Program, Inputs and Calibration of the Direct Traffic Estimation Method

GEOFFREY J. H. BROWN, Washington, D.C., and
ALBERT H. WOEHRLE, JR., Tri-State Transportation Commission

A new approach to traffic volume estimation is being developed by the Tri-State Transportation Commission. The Direct Traffic Estimation Method requires as inputs a coded highway network and a file of trip ends per unit area, all items of data being identified by X-Y coordinates. Output data available include link volumes, turning movements, and minimum cost trees.

The general concepts of the method are reviewed. The inputs and methods of computation are described and the initial calibration results are presented.

•THE Tri-State Transportation Commission is responsible for long-range land use and transportation planning in the 28 county New York Metropolitan Region. While the region contains 18 million residents in its 8, 000 square mile area, over 16.3 million persons live in the continuously urbanized 3, 600 square mile core (the cordon area).

In common with all urban transportation studies, Tri-State has the task of making facility traffic volume forecasts for given land use distributions and transportation networks. Typically, this work takes two distinct forms: (a) mainline volume estimates are made for various land use—transportation plan alternatives; and (b) once the basic network has been adopted, most of the subsequent work is involved with determining the effects of local changes in the network and in the land use patterns.

In this latter project-oriented work, full-scale network-wide traffic assignments are rarely justified, since the effects under study (i. e., changes in traffic volume) are usually limited to that part of the region in which the changes have taken place. Furthermore, in dealing with a region the size of Tri-State, a full-scale assignment is very costly. For this reason, and in order to permit an arbitrary degree of precision in describing the region, a new technique of traffic estimation was developed.

This new technique, Direct Traffic Estimation, permits traffic estimates to be made on any or all links in the network during a single run of the program. It thus eliminates the need to make region-wide zone to zone movements to test the consequences of localized network changes such as the addition or deletion of an interchange. Another reason for the decision to develop this approach to traffic estimation was the desire to eliminate the lumpiness in link volumes that results from placing all trips originating in a zone on the network at one point. Each square mile might be considered to be a "zone" using this approach.

Within limits, the program is indifferent to the size of the region and to the number of links in the network. Areas of special interest may therefore be described in a very fine level of detail, with a consequent improvement in sensitivity.

Paper sponsored by Committee on Origin and Destination and presented at the 47th Annual Meeting.

METHOD

The method of Direct Traffic Estimation was developed by Morton Schneider during 1965-1966 and has since been implemented as a computer program for the IBM System/360.

The traffic volume at any point in a highway network is derived as a function of the cost of travel and the density of trip-ends in the vicinity of the point of interest. Vehicles on the link of interest are taken to distribute themselves among possible destinations according to the function

$$dV = a \int_0^\infty \frac{1}{K^2} e^{-\left(KC + \frac{a}{K}\right)} dK \, dT \qquad (1)$$

in which dT is some increment of trip-ends at a distance (travel cost) of C cents from the point of interest. The constant a is related to the average trip length in the region and is measured in cents^{-1}, and K is the constant of integration. (This formula supersedes the distribution function given in reference 1 and its derivation may be found in reference 2.)

The two-way average daily traffic volume at the point of interest is obtained by integrating the function (Eq. 1) over various domains around the point of interest and then by solving

$$Q = \frac{I_n I_s{}' + I_s I_n{}'}{I_n + I_s}$$

in which the I's are the "domain integrals." The domains themselves are defined as follows. First, the region is divided into two domains by a line through the point of interest such that all points in one domain are most easily reached by going in one direction on the link of interest, while points in the second are most easily reached by going in the other. To put it another way, if in Figure 1 a vehicle is moving north on the link of interest then its destination lies somewhere in domain N. Conversely, if it is moving south it is destined for a point in domain S.

