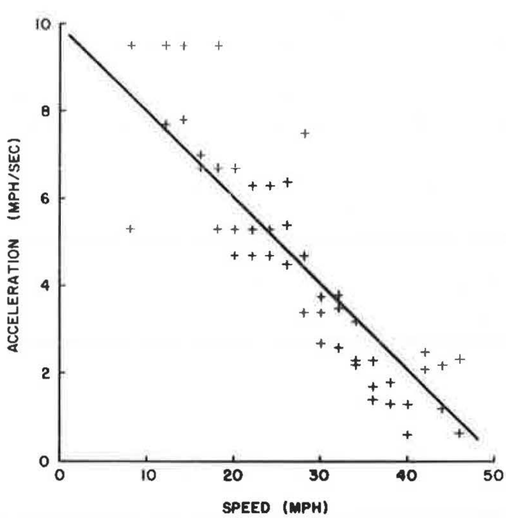


Figure 5. Speed-acceleration characteristics on the Newcastle entrance ramp (5 percent upgrade).

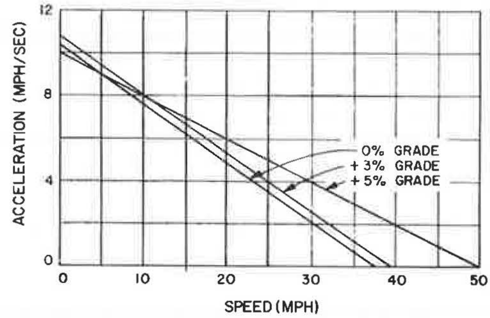


Figure 6. Effect of ramp grade on speed-acceleration characteristic of passenger car.

TABLE 1

NUMBER OF VEHICLES HAVING A GIVEN TRAVEL TIME BETWEEN THE RAMP SIGNAL AND THE MERGE DETECTOR

Travel Time, Class Midpoint, Seconds	Woodridge		Mossrose		Griggs		Wayside		Telephone		Dumble	
	First Day	Second Day	First Day	Second Day	First Day	Second Day	First Day	Second Day	First Day	Second Day	First Day	Second Day
4.25			3									
4.75			3	2							2	2
5.25		3	8	9			1			1	4	3
5.75	2	1	18	19	6	3	2	1	1	2	5	4
6.25	5	6	22	26	9	5	3	1	3	7	4	6
6.75	6	10	25	29	9	10	8	2	9	9	4	13
7.25	6	7	15	20	14	14	7	6	9	22	11	16
7.75	12	6	9	11	9	11	12	7	14	19	11	12
8.25	4	11	8	6	6	6	18	10	11	18	8	10
8.75	2	4	4	5	3	8	22	12	9	9	8	8
9.25	3	2	1	2	6	3	18	4	10	5	7	4
9.75	4	1	2	0	2	2	11	6	3	1	4	3
10.25	1		2	2	2	0	3	2	5	2	2	0
10.75					2	1	0	0	2	2	2	0
11.25					1	1	0	0		3	0	1
11.75					1		0	1		0	1	1
12.25							1			1		
Total	45	51	120	131	70	64	106	52	76	101	73	83
Mean	7.71	7.39	6.70	6.76	7.67	7.63	8.39	8.40	8.16	7.88	7.79	7.50
Variance	1.32	1.04	1.39	1.04	2.05	1.32	1.23	1.23	1.32	1.51	2.25	1.64
χ^2 Calc.	6.89	4.69	6.07	8.67	8.29	3.59	7.72	3.08	3.29	9.85	4.00	2.17
χ^2 0.05	9.49	7.82	12.59	12.59	12.59	9.45	11.07	9.49	11.07	12.59	12.59	11.07
t Statistic	1.45		0.43		0.18		0.05		1.54		1.30	
F Statistic	1.27		1.34		1.55		1.00		1.14		1.37	

Data Analysis—The data, classified into frequency tables for each of two days for each ramp, are given in Table 1. Also given are the means and variances of the data, as well as three different test statistics. The chi-square test was performed on each data set to test the goodness of fit with a normal distribution having the same mean and variance. By comparing the calculated chi-square values with the chi-square value for the 5 percent confidence limit, it was concluded that each of the twelve data sets was drawn from normal populations, allowing a 5 percent chance that this conclusion is in error.

In order to determine if the travel times measured on one day were significantly different from those measured on another day, t-tests and F-tests were performed on the means and variances of the data, testing respectively the hypotheses that the means on two different days are the same and that the variances measured on the two different days are the same. The calculated t- and F-statistics (Table 1) lead to the conclusion that, with 5 percent confidence, the distribution of ramp travel times did not change from day to day on the same ramp.

Of primary interest, of course, is the ability to predict the ramp travel-time distribution for a given signal location. Toward this end, the data taken on two days, for each ramp, were grouped together and the means and variances calculated. The following table gives the mean and standard deviation of ramp travel time together with the length of ramp over which travel times were measured:

	Woodridge	Mossrose	Griggs	Wayside	Telephone	Dumble
Length (feet)	145	108	118	168	146	143
Mean travel time (seconds)	7.54	6.73	7.65	8.39	8.00	7.63
Variance	1.19	1.21	1.69	1.23	1.44	1.93

These travel times include the starting delays of drivers at the signal. By plotting these data points on time and distance coordinates and superimposing the plot over time-space graphs, the starting delay and the acceleration rates can be estimated. By assuming non-uniform acceleration and using the parameter of a = 10 mph/sec and b^{-1} = 4 seconds derived earlier, the starting delay was estimated to be 2.4 seconds. For the sake of simplicity a uniform acceleration of 10 ft/sec, together with a starting delay of 2.4 seconds, can be used. For all practical purposes, over the range of ramp lengths of from 80 to 200 feet, the ramp travel time of the average vehicle on a level ramp can be considered as being given by

$$\text{Travel time} = 2.4 + \sqrt{\frac{\text{distance}}{5.0}}$$

The variances measured at the six different ramps as given in the table indicate that, over the distances covered (110 feet to 170 feet), there is apparently no relationship between the distance traveled and the variance in travel time.

Conclusions—The following conclusions are warranted:

1. For purposes of predicting the ramp travel time on a controlled ramp, it makes practically no difference whether uniform or non-uniform acceleration is assumed over the normal range of ramp signal locations.
2. If uniform acceleration is assumed, then it is reasonable to assume that the average vehicle has a starting delay of 2.4 seconds and an acceleration of 10.0 ft/sec^2; if non-uniform acceleration is assumed, then it is reasonable to assume that the average vehicle has a starting delay of 2.4 seconds and a maximum acceleration of 10 mph/sec, which decreases linearly by ¼ mph/sec for every 1-mph increase in vehicle speed.
3. Ramp travel times, on a controlled ramp where all vehicles are brought to a stop before they are released for a gap, are uniformly distributed, with the mean related to the distance traveled as mentioned in conclusion 2.
4. The distribution of ramp travel times on the same ramp does not change from day to day.
5. The ramp travel times were not markedly more variable over the longer ramps than over the shorter ramps studied.

The Value of Hitting a Gap

The foregoing data pertained to vehicles whose progress up the ramp was unimpeded and, therefore, were able to enter the freeway smoothly, performing ideal merges. The provision of ideal merging opportunities is one of the two objectives of merging control systems (5). In many cases, when a ramp driver misses his assigned gap he is forced to stop on the acceleration lane and wait for an acceptable gap. Since the relative speed is now greatly increased, the size of the acceptable gap is also increased, necessitating a longer wait for the stopped driver as well as for other drivers desiring to enter the ramp (6). This increased delay reduces the service volume of an entrance ramp, as illustrated by the capacity curves developed by Brewer et al (7). The value of fitting a vehicle into a gap under such conditions is evident.

When the freeway volume is fairly low, chances are that a ramp vehicle missing its assigned gap may find another gap to enter into without having to come to a complete stop. Even under these conditions, however, missing the assigned gap has a detrimental effect on ramp operation. This is illustrated by Figure 7, which shows the time-speed traces recorded by an instrumented vehicle under three conditions when it hit its assigned gap and under three conditions when it did not. The detrimental effect of not hitting its assigned gap is immediately evident.

As a quantitative measure of this effect, the acceleration noise experienced by the driver was calculated for eight runs in which the driver performed an ideal merge and five runs in which he missed the gap but still entered the freeway without having to stop.

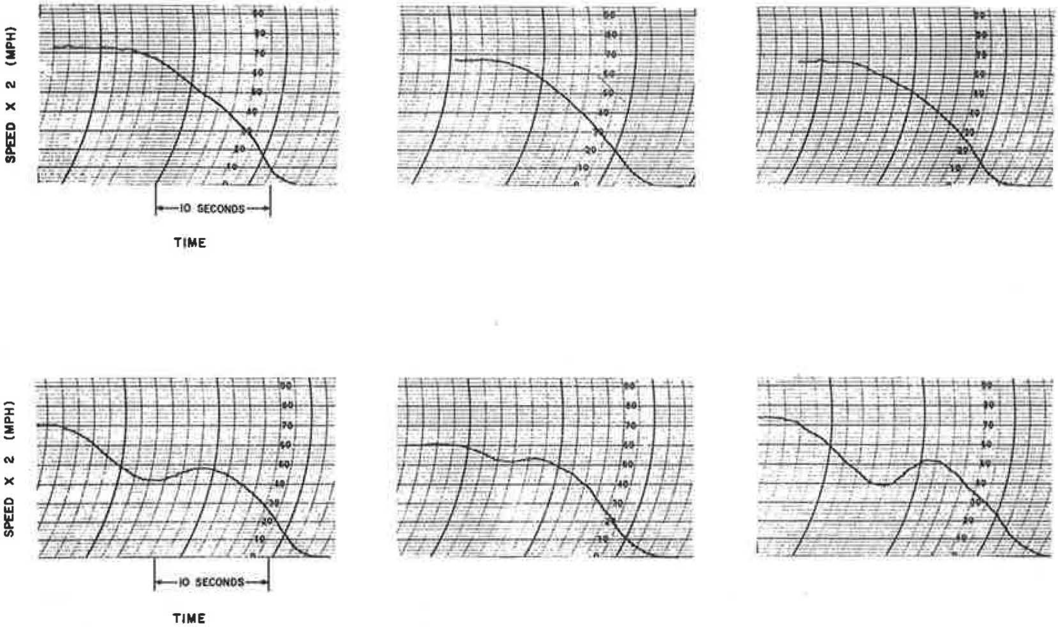


Figure 7. Recording speedometer traces of ramp vehicle: (top) matched with acceptable gap; (bottom) not matched with acceptable gap.

The average acceleration noise for the eight ideal merges was 2.23 mph/sec, and for the five merges in which the driver missed his assigned gap, it was 3.91 mph/sec.

Number of Vehicles Hitting Gaps—In order to study the effect of gap-acceptance control on merging behavior, studies were carried out on five entrance ramps during the periods of afternoon control by digital computer. During the afternoon, it is considered that freeway capacity is not a problem. The first objective of a merging control system, that of prevention of congestion, is therefore not a factor in afternoon control. However, ramp volumes are generally higher during the afternoon peak than during the morning peak. These studies therefore isolate the second objective, that of improving merging conditions.

In these studies, observers classified all ramp vehicles first into two groups—those that used the ramp signal correctly (called non-violators) and those that "ran" the red signal (called violators). The non-violators were then further classified into one of four groups—those that "hit" gaps (assumed to be their assigned gaps) and were thus able to perform smooth ideal merges; those that were matched with their assigned gaps but refused them and accepted another gap without first stopping; those that missed their assigned gaps and entered another gap without stopping; and finally, those that stopped on the ramp or acceleration lane before merging either because no gap was available or because another vehicle was stopped in the merging area. Violators were classified into one of three groups—those that met gaps and performed smooth merges; those that did not immediately meet a gap but were still able to merge without stopping; and those that stopped before merging. Studies were made for a period of 1 hour during the afternoon control, which lasts $1\frac{1}{2}$ hours. Data were collected for each 5-minute period during the study. Altogether, 31 studies were made, covering 370 five-minute periods. Observations were taken on 9831 ramp vehicles.

In order to investigate whether the percentage of vehicles in each group was affected by the freeway volume, plots were drawn up of the percentage in each group, based on 15-minute totals, against the freeway outside lane volume, based on 15-minute counts

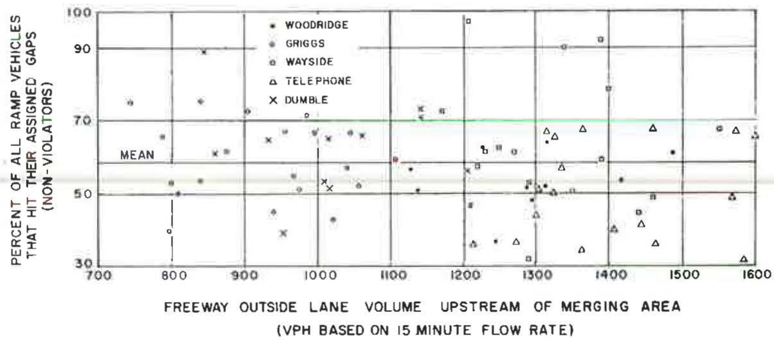


Figure 8. Relationship between freeway volume and frequency of smooth merges for gap-acceptance control by digital computer.

by the digital computer. (Freeway counts were not available for some of the study days.) Figure 8 shows the plot of non-violators performing smooth merges as a percentage of all ramp vehicles, against the freeway outside lane flow rate. This plot is typical of all the other groups (except for the magnitude of the percentage). No trend is evident, and the conclusion is, therefore, that the percentage of ramp vehicles in each of the aforementioned groups is not affected by the freeway volume. Because of this conclusion, all the data for each ramp were grouped together. The results are shown in Figure 9. Since the percentages of vehicles that missed or refused gaps but still entered without stopping were relatively small, both violators and non-violators are shown in two groups only—those that merged smoothly and those that did not meet a gap. The full distinction is made in the column representing the results of all data grouped together.

The vehicles in each group, as a percentage of non-violators and of violators, are shown in Figure 10. Violators, being generally the more aggressive drivers, did not perform much worse than non-violators. However, both groups performed much better than on uncontrolled ramps. The uncontrolled data are based on five studies performed on two of the ramps before control was initiated.

In interpreting the results, it should be borne in mind that these are the results of actual studies and that they do not represent the best performance that can be expected

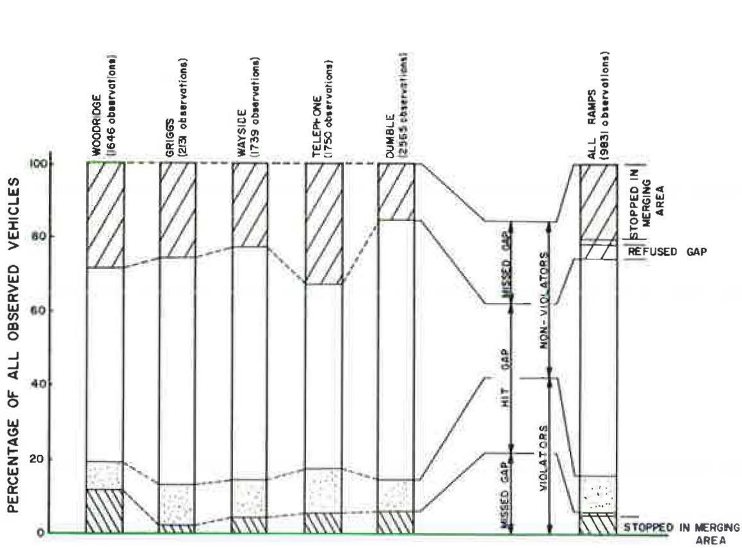


Figure 9. Results of gap-acceptance studies during afternoon control.

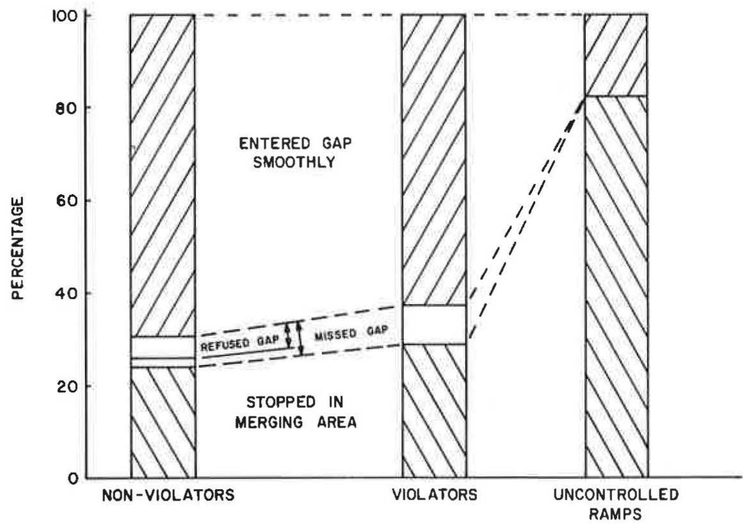


Figure 10. Merging characteristics of non-violators and violators compared with an uncontrolled ramp.

of gap-acceptance control. For example, many vehicles stopped in the merging area actually would have hit their gaps but were forced to stop because the preceding vehicle stopped in the merging area. This happens because the merge detector is located so far from the ramp signal that invariably a second vehicle will be released before the merge override is called in (4). This situation should be practically eliminated by the newer, second-generation controllers. The non-violators hitting gaps plus those refusing gaps give a better indication of the capability of the controllers to match ramp vehicles with gaps. However, a number of vehicles that were actually matched with gaps and refused these gaps were forced to stop before merging. These are therefore included in the latter group. Furthermore, the percentage of drivers violating the ramp signal is quite large on all ramps. With some degree of enforcement, the situation could be substantially improved.

Signal Violations—The foregoing data show a high proportion of violators on all ramps. Signal violations have been a problem in the afternoon control from the start. In six months of control on a day-to-day basis, the percentage of violations has decreased only slightly. For one week, policemen were stationed in the vicinity of the ramp signals. During this week, violations decreased to practically nothing, but returned to the original percentage after the policemen left. During the morning control, when high densities on the freeway make the need for the ramp control much more evident, the proportion of violations is much lower. The following table, based on observations taken during five to eight control periods on each ramp, gives the average percentage of ramp signal violations:

	Woodridge	Wayside	Telephone	Griggs	Dumble
Afternoon control	19.6	14.9	18.1	13.9	15.1
Morning control	5.4	7.7	5.4	8.3	6.0

Ramp Travel Times by Vehicle Types

Where loaded trucks and other heavy or slow vehicles form a substantial part of the ramp traffic, it may be important to design the ramp control system so as to give special consideration to these types of vehicles. Such special consideration may, for example, take the form of allowing for a longer ramp travel time, releasing the vehicle for a bigger gap, and holding the signal in red until the vehicle has cleared the merging area. This may eventually involve the automatic recognition of vehicles with different weight/horsepower ratios. In order to develop special considerations for these vehicles, their relative operating characteristics must be known. One characteristic needed is the ramp travel time from a stopped position at the ramp signal.

Study Procedure—About 100 ramp travel times were measured for each of four different types of vehicles. Because the truck traffic on any of the Gulf Freeway ramps is very light, the data from Woodridge, Dumble, Wayside, and Telephone ramps were grouped together. In each case, the travel time was measured from the instant that the ramp signal turned green until the instant that the vehicle reached a point about 200 feet downstream from the signal. This point is further down the ramp than the merge detector to which the travel times analyzed earlier were measured. It is located approximately at the painted nose of the ramp, whereas the merge detectors are usually located at the physical nose of the ramp.

Results—The data collected were grouped, by class, into frequency tables and their means and variances calculated as given in Table 2. Also shown are the calculated and theoretical chi-square values obtained by fitting the data to normal distributions. The chi-square tests indicate that the travel times for each class of vehicle can statistically be considered as being normally distributed. The results of t- and F-tests in Table 2 lead to the conclusion that different classes of vehicles have significantly different mean travel times and that the variation within one class does not differ significantly from the variation within the adjacent class.

Conclusions—With regard to different vehicle types, it can be concluded that:

1. The ramp travel times of different classes of vehicles are normally distributed within each class;

2. The average ramp travel time of a certain class of vehicles is significantly shorter than the average travel time of any class of larger vehicles; and

TABLE 2

RAMP TRAVEL TIMES FOR DIFFERENT CLASSES OF VEHICLES

Travel Time, Class Midpoint, Seconds	Frequency of Occurrence			
	Compact Cars	Passenger Cars	Single-Unit Trucks	Tractor-Trailers
5.75	3			
6.25	1	1		
6.75	1	3		
7.25	14	9	3	
7.75	8	8	4	
8.25	21	16	5	
8.75	13	10	6	
9.25	17	14	14	2
9.75	7	9	11	4
10.25	7	11	13	10
10.75	1	6	8	3
11.25	1	5	12	12
11.75		3	3	6
12.25			6	13
12.75			1	9
13.25			3	5
13.75				7
14.25				10
14.75				4
15.25				4
15.75				1
16.25				1
Total	94	95	89	91
Mean	8.42	8.98	9.99	12.37
Variance	1.17	1.77	1.96	2.86
χ^2 Calc.	10.30	7.40	7.75	9.11
χ^2 0.05	11.07	14.07	14.07	14.07
t Statistic	3.18 (Compact–Passenger)	5.02 (Passenger–Single-Unit)	10.29 (Single-Unit–Tractor)	
F Statistic	1.51 (Compact–Passenger)	1.11 (Passenger–Single-Unit)	1.46 (Single-Unit–Tractor)	

TABLE 3

SUMMARY OF DATA ON STOPPING POSITION AT RAMP SIGNAL

Statistic	Ramp Location		
	Telephone	Dumble	Mossrose
Lower limit	-15	-16	-16
Upper limit	+5	+4	+4
Mean	-2.68	-5.10	-2.81
Standard deviation	4.01	4.83	3.91
Sample size	488	297	430

3. The variations among travel times of vehicles of a certain class do not differ significantly from the variation among travel times of vehicles of the next larger or smaller class; only the variation among tractor-trailer combinations differs significantly from the variation in ramp travel times of vehicles in general.

Stopping Distance From Ramp Signal

The position with respect to the ramp signal at which drivers stop their vehicles on a controlled ramp is of primary importance in locating the check-in detector. The detector should be so located that all vehicles stopping at the signal will actuate it. A single vehicle failing to actuate the check-in detector can cause considerable problems, because the signal indication will remain red.

Study Procedure—In studying the behavior of vehicles at the ramp signal, video tapes were made using a portable television camera mounted on a tripod placed across the frontage road from the signal, well-hidden from the view of ramp drivers. The tapes were then played back over a television monitor and the stopped position of vehicles read off by means of marks on an acetate overlay placed on the face of the monitor.

Data were collected at the Telephone and Dumble entrance ramps during two peak hours each and at the Mossrose entrance ramp during one peak hour. The distance that the front wheel of a vehicle stopped from the signal was measured to the nearest foot, with distances in front of the signal designated as negative and distances past the signal as positive.

Results—A summary of the results of the experiments is given in Table 3. The data on two different days at both Telephone and Dumble were quite similar. The Kolmogorov-Smirnoff goodness of fit test indicates that the observed differences between different days are simply due to chance.

The total data sample at each ramp is shown in Figure 11; as can be seen, the distributions at Telephone and Mossrose are almost identical. However, at Dumble, drivers seem generally to stop 2 to 3 feet earlier than at the other two ramps. Note also that the stop-line 5 feet in front of the signal at Telephone seems to have no effect as compared to the behavior at Mossrose.

Application—The sum of the two days' data for the Telephone ramp is plotted in Figure 12. The frequency distribution is positively skewed with a finite upper limit, reflecting the influence of the signal. Superimposed on the figure are scale drawings of a compact vehicle in the two critical positions. Since all vehicles are to be detected,

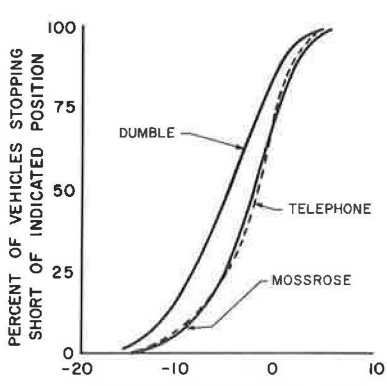


Figure 11. Stopping positions of vehicles at ramp signal.

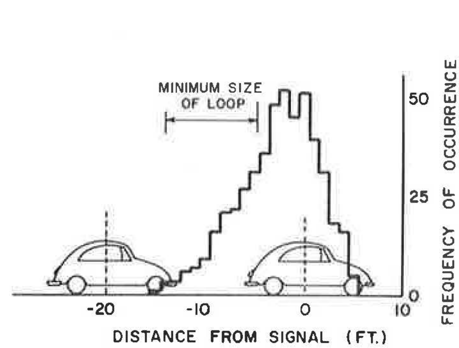


Figure 12. Frequency distribution of stopping distance between ramp signal and the front wheel at the Telephone entrance ramp.

