DYNAMIC TRAFFIC ASSIGNMENT FOR URBAN ROAD NETWORKS

BRUCE N. JANSON
Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6366, U.S.A.

(Received 29 May 1989)

Abstract—This paper presents a nonlinear programming formulation of the dynamic user-equilibrium assignment problem (DUE) for urban road networks with multiple trip origins and destinations. DUE is a temporal generalization of the static user-equilibrium assignment problem (SUE) with additional constraints to insure temporally continuous paths of flow. In DUE, the full assignment period of several hours is discretized into shorter time intervals of 10-15 minutes each for which trip departure matrices are assumed to be known. This formulation of DUE includes SUE as a special case in which there is only one time interval for the full assignment period. The assumption of steady-state flows allows SUE to have all linear constraints, but DUE requires nonlinear flow continuity constraints. Whereas SUE is typically solved by methods of linear combinations, these methods create temporally discontinuous flows if applied to DUE. A dynamic traffic assignment heuristic (DTA) is presented that generates approximate solutions to DUE in an efficient manner for large networks. DTA is not a convergent solution algorithm for DUE, but was designed instead to produce assignments that approximate the DUE optimality conditions. An overview of alternative dynamic assignment approaches is given, including the limitations of other optimization and simulation approaches. Test results presented in this paper show that DTA generates both static and dynamic assignments that approximately satisfy the user-equilibrium conditions of these problems.

1. INTRODUCTION

Methods of performing dynamic traffic assignment are needed to simulate traffic conditions on urban highways during peak periods of congestion. Valid comparisons of alternative transportation plans under conditions of time-varying travel demand cannot be made within the steady state assumptions of static assignment procedures or the network size restrictions of existing dynamic models. Dynamic assignment results are needed to evaluate regional transportation plans and to examine the effectiveness of alternative traffic management plans during emergencies and special events such as major reconstruction projects. Dynamic assignment models can also be used to investigate whether, and to what extent, urban traffic congestion and related impacts, such as fuel consumption and pollution, can be reduced by shifting work trip departure times. Dynamic assignment procedures will also play a role in the development of real-time in-vehicle route guidance systems.

Many temporal analyses of work trip departure times have been conducted in recent years, but research efforts to date on modeling time-varying traffic flows on congested networks have been mainly directed towards developing closed-form analytical models or simulations of small, idealized networks. Dynamic network models have not been extensively developed or implemented. As stated by Carey (1987):

Almost all work to date on network flows over time has considered only one or two arcs or intersections, has been heuristic, has assumed that the travel time on each arc of the network is independent of the flow rate or volume on the arc, or has used simulation rather than optimization.

Two of the earliest works on this topic were produced by Robillard (1974) and Yagar (1976), although their models were very limited in scope. Merchant and Nemhauser (1978a) formulated a nonconvex system-optimal version of the dynamic traffic assignment problem in which there can be multiple origins but only one destination. Merchant
and Nemhauser (1978b) and Ho (1980) explored ways of solving that problem, for which a globally optimal solution is rather difficult to obtain. Carey (1986) shows that the Merchant and Nemhauser formulation does satisfy a certain constraint qualification needed for Kuhn-Tucker optimality conditions to exist at the optimum. Carey (1987) presents an alternative formulation of the Merchant and Nemhauser problem with only one destination that is convex and can be made piecewise linear for solution purposes.

Other analytical models were developed by Hendrickson and Kocur (1981), Mahmassani and Herman (1984), and Ben-Akiva et al. (1986) to examine the congestion and delay effects of varying work trip departure times. While these formulations have improved our understanding of how dynamic travel demands affect daily traffic flows and impedances on alternate routes, their models are limited to restricted network configurations. None of the above research efforts has produced a model that is compatible with the aggregate transportation planning procedures of U.S. metropolitan planning organizations. Leonard et al. (1978, 1982) and Van Vliet (1982) describe simulation approaches used in CONTRAM and SATURN, respectively, to model subarea traffic schemes in greater detail including intersection movements, queuing delays, and time-varying flows. As reviewed by Van Aerde (1987), both models are prohibitive for large networks in which links and nodes number in the tens of thousands.

This paper presents a math programming formulation of the dynamic user-equilibrium assignment problem (DUE) and describes a dynamic traffic assignment procedure (DTA) that can be applied to large networks to generate approximate solutions. The dynamic traffic assignment problem is defined in this paper as follows: Given a set of zone-to-zone trip tables containing the number of vehicle trips departing from each zone and headed towards each zone in successive time intervals of 10 to 15 minutes each, determine the volume of vehicles on each link of the network connecting these zones in each time interval. Link lengths that determine the paths to which vehicles are assigned are computed on the basis of impedance functions of each link's volume in each time interval. Typical assumptions made in solving conventional transportation planning models, such as methods used to compile trip tables of vehicle equivalents from survey data and the form of each link's impedance function, are also taken as given.

DUE is a temporal generalization of the static user-equilibrium assignment problem (SUE) with additional constraints to insure temporally continuous paths of flow. Previous formulations of DUE presented by Boyce (1989) and Hamerslag (1989) do not insure temporally continuous flows. Linear conservation of flow constraints in SUE are nonlinear in DUE, and the temporal flow continuity constraints in DUE are nonlinear, integer, inequalities. Although the additional constraints greatly complicate the solution approach to DUE, the necessary conditions for dynamic user equilibrium are derived from a qualified statement of DUE in which all temporally continuous paths in the solution to the general problem are predefined.

A simple example is given to show how methods of linear combinations, such as the Frank-Wolfe (F-W) or PARTAN techniques, create temporally discontinuous flows when applied to DUE. Although these methods may converge to approximate solutions, they require large increases in computation and memory for even moderate size applications. For these reasons, DTA was developed to generate approximate solutions to DUE in an efficient manner. DTA does not converge toward a dynamic user-equilibrium solution. Instead, DTA incrementally assigns trips departing in each time interval using a tree-by-tree assignment that anticipates future link volumes. The formula used in DTA to project future link volumes in each time interval will be explained later.

Several measures are used to evaluate the extent to which static and dynamic assignments from DTA satisfy the user-equilibrium conditions of those problems, and the amount of internal variation present in the results. Comparisons of static assignments produced by the F-W and DTA procedures show that DTA compares favorably with F-W in producing SUE solutions, and test cases with dynamic travel demands show that DTA can successfully produce assignments that approximate dynamic user-equilibrium conditions. The memory requirements of DTA are quite manageable for large networks. The memory requirements of a linear combinations approach to DUE would be at least
four times greater than that of DTA. The major computational difference between F-W
and DTA is that DTA requires several shortest path trees to be found and loaded for
each time interval of the full analysis period.

A mathematical formulation of DUE is presented in the next section, and the DTA
procedure designed to produce an approximate solution to DUE is explained in Section
3. Section 4 presents the results of applying DTA to both static and dynamic travel
demands, and examines the degree to which optimality conditions of those problems are
satisfied. Test cases were performed on the well-known Sioux Falls network (24 zones, 24
nodes, 76 one-way links) and a Pittsburgh network containing 30 zones, 372 nodes, and
807 one-way links. Potential applications of the DTA procedure for other uses, such as
trip departure time estimation and real-time traffic management, and the potential use of
parallel computing for large-scale applications are discussed in Section 5.

