Implementation of Service-Area Concepts in Single-Route Ridership Forecasting

ALAN J. HOROWITZ and DAVID N. METZGER

ABSTRACT

Transit service area is the basis of several direct-demand models of single-route ridership. This service-area concept has not been integrated into the more recent single-route models that have evolved from the Urban Transportation Planning System and similar four-step simulation procedures. How the service-area concept was incorporated into the Transit Ridership Forecasting Model is explained and the application of the concept on routes selected from the Milwaukee urban area is illustrated. It is shown that use of the service-area concept removes a serious tendency for four-step models to underestimate ridership on marginal routes. It is also shown that proper application of the service-area concept can reduce both computation time and data requirements.

Extensive work has been done on modeling ridership for entire transit systems. Those models, notably those of the Urban Transportation Planning System (UTPS), have been in use for nearly a decade and are particularly good for long-range forecasts where there are major system changes. Parallel, but much less recognized work has been done on estimating ridership for single routes. These models are designed for short-term and mid-term planning where changes to the system are confined to only a few routes.

Early single-route ridership models were basically of two types: direct demand and special-purpose simulation. The direct-demand models (1-5) were essentially statistical or elasticity equations that estimated total route ridership on the basis of service area, dwelling-unit density, and aggregated service characteristics. They had the advantages of simplicity and explicit consideration of the existing state of the system and service areas of routes. Simulations (6) applied the same four-step approach (i.e., trip generation, trip distribution, mode split, and trip assignment) as found in the full-system models to special single-route cases, such as many-to-one and park-and-ride routes. It appeared that the four-step approach had particular advantages in terms of flexibility and accuracy across a variety of route types, so development of a new single-route model proceeded in that direction. The Transit Ridership Forecasting Model (TRFM) (7-9), the model that forms the basis of the research described here, illustrates one way of applying the four-step approach.

TRFM, like the direct-demand models, maintains consistency with service areas used for other route-planning activities. Implementation of the service-area concept permitted significant simplifications in the four-step approach. The program size is smaller than that of the Transit Operations Planning (TOP) Model (10), which is a highly sophisticated direct-demand model, yet through the windowing and focusing concept (11), it can handle routes in systems of nearly any size--even on a small microcomputer. The windowing and focusing concept was implemented with two simplifying assumptions: (a) the possibility of multiple transfer trips can be ignored and (b) small travel zones are required only on the route of interest. These assumptions were subjected to sensitivity tests (8) and were not found to introduce significant error. Implementation of the windowing and focusing concept has accuracy advantages as well. It is possible to represent the route of interest in far more detail than would be practical with full system models. Earlier tests of TRFM showed that once the model has been calibrated on a single route in a radial network, it can accurately (within 5 percent) forecast ridership on other existing routes in the same network. A detailed description of TRFM may be found in the references cited earlier.

The standard definition of a service area is the region within 0.25 mi of the route or set of routes in a system. Furthermore, the service-area concept has frequently been extended to account for natural and man-made barriers to travel (such as school district boundaries) and temporal constraints on service (such as service only during rush hours). Any such service area is arbitrary in that some of those who live outside the service areas still use transit and others refuse to walk even a short distance to a route. Nonetheless, the service-area concept persists as one of the most useful and used tools for transit route planning.

Implementation of the service-area concept is relatively easy in radial systems with little circuity. In these systems, almost all possible transit trips can be satisfied with at most one transfer. Thus it is only necessary to first determine which trips occur exclusively within the service area of the route of interest and then determine which trips go between that route's service area and the service areas of immediately connecting routes. All other possible transit trips can be ignored.

Serious problems with the service-area concept develop when there is substantial circuity in the network. Then it becomes difficult to allocate some land parcels exclusively to service areas of single routes, especially parcels near transfer points. In gridded networks, where there is substantial circuity, even single-transfer trips have alternative paths through the transit network. The foregoing simple procedure for handling service areas in radial systems fails to adequately describe transit trip making in gridded systems.

TRFM was not originally designed for large, gridded transit systems. When the model was upgraded for this type of system, it was deemed necessary to retain TRFM's simplicity of operation, to not increase the amount of data preparation, and to maintain its implementation on a small microcomputer. Much of this was accomplished by exploiting to the maximum extent the interactive graphics features of the program (9).

OVERVIEW OF THE SERVICE-AREA ALLOCATION PROCEDURE

To properly allocate land to service areas in gridded networks it was found necessary to implement a version of multiple-path trip assignment. It is interesting to note that multiple-path trip assignment has not been implemented on UTPS. Such an implementation would be difficult without placing severe restrictions on the manner in which zones can be defined, because transit riders' choice of routes depends heavily on walking distance. So that the highway and transit networks can be consistent, UTPS permits zones of arbitrary shape, that is, having parts both inside and outside service areas of routes. But for a multiple-path assignment to work well the algorithm must know where service areas of connecting routes overlap, that is, where people have a true choice between routes. This would not be possible unless zonal boundaries closely corresponded to service-area boundaries. Consequently, UTPS does not check for overlapped service areas but instead permits planners to adjust walking times to maintain a reasonable split between routes. As will be seen later in this paper, such ad hoc methods of handling multiple paths have serious consequences when ridership on routes with poor service characteristics is estimated.

