Ms. Louise Lunder, TSC-8311
Transportation Systems Center
Room 845 M-Building

REPORT NO. DOT-TSC-RSPA-78-10

COMPUTER-ASSISTED TRAFFIC ENGINEERING USING ASSIGNMENT, OPTIMAL SIGNAL SETTING, AND MODAL SPLIT

Stanley B. Gershwin
John D.C. Little
Nathan Gartner

Massachusetts Institute of Technology
Cambridge MA 02139

MAY 1978
FINAL REPORT

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
RESEARCH AND SPECIAL PROGRAMS ADMINISTRATION
Office of Transportation Programs Bureau
Office of Systems Engineering
Washington DC 20590

Technical Report Documentation Page

1. Report No. DOT-TSC-RSPA-78-10	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle COMPUTER-ASSISTED TRAFFIC ENGINEERING USING ASSIGNMENT, OPTIMAL SIGNAL SETTING, AND MODAL SPLIT		5. Report Date May 1978
		6. Performing Organization Code
7. Author(s) S.B. Gershwin,* J.D.C. Little,** Nathan Gartner**		8. Performing Organization Report No. DOT-TSC-RSPA-78-10
9. Performing Organization Name and Address Massachusetts Institute of Technology† *Electronic Systems Laboratory, **Operations Research Center, Department of Electrical Engineering and Computer Science, Cambridge MA 02139		10. Work Unit No. (TRAIS) OS750/R8526
		11. Contract or Grant No. DOT-TSC-849
12. Sponsoring Agency Name and Address U.S. Department of Transportation Research and Special Programs Administration Office of Transportation Programs Bureau Office of Systems Engineering Washington DC 20590		13. Type of Report and Period Covered Final Report April 1975-Sep 1977
		14. Sponsoring Agency Code
15. Supplementary Notes †Under contract to: U.S. Department of Transportation Transportation Systems Center Kendall Square Cambridge MA 02142		

16. Abstract

Methods of traffic assignment, traffic signal setting, and modal split analysis are combined in a set of computer-assisted traffic engineering programs. The system optimization and user optimization traffic assignments are described. Travel time functions are presented for freeways, freeway entrance ramps, and signalized streets. Both single-vehicle and multiple-vehicle class (cars, car pools, and buses) formulations are described. Energy optimization is treated, and gasoline consumption functions for cars and buses are shown. Modal split analysis is described and integrated with the assignments so that the effect of favoring buses and car pools with "diamond lanes" can be measured. The procedures are described and numerical examples are presented. Some areas of required research are identified; and it is concluded that this approach to traffic engineering is practical and computer programs can be written to analyze major parts of the traffic networks of U.S. cities.

17. Key Words Traffic Assignment, Modal-Split, Multiple Class Assignment, Energy Optimization, Optimal Signal Setting	18. Distribution Statement DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
UNCLASSIFIED	UNCLASSIFIED	100	

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

PREFACE

This research was conducted by the Transportation group at the MIT Electronic Systems Laboratory and the MIT Operations Research Center and sponsored under contract DOT-TSC-849.

The influence of Dr. Paul Ross of the Federal Highway Administration is greatly appreciated. His advice, encouragement, but mainly his intense interest, have helped to advance the research reported here.

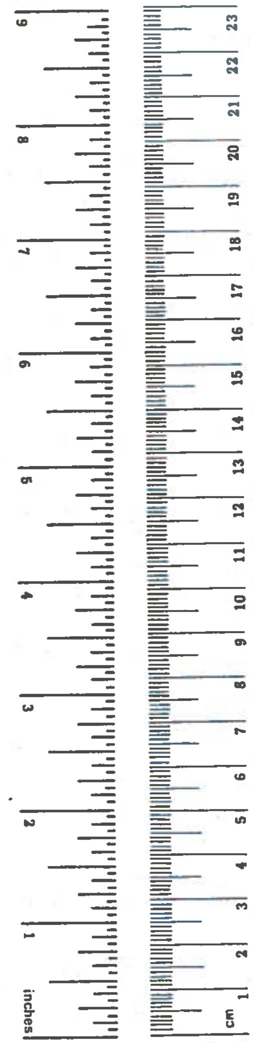
This work has benefited from the criticism and support of Professor Michael Athans. Mr. Michael McIlrath helped formulate the multiple class optimization problem. Dr. Joseph Defenderfer contributed his EF algorithm and Mr. Douglas Looze converted it to a transportation network program. Mr. Sam Chiu coded some of the cost functions and participated in designing the network. Mr. Han-Ngee Tan, with much patience and many late nights, was able to focus all this into a set of working procedures.

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

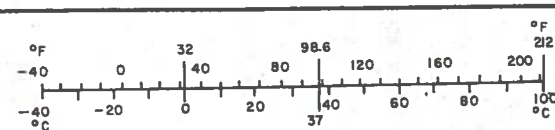
Symbol	When You Know	Multiply by	To Find	Symbol
	LENGTH			
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
	AREA			
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
	MASS (weight)			
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
	VOLUME			
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
	TEMPERATURE (exact)			
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
	LENGTH			
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
	AREA			
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
	MASS (weight)			
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
	VOLUME			
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
	TEMPERATURE (exact)			
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



CONTENTS

SECTION		PAGE

ILLUSTRATIONS

TABLES

1. INTRODUCTION

1.1 Purpose

A controversial approach to the reduction of energy consumption in urban roadway networks is the "diamond lane." This is the restriction of automobiles with fewer than a fixed number of occupants from specified lanes in freeways. This report describes the results of a pilot study aimed at determining whether a diamond lane actually does save energy.

A larger purpose of this research has been to investigate the feasibility of using large scale optimization techniques in a program of computer assisted traffic engineering. The diamond lane study is a special case of this. We have determined that this approach is feasible, and that there are circumstances under which diamond lanes can save energy.

The problem of traffic management (including energy conservation in networks) is difficult because of its complexity. There are a large number of cars, streets, and traffic signals in a network, and the existence of each has an effect on the others. For example, in assessing a diamond lane, one must recall that it is not the only control in the network. It may be necessary to change the settings of traffic signals, to obtain optimal energy consumption, after a diamond lane has been established. Furthermore, these changes are likely to alter the locations of congestion and points of delay. Drivers are likely to respond by changing their routes. Finally, diamond lanes favor buses and carpools over cars. Commuters will consider this and may change transportation modes as a result.

Consequently, all of these controls and effects must be considered simultaneously,and this was the approach taken in the pilot study reported here.

1.2 Outline of Approach

At the heart of the procedure is a traffic assignment program: an algorithm that calculates the distribution of traffic flow in a network when origin-destination demand data is specified and broken down by mode. Such a program requires two things: an optimization technique and a set of traffic models.

The optimization techniques used here were formulated for a similar problem in computer communication networks. It very efficiently exploits the special attributes of the network structure.

The traffic models calculate the cost, according to each of the criteria considered, of traveling on each roadway link. To do this, they include detailed functional relationships between flow, velocity, signal settings, delay, and energy consumption. Such relationships are obtained from mathematical modelling and empirical data.

Two versions of the assignment program are described: one for single-vehicle-class problems and one for multiple-vehicle-class problems. The former is appropriate when all vehicles in the network are treated the same, and controls, such as traffic signals, apply to all vehicles equally. The latter is appropriate when some are favored at others' expenses, for example in diamond lanes.

There is a large body of literature dealing with <u>optimal signal setting</u>: the calculation of green splits, cycle times, and other quantities, given the flows of traffic through an intersection or in a network. Relatively little study has been directed toward the choice of signal settings given their effect on network flows. In this report, an iterative heuristic procedure is proposed to calculate settings taking this into account.

The study of <u>mode split</u> seeks to find how the population of travellers chooses among the available transportation modes. We treat travel time as the most important <u>variable</u> parameter* that travellers consider. Travel time, for each mode, depends on flows on each link and on all traffic controls, including diamond lanes. We propose another iterative, heuristic procedure to simultaneously treat all these effects.

1.3 Summary of Capabilities

Different versions of many of the elements of this procedure were constructed. These are briefly summarized here, and discussed in detail

* That is, we treat other attributes, such as comfort or crowding as different for different modes, but independent of traffic flows.

elsewhere in the report.

1) The assignment may be according to system or user optimization. That is, it can be best from the system's point of view or it can predict how drivers actually choose routes. The experimental user optimization method discussed here is new and appears to represent a significant improvement in flexibility and accuracy over existing techniques.

2) If the assignment is system optimized, the criterion can be either average travel time or total fuel consumption rate or any positive linear combination of the two.

3) Both fuel consumption rate and total travel time rate (in vehicle-hours per hour or passenger-hours per hour) are evaluated, regardless of assignment optimization principle or criterion.

4) One version (single class) of the assignment program assumes that all vehicles are the same. The other (multiple class) assumes that vehicles can be separated into three distinct classes, with different passenger occupancies, impact on delay, and energy consumption.

5) Multiple class system optimized runs can reserve lanes if a network with the proper structure is analyzed. In general, any link can be reserved by the program.

6) The programmer can specify reserved links in a user optimized multiple class run.

7) Signal settings can be calculated in two ways: a computationally fast, approximately optimal formula can be used for either isolated or widely spaced signals. In addition, a program is available to calculate cycle time and phase offset for coordinated signals.

8) Mode split can be calculated as a function of discrepancies in travel time among the modes.

1.4 Outline of Report

In Section 2, we describe the formulation of the model. This includes a presentation of fuel consumption and delay functions at signals, ramps, and freeways, a description of user and system optimization, and a demonstration of mode split. In Section 3, we discuss the methods used to solve the problem. Section 4 contains numerous examples to illustrate the approach described here. Section 5 concludes this report with a set of recommendations. These recommendations are chosen to improve the convenience, generality, and computational efficiency of these methods, and no great difficulties are anticipated in carrying them out. Appendix A contains a description of the MITROP program, which calculates optimal settings of coordinated traffic signals. Appendix B shows how the signal delay calculation of MITROP is adapted for use with the assignment program. We discuss in Appendix C the way both system and user optimization behaviors are exhibited simultaneously by actual networks.

1.5 Main Conclusion

The fusion of assignment, signal calculation, and mode split is a highly complex, large scale optimization problem. We present a systematic approach to the solution of this problem and conclude, by means of the pilot study reported here, that this approach is a feasible one.

2. MODEL FORMULATION

2.1 Introduction

In this section three important problems of traffic engineering are formulated: assignment, the optimal choice of traffic signal parameter values, and modal split.

These problems are approached assuming a steady state (or static) model of traffic flow. The vehicle traffic in a network is never at rest, but there are quantities that remain nearly constant and steady state relationships among them that are nearly valid over relatively long periods of time. Such relationships have traditionally been used to simplify the analysis of transportation networks [2] through [16].

We assume here that the flow on a link, in vehicles per hour, is a parameter that remains approximately constant over periods that are long compared to, say, the time a given vehicle spends on the link. The same is true for all other quantities discussed here.

The assignment problem is that of calculating the distribution of vehicles in a traffic network, given the travel demand and the network structure.

The mathematical formulation of this problem requires several pieces of information. Delay on each roadway link must be specified as a function of traffic flow and other parameters (such as link capacity and traffic signal variables). The assignment principle must be defined: do drivers choose the shortest paths available or does a central controller dictate route choices to minimize a system wide cost (such as energy consumption)? In this section we describe delay and cost models for different kinds of network links. Assignment principles are discussed in section 2.4.

Two versions of the assignment problem are presented. In one, all vehicles are treated as if they are exactly the same. In the other, we distinguish among classes of vehicles: for example, single passenger (i.e., non-pooled or private) cars, car pools, and buses. We consider three characteristics of vehicles: the impact on link delay, the average passenger occupancy, and the energy consumption per vehicle. Private cars and car pools have the same energy consumption and the same effect on travel delay on each link.

Car pools have a higher passenger occupancy than private cars. Buses have a greater impact on the delay on a link, they have greater passenger occupancy, and they have greater energy consumption than cars or car pools. These concepts are quantified below.

The values of traffic signal parameters, such as cycle time and green split (the fraction of a cycle that the signal is green), have an important effect on delays on signalized links and thus on the assignment. The problem of <u>optimal signal setting</u> is that of calculating the parameters that minimize delay or other cost function. This requires models of traffic signal delays and an optimization algorithm. In this section, models of signal delays are presented. In the following section, methods of calculating signal parameters are discussed.

In the multiple mode assignment problem, different classes of vehicles may experience different levels of service. For example, if cars are prohibited from traveling on some links, they will probably encounter longer travel times than buses. Travelers can be expected to respond to such discrepancies by choosing more desirable transportation modes. The calculation and prediction of this behavior is the problem of <u>modal split</u>.

2.2 Flows and Constraints

In this section, we define <u>traffic flow</u> for the single and multiple vehicle class formulations. Traffic flow is never negative and, in steady state, is conserved at network modes. These constraints are stated precisely here.

2.2.1 Single-Vehicle Class Formulation

We define ϕ_{ij} to be the <u>flow</u>, in vehicles per hour, of traffic on link i destined for node j.

<u>Positivity</u>: All flows are positive, i.e.,

$$\phi_{ij} \geq 0, \text{ all } i,j. \tag{2.1}$$

<u>Conservation of flow</u>: The total flow entering a node equals the total flow leaving the node, and this is true when the flow is broken down by

destination. Assume that in Figure 2.1 links $i_1,\ldots,i_\alpha$ carry flow toward node k and links $\ell_1,\ldots,\ell_\beta$ carry flow away from node k. Define r_{kj} to be the flow requirement, or origin-destination demand data. This is the rate, in vehicles/hour, of vehicles appearing at node k whose drivers want to go to node j. Such traffic could be arriving from outside the system, from parking spaces or parking garages, or other sources. Then the total flow leaving node k and destined for node j is $\phi_{\ell_1 j} +\ldots+ \phi_{\ell_\beta j}$ and the total flow entering and destined for node j is $\phi_{i_1 j} +\ldots+ \phi_{i_\alpha j} + r_{kj}$. If $k \neq j$, we have

$$\phi_{\ell_1 j} + \ldots + \phi_{\ell_\beta j} - (\phi_{i_1 j} +\ldots+ \phi_{i_\alpha j}) = r_{kj} . \qquad (2.2a)$$

If $k = j$, then

$$\phi_{\ell_1 j} + \ldots + \phi_{\ell_\beta j} - (\phi_{i_1 j} + \ldots + \phi_{i_\alpha j}) = -\sum_{m \neq j} r_{mj} , \qquad (2.2b)$$

where the term on the right is the negative of the total flow destined for node j.

Define the <u>total flow</u> on link i to be

$$\phi_i = \sum_j \phi_{ij} . \qquad (2.3)$$

2.2.2 Multiple-Vehicle Class Formulation

The multiple class flow problem [10] is similar, but there is one variable for each vehicle class on each link.

Let $\phi_{ij}^{(n)}$ be the flow, in vehicles per hour, of traffic onlink i whose destination is node j and which is <u>in class n</u>. In this report, n=1 refers to single-passenger cars; n=2 refers to car pools; and n=3 refers to buses.

As before, we have <u>positivity constraints</u>, i.e.,

$$\phi_{ij}^{(n)} \geq 0; \text{ all } i,j,n, \qquad (2.4)$$

and <u>conservation of flow constraints</u>:

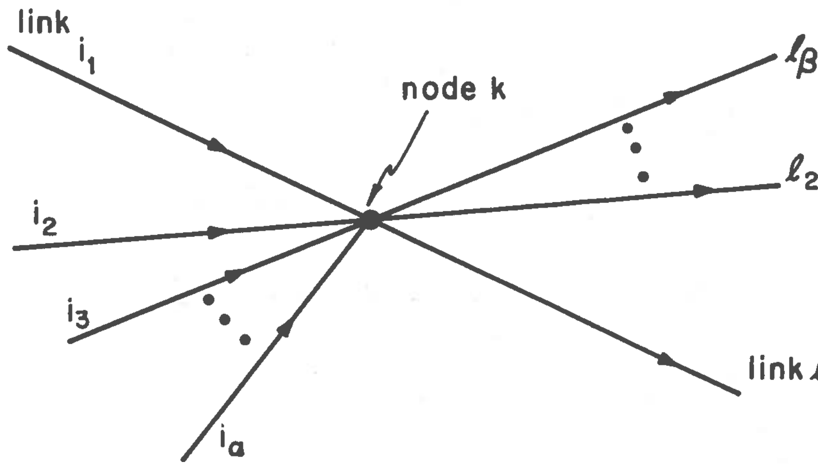


Figure 2.1 A Link-Node Configuration

$$\phi_{\ell_1 j}^{(n)} + \ldots + \phi_{\ell_\beta j}^{(n)} - (\phi_{i_1 j}^{(n)} + \ldots + \phi_{i_\alpha j}^{(n)})$$
$$= r_{kj}^{(n)}, \; k \neq j, \tag{2.5a}$$

$$\phi_{\ell_1 j}^{(n)} + \ldots + \phi_{\ell_\beta j}^{(n)} - (\phi_{i_1 j}^{(n)} + \ldots + \phi_{i_\alpha j}^{(n)})$$
$$= -\sum_{m \neq j} r_{mj}^{(n)}, \; k = j. \tag{2.5b}$$

The total passenger flow rate P_i on link i is

$$P_i = \sum_j \sum_n \bar{w}^{(n)} \phi_{ij}^{(n)}, \tag{2.6}$$

where $w^{(n)}$ is the average passenger occupancy, i.e. the average number of people per vehicle, of vehicles of class n. The total passenger flow rate is the number of passengers per hour that travel on link i.

The total passenger-car-equivalent flow ϕ_i on link i is

$$\phi_i = \sum_j \sum_n e^{(n)} \phi_{ij}^{(n)}. \tag{2.7}$$

We assume that a mixture of vehicles produces a delay on a link which is the same as that produced by an equivalent number of passenger cars. Then $e^{(n)}$ is the ratio of the increase in delay produced by one more vehicle of class n on a link to that produced by one more passenger car: see references [31] and [66]. In the single vehicle class case, $e^{(1)} = 1$ and (2.7) reduces to (2.3).

2.3 Travel Time and Energy Consumption Cost Functions

2.3.1 Travel Time

Define τ_i to be the average time a vehicle spends on link i. The functional expression for τ_i depends on whether link i is a freeway segment, a freeway entrance ramp, or a signalized arterial. A vehicle spends its time on a link either traversing the link (possibly slowly because of congestion) or waiting in queues (i.e., actually at rest). The time τ_i depends on ϕ_i, possibly the total flow on other links, and possibly on some traffic signal parameters.

