TRANSPORTATION RESEARCH RECORD

No. 1453

Highway Operations, Capacity, and Traffic Control

Intelligent Transportation Systems: Evaluation, Driver Behavior, and Artificial Intelligence

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1994

Transportation Research Record 1453
ISSN 0361-1981
ISBN 0-309-06061-3
Price: $26.00

Subscriber Category
IVA highway operations, capacity, and traffic control

Printed in the United States of America

Sponsorship of Transportation Research Record 1453

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Chairman: Thomas F. Humphrey, Massachusetts Institute of Technology

Transportation Forecasting, Data, and Economics Section
Chairman: Mary Lynn Tischer, Virginia Department of Transportation

Committee on Transportation Supply Analysis
Chairman: Hani S. Mahmassani, University of Texas at Austin
David E. Boyce, Yupo Chan, Carlos F. Daganzo, Mark S. Daskin, Michel Gendreau, Theodore S. Glickman, Ali E. Haghani, Randolph W. Hall, Rudi Hamerslag, Bruce N. Janson, Haris N. Koutsopoulos, Chryssi Malandraki, Eric J. Miller, Anna Nagurney, Earl R. Ruiter, K. Nabil A. Safwat, Mark A. Turnquist

GROUP 5–INTERGROUP RESOURCES AND ISSUES
Chairman: Patricia F. Waller, University of Michigan

Committee on Artificial Intelligence
Chairman: Michael J. Demetsky, University of Virginia
M. Hadi Baaj, Anselmo Osvaldo Braun, Edmond Chin-Ping Chang, Louis F. Cohn, David J. Elton, Ardeshir Faghri, Jerry J. Hajek, Chris T. Hendrickson, Shinya Kikuchi, David Martinelli, Olin W. Mintzer, Prahlad D. Pant, Ajay K. Rathi, Shashi Kumar Sathisan, Gary S. Spring, James A. Wentworth, Charles A. Wright

Committee on Intelligent Vehicle-Highway Systems
Chairman: Daniel Brand, Charles River Associates Inc.
Stephen Albert, Sadler Bridges, Melvyn Cheslow, Patrick J. Conroy, Randolph M. Doi, Eugene Farber, Ronald J. Fisher, Robert L. French, William J. Harris, Jr., Thomas A. Horan, Eva Lerner-Lam, Hani S. Mahmassani, Joel E. Markowitz, Stephen Edwin Rowe, Donald A. Savitt, Lyle Saxton, T. Russell Shields, Steven E. Shladover, William M. Spreitzer, Joseph M. Sussman, Richard J. Weiland

Transportation Research Board Staff
Robert E. Spicher, Director, Technical Activities
James A. Scott, Transportation Planner
Nancy A. Ackerman, Director, Reports and Editorial Services
Naomi Kassabian, Editor

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1993.

Transportation Research Record 1453

Contents

Foreword

The papers in this volume focus on various aspects of intelligent transportation systems (ITS) [formerly intelligent vehicle-highway systems (IVHS)]. Brand and Underwood and Gehring discuss criteria for evaluating IVHS. The approach by Havinoviski et al. is regional; they discuss a study in Orange County, California. Stevens defines concepts for automated highway systems (AHS). Abdel-Aty et al. and Polydoropoulou et al. focus on the influence of traffic information on driver behavior. Daganzo and Lin present a solution to the hydrodynamic model for a freeway carrying morning commuters to a single destination.

In the area of artificial intelligence (AI), Chang and Wang use fuzzy set theory to improve freeway incident detection, Hua and Faghri apply artificial neural networks to IVHS, Khasnabis et al. use AI in urban rail corridor control, and Smith and Demetsky use neural networks in short-term traffic flow prediction.

Criteria and Methods for Evaluating Intelligent Transportation System Plans and Operational Tests

DANIEL BRAND

An evaluation process for the preparation of intelligent transportation system (ITS) plans that is sensitive to the differences between ITS and conventional transportation improvements is described. [The term "intelligent transportation system" replaces "intelligent vehicle-highway system (IVHS)."] A relatively complete set of evaluation criteria for ITS improvements is presented that is structured to clarify the confusion between the supply and demand impacts of ITS. This separation between "efficiency" and "output" measures means that it is possible to distinguish between ITS technology efficiency benefits and the individual and corporate demand responses to ITS that actually increase output (benefits) over those produced by the technology alone. The proposed criteria structure also incorporates the time scale of the impacts. This highlights certain fundamental correlations between the criteria that can lead to double counting of benefits and to highly correlated outcomes, which are not helpful in choosing between alternatives. The criteria structure facilitates selection by decision makers of greatly reduced criteria sets to simplify ITS evaluations. By recognizing the separate supply (efficiency) and demand (increased output) impacts of ITS, it is also possible to avoid dramatically underestimating the benefits of the new technology and to avoid serious mistakes in assessing the safety, environmental, and energy impacts of ITS alternatives. Default values to evaluate ITS improvements for inclusion in transportation system plans are provided. The criteria and default values highlight where research and operational tests can provide improved values and information that will most quickly advance the state of the art of ITS evaluation.

Developing and evaluating intelligent transportation system (ITS)—formerly intelligent vehicle-highway system (IVHS)—plans require a methodology that meets the following requirements:

1. Is fully sensitive to differences between ITS and conventional transportation improvements;
2. Recognizes that many criteria are measures of the same benefits and therefore aggregates these evaluation criteria to minimize double counting and misplaced higher implied weights given to the same consequences under different names;
3. Is sensitive to the needs of various groups in society and areas within a region or state to benefit from the program;
4. Provides strategic direction (where should we head and is it really worthwhile to undertake ITS projects to get there?);
5. Emphasizes accurate and sensible results (subject to face validity checks) rather than (false) precision;
6. Avoids criteria specific to individual actions that promote their adoption in a "self-fulfilling" evaluation; and
7. Focuses as much as possible on site-specific results (rather than hoped-for achievement of benefits in a generic type of setting).

To satisfy Requirements 1 and 2 it is necessary to do the following:

Charles River Associates, 200 Clarendon Street, Boston, Mass. 02116–5092.

- Avoid underestimating the mobility and other personal and corporate economic benefits from ITS; and
- Recognize the occurrence over varying periods of time of the same impacts under different names.

Requirements 3 and 4 lead to a three-stage evaluation process:

1. Stratification of projects by location (e.g., by geographic area within a region or a state);
2. Grouping of projects by their relative merit within strata; and
3. Evaluating the absolute worth of candidate ITS projects for inclusion in a system plan or reporting the results of an operational test.

ITS EVALUATION PROCESS

Figure 1 is a flowchart of the ITS plan development and evaluation process. The process starts with development of program goals and a set of candidate projects responsive to these goals. As shown in the flowchart, the projects can be stratified by geographic area (location) within the region or state for which the ITS plan is being developed. Project impacts and costs are then assessed relative to a set of evaluation criteria developed as described below, and the projects are grouped by their relative merit within strata. A budget constraint can be developed for each stratum, and projects from the groups having the most merit can be included in programs of projects up to the budget limits for each stratum.

Finally, for them to be included in an ITS system plan, the projects must meet not only the budget test but also an absolute-worth test. Therefore, the last stage of the evaluation process involves carrying out program-level benefit-cost analyses. If the programs of projects have benefits that exceed their costs, the program can be recommended for implementation. If a program fails the benefit-cost test, the process can be repeated, at least to the step of revising the allocated program budget for the relevant project stratum. The process may also require redefinition of the evaluation criteria or program goals, or both.

In summary, this evaluation process implies that there will be a cutoff of projects in the plan on the basis of adhering to known or assumed budget limits, as well as meeting certain benefit-cost thresholds. In addition, this evaluation process breaks considerable new ground. A comprehensive screening and evaluation of a very widely cast net of candidate ITS projects based on their site-specific benefits and costs had not been carried out earlier before the application of this methodology in the preparation of the Washington State and metropolitan Boston ITS strategic plans (*1*). In the past,

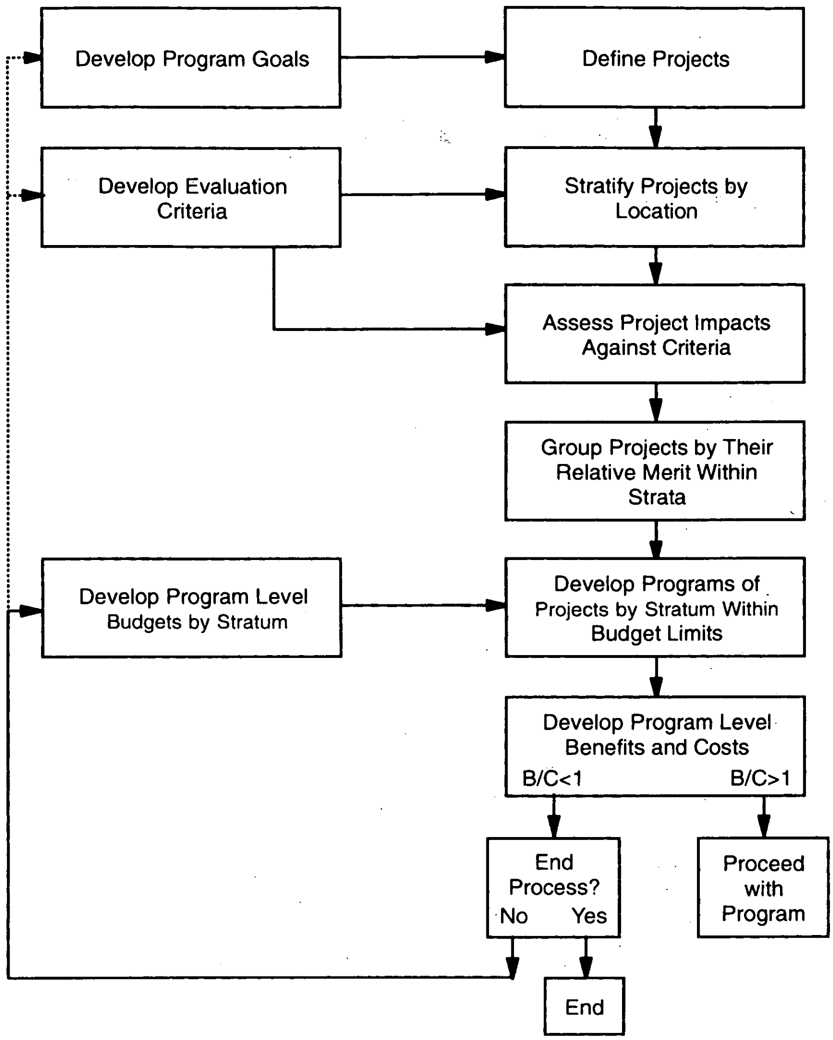


FIGURE 1 ITS evaluation methodology.

ITS plans have consisted of lists of projects deemed worthwhile on the basis of hoped-for results. Although this method is entirely acceptable for planning a research program whose payoff cannot be known in advance, ITS is now advancing, as it should, into its production mode. This puts severe demands on the current status of knowledge of ITS impacts.

Similarly, the evaluation of ITS operational field tests requires identifying and anticipating the impacts of ITS to be able to measure them as part of the operational test evaluation. Since only now is it becoming possible to recognize the existence of very important differences between the impacts of ITS and those of conventional transportation improvements, there is considerable uncertainty in quantifying many, if not most, ITS benefits. These limitations point up the urgent need for systematic evaluations of ITS operational tests to be able to quantify the impacts of ITS projects and develop ITS system plans that provide net benefits to society. The first step in this process is to develop an appropriate set of evaluation criteria that allows one to anticipate and evaluate the important impacts of ITS projects.

PROPOSED EVALUATION CRITERIA

A comprehensive list of appropriate ITS evaluation criteria is presented in Table 1. Except for the costs of ITS, most of the criteria in Table 1 are positively worded, which is not intended to imply that ITS projects have only positive impacts. Certainly there will be projects that score negatively with respect to various criteria.

The major structure of the criteria in Table 1 is along the following two dimensions:

	Time Scale		
Criterion Type	*Short Term*	*Medium Term*	*Long Term*
Increased operational efficiency (supply)			
Demand adjustments that further increase output			

This structure deals head-on with the great confusion between supply and demand impacts in ITS evaluation. The separation between efficiency (Criterion 1.0 in Table 1) and output (Criterion

TABLE 1 Comprehensive List of ITS Evaluation Criteria

1. Increased Operational Efficiency (supply-side efficiency, meaning more output per unit of input)

   1.1 Short Term: Transportation System Operation

   1.1.1 Infrastructure Efficiency
   - Increased throughput or effective capacity
   - Increased speeds
   - Reduced stops
   - Reduced delay at intermodal transfer points
   - Reduced operating costs (e.g., from ETTM or information for incident response, etc.)
   - ITS O&M cost

   1.1.2 Vehicle Efficiency

   1.1.2.1 Private Autos
   - Increased vehicle occupancy
   - Reduced operating costs (including wear and tear)
   - ITS O&M cost

   1.1.2.2 Transit
   - Reduced operating costs
   - Increased usage (i.e., volume of people moved)
   - Facilitate fare collection and fare reduction/equity strategies
   - APTS O&M cost

   1.1.2.3 Freight
   - Reduced operating costs
   - Increased throughput (i.e., volume of goods moved by the existing fleet)
   - CVO O&M cost

   1.2 Medium Term: ITS Costs
   - Capital costs of ITS
   - Liability costs of ITS

   1.3 Long Term: Investment Costs
   - Reduced capital costs of new infrastructure
   - Improved data for more cost-effective transportation investment planning
   - Improved data for concurrency planning

(continued on next page)

2.0) measures means that it is possible to separate the ITS technology benefits from the individual and corporate demand responses to ITS that actually increase output (benefits) over those produced by the technology alone. This separation also makes it possible to evaluate induced travel. Induced travel has its negative physical (travel volume, congestion, and flow-related environmental) impacts and positive mobility and economic benefits; each one is dealt with separately (if imperfectly, given today's demand models).

The structure also deals explicitly with the time frame of the impacts. Some impacts occur quickly, typified by travel behavior responses to ITS changes (Criterion 2.1). Some take more time to occur, as exemplified by ITS technology investments (Criterion 1.3) and investments in other plants and equipment (Criterion 2.2) to increase the productivity of the economy. Finally, there are the long-term impacts such as infrastructure cost savings and changes in long-run demand (Criteria 1.3 and 2.3, respectively). In most cases, the impacts that occur over various lengths of time are responses to the same underlying benefits of ITS. Therefore, the same benefits may be considered (double counted) a number of times. Organizing the impacts according to their time scale highlights certain fundamental correlations between the criteria and helps simplify the evaluation process. This is discussed further.

More generally, the evaluation criteria in Table 1 are not ITS strategy or technology specific. For example, "reduced delay at border crossings" is not included as a separate criterion because it suggests a certain set of actions; rather, "reduced delay" and reductions in the various personal, shipper, and corporate user costs should suffice. Similarly, the following are not included on the list:

- Reduced delay from improved incident detection,
- Reduced incident response times,
- Reduced accidents from improved . . . ,
- Benefits from 911 emergency services,
- Improved air quality by smoothing traffic flow, or
- Information to agencies to improve system operation.

These criteria are all specific to individual ITS options or strategies. They are not included in Table 1 because they are biased to the options and would lead to "self-fulfilling" evaluations.

TABLE 1 (*continued*)

2. Increased Output (demand adjustments that further increase output or benefits from ITS improvements) 2.1 Short Run: Mobility 2.1.1 Personal (passenger) • Increased travel opportunities (trip end benefits) • Decreased costs (disutility) of travel (including travel and delays to unfamiliar drivers/travelers). Includes: – Increased awareness, and ease of use of transit and ridesharing – Travel time (and its various components) – Travel time reliability – Travel cost (and its various components) – Comfort, stress, fatigue, confusion, etc. – Safety and personal security • Increased sense of control over one's own life from predictable system operation (including toll and transit fare charges) 2.1.2 Freight • Decreased cost of freight (goods) movement to shippers, including: – More reliable "just in time" delivery – Travel time – Travel cost – Driver fatigue, stress, etc. – Cargo security – Safety (e.g., from tracking hazardous material) – Transaction costs 2.2 Medium Run: Economic Development • Increased access to – Labor – Materials – Markets • Increased industrial output • Reduced costs • Increased investment in plant and equipment • Opportunities for new services/product innovation • Opportunities for public/private partnerships • Increased international competitiveness 2.3 Long Run: Personal Adaptations • Lifestyle changes • Land use (settlement) pattern changes (to internalize or otherwise be "informed by" congestion and other social costs of private travel and location decisions)

(*continued on next page*)

WHY SEPARATE DEMAND SIDE FROM SUPPLY SIDE?

ITS differs from conventional transportation improvements in the way information is communicated and used to increase the benefits from travel and transportation system operation. Information is communicated in real time to the traveler on the transportation system status and operation and on travel services and trip end opportunities. Information is also communicated in real time to the system to improve its operational control capabilities and its ability to provide the most helpful information to the traveler.

When the pre-ITS concern was to improve the physical transportation infrastructure, improvements were evaluated on the basis of the use of the network. Aggregate observable VMT on the network, and congestion and travel times on links were the measures of interest. With the development of a parallel information infrastructure, parallel emphasis must be on the use of the information; how individual travelers use the information to make their personal travel decisions and how firms use the information to ship their product. Mobility, which is what is being sought, is measured by the opportunities for, and the benefits from, travel. One can anticipate that there will be significant individual and corporate demand responses (Criterion 2.0 in Table 1) to the ITS information that will increase the benefits of ITS systems over and above their improved system operational efficiency (Criterion 1.0). These mobility benefits of ITS must be measured at the level of the individual tripmaker or firm instead of being based on aggregate flow volumes or travel times on the network (*2*).

For example, by providing reliable attraction location information and travel directions to unfamiliar drivers, the 1992–1993

TABLE 1 *(continued)*

3. Safety • Increased personal security • Reduced number and severity (cost) of: – P.D. accidents – P.I. accidents – Vehicle thefts • Reduced fatalities
4. Environment and Energy (physical impacts) 4.1 Environment • Reduced vehicle emissions • Reduced noise pollution • Reduced right-of-way requirements • Neighborhood traffic intrusiveness (affecting community acceptance) 4.2 Energy • Reduced fuel consumption
5. Implementation 5.1 Ease of Implementation/Deployment • Technical feasibility (including standards issues) • Regulatory support • Revenue and financial feasibility • Equity impacts • Privacy impacts • Availability of staffing/skills • O&M requirements 5.2 Agency Cooperation/Coordination • Increased sharing of incident/congestion information • Reduced information-gathering costs • Increased coordination/integration of network operation, management and investment • Agency commitment to ITS system 5.3 Technology Flexibility • Ability to evolve with changes in system performance requirements and technology

Travtek demonstration in Orlando was intended to minimize the time spent lost in a strange city. In addition, the information the system provided was likely to encourage tourists to visit more attractions and increase the entertainment value of their vacations. Aggregate VMT and time spent traveling might increase, but mobility and user benefits would increase even more. It is reasonable to conclude that the user benefits of ITS will be much greater than those resulting from reductions (if any) in aggregate travel time and delay.

