1983

A dynamic space-time network flow model for city traffic congestion

Daniel J. Zawack
Carnegie Mellon University

Gerald Luther Thompson

Carnegie Mellon University.Design Research Center.

A DYNAMIC SPACE-TIME NETWORK FLOW MODEL FOR CITY TRAFFIC CONGESTION

by

D.J. Zawack & G.L. Thompson

December, 1983

DRC-70-18-83

A DYNAMIC SPACE-TIME NETWORK FLOW MODEL

FOR CITY TRAFFIC CONGESTION

by

Daniel J. Zavack

and

Gerald L. Thompson

October, 1983

This report was prepared as part of the activities of the Management Science Research Group, Carnegie-Mellon University, under Contract No. N00014-82-K-0329 NR 047-048 with the U. S. Office of Naval Research.
Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

Management Sciences Research Group
Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

A DYNAMIC SPACE-TIME NETWORK FLOW MODEL FOR CITY TRAFFIC CONGESTION

by

Daniel J. Zawack and Gerald L. Thompson

ABSTRACT

A space-time network is used to model traffic flows over time for a capacitated road transportation system having one-way and two-way streets. Also, for the first time, traffic signal lights which change the network structure are explicitly incorporated into the model. A linear (time) cost per unit flow is associated with each arc, and it is shown that under the model structure, travel time on a street is a piecewise linear convex function of the number of units traveling on that street. Hence congestion effects are explicitly considered while maintaining the linear nature of the model. Two efficient solution methods are proposed. A network flow solution for a multiple source single destination network and a shortest path solution for a single source single destination network.

Two examples are presented. The first example has one source and one sink. There is a unimodal buildup of traffic at the source (say a factory) which enters the street network as quickly as its capacity permits and proceeds through the network, stopping at red lights when necessary, towards the sink (a residential area). Computations with this example show that the arrival rate has multiple peaks which are induced by the stop lights.

In the second example there are multiple sources and one sink. The results here are similar except that the arrival rate has a single broad peak which is due to the extreme symmetry of the constraints of the problem.

KEY WORDS: Dynamic traffic assignment model, Space-time network flow model, Network traffic flow, Traffic congestion

I. INTRODUCTION

A fundamental problem faced when planning for the construction of new roads and the improvement of existing roads is to predict in detail the resulting traffic patterns. A model which can estimate, with reasonable accuracy, future patterns will assist in avoiding the creation of new bottlenecks. The resulting problem is generally known as the traffic assignment problem.

In the literature the problem of estimating traffic flows in road networks is generally formulated as an equilibrium problem. Knight [18] has given the following intuitive description underlying this equilibrium approach:

> "Suppose that between two points there are two highways, one of which is broad enough to accommodate without crowding all the traffic which may care to use it, but is poorly graded and surfaced, while the other is a much better road, but narrow and quite limited in capacity. If a large number of trucks operate between the two termini and are free to choose either of the two routes they will tend to distribute themselves between the roads in such proportions that the cost per unit of transportation, or effective return per unit investment, will be the same for every truck on both routes. As more trucks use the narrower and better road, congestion develops, until at a certain point it becomes equally profitable to use the broader but poorer road."

Wardrop [25] formalized the notion of a network equilibrium into the following two principles:

1. "The journey times on all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route."

(2) "The average journey time is a minimum.

Waldrop goes on to explain that

> "The first criterion is quite a likely one in practice, since it might be assumed that traffic will tend to settle down into an equilibrium situation in which no driver can reduce his journey time by choosing a new route. On the other hand, the second criterion is the most efficient in the sense that it minimizes the vehicle hours spent on the journey."

The first principle implies that each individual selects the shortest route available under given traffic conditions. The resulting pattern is generally called user optimal. The second implies a central authority imposing a set of routes which minimizes total travel time. The resulting pattern in this case is generally called system optimal.

In Wardrop's description of the problem the required number of trips between each origin destination pair is a specified constant. This problem is known as the fixed demand traffic assignment problem. The elastic demand traffic assignment problem was first enunciated by Beckman, McGuire and Winston [1]. In this problem.the number of trips between each origin destination pair is a function of the travel time between each origin destination pair.

Both of these formulations have been studied extensively. A survey of work done on problem formulation is given by Beckman [2]. The basic paper on the constant demand model was written by Dafermos and Sparrow [3]. This work has been extended by Dafermos [4,5,6,7], and Florian [12]. Nguyen [22] has developed efficient algorithms that are able to treat large scale problems encountered in practice. Florian and Nguyen [11] present the results of a validation study using the constant demand model. Gartner [13,14] provided a survey of work done on the elastic demand problem. Leblanc and Farhangian @192 and Dafermos [8] have made recent extensions. An introduction to the field is provided by Potts and Oliver [23].

In both the constant demand and elastic demand models a static set of origin destination demands is assumed to exist. This may be a valid long run assumption but it is not in the long run that road network congestion occurs. The bottleneck situations which arise are short term dynamic phenomenon. Rush hour does not occur on the average or in the long run and yet these types of peak load conditions are the most important to analyze in judging traffic patterns on a street network. By assuming traffic is in equilibrium one assumes away the dynamics of demand which is the crux of the problem which planners need to resolve. Hence the appropriateness of the equilibrium assumption must be questioned in many applications.

Recent papers by Heydecker [15] and Steinberg and Zangwill [24] also indicate that equilibrium solution concepts may not always be applicable to real traffic situations.

Traffic assignment models are often intended for analysis of rush hour situations in which the rate of traffic flow into the network will be increasing for some period until a peak rate is achieved and possibly maintained for a brief period, followed by a decrease. There may be multiple peaks in the rate of traffic flow as various factories or office buildings release their workers. Furthermore the road network itself is dynamic. As traffic signals change available routes are being opened and closed. Hence it is not likely that there will ever be an equilibrium state achieved in the network.

One model has been proposed by Merchant and Nemhauser [20] which considers dynamic flows. It allows multiple origins and a single destination and treats congestion explicitly. The model is nonlinear and nonconvex. Proof is given that it can be solved for a global optimum using a one pass simplex algorithm. The model minimizes the total cost of the flows. An efficient solution procedure for the model is presented by Ho [16].

In the present paper we propose a space-time network flow model which explicitly incorporates all the characteristics of road networks generally considered in the literature together with dynamic aspects not usually considered. Congestion is modeled by a linear space time network and it can be shown that, from a given start time, the time to travel a street is a piecewise linear convex function of the flow on the street at the start time. Each route has a capacity at any given time which is determined by the physical space available at that time on that route. The model accepts a variable set of inputs over time. Finally the dynamics of a road network caused by the changing of traffic signals are for the first time explicitly incorporated in a road network flow model.

Two types of solutions are proposed. The first for a multiple source single destination space time network is termed a network flow solution and requires a set of dynamic flows over the time horizon of the model which minimize total travel time. This solution is computed by use of the network code NETFLO, due to Kennington and Helgason [17], which is very fast. The second, for a single source single destination space time network termed a shortest path solution, gives a set of dynamic paths from a single source to a single destination which are the shortest available at a given time relative to traffic leaving the source in all prior time periods. This solution is obtained by repeated use of a shortest path algorithm which is known to have time complexity $0(n^2)$ where n is the number of nodes of the network.

In the second section of this paper the model is stated as a space time network model. The space time network model itself has more general application than simply traffic flow analysis, for instance, to the Job Shop Scheduling problem. In section three the model is a-plied specifically to the dynamic traffic flow problem. In sections four and five the space time network flow and shortest path solution methods are discussed. Finally in section six two hypothetical street networks are given and solutions are determined for each situation.

II. A TWO ATTRIBUTE DYNAMIC NETWORK FLOW MODEL

In this section a two attribute network flow model which is dynamic, i.e., changes over time, will be developed. The model is dynamic in two senses. First the input and output of the model can vary over time. Second, the network itself may change over time. In the present setting the other attribute beside time is space, but applications are under development where the second attribute may be other than space. For clarity in the description, the model attributes will be referred to as space and time, but it should be remembered that the space attribute could be interpreted as other attributes, also in other models there could be other attributes.

In the space-time setting the network is constructed from nodes and directed arcs having the following properties. Each node represents both time and location. Each directed arc represents a directed path of length k time units connecting a pair of nodes.

There are two types of arcs in the network model. The first type is a capacitated arc called a green arc. The terminology given indicates that as flows traverse these arcs they are in a go state, progressing from one location to another. A green arc connects two nodes (i,t) and (j,t+k(i,j)), j *4* i where node (i,t) represents location i at time t and node (j,t+k(i,j)) represents location j at time $t + k(i,j)$ where $k(i,j)$ is the length of time required under normal conditions to travel from location i to location j. The arc $[(i,t), (j,t+k(i,j))]$ connecting node (i,t) to $(j,t+k(i,j))$ indicates that it is feasible to travel from location i to location j in $k(i,j)$ time units starting at time t. The upper bound on the quantity of flow per unit time which may enter and traverse the arc $[(i,t),(j,t+k(i,j))]$ is called the capacity of the arc. If this capacity is exceeded the remaining units must wait to use green arc $[(i,t+1),(j,t+1+k(i,j))]$ or take another route. The capacity of green arcs connecting two locations i and j may vary over time.

The second type of arc is also capacitated and is termed a red arc. The terminology red indicates that as flows traverse these arcs they are in a delay state. Flows which traverse red arcs are simply passing time at a fixed location. A red arc connects two nodes (i,t) and $(i,t+1)$ where node (i,t) represents location i at time t and node $(i,t+1)$ represents location i at time $t+1$. The flow on arc $[(i,t),(i,t+1)]$ connecting node (i,t) to node $(i,t+1)$ represents flow units that are not allowed to flow on some green arc from node (i,t) and are delayed during the time interval $[t,t+1]$.