On the other hand, TRFM forces planners to construct zones by following service-area boundaries and to construct networks according to a rigid set of rules. The rules are so specific that the eventual location of nodes and links in a TRFM network, as displayed on a CRT screen, contains all the information necessary to completely regenerate the underlying zone structure. If necessary, the computer can be directed to find every overlapped service area without explicitly knowing the zonal boundaries. The following example helps illustrate how TRFM networks are constructed and how the program interprets them.

Figure 1 shows a hypothetical gridded transit network with 13 routes and 42 transfer points. Route B is the route of interest, that is, the route for which ridership is estimated. A TRFM network focuses on the route of interest; immediately connecting routes are shown in far less detail. If it is assumed that riders do not make multiple transfers, the TRFM network and zone system look like those in Figure 2. The routes that do not immediately connect to the route of interest are shown as dashed lines. All trips that use the route of interest must be on this network.

There is a possibility of two paths between a point on a connecting route and a point near a transfer node along the route of interest. For example, a trip from x to y could be satisfied by following Route I and then Route B or by following Route F and then Route K. This latter path does not use the route of interest. The same is true for the reverse trip. The only people who have these two potential paths are persons traveling between the overlapped service areas of Routes B and K and the overlapped service areas of Routes F and I. This choice between two paths should be presented to only these few riders, not all persons traveling to or from zone x.

Multiple-path procedures for highway traffic assignment have been available for some time, rising from heuristic procedures of the 1960s to efficient, precise stochastic assignment procedures in the

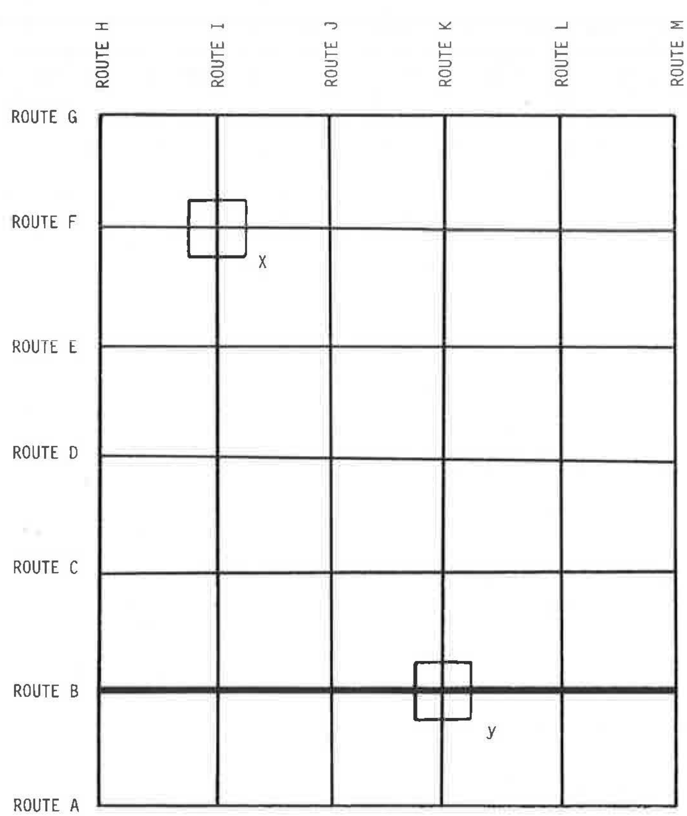


FIGURE 1 Hypothetical gridded network.

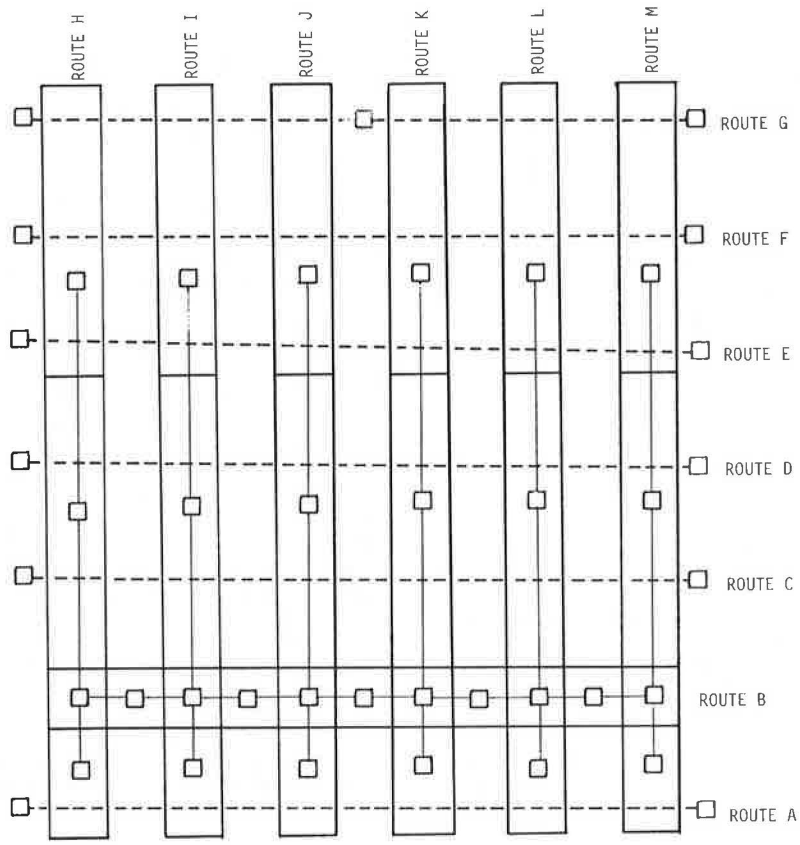


FIGURE 2 TRFM network and zone system for the hypothetical network.