In the more general case, travel decisions involve a series of tradeoffs between the times and costs of travel on all available alternatives and the benefits of travel from engaging in activities at the trip ends. Without adding capacity, the information from ITS will increase the informed nature of these tradeoffs and all of the adjustments people make to minimize their cost of travel (e.g., to avoid congestion) and maximize their benefits from travel. For example, with reliable travel time information, travelers for whom the benefits of certain trips are small may choose to travel shorter distances, change modes, or forgo or defer trips when congestion is heavy. Others may choose to travel to higher-value destinations that are farther away or make more frequent trips with the confidence that they will not be caught in heavy congestion. Trip end information on the availability (in real time) of goods and services at specific prices and locations (e.g., stores) would eliminate searches involving travel to obtain the same information. ITS systems likely will result in higher-value use of personal time and resources for work and leisure activities and more productive use of commercial and industrial resources. The net increase in user travel benefits may be substantial, yet aggregate observable reductions in VMT and travel time are not likely to reflect these benefits. In fact, the aggregate reductions in VMT and travel time are likely to be small.

This means that estimates of the mobility and other personal and corporate economic benefits of ITS that are based on aggregate observable flow volumes and travel times on the network are likely to seriously underestimate these benefits. Instead, it is necessary to measure the demand-side benefits of ITS at the individual tripmaker and firm level, rather than base them on aggregate measures of flow volumes and travel times on the network. This is the reason it is necessary to separate the demand-side criteria, measurable at the individual (disaggregate) level (Criterion 2.0 in Table 1), from the more familiar aggregate observable supply-side efficiency criteria (Criterion 1.0) in evaluating ITS plans and operational field tests.

WHY SEPARATE ITS IMPACTS BY TIME FRAME OF OCCURRENCE?

As noted earlier, the proposed evaluation structure deals explicitly with the time frame of the impacts. On the supply side (Criterion 1.0), certain impacts occur quickly, such as increased vehicular speeds and throughput, reduced commercial vehicle operating and maintenance costs, and improved transit operating characteristics (travel times and delays). On the other hand, possible cost savings from reduced infrastructure construction can take a long time to occur. It is also not proper to count these construction cost savings as benefits, because if the money actually were spent and new facilities were constructed, the new capacity would provide additional benefits over and above what the original ITS investment provided. Therefore, the original operational benefits of ITS that led to the long-run infrastructure cost savings are really the relevant benefits to use in the evaluation. Counting the infrastructure cost savings double counts the benefits that give rise to the cost savings (and also ignores other benefits of the additional expenditure).

The most important reason for organizing the criteria by the time frame of their occurrence is to highlight certain fundamental correlations between the criteria and to simplify the evaluation process. In most cases, the impacts that occur over various lengths of time are responses or adjustments to the same underlying benefits of ITS. For example, with regard to the demand impacts (Criterion 2.0), there is a well-known hierarchy of short-run (travel) to long-run (land use) behavioral responses to transportation system changes for which separate forecasting models and relationships are (mistakenly) used to evaluate impacts (*3*). (Also, these models have been developed only to forecast the short- and long-run demand consequences of transportation capacity increases, not the information infrastructure that sets ITS apart from conventional transportation improvements.) In any event, it is important to understand that the longer-run behavioral responses to transportation system changes, namely the land use and productivity/economic development (Criteria 2.3 and 2.2, respectively) impacts, usually involve double-counted short-run mobility (Criterion 2.1) benefits to both passengers and shippers. Why this is so is discussed in the following paragraphs.

The common origin of both the short-run (travel) and long-run (land use) behavioral responses to ITS improvements is shown in Figure 2 (*4*), the chain of causality in a model of individual behavior that incorporates ITS information on activity and travel opportunities. The individual utilizes information on opportunities to engage in activities at various locations, some or all of which may involve travel. The individual also has needs: to work, shop, play, be safe, and have a home. These determine how the individual chooses from among the various activity opportunities. The individual also has resources (e.g., time and money) that affect his or her response to opportunities to travel and engage in activities at various places and prices.

Figure 2 highlights the lack of a direct causal relationship between land use and travel. Both impacts stem from a third variable, namely individuals responding to information on opportunities, needs, and resources to "consume" both land and travel. Empirically, the presence of third variables driving both land use and travel has been amply demonstrated; individuals consume both more land and more travel as their income increases (*5*). Understanding that ITS information has the potential to affect both allows one to better assess these impacts in this evaluation.

In general, land value increases from increased transportation accessibility are considered to be the capitalized stream of short-run user travel benefits from the improvement. Even if ITS leads to longer trip lengths and "sprawl," one can assume that the benefits from trip length increases resulting from higher-value residential locations and other activities at the trip ends are equal ("at the margin") to the added travel time/cost of these longer trips. Therefore, it is possible to avoid making value judgments on the worth of various land use distributions and use the results of individual choice behavior first to determine the numbers of trips made with each length and mode and second to value these trips at the values used by the individuals in making their travel decisions. This means that the travel time and cost impacts can be valued at the travel time versus out-of-pocket travel cost "utility" values of travelers. These utility function values of time are fairly well researched in urban travel demand forecasting. For example, for daily trips to work,

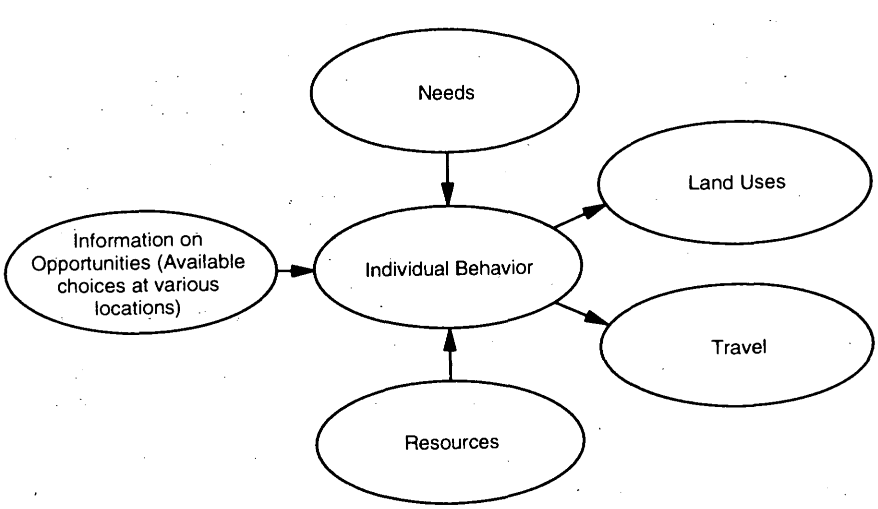


FIGURE 2 Paradigm of individual behavior incorporating ITS information.

travel time is valued on average at approximately 40 percent of the wage rate. One can use this short-run travel time value to value both the short-run travel and the long-run land use distribution impacts of ITS investments.

In the long run, the behavioral response of travelers and shippers to ITS should substantially benefit the "efficiency" of the land use or settlement patterns. Without ITS, the automobile-highway system is a classic example of a system governed by individual choice that puts private interests over the public interest. Every time a person drives his or her car onto a congested roadway, far more aggregate delay is imposed on others—on the system—than on the driver. In economic terms, the marginal private cost of highway travel is much lower than the marginal social cost of travel on an already congested highway system. In fact, the more congested the highway corridor, the greater the difference between the marginal social and private costs of making a trip by automobile (*6*).

Congestion is also the price the current transportation system imposes on its users as a result of individual private decisions to locate on larger plots of land, farther away from work and shopping. And as increasing amounts of money are spent on housing, the transportation price that individual lifestyle decisions impose on everyone else is not known by the individual when he or she makes those decisions. Investments are made by individuals in expensive housing without consideration of the total cost of their location decisions on society. These decisions lead to real inefficiencies: the system has lost its ability to confront consumers with the real costs of their decisions. This is as true in the long run for land use location decisions that generate congestion as it is in the short run for the individual travel decisions described in the previous paragraph.

ITS systems can lead to more efficient lifestyles in the sense that it should be possible to provide more accurate information on the real costs of travel and land use location decisions with than without ITS systems. More accurate information will lead to a more predictable travel environment in which travel costs are internalized before the fact to influence travel decisions, rather than after the fact when the traveler who is stuck in traffic cannot do anything about it. Once again, however, to the extent these impacts are internalized to the travelers, they can be calculated as lower travel (user) costs, rather than higher-value land use location impacts of ITS that are separate from the travel costs.

These examples are intended to show the prevalence of double counting between the short-run and long(er)-run criteria under the first two categories of impacts (criteria 1.0 and 2.0) in Table 1. Criteria that are highly correlated in their outcomes are not helpful in choosing between alternatives. However, although double counting should be avoided, different decision makers may value (weight) differently different manifestations of the same impacts. Ultimately these decision makers also must decide which manifestations of the same benefits represent real value added to society and which can be combined or eliminated in the evaluation. Therefore, all criteria are included in Table 1, subject to their being grouped together to facilitate the evaluation process as described below.

SAFETY, ENVIRONMENTAL, AND ENERGY CRITERIA

The safety and environmental and energy criteria (criteria 3.0 and 4.0 in Table 1) are related to the aggregate flow volumes and conditions on the network. Normally, in the evaluation of conventional transportation improvements, these "flow-produced physical impacts" are directly related to user benefits because increased user benefits from a transportation improvement lead to more travel, which then gives rise to more physical impacts of this travel. This direct relationship for conventional transportation improvements does not hold for ITS improvements, as explained earlier.

Therefore, the flow-produced physical impacts (Criteria 3.0 and 4.0) of ITS should be related to the flow volumes and operating characteristics included in the efficiency group (Criterion 1.0) rather than the measures included in the demand group (Criterion 2.0). These events produce the chain of causality for forecasting these ITS impacts, which is shown in Figure 3 (*2*).

IMPLEMENTATION IMPACTS

Implementation impacts (Criterion 5.0) include the following:

- 5.1 Ease of implementation;
- 5.2 Agency cooperation/coordination; and
- 5.3 Technology flexibility.

These implementation impacts are most readily described by the detailed bulleted criteria under each of these headings in Table 1.

COMPARISON OF CRITERIA WITH GOALS IN 1992 U.S. DEPARTMENT OF TRANSPORTATION ITS STRATEGIC PLAN

The six national goals in the 1992 U.S. Department of Transportation ITS Strategic Plan (*7*) map very well on the criteria structure in Table 1. The six national goals (with their corresponding criterion numbers from Table 1) are as follows:

1. "Improve the safety of surface transportation" (Criterion 3.0);
2. "Increase the capacity and operational efficiency of the surface transportation system" (Criterion 1.0);
3. "Enhance personal mobility and the convenience and comfort of the surface transportation system" (Criterion 2.0);
4. "Reduce the environmental and energy impacts of surface transportation" (Criterion 4.0);
5. "Enhance the present and future productivity of individuals, organizations, and the economy as a whole" (Criterion 2.2); and

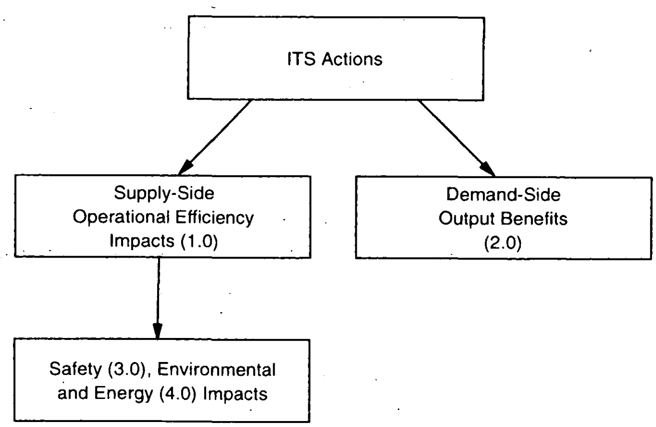


FIGURE 3 Relationship between ITS impacts.

6. "Create an environment in which the development and deployment of ITS can flourish."

The first four national goals are the slightly reordered 1.0 through 4.0 in Table 1. The reordering in this paper is methodologically based on the need to evaluate first the aggregate flow and efficiency impacts (1.0) of ITS options, which are then used to assess the safety consequences (Criterion 3.0) of ITS options (see Figure 3).

The fifth national goal (enhance productivity) is the medium-run (Criterion 2.2) set of demand adjustments to increase output in Table 1. The criteria organization facilitates producing information and helps avoid double counting during the evaluation process. Understanding the differences between the national goals and the structure in Table 1 can advance the state of the art of ITS evaluation.

Finally, the last national goal (create a U.S. ITS industry) is legitimate at the national level but not at the local level, except as it is included in the economic development/productivity criterion (2.2). Creating a technology for its own sake should not be relevant at the local level in developing an ITS plan.

TREATMENT OF AFFECTED GROUPS

The criteria in Table 1 generally are applicable at the urban, rural, and intercity levels, as well as to many groups in society. Table 2 shows the headings from the benefits "taxonomy" in the Mobility 2000 Benefits Report (*8*). In the example evaluations shown in this paper, the impacts on particular groups can be considered in the grouping of criteria for evaluation discussed in the next section or in the weighting of the criteria, which can vary by location. The location strata used in the evaluation process highlight the varying importance of various groups in the different locations. Projects can be included in an ITS plan that provide benefits to rural and intercity travelers as well as to various affected groups in large, midsized, and smaller urban areas. It is important in any evaluation to highlight the tradeoffs that decision makers need to make with the information provided in an ITS evaluation process (*9*).

GROUPING CRITERIA FOR EVALUATION

The considerable uncertainty in forecasting the impacts of ITS projects requires sensible impact assessments (forecasts), subject to strong face validity checks, rather than false precision. The full set of criteria in Table 1 is structured to highlight the inherent correlations of the criteria. This allows grouping the criteria to assess candidate ITS projects using greatly simplified evaluation matrices that can be developed and used by decision makers.

It is strongly recommended that grouping the evaluation criteria and carrying out the evaluation of relative project merit (described in the next section) be done by a diverse group of the highest-level decision makers who can be assembled for at least a half-day process. Ideally, agency heads from all the modal transportation agencies in the study area should be involved, in addition to planning agency and citizen and environmental group representatives. A facilitator who is familiar with the entire process keeps the entire group as focused and productive as possible in a high-energy process, limited only by the time and attention span of the often-nontechnical attendees.

Table 3 shows an example list of grouped benefit-related evaluation criteria. The grouped criteria reflect the transportation program goals developed for a particular region. By using the same numbering system as that in Table 1, Table 3 shows which Table 1 criteria were used, either singly or grouped. Table 4 provides an easier-to-use list of the grouped evaluation criteria shown in Table 3.

The grouped evaluation criteria shown in Tables 3 and 4 are not meant to be all inclusive but rather to reflect an example set of objectives and understandings of decision makers. For example, "land use" (Criterion 2.3 in Table 1) as an evaluation criterion is missing from the example list. One reason this may happen is that it is not clear what the land use impacts of most ITS projects will be in the next decade or so. Another important reason is that many (but not all) decision makers are reluctant to take a position on which parts of a region should grow "at the expense" of land values in other parts of a region. Rather, the principal objective of many decision makers is to promote the mobility of the population by the information or travel options that ITS can provide to residents and firms. Any land use impacts will be highly correlated with the ITS mobility impacts and will likely be independent of VMT changes caused by ITS. To the extent that decision makers want to single out particular groups living in specific locations for special consideration in the evaluation, this can be done by defining separate criteria for those groups or even separate locational strata (see Figure 1). For example, in Table 3 the personal mobility (Criterion 2.1.1) of residents and nonresidents is selected as separate criteria in the evaluation.

Table 5 shows that there are important relationships between the criteria in Tables 3 and 4. Consistent with Figure 3, environment and safety are related to the primary supply-side operational efficiency impacts [congestion and Single-occupancy-vehicle (SOV) reduction], whereas the demand-side mobility affects drive economic development, which is a longer-run consequence of the same ITS user benefit. Conversely, as shown in Table 5, an improvement that is primarily safety oriented can reduce incidents, which in turn can reduce congestion as a secondary impact.

TABLE 2 Potentially Affected Groups

Categories of Benefit	Impacted Groups					
	Users (Groups)	General Population		Organizations		
	Urban, Rural, Elderly, Suburban, Commuter, etc.	Other Transportation System Users	Nonusers	Public Sector	Private Sector	
				Operators	Operators	Industry

TABLE 3 Example of Grouped Criteria Used to Evaluate ITS Projects

1. Increased Operational Efficiency (supply-side efficiency, meaning more output per unit of input)

    1.1 Short Term: Operational Efficiency

        1.1.1 Infrastructure
        - "Decreased Congestion"

        1.1.2 Vehicle Efficiency ("Increased alternate mode share")

            1.1.2.1 Private Autos
            1.1.2.2 Transit
            - Single Occupant Vehicle (SOV) Reduction

2. Increased Output (demand adjustments that further increase output or benefits from ITS improvements)

    2.1 Short Run: Mobility

        2.1.1 Personal
        - Mobility of Residents
        - Mobility of Nonresidents ("support tourism")

        2.1.2 Freight
        - Mobility of Commercial Vehicles ("facilitate efficient goods movement")

    2.2 Medium Run: Economic Development
    - Economic Development

3. Safety ("Improve Highway Safety")
    - Safety

4. Environment and Energy ("Improve Environment")
    - Environment

5. Implementation
    - Ease of implementation/agency commitment

Note: The table shows the grouping of criteria from Table 1; the example criteria are the bulleted items.

EVALUATION OF RELATIVE PROJECT MERIT

When transportation improvements are under evaluation, weights often are assigned to evaluation criteria. Multiplying the quantitative weights times the (interval) scaled measures of criteria "attainment" allows the evaluator to produce an overall weighted measure of merit of each project.

This stage of the evaluation process shown in Figure 1 can begin by having the decision makers assign weights to a larger number of the Table 1 criteria than those shown in Tables 3 through 5. This can help in deciding on the final criteria shown in the latter tables. The result of the criteria weighting process can be different relative weights of the criteria on a one-to-five scale for each of, say, five geographic areas (strata). Example weights are shown in Table 6, in which safety receives the highest weight (5) in every area, whereas ease of implementation receives a 1 in every area. Safety is always of great importance in transportation, but not all ITS projects affect safety to any great extent. Conversely, ease of implementation is weighted low, but ITS projects vary widely in their outcomes with respect to this criterion. The criterion probably should be weighted low both because it may be seen as a tiebreaker ("We want early successes") and because a project should not be implemented without benefits other than being "easy to do."