Consider the cause for this delay and the impact of delay on the time to travel between two locations i and j. The reason for delay at location i starting at time t is that the capacity of green arc $[(i,t),(j,t+k(i,j))]$ over which flow is desired has been reached preventing further flow through this arc starting at time t. If the flow is delayed then it remains at location i until time $t+1$ and, if there is capacity available on green arc $[(i,t+1),(j,t+1+k(i,j))]$, the flow may move forward, otherwise it will continue to be held at location i until time $t+2$.

Let $c(i,j)$ be the capacity of flow on any green arc connecting i to j and let $k(i,j)$ be the normal flow time from i to j. Then the process described above implies that the total time required for X flow units to travel from i to j given they start at time t is given by the following function:

$$F[X((i,t),j)] = k(i,j)\cdot X((i,t),j) \text{ if } 0 \le X((i,t),j) < c(i,j) \qquad \text{(a)}$$

$$= k(i,j)\cdot c(i,j) + [k(i,j)+1][X((i,t),j) - c(i,j)]$$
$$\text{if } c(i,j) \le X((i,t),j) < 2c(i,j) \qquad \text{(b)}$$

$$= k(i,j)\cdot 2c(i,j) + c(i,j) + [k(i,j)+2][X((i,t),j) - 2c(i,j)]$$
$$\text{if } 2c(i,j) \le X((i,t),j) < 3c(i,j) \qquad \text{(c)}$$

$$\vdots$$

$$k(i,j)\cdot nc(i,j) + \frac{n(n-1)}{2}c(i,j) + [k(i,j)+n][X((i,t),j) - nc(i,j)]$$
$$\text{if } nc(i,j) \le X((i,t),j) < B \le (n+1)c(i,j) \qquad \text{(d)}$$

Expression (a) states that for flow quantities $X((i,t),j)$ which are less than the capacity of the green arcs between i and j the total time to travel from i to j for all units leaving at time t is the normal travel time between i and j times the number of units travelling. Expression (b) states that if the quantity of flow $X((i,t),j)$ is greater than the capacity of green arc between i and j but less than twice that capacity then the total time to travel from i to j for all units leaving at time t is the normal time $k(i,j)$ for the first $c(i,j)$ units and then the normal time plus one for the units exceeding $c(i,j)$. Expression (d) states that if the quantity of flow is greater than n times the capacity of the green arcs between i and j but less than B

a bound on X((i,t),j) which is less than (n+1) c(i,j) then the total time to travel from i to j for all units leaving at time t is the sum of these terms. The first term is the normal travel time fcr the first n*c(i,j) flow units. The second term is the sum of the first n-1 integers multiplied by the capacity c(i,j) of the green arcs between i and j. This term is the additional time required for the (n-l)#c(i,j) units which exceed the normal capacity between i and j. The last term is the marginal travel time for the units exceeding n»c(i,j) multiplied by the number of units in excess. That is the flow units in the interval [n*c(i,j)+l,B] require the normal travel time plus n time units to travel from i to j. In expression (d) the upper bound Bon X((i,t),j), the flow starting at time t from location i to j, is the capacity of the red arc at location i (this capacity does not appear in the expression) plus the capacity c(i,j) of the green arc connecting i and j. The red arc capacity can be interpreted as the maximum length of a queue before the green arc.

Note that k(i,j) and c(i,j) are nonnegative. Hence the function F is nonnegative. Also note that for each interval [m c(i,j),(m+l)c(i,j)} for m = 0,1...n the slope of the function is a positive constant k(i,j) + m and as m increases the slope increases. Clearly the function F is piecewise linear convex. Hence the next theorem follows immediately.

THEOREM: The time required for X units of flow starting at time t to travel from i to j in the space time network is a piecewise linear convex function of X given by the function F above.

A dynamic space time network is constructed for a graph P = (S,N,D,A), where S is a set of source locations, N is a set of intermediate locations, D is a set of destination locations and A is a set of ordered pairs (i,j) i *4* j i e S U N , j e N U D . For each ordered pair (i,j) e A, let k(i,j) be the

normal travel time from i to j. This defines the length of arc (i,j) in P. Let c(i,j) be the flow capacity in flow units per unit time for arc (i,j) *e* A. Let T be the length of the time horizon over which the dynamic flow model will be constructed. Let u(i,j) be the length of the shortest path from i e S to j e SUNUD. Let u(*,j) - min u(i,j) j e S U N U D.
ieS

Based on this data the space time network G • *(D,N,G,R,V)* can be constructed where D is the set of destination locations of P, *hi* is the space time node set of G, 6 is the green arc set of G, R is the red arc set of G, and *V* is a dummy arc set of G

hi = {(i,t)/i e S U N J D , t« 0,...,T}

R •- U(i,t),(i,t+l)]/i e S JN, t - u(**>•-..,T-l}

G - {[(i,t),(j,t-Hc(i,j))]/(i,j) e A, t - u(*,i),...,T-k(i,j)}

V = {[(i,t),i]/i *e* D, t « u(*,i),...,T}

The arcs in *V* all have length zero and infinite capacity. They allow the flow to arrive at a destination at various times.

The length of each arc [(i,t) ,(j ,t4k(i,j))3 -c *G* is k(i,j). The length of each arc [(i,t),(i,t+l)] *e* R is 1. The capacities of the green arcs c[(i,t), (j,t+k(i,j))] and of the red arcs c[(i,t), (i,t+l)] are the maximum rates in units/unit time at which units may enter an arc. These rates can be specified in various ways depending on the specific application.

As described the graph G may have multiple sources and multiple destinations. In the applications of interest in this paper the results of solving a network flow problem associated with this network can only be meaningfully interpreted when quantities sent from a source can be identified as to their final destination. Hence if there are multiple destinations in the set D the associated network flow problem must be formulated as a multicommodity network flow model to insure units leaving a specific source arrive at a specific destination. The multiconmodity case will be treated in a later paper. In the rest of this paper the focus will be on situations where D is a singleton. Then a standard network flow problem can be solved while insuring all units arrive at the appropriate destination.

Next the network flow problem associated with G will be defined.

Let P = (S,N,D,A) be as the above graph where D • {d} and let G = *(D,N,G,R,V)* be the associated space time network.

Let x[(i,t),(j,t+k(i,j))] be the flow in arc [(i,t),(j,t4k(i,j))] *e G.*

Let x[(i,t),(i,t+l)] be the flow in arc [(i,t),(i,t+l)] e R

Let x[(d,t),d] be the flow in arc [(d,t),d] *e V.* This is the flow arriving at the destination at time t. It is an instantaneous flow from location d at time t to a timeless representation of location d.

Let b(s,t) be the supply of flow units at source seS at time t.

Let G(s,t) = {[(s,t)(j,t+k(s,j))] e G} for seS and fixed t.

FG(i,t) - {[(i,t),(j,t+k(i,j))] e 5} for ieN and fixed t.

TG(i,t) = {[(j,t-k(j,i)),(i,t)] c G} for IeN and fixed t.

FR(i,t) -' {[(i,t),(i,t+l)] e R} for ieN and fixed t.

TR(i,t) = {[(i,t-l),(i,t)] c R} for ieN and fixed t.

G(d,t) - {[(j,t-k(j,d)),(d,t)] cG} for deD and fixed t.

The network flow problem is defined by the following statements. A brief discussion of them follows

(1) Min $\sum_{G}$ k(i,j)-X[(i,t),(j,t+k(i,j))] + $\sum_{R}$ X[(i,t),(i,t+1)]

Subject to

$$(2)\quad \underset{G(s,t)}{E}\; X[(s,t),(j,t\text{-}Hc(s,j))] + X[(s,t),(s,t+1)] = b(s,t)$$

for seS, $t = 0,\ldots,T-1$

$$(3)\quad \underset{TG(i,t)}{E}\; X[(j,t-k(j,i)),(i,t)] + \underset{TR(i,t)}{E}\; X[(i,t-l),(i,t)]$$

$$- \underset{FG(i,t)}{E}\; X[(i,t),(j,t4k(i,j))] - \underset{FR(i,t)}{E}\; X[(i,t),(i,t+1)] = 0$$

for ieN, $t = u(*,i)\ldots\ldots T-l$

$$(4)\quad \underset{G(d,t)}{E}\; X[(j,t-k(j,d)),(d,t)] - X[(d,t),d] - 0$$

for $\{(d,t)|\ t = u(*,d),\ldots,T\}$

$$(5)\quad \underset{t(d,t),d]\ e\ V}{E}\; X[(d,t),d] = \underset{seS}{E}\ \overset{T-l}{\underset{t=0}{E}}\ b(s,t)$$

$$(6)\quad 0 \wedge x[(i,t),(j,t+k(i,j))] \leq c[(i,t),(j,t+k(i,j))]$$

for $[(i,t),(j,t+k(i,j))]$ c G

$$(7)\quad 0\ {*}x[(i,t),(i,t+1)] \leq c[(i,t),\ (i,t+1)]$$

for $[(i,t),(i,t+l)]$ e R

$$(8)\quad x[(t,d),d]\ \wedge 0 \quad \text{for } \{(t,d)\ |\ t = 0,\ldots,T\}$$

The objective function (1) calls for the minimization of the sum of total flow time in the network. The constraint set specifies the following. Equation (2) says that the sum of flows leaving a source location s at time t via all green and red arcs emanating from that source must equal the supply b(s,t) at the source node. This must be true for all source locations and times. Equation (3) says that the sum of flows into any node other than a source or destination node must be equal to the sum of flows out of the node. Equation (4) indicates that the sum of flows into the destination at any specific time t must be equal to the flow out of the destination at that time t and into the dummy destination. Equation (5) states that the sum of the flows into the dummy destination must equal the sum of all supplies from all source locations in all time periods. Finally (6) and (7) specify that the flows on all green and red arcs are non-negative and capacitated.

As stated at the outset of the section, the model is dynamic in two aspects. The first is that inputs at source locations can be varied in any fashion over time. The output is also free to arrive at the destination in any feasible manner where feasibility is determined by the arc capacities. The second dynamic aspect is that the network itself can be made to vary over time. This can be done by changing the capacities or. the arcs. An arc can be closed to flow at a specific time simply by setting it's capacity at zero. Similarly if conditions change so that at a specific time capacity on a path in the network is reduced then the capacity of the corresponding arcs can be reduced accordingly.