1970s (12,13). But, contrary to conventional wisdom, trip assignment on transit routes is not the same problem as traffic assignment. In highway networks, the number of possible paths is huge and many of the paths have nearly identical disutilities. For transit trips, the reluctance to transfer between routes is so strong that optional paths are few. In well-laid-out transit systems, where overlaps between service areas of different routes have been minimized, the optional paths are obvious to most riders and should be obvious to a properly constructed assignment algorithm. For example, most bus trips are made without transferring, and these riders simply do not consider possible paths that involve transfers, even in rare cases where there may be small travel time savings (14-16). Many of the remaining riders find that only one path can satisfy their trip within one transfer. Very few transit riders have a choice of more than two paths. All that is really needed is a bipath algorithm. Thus, multiple-path trip assignment for transit networks must be handled differently from that for highway networks. Not only are there fewer path choices, but as will be seen, the way in which zone boundaries are defined and used becomes critical.

The bipath algorithm, in theory, is quite straightforward. It has to determine where multiple paths exist, compute the overlapped service areas, determine the fraction of trips that have a choice of two paths, and assign the trips on the basis of relative service characteristics of the two paths. However, in practice such an algorithm can become complex while dealing with all the possible shapes of zones and overlaps of service areas.

In order to easily discuss the various aspects of how the service-area concept is implemented with a bipath algorithm, it is necessary to first define a few terms. The route of interest is here called the primary network (Route B in Figure 2). The immediately connecting routes and the primary network together form the secondary network (Routes B, H, I, J, K, L, and M in Figure 2). The secondary network must be a tree; that is, it must not contain circuits. Routes that intersect the secondary network but not the primary network are called tertiary networks (any of Routes A, C, D, E, F, and G). Each tertiary network is also a tree. There may be several tertiary networks associated with a single secondary network. If all tertiary networks were to be overlaid on the secondary network and all intersections between them were to be explicitly designated as nodes, the combined network would look very much like those created for UTPS and other full-system models. However, intersections between tertiary networks and the secondary network are not explicitly indicated, so they are called virtual nodes. Sections of routes between adjacent virtual nodes on tertiary networks are called virtual links.

Tertiary networks are drawn on the CRT display in the same manner as secondary networks. Explicit nodes are drawn first and then explicit links are connected to pairs of explicit nodes. A sample CRT display with both secondary and tertiary networks is shown in Figure 3. Only the explicit links are important to the bipath algorithm. They are of three types: (a) links that always form virtual nodes at intersections with the secondary network, (b) links with a one-way characteristic that form virtual nodes but only allow these nodes to be used by trips going in one particular direction, and (c) links that do not form virtual nodes at intersections with the secondary network. Most tertiary networks are composed of links of the first type, and many of these networks have only a single explicit link. However,

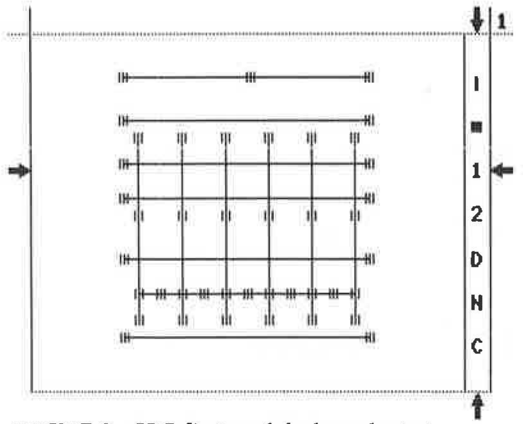


FIGURE 3 CRT display of the hypothetical network.

the latter two types of links can be used where it is clear that certain paths should be excluded from choice sets of riders.

TRFM requires that its various networks be trees for four reasons. First, a more general network configuration adds little or nothing to the accuracy of forecasts and requires substantially more data preparation time. Second, trees can be much more compactly stored in the computer's memory. Third, algorithms for analyzing trees are far faster than those for general networks. And fourth, it is much easier to regenerate the underlying zone structure of a tree.

The decomposition of general networks into trees is not an unusual practice, given that it is a standard technique of computerized network analysis. TRFM, however, makes planners aware of this process, giving them complete control over how the decomposition is to be accomplished and allowing them to optimize their data requirements around it. Because of the tree representation, TRFM is able to dispense (without sacrificing generality of network representation) with most zonal centroids, walking links, transfer links, and other artificial network elements that are required by full-system models. Equivalent network elements are generated internally; the planner need not be aware of their existence.

TRFM's capacity gives an idea of the compactness of network representation that can be achieved when networks are created by superimposing trees and exploiting service areas. TRFM runs on a 64 K microcomputer (an Apple II+/IIe), which is small by current standards. It can easily handle networks with 160 explicit nodes, 320 explicit one-way links, 800 virtual nodes, and 1,600 virtual links without memory swapping. Because TRFM requires few artificial network elements, its effective size is even larger. It can handle single-route (and even some multiple-route) problems that previously could only be analyzed on main-frame computers or large minicomputers. An implementation of TRFM on a slightly larger microcomputer would permit analysis of systems in the largest of cities.

