STEADY STATE MINIMIZATION OF TRAVELER COST

FOR FREEWAY CORRIDOR SYSTEMS

BY

Michael B. McIllrath

Electronic Systems Laboratory
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

STEADY STATE MINIMIZATION OF TRAVELER COST

FOR FREEWAY CORRIDOR SYSTEMS

by

Michael B McIlrath

June, 1976

ABSTRACT

An approach to the steady state minimization of travel
time on a freeway corridor system by the assignment of traffic
to routes and the control of signal settings is presented. A
model of traffic behavior is developed from which the total
travel time, or cost, on any part of the system can be esti-
mated. It is demonstrated that a general optimization pro-
ceedure for the solution of the problem is inefficient, and
an alternative one is suggested.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

CHAPTER I. INTRODUCTION

I.1 General Introduction--the Freeway Corridor System

A freeway corridor system is a network of roads consisting of one or more limited access highways, the other highways and major streets which parallel them, and the associated connecting roads. A typical example of such a system is a road network connecting the major business center of a metropolitan area with its suburbs. The demand on such systems, in terms of traffic which must be carried, has been continually increasing and will do so for the forseeable future. Further, in many areas it is either impossible or economically unfeasible to make a significant increase in the capacity of a system of roads by building additional roads. Thus, there is a strong need to determine ways of more effectively utilizing existing road systems.

I.2 Survey of Related Publications and Research

Since the publication of a fundamental paper by Wardrop [1] in 1952, considerable attention has been paid to various mathematical and engineering aspects of traffic problems. Among other issues, Wardrop addressed the problem of determining the most preferable assignment of traffic to routes on a system of roads (assuming that there would be a choice of several routes for the traveler to take to his destination) according to two

principles, or decision rules: first, the 'system-optimizing' rule, i.e., the total cost (usually travel time) of all vehicles on all routes of the system is to be minimized; and second, the 'user-optimizing' rule, according to which traffic is assigned to routes so that no traveler will be able to reduce his own cost by taking a different route. It is assumed that the cost incurred by any assignment of traffic to routes can be determined. The existence of user- and system-optimizing solutions, and the relation between them, has been extensively studied by Dafermos and Sparrow [2], [3]. The user-optimizing problem for networks containing freeways was considered by Payne and Thompson [4].

Lighthill and Whitham [5] and later Greenberg [6] and Preparta [7] observed correspondences between traffic flow and the flow of a compressible, continuous fluid and developed a model of freely flowing highway traffic which was experimentally observed to be plausible except at low traffic density. Effects of disturbances, or interruptions of the traffic flow, are also considered in [5] and [7]. Numerous efforts have been made to understand the behavior of traffic at intersections and in other circumstances where queues develop and a fluid model is inapplicable or incomplete, particularly in regard to the probabilistic nature of traffic behavior is such situations. In particular, signalized intersections are considered by Wormleighton [8] and de Smit [9]. Delays at intersections due to vehicles turning left are studied by Hellinger [10]. Grafton and Newell [11], Sako and Zundlevich [12], and Allsop [13] address the problem

of determining traffic signal settings so as to minimize total delay to all vehicles. Queueing at unsignalized intersections is studied by Hawkes [14]. Freeway entrance ramps and the optimal control of queues that develop on them is the subject of Shaw's article [15]. Gaps between venicles and the effects of venicles moving at different speeds in nighway traffic flows are considered by Ashton [16] and Daganzo [17].

Attention has been given to the control of traffic on networks, particularly dynamic control involving real-time data collection and the use of a central controlling computer; for example, by Miller [18], Rosdolsky [19], and Gartner, Little, and Gabbay [20], and also by Nguyen [21].

A comprehensive review of the theory of traffic modeling and control has been made by Gazis [22], including an extensive survey of the literature.

Each of the control studies mentioned above, except for that of Nguyen, emphasizes the control of traffic signals. However, in most urban areas of the United States today, a major portion of the traffic is carried by limited access highways, which are not generally controlled directly by traffic lights. Instead, the amount of traffic already on a freeway determines the maximum rate at which additional traffic may enter it. Furthermore, as previously noted, a typical urban road network will consist of one or more freeways and a number of additional streets and highways. Control of such a network requires some means of determining an optimal routing policy and a means of

assigning traffic to routes.

Ways of developing a control system for urban road net-
works are presently being studied at the M.I.T. Electronic
Systems Laboratory, under contract from the U.S. Department of
Transportation [23]. Information about traffic flow is gathered
by sensors [24] and processed by an estimation/detection scheme
[25]. Dynamic control is facilitated by a feedback system [26].
The assignment of traffic to routes is performed by a static
optimization program [27], which is called periodically and
receives information about the current state of the network
from the estimation/detection system. The static optimization
program also determines traffic signal settings on the network.
One method of performing this static, or steady state, optimiza-
tion by attempting to minimize the total travel time for all
vehicles, is considered by Gershwin in [27].

I.3 Objectives and Summary

The goal of this paper, which is part of ongoing research
at the Electronic Systems Laboratory, is to point out some pos-
sible methods for determining an assignment of traffic to routes
in a freeway corridor system so as to minimize total travel time
for all travelers, as opposed to travel time for vehicles, since
different vehicles may carry different numbers of travelers. It
is assumed that there will generally be more than one route for
a vehicle to take from its origin to its destination. It is
further assumed that there will be some means of directing

vehicles to the proper routes.

A model of traffic flow and associated travel times on a freeway corridor system is presented in Chapter II. This model considers only the steady state, or stationary, distribution of flow on the network; that is, traffic flow is assumed to be essentially constant over some (sufficiently long) interval of time. This is not a highly detailed model, but is a sufficient approximation for the purpose of simulating some of the most important types of vehicle behavior on a freeway corridor system. In particular, the effects of traffic density on average velocity of vehicles, and the delays incurred on freeway entrance ramps and at traffic signals, are considered. A cost function, representing the total travel time for all travelers on the network, is developed.

Two optimization techniques for minimizing the cost by adjusting the traffic flows and signals, are presented in Chapter III. The first of these is a general optimization algorithm, called accelerated gradient projection [28], which does not take any special properties of the traffic problem into account. The second algorithm, a decomposition method, was developed for use on communications networks [29], where problems similar to those encountered in traffic systems arise. It is shown how both methods can be adapted for use on vehicular traffic problems, and their respective computational efficiencies are compared.

Some preliminary numerical results achieved by the use of a computer are presented in Chapter IV. The applicability of

both the model and the optimization methods to real-time solu-
tion of the problem on a freeway corridor system are discussed.
in Chapter V.

CHAPTER II. STEADY STATE TRAFFIC MODEL

II.1 Introduction

In this chapter, a model of traffic behavior on a road
network will be presented. The model will not describe all
aspects of traffic behavior in detail, but will represent some
of the phenomena that are most important in the freeway corridor
system problem.

Elementary graph theory will be used to give a precise
definition of a road network. The flow of moving traffic will
be considered by analogy to the flow of a continuous compress-
ible fluid, and the effects of traffic signals and freeway
entrance ramps will modeled as single server queues. In all
cases, it will be a steady state model that is used. Finally,
a cost function, representing the travel time in accordance
with the model presented, will be developed.

II.2 Networks and System Parameters

A road network is defined to be a _directed graph_; that is,
a finite set η of nodes, and a set $\mathcal{L}$, whose members are ordered
pairs of nodes, called _links_. Nodes will be numbered 1, 2, 3,
etc. If n_i, n_j are nodes, then (n_i, n_j) is the link _connecting_
$\underline{n_i}$ $\underline{to}$ $\underline{n_j}$, and is distinct from (n_j, n_i), which connects n_j to n_i.
In general, there will not be a link connecting every pair of
nodes.

