# A Direct Approach to Traffic Assignment

MORTON SCHNEIDER, Systems Research Chief, Chicago Area Transportation Study

• THE MOST straightforward way of obtaining the distribution of traffic in a road network is simply to distribute it: to resolve the traffic into trips, each with an origin and a destination, then find the best path through the network for each trip and note the aggregate appearances on every network member of trips following these paths. This technique is known as assignment, and is the conventional approach to the distribution problem. Though the convention is hardly an antique one, partial and minced assignments have been attempted for some time; however, the first full-scale assignment to a fairly complete metropolitan street system was performed only a few years ago. This was done by the Chicago Area Transportation Study, in the summer of 1958, on an IBM 704 computer. The computer program was then rewritten a year later to allow simpler coding and mapping procedures, and to incorporate sensitivity to congestion of streets, as well as other features. It was modified early in 1961 for the IBM 7090.

The apparently reasonable procedure of assignment has, however, several flaws, both fundamental and practical, which give rise to rather bizarre results here and there. (Also to variations on the basic method. These variations sometimes have a tenuous plausibility, are very elaborate computationally, can be adjusted to give any desired answers, and are called models.) In the first place, very little more is understood about trip-end distribution than about traffic distribution itself. Besides, what constitutes the best path (it is usually taken to be the path of least travel time), and do all trips have the same idea of best? What happens as trips interfere with each other? This latter point is widely held to be most crucial, and one ambition of nearly every neo-assignment is to account for it.

There is also the matter of zones, less widely held to be crucial. Computers can handle only so much detail, and the assignment method is, in a sense, too unsophisticated to give the computer any help. Among other reductions, the study area must be divided into arbitrary zones, and all trips emanating from a zone must be construed to have their origin at a single point representing that zone. Thus, the zonal interchange (the number of trips moving from some one zone to some other zone) replaces the trip as the elementary unit. At the same time, the synthetic and warped problem of predicting these zone-to-zone movements becomes dominant, blunting what little insight there is into trip behavior; even a perfect theory of trip distribution could correctly calculate interchanges only with a difficult explicit treatment of zone geometry. More than that, though, this condensation into points instantly shatters the locational precision of the system. Just how much does an area of some hundreds of points, all bursting with trips, spaced among thousands of street intersections resemble a metropolitan region? The errors introduced by zoning are surely considerable, possibly greater than all others. Perhaps the simulation refinements added to assignments to correct for congestion, diversion from best path, etc., are really, in their operation, nothing more than mechanical devices that spread out zone-connected errors, which is not at all the same as eliminating those errors.

Although assignments, once their peccadilloes are understood and forgiven, can be made fairly useful in a rough or over-all way, they tend to be lumpy and irresponsible in detail. Yet there is a growing demand for minor information—the effect of building a bridge here, changing the location of a ramp there. Even if assignments were very good at this sort of thing, they are too expensive and arduous to undertake for ordinary details. Certainly a small piece of area cannot be dissociated from the large region in which it resides, but that does not mean a tolerable estimate cannot be made without a

Paper sponsored by Committee on Origin and Destination.

ponderous total solution. The traffic load on a particular street section is generally not much influenced by anything more than a few miles away.

There are other odds and ends of difficulties with the assignment technique, notably boundary problems, but this paper really has nothing personal against assignment. Some if its best friends are assignments (for example, the programs mentioned earlier). The object here is to suggest a new, less stiff research posture.

A highly simplified situation is used as an example. There is a square region of unlimited size, and the rate of trip origins per day per square mile is the same everywhere in this region. Three road networks of distinctly different quality overlay the region; each network is a uniform grid. The first, highest quality network—it might even be called expressway—connects with the second, or arterial, network at every intersection of the two, but does not connect at all with the third, or local. The arterial system in turn connects with the local at every intersection of the two. All trips originate and terminate on the local system. The region is so large that perturbances due to its borders may always be ignored. Then the total length of road in each network is $2L^2/z_i$, in which L is the side of the region, and $z_i$ is the grid interval of the particular network (the spacing between parallel members). The total number of vehicle miles per day driven in the region is by definition $\rho L^2\bar{r}$; $\rho$ is the density of trip origins, and $\bar{r}$ is the average over-the-road trip length.

It is assumed that the daily volume of traffic is the same on every link of a particular network. The total vehicle-miles driven on each network is, therefore, just the network volume times the miles of road in the network:

$$2L^2(V_1/z_1 + V_2/z_2 + V_3/z_3) = \rho L^2\bar{r} \tag{1}$$

or, eliminating the size of the region,

$$V_1/z_1 + V_2/z_2 + V_3/z_3 = \rho\bar{r}/2 \tag{2}$$

$V_i$ is the volume found on the i network.