The bipath algorithm is not applied to the roughly two-thirds of trips that occur exclusively on the primary network (i.e., within the service area of the route of interest). Of the remaining third or so, considerably less than half will be assigned to the tertiary networks and thus discarded from ridership estimates on the route of interest. This is done by (a) determining which trips can feasibly use tertiary networks by looking for overlapped service areas and (b) splitting them between the secondary and tertiary networks on the basis of service characteristics. The mathematical aspects of the algorithm will not be described in detail. Rather, the remaining portions of this paper will concentrate on discussing how well the concept works and how it may be applied to full-system simulation.

TRFM'S IMPLEMENTATION OF THE BIPATH ALGORITHM

In TRFM's specific implementation of the bipath algorithm the following six assumptions are made:

1. Boundaries of service areas along routes parallel those routes at a fixed distance, typically 0.25 mi.
2. The CRT drawing of the network is to scale.
3. Bus running time between two points following a tertiary network is identical to the bus running time between the same points on the secondary network. Consequently, the only difference between the level of service on these two paths relates to out-of-vehicle time. Of course, paths along tertiary networks that are obviously poor choices can be selectively deleted by the planner.
4. Transfer coordination does not exist at virtual nodes. The transfer time is taken to be the mean waiting time of the appropriate intersecting route.
5. Choices between paths exist only where the service areas of two paths overlap.
6. The probability of choosing a particular path is provided by a logit model as suggested by Dial (12).

The second through sixth assumptions were made to save data preparation time, and they could easily be made more rigorous if it were found necessary. The analysis described later in this paper suggests that more rigor is not necessary. In fact, a good argument can be made for further simplifications to the bipath algorithm in the single-route case.

In order to perform the bipath algorithm, the program must go through the following steps:

1. Identify link segments associated with each zone (the program must generate its own links not explicitly designated by the planner at the ends of routes);
2. Find the areas of zones;
3. Identify the tertiary networks;
4. For each virtual node, determine its location, intersecting links, directionality of intersecting links, and the closest transfer node on the primary network;
5. For each origin-destination pair, determine by using overlapped service areas the fraction of trips that can possibly use a tertiary network;
6. Split these trips between secondary and tertiary networks by comparing the various components of out-of-vehicle time, thereby determining the fraction that does not use the secondary network;
7. Compute the number of trips that use the secondary network from the fractions determined in the previous two steps and the number of trips for each origin-destination pair;
8. Resolve any double counting of trips due to multiple overlaps of service areas; and
9. Assign the trips to the secondary network and count them in total ridership.

Because of the underlying tree structure of the networks, this algorithm can be efficiently executed. Steps 1 to 8 take about 140 sec on a typical network with 200 virtual nodes.

An earlier version of the algorithm further subdivided overlapped service areas into 400 smaller parcels in order to better account for differences in walking distances to the intersecting routes. It was found that provision of this level of detail added almost nothing to model accuracy but much to computation time, so the step was abandoned.

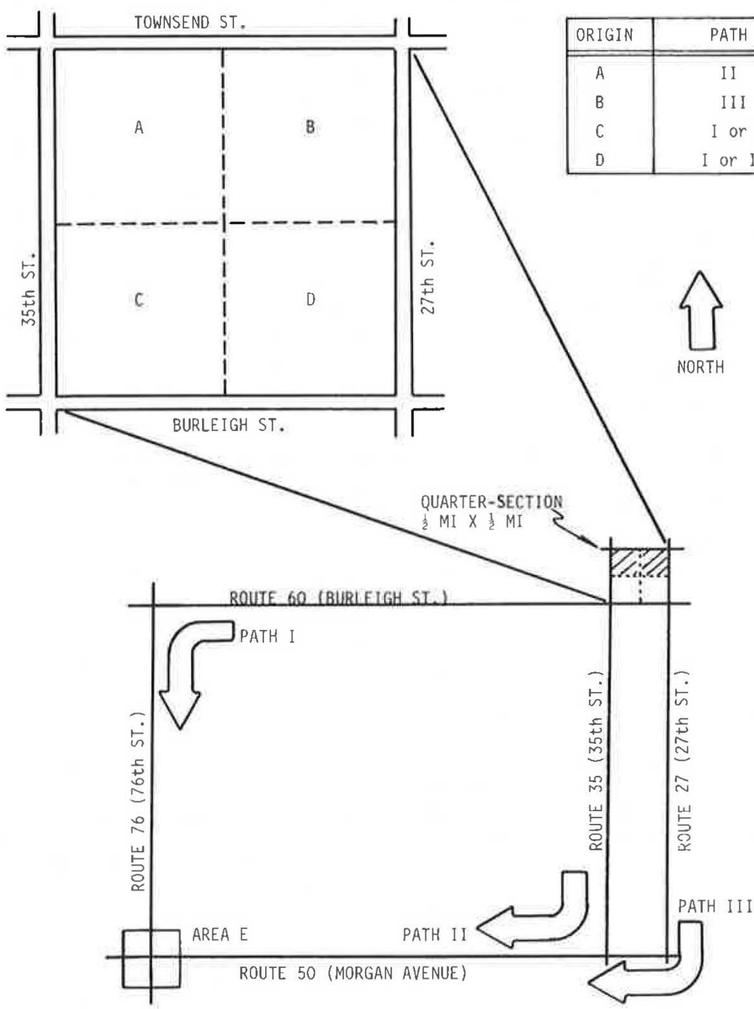
COMPARISON WITH OTHER AVAILABLE MODELS

A direct comparison with UTPS is not possible, because multiple-path assignment has not been implemented on it. At best it can be discussed why an all-or-nothing trip assignment, which ignores service-area considerations, is inadequate for projections of ridership on individual routes. This will be done by example.