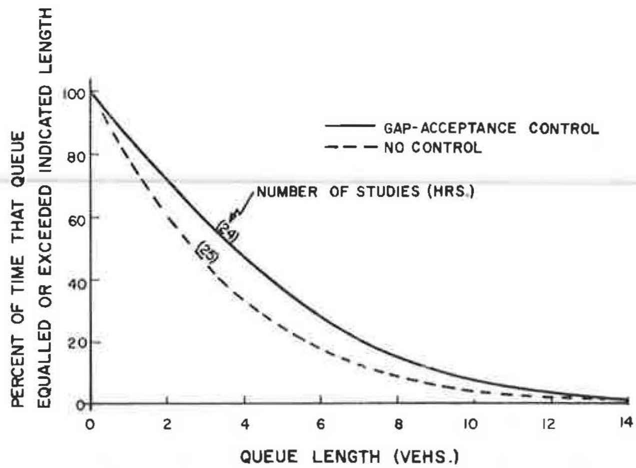


Figure 13. Number of vehicles between back of queue and merge point.

the location and size of the check-in loop should be based on the smallest vehicle stopped at the two extremes of the frequency distribution. By allowing a 1-foot overhang into the loop in either position, it is clear that, based on the Telephone data, the minimum length of the loop should be 9 feet and the loop should be located no further than 5 feet from the signal. Based on the Dumble data, the loop should be of the same length, but the trailing edge would be 6 feet from the signal. When all three ramps studied are considered together, a loop 10 feet long, placed 5 feet from the signal, would be actuated by all the vehicles. Experience indicates that a loop width of 6 feet is satisfactory.

Queuing Characteristics

It can generally be expected that longer queues will form on a controlled ramp than on an uncontrolled ramp and that drivers will suffer longer delays. To investigate these aspects, studies were made of the operation on ramps with and without control. In these studies, observers watching the closed-circuit television monitors on the Gulf Freeway used push-buttons to signal the digital computer when each ramp vehicle (a)

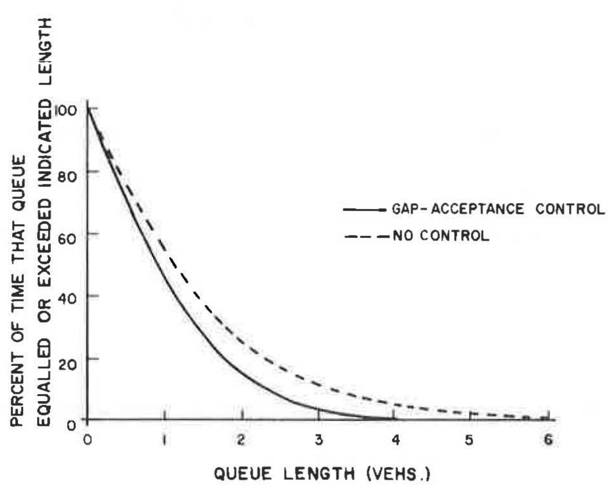


Figure 14. Number of vehicles between ramp nose and merge point.

TABLE 4

AVERAGE QUEUE LENGTHS

Location	In Ramp System		On Acceleration Lane	
	Uncontrolled	Controlled	Uncontrolled	Controlled
Woodridge	3.53	3.76	1.46	0.68
Griggs	3.86	4.93	1.21	0.57
Wayside	1.74	4.00	0.67	0.68
Telephone	2.33	3.44	0.81	0.64
Dumble	2.49	4.41	0.73	0.69

arrived in the ramp queue, (b) passed the ramp signal, (c) passed the physical nose of the ramp, and (d) merged into the traffic stream. From this information, the computer calculated the travel time distributions in each of the subsystems as well as the percent of time that a given number of vehicles were in each subsystem. In this manner, five of the inbound ramps were studied during the afternoon control as well as before control was initiated on May 8, 1968. Each study period lasted for four 15-minute periods, usually from about 4:15 p.m. to 5:15 p.m. Altogether, 25 studies were carried out before and 30 studies after control was initiated. Observations were taken on more than 15,000 ramp vehicles.

Queue Lengths—For each 15-minute period of each study, the number of seconds during which there were 0, 1, 2, etc., vehicles between the back of the queue and the merge point and between the ramp nose and the merge point were tabulated. The data showed a large variability from one 15-minute period to another and from one day to another. The averages of a number of representative studies are shown in Figure 13. The graphs show, for all ramps grouped together, the percent of time during which there were a given number, or more, vehicles between the back of the queue and the merge point. Also shown are the number of 1-hour study periods on which each curve is based. As can be seen, queues were generally longer during control.

One of the advantages of entrance ramp control is the prevention of long queues forming on the acceleration lane, thus adversely influencing freeway traffic. This is illustrated by the results shown in Figure 14. These are again averages calculated over a number of representative study periods and over all ramps studied. Each graph shows the percent of time during which there were a certain number, or more, vehicles between the ramp nose and the merge point. The average queue lengths derived from all the data for each ramp grouped together are given in Table 4. These averages were derived by calculating the total travel time (vehicle seconds) in the subsystem divided by the total time of the studies. Because of the variability in the data, no relationship between average queue lengths, ramp volumes and freeway volumes could be established.

Ramp Delays—For each 15-minute period of each study, the frequency distribution of travel times (or delays) in each of the four subsystems was tabulated. By grouping all the data for all ramps together, the cumulative frequencies of delays in the overall ramp system (back of queue to merge) were determined; these are shown in Figure 15 for both controlled and uncontrolled conditions. As may be expected, ramp delays under control are generally longer than under no control.

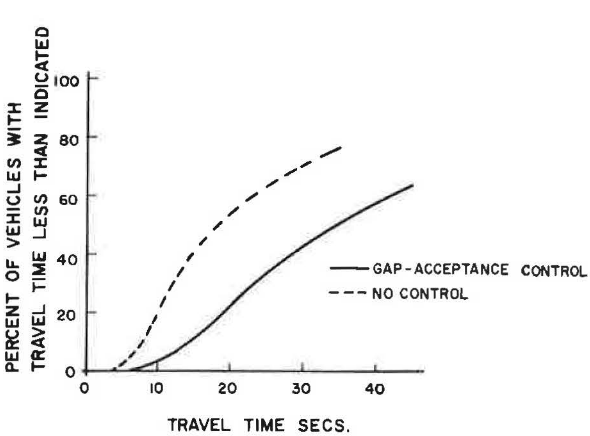


Figure 15. Travel times in ramp system (back of queue to merge point).

TABLE 5

AVERAGE DELAYS TO RAMP VEHICLES (SECONDS)

Location	In Ramp System		On Acceleration Lane	
	Uncontrolled	Controlled	Uncontrolled	Controlled
Woodridge	32.9	46.5	13.6	8.3
Griggs	25.5	53.9	8.0	6.2
Wayside	19.8	56.6	7.7	7.6
Telephone	29.5	42.6	10.3	7.0
Dumble	20.3	42.2	6.0	5.7

However, the delays suffered in the merging area are generally shorter during control than during no control, as evidenced by Figure 16. The average delays are shown in Table 5. These averages were derived by calculating the total travel time (vehicle seconds) in each subsystem, divided by the total number of vehicles.

FREEWAY TRAFFIC CHARACTERISTICS

All freeway traffic characteristics are, of course, of interest in freeway surveillance and control. Some of these characteristics directly influence the design of a merging control system or are influenced by the control system. In order to study some of these characteristics, data were collected with three pairs of 6- by 6-foot loop detectors, placed on the outside lane of the Gulf Freeway, upstream from the Dumble inbound entrance ramp. This three-lane freeway section is level and straight. The loop pairs were installed at the ramp nose, at 520 feet and at 770 feet upstream from the nose. The leading edges of each loop pair were 18 feet apart. The data were collected with computer by recording the time, to the nearest hundredth of a second, at which a vehicle actuated each detector, and, by keeping track of separate vehicles (with the aid of an operator using the closed-circuit television system), gap sizes, speeds and travel times were calculated. These data were then further analyzed to study the characteristics discussed in the following.

Speed Stability

The changes that occur in vehicle speeds between the gap detector and the merge point can be defined as speed stability and can be quantitatively described by the mean change, the variance of the changes, and the distribution of the changes and should be related to the distance over which the changes occur. This speed-stability characteristic is of importance in gap-acceptance control, because it reflects the ability and consistency with which a merge controller can predict the rate of progress of a gap as it approaches the merging area. Speed stability was investigated by measuring the speed of the same vehicle at the three locations upstream from the merge area.

Some of the results are shown in Figure 17, where the speed at the nose is plotted against the speed of the corresponding vehicle at the first measuring station, 770 feet upstream. Also shown is the line for maximum stability. Points that fall on this line indicate no change in speed. It appears that speeds are generally lower at the nose than at the upstream point, especially so at low speeds. This may be the effect of the entrance

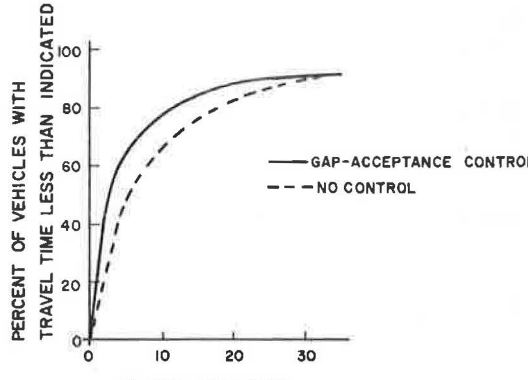


Figure 16. Travel times on acceleration lane (nose to merge point).

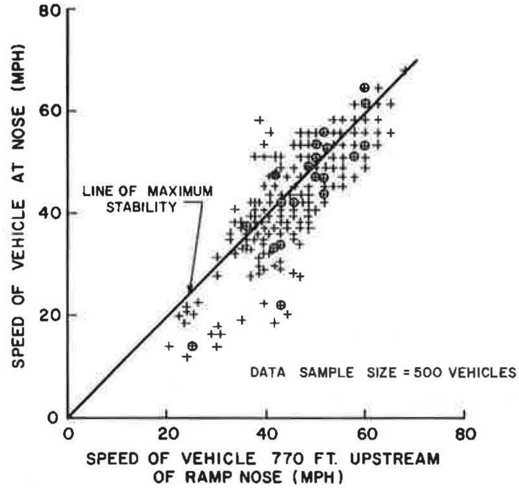


Figure 17. Speed of freeway vehicle 770 feet upstream from nose vs speed of the same vehicle at the nose.

ramp. Although some of the data points are slightly more scattered at around 40 mph than at other speeds, it is apparent that the variances do not increase with the speed. Some of the data cells are circled to indicate that these cells contain trucks, which are in this case defined as vehicles longer than 30 feet. It is apparent that the speed-stability characteristic of trucks is no different from that of passenger vehicles.

Figures 18 and 19 show the difference in upstream and downstream speeds plotted against the upstream speed, as well as the distributions of these differences. Most speeds do not change by more than 10 mph, and the distance apparently has no influence on the variance of the speed changes.

In summary, it can be said that, at the location studied, the average freeway driver reduced his speed slightly as he approached the nose. This speed reduction does not appear to be greater at high speeds than at low speeds. Most of the speed reduction appears to take place several hundred feet upstream from the nose. Since data were collected at only two points upstream from the nose, a reliable relationship between distance from the nose and average speed reduction could not be established. The variability in speed changes among drivers does not seem to be greater at high speeds than at low speeds and does not seem to be influenced by the distance studied. The speed-stability characteristics described may well be different at different geometric configurations and may be influenced by ramp demand.

Speed and Travel Time

In order to meter vehicles from a ramp into gaps on the freeway, the gaps necessarily have to be measured some distance upstream from the merging area. The travel time of this gap to the merge point then has to be estimated, using some speed measurement, in order to determine the exact moment at which to turn the ramp signal green. Since it has been shown that, on the average, speeds change somewhat between an upstream point and the nose of the ramp, it is of interest to investigate how this affects travel times and how the measured speed can best be used to estimate the travel time.

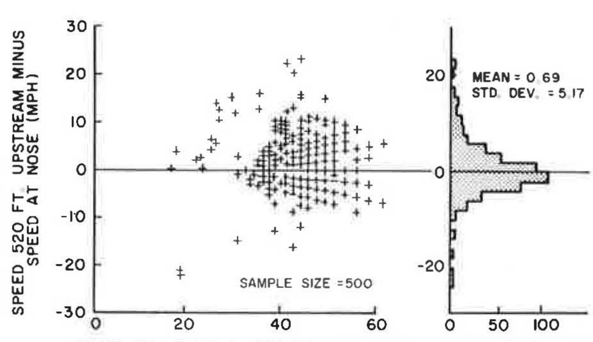


Figure 18. Speed difference between ramp nose and point 520 feet upstream.

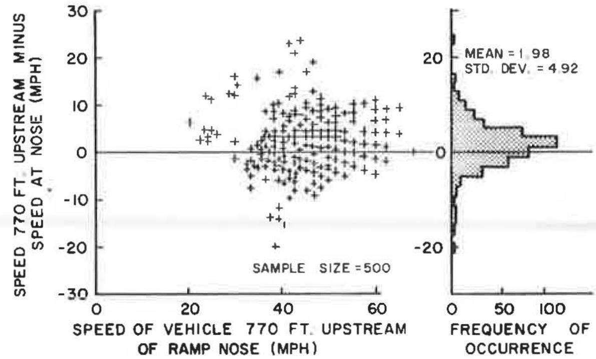


Figure 19. Speed difference between ramp nose and a point 770 feet upstream.

Figure 20 shows the measured travel times plotted against the travel times estimated from the speeds measured 770 feet from the nose. Also shown is the line of no change. As expected, the actual travel times are generally higher than the calculated travel times. The higher travel times (low speeds) are generally more variable than the low travel times (high speeds), but this is to be expected since the same speed change would make a bigger difference in the travel time at low speeds than at high speeds. The data collected at 520 feet show essentially similar characteristics.

The fact that actual travel times were generally lower than calculated travel times can be compensated for by the control system. What cannot be compensated for is the variability of this difference. It is, therefore, of interest to investigate the distribution of the difference between actual travel times and travel time calculated from the measured speeds. This distribution appears normal with a mean of 0.81 seconds and a standard deviation of 0.88 seconds, but does not provide a statistically significant fit with a normal distribution, mainly because of the high modal peak. The variability in these differences would introduce an error into the accuracy with which the appearance of a gap in the merging area can be predicted. This error is such that almost 80 percent of the travel times could have been estimated to within $1\frac{1}{2}$ seconds. In a control system, the speeds at the gap detector can be measured much more accurately than in this experiment, so that a better accuracy in predicting travel times can be obtained, if so desired.

The data measured at 520 feet showed a mean difference of 0.44 seconds with a standard deviation of 1.04 seconds. From the data collected, it does not appear as if the variation in the difference increases with distance.

Gap Stability

The gap acceptance controller should estimate not only the time of arrival of a gap in the merging area, but also the size of the gap in the merging area based on its size as measured upstream. This estimation needs to take gap stability into account. In an earlier work on gap stability, models were developed and fitted to data collected at a particular site and the application to freeway control was

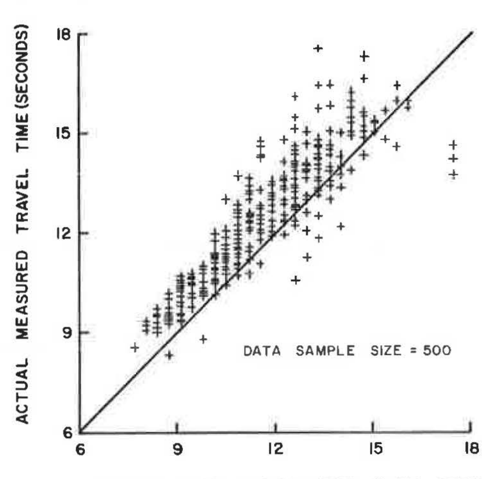


Figure 20. Comparison of actual and estimated travel times over 770 feet.

discussed (8, 9). It is, therefore, mainly of interest to compare the stability of gaps found in this study to that found earlier, using the same models.

The model for the mean of the gap stability distribution is

$$T_d = b_1 G + b_2 U + b_3 GU$$

where

G = gap size measured at detector placed d feet upstream of ramp nose;
T_d = size of gap G at nose;
U = speed measured at gap detector in miles per hour; and
b_i = regression coefficients.

By applying this model to the data obtained in this experiment, the results given in the following table were found:

d	b_1	b_2	b_3
520	1.2942	-.00073	-.00619
770	1.1556	-.00616	-.00375

This model, for 770 feet, is plotted in Figure 21 for speeds of 20 mph and 70 mph, together with the data. Also shown on the figure (dashed lines) are the results found earlier by Buhr, extrapolated to the same distance (8, 9). The gaps studied in this experiment changed much less, on the average, than those in the earlier study, which was performed at a different location. However, the general trend is similar in that both studies indicated that, on the average, gaps tend to increase at low speeds and to decrease at high speeds.

The variation in the changes in gap size found in this study is shown in Figure 22, together with that found earlier. Again it is evident that the trend is the same, but that the gaps were much more stable than those studied earlier. By splitting the data into speed groups it was found, as earlier, that the variation in the changes in gaps did not change with speed—i.e., gaps are not more stable at one speed than at another.

When control is exercised with a digital computer, gap stability can be taken into account in the gap-projection process. This is done by projecting vehicles rather than gaps, while the computer takes the probable change in gap size into account (10). If, for example, the lead vehicle of a gap is slower than the lag vehicle, this procedure will reduce the gap size as indicated by the relative speed.

To determine the effectiveness of this procedure, the data collected in this experiment were used as follows: From the speed of each of two consecutive vehicles, the travel time of each vehicle to the ramp nose was calculated; the difference in the calculated travel times of two consecutive vehicles was added to the gap between them as measured at the gap detector (this would be the gap size estimated by the computer for control); this gap size was then subtracted from the actual gap size measured at the nose, yielding a set of differences between actual and estimated gap sizes. The average difference was 0.01 seconds,

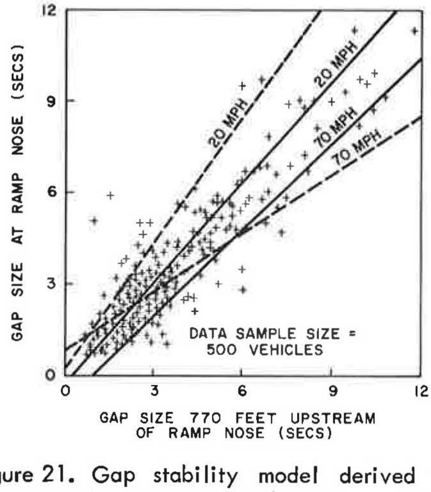


Figure 21. Gap stability model derived from present data compared with earlier model by Buhr (dashed lines).

and the standard deviation was 0.64 seconds. This represents a tremendous improvement over the actual stability, as shown earlier in Figure 22. As a matter of interest, the differences provided a statistically significant fit with a normal distribution having the same mean and variance.

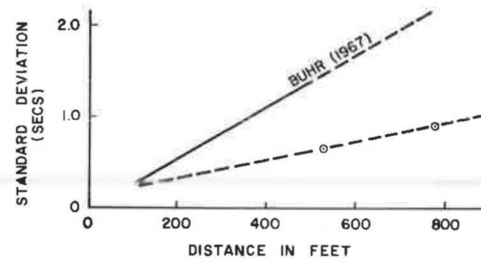


Figure 22. Standard deviation of gap stability distribution.

Total Travel and Total Travel Time

The objective of freeway control is often misstated as that of relieving freeway congestion or maximizing level of service. In actual fact, the objective is to optimize freeway operation or level of service. Just what this optimum is, is still open to question. Some authorities on the subject believe that optimum operation is reached when the flow smoothness is maximized (acceleration noise minimized). Others find strong arguments in favor of maximizing throughput—i.e., moving the maximum amount of traffic in the minimum amount of time. However, whatever the criteria, it is necessary to have certain measures of effectiveness that can be evaluated in order to establish how well a certain freeway is operating or how well a certain control scheme is functioning. It is also necessary to know, or be able to establish, the optimum values of these measures of effectiveness.

The volume count, for example, is an excellent measure of effectiveness. Most traffic engineers have a good feel for the optimum volume count. Volume is, however, a point measure and does not necessarily reflect the operation in the entire system. Two measures that have often been used in the past are total travel (vehicle-miles) and total travel time (vehicle-hours).

The measure of total travel time has often been used with the implication that the system is improved if the total travel time is reduced (11). This is usually true as far as a congested freeway is concerned but need not necessarily always be true. To use a somewhat facetious example, by closing the freeway down completely, the total travel time can be reduced to zero.

Total travel time is related to total travel in somewhat the same way that density is related to volume. This relationship is illustrated in Figure 23 for a 2½-mile section of the Gulf Freeway. The data were collected for a 1-hour period, during both morning and afternoon control. If the freeway is congested, operation will be on the right-hand side of the curve. Under these conditions, it is true that the operation can be

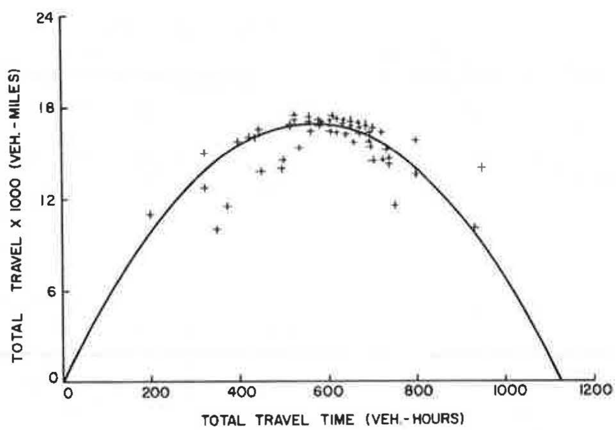


Figure 23. Operating characteristic for a 2½-mile section of the Gulf Freeway control system.

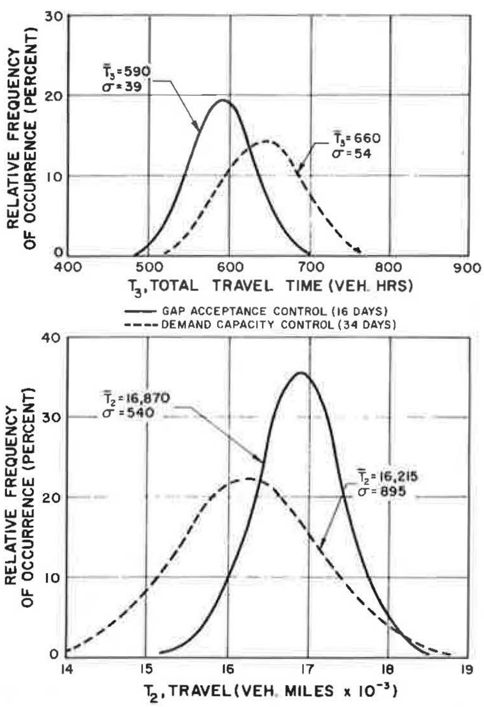


Figure 24. Operating characteristics on a 2½-mile section of the Gulf Freeway during the morning peak hour control (7-8 a.m.).

improved by decreasing total travel time, provided that total travel also increases. Under free and stable flow conditions, operation will be on the left-hand side of the curve. In order to optimize operation under these conditions, total travel time has to be increased, provided that total travel also increases. Since the problem is actually a lack of demand, however, a freeway control system cannot force operation toward an optimum under these conditions. In some respects, therefore, this relationship can be considered as a warrant for control, because the operation can be optimized if it falls on the right-hand portion of the curve but not on the left. However, the secondary objective of freeway control is to improve merging operations. This aspect is not reflected by the relationship shown in Figure 23. It is thought that this type of relationship will lead to warrants for the installation of a control system, but not necessarily to warrants for control, the difference being that the installation of a control system may be justified by the morning peak (or afternoon peak), while control of the other peak period may be justified by the improvements in merging operation, since the system then already exists.

In order to evaluate the effectiveness of two of the control systems presently installed on the Gulf Freeway, the total travel and total travel times were accumulated for 1 hour during the morning under demand-capacity control and under gap-acceptance control. Demand-capacity control was exercised on all eight inbound ramps by analog controllers. Gap acceptance control was exercised by the digital computer on six of the inbound ramps while the remaining two ramps were under demand-capacity analog control. The results from 34 days of demand-capacity control and 16 days of gap-acceptance control (these days were selected because of the absence of accidents or incidents) were classified into frequency distributions and normal distributions fitted to them. The resulting normal distributions are shown in Figure 24. The total travel times fit the normal distributions quite well, but the total travel was not normally distributed. The distributions in this case are used mainly for illustrative purposes. From the figures, it is clear that gap-acceptance control not only provided better operation, but provided consistently better operation as evidenced by the reduced standard deviations. When the coordinates of the mean values from Figure 24 are compared with the operating characteristic of Figure 23, it can be seen that the gap-acceptance control results in operation very close to the optimum, while the demand-capacity control results in operation to the right (congested) of optimum.