Link Traversal Time $t_i(\phi_i)$

A common link traversal function that is used in assignment studies is the fourth power polynomial [2]

$$t_i(\phi_i) = t_{oi}(1 + .15\,(\phi_i/C_i)^4), \tag{2.8}$$

where t_{oi} is the traversal time when $\phi_i = 0$ and C_i is the link capacity.

Alternatively, one may use

$$t_i(\phi_i) = \ell_i/v_i, \tag{2.9}$$

where ℓ_i is the length of link i and v_i is the velocity on link i. The velocity is obtained from

$$\phi_i = \rho_i v_i, \tag{2.10}$$

and the fundamental relationship between flow ϕ_i and density ρ_i (see references [1] and [18]- [22]).

$$\rho_i = k_i(\phi_i), \tag{2.11}$$

as illustrated in Figure 2.2.

The density can be represented by a polynomial in the flow. Coefficients can be chosen as in (2.8), or to approximate (2.11) or Figure 2.2, or to represent cases in which velocity (and thus traversal time) is constant, i.e., independent of flow.

Freeway Links

There are no sources of delay on freeways except link traversal time. Thus, if i is a freeway link,

$$\tau_i = t_i(\phi_i)\ . \tag{2.12}$$

Freeway Entrance Ramps

If link i is a freeway entrance ramp, then the travel time on link i depends on ϕ_i and also ϕ_ℓ, where ℓ is the adjacent freeway link.

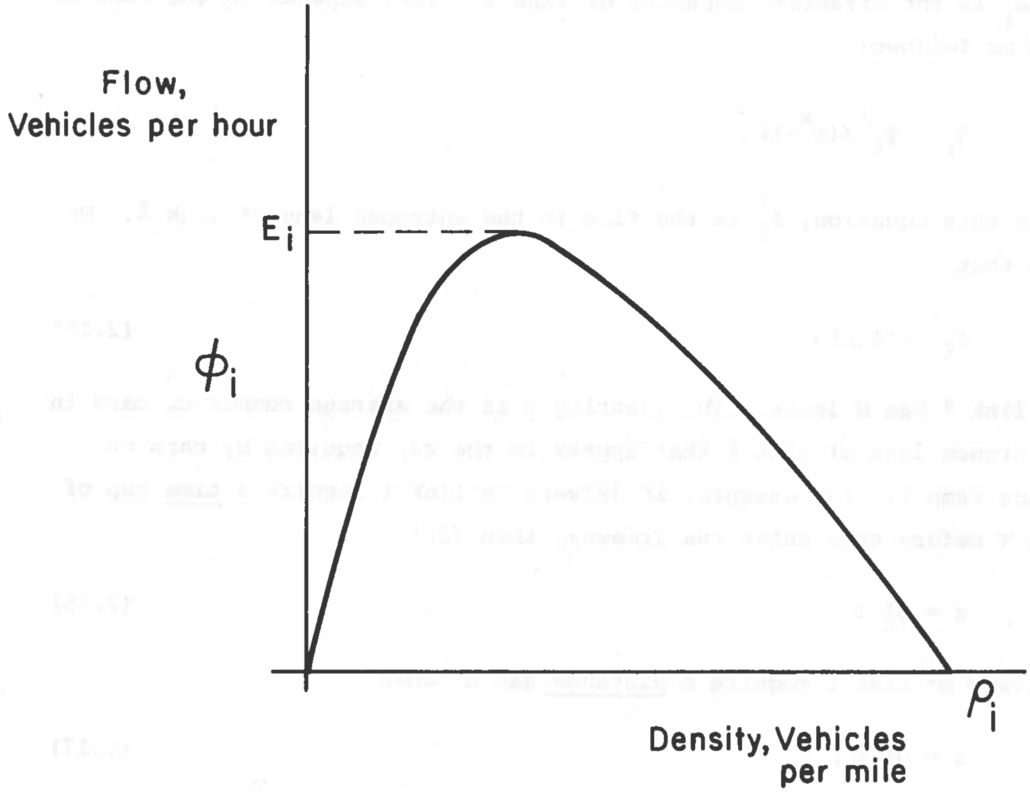


Figure 2.2 Fundamental Diagram

(See Figure 2.3). This is because a driver on link i must wait longer for an acceptable gap in traffic on link ℓ if the freeway traffic flow increases.

A function that expresses this is

$$\tau_i = \frac{1}{E_i - \phi_i} + t_i(\phi_i), \tag{2.13}$$

where E_i is the effective capacity of link i. This depends on the flow on link ℓ as follows:

$$E_i = \phi_\ell' / (e^x - 1). \tag{2.14}$$

In this equation, ϕ_ℓ' is the flow in the entrance lane of link ℓ. We assume that

$$\phi_\ell' = \phi_\ell / N, \tag{2.15}$$

where link ℓ has N lanes. The quantity x is the average number of cars in the entrance lane of link ℓ that appear in the gap required by cars on entrance ramp i. For example, if drivers on link i require a time gap of length T before they enter the freeway, then [32]

$$x = \phi_\ell' \, T. \tag{2.16}$$

If drivers on link i require a distance gap D, then

$$x = \rho_\ell' \, D, \tag{2.17}$$

where

$$\rho_\ell' = \rho_\ell / N. \tag{2.18}$$

Equation (2.17) is used in the computer program with D=160 ft. = 0.0303 miles.

The function $t_i(\phi_i)$ is the average traversal time and may, for example, be constant. Note that in equation (2.13) the travel time on link i depends on the flow on link ℓ.

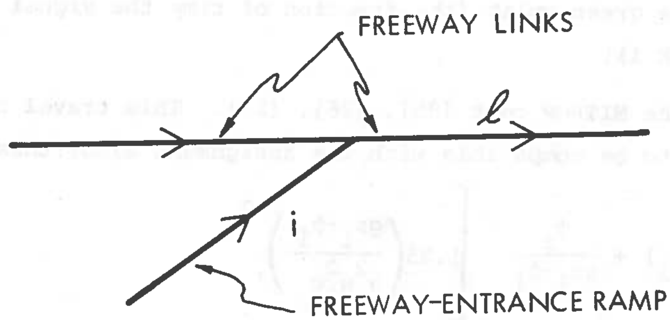


Figure 2.3 Freeway-Entrance Ramp

Signalized Arterials

Two signalized arterial travel time functions have been considered in this study. One is the Webster formula [23], [24], [61] and is most appropriate for isolated intersections, i.e. where signalized intersections are relatively far apart. The travel time is given by

$$\tau_i = t_i(\phi_i) + .45 \left\{ \frac{c(1-g)^2}{1-\phi_i/s_i} + \frac{\phi_i}{gs_i(gs_i-\phi_i)} \right\}, \tag{2.19}$$

where $t_i(\phi_i)$ is the traversal time discussed above, c is the cycle time, s_i is the saturation flow (the capacity of link i if the signal were always green), and g is the green split (the fraction of time the signal is effectively green to link i).

The other is the MITROP cost [25], [26], [27]. This travel time function has been rewritten to be compatible with the assignment algorithms as follows

$$\tau_i = t_i(\phi_i) + \frac{\phi_i}{gs_i-\phi_i} \left[1.25\left(\frac{gs_i-\phi_i}{g^2 s_i^2 c}\right)^2 + 2.25\, \frac{\phi_i}{gs_i}\, \frac{gs_i-\phi_i}{g^2 s_i^2 c} + .008\left(\frac{\phi_i}{gs_i}\right)^3 \right] + \frac{\gamma_i^2}{2p_i\left(1-\frac{\phi_i c}{p_i s_i}\right)}, \tag{2.20}$$

where γ_i is the offset between the signals at the ends of link i and p_i is the platoon length on link i. These variables, as well as g_i and c, are obtained from the MITROP program. The MITROP program is discussed in Appendix A and equation (2.20) is derived in Appendix B.

2.3.2 Energy Consumption

In this study we also investigate the total energy consumed per hour by the vehicles on the network. Again, there are different functional forms for different kinds of links.

Freeway Links

Extensive research was performed by Claffey [28] who found data for G(v), the gasoline consumption per vehicle per mile as a function of velocity. We fitted a polynomial to his tabulated results for automobiles and used

$$F_i = \ell_i \phi_i G(v_i), \tag{2.21}$$

as the total fuel consumption per hour on link i in the single vehicle class formulation. The velocity v_i is obtained from (2.10) and (2.11).

In the multiple vehicle class formulation, the total fuel consumption on link i is given by

$$F_i = \sum_n \ell_i \phi_i^{(n)} G^{(n)}(v_i), \tag{2.22}$$

where

$$\phi_i^{(n)} = \sum_i \phi_{ij}^{(n)} \tag{2.23}$$

is the total flow of class n vehicles on link i. The function $G^{(n)}$ is the fuel consumption rate of class n vehicles in gallons per mile. When n=1 or 2 (cars or car pools) we use Claffey's results for automobiles ([28], Table 6, page 17). Claffey has no results for buses, but for n=3 we use his results for two axle, six tire trucks ([28], Table 13, page 24).

The gasoline consumption functions $G^{(n)}(v)$ used here are polynomial fits to Claffey's results. For automobiles, n=1 or 2 (for single passenger cars and car pools) and we use

$$G^{(n)}(v) = .121415 - .674656\text{x}10^{-2}v + .205562\text{x}10^{-3}v^2$$
$$-.260145\text{x}10^{-5}v^3 + .126133\text{x}10^{-7}v^4.$$

For buses, using data for six-wheel, two axle trucks, we use

$$G^{(3)}(v) = .129518 - .842106\times10^{-2}v + .331912\times10^{-3}v^2 - .485759\times10^{-5}v^3 + .271696\times10^{-7}v^4 .$$

Entrance Ramps

Claffey [28] has results that are of value for entrance ramp fuel consumption. The gasoline consumed $H(v)$ by a vehicle in a stop-go cycle, i.e. in changing its speed from v to zero and back to v is tabulated*. The fuel consumed, F_i, by vehicles on entrance ramp i in the single class formulation is given by

$$F_i = \phi_i \left(H(v_i) + \frac{I}{E_i-\phi_i} \right), \tag{2.24}$$

where I is the gasoline consumption rate in idling and $(E_i-\phi_i)^{-1}$ is the average time a driver on entrance i must wait for an acceptable gap on link ℓ to appear. See Figure 2.3 and equation (2.13).

In the multiple class formulation, the fuel consumed on entrance ramp i is given by

$$F_i = \sum_n \phi_i^{(n)} \left(H^{(n)}(v_i) + \frac{I^{(n)}}{E_i-\phi_i} \right), \tag{2.25}$$

where for n=1 and n=2 we use a polynomial fit $H^{(n)}$ to Table 7, page 18 of [28] and for n=3, we use Table 14, page 26 of [28]. These are, for n=1 and 2,

$$H^{(n)}(v) = -.777448\times10^{-2} + .127548\times10^{-2}v - .402004\times10^{-4}v^2 + .694154\times10^{-6}v^3 - .40959\times10^{-8}v^4 ,$$

and

---

* Actually $H(v,t)$ is tabulated. $H(v,t)$ is the gasoline consumed by a vehicle in a stop-go cycle and idling for t seconds. There is evidently an error in Table 7, page 18 of [28] which displays $H(v,t)$ for automobiles. The tabulated data is not consistent with the specified idling rate of .58 gallons per hour. The correct values are given by $H(v,t) = H(v,0) + .58t$, where $H(v,0)$ is found in the first column of Table 7.

$$H^{(3)}(v) = -.170478\text{x}10^{-4} + .235099\text{x}10^{-3}v + .117785\text{x}10^{-4}v^2$$
$$+ .138942\text{x}10^{-6}v^3 - .504837\text{x}10^{-8}v^4 .$$

Also, $I^{(1)} = I^{(2)} = .58$ gallons per hour and $I^{(3)} = .65$ gallons per hour (Table 9, page 19 of [28]).

Signalized Arterials

Based on the work of Evans and Herman [29], we calculate the fuel consumption on a signalized link in the single vehicle class formulation by

$$F_i = (A\ell_i + B\tau_i)\phi_i, \tag{2.26}$$

where τ_i is given by (2.19) or (2.20), and where

A = 0.0434 gallons/vehicle/mile

B = 0.000258 gallons/vehicle/second.

Unfortunately, Evans and Herman do not consider the fuel consumption of buses. Therefore we assume that bus fuel consumption on streets can be approximated by a constant K times autombile fuel consumption on streets. We obtain K by taking the ratio of bus (i.e., two-axle, six-tire truck) to automobile fuel consumption at 30 mph as calibrated by Claffey [28].

Thus, in the triple vehicle class formulation,

$$F_i = F_i^{(1)} + F_i^{(2)} + F_i^{(3)} = (A\ell_i + B\tau_i)(\phi_i^{(1)} + \phi_i^{(2)} + K\phi_i^{(3)}) \tag{2.27}$$

where

$$K = \frac{G^{(3)}(30 \text{ mph})}{G^{(1)}(30 \text{ mph})} = 1.523. \tag{2.28}$$

2.3.3 Total Cost Functions

Given the average travel times and energy consumption rates on each link of the network, it is of interest to evaluate such quantities for the network as a whole. In this section we define system-wide cost measures.

Vehicle-Average Travel Time

The average time a vehicle spends in the system is

$$\frac{\sum_{\text{all links } i} \tau_i \phi_i}{\sum_{\text{nodes } j,k} r_{kj}},$$

where the summation in the numerator is taken over all links in the network and the summation in the denominator is over all pairs of nodes. The denominator is specified and is equal to the total vehicle flow rate through the network. Thus, we can measure this average by calculating

$$C_v = \sum_i \tau_i \phi_i . \tag{2.29}$$

Passenger-Average Travel Time

Similarly, we can measure the average time a passenger spends in the system in the multiple vehicle class formulation by calculating

$$C_p = \sum_i \tau_i P_i . \tag{2.30}$$

Total Energy Consumption

$$C_e = \sum_i F_i \tag{2.31}$$

is the total energy consumption in the network, where F_i is given by (2.21), (2.24), and (2.26) or (2.22), (2.25)and (2.27).

Total Cost Function

In section 3 we discuss a computer program that minimizes a linear combination of time and energy given by

$$C = W_1 \sum_i F_i + W_2 \sum \tau_i P_i , \tag{2.32}$$

where W_1, W_2 are constants specified by the programmer. If W_1=0, the program minimizes average or total travel time in the network. If W_2=0, the program minimizes total energy consumption. If both are positive, they can be thought of as the dollar cost of fuel and time [30] and C is a total dollar cost per hour.

2.4 Assignment Principles

The assignment principle is the rule by which network attributes, which depend on flow values, influence the distribution of flow in a network traffic assignment.

Two assignment principles were formulated by Wardrop [1]. The system optimization assumption is that traffic is distributed in a way that minimizes some criterion function. (Wardrop suggested average travel time.) The user optimization assumption attempts to describe the actual behavior of drivers. Drivers are assumed to choose the shortest paths available. As a consequence, trip times on different paths traveled between the same pair of points tend to equalize.

2.4.1 System Optimization

The system-optimized allocation of vehicles is found by specifying and solving a mathematical programming problem. A mathematical programming problem is one in which values of independent variables must be found to minimize a specified function, and in which the variables are subject to specified constraints. For example, in the network optimization problem, the independent variables are the flows on each of the roadway links. The constraints require that all flows be positive or zero, that conservation of flow is maintained at all nodes, and that arriving flows have certain specified values. The cost function, which is to be minimized, can be any cost function in section 2.3.3, or any combination thereof.

The system optimizing assumption leads to a set of necessary conditions (the Kuhn-Tucker conditions [17]). The computer algorithm discussed in Section 3 is used to find an assignment that satisfies these conditions.

2.4.2 User Optimization

A user optimization traffic assignment is one in which the travel times on all paths actually utilized between the same pair of nodes are the same. Other paths between that pair of nodes carry no flow, and have longer travel times. Most authors ([4] - [14]) satisfy user optimization conditions by constructing an associated system optimization problem and solving that. This approach has certain limitations which we discuss in Section 3. We propose in Section 3 an alternative approach. While this approach has not yet been carefully studied, preliminary results are encouraging.

To calculate single class user optimization assignments, we use the travel time expressions (2.8) - (2.20). For multiple class assignments, the same expressions are used with one exception: if class n vehicles are excluded from link i, then the travel time on link i for those vehicles is set to an arbitrary, very large number. The exclusion of vehicles from links (such as diamond lanes or selected entrance ramps) is controlled by the programmer.

2.5 Modal Split

In investigating mode split, we use the logit model [33], [52]. Define the passenger demand $R_{ij}^{(n)}$ to be the rate of arrivals of passengers at node i who wish to go to node j on mode n. Note that

$$R_{ij}^{(n)} = r_{ij}^{(n)} w^{(n)}, \tag{2.33}$$

where $r_{ij}^{(n)}$ is the rate of arrival of vehicles of class n at node i whose drivers wish to go to node j, and $w^{(n)}$ is the average passenger occupancy of vehicles of class n.

Define the total passenger demand, R_{ij} from i to j as

$$R_{ij} = \sum_n R_{ij}^{(n)}. \tag{2.34}$$

Then the logit model is defined by

$$R_{ij}^{(n)} = R_{ij} \frac{e^{-\theta u_{ij}^{(n)}}}{\sum_m e^{-\theta u_{ij}^{(n)}}}. \tag{2.35}$$

In this expression, θ is a constant which reflects the strength of preference of travelers for modes. The quantity $u_{ij}^{(n)}$ is the disutility or unattractiveness of using mode m and this depends on network congestion. If θ is large, travelers are very concerned about differences in service among different modes and make their choices accordingly. If Θ is small, they are indifferent and divide themselves nearly evenly among them.

In this study we use

$$u_{ij}^{(n)} = \alpha^{(n)} + \beta^{(n)} T_{ij}^{(n)}, \quad (2.36)$$

where $\alpha^{(n)}$ and $\beta^{(n)}$ are constants and $T_{ij}^{(n)}$ is the travel time from i to j on mode n. We describe the use of this model in section 3 and the values of parameters θ, $\alpha^{(n)}$, and $\beta^{(n)}$ in section 4.

2.6 Possible Future Improvements

In Section 5 we discuss which of the energy and delay models we use have been validated and which have not. Clearly the accuracy of the numerical results will improve as more faithful representations of reality are used.

