An Iterative Assignment Approach to Capacity Restraint on Arterial Networks

ROBERT SMOCK, Detroit Area Traffic Study

• ONE OBJECTIVE of contemporary traffic planning research is to devise a method for the computer assignment of interzonal volumes to major street networks in a way that respects the capacity of the streets. The Detroit Area Traffic Study has developed a program for a medium-size computer that represents one solution to this problem. Before describing the new procedure, this paper discusses some of the background of the problem, and, after describing the procedure, the paper mentions the place for arterial assignment in the total process of transportation planning.

It has been common practice for transportation planners to consider the results of an assignment of partial interzonal volumes to a single set of best paths through a proposed freeway network. (That part or percentage of each interzonal volume to be assigned to the freeway network is determined by a "diversion curve" which causes the assigned percentage to vary with the varying merit of a freeway route as compared to a non-freeway route.) (1). Volumes thus assigned have two uses. They aid in providing economic justification for the proposed network when they indicate appropriately high future traffic volumes on it. They also aid in the geometric design of the freeway suggesting the number of needed lanes.

Volumes thus assigned have definite limitations with which planners are familiar, and for which they must compensate in some way. Freeway "desire" assignments have been utilized regardless of their shortcomings because there has been no better way to perform the essential task of estimating the volume to be expected on the various parts of proposed facilities. The nature of their chief shortcoming may be noted when "desire" assignments are described.

In this paper, any assignment which places all of each assigned interzonal volume on a single "best" path will be called a "desire assignment." (Whether or not the assignment first discards a part of each volume on the basis of a diversion curve, the main point is that only one set of best paths is computed. A "best" path is that series of links which constitutes the route through the network for an interzonal trip which is fastest, or shortest, etc.) The fundamental limitation of such an assignment for planning purposes appears when volumes thus assigned exceed the capacities of parts of the tested network. The volumes on all other parts of the network then become questionable because it is not possible to anticipate the consequences for the whole network of the diverting of excess volumes from the over assigned parts.

Detailed information suitable for either economic justification or geometric design, therefore, cannot be provided with guaranteed reliability by desire assignments, if parts of the network are assigned volumes beyond their capacity. The efforts to develop more elaborate assignment procedures have not arisen from this problem alone, however, but also from the problems involved in linking a freeway network to a system of existing or proposed major streets. There can be no doubt that an existing street system will have to have its capacities adjusted upward in some places and perhaps downward in other places, if it is to link efficiently with a new freeway network. These problems call for research to create a computer program that will assign realistic traffic volumes to large arterial networks that include both freeways and other major arteries. In addition to helping with the linkage problem, such a program would have the value of providing estimates of future traffic on surface major streets so that their improvement might be better planned independently of freeways.

The research task has divided into two parts: the computation of best paths through "large" networks, and the assignment of "realistic" volumes. Not many years ago, the computation by an electronic computer of best paths through a freeway network as large as fifty interchanges was impossible. Yet there are at least a hundred intersections in the major street systems of even medium-size cities. Only within the last few years has it become possible to work with sufficiently large networks. Experimental programs were in use in various places earlier, but the first manual for an assignment procedure for large networks was not published until 1960 (2).

As an example of one of the earlier programs, a desire assignment was made to the 1,000-intersection major street system (including two freeways) for Detroit as of 1958. This produced the useful information of a freeway desire assignment in a much more extensive form, but it had all the limitations of a desire assignment, too. For example, it assigned 300,000 vehicles per day to parts of the freeway network which could handle only about one-third that volume, and indicated no traffic volume at all on Detroit's "main street" approaching the CBD. These volumes were too unrealistic to help plan the adjustments in the surface street system called for by the expected traffic-pattern changes created by the freeways.

The production of "realistic" assigned volumes has proved to be a more difficult problem. A number of approaches to its solution have been tried at a number of research centers (3). The iterative (i.e., repeated-assignment) solution described below has three characteristics that particularly recommend it. First, it produces significantly fewer assigned volumes that are in excess of capacity than does a desire assignment. Second, it produces assigned volumes that are significantly closer to traffic counts than does a desire assignment. Third, it includes a particularly efficient computer program that allows a capacity-restrained arterial assignment to be made at a reasonable cost.

