TRANSPORTATION RESEARCH RECORD

491

Formerly issued as Highway Research Record

Interactive Graphics and Transportation Systems Planning

7 reports prepared for the 53rd Annual Meeting of the Highway Research Board

subject area
84 urban transportation systems

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

Washington, D. C., 1974

NOTICE

These papers report research work of the authors that was done at institutions named by the authors. The papers were offered to the Transportation Research Board of the National Research Council for publication and are published here in the interest of the dissemination of information from research, one of the major functions of the Transportation Research Board.

Before publication, each paper was reviewed by members of the TRB committee named as its sponsor and accepted as objective, useful, and suitable for publication by the National Research Council. The members of the review committee were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the subject concerned.

Responsibility for the publication of these reports rests with the sponsoring committee. However, the opinions and conclusions expressed in the reports are those of the individual authors and not necessarily those of the sponsoring committee, the Transportation Research Board, or the National Research Council.

Each report is reviewed and processed according to the procedures established and monitored by the Report Review Committee of the National Academy of Sciences. Distribution of the report is approved by the President of the Academy upon satisfactory completion of the review process.

The National Research Council is the principal operating agency of the National Academy of Sciences and the National Academy of Engineering, serving government and other organizations. The Transportation Research Board evolved from the 54-year-old Highway Research Board. The TRB incorporates all former HRB activities but also performs additional functions under a broader scope involving all modes of transportation and the interactions of transportation with society.

TRR 491
ISBN 0-309-02279-7
LC Cat. Card No. 74-12741
Price: $3.40

Transportation Research Board publications may be ordered directly from the Board. They are also obtainable on a regular basis through organizational or individual supporting membership in the Board; members or library subscribers are eligible for substantial discounts. For further information write to the Transportation Research Board, National Academy of Sciences, 2101 Constitution Avenue N.W., Washington, D.C. 20418.

CONTENTS

FOREWORD

This RECORD contains reports on the application of interactive graphics to systems planning and design and a series of papers focusing on techniques and tools to improve the transportation planning process.

Mead proposes an operational definition of the system planning process that can be used by highway agencies and state governments to display the state transportation planning process design to all interested groups.

LeBlanc, Morlok, and Pierskalla describe a new solution technique for the equilibrium traffic assignment problem. As stated by the authors, the algorithm promises to be efficient for finding the equilibrium on a network with hundreds of nodes inasmuch as its most difficult computational requirements are identical to those of the iterated capacity restraint simulation models of traffic assignment currently used.

Hocking describes a time-staged strategy approach as a means to improve the planning process. The approach guides transportation development in a given time sequence by seeking to develop first-stage actions that maintain a range of choices for future actions and that employ elements of decision theory that consider the potential relationship between series of actions over a long period of time.

Ruiter describes network equilibrium procedures now being made operational as a part of the UMTA transportation planning system. Procedures are described in the light of their theoretical and mathematical background.

Tomazinis discusses research under way that is aimed at suggesting new approaches in urban transportation planning and basically a new type of plan based on studies of efficiency, productivity, and quality. The author offers this as essential to transportation planning for the seventies.

Gur discusses INTRANS, a man-computer interactive system designed for real-time analysis of transportation and urban data. The use of INTRANS as an aid in the evaluation of alternative plans for the Chicago Area Transportation Study is described by the author.

Schneider and Porter discuss the use of an interactive graphics computing system in designing a bus rapid transit system. Five teams of students used interactive graphics to design a transit system, and, though none of the teams found a wholly satisfactory design, the average improvement in performance (design quality) for all teams was close to 50 percent (relative to their initial designs).

A FRAMEWORK FOR SYSTEM PLANNING PROCESS DESCRIPTION

Kirtland C. Mead, MITRE Corporation

In response to increasing process-oriented federal directions, such as the FHWA process guidelines, this paper proposes a definition of system planning in process terms. System planning is the collection of planning institutions, funding sources, and programming procedures that interact with society continuously to produce transportation policy and investment decisions over time. The paper then presents a framework for describing the system planning process, which operationalizes this definition. The paper concludes with a description of the system planning process in California in terms of this framework.

•IN THE PAST FEW YEARS many major highway projects proposed by state highway agencies have encountered serious public opposition. When opposition first appeared, highway planners believed that improved highway design would satisfy community demands. Since then it has become clear that many interest groups object as much to being denied a role in the decision-making process as to the actual decisions. In other cases, opposing a highway project is the only way a community can publicly deplore the exclusion of effective transit planning from the decision-making process. In short, the process by which transportation decisions are made has become a major transportation issue in America.

The federal government has taken a leading role in building process consciousness. The Federal Highway Administration (10) has realized that earlier and more thorough community participation in highway planning cannot simply be demanded of the state highway agencies. Participatory planning is in fact only possible when significant changes are made in the state-level decision-making process. Although the FHWA process guidelines do not challenge the restriction of many state gas taxes to use on highways or demand institutional changes, they do require early involvement of citizen groups, consideration of a wide range of impacts, and consideration of the "do-nothing" alternative.

Partly as a result of the guidelines, the design of the decision-making process at the state level is in a state of flux more today than at any time since the 1956 Federal-Aid Highway Act. Legislatures are considering new regional transportation institutions, highway agencies are developing corridor studies, and in Washington, D.C., the Highway Trust Fund has at last begun to provide a dribble of transit money.

Intelligent modification of existing state-level transportation decision-making processes will be a difficult, confusing, and time-consuming business. At least partly this is because few interested parties, even few highway engineers, really have a clear idea of what constitutes a decision-making process design in transportation. The purpose of this paper is to propose a definition of this process design that is sensitive to the directions changes seem to be taking in American transportation planning. The definition is in the form of a series of components that occur in any decision-making process design at the state level. Let us begin by stating some basic beliefs about the decision-making process.

We believe that transportation decision-making in the United States is dominated by the behavior of large public organizations, such as state highway departments, munici-

---

Publication of this paper sponsored by Committee on Transportation Systems Design.

palities, and environmental agencies. These organizations represent many different interests in society from at least the state level down to the neighborhood. Transportation decisions emerge from the maneuverings and negotiations of these system planning institutions (SPIs) with each other, with private groups, and with the general public over time.

We further believe that the behavior of these SPIs, and therefore the decisions they reach, is strongly influenced by the financial and legal structure constraining these negotiations. The transportation system itself is the result of accumulated decisions, and these decisions result from the incentives imposed on the SPI by this financial and legal structure. We can even define system planning as the SPI and the associated structure influencing their negotiations: System planning is the collection of system planning institutions, funding sources, legal structure, and programming procedures that interact with society to produce transportation policy and investment decisions over time.

In our view the decision-making process is the system planning process, and we will use the terms interchangeably. Just as a road can have different curve radii or different lane widths, the system planning process can have different SPIs and different funding sources. Just as different road designs result in different driver behavior, so different process designs result in different types of organizational behavior and decisions. If we could decide on what type of process behavior we would like to have, it might be possible to design actively a process to produce it. A paper by Mead treats this question in more depth (7). Here we merely present the components of the system planning process design, which are implied by the definition of transportation decision-making. After presentation of these components in the form of an abstract framework for system planning, we will describe the process design in California as a case study.

COMPONENTS OF SYSTEM PLANNING PROCESS DESIGN: A FRAMEWORK FOR PROCESS DESCRIPTION

The framework for describing the system planning process design consists of a list of components that occur in any process design. The framework operationalizes the definition of system planning given and is designed to facilitate comparison of different system planning process designs, to provide a process description of manageable length, and to isolate those points in the process design where leverage could be brought to bear to change the design.

System Planning Institutions

A great many public and private institutions participate in the system planning process. With interest directed to system planning by the public sector, it is expedient to emphasize in the framework those public institutions that actually have legal authority to make transportation decisions. The entire framework is oriented around these SPIs.

Examples of such institutions are the state legislature, transit districts, municipalities, or the state transportation or highway agency. In a state process, these are usually the key SPIs. The framework will often consider SPIs that physically overlap or whose business is not principally transportation, such as municipalities.

Each SPI tends to be responsive to interests in its area, but many respond only to some interests (e.g., chamber of commerce). Some institutions pursue specialized forms or modes of transportation (e.g., airport authorities, highway districts) to the exclusion of others. Different SPIs may have interdependent or independent funding.

The framework chooses to view all organizations and institutions that are not public institutions with legal authority as actors or interest groups, whose participation in system planning negotiation takes place within the SPI acting as a forum for dialogue.

The framework should identify the chief activities of each SPI in the system planning process. Some, such as a department of transportation, will be actively engaged in the programming and construction of transportation links. Others, such as municipalities or an environmental agency, may play the role of reviewer of department of transportation proposals.

In describing an existing process, the framework should also describe which interest groups or factions are represented by each SPI and identify to which interests the SPI decision-makers are responsive. If decision-makers are elected, or removable by elected officials, they tend to be responsive to pressures from interest groups and not just from their planning staffs. But, if they are appointed over long terms or are not responsible to elected officials, they are usually difficult to remove (often the case with highway commissioners) and less responsive to direct public pressures.

Finally, the framework should note which interest groups have no SPIs to represent them in the system planning dialogue.

Legal Structure

Many decisions made by SPIs are a function of the influences they exert on each other. But this exercise of power and influence goes on within the bounds of a legal structure of laws, statutes, and agency procedures, which constrain the strategies of the participating SPI. State and national legislatures generally play the major roles in defining the legal structure.

Certain SPIs (environmental control agencies, state DOTs, municipalities) are given review authority over projects proposed by other SPIs in their jurisdictions. In some cases this authority amounts to veto power over proposed projects. In others, review only guarantees the agency's right to include its comments with the proposal (the case with A-95 review agencies). Occasionally, low-level SPIs (e.g., municipalities) have review power over projects proposed by higher level SPIs (e.g., the state or its highway agency). But the most common review powers are the powers of program and budget review belonging to high-level funding agencies such as state highway agencies or the Federal Highway Administration. Often these powers are the major means by which the SPIs influence the decisions of regional or state-level implementing agencies.

Another important component of the legal structure is the authority granted to some SPIs (especially state legislatures and municipalities) to raise taxes for use in transportation system planning and construction.

Funding and Allocation Structure

Even though it is also obviously defined by legislation, the financial structure of a system planning process seems so important that it is treated separately in the framework. The following components of the structure can be defined:

Funding Sources—In general there are multiple sources of funding available. These funds may result from national or state taxing policies or bond issues or may derive from foundations. Sources are usually controlled by different SPIs, in many cases set up to administer the funds. The framework should indicate the important sources, the controlling SPI, and the approximate annual magnitudes.

Funding Restrictions—A fund may be restricted to use by specific organizations or for specific types of projects. Funds may or may not be available to ameliorate the adverse impacts of projects they pay for. The most notorious examples of restricted funds are the national and state highway trust funds, which are usually restricted to use for building and maintaining highways. Funds may be restricted to use on certain defined systems, such as a state highway system. There may also be minimum amounts that must be spent in a given SPI within given time periods.

Allocation Structure—In many cases, the SPI in direct control of a funding source does not actually spend the money for transportation projects but distributes or allocates the funds to a number of lower level SPIs. Because it determines the nature of the financial incentives operating on the lower level SPI, allocation is an important determinant of the behavior of a system planning process. Allocation generally assigns percentages of the total funds to each of the lower level institutions. These percentages then obtain for a number of budget periods, an interval known as the allocation period.

The most important aspect of allocation for a system planning process is the nature of the allocation mechanism used to divide up the money among competing lower level SPIs. Various methods are possible. Some are based on an analysis by the high-level SPI of programs proposed by the lower level SPI. Criteria for allocation based on programs include

1. Economic efficiency of proposed programs,
2. Benefit-cost ratios of proposed programs, and
3. Maintenance of consistent statewide levels of service across all lower level SPIs.

Allocation may often be based on formulas using socioeconomic data such as population, income, and miles of road. Allocation can also be on a project-by-project basis (essentially the method used by the FHWA in doling out federal aid).

Process Behavior Through Time

System planning, even when it is directed toward implementation of some ultimate master plan, is in effect a continuous activity. It is always characterized by certain periodic information or resource transfers between the SPIs. These transfers can be said to occur periodically at points in time called milestones. Milestones tend to "drive" the process because they encourage informal negotiation and communication between the SPIs before the actual information transfer occurs. These negotiations become more intense when important projects are reaching the end of important project development phases and when resource transfers are involved. For instance, a budgeting milestone can force planners to try to finish up a corridor study in time to allow budgeting of route location in the next budget period. If allocation is based on proposed programs, an allocation milestone can force an agency to try to finish the projects it said it would complete in the allocation period.

It may be difficult to say with any particular milestone how long or how intense the negotiation and bargaining between conflicting interests are before the actual information transfer; the negotiation may also depend on other factors. Nevertheless, it is clear that the number and periodicity of milestones are an important determinant of negotiations between SPIs and thus of process behavior generally. Possible milestones that might be mentioned in a process description are as follows:

1. The budget milestone (usually annual),
2. The allocation milestone (usually every few budget periods),
3. The process monitoring milestone (requiring the transfer of information giving the status of a lower level SPIs decision-making process from the lower level SPI to some higher level institution),
4. Process review milestones (there may be milestones requiring SPI to report evaluation of the structure of their system planning process to other institutions), and
5. Political elections.

The milestones imply periodic information flows between SPIs. The framework should list these formal transfers. A chart or diagram of process information flows constitutes perhaps the best one-page process summary.

Programming and Project Development

In many ways investment programming procedures are the heart of the process design. They dictate how most of the money gets spent in state highway agencies. Often they are not completely available on paper because of the political sensitivity involved in choosing major public projects.

Programming Procedures and Documentation—The framework should describe the procedures by which projects get into the programs of the SPI and how they evolve over time. Programming documents should be noted for their insight into how projects are chosen for development and formed into alternatives.

Criteria for Project Programming—The framework should note the criteria used by an SPI in deciding which projects to program. Some of these will be explicit, others implicit. Examples are benefit-cost ratios, political pressure, or predicted demand estimates and community acceptability.

Application of Resource Constraint, Programming Horizon, and Future Uncertainty—The framework should stipulate how (if at all) the reality of finite resources enters programming, how far ahead each SPI programs, and how it deals with future uncertainty. A related datum is the discount rate (if any) employed by the process.

Budgeting Documentation—Budgeting usually grows out of programming with the budget request bearing a strong relationship to the first year of the program. Budgeting of lower level SPIs by higher level SPIs normally takes place annually. The framework should explicate the structure of the budget enough to make clear the extent to which the proposed program and its projects (both planning and implementation) are "visible" within the budget request. The framework should also report the extent to which proposed projects are dependent on decisions by other SPIs or general "community acceptance": How much conditionality and future uncertainty are visible in the budget request?

Project Development Phases—These allow extended project development to be subdivided into shorter pieces that can be scheduled more easily in terms of the process periods and milestones. One or more planning phases (e.g., corridor study, route location) are usually followed by right-of-way acquisition and implementation. In case of controversy, an SPI may define study design phases to precede the project development phase. In evaluating a process design, the planning phases (including studies) are actually more important than the right-of-way and construction phases.

Important project milestones, such as public hearings and agreements between local jurisdictions and the implementing agency, usually occur at the end of the project development phases. Veto powers may relate to the right of review by a given SPI of the results of a particular development phase. For instance, a city may have veto power over the results of a detailed design study by virtue of veto power over implementation.

Process Monitoring and Process Review

Process Monitoring—Process monitoring is a term for ways in which SPIs review the character or quality of the decision-making process in the short run. It usually takes the form of process guidelines levied by one SPI on another. These guidelines specify rules for decision-making, such as early involvement of interest groups or use of interdisciplinary design teams. The FHWA's process guidelines are one example of process monitoring. The key to monitoring effectiveness is periodic checkup by the levying SPI backed up by some kind of effective incentive or threat such as curtailment of funds if the guidelines are not obeyed.

Process Review—This is a term for mechanisms the process may have for changing itself in the long run. Conceptually, at least, a system planning process might have built into it the capability to review and redesign itself to meet changing needs; process review might be institutionalized as a formal periodic activity. This capacity for periodic process review is very important, and the framework should mention it where it exists. Unfortunately, process review and change only occur in most present process designs in response to a crisis such as the freeway revolt or the environmental crisis.

Process review is generally done by a high-level SPI, such as the state legislature or DOT. Elements of the process design that might be changed in process review include

1. Project development phases,
2. Size of and restrictions on funding sources,
3. Spending minimums in a given SPI, and
4. Allocations mechanisms.

Support Models

Impact prediction is a major activity of transportation planning. Impact prediction and display models can have significant impact on the process through the assumptions they make about the world. Whatever these assumptions are, they inevitably bias both the predictions and people's views of the world. General information should be provided by the framework as well as more detailed information about transportation flow models, land use prediction, and other impact prediction models.

Included in general information are some of the following questions:

1. What is the range of impacts for which prediction tools exist? Do techniques exist for predicting both user and nonuser impacts? Which impacts?
2. How are prediction tools adapted to the needs of decision-makers? Are predic-

tions at a level of detail appropriate to the SPI using them? Is "turn around time" reasonable?

3. Are predictions a function of policy variables where appropriate?
4. Do prediction models give estimates of uncertainty in their predictions?

Network flow models should address the following questions:

1. Is the analysis multimodal?
2. Does flow prediction incorporate concepts of supply-demand equilibrium? Does trip generation depend partly on level-of-service variables delivered by the network (e.g., supply-demand equilibrium)? Does assignment treat network capacity as an input (capacity restraint) or as an output (uncapacitated assignment)? Are there feedback loops to facilitate level-of-service consistency at each step of the flow prediction: trip generation, trip distribution, modal split, assignment?
3. Are multiple level-of-service variables used in predicting demand (e.g., in and out of vehicle time, out-of-pocket and non-out-of-pocket costs)?
4. At what levels of detail does the process perform flow simulation: system-wide, corridor, subarea?

Land use prediction models should deal with the following items:

1. Is land use (traffic generator) prediction ad hoc, or are prediction models (land use models) used to some degree? Which models?
2. Are land use predictions partly a function of levels of service on the transportation network?
3. How far ahead are land uses predicted?
4. What land use, if any, "drives" the prediction (i.e., is predicted exogenously and assumed to cause development of other land uses)?

With other impact models, the framework should note the existence of models or techniques for predicting the following impacts:

1. Noise levels,
2. Air pollution levels,
3. Effects on local tax base and real estate values, and
4. Effects on local circulation patterns.

When display aids are used, the following questions should be answered:

1. What sort of display aids are used in meetings? In hand-outs?
2. What levels of detail are used?
3. Do displays encourage interest groups to participate?
4. Is the system planning process itself displayed? Do interest groups understand the decision-making process? Does the process display who the decision-maker is, when the decisions affecting them will be made, the points of view of other institutions with respect to a particular decision, and its own legal and institutional structure?

Informal Process Structure: Role Perceptions of the Planning Staff

We have discussed the formal structure of the system planning process design. We believe this structure is a major determinant of the behavior of planners working for the major system planning institutions of the process. It produces for them a set of roles to be assumed, a set of negotiation games to be played according to certain rules in arguing project decisions. For example, a city mayor is likely to be more interested in local transportation problems than in the general statewide performance of the transportation system. A state highway engineer is likely to emphasize state-level impacts over regional impacts because he is a state, not a regional, employee. A mayor will be easier for a state DOT to negotiate with if he has no local veto power over state projects. A highway district allocated funds based on a formula using socioeconomic data will not try so hard to sell roads as one whose allocation is based on a proposed construction program.

This informal process structure is often the most visible part of the process design to citizens involved in highway decisions. Interest groups may perceive a condescending, even arrogant attitude in a highway engineer and fail to see the dedicated funding sources he is trying to spend or the traffic flow predictions he is trying to maintain.

The formal structure of the process helps produce the planner's role perceptions and behavior, but these are also strongly influenced by the planner's own internalized professional standards and beliefs. A planner who believes in "optimum" system design and the desirability of growth and progress will probably behave differently from one who believes in the sacredness of the community's opinion and desires.

Even though they are explained partially by the formal process design, it is important to include an assessment of planners' role perceptions in the framework description. This is especially important for the chief implementing or funding institutions, such as state highway agencies. Role perceptions are evidently a major determinant of process performance.

TRANSPORTATION SYSTEM PLANNING PROCESS IN CALIFORNIA

The transportation system planning process in California offers a useful case study for the framework presented. The California process is perhaps the largest state process in America, spending almost $1 billion a year in state and federal money. Its formal process structure is more complex than that found in most states. It is a process encountering and responding to severe problems. In the research in several states necessary to produce this framework, the regional agencies and the California Department of Transportation stood out as institutions of unusual competence.

The California process is described as of September 1971. Changes since that time include the creation of a state DOT and regional planning agencies for transportation in each region entrusted with the creation of a new multimodal state highway plan. The FHWA guidelines will result in still more changes. The description is based on extensive contact with the California system planning process, directly both through field trips and through M.I.T. staff working on contract research in California. An intensive search of relevant documentation was also performed. The description contains only the essentials of the process design in the interest of brevity. More detailed descriptions are available.

System Planning Institutions

1. The legislature establishes state transportation funding such as the Highway Users' Tax Fund (HUTF) and State Highway Fund (SHF), the 16,000-mile state highway system, andthe freeway and expressway system within the state system. The interest groups represented are statewide lobbies such as the highway lobby, the Sierra Club, and local and regional interests who fail to gain access to the process at lower levels. The responsiveness of decision-makers is good.

2. The highway commission controls expenditures of the SHF. The interest groups represented are probably prohighway groups. The responsiveness of decision-makers is poor.

3. The division of highways is responsible for building and maintaining the state highway system with the contents of the SHF. The interest groups represented are state-level road interests and also regional and community interests failing to gain access to the process at the highway district level. The responsiveness of decision-makers is poor.

4. The highway districts are regional agencies of the Division of Highways who perform all planning, programming, and construction on the state system. The interest groups represented are real estate developers, mayors, and highway interests, and responsiveness of decision-makers is poor.

5. Among the councils of government are the Association of Bay Area Governments (ABAG) in the Bay Area, Southern California Association of Governments (SCAG) in Los Angeles, and the Comprehensive Planning Organization (CPO) in San Diego. The interest groups represented are pro-urban planning, antihighway, and environmental groups. The responsiveness of decision-makers is good.

6. The transit districts such as the Southern California Rapid Transit District in Los Angeles and the unusually powerful Metropolitan Transportation Commission (MTC) in San Francisco receive some state money from the new sales tax for transit.

7. The counties and cities receive direct grants from the HUTF for use on "select systems" of local highways. They negotiate exact location and design of state highways with the districts. The interest groups represented are development interests and community and neighborhood groups. The responsiveness of decision-makers is good.

Legal Structure

1. The legislature establishes gas and license taxes contributing to HUTF and reviews and revises master plans for the state highway and freeway and expressway systems. Master plans define links at the corridor level. The legislature also establishes gas taxes contributing to state transit funding.

2. The highway commission reviews division programs such as the multiyear financial plan (MYFP) based on district multiyear program proposals (MYPPs) and budgets, performs allocation of SHF to districts, and issues procedural guidelines to the division.

3. The division reviews district programs (planning program and MYPP) and budget and issues procedures to districts.

4. The district must obtain route adoption agreement from town and counties through which a proposed state highway passes, setting detailed corridor location. The district must then obtain the freeway agreement from these bodies based on a review of detailed design before it can close local roads for construction. The power to withhold the freeway agreement represents a local veto. Recent changes require a third agreement milestone, the cooperative or corridor agreement, which terminates a new corridor study phase, at least in urban areas, and precedes the route adoption agreement.

5. The regional MTC in the Bay Area reviews all transportation projects in its region and has veto authority over all transportation projects not deemed of statewide importance by the highway commission. Established by state law, MTC has compulsory membership of counties and towns in the Bay Area. MTC coordinates closely with ABAG.

6. ABAG, SCAG, and CPO have A-95 review power in their regions.

Funding and Allocation Structure

1. Among the funding sources, HUTF is based on gas and license taxes. Half of the fund goes to towns and counties in direct grants and half goes to SHF. State transit funding is based on ¼ percent sales tax on gasoline and is allocated directly to counties. Federal funds are available at various matching ratios for state roads (Interstate, aid to primary and secondary roads program).

2. In terms of funding restrictions, all counties but two in the mountains must receive at least $4 million over the allocation period of 4 years. These are defined as county minimums. SHF and federal program funds may be used for highway construction only. District programs must also satisfy district minimums as defined by allocation.

3. According to the allocation structure, allocation of direct grants from HUTF to cities and counties is by formula using socioeconomic data. Allocation of state transit funding is by formula using socioeconomic data. Sixty percent of SHF is allocated to Southern California and 40 percent to Northern California based on legislation. Within the north and the south, allocation of 70 percent of SHF is based on relative needs of the districts. Needs are based on flow model predictions of system size necessary to maintain statewide levels of service or average speed for each functional classification of roadway. This calculation determines district minimums over 4 years. The remaining 30 percent of SHF is allocated at the commission's discretion. The allocation period is 4 years.

Process Behavior Through Time

1. Milestones include budgets (district and division requests) on an annual basis, programming (MYPP, MYFP, planning program) on an annual basis, needs study and allocation (every 4 years), and recommended changes to state highway system (to legislature every 4 years).

2. Information flows are as shown in Figure 1.

Programming and Project Development

1. District programming is based on a needs study (based on model predictions of capacity required to maintain specified statewide levels of service on various classes of facilities and on local desires). District programming develops planning program containing 8 to 12 years of new construction and MYPP, sent annually for division review, containing 6 years of construction and programs for maintenance and local assistence. Division programming combines MYPPs from districts into MYFP, the division's program, for annual commission review. Budgets are handled similarly.

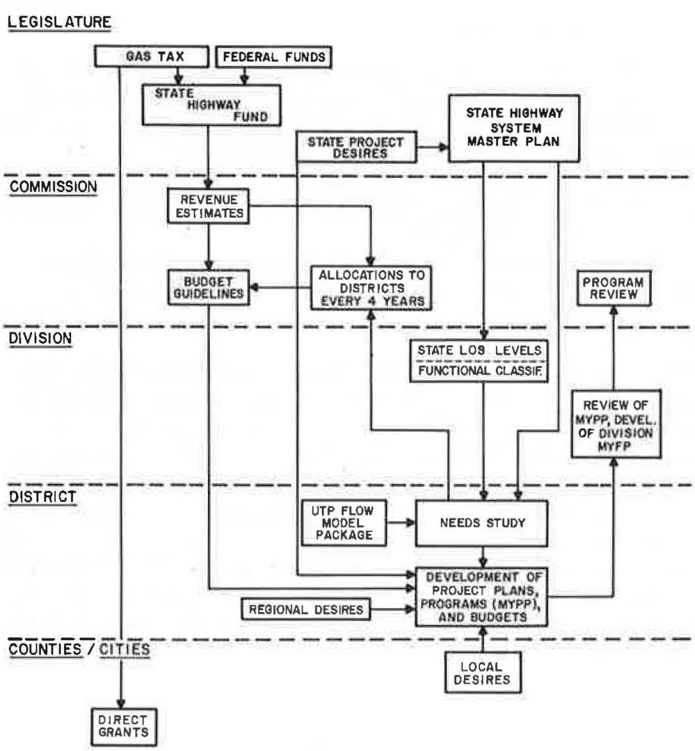
2. The criteria for project programming include credibility of need as given by the network flow model and interstate status and county and district minimums.

3. Resource constraint is applied after project sizing and location decisions have been made. The programming horizon is 8 to 12 years. There are no explicit consideration of future uncertainty and no discounting procedures.

4. In budgeting documentation, the budget is submitted in program form containing programs for maintenance and operations, improvements, local assistance, and general support. New construction dominates improvement programs and represents 75 to 80 percent of the entire budget.

5. Project development phases (and associated milestones) include a corridor study (recently added, ends with "cooperative agreement" between districts and cities containing the proposed corridor), route study (ends with route adoption), mapping and basic design (ends with freeway agreement), right-of-way acquisition, and construction.

Figure 1. Simplified information flow of the California system planning process.



Process Monitoring and Process Review

1. No systematic review or monitoring of the process exists.
2. Responses to present difficulties include community interaction units operated by the right-of-way division of the districts, new corridor study, and special planning policy for the coastal zone containing Calif-1.

Support Models

1. Network flow models deal with automobiles only (but transit mode soon to be included), no dependence of trip generation on level-of-service variables, 24-hour trips, and uncapacitated, "all-or-nothing" assignment to the network with link sizes as an output of the model.
2. Land use models include data for trip generation calculations based mostly on local predictions of socioeconomic indicators and work under way in San Diego and San Francisco on implementation of the PLUM land use model.
3. For other impact prediction models, information is not available.
4. Little information is available on display aids.

Role Perceptions of the Planning Staff

The key planners in the process remain the route planners employed by the Division of Highways and its districts. By and large most of these people define their roles in terms of completing the California freeway and expressway system. This perception demands no explicit reason for involving community groups in decision-making. They believe that political factors should not influence programming any more than necessary.

At the top management levels of the division some role perceptions have begun to accept more community participation. In fact, some of the division staff have adopted a complete "help the community make a decision" role perception. This is also true of the staff of the district- and division-level Community and Environmental Factors Units. California is a land of extremes, and this is evident in the wide range of role perceptions that coexist in the Division of Highways and its districts.

CONCLUSION

The transportation decision-making process is changing faster today than ever before. This is partly because a wide variety of professional and citizen groups have become as interested in the design of this process as in the decisions it produces. Ultimately, they sense, the process design predetermines these decisions.

The growing concern with the process of reaching decisions finds reflection in recent federal policy such as the FHWA's process guidelines. These guidelines begin to specify what desirable process behavior should be.

Before the highway agencies will be able to respond fully to the guidelines, they need to understand the makeup of a system planning process. Similarly, interest groups, when they criticize highway decision-making, need to understand that the highway department is not the whole process.