2. THE DYNAMIC-USER EQUILIBRIUM ASSIGNMENT PROBLEM (DUE)

The dynamic user-equilibrium assignment problem (DUE) is a temporal generalization
of the static problem, of which the static problem is a special case. The static
user-equilibrium assignment problem (SUE) can be formulated equivalently in terms of
either path or link flows, but the path flow formulation allows for a clearer generalization
to dynamic assignment. The following formulation of DUE assumes that the full assign-
ment period of perhaps three hours has been divided into shorter time intervals of 10–15
minutes each. Instead of only one trip matrix for the full period, there must be given a
zone-to-zone matrix of trip departures for each time interval.

A temporal generalization of SUE requires that a time interval superscript be added
to each link flow variable and impedance function. A departure time superscript must
also be added to each element of the trip matrix \( Q \). Superscripts representing time inter-
vals of link use and trip departure must also be added to each path-link variable to
indicate whether trips on path \( p \) departing in time interval \( d \) use link \( k \) in time interval \( t \).
The dynamic user-equilibrium assignment problem (DUE) is given by eqns (1–9)

\[
\text{DUE Minimize } \sum_{k \in A} \sum_{t \in D} \int_0^{Q_t} f_k'(w) \, dw
\]

subject to:

\[
x_k^t = \sum_{p \in P} \sum_{d \in D} h_p^d \alpha_{pk}^{dt} \quad \text{for all } k \in A, \ t \in D
\]  

(2)

\[
q_r^a = \sum_{p \in P} h_p^d \quad \text{for all } r \in Z, \ s \in Z, \ d \in D
\]  

(3)

\[
h_p^d \geq 0 \quad \text{for all } p \in P, \ d \in D
\]  

(4)

\[
\alpha_{pk}^{dt} = (0,1) \quad \text{for all } p \in P, \ k \in K_p, \ d \in D, \ t \in D
\]  

(5)

\[
\sum_{r \in D} \alpha_{pk}^{dt} = 1 \quad \text{for all } p \in P, \ k \in K_p, \ d \in D
\]  

(6)

\[
b_m^d = \sum_{p \in P} \sum_{s \in S_p} f_s'(x_s^t) \alpha_{pk}^{dt} \quad \text{for all } p \in P, \ n \in N, \ d \in D
\]  

(7)

\[
[b_m^d - t \Delta t] \alpha_{pk}^{dt} \leq 0 \quad \text{for all } p \in P, \ n \in N, \ d \in D, \ t \in D, \ k \in A
\]  

(8)

\[
[b_m^d - (t-1) \Delta t] \alpha_{pk}^{dt} \geq 0 \quad \text{for all } p \in P, \ n \in N, \ d \in D, \ t \in D, \ k \in A
\]  

(9)
where,

\[ N = \text{set of all nodes.} \]
\[ A = \text{set of all links (directed arcs).} \]
\[ A_n = \text{set of all links incident from node } n. \]
\[ Z = \text{set of all zones (i.e. trip-end nodes).} \]
\[ P = \text{set of all paths between all zone pairs.} \]
\[ P_{rs} = \text{set of all paths from zone } r \text{ to zone } s. \]
\[ K_p = \text{set of all links on path } p. \]
\[ K_{pn} = \text{set of all links on path } p \text{ prior to node } n. \]
\[ \Delta t = \text{duration of each time interval (same for all } t). \]
\[ D = \text{set of all time intervals in the full analysis period (e.g. 18 ten-minute intervals for a 3-hour peak-period assignment).} \]
\[ x_{it}^l = \text{number of vehicle trips between all zone pairs assigned to link } k \text{ in time interval } t \text{ (variable).} \]
\[ f_k^t(x_{it}) = \text{travel impedance on link } k \text{ in time interval } t \text{ (variable).} \]
\[ h_{pm}^d = \text{travel time of path } p \text{ from its origin to node } n \text{ for trips departing in time interval } d \text{ (variable).} \]
\[ q_{rs}^d = \text{number of vehicle trips from zone } r \text{ to zone } s \text{ departing in time interval } d \text{ via any path (fixed).} \]
\[ h_p^d = \text{number of vehicle trips assigned to path } p \text{ that departed in time interval } d \text{ (variable).} \]
\[ \alpha_{pk}^d = \text{zero-one variable of whether trips departing in time interval } d \text{ and assigned to path } p \text{ use link } k \text{ in time interval } t \text{ (0 = no; 1 = yes) (variable).} \]

The above formulation of DUE assumes that a directed network \( G(N,A) \) is given, where \( N \) is the set of nodes and \( A \) is the set of directed arcs or links. Zones (denoted by the set \( Z \)) are nodes at which trips originate and/or terminate. Each individual path \( p \) is from a given origin zone \( r \) and to a given destination zone \( s \). The set \( P \) of all possible paths between all zone pairs is partitioned into subsets of paths between particular zone pairs (e.g. \( P_{rs} \) is the set of all possible paths from zone \( r \) to zones). The variable \( h_{p}^d \) indicates the amount of flow (i.e. number of vehicles) assigned to path \( p \) from zone \( s \) that departed in time interval \( d \). Flows between other zone pairs can use the same links as \( h_{p}^d \), but these flows are associated with distinct paths and departure time intervals. Equation (2) defines the total flow on link \( k \) in time interval \( t \) to be the sum of flows departing in any time interval on any path that uses link \( k \) in time interval \( t \) in order to formulate the objective function as given by eqn (1). Equation (3) constrains path flows to sum to the proper trip departure totals in each time interval between each zone pair, and eqn (4) requires all path flows to be nonnegative.

The above formulation assumes complete enumeration of all possible paths, such that every path-link variable \( \alpha_{pk}^d \) for all zone pairs is represented in these equations. Equation (5) defines each path-link variable to be a (0,1) variable that indicates whether path \( p \) uses link \( k \) in time interval \( t \) for trips departing in time interval \( d \). Equation (6) allows each path \( p \) to use each link \( k \) on path \( p \), denoted by the set \( K_p \), in only one time interval for each departure time.

Equations (1-5) exactly constitute SUE for the case of a single time interval in which each path-link variable \( \alpha_{pk}^d \) (disregarding time) could be prespecified to either 0 or 1. Since only continuous paths are defined in SUE, no additional flow continuity constraints are needed at the nodes. When actually solving SUE by a method of linear combinations, only paths to be assigned flow are found by the shortest path routine such that complete enumeration of all possible paths is obviated. An important difference to recognize between the static and dynamic problems is that the path-link variables in DUE are predefined constants in SUE.

In contrast to SUE, each path-link variable \( \alpha_{pk}^d \) in DUE cannot be prespecified with a fixed value because the time interval of link use is affected by travel times, which are affected by the link loadings. Each path-link variable in DUE is endogenous, causing
DUE to have nonlinear flow conservation constraints, and requiring DUE to have additional constraints to insure temporally continuous flows. In order to have temporally continuous flows, eqn (7) is added to sum the travel times to each node \( n \) along links in each path \( p \), denoted by the link set \( K_{pn} \). Equations (8) and (9) then force each path to use its links, given by the link set \( K_{pn} \), in the time intervals that are compatible with the travel times to each node.

According to eqns (7-9), links are used by paths in the time intervals that paths reach the tail nodes of links. For 10-minute intervals, interval 1 begins at 0 minutes, interval 2 begins at 10 minutes, interval 3 begins at 20 minutes, etc. If a path reaches a node at the exact beginning of a time interval (to the degree of floating point precision being used), then the solution algorithm can be coded to have the path use that link in that time interval rather than in the previous interval. Note that eqns (7-9) also work for the static case, albeit unnecessary, so long as the duration \( \Delta t \) of the single time interval exceeds the longest path in the network to which trips are assigned. Thus, eqns (1-9) are a complete formulation of both the static and dynamic assignment problems, with the static problem being a special case.