III. DYNAMIC TRAFFIC FLOW MODEL

Traffic flow has three features that car readily be captured by the dynamic two attribute network model.

First traffic flow is dynamic because the quantity of traffic entering a given street network varies over time. During rush hour the entering traffic may well exceed the normal capacity of the street causing long delays, where only a few minutes before the rush hour period traffic might be light. This characteristic of traffic flow is accounted for in the model by allowing varying source quantities to flow into the model over time.

Second as the quantity of traffic in the streets increases, congestion results, causing the travel time between any pair of locations to increase as a function of the number of vehicles on the road. This is precisely what occurs in the model. As green arcs between i and j reach their capacity traffic will be forced to utilize red arcs at i causing the travel time between i and j to increase as a function of the units traveling between i and j at a given time. Different travel time functions can be fitted by changing the capacity settings on the green arcs.

Third, when traffic signals change the street network changes its structure. When a signal is red in a given direction the intersection has been closed in that direction until the signal becomes green again. When the signal changes the intersection is reopened in the original direction and it is closed in the direction perpendicular to the original. As a traffic signal changes from one time period to the next the effect is to change the capacity of the routes controlled by the signal. For all routes using the direction in which the signal is red, the route capacity has been temporarily reduced to zero. Changes in route structure as a result of traffic signal operation can be incorporated into the model by varying the capacities of green arcs over time. This feature of the model will be demonstrated by an example.

Consider a two way intersection which normally requires one time unit to cross in any direction. Such an intersection can be represented by 8 nodes and 12 green arcs as in Figure 1.

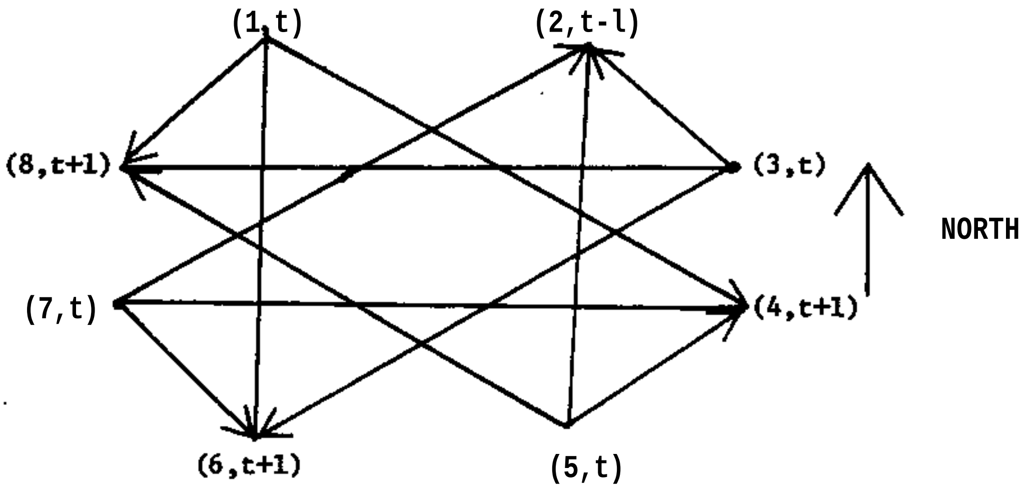


FIGURE 1. The network associated with the intersection of two two-way streets in which left and right turns and right-turn-on-red (RTOR) are permitted. If one or both of the streets is one-way then parts of this network should be eliminated.

Arcs	Nodes		Green arc capacities	
	From	To	North-South	East-West
1	1,t	6,t+1	20	0
2	1,t	8,t+1	10	5 (RTOR)
3	1,t	4,t+1	5	0
4	3,t	8,t+1	0	20
5	3,t	2,t+1	5 (RTOR)	10
6	3,t	6,t+1	0	5
7	5,t	2,t+1	20	0
8	5,t	4,t+1	10	5 (RTOR)
9	5,t	8,t+1	5	0
10	7,t	4,t+1	0	20
11	7,t	6,t+1	5 (RTOR)	10
12	7,t	2,t+1	0	5

TABLE 1. Green arc capacities for the network in Figure 1

Table 1 lists the 12 green arcs and their capacities. There are two sets of capacities associated with the arcs. In the column labeled green NS, capacities are specified as if the traffic signal were green in the north south direction and red in the east west direction. In the column labelled green EW, capacities specifying the opposite traffic signal condition are given. According to the capacities in the green NS column, traffic is prohibited from flowing east or west. Also left turns originating from either the east or the west are prohibited. Right turn on red (RTOR) is allowed. Traffic from the east or west may make a right turn, but the capacity of the right turn arcs have been reduced from their normal capacity of 10 to 5. Recall there are red arcs connecting (a,t) to (a,t+1) where a $\varepsilon\{1,3,5,7\}$ and (b,t+1) to (b,t+2) where b $\varepsilon\{2,4,6,8\}$.

Thus when traffic is delayed by a red signal it follows a red arc or if the signal is red for more than one time unit then it follows red arcs for as many time units as the signal is red after which it may resume its forward progress on a green arc. So by varying the capacities of green arcs which provide connections across intersections, all the workings of a traffic control device can be modelled.

Note that the 8 node 12 arc intersection representation of the previous example subsumes all intersection types in which two streets cross. For instance an intersection in which traffic is 2 way in one direction and one way in the other is formed out of a subset of the arcs and nodes of the example. Similarly if both cross streets are one way.

Next the construction of a dynamic traffic flow model from a street diagram will be discussed. First a graph $P = (S,N,D,A)$ is required. P is the graph formed from the map of streets and intersections. This graph has a node set S which includes all locations where traffic will enter the system. A second node set N which includes all intermediate locations where two or more streams of traffic join together or where one stream of traffic may separate into two or more streams. These nodes are associated with intersections on the map. A single traffic intersection where two streets cross may generate as many as 8 nodes if two way traffic is allowed on each street. As shown in Figure 1 of the previous example a location where 2 way traffic crosses has four points of entry to the intersection where traffic streams may separate and four points of exit from the intersection where streams of traffic may join. Partition the intermediate node set N into two subsets: NB is the subset of nodes in N which are beginning nodes of arcs making the connection across an intersection (such as nodes 1, 3, 5, and 7 in Figure 1). Similarly NE is the subset of nodes in N which are ending nodes of the same type arc (such as 2, 4, 6, 8 in Figure 1). Clearly N = NB U NE and NB f) NE ,= 0. The third node set is D which is the singleton destination of the traffic.

The final element of P is the set A of directed arcs partitioned into two subsets AI and AS. The set AI is the subset of arcs connecting nodes in NB to nodes in NE. These arcs carry traffic across intersections. Given AI define the complimentary set AS • A - AI. These arcs are associated with the streets.

In order to demonstrate these concepts consider the street diagram in Figure 2.

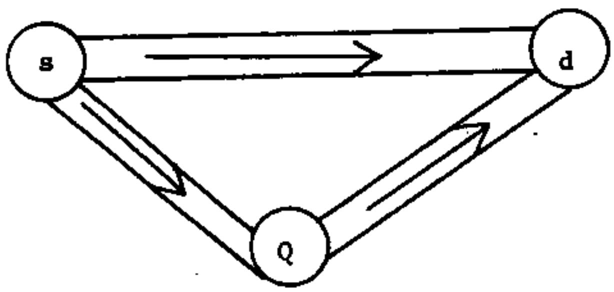


FIGURE 2. Simple street network

In this situation there are 3 streets, all one way. One can travel from s to d or from s to Q and from Q to d. Assume that at Q there is a conflicting stream of traffic (not shown) and hence there is an intersection to cross and a traffic signal at Q.

From this street diagram the graph P given in Figure 3 is formed with the sets S « {s}, NB - {q}, NE -' (r), D - {d}, AS -' {(s,d)(s,q)(r,d)} and AI - ((q,r)}.

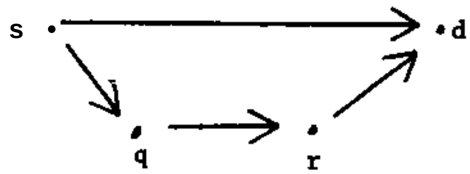


FIGURE 3. Graph associated with street network in Figure 2.

Normal travel times k(i,j) are associated with each arc (i,j). This is the time required to travel from i to j on arc (i,j) where there is no congestion on this arc. Also dynamic capacities c(i,j) are assigned to each arc (i,j) where the dynamic capacity of an arc (i,j) is the maximum number of vehicles which can depart per unit time and travel from i to j in the normal time k(i,j). Finally what might be termed the parking capacity of each street is required. This value will be utilized in establishing red arc capacities in the dynamic network. The parking capacity Z(i,j) of each street (i,j) e AS is the maximum number of vehicles which could be parked end to end within the lane structure of the street. For the example in Figure 3 the dynamic and parking capacities and normal travel times appear in Table 2.

	Arcs	Normal Travel Time	Dynamic Capacity	Parking Capacity
	(s,d)	k(s,d) = 2	c(s,d) = 5	2(s,d) = 25
AS	(s,q)	k(s,q) - 1	c(s,q) = 3	2(s,q) = 15
	(r,d)	k(r,d) =* 1	c(r,d) = 3	2(r,d) = 18
AI	(q,r)	k(q,r) - 1	c(q,r) = 3	

TABLE 2. Travel times and capacities for graph in Figure 3

Utilizing the specified data the graph G of a dynamic space time network can be constructed. First a time horizon T must be chosen. The node set is N-- {(i,t)|i e S JNB LINE JD, t = u(*,i,... ,T}. The red arc set is R = {[(i,t),(i,t+l)] |i e S UNB UNE, t = u(*,i),...,T-1} where the length of each red arc is 1. The green arc set is G = {[(i,t),(j,t+k(i,j))]|(i,j) e AI U AS, t = U(*,i),... ,T-k(i,j)} where u(*,i) is the shortest path from any s e S to i as in Section II. The associated $k(i,j)$ is the length of each green arc. *The* dummy arc set is V -- {[(i,t),i] |i e D t = u(*,i),... ,T} where each dummy arc has length 0 and infinite capacity.