PROCEDURE FOR ARTERIAL ASSIGNMENT

The procedure for arterial assignment reflects the assumption that a desire assignment will place volumes on freeways or other superior facilities that are in excess of the capacity of those facilities and, therefore, significantly different from what traffic counts would actually be. The goal of this procedure is to redistribute the excess volumes to realistic alternative paths in such a way as to produce assigned volumes that will be closer to actual traffic counts than were the desire-assigned volumes. The name given this procedure is the Wayne arterial assignment method.

Network

The assignment network consists of a number of nodes representing the intersections of major streets, and straight-line, two-way links between nodes (4). The nodes are identified by sequential numbers, and a link is specified by the statement of two such node numbers. The nodes are also coordinate coded so that the computer can determine distances between them; i.e., link lengths.

A generalized measure of capacity and a measure of "typical" speed are associated with each link. These measures vary with (a) the width of the facility, (b) the area, classified as CBD, intermediate, or outlying, and (c) facility type, classified as freeway, high-type areterial, or low-type areterial.

The capacity number represents the number of vehicles that can traverse the link in 24 hours under "typical" urban conditions including 10 percent signal failure at peak hour (5). A link's capacity is estimated by averaging the capacities of the intersections at its ends. This measure of capacity does not equal the absolute maximum number of vehicles that link could handle under conditions of greater congestion.

Certain of the nodes are designated "O-D intersections." They serve as "centroids" or points of trip end. Each traffic analysis zone is associated with a specific O-D intersection. Best paths are computed only between O-D intersections and not between all intersections, but the computer considers all intersections when it is determining best paths.

Volumes

The vehicular volumes to be assigned are contained in a nondirectional trip table; i.e., the volume going from A to B is combined with the volume going from B to A. Each interzonal volume is identified by two zone numbers. Intrazonal volumes are disregarded as if they were served by the local streets not in the major-street or arterial network.

Assignment Procedure

The program for securing an arterial assignment can be summarized in eight steps. Describing them requires that reference be made to five quantities designated V, X, R, e, and A.

1. A link value hereafter called V_1 is computed for each link. The quantity V_1 represents the time required to travel from end to end of the link (cf., 6, p. 110). It is estimated (by the computer) by dividing the length of the link as determined from its two coordinate codes by a "typical" speed which varies by type of facility and by area.
2. For each node designated an O-D intersection, a "tree" is computed. The paths that make up the tree represent best or shortest-time paths linking the O-D intersection of reference to every other O-D intersection. A best path is that sequence of links with the minimal sum of V_1.
3. The interzonal volumes are assigned to their paths. The sum of interzonal volumes accumulated on a link is said to be that link's assigned volume. It is designated X_1.
4. The program then determines R_1 or the ratio for each link, by dividing X_1 by the link's capacity. That is, R_1 for a link represents the assigned volume on that link after the first or desire-assignment pass, expressed as a ratio to the capacity of that link. It may be 0.0 if no volume is assigned to the link, or it may be 1.0 if the assigned volume is the same as capacity, or it may be 4.0 or 5.0 if the assigned volume exceeds capacity to that extent, or it may be any number between the extremes.
5. The value of each link (V_1) is changed on the basis of R_1. The formula by which the nature of the change is determined calls for an increasing "stretching" of the link value as R_1 increases, in the form of an exponential curve upwards. In other words, link travel time increases at an increasing rate as congestion increases. For measures of R_1 below capacity the formula allows a reduction of the link value in such a way that it cannot fall below about one-third of the original value. The exponential number determining the shape of the capacity-restraint curve is called e, and is approximately 2.7. The formula specifies that

$$V_2 = e^{(R_1 - 1)} V_1$$

The decision to use e in this particular formula was based on a combination of mathematical reasoning and trial-and-error experiment. The experimentation over an extended period that produced this formula, with the five axioms and three theorems that provide its mathematical rationale, are described by Smulick (7).