Figure 1 shows how the time interval of link use depends on the time at which a path reaches the tail node of a link. The link number is indicated by \( k \), and the time interval of link use is indicated by \( t \). Link 1 is used by this path in interval 1, although it overlaps into interval 2. Links 2 and 3 are used in interval 2, but link 4 is used in interval 3, since it begins exactly at the beginning of interval 3. The portion of link 1 that overlaps into interval 2 is discussed next.

In both DUE and the heuristic solution procedure explained later, a link is considered to be fully traversed within the time interval that a path reaches its tail node. This approximation is reasonable so long as link lengths are generally much shorter than the time interval duration (e.g. less than 20%). When links are short relative to the time interval, fractions of paths overlapping time intervals may not be significant. Suppose that a dynamic assignment has a mean link length of \( M \) minutes, and a time interval duration of \( N \) minutes. The probability is \( M/N \) that a link of \( M \) minutes will overlap time intervals in any path that uses it. When overlap does occur, the average amount of overlap into the next time interval is one-half the average link length, or \( M/2 \) minutes. Thus, the average amount that used links overlap time intervals is \( M^2/(2N) \), which is the one-half the average link length \( (M/2) \) times the probability that an overlap occurs \( (M/N) \). For example, links of 2 minutes in time intervals of 10 minutes will, on average, overlap intervals by 0.2 minutes or 10% of the link length. The overlapping percent of total trip length, which accounts for trip volumes on the links, was computed for the examples given later and was always found to fall below this formula's estimate. In networks with many O-D pairs, the estimation errors caused by these overlaps are expected to be unbiased if many different paths use each link.

The difficulty of insuring temporally continuous flows arises because we do not know how the time interval will change in the vicinity of each node with respect to flows originating from different origins in different time intervals. For example, link \( k' \) incident to node \( n \) in time interval \( t \) may have a short travel time so that its flow passes onto link \( k \) from node \( n \) in time interval \( t \). However, the travel time of link \( k' \) or links before it may grow longer due to congestion such that its flow can only pass onto link \( k \) in time interval \( t + 1 \). As a result, flows that satisfy constraint eqn (2), but not eqns (8) and (9), can skip over time intervals at any node or even enter nodes later than they exit.
free-flow time = 8 min. 
veh. capacity = 2000 /hr.

free-flow time = 6 min. 
veh. capacity = 2000 /hr.

3000 vehicles in time interval 1.
1500 vehicles each in time intervals 1 & 2.

Fig. 2. Example of temporally discontinuous flows.

A simple example is given to demonstrate the difference in solving DUE with and without eqns (6–9). Figure 2 shows a route from zone r to zone s taken by 3000 vehicles departing in time interval 1 between 7:00 AM and 7:10 AM. Assuming no trip departures during the only other time interval from 7:10 AM to 7:20 AM, then without eqns (6–9) to insure temporally continuous flows, the optimal solution is to subdivide the flow at node n into two parts of 1500 vehicles each. One part is assigned to link k in time interval 1, and the other part is assigned to link k in time interval 2.

Methods of linear combinations (e.g. Frank-Wolfe and PARTAN), which can be applied to problems with all linear constraints, are typically used to solve SUE by combining link volumes without regard to when they occur because they are assumed to occur continuously. Using the BPR impedance function with the free-flow travel times and capacities given for the links in Fig. 2, the F-W method will find the invalid solution to this example described above. Other examples with two or more routes are easily construed to show that methods of linear combinations will create temporally discontinuous flows when used to solve DUE. Once created, these invalid flows will never be eliminated from the solution as the algorithm proceeds.

One additional note is that the impedance functions in the objective function of DUE are not restricted to travel time. Travel times to each node of each path in eqn (7) must be in units of time to insure temporally continuous paths, but the impedance functions in eqn (1) can include travel cost factors other than travel time that affect route choice. The impedance functions in the objective function measure the disutility of travel that the solution attempts to equalize for alternative used routes.

Deriving dynamic user-equilibrium optimality conditions from the above formulation is complicated by the integer variables and nonlinear mixed-integer constraints. Although these optimality conditions cannot be derived from the general formulation with integer unknowns, they can be derived for a given set of path-link integer values that predefine all temporally continuous paths in the optimal solution to the general problem. The optimality conditions of this restricted problem derived for a given set of fixed integer values must be true of the global optimum.

Equation (10) is the Lagrangian of eqns (1–4) with linear constraints because of fixed integer values.

\[
L(X,H,\lambda,\mu) = \sum_{k \in A} \sum_{\omega \in D} \int_0^{x_k} f_k'(w) \, dw + \sum_{s \in Z} \sum_{r \in Z} \mu_{rs} \left[ q_{rs}^d - \sum_{p \in P} \sum_{d \in D} h_p^d \right]
+ \sum_{s \in Z} \sum_{r \in Z} \sum_{d \in D} \mu_{rs} \left[ q_{rs}^d - \sum_{p \in P} \sum_{d \in D} h_p^d \right] + \sum_{d \in D} \tau_p d (-h_p^d). \tag{10}
\]

The optimality conditions are given by eqns (11–13).

\[
\frac{\partial L}{\partial x_k \omega} = f_k'(x_k) = \lambda_k \tag{11}
\]

\[
\frac{\partial L}{\partial h_p d} = \sum_{k \in A} \sum_{\omega \in D} \lambda_k \alpha_{pk} \omega = \mu_{rs} + \tau_p d \quad \text{for all } r \in Z, s \in Z, p \in P, d \in D \tag{12}
\]
\[ \tau_p^d \cdot h_p^d = 0, \quad (\tau_p^d \geq 0) \quad \text{for all } p \in P, d \in D. \tag{13} \]

Over the domain of variable integer values, eqn (10) is nonconvex, and there are many local optima that are inferior to the global optimum. For any set of fixed integer values, the bordered Hessian matrix of eqn (10) is positive definite, which means that there is a unique global optimum with no local optima (Taha, 1982). The Hessian matrix is only positive definite so long as each impedance function is a monotonically nondecreasing function of flow on link \( k \) in time interval \( t \) alone, which is assumed to be true in the above formulation of DUE.

Equations (11) and (12) result from setting the first derivatives of eqn (10) in terms of \( x_d^r \) and \( h_p^d \) equal to zero. The last part of eqn (10) ensures nonnegative path flows and results in a third optimality condition given by eqn (13), which requires \( \tau_p^d \) to be zero if flow is assigned to path \( p \) for departure time \( d \), and nonnegative otherwise. According to eqn (11), the optimal solution has a unique equilibrium impedance for each link in each time interval. According to eqns (12) and (13), for any given pair of zones and departure time, all used paths have the same travel impedance, and any unused path for a given departure time cannot have a lower impedance than an alternative used path.

