A variety of applications of computer graphics to the field of transportation planning has been presented. The benefits realized through the work illustrated here include the ability to have traditional transportation planning programs produce graphics for analytic purposes; the capability for non-computer-oriented planners to select from a menu list and generate graphics for varied and specialized purposes; the potential for the experienced system analyst to interactively manipulate and analyze information in a manner not addressed by the standard models; and, finally, the production of graphics that are used for analysis and that may also be used in presentations and publications.

Even though there is a definite need for the development of specialized interactive graphics tools, especially for operations analysis, it is thought that low-cost, commercially available graphics packages can be effectively used to enhance the knowledge of transportation planners and their audience. The planning field is characterized by the complexity of its data sources and structures. The professional should make use of all tools available in order to realize the full potential of personnel, data, and models.

In conclusion, it is recommended that every effort be made to explore potential applications of existing software packages. In addition, efforts should be made to train transportation professionals in basic uses of computer graphics, both batch and interactive, in order to take the best advantage of the enormous potential.

REFERENCES


Publication of this paper sponsored by Committee on Computer Graphics and Interactive Computing.

Design of a Single-Route Ridership Forecasting Model

ALAN J. HOROWITZ

ABSTRACT

The transit ridership forecasting model (TRFM) has been designed to overcome many of the serious obstacles to implementation presented by previous methods of forecasting ridership on a single route. TRFM simplifies, optimizes, and repackages conventional ridership forecasting techniques to make the job of the planner as easy as possible. The model exploits the advantages of a popular, modest-sized microcomputer (e.g., animated color graphics), but it also deals effectively with inherent limitations of microcomputers (e.g., small memory and slow calculation speed).

Travel demand estimation is considered an integral part of transportation planning. However, despite two decades of model development, few transit planning agencies use the best available methods for ridership estimation. Transit planners instead often substitute rules-of-thumb or intuition in determining the impacts of service changes. There are many reasons for this gap between theory and practice, but one major reason is that virtually all available computer packages for ridership estimation require more data, more computer expertise, more equipment, and more time than planners generally possess.

The transit ridership forecasting model (TRFM) attempts to put sophisticated forecasting methodology into the hands of transit planners. TRFM greatly simplifies ridership estimation from the planner's viewpoint and attempts to retain the accuracy of mainframe models. In TRFM, simplification takes three forms: (a) eliminating mathematical procedures that are unnecessary to its sole objective of single-route ridership estimation; (b) designing the input procedures to remove, as much as possible, the burden of data preparation; and (c) organizing the program so that planners may easily customize the model to their needs.

TRFM is a fully interactive, color graphics program that just estimates ridership on a single route. The design of TRFM retains the salient parts of mainframe models such as the urban transportation planning system (UTPS) but exploits the advantages of microcomputers, specifically the Apple II+/IIe.
However, the limitations of the Apple II and similar computers (limited memory, slow computation speed, long disk access time, and limits to the resolution of color graphics) have resulted in several compromises. TRFM serves as a good reference point for what may and may not be accomplished on a modest-sized microcomputer and indicates the types of trade-offs that must be made.

The research project that generated TRFM had two purposes: the first was to design a microcomputer program that would overcome the implementation obstacles of previous models; the second was to determine whether planners would adopt the methodology when it was made available. The first purpose has been satisfied, but the second is only now being pursued.

MINIMUM DESIGN STANDARDS

A list of "musts" was developed at the beginning of the project to develop TRFM. It was believed that if the model failed to accomplish any one of these points, its applicability would be so seriously limited that few planning agencies would want to use it. For TRFM to be a useful tool rather than merely an academic exercise, the following had to be accomplished:

1. Sensitivity to service changes on the route of interest, with the implication that ridership estimates would be based on a mode-choice model of sufficient breadth to reflect virtually all service changes contemplated by planners.
2. Familiarity to quantitatively oriented planners; it was believed that if TRFM were too experimental it would not become widely accepted as a planning tool.
3. Compatibility with the 1980 Census because few transit planning agencies have the resources to collect and analyze vast amounts of travel and demographic data; TRFM needed to efficiently use existing data sources.
4. Quick computation with an outside limit of 30 minutes. Any program running longer than that could hardly be called interactive.
5. Easy-to-understand structure, so that even a planner with only passing knowledge of ridership forecasting could produce reasonably good results.
6. Enough sophistication to satisfy planners who possess extensive knowledge of the principles of ridership estimation; the model could not insult their intelligence by making too many assumptions nor by being overly restrictive in how calculations were handled.
7. Attractive packaging—TRFM had to measure up to other software products in terms of user friendliness and polish. Much of the commercial software now available is so well written that expectations of owners of microcomputers are rapidly rising. Line-by-line input procedures that once were acceptable on multiple-user systems are now unacceptable on single-user systems.
8. Performance of the preceding seven items entirely within the restrictions of a modest-sized microcomputer, because this is what is most readily available to transit planners.

These were the minimum requirements; all of them were met. There was also an additional list of goals, many of which could not be fully realized:

1. Color graphics to support all data input. Proper use of color graphics can make the tedious and boring process of data entry more enjoyable. This goal was fully achieved.
2. Parameters that are readily understood, so that they can be grasped by people not familiar with ridership estimation. This would be necessary if the model were to be trusted by users or explained to others. This goal was achieved for all but two parameters, which ended up with units of "utils/minute."
3. Applicability to cities of every size and geography. Although a microcomputer model might be most helpful to transit agencies in small cities, it probably would be first adopted by larger systems that had at least one full-time planner. This goal has been met, although systems in large cities are required to provide (proportionally) slightly more data than systems in small cities.
4. Computation time of 5 minutes or less. A short computation time enables planners to adjust default parameters by repeatedly running the model until ridership estimates consistently match current levels. Formal, statistical calibration can thereby be avoided. This goal was achieved only for shorter-than-average routes.
5. Accuracy of results to within 1 percent of those obtained from mainframe models. This goal was surpassed; in many planning situations TRFM is just as accurate as mainframe models.

In the following sections how well the minimum requirements were met and the extent to which the additional goals were achieved are discussed more fully. But first it is necessary to describe some of the mathematical portions of TRFM.

ANALYTICAL STRUCTURE OF TRFM

Mathematically, TRFM is an implementation of many of the results from a major investigation of riders' evaluations of the time spent in travel (1,2). Psychological scaling was used to determine the relative values placed on various elements of bus transit, walking, and automobile trips. In addition to providing many default parameters for TRFM, this study identified which elements could be ignored and which must never be ignored in a ridership estimation procedure. Those that can be ignored, resulting in a more simplified model, will be discussed first.

Network Simplifications

Current microcomputer models for forecasting ridership evolved from models to estimate highway traffic volumes. But people use transit networks very differently from highway networks. An automobile driver has nearly unlimited freedom to choose a path that minimizes travel time between origin and destination. A transit rider has a very limited choice of paths. It has been established that riders dislike transferring and avoid transfers whenever possible (2). If an assumption is made that, by choice, riders make at most one transfer, then there is only one possible path for the vast majority of trips that are known to use one particular route. Consequently, it is necessary to produce only a network containing the route of interest and all immediately connecting routes. Furthermore, in gridded and radial systems these connecting routes almost never intersect the route of interest more than once, so circuits in the network can be avoided. TRFM insists that such circuits are not present, so that overall network trip times can be rapidly calculated.

The process of creating a network for TRFM is similar to Dial's (4) windowing and focusing concept. The network focuses on the route of interest, which is shown in considerable detail; connecting
routes are shown in far less detail. A good network focusing on a single route can almost always be shown in less than 80 nodes, TRFM's current maximum. Consequently, a TRFM network requires a small fraction (5-10 percent) of the data that would be necessary without windowing and focusing.