For the volume to be the same on every link of a network, it is necessary (though not sufficient) that the trips leaving a network at any intersection with another network be equal to those entering at that intersection. If $p_{ij}$ is the probability of transferring from the i to the j network,

$$p_{12}V_1 = p_{21}V_2 \text{ and } p_{23}V_2 = p_{32}V_3 \tag{3}$$

or

$$V_2 = \frac{p_{12}}{p_{21}} V_1 \text{ and } V_3 = \frac{p_{23}}{p_{32}} V_2 \tag{4}$$

Setting

$$\frac{p_{ij}}{p_{ji}} = U_{ij} \tag{5}$$

Eq. 2 may be rewritten

$$V_1\left(1/z_1 + U_{12}/z_2 + U_{12}U_{23}/z_3\right) = \rho\bar{r}/2 \tag{6}$$

It should be remembered throughout the argument that, inasmuch as the situation is set up to be entirely uniform, there is no reason for corresponding quantities at different places in the region to be anything but the same.

Here a little definition is in order. Among other things, the phrase "networks of distinctly different quality" is taken to mean that a trip that once moves from a higher to a lower quality network will never again return to the higher network—it has used the higher street as much as it profitably can and is on the way down to its destination. This is a more important construction than it may seem, because it permits a particularly simple line of attack on the $p_{ij}$'s.

With this delimitation, the total number of trips using a network for any part of their journey must exactly equal the number that turn from that network to a lower one, except, of course, for the local network. Once a trip enters a network, it must—whatever else it does—turn down, and it can do that only once. So the number of trips that use expressways is the number that leave an expressway at each arterial intersection, times the number of intersections; that is, $p_{12}V_1R/z_2$. Here, R is the total length of expressways. But the number of trips using expressways may also be measured as the vehicle-miles on expressways divided by the average distance traveled on expressways by those trips that use expressways at all. Thus,

$$p_{12}V_1R/z_2 = RV_1/\bar{r}_1 \tag{7}$$

Solving gives

$$p_{12} = z_2/\bar{r}_1, \text{ and similarly for arterials, } p_{23} = z_3/\bar{r}_2 \tag{8}$$

The expressway average, $\bar{r}_1$, may be gaged by choosing a convenient approximation. Although trips do not appear to be strictly governed by a distance distribution, especially in their detailed behavior, the force of distance is quite strong. Certainly, data from several regions show a consistent decay in number of vehicle trips as trip distance increases and suggest, moreover, that this decay is exponential. Probably no simple-minded distribution concept fits the data better than

$$dn/dr = ae^{-br} \tag{9}$$

The constants, a and b, are established from the natural conditions of distributions,

$$\int dn = 1 \text{ and } \int r dn = \bar{r} \tag{10}$$

yielding

$$\frac{dn}{dr} = \frac{1}{\bar{r}} e^{-r/\bar{r}} \tag{11}$$

Now this function has an amiable property: the average remaining length of a trip which has already traveled any distance, C, is just the over-all average length plus the minimum remaining length, m. From Eq. 11, the number of trips longer than C + m is

$$\int_{C+m}^{\infty} \frac{1}{\bar{r}} e^{-r/\bar{r}} dr = e^{-(C+m)/\bar{r}} \tag{12}$$

and the sum of the remaining lengths of these trips is

$$\int_{C+m}^{\infty} \frac{1}{\bar{r}} (r - C) e^{-r/\bar{r}} dr = (\bar{r} + m)e^{-(C+m)/\bar{r}} \tag{13}$$

Therefore, the average (Eq. 13 divided by Eq. 12) emerges as $\bar{r} + m$. A trip just entering an expressway may be regarded as having already traveled its total approach distance, both at origin and destination, whereas its minimum remaining length is merely the distance to the next exit, $z_2$. The average expressway length of expressway trips is

$$\bar{r}_1 = \bar{r} + z_2 \tag{14}$$

There is one difficulty, however. Although C, the distance already traveled, need not be constant, it must not be a function of total trip length or the calculation of the average breaks down. That is to say, expressway trips must use the expressway network as much as they can, using lower networks only as approaches to the highest, rather than as alternate routes. This further augments the definition of quality. (Further, the larger the expressway spacing, the more C tends to become related to the total trip length.)

Simple-mindedly adopting the preceding distribution gives a working evaluation of $\bar{r}_1$. But the arterial average, $\bar{r}_2$, is a different matter, because trips clearly do not use arterials as much as they might. If an arterial link is defined as a piece of arterial between two expressways, and then if every trip that used an arterial passed, say, the midpoint of a link once and only once, the total number of trips using arterials would be the arterial volume times the number of links; i.e., $V_2R/z_1$. (Earlier R was the total length of expressways; here it is the total length of arterials.) Actually, this quantity includes some trips counted at more than one link and neglects some trips that pass no link midpoint. Therefore, the volume times the number of links equals the total trips plus a residual; referring to Eq. 7, this may be written

$$V_2R/z_1 = V_2R/\bar{r}_2 + (D - E)R/z_1 \tag{15}$$

or

$$1/z_1 = 1/\bar{r}_2 + (D - E)/z_1V_2 \tag{16}$$

D is the number of double-counted trips at each link midpoint, and E is the number of trips per link that are never counted. It is natural to hope that D and E cancel each other out, or at least that (D - E) is small compared to $V_2$. It also seems reasonable when expressway spacing is in the range usually discussed. At any rate, a first approximation is better than nothing; therefore, letting the rightmost term in Eq. 16 drop out,

$$\bar{r}_2 = z_1 \tag{17}$$

This is really more an assertion than a derivation. But Eq. 16 gives some clue to the order of error.