Other criteria will vary in their relative weights, depending on their local importance. As indicated in Table 6, reducing congestion is most important in the example major urban area, but congestion reduction, by itself, may not be of the greatest importance any-

TABLE 4 Grouped Criteria Used to Evaluate ITS Projects in Example Study: Simplified Numbering System

1. Congestion
2. SOV Reduction
3. Mobility — Residents
4. Mobility — Nonresidents
5. Mobility — Commercial
6. Economic Development
7. Safety
8. Environment
9. Implementation Ease

TABLE 5 Relationships Between Evaluation Criteria

Primary	Secondary
1. Congestion	7. Safety
2. SOV Reduction	8. Environment
3. Mobility -- Residents 4. Mobility -- Nonresidents 5. Mobility -- Commercial	6. Economic Development
7. Safety	1. Congestion
9. Implementation Ease	

where. SOV reduction may be more important in the major urban area as a measure of increased alternate mode share, which may be a major ITS program goal. Economic development may always be very important, but relatively speaking may be most important in the small urban areas and in rural areas.

Many ITS plans will be developed for only one region, in which case the added complication of different weights and separate impact assessments for different areas (or strata) will not be needed. This can greatly reduce the time and effort needed to be expended by the decision making group in the process.

The next step in the evaluation of *relative* project merit is the impact assessment step. *Consumer Reports*-type measures of impact assessment on a plus or minus five- or ten-point scale can be used for each criterion that is affected by an ITS project. (A 0 can be used for no impact.) The impact assessment portion of the evaluation should be a screening and informing process that allows the decision-making participants to become comfortable with the criteria, the assessment of project impacts in terms of these criteria, and the grouping of projects by relative merit on the basis of their impacts. The actual assignment of impact values for each project is done by the members of the decision-making group with strong input from the technical facilitator who is familiar with the likely impacts of the ITS options. The process has been shown to be an excellent way to educate and obtain the support of decision makers for the conclusions of the evaluation.

Table 7 summarizes example relative merits of a series of example candidate ITS projects applied to several locations, whereas Table 8 provides a more generic example of a project evaluation rating sheet for one area (or stratum). Note that the candidate projects (rows) differ between the two tables, and the criteria (columns) in Table 8 omit two of the criteria in Tables 3 through 6. Table 8 leaves blank spaces for three additional criteria and illustrates that there will be site-specific differences in the selected criteria, criteria weights, candidate ITS projects, and their relative impacts between any two locations (strata). The righthand "weighted-sum" column in Table 8 is filled out by the evaluators for each candidate ITS project. This column illustrates how the weighted measures of merit summarized in Table 7 are calculated (for a different set of ITS projects).

Project cost/effectiveness or cost/merit is not calculated at this stage because projects often can be subdivided into lower-cost projects that still have high relative merit. Also, many candidate projects are interdependent, relying on the same infrastructure (e.g., traffic speed monitoring for ATMS and ATIS). It is not always easy

TABLE 6 Example of Weighting of Evaluation Criteria by Geographic Area (Strata) Within State

	Geographic Area				
	Major Urban Region	Mid-sized Urban Regions		Small Urban	Intercity/ Rural
Criteria		A	B		
1. Congestion	3	2	2	1	1
2. SOV Reduction	4	2	3	1	1
3. Mobility — Residents	3	3	3	4	4
4. Mobility — Nonresidents	1	2	2	2	3
5. Mobility — Commercial	4	3	4	2	4
6. Economic Development	4	4	4	5	5
7. Safety	5	5	5	5	5
8. Environment	4	3	3	2	1
9. Implementation Ease	1	1	1	1	1

TABLE 7 Example of Relative Merit of Candidate Projects

User Service Project	Merit				
	Major Urban Region	Mid-sized Urban Regions		Small Urban	Intercity/ Rural
		A	B		
Public Transit/TDM	50	27	32	17	13
TDM Support	33	26	30	14	-
HOV Priority Support	19	-	-	-	-
Transit Vehicle Management	33	24	26	28	26
Congestion Pricing Support	57	-	-	-	-
Traveler Information					
Trip Planning — Pre-Trip	97	51	71	42	30
Trip Guidance — En-Route	77	53	59	32	43
Traffic Management					
Incident Detection & Management	77	47	51	42	46
Freeway Ramp Metering	73	36	48	26	-
Traffic Control	99	89	97	70	20
Freight and Fleet Management					
Intermodal Port Transfers	25	-	25	-	-
Regulatory Support/Borders	-	-	-	-	37
Hazardous Materials	12	11	11	10	9
Additional Services					
Emergency Service Management	15	16	15	17	14
Enforcement System	21	16	15	15	16
Mayday Test	-	-	-	-	28

to decide which project comes first and should be charged the common infrastructure cost in a calculation of cost per relative merit.

Once candidate projects are grouped by their relative merit within strata, programs of projects can be developed that exhaust the budget limits for each stratum (see Figure 1). Conversely, budget limits can be decided after the absolute worth of the programs of projects has been assessed in a benefit-cost analysis (see next section). In either event, the development of programs should avoid the scenario in which the most beneficial but expensive project uses up the entire program budget for the strata. This problem can be avoided by selecting projects from the highest ranked groups that provide significant benefits themselves and contribute to the infrastructure needed to implement other highly ranked projects (e.g., traffic speed monitoring for ATMS and ATIS).

ASSESS ABSOLUTE WORTH OF ITS PROGRAMS IN BENEFIT-COST ANALYSIS

The final stage of the evaluation process involves a more formal program-level benefit-cost analysis of the selected program of projects for each location or stratum that exhausts the funds available for that location. Costs already may be estimated for all projects reaching this stage of evaluation because of the budget constraint test. On the benefit side, the largest portion of the quantifiable benefits will likely be the combined user benefits from the time savings associated with the first criterion, congestion reduction (Criterion 1), and the mobility increases (Criteria 3 through 5 in Tables 4 through 6). Given the state of the art, it is difficult to quantify the proportion of these benefits that will be aggregate observable time savings on the (multimodal) network (that is, at the equilibrium levels of congestion and travel volume, including induced travel, that result from the ITS projects), and how much will be attributable to mobility increases that are measurable only with appropriate behavioral models at the individual traveler or firm level. Devoting considerable resources to attempting to quantify the value of one versus the other is certainly false precision, given the current state of the art.

Similarly, distinguishing between the longer-run economic development (Criterion 6, Tables 4 through 6) and short-run (travel) impacts of ITS improvements is beyond the current state of the art. The best current (travel) demand models forecast travel directly by mode between origins and destinations as a function of the activity systems at the origins and destinations, and the price and level-of-service conditions by the travel mode and all its substitutes (*10*). These direct demand models are partial equilibrium models that describe how travelers behave so that they will be in equilibrium with the rest of the system. They model the behavior of the tripmaker, who considers all trip end opportunities to be fixed. Direct demand models (estimated with cross-sectional data) are themselves simplifications of "general equilibrium" models that attempt to explain how land use and travel vary simultaneously with transportation improvements (*3*). The demand relationships embedded within them (e.g., the elasticities and values of time) reveal something about both long- and short-run behavior.

Also as noted earlier, even if ITS leads to longer trip lengths and "sprawl," one can conservatively assume that the benefits from trip length increases, resulting from higher-value residential locations and other activities at the trip ends, are equal ("at the margin") to the added travel time/cost of these longer trips. Therefore, the results of appropriate demand models can be used to value congestion reduction and mobility improvement at the values used by the individuals in making their travel decisions (i.e., 40 percent of the wage rate for daily trips to work). For example, an average hourly wage

TABLE 8 Example of IVHS Project Evaluation Rating Sheet

IVHS Function/Project	Decreased Congestion	SOV Reduction	Improved Mobility – Residents		Improved Mobility – Commercial	Economic Development		Safety		Implementation Ease (Institutional)	Weighted Sum
weight											
Traveler Information											
Incident, etc. Advisories – Enroute											
Incident, etc. Advisories – Pre–Trip											
Time/Cost Advisories – Enroute											
Time/Cost Advisories – Pre–Trip											
Route/Mode Guidance – Enroute											
Route/Mode Guidance – Pre–Trip											
Yellow Pages/Services Info.											
Traffic Management											
Incident Management											
Ramp Metering/Access Control											
Street Traffic Control											
Demand Management (Cong. Pricing)											
Construction Management											
Freight and Fleet Management											
HAZMAT Monitoring and Tracking											
Weigh–in–Motion (WIM)											
Regulatory Support											
Public Transport											
Transit Vehicle Monitoring											
Transit Vehicle Priority											
Smart Fare Payment											
Dynamic Ride Sharing											
Safety/Security Systems											
Electronic Toll Collection											
ETTM											

rate of $14.00/hr results in a value of time for work trips of $5.60/hr. The generally accepted figure for nonwork (offpeak) trips is half this value.

In summary, the quantifiable user benefits from ITS projects are bundled in an impossible-to-disentangle ball of short- and long-run measures under Criteria 1 and 3 and 4, through 6. However, until the results of new demand modeling research are available, it is possible to approximate their value without disentangling them. It is possible to value them using the observed aggregate time savings for similar ITS projects (if any) implemented elsewhere, times a multiplier on this lower bound estimate of user benefit to account for the mobility benefits individuals and firms experience from the tradeoffs they make to maximize their net benefits from travel (i.e., their tradeoffs between the times and costs of travel on all available alternatives they are informed about, and the benefits of travel from engaging in activities at their trip ends).

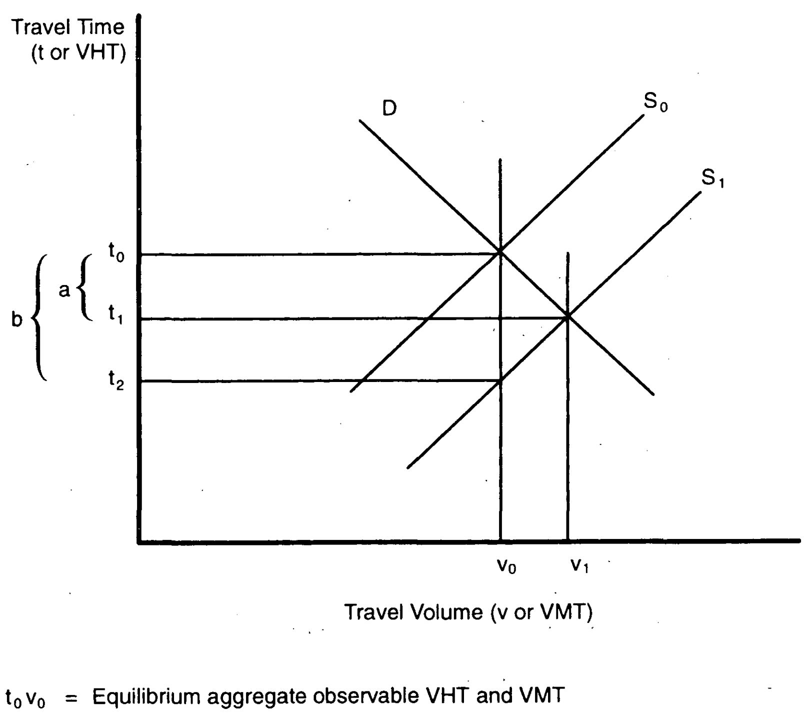
The following multipliers may be applied for projects involving both ATIS (information) and ATMS elements:

Travel Segment	*Multiplier*
Peak period person travel	2.0
Offpeak person travel (including tourists)	1.3
Commercial vehicle travel	1.5

These (default) multipliers should be applied only to projects that include an important element of information to travelers (real-time congestion information to residents and commercial vehicles; static destination and route information to nonresident tourists).

These multipliers were derived using long-run automobile person trip demand elasticities of −0.8 and −1.0 for work and nonwork trips, respectively (*10*), and speed elasticities of −0.75 and −0.375 with respect to peak and offpeak volumes (*11*). Commercial vehicle demand elasticities were assumed (conservatively) to be lower than automobile work trip demand elasticities. The multipliers measure the relationship between the total change in consumer surplus (user benefit) from a transportation improvement and the aggregate observable benefit at the intersection between the supply and demand curves that defines equilibrium flow (including induced travel) after a transportation improvement. In economic terms, the multiplier is the amount by which the entire utility (γ-axis) difference of the consumer surplus rectangle (assuming a fixed-trip table) exceeds the smaller γ-axis utility difference of the consumer surplus triangle relating to induced travel. Figure 4 illustrates the multiplier for the (work trip) situation when the slopes of the supply and demand curves are equal.

It is probable that the multipliers on aggregate observed time savings given are underestimates of the actual total user benefits from ITS improvements. The reason for this is that although the −0.8 automobile person work trip elasticity given is a long-run elasticity incorporating "demand shifts," it was not estimated for transportation improvements involving ITS. The long-run demand



$t_0 v_0$ = Equilibrium aggregate observable VHT and VMT before ITS improvement

$t_1 v_1$ = Equilibrium aggregate observable VHT and VMT after ITS improvement

t_2 = Disutility of travel after ITS improvement measured at the individual traveler level (assuming no net change in travel volume)(2).

b = 2a when slopes of Demand and Supply curves are equal

FIGURE 4 Illustrative derivation of user benefit multiplier.

elasticity for transportation improvements involving ITS is likely to be larger than −0.8 because the number of known alternatives is larger and the possible substitutions are greater. This will result in a flatter demand curve, more induced travel, less aggregate observable congestion relief, and a larger multiplier. In Figure 4, these results may be diagrammed as a flatter demand curve, *D*, rotating counterclockwise around t_0v_0, causing the new observed aggregate equilibrium point t_1v_1 to be higher on the supply curve S_1 than before. The result is that *a* in Figure 4 is a smaller fraction of *b,* and the multiplier increases, possibly substantially. This is why, as stated earlier, aggregate observable reductions in congestion (travel time) from ITS improvements are so likely to be dramatic underestimates of ITS user benefits.

Whatever multipliers are used, the estimates of user benefits should be made separately for peak and offpeak (including tourist) passenger and all commercial (freight) movements and applied to the volumes of travel affected by the ITS improvements. For peak and offpeak passenger trips, the total user benefit can be valued at 40 and 20 percent of the real hourly wage rate, respectively, at the forecast year(s) for which the ITS improvements are being evaluated. In addition, because the observable changes in aggregate vehicle miles traveled (VMT) are only part of the user benefits, the separate calculation of passenger vehicle operating cost savings resulting from VMT changes will be small and can be ignored in most circumstances.

For large trucks, $60/vehicle-hr can be used as the value of time saved. This includes labor, vehicle operating and maintenance and depreciation for an 18-wheeler carrying 20 tons of cargo, plus an economic development impact (Criterion 6) of $10/hr. A fee of $10/hr is an inventory carrying charge calculated on the assumption that manufactured goods are worth an average of $5.00/lb, a 10 percent annual interest rate (i.e., 20 tons × 2,000 lb/ton × $5/lb × 0.10 ÷ 3,000 hr/year trucking time = $6.67/hr) and a 50 percent premium to account for the value of travel time reliability for "just-in-time" delivery.

The smaller portion of the quantifiable benefits for calculating the absolute worth of candidate ITS programs is composed of the savings in the social costs of travel attributable to the safety, environmental, and energy benefits that accompany the aggregate observable reductions in VMT on the network. As shown in Table 5, these "flow-produced physical impacts" are related to the aggregate observable flow volumes and operating characteristics included in the efficiency group of criteria (operating speed and SOV reduction), rather than to the measures included in the demand (mobility) group. To quantify them, the first portion of the user benefit estimate described above is used (i.e., without the multiplier), namely the VMT reduction observed for similar ITS projects (if any) implemented elsewhere. More specifically, only the portion of the VMT reduction caused by SOV reduction should be used for quantifying the environmental impacts. This is because the environmental impacts are caused primarily by SOV reductions, whereas the energy and safety improvements may be estimated on the basis of both impacts (see Table 5).

The safety, environmental, and energy benefits to society may be valued by using the following (default) values per unit VMT saved:

Measure (reference)	*Value per VMT ($)*
Safety (*12*)	0.022
Air pollution (*12*)	0.03
Noise pollution (*12*)	0.0035
Energy (*13*)	0.0025

As an approximation, these unit values per VMT should be applied only to the SOV volume reductions achieved as a result of the ITS improvement. These savings apply only to aggregate reductions in observed (or observable) trips caused by the ITS improvement. In the case of air pollution, this is because most air pollution is a result of vehicle trips rather than trip length. In the case of safety, this is because of the difficulty of relating accident rates to vehicle flow rates on a given type of roadway. In the case of energy consumption and noise pollution, the very low dollar values of these impacts makes the added precision of calculating these impacts as a function of vehicle speeds not worthwhile.

CONCLUSION

ITS plans and operational tests should be developed and evaluated in a way that is sensitive to the differences between ITS and conventional transportation improvements. By recognizing the separate supply (efficiency) and demand (increased output) impacts of ITS, it is possible to avoid dramatically underestimating the benefits of the new technology. The methodology and criteria presented in this paper provide the required structure and default values to evaluate ITS improvements for inclusion in transportation system plans. The criteria and default values provided in the paper also highlight where research and operational tests can provide improved values and information that will most quickly advance the state of the art of ITS evaluation.

ACKNOWLEDGMENTS

The author is grateful to Jack Kay, Don Samdahl, and Lou Neudorff of JHK & Associates, Les Jacobson and Peter Briglia of Washington State DOT, and Steve Pepin of the Massachusetts Highway Department for their helpful suggestions during the course of the ITS system planning work during which the ideas in this paper were developed.

REFERENCES

1. JHK & Associates; Charles River Associates; et al. *ITS Strategic Plan for Washington State: VENTURE Washington.* Washington State DOT, Nov. 1993.
2. Brand, D. An ITS Benefits Assessment Modelling Framework. In *Transportation Research Record 1408,* TRB, National Research Council, Washington, D.C., 1993.
3. Brand, D. Theory and Method in Land Use and Travel Forecasting. In *Transportation Research Record 422,* TRB, National Research Council, Washington, D.C., 1973.
4. Brand, D. Research Needs for Analyzing the Impacts of Transportation Options on Urban Form and the Environment. In *Special Report 231: Transportation, Urban Form, and the Environment,* TRB, National Research Council, Washington, D.C., 1991, pp. 103, 107.
5. Deakin, E. A. Jobs, Housing and Transportation: Theory and Evidence on Interactions Between Land Use and Transportation. In *Special Report 231: Transportation, Urban Form, and the Environment,* TRB, National Research Council, Washington, D.C., 1991, pp. 25–42.
6. De Corla-Souza, P. Suburban/Urban Transportation Investment—What Will It Cost and Who Will Pay. FHWA, U.S. Department of Transportation, 1990.
7. *ITS Strategic Plan.* U.S. Department of Transportation, 1992, pp. 17–19.
8. Mobility 2000. *Benefits Report,* March 1990, Dallas.