In 1981 the Milwaukee County Transit System, in cooperation with the Southeastern Wisconsin Regional Planning Commission, ran several projections of system ridership by using UTPS. They had originally hoped to use the projections for route-level planning but backed off when, in spite of all best efforts, it was found that the current year projections failed to consistently match actual passenger loads. The root-mean-square (RMS) error in route ridership was 43 percent of the average ridership, even though projected total ridership for the system was in good agreement with actual figures.

It stands to reason that multiple-path trip assignment would have somewhat improved the route-level results had such an algorithm been available. But any multiple-path trip assignment algorithm would have been greatly handicapped by a zone structure that was originally designed for highway planning. Zones were based on quarter sections and many zones were as large as full sections (i.e., 1 mi^2). Major arterials, and consequently bus routes, follow quarter-section boundaries. Thus, zonal centroids were as much as 0.50 mi from transit routes. Under these conditions it would be nearly impossible for an algorithm to determine which portions of zones are within the service areas of connecting routes.

Figure 4 shows a portion of the UTPS network in Milwaukee. The origin zone is a quarter section and is bounded by three routes: Route 60 running east and west and Routes 27 and 35 running north and south. Those who wish to reach Area E on Morgan Avenue (Route 50) have an apparent choice of three paths, even though at most two choices are available to anyone. Route 50, as will be discussed later, provides poor service, so an all-or-nothing assignment would throw all trips onto Path I. However, it is highly unlikely that persons in the hatched areas would use this path. Rather, those at A will use Path II, those at B will use Path III, and those at C and D will choose their path on the basis of relative service characteristics. Similar mistakes in allocating trips to paths are repeated at hundreds of other zones. For a route like Route 50, the mistakes are cumulative, leading to a systematic under-



ORIGIN	PATH
A	II
B	III
C	I or II
D	I or III

FIGURE 4 Partial UTPS network, Milwaukee, Wisconsin.

counting of ridership and a redistribution of link loads. This tendency for all-or-nothing assignment to be particularly unkind to routes with relatively poor service characteristics was repeatedly seen in the Milwaukee case study. The 14 routes with the lowest actual volumes had an RMS error of 75 percent of average ridership. Most of this error was due to underprediction. Specifically, UTPS allocated to Route 50 only 9 percent of its actual ridership and to Route 15 (a considerably larger route to be discussed later) 59 percent of its actual ridership. The magnitude of this error is such that a comparison of predicted link loads to actual link loads would be meaningless.

From the viewpoint of total system ridership these problems are of little importance. In this case the planner would only be interested in the fraction of all trips that use transit, and a misallocation of transit trips away from poorer routes would affect total ridership only if levels of service varied greatly among alternative paths. However, at the route level these misallocations are costly. For the previously cited example, because it has lower headways, Path I is allocated all of the trips, although it should be allocated at most 50 percent of possible trips.

These problems could be mitigated by a much finer division of zones, but the issue persists of what to do about overlapping service areas. An improper placement of an important trip generator could have major implications for predicted riderships on nearby routes. It is desirable that multiple-path trip assignment be implemented in full-system models, but the algorithm will not be effective unless the zone system closely matches service areas of routes. This is not as easy as it sounds because a minor realignment of a single route could necessitate a complete reconstruction of the entire zone system. TRFM avoids these problems by requiring a custom zone system for each route that is analyzed. Such a procedure would be impractical when a full-system model was used in a large city.

A more practical, multiple-path algorithm for a full-system model could be implemented by following essentially the same steps as those of TRFM. First, there would necessarily be a hierarchy of three zone systems: very fine, coarse, and medium. The very fine zone system would be used for organizing demographic data and perhaps for calculating trip generation. The coarse zone system would be needed for trip distribution and mode split and would be necessarily consistent with the highway network. The medium zone system would be used for trip assignment. The size of the medium zones would be small by current standards, because each zone has to be unambiguously assignable to the service area of a single route or assignable to the overlapped service areas of intersecting routes. On the basis of experience with TRFM, the zone would be approximately 0.25 mi^2 and would be centered on routes.

Second, trips would be assigned by order of the number of transfers required. Trips that can be satisfied without transferring would be assigned first. Almost all these trips (e.g., those with both ends exclusively within the service area of a single route) have only one possible path. The few remaining zero-transfer trips can then be split among the routes that share the two trip ends. Trips that can be satisfied with a single transfer, having at least one end exclusively within the service area of a single path, would be assigned to that path. The remaining single-transfer trips can be split among the small number (usually two) of alternative single-transfer paths. Trips that can only be satisfied with more than one transfer can be assigned to the shortest path, because there are so few multiple-transfer trips.

ILLUSTRATION OF THE SERVICE-AREA CONCEPT

The first transit system analyzed by using TRFM was that located in Racine, Wisconsin (8). The bus system in this small community was of the radial type, with little consequential overlapping of service areas. All but one route met in the downtown at a common transfer point. Because of the structure of the Racine system, all-or-nothing trip assignment was more than adequate. Calibration of the model was accomplished by varying a single parameter in the logit mode-split equation. This single parameter is a coefficient on total trip disutility that has been converted to units of in-vehicle time. It would have the same value as the coefficient on the in-vehicle time variable in a more traditionally calibrated logit equation. All other parameters were taken from other published studies that were conducted in or around southeastern Wisconsin. Once this minimal calibration exercise was performed, TRFM produced consistently good estimates of actual ridership in all the test networks. Details of the Racine tests may be found elsewhere (8).