The purpose of this paper has been to propose an operational definition of the system planning process that can be used by highway agencies and state governments to display the state transportation planning process design to all interest groups.

It is time that federal agencies such as FHWA and UMTA require process descriptions as part of the state applications for federal funds. FHWA should require complete process descriptions as a follow-on to the process guidelines. State decision-making process designs should be on file in updated form and should be available to all for comparison and review. States should compare process designs the way they compare their populations, economies, and highway systems. Only such open discussions will enable us to make intelligent changes in these designs to ensure that they continue to answer the country's evolving transportation needs.

ACKNOWLEDGMENTS

This paper is the result of research performed at the Urban Systems Laboratory, Massachusetts Institute of Technology, for the State of California, Transportation Agency, Department of Public Works, Division of Highways and the U.S. Department of Transportation, Federal Highway Administration.

The author would like to express his thanks to all the members of the Transportation and Community Values Group of the Urban Systems Laboratory for their many inputs to this paper. Thanks are also due to the MITRE Corporation for its support of the author's academic program at M.I.T.

The opinions, findings, and conclusions expressed in this paper are those of the author and not necessarily those of the State of California or the Federal Highway Administration.

REFERENCES

1. Colcord, F. C., Jr. Transportation System Planning in California: Institutional Arrangements of State and Local Governments. Urban Systems Laboratory, M.I.T., Cambridge, USL Rept. 72-7, June 1972.
2. Downs, A. Inside Bureaucracy. Little, Brown, 1967.
3. Friedman, R. S. State Politics and Highways. In Politics and the American States (Herbert, J., and Vines, K., eds.), Little, Brown, 1971.
4. Long, N. E. The Local Community as an Ecology of Games. In The Search for Community Power (Hawley, W. D., and Wirt, F. M., eds.), Prentice-Hall, 1968.
5. March, J. G., and Simon, H. A. Organizations. John Wiley, 1958.
6. Mead, K. C. Design of a Statewide Transportation System Planning Process: An Application to California. Dept. of Civil Engineering, M.I.T., Cambridge, PhD thesis, June 1973.
7. Mead, K. C. Resource Allocation and the System Planning Process. Highway Research Record 467, 1973.
8. Pecknold, W. M., et al. Transportation System Planning and Community and Environmental Values. Urban Systems Laboratory, M.I.T., Cambridge, USL Rept. 72-3, June 1972.
9. 1971 Annual Highway Planning Report. California Dept. of Public Works, Aug. 1971.
10. Policy and Procedures Memorandum 90-4, Process Guidelines. Federal Highway Administration, U.S. Dept. of Transportation, 1972.

AN ACCURATE AND EFFICIENT APPROACH TO EQUILIBRIUM TRAFFIC ASSIGNMENT ON CONGESTED NETWORKS

Larry J. LeBlanc, Southern Methodist University;
Edward K. Morlok, University of Pennsylvania; and
William P. Pierskalla, Northwestern University

This paper describes a new solution technique for the equilibrium traffic assignment problem. After existing methods of solution are reviewed and difficulties that have been experienced with current techniques are discussed, a mathematical programming model for the equilibrium traffic assignment problem is presented. The solution technique for this programming model is one that has been proved to converge every time and that rapidly closes in on the equilibrium flows without excessive computational requirements. It is noted that the computational requirement of the proposed approach is very similar to those of currently used solution methods, which clearly indicates the feasibility of using the proposed approach to find the equilibrium flows on networks with hundreds of nodes. Numerical results for the proposed solution technique on a test network having 76 arcs and 24 nodes are given. A computing time (central processing unit) of 6 seconds on the CDC 6400 computer is reported for accurately computing the equilibrium flows on the test network.

•THIS PAPER describes an efficient method for finding the equilibrium traffic flows on urban transportation networks. The problem is as follows: We are given a system of streets and zones representing a particular urban area, and we have estimates of the number of travelers (amounts of flow) who will drive between each pair of zones. It is well known (2) that the travel time along any street experienced by each driver depends on the number of vehicles flowing along the street. We assume that each driver will take the shortest (quickest) route between his origin and destination, and we wish to determine the traffic density on each street that results from the interaction among drivers as they congest the streets by traveling to their destinations. An equilibrium exists when a driver (increment of flow) cannot reduce his travel time by switching to another route between his origin and destination. Thus we wish to determine how the traffic between the zone pairs will be distributed over the streets of the city.

The equilibrium traffic assignment problem is an especially important one inasmuch as every metropolitan area experiences to some degree the serious problem of traffic congestion, notably during peak hours of movement. To improve an urban transportation system to meet projected demands for trips between each pair of zones in some future period requires that a model be developed for testing the proposed improvements. Alternatively, we may wish to determine whether the existing system can accommodate future increases in traffic without excessive congestion.

A system of streets and expressways is usually modeled by a network whose nodes represent major intersections and interchanges; the nodes are connected by directed arcs so that a two-way street is modeled by two arcs in opposite directions. The network is generally used to represent only the major streets of an urban area, whereas minor roads such as side streets in housing areas are usually not included.

An urban area is typically divided into zones. We assume that a matrix is available that specifies the expected number of trips between the various zones during the time

Publication of this paper sponsored by Committee on Transportation Systems Design.

period being studied. This matrix is called a trip table; the (i,j) entry equals the number of vehicles that must depart origin i to arrive at destination j. Each zone is identified by a node in the network. A node will not necessarily have any demand for trips associated with it; it may simply represent an intersection of two streets. All traffic that leaves any zone emerges into the network through its associated node or nodes, and traffic entering the zone leaves the network through the associated nodes. In this way, origin-destination estimates based on urban zones are transferred into origin-destination estimates based on nodes in the network. The model assumes that all traffic enters and leaves the network through the nodes.

The travel time experienced by a user of any road or arc, called the average travel time function or the volume-delay curve, is a known function of the total volume of flow along the road. Let $A_{ij}(x_{ij})$ denote the travel time experienced by each user of arc (i,j) when x_{ij} units of vehicles flow along the arc. For example, if arc (i, j) is 1 km long and vehicle speed is 30 km/h when the volume of flow on the arc is $\bar{x}_{ij}$, then $A_{ij}(\bar{x}_{ij}) = 2$ min. Almost all recent studies have recognized the effect that congestion of an arc has on travel time and have used nonlinear, increasing travel time functions. We assume that $A_{ij}(x_{ij})$ has continuous derivatives; this assumption is not at all restrictive. FHWA uses polynomial functions that have this property.

The travel time functions used by FHWA are shown in Figure 1 (2). The shape of the function $A_{ij}(x_{ij})$ is intuitive. As in the figure, the travel time per user increases very slowly at first; it remains almost constant for low levels of flow. However, as the flow begins to reach the level for which the arc (street) was designed, the travel time experienced by each user begins to increase rapidly. The a_{ij} and b_{ij} are empirically determined parameters for each arc, which depend on the arc's length, speed limit, and number of lanes and traffic lights. If there is a significant delay in making a left turn at an intersection (node), then turn penalties can be incorporated by using dummy arcs to represent the delay in making the turn.

Wardrop (10) has formulated two conditions that together formally characterize a network equilibrium. A set of flows along the arcs of a network is said to be at equilibrium if the following two conditions are satisfied for every origin-destination r-s pair:

1. If two or more routes between nodes r and s are actually traveled, then the cost to each traveler between r and s must be the same for each of these routes; and
2. There does not exist an alternative unused route between nodes r and s with less cost than that of the routes that are traveled.

The assumption is made that each user of the network seeks to minimize his own travel cost and that he experiments with different routes, eventually finding the least cost one. It is clear that, if 1 or 2 were not true, some drivers would switch to the cheaper routes, congesting them and causing a new flow pattern to evolve. Equilibrium is the aggregate result of individual decisions; at equilibrium, no single driver can reduce his own cost by choosing an alternative route in the network.

EXISTING SOLUTION TECHNIQUES FOR THE EQUILIBRIUM TRAFFIC ASSIGNMENT PROBLEM

The majority of solution techniques used today for finding the equilibrium flows on a network are simulation models, heuristic in nature, involving the concept of the shortest route between two nodes. One of the earliest techniques for this traffic assignment problem is called the "all-or-nothing" assignment technique. This method assumes that the travel time experienced by each driver on any arc in the network is a simple constant, independent of the flow level along the arc, and thus it completely ignores the very real problem of traffic congestion. The all-or-nothing technique is first to determine the shortest path between each origin-destination pair and then to assign all of the trips between this node pair to the shortest path. The all-or-nothing assignment technique is unstable: A slight change in the demand matrix can cause radical changes in the predicted arc volumes. Changing the demand has caused an arc's volume to change from the heaviest in the network to too few trips to justify its con-

Figure 1. Travel time functions used by the U.S. BPR.

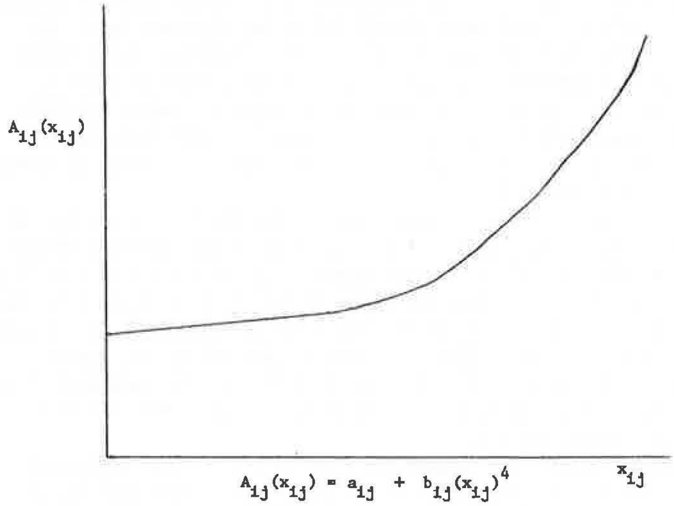
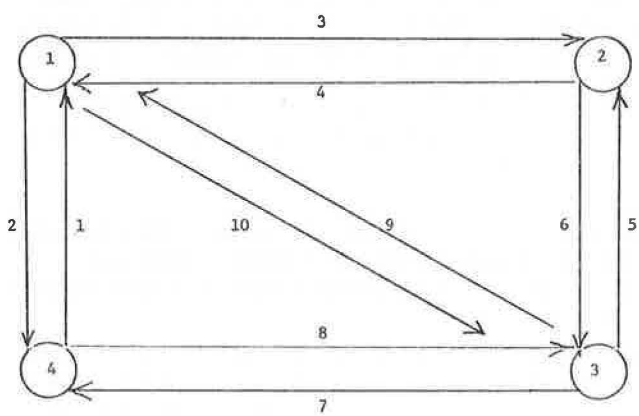


Figure 2. Test network.



struction. In view of the approximate nature of the demand estimates and of the travel time functions, this instability can seriously hinder the model's utility (3).

A natural extension of the all-or-nothing technique is to recognize the effect of congestion on the travel time on any arc. The effect of congestion is incorporated in the capacity restraint simulation models; these are traffic assignment procedures in which the travel time along an arc is adjusted according to a functional relationship between the design capacity and the volume of flow along the arc. An example of a capacity restraint model is the traffic assignment technique developed by the Chicago Area Transportation Study. In the CATS technique, one node is randomly selected from the network, and the minimum paths are determined from this node to all other nodes. All trips from the selected node are then assigned to the corresponding minimum paths. The network is then updated with new travel times calculated for the arcs in the minimum paths. The procedure is repeated with the random selection of another node and the computation of the new shortest routes.

The CATS procedure does not involve any iteration because the network is updated before the computation of each set of shortest routes, and thus the travel times are directly related to the volumes of flow along the arcs (6). However, the CATS procedure gives different flow patterns, depending on the order of selection of the origin nodes (2). Also the flow pattern produced by CATS does not really ensure that all users follow the shortest path between their origin and destination. For example, if node s is randomly selected during the early stages of the procedure, then all trips originating at node s are assigned to the shortest paths between node s and each destination by using the current arc travel times. But other nodes are subsequently generated, their trips are assigned to routes in the network, and the arc travel times are changed. The result is frequently that the paths used by travelers originating at node s are no longer the shortest paths to their destinations.

In an attempt to find the true equilibrium, iterative procedures are often used. Iterative procedures are simply continuations of the previous model; after all nodes have been generated, the model continues to generate nodes again. The rationale of these iterative models is similar to Charnes' game theoretic interpretation of the problem. Charnes associated with each origin a player who tries to choose a set of routes such that the correct number of vehicles will travel from the origin to each destination at minimum travel time. Because vehicles from the various origins interact, the travel times as seen by a given player depend on the actions of the other players. Thus an iterative technique is used to determine the equilibrium flows on the network. Each player chooses his routes in turn; after all the players have made their decisions, the resulting times along each arc are revealed to all the players, and they again take turns in revising their routes (4).

If the procedure described above is iterated enough times, the sequence of flow vectors may converge to an equilibrium. One possible termination criterion is to stop when the maximum percentage change between the components of two consecutive flow vectors is less than some specified amount. However, this iterative technique does not always produce a convergent sequence; examples are known where the sequence oscillates around a flow pattern that is not in equilibrium. In actual applications of large-scale problems, the practice is usually to terminate after four iterations (2).

The incremental assignment technique is a variation of the all-or-nothing method in which only a small increment of the total number of trips between any two nodes is assigned to the minimum path between the two nodes. In this technique, a node pair is randomly selected and the shortest path between these two nodes is determined. The length of each arc is set equal to the value of the arc's volume-delay function evaluated at the current level of flow along the arc; initially, the flow is zero. Then a small percentage of the total required flow is sent along this path, the flow level for each arc in the shortest path is incremented, and the new lengths of the arc are determined. Another node pair is then selected, and the process repeats itself until all traffic has been assigned. This represents an attempt to load the network in a balanced manner so that all arcs of the network approach the fully loaded condition at the same time. Thus the effect of congestion becomes more significant, and there is a better chance of achieving the conditions of network equilibrium. However, the incremental assignment technique is time-consuming from a computational point of view inasmuch as a shortest route problem must be solved many times for each distinct origin-destination pair in the network (5). Also, it has not been proved that the incremental assignment technique converges to the equilibrium flows; thus for a particular problem, it may not produce the equilibrium flows.

As this brief review has indicated, there are a number of solution techniques in use for the traffic assignment problem. However, there have been difficulties with existing solution techniques, as a recent study in Lancaster, Pennsylvania (9), has shown. The study used data for the existing network in an attempt to replicate observed flows along the arcs by the unrestrained all-or-nothing technique and capacity restraint versions of this procedure involving one, two, three, and four iterations. To determine which of these five algorithms produced flows most closely resembling the observed flows, we made extensive comparisons of each assignment algorithm with the observed ground counts. A chi-square index was used as one means of comparison:

$$\sum_{i=1}^{n} (G_i - A_i)^2 / n$$

where

G_i = observed flow on street i,
A_i = flow on arc i predicted by the assignment algorithm, and
n = number of arcs in the network.

The authors used the chi-square index as an intuitive means of comparison; however, they report that, if the values of the chi-square were used in a statistical test, then "... all of the assignments would be rejected according to a Chi-Square test, since all of the values are significantly different from the ground count." As given below, the chi-square index actually increased after the first iteration, and, even after four iterations, it was still larger than the index of iteration number one:

Technique	Chi-Square Value
Unrestrained assignment	19,035
Iteration No. 1	12,597
Iteration No. 2	16,616
Iteration No. 3	14,599
Iteration No. 4	14,187

No mention of a confidence level is given. The authors' conclusion emphasizes their problem with lack of convergence: "It is not recommended that additional iterations of capacity restraint be made utilizing the same function or model because it has been concluded that the fourth iteration is only the third best assignment."

A chi-square test was also used in the National Cooperative Highway Research Program study (7) to check the all-or-nothing algorithm and several variations of the capacity restraint algorithms. Here it is also reported that "all of the values are significantly different from the ground count estimates, indicating that the difference in assignment is more than can be expected by chance alone."

Another criterion for comparing the output of each algorithm with the observed flows is that of total vehicle-miles in the network. In their report (7), Huber, Boutwell, and Witheford tested the ground count vehicle-miles and each algorithm's predicted vehicle-miles for the Pittsburgh network. The hypothesis tested was that the observations from the ground count and from four algorithms are all from the same normally distributed population. No justification of the assumption of normality is given. The authors again conclude that the hypothesis must be rejected. In fact, they reach the remarkable conclusion that all of the assignment algorithms are equally poor, stating that "... the various assignment techniques gave results which were closer to each other than to the ground count results." This clearly indicates the need for an improved equilibrium traffic assignment algorithm.

In the next section we will present a different model for the equilibrium traffic assignment problem. A solution technique that has been proved to converge every time and that rapidly closes in on the equilibrium flows without excessive computational requirements will then be described.

MATHEMATICAL PROGRAMMING MODEL FOR LARGE-SCALE NETWORK EQUILIBRIUM PROBLEMS

Consider a fixed network with n nodes, and assume that nodes 1, 2, . . ., p, $p \leq n$ are origins and destinations. Define A as the set of arcs (i, j) in the network. Let x_{ij} denote the total flow along the arc (i, j), let x^s_{ij} denote the flow along arc (i, j) with destination s, and let D(r, s) denote the fixed amount of flow required between nodes r and s. Obviously,

$$x_{ij} = \sum_{s=1}^{p} x^s_{ij}$$

As above, we let the average travel time function for arc (i, j) be denoted by $A_{ij}(x_{ij})$. Now define

$$f_{ij}(x_{ij}) \equiv \int_0^{x_{ij}} A_{ij}(t)dt$$

Using the definition $A_{ij}(x_{ij})$ in Figure 1, we see that

$$f_{ij}(x_{ij}) = \left[a_{ij}\left(\sum_{s=1}^{p} x^s_{ij}\right) + (b_{ij}/5)\left(\sum_{s=1}^{p} x^s_{ij}\right)^5 \right]$$

Then the optimal solution for the nonlinear programming problem

$$\text{(NLP)} \min \sum_{(i,j)\epsilon A} f_{ij}\left(\sum_{s=1}^{p} x^s_{ij}\right) = \min \sum_{(i,j)\epsilon A} \left[a_{ij}\left(\sum_{s=1}^{p} x^s_{ij}\right) + (b_{ij}/5)\left(\sum_{s=1}^{p} x_{ij}\right)^5 \right] \quad (1)$$

$$\text{s.t.} \sum_{i \epsilon B(j)} x^s_{ij} + D(j, s) = \sum_{k \epsilon A(j)} x^s_{jk} \quad (2)$$

for j = 1, . . ., n and s = 1, . . ., p and

$$x^s_{ij} \geq 0 \ (i, j) \ \epsilon A \quad (3)$$

for s = 1, . . ., p constitutes the equilibrium flows. The objective function, Eq. 1, is the sum of the integrals of the average cost functions. In Eq. 2, B(j) is the set of nodes with arcs leading into node j (before j) and A(j) is the set of nodes with arcs leading into them from j (after j). The constraints of Eq. 2 are conservation of flow equations that state that, for each destination s, the sum of the flows into each node destined for s plus the flow originating at that node destined for s equals the sum of the flows out of that node destined for s. Constraints of Eq. 3 are simply the nonnegativity requirements.

Problem NLP is closely related to the work done by Kirchoff in electrical networks. Beckmann (1) seems to have been the first to apply the idea to transportation networks;

he proves that the solution to NLP is the desired equilibrium. Unfortunately, this problem appears very much harder to solve than current simulation models. The number of constraints in NLP is enormous—the number of conservation of flow constraints of Eq. 2 equals the product of the number of nodes in the network with the number of destinations in the network. Thus this nonlinear programming problem for a network with 200 nodes, 100 of which are destinations, would have 20,000 conservation of flow constraints. In addition there are the nonnegativity constraints. Potts and Oliver (8) state that computational success has been limited to small versions of problem NLP, and so the approach appears useless for realistically sized equilibrium traffic assignment problems.

However, a rigorous examination and exploitation of the structure of problem NLP reveal that this is not at all the case. LeBlanc (5) used the Frank-Wolfe algorithm (11) to solve the equilibrium traffic assignment problem, e.g., problem NLP. This solution technique has proved to be remarkably accurate and efficient. In the Frank-Wolfe algorithm, Eq. 1 is replaced with a very simple linear approximation, and the linear programming problem of minimizing this linear approximation subject to Eqs. 2 and 3 is solved. The optimal solution to this linear programming problem is then used to define a search direction in which to minimize Eq. 1; the result of this search is an estimate of the equilibrium flows. After the search is completed in this generated direction, a new linear approximation is obtained, the linear programming problem is resolved with the new objective function to obtain a different direction of search, and a better estimate of the equilibrium flows is obtained by searching in this direction. The algorithm continues to iterate in this manner, solving one-dimensional searches and linear programming problems that minimize successively better linear approximations to the nonlinear objective function of Eq. 1.

An alternative procedure for solving NLP would be the usual method of linearization—approximate Eq. 1 with a piecewise linear function and use the simplex method to solve the resulting linear programming problem. The solution of problem NLP by linearization has been attempted in the past. Because linearization uses a more accurate linear approximation and hence does not solve a sequence of linear programming problems, it may seem that linearization is more efficient than the iterative technique from a computational point of view. However, this is not the case. Both solution techniques were coded on the CDC 6400 computer for a test network, and the computing time for the iterative technique was less than that of the usual linearization procedure by orders of magnitude. These numerical results are reported in the next section.

The key reason for the computational success of the Frank-Wolfe algorithm described above is that each of the linear programming problems has such an extremely simple structure that it can be solved by a shortest route algorithm. This means that all of the conservation of flow equations and nonnegativity constraints can be ignored; they are automatically satisfied by definition of a route between two nodes. It is well known in operations research literature that the computational requirements of a shortest route algorithm are trivial as compared to the requirements of a linear programming problem. The net result is that when problem NLP is solved by the Frank-Wolfe algorithm, the computational requirements of several shortest route problems and one-dimensional searches are vastly less than the computational requirements of the simplex method for solving NLP by linearization.

In the preceding section, a currently used iterative simulation technique based on Charnes' game theoretic interpretation of the equilibrium traffic assignment problem was described. LeBlanc (5) showed that each iteration of this simulation technique involves solving a shortest route problem that is identical to the shortest route problem solved at each iteration of the algorithm suggested in this paper for solving problem NLP. Thus we have the remarkable conclusion that the computational requirements of the proposed approach are not significantly different from the computational requirements of currently used simulation techniques for the equilibrium traffic assignment problem; so it is obvious that the proposed approach will be efficient for large problems.

The basic difference between the simulation technique and the algorithm suggested in this paper is that the proposed algorithm solves the shortest route problem to determine a direction of search and then minimizes the objective function in this direction to obtain a new estimate of the equilibrium flows. The simulation procedure, on the other hand, uses the solution to this same shortest route problem itself as a new vector of flows. This leads to completely distinct flow vectors. These two approaches to the problem are fundamentally different. One is a simulation technique based on heuristic assumptions about the system; it frequently does not converge. The other is a rigorous application of a convergent algorithm to an NLP problem whose optimal solution is proved to be the equilibrium.

COMPUTATIONAL RESULTS

The Frank-Wolfe algorithm described was programmed in FORTRAN IV on a CDC 6400 computer; the test network shown in Figure 2 was used initially for debugging of the computer programs only. The network and travel time functions were chosen so that the equilibrium flows could be determined by inspection. Because the Frank-Wolfe algorithm converges to the equilibrium solution only after an infinite number of iterations, the primary concern was to determine how many iterations of the procedure are required for a reasonably accurate answer. The results are given in Table 1; after eight or 10 iterations, the flow values are probably more accurate than the data that are input to this type of model.

The algorithm was then run on a larger network consisting of 76 arcs and 24 nodes, each of which was both an origin and a destination. This network was used to model Sioux Falls, South Dakota, a city of approximately 125,000 residents. The network, trip table, and volume-delay functions are shown in Figures 3 and 4. Because the trip table in Figure 4 is symmetric—the number of trips between node i and node j equals the number of trips between nodes j and i—the equilibrium flow values will also be symmetric. In other words the equilibrium flow on arc (i, j) will be equal to the flow on arc (j, i), and thus we really need compute only 38 flow values. In this problem, one unit of flow was chosen to be 1,000 vehicles per day. Problem NLP for this network had 1,824 variables, 576 conservation of flow constraints, and 1,824 nonnegativity constraints. Because this is a general nonlinear programming problem, it is impossible to determine the exact solution in a finite amount of time. However, examination of the sequence of flows in Figure 5 shows that, after 20 iterations, only two variables changed by more than 5 percent; the majority changed by less than 2 percent. Thus the final flow vector appears to be a highly accurate estimate of the equilibrium solution. Computing time for 20 iterations, excluding 3 seconds of compilation time, was 6 seconds. If the termination rule had been to stop when the maximum percentage change in components was less than 8 percent, the procedure would have terminated after 16 iterations. After 16 iterations, the maximum percentage change in the components was 7.7 percent; computing time was 5 seconds.

The most encouraging computational result was the very small increase in the number of iterations required by the Frank-Wolfe algorithm for the two example problems. There were 12 conservation of flow constraints and 40 nonnegativity constraints for the initial network used for debugging, whereas the problem for the larger network in Figure 3 consisted of 576 conservation of flow equations and 1,824 nonnegativity constraints. Nevertheless, the required number of iterations increased from approximately eight or 10 to only 16 or 20. The number of iterations appears to be related to the number of nodes in the underlying network rather than to the number of constraints in the NLP problem. Increasing the number of nodes by a factor of six only doubled the number of iterations; this indicates that the algorithm will be computationally efficient for problems as large as several hundred nodes.

As mentioned earlier, problem NLP can also be solved by a more common form of linearization in which Eq. 1 is approximated by a piecewise linear function. This approach was also attempted on problem NLP for the network in Figure 3. However, the computing time was much greater: For this technique, the Optima package on the CDC 6400 required 700 seconds to solve NLP—more than 100 times as long as the Frank-Wolfe technique.

Figure 3. Sioux Falls network.

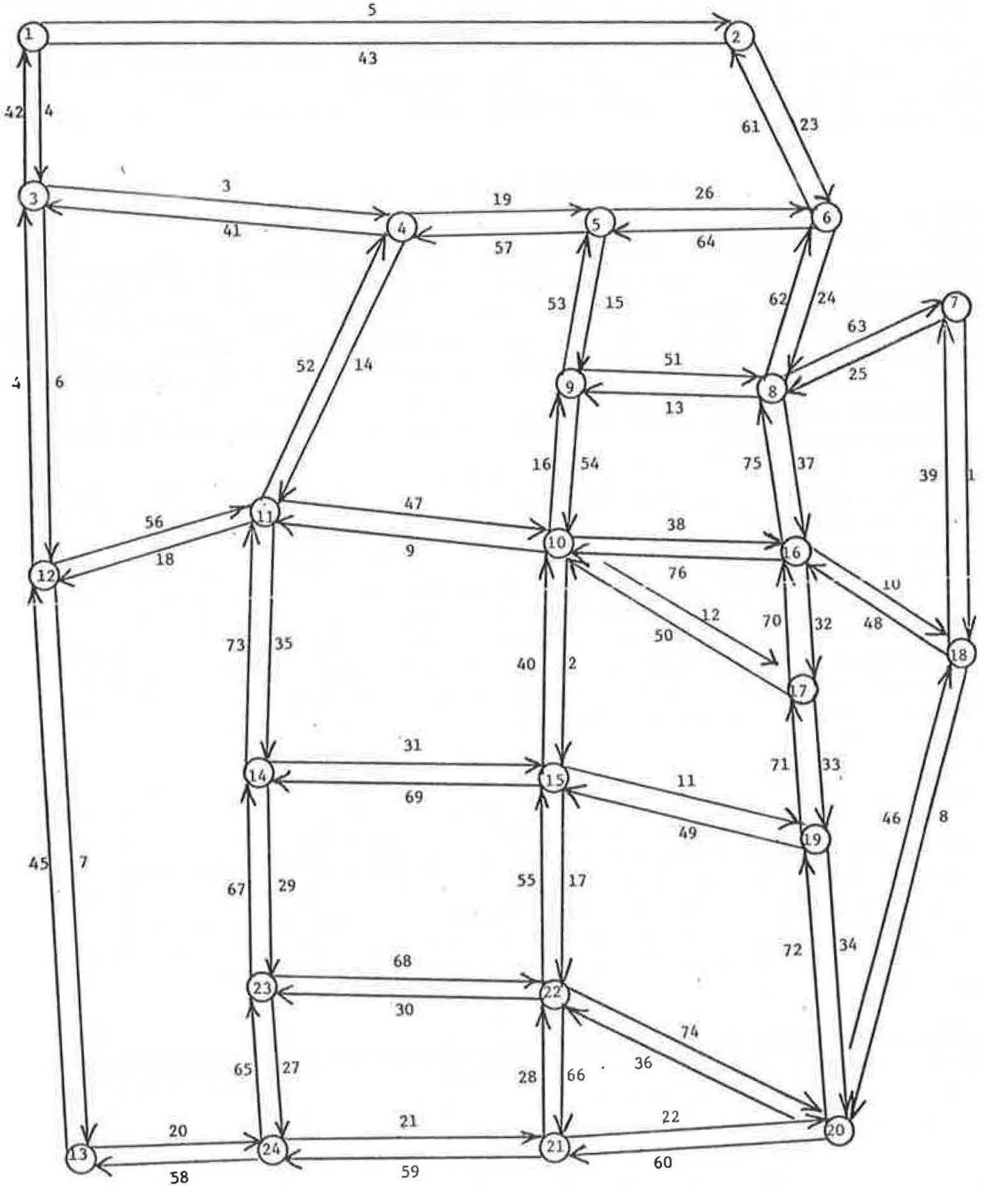


Table 1. Sequential estimates of equilibrium flows on 10 arcs of Figure 1.