The optimality conditions for dynamic user equilibrium can be stated very similarly to Wardrop’s (1952) statement of necessary conditions for static user-equilibrium. For trips from zone \( r \) to zone \( s \) departing in time interval \( d \), let \( u_p^d \) be the travel impedance of path \( p \) equal to the left side of eqn (12), and let \( \mu_r^s \) be the equilibrium travel impedance of used paths as indicated by the right side of eqn (12). Also, let \( h_r^s \) be the equilibrium flow on path \( p \) from zone \( r \) to zone \( s \) that departs in time interval \( d \). At equilibrium, all paths from zone \( r \) to zone \( s \) used by trips departing in time interval \( d \) have impedance \( \mu_r^s \), and no unused path between these zones for that departure time can have a lower impedance. These two conditions are given by eqns (14) and (15).

\[ u_p^d = \mu_r^s \quad \text{if } h_r^s > 0 \text{ and } \tau_p^d = 0 \quad \text{for all } d \in D, p \in P, r \in Z, s \in Z \tag{14} \]
\[ u_p^d \geq \mu_r^s \quad \text{if } h_r^s = 0 \text{ and } \tau_p^d \geq 0 \quad \text{for all } d \in D, p \in P, r \in Z, s \in Z. \tag{15} \]

Equations (14) and (15) are equivalent to static user-equilibrium conditions if there is only one time interval for the full analysis period so that the departure time superscripts can be removed from all terms. When SUE is solved by a method of linear combinations, one measure of how close a solution is to equilibrium is the sum of trip impedance differences from shortest path impedances between all zone pairs in each iteration. This measure of convergence is often referred to as the duality gap or impedance gap. In solving DUE, this gap is the sum of trip impedance differences from shortest path impedances between all zone pairs for trips departing in each time interval, which is one way in which the example results presented later are evaluated.

3. A DYNAMIC TRAFFIC ASSIGNMENT PROCEDURE (DTA)

The DTA procedure described next assigns trips departing in each time interval so as to approximately satisfy the conditions for dynamic user equilibrium given by eqns (14) and (15). A key element of dynamic traffic assignment is the way in which vehicle trip flows are tracked through a network on a link-by-link or node-by-node basis in successive time intervals. Flows must be tracked across the network by time interval, and possibly by the destination to which they are headed, depending on the assignment approach used. The major difficulty is that both travel time and network flows are continuous, but they must be discretized in some manner in order to represent them within the assignment process.

A conventional zone-to-zone trip table represents numbers of trips traveling from each origin to each destination within a given analysis period (e.g. 1 hour, 3 hours, or 24 hours). However, a trip table does not typically contain any departure time or arrival time information. Once a trip table has been compiled from survey data, it is unknown
when trips actually travel within the period or whether they both depart and arrive within
the period. If a multi-hour trip table is subdivided into shorter time intervals, then many
trips may not be completed within their departure interval. Hence, trip tables for short
time intervals are actually “trip departure” matrices of trips departing from each zone
and to which zone they are headed.

In assigning trips over time intervals, an option exists whether to load each path
from origin to destination over the set of predicted time intervals, or to load each path
for only the distance traveled in one time interval. Loading only the links traveled by
each path in each time interval would allow trip paths to be revised enroute during the
assignment process. However, that approach requires storing and updating matrices of
trips by destination at each link or node in each time interval, which requires a prohibitive
increase in computer memory for large networks. DTA is designed to produce assign-
ments that approximate dynamic user-equilibrium conditions “in the aggregate" regard-
less of whether route choice decisions and revisions are made enroute or at trip departure
times only.

The DTA approach is to find and load shortest path trees based on projected link
volumes in future time intervals so as to obtain an approximate solution in DUE without
having to store flows by their destinations at the links or nodes. DTA assigns trips
departing in each interval to complete paths in the time intervals that paths are projected
to traverse the links, where link use intervals are based on shortest path travel times
determined from projections of current link volumes into future time intervals. Thus,
trip routes and time intervals of link use are fixed once trips have been assigned to the
network. In addition, only link volumes are stored by time interval, since link impedances
can be computed directly from them during each shortest path tree search.

The reason why current link volumes must be projected into future time intervals for
the purpose of finding assignment paths is that some trips departing in later intervals
from other origins will use links concurrently with trips from the present origin for which
paths are being found. In DTA, link volume projections are made for future time inter-
vales on the basis of current link volumes and ratios of travel demands in the future and
current intervals. The assumption being made here is that satisfactory estimates of future
link volumes can be obtained by multiplying current link volumes by factors that account
for travel demand changes in future time intervals. These projections are made only for
finding paths and are not factored into the actual assignment of link volumes.

The following notation is used to describe the DTA procedure:

\[ x_{ht} = \text{assigned volume on link } k \text{ in the current interval } t \text{ after trips departing in time intervals 1 through } t - 1 \text{ have been assigned, and while trips departing in time interval } t \text{ are being assigned.} \]

\[ x_{ht-1} = \text{assigned volume on link } k \text{ in time interval } t - 1 \text{ (i.e. just prior to the current time interval) after all trips departing in time intervals 1 through } t - 1 \text{ have been assigned.} \]

\[ y_{kt+n} = \text{projected volume on link } k \text{ in interval } t + n \text{ (where } n \geq 0 \text{ and } t + n \in D) \text{ after trips departing in time intervals 1 through } t - 1 \text{ have been assigned, and while trips departing in interval } t \text{ are being assigned.} \]

\[ f_{kt+n}(y_{kt+n}) = \text{projected impedance of link } k \text{ in the current or future time interval } t + n \text{ computed directly as a function of the projected volume } y_{kt+n}, \text{ where } n \geq 0 \text{ and } t + n \in D. \]

\[ Q' = \text{total number of trips departing from all zones in time interval } t \text{ (i.e. total inflow to the network in time interval } t). \]

The superscript of each link volume term always indicates the time interval of link
use, whereas the superscript of each trip matrix term always indicates the time interval of
trip departure. The DTA procedure assumes that all trip departures are known and fixed
for all time intervals and O-D pairs. However, only the trip departure matrix of the
Traffic assignment for urban road networks

current time interval must be stored in random access memory during each iteration of the DTA procedure. The other trip departure matrices can reside in permanent memory until needed.

Steps of the DTA procedure

1. Read network data (link-node incidences, free-flow impedances, and practical capacities for the link impedance functions) adjusted to the time interval duration \( \Delta t \). Initialize the link volumes in each time interval to 0, or read in a set of starting volumes if known. Specify as \( NTREES \) the number of shortest path trees to be assigned trips from each origin zone in each time interval. Initialize the current time interval \( t \) to 0.

2. Increment the current time interval counter \( t \) to \( t = t + 1 \). Read in matrix of trip departures between each O-D pair in time interval \( t \).

3. Randomly select an origin zone for which all trips departing in the current time interval \( t \) have not yet been assigned. Find a shortest path tree from this origin zone to all other zones based on the projected link impedances in the current and future time link intervals. The path search routine finds shortest paths based on link impedances as they are projected to exist during the time intervals in which they are traversed. To minimize array space allocation, each projected link volume and link impedance can be calculated as it is needed in the shortest path routine. The projected link volumes are calculated according to eqn (16), where \( \theta' \) is the percent of \( Q' \) that has not yet been assigned to the network. Define \( \omega_{-1}^{t+n} = Q'^{t+n}/Q'^{-1} \) as a measure of systemwide travel demand and projected traffic volumes in interval \( t + n \) relative to \( t - 1 \), and similarly for \( t + n \) versus \( t \). The projected volume on link \( k \) in time interval \( t + n \) is equal to:

\[
y_{k}^{t+n} = \theta' \omega_{-1}^{t+n} x_{k}^{t+1} \quad (1 - \theta) \omega_{0}^{t+n} x_{k}^{t} \quad \text{for all } k \in A, n \geq 0, t \in D. \quad (16)
\]

Hence, for each link, the current and projected link volumes are estimated as weighted combinations of the final volume assigned to that link in the previous interval \( t - 1 \) and the volume assigned thus far in the current interval \( t \), weighted by ratios of total trip departures from all origins in intervals \( t - 1, t, \) and \( t + n \). Each projected link impedance is computed directly from its projected volume in its impedance function.