(Links may also be numbered 1, 2, 3, etc., and referred to by number as long as it is clear which ordered pair of nodes is implied.) The link (n_i, n_j) is understood to represent a roadway on which traffic may flow from n_i to n_j. Some nodes may have the flow entering (leaving) them specified, such nodes are called destination (origin) nodes. A <u>chain</u> is an ordered set of links which traffic may take from one node to another. Figure 1 shows a sample network, of four nodes and five links.
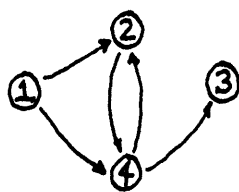
Nodes	Links
1	(1,2)
2	(1,4)
3	(2,4)
4	(4,2)
	(4,3)

Figure 1. Sample Network

On the network a <u>distance function</u>, or <u>metric</u>, $\mathcal{L}$, is defined, which assigns a positive number to every link. That is, if (n_j, n_k) is a link, then $\mathcal{L}(n_j, n_k)$ is the distance from n_j to n_k along that link; or the length of the link. If (n_j, n_k) is referred to as link i, then $\mathcal{L}_i$ is the length of link i.

The following quantities will be of concern in the model presented here: the traffic flow Φ, the traffic density ρ, the average velocity $\underline{v}$ of a vehicle in a traffic stream, the distance that a vehicle travels, and the travel time τ. Table 1 summarizes the symbols used and their units.

Symbol	Meaning	Units
Φ_i	traffic flow on link i	vehicles/hour
ρ_i	traffic density on link i	vehicles/mile
v_i	average velocity of vehicles on link i	miles/hour
ℓ_i	length of link i	miles
τ_i	average travel time for vehicles on link i	hours

Table 1. Symbols and Units

II.3 System Equations and Inequalities

There are two basic principles of flow on a network:
1) all flows are nonnegative, and 2) flow is conserved. That is,
at all nodes that are neither origins (sources) nor destinations
(sinks), the flow entering is equal to the flow leaving the node,
and the total flow which enters at origins is equal to the total
flow leaving the network at destinations.

In addition it is required that the flow on each link be
less than a fixed maximum, called the capacity of the link.
(Different links may have different capacities.) The capacity
of link i is denoted by $\Phi_{i\max}$.

Finally, the demands, or traffic entering and leaving the
network at origins and destinations, are specified.

This leads to the following set of equations and inequalities, called <u>constraints</u>:

$$\sum_i \Phi_i - \sum_j \Phi_j = 0 \qquad (1)$$

where the first sum is taken over all links i that enter a given node which is not an origin or destination, and the second sum is over all links j which leave the node.

$$\Phi_i = D_i \qquad (2)$$

where D_i is the specified demand on link i; i.e., the traffic which enters or leaves the network via link i.

$$\Phi_i \geq 0 \qquad (3)$$

for all links i.

$$\Phi_i \leq \Phi_{i\,max} \qquad (4)$$

for all links i.

II.4 <u>Traffic Flow, Density, and Velocity--A Fluid Model</u>

Several authors have applied fluid mechanics principles to the problem of moving traffic; in particular, Lighthill and Whitham [5] and Greenberg [6]. The following approach, due to Greenberg, assumes that traffic behaves as if it were a continuous fluid. Data taken by Greenberg indicates that such a model is a good representation of the macroscopic behavior of traffic above a minimum density.

The equation of motion of a one-dimensional fluid of density ρ and moving with velocity v is assumed:

$$\frac{dv}{dt} = -\frac{c^2}{\rho} \frac{\partial \rho}{\partial x} \qquad (5)$$

where <u>x</u> is the distance variable in the direction of motion, and

where $\underline{c}$ is a constant which depends on the fluid. This states that the acceleration of an average driver in the traffic stream is proportional to the concentration (density) gradient $\partial\rho/\partial x$ and inversely proportional to the traffic density. If the velocity is a function of location and time, then the equation of motion is

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + \frac{c^2}{\rho} \frac{\partial \rho}{\partial x} = 0 \tag{6}$$

By continuity, or conservation of flow,

$$\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = 0. \tag{7}$$

It is further assumed that

$$\phi = \rho v \tag{8}$$

and that the average velocity at a point is a function only of the density at the point, giving

$$\frac{\partial v}{\partial x} = \frac{dv}{d\rho} \frac{\partial \rho}{\partial x} \quad , \qquad \frac{\partial v}{\partial t} = \frac{dv}{d\rho} \frac{\partial \rho}{\partial t} \tag{9}$$

Solving the equations of motion and continuity using (9) gives

$$\frac{dv}{d\rho} = - \frac{c}{\rho} \tag{10}$$

Integrating, and setting v = 0 when $\rho = \rho_{max}$, the maximum density, gives

$$v = c \, \log\left(\rho_{max}/\rho\right) \tag{11}$$

and

$$\phi = c \rho \, \log\left(\rho_{max}/\rho\right) \tag{12}$$

- 15 -

These are Greenberg's equations for velocity and flow as a function of density. Equation (12) leads to a graph of the form shown in Figure 2.
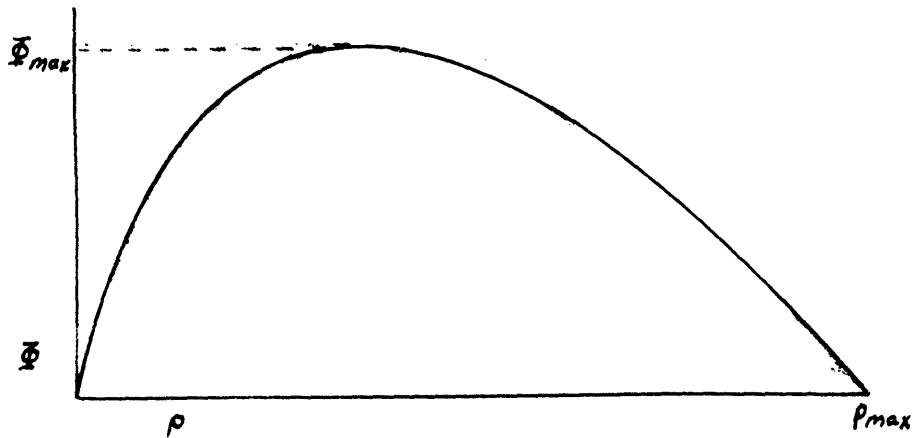
Figure 2. Flow vs. Density for One-dimensional Continuous Fluid

Differentiating (12) yields

$$\frac{\partial \phi}{\partial \rho} = c \left(log \left(\rho_{max}/\rho \right) - 1 \right) \tag{13}$$

Solving for ϕ_{max}, the maximum value for ϕ, by setting (13) equal to zero gives

$$\rho = \rho_{max}/e \tag{14}$$

That is, ϕ is maximum when ρ is given by (14).