Sequential Vectors of Equilibrium Flow	Arc 1	2	3	4	5	6	7	8	9	10
$x^{(1)}$	37.2	28.3	37.2	28.3	28.3	37.2	37.2	28.3	14.4	14.4
$x^{(2)}$	32.2	32.6	32.2	32.6	32.6	32.2	32.2	32.6	20.6	20.6
$x^{(3)}$	33.9	29.9	33.9	29.9	29.9	33.9	33.9	29.9	20.4	20.4
$x^{(4)}$	32.0	31.6	32.0	31.6	31.6	32.0	32.0	31.6	22.6	22.6
$x^{(5)}$	29.9	33.0	29.9	33.0	33.0	29.9	29.9	33.0	22.2	22.2
$x^{(6)}$	31.0	31.8	31.0	31.8	31.8	31.0	31.0	31.8	23.6	23.6
$x^{(7)}$	32.5	29.9	32.5	29.9	29.9	32.5	32.5	29.9	23.0	23.0
$x^{(8)}$	31.7	30.6	31.7	30.6	30.6	31.7	31.7	30.6	23.0	23.9
$x^{(9)}$	29.9	32.1	29.9	32.1	32.1	29.9	29.9	32.1	23.3	23.3
$x^{(10)}$	31.7	29.9	31.7	29.9	29.9	31.7	31.7	29.9	22.7	22.7
$x^{(11)}$	30.0	31.4	30.0	31.4	31.4	30.0	30.0	31.4	22.3	22.3
$x^{(12)}$	31.2	30.0	31.2	30.0	30.0	31.2	31.2	30.0	22.0	22.0
$x^{(13)}$	30.0	31.0	30.0	31.0	31.0	30.0	30.0	31.0	21.8	21.8
$x^{(14)}$	31.0	30.0	31.0	30.0	30.0	31.0	31.0	30.0	31.6	21.6
$x^{(15)}$	30.0	30.9	30.0	30.9	30.9	30.0	30.0	30.9	21.5	21.5
Equilibrium flow	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	20.0	20.0

Figure 4. Trip table and arc parameters for Sioux Falls network.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	1	1	5	2	3	5	8	5	13	5	2	5	3	5	5	4	1	3	3	1	4	3	1
2	1	0	1	2	1	4	2	4	2	6	2	1	3	1	1	4	2	0	1	1	0	1	0	0
3	1	1	0	2	1	3	1	2	1	3	3	2	1	1	1	2	1	0	0	0	0	1	1	0
4	5	2	2	0	5	4	4	7	7	12	14	6	6	5	5	8	5	1	2	3	2	4	5	2
5	2	1	1	5	0	2	2	5	8	10	5	2	2	1	2	5	2	0	1	1	1	2	1	0
6	3	4	3	4	2	0	4	8	4	8	4	2	2	1	2	9	5	1	2	3	1	2	1	1
7	5	2	1	4	2	4	0	10	6	19	5	7	4	2	5	14	10	2	4	5	2	5	2	1
8	8	4	2	7	5	8	10	0	8	16	8	6	6	4	6	22	14	3	7	9	4	5	3	2
9	5	2	1	7	8	4	6	8	0	28	14	6	6	6	9	14	9	2	4	6	3	7	5	2
10	13	6	3	12	10	8	19	16	28	0	40	20	19	21	40	44	39	7	18	25	12	26	18	8
11	5	2	3	15	5	4	5	8	14	39	0	14	10	16	14	14	10	1	4	6	4	11	13	6
12	2	1	2	6	2	2	7	6	6	20	14	0	13	7	7	7	6	2	3	4	3	7	7	5
13	5	3	1	6	2	2	4	6	6	19	10	13	0	6	7	6	5	1	3	6	6	13	8	8
14	3	1	1	5	1	1	2	4	6	21	16	7	6	0	13	7	7	1	3	5	4	12	11	4
15	5	1	1	5	2	2	5	6	10	40	14	7	7	13	0	12	15	2	8	11	8	26	10	4
16	5	4	2	8	5	9	14	22	14	44	14	7	6	7	12	0	28	5	13	16	6	12	5	3
17	4	2	1	5	2	5	10	14	9	39	10	6	5	7	15	28	0	6	17	17	6	17	6	3
18	1	0	0	1	0	1	2	3	2	7	2	2	1	1	2	5	6	0	3	4	1	3	1	0
19	3	1	0	2	1	2	4	7	4	18	4	3	3	3	8	13	17	3	0	12	4	12	3	1
20	3	1	0	3	1	3	5	9	6	25	6	5	6	5	11	16	17	4	12	0	12	24	7	4
21	1	0	0	2	1	1	2	4	3	12	4	3	6	4	8	6	6	1	4	12	0	18	7	5
22	4	1	1	4	2	2	5	5	7	26	11	7	13	12	26	12	17	3	12	24	18	0	21	11
23	3	0	1	5	1	1	2	3	5	18	13	7	8	11	10	5	6	1	3	7	7	21	0	7
24	1	0	0	2	0	1	1	2	2	8	6	5	7	4	4	3	3	0	1	4	5	11	7	0

ARC	A	B	ARC	A	B	ARC	A	B	ARC	A	B
1	.02	.00000001	20	.04	.00000893	39	.02	.00000001	58	.04	.00000893
2	.06	.00000027	21	.03	.00000790	40	.06	.00000027	59	.03	.00000790
3	.04	.00000007	22	.06	.00001373	41	.04	.00000007	60	.06	.00001373
4	.04	.00000002	23	.05	.00001241	42	.04	.00000002	61	.05	.00001241
5	.06	.00000002	24	.02	.00000521	43	.06	.00000002	62	.02	.00000521
6	.04	.00000002	25	.03	.00000119	44	.04	.00000002	63	.03	.00000119
7	.03	.00000001	26	.04	.00001001	45	.03	.00000001	64	.04	.00001001
8	.04	.00000002	27	.02	.00000451	46	.04	.00000002	65	.02	.00000451
9	.05	.00000075	28	.02	.00000401	47	.05	.00000075	66	.02	.00000401
10	.03	.00000003	29	.04	.00001020	48	.03	.00000003	67	.04	.00001020
11	.03	.00000010	30	.04	.00000960	49	.03	.00000010	68	.04	.00000960
12	.08	.00001930	31	.05	.00001085	50	.08	.00001930	69	.05	.00001085
13	.10	.00002306	32	.02	.00000401	51	.10	.00002306	70	.02	.00000401
14	.06	.00001550	33	.02	.00000554	52	.06	.00001550	71	.02	.00000554
15	.05	.00000075	34	.04	.00000958	53	.05	.00000075	72	.04	.00000958
16	.03	.00000012	35	.04	.00001061	54	.03	.00000012	73	.04	.00001061
17	.03	.00000052	36	.05	.00001130	55	.03	.00000052	74	.05	.00001130
18	.06	.00001550	37	.05	.00001157	56	.06	.00001550	75	.05	.00001157
19	.02	.00000003	38	.04	.00001080	57	.02	.00000003	76	.04	.00001080

Figure 5. Sequence of vectors of flow on (a) arcs 1 through 19 and (b) arcs 20 through 38.

(a)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	17.2	21.5	15.9	9.4	2.9	14.5	14.1	9.4	17.4	14.6	16.1	14.9	16.9	12.7	19.4	26.4	25.2	12.8	21.0
2	21.3	22.2	18.9	8.8	3.4	16.7	14.6	17.6	18.5	23.6	16.2	11.2	12.7	9.5	21.6	28.6	21.1	9.7	24.3
3	21.9	26.4	17.2	9.6	4.4	16.8	15.0	20.0	20.4	24.1	15.9	12.3	10.9	9.2	19.8	26.8	25.0	10.8	22.3
4	19.9	25.7	20.0	11.9	7.1	19.7	16.3	21.5	17.4	20.3	20.8	10.8	8.1	9.4	20.6	28.3	21.7	7.2	23.3
5	20.4	25.8	18.1	12.1	6.9	19.7	17.3	22.9	20.7	22.6	19.7	10.5	8.3	9.0	19.2	26.2	22.1	10.6	21.0
6	16.4	26.4	23.4	10.1	6.0	22.0	17.5	21.7	17.0	18.7	19.2	9.3	8.5	8.6	20.4	28.7	18.4	5.3	24.6
7	17.2	25.9	20.7	10.9	6.4	21.2	16.9	22.0	20.5	19.9	18.8	9.3	8.6	8.6	18.5	27.1	18.4	9.0	21.8
8	17.4	26.0	19.3	11.3	6.9	19.1	17.2	23.2	18.9	18.8	20.5	9.2	9.0	7.9	18.5	25.9	19.4	9.4	20.6
9	16.7	24.5	18.4	10.3	6.3	16.9	16.5	22.8	20.6	19.1	18.8	8.9	7.2	7.6	17.2	25.0	20.6	9.6	20.7
10	17.6	25.1	19.3	10.6	6.1	17.0	16.7	22.8	19.3	19.6	19.0	8.9	8.5	7.3	18.8	25.6	19.9	9.3	21.8
11	17.6	25.9	17.2	11.9	7.6	15.1	16.1	23.3	20.8	19.1	19.7	8.8	7.3	6.8	17.4	24.8	19.6	9.6	20.0
12	17.3	25.7	18.1	11.9	7.2	15.8	16.3	22.8	19.6	19.1	19.4	8.8	8.2	7.3	18.5	24.7	19.1	8.8	20.9
13	17.3	24.5	17.9	11.4	6.7	15.1	16.2	22.9	19.3	18.9	19.4	8.8	7.5	7.2	17.8	24.3	19.2	9.2	20.7
14	17.9	25.1	18.4	11.3	6.7	15.3	16.1	22.6	18.8	19.3	19.6	8.8	8.3	7.0	18.9	24.9	19.1	8.7	21.4
15	17.6	25.4	18.2	11.0	6.4	14.6	16.0	22.4	19.2	19.0	19.5	8.7	7.8	7.9	17.2	23.4	19.0	9.2	20.4
16	17.5	25.2	18.6	11.0	6.4	14.9	16.0	22.5	18.9	18.9	19.7	8.7	8.4	7.7	18.4	23.8	19.1	8.7	21.1
17	17.5	25.2	17.6	10.7	6.8	14.4	16.0	22.7	20.1	18.8	19.5	8.7	8.1	7.4	17.9	23.7	18.4	9.1	20.5
18	17.5	25.2	18.0	10.7	6.6	14.6	15.9	22.5	19.4	18.3	19.7	8.7	8.5	7.3	18.6	24.0	18.8	8.7	20.9
19	17.5	25.5	18.0	10.5	6.3	14.2	15.8	22.8	19.3	18.3	19.5	8.7	7.9	7.1	18.1	23.8	19.2	8.9	21.0
20	17.4	24.9	18.1	10.5	6.3	14.5	15.9	22.7	19.1	18.1	19.6	8.7	8.3	7.2	18.4	23.8	18.9	8.6	21.1

(b)	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
1	13.1	8.8	7.1	8.4	11.8	14.6	7.3	10.6	9.4	14.0	13.8	11.6	19.2	23.7	10.3	12.7	10.6	16.0	11.1
2	12.1	11.8	12.8	6.0	14.6	18.2	10.2	11.7	11.3	11.0	10.4	11.6	20.5	17.9	12.6	10.0	8.0	12.1	16.1
3	12.2	10.4	11.0	6.4	14.8	18.5	10.1	12.2	11.1	10.4	12.3	10.2	18.4	15.3	10.7	10.0	13.2	10.3	13.8
4	12.2	11.5	11.5	7.9	15.5	16.0	9.5	11.6	12.6	10.5	8.9	12.0	14.4	13.5	14.3	13.3	8.8	9.9	11.1
5	12.6	9.9	9.4	7.3	13.7	16.0	8.3	11.8	12.1	9.7	10.9	10.8	13.9	11.8	11.8	11.6	11.7	8.1	13.0
6	12.3	12.3	8.9	7.0	15.5	15.0	11.6	10.4	11.4	10.2	7.1	11.2	14.5	11.7	8.3	8.9	8.6	10.7	11.5
7	11.4	10.7	8.1	7.0	14.1	15.1	9.9	10.3	10.6	10.7	9.6	9.9	13.2	11.3	8.7	9.9	9.0	9.1	12.9
8	12.3	10.1	8.2	7.2	13.3	14.6	8.5	10.5	9.9	9.2	11.0	10.7	11.9	10.8	11.1	10.6	8.7	9.2	11.2
9	12.2	11.3	7.9	7.1	14.7	13.7	10.2	9.3	10.5	8.7	10.4	10.0	12.9	11.1	8.2	11.2	9.6	9.5	12.3
10	12.1	10.6	7.8	6.7	13.8	15.0	9.5	9.4	9.9	9.0	10.2	9.8	12.3	10.9	9.0	10.6	9.4	8.6	11.8
11	11.9	11.1	8.1	7.8	14.6	14.4	8.7	8.6	9.5	8.3	10.1	10.8	11.3	9.4	11.0	10.5	8.8	8.7	11.4
12	12.0	10.3	7.8	7.4	13.8	14.2	8.3	8.6	9.3	9.0	9.9	10.1	11.7	9.8	10.1	10.4	8.9	9.0	11.5
13	12.1	11.4	7.7	7.2	14.2	13.8	9.1	9.7	10.1	8.8	11.5	9.7	11.8	9.9	9.3	11.1	8.7	9.0	11.9
14	12.0	10.9	7.7	7.1	13.7	14.8	8.6	9.3	9.7	8.9	10.6	9.5	11.8	9.9	9.4	10.7	8.4	8.3	11.7
15	11.8	10.7	7.5	6.9	14.1	14.6	9.3	8.6	9.5	8.9	9.7	10.4	12.2	10.4	9.0	10.2	8.3	8.9	11.8
16	11.9	10.4	7.5	6.8	13.6	14.5	8.8	8.8	9.4	8.7	10.5	10.1	11.8	10.1	9.5	10.4	8.3	9.0	11.6
17	12.0	10.9	7.5	7.3	14.0	14.3	8.7	8.4	9.6	9.2	9.6	9.3	11.9	10.0	9.1	10.2	8.2	8.9	12.0
18	11.9	10.9	7.5	7.1	13.6	14.3	8.5	8.6	9.6	8.9	10.4	10.1	12.0	10.1	9.2	10.4	8.3	9.0	11.7
19	11.8	10.1	7.5	6.9	14.0	14.1	9.2	8.1	9.0	8.4	10.2	9.4	12.0	10.0	8.9	9.9	8.2	8.9	12.0
20	11.9	10.5	7.5	6.8	13.7	14.1	8.8	8.4	9.1	8.5	10.3	10.1	11.7	10.0	9.4	10.6	8.2	9.0	11.8

CONCLUSIONS AND RECOMMENDATIONS

In this paper we have addressed the problem of finding the equilibrium traffic flows on urban networks. In particular, we have looked at the computational aspects of large-scale equilibrium problems. The algorithm that was discussed above promises to be efficient for finding the equilibrium on a network with hundreds of nodes, since its most difficult computational requirements are identical to those of the iterated capacity restraint simulation models of traffic assignment currently used. And yet the above algorithm is a rigorous one; it is proved theoretically that it always converges to the exact equilibrium. This algorithm has demonstrated its capability by solving a large-scale nonlinear programming problem (576 linear constraints and 1,824 variables and nonnegativity constraints) in 6 seconds on the CDC 6400 computer. Even this small computing time could certainly be reduced by examining the computer program in detail and by making it more sophisticated and more efficient. Further research is needed, however, to determine exactly how large a network can be handled in a reasonable amount of computer time. This question is of particular interest to transportation planners inasmuch as the assumption that a true equilibrium is achieved is almost universally used in system models used to support such planning. Current methods of network equilibrium analysis are known to be inaccurate (in that they often do not converge to an equilibrium) and very costly, making the potential payoff from research on better methods very substantial. The problem of selecting a suitable test network and appropriate data must also be addressed. Finally, suitable comparisons for the outputs of different assignment models must be chosen.

REFERENCES

1. Beckmann, M., McGuire, C. B., and Winsten, C. Studies in the Economics of Transportation. Yale Univ. Press, 1956.
2. Comsis Corporation. Traffic Assignment: Methods-Applications-Products. Office of Planning, Federal Highway Administration, preliminary rept., 1972.
3. Dial, R. B. Probabilistic Assignment: A Multipath Traffic Assignment Model Which Obviates Path Enumeration. Dept. of Civil Engineering, Univ. of Washington, PhD dissertation, 1970.
4. Hershdorfer, A. M. Predicting the Equilibrium of Supply and Demand: Location Theory and Transportation Network Flow Models. TRF Papers, 1966, pp. 131-143.
5. LeBlanc, L. J. Mathematical Programming Algorithms for Large Scale Network Equilibrium and Network Design Problems. Dept. of Industrial Engineering and Management Sciences, Northwestern Univ., PhD dissertation, 1973.
6. Martin, B. V., and Manheim, M. L. A Research Program for Comparison of Traffic Assignment Techniques. Highway Research Record 88, 1965.
7. Huber, M. J., Boutwell, H. B., and Witheford, D. K. Comparative Analysis of Traffic Assignment Techniques With Actual Highway Use. NCHRP Rept. 58, 1968.
8. Potts, R. B., and Oliver, R. M. Flows in Transportation Networks. Academic Press, New York, 1972.
9. Starasinic, F. N., and Schuster, J. J. Comparative Analysis of Capacity Restraint Traffic Assignments. Traffic Engineering, 1972, pp. 48-52.
10. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. Proc. Institution of Civil Engineers, Vol. 1, Part II, 1952, pp. 325-378.
11. Zangwill, W. Nonlinear Programming: A Unified Approach. Prentice-Hall, 1969.

TIME-STAGED STRATEGY IN THE TRANSPORTATION PLANNING PROCESS

Richard J. Hocking, Barton-Aschman Associates, Inc., Chicago

Transportation planning projects are often the subject of controversy. This can be avoided or reduced by improving the planning process. The public has become sensitive to how a project is planned in addition to what is proposed. One of the components of the planning process is plan staging. It is recommended that the process be improved through use of an approach called a time-staged strategy. A strategy is a series of actions designed to achieve certain system states in a given time sequence. The choice of actions conforms to an overall strategy that, in turn, is an interpretation of planning goals and objectives and their relative priorities (determined by community participation programs). A series of alternative single-purpose strategies is initially designated and converted in representative physical plans. This involves an evaluation of how well the elements of the recommended long-range plan comply with a given strategy. These single-purpose strategy plans are merged into a composite plan to identify those elements serving multiple strategies. These results are further evaluated by considering the ranking of the alternative strategies in order to determine what plan elements are most important in a staging sequence. The composite strategy plan and the top four strategies represent an overall strategy. With this as a guide, a series of activities can be devised that will improve the system achieving the goals and objectives valued most highly and that will achieve certain system states at given times. Coordination of actions must be flexible to respond to technical uncertainties. By using a time-staged concept, overall control of action can be achieved.

•ACROSS THE COUNTRY, transportation projects are involved in heated controversy, with many groups arguing either side of the question. The most clamorous, however, usually seem to be those challenging the need for a given project.

Although this paper does not address the causes of these situations, factors that seem to create controversy include the following:

1. Mobility or the ability to travel is a major facet of our life-style; therefore, projects that affect it are of significant personal interest.
2. As a land use, transportation facilities can have harsh impacts on contiguous land uses by altering the existing environment. Often, the people who bear some of the costs receive none of the benefits.
3. Transportation is sometimes viewed as a negative environmental force that can have effects considered too severe by our ecology-sensitive value judgments.
4. As urban areas grow more complex, the interrelationships among population groups, their activities, and facilities become more difficult to understand; the one aspect that appears certain is that, as individuals, we are more dependent on one another.

The inherent conflict is that we need and demand transportation service; however, many times this conflicts with environmental objectives. Such conflicts usually arise

Publication of this paper sponsored by Committee on Transportation Systems Design.

with only segments of the population, for the negative impacts tend to have a localized character. This gives rise to many proponents and opponents of a transportation project.

These factors were noted to establish a framework for the considerations discussed in this paper. If the goal of professionals is to plan and guide the implementation of needed transportation services and facilities, these factors need to be addressed.

The difficulties are a challenge for the transportation planner. However, development of more innovative and imaginative solutions is not enough. Resolution of the conflicts and provision of needed facilities must begin with improvements in the planning process.

This call for a significant improvement in the planning process is based on a need to increase sensitivity to the transportation planning issues described. The success or failure of a given planning project will rest on how the project is done in addition to which project is recommended. This concern has already been reflected in major legislation—the National Environmental Protection Act. The preparation of environmental impact statements (EISs) should be the outgrowth of an improved planning process. The EIS should be a summary of a planning process that has been environmentally sensitive throughout its course. The game is changing, and we planning professionals need to change our game plan.

PLANNING PROCESS

Generally, the planning process is the series of activities related to the planning and designing of a transportation project, from recognition of a need (existing or future) to an acceptable implementation plan. As shown in Figure 1, the planning process has several standard components, each of which could be improved relative to technical methodology, interpretation of results, and so on. One component, the staging plan, which is the subject of this paper, should receive particular attention. It is central to the issues noted previously and can have significant bearing on improving the process.

Why Improve Staging?

Staging deals with the translation of general or long-range plans and concepts into a sequence illustrating desired system states at given time periods. Staging converts the planning process from the more abstract network plans into physical systems. It addresses the basic question of what should be done first. This is particularly important because initial steps may establish a commitment to a given system or may even be the only ones implemented and, hence, the only means to achieve desired benefits.

An Approach to Staging

The argument that improvement in staging of the planning process presumes that there are improvements to be made—i.e., that the traditional approach leaves something to be desired—is true.

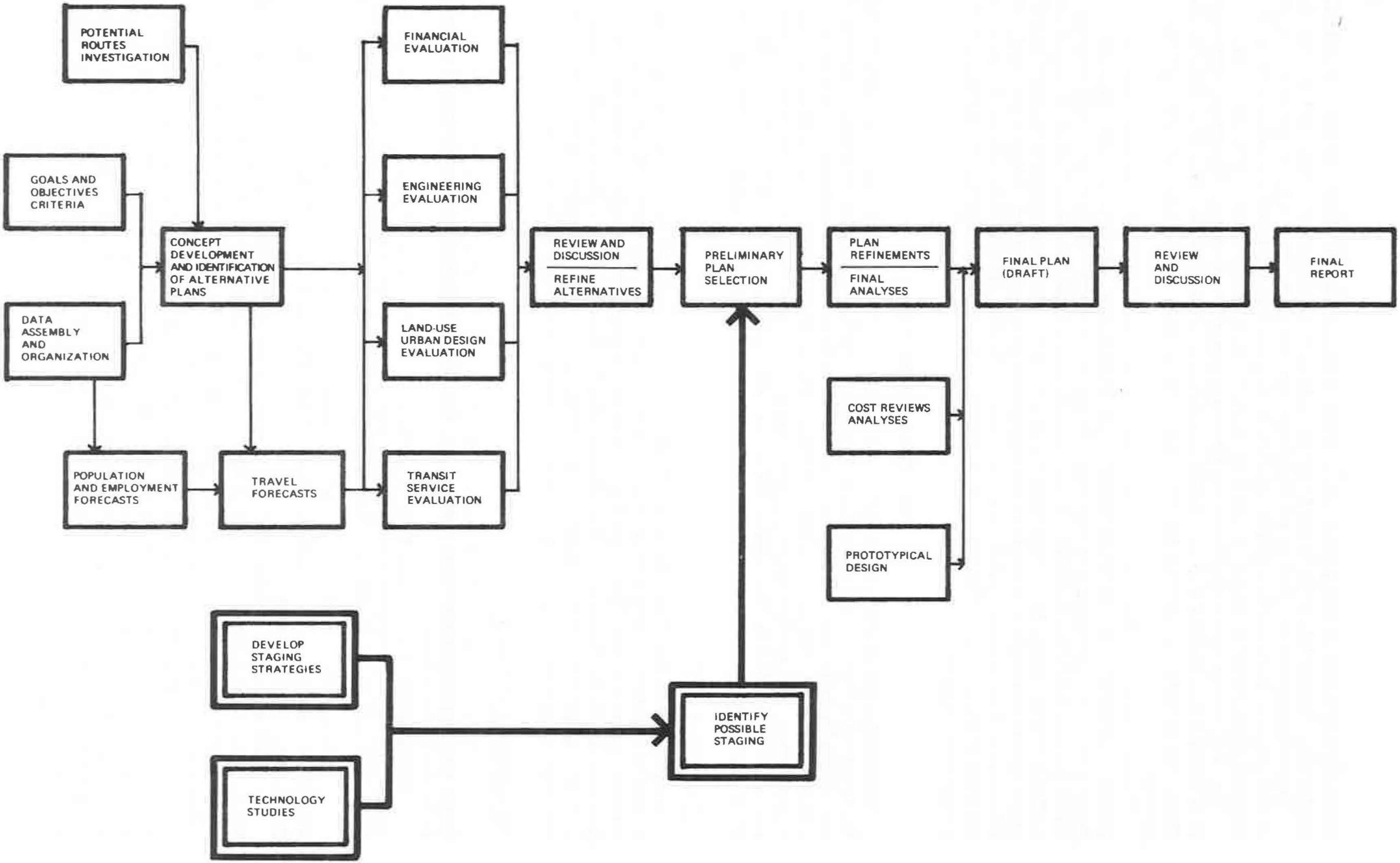
Traditionally, staging has dealt in purely physical terms. It has segmented a long-range plan into manageable components in a sequence that satisfied a logical pattern. This approach is not totally invalid, but it has become inadequate for several reasons.

1. It has not had a strong or explicit goals orientation. Goals and objectives were part of the abstract or conceptual phase of the study and too often were not explicitly expressed in action plans.

2. Plan staging has normally reflected geographical rather than population considerations; i.e., plan elements are selected to serve a particular part of the area or to implement route A first, rather than being selected to serve, for example, the dependent labor force first or to provide service to all major medical centers.

3. As citizen participation increases, the role of the professional planner changes. Rather than provide a final recommendation for a transportation improvement program or facilities, the professional may be asked only to evaluate alternatives and their consequences; then the community makes the final recommendation. This affects the staging plan in a similar way. Staging needs to be developed in a manner such that a public decision-making process can be employed. That is, if the staging plan is to be de-

Figure 1. Study work sequence.



veloped as suggested in item 2, community values and priorities must be identified and utilized.

4. As urban areas become increasingly complex, the ability to predict transportation needs becomes more precarious. Coupling this with the fact that the scope and cost of most transportation systems have become very high indicates the high risk of overcommitment to the wrong solution. Such a situation should be reflected in alterations to the traditional staging process to provide a more flexible process that maintains options for future decisions.

5. Increased urban complexity also means that there is a need to forecast more conditions or relationships. This becomes difficult for long-range planning and suggests the needed capability of developing plans and stages into functional entities that can be evaluated separately as well as in whole systems. By this means, the planner can ensure that the incremental value of stages is in balance with the cost required to achieve them.

This review of the traditional staging process provides an approach to improve techniques. It should have a goals orientation, reflect priorities, achieve identifiable functions and services, and be flexible.

BASIS FOR STAGING

The plan-staging process described in this paper was developed in relation to the public transportation (transit) planning process. The basic approach, however, is applicable to other planning projects. The essential change represented by this approach is use of staging in a time-staged strategy.

Why a Strategy

The word staging implies "the state or condition" of a system at a given time. A strategy, on the other hand, is a series of actions taken over a given time span in a specific sequence to achieve certain objectives or end states (stages). This concept has a significant characteristic, i.e., choice. The advantages of a strategy approach are as follows:

1. It has a goals orientation; hence, objectives are explicitly stated in terms of action.
2. It creates a framework for staging that allows flexibility in selecting activities so that more than one set of actions might be used to achieve a desired end state.
3. The approach emphasizes the dynamic path to a plan that is action oriented rather than a static end state; it provides a mechanism for relating "end-state" planning to "means-state" planning.
4. Because a strategy relates decisions to a shorter time frame, it is more relevant to community participation and plan-staging actions can be evaluated in relation to community values and priorities.

The general organization of this strategy approach is shown in Figure 2. The process contains several steps that yield two end products, i.e., a selected overall strategy and a series of actions aimed at achieving certain system stages.

Method

The time-staged strategy process uses the following steps (Fig. 2):

1. Develop alternative strategies,
2. Prepare single-purpose strategy plans,
3. Select a strategy (combination), and
4. Develop action program.

These steps are based on three important products of the planning process that would normally precede the staging activity: the statement of goals and objectives for the planned transportation system, the relative priority of these statements, and the preliminary conception of the long-range (overall) plan.

The four steps in the staging process are described as follows:

1. Alternative strategies—This is a conceptual step in which goals and objectives are used to identify the widest reasonable range of possible strategies. Goals suggest the various services or impacts to be achieved by the plan. These can be equated to a strategy. For example, a potential strategy could be a plan that seeks to serve transit-dependent areas as the primary objective. All action included in such a strategy would be focused in this single purpose.

2. Single-purpose strategy plan—This step is an evaluation of how well the elements in the long-range or overall plan achieve each of the potential strategies. That is, it is necessary to determine what elements would be effective if any of the strategies were selected as the single purpose of the system. This assumes nearly perfect performance of each plan element.

3. Selection of a strategy—The stated goals and priorities can be used to rank alternative strategies by their functional importance, i.e., priority ranking. However, because the strategies are developed relative to single purposes, it is unlikely that only one would be desired to guide the staging process. A series or combination of strategies would allow the entire set of planning goals and objectives to be achieved to some extent. Hence, a combination strategy would be developed by building a composite strategy plan based on step 2.