APPLICATION TO FREEWAY CONTROL

The application of the characteristics derived in this paper to the design and operation of a freeway control system has been discussed under each heading. These are by no means all the traffic information required to design, evaluate, and calibrate a control system. These characteristics merely begin to fill the void left by the rapid expansion of freeway control technology. It is hoped that the characteristics presented in this paper will be of some help to the traffic engineer charged with responsibility in this field and that they will prompt further investigations.

REFERENCES

1. Matson, T. M., Smith, W. S., and Hurd, F. W. Traffic Engineering. McGraw Hill, New York, 1955.
2. Traffic Engineering Handbook, Third Edition, 1965.
3. Knox, D. W. Merging and Weaving Operations in Traffic. Australian Road Research, Vol. 2, No. 2, Dec. 1964.
4. Buhr, J. H. McCasland, W. R., Carvell, J. D., and Drew, D. R. Design of Freeway Entrance Ramp Merging Control Systems. Research Report 504-3, Texas Transportation Institute, 1968.
5. Drew, D. R., Brewer, K. A., Buhr, J. H., and Whitson, R. H. Multilevel Approach to the Design of a Freeway Control System. Paper presented at the 48th Annual Meeting and included in this Record.
6. Drew, D. R., LaMotte, L. R., Wattleworth, J. A., and Buhr, J. H. Gap Acceptance in the Freeway Merging Process. Highway Research Record 208, p. 1-36, 1967.
7. Brewer, K. A., Buhr, J. H., Drew, D. R., and Messer, C. J. Ramp Capacity and Service Volume as Related to Freeway Control. Paper presented at 48th Annual Meeting and included in this Record.
8. Buhr, J. H. Freeway Entrance Ramp Merging Control Systems. Dissertation, Texas A&M University, Jan. 1967.
9. Buhr, J. H. Gap Stability and Its Application to Freeway Merging Control Systems. Traffic Engineering, March 1968.
10. Buhr, J. H., Drew, D. R., Gay, J. N., and Whitson, R. H. Design Specifications for Freeway Surveillance and Control Systems. Research Report RF 504-7, Texas Transportation Institute, 1968.
11. Wattleworth, J. A., Courage, K. G., and Carvell, J. D. An Evaluation of Two Types of Freeway Control Systems. Research Report 488-6, Texas Transportation Institute, 1968.

Some Electronic Measurements of Macroscopic Traffic Characteristics on a Multilane Freeway

KENNETH G. COURAGE, Senior Transportation Engineer, Kelly Scientific Corporation

This report deals with the measurement of the commonly sought characteristics of macroscopic traffic flow by means of automatic sensing equipment using a high-speed digital computer for data collection and analysis. Independent measurements of speed, density, flow, and kinetic energy are obtained from ultrasonic presence detectors and from an optical speed trap. The results are combined to produce composite measurements where possible. The ability of each measurement technique to produce reasonable values is analyzed by linear regression, and the conformance of the data to accepted traffic flow theories is discussed. Finally, some thoughts on the application of these results are explored.

•TRAFFIC operational studies usually require some form of data input. In general, the more complex the study, the greater the need for automation in the collection and processing of data. The digital computer has solved most of the processing problems; however, the techniques of collecting much of the required information are still somewhat primitive. The measurement system described in this report is indeed primitive; nevertheless, it does offer some potential in traffic measurements.

The traditional macroscopic descriptors of traffic movement are speed, flow rate, and density. These variables are related by definition such that any one quantity may be calculated in terms of the other two. This relationship is

$$Q = KU$$

where

Q = the number of vehicles passing a given point on the freeway per unit of time,
K = the equivalent density of the traffic stream, and
U = the space mean speed of the traffic stream.

Thus, to obtain a complete picture of the traffic stream, it is necessary to measure flow and speed, flow and density, or density and speed. In this section two measurement techniques will be considered. The first employs "off-the-shelf" presence detectors and involves measurement of flow and density (estimated from detector occupancy) to arrive at a calculated speed. The second uses an optical speed trap from which measurements of speed and density (also estimated from occupancy) are obtained. In this case the calculated variable is flow.

Optical speed traps have been used successfully in the past in single-lane traffic flow measurements. Speeds and densities have been measured in the Holland Tunnel and used as a basis for evaluation of operation and control of access (1). A similar system has also been employed in the calibration of a speed trap using two magnetic loop detectors (2).

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 48th Annual Meeting.

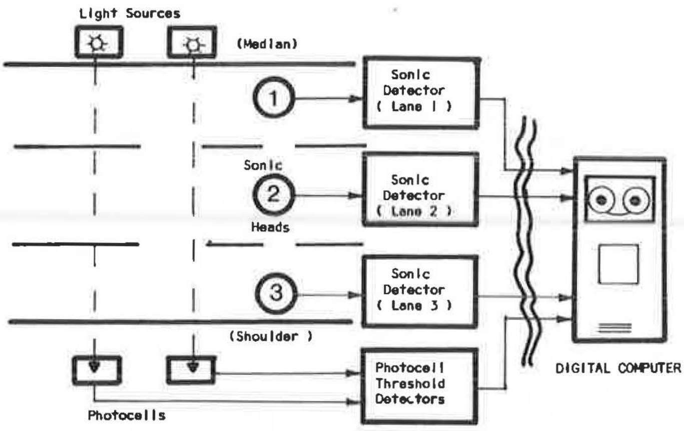


Figure 1. Layout of measurement system.

In approaching this problem, the operation of the measurement station will first be considered. This station will comprise one optical speed trap and three presence detectors (one per lane) of the ultrasonic type. The potential applications in the measurement of a number of parameters will then be examined and some experimental results will be presented. A more detailed analysis of these results will follow, and, finally, some surveillance and control applications will be discussed.

THE MEASUREMENT SYSTEM

The measurement system used in the collection of data is shown in Figure 1. The sensors were installed on the John C. Lodge Freeway at the Calvert Avenue overpass.

Figure 2 shows the time relationships of the relay closure signals generated by the interruption of the two beams. The following terminology is established as a basis for further discussions:

S = spacing between upstream and downstream detectors (ft);
T_i = occupancy time for i th platoon on upstream detector (sec);
T'_i = occupancy time for i th platoon downstream detector (sec);
α_i = travel time between detectors for leading edge of platoon i (sec);
β_i = travel time between detectors for trailing edge of platoon i (sec); and
G_i = length of gap following T_i (sec).

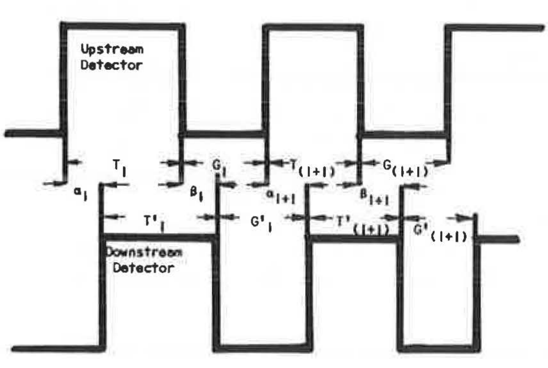


Figure 2. Time relationships for optical speed traps.

These definitions apply to individual measurements on the i th platoon. For the purposes of this analysis, each separate interruption of the beam by one or more vehicles is considered to constitute a distinct platoon.

The system under consideration will provide independent measurements of all three of the macroscopic variables, and will afford a means of validating a number of calculations based on the known relationships between these variables.

Some further definitions describing the sample measurements are required:

M (min) = sampling period;
U_s (mph) = speed as measured by sonic detectors;

times the square of the velocity" (4). In this case, the mass is proportional to the density of the stream and, therefore, the energy may be estimated using the relationship

$$E = KU^2$$

or, in terms of the measurement system under consideration,

$$E_C = Q_S U_O$$

Independent energy calculations may be obtained by using detector occupancy as an indicator of density. Separate estimates may be calculated from the optic and sonic detectors, and the appropriate calculations are

$$E_S = \frac{{Q_S}^2}{O_S} \text{ (sonic)}$$

and

$$E_O = {U_O}^2 \cdot O_O \text{ (optic)}$$

Acceleration Measurements

Acceleration is somewhat more difficult to calculate, for two reasons. First, extreme accuracy is required in the measurement of occupancy and elapsed times because even the largest practical values of acceleration would only produce minute differences in speed over a reasonable length of speed trap (5). Second, the computed values would only be valid for platoons containing a single vehicle because the basis for computation would be the change in speed between the leading and trailing edges of the platoon. Because of this second problem, some additional logic would be required to eliminate multiple-vehicle platoons and, of course, this would reduce the sample size considerably, particularly during periods of congested operation. For these reasons, acceleration measurements are considered beyond the capability of the system under consideration.

PLATOON CHARACTERISTICS

Prior to evaluation of the various measurements obtained from the system, it was necessary to examine the platoon characteristics observed under high densities to ensure that the multiple-lane operation would not seriously reduce the sample sizes involved in the speed calculations. It was observed that, under capacity flow conditions in the order of 100 vehicles per minute (three lanes), platoon counts ranging from 60 to 80 per minute were obtained. Under heavily congested operation with average speeds close to 10 mph, the flow rate dropped to approximately 60 vehicles per minute and the platoon counts were reduced to about 30 per minute. The standard deviation of the individual speeds was fairly consistent at about 5 mph, giving a 95 percent statistical confidence interval of approximately ±0.5 mph at capacity flow. This figure increased to about ±1 mph under severe congestion.

OBSERVATIONS

The data obtained under two separate days of operation (one "wet" and one "dry") are shown in Figure 3 in the form of an energy-momentum comparison. The actual measured values of speed and flow were used in this comparison as opposed to the calculated values (i.e., U_O and Q_S as opposed to U_S and Q_O). The energy value was calculated as the product of speed and flow.

It is noted that the observed relationships between the variables agree closely with the theoretical relationships developed in more detail by Drew and Keese (4).

U_O (mph) = speed as measured by optic detectors;
Q_S (veh/min) = flow as measured by sonic detectors;
Q_O (veh/min) = flow as measured by optic detectors;
O_S (percent) = detector occupancy as measured by sonic detectors;
O_O (percent) = detector occupancy as measured by optic detectors;
E_S (veh-mi/hr^2) = kinetic energy as measured by sonic detectors;
E_O (veh-mi/hr^2) = kinetic energy as measured by optic detectors;
E_C (veh-mi/hr^2) = kinetic energy as calculated by composite measurements; and
K (veh/mi) = density as calculated by composite measurements.

Flow Measurements

Direct measurement of flow (Q_S) may be obtained by summing the output pulses received from the presence detectors over the sampling period. An independent estimate of flow (Q_O) may be obtained from the speed trap by calculation, using the product of the measured speed by the density (estimated from occupancy).

Speed Measurements

The speed trap formed by the two beams will produce accurate speed measurements. Individual speeds may be calculated as

u_{Li} = speed of leading edge = S/α_i ft/sec; and
u_{Ti} = speed of trailing edge = S/β_i ft/sec.

Mean speeds over a given sampling period are more likely to be required in connection with the operation of a surveillance system than individual speeds. Again, two independent estimates are possible. The first is a direct measurement of space mean speed from the speed trap (U_O) where

$$U_O = \frac{QS}{\sum_{i=1}^{Q} \alpha_i}$$

The second is calculated from the presence detectors using the relationship

$$U_S = \frac{Q_S}{O_S}$$

Density Measurements

In the case of density measurements, three independent estimates are possible. The detector occupancy values obtained from both the optic and sonic sensors (O_O and O_S, respectively) can provide reasonable indications of density. A further measurement, K, may be obtained, using independent values of speed and flow, from the relationship

$$K = \frac{60\,Q_S}{MU_O} \text{ (vehicles per mile)}$$

It is important to note that, although density is expressed in terms of vehicles per mile, these dimensions represent an equivalent "point" density over a finite time interval as opposed to an instantaneous density over a finite space interval. This concept is explained in greater detail by Edie (3).

Energy Measurements

Using the analogy between traffic flow and hydrodynamic flow, the energy or "dynamic pressure" of the traffic stream may be expressed in terms of the classical "mass

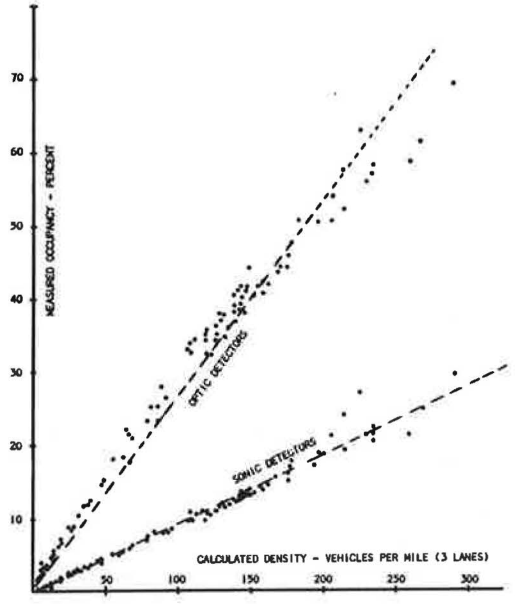


Figure 4. Measured occupancy vs calculated density for optic and sonic detectors.

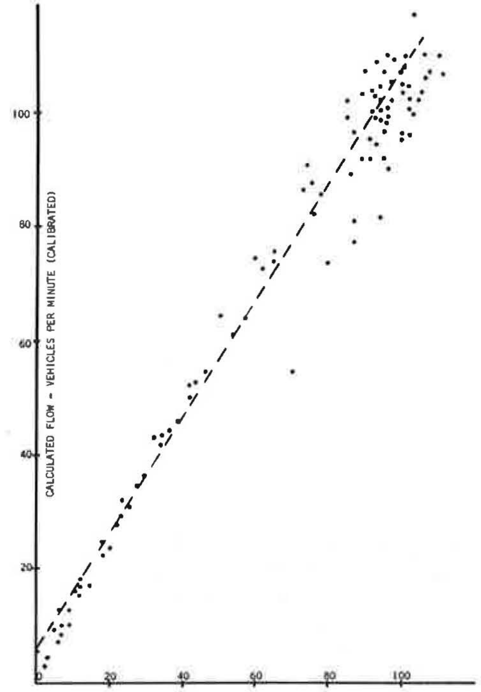


Figure 5. Calculated flow (optic detectors) vs measured flow (sonic detectors).

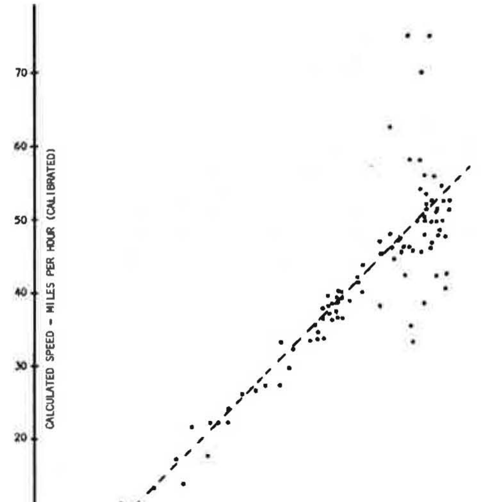


Figure 6. Calculated speed (sonic detectors) vs measured speed (optic detectors).

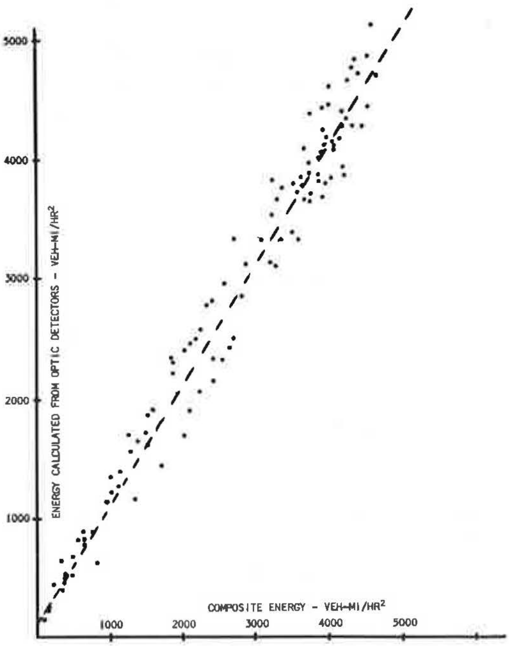


Figure 7. Energy calculated from composite measurements vs energy calculated from optic detectors.

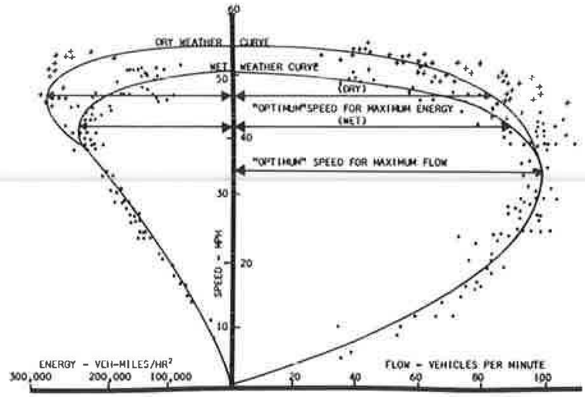


Figure 3. Energy-momentum comparison for wet and dry pavement.

The four measured quantities (U_O, Q_S, O_O, and O_S) provide the basis for calculation of the remainder of the ten quantities previously described. A summary showing the manner in which each quantity was derived is given in Table 1.

A number of these values involve independent measurements of the same quantity, and direct comparison should therefore provide an evaluation of these measurements. Linear regression analysis was performed on six sets of measurements expecting, of course, a straight line emanating from the origin. The results of these analyses are given in Table 2 and Figures 4 to 8, inclusive. Table 2 indicates, for each comparison, the intercept and correlation index, r^2. The figures show the actual data points along with the regression line.

It is observed that excellent correlation is obtained in all cases except one, optic speed vs sonic speed. The reason for the lack of agreement is apparent in Figure 6, in which a wide range of sonic speeds is obtained over a very narrow range of optic speeds at the upper limit. There are two reasons for this. First, lower flow rates are obtained in the upper speed ranges and the reduced sample sizes do not "average out" as well. Second, the sonic units employ distinct pulses (20 per second) in the detection of vehicle presence and, therefore, a fixed scanning error is introduced, which becomes clearly more significant as speed increases. This problem is treated in detail by Weinberg and Deleys (6).

No attempt was made in this case to force the linear relationships through the origin, since the primary interest lies in

TABLE 1

DETERMINATION OF INDIVIDUAL TRAFFIC CHARACTERISTICS

Quantity	Significance	Methods of Determination
U_o	Speed from optics	Measured
Q_s	Flow from sonics	Measured
O_o	Occupancy from optics	Measured
O_s	Occupancy from sonics	Measured
U_s	Speed from sonics	Q_s/O_s
Q_o	Flow from optics	$U_o \cdot O_o$
K	Calculated density	Q_s/U_o
E_o	Energy from optics	$O_o \cdot U_o^2$
E_s	Energy from sonics	Q_s^2/O_s
E_c	Composite energy	$Q_s \cdot U_o$

TABLE 2

COMPARISON OF MEASURED AND CALCULATED VALUES

Comparison	Intercept (% of upper limit)	Correlation Index (r^2)
K vs O_o	4.3	.976
K vs O_s	-0.6	.980
Q_s vs Q_o	5.0	.972
U_o vs U_s	-4.8	.640
E_c vs E_o	3.2	.972
E_c vs E_s	-0.4	.990

$$E = K\left(U_f^2 - \frac{2\,KU_f^2}{K_j} + \frac{K^2U_f^2}{K_j^2}\right)$$

$$= KU_f^2 - \frac{2\,K^2U_f^2}{K_j} + \frac{K^3U_f^2}{K_j^2}$$

$$= U_f^2\left(K - \frac{2\,K^2}{K_j} + \frac{K^3}{K_j^2}\right)$$

To maximize E

$$\frac{dE}{dK} = 1 - \frac{4K}{K_j} + \frac{3K^2}{K_j^2} = 0$$

Using the quadratic formula, the roots will be located at

$$E = \frac{\dfrac{4}{K_j} \pm \sqrt{\dfrac{16}{K_j^2} - \dfrac{12}{K_j^2}}}{\dfrac{6}{K_j^2}}$$

$$E + \frac{\dfrac{4}{K_j} \pm \sqrt{\dfrac{4}{K_j^2}}}{\dfrac{6}{K_j^2}} = \frac{\dfrac{4}{K_j} \pm \dfrac{2}{K_j}}{\dfrac{6}{K_j^2}} = \frac{6}{K_j} \cdot \frac{K_j^2}{6} = K_j \text{ (trivial)}$$

and

$$\frac{2}{K_j} \cdot \frac{K_j^2}{6} = \frac{K_j}{3}$$

which we will define as K'_m. The corresponding speed value will be $U'_m = \frac{2}{3}\,U_f$.

Thus, the maximum energy is obtained at $U'_m = \frac{4}{3}\,U_m$ and $K'_m = \frac{2}{3}\,K_m$. Its maximum normalized value is

$$E = \tfrac{2}{3}\,K_m \cdot \left(\tfrac{4}{3}\,U_m\right)^2$$

$$= \tfrac{27}{32}\,K_mU_m^2$$

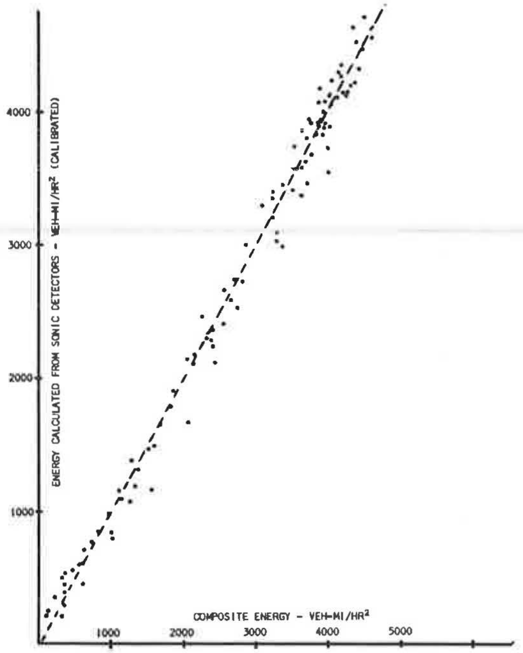


Figure 8. Energy calculated from composite measurements vs energy calculated from sonic detectors.

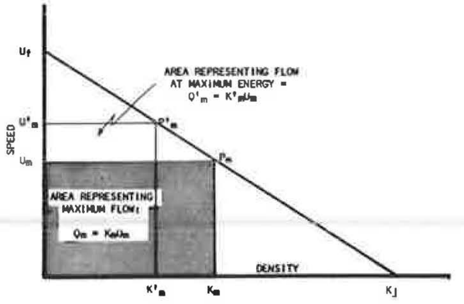


Figure 9. Linear speed-density model.

the middle of the occupancy range. It is noted, however, that in all cases the calculated intercept is very close to zero.

THE LINEAR SPEED-DENSITY MODEL

A single trip on the freeway is sufficient to decide that speed decreases with increasing density, and a number of models describing this relationship have been formulated. The speed-density model to be used in this analysis was first proposed by Greenshields (7). This model assumes a linear relationship given by

$$U = U_f(1 - K/K_j)$$

where U and K have already been defined as the appropriate speed and density values; U is the "free speed", i.e., the speed at which traffic would theoretically move no density or legal constraints; and K_j is the "jam density", i.e., the density at which traffic would (again theoretically) grind to a complete halt. Neither of these operating points are likely to be observed in practice; however, the linear model will be shown to provide useful application over a fairly wide range of values in between.

Figure 9 shows the linear model in graphical form and illustrates that, for any point, P at K and U, the flow Q is represented by the area OKPU (since Q = KU). It is also apparent that the maximum Q will be attained at $K - \frac{1}{2} K_j$ and $U = \frac{1}{2} U_f$. Using the subscript m to denote these values, the relationship $Q_m = K_m U_m$ establishes a means for estimating the capacity at a particular location.

It should be noted that all of these values apply only to maximization of flow. A corresponding set of points could, however, be derived for the maximization of energy using this model. Remembering that

$$U = U_f\left(1 - \frac{K}{K_j}\right) = U_f - \frac{KU_f}{K_j}$$

and

$$E = KU^2 = K\left(U_f - \frac{KU_f}{K_j}\right)^2$$

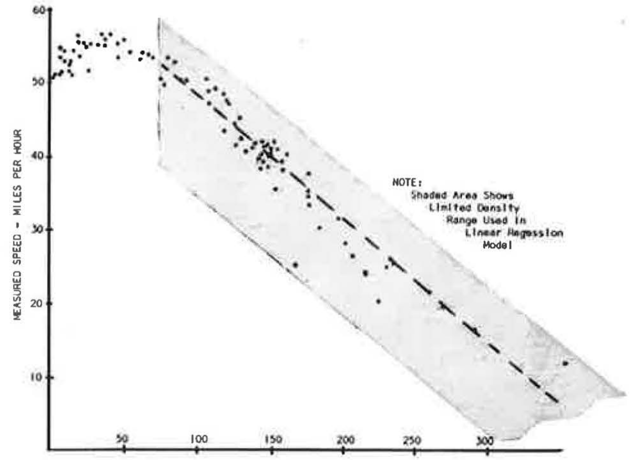


Figure 10. Speed-density curve for composite measurements.