Next the capacities on the green arcs and the red arcs must be specified. In order to specify the green arc capacities it is useful to consider two disjoint subsets of green arcs. The first being the subset of green arcs which correspond to streets. Call this set *GS* where *GS* - {[(i,t),(j,t-Hc(i,j))l| (i,j) e AS, t - u(*4)...T-k(i,j)}. The capacity c[(i,t),(j,t+k(i,j))] of arc [(i,t),(j,t+k(i,t))] *e GS* is c(i,j). It is the maximum number of vehicles which can depart per unit time and travel the street in normal time. The second subset of green arcs contains those arcs such as arc [(l,t),(2,t+l)] in Figure 1 which correspond to crossing an intersection. Call this set GI where GI * {[(i,t),(j,t-Hc(i,j))]|(i,j) *e* AI, t -u(*$_f$i)...T-k(i$_f$j)}. The capacities on these arcs are a function of the traffic signal pattern. In period t if traffic is not restricted by a signal from leaving node i then the capacity c[(i,t),(j,t-Hc(i,j))] of arc [(i,t),(j,t-Hc(i,j))] *e* GI is c(i,j) where c(i,j) is as defined above. In period t if traffic is stopped by a signal from leaving node i to go to node j then the capacity of the corresponding arc is zero. If on the other hand traffic in period t is restricted by a signal from leaving node i to go to j, for example in the case of a right turn on red then the capacity of the corresponding arc is reduced from c(i,j) in an appropriate manner.

The red arcs permit queues to form. Three types of queues may be visualized. The first are queues which may form at the source. If at time t there are more vehicles desiring to enter the system from a source location than capacity allows at time t then a queue forms of vehicles which must wait until time $t+1$ to enter the system.

The two other types of queues form on streets that begin and end at intersections. The first of these arises from the build up of vehicles waiting at a red traffic signal. These are vehicles at the end of the street waiting to cross an intersection. The second of these queue types is formed by the build up of traffic on the street as it becomes congested. This may be viewed as traffic waiting at the beginning of the street for its opportunity to travel the street in normal time.

In order to specify the red arc capacities the red arc set is partitioned into three disjoint subsets corresponding to the three queue types discussed above. The first is the subset RS of red arcs formed from the nodes in S where $KS - \{[(i,t),(i,t+l)] \mid ieS,\ t = 0,...,T-l\}$. These are the red arcs in which a queue may form at a source location. The capacity of each of these arcs is infinity because there is no restriction on the number of vehicles which may be waiting at a source location at any given time to enter the network.

The other two sets of red arcs form the queues at the beginning and end of streets which begin and end at an intersection. The capacities of these red arcs are physically constrained by the parking capacity of the street. The sum of vehicles at any given time in these red arcs plus the number of vehicles travelling on the street cannot exceed the parking capacity $Z(i,j)$ of street $(i,j)eAS$. In order to specify these red arc capacities for a street (i,j) an expression called $MVT(i,j)$ is developed for the maximum number of vehicles travelling on a street at any given time. Once $MVT(i,j)$ is known, the remaining parking capacity may be assigned to the red arcs at the beginning and end of the street. The value assigned to $MVT(i,j)$ is the maximum number of vehicles which can be admitted to the street per unit time (capacity) multiplied by the number of time units normally required to travel the street. For example if 10 vehicles per time unit may start and travel a street in 3 time units there may be a maximum of 30 vehicles travelling the street at a given time. Hence for a street $[((i,j),(j,t+k(i,j))] \in GS$, $MVT(i,j) = k(i,j) \cdot c[(i,t),(j,t+k(i,j))]$. Thus the total number of vehicles allowed to be waiting at the beginning and at the end of arc $[(i,t),(j,t+k(i,j))] \in GS$ is $Z(i,j) - MVT(i,j)$.

Given $Z(i,j) - MVT(i,j)$ for $(i,j) \in AS$ the total capacity available to be assigned to red arcs at the beginning and end of a street is known. The subset RB of red arcs at the end of the street is formed from nodes in NB where $RB = \{[(j,t),(j,t+1)] \mid j \in NB,\ t = u(*,j),\dots,T-1\}$. These are the red arcs in which a queue may form in front of a red light. If L is defined to be the length in time units of the green cycle of the intersection at j and t^* is a fixed time when the traffic signal at j is green then the capacity of these arcs is

$$\min\{L \cdot \sum_{[(j,t^*),(a,t^*+k(j,a))] \in GI} c[(j,t^*),(a,t^*+k(j,a))],\ k(i,j) - MVT(i,j) - c(i,j)\}.$$

The first expression in the minimization is the number of vehicles which can pass through the intersection at the end of street $(i,j) \in AS$ during a green cycle of the traffic signal. This is found by summing the capacities of all arcs leaving location j during a green phase of the traffic signal. The second expression is capacity available to the red arcs at the beginning and end of street (i,j) less the capacity of (i,j). Hence the capacity of the red arcs in RB is the maximum size of a platoon that can pass through the intersection from j during the next green cycle unless this amount is greater than the total number of vehicles allowed to be in the red arcs at the beginning and end of street (i,j) less $c(i,j)$. This insures that there will always be room for at least $c(i,j)$ vehicles at the beginning of the street.

Remaining is the subset RE of red arcs formed from the nodes in NE where $RE = \{[(i,t),(i,t+1)] \mid i \in NE,\ t = u(*,1),\ldots,T-1\}$. These are the red arcs in which queues may form at the beginning of the street (i,j). The capacity of these arcs is the physical space remaining for vehicles after considering the maximum number of vehicles traveling the street plus the number allowed to be waiting at a signal. Specifically

$c[(i,t),(i,t+1)] = Z(i,j) - MVT(i,j) - c[(j,t+k(i,j)),(j,t+k(i,j)+1)]$ for $[(i,t),(i,t+1)] \in RE$. Note that j is uniquely determined because all the traffic at node (i,t) where $i \in NE$ is entering a unique street. G is now completely defined.

Figure 4 shows the graph G associated with the graph P of the example. The time horizon T has been selected to be 5. Table 3 contains a listing of the arcs (except for the dummy arcs) with their associated lengths and capacities. Note that the intersection formed by arcs from location q to location r has a green traffic signal in period 2 and a red traffic signal in periods 1, 3.

GREEN ARCS

ARC	LENGTH	CAPACITY
(S,O)(q,l)	1	3
(s,i)(q,2)	1	3
(S,2)(q,3)	1	3
(S,3)(q,4)	1	3
(S,0)(d,2)	2	5
(S,l)(d,3)	2	5
(S,2)(d,4)	2	5
(S,3)(d,5)	2	5
(q,D(r,2)	1	0
(q,2)(r,3)	1	3
(q,3)(r,4)	1	0
(r,2)(d,3)	1	3
(r,3)(d,4)	1	3
(r,4)(d,5)	1	3

RED ARCS

ARC	LENGTH	CAPACITY
(S,O)(S,1)	1	∞
(S,1)(S,2)	1	∞
(S,2)(S,3)	1	∞
(q,D(q,2)	1	3
(q,2)(q,3)	1	3
(r,2)(r,3)	1	15
(r,3)(r,4)	1	15

TABLE 3

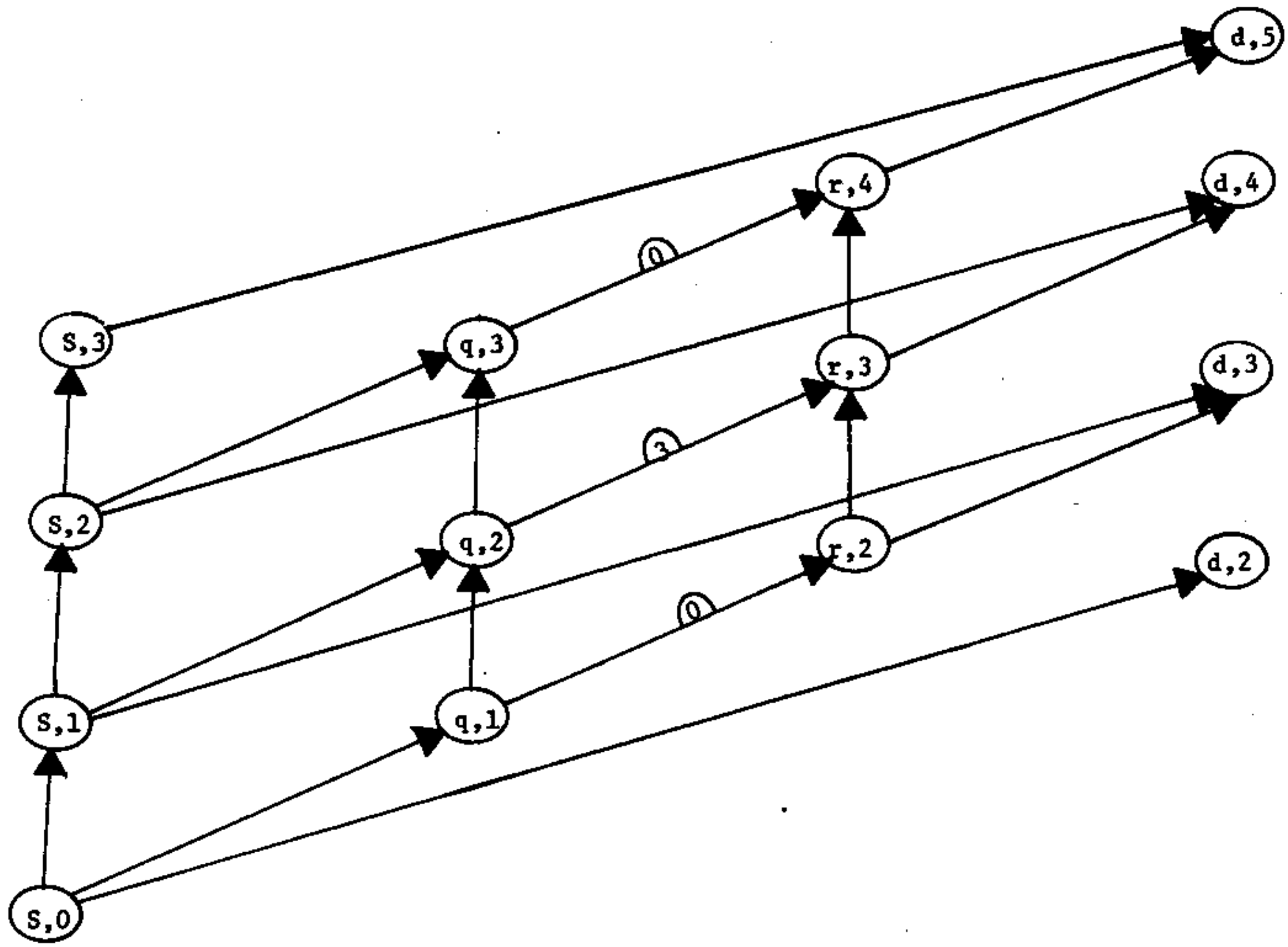


FIGURE 4. Dynamic graph for the example. Arcs from q to r have varying capacities as shown depending on the stop light settings.