4. Action program—With a selected combination strategy, a series of specific actions can be developed that would maximize the system benefits compatible with community priorities. Activity selection also requires an integration of other factors such as operational objectives, physical constraints, and general financial conditions. Because the result may be somewhat complex, the actions are organized into a time-staged decision or management network.

Nature of the Process

Besides an understanding of the basic methodology for the time-staged strategy process, other aspects need to be discussed to completely present the concept. These aspects generally deal with the nature of the process in terms of content, elements of choice, and strategy objectives.

Content—As noted earlier, staging plans have traditionally dealt with the configuration of the plan, i.e., the various routes, lines, or links that make up the total system. This aspect is valid for a strategy, except that there are added aspects. A strategy for a transit plan must consider service (operations) and transportation technology. The former is an aspect relating to both plan configuration and vehicles, whereas the latter is concerned with vehicle systems and their change over time.

Configuration—To develop the strategy relative to plan configuration requires that uncertainty be considered. This may be present in the following ways:

1. Estimation of transit demand—This is probably more uncertain than would be the case for the forecasting of highway traffic volumes. It deals with new or expanded transit systems and significantly altered levels of service. However, because the analytical models are based on existing conditions and relationships, there is no assurance of how accurately these relationships can predict public response to new service or systems. Similar technical questions can be raised about land use and activity forecasts on which future trip estimation is based. Questions of feasibility and system capacity should, therefore, be subject to continuing examination as implementation of the plan proceeds.

2. System availability—This aspect pertains to the ability of the local government to implement the system as needed. Because the plan includes new routes and facilities, its implementation will involve community interaction and proper programming of financial resources. How can future public acceptance be forecast, and what assurance is there that sufficient funds will be available? Will new rights-of-way be available?

3. Community priorities—This is related to item 2. The funds may be technically available, but the community may have different priorities causing a reallocation of the funds. Will the public continue to accept a major commitment to transit and, possibly, reduce highway expenditures after initial plan approval? Much has been said about the

need for transit, but will the community support a significantly different priority? Do social planning objectives that increase the need for transit have significant priority?

4. Public policy—There are certain public policies that have a major impact on public transportation but that the public has no clear position on. Potential policy matters of this type are (a) citywide land use or development policy, (b) use of direct public action to implement development policy, (c) proof of financial feasibility of a transit plan, (d) attitude toward transit as an alternative to the automobile (rather than the attitude of "letting someone else use transit so that I can use the freeway"), and (e) acceptable order of magnitude of transit improvement costs.

Because of these uncertainties, a strategy needs to be flexible. Good strategies retain as many options as possible.

Technology—In a similar way, transportation technology needs to be considered so that it can be staged in an overall strategy. Because the use of technology involves many decisions, a general approach to it—its role and what hardware to select—is an important part of a strategy. There are three principal aspects to such an approach.

1. Goals. Among the goals are use of public transportation as an effective alternative to the private automobile, to affect land development patterns, and to increase mobility for all segments of the population. In terms of technology, these goals raise issues relative to the "automobile" character of transit, i.e., convenience, security, comfort, the need for permanence relative to development patterns, and the need for ubiquitous service capability. Each of these achievements is specific, but they all pertain to practical and important features of a public transportation system.

2. Timing. The implementation of transit may be considered in the context of short- or long-range needs, and long-range needs are at the heart of the problem for technology selection. That is, how can technology systems be selected so that they can serve short-term and long-term goals? It would appear that technology selection for short-term needs is an easier task. It is necessary to determine what is available or could be available to satisfy short-term transit needs. For comparison, long-term needs are complicated by many questions and issues that are difficult to answer. Technology selection is thus based on less solid ground and becomes quite uncertain. Because of these uncertainties, a technology approach must be flexible.

3. Environment. A final aspect for consideration in a technology approach concerns the long-term environment. That is, beyond the three goals discussed, the ultimate goal of a public transportation program is to support and stimulate the creation of a new urban living environment. That is, in the long-range future, a "new world" ought to be created that is substantially better than today's world. Transportation of the future should not be troubled by the same problems we experience today, e.g., intolerable congestion, pollution, lack of mobility. Thus, with an optimistic view of the distant future, the potentially large investments in a new public transportation system should support the development of a better environment. This is a challenge to be innovative, to have an ultimate view about the desired future environment, and to devise a plan that maximizes benefits for a long duration rather than being guided only by short-range problems.

Elements of Choice—The second major aspect of the strategy process is the elements of choice. Given a transit plan, there are various components that can be changed or manipulated within the context of a strategy. These are the matters that should receive maximum analytical attention by the professional planner and community. For a transit plan, the choice elements are as follows:

1. Level of service is perhaps the primary element of choice. The variations are many and include (a) feeder service having a dial-a-bus concept and a CBD distributor system (minimum walking distance and transfers), (b) corridor express service with feeder transit on each end, (c) park-and-ride service with express transit, (d) free fare in special districts, (e) varying headways depending on function, and (f) bus priority operations on streets and freeways.

2. The location of transit corridors is described in the plan. However, there are choices in total network location in terms of (a) timing, (b) extent (extension in a corridor),

(c) pace of station development, (d) facilities at station, and (e) creation of route structure to achieve particular service.

3. Technology or transit hardware can be varied by corridor or within major activity centers. The choice elements in technology selection include the following features: (a) vehicle type and character, (b) vehicle size, (c) guideway type, (d) control system and operating mode, and (e) accessibility of service.

Strategy Objectives—The final aspect of the strategy process pertains to a set of objectives needed to guide strategy development. These objectives are somewhat abstract. They apply to a public transportation plan. Possibly, for a specific planning area, some items would not be applicable or others would need to be added. In any event, a list of this type is needed as a basis for developing a specific strategy. The creation of this list is very similar to the task of converting planning goals and objectives into a set of criteria. The latter are end-state oriented, whereas strategy objectives tend to be means-oriented. However, they still represent an interpretation of the planning goals and objectives.

1. Showcase a new level of service as soon as possible;
2. Serve transportation needs that are created by new or desired land use development and that cannot be met by increased highway supply;
3. Tap markets not using the existing bus system;
4. Place transit in existing land use corridors in which highway capacity shortage is evident and there is no possibility to upgrade highway service;
5. Preserve or save right-of-way opportunities;
6. Demonstrate compatibility with environment;
7. Provide a means to gather information about travel response to new service;
8. Ensure immediate availability of financial resources (i.e., take advantage of grants, other special funds, etc.);
9. Preserve flexibility in the long-range system;
10. Increase transportation service relative to social objectives and their order of priority, i.e., service to particular social, ethnic, or other groups;
11. Generate the most logical route structure; and
12. Disperse transit benefits to various geographical districts.

STRATEGY DEVELOPMENT

The Dallas Public Transportation Study can serve as a case study for a time-staged strategy.

The Dallas study, completed in 1973, sought to develop a long-range transit plan with three basic goals:

1. Develop an alternative to an automobile-only transportation system,
2. Use public transportation as a land use and development-shaping force, and
3. Increase the mobility of all members of population.

The study site (Fig. 3) was the Dallas subregion, an area equivalent to Dallas County. The population was 1,327,000 in 1970 and is forecast to be 2,316,000 by 1990. Similarly, employment is forecast to grow from 647,000 in 1970 to 1,161,000 in 1990.

Dallas has enjoyed a development boom in recent years. This now is being spurred by the Dallas/Ft. Worth Airport. Based on this growth, the community concluded that there may be a need for a vastly expanded public transportation system. This attitude is based on the recognition of several issues:

1. The Dallas/Ft. Worth Airport opened in 1973. Within 10 to 15 years, traffic generated by the airport may exceed available highway capacity.
2. Building construction in the Dallas CBD is continuing, but access freeways are approaching capacity.
3. Concern for the environment, community preservation, and citizen involvement is becoming more significant. The latter especially includes a stronger voice for minority groups.
4. The new airport, in addition to normal economic pressures, will generate sub-

stantial growth in the midcities subregion, western Dallas and eastern Tarrant Counties. These patterns could be affected by the provision of increased accessibility at desirable locations.

Alternative Strategies

The strategy approach attempts to translate the components of the basic plan into a priority description in which each component is evaluated by its ability to achieve a certain strategy. Once so identified, components can be combined to maximize the achievement of those strategies deemed most important. Such combination would, of course, be also tested in terms of physical and operational logic.

For Dallas, a major goal was to closely coordinate the land use or development concept with that for public transportation. Figures 4 and 5 illustrate the recommended development concept and long-range transit plan.

The development concept reflects a compromise between a trends plan and satellite cities concept. Significant outlying multipurpose activity centers and high accessibility corridors are envisioned. The Dallas CBD would continue to be a dominant center. The transit plan parallels the concept. Strong CBD access is provided, but various crosstown routes are added to give the network a grid-like character.

These plans and the strategy objectives were used to formulate a set of alternative strategies. Each has a single major purpose; however, the strategies are not mutually exclusive. A certain amount of duplication exists indicating support between various strategy objectives. This approach to alternative strategies is used so that a wide range of possibilities is considered.

Strategy 1. Access to High Activity Centers—This focuses access on specific development nodes, the Dallas CBD and the airport. Access to these centers is provided for major travel groups, i.e., air travelers, CBD shoppers, and CBD employees.

Strategy 2. Modal Split—This strategy attempts to achieve automobile trip diversion. Transit is needed in corridors in which travel is primarily by commuters, freeways are at or near capacity, or constraints have been placed on further highway expansion.

Strategy 3. Showcase—This focuses on attempts to use transit in innovative ways to show potential and to reflect Dallas' reputation as a progressive urban area. Service is provided to unique high-accessibility corridors, new technology is used for distribution, the multiple land use concept is implemented, and right-of-way opportunities are conserved.

Strategy 4. Social Objectives—This focuses on providing transit service to transit dependents with specific service to black and Mexican-American areas. Transportation access is created between employment and residential areas.

Strategy 5. Technology Evolution—This strategy attempts to apply new technology at an early time. It employs a demonstration project approach and seeks to establish orderly evolution of transit hardware and other facilities. This also means that future options, in terms of vehicles and guideway, would be preserved as new systems become available.

Strategy 6. Land Use Concept—This strategy seeks to support a subregional development concept through variations in accessibility. The strategy attempts to complement the multipurpose centers concept, the complete communities concept, and the use of transit to encourage balanced growth patterns.

Strategy 7. Political Support—The strategy focuses on approval of implementation by the primary political decision-makers. Transit service may have to be widely distributed so that benefits are proportional to political influence. This also considers regional versus local interests.

Strategy 8. Environmental Protection—Transit might be viewed as a means to reduce negative environmental impacts because of reduced right-of-way needs (compared to freeways) and air pollution by redirecting urban development trends away from sensitive open space areas.

Strategy 9. Industrial Growth—The focus of this strategy is on the provision of transit access to existing and emerging industrial districts as a means to enhance economic growth. Particular emphasis is given to new districts or those proposed in the development concept.

Figure 2. Time-staged strategy of development process.

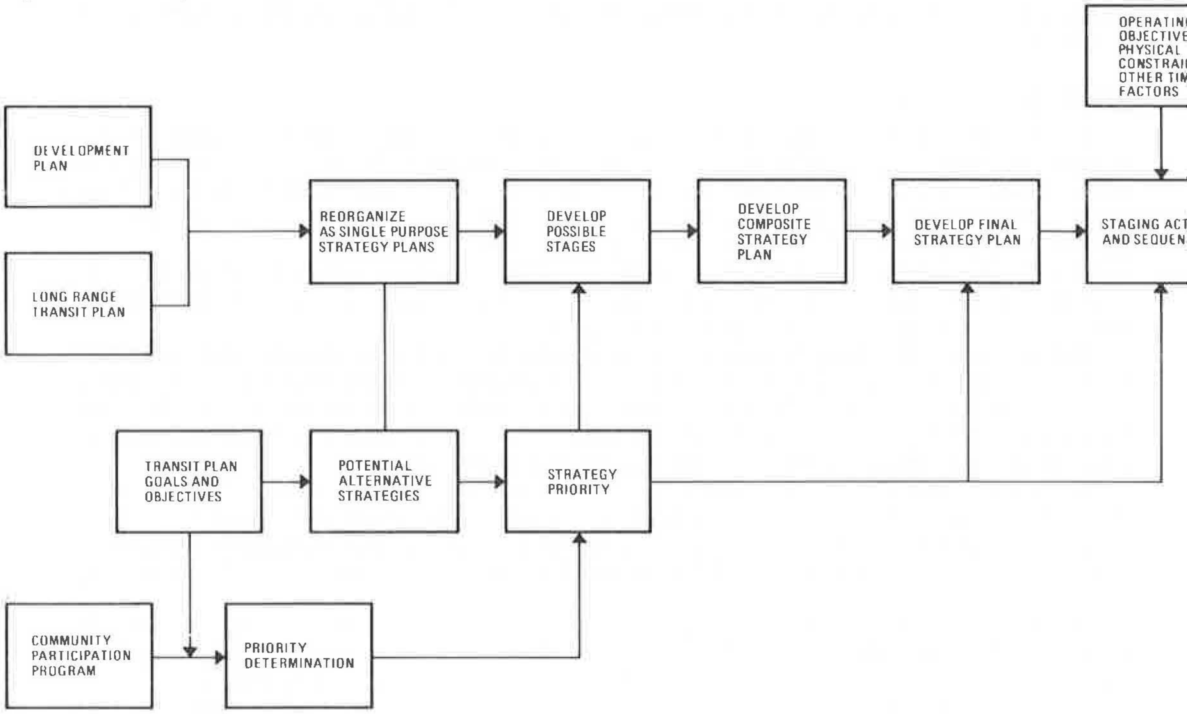


Figure 3. Dallas study area.

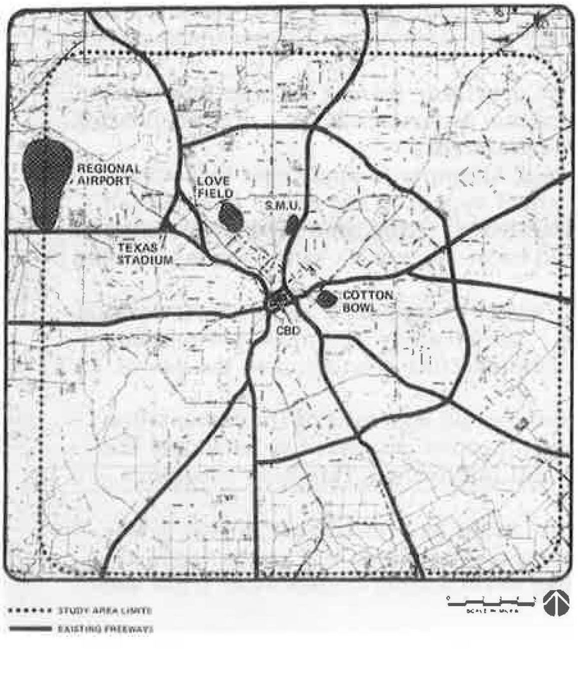
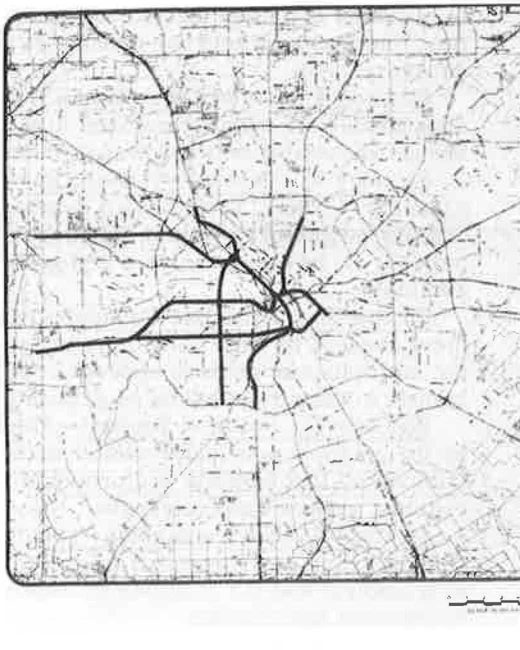


Figure 4. Dallas development concept.



Strategy 10. Facilitate Implementation of Current Plans—This strategy provides transit to support current proposals including the North Central Busway, urban tracked-air-cushioned vehicles to the airport, and the CBD transportation center plan.

Strategy 11. Economic—Under this strategy, realistic consideration is given to the availability of economic resources. The proposed transportation improvements need to be balanced with resources in terms of total amount and pace of availability.

Application of Alternative Strategies

Each alternative strategy represents a means concept; each creates a focus for continuing planning and implementation activities. The first step in translating the strategy into a specific action plan involves the conversion of the strategy into physical terms. This is accomplished by evaluating each element of the overall long-range plan in terms of its functional role to achieve the objective of each strategy. Combining various elements that achieve a given function yields a new plan. Each "strategy plan" represents a special version of the long-range plan wherein only certain portions are used. Each one is further evaluated relative to plan goals, objectives, and priorities to identify potential first and second stages for each strategy plan.

For Dallas, a series of alternative strategy plans was developed. A portion of these plans is shown in Figures 6, 7, and 8. The following brief notes attempt to describe the rationale for each:

1. Strategy 1—Social objectives (a) provide transit service to link labor force with employment from residential areas in the south, southwest, and west sections to the CBD, Stemmons Freeway commercial district, Redbird industrial district, and Grand Prairie industrial area; and (b) provide access to social service and medical institution areas (University of Texas Medical Center, Baylor University Medical Center).
2. Strategy 2—Technology evolution (a) uses North Central Expressway as a location for a busway; (b) uses Texas-183 as a route for special airport transit service; (c) uses Love Expressway for freeway flyer service; (d) provides dial-a-ride service for north Dallas; and (e) develops park-and-ride facilities at intercepting locations along three freeways serving the CBD.
3. Strategy 9—Industrial growth (a) provides high level of transportation access to Redbird, midcities, North Stemmons Freeway, and Fair Park industrial areas; and (b) uses freeway corridors extended toward the Flower Mound and Plano development areas as locations for new industry.

STRATEGY SELECTION

The preceding steps of the time-staged strategy will yield alternative strategy plans that incorporate elements of the long-range plan in a manner to achieve the planning of goals and objectives. These alternatives must be evaluated to produce one time-staged strategy as a basis for staging the long-range plan. This is strategy selection.

The selection step is composed of two parts. First, each single-purpose strategy plan is compared to determine which element of the plan is common to several strategy plans. This allows the assessment of the role each plan element could play in achieving several planning goals and objectives. The second part involves combining those plan elements that achieve high-ranking goals and objectives, i.e., achievement of strategies in proportion to ranking. The result is a time-staged strategy for the plan.

Composite Strategy Plans

A plan overlay technique is used to compare the alternative strategy plans to identify common elements. The result is a composite strategy plan, as shown in Figure 9.

For Dallas, the composite plan was mapped to show potential first stage elements of each strategy plan. The result identifies plan elements that would show a significant performance in achieving multiple goals and objectives. Plotting potential first stage elements attempts to focus attention on those that should be candidates for the first stage.

Figure 5. Dallas transit plan.

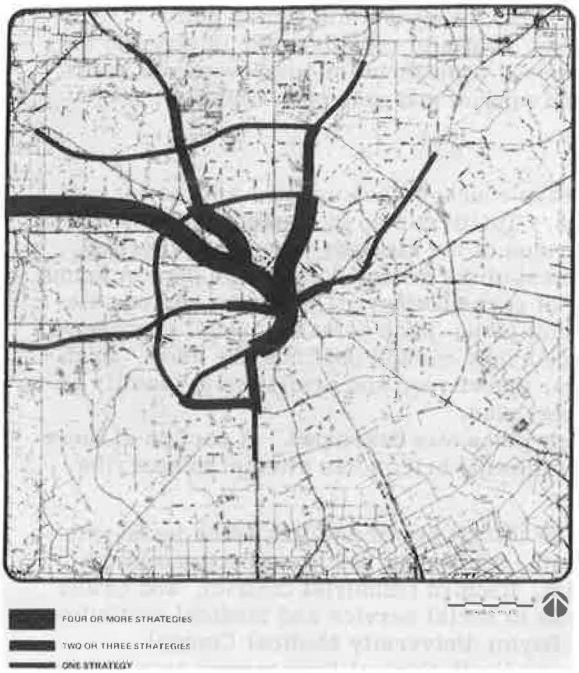


Figure 6. Strategy 1: social objectives.

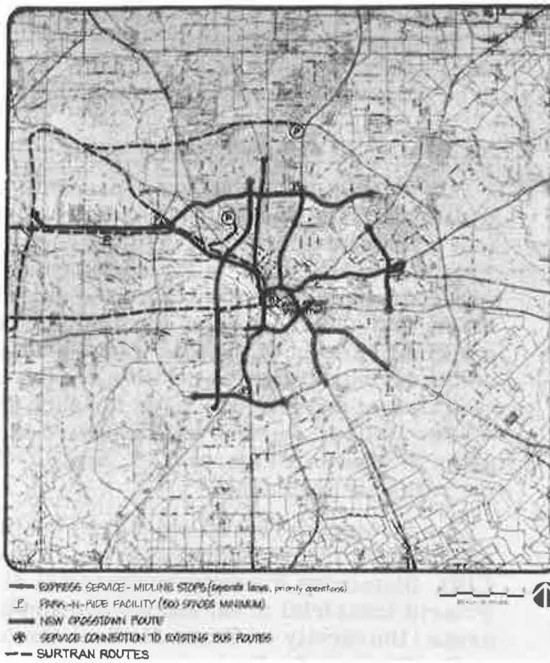


Figure 7. Strategy 2: technology evolution.

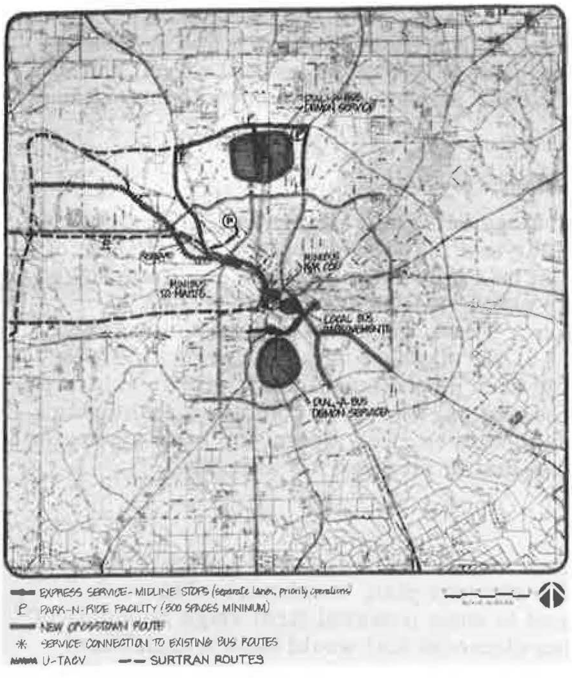
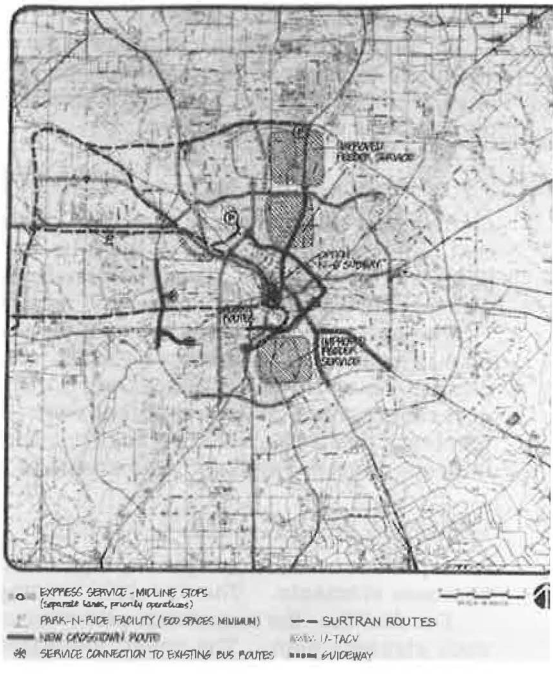


Figure 8. Strategy 9: industrial growth.



A Time-Staged Strategy

If all strategies have equal importance, the composite strategy plan can be used directly as the final strategy. However, this equality is not normally present because of priority ranking. This rank is determined by community values and judgments developed in the community participation process. When this aspect is introduced, the composite plan can be evaluated by weighting the importance of plan elements according to the ranking of strategies. The result is representative of a final strategy.

Recommended Strategy and Staging Plan—For the Dallas project, the attitudes expressed by local planning officials and citizens concerning the relative priority of planning objectives indicated that strategies 1 and 4, high activity center access and social objectives, had the highest priority. Strategies 2 and 5, modal split and technology evolution, had second highest priority.

With this finding and the composite plan analyses results, the relative value of each element of the plan was assessed; i.e., each one was described in a priority listing for consideration in successive plan stages. On this basis, preliminary staging for the Dallas subregional transit plan was devised. This is shown in Figures 10, 11, 12, and 13.

The overall concept for the recommended staging plan is that major improvements would begin in the central parts of the city and move outward by corridor according to apparent trends in urban growth. The provision of increasing transit service levels employs a pattern of establishing transit service first with buses, then with the more expensive guideway. This attempts to reduce the uncertainty and risk in major investments. The general sequence of staging can be summarized as follows:

1. Stage 1: Upgrade level of transit service by using buses on freeways, provide improved CBD access including initial phases of a transit mall, introduce satellite parking concept, provide new crosstown bus routes for access to Stemmons Freeway district and medical center and from north Dallas to Texas 183 corridor, and introduce special bus service to new regional airport.
2. Stage 2: Extended freeway express bus service to include all corridors within I-635, add more crosstown bus service, develop and implement demonstration dial-a-ride bus in different socioeconomic areas, introduce secondary transit service in CBD and Stemmons Freeway corridor, and implement urban tracked-air-cushioned vehicle (demonstration project) between CBD and regional airport.
3. Stage 3: Develop first sections of guideway system along North Central Expressway and south into Oak Cliff including first CBD subway along east-west corridor, add more crosstown service, develop permanent collector system in north and south corridors based on dial-a-ride results, and improve secondary transit in the CBD.
4. Stage 4: Develop north-south subregion in CBD and extend east and west subway to Fair Park, Baylor, and West Dallas areas; add stations on urban tracked-air-cushioned vehicle route; expand collector systems; and extend transit coverage in outer areas of region.

Time-Staged Decisions—The recommended plan stages and the individual projects included represent the general state of the transit service system at various points in time. These states can be achieved by undertaking a series of actions. These represent the final part of the time-staged strategy. The selection of actions would be guided by the strategy or combination strategy and would be aimed at the physical or operational performance illustrated by the plan stages. The actions are the means, and plan stages are the ends.

Further, the recommended plan stages are an outgrowth of the best knowledge available at the time. However, as observed earlier, there is a significant degree of uncertainty in both technical and public policy matters. Therefore, the staging plan and the implementation actions must be viewed as flexible products of the planning process.

Flexibility or a propensity to change plan recommendations could produce confusion and disorder in the implementation period that could destroy any chances to achieve the desired goals and objectives. Flexibility needs to be controlled in an orderly way. The selected strategy creates an overall framework for such control. This is supplemented by a management process that guides and coordinates the various actions undertaken during each plan stage.

Figure 9. Composite strategy plan.

Figure 10. Preliminary Dallas transit plan—first stage.

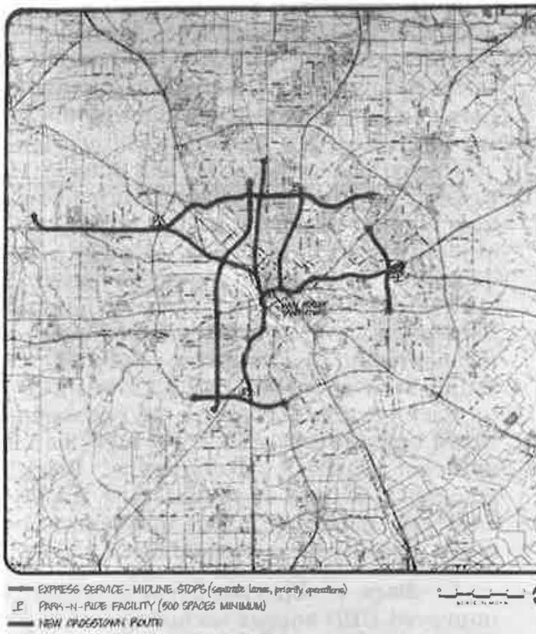


Figure 11. Preliminary Dallas transit plan—second stage.

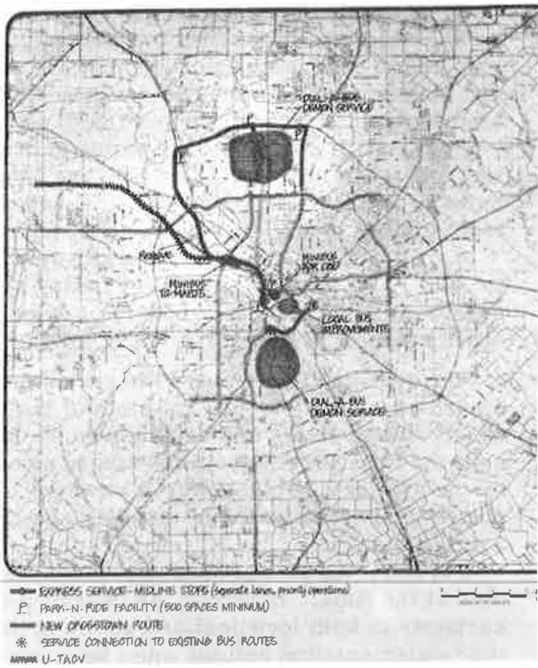


Figure 12. Prelimininary Dallas transit plan—third stage.