TRFM does not require a highway network. It has been shown that automobile speeds are roughly constant multiples of bus speeds, regardless of traffic conditions (5). When it is reasonable to assume that automobile trips follow the path of bus routes, TRFM calculates automobile trip time as a fixed fraction of bus running time. This permits eliminating most highway network data—data that are generally not available to transit planners.

There are two important instances in which automobiles cannot be assumed to follow bus routes: when there are "U"s or large one-way loops. Automobile drivers will generally travel shorter distances than bus riders between two points on a loop or a U. TRFM employs a scale drawing of the network, produced with the assistance of graphics routines, to detect the existence of a loop or a U and then shortens automobile trips by the appropriate amount.

In gridded networks, as opposed to radial networks, a significant number of riders may reach areas around the route of interest without ever going on that route. To allow for this behavior, TRFM discards trips that start at a point on a connecting route and end at the transfer point between that route and the route of interest. TRFM also permits the planner to further discard a fraction of all trips that start or end at a point on a connecting route. This fraction is normally zero in small cities, but it can be as high as 0.5 for systems with closely spaced or perfectly gridded routes. The effect of discarding trips is similar to what would be achieved with stochastic traffic assignment.

In the significant savings in required data, these network simplifications have computational advantages. The description of the network can be compactly stored, preserving memory and disk space. Algorithms for finding node-to-node trip times and link loads can be optimized for a network without circuits. This is important because of the need for quick results.

Trip Distribution

Trip distribution equations vary greatly in complexity but have similar effects on estimates of total route ridership. A doubly constrained gravity equation (e.g., where the trip table must be consistent with trips attracted to every zone as well as trips produced at every zone), which is preferred by highway planners, has more than twice as many parameters as nodes. These parameters must be iteratively determined each time the model is run. A singly constrained gravity equation (i.e., where only trip productions need be consistent with the trip table) does not require iterative recalibration. Consequently, the equation selected for trip distribution was singly constrained.

In addition to speed, singly constrained trip distribution equations have other advantages because total trip production need not agree with total trip data (as it is the practice in highway volume estimation) just to have a slightly more accurate trip table. Limiting TRFM to a single trip purpose has obvious computational advantages, but it also greatly reduces data requirements at the trip generation step.

Trip Generation

Single purpose trip productions and trip attractions are based on demographic and employment information. The default trip generation equations are patterned after those developed by the Southeastern Wisconsin Regional Planning Commission (7). Trip production is based on numbers of households falling into specified size and automobile availability categories, in a form compatible with 1980 Census data. Trip attraction is derived from the number of service employees, nonservice employees, and students.

Mode Split

TRFM, following standard practice, divides transit riders into two categories: captive and choice. Captive riders are assumed to be insensitive to service variables and constitute a fixed fraction of all trips produced in the service area. Choice riders are divided between transit and automobile by a binary logit equation that has terms for initial waiting time, walking time, waiting penalty, and transfer penalty. The selection of terms was based on the previously mentioned study of evaluations of time spent traveling. Because of the availability of transferable parameters (8), the binary logit equation was preferred to other mode-split equations.

INTERACTIVE COLOR GRAPHICS AND OTHER VISUAL DISPLAYS

Graphics has been a positive force in transit route planning, dating back to IGTDs (9) and NOPTS (10). However, extensive use of graphics was impractical until the introduction of the microcomputer because of the specialized hardware that was previously required. Nonetheless, these early efforts demonstrated that an interactive graphic capability could make data preparation and manipulation easier, faster, and more pleasurable. The design of TRFM's graphics routines was heavily influenced by these pioneering efforts.

The original goal of the graphics routines in TRFM was to allow the planner to enter a drawing of portions of the transit network. Then, whenever TRFM needed a piece of data, the program would prompt the user by pointing to the right node or link. Conversely, when the planner wanted to see a piece of previously entered data or a calculated result, he could get it by pointing to the correct link or node on the display. The drawing becomes a communications device, eliminating the need for explicit link and node numbers (or similar identification).