6. The second pass of an iterative arterial assignment begins with the computation of V_2 for each link. The network for the second pass differs from the network at the time of the desire assignment in that all links that have been desire-assigned volumes beyond their capacity now require more time to travel, and links desire-assigned volumes less than their capacity now require less time to travel. The trees of best paths computed through the network in which each link is described by V_2 differ significantly from the trees for the desire assignment. Assigning the interzonal volumes to the new trees produces a new volume, or X_2, on each link.
7. The computer averages X_1 and X_2 to produce A_2. The quantity A (or average assigned volume) at the conclusion of any pass may be considered the end result of the procedures, or each A_2 may be divided by link capacity to produce R_2. For each link a third link value may be computed according to

$$V_3 = e^{(R_2 - 1)} V_1$$

8. By averaging assigned volumes, computing R, revising V, and re-assigning, an indefinitely large number of passes could be made. However, at the conclusion of the desire assignment the R's range from 0.0 to 4.0 or 5.0, and the capacity-restraint formula works to alter link values more extensively the farther R is from one. If the second pass is successful, then the assigned volume at its conclusion, or A_2, is nearer to capacity than was the desire-assigned volume. Therefore, the difference between V_1 and V_3 will be less than the difference between V_1 and V_2. As subsequent passes produce volumes even closer to capacity, subsequent V's return even closer to V_1. This means that the assigned volume on a link comes closer, with each subsequent pass, to the level at which the typical speed of that link is actually possible; i.e., volume and speed approach a balance. It also means that after some particular pass, the majority of trips will be back on their original best paths and relatively little change from pass to pass is to be expected thereafter. This is one basis for considering the capacity-restrained arterial assignment to be complete. Another basis would be a pass which produced volumes little changed from those of the preceding pass.

To summarize, the eight steps are (a) to determine link travel time, (b) to trace best paths between all zones, (c) to assign volumes, (d) find the percentages of capacity represented by the assigned volumes, (e) to compute adjusted link travel times, (f) to compute new paths and assign volumes again, (g) to average assigned volumes, find percentage of capacity, and readjust link travel time, (h) to repeat assignments until a predetermined approximation of balance is attained.

Testing the Procedure

An arterial assignment of this type has been made to the major street system of the City of Flint, Mich., as that system existed in 1950. The system consists of 143 links and 99 intersections and includes no freeways. At the conclusion of the desire assignment to this network, 10 percent of its links had been assigned volumes which were more than 150 percent of the link capacity including some loaded with four or five times the volume they could actually carry. Another 25 percent of the links were loaded to less than one-half their capacity, including some assigned no volume at all.

A third of the interzonal paths computed for the desire assignment called for less than 10 min of travel time. Another third required between 10 and 15 min, and the final third required more than 15 min but less than ½ hr. These travel times assume that all vehicles could travel at speeds typical of city streets. Using the same typical speeds as a basis for comparison, best paths for the second pass were significantly longer. Forty-nine percent of those paths required 15 min or more to travel, and about 10 percent required 30 to 40 min, in terms of the original travel times. Only 25 percent of the paths remained the same as on the first pass.

Dividing volumes between the paths of the first and second passes and applying the capacity restraint formula on the basis of the new volume-capacity ratios produced a new set of link values or V_3, some of which were closer again to V_1 than were the second set of travel times. The net effect of this change was that more than one-half of the paths of the third pass were the same as the desire paths. At this point the assignment could be considered complete, but three additional passes were made as a test of the underlying assumptions. Only the expected small variations were observed.

Averaging the assigned volumes of the first three passes produced a significantly more realistic assignment. The volumes on about two-thirds of the links had shifted closer to capacity, and almost all of the volumes not shifting closer to capacity were those which were within 50 percent in the first place. Ninety-seven percent of the links were loaded to less than 150 percent of their capacity, which means that they could actually carry the volumes assigned to them although some of them would be congested.

In this particular case then, a solution to the problem of desire assignment was found. Unfortunately, however, there are a large number of "solutions" that would create a final balance. For example, a number of crosstown trips could merely be routed around the outer edge of the city. Congestion on links in the city's center would be relieved, a balance between volume and speed could be demonstrated, and a solution would have been found. The assigned volumes would be useless for traffic planning purposes, however, because they would not represent the unique solution that approximates actual traffic counts.

Part of the reason that the 1950 Flint network was selected for these experiments was that an actual 24-hr, midweek traffic count was available for all of its links. Some of the counts were below capacity and some above; they averaged 80 percent of capacity. At the conclusion of the desire assignment one-half the assigned volumes were more than 50 percent off from count, about equally divided between being too low and too high. During the second and third passes two-thirds of the assigned volumes shifted closer to traffic count. Almost all of the volumes not shifting closer to count were those which were within 50 percent in the first place. Of the final assigned volumes, three out of four were within 50 percent of traffic count (compared to one-half after the desire assignment). Only a few links were loaded significantly low. Twenty-three percent of the links continued to be loaded to more than 50 percent above count after the third pass, but this might be expected inasmuch as the capacity measure that was the basis of the adjustment tended, on the average, to be higher than count. This would mean that values of links overloaded on the desire assignment sometimes would not be adjusted enough to push volumes down to count even when properly adjusted in terms of capacity.