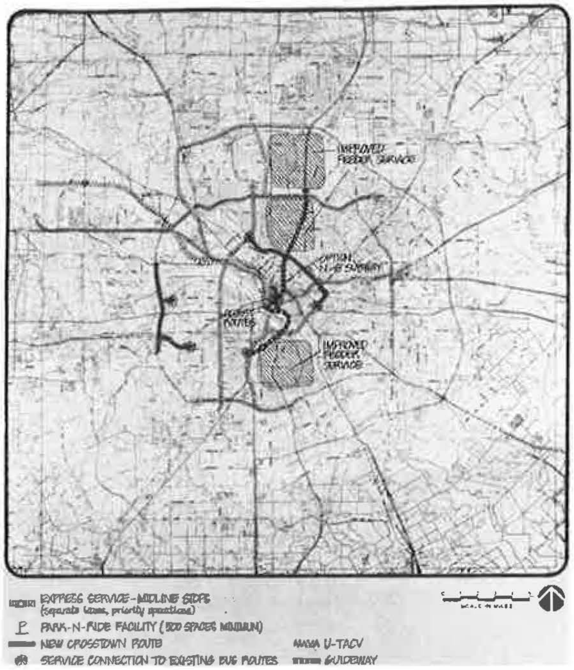
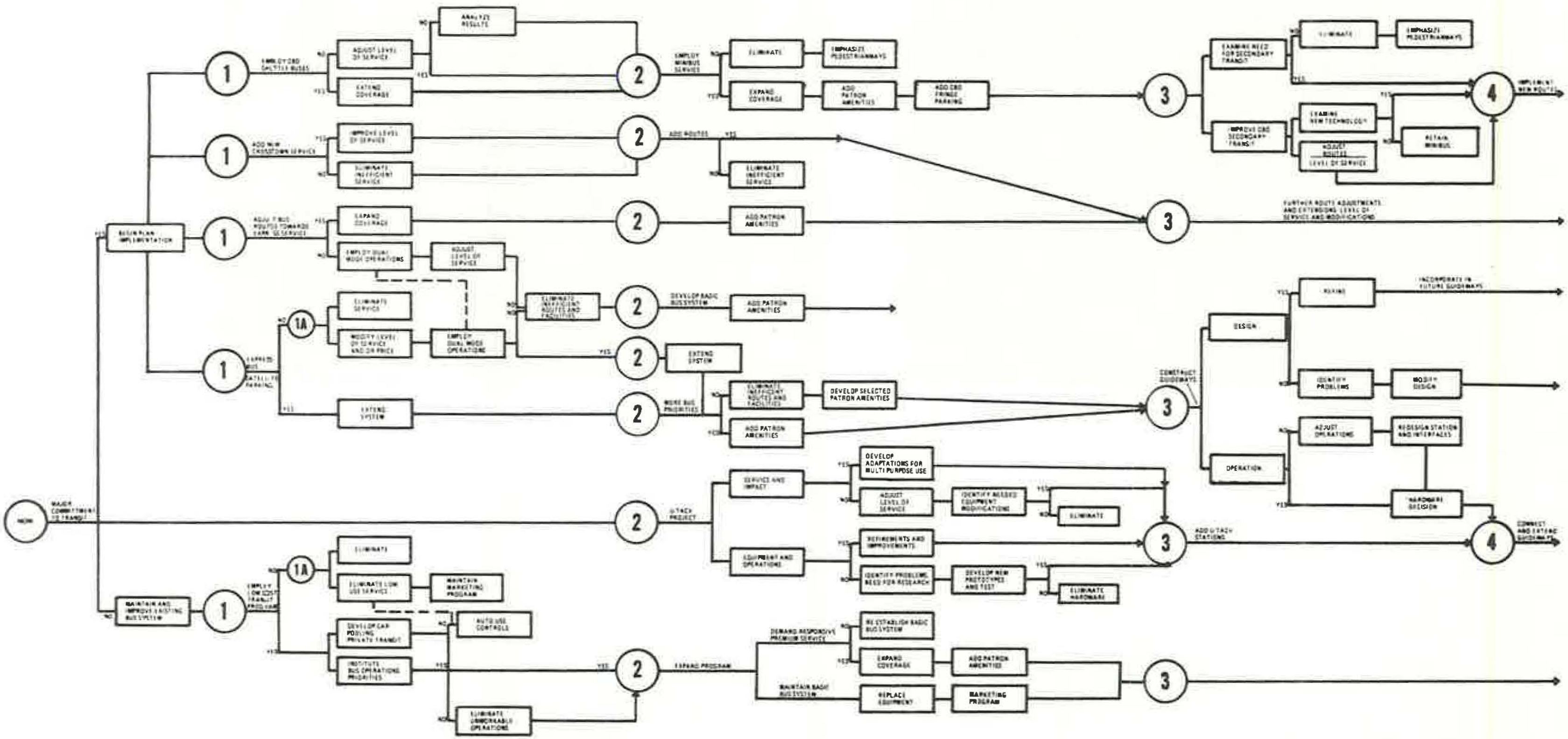


Figure 13. Preliminary Dallas transit plan—fourth stage.

Figure 14. Time-staged decisions.



The management device suggested is a time-staged decision system. A general outline of this system is shown in Figure 14. The time-staged decision concept provides for a logical sequence of detailed activities occurring in each stage.

As suggested by the figure, the action program for each stage is segregated into various sets. Each one is organized according to a major functional subsystem of the transportation plan. For Dallas, these were the guideway system, crosstown service, collector-distributor service, urban tracked-air-cushioned vehicle demonstration project, and a low capital investment bus system (as an alternative to the high investment program).

With each action set, there are identifiable sequences of activities. They attempt to deal with the various technical uncertainties and the continuing influence of a community participation program. The decision concept is derived from the multiple-decision points in this management system. There are several go/no-go points or decisions relative to a general course of action. Whatever the actions or decisions in this network, they would be guided by the time-staged strategy selected for the planning program. Hence, flexibility is provided but with consistent direction.

Because community priorities can significantly change over time, the selected strategy needs to be tested or evaluated on a regular basis. Within the decision system, this would be done at or near the completion or beginning of any stage, before commitment was made to a high investment project.

CONCLUSIONS

The time-staged strategy approach offers a means to improve the planning process. It provides a means to maximize the effect of preselected goals, objectives, and community participation on implementation action. The actions selected for the staging program are then a more explicit interpretation of citizens' desires for their community.

The use of strategy creates an overall framework that guides the selection of actions. The accent on actions reflects the emphasis on implementation. Stages are merely system states during the implementation process. This orientation gives emphasis to achievement of goals and objectives. Further, by creating a framework, the process has flexibility. Specific actions can be modified as values change or as new data become available.

These attributes are needed in the planning process. Urban problems are difficult; in many cases, it is very difficult to develop confidence in long-range plans. Such plans should be developed to guide short-range actions. However, if long-range plans are questionable, short-range action still cannot be forestalled because of the pressing need for transportation service. The strategy approach is workable in this context and allows the planning-implementation process to move forward.

IMPLEMENTATION OF OPERATIONAL NETWORK EQUILIBRIUM PROCEDURES

Earl R. Ruiter, Cambridge Systematics, Inc., Massachusetts

Operational network equilibrium procedures are being developed for the fixed-demand single-mode case. The basis of these procedures is described in the light of the historical development of approaches to the problem of predicting equilibrium flows in transportation networks. The procedures are described as capacity restraint methods that have the following advantages over traditional approaches: On each iteration, improvement of the solution is ensured; and following each iteration, a measure that indicates the maximum amount of error remaining can be calculated. This paper describes network equilibrium procedures being made operational as a part of the UMTA Transportation Planning System. These procedures are described in light of their theoretical and mathematical background. Although significant theoretical work has been done on the variable-demand network equilibrium problem, the first developmental step being taken is to provide an efficient fixed-demand equilibrium procedure. It is expected, however, that expansion to the variable-demand case will be possible within the general algorithmic framework being developed. The paper begins by stating the general (variable-demand) network equilibrium problem. This problem is then formulated mathematically, and the nature of its solution is discussed. Previous work to develop efficient solution techniques is discussed. The results of much of the previous work are summarized as a general equilibrium algorithm for the fixed-demand problem. Finally, based on this general algorithm, current development work is described.

•THE PROBLEM of predicting flow equilibrium in transportation networks is in determining the values of interzonal flows and costs and link flows and costs. (Cost is used in a very general sense to represent in a single variable a combination of things such as travel time, fares, operating expenses, and discomfort.)

These are the output variables; the inputs are the structure of the transportation network, sets of link supply and interzonal demand functions, and flow distribution rules. Because deterministic and static, or steady-state, inputs are used, the output variables are also deterministic and static. They therefore represent constant or average conditions over a period such as a peak hour or an average day.

The components of the inputs to the flow equilibrium problem listed above are described as follows:

1. The network—A network is composed of nodes and ordered pairs of these nodes termed links. Links connect two nodes and allow flow to occur in only one direction between them. Some of the nodes are zones at which trips enter and leave the network.
2. Supply functions—Each link has associated with it not only a flow but also an impedance to flow in the form of a travel time or generalized cost. The relationship between link flow and link cost is expressed by a supply function that indicates how cost increases as flow increases. Typically, the supply function for each link may have an asymptote at a maximum flow level or capacity.

Publication of this paper sponsored by Committee on Transportation Systems Design.

3. Demand functions—Each zone pair has associated with it a demand function that relates the origin-destination (O-D) travel cost to the volume of travel that will flow from origin to destination. In the variable-demand case, this volume of travel decreases as the cost increases. In the fixed-demand case, the volume of travel remains constant for all levels of cost.

4. Flow distribution rules—A flow distribution rule that describes how travelers route themselves over links to move from an origin zone to a destination zone is assumed to exist. This rule can imply either individual route choice, systemwide control of route choice, or some combination of these. For the representation of highway travel by private vehicles, the common assumption is that individuals choose a minimum cost route. The results of this flow distribution rule are that all routes chosen from any origin to any destination will have equal travel costs and that all other routes will have higher travel costs. These results are termed Wardrop's first principle (26) or a user-optimized flow pattern (4).

For the representation of travel by vehicles belonging to a single authority, such as a railroad providing freight service, a logical assumption is that the single authority wishes to maximize its total consumer's surplus and that its flow distribution rule is to make routing decisions with this objective in mind. The result of this flow distribution rule has been termed Wardrop's second principle or a system-optimized flow pattern.

Our concern is with the prediction of user-optimized flow patterns, although the relationships of these two flow patterns will also be explored. When user-optimized flow patterns are obtained, Wardrop's first principle states that there will be a unique travel cost for each zone pair.

MATHEMATICAL FORMULATION

The mathematical relationships that exist between the components of the equilibrium problem will be detailed here for the user-optimized problem. This has been done in the literature in a number of ways, based on alternative mathematical descriptions of network flows. The approach used here has been borrowed largely from Kulash (11).

The following notation is used:

a = a typical link connecting two nodes,
k = a typical O-D pair,
m = a typical path for a given O-D pair, and
P_{km} = $(a_1, \ldots, a_n)$ = the set of links on path m connecting O-D pair k.

The links included in each P_{km} constitute a single path from the origin to the destination of k. This path must be free of loops, and all links included in the path must be used in proceeding from origin to destination.

This notation can be used to define the following variables:

f_a, c_a = flow and cost on link a,
f^{km}, c^{km} = flow and cost on path m for O-D pair k, and
f^k, c^k = flow and cost for O-D pair k.

The relationships between these variables are the following:

1. The network structure gives rise to flow relationships for interzonal flows:

$$f^k = \sum_{\text{all } m} f^{km} \tag{1}$$

for all k; for link flows:

$$f_a = \sum_{\text{all } k,\, m} f^{km} \tag{2}$$

for all a and for which $a \in P_{km}$; and for path costs:

$$c^{km} = \sum_{\text{all a}} c_a \tag{3}$$

for all k,m and for which $a \in P_{km}$.

2. The supply relationships are

$$c_a = s_a(f_a) \tag{4}$$

for all a where s_a is a function.

3. The demand relationships are

$$f^k = d^k(c^k) \tag{5}$$

for all k where d^k is a function.

4. The flow distribution rule, for a user-optimized flow pattern, gives rise to the following equilibrium relationships:

$$c^{km} \begin{Bmatrix} = c^k & \text{if } f^{km} > 0 \\ \geq c^k & \text{if } f^{km} = 0 \end{Bmatrix} \tag{6}$$

for all k.

The properties of the solution to the user-optimized network equilibrium problem can be obtained by defining an equivalent optimization problem. This can be done as follows:

1. For each demand function d^k (Eq. 5), define the inverse function g^k such that

$$c^k = g^k(f^k) \tag{7}$$

2. Define a new link function, $\underline{S}_a$, as follows:

$$\underline{S}_a(f_a) = \int_0^{f_a} s_a(x)dx \tag{8}$$

3. Define a new interzonal function, $\underline{G}^k$, as follows:

$$\underline{G}^k(f^k) = \int_0^{f^k} g^k(x)dx \tag{9}$$

The equivalent optimization problem is then

$$\text{Maximize } Z = \sum_k \underline{G}^k(f^k) - \sum_a \underline{S}_a(f_a) \tag{10}$$

subject to Eqs. 1 to 3.

This equivalence is proved by a number of mathematicians, including Gibert (9) and Murchland (17) for the general case and by Dafermos (5) for the fixed-demand case. In the fixed-demand case, the function g^k cannot be obtained. There is, however, an analogous optimization problem:

$$\text{Minimize } Z = \sum_a \underline{S}_a(f_a) \tag{11}$$

subject to Eqs. 1 to 3.

After the equivalency of the two problems has been demonstrated, the mathematics of nonlinear convex programming was used to prove that the solution of both problems exists, is unique, and is stable.

REVIEW OF EQUILIBRIUM APPROACHES

Problem Formation

It appears that the first recognition of the difference between user-optimized and system-optimized network flows was described by Pigou in 1920 (20), who demonstrated for a simple two-link, two-node network. Current interest in the problem, however, dates from Wardrop's statements of the two kinds of problems in 1952 (26).

Subsequent work on the formulation of the network equilibrium problem was done by Beckmann, McGuire, and Winsten (1), Prager (21), and Jorgensen (13). Jorgensen showed that if the supply functions (Eq. 4) are used to define a new set of functions $S_a^{*}(f_a)$ by using the relation

$$S_a^{*}(f_a) = 1/f_a \int_0^{f_a} S_a(v)dv \qquad (12)$$

then any flow pattern that is user-optimizing with respect to the set of cost functions $S_a(f_a)$ is at the same time system-optimizing with respect to the set of cost functions $S_a^{*}(f_a)$.

Solution Procedures

Based on the foundations laid in the 1950s and early 1960s, solution procedures have been developed that can be divided into four general classes: traffic assignment approaches, mathematical programming approaches, algorithmic approaches with fixed demands, and algorithmic approaches with varying demands.

Traffic Assignment Approaches—This class of solution procedures has by far predominated the other classes in actual application and in number of variants. [For an early survey, see Martin, Memmott, and Bone (15). The most common methods are described in the FHWA Traffic Assignment Manual (8).] Here, it is only necessary to note the major deficiencies of these approaches as methods for solving the network equilibrium problem:

1. Link travel times have often been kept constant, thereby ignoring the existence of link supply functions;
2. Origin-destination trips have often been kept constant, thereby ignoring the existence of travel demand functions;
3. The number of paths traveled between each origin and destination has often been limited to one, making it impossible, normally, to satisfy Wardrop's first principle;
4. The accuracy of the approaches as approximations of equilibrium has not been determined (this includes both their convergence properties, if they involve iterations, and their expected errors upon completion).

These deficiencies are not inherent in the traffic assignment process, and all of them are not true for each assignment procedure. Indeed, the procedure developed by Martin and Manheim (14), and implemented in transportation analysis systems at M.I.T. (14, 22), has only the last deficiency mentioned. Similarly, the package of assignment programs developed by Wigan (27, 28) includes procedures that have all features listed above except proven convergence properties.

Mathematical Programming Approaches—Charnes and Cooper (3) have developed linear programming solutions to network equilibrium problems with fixed demands. Their contribution is the multicopy assignment algorithm, which takes advantage of the specific structure of the linear program they formulate.

Yang and Snell (30) formulated a nonlinear equilibrium problem with fixed demands and developed a solution algorithm based on the maximum principle of Pontryagin. Tomlin (24) formulated a quadratic programming problem involving both the assignment of traffic and the distribution of trip ends over all destinations by using a gravity model. The major problem with all mathematical programming approaches is the prohibitive solution cost for real-sized problems.

Algorithmic Approaches With Fixed Demands—Three major efforts are known that have led to the development of network equilibrium algorithms for the fixed-demand case. These algorithms are significantly more efficient than the mathematical programming approaches. In each case, the improvement over programming approaches is obtained by using each of the following features of the network equilibrium problem:

1. The relationship between the system-optimizing and user-optimizing problems;
2. The theorems of mathematical programming, which are applicable because of the first feature; and
3. The process actually used by travelers to progress to equilibrium.

Expanding on Jorgensen's work, Mosher (16) was the first to formulate the user-optimizing equilibrium problem explicitly and to develop a solution algorithm that can be shown to converge. Dafermos and Sparrow (6) and Dafermos (4, 5) have developed more general algorithms. These algorithms are not limited to linear functions and have been extended explicitly to cases where the supply functions are of the following form:

$$c_a = S_a(f_1, \ldots, f_J) \tag{13}$$

where $f_1, \ldots, f_J$ are a subset of all network links. This extension is useful for representing delays due to two-way traffic on facilities and to intersection flows. A second extension involves the definition of multiple user groups, which can represent different vehicle types or users of different modes. A third set of algorithms for the fixed-demand case has been developed by Bruynooghe, Gibert, and Sakarovitch (2, 10). Their major advance is the elimination of the need to specify paths prior to the start of the procedure. New paths are found as the algorithms progress by using a minimum path procedure.

Algorithmic Approaches With Varying Demands—A number of algorithms have been developed to obtain solutions to the general problem of user-optimized network equilibrium when both demands and supplies vary with travel cost. These are very recent developments developed since 1967.

As an extension of the final fixed-demand algorithm described previously, Gibert (9) developed what appears to be the first variable-demand algorithm with proven convergence properties. Expanding on the work of Gibert, Murchland (17) has described the network equilibrium problem with varying demands in a way that explicitly brings out the relationships between the system- and user-optimized problems. Rather than specify exactly the steps of an algorithm, Murchland gives four principles for their development and states that a number of algorithms should be developed based on these principles and then tested to determine the most efficient one. The principles stated are the following:

1. The algorithm should have as its goal the minimization of either the equivalent system-optimizing problem or its dual. Murchland suggests the use of an error indicator, δ, which is the difference between the objective functions for these two problems.
2. Because these two objective functions are equal at equilibrium, the algorithm can be stopped when δ is sufficiently small.
3. As the algorithm continues, δ can be minimized by forming linear combinations of old and new flow patterns.
4. Because the final solution will typically have flows on a number of paths between all origins and destinations, any single iteration method that will assign flows to a number of paths should improve the speed of convergence.

Murchland has used these principles to develop a research-oriented network equilibrium computer program.

Two approaches to network equilibrium with varying demands have been developed in the United States. The first, by Wilkie and Stefanek (29), applies control theory to the user-optimized equilibrium problem. The second approach, by Kulash (11), is an initial effort involving only linear equations.

Finally, an algorithm has been developed by Netter and Sender (18, 19) that addresses explicitly the multiple user group, multiple dimensioned supply function (as in Eq. 13), and multiple dimensioned variable demand function problem. Netter and Sender show that multiple solutions exist unless the supply functions have a very restricted form. The algorithm is shown to converge to one of the multiple solutions; which one depends on the starting point chosen.

A GENERAL EQUILIBRIUM ALGORITHM FOR FIXED DEMANDS

The previous work done in developing network equilibrium solution procedures can be summarized by stating the features of these procedures that are essential to ensure convergence to the equilibrium solution, as agreed on by a number of authors, and that represent a minimum departure from existing traffic assignment procedures. (The restriction to minimum departures from existing procedures eliminates further consideration of approaches involving significantly more flow variables than used in traffic assignments.) These will be presented within the general algorithmic framework developed by Murchland (17), inasmuch as it can be applied to the fixed-demand problem (all f^k's fixed). After the elements of this framework are listed, some of the options available for each element will be described, emphasizing the suitability of existing production-oriented procedures as parts of equilibrium algorithms.

1. Step 1—Develop an initial network solution, S.
2. Step 2—Determine the best direction in which to proceed to obtain a new trial solution.
3. Step 3—Develop a trial solution, S_t.
4. Step 4—Use an optimization procedure to obtain the best next solution, as a combination of S and S_t. Symbolically, $S = C(S, S_t)$ where C is some combination.
5. Step 5—Determine whether S is a satisfactory final solution. If it is not, return to step 2.

Step 1—Initialization

Because any solution for which the flow conservation relationships hold is appropriate, this step can be accomplished very efficiently by assigning total demands in an all-or-nothing manner to the minimum cost paths corresponding to zero flow. This step concludes with an updating of all link and O-D cost variables. Normally, O-D cost variables will be set equal to the cost on the new minimum path for the O-D pair.

Step 2—Direction for Trial Solution

A new demand level for each O-D pair can best be obtained by adopting a value that equals the old value plus a fraction of the difference between the old value and the value predicted by the demand function at the current minimum path cost.

A number of authors show that the path over which new travel should occur for each O-D pair is the minimum cost path; its choice is assumed in the proofs of convergence. As an alternate, a multiple-path approach, using the link travel costs on the previous solution, can be used. A multiple-path solution for which the average travel cost is less on these new paths than on the old paths, using the old set of link costs, will also be satisfactory.

An important option for the whole algorithm is whether new solutions are developed separately for each O-D pair or at one time for the entire system. The choice of this option will determine whether steps 2, 3, and 4 are done in sequence separately for each O-D pair, or just one time, with an O-D pair loop within each step.

Step 3—Develop Trial Solution

With the directions developed in step 2, the trial solution can be developed by using standard loading and link cost updating procedures to determine all flow variables (new f_a and updated c_a and c^k) associated with this trial.

Step 4—Combine S and S_t to Obtain a New Solution

This is the critical step, because it is here that all existing capacity restraint methods fall short of being network equilibrium procedures with proven convergence properties. For convergence, it is necessary that the proportions of old and trial solutions be determined by the procedure itself rather than by the analyst.

A number of options exist with respect to the nature of the combination method, including the characteristics of the function itself, and the procedure for choosing the parameter of this function:

The Combination Function—If O-D pairs are considered separately, which requires saving the route of each path through the network and the corresponding path volume, then two combination functions are suggested:

1. A transfer of volume from the longest path for an O-D pair to the shortest as suggested by Dafermos and Gibert.
2. An increase in volume on the shortest path (a fraction of the trial solution) plus a proportional decrease on all previous paths as suggested by Murchland.

If only systemwide flow changes are made, no path volumes and routes need be saved. Then the only feasible combination function appears to be one corresponding to 2 above, a linear combination of the trial solution and the former solution.

The Combination Function Parameter—If O-D pairs are considered separately and the combination method of 1 above is used, the amount of volume shifted can be calculated based on maximizing the improvement to the objective function, Z (Eq. 11).

If combination method 2 is used or if systemwide flow changes are made, the fraction of the new flow to add to the remaining portion of the old flow can be obtained either by maximizing the change in the objective function, Z, or by minimizing an error measure for the new solution. The details of the former approach are described in the next section. After a new solution is obtained, all link and O-D cost variables should be updated to represent the new flows.

Step 5—Apply Stopping Rule

A number of stopping rules can be envisioned. These will take different forms depending on the method used to determine the combination parameter in step 4:

1. Stop when the change in the objective function Z is small compared to Z itself: $(\Delta Z/Z) \leq \epsilon$.
2. Similarly, if an error function is used, stop when the change in the function is small compared to the function value itself.
3. Stop when a specified number of iterations have been performed.
4. Stop when a specified computing cost, measured in dollars or CPU minutes, has been spent.

Whichever stopping rule is used, the final printout should include measures of the remaining error.

It is useful to summarize the various components that can be used to provide the options discussed and to state their availability.

1. Efficient minimum path, link loading, and link updating capabilities are available in a number of traffic assignment packages. One of these, Dial's STOCH procedure (7), provides an efficient multiple-path assignment capability.
2. A variant of the ability to form linear combinations of two sets of link loadings is included in Wigan's system.
3. The ability to obtain an error measure for any flow pattern that indicates its maximum variation from an equilibrium solution and the nature of such measures are discussed by Murchland and Wigan.
4. The ability to determine the fraction that should be used in forming a linear combination of two flow patterns so as to minimize the error measure is discussed by Murchland, Gibert, and Dafermos.

AN OPERATIONAL ALGORITHM

A number of the options described previously are being investigated, in preparation for specifying additions to the UTP system to incorporate network equilibrium. The basic algorithm serves as the standard of comparison for the efficiency and accuracy of all options developed. This algorithm is basic in that it makes maximum use of available assignment procedures and data structures. Alternatives to this basic algorithm will be judged by comparing their benefits—in terms of increased efficiency and accuracy—to their costs in terms of extra development time and, in some cases, computer storage requirements.

The basic algorithm is described in this section, and the following notation is used:

P^i = set of minimum paths between all zone pairs k, for iteration i;
F_i = set of link flows for all links a, for iteration i;
C_i = set of link costs for all links a, for iteration i;
M_i = set of link supply function slopes for all links a, at the flow levels F_i;
ΔZ_i = change in value of the objective function (Eq. 11); and
$S_a(x)$ = supply function for link a, evaluated at flow level x and

L_1, L_2, L_3, L_4 are analyst-supplied parameters.

Step 1—Initialization

Perform an all-or-nothing assignment to the minimum paths corresponding to zero flows on all links [P^0 based on $C^0 = S_a(0)$]. The result will be F_1. Then, update all link costs to correspond to F_1, yielding C_1. At the same time,

1. Estimate supply function slopes at the current flow levels by performing the following calculation for each link a:

$$m_{1a} = \frac{S_a(1.01\, f_{1a}) - c_{1a}}{0.01\, f_{1a}} \qquad (14)$$

2. Estimate the initial value of the objective function, Z_1.

$$Z_1 = \tfrac{1}{2} \sum_a f_{1a}(c_{1a} + c_{0a}) \qquad (15)$$

Set i = 1 and α, the initial combination size, equal to L_α. Finally, print i and Z_1.

Step 2—Determine Trial Solution Direction

Find new minimum paths, P^{i+1}, based on the link costs C_i.

Step 3—Develop Trial Solution

Assign all travel to the paths P^{i+1}, yielding flows F_t for the trial solution.

Step 4—Obtain New Solution

The parameter λ is determined to (approximately) minimize the (positive) change in the objective function. As derived in the Appendix, the expression for λ (Eq. 25) is

$$\lambda = - \frac{\sum_a c_{ia} \Delta f_a}{\sum_a m_{ia} (\Delta f_a)^2} \qquad (16)$$

where

m_{ia} = slope of the supply function for link a at flow level f_{ia} and
$\Delta f_a = f_{ta} - f_{ia}$.

As discussed in the Appendix, λ must be limited to the range $0 < \lambda < 1$. It is shown in the Appendix that Eq. 16 cannot result in a value of λ less than zero. If $\lambda > 1$, λ should be set equal to 1.

Form a new solution, F_{i+1}, by combining F_i and F_t. For each link, this involves

$$f_{i+1,a} = (1 - \lambda)\, f_{ia} + \lambda f_{ta} \quad (17)$$

Update all link costs to correspond to F_{i+1}, yielding C_{i+1}. At the same time, reestimate supply function slopes between solutions i and i+1 by performing the following calculation for each link, a:

$$m_{i+1,a} = \frac{c_{i+1,a} - c_{ia}}{f_{i+1,a} - f_{ia}} \quad (18)$$

Also, calculate the final estimate of the change in the objective function, ΔZ, and the new value of the function Z_i.

$$\Delta Z = \tfrac{1}{2} \sum_a (f_{i+1,a} - f_{ia})(c_{i+1,a} + c_{ia})$$
$$Z_i = Z_{i-1} + \Delta Z \quad (19)$$

Print i, ΔZ, Z_i, and λ; and set i = i+1.

<u>Step 5—Apply Stopping Rules</u>

The procedure is stopped and the desired assignment outputs are generated if any of the following are true:

1. $-(\Delta Z/Z) \leq L_1$,
2. $-\Delta Z \leq L_2$,
3. $i = L_3$, or
4. CPU minutes $\geq L_4$.

If none of these is true, return to step 2.

CONCLUSIONS

A review of the literature on the network flow equilibrium problem indicates that the problem has a number of interesting properties that are useful in developing solution algorithms. Included are the existence, uniqueness, and stability of a solution and the equivalency of the user-optimized problem and a system-optimized problem. A number of solution algorithms have been developed, and their convergence to a true equilibrium solution can be proved. A number of these algorithms can be made operational by putting together standard components of transportation network analysis systems and simple new evaluation tools. The kinds of computations to be performed by these tools are described in operational terms. The computation costs of these algorithms are expected to be comparable to those of existing restrained capacity assignment procedures.

A basic algorithm is described that involves minimal departures from existing capacity restraint procedures. Efficiency and accuracy results obtained for this algorithm are being used as a base point against which to compare more innovative algorithms.

ACKNOWLEDGMENTS

The work described in this paper was supported by the Software Systems Development Program, Urban Mass Transportation Administration, U.S. Department of Transportation. The interest and support of this agency, and especially its head, Robert B. Dial, are greatly appreciated.