When the mathematical model steps of TRFM were refined, another important use was found for the graphics information. As previously mentioned, TRFM
does not use a highway network; it infers automobile trip times from bus running times. Consequently, it is essential for TRFM to know when automobile trips roughly follow bus routes and when they do not. The graphics display, if the network is drawn to scale, has sufficient information to make this determination and to allow for calculation of any necessary corrections in automobile trip time.

The standard Apple II+ has been chosen for its popularity among transit agencies, not because it has outstanding graphics features. Apple II+ graphics capabilities are the minimum necessary for network display (though the Apple is better in this regard than many other popular brands of microcomputers). The design of the graphics routines had to overcome the limited color resolution (140 horizontal x 192 vertical), the slowness of the BASIC graphics statements, aggravating and unexpected color changes, and the lack of a cursor or cross hairs. In addition, the graphics routines had to work on a monochrome display as well as a color one and make allowances for computers without light pens, paddles, or joysticks.

Figure 1 shows the design of the graphics display. The arrows indicate a position within the display. They can be animated by paddles or keyboard. The strip of symbols along the right side of the display serves as the menu from which various functions can be selected: plotting nodes (square dot), plotting links (vertical line), deleting nodes and links (D), starting a new network (W), continuing to the numerical data input step (C), and printing the display (1 and 2).

Getting the graphics to work was straightforward; BASIC is a good language for graphics if speed is not important. Getting the graphics to work quickly required substitution of machine language subroutines for BASIC statements plus the use of some little-known tricks (e.g., hiding some of the program from itself) to dramatically improve BASIC’s execution time.

When the network has been drawn, data can be entered for each node and link. TRFM prompts the planner for information by pointing to each link with an arrow and highlighting each node in a contrasting color. As the data are assembled, nodes turn different colors depending on their status (route of interest, connecting route, transfer point) and oneway links are marked.

The graphics routines have the feel of an arcade game. Ridership estimation is serious business, but it need not be boring. The use of color, sound effects, and game paddles with a fire button make plotting the network far more enjoyable and no less accurate than it would be if the drawing were produced by inputting screen coordinates onto a quiet, monochrome display.

Other Aids to Data Preparation

TRFM employs 10 on-screen work sheets to aid data preparation. Any time TRFM prompts for a number, a work sheet is available. At the minimum, a work sheet displays the current value and gives the planner an opportunity to change it. All work sheets accept both numbers and arithmetic expressions as inputs. Some work sheets (e.g., those used for calculating trip production and trip attraction) allow the planner to reference statistical equations. When the calculations on a work sheet call for parameters, a "parameter page" is readily available. It displays the necessary parameters and gives the planner an opportunity to change them. TRFM is provided with a complete set of default parameters. Examples of a work sheet and an associated parameter page are shown in Figures 2 and 3.

Some model designers have adopted the attitude that parameters should be difficult or impossible to change. They fear that, if parameters can be easily modified, inexperienced planners will botch forecasts. The philosophy behind TRFM is considerably different. Users are encouraged to adjust default parameters as needed. TRFM remembers any changes and incorporates them in subsequent forecasts. Over time, and without formal calibration, parameters can be refined until highly accurate forecasts are consistently achieved for a given system.

The work sheets and parameter pages were patterned after those of VisiCalc. The popularity of VisiCalc stems largely from two features: the ability to see all numbers as if they were written on a sheet of paper and the instantaneous recalculation that occurs when any number is changed. These capabilities were incorporated into TRFM’s work sheets.
When data have been entered, TRFM allows the planner to quickly view the entries, change any value that might have been typed improperly, and modify the network. A network can be saved on disk for later reference.

**COMPUTATION TIME**

It should be clear from earlier discussions that computation time was of great concern. An initial goal was a computation time of 5 minutes. A substantially longer computation time would inhibit the refinement of parameters and of data preparation techniques by simply running the model repeatedly. When the first draft of TRFM was prepared, calculation for a full-sized network took 3 hours. This greatly exceeded the maximum computation-time standard set at the beginning of the project.