IV. NETWORK FLOW SOLUTION

One solution to the problem of finding a set of optimal traffic flows in a dynamic space time traffic flow network G is obtained by solving the network flow problem defined in Section II by statements (1) through (8). This problem is a capacitated network flow problem and is easily solved by available network codes [17] . Examples of such solutions are given in Sectior VI.

The average travel time for vehicles in the network is the sum of all vehicle travel times divided by the number of vehicles. Since in the network flow problem the number of vehicles is fixed and the sum of vehicle travel time is minimized the average travel time is also minimized in the network flow solution for vehicles in the network. In order to interpret a network flow solution it is assumed that all travelers in the network are willing to cooperate to achieve a traffic flow pattern which minimizes the average travel time for all vehicles. Achieving this objective may require that some vehicles do not follow a path currently available to them having travel time (for them) shorter than that given by the path chosen (for them) by the network flow model. The latter occurs when the input exceeds normal capacity of the network for a number of consecutive periods and the path currently available, though shorter in total duration, involves waiting at some location. The fact that normal network capacity is exceeded means that for every space in the system which opens there is a vehicle waiting to fill that space. Hence any unnecessary waiting by an earlier vehicle only results in increased delay to later vehicles. By continuing on even a longer path and not waiting for the shorter path to open, delay for following units is reduced, thus reducing average travel time.

The dynamic space time traffic flow network has been described as a multiple-source-single-destination network. The solution of the associated network flow problem specifies a set of flows which will bring all vehicles from all source locations and all start times to the specified destination. This problem can be viewed as finding optimal shipping routes for a single commodity from multiple sources to a single destination.

Intuition might lead one to conclude that the symmetric problem having a single source and multiple destinations could also be solved by a simple network formulation; but it cannot. In the single source multiple destination problem vehicles will be scheduled to leave the source at a specific time with a specific destination. The vehicles required at a given destination will be the sum of vehicles sent from the source in all time periods bound for that destination. If the simple network flow problem associated with the single source multiple destination network is solved the solution will fulfill the requirements at each destination with the required units coming from any set of source times such that the overall solution is optimal. The required units need not have come from the correct starting locations or at the correct starting times. The only way to insure the solution corresponds to the original requirements is to solve the problem as a multicommodity network flow problem where each of the sets of vehicles leaving from a given location and bound for a given destination comprise a separate commodity.

V, THE SHORTEST PATH SOLUTION

The shortest path solution will be defined only for a single source single destination street graph P. The reason why it cannot be defined more generally will be made clear after the algorithm for finding a shortest path solution is stated.

The intuition behind a shortest path solution is that all users desire to reach the destination as quickly as possible. This implies the optimal route for a user is the shortest path from source to destination, but since the system is dynamic and can become congested, the path which will arrive at the destination as early as possible depends on both the level of traffic and the traffic signal sequence ahead.

The algorithm for determining a shortest path solution first dispatches all traffic arriving at the start node at the first period. When all of the first period traffic has been routed from source to destination the traffic arriving at the start node at the second period is then routed from source to destination, etc. Traffic is always dispatched along the shortest route having unused capacity until the capacity $c[(i,t),(j,t+k(i,j))]$ of some arc in the route is reached, and the route is saturated. Then a new dynamic space time traffic flow network is generated. The new network is the same as the old network except that the capacities of all arcs in the route along which traffic was dispatched are reduced by $c[(i,t),(j,t+k(i,j))]$, the quantity of traffic dispatched. Then the shortest path in the new network with available capacity is found and traffic dispatched along that route until the capacity $c[(i,t),(j,t+k(i,j))]$ of some arc in the route is reached. Then another new graph with reduced arc capacities is generated and the process repeated. Note that in this process the same physical path may be followed by successive dispatches of vehicles with congestion delay added as vehicles follow red arcs. Therefore these paths will require a longer time to traverse.

In order to interpret this solution two assumptions about user behavior are made. First, users wish to arrive at the destination as soon as possible after leaving the source. Second, the user has perfect knowledge of traffic conditions and traffic signal sequencing ahead and with this knowledge selects the shortest available path.

The algorithm for finding the shortest path solution is given in Appendix A.

The paths which are found can be used to calculate the average travel time for all vehicles and the variance in travel time. The actual routes traveled from source to destination are also available.

At the beginning of the section it was stated that this algorithm is only defined for a single source single destination P graph. In this setting traffic is always dispatched so as to give priority to vehicles in order of their departure times. If there are multiple sources and a single destination then dispatching by time priority alone does not uniquely define a solution. To see the difficulty note that vehicles may leave from two separate sources at the same time and arrive at some intermediate location at the same time. If the route ahead will not admit both vehicles due to capacity constraints, it is no longer clear which vehicle should go and which should wait.

The algorithm can also be applied to a network flow solution for a single source single destination network. Dispatching the solution on a time priority basis yields a set of routes and quantities for those routes which are consistent with the network flow solution. Comparison of the shortest path solution and a time priority dispatch of a network flow solution allow one to observe when a route other than the shortest available is utilized by the network flow solution in order to reduce the sum of travel times.

In order to find the shortest path solution corresponding to network flow solution, a capacitated network defined by the network solution is utilized. The flows assigned to the arcs in the network solutions are used to redefine the arc capacities. Arcs with zero flow in the solution are assigned zero capacity. Arcs with positive flow values in solution are assigned their respective flow values as capacities. The algorithm for finding a shortest path solution can be applied to the resulting network to find a consistent shortest path solution. In order to interpret the resulting solution, it is assumed that users follow shortest paths given, they are constrained to be the actual flows in the network flow solution.

VI. COMPUTATIONAL EXAMPLES

The model has been applied to two hypothetical street diagrams. The first one has a single source, 23 intersections and a single destination. The street diagram appears in Figure 5. This diagram was transformed into a P graph and then a G graph. In the P graph each intersection in the diagram is exploded into the appropriate subset of the 8 node 12 arc representation of intersection movements in Figure 1. Each node of the diagram in Figure 5 becomes anywhere from 6 nodes and 7 arcs to 3 nodes and 2 arcs in the P graph. For instance node 33 requires 6 nodes and 7 arcs to represent it in the P graph where as node 12 requires only 3 nodes and 2 arcs in the P graph. Also in Figure 5 arcs with arrowheads at both ends represent streets which allow two way traffic. Hence these arcs are represented by two arcs in the corresponding P graph. The resulting P graph has 101 nodes and 145 arcs.

All intersections require one time unit to cross. It takes one time unit to travel from the source to any of the three intersections reachable from the source. All east-west streets require two time units to traverse. In the south direction three time units are required to traverse the streets with intersection numbers in the thirties. It takes two time units to traverse the streets with intersection numbers in the twenties or in the forties.

Four time units are required to traverse streets with intersection numbers in the tens or in the fifties. It also takes four time units to travel between intersections 14 and 25 and between 54 and 45. Finally, two time units are required to travel from 25 or 45 to the destination 60.

The capacity from the source to each of the three intersections reachable from it is 100 units/time unit. The capacity of all east west streets is 15 units per time unit. The capacity of all south bound streets except for those with intersections having numbers in the twenties or forties is also 15 units per time unit. South bound streets with intersection numbers in the twenties or forties have capacity 30 units/time unit. The streets from intersections 25 and 45 to the destination have capacity of 150 units/time unit. Arcs crossing intersections had their capacities set as the smaller of the capacity of the street from which traffic is exiting or the capacity of the street to which traffic will enter.

The G graph was created from the P graph for a 37 period horizon. The resulting G graph contains 2958 nodes and 6069 arcs. Only intersections having potentially conflicting traffic were signal controlled. For instance, intersections 51 and 52 were not signaled since traffic simply merges at 51 and separates at 52. There are 15 intersections where traffic may actually cross and these intersections were signal controlled. All the signals were on a simultaneous 3 period cycle in which for 2 periods the signal allowed traffic to flow south while it was stopped east and west except for right turns. Then for one period the situation was reversed.

Figure 6 shows the graph of the number of departures from the source in each period from time 0 to time 10. The arrivals at the destination for the network flow and shortest path solutions are also shown in Figure 6. Observe that while the input increases to a peak and then decreases, the output is very irregular. The restrictions caused by the traffic signals are the cause of the irregular output. This is evidenced by the fact that the time from trough to trough and from peak to peak in the arrival graphs is 3. This is the same as the traffic signal cycle time.