REFERENCES

1. Beckmann, M. J., McGuire, C. B., and Winsten, C. B. Studies in the Economics of Transportation. Yale Univ. Press, 1956.
2. Bruynooghe, M., Gibert, A., and Sakarovitch, M. Une Methode d'Affectation du Trafic. In Fourth Symposium on the Theory of Traffic Flow (Leutzbach, W., and Baron, P., eds.), Strassenbau und Strassenverkehrstechnik, Bonn, 1969.
3. Charnes, A., and Cooper, W. W. Multicopy Traffic Network Models. Proc. Symposium on Traffic Flow (Herman, R., ed.), 1961.
4. Dafermos, S. C. An Extended Traffic Assignment Model With Applications to Two-Way Traffic. Transportation Science, Vol. 5, No. 4, Nov. 1971.
5. Dafermos, S. C. The Traffic Assignment Problem for Multi-Class User Transportation Networks. Transportation Science, Vol. 6, No. 1, Feb. 1972.
6. Dafermos, S. C., and Sparrow, F. T. The Traffic Assignment Problem for a General Network. Jour. of Research, National Bureau of Standards, Vol. 73B, No. 2, 1969.
7. Dial, R. B. A Probabilistic Multipath Traffic Assignment Model Which Obviates Path Enumeration. Transportation Research, Vol. 5, 1971, pp. 83-111.
8. Urban Transportation Planning: General Information. Federal Highway Administration, U.S. Department of Transportation, March 1972.
9. Gibert, A. A Method for the Traffic Assignment Problem When Demand Is Elastic. Transport Network Theory Unit, London Business School, LBS-TNT-85, July 1968.
10. Gibert, A. A Method for the Traffic Assignment Problem. Transport Network Theory Unit, London Business School, LBS-TNT-95, Aug. 1968.
11. Kulash, D. A Transportation Equilibrium Model. Urban Institute, Paper 708-45, Sept. 1971.
12. Manheim, M. L., and Ruiter, E. R. DODOTRANS I: A Decision-Oriented Computer Language for Analysis of Multimode Transportation Systems. Highway Research Record 314, 1970.
13. Jorgensen, N. O. Some Aspects of the Urban Traffic Assignment Problem. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, graduate rept., 1963.
14. Martin, B. V., and Manheim, M. L. A Research Program for Comparison of Traffic Assignment Techniques. Highway Research Record 88, 1965.
15. Martin, B. V., Memmott, F. W., and Bone, A. J. Principles and Techniques of Predicting Future Demand for Urban Area Transportation. M.I.T. Press, Cambridge, M.I.T. Rept. 3, 1961.
16. Mosher, W. W. A Capacity Restraint Algorithm for Assigning Flow to a Transport Network. Highway Research Record 6, 1963.
17. Murchland, J. D. Road Network Traffic Distribution in Equilibrium. Paper for the Tagung ueber "Mathematische Methoden in den Wirtschaftswissenschaften," Mathematisches Forschungsinstitut, Oberwolfach, October, 1969.
18. Netter, M. Equilibrium and Marginal Cost Pricing on a Road Network With Several Flow Types. In Traffic Flow and Transportation, American Elsevier, New York, 1972.
19. Netter, M., and Sender, J. G. Equilibre Offre-Demande et Tarification sur un Reseau de Transport. Institut de Recherche des Transports, Arcueil, France.
20. Pigou, A. C. The Economics of Welfare, First Ed. Macmillan, London, 1920.
21. Prager, W. Problems of Traffic and Transportation. Proc. Symposium on Operations Research in Business and Industry, Midwest Research Institute, Kansas City, Mo., 1954, pp. 105-113.
22. Ruiter, E. R. ICES TRANSET I: Engineering Users Manual. Civil Engineering Systems Laboratory, M.I.T., Res. Rept. R68-10, March 1968.
23. Ruiter, E. R. The Prediction of Network Equilibrium—The State of the Art. Internat. Conf. on Transportation, Bruges, Belgium, June 1973.
24. Tomlin, J. A. A Mathematical Programming Model for the Combined Distribution and Assignment of Traffic. Transportation Science, Vol. 5, No. 2, May 1971.
25. UMTA Transportation Planning System: Reference Manual. U.S. Department of Transportation, Sept. 1972.

26. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. Proc. Inst. of Civil Engineers, Pt. 2, Vol. 1, 1952, pp. 325-378.
27. Wigan, W. R. Benefit Assessment for Network Traffic Models and Application to Road Pricing. Road Research Laboratory, Crowthorne, Berkshire, England, Rept. LR 417, 1971.
28. Wigan, W. R., and Bamford, T. J. G. A Perturbation Model for Congested and Overloaded Transportation Networks. Road Research Laboratory, Crowthorne, Berkshire, England, Rept. LR 411, 1971.
29. Wilkie, D. F., and Stefanek, R. G. Precise Determination of Equilibrium in Travel Forecasting Problems Using Numerical Optimization Techniques. Highway Research Record 369, 1971.
30. Yang, T. C., and Snell, R. R. Traffic Assignment by the Maximization Principle. Jour. Highway Div., Proc. ASCE, Vol. 92, No. HW2, 1966.

APPENDIX

DERIVATION OF THE COMBINATION PARAMETER λ

The combination parameter λ is to be determined to approximately minimize the positive change in the objective function, Z (Eq. 11). This change, ΔZ, is made up of components for each link a, such as the shaded area shown in Figure 1. When Δf_a is positive, as shown in the figure, the contribution to $\Delta Z (\Delta Z_a)$ is positive. Similarly, when Δf_a is negative, ΔZ_a is negative.

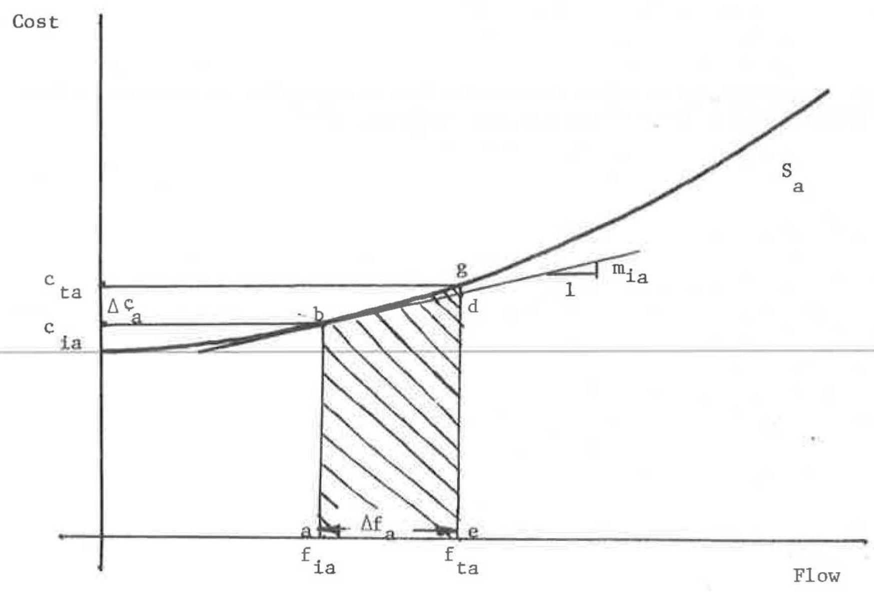
To avoid the necessity of determining c_{ta}, we approximate point g by point d, which can be determined from the following relationship:

$$\Delta c_a = m_{ia} \Delta f_a \tag{20}$$

Then, using the area abde as an approximation for ΔZ_a, we obtain

$$\Delta Z_a = \tfrac{1}{2} \Delta f_a (2\, c_{ia} + m_{ia} \Delta f_a) \tag{21}$$

Figure 1. Relationships used to calculate ΔZ_a for a typical link.



This quantity can be summed over all links to obtain the total change.

$$\Delta Z = \sum_a \tfrac{1}{2}\, \Delta f_a (2\, c_{1a} + m_{1a} \Delta f_a) \qquad (22)$$

To find λ requires that ΔZ be defined as a function of λ. This can be done by replacing Δf_a in Eq. 22 with $\lambda \Delta f_a$, resulting in

$$\Delta Z(\lambda) = \sum_a \tfrac{1}{2}\, \lambda \Delta f_a (2\, c_{1a} + \lambda m_{1a} \Delta f_a) \qquad (23)$$

The valid range for λ is $0 \le \lambda \le 1$.

The optimum value for λ can be found by differentiating ΔZ(λ) with respect to λ and setting the derivative equal to zero.

$$0 = \frac{\delta \Delta Z(\lambda)}{\delta \lambda} = \sum_a c_{1a} \Delta f_a + \lambda \sum_a m_{1a} (\Delta f_a)^2 \qquad (24)$$

Solving for λ gives the following expression:

$$\lambda = - \frac{\sum_a c_{1a} \Delta f_a}{\sum_a m_{1a} (\Delta f_a)^2} \qquad (25)$$

Note that the denominator is always positive. If the numerator is not negative, λ will be negative. This will only occur if the trial solution, evaluated at the former costs C_{1a}, is not so good as solution i. This cannot occur because flows are being shifted from higher to lower cost paths—at the current costs—in steps 2 and 3.

If the value of λ from Eq. 25 is greater than 1, then λ should be set equal to 1. This implies that all of the former solution is being replaced by all of the trial solution t.

PLANNING URBAN TRANSPORTATION SYSTEMS FOR PRODUCTIVITY, EFFICIENCY, AND QUALITY OF SERVICES

Anthony R. Tomazinis, Transportation Studies Center, University of Pennsylvania

This paper reports on current work and analysis of the problem carried on at the Transportation Studies Center of the University of Pennsylvania. Although the work has not yet been completed, the work undertaken enables the author to suggest that a new approach in urban transportation planning and a new type of urban transportation plan, based on studies of efficiency, productivity, and quality, may prove to be what the field needs for the 1970s. The urban transportation system is segmented into the network, the primary services offered, and auxiliary services.

NUMEROUS STUDIES in productivity and efficiency have been conducted for most sections of the economy. For the last 15 years even studies on productivity of federal services have been repeatedly undertaken and since 1970 the concern for efficiency and productivity of local government functions has grown. The studies by Kendrick (1) and Fuchs (2, 3) suggest the importance attached to productivity in the private and governmental sectors. Also the recent studies of the Urban Institute (4, 5) provide an indication of the significance attached to productivity of local government services.

On the other hand, urban transportation planning has been going on in most metropolitan regions of the country in an intensive manner since the early 1950s. In many cases the transportation planning effort resulted in the publication of impressive reports and study documents that purported to present evidence for "optimized" regional transportation plans (6, 7). Curiously enough all this effort was taking place while no overt attention was being paid to issues and problems of productivity and efficiency of the proposed systems.

Evidence clearly suggests that the primary concern in the major studies of the last 2 decades followed a long-established trend of expanding major facilities to new areas of development and of proposing new major facilities, usually highways, within the already developed part of the region. Usually, the recommendations were formed within a framework of user cost minimization as measured on a systemwide basis. Travel cost savings were then pitted against systemwide capital investment by using some of the most simplistic economic techniques, e.g., simple benefit-cost ratio, a least total cost measure, or an incremental rate of return determination. With regard to quality of service and the quality of the systems themselves, practically all major studies were concerned with only one index, that of average speed on a daily or rush-hour basis.

At the end of this prolific era of urban transportation planning, the realization has slowly emerged that the permanent accomplishments of this period have been limited indeed. The ephemeral enthusiasm of the mid-1960s gave way to the prevailing concern about the significance of what was produced at the height of the effort. Two major factors emerged in the ensuing years.

1. A vast number of individuals within urban regions discovered that the proposed transportation plans and programs included little of the quality of service of which the people were in need. In most cases plans and programs tended to ignore possible harm that the plans would produce for many individuals and whole communities and to

Publication of this paper sponsored by Committee on Transportation Systems Design.

emphasize benefits (or quality characteristics) of services that had little or no appeal to those concerned.

2. There was greater appreciation of the significance of high productivity and efficiency measurements in whatever is being done by private or public funds. A planner can no longer, with impunity, avoid issues of productivity because the investment is made through public funds. Nor can he act wisely by ignoring the efficiency rates of each major component of a complex system simply by proclaiming that the regional, total, plan is acceptable.

The three items, productivity, efficiency, and quality of services, emerge as the focus of planning activity for metropolitan transportation systems for the coming years. Their importance and centrality are indeed apparent in the midst of increasing general concern for the unit cost of all types of services produced by governmental or private organizations and for the quality of services the consumer receives. Their inseparability is also rather obvious. In fact only a transportation system that achieves a measure of all three objectives can be considered a distinct improvement over what we have and over what we have been planning.

DISSECTING THE TRANSPORTATION SYSTEM FOR PRODUCTIVITY AND EFFICIENCY STUDIES

The urban transportation system (UTS) is a complex entity. Early studies on productivity, efficiency, and quality of such a system indicate that it is imperative to conceive the system in its totality and then to dissect the overall system appropriately.

The urban transportation system is composed of, and then divided into, the network, the services, and the auxiliaries. This division differs considerably from previous breakdowns of UTS into distinctive submodes, nodes and links, or lines and terminals. The proposed new division directs analytical efforts along new lines of thought and investigation.

The Network

The analyst of the network of a transportation system emphasizing productivity, efficiency, and quality of the system has to investigate three aspects of the system:

1. The geographic location, distribution, and linkages of the various nodes and links of the network with regard to every combination of origin and destination points;
2. The magnitude, sequence, and consistency of the various attributes of each node and link (such as capacity and safety); and
3. The interrelationships between network characteristics and the surrounding elements of the other urban systems (land use, utilities, facilities).

Several studies, of course, examined networks from other points of view. Geographers' studies are well known (8, 9) as are the ones on electric network theory (10, 11) and graph theory (12, 13). Nonetheless, seemingly there has been very little thinking concerning network analysis with productivity, efficiency, and quality issues in central focus. For instance, it is obvious that a network design with 200 miles of links and 50 nodes of which only one is a central node permitting complete transfer from one section of the network to another would facilitate movement and interchanges in a much more limited manner than design with the same 200 miles of links and 50 nodes of which more than a half dozen are multiple-transfer nodes facilitating transfers from one section to another. Whatever services can be provided on such a network, the basic efficacy of transfers built in the network design will affect all measurements of productivity, efficiency, and quality.

The Services

The analyst of services that an urban transportation system offers over its network with emphasis on productivity, efficiency, and quality of the overall system has to separate the system into several components, preferably by mode, but also by link and node. The cardinal rule seems to be the closest possible matching of the demand for services to the supply. Some parts of an urban transportation system have special flexibility in providing services, whereas others have a fixed provision (capacity) regardless of demand variation. For instance, a transit system can contract or expand the provision of services by varying the frequency of vehicle departures, by increasing the size of the trains, by increasing the legal number of standees on each vehicle, or by any similar combination of actions. A highway system has much more rigid characteristics, although on several occasions flexibility can be achieved by limiting curb parking, reversing a central lane, reserving a lane for special vehicles, altering signalization, opening a bypass, or, more recently, by electronically controlling the inflow of vehicles into the traffic stream. In all cases the essential objective is to match the supply of services to the demand for services.

Several complications appear from the outset in these efforts. First, services may be provided automatically or by simple regulation as in the highway system, or by special provision of facilities and crews as in mass transit. Second, services may not be provided in a manner consistent with the demand and supply. A good example follows. The demand profile may have an extraordinary peak followed by a low point in terms of both location and time. The profile of the supply, therefore, would have to provide (if it is to be matched well) for such an extraordinary variation of peaks and valleys. On the other hand, the supply mechanisms have limitations; e.g., a train cannot add cars beyond the length of the station platforms, nor can it drop cars between major stations. Also, highways cannot, as a rule, reverse or reserve more than one lane.

The matching of the profiles of demand and supply is one component of efficiency analyses; another deals with the flexibility and feasibility of adding, reducing, and shifting services within and among the various parts of the network. A third component deals with the matching of network characteristics and service requirements or objectives. This constitutes a bridge between requirements and capabilities of the network and the services taken together.

The Auxiliaries

In many cases the productivity, efficiency, and quality of the network and the services depend on the performance of the auxiliary services. Recent statistics indicate that there is a vast difference among cities in the average number of hours that buses stay in the shop for repairs and in the number of buses that are available for assignment from system to system. For example, in one city it was found that as many as one-third of the buses were inoperable on a random day as opposed to only 5 percent of the buses in other systems. Similarly, many highway sections and intersections can be kept out of use for repairs and modification throughout the year, which reduces the efficiency of all neighboring facilities.

The condition of the auxiliary services of UTS can be evaluated from several perspectives:

1. The absolute size of the auxiliary services,
2. The relative size of the auxiliary services with regard to the size of the network and the primary services of the system, and
3. The composition of the auxiliaries.

What are, for instance, the clearly supportive services, and what are the extra services that the system provides? Also, what are the services that provide for past obligations, present needs, and future plans and expectations? A fourth perspective is the impact of the auxiliary services on the network and the primary services of the entire system or any part of it.

Not all parts of the UTS have auxiliaries of equal significance. For instance, an urban transit system has more auxiliaries than a highway system. Whether it needs all of these auxiliaries is, of course, a question that should be answered by an efficiency analysis. Also, there is increased emphasis on flexibility and management of highway systems and the associated growth in importance and size of auxiliaries in highway systems. The trend started with the provision and management of service (reversal of lanes, curb parking, reserved lanes) and is now characterized by the introduction of complete systems for urban freeway surveillance and control. In all cases the investigation should leave the question of the function and utility of the auxiliaries open inasmuch as services and subsystems can be found that are pro, con, or completely neutral to the objectives of better efficiency, productivity, and quality.

SYSTEM ASSESSMENT AND SYSTEM PARTS

A second major understanding seems necessary. The concepts of productivity, efficiency, and quality of service should be examined together. On the other hand, the UTS should also be divided into three major parts. The two sets form a symmetric matrix with nine cells (Fig. 1).

Based on the matrix shown in Figure 1, different levels of association among the three concepts and the three system parts can be discussed.

Productivity

The concept of productivity is primarily concerned with total system inputs and outputs; therefore, the unit measurement of this concept (number of units of output per unit of input, such as thousands of travelers per man-hour, per dollar, per mile, per bus) must be expressive of the total output of the system. No partial productivity measure really makes sense. Of course, if an urban transportation system consists of a single part, (e.g., network), the measure of productivity of the system is also the productivity of the network. In all other cases the mutual dependence of the parts of the systems precludes any meaningful measure of productivity by part or by subsystem.

Efficiency

As Figure 1 shows the situation with regard to efficiency analyses is quite different. In this case it is rather meaningless to discuss the efficiency of a complex, multiple system that carries components with various oscillating rates of efficiency. The concept of efficiency deals with the rate of success of a specific process in recovering expended resources. In this respect studies in efficiency would need to dissect the system into the largest possible number of distinct, complete subprocesses and measure the efficiency of each in detail. In this division of the total system efficiency studies should focus first on the network, then on the primary services, and finally on the auxiliaries. Further, each of these parts should be divided into submode aggregates, such as efficiency of the highway network, transit network, and railroad network and the efficiency of transit service, highway service, special terminals, bridge crossings, and the like. Third, the efficiency with which auxiliary services and subsystems are made available and serve the primary services, the network, and, in general, the system itself, should be examined.

Efficiency studies usually need to be detailed if they are to be used in planning and managing a UTS. In fact, the efficiency of some key components of a subsystem is of central concern in more cases than the overall inefficiency of services and networks. Inefficiencies of the latter type are usually obvious and soon become the topic of newspaper editorials and the subject of political controversy. As a result they are subject to elimination soon after they have been discovered and discussed. Inefficiencies of the first type, however, although numerous and frequent in many a system, are difficult for the public to locate, magnify, and subsequently force out of the system. Usually they take a technical form, a residual of technology application and an unavoidable character that defies gross actions and generalized solutions. To eliminate this type of inefficiency requires technical studies. Systematic and detailed analysis of each system component and of each factor and relationship that affects system performance needs to be in central focus. This approach is advisable for efficiency studies on all three system components.

Efficiency studies must be made not only on each unit of the system (i.e., a major link, a major node, a major transit line) but also on a complete process. The first type concerns the producer of services; the second concerns the user of services. Efficiency of the operation of the unit of the system is directly related to the productivity of the unit and the system as a whole. Such an efficiency measure says very little, however, about the case with which a particular trip is made. This is of direct concern to the consumer (the user) of the system. Therefore, it is important that studies on efficiency include measures of the efficiency with which complete (from the origin to the destination) representative trips are being made. On aggregation, efficiency of whole corridor movements should be studied. This is where major deficiencies may produce total elimination of a trip or, in the long run, a substantial change in the travel patterns of the region.

Efficiency studies can take several forms. Although systems analysis approaches may prevail, in many cases an efficiency study would clearly be drawing much from traffic engineering, in other cases from management sciences, and in many cases from straight economic theory, especially from the theory of the firm, and the consumer's behavior theory. The main issues are efficiency of production, consumption, and distribution. The analyst can be surprised when he realizes that he moves rapidly from one field to another as he traces the efficiency of the various components. TOPICS, for instance, was nothing more than a crude attempt to study efficiency problems in the highway network. Similar programs with approximate crudeness are currently in effect in the transit field: airports, harbors, and turnpike and bridge authorities.

Finally, efficiency studies do not have to be limited to existing systems. They can be of great use in planning new systems, and they should guide the planner in assessing the technical proficiency, in succession, of the proposed networks, the new service patterns, and the new combinations of auxiliaries.

Quality

The significance of attaching quality studies of the transportation system to any set of efficiency and productivity studies becomes apparent if one considers the rather obvious trade-offs between efficiency and quality of service. It is in fact the presence and feasibility of these trade-offs that foster one of the major controversies in the field of urban transportation. For the supplier of the system, the quality of the system is measured along the dimensions of efficiency and productivity. The more efficient and productive a system is, the better this system is considered by the supplier. Although aspects of efficiency also have appeal to the consumer, his concerns far exceed those of efficiency. For instance, in a recent study of the significance that consumers place on transportation system quality (14), it became rather clear that the service quality items rather than the efficiency items received top rating. Among 32 quality attributes that were included in this study, the entire population and three major subgroups (under 20 and single, elderly, and low income) chose items such as arriving when planned, having a seat, no transfers, less waiting time, shelters at pick-up points, and longer service hours instead of the traditional emphasis on items such as faster trips and more direct routes.

Unless the systems produce service that can be consumed it makes no difference how productive and efficient the system is. The most efficient service is the service that not only has the best matching of its supply profile over time and over space with the profile of demand but also meets the quality characteristics that the consumers impose. Otherwise, the consumers would not use the services of the system. Thus, the paradox can be seen of a system most efficient from the suppliers' point of view which is both going bankrupt and also castigated as completely inefficient from the user's point of view.

The quality of service of a transportation system can be thought of as a matrix. Quality can then be divided into two groups, those associated primarily with each trip (immediate factors) and those associated primarily with long-range considerations of trip patterns. The first group includes convenience, comfort, frequency, and familiarity with the system.

The second group of quality attributes is more pervasive in nature. These attributes are reliability of current and long-range system performance, availability of service for any purpose at any time, security provided by the system, travel cost and travel speed, and level of privacy and individualism in services that the system offers.

Quality of service analyses must be related to the efficiency of each component of the system and, if possible, to the productivity of the entire system.

igure 1. Matrix of associations.

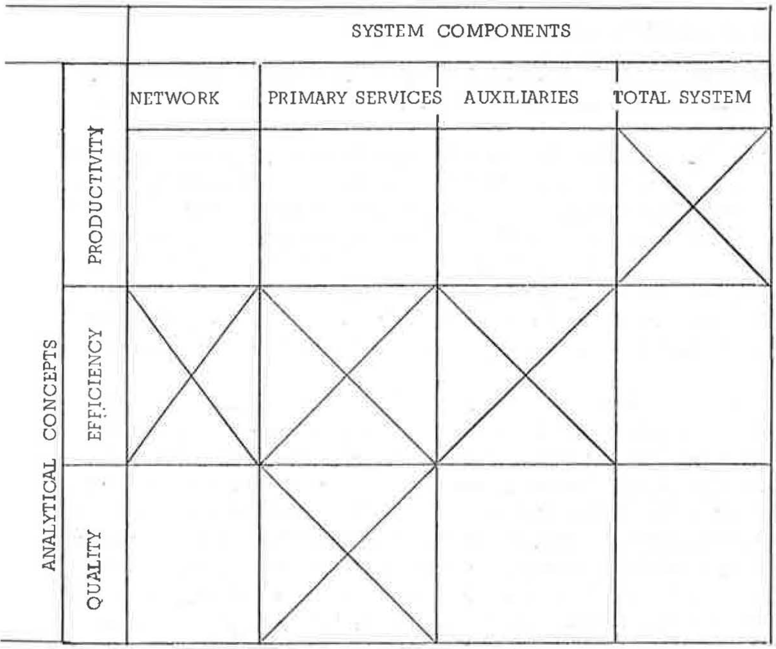
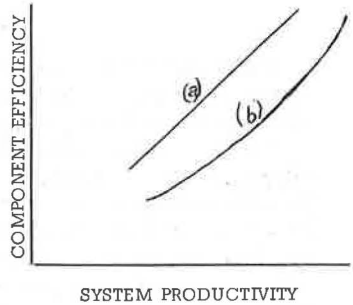


Figure 2. Efficiency improvements versus productivity improvements.



igure 3. Efficiency improvements versus uality of service improvements.

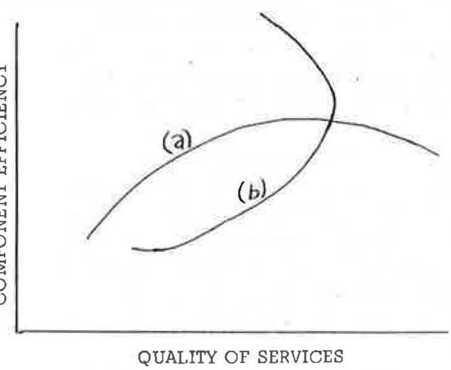
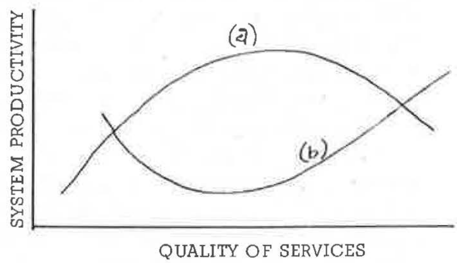
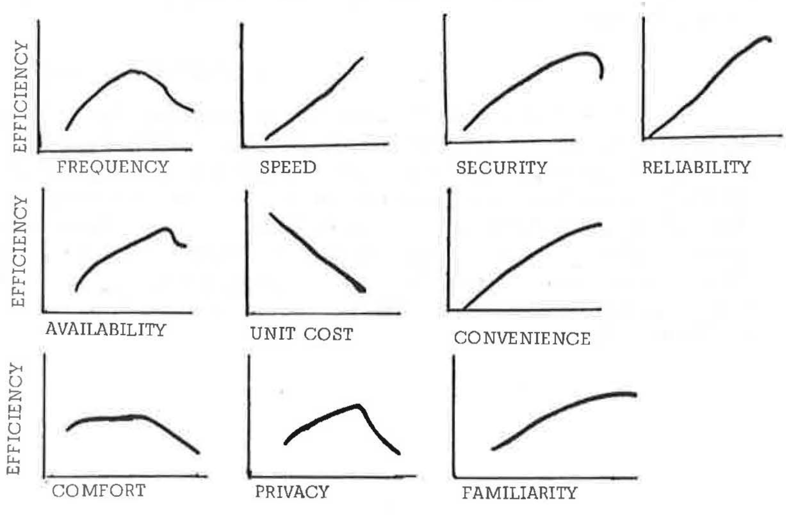


Figure 4. Productivity improvements versus quality of service improvements.



igure 5. Efficiency versus quality of ttributes.



RELATIONSHIPS AND TRADE-OFFS

The relationships and trade-offs among the three concepts (efficiency, productivity, and quality) for urban transportation systems loom from the outset as potentially complex and on occasion undefinable.

Figure 2 shows the two general forms that one would expect the relationship between efficiency and productivity to take. Clearly, as efficiency of the various system components increases so does the overall productivity of the entire system. This relationship may have one-to-one correspondence (curve a), or it can have a correspondence smaller or greater than one, depending on the specifics of the application (curve b). Based on the division of the UTS into the network, services, and auxiliaries and the variability of the circumstances prevailing in each system, the variable correspondence between partial efficiency improvements and total system productivity improvements is plausible.

Figure 3 shows the whole variation that the relationship between efficiency improvements and quality improvements can take. Normally initial improvements in efficiency measures are expected to correspond to improvements in quality measures and vice versa. However, after a particular point, improvements in one set of measures may correspond to deterioration of the other set of measures. Both curves of Figure 3 indicate this reversal of the correspondence between quality and efficiency improvements.

Figure 4 shows two other forms of the potential relationship between system productivity and quality of service. Curve a suggests an increase of productivity as quality of services improves to a certain point, beyond which the reverse takes place. This relationship can be seen within the context of consumer reaction to available services. As the quality of services improves, the consumer makes greater use of the system and, thus, more usable service is "bought" by the public. Beyond a certain point consumer response may not be so extensive as continual improvements in quality may be, and, therefore, the overall productivity of the system may decline (with respect to either labor or capital). Curve b represents the reverse sequence of events, and its plausibility can easily be constructed for each stage. Obviously, the analyst would have to carefully establish the exact point and type of relationship between productivity and quality and, further, explore the change that may occur in productivity by any measure involving change in the quality of the services offered.

In exploring in detail the potential relationships and trade-offs between quality of service and efficiency of operations, the analyst may have to investigate these relationships as they emerge with each of the 10 factors of quality that were discussed earlier. Figure 5 shows a plausible form of the trade-offs between efficiencies and each quality attribute. As can be seen, the relationship depends on both the level of efficiency already achieved and the nature of the quality attribute. In most cases efficiency (or productivity) would cease beyond a certain level regardless of quality improvements. In other cases, efficiencies will clearly decrease for any increase of quality of operations. Again the analyst would have to focus on the particular quality attribute that is explored and its specific impact on the operations of the specific system component that is going to be affected.