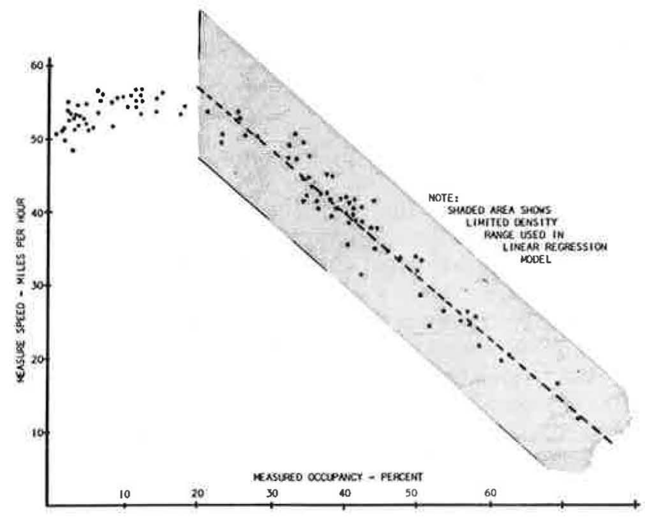


Figure 11. Speed-occupancy curve for optic detectors.

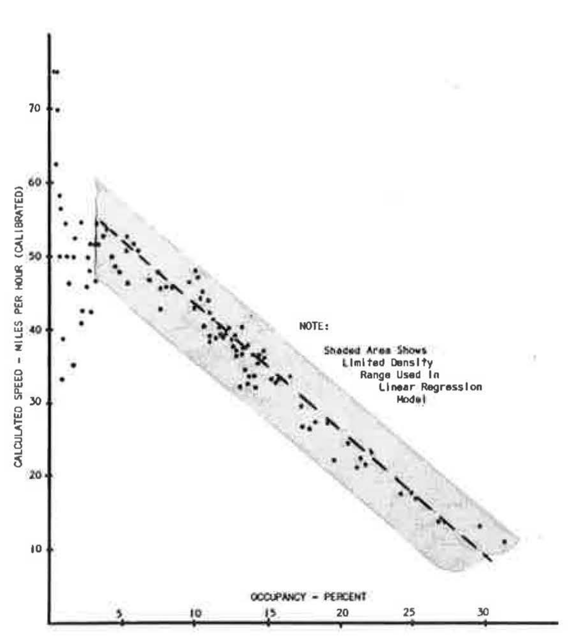


Figure 12. Speed-occupancy curve for sonic detectors.

TABLE 3

SUMMARY OF MACROSCOPIC CHARACTERISTICS ESTIMATED BY LINEAR REGRESSION

Characteristic	Units	K vs U_o	O_o vs U_o	O_s vs U_s
Free speed	mph	65.6	74.4	60.0
Jam density	veh/lane-mi	128	110	125
Optimum speed	mph	32.8	37.2	30.0
Optimum density	veh/lane-mi	64	55	62.5
Capacity	veh/min	105	107	96
Maximum energy	veh-mi/hr^2	58,000	64,000	47,500
Correlation index	—	.942	.948	.948
Figure	—	10	11	12

Applications of the Linear Model

A linear regression analysis was performed on the independent measurements that relate to the speed-density curve.

Figures 10 through 12 show the individual data points plotted for the three relationships, U_O vs O_O, U_S vs O_S, and U_O vs K. These relationships are clearly not linear, and this is not surprising because the speeds are expected to be independent of density in the lower density ranges (i.e., near the speed limit).

More meaningful conclusions can be drawn from a limited density model that appears to be fairly linear throughout the usual range of operation of a surveillance and control system. The regression analysis was, therefore, performed using only those points with density greater than 25 vehicles per lane mile.

Table 3 gives the results of the speed-density regression in terms of free speed, jam density, capacity, optimum density, optimum speed, and correlation index. It is observed that, within the limited density, the model in all three cases appears to fit the assumption of linearity, giving reasonable values for optimum speed, optimum density, and capacity, and exhibiting fairly good correlation. In the case of the sonic speed vs sonic occupancy, a similar scattering of points is noted at higher speeds as was previously observed.

TABLE 4

NORMALIZING VALUES FOR CALCULATIONS USED IN THE FREEWAY SURVEILLANCE AND CONTROL LOGIC

Estimated Quantity	Actual Units	Freeway Location				
		Edsel Ford	Seward	Chicago	Calvert	Glendale
Free speed, U_f	Vehicles per minute per percent occupancy	5.10	9.48	6.42	10.02	10.38
Optimum speed, U_m	Vehicles per minute per percent occupancy	2.55	4.74	3.21	5.01	5.19
Jam density, K_j	Percent occupancy	55	33	54	39	38
Optimum density, K_m	Percent occupancy	27.4	16.6	28.0	19.5	18.9
Capacity, Q_m	Vehicles per minute	70	105	90	97	98
Maximum energy, E_m	Vehicles^2 per minute^2 per percent occupancy	211	587	335	580	593
Flow at E_m, Q'_m	Vehicles per minute	62	93	80	87	87

This linear model is not intended to negate the more advanced theories that have been proposed to describe the speed-density relationship. It is not felt, however, that sufficient data have been collected to explore the various "multi-regime" theories (8), nor was the sampling interval short enough to examine such concepts as hysteresis (9). The investigations reported herein were undertaken primarily to support certain other research activities and have proved useful in this regard.

As a direct application of these techniques to the John C. Lodge Freeway surveillance and control system, Table 4 gives the values of the various macroscopic parameters derived for all of the measurement stations on the Freeway. The figures given for the sonic detectors are uncalibrated; i.e., the speeds are represented by the proportional value of Q/O. No calibration is required in this case because the surveillance system employs only normalized speed values in the computations. In the case of the two optic units, calibration by manual counting was required to give an actual flow value since compatibility with other flow measurements was essential to the computations.

SUMMARY AND CONCLUSIONS

The simultaneous measurement of speed, flow, and occupancy at the same point on the freeway using the two techniques described herein supports the following conclusions:

1. A very close relationship exists between equivalent vehicular density (vehicles per mile) and percentage of roadway occupancy measured by both methods;
2. This relationship may be used to calculate reasonable mean values of flow from the optic speed measurements and speed from the sonic flow measurements over a 1-minute sampling period, within the density range normally experienced during peak period operation;
3. Reasonable estimates of mean kinetic energy may also be obtained using either measurement technique, over the same time interval; and
4. Within the limited depth of the analysis, the linear speed-density model, when applied to the observed data, produced results that appear to be useful to the operation of the present freeway surveillance and control system.

ACKNOWLEDGMENTS

This study was carried out under NCHRP Project 20-3 by the Texas Transportation Institute, where the author, who was the principal investigator, worked as an assistant research engineer. He wishes to thank the Detroit-based staff of the Texas Transportation Institute as well as the Michigan Department of State Highways staff assigned to the surveillance project for their help in performing these studies. Messrs. Gordon Paesani, Terry Cox and Ed Podany did a commendable job of computer programming, and Mr. Darrell Campbell's efforts in the installation of equipment were greatly appreciated.

Special appreciation is extended to Dr. Joseph A. Wattleworth, Vice President of Kelly Scientific Corporation and formerly head of the Traffic Systems Department of the Texas Transportation Institute, for his frequent advice and assistance.

REFERENCES

1. Foote, R. S., and Crowley, K. W. Developing Density Controls for Improved Traffic Operations. Highway Research Record 154, p. 24-37, 1967.
2. Stern, S. Traffic Flow Data Collection Using Magnetic-Loop Vehicle Detectors. Highway Research Record 154, p. 38-52, 1967.
3. Edie, L. C. Discussion of Traffic Stream Measurement and Description. The Port of New York Authority, Rept. 63-1, 1964.
4. Drew, D. R., and Keese, C. J. Freeway Level of Service as Influenced by Volume and Capacity Characteristics. Texas Transportation Institute, Research Rept. 24-3, 1965.

5. Matson, T. M., Smith, W. S., and Hurd, F. W. Traffic Engineering, McGraw-Hill, 1955.
6. Weinberg, M. I., Deleys, N. J., and Schneeberger, R. R. Surveillance Methods and Ways and Means of Communicating With Drivers. NCHRP Report 28, 1966.
7. Greenshields, B. D. A Study of Traffic Capacity. HRB Proc. Vol. 14, p. 448-477, 1934.
8. Drake, J. S., Schofer, J. L., and May, A. D., Jr. A Statistical Analysis of Speed Density Hypotheses. Highway Research Record 154, p. 53-87, 1967.
9. Blunden, W. R. Capacity of Highways—Some Fundamental Considerations. Proc. 35th Annual Meeting, Institute of Traffic Engineers, 1965.

Discussion

R. DUSTIN ARNOLD, Texas Instruments Incorporated—The data collection system described by Courage makes use of a multichannel communication link and a general purpose digital computer, something which many researchers cannot afford. Such basic traffic data have great utility and, if they were more easily obtainable, sufficient data could be obtained in many cities on many facilities. Such traffic data have long been sought by traffic engineers and researchers. The following discussion presents an alternative data collection system designed for this same purpose, and represents a significant reduction in cost while still having a fair amount of flexibility. It provides speed, flow rate, and density measurements over variable time periods and is small enough so that it can be moved from site to site in a small van. Final output format as presently conceived is printed paper tape, which can be accumulated and analyzed when convenient.

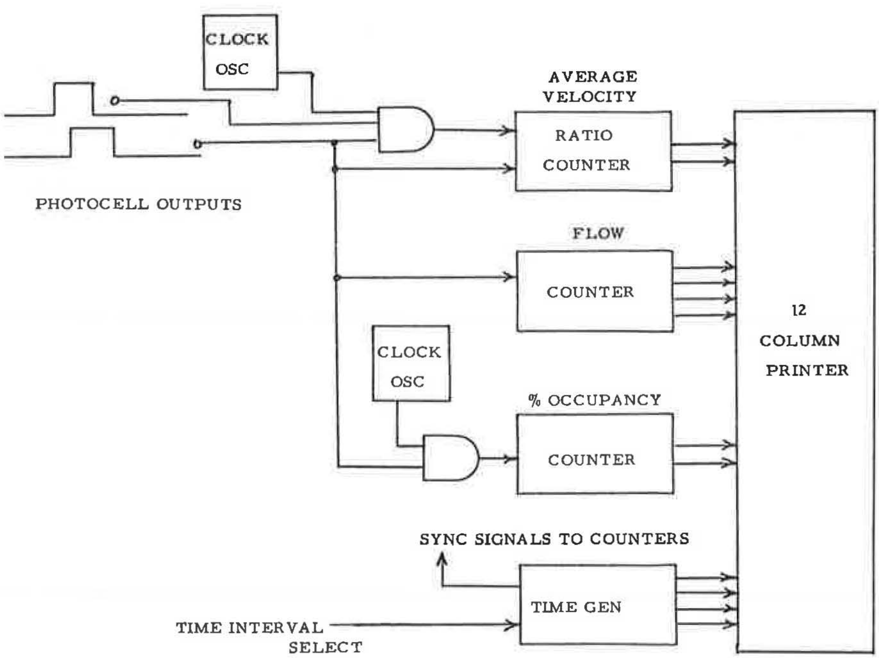


Figure 13. Measurement system.

The Measurement System

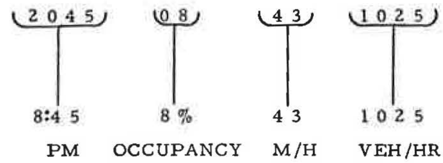


Figure 14. Typical data printout.

The suggested measurement system shown in Figure 13 uses the outputs of of photocell detectors to derive flow, density, and speed. The photocell outputs control general-purpose electronic counters whose outputs are connected to a paper tape printer, the final data being properly identified and timed. Density is expressed as percent occupancy, flow as vehicles per hour, and average velocity as miles per hour. The measurement period suggested here is 5 minutes and 1 minute (switch selectable), but any shorter or longer periods could be used with the appropriate change in the time base and clock oscillators. The resultant typed answers on the paper tape are a permanent record of the traffic movement in units that the traffic researcher can easily use. A typical typed output format is shown in Figure 14.

Traffic Flow

The photocell output is used to trigger an ordinary event counter to register total vehicles per sample period. An external time base is used to control gate period, allowing averaging over whatever time period the researcher desires.

Speed Measurement

Speed measurements would be made in a manner similar to that suggested by Courage, utilizing the overlapping portion of the two photocell outputs to gate on a clock frequency. The total clock counts are summed in the B portion of a ratio counter, while total vehicle passages are summed in the A portion, the resultant output being A/B, which is proportional to velocity. The setting of the clock frequency introduces a constant factor that takes into consideration the photocell spacing. The following relationship holds true:

$$V = \frac{S\ J\ 10^5}{C}$$

in which V = average velocity, mph; S = photocell spacing; J = number of vehicles in measurement period; and C = sum of counts of gated clock. Clock frequency is as follows:

$$f_{clock} = \frac{5280}{3600} \times \frac{10^6}{S}$$

where S = photocell spacing in feet. The velocity readout is then the average velocity of all the vehicles that passed by within the selected sample period.

The very portable system described here could be used to measure actual speed, flow, and occupancy at any location the traffic researcher desires. The overall cost of the system would be around $5000 and would provide plentiful and accurate data. It is hoped that the author accepts this suggested system as a flexible and economic means of continuing research in this area.

KENNETH G. COURAGE, Closure—The author would like to thank Mr. Arnold for his interest in the data acquisition system and for his constructive thoughts on a portable and practical measurement technique. The data on which the studies were based were collected using a multipair cable connected directly to a digital computer. This method has many advantages, including the capacity for real-time use of the information in the ramp metering control logic. (This feature has now been incorporated into the control system.) There is no question, however, that a less costly system would be extremely useful in a variety of basic and applied research studies.

Regarding the measurement of flow, the technique proposed by Mr. Arnold would simply sum the output pulses from one of the photocell units. Since we are generally dealing with a multilane freeway, the summation of pulses would only indicate how many platoons crossed the detector as opposed to the actual vehicle count. In the studies described in the paper, the flow was estimated from the product of speed and occupancy, and the results were verified by comparison with the sonic detector counts (see Fig. 5). It is unlikely that the necessary hardware multiplication would prove economical in a special-purpose digital device of the type proposed by Mr. Arnold; however, the necessary calculations could easily be performed off-line on a larger computer using the speed and occupancy information as inputs. The pulse counter could then be used for calibration of the output. A few hours of manual count data (collected in this case by connecting a hand switch to the pulse counter) could be used to determine the constants of a proportionality by linear regression. Experience with the John C. Lodge surveillance system suggests that 1 hour of manual data obtained separately under low, medium, and high densities would be adequate.

Once a calibrated model was available for a particular location, large amounts of data could be collected automatically for a variety of purposes, including bottleneck capacity estimation, investigation of the operational effects of geometrics and environmental conditions, and evaluation of traffic engineering improvements.

A Moving Vehicle Merging Control System

J. H. BUHR, M. L. RADKE, B. M. KIRK, and D. R. DREW, Texas Transportation Institute, Texas A&M University

•THREE research projects have been undertaken to design a system of optimal freeway control. The projects studied the John Lodge Freeway in Detroit, the Eisenhower Expressway in Chicago, and the Gulf Freeway in Houston. Each project has met with a certain degree of success, although each has developed a different control strategy. The one common factor among these projects is the concept that, in order to relieve freeway congestion, the rate of entry of vehicles from the entrance ramps has to be controlled in some manner.

The Houston project is the only one that assumed the secondary objective, apart from relieving freeway congestion, of aiding the ramp driver in performing the merging maneuver. This led to the development of the gap acceptance method of control. In this method, large gaps representing units of capacity are detected on the freeway lane adjacent to the on ramp, and vehicles are then released from the ramp in an attempt to fit them into the gaps. Since there would often be more vehicles trying to use the ramp than there are available gaps, it is necessary to delay vehicles on the ramp until a large enough gap becomes available. Stopping vehicles on the ramp also has an added advantage in that the travel time of a ramp vehicle from the stopped position to the merge point can be easily predicted (1), thus greatly simplifying the hardware designed to fit ramp vehicles into freeway gaps (2).

This form of control has been very successful in achieving both its primary objective of relieving freeway congestion and its secondary objective of aiding ramp drivers. Because of the primary objective, such controls have been operated only during peak hours when freeway congestion constitutes a problem. However, the secondary objective is thought to be of such importance that consideration was given to operating the controllers also during off-peak period when freeway capacity does not constitute a problem. Under such conditions there would likely be many more large gaps on the freeway than there are ramp vehicles to fill them. As a result many or perhaps even most ramp drivers would not experience a problem in finding an acceptable gap after they reach the merging area, and it would therefore be unnecessary to delay all vehicles on the ramp. Consequently, it was necessary to develop a gap acceptance controller that would not require all ramp vehicles to come to a stop. Such a control system has been designated a "moving vehicle merging control system" as opposed to the "stopped vehicle merging control system" described earlier.

THEORY

Model Description

A control system designed to fit ramp vehicles into freeway gaps has three basic requirements:

1. It must be able to measure or otherwise determine the estimated time-space trajectory of a ramp vehicle;
2. It must be able to measure and estimate the time-space trajectory of a freeway gap; and
3. It must be capable of controlling or affecting in a known manner at least one of the first two characteristics.

Paper sponsored by Committee on Traffic Control Devices.

The first two requirements are relatively simple because vehicle presence and speed can readily be detected with any of various types of sensors. For various reasons, it is more desirable to control the time-space trajectory of a ramp vehicle rather than that of a freeway vehicle. The stopped merge system controls the ramp vehicle by stopping all vehicles at a known point. In the first approach to the development of a moving merge control system, consideration was given to providing a changeable speed message sign on the ramp to meet the third requirement. The controller would then sense the presence and speed of a freeway gap, sense the presence and speed of a ramp vehicle, solve the equations of motion for the speed requirements of the ramp vehicle in order to match the gap, and then display this speed on the variable speed sign. The driver seeing the displayed speed would then have to adjust his speed accordingly in order to merge into the assigned gap. There are at least two serious objections to such a system: (a) in order to perform the control functions, freeway and ramp characteristics have to be sensed very far upstream of the merging area; and (b) it is doubtful that the average driver would react to the speed sign in the desired manner, if at all. The first objection would limit the application of the control system to very long ramps entering a freeway with no on or off ramps for a considerable distance upstream of the controlled entrance. The stability of the measured parameters would also be questionable over such a long distance. For reasonable distances the required speed changes would in some cases be too abrupt or of too large a magnitude. This approach was therefore abandoned for the one described next.

The freeway conditions under which the moving merge model would be applicable would connote numerous large gaps on the freeway. Many ramp vehicles would be able to enter these gaps without the aid of a control system and it would therefore be best if the control system does not affect them. Other ramp vehicles would not be so fortunate but would encounter a "block" on the freeway, making the merging maneuver quite difficult and forcing the driver to slow down or even stop. These then are the ramp vehicles that should be affected by control. Some of these vehicles can be matched with large gaps by a small adjustment in speed while others would require an unreasonably large speed adjustment. The latter group would therefore have to be stopped to await the arrival of an acceptable gap. Such a control system can be effected by a series of traffic signals on the ramp (Fig. 1). This system uses green and red "bands" that, through signal progression, lead into acceptable and unacceptable gaps on the freeway. Since vehicles that would otherwise be unable to match their paths with those of large gaps would be stopped by the first signal, the signal progression has to be designed for a stopped vehicle. It is therefore necessary to consider the equations of motion of a vehicle accelerating from a stopped position.

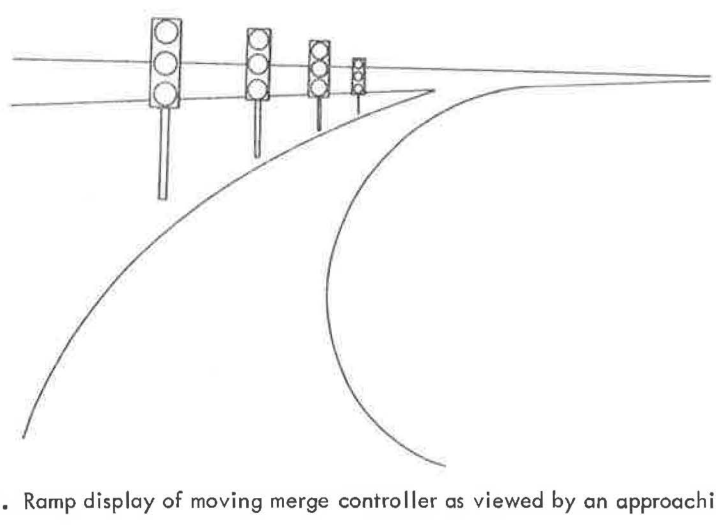


Figure 1. Ramp display of moving merge controller as viewed by an approaching driver.

Equations of Motion

The operation of the controller in terms of the proper instant in time to initiate a color band and the correct timing sequence through the signals depend entirely on the assumed traffic characteristics and the geometric layout of the freeway detectors and ramp signals.

It can be assumed that the acceleration of a vehicle is linearly related to its speed (1) as expressed by

$$\frac{du}{dt} = a - bu \tag{1}$$

where u is the speed of the vehicle, t is time, and a and b are constants.

This relationship leads to other vehicle operating characteristics that are determining factors in the operation of the controller. A ramp vehicle starting from a stopped position will therefore have the following characteristics. The time required to accelerate to speed u is given by

$$t = -b^{-1} \ln\left(\frac{a - bu}{a}\right) \tag{2}$$

Its speed after time t is given by

$$u = \frac{a}{b}\left(1 - e^{-bt}\right) \tag{3}$$

The distance covered in time t is given by

$$x = \frac{a}{b} t - \frac{a}{b^2}\left(1 - e^{-bt}\right) \tag{4}$$

The distance covered in accelerating to speed u is given by

$$x = b^{-1}\left[u + \frac{a}{b} \ln\left(\frac{a - bu}{a}\right)\right] \tag{5}$$

Given these relationships, the operating characteristics of ramp vehicles depend only on the assumed values of a and b, where a is the maximum acceleration used by drivers under "normal" conditions and b is the rate of change of acceleration with speed. These two parameters affect the projection time of gaps and the timing sequence of the signals. Their effect is treated in greater detail in another later paper (1). For purpose of illustration, it will be assumed that $a = 8$ ft/sec^2 and $b^{-1} = 18$ sec. It is desirable that the control equipment allow some flexibility in the assumptions regarding the values of these parameters.

The time-space trajectory of a vehicle that starts from a stopped position and uses normal acceleration as assumed is shown in Figure 2. The speeds reached at various traveled distances are also indicated.

Geometric Configuration

To provide drivers with a "follow the rabbit" type display, thus leading them into gaps, it would be desirable to have signals every few feet along the ramp and acceleration lane, starting far enough up the ramp so that even a standing vehicle can easily

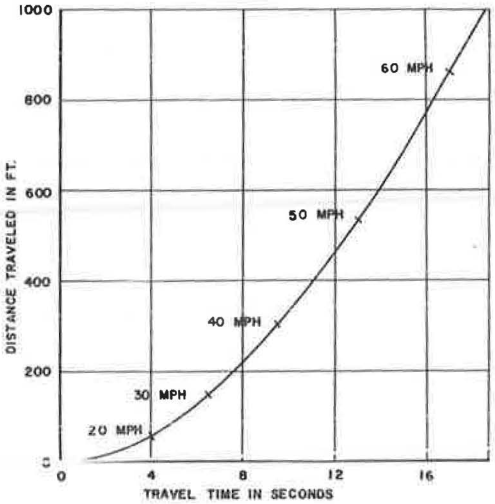


Figure 2. Time-space trajectory of a vehicle utilizing normal acceleration: $a = 8$ ft/sec^2, $b^{-1} = 18$ sec.

reach the highest anticipated freeway speed, and terminating at the end of the acceleration lane. Since the control system can never provide a 100 percent certainty that an acceptable gap will be available when the ramp vehicle gets to the merging area, the ramp display should not give the impression that it provides absolute right-of-way. The final decision of whether to merge or not must be left up to the driver. Furthermore, a vehicle that enters the signalized section of the ramp but travels too slowly will be stopped by one of the signals and will then have to accelerate from a standing start over a short distance before it has to merge. For these reasons, the last signal has to be placed some distance upstream of the ramp nose. In the prototype model described later, it was decided to place the last signal 150 ft from the nose. This also enhanced the adaptability of the controller to the stopped mode of control in which this last signal will act as the ramp control signal. The same controller could therefore operate on a 24-hour basis—operating in the stopped vehicle mode during peak hours and in the moving vehicle mode during off-peak periods.