The Racine networks could only serve as a partial test of the service-area concept. Ridership on all the routes was of roughly the same magnitude and there was too little circuity. A far more complex test for the concept, to be reported here, was judged necessary. The Milwaukee County Transit System (MCTS) provided such a test. MCTS is a gridded system with a few radial routes, emanating from the central business district (CBD). It serves a large, heterogeneous metropolitan area, with performance of its various regular routes ranging from excellent (30,000 riders per day) to poor (less than 500 riders per day). The MCTS network is as difficult a test as could be devised for a ridership forecasting model.

Two routes on opposite sides of the performance spectrum were randomly selected from the Milwaukee system. The first, Route 50 along Morgan Avenue, is at the fringe of the urban area and largely serves residences, a large high school, and a subregional shopping center. Its total ridership is low (about 1,200 riders per day), barely meeting MCTS's minimum performance standard of 22 passengers per bus hour. As might be expected, Route 50 operates on long headways, typically 30 min. The second route, Route 15, is a larger-than-average route (12,000 riders per day). It serves the heart of the CBD, the University of Wisconsin--Milwaukee, a regional shopping mall, major industrial areas, strip commercial areas, and residences. Route 15 has 19 connecting routes. Being able to estimate ridership on both Routes 15 and 50 by using exactly the same set of parameters and with a minimum of recalibration would lend strong support to the service-area concept.

To the extent possible, data were prepared according to the procedures established in the Racine case study. Only once was it necessary to deviate from those procedures--in handling trips arriving at the high school on Route 50. Milwaukee has court-ordered busing to implement desegregation in its public schools. As a result, many of Route 50's patrons are captive riders, whose distribution of origins is dictated by artificially contrived boundaries. To accommodate this situation in TRFM, a separate subnetwork was created to specifically carry these riders. The appropriate number of captive riders was forced onto this subnetwork, yielding an estimate of the number of captive riders on each link of Route 50. Consequently, the trip generation, trip distribution, and mode-split steps of the model were overridden for these riders.

Ridership was estimated for both routes by using the original Racine parameters, producing results that were approximately 35 percent too low in both cases. This underestimation was expected, having been seen in other tests on the Milwaukee system. Racine is a much smaller city, has relatively fewer individuals of the type that normally patronize transit, places less dependence on its transit systems for busing students, and has a newer transit system with a lower overall level of service. In short, people in Racine have less of a predisposition to travel by transit. Nonetheless, this completely hands-off application of TRFM to Milwaukee by using another city's parameters did significantly better on the sample routes than did UTPS, which was specifically calibrated for Milwaukee.

In order to provide a better fit to the actual data in Milwaukee, the single mode-split parameter was adjusted until the percentage error in total ridership on both routes was minimized. The other 29 parameters in the model were left alone. Because of this adjustment, there is less interest in how well TRFM predicts total ridership than in seeing how well the model represents link loads and in observing consistencies between the two routes. A good match between actual and estimated link loads can only be achieved when boardings and alightings are accurately predicted everywhere along the route.

The actual and estimated load profiles for Routes 50 and 15 are shown in Figures 5 and 6, respectively. These link loads are the average of the two directions of travel. Overall, the model matched the load profiles well. The peak load points are properly located, and the peak link loads are accurate. The RMS error in link load for Route 50 was 21 percent; the RMS error in link loads for Route 15 was 9 percent. These are small errors, considering the low level of data aggregation represented by link loads. The largest errors in link loads occur near the ends of Route 50. These errors can be attributed to the peculiar behavior of riders on routes with very long headways in cities with cold weather. Many riders who arrive at their stop early catch their bus going in the opposite direction and ride through the layover. Available on-off counts revealed the number of riders behaving in this fashion, so the double counting of these riders could be eliminated from

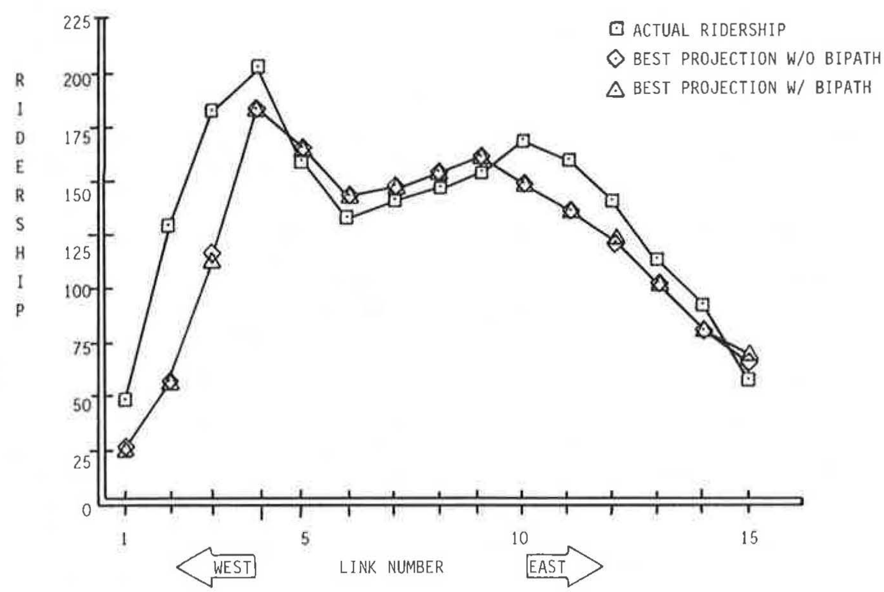


FIGURE 5 Ridership comparisons for Route 50 (Morgan Avenue).