The important fact is that more than 90 percent of assigned volumes were either within 50 percent of count in the first place or shifted closer to count during the capacity-restrained assignment passes. It is this fact that would give validity to 1980 volumes assigned by this method.

Practicality of Procedure

It has been demonstrated, then, that this method achieves its goal of placing the excess volumes of a desire assignment on realistic alternative paths in the sense that its final assigned volumes approximate traffic counts significantly more closely than desire-assigned volumes. Its practicality for adoption by highway planners depends, however, on its costs for personnel and machine time, as well as its validity. Freeway desire assignments are performed for traffic engineers less often then their usefulness warrants, in part because of their cost. At first glance, therefore, a program for arterial assignment that calls for a detailed description of the total major street system, and then prescribes assignment by repeated passes, appears to be impractical.

This is not necessarily the case, however, because of two factors. The first is the simplification of network description previously suggested. All necessary information may be provided by (a) a map of major streets and (b) an inventory specifying pavement width and the location of one-way streets and divided highways. This presumes the existence of a trip table and a system of coordinates from an O-D survey, of course. The other factor in reducing costs is the shorter running time of an improved computer program for finding minimum paths and assigning volumes to a network. The improved program is called the branch method for arterial assignment and it includes the following five features:

1. Basic to the branch method is the concept of branch paths. A branch path is a path from an intervening node to the destination contained in a path from origin to destination. A branch path of a minimum path is also a minimum path. Consider, for example, the minimum path from origin node i to node j found in the build-up of tree i. If the minimum path from i to j passes through three intersections (j_1, j_2, and j_3), then, while determining the path from i to j, three other minimum paths are defined which are branch paths. The branch paths are from node j_1 to node j, from j_2, to j, and from j_3 to j. In the branch method, during the tracing of paths for any tree all branch paths are traced and information which can be used during the computation of later trees is stored. In fact, many later trees are completely defined during the build-up of the first few of them.

2. A second feature was suggested by an investigation of the branch pat concept. This is the technique of tracing and recording individual paths during tree build-up,

rather than after a tree is finished. This in itself saves considerable machine time, and in the branch method volumes are assigned to links at this point, rather than afterwards, resulting in a further time saving.

3. Another technique that contributes a time saving is the tracing of one-way paths rather than two-way paths, because link values are symmetrical; i.e., the value of a link joining two nodes is the same in one direction as in the other.

4. Sequencing of nodes during tree build-up, in the sense of putting them in the most convenient numerical order, has had considerable application elsewhere (2, 8) and was found important for the branch method.

5. The final technique which is incoporated had long been published (9) but had found little practical application. It involves keeping track of the adjacent node in a path from destination to origin during tree build-up for use in later path tracing.

By means of these various techniques the branch method for arterial assignment achieves a reduction in machine time great enough so that the running of several passes is feasible in terms of cost. A three- or four-pass assignment can be made for no more than twice the cost of a single freeway assignment. Because more than twice as much information is provided, this may be considered an entirely reasonable cost level.

The branch method is now programed for the IBM 650-Ramac, and can handle a network of 240 intersections. This is judged sufficient for all Michigan cities except Detroit. The program includes an output-analysis subprogram which can print out, at the conclusion of any pass (a) the value for each link; (b) the assigned volume (X and A) on each link; (c) the capacity, and also the assigned volume as a percentage of capacity, for each link; (d) the intersections composing all best paths; and (e) the total travel time (or sum of V_1) for each path. The branch method is now being programed for the IBM 7070 to handle a network of 1,000 intersections.

ARTERIAL ASSIGNMENT AND TRANSPORTATION PLANNING

On the basis of experience in the recent past, it should be assumed that this and any other existing arterial assignment procedures will be refined and improved rapidly, in the next few years if not months. For example, experiments on the program described in this paper are being continued to increase its flexibility, and perhaps to improve the equal-share way of dividing interzonal volumes over alternate paths. Even after an arterial assignment method is valid, reliable, and not too costly; however, it will not be widely used unless there is some agreement as to its importance for transportation planning.