The first signal encountered by a driver should be placed far enough from the nose so that a vehicle can reach the maximum desirable speed from a standing start. This distance is also limited by the ramp length. It is believed that there is little point in attempting to encourage a ramp vehicle to accelerate faster than 50 mph by the time it reaches the nose. It was therefore decided to place the first signal of this prototype installation 600 ft from the nose. Under the assumption made earlier, this would allow ramp speeds of approximately 52 mph at the nose (Fig. 2). This signal could, of course, be placed further from the nose if desired, as far as the length of the ramp will allow. Between the first and the last signal can be placed as many signals as desired. A great number of closely spaced signals would give drivers a better impression of the "follow the rabbit" type display. However, this somewhat unconventional display may create confusion especially to drivers trapped by a red band. It was therefore felt that the intermediate signals should be a sufficient distance apart so that a vehicle trapped in the signalized section could be given a leading green while still maintaining the basic sequence of signal indications.

The speeds of vehicles on the freeway as well as the size of the gap between them are measured by a single pair of closely spaced loop detectors placed in the outside lane of the freeway, some distance upstream of the ramp nose. The placement of these loop detectors will affect the projection time; this is discussed in another report (7).

Operation of Controller

The primary input to the moving merge controller is from a detection station on the freeway in the lane adjacent to the controlled entrance and some distance upstream of the merging area. At this detection station the controller measures freeway speeds and gaps between vehicles to determine whether the gaps are smaller or larger than a predetermined acceptable gap size, G. The controller is primarily concerned with changing the color indication on the first signal encountered by a ramp driver. The other signals on the ramp change in a consecutive sequence after the primary signal changes. This sequence may be predetermined or it may be established by the controller in accordance with the measured freeway speed. The action taken by the signal controller depends on whether the measured gap is greater or less than the acceptable gap, G, and on the present state of the primary ramp signal, i.e., green, amber, or red. Henceforth, the amber shall be considered as part of the green phase. If the primary signal is green, the controller searches for gaps smaller than G. When such a gap is detected, a red band is initiated at the proper time. If the primary signal is red, the controller searches for gaps bigger than G, and when such a gap is found a green band is initiated at the proper time. The action to be taken by the controller is as follows:

State of Signal 1	Gap ≤ G	Gap ≥ G
Red	None	Turn signal green
Green	Turn signal red	None

The proper instant to change the primary signal indication is such that the travel time of the leader of the gap to the merging area is the same as that of a ramp vehicle stopped at the first ramp signal. This instant is designated the "decision point." The action just described is taken by the controller regardless of whether there is a vehicle waiting at the first signal or not.

The controller also receives indications from detectors placed on the ramp in front of each signal except the first signal. If one of these detectors indicates a vehicle trapped in the signalized section and thus stopped at one of the signals, the controller will modify the usual signal progression sequence by providing a leading green at this location.

It may be desirable to inform ramp drivers of the freeway speed in effect at the time in order to encourage them to accelerate to that speed. In this case, the controller can continually update the speed indication on a variable message speed sign placed on the ramp.

Projection Time

Every vehicle crossing the freeway detection station is projected for at least a short time. The term "projection time" as used here means the delay between the instant that a vehicle is first detected and the decision point.

The projection time depends essentially on (a) the location of the gap, (b) speed detectors, and (c) the travel time of a ramp vehicle, as determined by its location, the normal acceleration time to freeway speed, and the assumption that this speed is maintained until merging with the freeway. Under the assumption made earlier, the time required for a vehicle to accelerate to speed u is given by Eq. 2 as

$$t = -18 \ln \left(1 - \frac{u}{144}\right)$$

In accelerating to this speed, the driver would have covered a distance given by Eq. 5 as

$$x = -18 \left[u + 144 \ln \left(1 - \frac{u}{144}\right)\right]$$

If the first signal is located 600 ft upstream of the merge point, then the remaining distance to be covered is equal to 600 - x, which the driver is assumed to traverse at a steady speed of u. The travel time of a ramp vehicle from a stopped position in front of the first signal to the merge point is therefore given by

$$t = \frac{600}{u} + 18 + \left(\frac{2592}{u} - 18\right) \ln \left(1 - \frac{u}{144}\right) \tag{6}$$

where u is in ft/sec and t in seconds, $u \leq 52.5$ mph. Solving this equation for various values of u yields the times given in Table 1.

The freeway detector station has to be located a sufficient distance upstream of the ramp nose so that the travel time of a freeway vehicle, traveling at the highest speed anticipated, is equal to the ramp travel time given in Table 1, plus the largest gap to be measured, plus the starting delay of a ramp driver. The ramp vehicle should arrive in the merging area just shortly after the lead vehicle of an acceptable gap. If the green band is therefore set to lead into the merging area L seconds ahead of the lead vehicle, a ramp vehicle stopped at the first signal at the beginning of the green band should arrive in the merge area at a time equal to its starting delay minus L seconds after the lead vehicle. This lead time, L, shall be referred to as the leading green time. This would also be approximately true for a moving vehicle that just catches the beginning of the green band since the signal has to turn green while the vehicle is still some distance away so that it is not forced to slow down. The freeway detector should thus be placed so that the minimum freeway travel time between it and the nose equals the corresponding ramp travel time plus the maximum G plus L seconds.

The minimum distance that must be provided between the nose and the freeway detector is given in Table 2. To control ramp vehicles effectively when the freeway speeds are 60 mph, it is necessary that the freeway detectors be located at least 1750 ft upstream of the nose of the ramp. The projection time at various freeway speeds would therefore be given by

$$\frac{1750}{u} - (G + L) - \left[b^{-1} + \frac{600}{u} + b^{-1}\left(\frac{u}{ab} - 1\right) \ln \left(1 - \frac{ub}{a}\right)\right] \tag{7}$$

where u is in ft/sec.

Under the assumptions regarding a and b made earlier and assuming that (G + L) is set at 6 sec for all speeds, the relationship reduces to

$$\text{Projection time} = \frac{1150}{u} - 24\left(\frac{2592}{u} - 18\right) \ln \left(\frac{1 - u}{144}\right) \tag{8}$$

The timing operation of the controller is shown in Figure 3. To provide full flexibility as a research instrument, the only characteristic that should be considered fixed is the freeway travel time. All the other times should be adjustable by means of settings on a controller designed for research. This would enable adjustments not only

TABLE 1

RAMP TRAVEL TIME FOR VARIOUS FREEWAY AND RAMP SPEED CONDITIONS

Freeway Speed (mph)	Ramp Speed Reached (mph)	Ramp Travel Time (sec)
25	25	18.87
30	30	16.71
35	35	15.36
40	40	14.52
45	45	14.04
50	50	13.83
55	52	13.80
60	52	13.80

TABLE 2

MINIMUM DISTANCE REQUIRED BETWEEN RAMP NOSE AND FREEWAY DETECTOR

Freeway Speed (mph)	Required Freeway Travel Time (sec)	Freeway Detector Location (ft)
25	24.87	915
30	22.71	1000
40	20.52	1208
50	19.83	1458
55	19.80	1600
60	19.80	1748

for different acceptable gaps, at distinctive freeways speeds, but also for different ramp operating characteristics.

Upon turning the signals from green to red, the timing operation remains essentially the same. In this case, a leading red time is desired similar to the leading green time L. However, the operation of the controller could be such that the leading red time is equal to the leading green time. The projection time would therefore remain unchanged. Instead of turning the primary signal, it should remain amber at this point and turn red only after an extended amber period has expired. The other signals should turn red on schedule as if the first signal had turned red at the decision point, with the amber period inserted just prior to this time. In other words, the signals are turned red on the same timing sequence as they are turned green, except that the red on the primary signal is delayed by an amount equal to the additional amber time. It is thought that this procedure would encourage drivers who approach the signalized section just before a red band to speed up, thus still enabling them to enter the gap. An extended amber period is also desirable to prevent sudden stops.

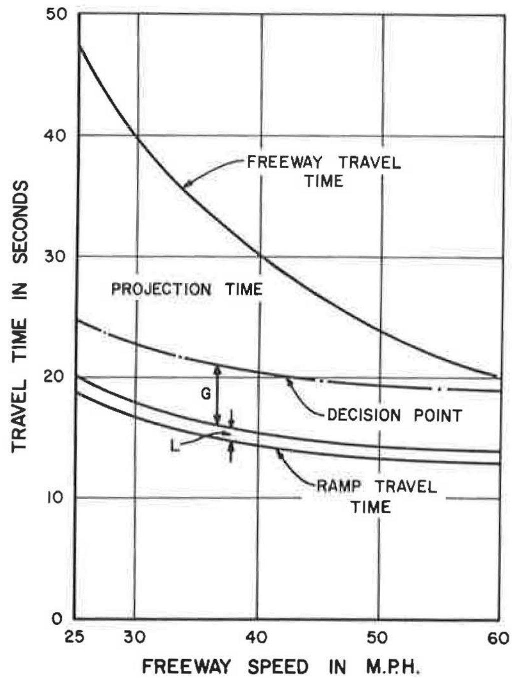


Figure 3. Controller timing operation.

Signal Sequence Timing

As explained earlier, the timing sequence at which consecutive signals change their color should be based on the time-space trajectory of a vehicle accelerating from a stopped position to the current freeway speed, utilizing normal acceleration, and then maintaining the freeway speed. The exact offsets will depend not only on the ramp vehicle characteristics but also on the placement and spacing of the signals. In more refined equipment, the offsets, especially that of the last signal, should change with freeway speed. However, in the experimental prototype described later, all that was required was an adjustable timing device by which the offsets could be set, since the freeway speeds could be controlled. The offsets required can easily be determined for any particular set of vehicle operating characteristics and signal locations with the aid of graphs such as Figure 2.

Figure 4 shows typical green and red bands for a freeway speed of 50 mph. It can be seen that only vehicles stopped at the first signal, or vehicles arriving just before the red phase, will be encouraged to change their speed. Vehicles arriving during the green phase may proceed through the controlled section undisturbed, which is desirable since they would have found the merging area empty anyway. The desirability of a curved time-space trajectory at the beginnning of the green band, as opposed to a straight-line (constant-speed) relationship, is evident since some vehicles may be stopped at the first signal. Furthermore, vehicles approaching the first signal near the beginning of the green band will only have to slow down without actually having to stop. The desirability of a curved time-space relationship at the beginning of the red band is shown by Figure 4. A vehicle entering the controlled section just before the red band at too slow a speed will be encouraged to speed up to the freeway speed. Otherwise, the slow vehicle will be stopped by one of the signals. However, in order to permit a vehicle traveling at 35 mph to maintain its speed and still enter the acceptable gap, the first signal should be delayed somewhat, resulting in a slightly different time-space trajectory for the red band than for the green band.

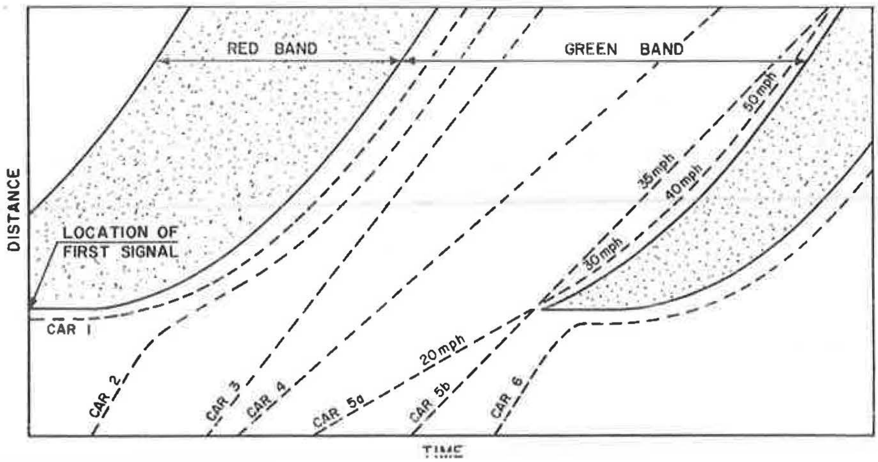


Figure 4. Illustration for 50-mph freeway speed.

TESTING

Test Facilities

To investigate the feasibility of the moving vehicle merging control system described, it was decided to test it in a controlled environment rather than under real-life conditions. For this purpose, a test track was laid out on the concrete runways of the Texas A&M Research Annex, formerly the Bryan Air Force Base (Fig. 5).

The test track consisted of three 12-ft lanes with a 12-ft shoulder on each side. The shoulders were sprayed with asphalt emulsion in order to create a more realistic freeway environment. An entrance ramp was located on a straight section of the simulated freeway. The entrance ramp, laid out with a 3-degree curve, was 14 ft wide and 600 ft long with an acceleration lane of 650 ft.

Control System

Four traffic signals were used in testing the moving merge model. The first signal was placed 600 ft from the nose and the others at 150-ft intervals so that the last signal was located 150 ft from the ramp nose. The signals were controlled by a standard fixed-time signal controller with the cams stacked to provide the desired timing sequence. An advisory speed sign was placed on the ramp 525 ft upstream of the nose, i.e., between the first two signals. The speed indication was changed manually. A radio communication system was obtained from the Bureau of Public Roads for communication with the ramp driver. This radio system has been successfully used in previous research. A receiver was placed in the ramp vehicles and prerecorded messages regarding the freeway speed were transmitted to the ramp driver.

Data were collected by 13 loop detectors taped to the pavement in front of each signal and at various locations on the freeway and acceleration lane. When a vehicle was detected it was recorded on a 20-pen event recorder, which permitted the correlation of ramp and freeway vehicle locations, ramp driver compliance to the controls, ramp vehicle travel time, the merge speed, and the exact location of any ramp vehicle that might have been stopped by one of the signals.

In addition, an instrumented vehicle was obtained on a loan basis from the Ford Motor Company. This instrumented vehicle allowed the measurement of physiological factors and vehicle control movements, such as fine steering reversals required for tracking, gross steering reversals required for lane changes and turning movements, acceleration reversals, brake applications, driver heartbeat rate, and galvanic skin response.

Experimental Design

In evaluating the feasibility of the moving merge concept, it must be recognized that all possible variables cannot be included in the investigation. A few of the variables must be selected for detailed study and the remainder must be assumed sufficiently random as to have no effect on the results and thus can be ignored. This approach restricts the extrapolation of the results to real-world behavior, but should indicate the ramp driver's acceptance of the controls imposed.

The variables chosen for study were divided into two groups: (a) those that were held constant and (b) those that were deliberately varied. The constant variables were ramp geometrics, condition of pavement, sex of driver, educational background of driver, and ramp vehicle characteristics. The experimental condition of the variables to be

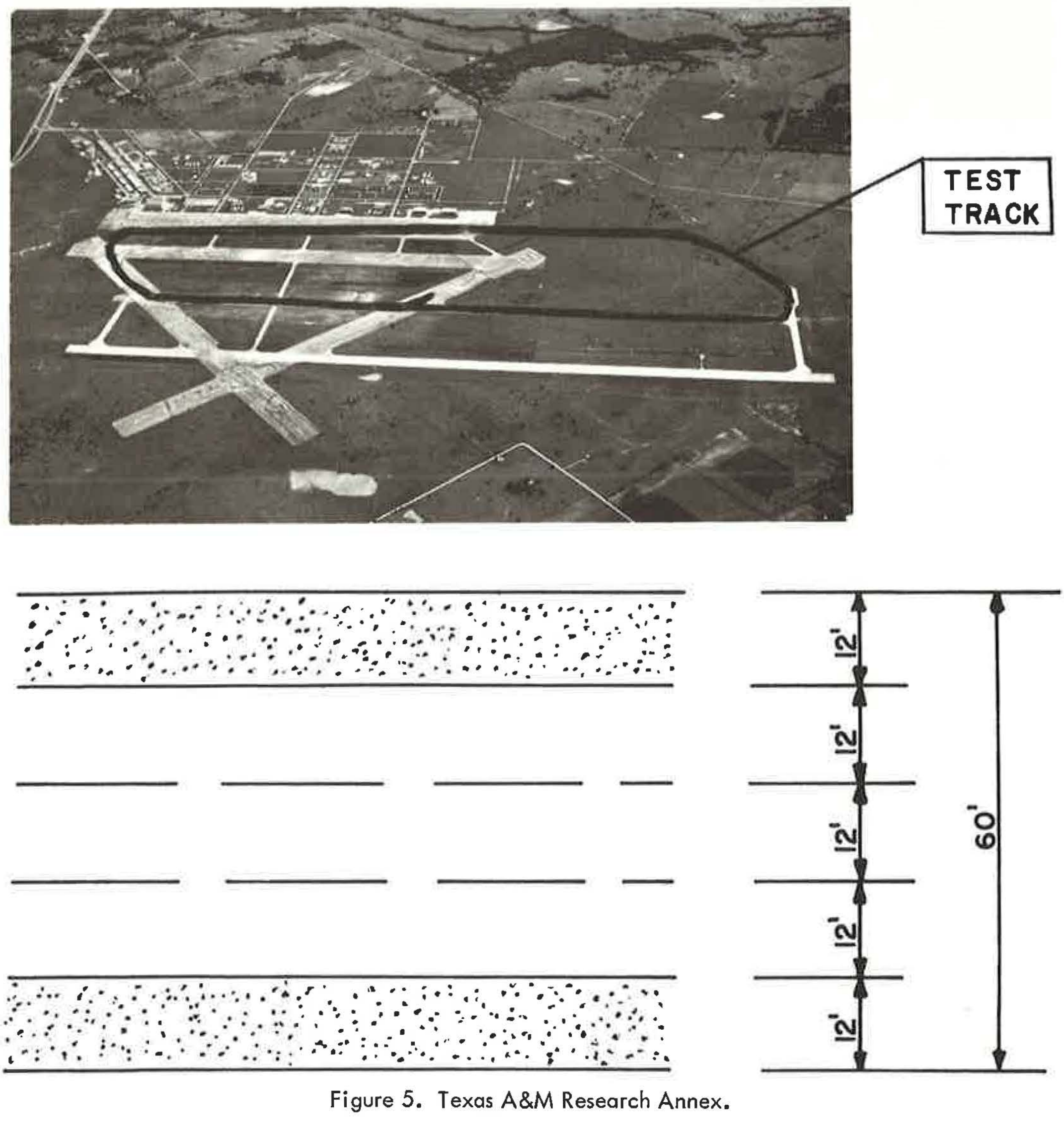


Figure 5. Texas A&M Research Annex.

Figure 6. Ramp vehicle with blackout panels.

altered was established before developing the experimental design. These variables and their conditional levels are as follows:

1. Factor A—freeway speed: level 1, 40 mph; level 2, 50 mph; and level 3, 60 mph.
2. Factor B—ramp approach speed: level 1, 25 mph; level 2, 30 mph; and level 3, 35 mph.
3. Factor C—status of ramp signals as seen by driver approaching ramp (Fig. 4): level 1, signals not in operation; level 2, signals red, symbolizing a "block" on freeway, sequencing to green after pause; level 3, first signal green, balance sequencing to green, signifying approach at the start of the green band (Fig. 4); level 4, first and second signals green, balance sequencing to green representing an approach at the beginning of the green band; level 5, all signals green, depicting an arrival during the green phase; and level 6, first signal amber, balance sequencing to red, illustrating arrival at the beginning of the red band.
4. Factor D—advisory speed information to ramp driver concerning the freeway speed: level 1, none; level 2, visual presentation; level 3, audio presentation; and level 4, visual and audio presentation.
5. Factor E—backsight (no outside rearview mirror): level 1, ramp driver's vision unobscured; level 2, ramp driver's vision for estimating velocity and distance distorted by clear plastic panels with grid on the left side of vehicle; and level 3, ramp driver's vision obscured by blackout panels on the left side of the vehicle (Fig. 6).
6. Factor F—ramp drivers: level 1 to N, each level is represented by a different subject functioning as the ramp driver.

Each test run was evaluated on the basis of whether the ramp vehicle had been placed in a position to merge or not. This evaluation was utilized in both the first series with a standard automobile as the ramp vehicle, and in the second series with an instrumented car as the ramp vehicle. The physiological factors of heartbeat rate and galvanic skin response were obtained by using the instrumented car as the ramp vehicle.

Because of the large number of variables involved, the experimental design selected was a random balance design (3). This technique permits the screening of a large number of possible contibuting factors in an experiment involving a limited number of test runs. The results allow the factors of importance to be isolated for further study.

Table 3 gives the design matrix for the series of tests with the standard car. The elements of the matrix are determined with a random sampling process. All factors and levels are considered by choosing at random the level of each factor to be used in a particular treatment combination. The assumption can therefore be made that the interpretation of a factor is independent of every other interpretation made in the investigation.

It was recognized that the variation of driver characteristics would have to be taken into account. This was accomplished for the standard car by the random choice of drivers. In the series of tests involving the instrumented car, the ramp drivers were taken as factors with each individual considered as a conditional level.

TABLE 3

DESIGN MATRIX, STANDARD CAR SERIES

Test Run	Variable Levels				
	A	B	C	D	E
1	1	1	2	2	3
2	3	1	3	2	2
3	1	3	4	4	3
4	1	2	4	3	2
5	2	3	2	4	3
6	1	2	6	2	1
7	3	2	6	4	2
8	1	2	3	2	3
9	1	2	1	4	1
10	1	2	3	2	2
11	1	1	3	4	3
12	3	1	1	4	2
13	1	2	5	2	1
14	1	2	6	2	1
15	3	2	1	4	1
16	3	2	4	2	3
17	3	1	5	4	2
18	1	1	2	3	2
19	1	1	1	1	1
20	3	2	3	4	3
21	2	1	5	1	2
22	3	1	4	3	3
23	3	3	2	3	1
24	2	2	1	1	3
25	2	3	6	2	3
26	3	1	3	3	2
27	2	3	1	1	2
28	3	3	6	1	1
29	3	3	4	3	2
30	2	3	2	1	3
31	2	1	6	3	1
32	2	1	5	1	1
33	2	3	2	3	3
34	2	3	4	1	1
35	2	3	5	3	1
36	2	3	5	1	2

Experimentation

A sample of undergraduate students was selected as ramp drivers for the first series of tests. The subjects averaged 21 years of age with 5½ years of driving experience. All had taken a driver's education program while in high school.

The subjects were instructed to perform the merge operation in accordance with the conditions presented to them. The only specified condition was that the vehicle was to approach the ramp at a given speed for each test. Vision to the left-rear was varied through the application of special panels on the driver's side of the ramp vehicle (Fig. 6). Two methods were used to present the freeway speed to the ramp driver: (a) the freeway speed was displayed on a standard type highway speed sign, or (b) an audio message was transmitted through the induction radio system. The status of the ramp signals as seen by the ramp driver as he approached was accomplished by the manual operation of the signal controller. The signal sequence was initiated so as to lead the ramp vehicle into a fixed gap on the freeway.

A freeway gap of approximately 4 sec was chosen for all test runs. The lead vehicle on the freeway maintained the freeway speed at the velocity designated by the design matrix, and the lag vehicle was spaced accordingly.

Thirteen detectors were monitored on the ramp, acceleration lane, and outside freeway lane. Actuations of the loop detectors were recorded by a 20-pen recorder during each test run. Data reduced from this tape permitted the plotting of time-space diagrams of each merge (Fig. 7) as well as the evaluation of the merge and calculation of ramp accelerations and velocities.

The second series of tests involving an instrumented vehicle were performed with graduate students as ramp drivers. The average driver in this group was 26 years old with 12 years of driving experience. Only 25 percent had received formal driver education training in high school.

The additional data obtained, heartbeat and galvanic skin response, were recorded directly on magnetic tape in the instrumented vehicle. A special computer program was used to evaluate and print these data. The loop data tape was processed in the same manner as that of the first series.

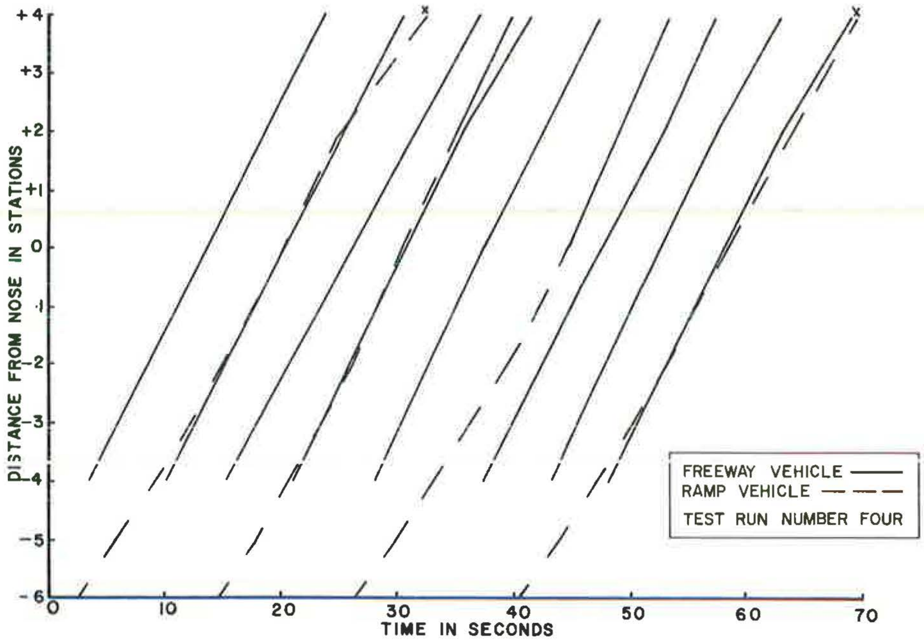


Figure 7. Time-space diagram—test run No. 4.