4. Assign \( 1/NTREES \) of the trips departing in interval \( t \) from the current origin to the shortest path tree found in Step 3. Store the assigned link volumes \( x_{k}^{t+n} \) by time of link use for all \( n \geq 0 \) and \( t+n \in D \). If all trips departing from all zones in time interval \( t \) have been assigned, go to Step 5. Otherwise, return to Step 3 for the next origin zone.

5. If all departure time intervals in the analysis period have been processed, STOP the program. Otherwise, write out the assigned link volumes for the current time interval \( t \) to a disk file, and copy the link volumes of each future time interval into the link volume array space of the preceding interval in order to economize on array space. Return to Step 2.

In making trip assignments from a given origin, it is unknown how current and future link volumes will be affected by each assignment of trips from another origin. After each incremental assignment of trip departures from a given origin in the current time interval, projections must be made of currently assigned volumes into future time intervals. The technique used in DTA to project current link volumes into future time intervals is to use a weighted combination of current link volumes assigned thus far and final link volumes from the previous interval. This combination is weighted 100% towards the just previous link volumes when paths are found for trips from the first origin, and changes to 100% of the current link volumes after trip departures from the last origin have been assigned. In between, the weight given to current link volumes equals the percent of total trip departures that have been assigned thus far. Other formulas for projecting future link volumes were considered and tested, but none were found to perform as well as eqn (16).

Trips are assigned from origins chosen in a geographically random order so as to
randomize the order of link loadings. A strategy that reduces random variability in the link volumes from one time interval to the next (as opposed to variations between intervals caused by changes in travel demand) is to find NTREES shortest path trees from each origin in each time interval, and to assign 1/NTREES of the trip departures from each origin to each tree. For each time interval, steps 3 and 4 are executed for each origin chosen randomly without replacement, but only 1/NTREES of the trip departures from each origin zone are assigned to the network until NTREES shortest path trees are loaded from each origin. Steps 3 and 4 are repeated in each time interval until all trips from all origins have been assigned in random order.

The DTA procedure was tested on two networks described later with NTREES values of 2, 3, and 4. As expected, the DTA assignments always improved in terms of satisfying the desired user-equilibrium conditions as NTREES was increased. For these networks, loading 3 trees from each origin in each time interval produced significantly better results than loading only two trees from each origin in each time interval. However, increasing NTREES from 3 to 4 achieved a much smaller improvement, which indicates a decreasing marginal rate of improvement for the additional burden of finding more trees from each origin in each departure interval. All of the DTA assignments presented later were generated with NTREES set equal to 3.

The path search routine in DTA finds shortest path trees from each origin on the basis of when each tree is projected to traverse each link, but this does not require comparing link impedances in multiple time intervals. The main computational increase of this routine over shortest path routines for static impedances is that the time interval in which each tree uses each link must be recorded. To avoid storing a link impedance array for each time interval, the impedance of a link is recalculated each time the link is considered for inclusion in a tree. A small proportion of link impedances are calculated more than once per tree because of the label-correcting routine being used, but this adds marginally to the total shortest path computations. One computational savings of DTA over methods of linear combinations is that there is no line search for a link volume combination factor in each iteration.

The DTA procedure must have an initial set of link loadings with which to find paths for trips departing in the first interval of the analysis period. An inefficient strategy would be to have the analysis period begin at 3 A.M., when virtually all links are operating at free-flow volumes such that existing flows on the links are negligible for modeling purposes. Alternatively, we might base the initial link flows on another assignment for the same analysis period. As explained later, the method used in the examples to initiate the DTA procedure was to assign several time intervals of suitable trip departure rates prior to the analysis period in order to obtain initial link volumes.

DTA is much less burdensome than methods of linear combinations in both its memory requirements and computational demands. The DTA procedure requires that one link volume array be stored for as many time intervals as required to accommodate the longest path assigned any trips during execution. For example, with 10-minute intervals, if every path is shorter than one hour, then only seven link volumes arrays are required, adding one for the preceding interval. In applying the F-W algorithm to DUE, two link volumes arrays must be stored for each time interval of the analysis period, and for all time intervals prior to and after the analysis period with trips that extend into the analysis period. A one-hour assignment with 10-minute intervals and trip lengths up to one hour long would require 32 link volume arrays to be stored simultaneously throughout execution of the F-W algorithm (48 arrays for the PARTAN technique).

An appealing property of the F-W algorithm for solving SUE is that each iteration always moves the solution closer to equilibrium (or at least not away from it). Of interest is how DTA assignments with steady-state travel demands compare to F-W solutions after comparable amounts of computational effort. Static assignments from the F-W and DTA procedures are compared for two networks in the next section, with comparisons to observed traffic counts and travel times for the Pittsburgh network. Other tests show the results of applying DTA to the Pittsburgh network with dynamic travel demands.
DTA results for steady-state travel demands are first compared to an equilibrium assignment for the well-known Sioux Falls network having 76 one-way links, 24 nodes, and 24 origin-destination zones (LeBlanc, 1975). Equilibrium assignment results for this network using the standard Frank-Wolfe (F-W) algorithm have been described by numerous authors, including Fukashima (1983), LeBlanc et al. (1985) and Rose et al. (1988). The standard F-W algorithm was also used to generate the equilibrium assignments to which DTA results are compared in this paper.

DTA link volumes will exhibit some variation from one time interval to the next due to the nature of the procedure, even when executed for steady-state travel demands. Thus, for static assignment cases, it is important to (1) examine the degree to which DTA link volumes vary between time intervals, and (2) to compare DTA link volumes over the assignment period to final F-W link volumes. Two measures of variation in DTA link volumes (called APV1 and APV2) are defined for this purpose. The degree to which DTA link volumes vary between time intervals is measured by their average percent variation (APV1) from their means as given by eqn (17)

\[
APV1 = \frac{100}{TX} \sum_{k \in A} \sum_{t \in D} | x_{k}^t - m_k |
\]

where,

- \(APV1\) = the average percent variation between time intervals of DTA link volumes from their means.
- \(x_{k}^t\) = DTA volume on link \(k\) in time interval \(t\).
- \(m_k\) = mean DTA volume on link \(k\) over the time intervals in \(D\).
- \(TX\) = total DTA link volume assigned in all time intervals = \(\sum_{k \in A} \sum_{t \in D} x_{k}^t\).

A second measure of DTA link volume variation (APV2) is used to evaluate the disparity between DTA and F-W link volumes. APV2 is the average percent variation or absolute difference of DTA link volumes from final F-W link volumes. APV2 is computed in the same way as APV1, except that:

1. \(m_k\) in APV1 is replaced in APV2 by the final F-W volume on link \(k\), where each F-W link volume is divided by the number of time intervals in \(D\) so as to represent the same time units as \(m_k\).
2. \(TX\) in APV1 is replaced in APV2 by the sum of F-W link volumes for the full assignment period.

APV2 is expected to exceed APV1, since APV2 indicates variation from values not derived from the DTA link volumes. APV2 is affected by both the between-interval variation of the DTA link volumes and the differences between DTA mean volumes and F-W volumes. Both APV1 and APV2 are reported for each DTA static assignment.