Equation (11) clearly does not correspond to real traffic flow at low density, since v increases without bound as the density ρ goes to zero. We wish to modify this so that v is always less than some maximum velocity, v_{max}. Then (11) is valid for

$$c \log \left(\rho_{max}/\rho \right) < v_{max} \tag{15}$$

or

$$\rho > \rho_{max} \, e^{-v_{max}/c} \tag{16}$$

Assuming that $v = v_{max}$ for all values of ρ below that given by (16), we can replace (11) by a linear function for small ρ and obtain a graph of the form shown in Figure 3.
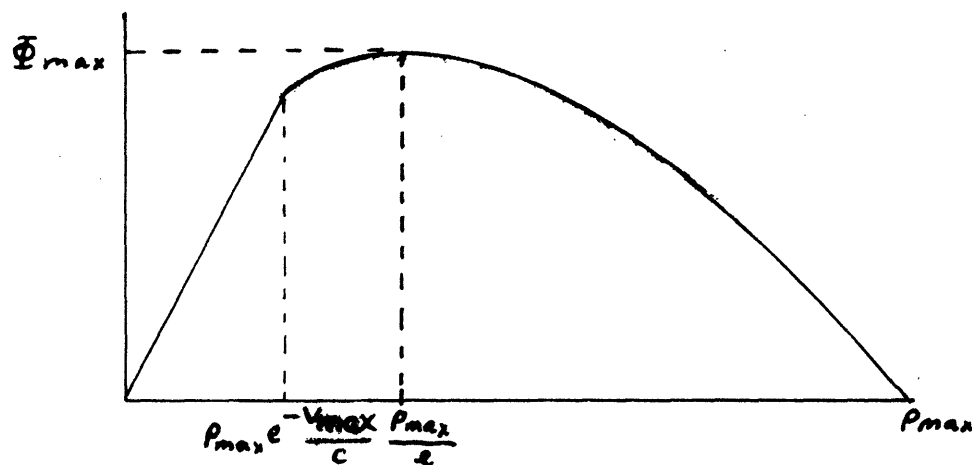
Figure 3.   Flow vs. Density, with Velocity Bounded

In real traffic problems, it is flow, not density, which must be considered as the independent variable from the standpoint of the assignment of traffic to routes, since the demand will be in terms of moving some number of vehicles per hour. Thus, an expression for ρ in terms of Φ is needed.

Consideration of a graph of the form shown in Figure 3 indicates that, for every flow $\Phi < \Phi_{max}$, there are two corresponding values for the density ρ . Clearly, the smaller value for ρ is more desirable, as it corresponds to a higher average

velocity and thus reduced travel time. The larger value for ρ indicates congestion. In fact, consideration of the quantity $\partial \Phi / \partial \rho$ (the <u>wave</u> <u>velocity</u>, or rate at which a <u>disturbance</u> is <u>propagated</u>), indicates that the right hand side of the graph is <u>unstable</u> in the sense that an increase in the density at some point will propagate backward, since $\partial \Phi / \partial \rho$ is negative, resulting in further congestion; whereas at the lower density (left hand side of the graph) an increase in the density propagates forward, so that if the increase is not too large, the density returns to its former level, and hence the situation is <u>stable</u>. Wave effects in traffic flow are extensively considered by Lighthill and Whitham [5], and also by Preparta [7].

Assuming that the velocity at maximum flow given by (11) is less than v_{max}, we have, at Φ_{max}, by substituting in (12),

$$\Phi_{max} = c \left(\frac{\rho_{max}}{e} \right) \log \left(\frac{\rho_{max}}{\rho_{max}/e} \right) \tag{17}$$

or

$$c = e \left(\frac{\Phi_{max}}{\rho_{max}} \right) \tag{18}$$

To be compatible with work done by other members of the M.I.T. Electronic Systems Laboratory freeway corridor research team, Φ_{max} was chosen to be 2000 vehicles/hour/lane, and ρ_{max} to be 225 vehicles/mile/lane [27]. This gives for the fluid constant

$$c \doteq 24.16 \text{ miles/hour} \tag{19}$$

We also chose $v_{max} = 55$ miles/hour, the present maximum legal speed.

Using (16), the minimum value of ρ such that $v = v_{max}$ is

$$\rho_{max} e^{-v_{max}/c} \doteq 23.10 \text{ vehicles/mile} \qquad (20)$$

Thus a graph of velocity vs. density has the form shown in Figure 4.
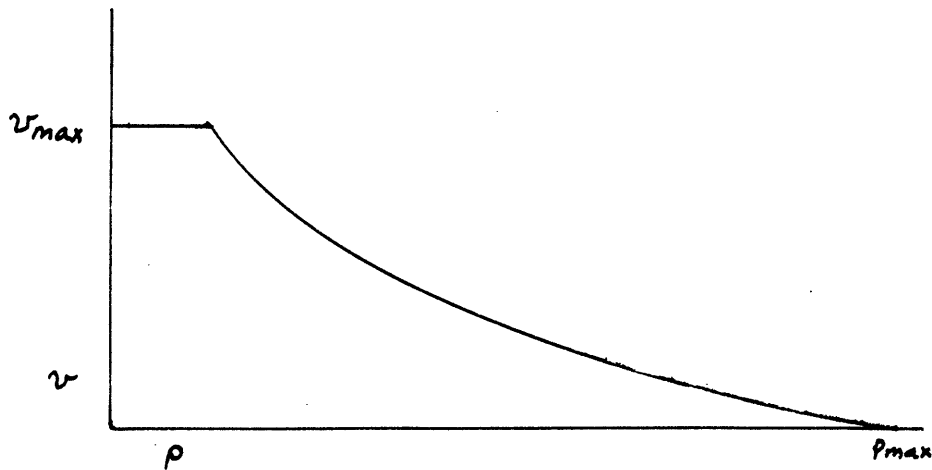
Figure 4. Velocity vs. Density for Traffic Flow

The travel time τ_i for an average vehicle on link i, where traffic is flowing as described above, is

$$\tau_i = l_i / v_i = l_i \left(\frac{\rho_i}{\Phi_i} \right) \qquad (21)$$

where l_i is the length of link i.

Thus we need an expression for ρ in terms of Φ. The requirement that $\Phi \leq \Phi_{max}$, as given in section II.3, together with (12), gives Φ as a (single-valued) function of ρ. Given Φ, one method of determining ρ would be to invert (12) numerically. However, this would be time-consuming and inefficient since it must be done many times in the process of solving a problem of finding an optimal traffic assignment. Hence a polynomial

approximation $P(\bar{\phi})$ is used to obtain ρ in terms of $\bar{\phi}$. $P(\bar{\phi})$ has the form

$$P(\bar{\phi}) = \bar{\phi}/55 \quad + \quad K \bar{\phi}^7 \tag{22}$$

This polynomial is an approximation to the inverse of (12), con-strained by $v \leq v_{max}$. Note that

$$\lim_{\bar{\phi} \to 0} v = \lim_{\bar{\phi} \to 0} \frac{\bar{\phi}}{P(\bar{\phi})} = v_{max} = 55 \tag{23}$$

The travel time function given by (21) is taken to represent the average time spent by a vehicle on a link where traffic is always flowing (there are no additional delays due to stopped traffic.) For this model, such links will generally rep-resent sections of a freeway. Links on which traffic may stop and vehicles be subject to queueing delays are considered in the next section.

II.5 Queueing at Merges and Intersections

There are two general classes of delays which occur in real traffic situations:

(1) delays which are a function only of distance to be traveled and the amount of traffic on the road, where traffic is actually moving at all times.

(2) delays which occur in situations where a vehicle must stop and wait for some event before proceeding (e.g., for a traffic light to change.)

As indicated in section II.4, the travel time when only

type (1) delays are encountered is given by

$$\tau = \ell/_{\nu} = \ell \left(\rho/_{\phi} \right) \tag{24}$$

In general, subjecting vehicles to delays of type (2) will
result in the formation of queues on the roadway. Thus, delays
of type (2) will be referred to as queueing delays, and the time
a vehicle spends in such a situation as the queueing time.

Queueing due to congestion is eliminated in this model by
the requirement that all flows be less than capacity, as noted
in section II.4. Two types of queueing will be considered: de-
lays at traffic lights, or signalized intersections, and delays
on freeway entrance ramps, also known as on-ramps or merges,
where vehicles may have to wait in order to enter the freeway.
In both cases, it is assumed that the amount of space occupied
by the queue is negligible compared to the length of the link on
which it occurs.