Automatic Controller

In the development of an automatic controller to perform the moving merge control function, work has proceeded on the modification of the first prototype merging controller developed on the Gulf Freeway Surveillance and Control Project (8). Since stopped merge control and moving merge control are similar in many respects, this modification was fairly easy to achieve. However, modification was completed too late for inclusion in this paper.

EVALUATION

Merging Maneuver

This investigation demonstrated that ramp vehicles can be place in position to merge with a moving vehicle merging control system (Fig. 8). However, having been placed in position, there was a great difference in the acceptance of the gap by the subjects. The more experienced drivers accepted the gap with higher frequency and with less use of the acceleration lane than those with less experience.

In the evaluation of each series of tests, values were assigned to the merge situations as follows:

1 Ramp vehicle in position to accept gap at the nose of the ramp;
0 Ramp vehicle not in position to accept gap at the nose of the ramp;
- Test run data disregarded as all levels of treatment combination were not correct.

The evaluation matrix for the standard car series is given in Table 4.

Included in the random balance design is the assumption that each factor can be considered individually. Therefore, linear statistical models can be assumed for each factor. The first series of tests had an additional effect in that each driver provided one observation for each treatment combination. Consequently, in the analysis of variance this replication must be taken into account.

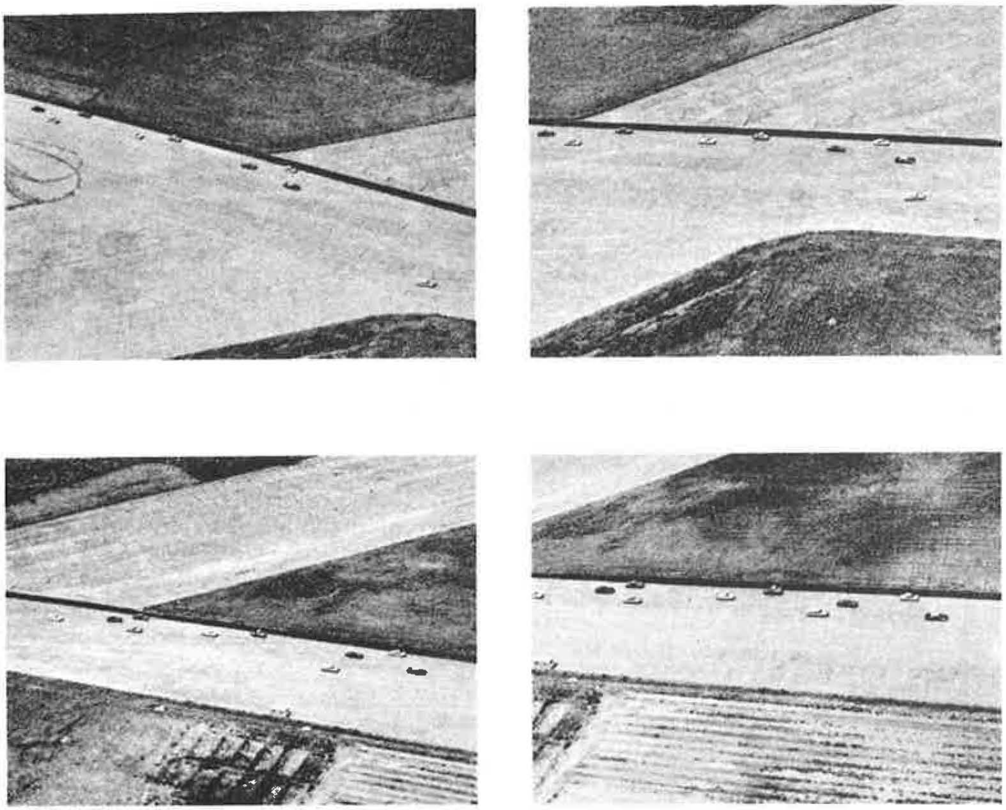


Figure 8. Time sequence.

The statistical model for the first series is represented by

$$Y_{ij} = m + r_i + f_j + e_{ij} \qquad i = 1,\ldots b$$
$$j = 1,\ldots t$$

where

m = the true mean effect,
r_i = the true effect of the ith replication (driver),
f_j = the true effect of the jth factor, and
e_{ij} = the true effect of the experimental unit that is subject to the jth factor in the ith replication.

The analyses of variance for the factors that show a significant difference between levels are itemized in Table 5. This significance would lead one to conclude that these factors are important contributors to placing the ramp vehicle in a position to merge.

It should be noted that individual drivers are not a factor in the first series, but account for more variation than the factors being considered. Drivers as well as freeway speed, status of ramp signals as seen by the ramp driver, and backsight should be given further study.

To determine if bias had been created by the grouping of data, it was necessary that the effect of the factors within drivers be considered. Only Factor C, status of ramp signals as seen by the ramp driver, shows significance. This indicates that the effect of signals differs among the drivers, which is to be expected. Table 6 presents the analysis of variance for this factor.

TABLE 4

EVALUATION MATRIX, STANDARD CAR SERIES

Test Run	Driver			
	1	2	3	4
1	1	1	1	0
2	1	1	1	0
3	1	1	1	1
4	1	1	1	0
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	0	1	1	0
9	-	-	1	-
10	1	0	1	0
11	1	1	1	1
12	-	1	0	0
13	-	1	1	-
14	1	1	1	1
15	0	1	1	-
16	1	0	-	0
17	0	1	1	0
18	1	1	1	1
19	1	1	1	0
20	1	0	-	1
21	0	1	1	-
22	0	1	-	0
23	0	-	1	0
24	1	1	0	-
25	1	1	1	1
26	0	-	-	0
27	1	1	1	0
28	1	1	1	-
29	0	-	-	-
30	1	1	1	0
31	1	1	1	-
32	0	0	0	0
33	-	1	1	1
34	0	-	0	0
35	0	1	1	0
36	1	1	1	0

TABLE 5

ANALYSIS OF VARIANCE, FIRST SERIES

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Total	117	25.78		
Drivers	3	5.16	1.72	10.13[a]
A	2	1.59	0.79	4.67[b]
Error	112	19.03	0.17	
Drivers	3	5.16	1.72	10.54[a]
C	5	3.82	0.76	4.67[b]
Error	109	20.23	0.16	
Drivers	3	5.16	1.72	9.47[a]
E	2	1.27	0.63	3.48[b]
Error	112	20.35	0.18	

[a]Significant at the 1 percent level.
[b]Significant at the 5 percent level.

TABLE 6

ANALYSIS OF VARIANCE, FIRST SERIES GROUPING

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Total	117	25.78		
Drivers	3	5.16	1.72	5.57[a]
C within drivers	20	6.18	0.31	1.88[b]
Error	94	15.43	0.16	

[a]Significant at 1 percent level.
[b]Significant at 5 percent level.

In the second series of tests, with the change of variables, there was no indication of significance between levels of any variables considered; that is, even the difference between drivers was not sufficient to be other than a random difference.

Driver Stress

The time required for the ramp vehicle to approach, proceed up the ramp, and merge was less than a minute. Due to the delay in the galvanic skin response to a given situation, it was decided that this time was too short to provide any meaningful data. Thus the heartbeat rate was taken as a better variable to use in determining driver stress. An example of a driver's pulse rate during a test run is shown in Figure 9.

A significant change in the heartbeat rate is considered to have occurred when there is an increase or decrease of 10 beats per minute. Because the primary interest is stress, only changes above the normal rate are taken into account. Defining each time the driver's rate exceeds the point of significant change as a "stress-time", an index of driver comfort can be obtained. Table 7 is the stress-time matrix obtained from the second series of tests.

The heartbeat rate for each subject was taken for a full minute before and after his test runs. The mean of these observations was accepted as the normal rate for that driver.

The only factor that exhibited a significant difference between levels was Factor F (drivers). This significance is an indication that an additional study should be made in the area of driver stress in dynamic traffic situations. Table 8 is the analysis of variance for this factor.

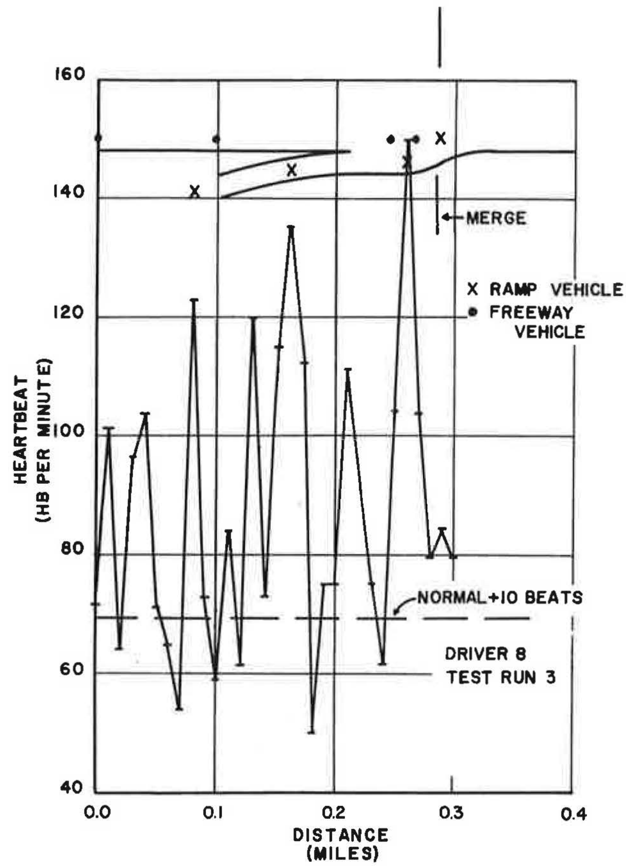


Figure 9. Typical driver's pulse rate during test run.

TABLE 7

STRESS-TIME MATRIX, INSTRUMENTED VEHICLE SERIES

Driver	Test Run 1	2	3	4	5	6
1	4	4	2	5	16	6
2	8	7	8	7	13	9
3	7	2	8	3	4	2
4		11	10	7	8	10
5	9	12	6	8	7	7
6	9	3	3	5	12	4
7	12	6	8	12	10	4
8	13	12	21	17	15	15

TABLE 8

ANALYSIS VARIANCE, SECOND SERIES

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Total	47	804.48		
Drivers	7	526.65	75.24	10.83[a]
Error	40	277.83	6.95	

[a]Significant at the 1 percent level.

CONCLUSIONS

The feasibility of the controlled moving merge was confirmed by both series of this study. The first series of tests placed the ramp driver in a position to merge 70 percent of the time. The instrumented car series was successful more than 50 percent of the time.

These tests have demonstrated that the controlled moving merge can be accomplished successfully and safely. The factors of backsight, driver, freeway speed, and ramp signals have been determined to be sufficiently important to warrant further research. The installation of a moving vehicle merging control system on a freeway would provide the opportunity for more detailed investigation as well as permit the study of the system under real-world conditions.

Subsequent research should also take into account principles of operation of a moving merge control system. It would appear necessary to determine under what conditions a controller should cease to function in the stopped vehicle merging control system and change to the moving vehicle system or vice versa. Some criteria should be developed for the determination of the benefits of 24-hour ramp control.

The second test series indicates that stress-time can be used as a factor in the indexing of driver comfort. If this approach to the evaluation of the level of service is to be accepted, further research into this area is required.

It is questionable as to the amount of driver stress desirable. Minimization of this stress might endanger the driver by reducing his level of alertness. These are some of the basic problems that should be considered in the future.

REFERENCES

1. Buhr, J. H., Whitson, J. H., Brewer, K. A., and Drew, D. R. Traffic Characteristics for Implementation and Calibration of Freeway Control. Paper presented at the 48th Annual Meeting and included in this Record.
2. McCasland, W. R., Buhr, J. H., Carvell, J. D., and Drew, D. R. Analog Merging Controllers for Freeway Entrance Ramps. Research Report RF 504-3, Texas Transportation Institute, 1968.
3. Satterthwarte, F. E. Random Balance Experimentation. Technometrics, Vol. 1, No. 1, No. 2, p. 111-137, May 1959.
4. Pinnell, C., Drew, D. R., McCasland, W. R., and Wattleworth, J. A. Inbound Gulf Freeway Ramp Control Study I. Research Report 24-10, Texas Transportation Institute, 1964.
5. Drew, D. R. Stochastic Considerations in Freeway Operations and Control. Research Report 24-5, Texas Transportation Institute, 1965.
6. Radke, M. L. Freeway Ramp Control Reduces Frequency of Rear-End Accidents. Research Report 24-21, Texas Transportation Institute, 1966.
7. Buhr, J. H., Drew, D. R., McCasland, W. R., and Whitson, J. H. Design Specifications for Freeway Surveillance and Control Systems. Research Report 504-7, Texas Transportation Institute, 1968.
8. Drew, D. R., McCasland, W. R., Pinnell, C., and Wattleworth, J. A. The Development of an Automatic Freeway Merging Control System. Research Report 24-19, Texas Transportation Institute, 1966.

Design of Freeway Entrance Ramp Merging Control Systems

JOHANN H. BUHR, WILLIAM R. McCASLAND, JAMES D. CARVELL, and DONALD R. DREW, Texas Transportation Institute, Texas A&M University

The central concept in the control of a freeway system is the control of each individual entrance ramp. This report deals with the control of individual entrance ramps—in particular, with the design and installation requirements of a gap acceptance merging control system. The operation of such a controller and the detection and display requirements are discussed in in sufficient detail to allow the writing of functional specifications and the design of a merging control system.

•IN March of 1966, the first prototype automatic traffic-responsive merging controller was installed on the Telephone Road inbound entrance ramp of the Gulf Freeway in Houston. Since then, considerable experience has been gained in the operation of this instrument and in the prediction of its effects on traffic behavior. The prototype controller has led to the development of eight first-generation controllers that have been installed on the inbound entrance ramps of the Gulf Freeway. These controllers are basically of two types—demand-capacity controllers and gap acceptance controllers. Although certain restrictive site characteristics and traffic-flow conditions may dictate the use of a demand-capacity controller, the gap acceptance controller is generally more desirable. Under contract with the Bureau of Public Roads, the Texas Transportation Institute has developed functional and operational specifications for gap acceptance control systems, some of which are discussed in this paper. Based on these specifications, three new second-generation controllers are presently under construction.

Because sufficiently detailed documentation has lagged far behind the rapid development of merging control systems, much of the associated technology is limited to relatively few researchers. The purpose of this paper is to narrow the gap between technology and documentation. It is hoped that it will enable the traffic engineer to write the functional specifications and design the installation of a merging control system suited to his own requirements.

THE CONTROL FUNCTION

Basically, the function of a freeway control system can be qualitatively stated as consisting of two objectives: (a) the optimization of freeway operation and (b) the optimal use of freeway gaps by merging ramp vehicles. Pursuant to these objectives, a multilevel design approach has been introduced as a means of decomposing the control function for a rational development of a freeway control system (1). This multilevel concept is directed toward establishing a hierarchy of levels of control, partly based on varying degrees of sophistication. The four suggested levels in ascending order of sophistication are the regulating, the optimizing, the adaptive, and the self-organizing control functions. This paper is concerned with controllers of up to the second, or

Paper sponsored by Committee on Freeway Operations.

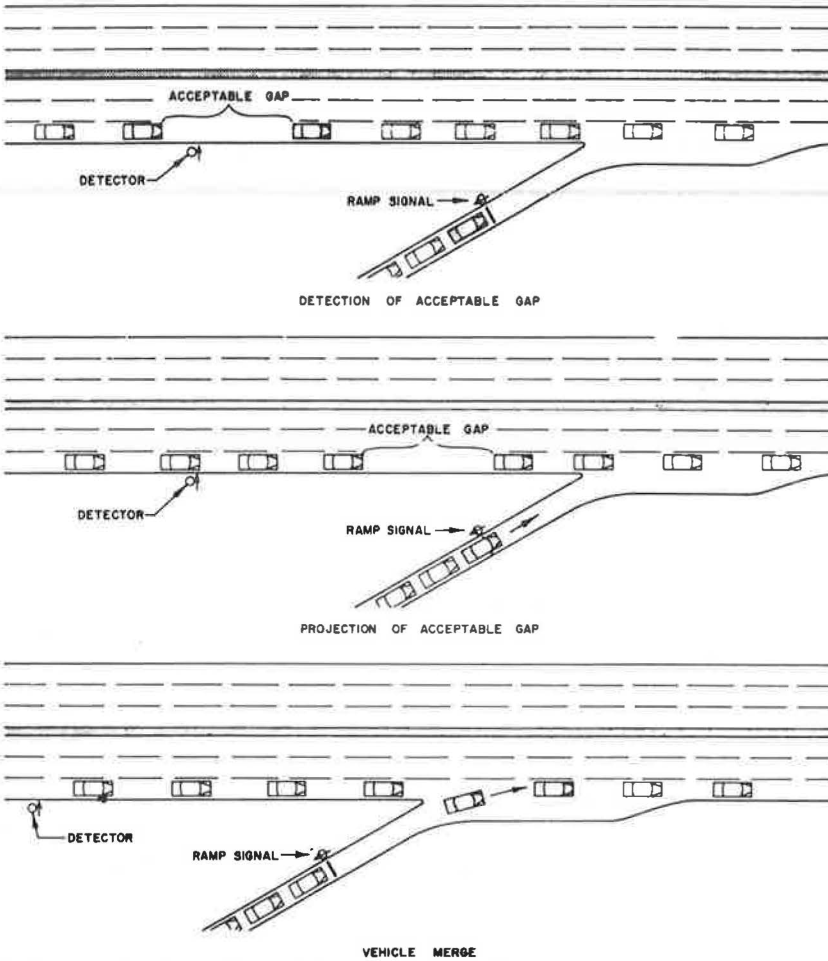


Figure 1. The three steps of entrance ramp control.

optimizing, level of control. Each controller operates an "isolated" merging area, or subsystem, without interconnecting or regard for the operation in other subsystems.

The gap acceptance concept is the basic or central principle of the multilevel control scheme. The higher levels of control exert their influence on the traffic process by modifying the basic control parameters, i.e., the service gap (2). Gap acceptance control is achieved by a controller that measures gaps and speeds in the outside freeway lane while a vehicle desiring to enter the freeway is stopped at a traffic signal on the ramp. When a gap large enough for the ramp driver to enter is found, the ramp vehicle is released by turning the signal indication to green at the proper instant so that the ramp vehicle will reach the merging area at the same time as the moving gap. This procedure, shown in Figure 1, can be thought of as consisting of three steps: gap detection, gap projection, and gap acceptance.

OPERATION OF CONTROLLER

To achieve the operation described, the controller must be able to measure the time headways between vehicles at a location upstream from the merging area to the extent that it can determine if a certain gap is larger or smaller than a predetermined gap. Furthermore, it must have the capability of measuring the speed of vehicles on the outside lane of the freeway, the capability of changing the indication of a ramp display such as a traffic signal, and the capability of determining through gap projection based on the speed measured the correct moment to change the signal indication. Each of these functions is discussed in greater detail later in this paper. These are the basic functions of the gap acceptance controller. However, to operate in a real-life environment, the controller must have a number of other auxiliary functions, the first of which is concerned with the operation of the ramp signal.

Operation of Ramp Signal

Assuming that the ramp signal is a conventional green-amber-red traffic signal, there are three ways that it can be operated: (a) the signal gives a green indication at the proper time whether or not there is a vehicle waiting; (b) the signal normally rests on green when there are no vehicles waiting; or (c) the signal normally rests on red when there are no vehicles waiting.

If, under the first alternative, a platoon of vehicles approaches the ramp signal, which is green but turns red just before the arrival of the lead vehicle, a hazardous condition will be created by an instrument designed to enhance safety. This operation has been observed in practice on several occasions and is deemed quite undesirable. It is therefore necessary to have a detector, the check-in detector, located immediately upstream from the signal in order that the controller might recognize the presence of a vehicle waiting at the signal.

The second alternative has been suggested to guard against failure of the check-in detector to indicate the presence of a waiting vehicle, either by virtue of detector failure or incorrect location of the sensor. This alternative means that only if a vehicle arrives should the signal turn red to stop the vehicle, and then it releases the vehicle in the normal manner. This operation is conceptually unappealing. Not only will it create the same hazardous conditions as the first alternative but it will also let many vehicles pass the signal uncontrolled unless an additional upstream detector is used, thus making the control system more complex than necessary.

The third alternative, with the signal resting on red and the controller providing a green signal only if there is actually a vehicle waiting, is the most desirable and is being used on the Gulf Freeway. This requires the use of a check-in detector, failure of which can be guarded against as mentioned later.

To turn the signal back to red after a vehicle has been released, a check-out detector, placed immediately downstream from the signal, was used at first on the Gulf Freeway. The signal was turned to red as soon as the released vehicle actuated this detector. Although conceptually appealing, problems were caused by the variability in the distance from the signal at which ramp drivers stopped (3). Sometimes a vehicle will stop with its front wheels past the signal, actuating the check-out detector, and, if the check-out detector is moved further downstream, two-vehicle platoons will often pass the signal. It was found far more desirable to turn the signal back to red, based on fixed lengths of green and amber phases, without using a check-out detector (3). As noted later in this paper, however, such a detector is required for the detection of slow vehicles. It also forms part of the detection system for the real-time evaluation of freeway operation (4).

Control Overrides

Based on the regulating function as described, the control system should theoretically operate adequately under all conditions. However, in a real-life installation, there are three extraneous conditions that will affect this operation: (a) vehicles stopping in the merge area, (b) ramp queues backing into other traffic systems, and (c) excessively long red phases. The effect of these three possible occurrences can be minimized by special override features that tend to maintain the desired operation, and, if the operation fails, they tend to return the control to the desired mode of operation.

There are a number of conditions that can cause drivers to stop in the merging area. When this happens, releasing another vehicle will simply compound the problem because the released vehicle will also be forced to stop and wait for an even bigger gap for which, in turn, another vehicle would have been released. It is therefore necessary to hold the ramp signal on red or to increase the service gap considerably until the merge area has been cleared. The occurrence of this situation should be detected with a presence detector—the merge detector placed in the merge area.

Furthermore, the ramp may be so located that an excessively long queue at the ramp signal will back into an intersection of the frontage road with a cross-street, thus adversely influencing an adjacent traffic system. To minimize this interference with off-freeway traffic systems, it is necessary to detect such an occurrence with a suitably placed presence detector (the queue detector), which, if continuously occupied for longer than a certain period, will cause vehicles to be metered at a somewhat faster rate by reducing the service gap. At locations where the ramp demand is fairly low and/or there is ample storage space for queues, this feature may not be necessary.

If a vehicle is delayed at the signal for longer than a certain period, the driver will probably assume that the signal is out of order and proceed past the signal, violating the control. This period varies among drivers, but is considerably shorter than at a traffic signal on a regular surface street intersection probably because of the somewhat unconventional location of the signal and the absence of any immediate danger in violating the signal. More often than not the violating driver is forced to stop in the merge area. It is therefore necessary to have a maximum red phase, insuring that the signal will turn green at least once every so often. In practice, on the Gulf Freeway, a 20-second maximum waiting time is used. When a vehicle has occupied the check-in detector for longer than 20 seconds, it is given a green signal regardless of the availability of gaps. However, it is considered more desirable to reduce the service gap gradually as the waiting time at the signal increases (5). Furthermore, this feature should be based on the length of red phase rather than the occupancy of the check-in detector. This would protect the control against failure of the input detector to sense the presence of certain types of vehicles, such as motorcycles or high trucks.

These functions, considered as overrides or optimizing functions in the multilevel control approach, affect control by modifying the service gap given by the regulating function (5).

Trucks and Slow Vehicles

At locations where a relatively high percentage of the ramp traffic consists of trucks, it may be desirable to make special allowances for this type of vehicle. These allowances can be either releasing a truck for a bigger gap than a regular passenger vehicle or providing a longer ramp travel time by turning the signal green earlier, or both (3). This would require some specialized detection equipment to sense the presence of a truck at the signal. However, since the performance of some trucks, especially when empty, may approach that of regular passenger cars, it may be necessary to sense the weight/horsepower ratio of vehicles waiting at the signal. It is doubtful that this feature can be made cost-effective.

A better way of handling slow vehicles is to provide for the detection of vehicles at a point between the ramp signal and the merge detector. The travel time of a released vehicle from the signal to this detector can then be measured, and, if greater than a predetermined value, the signal can be held on red, or the service gap increased until the vehicle has cleared the merge detector or until the merge detector overrides the operation.