The long computation time was caused by the BASIC interpreter (a slow way to execute any program) and exacerbated by structured programming techniques. Compiling the program produced a sevenfold improvement in speed. Additional time savings were achieved by three measures: extensive rewriting to combine as many steps as possible, substitution of integer arithmetic for floating-point arithmetic, and reduction of the dimensionality of frequently referenced arrays. These actions reduced maximum computation time to 18 minutes. The set of test networks for Racine and Milwaukee, Wisconsin, averaged about 13 minutes. Further improvements in computation time were achieved by installing some additional hardware (a widely available 8088 coprocessor board) in the Apple, which cut average computation time to 8 minutes. Brute-force refinement of parameters was made practical by these improvements in computation time.

Unfortunately, these speed increases carry a price tag. It is now more difficult for a user to customize those portions of the program that have been compressed and compiled.

**CALIBRATION ISSUES**

All forecasting models need some calibration. In recent years, calibration has become synonymous with statistical estimation, but other methods do exist. Statistical calibration is essential when models become so large that only a portion of parameters can be determined at one time. But if a model has only a few parameters that require minimal adjustment, the model can be run repeatedly until the planner is comfortable with the accuracy of the forecasts.

This brute-force calibration will not work properly unless the program has been specifically designed for that purpose. TRFM includes several features to facilitate brute-force calibration. The TRFM runs quickly; (b) TRFM allows for transferable parameters so that the number of truly unknown parameters can be held to a minimum; (c) parameters are easily accessed and well explained, both on screen and in written documentation; (d) where possible, parameters are presented in ways that have physical, economic, or behavioral meanings; (e) default parameters are provided to serve as a reasonable starting point; and (f) results are consistent with the way transit agencies collect ridership data (total ridership, revenue ridership, on-off counts, and check-point loads) to aid comparison between model results and data from the existing system.

**ERROR ANALYSIS**

TRFM was subjected to extensive error analysis to determine if the simplification assumptions could substantially affect the results. The error analysis will only be summarized here. Tests were performed on three routes in Racine, Wisconsin, to determine if the following four procedures would lead to significant error: (a) adopting a no-multiple-transfer rule, (b) approximating automobile trip times as fractions of bus running times, (c) correcting for loops and U's from the scale drawing of the network, and (d) showing connecting routes in less detail than the route of interest. In each case, errors were less than 0.1 percent in total ridership, or less than one rider.

More worrisome are errors due to misspecification of model parameters. All forecasting models, not just TRFM, are vulnerable to inappropriately selected parameters. However, TRFM confronts this problem directly by presenting parameters in easy-to-understand terms and by giving the planner ample opportunity to study and modify them as necessary.

Another source of error, almost never discussed, is sloppy work. It is only human nature to be more careful in preparing a few pieces of data than in preparing many pieces of data. Consequently, a model that is optimized in its data requirements, like TRFM, may be much more accurate than a complex model, which appears to be better on paper.

**COMPARISON WITH OTHER MODELS**

The impetus for development of TRFM came from transit operators who were disappointed that the interactive graphics transit design system (IGTDS) could not be implemented on readily available hardware. IGTDS forecasts ridership and other impacts of bus rapid transit. In particular, it handles situations in which there are many origins of trips but only one destination, such as park-and-ride or freeway flyer service. In structure IGTDS is similar to TRFM, employing graphics to facilitate data input and basing ridership on the traditional four-step modeling procedure (trip generation, trip distribution, mode split, trip assignment). TRFM is much broader than IGTDS in the types of routes it can analyze, but TRFM does not provide as comprehensive an evaluation of route performance.