In Eq. 2, I_n and I_s are the results of integrating Eq. 1 over each main domain in turn. The quantities $I_s{}'$ and $I_n{}'$ measure the effect of links which compete with the link of interest for traffic crossing the main domain boundary. As the effect of competing links increases, I'_s and I'_n and consequently Q, becomes smaller. Conversely, if the main domain boundary is some physical barrier, such as a river, crossed only by the link of interest, then the competition effect is zero and I'_s and I'_n assume their maximum values of I_s and I_n, respectively. The traffic volume Q then becomes

$$Q = \frac{2 I_n I_s}{I_n + I_s}$$

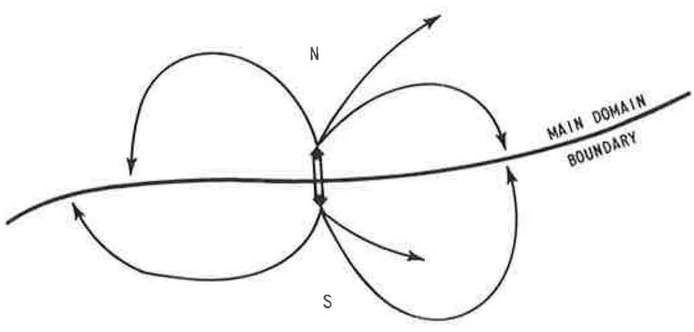

Figure 1.

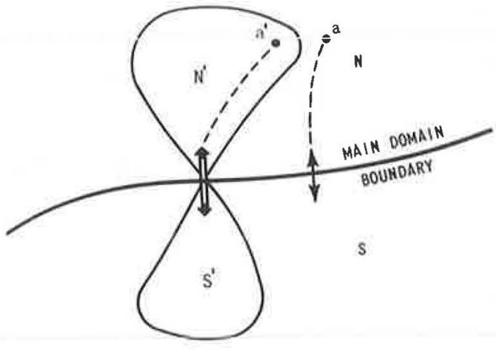

Figure 2.

The domains over which I_n' and I_s' are computed (referred to as prime domains) define the areas in the main domains which contribute traffic uniquely to the link of interest. A point is said to be in the north prime domain if and only if the link of interest lies on the least cost path between the given point and every point in the south main domain. The south prime domain is similarly defined (Fig. 2).
It follows that the prime domains exist only if the link of interest offers some travel cost advantage over competing links which cross the main domain boundary. It is seen from Eq. 2 that unless at least one prime domain exists, the traffic volume on the link of interest is zero.

A rigorous derivation of Eq. 2 is given elsewhere (1). Here, an attempt is made only to show that the traffic volume formula behaves in a reasonable way.

MAIN FEATURES OF THE PROGRAM

The computer program by which the method is implemented uses two basic files of input data. The first of these is a description of the regional highway network and is used to determine a travel costs and domain boundaries associated with each volume estimate. The second file is a description of the trip end density throughout the region, given in terms of a square mile grid, and is used in the computation of the domain integrals.

The program provides the user with two important options:

1. The number of links to be estimated is completely variable and directly determines the running time of the program. In other words, the user pays only for the output information he needs.

2. The area for which network and density data are read by the program may be a rectangle of any size (subject to core storage limitations). This allows the user to maintain input files containing large (unlimited, in fact) amounts of detailed information, but to cause the program to read only those data items pertinent to the problem at hand.

It may be noted here that the distribution function (Eq. 1) behaves in such a way that it is usually unnecessary to consider data much further away than 30 miles from the area in which estimates are being made, and that satisfactory results can often be obtained by limiting the input data to an even smaller radius.

In addition to computing link traffic volumes, the program is able to compute the turning volumes at each end of the link (i. e., the way the link volume splits among the connecting links at each end). An optional output for each specified link is the minimum-cost tree, originating at the mid-point of the link, which is used to determine the travel costs and domain boundaries associated with the link.

The program is currently being run on System/360, models 50 and 65 with 256k and 512k bytes of core storage, respectively. Table 1 summarizes the major statistics for each version of the program.

INPUT DATA FILES

The input data files to the program (DTE) are contained on a single reel of tape, which is prepared off-line from highway network and trip-end inventory files.

For each link in the highway network there is a single record in the network file which includes the X-Y coordinates of each node (link network level code), link cost, a flag to indicate if the link is one-way, and a flag to indicate if turning movements are to be computed for the link. The link coordinates are given to the nearest one hundredth of a mile and may range from 0 to 255.99 miles.