Multi-Vehicle Metering

When the ramp signal turns green, it stays green for a fixed period long enough to release a single vehicle. The signal then turns amber (where allowed by state law, it may be desirable to omit the amber phase) for a short fixed period and then turns back to red. It is necessary that it stay red at least long enough to give the second vehicle in line time to pull up to the signal. This minimum cycle length should be about 5 sec but not less than 4 sec. As a result, the maximum number of vehicles that can be metered under ideal conditions is approximately 720 per hour. When ramp demand approaches this value and there are many large gaps available on the freeway, it is desirable to employ some form of multi-vehicle metering that permits more than one vehicle to pass the signal on the same green phase.

Multi-vehicle metering has not been tested extensively on the Gulf Freeway, but some experiments are now being planned. One experiment simply involves the use of longer fixed-time green phases, designed to let two-vehicle units or three-vehicle units pass the signal in one phase. The other experiment involves an additional display, to be used in conjunction with longer green phases that would indicate the number of vehicles that should proceed down the ramp. If multi-vehicle metering is planned, the controller should have the additional capabilities of recognizing gaps larger than several predetermined values and of providing several green phase lengths and perhaps amber phase lengths. The use of multi-vehicle metering will also affect the location of the gap detector, as discussed later.

Service Gap

The service gap is defined as the smallest gap for which the controller will release a vehicle. Unless one of the previously discussed overrides is in operation, this gap would normally not be smaller than the "acceptable" gap of ramp vehicles. The service gap should increase as the freeway speed decreases (2, 6). Once the freeway reverts to forced flow conditions and speeds drop below 20 mph, the service gap should be sharply increased, because even a small space between vehicles constitutes a large time headway. For example, if the freeway should come to a complete stop, the gap over the gap detector will look like an infinitely large gap, and the controller will accordingly release many vehicles off the ramp unless it is designed to handle such an eventuality. In the local controllers presently being used, the regulating function was such that, if the speed dropped below a preset value, vehicles were metered at a fixed rate. This critical speed was based simply on the subjective evaluating of the day-to-day operation of the freeway and was usually set at about 25 mph. It was felt that under forced-flow conditions there was little need for projecting gaps, since they would probably be so unstable as to make it impossible to fit vehicles into specific gaps. On the other hand, according to the latest thinking, it may still be preferable, even under these conditions, to attempt to fit vehicles into specific gaps that represent, in effect, units of capacity. Even if the time-space trajectories and the vehicle and the gap are not matched exactly, the unit of capacity will still be in the vicinity of the merging area to absorb the unit of demand. The function defining the gap size as based on freeway speed is presented in another report (5).

Apart from the functional requirements already discussed, the controller should have the capability of determining the size of the service gap, based on the speed measured on the freeway. In the existing local controllers, this is achieved by four level monitors, defining five speed ranges. The speed ranges can be defined by adjustable dials. When speeds are in the lower range, the controllers operate on a fixed-time basis. For each of the other four speed ranges a service gap can be set by adjustable dials. In this manner, the continuous service-gap speed-function can be approached by a four-step function. This appears to be a satisfactory manner of determining the service gap, although use of the continuous function would be more desirable.

DETECTION REQUIREMENTS

Number of Detectors

Figure 2 shows the functional components of the typical merging control system. As shown in the figure and as discussed, there are five detectors involved: the gap/speed detector (D_A), the check-in detector (D_i), the merge detector (D_m), the queue detector (D_q), and the slow-vehicle detector (D_t). All of these detectors would normally be used, but, as mentioned earlier, the queue detector may not be required at certain locations. Furthermore, the designer may decide not to use the slow-vehicle functioning, consequently omitting D_t. It should be kept in mind, however, that for system control of the freeway this detector doubles as a counting detector and is therefore quite important. The other detectors on the ramp are all presence detectors that cover a large area and are unsuitable for counting purposes.

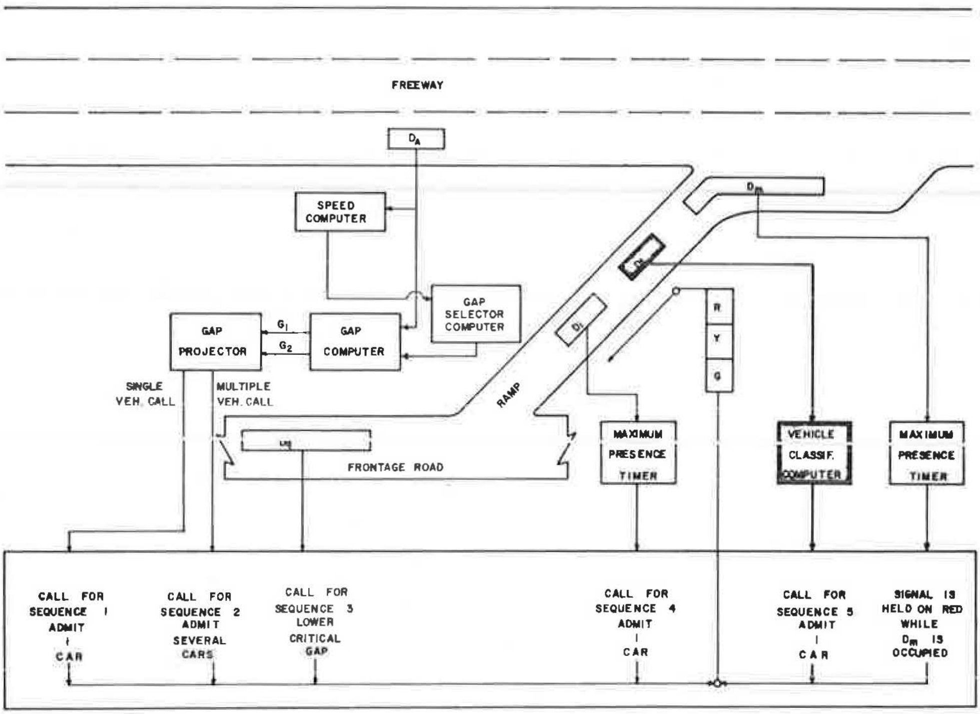


Figure 2. Functional components of the typical merging control system.

The detection of gaps and speeds can be accomplished with a single detector. However, since the speed is also used for estimating the arrival time of a gap in the merging area, the speed must be measured quite accurately. To achieve the desired accuracy, a set of two closely spaced detectors must be used at his location. Depending on the local requirements, therefore, the number of detectors used with a merging control system can vary between three and six.

Size and Location of Detectors

The size and location of a detector and the behavior of a driver are critical factors to be considered before most detectors can be expected to perform their intended functions. The intended function of the check-in detector is to indicate the presence of a vehicle at the signal. If a vehicle does not stop over this detector, the signal will remain red until the maximum red phase override is called in. The result of such a situation is the formation of long queues and the creation of long delays in the system. Studies indicate, however, that if the detector were of a minimum size of 10 ft long and 6 ft wide and no further than 5 ft from the signal, this situation would be all but eliminated (3). These are minimum dimensions. Larger values would be more desirable.

The size and location of the merge detection area also depend on driver behavior, which is largely influenced by the geometric design of the ramp and acceleration lane. Ideally, the merge detector should sense all vehicles stopped in the merging area. However, in practice, the position at which drivers stop was found to be so variable that even quite a large detection area fails to detect all stopped vehicles. Some drivers will stop 50 ft or more upstream from the ramp nose, while others will travel right down to the end of the acceleration lane. The latter type of operation does not constitute much of a problem because as long as a vehicle does not block entry to the freeway, the need for holding the signal on red is decreased. The best location of the merge detector can be fairly well established by observing the operation of an entrance ramp during a peak hour. In general, the detection area should stretch from some distance upstream from the ramp nose to about 40 ft downstream from the ramp nose.

The queue detector would normally be located on the frontage road. Its location depends entirely upon the geometric layout. In general, it should be located no closer than about 4 or 5 vehicle lengths from the traffic system that it is designed to protect. Its size should be so that the detection would not drop out between vehicles in a slowly moving queue—i.e., the next vehicle should normally enter the detection area before the previous one leaves it. This requires a length of at least 12 ft. However, if it is desired that the queue detector double as a counting detector for counting frontage road traffic, a point sensor can be used. In this case, however, the controller should have the additional capability of, in effect, extending the size of the queue detection area by a delay function. Such a function would hold a call from the sensor for a predetermined period after a vehicle leaves the detection area.

The purpose of the slow-vehicle detector, D_t, is to determine if a ramp driver's travel time, from the instant the signal turns green to the instant the vehicle actuates D_t, exceeds a predetermined value. When this happens, the controller should hold the signal on red until the merge override can be called in. The reason for using this function is that the merge detector is usually located so far from the signal that, by the time the merge override is called in, a second vehicle will often have been released. This second vehicle is then also forced to stop, perhaps in advance of the merge detector. When the first vehicle then leaves, the average override drops out and a third vehicle is released while the second is still stopped in the merge area. The controller can have some difficulty in recovering from this situation. D_t therefore allows the prediction of a stop in the merging area and tends to avoid the situation of two vehicles stopped in the merge area. As such, it is a highly desirable feature of the system.

To perform its intended function properly, D_t must be located so that a ramp vehicle, using the correct acceleration to hit the gap, would actuate it within a time interval representing the minimum signal cycle length. Furthermore, it is necessary that this sensor distinguish between separate vehicles. It therefore has to be a point sensor longitudinally and cover a lateral width of at least 6 ft. The location of D_t with respect to the signal thus depends on ramp-vehicle operating characteristics (3). When this feature is used, it may be desirable to have a cutoff period because sometimes a driver will miss his green signal and not proceed down the ramp. This would cause the slow-vehicle feature to be called in when there is actually no vehicle on the ramp. However, since the slow-vehicle feature has a built-in cutoff period, a special cutoff feature will be of value only with certain limited geometric layouts.

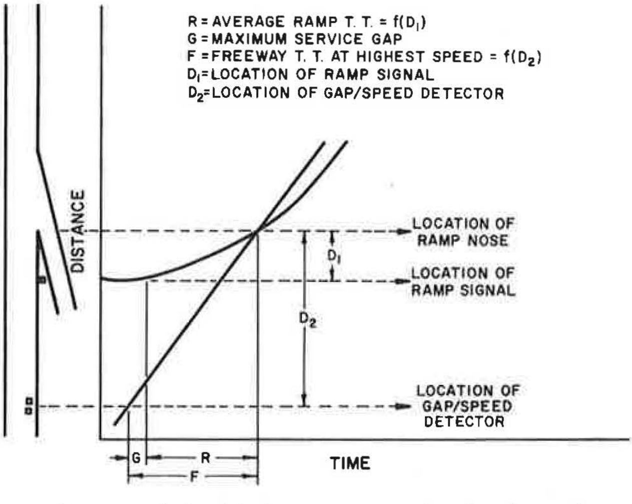


Figure 3. Time and space relationship between ramp signal and gap detector locations.

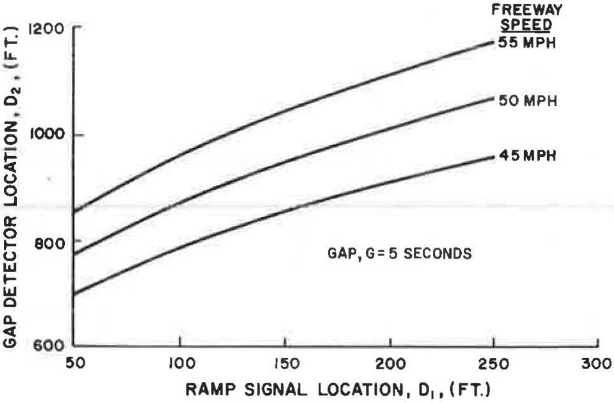


Figure 4. Gap detector and ramp signal locations for metering passenger vehicles.

The location of the gap/speed detector, D_A, should be determined in conjunction with the location of the ramp signal. The distance between the gap detector and the ramp nose should not be less than the distance covered by a freeway vehicle, at the highest speed expected during control, during the time required for a ramp vehicle to travel down the ramp plus the time required to measure a gap. In other words, the closest permissible location of the gap detector is such that the freeway travel time equals the average ramp travel time plus the maximum service gap. This relationship is shown in Figure 3. The ramp travel time depends on vehicle operating characteristics and is a function of the ramp signal location (3). When the truck control function is used, the ramp travel time should be that of the average truck and also the gap, G, should be the acceptable gap for trucks. The use of multivehicle metering will also increase the maximum service gap, requiring that the gap detector be located further upstream.

Figure 4 shows the trade-off between ramp signal location, gap detector location, and freeway speed based on the ramp travel time characteristics determined on the Gulf Freeway (3). In many cases the location of the gap detector will be limited by an upstream off-ramp, thus determining the location of the ramp signal. In other cases, the location of the ramp signal may be limited by a short on-ramp, thus determining the position of the gap detector. Certain geometric configurations may even preclude gap acceptance control, except at low freeway speeds. At locations where no limitations exist, the position of the ramp signal should be established first and the gap detector then positioned accordingly. In general, the ramp signal should be so located that the check-in detector will not be actuated by vehicles that travel along the frontage road without entering the ramp. The signal should be as close to the nose as possible in order to reduce the variability in ramp travel times, but should be far enough away that ramp vehicles can attain fairly high speeds. A ramp signal located 150 to 200 ft from the nose yields quite satisfactory operation.

As discussed, the gap detector location is based on the highest freeway speed at which control is to be exercised. When the speed drops below this value, the travel time between the gap detector and the ramp nose increases, thus introducing an additional time factor. This additional time, termed the "delay" or "projection time", discussed later in this paper, is the time between the instant at which the gap is determined to be larger than the service gap and the instant at which the ramp signal indication should be turned to green.

On the Gulf Freeway, the gap sensors are 6 ft by 6 ft loops. Where double loops are used to sense speed, the leading edges of the loops are spaced at 18-ft intervals. These sizes operate fairly satisfactorily, but problems can be created when a lane-change takes place over the loop pair. Furthermore, where a 6 ft by 6 ft loop doubles as a counting detector, it will undercount when stop-and-go conditions occur because a vehicle may enter the detection area before the previous vehicle has left. Point sensors would therefore be preferable.

Because the gap detectors are located some distance upstream from the ramp nose, gaps and speeds in the merging area may differ considerably from their measured values (3). In order to minimize the effect of this instability, control systems involving four pairs of detectors upstream from the merging area are now being installed at three ramps on the Gulf Freeway. These systems will switch to a different loop pair when the freeway speed changes by a certain amount, thus always using a loop pair as close to the merging area as possible. The success of these controllers has not yet been evaluated.

GAP PROJECTION

In attempting to fit a ramp vehicle into a gap on the freeway, the controller must, after determining that a gap is at least of a certain size, "project" this gap downstream for a certain time period (determined by the speed of the gap) before turning the signal to green. In other words, upon observing an acceptable gap, the controller turns the signal green after a certain "projection time" or "delay". This delay depends, of course, on the ramp travel time, the location of the gap detector, the speed of the gap, and the size of the critical gap. Its magnitude, D, should theoretically be such that the travel time, T_f, of the lead vehicle of the gap, between the gap detector and the nose, equals the ramp vehicle's travel time, T_r, from the instant that the signal turns green until it reaches the nose, plus the size of the service gap, T, plus this delay—i.e.,

$$T_f = T_r + T_s + D \tag{1}$$

Since the ramp travel time is a constant and the service gap can be included in the projection time ($D = D + T_s$), the projection time is a function of the freeway travel time and, therefore, a function of the freeway speed and of the gap detector location:

$$D = f\,(T_f) = f'\,(u_f, D_a) \tag{2}$$

For any fixed gap detector location, the projection time is, therefore, a continuous function of the freeway speed. However, since the gap detector location will change from one ramp to another, it is necessary that the controller be designed so that the gap projection vs freeway speed function can be adjusted to fit a certain detector layout, rather than be specifically designed to fit a certain detector layout, or, alternately, the detectors be placed so as to conform to the controller.

In the first-generation analog controllers now in use, this projection time, or delay, is handled by means of four level monitors. These same level monitors are used to define the service gap vs freeway speed function described earlier. Each speed range, defined by two level monitors, has associated with it a delay setting that can be adjusted with a dial. As a result, Eq. 1 is not always truly achieved, but is, in fact, approximated by a four-step function.

In attempting to fit the time-space trajectory of a ramp vehicle into that of a freeway gap, several errors can arise over which the system has a certain degree of control: (a) errors in the predicted freeway travel time caused by speed instability, (b) errors caused by the variability in ramp travel times, and (c) errors caused by approximating the delay function by a step function. Errors of the third type can be completely eliminated by using a digital computer or an analog controller that can solve the projection time vs speed function. It can also be reduced by using more level monitors. Errors caused by speed instability can also be considerably reduced by projecting vehicles rather than gaps (3). This, however, increases the complexity of the controller. In other words, the error can be reduced at increased cost. The designer should weigh the effectiveness of reducing these errors against the cost of reducing them. The best accuracy that can be attained is defined by the distribution of ramp travel times.

Studies performed on the Gulf Freeway have indicated that when a level monitor system is used the speed groups may, for all practical purposes, be equally spaced, with the projection time set for a speed at the midpoint of each speed group. Disregarding errors introduced by speed instability, a four-level monitor system, is, at best, about 80 percent as effective as a system using the theoretically correct projection time. The second-generation controllers now being installed on the Gulf Freeway do not make use of level monitors, but project gaps based on the actual speed of the lead vehicle of the gap.

DISPLAY REQUIREMENTS

The displays used in conjunction with merging control systems consist of two types: displays serving to indicate the operation of the controller itself and displays serving as driver communication devices. Displays of the first type are usually mounted on the face of the controller and are required to visually monitor the operation of the control system. In a strictly operational (as opposed to research) instrument, such displays can be minimized. However, it is advisable to have at least a small pilot light indicating the operation of each controller feature and the actuation of each loop. This greatly enhances troubleshooting and facilitates maintenance and checking out of the instrument. The driver communication devices used to date consist essentially of two types: ramp signals and advance warning signs.

Ramp Signals

The ramp signals used on all three freeway surveillance and control projects (Houston, Chicago, and Detroit) are essentially standard traffic signals. The signals used in Houston and Detroit are mounted on the left-hand side of the ramp and bear the legend "STOP HERE ON RED" (Fig. 5). This legend is quite important. On the Detroit project, the legend "STOP HERE ON SIGNAL" was used at first, but later changed because it apparently caused considerable confusion. On the Chicago project, two signals are used on each ramp, one placed on either side. These signals have only the red and green indications and bear the legend "WAIT HERE FOR GREEN". Operation without the amber phase is actually preferred, but some states require that the signals must conform to the code on uniformity of standard traffic control devices because they have all the appearances of conventional traffic signals. It is felt that a new type of ramp signal should be designed and tested—one that is completely different from any standard traffic control device and is simple yet effective. The ramp signal is in some respects the most important single component of the control system because it is the device that actually exerts the control over traffic operation. Surprisingly little research effort has been devoted to its design and effectiveness.

In Chicago and Detroit, stop-lines are used in conjunction with each signal. However, studies performed at several ramps on the Gulf Freeway revealed that stop-lines have

Figure 5. Ramp control signal used on the Gulf Freeway.

Figure 6. Advance warning sign used on the Gulf Freeway.

no effect on the stopping behavior of ramp vehicles. Consequently, stop-lines are not used on the Houston project.

The location of the signal with respect to the ramp nose has been discussed earlier. The signal head should be mounted so that it is approximately at a driver's eye height, and it should be turned so as not to be readily visible from the freeway.

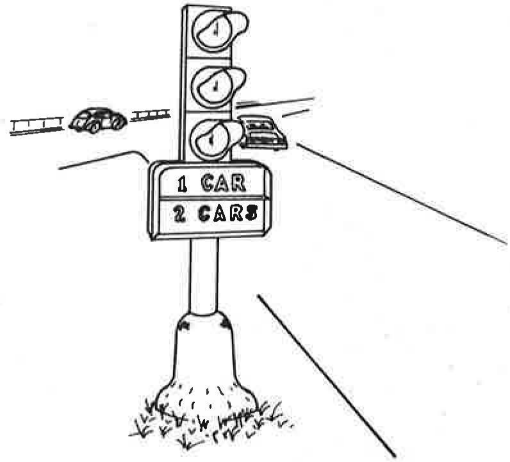


Figure 7. Experimental ramp control signal for multi-vehicle metering.

Advance Warning Sign

On all three projects mentioned, some type of advance warning sign is placed upstream from the ramp signal. The type used in Houston is shown in Figure 6; it consists of a black legend on a yellow diamond-shaped background, with two alternately flashing amber signals. The signs used in Chicago and Detroit are essentially similar, except that the legend on the Chicago signs reads "RAMP SIGNAL AHEAD." In the absence of any evaluation of the effectiveness of advance warning signs, it is felt that they are generally desirable.

Multi-Vehicle Metering

As mentioned earlier, the metering of multi-vehicle platoons has not been tested seriously to date. An experiment planned on the Gulf Freeway will involve the use of a ramp signal (Fig. 7). This signal still includes the conventional traffic signal, but will have an additional signal to indicate that one or two vehicles should proceed past the signal.

Figure 8 shows a suggested ramp control signal for multi-vehicle metering. It has a red "STOP" display and three green displays indicating that one, two, or three vehicles should proceed past the signal. When the controller is ready to meter a three-vehicle platoon, the signal will indicate "3 CARS" and, then, on a fixed-time basis, turn to "2 CARS", "1 CAR", and back to "STOP". This display is suggested as a simple, potentially effective, nonstandard signal.

Another proposed method of multi-vehicle metering involves the metering of vehicles alternately from two parallel queues. Such a method will usually require reconstruction of an existing ramp.

Figure 8. Suggested ramp control signal for multi-vehicle metering.

CONTROLLERS

The merging controllers developed and tested on the Gulf Freeway were all of the analog variety. The first prototype controller (Fig. 9) installed on the Telephone Road inbound entrance ramp of the Gulf Freeway in March 1966 was quite an elaborate instrument. It included three sets of level monitors with ten levels in each set, a gap computer, a gap projector, and a signal controller, together with various displays of speeds, volume, gap projection, and so forth. From this prototype controller, the first-generation controllers were developed (Fig. 10) and installed on the Gulf Freeway in October 1967. These controllers operate essentially as the pro-

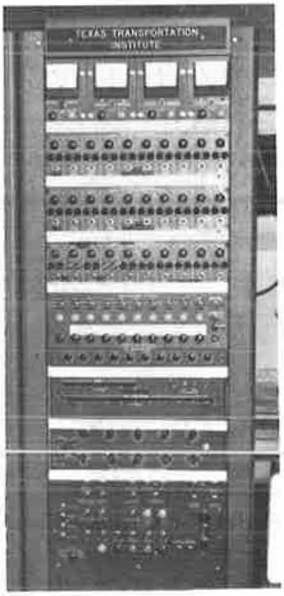


Figure 9. Prototype merging controller.

totype controller, but many features found unnecessary on the prototype have been eliminated. These controllers utilize four level monitors to select the service gap and projection time at various speed ranges, as discussed earlier. Furthermore, they consist of gap computer/projectors and signal controllers.

The construction of three new second-generation controllers was initiated in November 1967, but they had not been completed by the time this report was written. These new controllers will project gaps based on the actual speed of the lead vehicle of the gap, rather than on the speed range of the average freeway speed, as is the case with the first-generation controllers. Speeds and gaps will be measured on the freeway with four sets of closely-spaced detectors. The controller will automatically select one set to direct the operation at any one time, based on the freeway speed prevailing at the time. It will also include the multi-vehicle, slow-vehicle, and truck metering features discussed earlier. These controllers will be quite versatile in that all control variables will be adjustable and all optional features will be switch-controlled so that they can be turned off. As a result, this research instrument can be thought of as several controllers built into one chassis. It will be used to study the effectiveness of controllers of varying degrees of sophistication.

In July 1967, an IBM 1800 digital computer was installed in the control center (Fig. 11) in order to develop, implement, and evaluate system control measures on the freeway that would operate the freeway as a system rather than as eight independent entrance ramps. Computer control on a limited basis was initiated in March 1968. The versatility of this instrument has led to the hypothesis that even isolated entrance ramps can be better controlled with digital equipment. Since both the process inputs and the process outputs are digital signals (contact closures), the control of the merging process is by nature a digital problem. There are many small digital computers on the market (small in the sense of limited computing power and physical size) that are ideally suited to the job. These small computers are, however, generally too powerful and expensive for the control of a single entrance ramp, but are believed to be an economically feasible way of controlling three to six entrance ramps, or groups of ramps in the same interchange. It is believed that a small programmable binary digital computer can be developed that would not only be economically competitive with analog equipment, but also would be more versatile

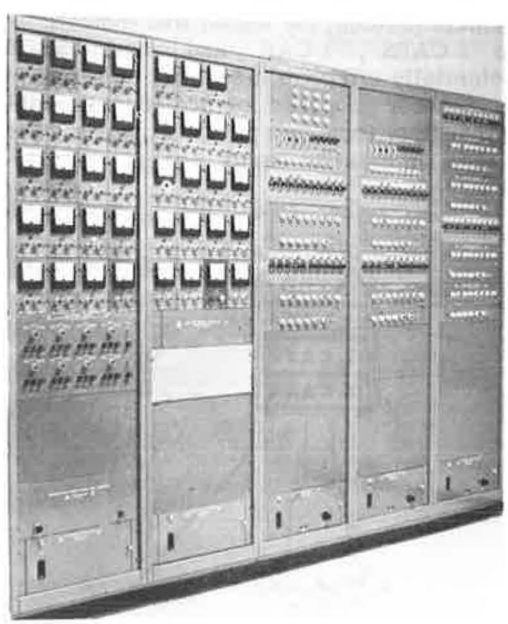


Figure 10. First-generation merging controllers.