In a street system with a single source and a single destination a set of streets whose removal from the system separates the source from the destination is called a separating set. In such a system, a separating set with minimal flow capacity determines the maximum flow capacity of the street system. In Figure 5 the maximum flow capacity is 105 given, for instance, by the capacity of the separating set $\{(13,14),(23,24),(33,34),(43,44),(53,54)\}$. From period one through period five the rate of departure exceeds the maximal flow capacity of the street system. In fact from period one through period eight a total of 900 units are trying to enter the network. This means there is input in excess of capacity for eight consecutive periods yet at the destination, node 60, no pattern of arrivals that repeats exactly can be observed.

Figure 7 shows the average travel times for each departing group under the network flow and shortest path solutions. For the network flow solution the arrival times were derived by assuming shortest paths were followed given the set of flows defined by the solution. The first departing group of 60 does better on the average under the shortest path solution, but after the first departing group the network flow solution travel times are better on the average for all other departing groups. To achieve this better average performance the first group of travelers do not take the shortest paths available. Rather some of them follow longer paths so those following will not be delayed in queues waiting to use shorter paths. Clearly, the individual and group goals are in conflict in this situation. This conflict may make it difficult or impossible to actually impose the network flow solution. In an actual application the observed solution will probably not be exactly the same as either solution.

The second street diagram has 12 sources, 8 intersections, and a single destination. The street diagram appears in Figure 8. As in the previous example, the street diagram was transformed into a P graph and then a G graph. First the P graph is formed by exploding the intersections in to the appropriate subset of the 8 node 12 arc representation of intersection movements. The resulting P graph has 74 nodes and 108 arcs.

All intersections require one time unit to cross. It takes three time units to travel from any source to its reachable intersection. It takes two time units to travel between any pair of intersections. Finally, the arcs leading to destination 33 are also two time units long.

The capacity of each of the arcs except the four arcs connecting to the destination is 10 units/time unit. The four arcs connecting to the destination each have a capacity of 30 units/time unit.

The G graph was created from the P graph for a 25 period horizon. The resulting G graph contains 2034 nodes and 4335 arcs. All intersections serve potentially conflicting traffic flows and therefore all intersections were signal controlled. All the signals were on a simultaneous 4 period cycle in which for two periods the signal allowed traffic to flow north and south while it was stopped east and west except for right turns. Then for two time periods the situation was reversed.

Figure 9 shows the graph of the total number of departures from all sources in each period from time zero to time five. There were five units sent from each source in periods zero and five, 10 units from each source in periods one and four and 15 units from each source in periods two and three. The total number of arrivals at the destination for the network flow solution are also shown in Figure 9. These arrivals are further divided into those entering the destination from the east and west and those entering from the north and south.

In this network the arrivals reach a steady state of seventy for periods 8 to 17. Also during periods 8 to 17 there is an oscillation of the flow quantities arriving from east and west versus those arriving from the north and south. The consistent arrival paths observed in this example are a result of the symmetry of the street network, in terms of travel times from source location, the traffic signal pattern, and the source flow quantities. A further contributing factor is the uniformity of street capacities. It is highly improbable that these four conditions would ever be satisfied in an application. It is to be expected that generally an irregular flow pattern will occur such as that observed in the first example rather than the steady state pattern observed in this example, because street lengths and capacities don't generally match exactly, source input quantities vary from source to source and traffic signals cannot usually be perfectly coordinated as in this case.

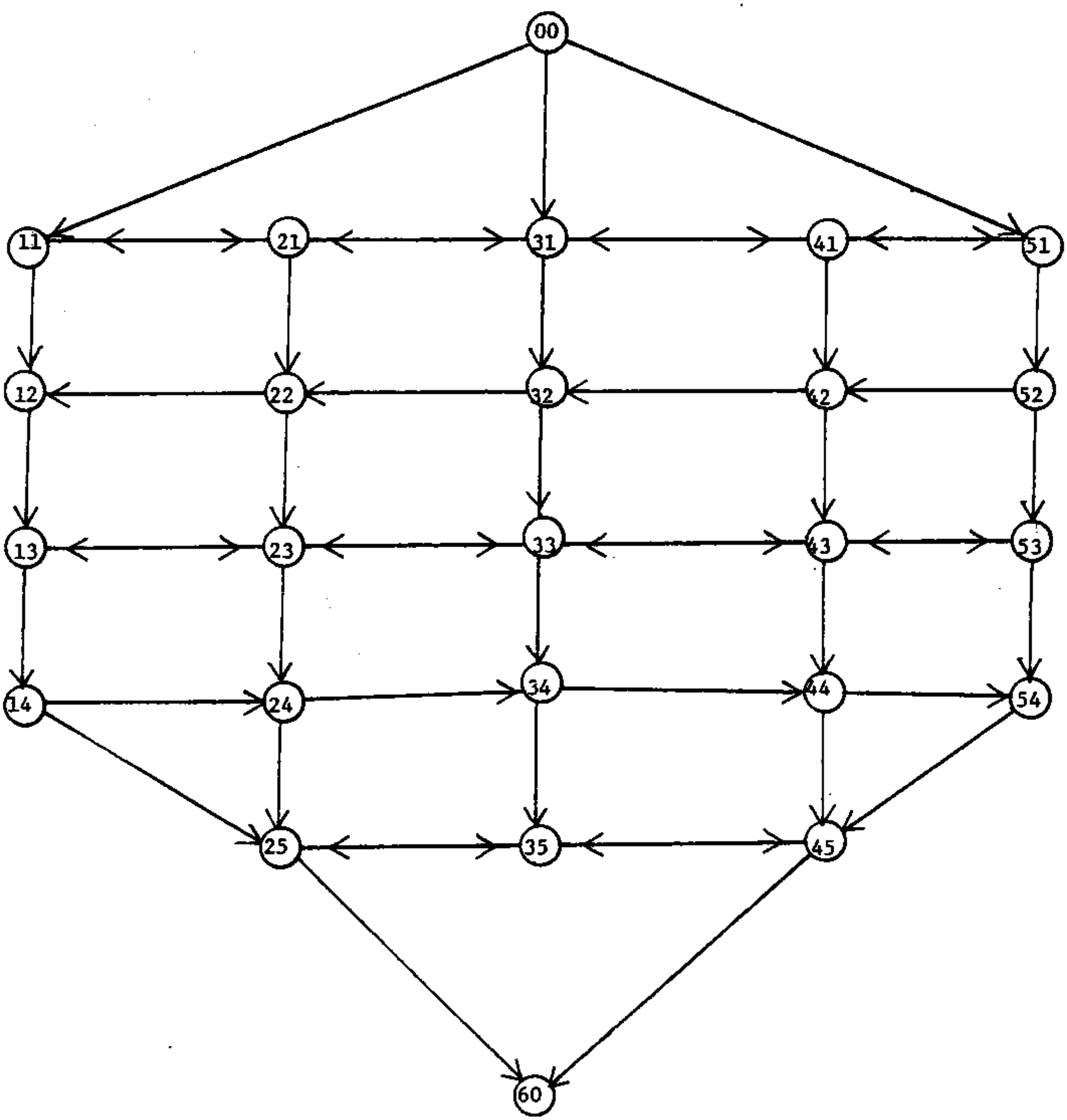


FIGURE 5. Street map for first example.