TABLE 3

CORDON AREA RANDOM SAMPLE

Ratio of $\frac{\text{DTE}}{\text{Ground count}}$ to Auto VMT

Function	Level 1	Level 2	Level 3	Total
K = 0.035	0.78	1.26	2.06	1.16
a = 0.027	1.04	1.37	2.33	1.37
a = 0.035	0.79	–	–	–
a = 0.040	0.68	1.07	1.69	0.98
a = 0.045	0.64	0.99	1.48	0.90

The network level code identifies the relative importance of the route of which a link is a part. All major activity centers are connected by Level 1 facilities. The Level 1 system includes all limited access roads. Level 2 highways include other important routes and, as a minimum, bisect the spacing between Level 1 facilities inside cordon area. Level 3 facilities include everything else that was inventoried and are generally limited to the cordon area. Approximately 30 percent of the coded mileage within the region is Level 1, 30 percent Level 2, and the remainder Level 3.

To allow for capacity restraint (as yet not programmed), link costs are broken into two components: initial time cost and fixed cost. The initial time cost reflects the time to travel the over-the-road length of the link at a speed which is a function of the posted speed limit and the number of signals. Time is converted to cost at 2.5 cents/min. Fixed cost is taken to represent accident and operating costs at the rate of 3.5 cents/min plus all link tolls. The limits of the two link cost components are 255 cents for time and 63 cents for the fixed element. Thus, the highway network coding is very similar to the coding for all commonly used assignment programs.

The trip-end density file is based on a mapping of the region into a set of one-mile grid squares. Each entry in the file represents a square area of ¼, 1, 4, or 16 sq mi, and is identified by the X-Y coordinates of the S W corner of the square. The file entry also contains the average trip end density in trips per square mile of the square, and a square size code.

The provision for varying sizes of square, permits the user to choose the square size appropriate to the particular part of the region under consideration. In the Tri-State region, Manhattan is described by ¼-sq mi entries, the remainder of the cordon area by 1-sq mi entries; the band between the cordon and the region boundary by 4-sq mi entries, and the area beyond the region boundary (where necessary) by 16-sq mi entries.

The last file on the input tape is a table of values of the distribution function (Eq. 1) for travel cost values up to $10.24. This table is prepared by a separate FORTRAN program.

METHOD OF COMPUTATION

The computation of the traffic volume on a single link entails summing up (integrating) the contribution to each of the four domain integrals of each trip-end density square in the region. The amount of this contribution is determined by multiplying the trip ends in the square by the value of the distribution function corresponding to the cost of travel between the square and the link of interest.

The first step in this process and one which is done once only for each run of the program, is to associate each density square corner with the (up to four) closest nodes in the highway network. In this way, when the minimum cost tree is built through the highway network, from the link of interest, the minimum travel cost to each density square corner can readily be found.

During the building of the least-cost tree, the main domains are defined. A node is defined (flagged) as being in the north domain if its minimum path enters the link of interest via its north node, and similarly for south domain nodes. At the completion of tree-building, every node in the network belongs to one domain or the other. Certain links have one node in each domain, and must, therefore, straddle the main domain boundary (Fig. 3). The next step is to determine the boundaries of the north and south prime domains.

Each of the boundary-crossing links (other than the link of interest), competes with the link of interest for traffic across the main domain boundary. To determine the extent of this competition, a second tree is built simultaneously from the midpoint of

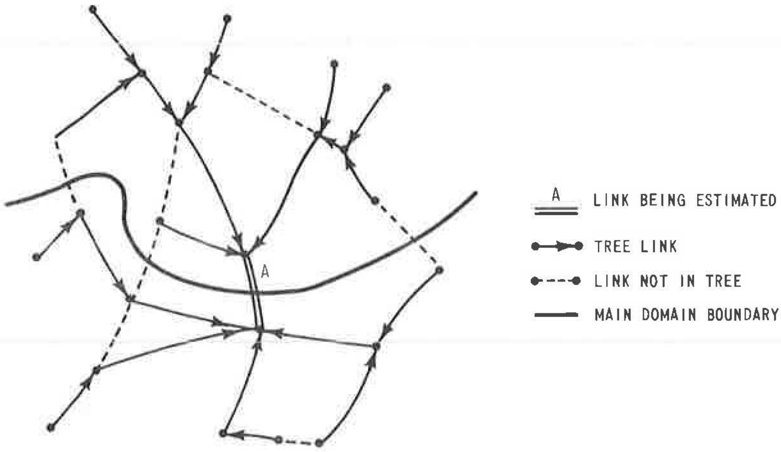

Figure 3.