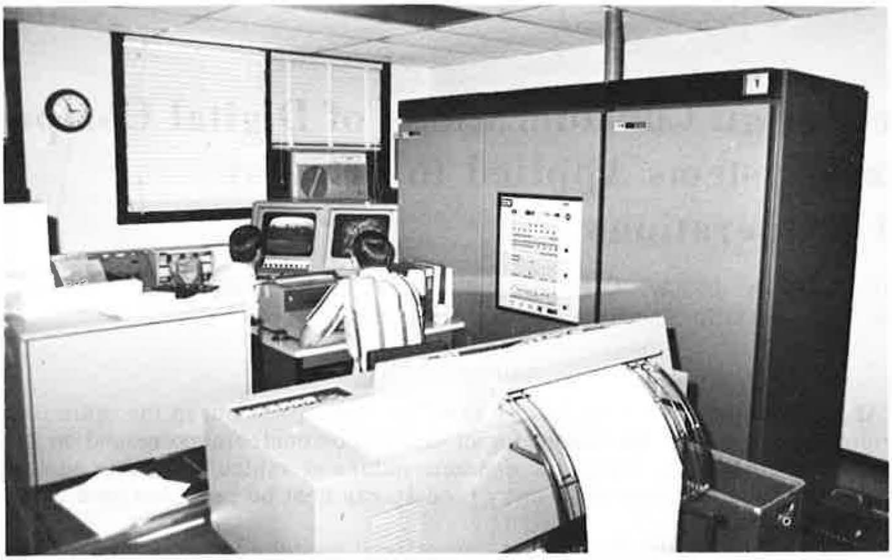


Figure 11. Digital computer data acquisition and control system in the Gulf Freeway Surveillance and Control Office.

and more effective. Research into the development of such a controller is strongly recommended.

REFERENCES

1. Drew, D. R., Brewer, K. A., Buhr, J. H., and Whitson, R. H. Multilevel Approach to the Design of a Freeway Control System. Paper presented at 48th Annual Meeting and included in this Record.
2. Brewer, K. A., Buhr, J. H., Drew, D. R., and Messer, C. J. Ramp Capacity and Service Volumes Under Merging Control. Research Report 504-5, Texas Transportation Institute, 1968.
3. Buhr, J. H., Whitson, R. H., Brewer, K. A., and Drew, D. R. Traffic Characteristics for Implementation and Calibration of Freeway Control Systems. Research Report RF 504-2, Texas Transportation Institute, 1968.
4. Whitson, R. H., Buhr, J. H., Drew, D. R., and McCasland, W. R. Real Time Evaluation of Freeway Quality of Traffic Service. Research Report RF 504-4, Texas Transportation Institute, 1968.
5. Buhr, J. H., Drew, D. R., Gay, J. N., and Whitson, R. H. Design Requirements of the Central Digital Computer in Freeway Surveillance and Control Systems. Research Report RF 504-9, Texas Transportation Institute, 1968.
6. Drew, D. R., LaMotte, L. R., Wattleworth, J. A., and Buhr, J. H. Gap Acceptance in the Freeway Merging Process. Highway Research Record 208, p. 1-36, 1967.

Some Design Considerations of Digital Computer Control Systems Applied to Freeway Traffic Operations

J. H. BUHR, D. R. DREW, J. N. GAY, and R. H. WHITSON,
Texas Transportation Institute, Texas A&M University

ABRIDGMENT

•BASICALLY, the objective of a freeway traffic control system is the optimization of freeway operation through the regulation of demand by controls exercised on the freeway entrance ramps. Because of the dynamic nature of vehicular traffic systems, the benefits of optimal utilization of freeway facility can best be provided by a real-time on-line control system.

This report concerns the design of a centralized digital computer control system for the surveillance and control of freeway traffic. It is primarily concerned with the analysis, definition, and development of system component requirements commensurate with the general nature of the control system specification.

The Basic Building Blocks

A freeway control system consists not only of the digital computer or central processing unit, but alsoof a variety of hardware devices. The equipment that constitutes an on-line control system can be divided into several classes, each consisting of various devices.

Data Processing Input/Output Units—The readers and printers serve as communication devices between the operator and the computer. The preferred freeway control system that is to serve almost exclusively as an operational (as opposed to research) device would have only a card reader and a typewriter as data processing input/output units.

Files—Control systems must store large amounts of information that can be rapidly retrieved for use in the control of the process. Such information is stored magnetically on tapes, drums, or disks. Of these, magnetic disks are the preferred files in the freeway control system.

Control Units—The control unit of the system is the central digital computer. It performs the calculations and logic and data manipulation and controls all the other parts of the system. At present, a single central computer will adequately serve the needs of control of all freeway control systems to be installed in the immediate future.

Communication Links—The transmission of data between the field sensors and the central computer forms a very important and highly specialized aspect of the control system. Its cost forms a substantial part of the total cost of a control system.

Process Control Input/Output Units—In order to control any process, the system must determine the status and environment of the process through its input units and then physically affect the process through its output units. In present freeway control systems, the input units are mostly loop detectors, and the output units are traffic signals placed on the entrance ramps.

Paper sponsored by Committee on Freeway Operations.

Some Control Considerations

The programming requirements of the control system can be thought of as consisting of four interacting subsystems: supervisory, control, data-acquisition, and analysis programs. These are discussed to acknowledge their existence and outline the purpose of each.

Supervisory Programs—In control systems it is often necessary to perform many functions simultaneously. There are communications to be handled, control parameters to be generated, and data to be recorded and displayed. All of this activity must continuously be scheduled and controlled by the computer as it is occurring. Priority decisions must be made moment by moment to decide what the computer will attend to next. This is the function of the supervisory program. It schedules the various parts of the control program and, when time and availability of storage permit, allows the compilation and execution of analysis programs not directly related to the control of the process. The supervisory program will also interrupt the execution of a certain task in order to service a higher priority job and, after servicing all higher priority interrupts, resume the execution of the original program.

Control Programs—The control programs consist essentially of the decision rules needed to improve the situation and the statements set up to change the status of the control devices designed to have the desired effect on the process.

Data Acquisition Programs—In order to achieve responsive control of a process it is necessary to continually measure certain parameters that describe the operation of the process. This is the duty of the data acquisition program. As such, it is a subsystem of the control program or the feedback required for the optimal control of a process. It also serves to evaluate and display the state of the operation for the benefit of the operating agency and can provide a permanent record of operation. The data acquisition subsystem further serves to detect equipment failures.

Response Time—Many of the considerations in the design of a process-control system are in some way connected with the response time of the system. Ideally, a short response time to all events is required. However, decreasing response times are directly associated with increasing costs. If a faster response time is desired, not only a faster computer, but also a larger computer and more rapidly accessible files may be required. In addition, faster response times may require additional transmission lines under certain circumstances.

Reliability and Standby Systems—The control system consists largely of a complex of sophisticated electronic equipment and as such is subject to periodic breakdown. Suitable arrangements must be made for this event by designing the system to modify its mode of operation when one of its components fail or, when necessary, to abandon a certain ramp to a simplified local controller which may provide for fixed-time control or simply a flashing amber sequence.

Design Considerations

Starting with a knowledge of the requirements, the objective of the design process is to find that configuration of functional elements which, translated into physical terms, meets the requirements as economically and efficiently as possible. Many factors enter into this choice of the proper configuration. The identification of the principal functions, the determination of their proper sequence, and the relative merits of centralization vs decentralization of control, for example, are but a few.

The design of a process control system is generally more complex than that of a conventional data processing installation. The hardware is more complex and many functions take place concurrently. The input to the system is not constant, orderly, and precisely predictable. The load on the system varies with time and occasionally reaches such peaks that processing is delayed. The penalties for errors in the design are also greater. If the load on a control system exceeds its capacity it may lose control of the process. In the unabridged paper, a step-by-step design procedure is outlined and each of the design requirements is discussed.

Implementation

The design and installation of a control system for purely operational purposes does not require lengthy data collection procedures by a large staff to establish a complete description of the operating characteristics of the freeway. With a knowledge of the requirements and procedures, a control system can be installed and made operational in considerably less than one year.

Ramp Control on Freeways in California

G. L. RUSSELL, California Division of Highways

ABRIDGMENT

•IN January 1969, three freeway ramp control projects were operating in California.

Harbor Freeway

The major project is on the 8-lane Harbor Freeway in Los Angeles. Traffic volume is about 190,000 vehicles per day. The control section is 5 miles long and includes six on-ramps. Five ramps are controlled by preprogrammed traffic signals and one ramp is simply closed during the peak period, from 3:30 to 5:30 p. m. During the rest of the day (and night), the signals remain green. Input from the controlled ramps was reduced by about 1500 vehicles during the two hours but input from the mainlane upstream was increased by about 1000. Net diversion to the city street system was only about 500 in 2 hours.

Freeway speeds have increased from about 20 mph to about 40 mph. Freeway users are saving about 1000 vehicle hours per day against a loss, or increased travel time, for diverted or delayed ramp traffic of about 130 vehicle hours. The effect of the added load on streets is too small to notice or measure.

Capacity (7200-7500 vph) was not increased, but throughput upstream of bottlenecks was increased. Nonrecurrent incidents affect the system about one day out of two, but most of them are cleared up soon enough to avoid spoiling the whole operation. It is planned to go to traffic-responsive control in order to take advantage of the unused capacity downstream of incidents when they occur.

Some unusual techniques being employed are (a) preferential treatment for buses—they are allowed to bypass ramp queues and to make left turns where other traffic is not; (b) storage of queued vehicles on a frontage road—it is carried back across an intervening major thoroughfare by timing the intersection signal so that the intersection itself remains clear; (c) two-abreast release of vehicles at the ramp signal at one location increases storage room and makes possible high metering rates.

Total cost of the Harbor Freeway project was $52,000. Annual operating cost is nominal (electricity and signal maintenance); annual benefits about 200,000 vehicle hours.

Hollywood Freeway

The Hollywood Freeway project consists of one metered ramp and one ramp that is closed by barricades during the peak period. The metered ramp is operated by a pretimed signal. Travel time savings are about 450 vehicle hours per day, consisting of 500 hours savings for freeway users upstream of the bottleneck less 50 hours for loss to diverted or delayed ramp traffic.

Chula Vista Project

Four ramps on 3.2 miles of a 4-lane freeway in Chula Vista are being controlled. Peak-hour input from these four ramps was reduced by 580 vehicles, and input to the mainline upstream of the control section was increased by 540 vehicles. Speed on the freeway increased from 26 mph to 43 mph for the higher volume. Traffic on parallel

---

Paper sponsored by Committee on Freeway Operations.

streets increased 225 vehicles. Controllers were designed and fabricated in the district shop. Total cost of the project was less than $5,000. This project was surveyed, designed, and executed by three men on the regular district traffic department staff in addition to their other duties.

Airport Demand Characteristics Affecting Freeway Operations

PAUL N. BAY, Wilbur Smith and Associates

ABRIDGMENT

•BELATED recognition of the magnitude of growth in automobile trip generation at major airports has caused some recent limited research at individual airports directed primarily toward solution of immediately pressing problems, such as parking requirements and terminal curb use. However, the studies to date have not sufficiently evaluated the question of impact on freeways and arterials, nor defined the problem adequately to predict future freeway demand levels at airports with much confidence. Further research is badly needed, but must be more comprehensive in nature. First, the airport must be recognized as more complex than past studies have indicated, and the separate functions occurring at the airport must be separately analyzed.

For example, airports are major employment centers, and the travel demands and trip characteristics for this function are entirely different from and only indirectly related to the primary function of transportation transfer terminal. A third function of major airports is as a place of commerce, sales, services, and recreation, and these trips are also different in character and should be treated separately.

Airport demand studies should also be directed toward the following problems:

1. The present relationships between daily air passenger and cargo volumes at individual airports and the volumes of vehicles required to serve those activity levels. Care must be taken in developing these parameters to evaluate sensitivity of such ratios to other factors, including the types of air trips served, the percentage of air transfers, the seasonal variations in air travel, and the role of the airport in the region. In some respects, each airport is unique. But what families of curves could be developed?
2. The time distribution of travel to airports (for each airport function) and the relationship of the time distribution to overall aviation patterns and air trip purposes. Can peaks be spread?
3. The possibility of separating airport functions spatially, to distribute vehicle demands to more than one interchange or freeway facility.
4. The sensitivity of all the foregoing relationships to new aviation technology, including "jumbo" aircraft, VSTOL development, supersonic travel, and improved ticketing and baggage handling.

The growth in air travel has made airports some of the most important traffic generators in metropolitan regions. To understand fully the implications for freeway travel of this growth and the future technology in aviation, individual airports must be analyzed by their separate functions, and their travel patterns thoroughly understood. Even though individual airports will vary in their demand characteristics, parameters should be established based on continuing research that will enable freeway planners to better foresee the needs of the future.

Paper sponsored by Committee on Freeway Operations.

Analysis of Wrong-Way Incidents on Michigan Freeways

CHARLES L. RICHARD, Michigan Department of Highways

ABRIDGMENT

•THIS study covers a cooperative data-gathering program between the Michigan Department of State Highways and the Michigan State Police to record and analyze wrong-way incidents and accidents on Michigan's rural freeways in the southern portion of the state. The report provides an analysis of 200 wrong-way incidents and 44 wrong-way accidents that occurred in 1966.

A summary of the results is as follows: 50 percent of the wrong-way accidents were caused by drinking drivers; non-drinking drivers appeared to drive wrong-way equally both day and night, and at times of higher volume periods; drinking drivers appeared to drive the wrong-way during the dark hours and at the lower volume periods; high nighttime and weekend incidents were related to the drinking driver incidents; 42 of 100 traceable wrong-way incidents occurred at a diamond-type interchange—this is too small a sample to make a definite statement as to interchange type vs wrong-way incidents; during weekdays, the age group 55 and over experienced significantly more difficulty in daylight hours, and the age group under 25 appeared to be less capable during dark hours; no definite pattern has been set by accidents; 80 percent of the wrong-way accidents reported occurred in dark hours; and 50 percent of the accidents were caused by drivers with 10 or more driver violation points at the time of the accident.

The data collected and analyzed here are additional proof that, although wrong-way accidents are a small percentage of the total accident picture, they are severe and dangerous. It is also apparent that a large number of wrong-way incidents occur that escape detection and accident involvements. All efforts should be made to eliminate these potential wrong-way incidents and/or accidents.

Paper sponsored by Committee on Freeway Operations.

An Analysis of Short-Term Implementation of Ramp Control on the Dan Ryan Expressway

ROY D. FONDA, Chicago Area Expressway Surveillance Project,
Illinois Division of Highways

ABRIDGMENT

•THE Chicago Area Expressway Surveillance Project undertook a ramp metering study on the northbound Dan Ryan Expressway with objectives of supplementing past experience in the general application of control techniques and gaining additional knowledge of the relative effects of various geometric design features on ramp control.

An interplay between the expressway and frontage street resulted in the generation of exceedingly high entrance ramp demands at the point where the expressway curves away from the frontage street. Congestion was triggered by high volumes force-merging with a near-capacity expressway, while the frontage street, through its discontinuity, was directly involved in sustaining the cause of congestion and delaying the local recovery from congested operations.

Ramp control, in the form of one-vehicle-at-a-time metering utilizing manually-operated portable equipment, was used on four successive entrance ramps to adjust merge demands to a level that could be accommodated.

Control of the four entrance ramps did not eliminate congestion, but significantly reduced the extent and duration. The severity of congestion was reduced such that individual motorists saved up to 5 minutes in traversing the 3.6-mile study section. A daily average of 627 vehicle hours of expressway travel time was saved during control, while the peak-period vehicle-miles of expressway travel increased by 5 percent.

Delays incurred were at the expense of vehicles waiting in queues at the controlled ramps. Even though motorists using the metered ramps were subjected to individual delays of up to 7 minutes in queues, compliance with the one-vehicle-at-a-time scheme averaged over 90 percent. There was negligible delay to surface street traffic in the corridor.

Ramp metering at just four ramps did not produce enough diversion to downstream ramps and/or surface street routes to completely prevent expressway overloading from occurring in the study section, but shifted the point of initial overloading upstream. The benefits derived from this study appear to be primarily attributable to a small amount of diversion to downstream ramps and to a reduction of merging turbulence produced by one-at-a-time entry at a decreased rate.

This initial study pointed out the need to extend control to additional upstream entrance ramps and to adjust the level of control on a traffic-actuated basis in order to further improve expressway operations.

Portable metering devices again proved valuable in providing the means for conducting short-term studies without the expense and delay inherent to permanent installations.

Paper sponsored by Committee on Freeway Operations.

Automatic Freeway Surveillance and Evaluation Techniques

JOSEPH M. McDERMOTT, Chicago Area Expressway Surveillance Project, Illinois Division of Highways

ABRIDGMENT

•SINCE the original installation of the unique pilot freeway detection system in 1962, the Chicago Area Expressway Surveillance Project has refined this experimental detector-computer system to provide automatic surveillance, control, and evaluation for a total of 24 directional miles on two Chicago area expressways. Critical surveillance-evaluation detectors transmit data over leased phone lines to a process-control computer in the project office. The digital computer processes detector data into on-line, real-time, freeway surveillance, entrance ramp metering control and evaluation outputs.

Some operational surveillance and control outputs are visually displayed on a map panel depicting current freeway traffic conditions and on a console depicting current control modes. Most evaluation outputs, however, are preserved on punched paper tape or on printouts developed on-line with two automatic typewriters. This report primarily describes how the electronic surveillance system is used to automatically evaluate freeway traffic operations throughout each rush period.

On one automatic typewriter, an on-line lane occupancy plot is developed by the digital computer. The average 5-minute occupancy of the detector in the second lane from the median at each freeway surveillance station is recorded via on-line printout each 5 minutes throughout each rush period, thereby providing a distance-time picture of freeway congestion.

On the second automatic typewriter, an "excess vehicles" plot is developed by the digital computer from a closed system of freeway and ramp detectors. Using the input-output technique of computing subsystem density changes, the excess total travel time experienced by rush-period traffic is determined.

Examples are presented in the report to illustrate the extremes of rush period freeway operations, as well as additional refinements in the printout format. The data of each printout have been summarized in one tag-type parameter, which resulted in a correlation coefficient of 0.78 when each lane occupancy printout was compared with each excess vehicle printout for 30 successive rush periods. Several applications for the refined surveillance/evaluation system are suggested. Further research possibilities with data being catalogued on a routine basis are pointed out.

Paper sponsored by Committee on Freeway Operations.

Development and Evaluation of Experimental Information Signs

GERALD C. HOFF, Chicago Area Expressway Surveillance Project,
Illinois Division of Highways

ABRIDGMENT

•AS freeway entrance ramp metering becomes a more widely accepted operational tool of the traffic engineer, the desire to reduce those adverse effects resulting from its employment increases. Studies conducted by the Chicago Area Expressway Surveillance Project have shown that the major portion of the delay is confined to the queues that develop on a controlled ramp.

This paper summarizes the development and evaluation of one technique that can be used to reduce the effects of queues. The potential expressway user has only a modicum of real-time information concerning traffic conditions existing along his route. If the driver were made aware in advance of the amount of delay that might accrue to him if he used the nearest controlled entrance ramp, he might divert to another entrance ramp or use the arterial street system for his entire trip. The efficiency of the highway system could be improved if his extra travel time were less than the delay he would have experienced at the nearest ramp.

Four color-coded changeable message signs were installed to provide the expressway-bound motorist with current expressway traffic information. The sign face is in the form of a map representing the arterial streets, entrance ramps, and the freeway. An arrow-shaped opening is cut out of the face of the sign at each location where traffic conditions are to be displayed. A translucent color wheel, divided into red, yellow, and green sectors, is placed inside the sign between each opening and a flood lamp. Each sign has four openings and displays information for the ramp and freeway at the nearest and next downstream entrance ramp. A digital computer and control system causes the appropriate sector of the color wheel to appear in the opening.

The evaluation of the sign's effect on the driver's decision process was conducted in two phases. The first phase was an attempt to estimate the alteration in traffic patterns caused by the signs. The policy under which the signs were operated did not allow erroneous information to be knowingly displayed, and only a limited amount of data could be obtained. The data indicated that the maximum diversion attributable solely to the signs could constitute about only one-fourth of the total diversion.

The second phase of the evaluation utilized a questionnaire to determine comprehension of and response to the sign. In general, most motorists that passed the sign understood its purpose. A large portion (50 percent) of the sample indicated that they made use of the displayed information in selecting their route.

Thus, two different techniques for determining the use of the sign information have produced somewhat divergent results. Additional investigations in the area of driver information are needed before dissemination systems can be planned on a rational basis.

Paper sponsored by Committee on Freeway Operations.

Lane-Changing in Multilane Freeway Traffic

R. D. WORRALL, A. G. R. BULLEN, and Y. GUR, Northwestern University

ABRIDGMENT

•A SERIES of analyses are presented indicating that lane-changing in multilane freeway traffic may be effectively described as a random process.

Assuming first that lane-changing may be treated as an isolated, independent event within the traffic stream, it is shown that the number of lane changes, X_{ij}^{kt}, between lanes i and j within a length of freeway k and a time span t may be estimated as the outcome of a simple Poisson process $\left\{X_{ij}^{kt};\ t > 0\right\}$, where

$$\text{Prob.}\ \left\{X_{ij}^{kt} = N\right\} = \frac{\exp\left(-\lambda_{ij}^{k}\ t\right)\left(\lambda_{ij}^{k}\ t\right)^{N}}{N!}$$

and λ_{ij}^{k} = Avg. No. lane changes between lanes i and j per unit time within section k.

Empirical data are presented to show that the value of λ_{ij}^{k} varies systematically with both traffic speed and traffic volume, and with proximity to entrance and exit ramps. Further, the assumption of randomness is shown to break down during medum-heavy flow periods (flows equivalent to 62 to 87 veh/min/3 lanes and at points immediately downstream from freeway entrance ramps.

It is similarly shown that the pattern of lane changes with the k'th section of length L may be modeled as a finite Markov process $\{X(k);\ L > 0\}$, where

$$\text{Prob.}\ \{X(k) = j \mid X(0), \ldots, X(k = 1);\ L\} = \text{Prob.}\ \{X(k) - j \mid X(k - 1);\ L\}, \text{ and}$$

Prob. $\{X(k) = j\}$ = Prob. that a given vehicle is in lane j as it leaves section k. A simple Markovian model based on this structure is calibrated from field data collected on a 6-lane freeway in Chicago, and extended to cover the situations of lane-changing in the vicinity of freeway entrance and exit ramps and within a complex weaving section. In each case the model effectively replicates the observed maneuver pattern.

Paper sponsored by Committee on Freeway Operations.

A Simulation Model of Lane-Changing On a Multilane Highway

A. G. R. BULLEN, R. D. WORRALL, and S. ROBERTSON,
Northwestern University

ABRIDGMENT

•A DIGITAL simulation model of lane-changing on a multilane highway is described. The model covers 2-, 3-, and 4-lane sections of freeway, and provides for scanning of vehicle trajectories once every second.

Vehicles are input randomly into each lane, with desired speeds sampled from a normal distribution and headways from shifted exponential distribution. The parameters of these distributions are treated as exogenous input to the model, identified by lane. The motion of each vehicle is controlled by a set of simple car-following laws, including specification of maximum acceleration/deceleration rates, an "uncomfortable headway" distribution, and an exponential gap acceptance function. Again, the parameters of these distributions are treated as exogenous input.

Lane changes are generated whenever a vehicle is forced to travel at an "uncomfortable headway" and finds an "acceptable gap" available in an adjacent lane.

Output of the model includes lane-change counts, total and average lane-changing delay, average speeds, and the distribution of accepted gaps. The paper discusses the calibration and validation of the model based on Chicago area data, and describes the results of an extensive program of sensitivity testing. These latter tests suggest that the pattern of lane-changing is most strongly dependent on the desired speed distribution, and is not unduly sensitive to gap acceptance or acceleration/deceleration inputs. The model effectively replicates observed lane-changing behavior.

Paper sponsored by Committee on Freeway Operations.