9. Foster, E. Cost-Benefit Analysis Applied to ITS. University of Minnesota, Dec. 1993.
10. Domencich, T. A., G. Kraft, and J-P. Valette. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. In *Highway Research Record 238,* HRB, National Research Council, Washington, D.C., 1968.
11. Loudon, W. R., and D. A. Dagang. Predicting the Impact of Transportation Control Measures on Travel Behavior and Pollutant Emissions. Presented at 71st Annual Meeting of The Transportation Research Board, Washington, D.C., 1992.
12. Charles River Associates. *Study of the Benefits of the Diversion To Transit of a Peak-Period Auto Commuting Trip to the CBD of a Typical Large U.S. City.* FTA, U.S. Department of Transportation, n.d.
13. *The External Costs of Oil Used in Transportation.* Congressional Research Service, Environment and Natural Resources Policy Division, June 1992.

Publication of this paper sponsored by Committee on Intelligent Vehicle Highway Systems.

Framework for Evaluating Intelligent Vehicle-Highway Systems

STEVEN E. UNDERWOOD AND STEPHEN G. GEHRING

Evaluation of intelligent vehicle-highways systems (IVHS) is a relatively new activity. An evaluation framework useful for those interested in IVHS is presented in which each component of IVHS evaluation offers new challenges. These challenges involve public and private benefits, new product functions, market penetration, abundant data generated from new systems, human interaction, and multisite deployments. Although many methodologies for evaluation exist, some significant challenges need to be addressed to properly evaluate IVHS.

The effects of intelligent vehicle-highway systems (IVHS) are diverse, ranging from the personal security provided by in-vehicle communication systems to improvements in public health provided through more efficient dispatch of emergency vehicles. Although some benefits are available to the public, such as shorter delays at incidents through the use of highway advisory radio, others are enjoyed by those who buy a service, such as the routing improvements of autonomous navigation.

The variety of systems subsumed under the umbrella of IVHS no doubt contributes to this diversity. IVHS includes an ever-increasing list of systems: motorist information systems, route guidance systems, in-vehicle navigation systems, collision warning and avoidance systems, vehicle control and platooning systems, traffic control systems, traffic-monitoring systems, automated toll and vehicle identification systems, and commercial vehicle location systems, to name a few. Each of the foregoing systems has a unique set of impacts and evaluation requirements that makes IVHS evaluation a challenging prospect.

FRAMEWORK FOR BENEFITS ASSESSMENT

The potential for IVHS is uncertain at this time, and decisions on whether to support the development of various systems require data and careful analysis. This is where evaluation helps. The purpose of operational field tests is to provide the participating organizations with a realistic setting in which they can assess the potential benefits of IVHS without assuming the risks of full deployment. Models and other analytical approaches can supplement field tests to assess the prospects of alternative deployment scenarios. The evaluation should address the interests of the general public (with measures such as congestion mitigation, enhanced safety, more efficient travel, etc.) and the traveler (through easier and more efficient travel), as well as manufacturers and suppliers (through market potential, enhanced transportation of goods, etc.). In this sense evaluation is a decision support tool that provides information about the potential benefits and risks of system development. Figure 1 indicates which evaluation methodologies are most appropriate for a particular evaluation component.

Evaluation is a tool to aid decision making. An impact is the product of the interaction between IVHS and society. Direct impacts are those effects directly attributable to IVHS; higher-order impacts are the products of direct effects. A benefit is an advantage, privilege, or cost, or all of these, that accrues to the traveler through use of IVHS compared with use of the system without IVHS. These benefits are often evaluated as a decrease in user cost.

Figure 2 presents four types of benefits that are useful in the evaluation of IVHS. The graph plots the average trip time in minutes under three scenarios: (*a*) the current technology baseline, (*b*) with IVHS deployed, and (*c*) under optimal conditions. Potential benefits are the difference between the optimal average travel time and the current technology baseline. Expected benefits are the difference between the implementation of IVHS and the current technology baseline. Current benefits are those that are experienced at the current time. Future benefits are projected on the basis of demographic trends and forecast diffusion for some time in the future.

Figure 3 enhances Figure 2 by showing several different types of evaluation activities. Current benefits may be measured empirically through evaluation of operational field tests and other direct approaches. Comparative evaluations may be deployed in the field to determine which of several systems is the most cost-effective. Prescriptive evaluation takes a more formal mathematical approach.

Figure 4 presents the various types of evaluation models that are available. The optimization model compares two points in time along the optimal condition curve. The forecast model can be used to compare the various conditions at a future point in time on the basis of demographic, market, and land use models.

CHALLENGES OF IVHS EVALUATION

Evaluation methods are not new with the arrival of IVHS; they have been around for some time. Therefore, a question one might ask is, What is special about IVHS evaluation that requires the development of new methods? For the last few months that is the question that has been raised by those who work with IVHS. In most cases the answer has been there is little difference between IVHS and other evaluation applications, with a few exceptions. The conclusion of this study is that in most cases it is sufficient to apply existing evaluation methods to assess the benefits and costs of IVHS. Only in those cases where there is an exception to standard evaluation conditions is there a reason to consider the development of new methods. This section focuses on those unique characteristics of IVHS that generate a special need for methodological development. The development of new approaches, methods, and tools are needed for the reasons discussed next.

Program in Intelligent Vehicle Highway Systems, University of Michigan, Ann Arbor, Mich. 48109.

	A System Technical Performance	B Safe and Efficient User Interface	C Indiv. Driver Behavior & Impact	D User Attitudes and Preferences	E Traffic Behavior and Impacts	F Environment and Energy Impacts	G Institutional & Socio-eco-nomic
1 Field exper-iment	●	●	◓	◓	◒	◒	○
2 Field quasi-experiemt	◓	◓	●	◓	●	●	○
3 Surveys	○	◒	◒	●	○	○	◓
4 Subject de-briefing	◒	◒	◒	●	○	○	◒
5 Cost anal-ysis	●	○	◓	○	●	●	○
6 Extrap. im-pact as-sess.	○	○	●	○	●	●	○
7 Narrative case stud-ies	◓	◒	◒	◒	◒	◒	●

¶

Most appropriate	Often appropriate	Occasionally appropriate	Least appropriate
●	◓	◒	○

FIGURE 1 Evaluation methodologies most appropriate for evaluation components.

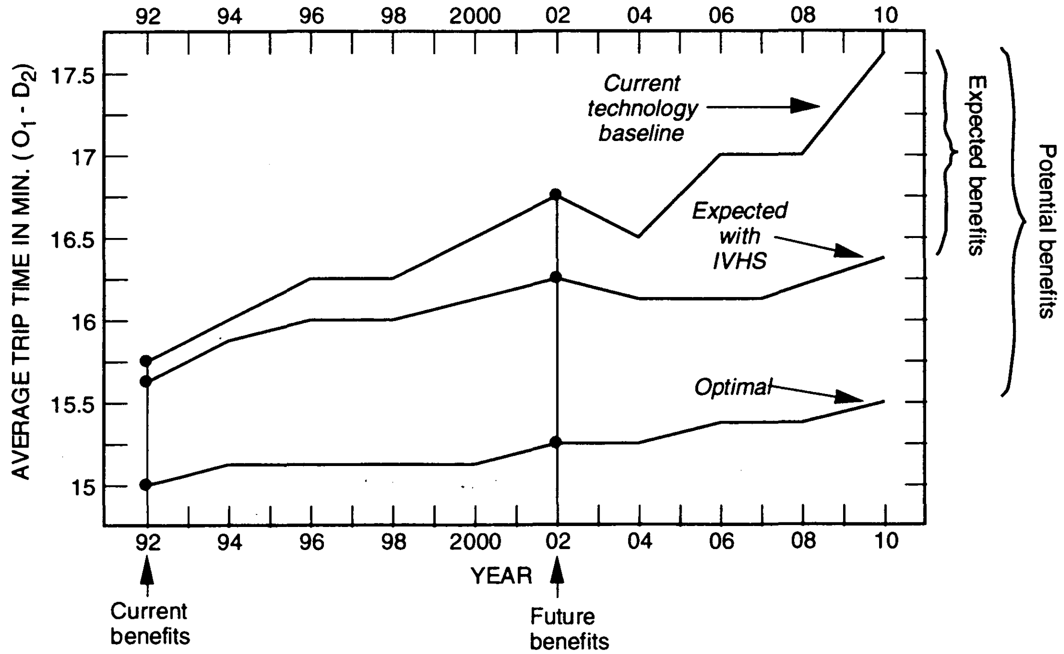


FIGURE 2 Types of benefits.

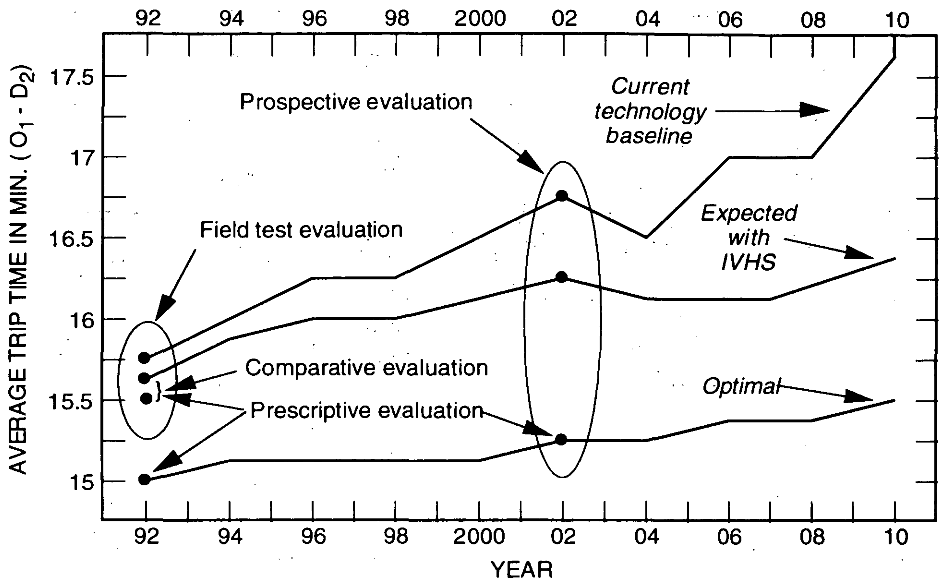


FIGURE 3 Types of evaluation.

Public and Private Benefits

Most IVHS provide both private and public benefits. It is useful at this point to reflect on why one should evaluate the potential benefits of IVHS in the first place. A number of responsible people advocate a Darwinian approach to IVHS deployment in which the survival of the fittest IVHS will ensure that the user gets the best product and service possible. Although this approach may be effective in a perfectly competitive market, the mix of public and private benefits expected from IVHS renders the "natural" selection process less than completely effective. This competitive process may be appropriate in a perfectly competitive market. However, the mixture of public and private benefits requires a more formal and deliberate selection process.

Public decisions will be made about whether to continue to support the development of IVHS or to support some other worthwhile public program. In this type of arena IVHS will have to be justified on the expected benefit to the public. For example, an in-vehicle dynamic route guidance system may guide the vehicle owner to the shortest time path through the network. The time saved is a private benefit. Diverting this vehicle away from the congested area, however, will also reduce the area's congestion by a single vehicle. Therefore, the vehicles without the route guidance systems also benefit through a marginal reduction in congestion. The public benefit is small in this single-vehicle example. However, when 5 or 10 percent of the drivers adopt route guidance, the public benefits may become significant. Within this context, there is a need for methods that discern both public and private benefits (and drawbacks) that integrate these benefits into a logical consistent manner.

For example, improvements in traffic management and safety are by-products of driver information and control systems; they benefit the general public as well as consumers who buy the information and control products for their vehicles. If left to market forces alone, IVHS services may be underproduced because there is no market for IVHS externalities.

New Products and Functions

Most IVHS functions are new and unfamiliar to potential users. Procedures to educate potential users may be required before preference, perception, and performance measures are of any value. For example, 20 years ago it would have been difficult for most people to say with any degree of confidence how frequently they would use an automatic teller machine, a video telephone, a video cassette recorder, or even a notebook computer. It would have been difficult for them to say what they would be willing to pay for the new functionality provided by these products. If they can use the system for a while under natural conditions, then their responses may have more value. Similarly, potential users cannot be expected to operate the systems adequately without having time to learn the system. All operational performance measures need to include time of exposure measures to account for learning effects.

A serious drawback of allowing potential customers to use prototypes is that the devices and user interfaces may have idiosyncratic characteristics that are not of design intent and are not essential to the system's design. Nevertheless these characteristics are a source of distraction to the user. For example, if a prototype system is not as reliable as it should be, the use of the prototype may negatively bias the consumer's reaction. It is critical in all user acceptance and performance studies of new product functions to weigh this trade-off between the positive and negative sides of user exposure.

Systems Management Evaluation

Subtle system improvements require detailed models and sensitive measures. Most IVHS, like motorist information systems and traffic control systems, are likely to produce subtle improvements in the flow of network traffic. For example, once congestion builds to a critical point on the primary highway system, some of the traffic may be diverted to less congested secondary roads. This method

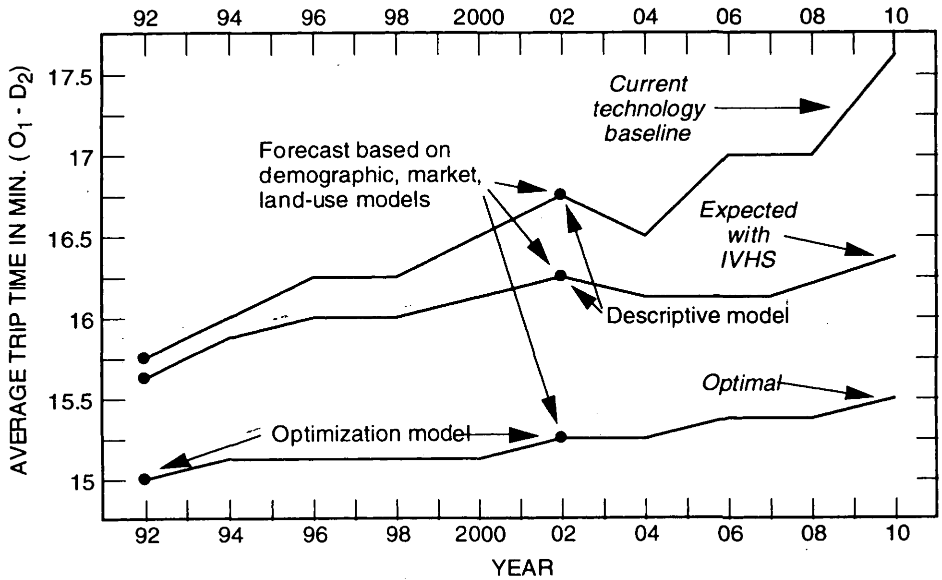


FIGURE 4 Types of evaluation models.

should improve travel conditions for both diverted and nondiverted drivers. The common notion is that travel demand varies by the time of day; therefore, when traffic builds it makes sense to spread the traffic over a larger geographic area to stem the growth of recurrent and nonrecurrent congestion.

Most existing traffic models were developed to evaluate road construction or traffic control options that are not so subtle, and the traffic modelers could make do with assumptions of static demand, static equilibrium assignment, optimal routing of all vehicles, subnetwork optimization, separation of freeways and surface streets, and macroscopic representation of vehicles. Also, most existing models do not have the ability to reflect IVHS functions such as various routing schemes, adaptive traffic control, driver-routing and departure heuristics, and the like. In other words, these existing models do not adequately represent time dependency and other essential details of IVHS and therefore are inadequate for IVHS evaluation.

In some cases the more detailed existing models may be upgraded, and every effort must be taken to make the most of what currently exists. In many cases however, existing models will have to yield to new models that incorporate these required features as fundamental elements of their structure. Therefore, new models are needed that incorporate both the functionalities of IVHS and their likely impacts on traffic. Furthermore, operational field tests must employ measures tuned to the benefits and costs expected from each of the respective systems. The field tests also should be designed as an opportunity to collect data to validate and calibrate the new models.

Market Penetration Effects

The benefits of IVHS should increase with increases in market penetration. The relationship between market penetration and benefits, however, may not always be linear. For example, imagine that the United States adopts a standard for radio broadcast data systems (RBDS), and the proportion of vehicles that come off the production line with FM sideband-capable radios increases steadily over a 10-year period. Also, assume that all vehicles equipped with the system receive the same filtered messages at any particular location. The early users of the system report significant improvements in their ability to avoid congestion and other traffic-related problems. As the proportion of vehicles equipped with RBDS continues to increase, however, the marginal benefits to individual drivers may drop because other drivers with RBDS are congesting the alternative routes. Furthermore, those drivers without systems seem to benefit more because there is less traffic on the original route. As full market penetration is approached, assuming that drivers become aware of these effects, it becomes harder to predict driver behavior.

This short example demonstrates two things. First, it may be difficult to predict what types of impacts that increases in market penetration will produce. Second, it may be inappropriate to extrapolate from operational field tests where there is a minuscule deployment of equipped vehicles to high levels of market penetration.

Abundant Data

Advanced traffic-monitoring systems may provide an abundance of data that can be used in benefits assessment. Probe vehicles may provide dynamic travel times, whereas wide area traffic detection will provide data on congestion, incidents, speeds, and queue lengths. As a result, the traditional transportation planning and evaluation models that were designed to function with sparse data may give way to newer models that take advantage of the new information inherent to the technology. The development of methods for sifting through the data is a priority at this time.