It would be of interest to investigate alternative ramp models, such as those in reference [32], and to calculate ramp speeds. We assume here that all vehicles come to rest on the ramp. This assumption is not realistic, and it does not give buses the priority they require.

This can be seen by examining [28]. Claffey, in Table 8 (p. 19) and Table 15 (p. 26), estimates "Excess Gallons of Gasoline Consumed per Slow-down Speed Change Cycle" for automobiles and two axle, six tire trucks. A slowdown speed change cycle is a change in velocity from a higher speed V_1 to a lower speed V_2 and back to V_1. The difference between car and truck fuel consumption in this maneuver is substantial, and is most pronounced as V_1-V_2 increases.

If our ramp model included V_1 and V_2 then clearly it would be in the system's interest to keep buses moving as fast as possible. Thus it may want to establish exclusive bus ramps so that the flow is small and V_2 is large.

Claffey's results, which we depend on heavily, do not apply precisely to buses. Furthermore, the automobile data are rapidly becoming obsolete as gasoline consumption rates improve. Evans' and Herman's results [29] were also not calculated for buses.

Lane changing is important when reserved lanes are considered. A simple model is discussed in Section 4. A more sophisticated model might improve the accuracy of these results.

Other possible improvements include more vehicle classes (to include trucks and rapid transit) and more general elastic demand problems.

2.7 Summary

In this section we have defined models and problems in many aspects of traffic engineering. We have obtained formulas for travel time and energy consumption for each kind of vehicle (private car, car pool, and bus) on each kind of roadway link (freeway, freeway entrance ramp, and signalized arterial). Since travelers base their modal choices on the service provided, we have described a modal split model.

Many of these formulas have been developed in the literature, and some have been calibrated with actual traffic data. Our contribution has been to bring them together and adapt them in a way that can be used for traffic control policy decisions.

3. COMPUTATION ALGORITHMS

3.1 Introduction

In section 3 algorithms for computation of numerical solutions to the problems posed in Section 2 are described. In 3.2 we discuss the assignment methods, which are at the heart of all the procedures presented here. An explanation of the schemes used to calculate signal settings appears in 3.3. The overall procedures used to integrate assignment, signal settings, and mode split are presented in section 3.4.

Subsection 3.5 suggests areas of possible future improvements to this set of algorithms. The remainder of 3.1 contains some general remarks intended for the reader not familiar with computational algorithms.

There are a great number of optimization algorithms available in the literature. There are some (e.g., see [34], [35]) that are intended for general problems. Such methods trade efficiency for breadth of applicability. That is, a general technique may solve a wide class of problems, but it is likely to be slow and to require a great deal of memory.

A specialized technique can be made more efficient for a smaller class of problems. The most famous example is linear programming. Methods for solving such problems ([34], [36]) are vastly superior to those for nonlinear problems in the sense that many more variables and constraints can be treated in the same amount of computer time. However, even linear programming methods can be made more efficient when there is special structure to the problem, e.g., finding minimal cost flows in networks, with linear cost [37], [38], [39], [40].

At the outset of this study, a general nonlinear programming technique was applied to a small freeway corridor network with a single destination and a single class of vehicles [63]. When the problem was modified to include two classes of traffic, the complexity and the computer expense of the algorithm became prohibitive [64]. The specialized technique discussed here was then applied, and computer times (per iteration) were reduced by a factor of over 100 [65].

For this reason, much of the research in traffic assignment has been in devising methods that exploit the structure of the problem. A widely used method is the Frank-Wolfe technique[50]. This method alternates between solving linear programming problems and performing one-dimensional searches. The latter operation is always fast and the former is fast because of the network structure. In 3.2 we describe an algorithm which is similar to the Frank-Wolfe scheme and which calculates the number of vehicles or passengers and the trip time on each route.

3.2 The MIT OPTFLOW Program

The Cantor-Gerla method was chosen to solve the assignment problem of Section 2. This approach is well suited to this problem and has been studied at MIT Electronic Systems Laboratory [39]. A version of the algorithm was available, and while it was being adapted to traffic assignments, other versions were written that were considerably improved in terms of computer expense.

System Optimization

The algorithm can be described geometrically with a very simple problem. In Figure 3.1, the solid lines define a feasible region. Points in the shaded region are not allowed. Since the feasible region is defined by straight lines, it can be represented by a set of linear constraints (e.g., equations (2.1) and (2.2)).

There is a nonlinear cost function and surfaces of constant cost are represented by dashed lines. As we move away from point M, the cost increases. (For example, the cost may be given by equation (2.29).)

Assume an initial guess G_1. Through G_1 is drawn the tangent to the constant cost curve and the normal in the direction of greatest decrease of cost. An approximate linear cost function may be defined which has surfaces of constant cost parallel to the tangent. That function, and the linear constraints, form a linear programming problem. The solution is the point in the feasible region which is farthest from G_1 in the normal direction. That point is E_1 (for extremal).

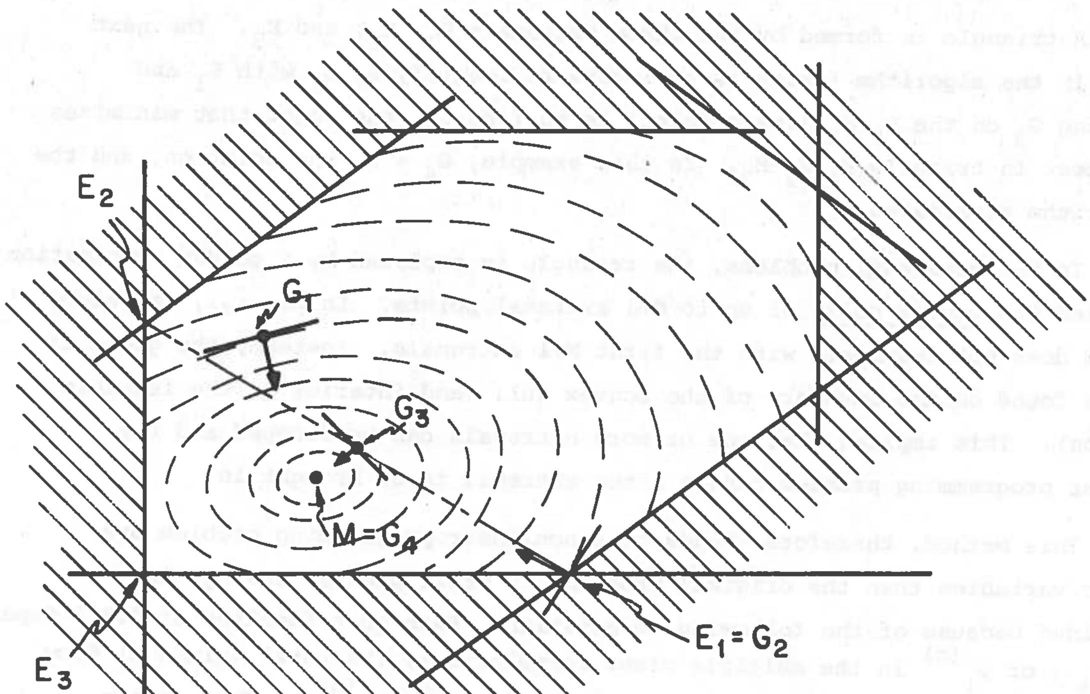


Figure 3.1 The Cantor-Gerla Algorithm

At E_1, since there is only one extremal, we designate $G_2 = E_1$, and follow the same procedure. This leads to E_2.

Now a line segment is drawn between E_1 and E_2 and the point with the minimum cost on that segment is found. This point is called G_3 and the linear programming procedure is repeated. This yields E_3.

A triangle is formed by the three extremals E_1, E_2, and E_3. The next step in the algorithm (which is analogous to identifying G_2 with E_1 and finding G_3 on the E_1-E_2 line segment) is to find G_4, the point that minimizes the cost in triangle E_1-E_2-E_3. In this example, $G_4 = M$, the solution, and the algorithm terminates.

In N-dimensional problems, the triangle is replaced by a convex combination (called the convex hull) of up to N+1 extremal points. In general, the algorithm does not terminate with the first N+1 extremals. Instead, the guess is often found on the boundary of the convex hull (and interior to the feasible region). This implies that one or more extremals can be dropped and the linear programming problem causes a new extremal to be brought in.

This method, therefore, leads to a nonlinear programming problem over fewer variables than the original problem. A still greater savings is obtained because of the following observation. Each cost function of 2.3.3 depends on ϕ_i (or $\phi_i^{(n)}$ in the multiple class formulation), the total (class n) flow on link i. This is a linear combination of ϕ_{ij} (or $\phi_{ij}^{(n)}$). Thus we can solve the linear programming problem in ϕ_{ij} (or $\phi_{ij}^{(n)}$) space and find the corresponding points in ϕ_i (or $\phi_i^{(n)}$) space. We can then perform the cost minimization over the convex hull of points in ϕ_i (or $\phi_i^{(n)}$) space.

Because of the network structure, each linear programming problem reduces to finding shortest route paths from each origin to each destination. Besides leading to great savings in computer time, this leads to a satisfying interpretation of this algorithm. At each master step (i.e., each time the minimum cost over the convex combination of the extremals is sought), the algorithm finds the best division of each origin-destination flow among the paths defined by the extremals. The algorithm terminates when the best set of extremals is found. Since the paths travelled are known, and the link flows are known, it is easy to calculate the travel time for each path, and thus the average travel time for each origin destination pair.

This discussion has been restricted to single vehicle class flows. The extension to multiple vehicle class flows is not difficult.

User Optimization

The preceding discussion is restricted to system optimization. We have experimented with an extension to user optimization with promising results.

The master step in the system optimization algorithm can be thought of as finding a solution to a set of nonlinear equations and inequalities. These are the Kuhn-Tucker conditions [17] which are the necessary conditions for a minimum in any optimization problem. Suppose instead, we satisfy a different set of nonlinear equations and inequalities: we find flows to equate travel times on paths defined by the extremals already generated. If this procedure converges, it creates flows that equate travel times on paths utilized between origins and destinations - the user optimization solution.

This may appear to be no more than an interesting mathematical abstraction; however, it is potentially of crucial importance. The literature of traffic assignment is filled with techniques that are only valid for separable travel time functions: where the travel time τ_i on link i depends on ϕ_i, the flow on link i, and possibly some parameters, but not on any other flows, ϕ_j, $j \neq i$. While many roadway links (e.g., freeways) are like this, many (such as entrance ramps (2.13) - (2.18)) are not. There may be relatively few of the latter, but they are located at crucial points. They could prevent a freeway from being as heavily utilized as a separable user optimization would predict. Similarly, weaving lanes can cause a bottleneck, but only when there is sufficient interacting traffic.

(Existing techniques work for symmetric delays as well, i.e., where $\partial\tau_i/\partial\phi_j = \partial\tau_j/\partial\phi_i$ for all i,j [8], [49]. This is more general than separable delays, but hardly more realistic.)

Yagar [16] shows that the traffic distribution measured in an actual network falls between the calculated system optimum and user optimization. It differs from the user optimization solution largely in that less actual traffic uses the freeway.

We propose a conjecture: his user optimization calculation failed to account for the reduction in effective capacity of entrance ramps due to freeway flow. This was unavoidable because of the separability assumption. Thus he underestimated ramp delays and overestimated ramp flow. This led to an overestimate of freeway flow. In this study we avoid this by including the effect of freeway flow on ramp delay in equations (2.13), (2.14).

3.3 Traffic Signal-settings

Procedures for determining traffic signal settings can be classified, broadly, into two categories: (a) single intersection; (b) interconnected intersections.

3.3.1 Single Intersection

In case of a single (isolated) intersection, the only variables to be determined are the cycle time c and the green splits g. The most common method for setting signals at isolated intersections is due to Webster [23], [61].

Assume that in Figure 3.2 road NS (north-south) and road EW (east-west) meet at a signal. Then the green splits satisfy

$$g_{NS} \geq 0; \; g_{EW} \geq 0, \tag{3.1}$$

and

$$g_{NS} + g_{EW} = 1 - \alpha, \tag{3.2}$$

where $\alpha > 0$ represents "lost" time. Assume links 1 and 3 are on NS and 2 and 4 are on EW. Define

$$y_i = \phi_i / s_i, \tag{3.3}$$

where ϕ_i is the total (passenger-car equivalent) flow on link i and s_i is the saturation flow on link i. Define

$$\left. \begin{aligned} y_{NS} &= \max\,(y_1, y_3) \\ y_{EW} &= \max\,(y_2, y_4) \end{aligned} \right\} . \tag{3.4}$$

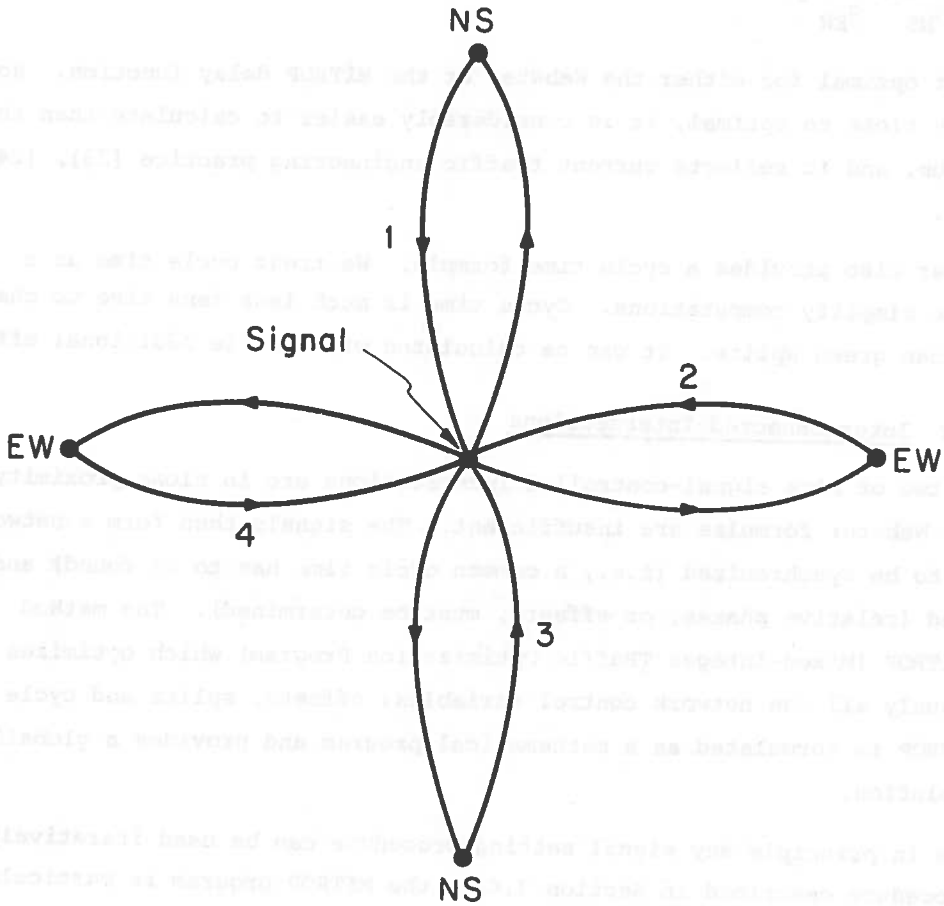


Figure 3.2 Signalized Intersection

Choose g_{NS}, g_{EW} to satisfy (3.1), (3.2), and

$$\frac{y_{NS}}{g_{NS}} = \frac{y_{EW}}{g_{EW}} . \tag{3.5}$$

This is not optimal for either the Webster or the MITROP delay function. However, it is close to optimal, it is considerably easier to calculate than the true optimum, and it reflects current traffic engineering practice [23], [24], [61], [62].

Webster also provides a cycle time formula. We treat cycle time as a constant to simplify computations. Cycle time is much less sensitive to changes in flows than green splits. It can be calculated with little additional effort.

3.3.2 Interconnected Intersections

When two or more signal-controlled intersections are in close proximity, the simple Webster formulas are insufficient. The signals then form a network which has to be synchronized (i.e., a common cycle time has to be found) and coordinated (relative phases, or offsets, must be determined). The method used is MITROP (Mixed-Integer TRaffic Optimization Program) which optimizes simultaneously all the network control variables: offsets, splits and cycle time. MITROP is formulated as a mathematical program and provides a globally optimal solution.

While in principle any signal setting procedure can be used iteratively in the procedure described in Section 3.4.1, the MITROP program is particularly suitable. This is because (a) it provides an explicit analytical expression for the travel time on the link as a function of the flow on that link (equation (2.20)); (b) it optimizes simultaneously all the traffic signal control variables in the network, including cycle time, green splits, and offsets; (c) it is capable of optimizing phase sequencing, in addition to the usual signal control variables, which is an important decision variable when rerouting of traffic is contemplated. See Appendixes A and B.

3.4 Overall Procedures

In this section we describe how the assignment, signal setting, and mode split calculations are integrated. It should be emphasized that in this pilot study, several formulations were considered. Those actually implemented are described in the discussion on numerical examples in section 4. Our goal here was to explore this problem formulation and determine feasibility of this solution approach. Many issues were raised which can only be resolved with further investigation.

3.4.1 Assignment and Signal-Settings

Figure 3.3 illustrates the overall program flow when an assignment and a set of signal settings are calculated together. Signal settings are guessed (1) and the corresponding assignment is calculated (2). Given the resulting flows, a new set of signal settings is found (3). A stopping criterion is tested, and if the procedure does not terminate, a new assignment is calculated (2) and the procedure repeats.

This is a heuristic procedure. We have no proof of convergence; however, it has converged in the examples considered.

It is clear [42] that signal settings have an effect on traffic assignment. The procedure we describe here is similar to others considered in the literature [43], [44]; however, we provide numerous options and we believe that our user optimization technique may be more accurate than existing methods.

Within this flow chart, there are the following options.

1) System optimized or user optimized assignments are possible.

2) If system optimized, any positive linear combination of fuel consumption and travel time can be minimized. See equation (2.32).

3) The assignment can be performed for single or multiple vehicle classes. The multiple class version includes non-pooled cars, pooled cars, and buses.

4) The assignment can be calculated assuming the Webster delay function (equation (2.19)) or the MITROP delay function (equation (2.20)) on signalized links.