A simple queueing model, the single server queue, with a
Poisson arrival process and an exponential server (also known as
M/M/1 queue [30]), will be used to represent both the signalized
intersection and the on-ramp. It is assumed that customers
arrive in accordance with a Poisson process and wish to be
served. If the server is already busy serving someone else, the
arriving customer must go to the end of the line, or queue.
Service is first-come, first-served. For a Poisson process,
the density function for interarrival times is given by

$$f(t) = \lambda e^{-\lambda t} \tag{25}$$

where λ is the parameter of the process. It is further assumed that the service time also has a negative exponential distribution, where the density function for interdeparture times (given that there is at least one customer to be served) is

$$f(t) = \gamma e^{-\gamma t} \tag{26}$$

That is, in time $[0,t]$ the expected number of customers to arrive is λt, and each has an expected service time $\bar{S} = \frac{1}{\gamma}$. Clearly λ must be less than γ (customers can be served at least as fast as they are expected to arrive) if the queue is to remain finite. Since it is a steady state model under consideration, we seek the steady state, or _stationary_, solution to the queueing problem; i.e., the behavior of the queue after sufficiently long time. (If $\lambda < \gamma$, such a solution exists [30].)

Let $\underline{W}_q$ be the expected time a customer spends waiting in the queue for service, and let $\underline{W}$ be the expected total time he spends in the queue system; i.e., his waiting time plus his service time. Then [31]

$$W_q = \frac{\lambda}{\gamma(\gamma - \lambda)} \tag{27}$$

and

$$W = \frac{1}{\gamma - \lambda} \tag{28}$$

Note that $W = W_q + \bar{S} = W_q + \frac{1}{\gamma}$. We may compare the expected service and waiting times as follows: if the wait time is greater than the service time, then $\frac{\lambda}{\gamma(\gamma-\lambda)} > \frac{1}{\gamma}$, or $\lambda > \gamma - \lambda$, implying

$$\lambda > \gamma/2 \tag{29}$$

-22-

In [27], Gershwin uses (27) rather than (28) to represent the total time lost due to queueing delays at merges and intersections. The significance of the difference between the two will depend on the arrival and service parameters and the proportion of the total travel time which is spent in the queues. In any event, it must be realized that the model is only a rough approximation, and that actual data must eventually decide the best model to be used for this particular problem.

By (28), the total queueing time becomes infinite as the arrival rate λ approaches γ, the service rate. Thus γ may be identified as the effective capacity of a link with queueing delay; i.e., the maximum possible flow on the link when the queueing delay is considered. The effective capacity is always less than or equal to ϕ_{max}, the maximum flow without queueing, and is not necessarily a fixed parameter but may depend on the flow on other links.

In the models presented here of traffic behavior at intersections and on entrance ramps, vehicles will be assumed to arrive in accordance with a Poisson process with $\lambda = \phi$, the flow on the link. For a traffic signal, it is assumed that the service rate is proportional to the green split, or fraction of time that the light is green, and that the service rate approaches ϕ_{max}, the maximum flow, as the green split approaches one. A signalized intersection is shown in Figure 5. We let g_i be the green split for link i, and require that

$$g_i + g_j = 1 \tag{30}$$
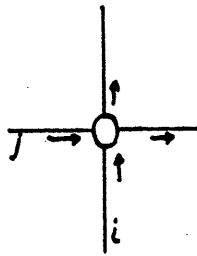
Figure 5. Signalized Intersection

We also assume that the service rate approaches zero as the green split g_i approaches zero. For simplicity, the linear relationship given below is used:

$$\gamma_i = \Phi_{imax} \cdot g_i \tag{31}$$

Clearly, this is not a complete representation of a traffic light, since a signalized intersection is not really a single server queue. Also no account is made of _cycle time_ (the time for the traffic light to complete one cycle). Further, if there is a sequence of traffic lights on a street that are sufficiently close together, the traffic stream may be broken up so that the Poisson process is no longer applicable. More detailed models have been proposed (see [8] thru [14]); however, they result in much more complicated expressions for the delay and are not suitable for use in this problem.

The additional time spent by a vehicle on link i due to the presence of a traffic signal is thus

$$\tau = \frac{1}{\Phi_{imax} \cdot g_i - \Phi_i} \tag{32}$$

and the total travel time, or time when the vehicle is moving
plus the queueing time, is

$$\tau_i \;=\; \ell_i\left(\frac{\rho_i}{\Phi_i}\right) \;+\; \frac{1}{\Phi_{imax}\cdot g_i - \Phi_i} \tag{33}$$

Note that this is an estimate of the average time for all
vehicles, including those that arrive when the light is green
as well as those that have to stop.

On freeway entrance ramps, the applicability of the
single server queue is more intuitive. Again, vehicles arrive
with rate $\lambda = \Phi$. The service rate now depends essentially on
the volume of traffic on the freeway. Specifically, it depends
most strongly on the traffic on the link fed by the on-ramp. In
Figure 6, links k and j represent freeway sections, while link
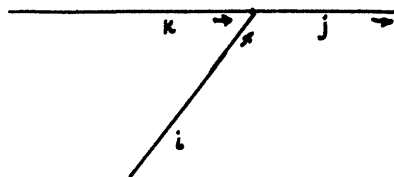i is the entrance ramp.

Figure 6. Freeway and Entrance Ramp

Even if there is no traffic on link k, traffic on link j which
has already arrived via link i may be sufficiently heavy to
cause delays to following traffic on link i.

If there is no traffic on link j $(\Phi_j=0)$ we set the ser-
vice rate, or effective capacity γ_i of link i to Φ_{imax}, the

maximum flow on link i. If, however, $\Phi_j = \Phi_{jmax}$, no more traffic from link i can enter the freeway. In this case, the flow capacity of link i is effectively zero. A linear equation which covers both of these cases is

$$\gamma_i = \Phi_{imax}\left(1 - \Phi_j\big/\Phi_{jmax}\right) \tag{34}$$

Then the queueing time on an entrance ramp is

$$\tau = \frac{1}{\Phi_{imax}\left(1 - \Phi_j\big/\Phi_{jmax}\right) - \Phi_i} \tag{35}$$

The total travel time on a link which includes an entrance ramp is therfore

$$\tau_i = \frac{1}{\Phi_{imax}\left(1 - \Phi_j\big/\Phi_{jmax}\right) - \Phi_i} + \frac{\ell_i \rho_i}{\Phi_i} \tag{36}$$

This model is most applicable if there is sufficiently heavy traffic to actually generate a queue. However, it is also during peak traffic periods that an optimization scheme of the type described in Chapter I would most likely be required. A more detailed model of the entrance ramp might consider, in addition, the problem of the waiting time for a gap of acceptable length in the freeway traffic for a vehicle on the ramp to enter the freeway. Gaps in road traffic are considered by Ashton [16], among others. Unfortunately, such models lead to much more complicated delay functions than could be handled adequately here.

II.6 Cost Function

The <u>vehicle cost function</u> C_i^V for link i is defined by

$$C_i^V = \phi_i \tau_i \tag{37}$$

that is, the cost on link i is the travel time per vehicle times the total flow on the link in vehicles per hour. Cost has the units of vehicle-hours per hour, or simply vehicles.