Human Factor

The effectiveness of advanced traveler information systems depends on how the traveler uses the information. One cannot assume

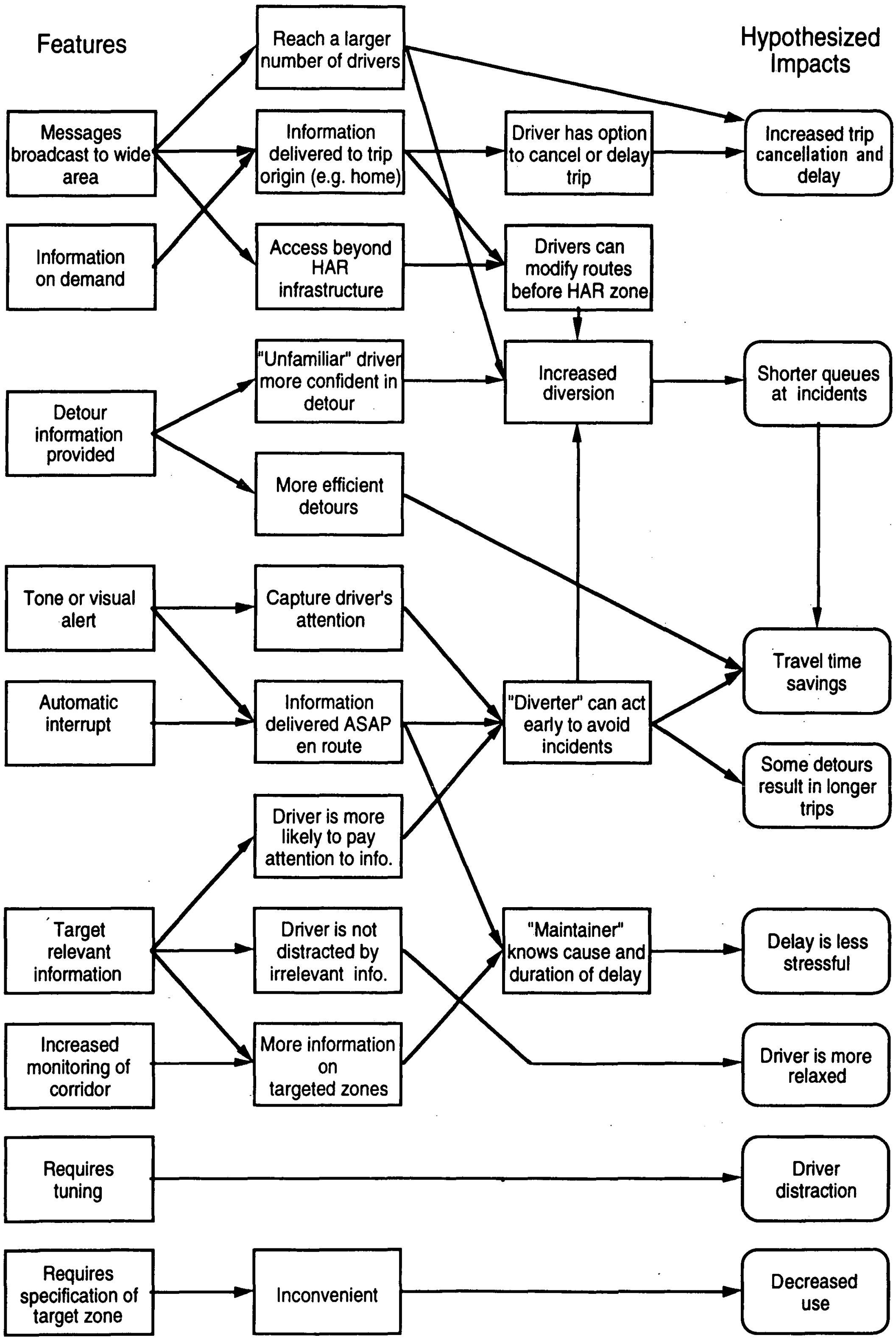


FIGURE 5 Features and hypothesized impacts of motorist information system.

that, because a driver has access to turn-by-turn route guidance instructions, while on a trip he or she will comply with these instructions. In a similar sense, one does not know how drivers will respond to traffic reports coming over RBDS or cellular telephones. Furthermore, benefits from the implementation of IVHS may be overshadowed by resulting induced travel demand behavior.

The driver's actual behavior is influenced by a host of mediating factors, including previous experience, knowledge of the area, and what other drivers are doing. There may be some benefits that do not require action. For example, a driver that chooses not to divert in response to a report may be satisfied to wait in congestion knowing what the options are. To predict the benefits one must know how travelers are likely to use the information and how they are likely to benefit, both in objective and subjective terms. Figure 5 shows the relationships between system features and expected benefits for a motorist information system. Figure 6 shows the variables that mediate the relationship between the system implementation and the level of satisfaction experienced by the user. These illustrate the importance of the human factor in evaluating the impact of IVHS.

The point of this figure is that the path from the information system box to the driver attitude box is indirect and mediated by numerous confounding variables. To really understand the relationship between implementing an information system and the satisfaction of the user requires a complex understanding of how the system works and how the driver is going to use the system. The only way to sort this out is to understand the process and to control for possible confounding variables.

Institutional Factors

IVHS will be deployed in institutional environments that may or may not support their intended function. These environments can be assessed to isolate features that either support or suppress the successful deployment and operation of IVHS. Quite separate from the institutional environment for deployment, which can be considered a supplemental enabling mechanism, are the socioeconomic impacts that may result from the widespread deployment of IVHS. These impacts include such things as land use development patterns, migration and population patterns, work and travel norms, mode shift effects, impacts on urban form and culture, and induced travel demand.

Many of the secondary socioeconomic impacts eventually feed back to the transportation sector through shifts in travel demand. These higher-order effects should be integral to long-range forecasting efforts. The inclusion of higher-order effects, however, increases the uncertainty of most forecasts, so it will become essential that the IVHS community base their long-run assessments of future developments on the development of multiple scenarios. These scenarios can be developed by assessing a reasonable range of strategic institutional factors, parameters for endogenous factors, and assumptions for endogenous factors.

Multisite Deployment

If different architectures providing the same IVHS functionality are to be compared in the field, then it is desirable to have all of the alternatives deployed at the same location, controlling for possible confounding effects such as differences in climate, network, time of day, and subjects. Given the political realities of IVHS deployment in the United States, however, it is unlikely that many of the operational field tests will involve comparisons of multiple competing systems at the same site. The deployment of different, competing systems at different sites is more likely. This deployment pattern will prevent valid comparisons between systems, especially if the unit of analysis is the network, such as in the case of motorist information and route guidance systems.

Multisite comparison becomes more reasonable when it is possible to control confounding variables. For example, a comparison of incident detection algorithms at different sites may be reasonable when weather, traffic flow, road type, detector deployment, and number of lanes are controlled. In most cases where comparison is the objective, however, single-site comparisons are superior.

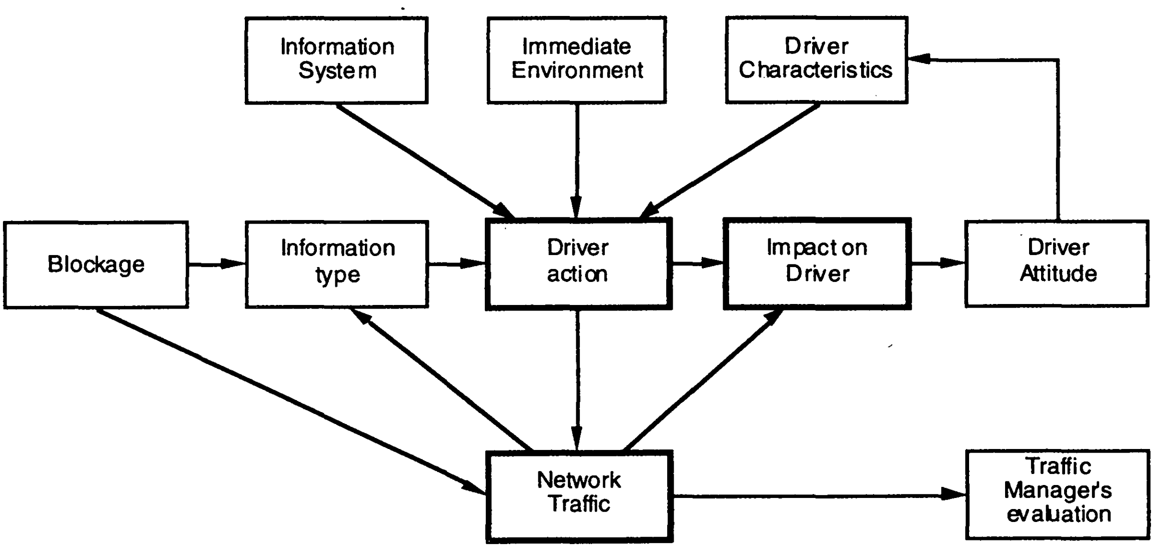


FIGURE 6 Determinants of driver perception and attitude.

Multisite deployment provides an opportunity for testing the robustness of system type under a variety of conditions. If competing systems are to be deployed at multiple sites, the assessment of these systems can focus on robustness, not on comparative effectiveness.

When comparisons are to be made it is best to deploy the alternative systems on the same site. In either case, whether the objective is to compare systems or to assess generalizability between sites, it is essential to develop and adopt a set of standards for evaluation measures, instruments, and methods that will facilitate cross-site synthesis at some later time.

Publication of this paper sponsored by Committee on Intelligent Vehicle Highway Systems.

Regional Approach to Strategic Intelligent Vehicle Highway System Planning in Orange County

GLENN N. HAVINOVISKI, BARBARA LEONARD, AND DEAN DELGADO

The Orange County, California, Intelligent Vehicle Highway Systems (IVHS) Study has developed a framework under which advanced technologies will be deployed to improve the operation of the county's highway and public transportation system. The most significant challenge of the study was to reconcile an overall high-level transportation vision with the needs, concerns, responsibilities, and financial limitations of all the local and regional agencies. The process used to develop a regional IVHS strategic plan is described and a review of the required areas of emphasis in developing such a plan is discussed.

The Orange County, California, Intelligent Vehicle Highway Systems (IVHS) Study has developed a framework under which advanced technologies will be deployed to improve the operation of the county's highway and public transportation system. Commissioned by the Orange County Transportation Authority (OCTA) and conducted by a team of consultants led by JHK & Associates, the study is a culmination of the following activities:

1. Identification of regional and local transportation goals;
2. Analysis of IVHS strategies and technologies that support these goals;
3. Identification of a transportation network for IVHS improvements;
4. Investigation of institutional issues associated with the implementation of IVHS;
5. Development of an IVHS master plan defining specific programs and an implementation strategy and an action plan that identifies specific projects and priorities, including estimated costs and funding availability; and
6. Preparation of final report that documents the project activities and findings.

These activities are illustrated in terms of the overall study process as shown in the flow diagram in Figure 1.

APPROACHES TO IVHS STRATEGIC PLANNING

The development of a strategic IVHS plan can be approached in various ways, on the basis of the scope of the project. For example, the work on the national level by IVHS America has focused on defining an overall direction for IVHS technologies and an overall approach to utilizing advanced technologies to solve transportation problems. Statewide studies such as those completed in Colorado and Washington State provide an overall vision and a general direction for development of programs. However, a key element in statewide IVHS planning is determination of the appropriate needs of specific regions. By nature, most states have several types of regions, ranging from major urban areas to smaller cities to rural areas. Although it is relatively elementary to identify an overall "high-level" IVHS vision and program for a region, a major issue is how the various local and regional agencies will be able to work together to implement the regionwide system, given their existing infrastructures and their inevitable limitations in the area of funding, staffing, and available expertise.

The most significant challenge of the Orange County IVHS study was to reconcile the overall vision of an integrated, multimodal transportation management and information system to serve the public, with the needs, concerns, responsibilities, and financial limitations of 31 cities plus several regional agencies. Important to the successful implementation and operation of IVHS in Orange County, as anywhere, is the realization that IVHS requires dedicated sources of funding and staff commitment for continued operations and maintenance. This need must be realized and met by public agencies responsible for the management and funding of transportation, politicians whose support is necessary to carry out programs, and the public.

GEOGRAPHIC AND DEMOGRAPHIC DESCRIPTION

Orange County, California, with 2.5 million people, is located between Los Angeles and San Diego along the Pacific Ocean. Historically an agricultural and later a predominantly residential area, the county has seen a considerable amount of growth since the 1950s, including substantial commercial, retail, and residential development. The county also contains a number of major tourist attractions, including Disneyland and Knott's Berry Farm. Other recreational trip generators include Anaheim Stadium and Arrowhead Pond (major league sports and concerts), as well as Irvine Meadows and Pacific Amphitheaters, and seasonally, the Orange County Fair and the beaches along the Pacific.

An extensive network of freeways and surface streets has been developed in the county, and the problem of major congestion during both peak and off-peak periods has been confronted. A countywide bus transit system, operated by OCTA, is being augmented by expanded commuter rail service in the Los Angeles-San Diego and Orange County-Riverside corridors. This expanding

G. N. Havinoviski, JHK & Associates, 300 South Harbor Boulevard, Suite 516, Anaheim, Calif. 92805. B. Leonard, Barbara Leonard Associates, 1125 Rosedale Drive, N.E., Atlanta, Ga. 30306-3914. D. Delgado, Orange County Transportation Authority, 550 South Main Street, P.O. Box 14184, Orange, Calif. 92613–1584.

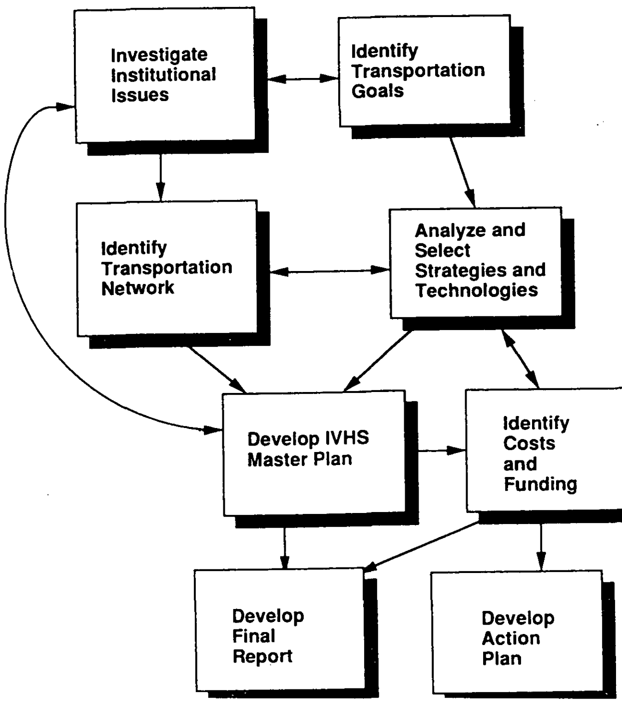


FIGURE 1 Orange County IVHS study process.

public transportation infrastructure, from a regional agency standpoint, is considered to be an increasingly valuable element of the overall transportation network.

IDENTIFICATION OF STRATEGIES AND TECHNIQUES

The first step in the development of the IVHS planning process for Orange County was the identification of local and regional transportation needs. Goals and objectives for the county's regional and local transportation systems were identified through interviews with state, county, and local agencies; neighboring agencies; and private institutions such as major activity centers, transportation management agencies, and the media. Additionally, national IVHS goals were assessed in relation to the county's needs. Eight primary goals emerged, the first five addressing transportation goals, and the remaining three addressing goals of the system architecture:

1. Increase efficiency,
2. Decrease emissions/energy use,
3. Enhance safety,
4. Support transportation operations and planning,
5. Improve quality of life,
6. Minimize cost,
7. Allow evolvability, and
8. Increase robustness.

The consultant team identified strategies that support the county's goals and related objectives (Figure 2). Finally, the strategies, which are independent of technology and type of improvement, were combined into sets of strategies similar in nature to assist in identifying user services, and their associated IVHS technologies and elements, which could be used to solve the various transportation problems in the county. Fifteen global strategies emerged, as presented in Figure 3. These strategies are correlated in this table to the IVHS user service categories, as defined by FHWA. These include traveler information, traffic management, freight and fleet management, public transport and emergency vehicle management, and additional services.

TRANSPORTATION NETWORK

To prioritize field improvements for the benefit of passenger vehicles, public transportation (e.g., buses, paratransit, rideshare), and commercial vehicles, an IVHS transportation network was identified. Those improvements that are more global in nature or are vehicle based, such as traveler information and Smart Bus operations, are detailed within the IVHS master plan, as discussed later in this paper. Orange County is fortunate in that a number of studies of the physical roadway network were conducted previously. The findings from these studies were incorporated in the analysis of the IVHS network, and the following classifications resulted (this list does not necessarily represent order of prioritization):

- Smart corridors: freeway segments with identified recurrent and nonrecurrent congestion and their arterial alternates;
- Smart streets: arterials located at regular intervals that have the ability to serve as freeway corridor replacements or freeway linkages; and
- Locally identified priorities such as
  - –Planned toll roads and
  - –Supplemental freeway segments: those freeway segments not identified as smart corridors.

Where the specific functions and nature of each of these categories of roadways (Figure 4) identified the need, various elements were recommended for implementation. These include "typical" traffic management system elements, including changeable message signs (CMS) and closed circuit television (CCTV) cameras, plus traffic control improvements. These include improved signal synchronization and adaptive control capabilities as well as integrated corridor signal and ramp metering operations. Various advanced surveillance elements, including video image processing, are identified for especially high-traffic or high-incident locations.

In addition to these facilities, public transit vehicles (buses, as well as fixed guideways, such as commuter rail lines) are recommended for deployment of vehicle location, data collection, and enroute information devices.

INSTITUTIONAL ISSUES

To assess the impact of institutional issues and establish a consensus with regard to the direction of IVHS programs in the county, the consultant team received direction and comments from a number of OCTA oversight groups at both a policy level and a technical level. In addition, interviews were held with public agencies and private institutions about transportation within the county. These interviews focused on a number of issues:

1. Signal pre-emption for emergency and transit vehicles,
2. Incident management and freeway construction projects,
3. Special event traffic management,
4. Interagency traffic management,
5. Transit and IVHS, and
6. Air quality and IVHS.

Goal Objective Strategy

1 INCREASE EFFICIENCY

1.1 Manage Demand
- 1.1.1 Transportation Demand Management
- 1.1.2 Spread the demand (Encourage non–peak travel)
- 1.1.3 Reduce Demand

1.2 Manage Flow
- 1.2.1 Decrease Turbulence
- 1.2.2 Manage Routing in recurring congestion
- 1.2.3 Manage Routing in Construction/Maintenance/Special Events
- 1.2.4 Provide Pre–Trip Information to Traveler
- 1.2.5 Provide Information to Motorist in Vehicle

1.3 Regain Capacity Following Incident
- 1.3.1 Preplan for Incidents
- 1.3.2 Detect Incidents
- 1.3.3 Identify/Verify Incidents
- 1.3.4 Respond to Incident
- 1.3.5 Clear Incident
- 1.3.6 Clear Incident–Caused Congestion

1.4 Increase Capacity
- 1.4.1 Add Facilities
- 1.4.2 Eliminate Bottlenecks

2 DECREASE EMISSIONS/ENERGY USE

2.1 Manage Demand
- 2.1.1 Restrictions on Travel when Air Pollution is High
- 2.1.2 Transportation Demand Management
- 2.1.3 Spread the demand (Encourage non–peak travel)
- 2.1.4 Reduce Demand

2.2 Encourage Fuel–Efficient/Clean–Running Vehicles
- 2.2.1 Economic Incentives/Disincentives
- 2.2.2 Mandates
- 2.2.3 Funded R & D into clean energy vehicles/subsystems
- 2.2.4 Fines for emissions
- 2.2.5 Highway Speed Emissions Monitor

2.3 Maintain Steady Speeds
- 2.3.1 Decrease Turbulence
- 2.3.2 Manage Routing in recurring congestion
- 2.3.3 Manage Routing in Construction/Maintenance/Special Events

3 ENHANCE SAFETY

3.1 Reduce the Number of Accidents
- 3.1.1 Eliminate Infrastructure Hazards
- 3.1.2 Decrease Turbulence
- 3.1.3 Prevent Unsafe Driving

3.2 Reduce Severity of Accidents
- 3.2.1 Eliminate Infrastructure Hazards
- 3.2.2 In–Vehicle Safety Measures

3.3 Avoid Secondary Accidents
- 3.3.1 Warn Driver
- 3.3.2 Respond to Incident
- 3.3.3 Clear Incident

FIGURE 2 Orange County IVHS architecture: goals, objectives, and strategies (*continued on next page*).