5) There are two ways of calculating signal settings (box 3 of the flow chart).

   a) Equalizing degree of saturation (i.e., satisfying (3.1) - (3.5)).

   b) The MITROP program [25] - [27]. When the MITROP delay function

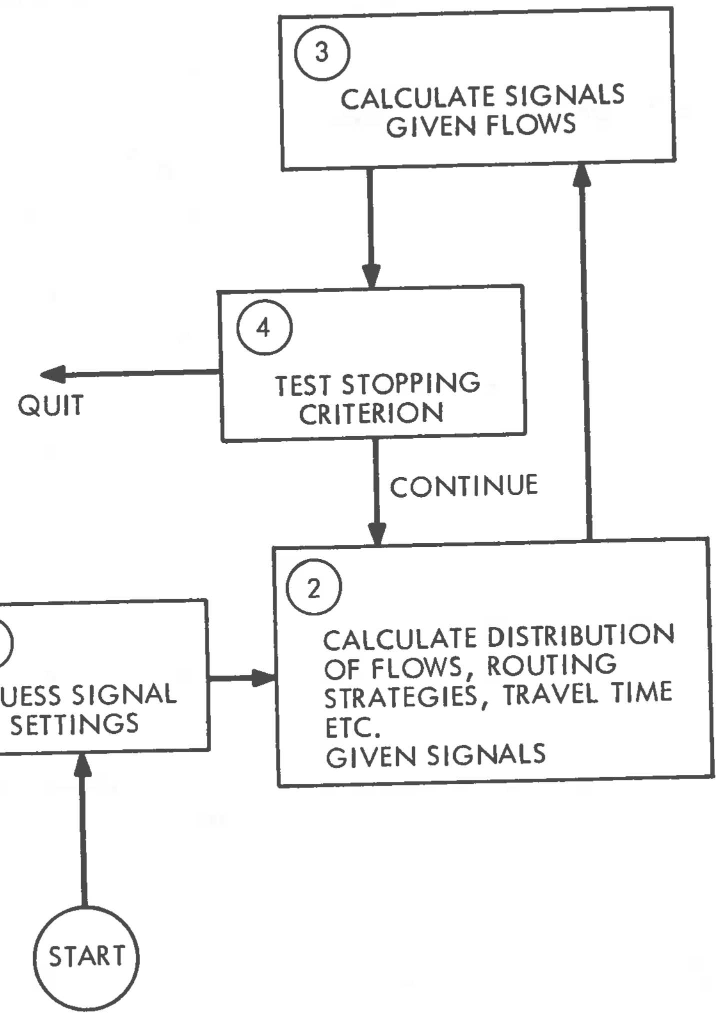


Figure 3.3 Flow Chart for Calculating Signal-Settings and Assignments

(2.20) is used in the assignment, the MITROP program may be used to calculate simultaneously green splits, cycle time, and phase offsets between adjacent signals in networks. The program, of course, requires considerably more computer time and storage than (3.1) - (3.5), but it produces much better results in terms of delay (and therefore of fuel). To save computer expense, the following options are available.

b.1) If multiple class assignment is required, the first few iterations can be performed with the single class version of the assignment. Clearly, great accuracy in assignment is not needed until the late stages, so some economy is possible in the beginning.

b.2) The MITROP program need not be used fully at every iteration. For example, the full MITROP, which optimizes integer variables, might only be required once or twice in a run. A restricted MITROP, with the integers held fixed, can be used for other iterations. This is much faster than the full program. Finally, for most iterations, we can hold cycle times and offsets constant and adjust green splits by equating degrees of saturation by (3.1) - (3.5). This is fastest of all.

6) Various stopping criteria can be employed. Are successive assignments sufficiently close? Are successive signal parameters sufficiently close? Have we exceeded some prespecified number of iterations?

3.4.2 Exclusive Ramps and Lanes

7) If the assignment is multiple class and system optimized, the network can have potentially exclusive lanes. For instance, the networks in section 4 are the same except that in one some structure is added to the freeways. The leftmost lane in both directions is separated from the two right lanes. This must be done carefully: there must be an opportunity for lane change between entrance and exit ramps, and the coefficients in the polynomial expression for $t_i(\phi_i)$ must be compatible with the number of lanes.

If it is optimal to do so, the assignment program will forbid a class of vehicles from the exclusive lane. Note that it is the programmer's option to make the lane <u>potentially</u> exclusive by separating it in this way. The program can do the same with entrance ramps or other links.

8) If the assignment is multiple class and user optimized, exclusive lanes can be separated as in item (7). Here, however, the programmer must specify which lanes are to be exclusive and which classes are excluded. This is because reserving a lane is for the good of the system, but the assignment is chosen by the users. With these constraints specified, the user optimized assignment is found.
A possible procedure for exploring user optimization with exclusive links is as follows: use item (7) to find exclusive lanes, entrance ramps, or other reserved links. Then perform the user optimized assignment with these reservations. This procedure has some limitations, and a proposed remedy appears in Section 5 and Appendix C.

3.4.3 Modal Split

Figure 3.4 contains the flow chart for integrating mode split with assignments and signal setting calculations. The mode split -- that is, the entire origin destination demand matrix broken down by mode ($r_{ij}^{(n)}$ of equation (2.5)) -- is guessed (1). The corresponding assignment and green splits are then calculated (2) according to one of the multiple class options described in subsection 3.4.1 and Figure 3.2. A new requirement matrix is calculated (3) from (2.28) - (2.31). A stopping criterion is tested, and if the procedure is not terminated, it proceeds to step (2). This is heuristic, as is the assignment-signal setting procedure, but it has performed satisfactorily.

Other authors have combined demand studies with assignment [14], [45], [46] and with modal split [47], [48], [49]. Florian [33] solves the mode split problem (with only automobile and transit traffic) by an iterative procedure. He guesses transit travel times and solves an elastic demand problem in the automobile sector. This leads to new transit travel times and the procedure repeats. The procedure presented here is unique in its flexibility and its approach to user optimization.

Numerous options are available.

1) Any multimodal and signal setting options described in subsection 3.4.1 may be used for step 2, as well as simply keeping signal settings fixed.

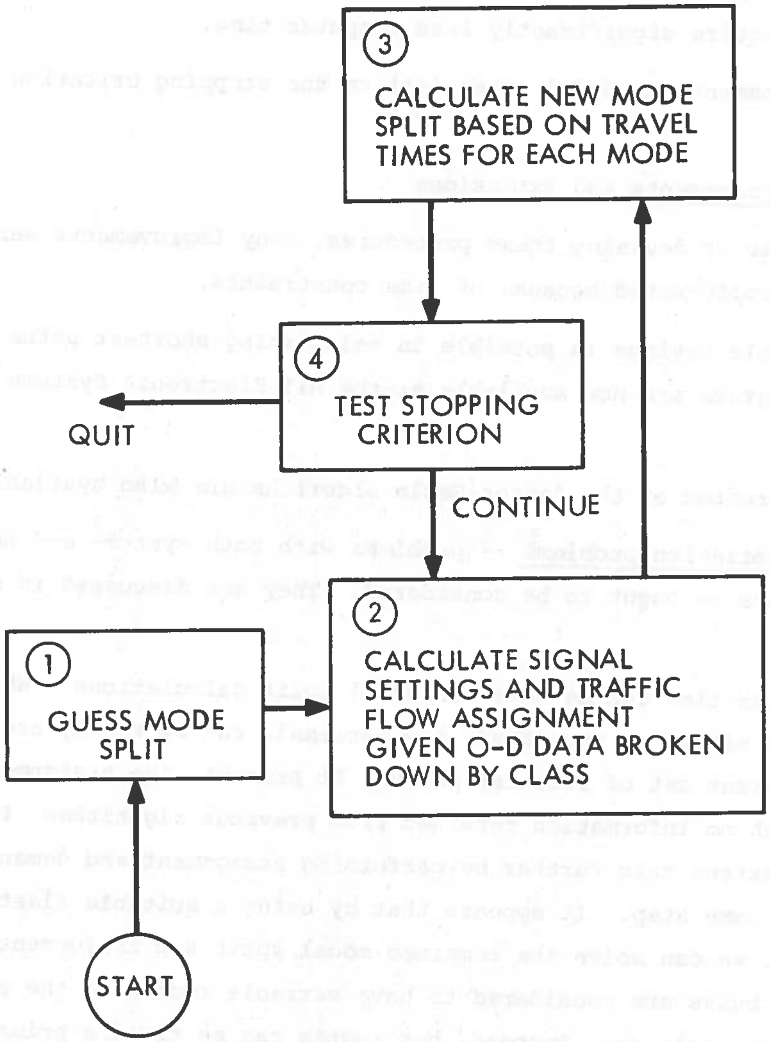


Figure 3.4 Flow Chart for Mode-Split Calculation

As suggested above, the less accurate options can be used for early iterations if they require significantly less computer time.

2) The comments in 3.4.1 (item (6)) on the stopping criterion apply here as well.

3.5 Possible Improvements and Extensions

In the course of devising these procedures, many improvements were considered but not implemented because of time constraints.

A considerable savings is possible in calculating shortest paths [39]. The required programs are now available at the MIT Electronic Systems Laboratory.

Improved versions of the Cantor-Gerla algorithm are also available [39].

Hybrid optimization problems -- problems with both system- and user-optimized elements -- ought to be considered. They are discussed in Section 5 and Appendix C.

Some computer time can be saved in modal split calculations. When a new requirements matrix is generated, new extremals can be easily created based on the current set of shortest paths. At present, the assignment is restarted with no information retained from previous algorithms It may be possible to extend this further by performing assignment and demand analysis in the same step. It appears that by using a suitable elastic demand function, we can solve the combined modal split and assignment problem.

At present buses are considered to have variable routes in the same sense that cars and car pools do. Instead, bus routes can be fixed *a priori*. This requires in advance a fixed, single extremal for the bus class.

The model can be made more accurate with the inclusion of bus stop delay, left turn delay, multiphase traffic signals, and other more realistic effects.

Other demand functions and problems where passenger (and not just vehicle) flow demand is to be found can be considered.

3.6 Summary

An optimization method, originally designed for communications network problems, has been extended to multiple-vehicle-class, system and user optimization transportation assignment problems. It has been coordinated with methods for calculating optimal traffic signal settings and a method for predicting modal split. Several improvements and extensions have been suggested.

4. EXAMPLES

Numerical results and computation expenses* are discussed in this section. Single class system optimized assignments with vehicle travel time and fuel consumption criteria are compared in Section 4.1. In Section 4.2, four triple class assignments are discussed. Two are system optimized with passenger travel time and total energy consumption as criteria, and two are user optimized with and without reserved freeway lanes. In Section 4.3 we repeat one of these assignments with an accident, reducing freeway capacity. The use of MITROP is demonstrated in Section 4.4, and a modal split run is discussed in Section 4.5.

These examples indicate that this approach is indeed feasible. They also indicate that energy savings is possible by using reserved lanes and they help to develop insight into traffic network behavior.

4.1 Single-Vehicle Class Assignments

Figure 4.1 displays the hypothetical network that we have considered in single class assignment runs. This network includes a three-lane freeway, freeway entrance and exit ramps, and signalized arterials. A schematic map of this network appears in Figure 4.2. The freeway links are links 35, 36, and 38 (carrying traffic west) and 39, 40, and 41 (east). Links 15 and 22 are entrances and 37 and 42 are exits. The remaining links are signalized arterials. We assume that this is a part of a much larger network which primarily carries traffic east and west.

Links join at indicated nodes. Traffic signals, numbered 1-8, are located at nodes 2 to 5 and 8 to 11. Delays at traffic signals are given by the Webster formula, equation (2.19). Traffic can originate at any node (that has one-way links which carry traffic away from the node) and travel to any other, but we have only chosen the external nodes (i.e., 1, 7, 6, 12, 15, and 21-28) to be origins or destinations. Note that we have grouped all three western-most destination nodes together into node 12, and all three

* Times and expenses are given for the MIT Information Processing Center IBM 370/168.

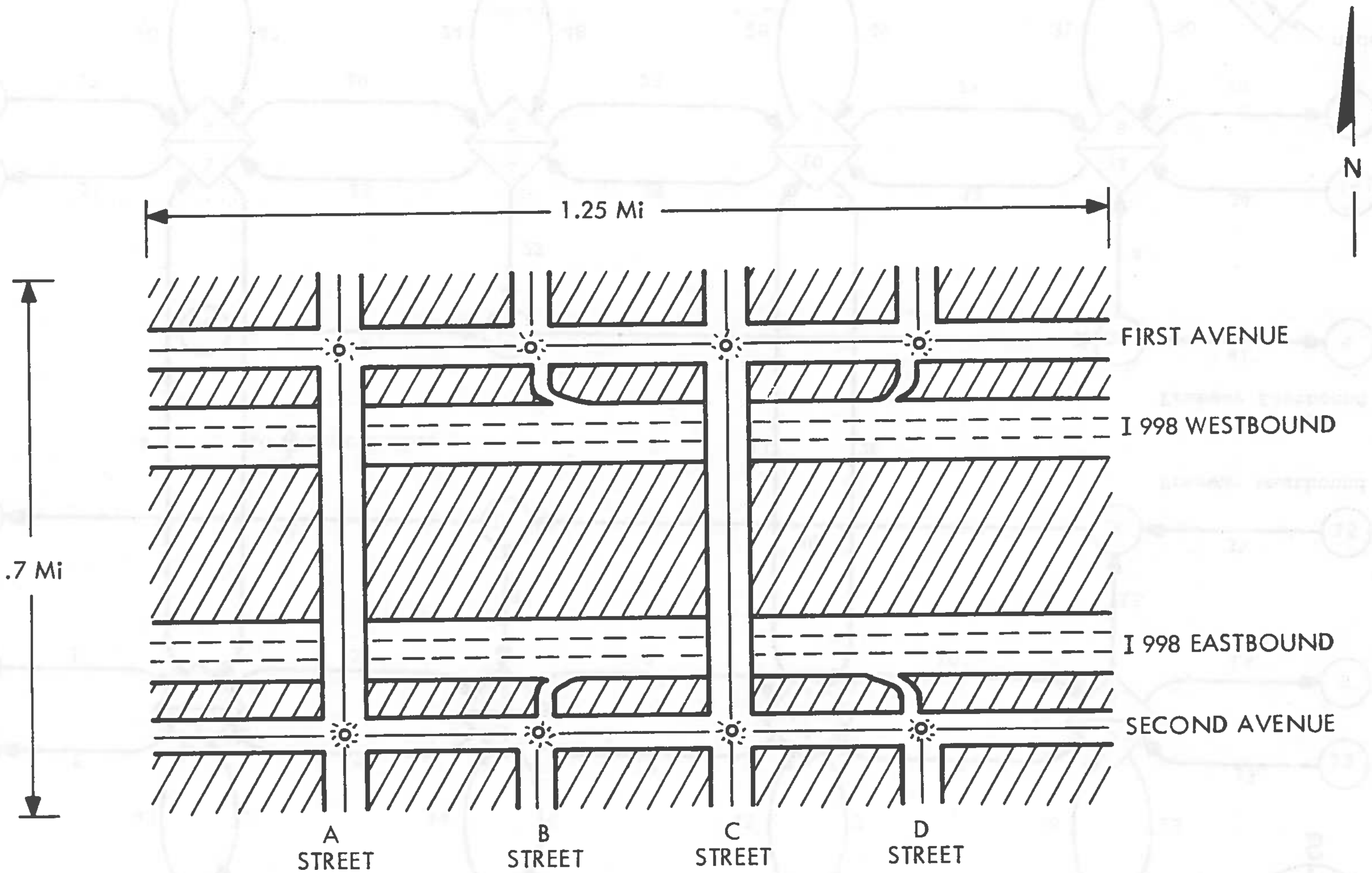


Figure 4.1 Sample Network: Physical Map

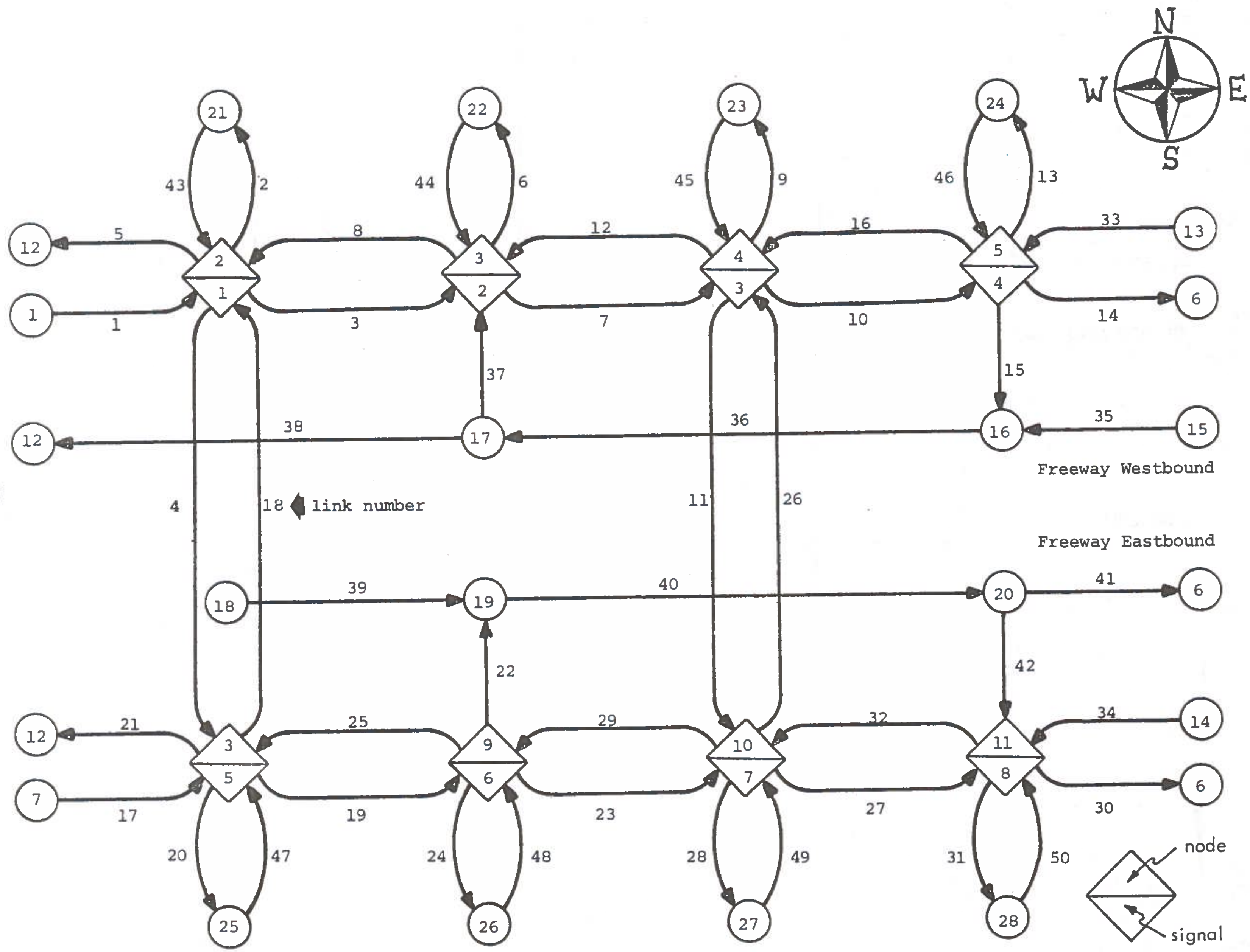


Figure 4.2 Roadway Network: Schematic Map

eastern-most destinations as node 6. This is because we have assumed that traffic destined for any node labeled 12 is traveling west and is indifferent as to which exit from this part of the network it uses. Network details, including link lengths, capacities, origin-destination data, and other parameters, appear in Appendix D.