For a freeway link,

$$C_i^V = \phi_i \left(\frac{\ell_i \rho_i}{\phi_i} \right) \tag{38}$$

or

$$C_i^V = \ell_i \rho_i \tag{39}$$

On any link on which queueing delays occur, the cost due to queueing is

$$C_i^V \text{ (queueing)} = \phi_i \left(\frac{1}{\gamma_i - \phi_i} \right) \tag{40}$$

with γ_i given by (31) for signalized links and by (34) for on-ramps. Thus the total cost on a link with queueing delay is

$$C_i^V = \ell_i \rho_i + \phi_i \left(\frac{1}{\gamma_i - \phi_i} \right) \tag{41}$$

The total <u>system vehicle cost</u> C^V is

$$C^V = \sum_i C_i^V \tag{42}$$

where the sum is taken over all links i of the network. This function represents the total cost in vehicles on the network.

The equations (37)-(42) represent cost in terms of vehicles. Consider, however, a real transportation problem, in which the goal is to move people, not vehicles. On any actual urban road network, there will be single-passenger cars, cars with several passengers, and buses, probably carrying between ten and forty passengers each.

The passenger cost function, to be defined below, will be in terms of total passenger-hours per hour on the network, rather than vehicle -hours per hour. However, the delay function will still be determined in terms of vehicles. A vehicle type will be determined by the number of passengers it is assumed to be carrying (e.g., one-passenger car, three-passenger car, twenty-passenger bus, etc.) Types of vehicles will be referred to by number (1, 2, 3, etc.) No distinction will be made in terms of contribution or sensitivity to delay according to vehicle type. That is, we assume that 1) all vehicles on a given link are subject to the same delays and hence the same average travel time on that link, and 2) the travel time does not depend on the vehicle mix on a given link, but only on the total flow.

In order to consider cost in terms of passengers, we require new parameters to represent passenger flow, the number of passengers carried by vehicles of different types, and the flow of vehicles of different types. The symbols for these new parameters are listed in Table 2.

Symbol	Meaning	Units
P_i	passenger flow on link i	passengers/hour
$\varphi_i{}^j$	flow of vehicles of type j on link i	vehicles/hour
w^j	passengers carried by vehicles of type j	passengers/vehicle

Table 2. Additional Symbols and Units

There are two basic relations:

$$\Phi_i = \sum_j \varphi_i{}^j \qquad (43)$$

and

$$P_i = \sum_j \varphi_i{}^j \, w^j \qquad (44)$$

That is, the total number of vehicles on link i is the sum of vehicles of all types on the link, and the total number of passengers on link i is the sum of all passengers carried by vehicles of different types.

The passenger cost function is now defined by

$$c_i^P = P_i \tau_i \qquad (45)$$

Thus for a freeway link

$$c_i^P = P_i \ell_i \left(\frac{P_i}{\Phi_i} \right) \qquad (46)$$

or

$$c_i^p = \ell_i \rho_i \quad \frac{\sum\limits_j \varphi_i^j w^j}{\sum\limits_j \varphi_i^j} \tag{47}$$

For a link with queueing delay, the cost is the cost of un-interrupted travel plus the queueing cost. The queueing cost is now

$$c_i^p \text{ (queueing)} = \frac{P_i}{\gamma_i - \phi_i} \tag{48}$$

or

$$c_i^p \text{ (queueing)} = \frac{\sum\limits_j \varphi_i^j w^j}{\gamma_i - \sum\limits_j \varphi_i^j} \tag{49}$$

The total passenger cost for the system is

$$c^p \text{ (total)} = \sum\limits_i P_i \tau_i \tag{50}$$

or

$$c^p \text{ (total)} = \sum\limits_i (\sum\limits_j \varphi_i^j w^j) \tau_i \tag{51}$$

The total cost now has the units of passenger-hours per hour, or simply passengers. This is the objective function to be minimized by adjusting the flows φ_i^j and the green splits g_i on the various links of the network.

We would expect the effect of considering passenger cost, as opposed to vehicle cost, to give preference to carpools and buses in the sense that the optimum solution would be one which put the most passengers on the lowest cost links. The w^j's may be thought of as 'priority ratings' which indicate how much preference a vehicle should be given: the more passengers, the higher the priority assigned to that vehicle type.

CHAPTER III.   OPTIMIZATION ALGORITHMS

III.1   Introduction

In general, a constrained optimization problem has the form: minimize (maximize) the scalar function F of k scalar variables x_1, . . ., x_k, subject to the condition that a given set of equations and inequalities involving the x_i, called constraints, must be satisfied.   The set of all $X=(x_1, . . ., x_k)$ that satisfy the constraints is called the feasible region.

Two methods for performing constrained optimization are presented, both of which can be used for traffic problems.   The first of these is a general method called accelerated gradient projection.   The second is a decomposition method which has been used on communications networks to solve a similar routing problem.

III.2   Accelerated Gradient Projection

A general optimization algorithm, known as accelerated gradient projection, was adapted for this problem.   The method was developed by Kelley and Speyer [28] for use in constrained optimization problems, using methods of Davidon [32] and Fletcher and Powell [33], which they invented for use in unconstrained optimization.   The accelerated technique is an improvement

to the usual gradient projection method for solving constrained optimization problems [34] , by using information about the first derivative of the objective function F to estimate its second derivative for the purpose of choosing a search direction.

The algorithm proceeds in two phases: first, given a guess for the solution vector X, a point in R^k, the algorithm checks to see if any constraints are violated. If any are, X is moved back inside the feasible region. The second phase picks a search direction, and performs a one-dimensional search for the minimum in that direction. The search direction chosen is the projection of the negative of the gradient of F(X) (the downhill direction, since the minimum is desired) on the feasible region.

Some stopping criterion is then checked, and if it is not satisfied, the algorithm returns to phase one and proceeds to take another step.

For this problem, the objective function F is the passenger cost function defined in section II.6. The vector of variables over which F is to be minimized are the vehicle flows $\varphi_i{}^j$ and the green splits g_i. As given in section II.3, there are equality constraints

$$G_1(X) = 0$$
$$\cdot$$
$$\cdot$$
$$\cdot$$
$$G_h(X) = 0 \tag{52}$$

where $\underline{h}$ is the number of nodes times the number of vehicle types.

That is, flow of each type is conserved at each node. There
are also constraints

$$g_i + g_j = 1 \tag{53}$$

where g_i and g_j are complementary traffic signals (section II.5).
The inequality constraints are

$$x_1 \geq 0$$
$$\vdots$$
$$x_k \geq 0 \tag{54}$$

That is, all flows and green splits are positive. Additionally,
it is required that all flows be less than the fixed maximum
for the link to which they are assigned (section II.3). Thus
we have

$$x_1 \leq \Phi_{1max}$$
$$\vdots$$
$$x_h \leq \Phi_{nmax} \tag{55}$$

where n is the number of links.

The steps of the algorithm are:

Step 0.   Initialize X to X_0 (the initial guess). Initialize
H to I, the identity matrix. H will later be used
to approximate the second derivative of F.

Constraint Restoration:

Step 1.   Evaluate all constraints at X. If no constraints are
violated, go to Step 7.

-33-

Step 2. Form a vector function $G(X)$, where G consists of the functions G_i that are equality constraints (52,53) plus any of the inequality constraints (54,55) that were found to have been violated by Step 1. The matrix G will, in general, be different at each iteration since different inequality constraints may be violated.

Step 3. Calculate the matrix $\frac{\partial G}{\partial X}$, the Jacobian of G, denoted by G_X.

Step 4. Calculate $\frac{\partial F}{\partial X}$, the gradient of F, denoted by ∇F.

Step 5. $X \leftarrow X - HG_X^T (G_X HG_X^T)^{-1}G - H(\nabla F)^T + HG_X^T(G_X HG_X^T)^{-1}G_X H(\nabla F)^T$

In this problem, the constraints are linear, so the feasible region is a convex polyhedron [35], and G_X is a constant matrix.