The interviews resulted in the development of an agency consensus, definition of a specific wish list of improvements, and identification of various constraints and concerns about the development of IVHS programs.

Agency Consensus

In general, it was felt the agencies can and do work together. However, it was felt that a greater degree of coordination between local and regional/state agencies was needed. At the same time, although most of the transportation problems in the county are interagency in nature, the communities are in fact diverse. Although several cities boast major commercial and retail development as well as residential areas, many communities are primarily residential and are sensitive to additional traffic flows or resultant congestion within their respective communities. Clearly, the various agencies wish to retain their autonomy and maintain control over their facilities even as part of a coordinated interagency transportation system.

Wish List of Transportation Management Improvements

A number of specific items or programs were desired by the agencies as part of the development of IVHS programs for the county. These programs included many institutionally related elements,

Goal Objective Strategy

3.4 Speed Emergency Response
3.4.1 Respond to Incident
3.4.2 Clear Incident
3.5 Enhance General Safety
3.5.1 Improve Emergency Vehicle Access
3.5.2 Support Civil Defense Plans
3.6 Minimize Impacts of Construction/Maintenance/Events/Incidents
3.6.1 Manage Routing in Construction/Maintenance/Special Events
3.6.2 Provide Pre-Trip Information to Traveler
3.6.3 Provide Information to Motorist in Vehicle
3.6.4 Preplan for Incidents
3.6.5 Detect Incidents
3.6.6 Identify/Verify Incidents
3.6.7 Respond to Incident
3.6.8 Clear Incident
3.6.9 Clear Incident-Caused Congestion
3.6.10 Support Rerouting

4 SUPPORT TRANSPORTATION OPERATIONS AND PLANNING

4.1 Collect data on system performance and usage
4.1.1 Real-time Data Base
4.1.2 O-D Data based on AVI/AVL/VIPS
4.1.3 Credible data analysis procedures for historical analysis
4.2 Facilitate Interagency Coordination
4.2.1 Data Base Accessible to All Agencies
4.2.2 Enhanced Interagency Communications
4.2.3 Single Facility for Interagency Activities
4.2.4 Open Architecture
4.2.5 Direct Computer-to-Computer Communications
4.3 Increase Productivity of City/Agency Staffs
4.3.1 Real—Time Information
4.3.2 Interactive/Intuitive Information Display
4.3.3 Decision Aids

5 IMPROVE QUALITY OF LIFE

5.1 Traveler Comfort
5.1.1 Assist Stranded Traveler
5.1.2 Manage Routing in recurring congestion
5.1.3 Manage Routing in Construction/Maintenance/Special Events
5.1.4 Provide Pre-Trip Information to Traveler
5.1.5 Provide Information to Motorist in Vehicle
5.1.6 Provide Consistent Travel Times
5.1.7 Provide Information for Tourists
5.2 Traveler Convenience
5.2.1 Transportation Alternatives
5.2.2 Mass Transit Schedules and Modes Readily Available
5.2.3 Decrease Turbulence
5.2.4 Manage Routing in recurring congestion
5.2.5 Manage Routing in Construction/Maintenance/Special Events
5.2.6 Provide Pre-Trip Information to Traveler
5.2.7 Provide Information to Motorist in Vehicle

FIGURE 2 *(continued).*

such as interjurisdictional cooperation in developing traffic management plans and improving real-time notification of incidents that may affect a specific community or roadway segment. Also important to local agencies was the availability of technical assistance to help operate newer traffic management elements that require a higher level of maintenance than do existing elements.

System improvements that were identified by the agencies included improved signal control and coordination for enhanced flow, improved real-time system monitoring, and the use of mobile CMSs featuring localized seasonal travel information (e.g., beach parking) in lieu of permanent CMSs in most communities. In general, public transit was a much greater concern at the regional level than at the local level, and an emphasis on improving public transit use was considered an important global function of the IVHS programs for the county.

Constraints and Concerns

The agencies identified various constraints and concerns about the implementation of IVHS programs, such as limited staff availability for operations, maintenance, and participation in regional meetings. Another concern involved the availability of measurable benefits relative to the estimated system expense. Finally, many areas exist in the county where the capacity of both primary and alternate routes is insufficient, thus frustrating attempts to reroute traffic.

Goal Objective Strategy

5.3 Equity regardless of socio-economic status, disabilities, etc.
5.3.1 Intelligence in Infrastructure rather than in vehicle
5.3.2 Multi-lingual, both audible and visual information
5.3.3 Wheelchair accessibility of mass transit
5.4 Equitable distribution of costs and benefits
5.5 Enhance Economic Vitality
5.6 Decrease Noise
5.6.1 Sound Barriers
5.6.2 Reduce Demand
5.6.3 Inspections
5.6.4 Noise Sensors Combined with AVI
5.7 Enhance Reliability of System
5.7.1 Computer-Based Training
5.7.2 Expert systems for Diagnostics/Maintenance
5.7.3 Technology Insertion & Upgrade Program
5.7.4 Computer Simulation

6 MINIMIZE COST

6.1 Analyze Life Cycle Cost for Range of Alternatives
6.2 Minimize Non-Recurring Costs
6.2.1 Minimize Infrastructure Costs
6.2.2 Minimize Detector Costs
6.2.3 Minimize Communication Costs
6.2.4 Reduce Data Processing Costs
6.2.5 Reduce Costs of Signage and Displays
6.3 Minimize Recurring Costs
6.3.1 Reduce Maintenance Costs
6.3.2 Reduce Surveillance and Monitoring Costs
6.3.3 Reduce Info Mgmt and Dissemination Costs
6.3.4 Reduce Response Costs
6.3.5 Reduce Costs of Toll Collection
6.3.6 Reduce Costs of Regulation

7 ALLOW EVOLVABILITY

7.1 Allow Expansion to Meet Future Demand
7.1.1 Open Architecture
7.1.2 Communications Capacity
7.2 Allow Expansion to Add Capabilities as Technologies, Funding Available
7.2.1 Open Architecture
7.3 Allow Modifications to Meet Future Political and Social Environments

8 ROBUSTNESS

8.1 Provide Operational Flexibility
8.1.1 Fault Tolerance
8.1.2 Open Architecture
8.1.3 Redundancy
8.2 Provide Maintainable System
8.2.1 Automatic Problem Identification
8.2.2 Redundancy
8.2.3 Modularity
8.3 Adapt to Changing Traffic Patterns
8.3.1 Modularity
8.3.2 Expandability
8.3.3 Relatively Load-Insensitive Design

FIGURE 2 *(continued).*

INTERAGENCY RELATIONSHIPS

Of primary importance to the development of the regional IVHS is (*a*) that jurisdictional responsibilities and autonomy are preserved and (*b*) that the organizational structure is oriented toward efficient planning, implementation, and operations. The proposed structure for Orange County allows for jurisdictional responsibilities to be kept intact and provides the agencies with an opportunity for direct and indirect input to program development and management. These opportunities were identified through the following:

- Development of an IVHS steering committee with representatives from various areas of the county, plus California Department of Transportation (Caltrans), California Highway Patrol, and

Strategy \ User Service	Traveler Information	Traffic Management	Freight and Fleet Management	Public Transport / Emergency Vehicle Management	Additional Services
Manage Congestion - Recurrent					
Manage Congestion - Non-Recurrent					
Reduce Traffic Turbulence					(Automated Vehicle Control)
Develop Decision Support & Response Mechanisms					
Manage Incidents - Detection/Verification					
Manage Incidents - Response					
Manage Incidents - Rapid Removal					
Provide TDM Tools					
Inform Travelers - Pre-Trip					
Inform Travelers - En-Route					
Support Technologies to Enhance Safety					(Automated Vehicle Control)
Provide Full Accessibility for All Travelers					
Provide Info & Accessibility for All Agencies					
Develop Features to Enhance Maintainability & Cost Effectiveness					
Develop Facilities & Technologies to Reduce Emissions, Energy Use and Noise					(Automated Vehicle Control)

FIGURE 3 Relation of IVHS user services to Orange County transportation strategies.

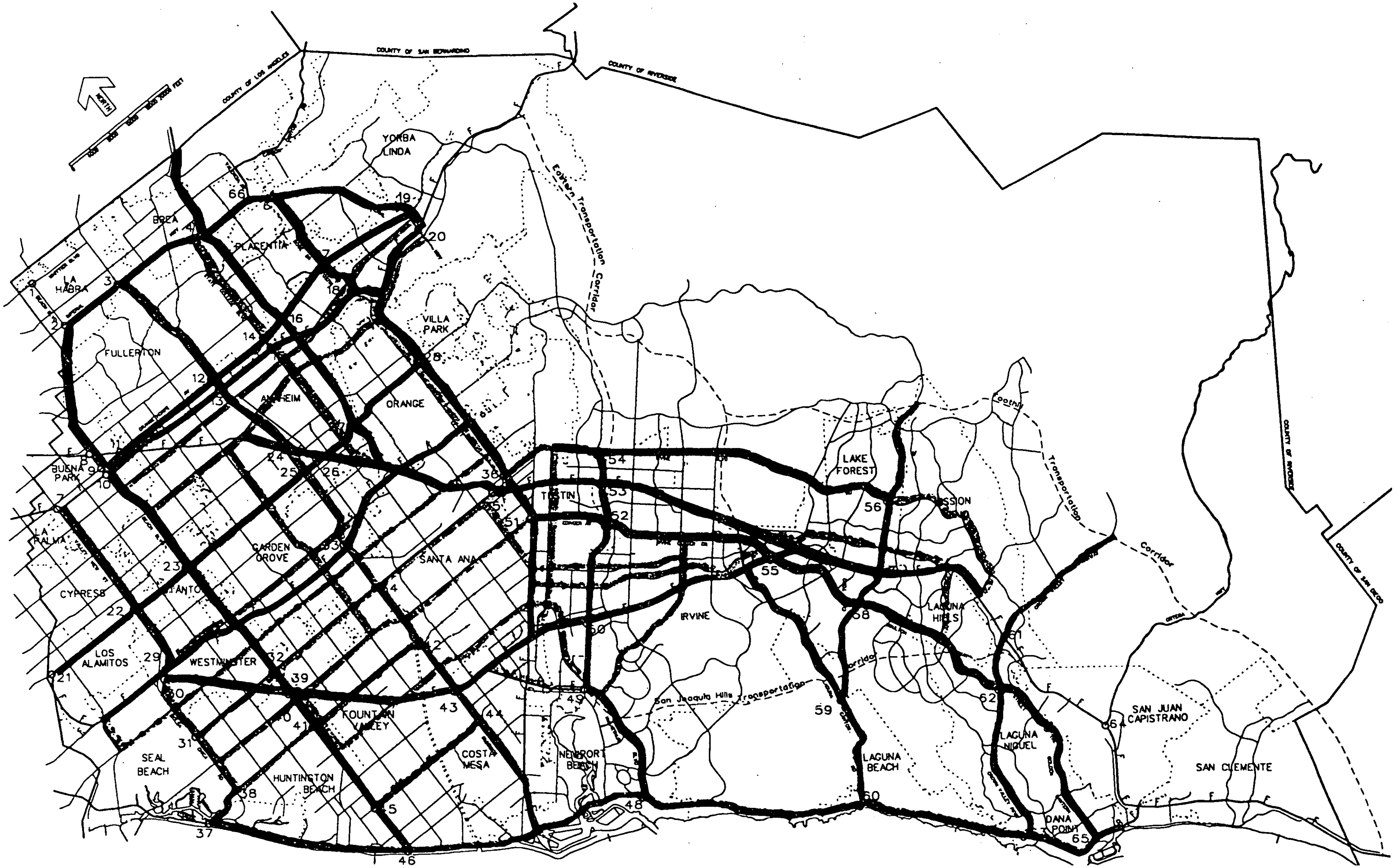


FIGURE 4 Preliminary IVHS network for Orange County.

the Automobile Club of Southern California. The responsibilities of the steering committee would include directing the development of IVHS in the county. These duties include the following:

–Identifying future programs and modifications to master plan,
–Securing funding,
–Implementing programs,
–Providing technical support,
–Developing traffic response/incident management plan, and
–Setting technical standards.

• Employment of an IVHS administrative staff by OCTA, the one agency within the county with responsibility for the entire transportation system, including streets and roads and transit. The administrative staff would carry out any and all of the administrative functions of the steering committee's responsibilities under the direction of the IVHS Steering Committee. Additionally, the administrative staff would be responsible for the following:

–Coordinating with multiagency growth management associations (GMAs),
–Providing support services to the steering committee,
–Maintaining draft agreements,
–Coordinating regional identification and formulating projects,
–Identifying and pursuing funding sources, and
–Coordinating projects.

The GMAs represent the cities as grouped in geographical subdivisions of the county. These areas were developed to implement transportation improvements in conjunction with Measure M, the county's ½-cent sales tax dedicated for transportation. It was recommended that the activities of the GMAs be expanded to incorporate subregional development, planning, implementation, and administration of IVHS programs. Local agencies, if they choose to coordinate through the GMAs, can work cooperatively as a subregion to further IVHS within their area. The latter is particularly important because much of the funding for IVHS on the regional, state, and federal levels places a high emphasis on the regional aspects of transportation improvements.

PROPOSED IVHS ARCHITECTURE

On the basis of an analysis of technologies and institutional implications of a countywide IVHS architecture, three alternative scenarios were considered, as indicated in Figure 5. These included a fully centralized architecture (centrally concentrated control, management, and dissemination), a decentralized architecture (similar to that of existing operations), and a hybrid architecture combining attributes of each of the above. The recommended hybrid architecture specifies the interconnection of local traffic management centers (TMCs) for local monitoring and control, a freeway traffic operations center (TOC), and a countywide multiagency traveler information center (TIC) for fusion of status data for distribution to travelers as well as the agency traffic managers (Figure 6).

The system centers around the collection, evaluation, and dissemination of data. The local TMCs and TOC receive data from whatever detection devices or other resources they use (e.g., loop detectors, CCTV, or police reports). These data are used to monitor the traffic in the jurisdiction and are also passed automatically to the TIC, where they are merged with data from throughout the county to form the countywide status. This status can then be called up by any TMC or by the Caltrans TOC. Furthermore, the TIC will alert any TMC or the TOC of incidents or events to which it should re-

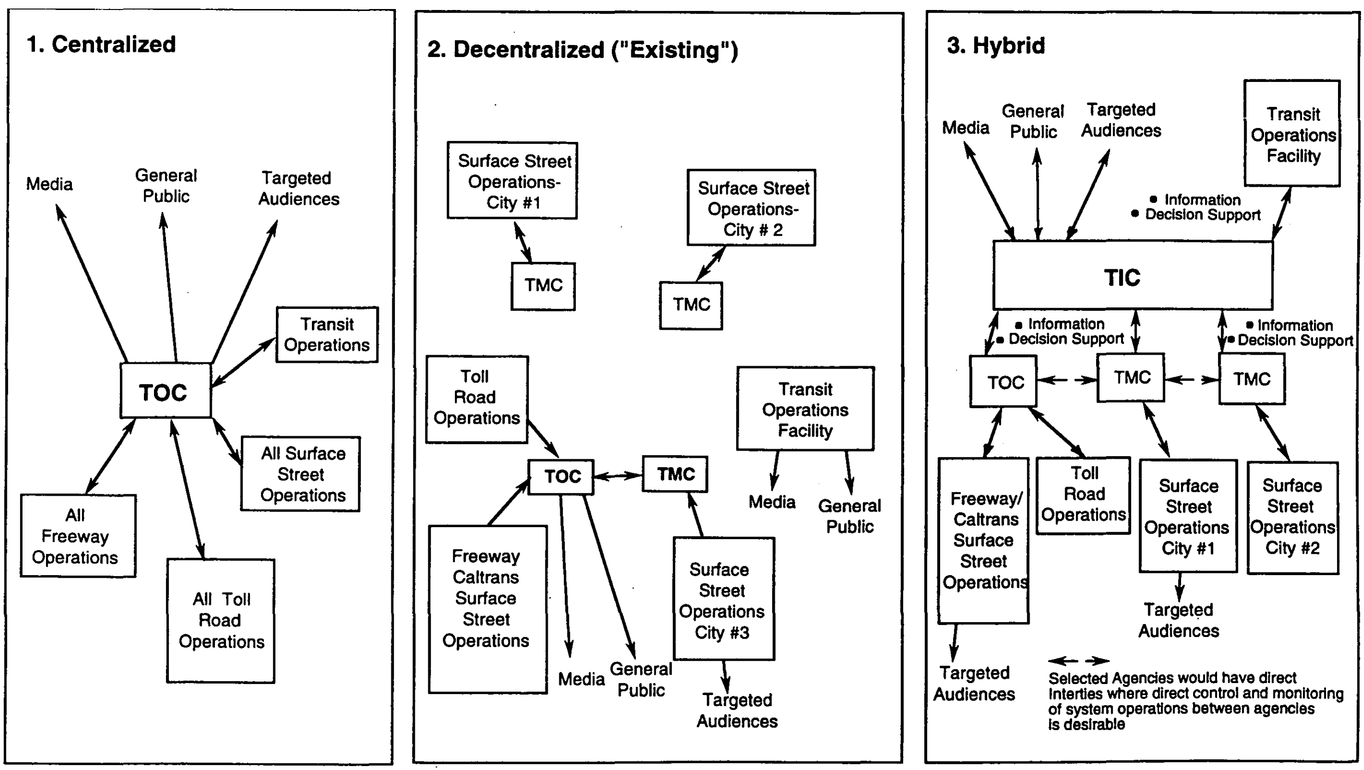


FIGURE 5 Comparison of IVHS organizational structures.