THE NEED TO SEPARATE MODES

These concerns notwithstanding, it seems imperative that an analytical effort on productivity, efficiency, and quality of urban transportation systems not be bogged down by conceptual generalities. The UTS is made up of three essential operational parts, the highway subsystem, the mass transit subsystem, and major multimode system terminals. The operational and technological differences among these parts are profound and frequently unbridgeable. Hence the analyst should recognize these differences and try to capitalize on them, rather than ignore them and presume an ability to establish concepts, methods, and units of measurement that are equally and universally usable for all three subsystems.

CONCLUSIONS

The approach of urban transportation planning that was developed in the 1950s and 1960s appears to be inapplicable for the 1970s. That approach produced a set of monumental plans and vastly expanded networks of highway facilities, with emphasis on accommodation of new highway trips. This produced widespread opposition to these plans and a deep concern about the normative values and optimal nature of the plans themselves. Currently most urban transportation planning teams are trying to rescue whatever parts of the regional plans seem feasible. Clearly, a change in approach and an essentially different type of urban transportation plan are imperative if transportation planners are to be effective in their efforts to improve travel conditions within urban areas.

What is proposed is a set of analytical studies of the entire transportation system of each urban region with emphasis on productivity, efficiency, and quality. Only at the conclusion of such studies, and in direct response to the needs to improve productivity, efficiency, and quality, would new facilities be suggested. Meanwhile, the system in its totality, as well as each subsystem, would be analyzed by focusing on improving its total system productivity and component efficiency. Such improvements would have to be introduced in any one or all of the major parts of a UTS: its network, its primary services, and its auxiliaries (services and subsystems).

As of now, no study is known to have been designed or undertaken following this new approach. The concern for efficiency, productivity, and quality has emerged in most studies indirectly. The work carried on currently at the University of Pennsylvania on which this paper is based is the only one known to this author. It is hoped that this initiative will soon be followed by others.

REFERENCES

1. Kendrick, J. W. Productivity Trends in the United States. Princeton Univ. Press, 1961.
2. Fuchs, V. R. Measuring Productivity of Federal Government Organizations. Bureau of the Budget, 1964.
3. Fuchs, V. R. The Service Economy. National Bureau of Economic Research, 1968.
4. Hatry, H. P., and Fisk, D. M. Improving Productivity and Productivity Measurement in Local Governments. Urban Institute, 1971.
5. Winnie, R. E., and Hatry, H. P. Measuring the Effectiveness of Local Government Services: Local Transportation. Urban Institute, 1972.
6. Chicago Area Transportation Study, Final Report. Vol. 1, 1959; Vol. 2, 1960; Vol., 3, 1962.
7. Penn Jersey Transportation Study. Vol. 1, 1962; Vol. 3, 1965; Vol. 5, 1969.
8. Kansky, K. J. Structure of Transportation Networks. Univ. of Chicago, 1963.
9. Garrison, W. L., and Marble, D. P. The Structure of Transportation Networks. Transportation Center, Northwestern Univ., 1961.
10. Caurer, W. Synthesis of Linear Communication Networks. McGraw-Hill, 1958.
11. Truxal, J. G. Automatic Feedback Control System Synthesis. McGraw-Hill, 1955.
12. Berge, C. Théorie des Graphes et ses Applications. Dunod Publications, Paris, 1958.
13. Wagner, H. M. Principles of Operations Research. Prentice-Hall, 1969.
14. Gustafson, R. L., Curd, H. N., and Golob, T. F. User Preferences for a Demand-Responsive Transportation System: A Case Study Report. Highway Research Record 367, 1971.

INTERACTIVE GRAPHICS AS A TOOL IN PLAN EVALUATION

Yehuda Gur, Creighton, Hamburg Inc., Chicago

INTRANS is a man-computer interactive graphics system designed for real-time analysis of transportation and urban data. During mid-1973, INTRANS was used by the Chicago Area Transportation Study as an aid in the evaluation of alternative plans for the 1955 major review. This paper describes the project. INTRANS has been proved effective as a tool for choice and design of figures, as a convenient medium for retrieving planning data, and in analysis of specific problems. A number of problems and difficulties, some of them marginal, have prevented more extensive and effective use of the system. It is concluded that interactive graphics systems in general, and INTRANS in particular, are cost-effective tools for transportation planners. Increases in their effectiveness and capabilities will come mainly through more extensive use by practicing professionals.

•INTRANS is a man-computer interactive graphics system designed for application in transportation and urban planning. INTRANS-BROWSE is a subset of INTRANS; it enables real-time graphical analysis of existing data sets (1). In mid-1973, the Chicago Area Transportation Study (CATS) performed a major review of alternative transportation plans for 1995 (2). INTRANS was used to support the planning and review process. The purpose of this report is to present the experience gained in this project.

INTRANS-BROWSE is an interactive graphics system that enables on-line graphical analysis of existing data sets. The system includes the following features:

1. Data management system—Data sets are structured in three levels: (a) files, one file for each alternative; (b) variables, one variable gives the number of transit trip ends in each zone; and (c) elements, the ith element in the variable is the number of trips from zone i. Variable size is limited to 2,500 elements, and there is no practical limit to the number of files or variables. In the CATS analysis there were seven files, each with approximately 100 variables of 1,700 elements.
2. Display routines—INTRANS enables the display of intensity maps, frequency distributions, and functional relationships of the variables.
3. Computation routines—These routines enable arithmetical manipulations of variables by using FORTRAN types of statements. Variables of different files can be used in the following routines: basic statistic routines, utility routines, such as Data Editor, and peripheral programs for data loading and maintenance.

INTRANS operates on an IBM 360/370 with full OS. It can operate by residing in dedicated core; however, the standard operating mode is under time sharing. The graphic terminal is a storage display tube, such as the Tektronix 4010. The terminal can be either hard-wired to the computer, or connected through a telephone line.

A device that prepares direct hard-copy image of the screen can be installed. Most of the figures in this report were prepared by this device.

The system is extremely inexpensive for both hardware and operation.

USE WITHIN THE MAJOR REVIEW

INTRANS was used in the major review as a tool for data retrieval, manipulation, and display. Various UTP models were run in the conventional batch processing mode;

Publication of this paper sponsored by Task Force on Interactive Graphics.

relevant input and output data from the runs were transferred into INTRANS files for interactive analysis. Within this scope, the major contribution of INTRANS was its aid in highlighting the differences among the alternatives, in terms of demand and performance characteristics. The amount of information accessible through use of INTRANS was much greater than that previously available (with reasonable effort and time). No less important, the INTRANS data files have become a permanent source for timely and detailed information regarding the Chicago region and the alternatives. Some examples of the use of INTRANS in this context are given in the paper.

Mainly because of administrative problems, INTRANS was available for use only in the evaluation stage of the planning process. Obvious potential applications in input preparation, model calibration, and plan formulation were not tried.

The maj : conclusion of this project is that interactive graphics systems in general, and INTRANS in particular, are cost-effective, operational tools for transportation planning. However, much work and experience are needed to use these systems to their full potential.

APPLICATIONS

Because of the lack of previous experience in the use of INTRANS within an operational planning process, this project was considered semi-experimental. On one hand, production use of INTRANS was made in rather straightforward and simple applications in which the final product was well-defined and directly usable within the planning process. At the same time, more complicated applications with potential benefits were tried, mainly to gain experience for future projects.

Four features of INTRANS have been of prime utility in the evaluation. First is the ability to produce intensity maps. By observing a map describing a variable, the analyst gets much more information in much less time than by scanning through a listing. For example, Figure 1 shows a map of estimated volume-capacity rates for the region. By examining the map, the analyst immediately sees a picture of location and extent of expected congestion. A second important feature of INTRANS is the computation capabilities. The ability to relate any two variables from the same alternative, or from different alternatives, is extremely useful. Figure 2 shows a map of the (interactively calculated) differences between estimated volumes and capacities. The analyst receives, through such a map, a comprehensive picture of highway deficiencies. As another example, Figure 3 shows differences between two alternatives in the transit travel time to a center of activity. Extent and locations of improvements are immediately apparent.

Another important feature of INTRANS is the organization of data and ease of access to them. In a period of intensive analysis, it is much cheaper and faster to access a specific piece of information through INTRANS than to look it up in a prepared listing. This feature is especially important when, during the analyis, it is necessary to examine carefully a specific issue. Without the use of INTRANS, data preparation for such a problem requires more time and is more expensive. A fourth useful feature of INTRANS is flexibility in designing graphical displays and the ease in their preparation. This feature has been instrumental in preparation of various reports.

It is significant to note that the hard-copy device has been proved essential. As a matter of fact, most of the hard analysis has been done on the hard copies; the interactive sessions have been used to try alternative displays and produce hard copies of those required for the later analysis. Even when interactive analysis has been tried, the hard copies efficiently replace the use of multiple screens.

Data for the Analysis

Most of the data used in the analysis were zonal data. They covered practically all phases of the modeling process. They include land use information, trip ends by purpose, average trip lengths, vectors of trip distributions and skim trees for selected zones, amount of travel generated by and going through each zone, network capacities (by mode) for each zone, mode and submode split information, and more. The information was transferred from output tapes of the model into the INTRANS files by a set

Figure 1. Estimated volume-capacity ratio displayed against expressway skeleton.

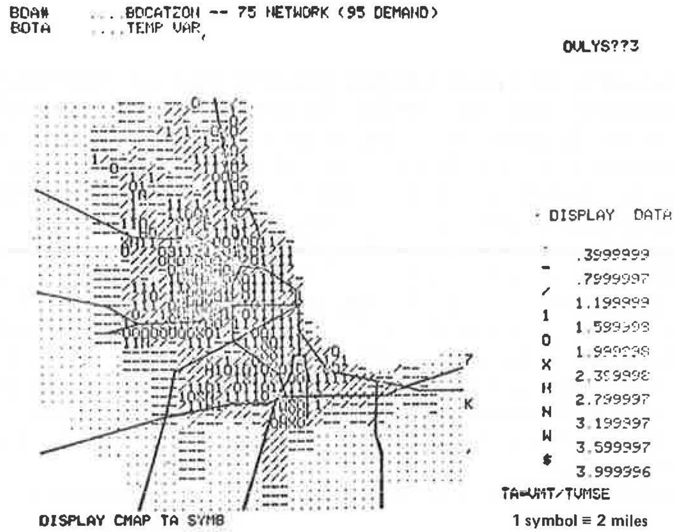


Figure 2. Highway deficiencies displayed against expressway network and major towns.

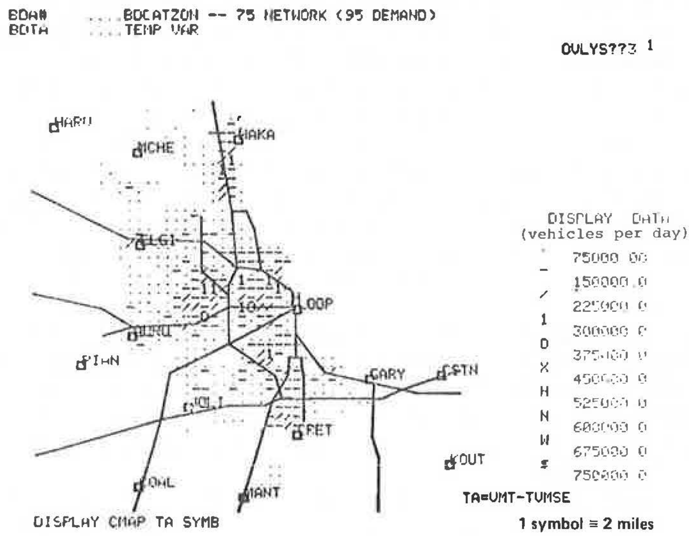
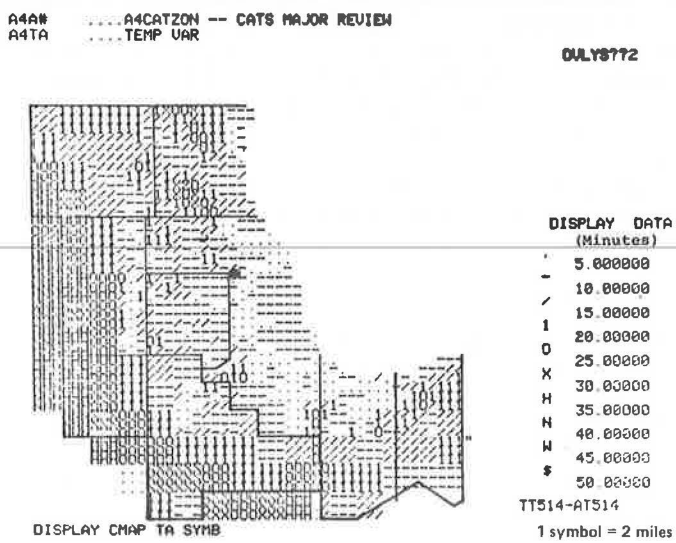


Figure 3. Differences in transit travel time to O'Hare between existing network and alternative D.



of peripheral programs that ensured smooth and fast transfer. It was possible to complete the data transfer process for any model run within half a day after the output tapes were available.

A special set of programs was written to retrieve data from the Department of Housing and Urban Development UTP transit analysis package. One set of programs was used to retrieve skim tree information such as travel times, access times, and priority modes for selected zones. Another set summarized zonal transit volumes (person-miles) and amount of service (seat-miles) by submode.

The INTRANS data files were stored on line for the period of intensive analysis and then deleted. Back-up data files were stored on off-line disks and tapes for future reference.

Types of Applications

Within the CATS major review process, INTRANS was used to prepare a plan reference report, design figures for the final report, and analyze special problems.

Plan Reference Report—A major product of the project was a plan reference report (3). The report included four types of information:

1. Variables describing the region in 1995 (e.g., trip generation rates),
2. Detailed description of demand and network performance for the do-nothing alternative,
3. Key variables describing the performance of each of the seven alternative plans, and
4. Displays of performance differences between each alternative and the do-nothing alternative.

The report described in detail the source and meaning of the variables being displayed, but included almost no analysis of specific displays. It included approximately 200 displays.

The plan reference report has been used as an easily accessible, permanent source of detailed information on the 1995 plan, mainly for internal use. It has also been used as a "shopping list" that summarizes for CATS planners the material and capabilities available to them through INTRANS. Consulting the material in the plan reference report has produced more effective special-purpose interactive sessions. The report might also be considered an alternative to direct use of INTRANS for creating displays that are likely to be needed repeatedly.

Design of Figures and Maps—Once decision to include a map in the report was reached, INTRANS was used for examination and design of the map. First, the maps were reviewed on the screen (or in the reference report) to decide whether they were effective. Second, various alternatives for scaling, group ranges, and so forth were tried to have the most effective design. Once the design was decided on, maps were produced in conventional methods. Use of INTRANS for this function saved a significant amount of manual work. The natural extension, i.e., direct use of INTRANS displays in the final report, was rejected because of the low graphic quality. The more important deficiencies were the coarse grain of the displays and small original size, which prevented reduction for reproduction, and lack of multiple-color capabilities. It should be noted, however, that INTRANS displays have been used extensively in technical reports and memos.

Analysis of Special Problems—During and after the evaluation process, specific questions were raised, in which the information required to answer them has not been readily available. One group of questions is related to subareas, or specific locations: What are the levels of transit service to O'Hare Airport in the various alternatives (Figs. 3, 4, 5)? Another question relates to specific projects or issues: What is the effect of changes in fare policies (alternative F) on rapid transit ridership (Fig. 6)?

The third question is more general: How do the relative levels of service of transit versus highway change among the alternatives (Fig. 5)? INTRANS has proved effective in supporting the analysis of such questions.

Previously, a standard procedure has been followed in the preparation and conduct of such analyses. A list of variables and relations that seem to be relevant to the

Figure 4. Transit travel time to O'Hare on existing network.

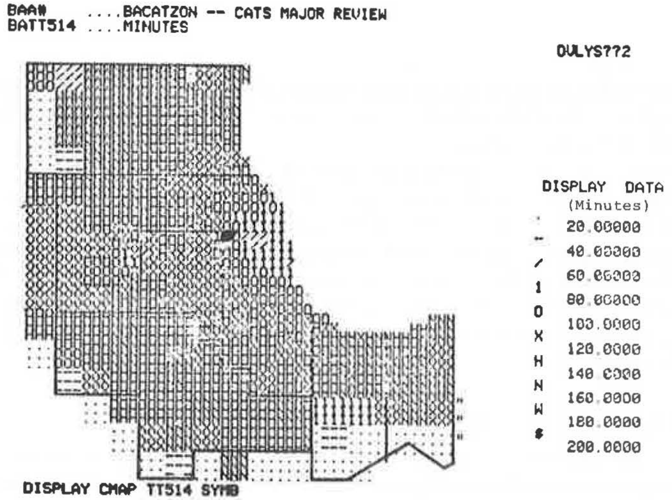


Figure 5. Differences in travel time to O'Hare between transit and car.

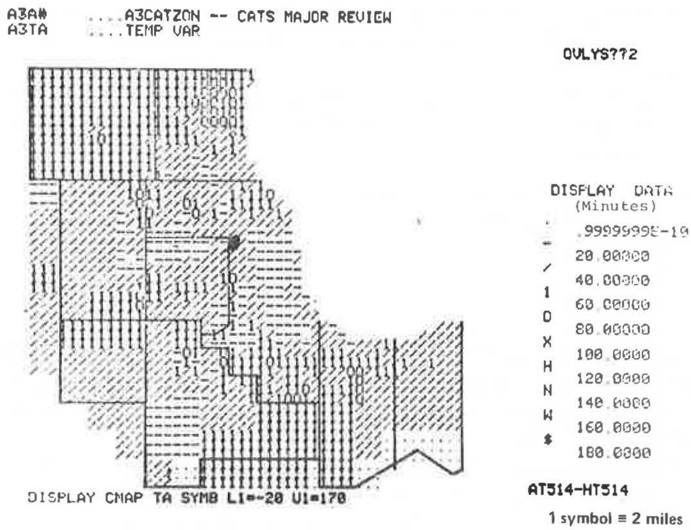
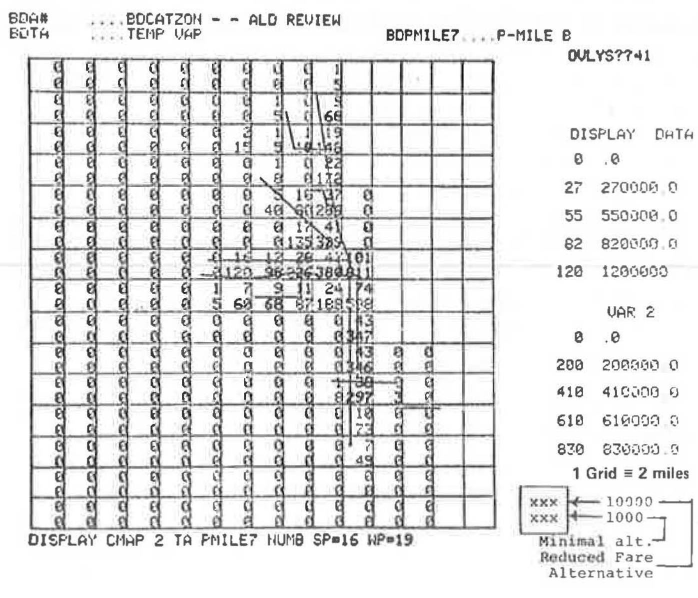


Figure 6. Total person-miles on rapid transit system displayed against transit network.



problem is prepared, together with a set of hypotheses (the plan reference report has been very useful in this stage). In an interactive session, which includes an operator and two or three analysts, the variables are examined, the hypotheses are modified, and the same process is iterated through. Hard copies, used as a skeleton for a short report summarizing the findings, are prepared of significant displays. In cases of more complicated problems, one session is usually not sufficient; when the displays of the first session are analyzed, more questions are raised that must be answered in a second or third session. Experience to date has shown that, unless the problems are extremely simple, a significant part of the analysis is not interactive. During the interactive session only rough ideas are formulated, but they are refined, finalized, and summarized in the relaxed environment of the designer's desk.

PROBLEMS AND DIFFICULTIES

For the CATS major review, the use of INTRANS was limited to a rather narrow range of applications. Within this range, however, the project demonstrated the significant effectiveness of IGS in planning. However the project exposed a set of difficulties that, if resolved, might increase the effectiveness of INTRANS substantially. (During fall 1973, many of the deficiencies discussed here were corrected. In particular, much more flexibility has been introduced to the interactive language and to the geographical ID methods.)

In the following paragraphs, the more significant difficulties are discussed. Part of the discussion is concerned with seemingly minute and uninteresting details that are, nonetheless, an important factor in the success or failure of IGS applications.

Interactive Analysis

Trials to perform a complete problem analysis in an interactive session have not been successful, especially when the problem has required significant amounts of concentration or mental effort. In most cases, the analyses during the interactive sessions have been limited to rough examination and acceptance or rejection of various hypotheses; major portions of the interactive sessions have been devoted to design and preparation of hard-copy displays. Detailed final analysis and drawing of conclusions have had to be done later by using the displays as a reference. Experiments in running complete interactive analysis sessions with the use of dictating machines, different group sizes, and other variables have not been very successful. More experimentation is needed to pinpoint the problems and to find effective procedures for interactive analysis.

Interactive Language

INTRANS gives the user desired flexibility in designing displays. However, to request nonstandard display design requires a lengthy command. The long commands have become quite bothersome in some cases, e.g., when similar displays have to be repeated as for all the alternatives. It is highly desirable to minimize this problem by enabling changes (during interaction) in the default display design.

More generally, it is highly desirable that the commands be as short as possible in terms of the number of characters typed in. The "natural language" commands that are now used by INTRANS make the initial training in the system very easy. However, long command words, with the resulting typing errors, become a drag to the experienced operator. A possible solution to this problem is to give the user freedom in defining his language or, more simply, to implement a number of language levels, including the option of using shorthand.

Geographical ID

The version of INTRANS used for the major review analysis required that area data be identified as a grid. This requirement did not cause many difficulties in treating CATS data. The areal ID of CATS has been based on townships that could be described as slightly distorted 6- by 6-mile grids. For the analysis, CATS zones were defined as squares that were parts of townships with varying sizes of ½, 1, 2, 3, and 6 miles.

When the data were loaded into INTRANS, the zonal data were transformed to a uniform 2-mile grid, which served quite well in most applications. However, in some cases it was found necessary to check data on specific CATS zones—a service that INTRANS could not provide. This capability seems to be useful for data editing, checks for errors, and so on.

A satisfactory solution to this problem is not easy to come by. The problem is complicated by the limitations of screen size, character-screen size ratio, size and number of zones that have to be displayed, and more. [Detailed discussion of the problems involved in creating intensity maps by IGS is given by Gur (4).] However, the problem of compatibility between the analysis zones and the display zones is important enough to justify more research and development. We are now experimenting with possible solutions to this problem. (The second version of INTRANS includes the option to store and analyze data for zones of any shape. Map displays are produced by using the proximity rule.)

Physical Environment of the Graphics Terminal

During an interactive session, analysts need, and are capable of handling, much more information than can be presented on the screen at one time. The environment of the terminal must be designed so that all this reference material can be easily viewed and handled. It seems unrealistic to plan meaningful interactive analyses with exclusive use of the CRT.

The Host Computer

CATS uses the state's computer center for data storage and simulation runs. INTRANS, on the other hand, runs on another computer. (Time-sharing is not available on the state's computer.) Therefore, data must be transferred by tapes between the two computers. This results in time delays, waste of computer time, and duplication of data files with the resulting updating problems. This problem can be best overcome if INTRANS is run on the same computer in which the data are stored.

Location of the Terminal

The INTRANS terminal was not located in CATS offices. This made the overhead on an interactive session rather high. As a result, the use of INTRANS as an occasional reference for a few pieces of information has become unfeasible. [Experience with an INTRANS terminal in the office proves the utility of short (5 to 15 minutes) sessions, which were not feasible when this paper was written.]

CONCLUSIONS

The project's main conclusion is that interactive graphics systems are already cost-effective as an operational tool for transportation planning. At the same time, it is clear that full utilization of this new tool can be realized only through much work and dedication of the people involved. The present contribution of interactive graphics is marginal and by no means revolutionary; at the same time it is apparent that this tool has tremendous potential. It seems that the rate at which the role of interactive graphics as a tool in transportation planning will grow depends mainly on the desire of practicing planners to learn how to use it for their benefit. Gaining experience in its use is now of more importance than trying to improve its sophistication and adding capabilities to systems such as INTRANS.

INTRANS has been proved to be effective mainly in the following areas:

1. Preparation of figures for subsequent use in reports, mainly by sorting alternative displays to choose the more effective ones, and detailed design of the chosen displays;
2. Examination of relationships between different variables and different plans (INTRANS makes it possible to analyze and display many relationships and pieces of information that previously were inaccessible);
3. Supplying a medium for examination and refinement of rough ideas and hypotheses regarding the attributes of the transportation system and of different plans; and

4. Supplying a convenient and fast access to large parts of the modeling data (access is instrumental for response to occasional inquiries on specific issues from within and outside of the agency).

INTRANS has been less successful in a number of areas. First, the graphic quality of INTRANS displays does not meet the standards of a high-quality final report. Manual drafting of figures for the report was required. Also, experiments to conduct complete interactive analysis of moderately complicated problems have not succeeded. The interactive session has been used most effectively to examine rough ideas and to prepare hard copies of relevant displays. Final formulation of ideas and conclusions requires further work after the interactive session.

In terms of logistics and the structure of the hardware system, the following improvements are suggested:

1. The IGS should operate on the same computer in which major data files are stored. Significant delays and inefficiencies occur when it is necessary to transfer data between computers and maintain two data files.
2. The terminal should be located within the agency's offices. Immediate access to the terminal to make fast inquiries on relatively small problems increases the utility of the system significantly.
3. In the interactive session, there should be an operator whose sole responsibility is to push the buttons based on the analyst's requests. If the analyst is required to concentrate on the operating details of the system, he will be prone to make mistakes, slowing down the analysis and ultimately losing much of the effectiveness.
4. An efficient, high-quality, and fast hard-copy device is mandatory. It is doubtful whether the system discussed would have been effective at all without such a device.

It is rather tempting to introduce a long list of new and improved capabilities that would be prestigious to add to INTRANS. Many of these capabilities would most likely be effective and will eventually be added. However, it is not apparent to me that more capabilities are of the utmost importance at present.

It is necessary, now, to study how existing interactive graphics capabilities can be used more effectively. The study has to concentrate on the man side of the man-machine system. How an interactive analysis should be prepared and conducted, the desired attributes, the required training of a good interactive analyst, and the types of problems that can best be analyzed by interactive graphics all must be determined.

Such research, together with increased use and further development in hardware and software, is likely to make interactive graphics a widely used analysis tool in transportation and urban planning in the near future.

ACKNOWLEDGMENTS

The work described in this paper was performed under a contract between the Illinois DOT and the University of Illinois at Chicago. The opinions and findings expressed herein are those of the author and not necessarily those of the supporting agencies.

The support and interest of G. Jones, A. Bicunas, H. Haack, and J. Miller of CATS and Illinois DOT have made this project possible.

C. Biddle, A. Vyas, T. Mulder, and I. Lowe, among others in CATS, helped in the data transfer and analysis. H. Stenson, S. Hess, F. Szienca, M. Lassaine, and G. Szuleck of UICC participated in performing the analysis. Computer services were supplied by the Computer Center, UICC.

REFERENCES

1. Gur, Y. INTRANS-BROWSE—An Interactive Graphics System for Planning Related Data Analysis. Highway Research Record 455, 1973.
2. Functional and Intermodal Evaluation of Alternatives for a 1995 Transportation System in the Chicago-Gary Region. Chicago Area Transportation Study, 1973.

3. Gur, Y. Presentation and Comparison of Alternative Long Range Plans for the Transportation System in the Chicago Area Using INTRANS. Chicago Area Transportation Study (in print).
4. Gur, Y. Display of Spatial Displays by Interactive Graphics Systems. Presented at Internat. Symposium on Applications of Computers and Operations Research, Washington, D.C., Aug. 1973.

ASSESSING THE UTILITY OF AN INTERACTIVE GRAPHIC COMPUTING SYSTEM: A TRANSPORTATION SYSTEMS DESIGN PROBLEM

Jerry B. Schneider and Dennis Porter,
Departments of Urban Planning and Civil Engineering, University of Washington

An experiment designed to assess the amount of improvement in the quality of a design that can be obtained by using an interactive graphic computing system was undertaken and is interpreted. The problem was design of a bus rapid transit system by five teams of students at the University of Washington. Each team used a man-computer system called UTRANS to search for a design for a BRT system that would satisfy 11 performance measures. None of the teams found a wholly satisfactory design, but the average improvement in performance (design quality) for all teams was close to 50 percent (relative to their initial designs). Further experiments of this type are needed to assess the utility of interactive graphic design tools in the transportation systems field.

•THE POTENTIAL UTILITY of interactive graphic computing systems as designing tools is still an issue of considerable debate (7). Recent hardware and software developments have made it possible to tackle a number of complex design problems with only modest investments in equipment. Still, there is far too little evidence that substantial payoffs (i.e., better designs) can be obtained by implementing an interactive graphic system.

How much improvement in the quality of a design can be expected from the use of an interactive graphic system? Our major hypothesis is that design improvements (using the initial or "first-cut" design as a base) on the order of 25 to 50 percent are likely and improvements between 50 and 100 percent are possible, given a well-trained systems operator and a well-developed interactive graphic computing system. It is probable that improvements of 50 to 100 percent need to be demonstrated before investments in interactive graphic design tools will become attractive to potential users. Improvement is defined, for our purposes, as a measure of increased performance over the initial or "first-cut" design taken as a base. It is recognized that the initial design is not an ideal base from which to measure improvement, and the limitations of using this procedure are discussed.