The matrix H, which will be modified later, is a symmetric positive definite matrix which defines a metric on R^k. The usual step taken here in accelerated gradient algorithms [28] is to move toward the feasible region along the shortest path according to the metric H. Step 5 is a modification by Gershwin [27] to take into account more information regarding the objective function F.

Step 6. Go to Step 1.

Minimization:

Step 7. Calculate F, G, and ∇F at X.

Step 8. Find the direction $\vec{D}$ of the negative F gradient, projected on the feasible region, given by

$$\vec{D} = -H(\nabla F)^T - HG_x^T(G_xHG_x^T)^{-1}G_xH(\nabla F)^T$$

Step 9. Approximate the Lagrange multiplier vector by

$$\lambda = -(G_xHG_x^T)^{-1}G_xH(\ F)^T$$

Step 10. Perform a one-dimensional search to minimize

$$F^*(\alpha) = F(X + \alpha D) + \lambda^T G(X + \alpha D)$$

On the plane of active constraints (equality constraints

plus inequality constraints imposed in Step 1), where

$G = 0$, $F^* = F$.

Fletcher and Powell [33] recommend cubic fit as a means

of performing the minimization in Step 10. In our

problem, we use an algorithm of Johnson and Myers [36]

which combines cubic fit with golden section as a means

of minimizing F^* as a function of α. Let $\alpha*$ be the minimum.

Step 11. $X \leftarrow X + \Delta X$,

where

$$\Delta X = \alpha^* D$$

Step 12. $H \leftarrow H + A - B$

where

$$A = \frac{\Delta X(\Delta X)^T}{(\Delta X)^T \Delta(\nabla F^*)}$$

and

$$B = \frac{H\Delta(\nabla F^*)\Delta(\nabla F^*)^T H}{\Delta(\nabla F^*)^T H (\nabla F^*)}$$

where $\Delta(\nabla F^*)$ is the change in ∇F^*, the gradient of F^*,

given by the change ΔX in X.

Step 12 is Fletcher and Powell's form of Davidon's [32]

update for H. If $\bar{X}$ is a vector such that $G_x\bar{X} = 0$, then

successive iterations should move $\bar{X}^T H^{-1} \bar{X}$ close to $\bar{X}^T F_{xx} \bar{X}$, where F_{xx} is the second derivative (Hessian) matrix $\dfrac{\partial^2 F}{\partial X^2}$.

Step 13. Test stopping rule. If satisfied, quit; otherwise go to Step 1, with X as the current guess. Fletcher and Powell recommend a number of possible stopping criteria for use in computer applications, one of which may be interpreted for this problem as to perform as many iterations as there are degrees of freedom in the system (variables minus equality constraints). Other possible rules we considered in the computational problem include a minimum change ΔX, and the ratio of the derivative along the search direction at X to the derivative at X_0, the initial guess.

For purposes of simplicity in calculating the first derivatives of the cost function (the gradient of F), the total flows ϕ_i were considered as independent variables, as well as the $\varphi_i{}^j$, and additional equality constraints of the form of (43), section II.6, were imposed.

III.3 Decomposition Method

Considerable improvement can be achieved over a general
method for an optimization problem if special characteristics of
the problem are taken into account. A decomposition method was
used by Cantor and Gerla [29] for optimizing the assignment of
routes for the transmission of information on the ARPA computer
network.

The largest amount of time in the execution of the accel-
erated gradient projection (AGP) algorithm was apparently spent
in inversion of the matrix $G_x H G_x^T$ in steps 5, 8, and 9, as de-
scribed in section III.2. Greater efficiency is achieved in the
Cantor-Gerla algorithm by replacing a large part of the non-lin-
ear optimization problem with a series of _linear_ programming
problems. The linear programming steps handle all of the system
constraints (52-55). When the master (nonlinear) routine is
called, optimization will be performed over only a small subset
of the feasible region at a time. It is the task of the linear
program to control this subregion by generating and removing
corner points; i.e., corners, or boundary intersections, of the
subregion. The corner points are generated in such a way that
the system constraints are satisfied, and also so that the sub-
region is moved toward, and eventually includes, the global
minimum. The nonlinear program finds a minimum in the subregion
as a convex combination of the corner points. This can be
performed by the gradient projection algorithm; however, it

will generally have to handle considerably fewer constraints than
if it is used directly, as described in section III.2. The
computation time for inverting an mxm matrix increases at least
as fast as m^2, so, since the size of the matrix to be inverted
depends on the number of constraints, a considerable savings of
time can be achieved. The linear program uses information gen-
erated by the master optimization step to generate new corner
points. When no new corner points can be generated, the sub-
region includes the global minimum (assuming that the objective
function is convex) which can then be found. This method is
called _decomposition_ because it decomposes the main problem into
a series of smaller ones.

Any feasible solution $X = (x_1, \ldots, x_n)$ to the routing
problem, where x_i is the flow on link i, can be written as the
convex combination of the corner points of the feasible region
ψ^i, where $\psi^i = (\psi_1^i, \ldots, \psi_n^i)$, as follows:

$$x_1 = \psi_1^1 Q_1 + \ldots + \psi_1^r Q_r$$
$$\vdots$$
$$x_n = \psi_n^1 Q_1 + \ldots + \psi_n^r Q_r \qquad (56)$$

where there are _r_ corner points, and the Q_i are scalars such
that

$$\sum_i Q_i = 1 \qquad (57)$$

and

$$Q_i \geq 0; \quad i=1, \ldots, r \qquad (58)$$

In fact, since there are n+1 equations, the system (56-58) has a solution such that at most n+1 of the Q_i are nonzero. Thus at most n+1 feasible points determine a subregion over which optimization can be performed. In the Cantor-Gerla algorithm, the master optimization is performed over a set with n+2 corner points, so that one of them can be eliminated and a new one generated by the linear program.

Decomposition Algorithm:

Step 0. Choose an initial set of ψ^i to be corner points.

The number __b__ of corner points initially used should be less than or equal to n+1. Cantor and Gerla use n+1 initial corner points; however, Deefenderfer [37] has been successful in using only one initial corner point, also working on a communication network problem.

Choose an initial basic solution $Q = (Q_1, \ldots, Q_b)$; for example, by setting $Q_1 = 1$ and all Q_i: $i \neq 1$ to zero. Let X be the initial feasible solution determined by the ψ^i and Q.

Master (nonlinear) Optimization:

Step 1. Minimize the objective function F over the subregion determined by the corner points ψ^i

Step 2. $Q \leftarrow$ the optimal solution $(Q_1, \ldots, Q_b)$ determined by Step 1.

Step 3. $X \leftarrow$ the new feasible solution determined by ψ^i and the new Q.

Step 4. Compute the vector $\nabla F = \frac{\partial F}{\partial X}$.

Linear Subproblem:

Step 5. If $b \leq n+1$, go to Step 6. Otherwise eliminate one of the corner points ψ^i. The corner point to which the smallest Q_i determined by the master program (step 1) corresponds is the one chosen to be eliminated.

Step 6. Use a linear programming method to find the solution to the problem of assigning traffic to the links of the network so that the shortest possible distance, according to the metric F, is traveled by the total traffic on the network. That is, assign a length $\frac{\partial F}{\partial x_i}$ to link i, and minimize the total distance traveled by all traffic. Note that this is a constrained optimization problem with a _linear_ objective function.

Step 7. Let $\psi^* = (\psi_1^*, \ldots, \psi_n^*)$ be the solution obtained by Step 6.

Step 8. Let $\Theta = \sum_{j=1}^{n} (\nabla F)_j (x_j - \psi_j^*)$. If $\Theta < \epsilon$, stop. Otherwise go to Step 9. (ϵ is some predetermined tolerance).