each boundary link. This has the effect of changing the path values of some nodes in the original tree while leaving others the same. Those nodes whose path values are unchanged during this second tree cycle are, by definition, in the prime domains.

When the domain of each network node has thus been established, the domain integrals are computed. For each trip-end density square the nodes yielding the least travel cost to each corner of the square are found. The trip ends in the square are divided among the four corners, and the contribution of each corner and the integral to which it contributes are then determined.

CALIBRATION AND OPERATION

The initial program version used a trip decay function of the form e^{-KC} instead of the Bessel function shown in the preceding sections. Based on a random sample of 75 links, a value of 0.035 was chosen for K. Link volume estimates were run for all links falling within seven areas, each inside the cordon line and 14 by 20 miles in size. Link volume estimates were compared to ground counts and mapped for review. VMT's were summarized by network level (Table 2).

All of the VMT summaries by network level showed the same pattern. Level 1 was low, Level 2 moderately high, and Level 3 quite high. The results on Levels 2 and 3 are partly explained by the fact that DTE is estimating a component of local street travel which is not represented in the ground counts. Correcting for local streets brings the total line of Table 2 to 0.73, 0.092, 1.07, and 1.10, respectively. Except for the lowness of Level 1, these are agreeable totals.

TABLE 2

Ratio of $\frac{\text{DTE}}{\text{Ground count}}$ Auto VMT, K = 0.035

Area (14 × 20 mi)	Level 1	Level 2	Level 3	Total
New Haven, Conn.	0.85	0.98	1.34	1.01
Norwalk, Conn.	0.85	1.57	2.11	1.27
Oyster Bay, N.Y.	0.78	1.51	2.04	1.11
White Plains, N.Y.	0.71	1.75	2.14	1.08
Newark, N.J.	0.78	1.66	2.35	1.28
New Brunswick, N.J.	0.53	0.89	1.38	0.84
Long Branch, N.J.	0.25	0.84	1.21	0.74
Total	0.73	1.33	1.93	1.10

The areas summarized were also rerun multiplying the time cost component by 1.2 and the fixed cost component (including tolls) by 0.75. This had little overall impact on the results. It was hoped that this would raise Level 1 volumes.

At this stage of development, considerable experimentation was done to test the sensitivity of the system to changes in network extent or service levels. Increasing speeds or the extent of the expressway system increased total VMT, since all trip ends could be reached with less cost. Estimated volumes across screen lines on Staten Island with and without proposed highway facilities showed constant total volumes. Adding the proposed expressways had the impact of placing the bulk of VMT on the expressways and the arterials leading to the interchanges. Throughout this phase of the testing, DTE showed a very logical response to network changes.

The e^{-KC} curve with a K = 0.035 that had been used has the characteristic of virtually eliminating the contribution to the main and prime domain integrals of trip ends more than $1.00 away from the link of interest. Since facilities providing high levels of service were being underestimated and low level facilities were being overestimated, it was concluded that the shape of the decay function needed adjustment. The contribution of nearby trip ends should be reduced while the impact of those further away needed increasing. This led to the selection of a Bessel function to replace e^{-KC}.

Figure 4 shows the shape of the e^{-KC} function for K = 0.035 and Bessel curves for several values of the parameter a (see Eq. 1). It is expected that the final value of a will be in the neighborhood of 0.03 to 0.04. While the curves are shown for a cost range of up to $1.20, the program utilizes the curves up to a cost of $10.24. The trip contribution ratio is multiplied by the trip ends in a square, dT, to yield dV in Eq. 1.