How well these assignments satisfy the desired static or dynamic user-equilibrium conditions must also be assessed. One measure of user equilibrium is the size of the SUE or DUE objective function, which is eqn (1) summed over one time interval in the static case. Equation (1) is applied to final F-W link volumes, whereas equation (1) is computed with the link volumes assigned by DTA in each time interval. For steady-state assignments, eqn (1) could be applied to average DTA link volumes over all time intervals, but this would produce a falsely lower objective function and make the DTA results appear more favorable than otherwise expected. Likewise, the average zone-to-zone trip impedances reported for DTA are always computed on the basis of actual trip paths and link use intervals and not by dividing the sum of link impedances by the sum of link volumes in any interval.
Another standard measure of user equilibrium for an assignment is the duality gap, which is the difference between the sum of assigned trip impedances and the sum of shortest path trip impedances based on the assigned link loadings (Rose et al., 1988). The time dimension of the duality gap (DG) defined by eqn (18) for a dynamic assignment can be disregarded in computing this measure for a static assignment with only one time period

\[
DG = \left(100/TC\right)\left[\sum_{r,s \in A} \sum_{d \in D} x_{rs} f_r(x_{rs}) \right] - \left[\sum_{r,s \in A} \sum_{d \in D} q_{rs} d c_{rs} d \right]
\]  

(18)

where,
- \(DG\) = the duality gap of a dynamic assignment.
- \(q_{rs} d\) = number of trips from zone \(r\) to zone \(s\) departing in time interval \(d\) via any path.
- \(c_{rs} d\) = shortest path impedance from zone \(r\) to zone \(s\) for trips departing in time interval \(d\) through a network of assigned link loadings.
- \(TC\) = total trip impedance if all trips were to use their shortest paths through a network of assigned link loadings = \(\sum_{r,s \in A} \sum_{d \in D} q_{rs} d c_{rs} d\).

The left-most bracketed term of eqn (18) is the total travel time of the solution of DUE or SUE. The right-most bracketed term, which equals \(TC\), is a strict lower bound on the optimal value of the DUE or SUE objective function based on a given feasible solution with no temporally discontinuous paths. \(TC\) is not a strict lower bound on the optimal value of the DUE or SUE objective function if it is based on an infeasible solution with temporally discontinuous paths. An equivalent calculation of the duality gap for any feasible solution is given by eqn (19)

\[
GAP1 = \frac{100}{TC}\left(\sum_{p \in P} \sum_{d \in D} h_{pd} \phi_{prs} \mid u_{pd} - c_{rs} d \right)
\]  

(19)

where,
- \(GAP1\) = the impedance gap of a dynamic assignment.
- \(h_{pd}\) = number of trips assigned to path \(p\) that departed in time interval \(d\).
- \(u_{pd}\) = solution impedance of path \(p\) for trips departing in time interval \(d\) through a network of assigned link loadings.
- \(\phi_{prs}\) = indicator of whether path \(p\) is from zone \(r\) to zone \(s\) (0 = no; 1 = yes).

\(GAP1\) equals the sum of absolute trip impedance differences from shortest path impedances for all zone pairs and departure times through a network of assigned link loadings. Whereas eqn (18) compares the total trip impedance computed from link volumes to the total impedance of all trips assigned to shortest paths, eqn (19) compares assigned trip impedances to shortest path impedances for each separate O-D pair and departure time. \(GAP1\) is divided by \(TC\), the sum of shortest path trip impedances, so as to equal the average absolute percent difference of assigned trip impedances from shortest path impedances.

\(GAP1\) is used in this paper to evaluate the degree of disequilibrium in approximate DUE solutions because temporally discontinuous paths can have lower impedances than the shortest temporally continuous paths. In this case, the impedance difference should increase the gap as it will \(GAP1\) in eqn (19), and not decrease the duality gap as it will \(DG\) in eqn (18). If \(u_{pd}\) is never less than \(c_{rs}\), which is true of any feasible solution to SUE or DUE in which all paths are temporally continuous, then \(GAP1\) can be calculated more easily with eqn (18). This condition, which is always satisfied by F-W solutions to SUE, allows eqn (19) without the absolute value signs to be written as eqn (18).

The duality gap of an equilibrium assignment decreases toward zero, although not
strictly monotonically, as the F-W algorithm is executed for additional iterations. The duality gap equals zero for a true SUE solution in which the impedance of every used path between each pair of zones equals the shortest path impedance. GAP1 also equals zero for a true dynamic equilibrium in which the impedance of every used path between each pair of zones equals the shortest path impedance for a given departure time.

For static assignments found with DTA, we can also evaluate the impedance gap between DTA trip impedances and final F-W shortest path impedances. The impedance gap of DTA trip impedances from final F-W shortest path impedances, denoted as GAP2, is calculated identically to GAP1 except that:

1. $c_{rs}$ from the DTA assignment is replaced by $c_{rs}$ from the F-W assignment, which is the shortest path impedance from zone $r$ to zone $s$ through the network of final F-W link loadings.
2. $TC$ is calculated on the basis of shortest paths through the network of final F-W link loadings.

As was the case with APV2 versus APV1, GAP2 is expected to exceed GAP1, since GAP2 measures differences from path impedances that do not result from the DTA assignment. GAP2 cannot be calculated for dynamic assignments, since equilibrium path impedances from a F-W assignment between all zone pairs for trips departing in time interval are unavailable. GAP2 is superior to GAP1 as a measure of user disequilibrium for a DTA static assignment, while GAP1 provides a low estimate of disequilibrium for DTA dynamic assignments. Note that both GAP1 and GAP2 account for all trip impedances and not just path impedances.

Although the convergence rate of the standard F-W algorithm can be improved, these improvements would not affect the comparisons in this paper because the F-W algorithm was run to high degree of convergence in each case. When applied to the Sioux Falls network to generate assignment results for this paper, the F-W algorithm was halted when the greatest single link volume change was less than 1% between iterations. This degree of convergence for the Sioux Falls network required 76 iterations of the standard F-W algorithm starting from free-flow impedances. Comparisons of F-W and DTA results can also depend on the initial link volumes or impedances used in the F-W algorithm. The outcomes reported in this paper would not be significantly affected by different F-W starting solutions because of the high degree of F-W convergence required in each case. Rose et al. (1988) found that final link volumes for the Sioux Falls network had less than a 0.5% coefficient of variation between solutions when they applied the 1% link volume change stopping criterion to the F-W algorithm with different starting solutions.

Before executing DTA, an important consideration is the time interval duration, or the number of time intervals in the analysis period. A time interval duration of 10 minutes was used in all of the following test cases. This interval duration was chosen after observing that the mean link impedance of the F-W assignment for the Sioux Falls network was 6 minutes. The DTA link volumes show less variation between intervals when the time interval duration is at least 4 to 5 multiples of the mean link impedance. The time interval duration would have to be 24 to 30 minutes to achieve this multiple for the Sioux Falls network. However, most transportation planning networks used in practice generally have shorter links, such as the Pittsburgh network used later in which the F-W mean link impedance is only 0.6 minutes.

As was mentioned in the previous section, the strategy used to obtain initial DTA link loadings was to execute DTA for a few intervals of average travel demand prior to the first time interval of the analysis period. In each execution of DTA, one-half hour (or three 10-minute intervals) were assigned in order to initially load the network. One-half hour was chosen after observing that the procedure required three time intervals of trip departures starting from zero flows for it to settle down to a relatively steady-state assignment for the test networks. Since the majority of trips in both the Sioux Falls and Pittsburgh networks have trip impedances within one-half hour, DTA had to assign
roughly one-half hour of trip departures for initial link loadings to be obtained. In each test case, DTA was run for three 10-minute intervals to obtain initial link loadings, then for the six time intervals of peak-hour trip departures, and for six more intervals to allow all peak-hour trips to clear the network.