Step 9. Add ψ^* to the set of corner points. If $b \leq n+1$ then $b \leftarrow b+1$.

Go to Step 1.

Note that steps 5 and 9 control the subregion over which the master optimization is performed.

Cantor and Gerla prove that this algorithm does converge to the optimal solution if the cost function is convex, has continuous, nonnegative first derivative, and depends only on

-40-

the total flow on the links. Further, the algorithm can be mod-
ified to include different classes of traffic, essentially by
performing the decomposition steps in parallel for each class.
For our problem, assuming that there are $\underline{m}$ vehicle types, there
will be m sets of scalars Q, m sets of corner points Ψ, and
m linear programming problems must be solved and m optimality
(stopping rule) tests performed at each iteration. The master
step performs the nonlinear optimization over the subregion
generated by the several linear programs. A further modifica-
tion of the algorithm is necessary for the determination of
optimum traffic signal settings. One possible solution is
to simply perform the algorithm with various assigned traffic
signal settings (green splits) and thereby arrive at an ad hoc
solution.

CHAPTER IV. COMPUTATIONAL RESULTS

The passenger cost function of section II.6 was used for the network shown in Figure 7. Although this network does not represent any particular freeway corridor system, or part of one, it does include some of the most important features that would appear in an actual traffic assignment problem for such systems, as described in Chapter II. The link types, capacities, and lengths are given in Table 3.

Link	Type	Flow Capacity	Length
(1,2)	Freeway	6000	.5 miles
(2,3)	Freeway	6000	.5
(2,5)	On-ramp	2000	.1
(2,8)	Street	2000	.15
(3,6)	On-ramp	2000	.1
(3,12)	Freeway	6000	.5
(4,5)	Freeway	2000	.5
(5,2)	On-ramp	2000	.1
(5,6)	Freeway	2000	.5
(6,3)	On-ramp	2000	.1
(6,12)	Freeway	2000	.5
(7,8)	Street	4000	.5
(8,9)	Street	4000	.5
(9,10)	Freeway	4000	.5
(9,11)	Freeway	2000	.05
(10,9)	Street	2000	.05

Table 3. Link Types, Capacities, and Lengths

Figure 7. Freeway Corridor System Network
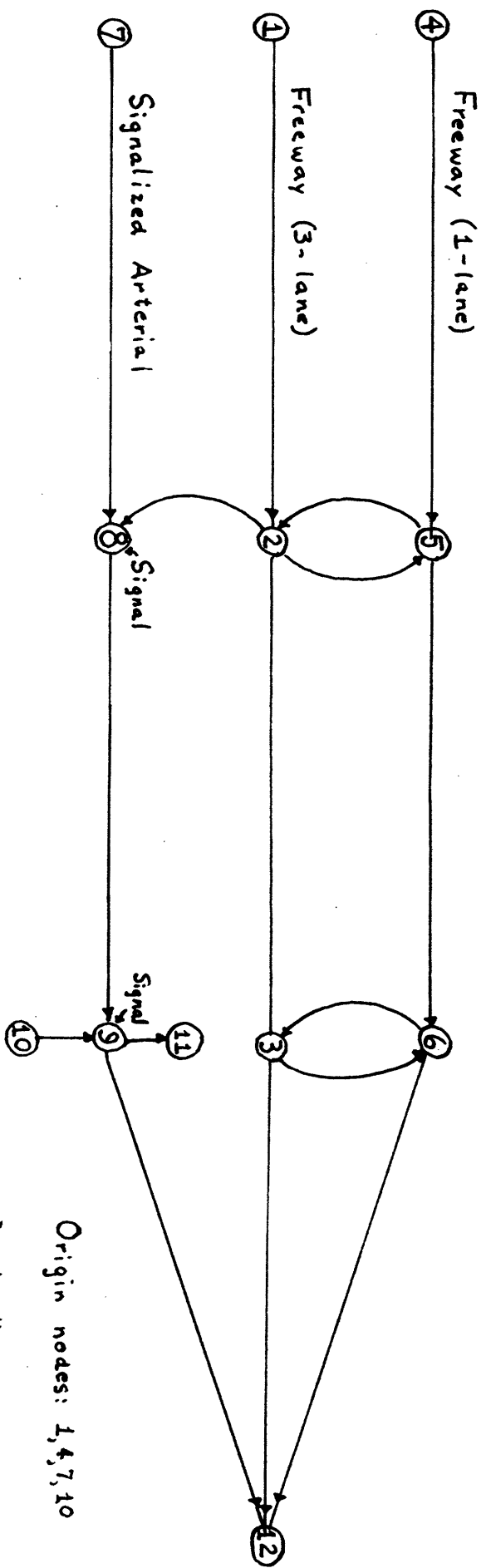
Origin nodes: 1,4,7,10

Destination nodes: 11,12

Traffic Signal Nodes: 8,9

Freeway Merges: 2,3,5,6

Freeway (1-lane)

Freeway (3-lane)

Signalized Arterial

Signal

Signal

The AGP method by itself, as described in section III.2, proved to be highly inefficient for this problem; in fact, so inefficient that it was impracticable to obtain converged optimal solutions to the problem by this method. Some examples of the results of attempting to use this method are summarized in Table 4.

Case	Problem	Variables and Constraints	Execution Time Per Iteration
1	1 vehicle type	36 variables 29 equality; 51 total constraints	7.41 sec
2	2 vehicle types 1 passenger/car 1 passenger/bus	52 variables 40 equality; 72 total constraints	25.59 sec
3	2 vehicle types 1.5 passengers/car 30 passengers /bus	same as case (2)	30.66 sec
4	same as case 3, with improved initial guess		24.46 sec

Although the computation time was somewhat dependent on the initial guess, for a small number of iterations, it was not possible to make much improvement by using the output from one run as the initial guess for the next, as was done in (3) and (4). Due to the expense of computer time, no more than five iterations of this method were performed on any run.

It seemed clear that the general AGP method was completely

unsuitable for this problem, and it was decided that a decomposition algorithm of the type described in section III.3 should be used. Unfortunately, it was not possible to code such an algorithm in time for this writing. However, Deefenderfer [37] has obtained the following results on a network of 11 nodes and 44 links, although with a different cost function, using a decomposition algorithm: convergence in thirty iterations, with a total computation time of .063 seconds.

CHAPTER V. CONCLUSIONS

As indicated in the Introduction, it is hoped that the approach to the freeway corridor system traffic assignment problem presented here will eventually find applicability as part of a control system which runs dynamically on an actual urban road network. The usefulness of the model presented in Chapter II will have to be determined by the collection of data, but will also depend on the controllability of the traffic flows and the accuracy with which data can be collected and processed in real-time. In particular, a means for distinguishing vehicle types will be necessary if passenger-cost minimization is implemented. Recent research at the Electronic Systems Laboratory indicates that it is possible to build electronic detectors that can, for example, distinguish between cars and buses. It is also possible that the model may have to be extended or modified to include areas not covered by the simplifying assumptions, as discussed in Chapter II.

Finally, it is concluded that if this method, or some similar approach involving a nonlinear cost function, is to be used as part of a real-time system, a decomposition algorithm similar to that of section III.3 be implemented for the purpose of determining an optimal traffic assignment.

REFERENCES

[1]   John G Wardrop, "Some Theoretical Aspects of Road Traffic
        Research", Proceedings of the Institute of Civil
        Engineers, 1952, Part II, pp. 325-378.

[2]   Stella C Dafermos and Frederick T Sparrow, "The Traffic
        Assignment Problem for a General Network", Journal
        of Research, U.S. National Bureau of Standards --
        B. Mathematical Sciences, Vol. 73B, No. 2, pp. 91-
        118.