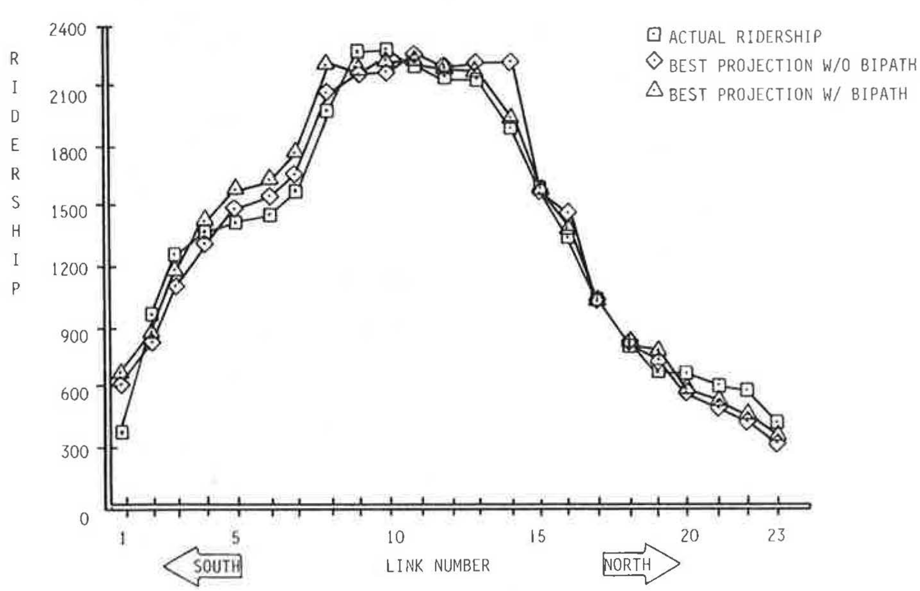


FIGURE 6 Ridership comparisons for Route 15 (Oakland Avenue).

total ridership, but it was impossible to determine where these riders actually boarded the buses. Consequently, it was not possible to properly simulate their behavior.

Of particular importance is the good fit to data on both routes by using an identical set of parameters. The model is unbiased with respect to the performance of route, demonstrating that the service-area concept is working well. Although only two routes were rigorously tested, these routes were so dissimilar that any such bias should have been obvious. As expected, the model is free from the misallocations that were so evident with all-or-nothing assignment in UTPS.

Much of the complexity in implementing the service-area concept stemmed from the need for multiple-path trip assignment. It is only logical, then, to question the importance of having this type of assignment. One way to do this is to eliminate tertiary routes from the networks and recalibrate the model by again adjusting the mode-split parameter. When this was done, the mode-split parameter increased by only 9 percent, far less than the change of 35 percent necessary to recalibrate the Racine model for Milwaukee. The load profiles (Figures 5 and 6) were virtually unchanged. On the basis of this exercise, two observations can be made. First, only about 20 percent of trips are affected by the bipath algorithm, so eliminating it does not obviously distort the results. Second, the bipath algorithm affects results in nearly the same way as an important parameter in the logit equation of the mode-split step. Thus, it would be difficult to determine the need for the bipath algorithm simply by observing how well it reproduces total ridership and link loads. The strongest arguments for the bipath algorithm are that (a) it is a logical extension of the service-area concept, (b) it is consistent with the way riders behave on transit networks, and (c) there is little extra data preparation required. The bipath assignment could be eliminated or downgraded, but at the expense of a bias in an important parameter, which could later have an unfavorable effect on forecast validity.

Routes 50 and 15 provided one additional test of the service-area concept: ease of data preparation. Both networks contained nearly every route in the Milwaukee system. Route 50 had 120 explicit nodes and 208 explicit links; Route 15 had 149 explicit nodes and 266 explicit links. Total data preparation time was approximately 2 person months. However, almost all of this time was spent reducing socioeconomic and demographic data for the entire Milwaukee urban area to a compatible level of spatial aggregation. Of course, the data previously prepared for the UTPS runs were totally useless. Each route took only about 3 days to complete once this initial preparation had been done. The portion of data specifically needed for the bipath algorithm required only about 2 hr of preparation time. This amount of data preparation is considerably less than that needed for a UTPS run, but TRFM certainly does not qualify as a quick-response technique. Data preparation time is more on the level of what would be needed for careful application of a direct-demand model.

DISCUSSION AND CONCLUSION

There has long been an inconsistency between the way in which transit planners designed service and the way in which the more sophisticated simulation models predicted ridership. This inconsistency can be attributed largely to historical accident; the simulation models were pioneered by highway engineers who found little use for the concepts of service area or transferring. In highway networks, either access existed or it did not. Where highway access did exist, highways were ubiquitous. Motorists had little impediment to switching streets in order to reach destinations as quickly as possible. The simulation models represented this situation rather well. Zones of nearly any size could be created at nearly any location and collapsed into a single point. Nodes could be placed at all arterial intersections to permit turning. And highly efficient traffic assignment algorithms could be written to find alternative paths through the network and estimate link volumes. However, transit networks behave in an entirely different way. They are neither ubiquitous nor encourage switching between routes, facts fully appreciated by transit planners.