Figure 4.3 shows the results of two runs with this network. The numbers displayed are the system optimized flows according to two criteria. The first number on each link is the energy optimal flow (W_1=1, W_2=0 in equation (2.32)) and the second is the time optimal flow (W_1=0, W_2=1). Table 4.1 displays the east-west green splits for these cases. We have assumed that at each intersection the cycle time is 60 seconds and that the sum of the east-west and north-south green splits is 0.9 (i.e., $\alpha = .1$ in equation (3.2)).

Table 4.2 lists the values of the cost functions. Since both runs are system optimizations, it is to be expected that the energy cost is smaller in the energy run and the travel time is smaller in the travel time run. What is striking is the very small differences in the costs.

We also observe that there is a greater discrepancy in the travel time cost. This appears to be part of a general pattern: the travel time is much more sensitive to changes in flows and green splits than energy consumption.

The energy run took 1.94 seconds of computation time to do 5 Cantor-Gerla iterations and 3 green split updates. This cost $9.89 of which approximately 90% was computer overhead and intermediate printing. The delay run took 2.51 seconds to do 8 Cantor-Gerla iterations and 3 green split updates for a cost of $10.20. Again, approximately 90% of this cost is due to overhead or extra printing.

The computation times and expenses we have quoted, while small, are even still inflated. In practice, a much wider stopping tolerance could have been used. In the travel time run, the travel time cost was down to 439.9 vehicle-hours per hour after 4 iterations and 1 green split update. The energy cost, at that point, was 815.9 gallons per hour. In the energy cost run, the energy cost was indistinguishable (to the accuracy presented in Table 4.2) from the final cost after 2 iterations and no green split updates. (The delay cost at

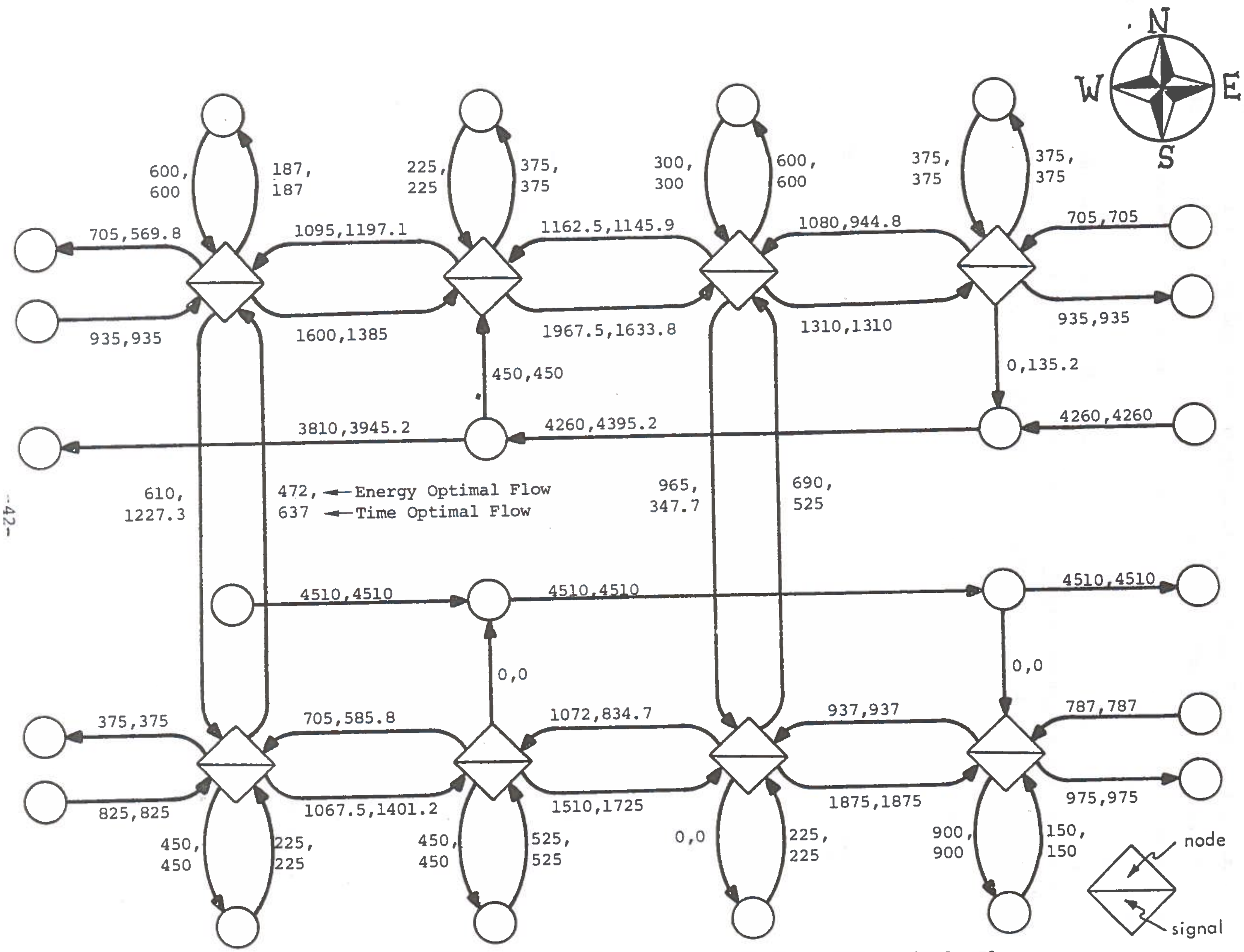


Figure 4.3 Energy and Delay Optimal Flows, Single Class

TABLE 4.1

GREEN SPLITS -- SINGLE-VEHICLE CLASS ASSIGNMENT

Signal	Green Split: Energy Optimization	Green Split: Time Optimization
1	0.581	0.587
2	0.576	0.546
3	0.666	0.681
4	0.700	0.700
5	0.517	0.362
6	0.604	0.655
7	0.549	0.749
8	0.833	0.833

TABLE 4.2

CRITERION VALUES -- SINGLE-VEHICLE CLASS ASSIGNMENT

Run	Energy Cost Gallons/hour	Travel Time Cost Vehicle-hours/hour
Energy Optimization	815.5	422.3
Travel Time Optimization	816.4	437.7

that point is not available, but after 3 iterations and 1 green split update, it was also indistinguishable from the result in the table, to the accuracy of the table.)

4.2 Multiple-Vehicle Class Assignments

Figure 4.4 displays a network used for multiple vehicle class assignments. This network is exactly the same as the one in Figure 4.2 with one exception: the freeway, in both directions, is separated. The two right lanes constitute one set of links (35, 37, 54, and 39 west and 40, 42, 56, and 43 east) and the left lane corresponds to another (36 and 53 west and 41 and 55 east). We assume that lane changing can take place only at certain points: at nodes 15, 29, and 12 for traffic flowing west and at 18, 30 and 20 for traffic flowing east. A detail of the physical map which illustrated this appears in Figure 4.5.

This means, for instance, that traffic entering the freeway at node 16 from entrance ramp 15 cannot reach the left lane until node 29 and link 53. This models the fact that buses which wish to use the left lane must interfere with right-lane traffic over some distance.

Note that some links and nodes are numbered differently from those in Figure 4.2.

Figure 4.6 displays the results of a system optimized assignment with fuel consumption as the criterion function. At each link, flows are broken into single-passenger car, car pool, and bus flows. We have chosen a requirements matrix $r_{ij}^{(n)}$ so that the required vehicle flow is equivalent to that of the single class assignments discussed in Section 4.1. That is, if we define

$$r_{ij} = \sum_{n} e^{(n)} r_{ij}^{(n)} , \qquad (4.1)$$

where $e^{(n)}$ is the passenger car equivalent of mode n, then r_{ij} is precisely the requirement matrix that led to the assignments in Figure 4.3. We use $e^{(1)} = e^{(2)} = 1.0$ and $e^{(3)} = 3.0$. Green splits are displayed in Table 4.3. Table 4.4 shows total fuel consumption in gallons per hour and travel time cost, in passenger-hours per hour. The latter is calculated assuming vehicle occupancies (i.e., $w^{(n)}$ of equation (2.33)) of 1.0 passengers for cars, 2.5

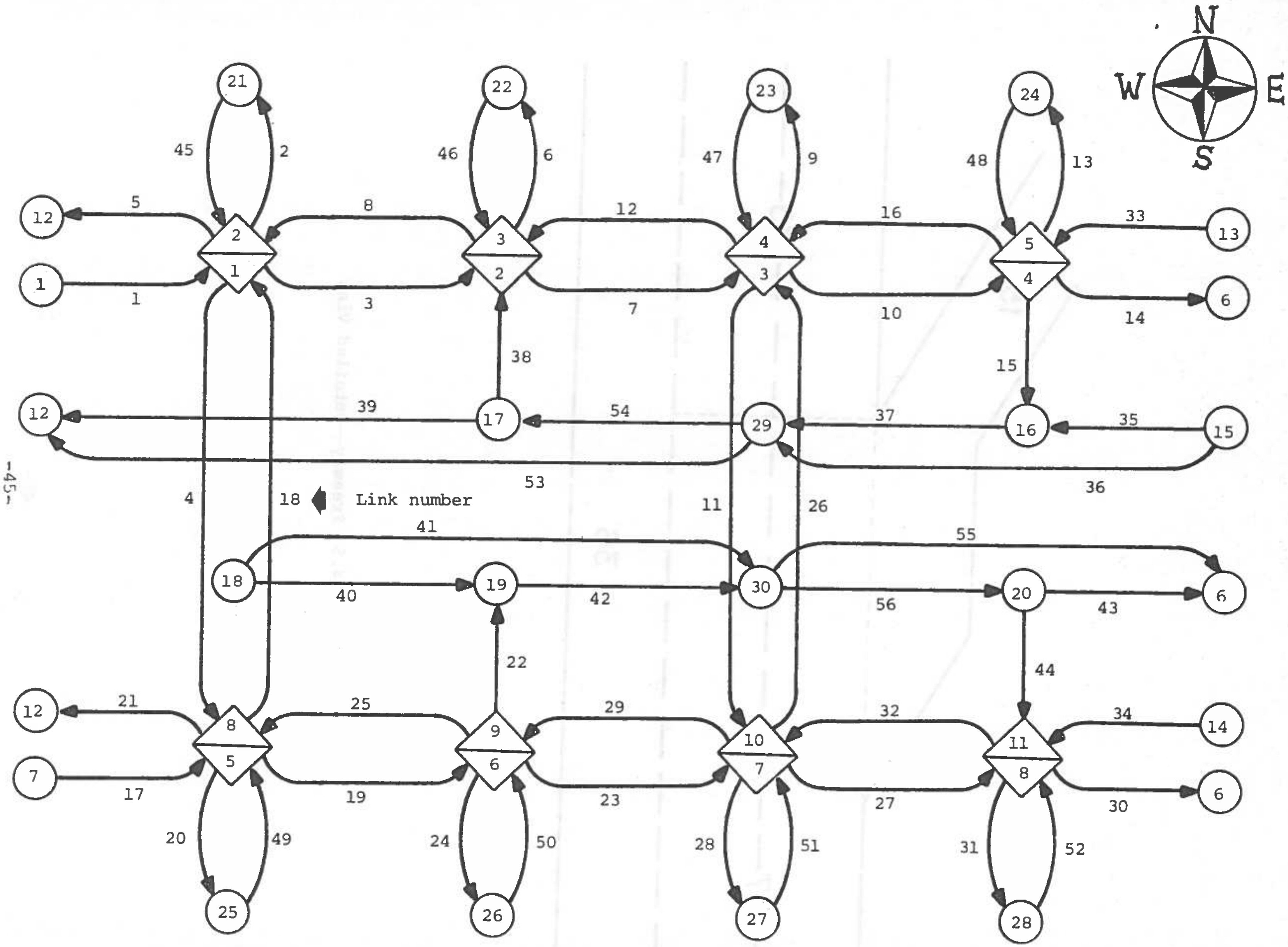


Figure 4.4 Network with Freeway Lanes Separated

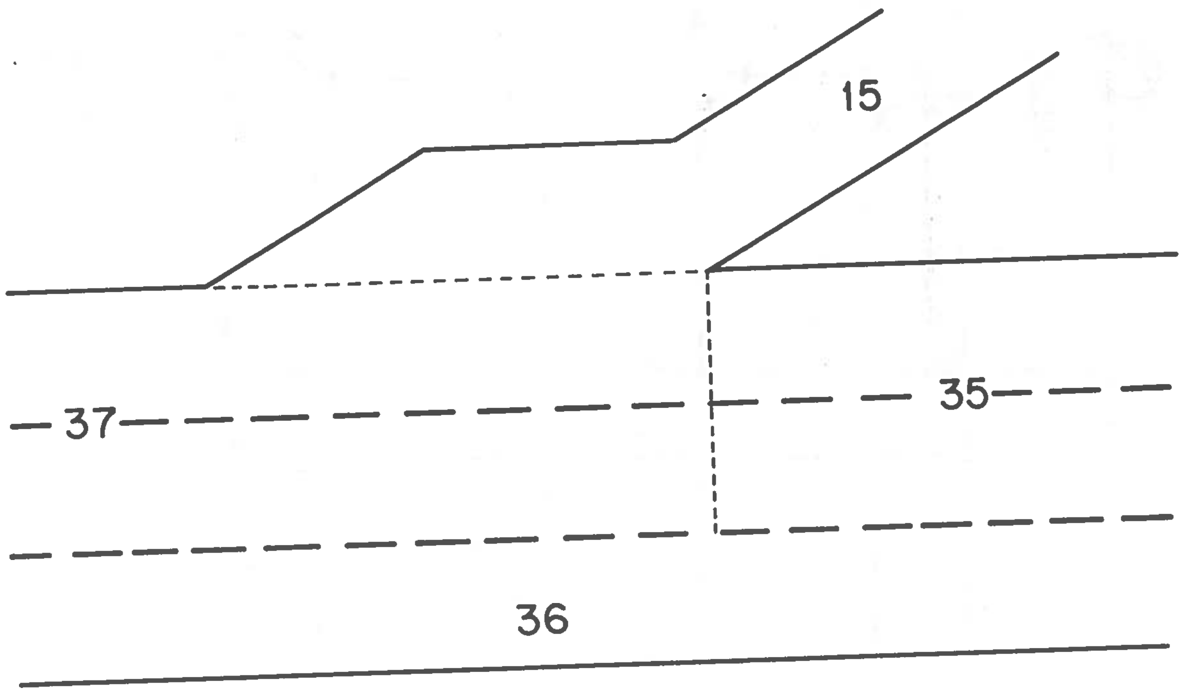


Figure 4.5 Freeway--Detailed View

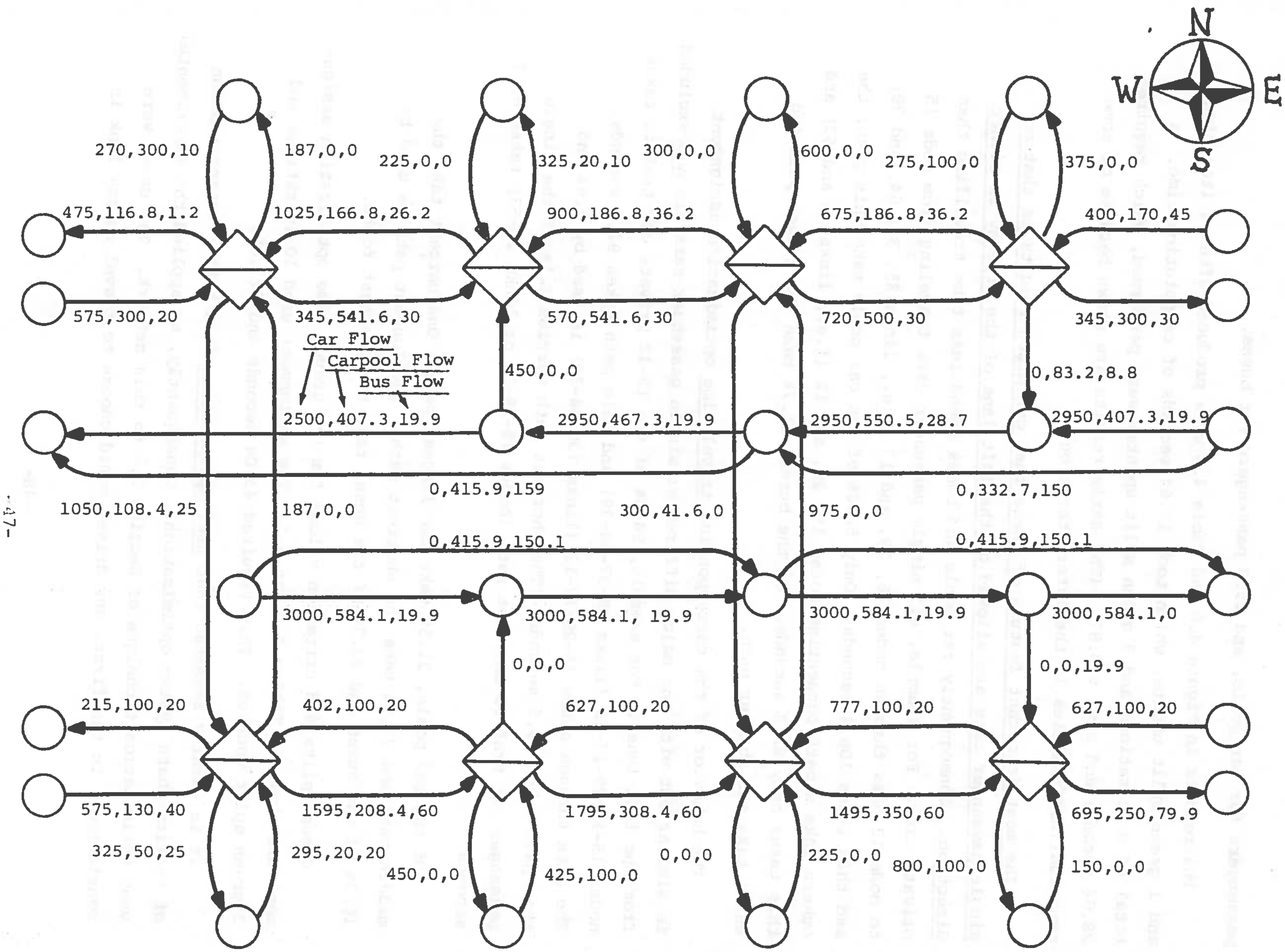


Figure 4.6 Energy-Optimized Assignment, Three Classes

passengers for car pools, and 25.0 passengers for buses.