A more inclusive sample of links was chosen for the calibration of the Bessel function. The number of links drawn by network level was made proportional to the cordon area mileage by network level. No links with tolls were included. A comparison of calibration results for various values of a, and K = 0.035, for this sample is given in Table 3—310 links are represented in each run, except for a = 0.035 for which only the sample's 85 Level 1 links were estimated.

The Bessel a = 0.040 had about the same results as K = 0.035 except that the high error of Level 3 was reduced considerably. Changing the function will not by itself solve the problem of increasing the accuracy of link volume estimates.

Each link volume estimate is statistically independent of the estimates on adjacent links. It depends on a reasonable description of network extent and cost, the trip end matrix, and proper definition of main and prime domains. Close examination of the outputs has shown that the relative magnitudes of the domain integrals can shift from link to link along a route. Thus, the sum of I_n and I_s will be fairly constant in a locale, but the split of this total between I_n and I_s can vary. Prime domains can also be quite sensitive to fluctuation.

The sum of I_n and I_s is also a useful measure of accessibility. Such measures have been prepared using trip ends, population, and floor space in the density matrix.

Effort is now concentrated on developing means of stabilizing the program's boundaries. Major differences in the boundaries result from very minor differences in accessing critical nodes during tree building. This must be accomplished before a final value of the a or K

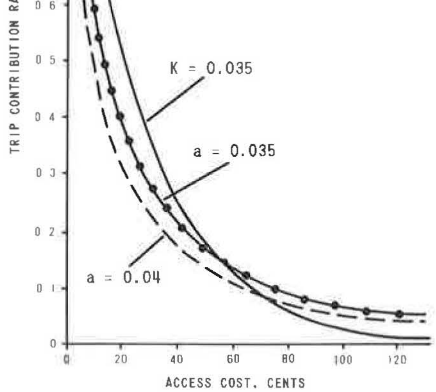

Figure 4.

TABLE 1
PROGRAM STATISTICS

Item	Model 50	Model 65
No. of 2-way links	5000	12500
No. of nodes	4095	8191
No. of density areas	5000	9135
Running time per link (sec)	5	3.3

parameter can be selected. The home interview file is also being processed to derive an empiric trip contribution ratio curve using trips that reported the use of specific East River crossings.

In making traffic volume estimates in response to highway department requests, the procedure that has been developed involves posting of I_n, I_s, I_n', and I_s' for each link. These values are smoothed and made consistent from link to link before mainline volumes are calculated.

DISCUSSION

The DTE program has been made operational and can accommodate very large networks. While the analyst can work with the outputs to make reasonable traffic volume estimates, more work needs to be done to improve the estimating capability and to reduce the amount of output manipulation. This work continues.

The reader will recognize that the principal elements of traffic assignment and distribution are present in DTE. Network description, trip ends by area, minimum path tree building and distribution (or decay) functions are a part of the working tools of any approach. What is new in this work is the combination of distribution and assignment into a single step. The use of the coordinate system provides great flexibility in the Tri-State application of the method, but is not a necessary condition for use elsewhere.

ACKNOWLEDGMENTS

No undertaking of this magnitude is possible without the support and participation of many individuals. This list must start with J. Douglas Carroll, Jr., William L. Mertz, and Edward F. Sullivan who provided many ideas and encouragement. Programming was done by G. J. H Brown, Malvin Ceren and John Bennett. Density inputs were prepared by Harold Deutschman and Stanley Feinsod. The networks were coded under the direction of A. Woehrle and Mark Roskin. Throughout all this permeates the spirit of Morton Schneider.

The preparation of this report was financed in part through Federal funds made available by the U. S. Department of Transportation, Federal Highway Administration, Bureau of Public Roads and the Federal grant from the Urban Renewal Administration of the U. S. Department of Housing and Urban Development, under the planning assistance program authorized by Section 701 of the Housing Act of 1954, as amended and in cooperation with the three states of Connecticut, New Jersey and New York.

REFERENCES

1. Schneider, Morton. Direct Estimation of Traffic Volume at a Point. Tri-State Interim Technical Report 4019-1320.
2. Schneider, Morton. Access and Land Development. Paper prepared for HRB Conference on Urban Development Models, June 1967 (HRB Special Report 97, 1968).