To obtain steady-state assignments from DTA, a uniform fraction of the one-hour trip matrix was assumed to depart in each time interval of the analysis period. In addition, each interval of trip departures was assigned incrementally to three shortest path trees (corresponding to an NTREES value of 3) by executing successive tree-by-tree assignments in each interval and loading one-third of each zone's origins to each tree. Hence, DTA could assign no more than three different paths between each pair of zones in each time interval. For this reason, fewer paths are assigned trips by DTA than by the F-W algorithm, and these paths are not assigned in a convergent manner.

Table 1 presents a comparison of the F-W and DTA static assignments for the Sioux Falls network. The total one-hour link volume of the DTA assignment (960) is slightly lower than the total one-hour link volume of the F-W assignment (969). For all trips departing in the one-hour assignment period, the mean DTA trip impedance of 15.84 minutes is 0.88 minutes greater than the mean F-W trip impedance of 14.96 minutes. The objective function values of these two assignments, which are directly comparable, are 50.05 and 50.74 for the F-W and DTA assignments, respectively.

It is important to mention that the F-W algorithm required 23 iterations to achieve a lower objective function value than the DTA procedure. However, the F-W algorithm is much faster per iteration than DTA per time interval. Using the Sioux Falls network and executing both procedures on the same 80386-based microcomputer with all output suppressed, the DTA procedure required 35 seconds to perform 15 time intervals of dynamic assignment, while the F-W algorithm required only 15.3 seconds to execute 23 iterations of equilibrium assignment. Thus, DTA required 2.3 times as much computational effort as the F-W method to generate a comparable one-hour static assignment for the Sioux Falls network. Both programs were coded by the author in Fortran using many of the same subroutines and compiled with the same compiler.

In order to assess the degree of steady-state equilibrium obtained by the F-W algorithm, Table 1 also shows the values of APV1, APV2, GAP1, and GAP2 as defined earlier. The average percent link volume variation between time intervals of the DTA assignment from its own link volume means is 3%, versus 5.88% variation from the final F-W link volumes. The impedance gap of trip paths assigned by DTA from final shortest path impedances in the DTA assignment is 1.51%, versus 2.69% from final shortest path impedances found by the F-W algorithm. The impedance gap of the F-W assignment from its own final shortest path impedances after 76 iterations was 0.2%.

Overall, DTA produced a static assignment for the Sioux Falls network that compared favorably to F-W results. Some tests were performed to see whether changing the random order in which origin zones were assigned by DTA had much effect. Changing the order from one run to the next, or even within each run, by changing the initial seed did not significantly alter the average results so long as each ordering was geographically random. However, simply assigning trips from origins in zone number order from one side of the network to the other did worsen the results.
A network of the Pittsburgh eastern travel corridor was used to examine the performance of the DTA procedure with both static and dynamic travel demands. This network contains 807 one-way links, 372 nodes, and 30 origin-destination zones. DTA was applied to the Pittsburgh network using 10-minute intervals. DTA was run for three time intervals to obtain initial link volumes, and then for 12 additional intervals to obtain a one-hour assignment. The F-W algorithm was run until every link volume varied by less than 3% between iterations, which required 33 iterations.

Table 2 compares the F-W and DTA static assignments for the Pittsburgh network. The total F-W link volume for all 807 links was 411178, but the total DTA link volume was only 408482. The objective function values and mean trip impedances of the two assignments are nearly equal, but the DTA assignment is at an advantage here because of assigning 0.65% fewer link volumes. Although the mean DTA trip impedance is not much greater than one time interval, many of the routes assigned by DTA were between 40 and 50 minutes in length. Table 2 also shows the link volume variations and impedance gap measures for the DTA assignment. The link volume variations are slightly greater than for the Sioux Falls network, while the impedance gaps are much lower. The impedance gap of the F-W assignment from its own final shortest path impedances after 33 iterations was less than 0.1%.

In other tests, the DTA procedure was applied to the Sioux Falls and Pittsburgh networks using 6-minute time intervals. As expected, the DTA procedure did not perform as well with 6-minute intervals on the Sioux Falls network because the average F-W link length was also 6 minutes, and DTA produces better assignments when the time interval duration is at least 4 to 5 times greater than the average link length. The DTA procedure produced very similar static assignments with both 6-minute and 10-minute intervals on the Pittsburgh network in which the average F-W link length was only 0.6 minutes. Thus, these few tests appear to indicate that DTA results are not greatly affected by moderate changes in the time interval duration so long as the interval duration remains several magnitudes greater than the average link length.

Next, the DTA procedure was applied to the Pittsburgh network with dynamic travel demands over the peak hour. Instead of assigning one-sixth of the peak-hour trip matrix to the network every 10 minutes, it was assigned in six successive portions equal to 12.5%, 16.5%, 21.0%, 21.0%, 16.5%, and 12.5% of the trip matrix. These trip departure percentages are based on travel data collected for a study of highway reconstruction impacts in the Pittsburgh eastern corridor (Hendrickson et al., 1982), and compared to percentages reported by Hendrickson and Plank (1984) for work trips commuting to Pittsburgh from the south. These percentages are used in the following example, but they may not be representative of the entire study region or any particular zone within it, since they are based on very limited data.

Comparisons between the Pittsburgh dynamic and static assignments are limited to a few of the evaluation measures. The value of GAP1 for the dynamic assignment is 0.483%, which is only slightly greater than the GAP1 value of 0.354% for the static assignment. The link volume variations (APV1 and APV2), and the second impedance gap measure (GAP2), are not meaningful with time-varying travel demands. The total one-hour link volume assigned in the dynamic case was 404845, which was slightly lower

<table>
<thead>
<tr>
<th>Evaluation Measure</th>
<th>F-W</th>
<th>DTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Link Volume (TX)</td>
<td>411178</td>
<td>408482</td>
</tr>
<tr>
<td>Mean Trip Impedance (min)</td>
<td>11.25</td>
<td>11.23</td>
</tr>
<tr>
<td>SUE or DUE Objective Func.</td>
<td>3884</td>
<td>3873</td>
</tr>
<tr>
<td>Volume Variation 1 (APV1)</td>
<td></td>
<td>3.84%</td>
</tr>
<tr>
<td>Volume Variation 2 (APV2)</td>
<td></td>
<td>7.28%</td>
</tr>
<tr>
<td>Impedance Gap 1 (GAP1)</td>
<td></td>
<td>0.354%</td>
</tr>
<tr>
<td>Impedance Gap 2 (GAP2)</td>
<td></td>
<td>0.551%</td>
</tr>
</tbody>
</table>

Table 2. Summary of Pittsburgh static assignment results
Table 3. Summary of Pittsburgh dynamic assignment results

<table>
<thead>
<tr>
<th>Time Interval</th>
<th>Time of Day</th>
<th>Trip Depart Prct</th>
<th>Number of Trip Departs</th>
<th>Mean Trip Time</th>
<th>Total Link Volume*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6:50</td>
<td>12.5%</td>
<td>3051</td>
<td>10.2 min</td>
<td>1180</td>
</tr>
<tr>
<td>2</td>
<td>7:00</td>
<td>16.5%</td>
<td>4027</td>
<td>10.9 min</td>
<td>1478</td>
</tr>
<tr>
<td>3</td>
<td>7:10</td>
<td>21.0%</td>
<td>5125</td>
<td>13.1 min</td>
<td>1948</td>
</tr>
<tr>
<td>4</td>
<td>7:20</td>
<td>15.0%</td>
<td>5125</td>
<td>14.8 min</td>
<td>2079</td>
</tr>
<tr>
<td>5</td>
<td>7:30</td>
<td>16.5%</td>
<td>4027</td>
<td>12.1 min</td>
<td>1699</td>
</tr>
<tr>
<td>6</td>
<td>7:40</td>
<td>12.5%</td>
<td>3051</td>
<td>10.5 min</td>
<td>1377</td>
</tr>
</tbody>
</table>

*Total link volume in each interval for the 10 screenline locations.

than for either the F-W or DTA static assignment. However, the mean travel impedance of the dynamic assignment was 12.22 minutes, which is one minute greater than either of the two static assignments due to the nonlinear link impedance function causing greater than linear increases in trip impedances with increasing travel demands. This nonlinear increase is also reflected in the higher DUE objective function value of 3922.