The results in Figure 4.6 and Table 4.3 were produced after 4 iterations and 1 green split update, which took 11.64 seconds of computation time. A total of 8 iterations and 3 green split updates were performed, which required 28.94 seconds and cost $18.87. (The early results are shown because of some numerical difficulties in the later iterations.)

The most important feature of these fuel optimized results is that no single passenger cars are allowed on the left lane of the freeway in either direction. Consequently car pools and buses spend less time travelling than private cars. For example, all single passenger cars traveling from node 15 to node 12 pass through nodes 16, 29, and 17 (i.e., links 35, 37, 54, and 39) and this takes 106.4 seconds. Only 58.4% of the car pools take this path; the others take a path connecting nodes 15, 29, and 12 (i.e., links 36 and 53) and this takes only 82.3 seconds. Of the buses, 11.7% take the longer path and 88.3% take the shorter path.

The behavior of the corresponding travel time optimization assignment is similar but with one major difference: single passenger cars are not excluded from the left lanes. For example, 94.3% of the 15-12 private car traffic takes nodes 15-16-29-17-12 (links 35-37-54-39) and this path takes 94.3 seconds. The path through nodes 15-29-17-12 (links (36-54-39) is used by 3.6% and that path takes 89.5 seconds. The shortest path carries 2.1% of the single passenger car traffic and that path (nodes 15-29-12 or links 36-53) takes 85.7 seconds.

Of the car pools, 21.5% take the longest path. One percent take the middle path and 77.5% take the shortest path. The longest path is used by 16.3% of the buses and 83.7% of the buses take the shortest route.

Green splits and criterion values for the travel time optimization assignment are shown in Tables 4.3 and 4.4. The assignment used 10 iterations and 3 green split updates. This required 43.04 seconds and $23.08.

It is usually asserted that user optimization is a better representation of reality thatn system optimization. Consequently, we applied the experimental user optimization technique of Section 3.2 to this network. Two cases were considered. In the first, any driver could choose to travel on any link in the network. In the second, we observed the prohibitions resulting from the energy optimized runs: no single passenger cars were allowed on any link from which single passenger cars are absent in Figure 4.6.

In the first case (no prohibitions), all vehicles traveling between nodes 15 and 12 take all possible routes (i.e., nodes 15-16-29-17-12, 15-16-29-12, 15-29-17-12, and 15-29-12) and each route takes between 96.2 and 96.5 seconds. (Convergence is evidently not perfect; if it were, all these travel times would be the same.)

In the second case (with prohibitions), all single passenger cars choose 15-16-29-17-12 which takes 97.5 seconds. No car pools or buses take this path. All buses and 96.3% of the car pools travel on 15-29-12 which takes 91.6 seconds. The rest of the car pools take 15-29-17-12, and this also takes 91.6 seconds. (Again, convergence was not complete. If it were, either no car pools would choose the longer path, or both paths would take the same length of time.)

Table 4.3 displays the green splits from these runs and Table 4.4 shows the values of the cost functions. It is interesting to see that the "diamond" lane - i.e., the prohibition of single passenger cars from the left lane - slightly reduces energy consumption and significantly reduces average travel time. This is evidently an instance of Braess' paradox [6], in which the addition of a link increases network costs.

It is not surprising that travel time is diminished by the diamond lane. The vehicle demand from node 15 to node 12 is 2500 vehicles per hour for cars, 800 for car pools, and 170 for buses. Using 1.0, 2.5, and 25.0 for average occupancies, the passenger demand is 2500, 2000, and 4250 per hour for each of the modes. It is clear that any change that can reduce travel time to car pools and buses and not increase single passenger car travel times by much lowers the average travel time.

4.3 The Effect of an Accident on an Assignment

The last entry in Table 4.4 refers to a user optimization assignment on the network in Figure 4.4 and in which there is an accident on link 39. One of the two right lanes on the westernmost link of the westbound side of the freeway is blocked.

TABLE 4.3

GREEN SPLITS-- TRIPLE-VEHICLE CLASS ASSIGNMENT

Signal	Green Split Energy Optimization	Time Optimization	User Equilibrium Without Prohibitions	User Equilibrium With Prohibitions
1	0.611	0.600	0.548	0.535
2	0.513	0.500	0.468	0.551
3	0.497	0.527	0.496	0.683
4	0.700	0.700	0.700	0.700
5	0.361	0.382	0.360	0.360
6	0.712	0.702	0.712	0.671
7	0.783	0.738	0.768	0.737
8	0.838	0.836	0.827	0.828

TABLE 4.4

CRITERION VALUES -- TRIPLE-VEHICLE CLASS ASSIGNMENT

Run	Energy Cost Gallons/hour	Travel Time Cost Passenger-hours/hour
Energy Optimization	762.0	914.5
Travel Time Optimization	787.6	877.9
User Optimization Without Prohibitions	792.3	908.9
User Optimization With Prohibitions	791.5	891.5
User Optimization Without Prohibitions and With Accident	786.8	1001.1

The increase in travel time due to the accident needs no explanation. The energy consumption decreases because freeway velocities decrease. In addition, some traffic takes link 5 instead of 39. Link 5 is slower, and thus more energy-efficient than link 39 before the accident.

4.4 Joint Assignment and Signal-setting with MITROP

In a numerical experiment, the joint assignment and signal setting procedure of Section 3.4 was run with the MITROP cost function (2.20) and program (Appendix A). The cases in 4.2, by contrast, use the Webster cost function (2.19) and equalization of degree of saturation ((3.1) - (3.5)) to set green splits. The MITROP procedure is more expensive, but it leads to a dramatic savings in average travel time.

In this experiment, the network of Figure 4.2 was used. The assignment minimized passenger-weighted travel time with those vehicle classes. The procedure described in Section 3.4 was exercised as follows. Green splits were chosen arbitrarily, and an assignment using the Webster function (2.12) (Assignment 0) was calculated. The passenger car-equivalent flows were used as input to the MITROP program (MITROP 1) which calculated the cycle time, green splits, and the offsets. These were used as inputs to the assignment program (Assignment 1) using the MITROP cost function (2.20) and the procedure (MITROP and MITROP assignment) was repeated until Assignment 3 was reached. Selected results are listed in Table 4.5.

The procedure has not completely converged, but it is expected that no further significant changes in flows or other parameters would occur after further iteration. (The green splits and flows presented were among the <u>most</u> variable of all parameters.)

<u>The cycle time and energy consumption change very little</u>. The MITROP cost, which measures vehicle delay at signalized intersections, decreases significantly. The assignment cost drops dramatically after the Webster assignment, and then appears to settle down.

The comments made earlier about computer expenses apply here as well. No attempt was made to reduce MITROP costs by freezing integers, holding cycle time constant, or calculating green splits by (3.1) - (3.5). Computa-

TABLE 4.5. JOINT ASSIGNMENT--MITROP RESULTS

Run	Travel Cost person-hrs/hr	MITROP cost vehicle-hrs/hr	Energy Cost gallons/hr	Cycle Time seconds	Signal #1 Green Split	Signal #6 Green Split	Links Flow vehicles/hr	Computer Cost
Assignment 0	875.6		790.0	60.00	0.592	0.683	1348.9	$15.87
MITROP 1		54.8		57.54	0.480	0.642		18.24
Assignment 1	812.7		790.7				1316.2	9.89
MITROP 2		47.3		58.04	0.536	0.644		70.00 (approx.)
Assignment 2	827.8		790.2				1276.0	14.28
MITROP		40.1		58.24	0.497	0.645		32.50
Assignment 3	811.9		790.8				1260.7	14.47

tional experience with the MITROP program has shown that substantial savings in computer running time (on the order of 50-75 percent) can be achieved by limiting the branch-and-bound search process. In the cases investigated so far, the optimal solution (or one that was very close to it) was obtained at a very early stage of the search; most of the search time was devoted to providing optimality. Thus, if one is willing to forego a proof of optimality, the running time of a MITROP iteration can be similar to that of an iteration of the assignment procedure.

4.5 Modal Split

Two examples of mode split are considered, and together they illustrate the effect of a diamond lane on energy consumption. In both examples, the assignment principle is user optimization and both use the Webster traffic signal delay function. The heuristic algorithm of 3.4 was followed and results appear in Table 4.6.

The passenger demand from each origin to each destination (R_{ij}) is the same in the two runs. The diamond lanes induce a discrepancy in travel times between cars and the other modes. This leads to a shift in demand toward those modes. The last entry in Table 4.6 is the total vehicle flow rate in passenger car equivalents, i.e.

$$\Phi = \sum_{i,j,n} e^{(n)} \; r_{ij}^{(n)} .$$

The decrease in this quantity is a measure of the effect of the diamond lane on mode choice. The first item in the table shows how energy consumption is reduced.

The passenger demand rate is less* in these examples than in the cases in Sections 4.2 - 4.4. The higher demand was originally considered, but it led to numerical difficulties. After one iteration, private car demand increased and other mode demand decreased. As a result, total vehicle demand exceeded capacity.

This indicates that the parameters chosen are not consistent with the demands in the earlier examples. The $\alpha^{(n)}$ and $\beta^{(n)}$ parameters are shown in Table 4.7, and $\theta = .0025 \text{ (second)}^{-1}$. It should be recalled that those demands and these parameters were chosen to be illustrative, and we do not pretend that they are based on data.

It is reasonable to ask why it was necessary to reduce passenger demand to make this procedure converge, and whether this indicates a flaw in the approach. We conclude it does not, but rather leads to an important insight into network management.

The values in Table 4.7 lead to a mode split that exceeds network capacity when the passenger demand rates of the earlier examples are used.

---

* By a factor of 60 percent.

TABLE 4.6

MODAL-SPLIT RESULTS

Performance Measure	Without Prohibitions	With Prohibitions
Energy Cost (gallons/hours)	646.6	598.9
Travel Time Cost (passenger-hours/hour)	540.2	496.1
Total Vehicle Flow Φ (car equivalents)	11,246.5	10,036.1

TABLE 4.7

MODAL-SPLIT PARAMETERS

Parameter	n=1, Private Cars	n=2, Car Pools	n=3, Buses
$\alpha^{(n)}$, seconds	0	180	300
$\beta^{(n)}$	1.0	1.1	1.3

Capacity is exceeded on many links, including streets. This causes large delays to all vehicles on those links.

These links carry car pools and buses as well as cars. Consequently, car pools and buses experience long travel times. The existence of diamond lanes on freeways reduces this somewhat, but these vehicles must travel on non-separated links to reach them. As a result, there is little incentive for travelers to switch to these modes.

If, however, there were diamond lanes on streets and entrance ramps, we could guarantee that such lanes are operated below capacity. This would guarantee that the more efficient modes have shorter trips than private cars.

We can then reach the following tentative conclusion: diamond lanes can reduce energy consumption, but only if they are sufficiently widespread so that neither car pools nor buses are obliged to travel on any oversaturated link.

The procedure used here to choose the diamond lane for the freeway can also be applied to diamond lanes elsewhere. The hybrid optimization formulation discussed in Section 5 and Appendix C is also appropriate.

4.6 Discussion

All the runs presented above produced satisfactory results, but there were some difficulties with the accident run (Section 4.3) and the mode split run (Section 4.5).

In the user optimized accident run, the flow on some links is very near capacity. This retards the convergence of the algorithm, and in our results, not all paths have equal travel times. We expect that additional study will correct this problem.

In the mode split run, we have already mentioned that the total demand had to be reduced. Even still, convergence was slow since some links were loaded close to capacity. We believe again that this difficulty can be resolved.

4.7 Summary

Several numerical examples were treated using the methods of Section 3.

The differences between system and user optimization, between single- and multiple-vehicle-class assignments, between energy and travel time minimization, between equalizing degree of saturation and MITROP, and between fixed and elastic mode split have been demonstrated. We have also shown how this methodology can be used to formulate and help decide policy questions.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The purpose of the effort reported here was to determine if an integration of the techniques of traffic assignment, modal split demand analysis, and signal optimization is feasible. On the basis of analysis and experimentation, we conclude that such an approach to computer assisted traffic engineering can indeed be made practical. In this section we present several recommendations for further study which will make this integration useful. Although they include research tasks, we feel that the effort required appears bounded and not open ended.

The major result of this research has been the creation of a general framework. On this framework can be added models for roadway structures and phenomena not discussed in this report: weaving lanes [31], multiphase traffic signals, extra delay due to left turns, [54], etc. To do this, one must use, for example, the standard techniques of queueing theory [51], other mathematical theories, or empirical data to calculate steady state delays, velocities, densities, fuel consumption rates, demand level, and other quantities of interest. There is a considerable body of literature which models traffic behavior at isolated structures. We provide a method for studying the system-wide interactions of such structures.

In the following sections, we describe several recommendations for actions that will make this approach more accurate, convenient, and useful to the traffic engineer.

We conclude that if these research and programming tasks are performed, a practical computer aided traffic engineering procedure will result. This procedure will deal with many issues now faced by traffic engineers, and will be flexible in the sense that new features can be added.

5.2 User Optimization

The procedure described in Section 3.2 for calculating user optimization should be investigated further. It has been used, in this research, as an experimental technique, and appears successful in most cases considered. There are many refinements that should be studied: analogs of Defenderfer's

[39] TR and CH algorithms; improvements to the master step and to the method of calculating shortest paths; possibly a redefinition of the shortest path problem to be solved. These refinements can save considerable computer expense and many be useful in proving convergence. Although we have experimented with this using the Cantor-Gerla technique, any existing optimization technique, such as Frank-Wolfe, can be adapted to this method.

5.3 Hybrid Optimization

An important contrast between the methods used to analyze exclusive lanes in 3.3 under system and user optimization assumptions should be apparent. Under system optimization, we simply model a lane which could be reserved, and the solution to the problem would tell us if it should be reserved. This is because traffic is assigned in a way that is optimal from the system's point of view. The same is true of ramp metering: the solution to a system-optimal assignment reveals the optimal level of ramp metering at each ramp.

Under user optimization, exclusive lane reservation requires programmer specification. The programmer has to specify, in a suitable network, which lanes are forbidden to which classes of traffic. That is, the program does not, automatically, indicate which lanes should be closed to which classes of vehicles. In order to find this, the programmer would have to perform a larger number of tests, trying out all plausible configurations of lane reservations. To save computer expense, a heuristic approach is suggested in Section 3.4 instead: first do a system optimized run to find the reserved links. Then find the user optimized flows with the system-optimized link reservations.

This contrast is due to the fact that the user optimized solution is not the best from the system's point of view. It is the operating agency, however, that decides on link reservations, and the decision is made from system-wide considerations such as minimizing average travel time or total energy consumption.

Similar considerations apply to metered entrance ramps and signalized intersections: the system, with system wide considerations, chooses control parameters such as metering rates, green splits, etc. In principle, those parameters should be found by performing a user optimization run with all possible combinations of parameters. This is absurd, and we have instead followed the heuristic algorithm of Section 3.4.

In general, an operating agency can control a limited number of links. If it could control all vehicles, system optimization would be the appropriate assumption. If it controlled the flow on no links user optimization would be correct.

Instead, we define a <u>hybrid</u> optimization assignment. Assume that the operating agency can control a <u>restricted set</u> of links by limiting the flow on them. On all paths, drivers equalize travel times. That is, user optimization prevails. On a controlled path (i.e., a path including controlled links), the agency restricts the flow so that the overall system-wide cost is minimized. This is discussed in more detail in Appendix C.

Additional research must be done to complete the statement of the hybrid optimization conditions. It is anticipated that computer algorithms can be constructed to implement these conditions. Such an algorithm may be an extension of the Cantor-Gerla algorithm [41] or Defenderfer's TR and CH algorithms [39] in a way analogous to that described above for user optimization. Convergence properties of such a procedure must be studied.

5.4 Model Credibility, Validation, and Calibration

Before a traffic engineer can make use of this computer assisted approach, he must be assured that the models used have been empirically verified or that they are part of the standard traffic engineering literature. This is true for most of the elements of the model described here.

1) A polynomial has long been used for a freeway delay function [2]. Other formulas (for example, those in [19], [20], [21], [22]) can easily replace this, or, a polynomial can be used to approximate such a formula to any required accuracy.

2) Both the Webster [23], [61] and the MITROP [25] formulas have been based on empirical data. Some compromises were, however, made in the MITROP program in dealing with turning traffic. MITROP has not been validated for low flow conditions, where a cost function based on bandwidth [62] may be more appropriate.

3) The ramp delay expression has appeared in the literature [32], but other maneuver models have as well (in [32], [53], [54], [55], [56]). A sensitivity analysis may be appropriate. In addition, we have pointed out (in Section 2.6) that a model with additional features would be useful for energy evaluation and minimization.

4) The energy consumption functions are entirely empirical [28], [29]. We must still caution that fuel consumption results are as credible as our velocity and delay models. In addition, neither [28] nor [29] provide data for buses so we were obliged to assume (i) that on freeways and ramps, bus fuel consumption is the same as six-tire, two-axle truck fuel consumption, and that (ii) on the signalized network, the bus fuel consumption rate is 1.523 times the automobile fuel consumption rate. Furthermore, technological change and federal law will render these data obsolete. For accurate results, fuel consumption studies such as [28], [29] should be repeated periodically.