[3]   Stella C Dafermos, "An Extended Traffic Assignment Model
        with Applications to Two-Way Traffic", Transporta-
        tion Science, Vol. 5, 1971, pp. 366-389.

[4]   Harold J Payne and Warren A Thompson, "Traffic Assignment
        on Transportation Networks with Capacity Constraints
        and Queueing", ORSA/TIMS Meeting, April 30 - May
        1-2, 1975.

[5]   M J Lighthill and G B Whitham, "On Kinematic Waves II, A
        Theory of Traffic Flow on Long Crowded Roads", Pro-
        ceedings of the Royal Society of London, Series A,
        Vol. 229, 1955, pp. 317-345.

[6]   Harold Greenberg, "An Analysis of Traffic Flow", Operations
        Research, Vol. 7, 1959, pp. 79-85.

[7]   Franco G Preparta, "Analysis of Traffic Flow on a Signal-
        ized One-Way Artery", Vol. 6, No. 1, Feb. 1972, pp.
        32-57.

[8]   R Wormleighton, "Queues at a Fixed Time Traffic Signal
        with Periodic Random Input", CORS Journal, Vol. 3,
        No. 3, Nov. 1965, pp. 129-141.

[9]   J H A de Smit, "The Transient Behavior of the Queue at a
        Fixed Cycle Traffic Light", Transportation Research,
        Vol. 5, Pergamon Press, 1971, pp. 1-14.

[10]   Steven J Hellinger, "Equilibrium Arrival Rates at a Par-
        tially Obstructed Intersection", Studies in Applied
        Math, Vol. 53, No. 1, March 1974, pp. 65-71.

[11]   R B Grafton and G F Newell, "Optimal Policies for the
        Control of an Undersaturated Intersection", in
        Vehicular Traffic Science, Leslie C Edie, Robert
        Herman, and Richard Rothery, ed., American Elsevier,
        New York, 1957, pp. 239-257.

[12]   Lloyd Sako and Samuel Zundlevich, "Proximal Decisions in
        Traffic Intersections", Transportation Engineering
        Journal, Aug. 1973, pp. 597-607.

[13]   Richard E Allsop, "Sensitivity of Delay at a Fixed Time
        Traffic Signal to Small Errors in the Observations
        Used for Calculating the Signal Settings", Traffic
        Flow and Transportation--Proceedings of the 5th
        International Symposium on the Theory of Traffic
        Flow and Transportation, G F Newell, ed., American
        Elsevier, 1972.

[14]   A G Hawkes, "Queueing at Traffic Intersections", in
        Proceedings of the 2nd International Symposium
        on the Theory of Traffic Flow, J Almond, ed.,
        Organisation for Economic Cooperation and Devel-
        opment, Paris, 1965.

[15]   Leonard Shaw, "On Optimal Ramp Control of Traffic Jam
           Queues", IEEE Transactions on Automatic Control,
           Vol. AC-17, No. 5, Oct. 1972, pp. 630-637.

[16]   Winifred D Ashton, "Distributions for Gaps in Road
           Traffic", Journal of the Institute of Mathematics
           and its Applications", Vol. 7, 1971, pp. 37-46.

[17]   Carlos F Daganzo, "Probabilistic Structure of Two-Lane
           Road Traffic", Transportation Research, Vol. 9,
           Pergamon Press, 1975, pp. 339-346.

[18]   A J Miller, "A Computer Control System for Traffic Net-
           works", in Proceedings of the 2nd International
           Symposium on the Theory of Traffic Flow, J.
           Almond, ed., Organisation for Economic Coopera-
           tion and Development, Paris, 1965, pp. 200-220.

[19]   H G Rosdolsky, "A Method for Adaptive Traffic Control",
           Transportation Research, Vol. 7, Pergamon Press,
           1973, pp. 1-16.

[20]   Nathan Gartner, John D C Little, and Henry Gabbay,
           "Optimization of Traffic Signal Settings in
           Networks by Mixed Integer Programming", Techni-
           cal Report No. 91, Operations Research Center,
           M.I.T., March 1974.

[21]   Sang Nguyen, "An Algorithm for the Traffic Assignment
           Problem", Transportation Science, Vol. 8, No.3,
           Aug. 1974, pp. 207-216.

[22]   Denos C Gazis, ed., Traffic Science, John Wiley and Sons,
           New York, 1974.

[23]   Paul K Houpt and Michael Athans, "Dynamic Stochastic
         Control of Freeway Corridor Systems, Volume I:
         Summary", M.I.T. Electronic Systems Laboratory,
         Report ESL-R-608, 1975.

[24]   Marcel Singleton and John E Ward, "Dynamic Stochastic
         Control of Freeway Corridor Systems, Volume V:
         A Comparative Study of Various Types of Vehicle
         Detectors", M.I.T. Electronic Systems Laboratory,
         Report ESL-R-612, 1975.

[25]   Daniel Ohrlac, Michael Athans, Jason L Speyer, and Paul
         K Houpt, "Dynamic Stochastic Control of Freeway
         Corridor Systems, Volume IV: Estimation of
         Traffic Variables Via Extended Kalman Filter
         Methods", M.I.T. Electronic Systems Laboratory,
         Report ESL-R-611, 1975.

[26]   Douglas Looze, Paul K Houpt, and Michael Athans,
         "Dynamic Stochastic Control of Freeway Corridor
         Systems, Volume III: Dynamic Centralized and
         Decentralized Control Strategies", M.I.T.
         Electronic Systems Laboratory, Report ESL-R-611,
         1975.

[27]   Stanley B Gershwin, "Dynamic Stochastic Control of
         Freeway Corridor Systems, Volume II: Steady
         State Optimal Traffic Assignment Using the
         Accelerated Gradient Projection Method", M.I.T.
         Electronic Systems Laboratory, Report ESL-R-609,
         1975.

[28]   Henry J Kelley and Jason L Speyer, "Accelerated Gradient
         Projection", in Lecture Notes in Mathematics, #132,

<p style="margin-left:3em">Symposium on Optimization, Balikrishnan, ed.,
Springer-Verlag, 1970, pp. 151-158.</p>

[29] David G Cantor and Mario Gerla, "Optimal Routing in a
Packet-Switched Computer Network", IEEE Transactions on Computers, Vol. C-23, No. 10, Oct. 1974.

[30] J W Cohen, The Single Server Queue, North-Holland,
Amsterdam, 1969.

[31] Sheldon M Ross, Introduction to Probability Models,
Academic Press, New York, 1972.

[32] W C Davidon, "Variance Algorithm for Minimization",
Computer Journal, Vol. 10, 1968, pp. 406-410.

[33] R Fletcher and M J D Powell, "A Rapidly Convergent
Descent Method for Minimization", Computer Journal,
July, 1963, pp. 163-168.

[34] J B Rosen, "The Gradient Projection Method for Non-
linear Programming, Part I: Linear Constraints",
SIAM Journal, March 1960; "Part II: Nonlinear
Constraints", SIAM Journal, March 1961.

[35] George B Dantzig, Linear Programming and Extensions,
Princeton University Press, Princeton, 1967.

[36] Ivan L Johnson, Jr. and Geraldine E Myers, "One Dim-
ensional Minimization Using Search By Golden
Section and Cubic Fit Methods", Manned Space-
craft Center Internal Note No. 67-FM-172,
NASA Manned Spacecraft Center, Houston, Nov 13,
1967.

[37]   Joseph Deefenderfer, M.I.T. Electronic Systems Laboratory,
         Personal Communication, May 1976.