Looking at Eq. 8, the probabilities of turning from the highest network to the intermediate, and from that to the lowest are determined, more or less, by

$$p_{12} = z_2/(\bar{r} + z_2) \text{ and } p_{23} = z_3/z_1 \tag{18}$$

One obstacle is left—the probabilities of turning from the lowest network to the intermediate, and then to the highest, $p_{32}$ and $p_{21}$. This will be hurdled by the assumption that the probability of a trip entering the i network from the j network is the same as the probability that a trip already on the i network will continue on it. For example, a trip approaching an expressway on an arterial is assumed to have just the same grounds for entering the expressway as a trip on the expressway has for staying there. In notation,

$$p_{ji} = 1 - p_{ij} \tag{19}$$

Putting this, together with Eq. 18, back into Eq. 5 gives

$$U_{12} = \frac{z_2}{\bar{r} + z_2} \Big/ \left(1 - \frac{z_2}{\bar{r} + z_2}\right) = \frac{z_2}{\bar{r}}$$

and

$$U_{23} = \frac{z_3}{z_1} \Big/ \left(1 - \frac{z_3}{z_1}\right) = \frac{z_3}{z_1 - z_3} \tag{20}$$

which fairly well completes the solution. (The $U_{ij}$'s can be greater than 1, although the $p_{ij}$'s, of course, cannot.) All that remains is to rewrite Eq. 6,

$$V_1 = \rho\bar{r}/ 2\left(\frac{1}{z_1} + \frac{1}{\bar{r}} + \frac{z_2}{\bar{r}\,(z_1 - z_3)}\right) \tag{21}$$

$$V_2 = \frac{z_2}{\bar{r}} V_1 \tag{22}$$

$$V_3 = \frac{z_3}{z_1 - z_3} V_2 \tag{23}$$

The average trip length, $\bar{r}$, must simply be taken at its empiric value—about six miles. A quick numeric example would be a region with a trip density of 10,000 vehicles per day per square mile, expressway spacing of 5 miles, arterial spacing of 1 mile, and local spacing of $1/10$ mile. The daily volumes past a point will then be expressways, about 75,000 vehicles; arterials, about 12,500 vehicles; and local streets, about 250 vehicles; which are not at all bad orders of magnitude.

The extension of these equations to more than three networks is trivial—the entire argument could have been carried out, unaltered, in generalized subscripts and open summations. The extension to more complicated connections among networks is less trivial, but more tiresome. Each connection introduces an equation of the form of Eq. 4, adds a degree of freedom in equations analogous to Eq. 8, and complicates network averages. The most elaborate situation that has been solved is one of five networks, each with as many as two connections to lower networks and two connections to higher networks. The final expressions are ugly nuisances. That kind of generalization probably drives the whole line of argument too far; the assumptions and suppositions—notably the assumption of uniform volumes on all parts of a network--begin to look doubtful even in hypothetical idealizations.

To a great extent, of course, all this really begs the question. Road systems cannot be decomposed into distinctly different networks, whatever that means; network averages are not really calculable; the assumptions of uniform volumes and complementary probabilities do not hold exactly. It is probably important, though, to carry elementary considerations to some sort of provisional conclusion, creating a platform from which enlargements can be launched.

There is no reason for this general point of view to be limited to whole networks interacting with each other. The turning probabilities could be associated with each intersection, irrespective of the kinds of networks involved. These probabilities are most interesting quantities. If they could be determined by some bold stroke of insight, they would implicitly define the entire realm of street quality and trip distribution. If these were understood then, conversely, the probabilities would be determined. A treatment of this sort might very well yield, without half-trying, the sensitivity of trip distribution to the transportation system. The deeper matter of trip generation is, at the moment, a little harder to fit into the grand vision.

In the meantime, on a more pedestrian level, some attempt ought to be made to relate the probabilities explicitly to a street quality measure. This could easily lead to expressing the $p_{ij}$'s as functions of the traffic volumes themselves, providing a truly elegant treatment of problems associated with trips interfering with each other. Further, trip distribution concepts must be sharpened, both to strengthen the base of the first-step uniform case and to probe the extremely important generalization to non-uniform densities.

Surprisingly, the simple expressions, as they are set down here and in somewhat modified form, are proving rather useful. There is no intention to document this as a realistic theory, but the results are realistic in a soft, fuzzy way — just as about as realistic, in fact, as assignment results. Non-uniformities must be largely ignored; and the network averages stated here — especially the arterial average — are valid only for a restricted family of spacings, which leads sometimes to patently spurious computations, most noticeable in the local network volumes. But the uses of traffic volume information are, right now, not too fine; orders of magnitude usually are enough or, at least, all that is expected. By abstracting the real street system and taking average local densities, minor questions can be answered roughly, and quickly. Certainly it would be wrong to pick this method up, as it stands, and run with it. Yet these simple equations have, perhaps, some value and some charm.