The substance of this paper is a report on some results of an experiment designed to assess how much improvement in design quality can be expected when the designer is aided by an interactive graphic computing system. The problem to be solved was design of a complex bus rapid transit (BRT) system. Five teams of students were asked to design a BRT system for a portion of the City of Seattle, Washington. Eleven objectives were specified for this problem, and the objective of each team was to find a design that would satisfy all 11 objectives. Each team was to use an interactive graphic computing system, known as the urban transit analysis system (UTRANS), to formulate and evaluate several alternative BRT system designs. The object of the experiment was to determine how much improvement could be obtained when man and computer work together, moving from an initial (unsatisfactory) design to one satisfying as many of the 11 design criteria as possible.

Publication of this paper sponsored by Task Force on Interactive Graphics.

The research design for this experiment was relatively simple and straightforward. The students were grouped into five teams of four or five persons each. One person from each team was given a short period of intensive training on how to operate UTRANS. All students were given some basic instruction on the theoretical and practical characteristics of the UTRAN system. All teams were assigned the same problem: Design a BRT system that will satisfy or exceed 11 objectives that relate to system performance. The definitions of these objectives were discussed in detail. Each team was to work independently, and competition between teams was encouraged. Each team was to have the same amount of computer time (in 2-hour blocks). The teams were formed with a random selection process except that each team had at least one person who had had some previous experience with computers.

In this setting, each team developed a first-cut design on paper and then began the interactive design process. This consisted essentially of looking for ways to improve the current design, making the changes, and then evaluating the modified design. Each team was expected to evolve some type of design strategy that would, more often than not, help the team find a series of successively better designs. This iterative process was to continue until the available computer time had been fully used, and the best design obtained (not necessarily the last design obtained) was to be presented to the class for discussion. Judgments on how to modify the design at each stage of the iteration were to be made by team members by using whatever decision-making procedure suited them.

Four specific evaluation objectives for the experiment were determined:

1. How similar or different are the five best designs, in both visual and quantitative terms?
2. How successful was each team in satisfying the 11 design objectives? How similar or different were the five teams in this respect?
3. How much improvement in performance over the initial design was achieved by each team and for the group as a whole?
4. What were the characteristics of the design strategies evolved by the successful teams? What were the reasons for unsuccessful efforts on the part of any team?

These questions will be discussed later in the paper.

DEFINITION OF THE BRT DESIGN PROBLEM

Network and Demand Data

The problem selected was the design of a peak-period BRT service from many residential origins to a single destination (i.e., a many-to-one service). The northern part of Seattle was chosen as the setting for the problem, and the corridor used is shown in Figure 1. The major destination is downtown Seattle, which is about 100 miles long and 5 miles wide and had a population of 200,000 in 1970. The network included all of the principal residential streets. It is bisected by Interstate 5, which has six major interchanges in the corridor. The density of people who live and work in the corridor is represented by a hexagon at each node in Figure 1. The size of the hexagon is proportional to the number of people who live near the node and who commute daily to downtown Seattle. These data, the network coded with travel times and the demand set, are the two basic elements for the problem.

Behavioral Assumptions

The heart of the UTRAN system is a modal-split model called the n-dimensional logit model (5). This model forecasts how any particular design may be expected to be used. It splits all travel to downtown Seattle among the three modes included in UTRANS: drive, park-n-ride, and walk-n-ride. All 11 performance measures are derived from the results of this modal-split forecast; thus it is most important for the planner (or designer) to understand how it works. Approximately 15 hours of instruction was devoted to this topic, and all of the participants obtained a good understanding of the mechanics of the model. The team's ability to fully understand how the model was formulated and how it worked was probably a major factor in its ability to formulate a successful design strategy.

Performance Measures Used for Evaluation of Alternate Designs

All evaluations of alternative designs were based on the 11 performance objectives. These performance measures, their definitions, and their desired values are given in Table 1. These measures relate to the travel cost, comfort, profitability, and the sociopolitical effects of a BRT system. Other important impact categories were not included in the interests of keeping the complexity of the problem within reason.

These performance measures are not an ideal set for use in this type of experiment inasmuch as some of them are interrelated. Measures that are more independent and, in some cases, less abstract would have been preferred. However, these measures were available, and generation of others would have involved additional programming. This was not possible because of resource limitations. These difficulties did serve a useful educational function in that the students learned about the problems of working with interdependent and abstract measures. While troublesome, these problems were not judged to be serious enough to invalidate the basic experiment.

Estimation of Weights for Performance Measures

When the performance measures had been defined and discussed, each student was asked to complete a partial paired comparisons matrix, with constant sums, to provide his own estimate of the relative importance of each of the performance measures. These results, when analyzed statistically, showed that there was no significant difference among the group average values of these weights. Even though these weights were about equal, some of the teams gave somewhat more thought to improving those measures that ranked highest in this analysis.

Estimation of Minimum-Maximum and Ideal Standards for Performance Measures

As a further step in the definition of the problem, each participant was asked to estimate what he thought the minimum, maximum, and ideal levels of each performance measure should be. These preferences were then analyzed, and a plot of the results for each performance measure for the group was made.

The results of this exercise revealed that some significant differences existed in the interpretation of the definitions of the performance measures. After these differences were identified and discussed by the class, the exercise was repeated, and satisfactory consensus on appropriate values was obtained. The average values for the group are given in Table 2. These values served as a perspective on the range of reasonable variation for the performance measures and were used to establish a common scale for comparison of the best designs of each of the five teams.

Constraints Imposed by the Instructor

A further definition of the problem was made by imposing a set of constraints on the design problem. Some of these constraints were due to hardware or software limitations. Others were included to represent typical environmental constraints that always appear in problems of this type. These constraints were as follows:

1. No more than 60 bus stops and 20 bus lines,
2. No more than 20 park-n-ride lots and no lot larger than 1,000 spaces, and
3. A downtown all-day parking fee of $1.50 and an average walk time downtown of 5 minutes.

This problem definition process together with the presentation of the UTRANS modal-split model required about 15 hours of instructional time and was considered essential to the conduct of the experiment.

CHARACTERISTICS OF THE PARTICIPANTS AND THE PROBLEM-SOLVING ENVIRONMENT

Characteristics of the Participants

Twenty-three students, from both undergraduate and graduate levels, took part in

Figure 1. Seattle study corridor.

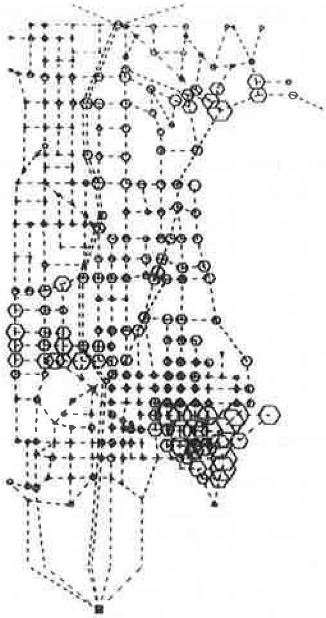


Table 1. Definitions and desired values of performance measures.

Performance Measure	Definition	Specified Satisfactory Level
Profit-loss of bus operations, dollars	Bus fare-box revenues less operating costs	0
Profit-loss of park-n-ride lots, dollars	Park-n-ride lot revenues less operating costs	-250
Walk-n-ride patronage share	Percentage of total demand using walk-n-ride mode	20
Park-n-ride patronage share	Percentage of total demand using park-n-ride mode	30
Bus load ratio, seated patrons	Systemwide efficiency factor of bus operation; measure of overall use of bus carrying capacity	0.80
Percentage of standing room used	Systemwide measure of crowdedness on bus	30
Park-n-ride lot load factor	Systemwide efficiency factor of park-n-ride lot operation; measure of overall use of park-n-ride lot parking capacity	0.80
Walk-n-ride accessibility index	Constructed accessibility measure for walk-n-ride mode	70.0
Park-n-ride accessibility index	Constructed accessibility measure for park-n-ride mode	120.0
Percentage of people within 5-min walk	Proximity measure of closeness of total demand to any bus stop	35
Percentage of people within 5-min drive of bus stop	Proximity measure of closeness of total demand to any park-n-ride lot	70

Table 2. Minimum-maximum and ideal levels of performance measures.

Performance Measure	Group Average Value	
	Minimum-Maximum	Ideal
Profit-loss of bus operation, dollars	-260.0	+72.7
Profit-loss of park-n-ride lots, dollars	-151.0	33.5
Walk-n-ride patronage share, percent	17.0	38.5
Park-n-ride patronage share, percent	22.0	40.7
Bus load ratio	0.63	0.94
Percentage of standing room used	35.0	2.00
Park-n-ride lot usage, percent	0.63	95.0
Walk-n-ride access index	58.4	96.6
Park-n-ride access index	77.0	121.1
Percentage of people within 5-min walk of bus stop	30.0	70.0
Percentage of people within 5-min drive of parking lot	37.7	70.5

this experiment. Of this number, only four had had any previous experience with man-computer systems, and only one had worked with the UTRAN system. They came from a variety of disciplines as given below:

Discipline	Participants
Urban planning	9
Civil engineering	6
Architecture	4
Psychology	2
Anthropology	1
Geography	1
Total	23
Level	
Undergraduate	8
Graduate	15
Total	23

None had had any experience with BRT system design. The most often expressed motivation to participate was the desire to obtain some "hands-on" experience with a man-computer system. Unfortunately, because of the large size of the group and the limited hardware available, only one person from each team was allowed to operate the computer. However, other members of the teams were able to observe the operation of the UTRAN system closely and to participate in the problem-solving (design) process (both on and off line).

UTRANS Operating Characteristics

The origins, evolution, and characteristics of the UTRAN system have been well-documented by Rapp (2, 3, 4), Gehner (1), and Schneider (8) and will not be discussed extensively here. Briefly, UTRANS is operated as shown in Figure 2. It has been structured to assist a planner in generating and evaluating alternative BRT system designs for service in urban activity centers. It is limited to cases where there are many origins and one destination. Two modes of operation are possible. The first is the manual mode where the planner makes a series of design decisions. The second is a partially automated mode, which relieves some of the decision-making burden. In the manual mode the planner is presented with a display of the street network, the demand pattern (i.e., the location of the people who desire to travel to the major activity center), and a display of land values superimposed on the street network. Then the planner lays out a first-cut transit system design by making a series of design decisions and entering them either on the graphic display or on the keyboard of the graphics terminal in the sequence shown in Figure 2.

Once the design decisions have been made, the first-cut transit system design is ready for computer evaluation. Each of these decisions is recorded by the computer and is input into a modal-split model. This model is designed to estimate the proportion of trip-makers that will use each of three modes of getting from their homes to their destination. The modal-split model constitutes the heart of this man-computer system inasmuch as the evaluation of alternative designs is derived wholly from its prediction of the expected performance of each design. This prediction procedure is based on the following assumptions:

1. Each trip-maker selects from among walk-n-ride, park-n-ride, or drive modes.
2. The trip-maker's choice depends on the relative difficulty (or impedance) that he perceives with each mode.
3. The total impedance of a mode is the sum of the impedances associated with the several elements of trip by that mode.
4. Each element of a trip is multiplied by a constant that is proportional to an esti-

Figure 2. Operation of UTRAN system.

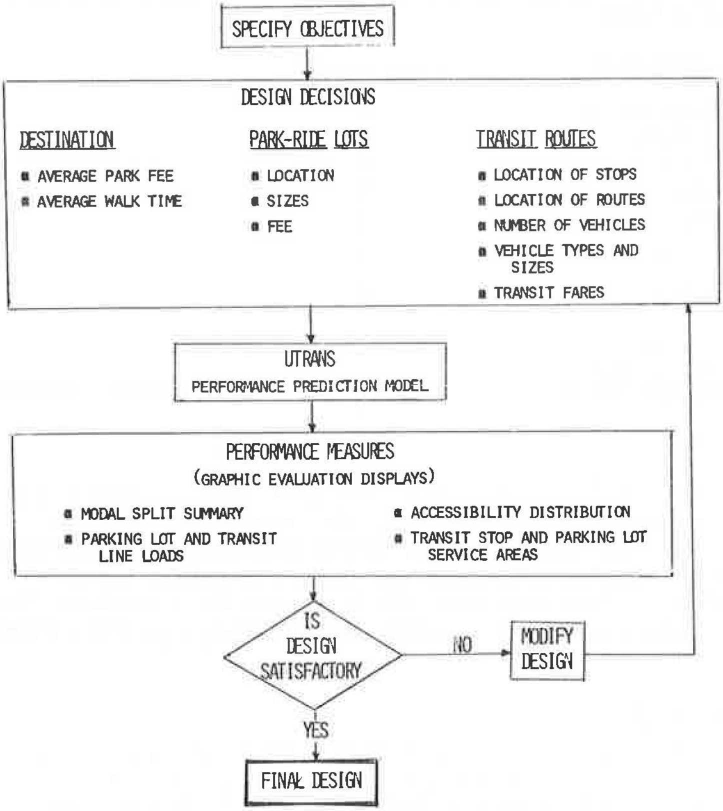
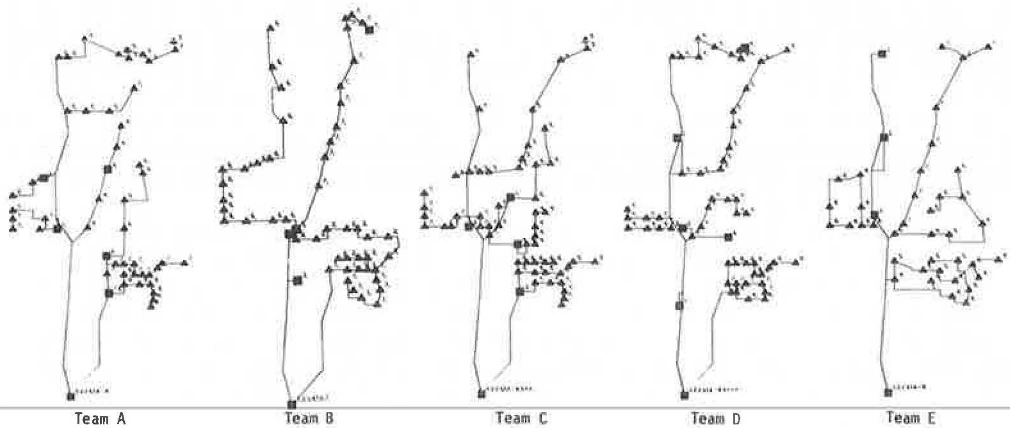


Figure 3. Final designs of each team.



mate of the relative disutility associated with that type of activity. The constants are called impedance coefficients.

5. The smaller the total impedance of each mode is, the more likely it is that a trip-maker will select it.

6. The share of the available patronage attracted by each of the three modes is inversely proportional (in a negative exponential fashion) to the mode's overall impedance.

The output of the modal-split model is as follows:

1. The percentages of all trip-makers using the three modes,
2. For each transit stop and park-n-ride lot, the volume of patrons who walk or drive (if the stop has a parking lot) to it,
3. For each transit line and parking lot, the costs and revenues of operation (including capital costs), and
4. The total system cost and revenue on a daily (24-hour) basis (the difference between these two figures is the overall daily profit or loss of the system).

This information is presented to the planner in tabular form on the scope face of the graphic terminal. Several additional evaluation displays are available to the planner, all of which are derived from the output of the modal-split model.

In the first cycle the planner structures a first-cut design, and the computer evaluates it and presents him with a variety of displays which he examines. The task then is to develop ideas that should improve the first-cut design by adding park-n-ride lots, changing parking fees, increasing the number of buses on a route, or modifying the original design in some other way. This revised or second design is evaluated by the computer and a new set of evaluations and a listing of performance parameters are displayed. This procedure is repeated until the planner achieves a satisfactory design or is restricted by time constraints. This is the process used by each of the five teams in the conduct of this experiment.

RESULTS OF THE EXPERIMENT

Implementation of the experiment was hindered by hardware malfunctions that interrupted it after the off-line instruction had been completed and the teams were ready to begin work on the UTRAN system. This delay probably reduced the effectiveness of the teams' ability to use the instruction and probably introduced a conservative bias into the results.

The results of the experiment have been grouped into four categories:

1. A description of the physical attributes of the five final designs,
2. An analysis of the performance levels of the five final designs,
3. Measures of the amount of improvement achieved by each of the teams, and
4. A general discussion of the problem-solving strategies used by the teams.

Physical Attributes of the Five Final Designs

One of the questions of interest in this research was whether five teams, working independently, would come up with similar design solutions. Comparisons among complex transit system designs can be made in at least two ways. One query is, Do they look alike? Another is, Are they composed of similar quantities of the various elements (i.e., buses, park-n-ride lots, and so on) that make up a design solution? As an answer to the first question, Figure 3 shows a comparison of the five final designs.

These five designs have a similar appearance. Most involve relatively complex networks, long bus lines, and close bus stop spacing. Park-n-ride lots tend to be located more in the lower half of the network, closer to the downtown destination. The design of team E is a minor exception to some of these qualitative observations. Most teams obviously attempted to provide the greatest service to the nodes of highest demand (Fig. 1). Each team, with the exception of team E, used all the major arterials in the corridor for one or more bus lines. It is possible that the similar appearance of the designs was structured by the dominance (accessibility-wise) of the single Interstate highway and the concentration of demand in the lower right part of the network.

Table 3 gives data relative to the second question. The first three items in Table 3 are quantitative in nature, whereas the other five are more qualitative. The physical elements and operating policies of the five designs are quite similar. This is probably not too surprising because all teams used fairly conservative approaches to the design problem and did little experimentation with unconventional designs. Most teams said that, if they had had more operating time on the computer, they would have tried more unconventional ideas in the design process.

Performance of Final Designs

Figure 4 shows the results of the final designs in terms of the 11 performance measures. Each performance measure is shown as a percentage of the difference between the minimum-maximum standard and the satisfactory level of each measure as given at the beginning of the experiment. As can be seen, none of the teams was able to satisfy all 11 objectives. Teams A and C satisfied eight of the objectives, two teams satisfied seven, and one team satisfied only six of the objectives. Figure 4 shows that all teams were able to oversatisfy some objectives, whereas others proved to be particularly troublesome to nearly all the groups. Data given in Table 4 show how often each of the 11 performance objectives was satisfied. As can be seen in Table 4, all teams were able to satisfy two objectives.

In terms of the 11 performance measures, the final designs were quite dissimilar. Table 5 gives the simple correlation coefficients for all possible pairs of designs. As can be seen, only three of these 10 coefficients are greater than 0.7 while six are less than 0.5. Thus, although the designs are quite similar in physical terms, they are very different in performance terms. The reasons for these variations in performance are many and varied and will be discussed in more detail later.

Measures of Design Improvement Obtained

As discussed previously, the major objective of the experiment was to determine how much improvement in the initial design could be achieved by each of the five teams. The measure used to address this question was as follows:

$$I \times 100 = \sum_j \frac{PM_{jf} - PM_{ji}}{PM_{ji}}$$

where

I = percentage of improvement obtained over initial design,
PM_{jf} = value of performance measure j in the final design f, and
PM_{ji} = value of performance measure j in the initial design i.

The average value of I for each team is given below.

Team	I (percent)
A	38
B	100
C	63
D	1
E	28

The average for all teams for all performance measures was 46 percent. The largest overall improvement was obtained by team B (100 percent), followed by team C with 63 percent. Teams A and E made substantial improvements (38 and 28 percent respectively), whereas team D made virtually no progress over its initial solution. The average improvement for all five teams and all 11 performance measures was 46 percent. This figure needs to be interpreted carefully.

It can be argued that the initial solution of a group of novice designers is not a reasonable base for measuring improvement because it was probably either very conserva-

Table 3. Team comparisons of transit system attributes.

Attribute	Team A	B	C	D	E
Number of bus lines	8	7	11	12	8
Number of bus stops	52	54	59	58	42
Number of park-n-ride lots	5	4	4	5	3
Varying bus headways	Yes	Yes	Yes	Yes	Yes
Varying bus types	Yes	Yes	Yes	No	Yes
Varying bus fares	Yes	Yes	Yes	Yes	No
Varying park-n-ride lot sizes	Yes	No	Yes	Yes	Yes
Varying park-n-ride lot fees	Yes	No	Yes	No	Yes

Figure 4. Comparative performance of final designs.

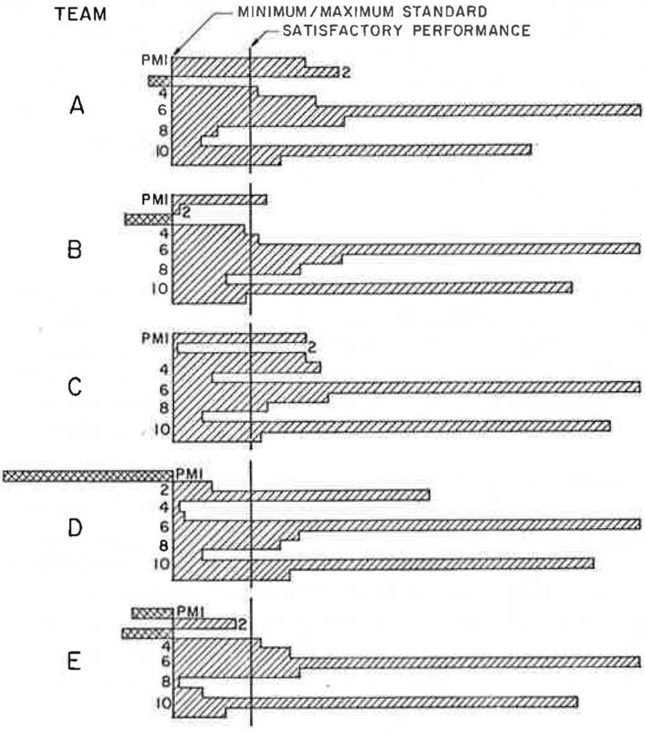


Table 4. Number of times performance objectives satisfied by five design teams.

Performance Measure	Number of Times Satisfied
Percentage of standing room used	5 of 5
Park-n-ride lot load factor	5 of 5
Profit-loss of bus operations	3 of 5
Park-n-ride patronage share	3 of 5
Bus load ratio, seated patrons	3 of 5
Walk-n-ride access index	3 of 5
Percentage of patrons within 5-min walk	3 of 5
Profit-loss of park-n-ride lots	1 of 5
Walk-n-ride patronage share	1 of 5
Park-n-ride access index	0 of 5

Table 5. Correlation coefficients for final designs.

Team	Team A	B	C	D	E
A	1.00	0.03	0.34	-0.68	-0.47
B		1.00	0.86	0.20	0.80
C			1.00	-0.30	0.41
D				1.00	0.74
E					1.00

tive or a wild guess. Although there is some truth in this position, all of the teams worked out their initial design on paper and gave it considerable thought. Still, if we assume that the inexperience of the participants resulted in poor initial designs, we should probably discount the 46 percent overall improvement figure by 10 to 20 percent.

On the other hand, several of the teams were confident that they could have improved their final design substantially if they had had more time to work on it. That two teams were able to achieve improvement proportions of 100 and 68 percent proves that this may be possible. Therefore, it seems likely that, given more time and more experienced designers, improvements in the 60 to 80 percent range would not be difficult to achieve. If this is true, then a general expectation of a 50 percent improvement in design performance seems to be justified in the context of the experiment being reported here. This amount of improvement is large enough to warrant further investments of time and resources in experiments of this type. If similar results are obtained, it will be possible to build a substantial case for the use of interactive graphic design tools in transportation systems design studies.

One further qualification of these results is needed. Each team spent approximately 8 hours working with the UTRAN computer system. Moreover, each team probably spent another 6 to 8 hours planning and discussing the design effort outside the computer room. In addition, some 15 hours of classroom time was used to prepare the participants for the design task. Because it will eventually be necessary to relate improvement obtained with time expended, future studies of this topic should be designed to incorporate this factor. It has not been possible to do so effectively in this study because so much time was spent in learning.

Problem-Solving Strategies

Each team devised a strategy that best matched its perception of the problem. No rigorous analysis of these strategies is attempted here, for they proved to be quite diverse and difficult to compare. Each team generated and evaluated seven or eight design solutions. All but one of the teams were able to improve their designs in successive iterations so that their best designs were their final effort. The most unsuccessful team (team D) generated eight designs but found that the second design was better than the final (eighth) design.

The two most successful teams (as measured by performance improvement achieved) both used an incremental approach to the design problem. They started with a fairly simple design and made those additions to it that they felt were most needed as a result of their evaluation of the previous design. On the other hand, the least successful team began with a very complex design and apparently became quite confused in the process of trying to improve the performance measures. This frustration with the problem led to a fairly negative attitude toward the experiment on the part of some members of this team, and the result (no improvement) was partly a function of this difficulty.

Generally speaking, the results were quite satisfactory when viewed in terms of the human element in the system. Our results tend to confirm those of Sackman (6), who found that access to interactive graphic systems tends to expand or accentuate individual differences. This simply means that a person who has some innate or learned capacity for design work will tend to be aided proportionally more by the computer system than will someone who has less of an innate or learned design ability. In other words, the availability of good interactive graphic design tools may be expected to increase the gap between good and poor designers while the whole spectrum of design quality will be shifted upward.

Much more research needs to be done on the strengths and limitations of the human mind in design situations like the one in this paper. Some work has been done along these lines (6, 9, 10), but much more work remains to be accomplished before we can expect to properly design man-computer systems and to adequately train people to make effective use of them.

CONCLUSIONS

1. The use of the UTRAN interactive graphic system enabled a group of novice designers to make design improvements estimated at close to 50 percent over the initial or "first-cut" solutions. This figure was greater than expectations and substantial enough to warrant further investigations of this type.

2. Four of five design teams were able to make steady progress toward improved performance from their initial design through seven to eight trials to their best design.

3. The designs of teams that began with a simple design and modified it incrementally performed better than other cases. However, the strategies employed by the various teams were quite diverse and are difficult to generalize. A more structured approach to the analysis of design strategies is needed.

4. The designs produced by the five teams were quite similar in appearance and highly similar in the physical elements of which they were composed, but design performance was different.

5. Difficulties with the definition of performance measures, interdependency among performance measures, computer malfunctions, and problem complexity all hindered the ability of the design teams to operate effectively.

6. Replications of the results obtained here are needed before a clear-cut case can be formulated to support the development and use of interactive graphic systems for aiding the transportation systems designer. These results should be regarded as preliminary and indicative of the need for more investigations of this type.

REFERENCES

1. Gehner, C. D., and Clark, J. W. The Urban-Region Transit Analysis System (UTRANS), Volume I: Documentation, Vol. II: User's Manual. Urban Transportation Program, University of Washington, Seattle, Res. Rept. 72-2, 1973.
2. Rapp, M. H. Man-Machine Interactive Transit System Planning. Socio-Economic Planning Sciences, Vol. 6, 1972, pp. 95-123.
3. Rapp, M. H. The Interactive Graphic Transit Simulator: A Tool for Planning Node-Oriented Transit Systems. Urban Transportation Program, Univ. of Washington, Res. Rept. 7, 1971.
4. Rapp, M. H. Planning Demand-Adaptive Urban Public Transportation Systems: The Man-Computer Interactive Graphic Approach. Urban Transportation Program, Univ. of Washington, Res. Rept. 71-4, 1972.
5. Rassam, P. R., Ellis, R. H., and Bennett, J. C. The n-Dimensional Logit Model: Development and Application. Highway Research Record 369, 1971, pp. 135-147.
6. Sackman, H., and Citrenbaum, R. L. Online Planning: Towards Creative Problem-Solving. Prentice-Hall, 1972.
7. Sackman, H. Man-Computer Problem-Solving: Experimental Evaluation of Time-Sharing and Batch Processing. Auerbach Publishers, 1970.
8. Schneider, J. B. Man-Computer Synergistics: An Aid to the Design of Urban Systems. Urban Transportation Program, Univ. of Washington, 16-mm movie, 1972.
9. Schneider, J. B. Solving Urban Location Problems: Human Intuition Versus the Computer. Jour. American Institute of Planners, Vol. 37, No. 2, March 1971, pp. 95-99.
10. Schneider, J. B., Symons, J. G., Jr., and Goldman, M. Planning Transportation Terminal Systems in Urban Regions: A Man-Machine Interactive Problem-Solving Approach. Transportation Research, Vol. 6, pp. 257-273.

SPONSORSHIP OF THIS RECORD

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Charles V. Wootan, Texas A&M University, chairman

TRANSPORTATION FORECASTING SECTION
E. Wilson Campbell, New York State Department of Transportation, chairman

Committee on Transportation Systems Design
Joseph L. Schofer, Northwestern University, chairman
Donald S. Berry, Daniel Brand, Thomas B. Deen, Ronald J. Fisher, Lonnie E. Haefner, John R. Hamburg, Thomas N. Harvey, Kevin E. Heanue, George C. Hemmens, Marvin L. Manheim, Elizabeth J. McLean, Joseph C. Oppenlander, Louis J. Pignataro, James E. Sale, Edward F. Sullivan, Martin Wachs, George V. Wickstrom, Richard D. Worrall, Dennis R. Young

Task Force on Interactive Graphics
Jerry B. Schneider, University of Washington, chairman
William G. Barker, Robert B. Cody, Robert B. Dial, Larry J. Feeser, Yehuda Gur, Hubert A. Henry, Edgar M. Horwood, George T. Lathrop, Peter S. Loubal, Ernest Nussbaum, Lloyd Pernela, Matthias H. Rapp, Earl R. Ruiter, Martin Wachs

James A. Scott, Transportation Research Board staff

Sponsorship is indicated by a footnote on the first page of each report. The organizational units and the chairmen and members are as of December 31, 1973.