We have assumed a flat straight freeway with an asphalt of high-type concrete surface and freely flowing traffic. Claffey [28] estimates corrections for grades, rough surfaces, and congestion. These corrections can further enhance the realism of this study, and would be easy to include.

5) A potentially vulnerable area in this study is that of modal split. Florian says, "In the development of models that serve to plan future transportation systems, the demand function seems at present to be the weakest link... each important study seems to produce a different form for the demand function." [33] The mode split model we use (equations (2.33) - (2.36)) is the same as in [33]. Others appear in [48], [49], [57], [58].

Model calibration is discussed in [58] and, in a different context, [59]. In both cases, the procedure was the same: hypothesize a mode split formula (several formulas in [59]), obtain data for all the mode attribute variables considered (e.g., travel time, comfort, etc.) and the demand, and perform a regression on the model parameters. This appears not to be a particularly difficult procedure, especially since there are only a limited number of parameters.

To calibrate a model such as that defined by equations (2.33) - (2.36) the following information is required: passenger origin-destination demand data and average travel times broken down by mode for at least a limited number of origins and destinations. In order to inspire confidence, there should be at least an order of magnitude more data than parameters. Since in our model there are seven parameters (Θ and $\alpha^{(n)}$ and $\beta^{(n)}$, n=1,2,3), there should be travel demand and time data for 70 origin-destination pairs.

5.5 Application to Large Networks

It is likely that some important applications of the procedure described here will involve much larger networks than in Section 4, for example, networks of 500 nodes or more. In this subsection, we consider the computer expense of such a run. We conclude that the expense for the full procedure (i.e., three vehicle classes, signal calculation, mode split analysis) may be an order of magnitude greater than that of a single class assignment.

Computer times and expenses for small networks are discussed in Section 4. Nguyen [5] applies a Frank-Wolfe technique to the single class assignment of traffic in the city of Hull, Canada. The network has 376 links, 155 nodes, and 690 O-D pairs, and an assignment took as little as 21 seconds (longer with a smaller tolerance).

Cantor and Gerla recommend a procedure for multiple class minimization which we use here. This is an efficient decomposition so that the linear programming step is replaced by three single class minimizations (for three classes) and the master step optimizes over three times as many variables. A similar procedure can be adapted to the Frank-Wolfe method. In that case, the master step can be replaced by a search over three variables, or three one-dimensional searches, or a single one-dimensional search. Analysis is required to establish which is the most efficient overall approach.

In any case, we can conclude that a triple class assignment is five to ten times as expensive as a single class assignment.

Florian and Nguyen [45] use a procedure similar to that of Section 3.4 for combined demand analysis and assignment. Florian [33] recommends a similar procedure for combined mode split analysis and assignment. This procedure takes approximately 25% longer than an assignment alone. While our procedure currently takes longer, it seems possible to reduce its time requirements to that level.

We anticipate, as well, that improved efficiencies are possible with signal setting calculations. As we point out in Section 3.4, a full MITROP computation need not be done more than twice per run.

This crude calculation indicates that a full triple-class assignment with signal settings and mode split analysis should be roughly ten times as expensive as a single class assignment alone. While this is not cheap, it provides at least an order of magnitude more information. Furthermore, we have not taken into account all the recommendations for saving time discussed in this report.

5.6 Programming and Algorithm Improvements

The existing set of computer programs is not in suitable condition for routine traffic engineering use. Some algorithm changes would reduce computation expense and some programming changes would make this convenient for the traffic engineer. We repeat some suggestions made earlier and add new ones.

1) Incorporate improved shortest path algorithms [39].

2) Incorporate the CH and TR algorithms [39].

3) Improve the master step for system and user optimization.

4) Investigate hybrid optimization.

5) Investigate time savings in modal split calculations. This may result in an algorithm similar to Florian's [33].

6) Put all the procedures together in a package that may be easily used. At present, there are different assignment programs for single and multiple classes, and Webster and MITROP cost functions. Also, there is a considerable amount of effort involved in interfacing with MITROP.

7) MITROP is, at present, an experimental code which uses a general purpose mixed integer linear programming package. Considerable computer expense can probably be saved by writing a special purpose routine which will make use of the special network features of the problem.

6. REFERENCES

1. J.G. Wardrop, "Some Theoretical Aspects of Road Traffic Research", Proc. Inst. Civil Engineers, Part 2, pp. 325-378, 1952.

2. Traffic Assignment Manual for Application with a Large, High-Speed Computer, U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, June 1964.

3. Traffic Assignment, Comsis Corporation, Wheaton, Maryland, Prepared for the Urban Planning Division, Office of Highway Planning, U.S. Department of Transportation, Federal Highway Administration, August 1973.

4. S.C. Dafermos and F.T. Sparrow, "The Traffic Assignment Problem for a General Network," J. of Research of the National Bureau of Standards, Vol. 37B, pp. 91-118, 1969.

5. S. Nguyen, "An Algorithm for the Traffic Assignment Problem," Transportation Science, Vol. 8, No. 3, pp. 203-216, 1974.

6. P. A. Steenbrink, Optimization of Transport Networks, Wiley, 1974.

7. P.A. Steenbrink, "Transport Network Optimization in the Dutch Integral Transportation Study," Transportation Research, Vol. 8, pp. 11-27, 1974.

8. S.C. Dafermos, "An Extended Traffic Assignment Model with Applications to Two-Way Traffic," Transportation Science, Vol. 5, pp. 366-389, 1971.

9. H.J. Payne and W.A. Thompson, "Traffic Assignment on Transportation Networks with Capacity Constraints and Queueing," 47th National ORSA Meeting/TIMS 1975 North American Meeting, 1975.

10. S.C. Dafermos, "The Traffic Assignment Problem for Multiclass-User Transportation Networks," Transportation Science, Vol. 6, pp. 73-87, 1972.

11. D.F. Wilkie and R.G. Stefanek, "Precise Determination of Equilibrium in Travel Forecasting Problems Using Numerical Optimization Techniques," Proceedings of the 1972 IEEE Conference on Cybernetics and Society, pp. 453-460, 1972.

12. S.D. Miller, M.J. Payne, and W.A. Thompson, "An Algorithm for Traffic Assignment on Capacity Constrained Transportation Networks with Queues," Presented at the Johns Hopkins Conference on Information Sciences and Systems, April 1975.

13. L.J. LeBlanc, E. Morlok, and W.P. Pierskalla, "An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem," Transportation Research, pp. 309-318, 1976.

14. M. Beckman, C.B. McGuire, and C. Winston, Studies in the Economics of Transportation, New Haven, Connecticut: Yale University Press, 1956.

15. R.B. Dial, "A Probabilistic Multipath Traffic Assignment Model Which Obviates Path Enumeration," Transporation Research, Vol. 5, pp. 83-111, 1971.

16. S. Yagar, "A Study of the Travel Patterns in a Corridor with Respect to the Assignment Principles of Wardrop," in Traffic Flow and Transportation, G.F. Newell, Ed., American Elsevier, 1972.

17. H.W. Kuhn and A.W. Tucker, "Nonlinear Programming," in Proceedings of the Second International Berkeley Symposium on Mathematical Statistics and Probability, Ed. by Jerzy Newman, University of California Press, 1951, pp. 481-492.

18. M.J. Lighthill and G.C. Whitham, "On Kinetic Waves II: A Theory of Traffic Flow on Long Crowded Roads," Proc. Royal Soc. (London) Ser. A, Vol. 229, pp. 317-345, 1955.

19. R.E. Franklin, "On the Flow-Concentration Relationship for Traffic," Proceedings of the Second International Symposium on the Theory of Traffic Flow, pp. 120-128, 1963.

20. J.G. Wardop, "Experimental Speed/Flow Relations in a Single Lane," Proceedings of the Second International Symposium on the Theory of Traffic Flow, pp. 105-119, 1963.

21. H. Greenberg, "An Analysis of Traffic Flow," Operations Research, Vol. 7, pp. 79-85, 1959.

22. W.S. Leung and C.S. Lau, "A New Study of the Traffic Flow/Concentration Relationship," Traffic Engineering and Control, September 1975, pp. 386-387.

23. F.V. Webster, "Traffic Signal Settings," Road Research Technical Paper No. 39, H.M. Stationery Office, London, 1961.

24. N.H. Gartner, "Area Traffic Control and Network Equilibrium," International Symposium on Traffic Equilibrium Methods, University of Montreal, November 1974; also MIT Operations Research Center Working Paper OR041-75, March 1975.

25. N.H. Gartner, J.D.C. Little, and H. Gabbay, "Optimization of Traffic Signal Settings in Networks by Mixed-Integer Linear Programming," Technical Report No. 91, Operations Research Center, MIT, March 1974.

26. N.H. Gartner, J.D.C. Little, and H. Gabbay, "Optimization of Traffic Signal Settings by Mixed Integer Linear Programming, Part I: The Network Coordination Problem," Transportation Science, Vol. 9, No. 4, pp. 321-343, 1975.

27. N.H. Gartner, J.D.C. Little, and H. Gabbay, "Optimization of Traffic Signal Settings by Mixed Integer Linear Programming, Part II: The Network Synchronization Problem," Transportation Science, Vol. 9, No. 4, pp. 344-363, 1975.

28. P.J. Claffey, "Running Costs of Motor Vehicles as Affected by Road Design and Traffic," National Cooperative Highway Research Program Report 111, Highway Research Board, 1971.

29. M. F. Chargo, L. Evans, R. Herman, and P. Wasielewski, "The Influence of Vehicle Characteristics, Driver Behavior, and Ambient Temperature on Gasoline Consumption in Urban Traffic," General Motors Research Laboratory, Research Publication GMR-1950, January 1976.

30. R. Winfrey and C. Zellner, "Summary and Evaluation of Economic Consequences of Highway Improvements," National Cooperative Highway Research Program Report 122, Highway Research Board, 1971.

31. Highway Capacity Manual Highway Research Board Special Report 87; National Academy of Sciences, National Research Council Publication 1328, 1965.

32. D.R. Drew, Traffic Flow Theory and Control, McGraw-Hill, New York, 1968.

33. M. Florian, "A Traffic Equilibrium Model of Travel by Car and Public Transit Modes," Universite de Montreal, Publication #32 --Centre de Recherche sur les Transports, Publication #216 --Departement d'Informatique, December 1975.

34. D.G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley, Reading, Mass., 1973.

35. W.I. Zangwill, Nonlinear Programming: A Unified Approach, Prentice-Hall, Edgemont Cliffs, N.J., 1969.

36. G.B. Dantzig, Linear Programming and Extensions, Princeton University Press, 1963.

37. T.L. Magnanti, "Optimization for Sparse Systems," MIT Operations Research Center, Technical Report No. 119, 1975.

38. A.A. Assad, "Multicommodity Network Flows --A Survey," MIT Operations Research Center, OR 045-75, 1975.

39. J.E. Defenderfer, "Comparative Analysis of Routing Algorithms for Computer Networks," MIT Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Electronic Systems Laboratory Report ESL-R-756, 1977.

40. M. Frank and I. Frisch, Communication, Transmission and Transportation Networks, Addison-Wesley, Reading, Mass., 1971.

41. D.G. Cantor and M. Gerla, "Optimal Routing in a Packet-Switched Computer Network," IEEE Transactions on Computers, Vol. C-23, No. 10, pp. 1062-1069, October 1974.

42. M.J. Maher and R. Akcelik, "The Re-Distributional Effects of an Area Traffic Control Policy," Traffic Engineering and Control, Vol. 16, No. 9, pp. 383-385, 1975.

43. J.A. Charlesworth, "Relations Between Travel-Time and Traffic Flow for the Links of Road Networks Controlled by Fixed-Time Signals," Transport Operations Research Group-Working Paper No. 13, University of Newcastle Upon Tyne, July 1975.

44. R.E. Allsop and J.A. Charlesworth, "Traffic in a Signal-Controlled Road Network -- An Example of Different Signal Settings Inducing Difference Routing," Second Conference on Traffic Safety and Traffic Engineering, Moscow Automobile and Road Construction Institute, 1976.

45. M. Florian and S. Nguyen, "A Method for Computing Network Equilibrium with Elastic Demands," Transportation Science, Vol. 8, No. 4, pp. 321-332, 1974.

46. M. Florian, S. Nguyen, and J. Ferland, "On the Combined Distribution-Assignment of Traffic," Transportation Science, Vol. 9, No. 1, pp. 43-53, 1975.

47. M.R. Wigan and T.J.G. Bamford, "An Equilibrium Model of Bus and Car Travel over a Road Network," Transport and Road Research Laboratory Report LR 559, Department of the Environment, Crowthorne, Berkshire, England, 1973.

48. J.C. Yu and R.K. Giguere, "Optimizing Transportation Facility Arrangements," Transportation Engineering Journal, November 1973, pp. 725-740.

49. D. Ghahrman, A.K.C. Wong, and T. Au, "Perturbations of a Multimodel Network Model for Urban Transportation Planning," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-4, No. 3, pp.241-248, May 1974.

50. M. Frank and P. Wolfe, "An Algorithm for Quadratic Programming," Naval Research Logistics Quarterly, Vol. 3, pp. 95-110, 1956.

51. D.R. Cox and Walter L. Smith, Queues, Methuen, London, 1961.

52. P.R. Stopher and A.H. Meyburg, Urban Transportation Modelling and Planning, Lexington Books, 1975.

53. J.C. Tanner, "The Delay to Pedestrians Crossing a Road," Biometrika, Vol. 38, pp. 383-392, 1951.

54. J.D.C. Little, "Approximate Expected Delays for Several Maneuvers by a Driver in Poisson Traffic," Operations Research, Vol. 9, No. 1, pp. 39-52, 1961.

55. R. Weiner, L.J. Pignataro, and H. N. Yagoda, "A Discrete Markov Renewal Model of a Gap-Acceptance Entrance Ramp Controller for Expressways," Transportation Research, Vol. 4, pp. 151-161, 1970.

56. A.G. Hawkes, "Queuing at Traffic Intersections," Proceedings of the Second International Symposium on the Theory of Traffic Flow, pp. 190-199, 1963.

57. R.T. Crow, K.H. Young, and T. Cooley, "Alternate Demand Functions for 'Abstract' Transportation Modes," Transportation Research, Vol. 7, pp. 335-354, 1973.

58. E.N. Dodson, "Demand Functions, Behavioral Analysis, and Cost Effectiveness in Urban Transportation," Transportation Science, pp. 139-148, 1975.

59. R.T. Crow and J. Savitt, "The Impact of Petroleum Shortages on Inter-City Travel and Mode Choice," Transportation Research, Vol. 8, pp. 383-397, 1974.

60. R. Wormleighton, "Queues at a Fixed Time Traffic Signal with Periodic Random Input," Canadian Oprs. Res. Soc. J., Vol. 3, pp. 129-141, 1965.

61. F.V. Webster and J.A. Hillier, "Traffic Control Systems," in Traffic Engineering Practice, Ernest Davies, Editor, Spon's Civil Engineering Series, London, 1968.

62. J.E. Baerwald, Traffic Engineering Handbook, Institute of Traffic Engineers, 1965.

63. S.B. Gershwin, "Dynamic Stochastic Control of Freeway Corridor Systems, Volume II: Steady State Optimal Traffic Assignment Using the Accelerated Gradient Projection Method," MIT Electronic Systems Laboratory, Report ESL-R-609, August 1975.

64. M.B. McIlrath, "Steady State Minimization of Traveler Cost for Freeway Corridor Systems," MIT Electronic Systems Laboratory, Technical Memorandum ESL-TM-670, July 1976.

65. D. Looze, "The MIT 'OPTFLOW' Program: Modification of the Cantor-Gerla Algorithm to Freeway Corridor Systems," MIT Electronic Systems Laboratory, Technical Memorandum ESL-TM-717, September 1976.

66. S. M. Thorne, "Relative Effect of Buses and Cars on the Speed of Traffic," Traffic Engineering and Control, June 1965, pp. 140-145.

APPENDIX A: MITROP TRAFFIC-SIGNAL OPTIMIZATION

MITROP is a program to minimize delay to traffic in a signal controlled road network. While the vehicle flows on each link are held fixed, the offsets, the splits of green time, and a common cycle time for the network are simultaneously treated as decision variables. The simultaneous treatment represents a significant departure from previous methods.

The traffic flow pattern is modeled as a periodic platoon. From this is deduced a link performance function, which is approximated by a piece-wise linear convex surface representing the delay incurred by the platoons. A further component of delay arises from a stochastic phenomenon. At flows close to capacity but still, on the average, below it, occasional fluctuations in platoon size lead to temporary overflow queues and consequent delays. As flow approaches capacity, average delay rises, slowly at first and then very rapidly. This is again represented by a piece-wise linear function.

The optimization problem becomes a mixed-integer nonlinear program. The integers enter from the loop constraints, which require that the sum of offsets around any loop of the network must be an integral number of cycle times. The piece-wise linear representations of the nonlinear functions convert the problem to a mixed-integer linear program.

In the following, we formally state the network synchronization and coordination problem as solved by MITROP. The notation used is that of references [25] to [27] and Appendix B and not that of the main body of this report.

Let S_j denote the traffic signal at node j and let (i,j) denote the link connecting nodes i and j. We define:

$r_{ij}(g_{ij})$ = effective red (green) time at S_j facing (i,j)

ϕ_{ij} = internode offset between S_i and S_j

ψ_j = intranode offset at S_j

f_{ij} = average flow on link (i,j)

z_{ij} = average disutility per vehicle on link (i,j) for traveling through S_j

Q_{ij} = average overflow queue at the stop line of S_j.