There has been some disagreement as to the wisdom of making major research investments in "capacity restraint" (3, p. 18ff). Part of the reason for this is a suspicion that a computer may not be able to "simulate" traffic as well as an experienced traffic engineer because his mind can take more of reality into account. However, this argument applies equally well to freeway desire assignments and they are performed anyway. The reason is that the wisest traffic engineer cannot do arithmetic as fast as a computer, and it takes considerable arithmetic to study the transportation system of a metropolitan area. But it also takes considerable interpretation and judgment before the computer's arithmetic can be turned into well-planned transportation, whether the figures are from desire or capacity-restrained assignments.

A specific objection to capacity restraint is that it could make any system "look good" in the sense of producing balanced volumes on a badly planned network. It is true that a tranportation system could be proposed that would call for freeways constructed diagonal to major desire lines, and that arterial assignment might load them to capacity anyway, and in this sense make the network "look good." However, trip times and distances through such a network, on the average, would be greater compared to those through a well-planned network. The Wayne method can determine trip times for a given network, and these can be compared to trip times in alternate networks. It is doubtful, therefore, that planners need to be deceived by a balanced assignment to a poorly-planned system. Instead, arterial assignment may play a useful role in determining the best-planned network among a number of proposals.

In highly generalized terms, the stages of transportation planning might now take this form:

1. The basic data for transportation planning, which are person and vehicle volumes moving between traffic analysis zones, are secured by an O-D survey. Alternately, basic data may be obtainable by means of synthetic models.

2. Interzonal volumes for some future year (selected for planning purposes) are provided, either by computer iteration based on projections of land uses or by some kind of model.

3. Desire lines may be determined by the rapid and inexpensive method of zone assignment (10) to help in the process of planning new facilities where needed.

4. New facilities in the form of public transportation, freeways, or improved streets and highways are planned and proposed as traffic engineers and planners consider them to be appropriate.

5. Arterial assignments can test the relative merits of alternative proposals, helping with both processes of economic justification and geometric design.

6. Capacity-restrained arterial assignments can assist in the planning of adjustments in the existing transportation system, as they are required by the changes brought about by new facilities.

ACKNOWLEDGMENTS

The development of this procedure has been under the general supervision of Albert J. Mayer, Director of the Detroit Area Traffic Study and Associate Professor of Sociology at Wayne State University. Sponsors are the Michigan State Highway Department, the Wayne County Road Commission, and the City of Detroit, in cooperation with the U.S. Bureau of Public Roads. Research and program development at the Computing Center of the University is under the general supervision of Walter Hoffman, Director, and under the immediate supervision of James Smulick with the assistance of Maggie Goodwin. The project as a whole is under the immediate supervision of Robert Smock, Assistant Director of the Traffic Study and Research Associate of the University, with the assistance of Patricia Ferman.

REFERENCES

1. "Traffic Assignment." HRB Bull. 61 (1952).
2. Mertz, W.L., "Electronic Computer Program for Assignment of Traffic to Street and Freeway Systems." U.S. Bureau of Public Roads (1960).
3. Mertz, W.L., "Review and Evaluation of Electronic Computer Traffic Assignment Programs." HRB Bull. 297, 94-105 (1961).
4. "Defining Arterial Segments for Traffic Assignment Purposes." Detroit Area Traffic Study (1957).
5. "Determining the Capacity of Arterial Segments." Detroit Area Traffic Study (1957).
6. "A Policy on Arterial Highways in Urban Areas." AASHO (1957).
7. Smulick, J., "A Mathematical Model for Travel Assignment." Master's essay, Department of Mathematics, Wayne State Univ. (1961).
8. Pollack, M., and Wiebenson, W., "Solutions of the Shortest-Route Problem—A Review." Operations Research, 8:No. 2 (1960).
9. Moore, E.F., "The Shortest Path Through a Maze." Proc., Internat. Symp. on the Theory of Switching, pt. 2 (April 2-5, 1957). In Annals, Computation Lab., Harvard Univ., Vol. 30 (1959).
10. Smock, R., "Traffic in Detroit in 1953 and 1958." Detroit Area Traffic Study (1961).