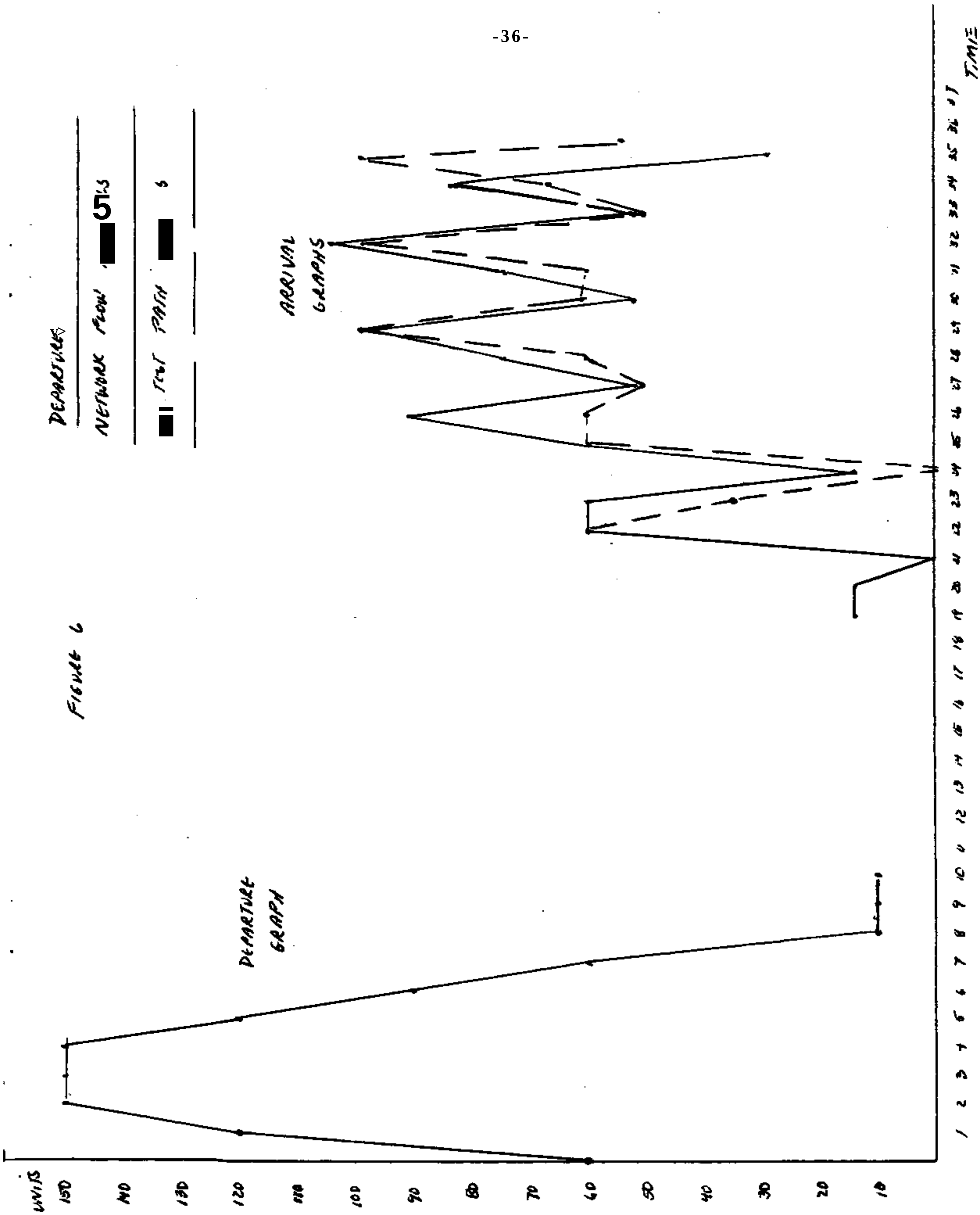
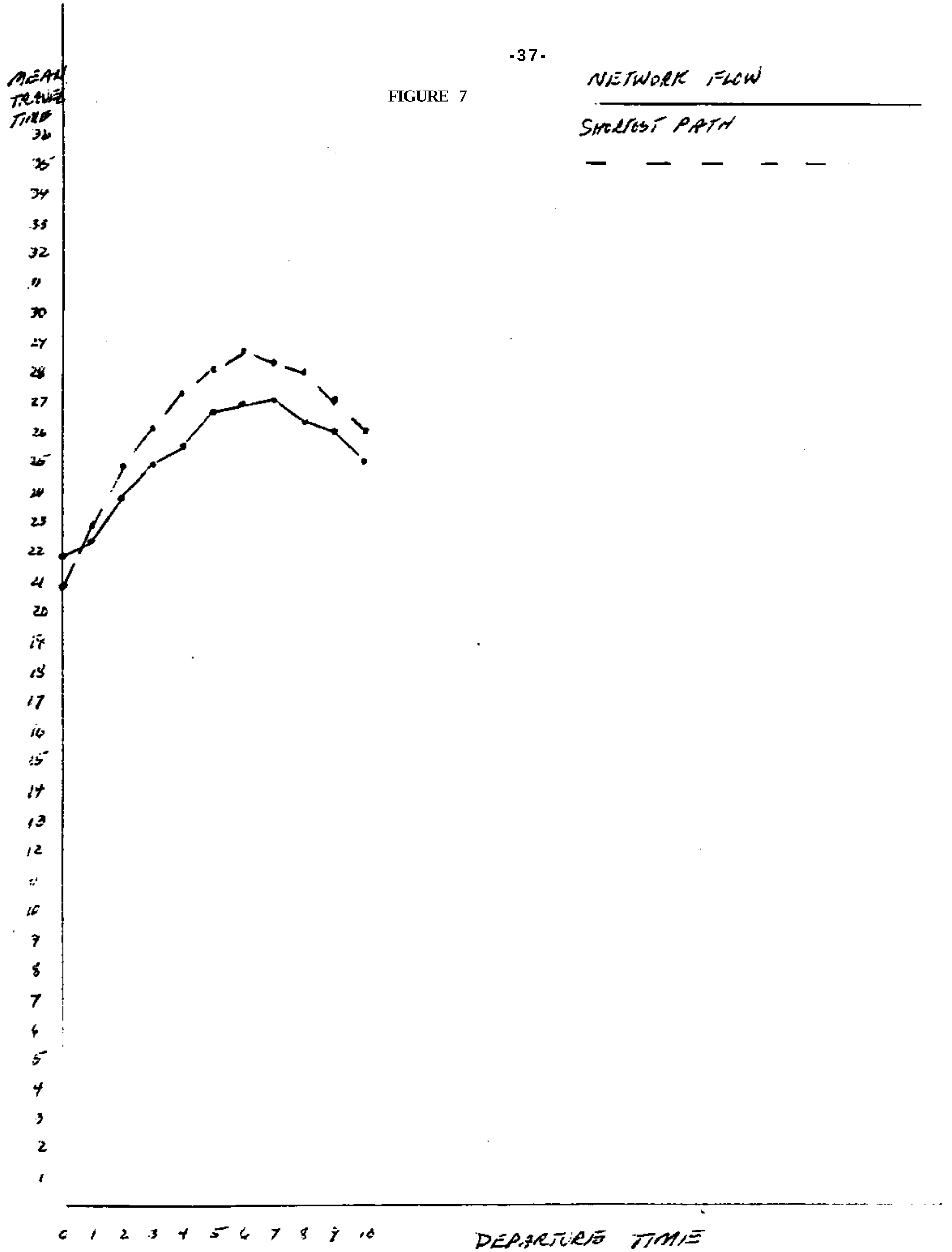

DEPARTURES
NETWORK FLOW [redacted]
TEST PATH [redacted]
ARRIVAL GRAPHS
DEPARTURE GRAPH
UNITS
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
TIME
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37

FIGURE 6

FIGURE 7


MEAN TRAVEL TIME
NETWORK FLOW
SHORTEST PATH
DEPARTURE TIME

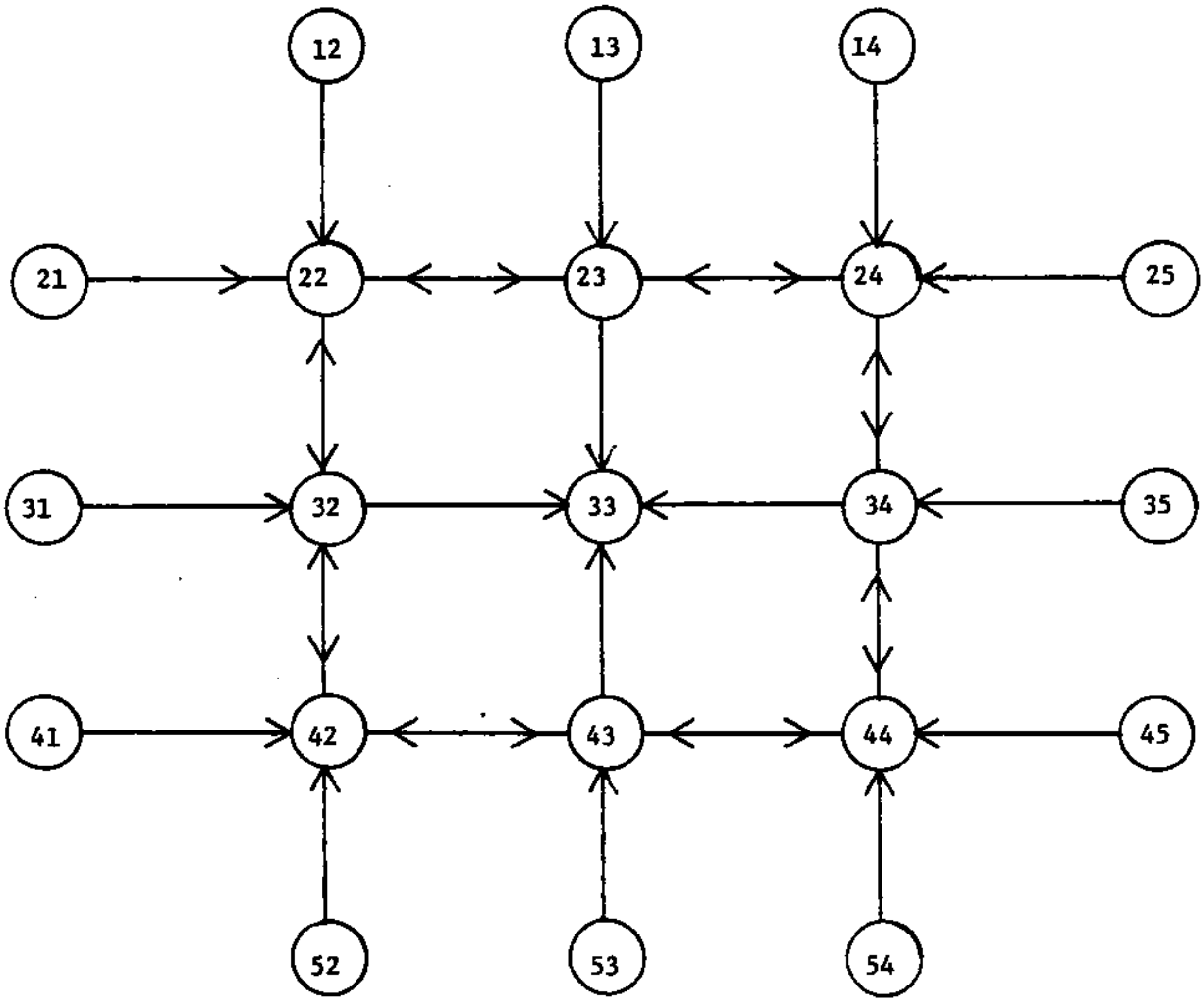
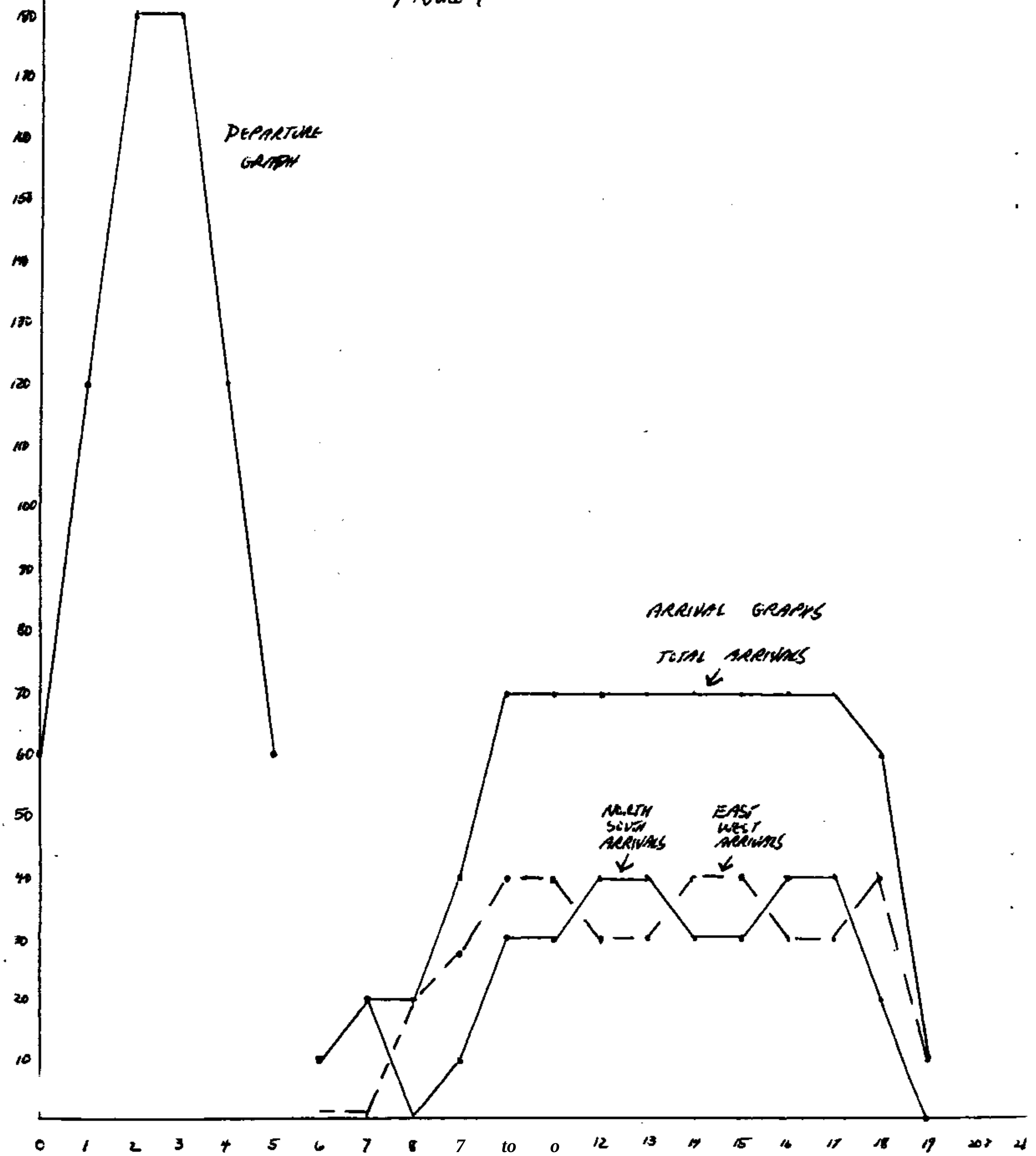