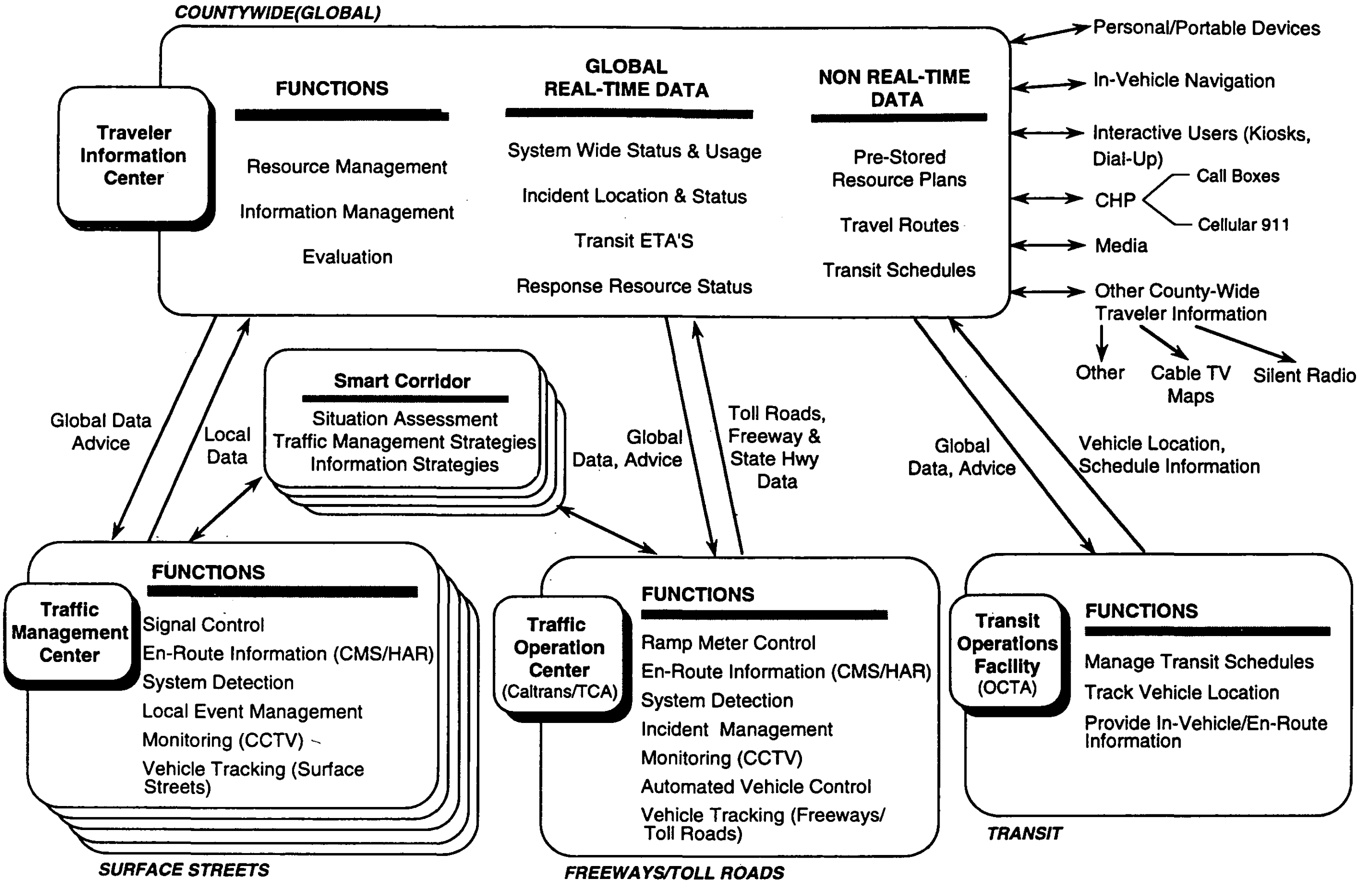


FIGURE 6 Overall structure: Orange County IVHS elements.

spond. Decision support systems (e.g., knowledge-based expert systems) will advise action. For example, a major accident on a freeway will cause one or more TMCs to be notified and asked to approve previously agreed upon multijurisdictional diversion plans.

IMPLEMENTATION STRATEGIES

The program development strategy leading to the IVHS master plan for Orange County was done in an incremental, building-block fashion. Two examples are presented below—one related to program development and the second related to an existing infrastructure. Unlike the "top-down" strategic development process, which is driven by a regional consensus and various high-level policy goals, the "bottom-up" program development process takes into account existing programs and constraints and utilizes these as the basis for development of an IVHS architecture.

Programmatic Example

In Figure 7 a programmatic example is given for the implementation of traveler information programs in the county. The programs were designed to build on existing elements, then incrementally develop elements that are capable of supporting the ultimate system. Two of the most critical near-term elements of the program are the development of a traveler information data base (TID), which serves as a clearinghouse for real-time traffic and transit information, and interagency links (interties), which allow the exchange of information between the agencies and the TID. These serve as building blocks on which the ultimate system is developed.

Systematic Example

The second example (Figure 8) is of the development of an arterial-based advanced traffic management system (ATMS) and is particularly relevant to the interrelationship of local and regional traffic operations. ATMS was a key element in the presentation to local agencies of how existing systems can be incorporated into an IVHS architecture. This example illustrates how the incremental development of an overall IVHS infrastructure can achieve specific objectives. These range from an isolated traffic signal (the most basic traffic management element other than a stop sign or striping) to real-time centralized system operations, to interagency coordination, and finally, to a regionwide system at the top level.

The regionwide system allows multiple agencies to share information and coordinate operations. Such a scheme is typified by the "smart corridor" concept, now being implemented in Los Angeles. Operator and traveler decision support systems (e.g., expert systems, interactive multimode traveler information systems) obtain data from local systems or a TID and provide tools for systemwide

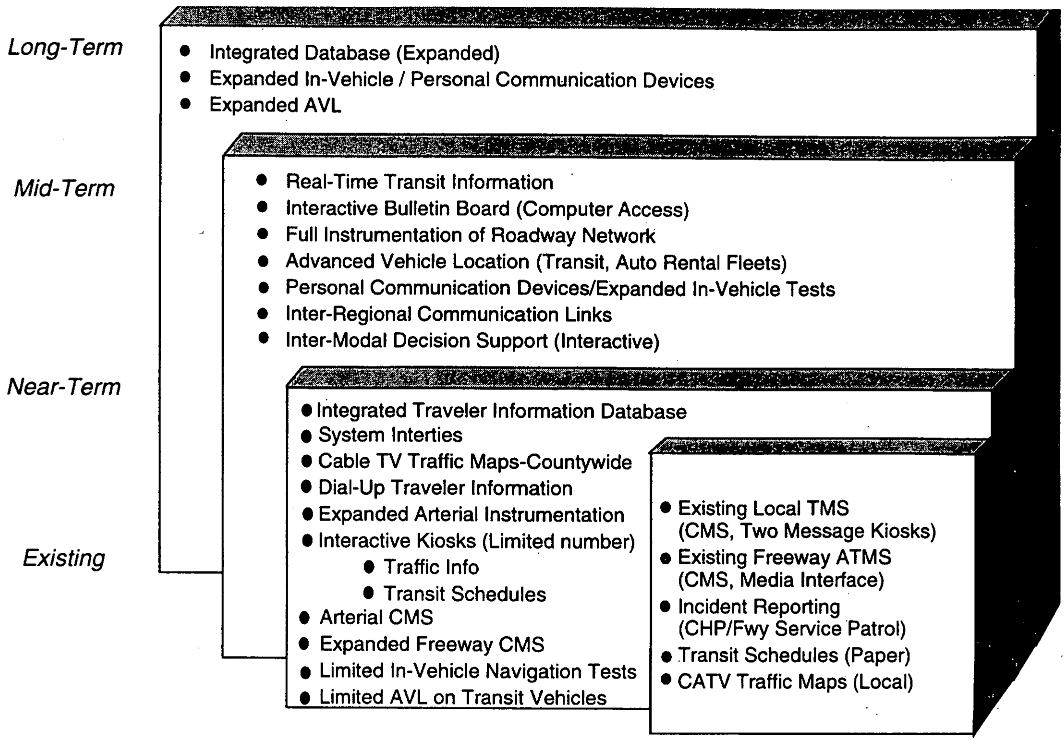


FIGURE 7 Programmatic IVHS development: traveler information programs for Orange County.

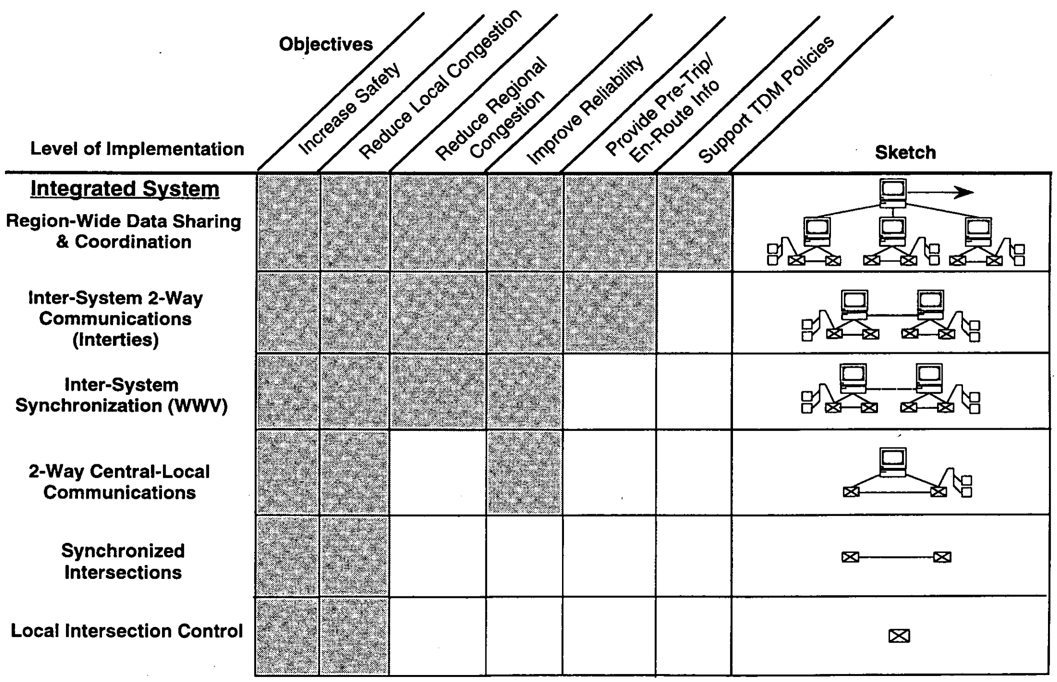


FIGURE 8 Implementation of arterial ATMS elements.

analysis. To support agency operations, coordinated interagency responses can be "downloaded" to the local systems for their confirmation and implementation. To support traveler information, appropriate guidance data (recommended mode, route, etc.) can be developed and displayed to travelers.

Existing Infrastructure

In Orange County a significant amount of effort has been placed into freeway management and signal synchronization. Caltrans is expanding the use of ramp metering, has been expanding CMS locations, and has been adding CCTV and loop detector stations to the existing surveillance system. It is also in the process of expanding its regional freeway TOC in line with a statewide TOC master plan.

Most of the local agencies in the county maintain some form of signal system. These range from Anaheim's centralized urban traffic control system to a number of closed-loop systems (smart field masters with central monitoring and download capability). Three cities have or are in the process of implementing CMS and CCTV elements as well as highway advisory radio.

However, the traffic operations backbone for more than half the agencies in the county consists of proprietary centralized traffic control systems that are vendor specific in terms of hardware and software. The nature of these systems over the years has been that they are incapable of sharing information with other systems. Activities are now under way by both the vendors and the agencies to permit the export of traffic information and signal status data.

The key of the IVHS architecture must then be to ensure that the information from the various agencies is compatible in both form and protocol. To that end, the study recommended the development of a countywide information protocol and linkages between local area networks located at various TMCs or encompassing various agencies. The Ethernet interface standard provides a solid foundation for standardized communications, as does the use of a protocol such as TCP/IP.

A number of current projects already have resulted in interties between Caltrans and local agencies to allow the sharing of local and freeway traffic information. Further improvements are resulting in the sharing of CCTV images between local agencies and Caltrans, such that an integrated approach to traffic monitoring and response can be implemented. In addition, the City of Anaheim has developed a traveler information system that provides interactive traffic information capabilities through dedicated kiosks, as well as a rotating color traffic flow map presented to local cable television subscribers on a community access channel.

IVHS MASTER PLAN

The recommended IVHS programs for Orange County include five of the following categories of development that are similar to the previously identified user services:

1. Traveler information,
2. Monitoring and data collection,
3. Traffic management,
4. Public transit/high-occupancy vehicles, and
5. Automated vehicle control.

The programs are described in Figure 9, along with estimated overall costs, including annual operations and maintenance.

ESTIMATED BENEFITS

The estimation of benefits relative to IVHS improvements is a difficult endeavor because many of the elements are new and are in the process of development or operational testing. Therefore, the analysis relied on the project team's experience from involvement in previous studies or projects that included a number of the traffic management or traveler information elements incorporated into the proposed countywide IVHS architecture. Elements evaluated included freeway incident management tools (incident detection, rapid response techniques using service patrols, and CCTV), traveler information elements (roadway-mounted CMSs, and limited usage of in-vehicle navigation tools), and traffic control (adaptive or traffic-responsive signal timing, corridor ramp metering). Because these are a subset of all the program elements recommended previously, the benefits assessment is thus conservative. At the same time, it was decided to compare this limited assessment of benefits with the total cost of IVHS improvements.

The analysis estimated the following benefits using the year 2005 as a base:

- Annual delay reduction benefit of $243 million (based on 34 million vehicle-hours saved at a $7.20/vehicle-hour delay cost as used by Caltrans) (*1*);
- Annual accident benefit of $48.9 million based on a 25 percent reduction in freeway accidents (12,000 accidents annually in year 2005 were estimated, at a cost of $16,300 on the basis of 1989 Caltrans data) (*2*); and
- Annual fuel consumption reduction benefit of $25.29 million (0.6 gal per vehicle-hr reduction at $1.25 gallon, based on above delay reduction)[2]

The conservative estimate of monetary benefit is $317 million annually, not including the impact of such IVHS-related improvements as real-time transit scheduling and information (and resultant impacts on mode split), as well as vehicle control and safety improvements. Given an estimated annual cost of $80 million for all except automated highway improvements and privately developed in-vehicle systems, this would result in a benefit-cost ratio of 4:1.

SUMMARY AND OBSERVATIONS

The Orange County IVHS Study, a multiagency, regionally oriented effort, has emphasized interagency consensus and the incorporation of an existing system infrastructure as the basis for a higher level of transportation management improvements. The following are the three major areas of emphasis identified by the consultant team;

- Countywide traveler information (pretrip and en-route),
- Integrated corridor traffic management, and
- Real-time management and information for public transit.

PROGRAM AREA / Programs	Functions
TRAVELER INFORMATION ($272 M)	**Provide transportation network information to public, media, transportation agencies**
Freeway Motorist Information Systems (FMIS)	Expanded roadside CMS & HAR on freeways
Arterial Motorist Information Systems (AMIS)	Roadside CMS & HAR on key arterials
In-Vehicle Information Support Infrastructure for On-Street Navigation (INVISION)	Provide information & vehicle-roadway communications infrastructure to support in-vehicle devices
Universal Traveler Information Program (UTIP)	Develop Traveler Information Center, databases and servers, interactive kiosks
Interagency Transportation Information Exchange (INTERTIE)	Develop distributed interagency communications and processing capabilities (standard interfaces)
Public Information Campaign	Provide public with information on the means to avoid delays through improved driving, travel habits
MONITORING AND DATA COLLECTION ($117 M)	**Provide real-time data for transportation and trip management as well as planning analysis**
Automatic Vehicle Location (AVL)	Equip vehicles with probes to obtain real-time location & operations data for use in travel monitoring & operations
Freeway Instrumentation	Detection, monitoring & surveillance for congestion measurement and incident detection on freeways (CCTV, VIP, detectors)
Arterial Instrumentation	Detection, monitoring & surveillance for congestion measurement and incident detection on surface streets (CCTV, VIP, detectors)
Detector Maintenance	Contracted technical support of local agencies for detector maintenance
TRAFFIC MANAGEMENT ($112 M)	**Enhance agency traffic operations capabilities and support both local and regional operations**
Traffic Management Centers	Build/Expand TMCs for management of state/local roads
Agency Traffic Operations Support (ATOS)	Maintenance Support for Local Agencies' IVHS Elements
Decision Support Systems	Expert Systems for real-time corridor traffic management
Emergency Priority System (EPS)	Testbed for interagency coordination of signal pre-emption through integration with TMSs for reduction of delays
Rapid Incident Clearance (RIC)	Expand Freeway Service Patrols and integrate reporting capabilities with UTIP program
Adaptive Signal Control and Signal Synchronization Program (ADAPT)	Enhance real-time traffic signal control (central and field improvements, software modifications)
Corridor Ramp Metering	Enhance real-time freeway flow through coordinated corridor metering strategies
Integrated Signal/Ramp Meter Control	Improve local signal/meter coordination to reduce impact of restrictive metering rates on surface street traffic
PUBLIC TRANSIT/HIGH-OCCUPANCY VEHICLES ($12 M, in combination with other programs)	**Support Transportation Demand Management policies through collection and dissemination of real-time transit/HOV information**
Public Transit/Smart Bus	Provide enhanced real-time transit information to public and for fleet management
Integrated Real-Time Rideshare (INTER-RIDE)	Interactive rideshare-matching through phone call-in, interactive UTIP terminals using rideshare database
Real-Time Intermodal Advisory (RITA)	Integration of transit and traffic information, development of real-time comparisons for travel times between modes
AUTOMATED VEHICLE CONTROL ($207 M public sector)	**Support future needs of automated control and central-to-vehicle real-time communications**
AVCS Operational Support	Support AVCS through communications servers and operations systems in conjunction with private sector investment for in-vehicle elements
Platooning Lanes	Provide infrastructure and civil engineering modifications (including new lanes) to accommodate automated control

SYSTEM TOTAL	**Estimated $601 Million**
OPERATIONS & MAINTENANCE	**$80 million annual**

FIGURE 9 Description of Orange County IVHS programs.

To satisfy these areas of emphasis, the study has focused in detail on developing an infrastructure (physical system as well as management) capable of integrating various information and management elements (both roadway and transit) and supporting the extensive interjurisdictional coordination required. Additional effort will be required to provide all of the agencies with the means (technical as well as financial) to support the operation and maintenance of IVHS elements. A key will be the continuous effort in identifying sources of funding that can be used for maintenance of the system.

Although the program emphasis is oriented more to the near term, the architecture developed in the study is highly appropriate for support of efforts toward in-vehicle navigation devices and automated vehicle control, efforts that are integral to the overall direction of the IVHS program. Nevertheless, to make the more advanced elements workable, it is necessary to develop a suitable backbone of real-time information and control capabilities. Thus, the nature of this regional IVHS strategic plan has been to emphasize the practicality of implementation as a key criterion.