The large number of attempts to forecast transit ridership with direct-demand models stands as a testament to transit planners' dissatisfaction with models developed for highway planning. The research reported in this paper demonstrates that there is a middle ground, at least for the single-route case. It is possible to retain the sophistication of the four-step models while giving full respect to the peculiarities of transit networks. In addition, it appears that this middle ground can exist for full-system models, too.

Unfortunately, the specific test of the bipath algorithm was inconclusive. Because of the small fraction of trips affected by the bipath algorithm in the single-route case, the algorithm would be extremely difficult to validate unambiguously in any system, including those with extensive origin-destination data. Even though an algorithm at this level of sophistication may not be essential to ridership forecasts on single routes, it is hard to imagine how a less sophisticated multiple-path algorithm could be successfully implemented in the full-system case. An accurate full-system model must at least be capable of determining which land parcels are in the service area of any given route, which land parcels are shared by routes, and what alternative paths exist for trips between parcels that are shared. This determination is made in the single-route case largely through the decomposition of the system into a hierarchy of trees; the bipath algorithm plays only a minor role in that determination.

The authors' experience in forecasting ridership by employing the service area concept confirms its utility and flexibility. Not only is it useful for simulating conventional fixed-route service, it can be extended to handle park-and-ride service, skip-stop service, downtown shuttles, and, as seen earlier, forced busing. However, implementing the concept may require a complete restructuring of existing data bases, as well as major enhancements to existing simulation models.

ACKNOWLEDGMENTS

Raw ridership data and UTPS results were provided by the Milwaukee County Transit System. Milwaukee Public Schools provided information on enrollments and district boundaries. The authors wish to thank Sam Seward, Senior Planner at MCTS, for his many insightful interpretations of the data, UTPS results, and TRFM results. They would also like to thank Ron Jensen-Fisher at UMTA for help in clarifying the role of UTPS in forecasting ridership on individual routes. This research was supported by UMTA through its University Research and Training Program and by the University of Wisconsin Extension.

REFERENCES

1. H.S. Chadda and T.E. Mulinazzi. A Transit Planning Methodology for Small Cities. Transit Journal, Vol. 3, No. 2, Spring 1977, pp. 19-40.
2. J. Hsu and V.H. Surti. Framework of Route Selection in Bus Network Design. In Transportation Research Record 546, TRB, National Research Council, Washington, D.C., 1975, pp. 44-57.
3. Peat, Marwick, Mitchell and Co. Analyzing Transit Options for Small Urban Communities--Analysis Methods. Report PPM.76.1.1. UMTA, U.S. Department of Transportation, Jan. 1978.
4. Peak, Marwick, Mitchell and Co. Simplified Aids for Transportation Analysis--Estimating Annual Ridership and Operating Expense for Fixed Route Bus Systems in Small Urban Areas. Report UMTA-IT-06-9020-79-3. UMTA, U.S. Department of Transportation, Jan. 1979.
5. J.H. Shortreed. Bus Transit Route Demand Model. In Transportation Research Record 625, TRB, National Research Council, Washington, D.C., 1977, pp. 31-33.
6. M.H. Rapp and G.D. Gehner. Criteria for Bus Rapid Transit Systems in Urban Corridors: Some Experiments with an Interactive Graphic Design System. In Highway Research Record 455, HRB, National Research Council, Washington, D.C., 1973, pp. 36-48.
7. A.J. Horowitz. Transit Ridership Forecasting: A Microcomputer Application. Transportation Quarterly, Vol. 38, No. 2, April 1984, pp. 297-310.
8. A.J. Horowitz. Simplifications for Single-Route Transit Ridership Forecasting Models. Transportation, Vol. 12, 1984, pp. 261-275.
9. A.J. Horowitz. The Design of a Single-Route Ridership Forecasting Model. In Transportation Research Record 980, TRB, National Research Council, Washington, D.C., 1984, pp. 112-117.
10. M.A. Turnquist, A.H. Meyburg, and S.G. Ritchie. Innovative Transit Service Planning Model that Uses a Microcomputer. In Transportation Research Record 854, TRB, National Research Council, Washington, D.C., 1982, pp. 1-6.
11. R.B. Dial. Demand Forecasting for New Options and Technology: Resource Paper. In HRB Special Report 143: Urban Travel Demand Forecasting, HRB, National Research Council, Washington, D.C., 1973.
12. R.B. Dial. A Probabilistic Multipath Traffic Assignment Model Which Obviates Path Enumeration. Transportation Research, Vol. 5, 1971, pp. 83-11.
13. R.L. Tobin. An Extension of Dial's Algorithm Utilizing a Model of Tripmakers' Perception. Transportation Research, Vol. 11, 1977, pp. 337-342.
14. A.J. Horowitz. Subjective Value of Time in Bus Transit Travel. Transportation, Vol. 10, 1981, pp. 149-164.
15. A.J. Horowitz and D.J. Zlosel. Transfer Penalties: Another Look at Transit Rider's Reluctance to Transfer. Transportation, Vol. 10, 1981, pp. 279-282.
16. E. Stern. A Cognitive Scalar for the Public Transport Connection Matrix. The Professional Geographer, Vol. 32, No. 3, Aug. 1980, pp. 326-334.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.