The objective of the network optimization procedure is to determine signal settings (offsets, splits and cycle time) that minimize the disutility encountered by the vehicles that travel through the signalized intersections. In MITROP this objective is composed of two components. The first component is associated with the mean of the traffic flow process and is represented by the link performance function (LPF). The second component is associated with the random variations about the mean and is represented by the saturation deterrence function (SDF). In general,

$$LPF_{ij} = f_{ij} z_{ij}(\phi_{ij}, r_{ij}, C), \tag{A.1}$$

$$SDF_{ij} = Q_{ij}(r_{ij}, C). \tag{A.2}$$

The constraints of the optimization program represent the street network structure as well as the interrelationships among the variables of the program for each link (i,j) we have,

signal aspect constraint: $r_{ij} + g_{ij} = C$, (A.3)

capacity constraint: $s_{ij} g_{ij} \geq f_{ij} C$, (A.4)

pedestrian clearance time: $r_{ij} \geq (r_{ij})_{min}$, (A.5)

bounds on cycle time: $C_{min} \leq C \leq C_{max}$. (A.6)

In addition to this we have for each loop ℓ in the network a loop offset constraint:

$$\sum_{(i,j)\varepsilon\ell} \phi_{ij} + \sum_{j\varepsilon\ell} \psi_j = n_\ell C , \tag{A.7}$$

where n_ℓ is an integer number associated with loop ℓ.

The traffic signal network optimization problem is then formulated as the following mathematical program:

Determine values of ϕ_{ij}, r_{ij}, C ,

$$\text{to minimize} \sum_{(i,j)} (LPF_{ij} + SDF_{ij}), \qquad (A.8)$$

subject to constraints (A.3) to (A.7) and r_{ij}, $g_{ij} \geq 0$; n_ℓ-integer.

The MITROP processor linearizes piece-wise the nonlinear components in the objective function, so that the program can be solved by mixed-integer linear programming. Commonly branch and bound techniques are used to solve such problems. In the present study, IBM's MPSX system is used for this purpose. Details on MITROP can be found in the references [25] to [27].

APPENDIX B: MITROP DELAY FUNCTIONS FOR SIGNALIZED LINKS

As described in the main body of the report, the overall procedure iterates between setting signals and assigning flows: starting from an initial setting of the signals, assignments are made that determine flows. Given these flows, new signal settings are made. In the more advanced versions of the optimization, MITROP is used to determine the signal settings. Under these circumstances, it is necessary to include the MITROP link delay functions in the traffic assignment part of the iteration, even though some or all of the actual signal settings are held fixed. The purpose of this appendix is to indicate what those delay functions are.

The delay has two components: one deterministic, one stochastic. The first is brought about by the stopping of all or part of a platoon at a signal. The process is modeled deterministically. The second is brought about by the variation in platoon size from one cycle to the next so that sometimes a few vehicles will be stopped by the signal even though a platoon of average size would pass through. This process is modeled stochastically. We take up each case in turn. The notation of this appendix is that of references [25] to [27] and Appendix A and <u>not</u> that of the rest of this report.

B.1 Deterministic Delay

Traffic on a link in the signalized network is modeled as a rectangular platoon: Let

p = platoon length (seconds)

γ = arrival time of the front of the platoon at the signal, measured from the start of green (seconds)

q = traffic flow rate within the platoon (vehicle/sec)

s = saturation flow (vehicle/sec)

$y = q/s$

z = delay/vehicle (seconds).

There is a considerable number of possible delay cases generated by differing relative signs of the parameters involved. These are analyzed in Gartner, Little, and Gabbay [26], [27]. However, for signal settings reasonably close to optimum for the link, a single case captures the major effect. For this case, the delay per vehicle is

$$z = \frac{\gamma^2}{2p(1-y)} \, . \tag{B.1}$$

Calculations with actual platoon data show the general quadratic shape. (See Fig. 4 of [26].) The parameters on the right are related to the signal settings as follows. Let

t = unimpeded travel time on link (seconds)

ϕ = extended internode offset, including travel time (seconds)

g = upstream green time (seconds)

k = platoon dispersion factor for link

f = flow on link (veh/sec).

Then

$$\gamma = t-\phi$$

$$p = kg$$

$$a = f/p \, .$$

For a given link s, k, and t are fixed constants. MITROP determines ϕ and g. Total deterministic delay per unit time for the link as a function of flow and green time is then

$$fz = f(t-\phi)^2 \, \{kg(1-[f/skg])\}^{-1} \, , \tag{B.2}$$

with only f and g varying in the traffic assignment part of the iteration.

B.2 Stochastic Delay

The stochastic component of delay recognizes an important physical phenomenon. At flows that are close to capacity but still, on the average, below it, occasional fluctuations in the size of the platoon can lead to temporary overflow queues that seriously degrade performance. This has a consequence that, as average flow approaches capacity, average delay increases, gradually at first, and then very rapidly. A representation of this effect is needed to prevent green time from approaching its lower bound too closely. It is, of course, possible to put a sizeable constraint on minimum green time, but such an approach misses the main idea of an optimization process. This is because of the tradeoff between capacity loss at short cycles and the inherently large delays of long cycles.

Wormleighton [60] has studied the effect in some detail motivated by experience with the Toronto traffic control system. Following his model, assume

1) Arriving vehicles come in platoons or other periodic function of time with an average arrival rate at time t of q(t).

2) Arrivals are a non-homogeneous Poisson process. Thus the number of vehicles in (t,t + C) is a random variable having a Poisson distribution with mean

$$A_c = \int_t^{t+C} a(u)\, du \ ,$$

for any t.

3) The service provided by the green time is deterministic with rate s(vehicles/sec) up to gs(vehicles/cycle).

If the state of the intersection is examined at the end of green, a bulk service queuing model is defined. Let

$Q(0)$ = number of vehicles in queue at the start of red (end of green) -- overflow queue

A_c = fC = average number of vehicles arriving in a cycle

S = gs = number of vehicles that can be served in a green time.

x = $A_c/S = fC/gs$ = utilization factor (degree of saturation).

Wormleighton finds the generating function of $A(0)$ and calculates $E\{Q(0)\}$ as a function of S and x over $S \in [5,55]$ and $x \in [.2,.975]$. Within these ranges, $E\{Q(0)\}$ varies from essentially zero to 18 vehicles.

MITROP takes the delay attributable to the overflow queue as a "saturation deterence function" in the optimization objective function. To make the function computationally convenient, Wormleighton's table has been approximated by the following expression:

$$Q(0) = \left(\frac{x}{1-x}\right)^2 [1.25\left(\frac{1-x}{S}\right)^2 + 2.25x^2\left(\frac{1-x}{S}\right) + 0.008x^3] , \qquad (B.3)$$

which gives the stochastic delay per unit time. The signal setting process determines C, and S is a link constant. Then equation (B.3) gives the delay as a function of f and g through $S=gs$ and $x=fC/gs$.

The stochastic delay per vehicle is $Q(0)/f$. Then the delay per vehicle is

$$z + Q(0)/f,$$

and the travel time on the link for each vehicle is

$$z + Q(0)/f + t.$$

This quantity, translated into the notation of Section 2, appears as (2.20).

APPENDIX C: HYBRID OPTIMIZATION

In the following, we describe a hybrid optimization problem for the simple network shown in Figure C.1. There is only one origin (O) and one destination (D) and N+1 links carry flow between O and D. Links 1,...,N are not controlled, but the system, by means of a traffic signal, can impose delay W on link N+1. This delay influences the flow in the network by making link N+1 less desirable than otherwise. The user optimizing behavior of drivers limits the traffic flow on this link.

Flows are represented by $\phi_1,\ldots,\phi_{N+1}$ and travel times are represented by $\tau_1,\ldots,\tau_N$ for the first N links and $\tau_{N+1}+W$ for the last link.

Assume that a total flow Φ travels between O and D. Define

$$T = \min (\tau_1,\ldots,\tau_N, \tau_{N+1} + W). \tag{C.1}$$

The positivity and conservation of flow constraints are

$$\phi_i \geq 0 , \quad i = 1,..,N+1, \tag{C.2}$$

$$\sum_{i=1}^{N+1} \phi_i = \Phi . \tag{C.3}$$

For each delay W, the user optimization flow is determined by

$$\left.\begin{array}{l} \phi_i(\tau_i - T) = 0, \; i = 1,\ldots,N \\ \phi_{N+1}(\tau_{N+1} + W - T) = 0 \end{array}\right\} . \tag{C.4}$$

Suppose the operating agency wishes to choose delay W to influence traffic to minimize a cost function $C(\phi_1,\ldots,\phi_{N+1}, W)$, which may be one of those discussed in Section 2.3.3. Then the hybrid optimization problem can be written

$$\underset{\phi_1,\ldots,\phi_{N+1}, W}{\text{minimize}} \quad C(\phi_1,\ldots,\phi_{N+1}, W),$$

subject to (C.2) to (C.4).

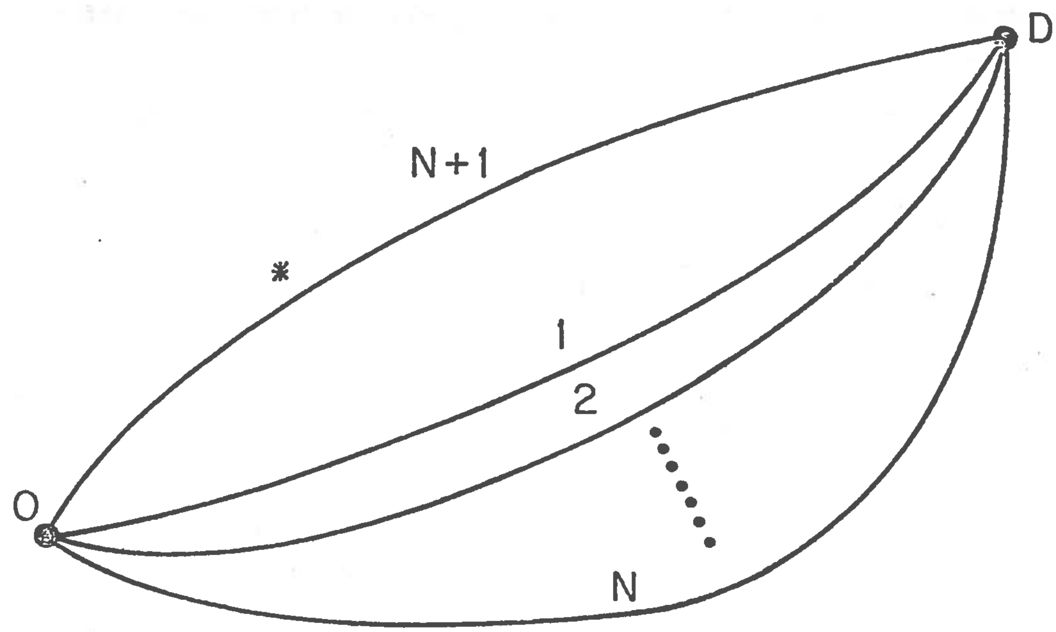


Figure C.1 Network with System Control Only on Link N+1

Further research is needed to formulate completely this problem in general networks to find necessary conditions for a solution, and to devise numerical optimization techniques.

APPENDIX D: NETWORK DETAILS

In Table D.1, details about the links in the examples of section 4.1 and 4.2 appear. Note that the density on a link is given as a function of the flow as follows

$$\rho_i = a_1\phi_i + a_7\phi_i^7 . \qquad \text{(D.1)}$$

The coefficient a_1 is the inverse of the free velocity.

The single class requirement matrix (the table of origin-destination demands) of the examples of section 4.1 appears in Table D.2. Table D.3 contains the multiple class requirements matrices for all other examples in section 4.1 with the exception of the mode split case in 4.5 in which the initial requirements were reduced to 60% of the values shown.

The link parameters for Figure 4.4 are shown in Table D.4. These parameters are used in all examples in and after Section 4.2 except for the accident case. In that case, the capacity of link 39 is reduced to 2000 and $a_7 = 7.71 \times 10^{-22}$.

TABLE D.1 LINK PARAMETERS -- ROADWAY NETWORKS

Link	Capacity	ρ max	Length	Free Velocity	a_7
1	4000	450	0.25	40	0
2	4000	450	0.1	40	2.14×10^{-24}
3	4000	450	0.25	40	0
4	4000	450	0.5	40	0
5	4000	450	0.25	40	2.14×10^{-24}
6	4000	450	0.1	40	2.14×10^{-24}
7	4000	450	0.25	40	0
8	4000	450	0.25	40	0
9	4000	450	0.1	40	2.14×10^{-24}
10	4000	450	0.25	40	0
11	4000	450	0.5	40	0
12	4000	450	0.25	40	0
13	4000	450	0.1	40	2.14×10^{-24}
14	4000	450	0.25	40	2.14×10^{-24}
15	2000	225	0.05	35	0
16	4000	450	0.25	40	0
17	4000	450	0.25	40	0
18	4000	450	0.5	40	0
19	4000	450	0.25	40	0
20	4000	450	0.1	40	2.14×10^{-24}
21	4000	450	0.25	40	2.14×10^{-24}
22	2000	225	0.05	35	0
23	4000	450	0.25	40	0
24	4000	450	0.1	40	2.14×10^{-24}
25	4000	450	0.25	40	0
26	4000	450	0.5	40	0
27	4000	450	0.25	40	0
28	4000	450	0.1	40	2.14×10^{-24}
29	4000	450	0.25	40	0
30	4000	450	0.25	40	2.14×10^{-24}

TABLE D.1 (continued)

Link	Capacity	ρ max	Length	Free Velocity	a_7
31	4000	450	0.1	40	2.14×10^{-24}
32	4000	450	0.25	40	0
33	4000	450	0.25	40	0
34	4000	450	0.25	40	0
35	6000	675	0.25	55	3.34×10^{-25}
36	6000	675	0.5	55	3.34×10^{-25}
37	2000	225	0.05	35	0
38	6000	675	0.5	55	3.34×10^{-25}
39	6000	675	0.25	55	3.34×10^{-25}
40	6000	675	0.5	55	3.34×10^{-25}
41	6000	675	0.25	55	3.35×10^{-25}
42	2000	225	0.05	35	0
43	4000	450	0.1	40	0
44	4000	450	0.1	40	0
45	4000	450	0.1	40	0
46	4000	450	0.1	40	0
47	4000	450	0.1	40	0
48	4000	450	0.1	40	0
49	4000	450	0.1	40	0
50	4000	450	0.1	40	0

TABLE D.2 REQUIREMENTS MATRIX

From \ To	6	12	21	22	23	24	25	26	28
1	560								375
7	600					225			
13		330		150			225		
14		375	187						
15		3810			225			225	
18	4510								
21	375								225
22							225		
23									300
24		375							
25					225				
26	375					150			
27				225					
28					150				

TABLE D.3 MULTIPLE-VEHICLE CLASS REQUIREMENTS

Class 1 -- Single-Passenger Cars

From \ To	6	12	21	22	23	24	25	26	28
1	200								375
7	350					225			
13		200		100			100		
14		215	187					225	
15		2500			225			225	
18	3000								
21	145								125
22							225		
23									300
24		275							
25	70				225				
26	275					150			
27				225					
28					150				

TABLE D.3 (continued)

Class 2 -- Car Pools

From \ To	6	12	22	25	28
1	300				
7	130				
13		100	20	50	
14		100			
15		800			
18	1000				
21	200				100
24		100			
25	20				
26	100				

TABLE D.3 (concluded)

Class 3 -- Buses

From \ To	6	12	22	25
1	20			
7	40			
13		10	10	25
14		20		
15		170		
18	170			
21	10			
25	20			

TABLE D.4

LINK PARAMETERS -- NETWORK WITH SEPARATE FREEWAY LANES

Link	Capacity	ρ max	Length	Free Velocity	a_7
1	4000	450	0.25	40	0
2	4000	450	0.1	40	2.14×10^{-24}
3	4000	450	0.25	40	0
4	4000	450	0.5	40	0
5	4000	450	0.25	40	2.14×10^{-24}
6	4000	450	0.1	40	2.14×10^{-24}
7	4000	450	0.25	40	0
8	4000	450	0.25	40	0
9	4000	450	0.1	40	2.14×10^{-24}
10	4000	450	0.25	40	0
11	4000	450	0.5	40	0
12	4000	450	0.25	40	0
13	4000	450	0.1	40	2.14×10^{-24}
14	4000	450	0.25	40	2.14×10^{-24}
15	2000	225	0.05	35	0
16	4000	450	0.25	40	0
17	4000	450	0.25	40	0
18	4000	450	0.5	40	0
19	4000	450	0.25	40	0
20	4000	450	0.1	40	2.14×10^{-24}
21	4000	450	0.25	40	2.14×10^{-24}
22	2000	225	0.05	35	0
23	4000	450	0.25	40	0
24	4000	450	0.1	40	2.14×10^{-24}
25	4000	450	0.25	40	0
26	4000	450	0.5	40	0
27	4000	450	0.25	40	0
28	4000	450	0.1	40	2.14×10^{-24}
29	4000	450	0.25	40	0
30	4000	450	0.25	40	2.14×10^{-24}

TABLE D.4 (continued)

Link	Capacity	ρ max	Length	Free Velocity	a_7
31	4000	450	0.1	40	2.14×10^{-24}
32	4000	450	0.25	40	0
33	4000	450	0.25	40	0
34	4000	450	0.25	40	0
35	4000	450	0.25	55	3.80×10^{-24}
36	2000	225	0.5	55	2.43×10^{-22}
37	4000	450	0.25	55	3.80×10^{-24}
38	2000	225	0.05	35	0
39	4000	450	0.5	55	3.80×10^{-24}
40	4000	450	0.25	55	3.80×10^{-24}
41	2000	225	0.5	55	2.43×10^{-22}
42	4000	450	0.25	55	3.80×10^{-24}
43	4000	450	0.25	55	3.80×10^{-24}
44	2000	225	0.05	35	0
45	4000	450	0.1	40	0
46	4000	450	0.1	40	0
47	4000	450	0.1	40	0
48	4000	450	0.1	40	0
49	4000	450	0.1	40	0
50	4000	450	0.1	40	0
51	4000	450	0.1	40	0
52	4000	450	0.1	40	0
53	2000	225	0.75	55	2.43×10^{-22}
54	4000	450	0.25	55	3.80×10^{-24}
55	2000	225	0.5	55	2.43×10^{-22}
56	4000	450	0.25	55	3.80×10^{-24}

APPENDIX E: REPORT OF NEW TECHNOLOGY

Although there are no inventions or other patentable items, the report does represent a novel achievement in providing an optimization procedure for steady-state traffic flow in a corridor network, where the assignment, signal-setting and modal-split effects are integrated into a single analysis and combined in a set of computer-assisted traffic engineering programs as described in Section 3.

110 Copies

U. S. DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER
KENDALL SQUARE, CAMBRIDGE, MA. 02142

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

U.S.MAIL

POSTAGE AND FEES PAID
U. S. DEPARTMENT OF TRANSPORTATION
513