FIGURE 8. Street map for second example.

FIGURE 9


DEPARTURE GRAPH
ARRIVAL GRAPHS
TOTAL ARRIVALS
NORTH SOUTH ARRIVALS
EAST WEST ARRIVALS

APPENDIX A

V. ALGORITHM FOR DISPATCHING TRAFFIC FLOWS TO MINIMIZE USER TRAVEL TIME.

Definitions and Notation:

$V = D \cup N$ the space time node set of G.

$E = R \cup G \cup \mathcal{D}$, the arc set of G.

d is the destination.

(s,t) is the source location of time t.

$S(s,t)$ is the supply of vehicles at the source at time t.

$P_k(t)$ is the kth path found from the source node (s,t) to destination d.

$E[(P_k(t))]$ is the arc set of path $P_k(t)$.

L is the length function where $L(\ell)$ is the length of arc $\ell \varepsilon E$.

$L(P_k(t))$ is the length of the kth path from the source node (s,t) to the destination d.

F is the flow function where $F(\ell)$ is the flow in arc $\ell \varepsilon E$.

$F(P_k(t))$ is the flow in the kth path from source node (s,t) to destination d.

CAP is the capacity function where $CAP(\ell)$ is the current capacity of an arc $\ell \varepsilon E$.

$CAP(P_k(t))$ is the capacity of the kth path

Path capacity is defined as

$$\min_{\ell \varepsilon E(P_k(t))} \{CAP(\ell)\}$$

$W(F(P_k(t)), L(P_k(t)), E(P_k(t)))$ be a vector valued function where the first element of the vector $F(P_k(t))$ is the flow in the kth path from source node (s,t) to destination d, the second element of the vector $L(P_k(t))$ is the length of the same path, and the remaining elements are the arcs in the path $P_k(t)$.

ALGORITHM

STEP 0 INITIALIZATION

Set $T = \{t : S(s,t) > 0\} = \{t_1, t_2, \ldots, t_n\}$ where $t_1 < t_2 < \ldots < t_n$

Set $F(\ell) = 0$ for $\ell \in E$.

Set $k = 1, j = 1$

Go to STEP 1.

STEP 1.

For (s,t,j) find the shortest path in G from (s,t,j) to d. Call this path $P_k(t,j)$. Let $L(P_k(t,j)) = L$. Let $E(P_k(t,j)) = P$.

Send $u = \min\{CAP(P_k(t,j)), S(s,t,j)\}$ units on path $P_k(t,j)$ by doing the following

For $\ell \in E(P_k(t,j))$ set $F(\ell) \leftarrow F(\ell) + u$.
Set $W(F(P_k(t,j)), L(P_k(t,j)), E(P_k(t,j))) = (V, L, P)$.

Update the arc lengths and arc capacities of G as follows:

For all $\ell \in E(P_k(t,j))$ set $CAP(\ell) \leftarrow CAP(\ell) - u$.

IF $CAP(\ell) = 0$ then set $L(\ell) = \infty$

Update the supply at node (s,t,j)

$S(s,t,j) \leftarrow S(s,t,j) - u$.

Go to STEP 2.

STEP 2.

IF $S(s,t,j) > 0$ then $k \leftarrow k + 1$ and go to STEP 1.

If $S(s,t,j) = 0$ then $j \leftarrow j + 1$ and go to STEP 3.

STEP 3.

If $j < n+1$ then $k \leftarrow 1$ and go to STEP 1.

If $j = n + 1$ STOP.

END OF APPENDIX A.

REFERENCES.

[1] M. J. Beckman, C. B. McGuire, and C. B. Winston, Studies in the Economics of Transportation, Yale University Press, New Haven, Conn., 1956.

[2] M. J. Beckmann, "On the Theory of Traffic Flows in Networks," Traffic Quarterly, 21:109-116, 1967.

[3] S. Dafermos and F. Sparrow, "The Traffic Assignment Problem for a General Network," Journal of Research National Bureau of Standards, 75B: 91-117, 1969.

[4] S. Dafermos, "An Extended Traffic Assignment Model with Application to Two Way Traffic," Transportation Science, 5:366-389, 1971.

[5] S. Dafermos, "The Traffic Assignment Problem for Multiclass-User Transportation Networks," Transportation Science, 6:73-87, 1972.

[6] S. Dafermos, "Traffic Equilibrium and Variational Inequalities," Transportation Science, 14(l):42-54, Feb., 1980.

[7] S. Dafermos, "Relaxation Algorithms for the General Asymmetric Traffic Equilibrium Problem," Transportation Science, 14(2):231-240, May, 1882.

[8] S. Dafermos, "The General Multimodal Network Equilibrium Problem with Elastic Demands," Networks, 12(l):57-72, 1982.

[9] M. Florian and S. Nguyen, "A Method for Computing Network Equilibrium with Elastic Demand," Transportation Science, 8:321-332, 1974.

[10] M. Florian, (ed)., Lecture Notes in Economics and Mathematical Systems: Traffic Equilibrium Methods, Springer-Verlag, 1974.

[11] M. Florian and S. Nguyen, :An Application and Validation of Equilibrium Trip Assignments Method," Transportation Science, 10:373-390, 1976.

[12] M. Florian, "A Traffic Equilibrium Model of Travel by Car and Public Transit Modes," Transportation Science, 8:166-179, 1977.

[13] N. H. Gartner, "Optimal Traffic Assignment with Elastic Demands: A Review Part II Algorithmic Approaches,Ir Transportation Science, l(2):192-208, May, 1980.

[14] N. H. Gartner, "Optimal Traffic Assignment with Elastic Demands: A Review Part I Analysis Framework," Transportation Science, 14(2):174-191, May, 1980.

[15] B. G. Heydecker, "Some Consequences of Detailed Junction Modelling in Road Traffic Assignment," Transportation Science, 17(1983) 263-281.

[16] J. K. Ho, "A Successive Linear Optimization Approach to the Dynamic Traffic Assignment Problem/' Transportation Science, 14(4):295-305, Nov., 1980.

[17] J. L. Kennington and R. V. Helgason, Algorithms for Network Programming, John Wiley and Sons, New York, 1980.

[18] F. H. Knight, "Some Fallacies in the Interpretation of Social Costs," Quarterly Journal of Economics, 38:582-606, 1924.

[19] L. Leblanc and K. Farhangian, "Efficient Algorithm for Solving Elastic Demand Traffic Assignment Problems and Mode Split Assignment Problems," Transportation Science, 15(4):306-317, Nov., 1981.

[20] D. K. Merchant and G. L. Nemhauser, "A Model and an Algorithm for the Dynamic Traffic Assignment Problems," Transportation Science, 12(3):183-199, Aug., 1978.

[21] D. K. Merchant and G. L. Nemhauser, "Optimality Conditions for a Dynamic Traffic Assignment Model," Transportation Science, 12(3), Aug., 1978.

[22] D. Nguyen, "An Algorithm for the Traffic Assignment Problem," Transportation Science, 8:203-216, 1974.

[23] R. B. Potts and R. M. Oliver, "Mathematics in Science and Engineering. Volume 90: Flows in Transportation Networks," Academic Press, New York, 1972.

[24] R. Steinberg and W. I. Zangwill, "The Prevalence of Braess" Paradox," Transportation Science, 17(1983) 301-318.

[25] J. G. Wardrop, "Some Theoretical Aspects of Road Traffic Research." In Proceedings, Institute of Civil Engineering, Part II, 325-378, 1952.