Table 3 shows how the total link volume and mean travel impedance varied over the six time intervals of the Pittsburgh dynamic assignment. The third and fourth columns show trip departures in each interval, both in numbers of trips and as percentages of the one-hour trip matrix. The fifth column shows the mean trip impedance of trips departing in each interval, while the sixth column shows the average link volume at ten screenline count locations in each interval. These ten screenline traffic counters are located on freeways and major arterials along a radius approximately 3 miles from the central business district (CBD) and essentially capture all significant volumes of traffic approaching downtown Pittsburgh from the eastern communities. The average screenline link volume is seen to change in proportion to the trip departure rate, which is one assurance that the projection formula based on future trip departures used to estimate link volumes in future time intervals is reasonable.

In a previous study using this network, Janson et al. (1986) found that equilibrium assignment produced reasonable estimates of link volumes but underestimated link impedances, both before and during a major freeway project. That study found that equilibrium assignment link volumes for this network differed from observed morning peak-hour link volumes at the ten screenline locations by an average percent variation of 16.2%. The level of convergence obtained in the Pittsburgh F-W assignment for this paper is greater than that obtained by Janson et al. (1986). As a result, the F-W link volumes found here differed from the observed morning peak hour link volumes along the screenline by an APV of 15.1%.

By comparison, the DTA static assignment link volumes for the full one-hour period differed along the screenline from observed traffic counts by an APV of 20.4%, and from F-W link volumes by an APV2 of 6.0%. The DTA dynamic assignment link volumes for the full one-hour period differed along the screenline from observed traffic counts by an APV of 21.6%, and from F-W link volumes by an APV2 of 7.6%. Thus, DTA produced one-hour link volumes that are marginally different from F-W link volumes and only slightly less accurate when compared to actual counts. Total screenline crossings from each assignment agreed with the total observed screenline counts to within 12%.

One additional comparison is the extent to which these two procedures achieve similar impedances for alternate routes used between a given origin-destination pair of zones. An examination of used trip paths revealed that four alternative routes connecting a residential zone east of Pittsburgh to the downtown CBD were assigned trips by the F-W algorithm. Routes A and B use all arterial streets, while routes C and D use arterials and portions of a major freeway called the Parkway East. Each route is roughly 8 miles, and contains between 19 and 24 links in the coded network. As expected of a highly converged F-W solution, all four routes had the same impedance of 12.3 minutes to one decimal place accuracy. The DTA static assignment resulted in average impedances of
12.1, 12.1, 12.5, and 12.2 minutes for routes A, B, C, and D, respectively, over the one-hour assignment period.

Figure 3 shows the travel times of these four alternate routes for trips departing in each interval resulting from the DTA dynamic assignment. Routes A-D exhibit similar changes during the peak hour, with differences being greatest in the middle of the peak hour when travel demand is greatest and most rapidly changing. The travel times all begin at around 11 minutes (slightly below the SUE travel time of 12 minutes due to the low initial departure rate of 12.5%), rise to between 13 and 14 minutes, and then decrease to their initial levels as travel demand falls off to a lower, steady-state level. Although observed travel times are not available for each 10-minute departure interval, the author's own experience with these routes, having lived in the area for 7 years, is that these times underestimate actual times, but do show relatively valid magnitudes of peak-period variability between intervals. In validating this network for a previous study, Janson et al. (1986) found that a SUE assignment from the F-W algorithm also tended to underestimate actual impedances along routes where travel time runs had been made.

5. CONCLUSIONS AND FUTURE RESEARCH

Data is unavailable for the networks used in this study to validate the DTA results against observed link counts and travel times in each time interval. The author is currently identifying a metropolitan planning organization that can provide network data and 10-minute traffic counts for validation purposes. The F-W and DTA procedures are shown to generate very similar assignments of steady state travel demands, and both procedures produce reasonable estimates of observed one-hour screenline traffic counts. The DTA procedure also produces assignments that approximate dynamic user-equilibrium conditions as measured by the impedance gap.

It is notable that assignments of this quality were obtained from a relatively straightforward heuristic, since DTA is in many respects a very incremental tree-by-tree assignment. The assignment procedure in CONTRAM is even more incremental in that it
sequentially assigns small "packets" of trips between individual zone pairs (Taylor, 1989). Comparisons have not been published between equilibrium assignments obtained from methods of linear combinations and very incremental tree-by-tree assignments in which link impedances are frequently updated after loading small numbers of trips from randomly selected origins.

DTA test assignments evaluated in this paper were made with the usual BPR impedance function. Research has been conducted to test whether other functions, such as the Davidson function, may be superior when used in aggregate assignment models (Boyce et al., 1981). Hungerink (1989) suggests a modification to the usual link impedance function as one method of approximating queuing delays in capacity-restrained assignments so that link volumes in excess of capacity affect the impedances of inflow links. Such impedance function alterations might be considered for dynamic assignment procedures, since travel times directly affect the temporal incidences of competing flows on links common to paths from different origins.

Previous formulations of DUE presented elsewhere for multiple origins and destinations did not prevent temporally discontinuous flows from entering the solution. An example was given earlier of how easily methods of linear combinations cause such flows to occur when applied to this problem. Hamerslag (1989) generated dynamic assignments by applying the Frank-Wolfe algorithm to a small network, but did not evaluate the solution's optimality or temporal continuity. Janson and Zozaya-Gorostiza (1987) show that the F-W algorithm introduces cyclic flows to an assignment that retard its convergence, and this problem will be compounded by the creation of temporally discontinuous flows in dynamic assignments. Any approximate solution to DUE generated by approaches such as DTA or an adapted F-W algorithm must be evaluated for its degree of optimality and temporal flow continuity.

Further research can undoubtedly lead to improvements in the DTA procedure or a similar approach. One possibility is to use DTA to provide a good initial assignment to a F-W algorithm for further improvement, since the F-W algorithm will tend to create temporally discontinuous flows early in the solution process when link loadings are very different from user-equilibrium. The DTA procedure can also be integrated with other travel forecasting procedures. For example, a way of estimating work trip departure times by fitting DTA link volumes to observed traffic counts in successive time intervals has also been described by the author (Janson and Southworth, 1989).

Whereas linear combination methods would require at least four times as much random access memory as DTA, the major computational increase of DTA relative to these methods is the number of path trees found and loaded. Computational advances, such as parallel and high-speed computing, will make it possible to run DTA on large networks in essentially real-time and to compare its link loadings to observed traffic counts as they are being recorded on selected links. Thus, DTA may be one approach to implementing real-time traffic assignment and route guidance systems on urban transportation networks.

Acknowledgements—The author appreciates the helpful insights of Frank Southworth at Oak Ridge National Laboratory during the early development of this paper. The author is also grateful for the comments of several anonymous reviewers.

REFERENCES