One recently developed computer program that in-

<table>
<thead>
<tr>
<th>Rate</th>
<th>0 Auto HH/S</th>
<th>1 Auto HH/S</th>
<th>2+ Auto HH/S</th>
<th>0 Person HH/S</th>
<th>1 Person HH/S</th>
<th>2+ Person HH/S</th>
<th>0/4 Person HH/S</th>
<th>1/4 Person HH/S</th>
<th>2+ 4 Person HH/S</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.71</td>
<td>.06</td>
<td>-.58</td>
<td>-1.425</td>
<td>-1.38</td>
<td>.47</td>
<td>.315</td>
<td>.897</td>
<td>.794</td>
</tr>
<tr>
<td>2</td>
<td>-.36</td>
<td>.47</td>
<td>.38</td>
<td>.3622</td>
<td>.5377</td>
<td>.288</td>
<td>.504</td>
<td>.571</td>
<td>.191</td>
</tr>
<tr>
<td>3</td>
<td>.71</td>
<td>.06</td>
<td>-.58</td>
<td>-1.425</td>
<td>-1.38</td>
<td>.47</td>
<td>.315</td>
<td>.897</td>
<td>.794</td>
</tr>
</tbody>
</table>

**FIGURE 3 Example of a parameter page: trip production.**
vites comparison with TRFM is the transit operations planning (TOP) model system developed by Turnquist, Meyburg, and Ritchie (12). Like TRFM, TOP is a microcomputer program that can estimate ridership on a single route; it is based on the traditional four-step procedure. TOP is more comprehensive than TRFM, because in addition to ridership estimation it estimates level-of-service variables and performance indicators. In structure, TOP is similar to mainframe models and requires a big microcomputer (an Apple III with 256K bytes of memory and three disk drives).

The most striking difference between TRFM and TOP is the emphasis placed on user needs. TOP assembles several sophisticated mathematical steps (e.g., an optimization procedure for reconciling trip tables with on-off counts, consideration of stochastic variations in bus arrival times, and a procedure for establishing equilibrium between levels of service and ridership) into a complex package. However, less effort was placed on streamlining the process for the planner. TOP appears to be best suited for planners who are well versed in the principles of ridership estimation. In contrast, TRFM dispenses with relatively less-important data and procedures, placing greatest emphasis on helping the planner enter the data and extract the results. In TOP a highly complex step is permitted only when both theory and existing data indicate the step is required for the specific task at hand.

DISCUSSION

Computers have not fulfilled their early promise in transit planning. They perform calculations quickly and they display intricate drawings, but they cannot collect and cull data or judge whether a particular equation is appropriate. There have been many attempts to create more elaborate mathematical models but relatively few recent attempts to improve the computer-human interface.

The introduction of microcomputers presents a new opportunity for model designers. Costs are now so low that a major percentage of computer resources can be devoted to making the job of the planner easier. Interactive techniques for doing this (e.g., color graphics, menus, and work sheets) are all well developed. They merely require adaptation to the particular needs of transit planners.

It is essential that planners understand any modeling used. However, the state of the art has now reached a point where only people well versed in operations research and statistics can properly apply the newest techniques. Transit agencies, weighing the advantages and disadvantages of mathematical models, often will opt to do without forecasting rather than use techniques that few of their people understand. If any model is to be widely adopted, its assumptions need to be made explicit; its rationale for selection of particular equations must be made obvious; and, desirable but less important, complexities should be either scrapped or presented as readily bypassed options. Furthermore, the interactive features of the computer program should serve to illuminate the process, both as data are entered and as alternatives are tested.

The acid test in determining if a program is sufficiently user friendly is its acceptance or rejection by transit operators. For this reason some emphasis has been placed on publicizing and distributing TRFM. Unfortunately, but understandably, there is considerable resistance among managers of transit systems to mathematical models of any sort. Efforts are being made to explain the benefits of ridership forecasting and the ways in which a microcomputer can help. Transit managers, in turn, are now providing valuable feedback on how this and future programs can aid and improve transit planning.

ACKNOWLEDGMENTS

The transit ridership forecasting model was developed as part of a comprehensive program of applied research and outreach sponsored by the University of Wisconsin—Extension. The author thanks Edward Beimborn, Robert Schmitt, and David Cyra for many ideas that found their implementation in TRFM. Test networks were developed by Manoochehr Adhami and David Metzger. Data were provided by the city of Racine, the Milwaukee County Transit System, and the Southeastern Wisconsin Regional Planning Commission.

REFERENCES