REFERENCES

1. JHK & Associates. *Bay Area Traffic Operations System Operational Procedures and Strategies Report.* California Department of Transportation, District 4, May 1992.
2. JHK & Associates. *Orange County TOC Study Final Report.* Orange County Transportation Commission, Oct. 1991.

Publication of this paper sponsored by Committee on Intelligent Vehicle Highway Systems.

Use of System Characteristics to Define Concepts for Automated Highway Systems

WILLIAM B. STEVENS

The automated highway system (AHS) is a surface transportation system program that uses modern electronics to instrument highways and vehicles to provide "hands-off" and "feet-off" vehicle operation. Research to date shows that AHS has the potential to double or triple the nation's highway efficiency and to dramatically increase highway safety. The impact of AHS on the nation's highways would be comparable to the impact the jet engine had on aviation 40 years ago. FHWA has established an AHS program that will (*a*) identify and analyze alternative AHS concepts; (*b*) demonstrate the potential feasibility of AHS in 1997; and (*c*) if feasible, select and document the evolutionary AHS concept to be operationally tested beyond 1997. An operational test, with public participation, would then be conducted. Analysis, modeling, simulation, and testing will be used in comparing and evaluating the concepts. To assist FHWA in defining the AHS program, an initial definition and assessment of alternative AHS concepts has been made. The process by which AHS concepts were defined using the primary system characteristics of an AHS, specifically the functional and physical characteristics of an AHS, are described, and those characteristics that may vary from one concept to another are identified. From this, those distinguishing factors, termed concept definition factors, are used to postulate the initial set of AHS concepts.

In response to the Intermodal Surface Transportation Efficiency Act (ISTEA) (*1*), FHWA has established a program to determine the feasibility of automated highway systems (AHS). An AHS uses modern electronics, sensors, and communications on highways and vehicles to provide "fully automated" vehicle operation; this means that as an AHS vehicle pulls onto the AHS lane of an expressway, control of the vehicle's steering, braking, and acceleration is assumed by the AHS to provide lateral and longitudinal control until the vehicle exits and the driver again assumes control of the vehicle.

Substantial research to date shows that AHS has the potential to double or triple the nation's highway efficiency (*2*) and to dramatically increase highway safety (*3,4*). This impact of AHS on the nation's highways would be comparable to the impact the jet engine had on aviation 40 years ago. The long-range goal of the program, then, is to significantly improve the safety and efficiency of the nation's surface transportation system through a national effort that best ensures the early, successful deployment of automated vehicle highway systems.

It is recognized that first and foremost, the AHS is a highway vehicle system; therefore, its design must be based on solid, state-of-the-art engineering of both the highways and the vehicles that operate on them. This paper focuses on just those aspects of the highway and vehicles that might pertain to AHS. Also, it is recognized that resolution of AHS institutional issues, such as tort liability, may have a significant impact on how an AHS is designed, implemented, and operated. This paper does not address these impacts.

The MITRE Corporation, 800 Maryland Avenue, S.W., Washington, D.C. 20024.

The near-term objective of the AHS program is to test AHS feasibility and, if positive, select an AHS concept that would be used for follow-on operational test and evaluation. It is believed that a feasible AHS will be a robust, affordable, user-friendly fully automated vehicle highway system that evolves from today's road system and has significantly better safety, efficiency of operation, and comfort than today's highways. Specific activities of the current program include a 1997 proof-of-feasibility demonstration, in response to the congressional direction in ISTEA; selection and documentation of the preferred AHS concept for further test and evaluation; and development of a plan for evolution from today's highways to AHS, possibly using products that incorporate partial vehicle control, such as collision avoidance.

Work has begun. An effort to develop human factor guidelines for the program is under way (5), and 15 contracts have been awarded to identify and analyze key AHS requirements, risks, and issues dealing with design, deployment, and operation (*6,7*). FHWA has issued a Request for Applications for a consortium to work in partnership with FHWA to (*a*) implement the 1997 demonstration and (*b*) select the AHS concept that will benefit the public and industry as it evolves from today's highway system, and is preferred for operational testing.

PURPOSE AND APPROACH

To assist FHWA in defining the AHS program, an initial definition and assessment of alternative AHS concepts has been made. This paper defines the process by which the AHS concepts were defined and describes an initial set of AHS concepts. In taking the first step toward defining AHS concepts, assumptions were made about potential approaches for AHS implementation. These assumptions were made with the intent of not excluding any AHS concepts that might conceivably be considered legitimate. As AHS research continues and full sets of alternative AHS concepts are defined, the material in this paper will need to be revised. An AHS must be designed to meet the overall goals established for the system. Table 1 provides an overall summary of some strawman AHS deployment and operations goals.

The AHS goals will not vary from one AHS concept or design approach to another (*8*). What varies are the design characteristics of the various system concepts or design approaches, or both, postulated to meet the goals. Accordingly, this paper describes the system characteristics that would be used to meet the AHS goals and identifies the AHS characteristics that may be used to distinguish one AHS concept from another, termed concept definition factors (CDFs). Then using the CDFs, an initial set of AHS concepts is defined and discussed.

TABLE 1 Overview of Major Deployment and Operations of AHS Goals

- **Improve Operating Effectiveness** — Increase throughput of people, goods, and vehicles, and improve operation in adverse weather.
- **Improve Transportation Service** — Provide a full range of services, reduce travel time, and improve travel reliability.
- **Improve User Desirability** — Improve safety, enhance personal mobility, increase comfort of highway travel, provide user-friendly service, reduce insurance costs, and ensure affordable cost.
- **Improve Community Desirability** — Reduce land use, property impact; reduce need for emergency support; reduce construction disruption.
- **Improve State Transportation Agency Desirability** — Provide a basis for long-term upgrade to major highways, enable smooth transition, enable smooth installation, enable practical operation, provide better cost/benefit ratio, and integrate with and support transit operations.
- **Provide Societal Benefits** — Strengthen the nation's economy, nurture the U.S. AHS industry, support national emergencies, reduce fossil fuel consumption, and reduce pollutants from vehicles.

DEFINITION OF AHS CONCEPT

For purposes of this preliminary investigation, an AHS concept is defined as a conceptual-level system configuration that is defined by a set of characteristics and is fundamentally different from other conceptual-level AHS system configurations. A fundamental difference is when both instrumented vehicles and roadways differ from other configurations to the extent that changing to another configuration would cause a major redesign of the system. An example of a fundamental difference would be an approach in which only specially designed narrow vehicles are allowed on narrower-than-normal AHS lanes. To change such an AHS system to a different approach using normal widths would mean (*a*) widening the AHS lanes or reconfiguring the roadways and (*b*) evolving away from the narrow vehicles.

AHS COMPONENTS

AHS goals can be met by a variety of conceptual-level system configurations or concepts. Both functional and physical characteristics may distinguish among AHS concepts.

For purposes of this paper, it is assumed that the functional and physical characteristics of an AHS can be described in the context of these major components. (There may be many different approaches for segmenting the AHS into major components; the segmentation in this paper is suitable for initial analysis; however, as the program progresses, adjustments will need to be made as more accurate or varied definitions or both, are developed.) The various segments are as follows:

- Vehicle,
- Roadway infrastructure,
- Command and control,
- Entry and exit infrastructure,
- Communications (could be included with command and control), and
- Operations and maintenance.

This paper considers only the major characteristics that distinguish among AHS concepts. The first three AHS components—vehicle, roadway infrastructure, and command and control—all have these major characteristics. The analysis upon which this paper is based shows that the variations in the communications component and operations and maintenance component, although interesting, will not distinguish one overall AHS concept from another. The entry and exit infrastructure component has some interesting characteristics and variables, but they do not seem to be fundamental differences; that is, changes in approaches will not change the way in which either the vehicle or roadway is designed, and the entry and exit strategies could be changed without changing the rest of the system. Following are some of the major entry and exit infrastructure variables:

- Time of check-in: On-the-fly check-in in buffer lanes versus slow/stop at check-in stations (such as toll booths); all concepts will have on-the-fly check-in as a goal, if technically possible. It is assumed that noninstrumented vehicles may be part of the check-in traffic stream.
- Queue for check-in: Intermixed vehicles versus separate lanes for certain types of vehicles (e.g., trucks); this variant would be decided by assessment of which is more efficient at a particular entry plaza; in fact a given system could employ both options. Either approach could be used by most concepts.
- Queue for merging: All vehicles enter sequentially intermixed into the same AHS lane versus vehicles separated by type or entering by platoon, or both, possibly into specialized lanes; this variant also would be decided by assessing which is more efficient at a particular entry plaza; in fact a given system could employ all of the options, and the approaches could all be used by most concepts.

On the basis of this analysis, it was concluded that the entry and exit infrastructure component did not have major characteristics that could distinguish one concept from another, and that most if not all fundamental differences among AHS concepts will be defined in the vehicle, roadway infrastructure, and command and control components.

VEHICLE CHARACTERISTICS

Primary Functional Characteristics

The primary function of the AHS vehicle is to carry the driver, passengers, and goods as they are moved through an AHS system. The AHS vehicle must (*a*) provide for controlled vehicle movement while in the AHS; (*b*) interact with the AHS roadway infrastructure component to obtain traction and support for operation while in the system, and to obtain lane boundary indications; (*c*) interact with the entry and exit infrastructure component to provide smooth and rapid entry and exit to AHS; (*d*) provide accurate control responses to directions received from the command and control component regarding vehicle braking, steering, throttle, and lights; (*e*) detect and maintain the status of critical vehicle functions; (*f*) support access to and from the command and control and communication components; and (*g*) interact with the driver, on a user-friendly basis, with respect to AHS status and driver directions.

Primary Physical Characteristics

For any vehicle to operate as part of an AHS, regardless of the concept, it must have certain physical characteristics.

AHS Capability Rating

Any vehicle produced for the U.S. market after a to-be-determined date will have one of three AHS capability ratings. Only the AHS-capable and upgraded AHS-compatible vehicles will be allowed on AHS roadways.

- AHS-capable Vehicle: The vehicle is capable, as it is produced, of fully automated operation on a standard U.S. AHS roadway.
- AHS-compatible Vehicle: The vehicle is capable of being upgraded for full operation on an AHS roadway. The vehicle must include those items that are most economically included at the factory; and the design must readily accommodate the AHS upgrade components (sensors, processors, etc.).
- Non-AHS-compatible Vehicle: The vehicle is not capable of being upgraded for full AHS operation on a reasonable basis.

Mode of Operation

AHS vehicles must be able to operate in a fully automated manner on AHS roadways and under manual control on nonautomated roadways. Instrumentation for other services on non-AHS roadways, such as autonomous intelligent cruise control, may be used to provide part of the AHS control; these other services may act as evolutionary steps toward AHS.

Vehicle Status

Vehicles must have, or be upgradable to include, sensors and diagnostics to detect (*a*) malfunctions in engine, cooling, electrical, and braking systems; (*b*) performance degradation, including but not limited to power train performance, tire inflation, and loss of traction; (*c*) status of the on-board AHS system components and their interfaces; (*d*) ability of the driver to resume control; and (*e*) low fluid levels for fuel, engine, transmission oil, coolant, and brake fluid. Additional/different sensors would be needed for electric vehicles.

AHS Instrumentation

An AHS vehicle must include the AHS instrumentation, which may encompass

- Longitudinal sensors: Provide warning for spacing (position keeping), collision avoidance, and obstacle detection;
- Lateral sensors: Provide sensing of other vehicles laterally for passing, lane changing, merging, and collision avoidance; and
- Lane boundary sensors: Provide sensing of the lane boundaries by interacting with the roadway infrastructure to provide for lane keeping.

Mounting for Communication, Command, and Control

The vehicle must provide for mounting communications and command and control components.

- AHS communications: Provide for mounting of, and interactions with, communications components for interacting with wayside and vehicles; mounting provisions must include electronics packages and antennae.
- AHS command and control processor(s): Provide for mounting of, and interactions with, command and control component electronics; interface provisions include interactions with sensors, the driver, actuators, the communications component, and other IVHS services.

Chassis Design

Special AHS chassis designs may include dimension restrictions; special crash protection such as special side, front, or rear crash protectors, or emergency towing or removal connectors, or both of the latter two. As the community moves toward AHS deployment, performance specifications for these sensors, actuators, and chassis provisions will be developed so that AHS-compatible vehicles can be produced and marketed.

Vehicle Variables

Some of the vehicle characteristics that distinguish one AHS concept from another are summarized next.

Vehicle Class

Vehicle class describes dimension and performance traits of the vehicles that can be accommodated by an AHS lane. The classes are defined on the basis of (*a*) maximum width of the vehicles that can be accommodated by an AHS lane; (*b*) minimum rate of acceleration to access the lane; and (*c*) minimum top speed in the lane. These variations are considered fundamental; for example, changing lane width or buffer lane length to accommodate vehicles that are wider or slower, respectively, could result in a major change to the AHS design. Other factors such as stopping distance, weight, or height also should be considered eventually.

Roadway Infrastructure Interaction

The roadway infrastructure interaction variable defines the physical vehicle-to-roadway interaction. Most (but not all) AHS approaches assume a rubber-tired vehicle riding on freeway-quality road surfaces. However, the interaction between the vehicle and the roadway infrastructure could be significantly different. Two variations are addressed in this paper. One variation would be the use of pallets (i.e., specialized trucks) with AHS instrumentation that would hold noninstrumented vehicles (*9,10*). The second (albeit unconventional) variation could be a specialized pallet that would use another form of roadway interaction that includes magnetic levitation and air cushion. This alternative is highly speculative but is theoretically possible; it is included for the sake of completeness. A pallet is assumed because vehicles may not be designed for this kind of operation.

Vehicle Power Source

Most AHS concepts assume that the vehicles have an on-board power source, such as fuel for internal combustion engines or batteries for electric motors. However, there is a variation in the way electric power is provided to electric vehicles while they are on the roadway (*9*). The vehicle would need to be designed to accommodate this power transfer. The straightforward way of providing the power would be through electric contacts with power rails along the roadway; however, there are concerns that this is not a viable technology to evolve through the twenty-first century. Other noncontact approaches for transferring electric power include microwave or induction; the viability of these approaches needs to be examined.

Vehicle Lateral Control Strategy

Lateral control strategy refers to the method by which the vehicle interacts with the roadway infrastructure to determine its lateral position in the AHS lane. There are at least three alternatives: (*a*) passive center lane markers such as magnets; (*b*) passive barriers or markers on the side of each lane, or both, or (*c*) active lane markers such as an activated embedded wire in the roadway or, in theory, radio frequency triangulation. The vehicle's lateral position sensors would need to be designed to accommodate these significantly different kinds of markers.

ROADWAY INFRASTRUCTURE CHARACTERISTICS

Primary Functional Characteristics

The function of the AHS roadway infrastructure is to (*a*) provide traction and support for vehicle operation, including vehicles operating properly and those that are malfunctioning; (*b*) enable safe vehicle operation by ensuring vehicle separation in case of severe system malfunction (e.g., separators and barriers); (*c*) provide connectivity for entering and exiting vehicles and connectivity to other AHS systems; (*d*) provide passive or active indication of lane boundaries; (*e*) provide sensing of environment or obstacles, or both; (*f*) support/enable command and control and communications access to AHS vehicles and to roadway conditions; and (*g*) support access to roadway by emergency and maintenance vehicles.

Primary Physical Characteristics

It is assumed that most AHS concepts will require a freeway type of AHS roadway (as defined by AASHTO), with the difference that the AHS lanes may be significantly narrower. The narrower lanes are possible because lateral position in the AHS roadway lane will be controlled automatically; therefore, lateral position can be held to much closer tolerances than they could with human operators. Because of this, more AHS lanes can be built on existing roadway surfaces.

The roadway infrastructure design as a whole must be tailored so that AHS roadway lanes are sufficient to respond to the traffic flow needs. The specifics of how the roadway may be constructed are not addressed in this paper. For example, the decision of where to add AHS lanes (median of existing roadway, shoulder of existing roadway, elevated over existing roadway, or separate from existing roadway) should depend on the existing roadway configuration, surrounding environment, and relative cost/benefits. Also not addressed are the design and placement of barriers and the type of roadway surface. These implementation options do not define alternative AHS concepts, although the AHS design must obviously consider them.

The AHS entry and exit plazas are defined as part of the entry and exit infrastructure component; however, buffer lanes are needed as part of the roadway infrastructure to accommodate vehicles that are accelerating and decelerating as they enter and exit the system. Provisions also have to be made to accommodate vehicles with malfunctions (flat tires, overheating, etc.). The roadway surface condition will be monitored as part of the command and control component. Conditioning of the roadway surface is required for all AHS installations. The only option is the extent to which the conditioning is automatic, such as heated surfaces on selected key bridges. Depending on the extent to which lane changing is allowed and the strategy for accomplishing that, some buffering may be needed between AHS roadway lanes.

The roadway infrastructure must have provisions for accommodating command and control sensors, processors, and communications links that may be located at roadside. Also, the infrastructure must be designed to accommodate the special AHS operations and maintenance systems and any specially designed emergency vehicles.

Roadway Infrastructure Variables

The variable roadway infrastructure component characteristics correlate directly to the variable vehicle component characteristics; the roadway infrastructure view of these variables is given below.

Lane Width

Lane width is defined by the maximum width of the vehicles that will operate on a given AHS roadway lane, and the expected performance tolerances of their automatic lateral control. Lane width correlates with vehicle class and implies that AHS systems can be designed so that different classes of vehicles are separated, either in different lanes, in separate platoons, or even by areas of access. For example, a congested urban area could restrict inner-city access by trucks to only "narrow" ones; or high-performance intercity lanes could be offered as an added service. The vehicle performance would affect the length of entry/exit buffer lanes and would affect the ability of one platoon to pass another in a restricted space.

Vehicle Interaction

Vehicle interaction is the variable that defines the physical vehicle-to-roadway interaction. As with the vehicles, this interaction can vary depending on the kind of vehicle to roadway interaction assumed for an AHS approach—either freeway road surface (primarily), pallet, or specialized roadway interaction. The pallet may not require any special roadway design but would require the entry and exit component to have provisions for pallet loading, unloading, recirculating, storage, and maintenance.

Roadway Power Source

If the roadway infrastructure component provides full or partial electrical power to some or all of the vehicles, its design would be significantly different from that of infrastructures in which no power is provided.

Roadway Lateral Control Strategy

Infrastructure design will differ significantly depending on whether the lateral control strategy is to use (*a*) passive lane markers that are magnets in the centers of lanes; (*b*) passive barriers or markers or both, on the side of each lane; or (*c*) active lane markers such as an activated embedded wire in the roadway or radio frequency triangulation.

COMMAND AND CONTROL CHARACTERISTICS

Primary Functional Characteristics

The primary control of AHS is provided by the command and control component. The approach for accomplishing the command and control functions will significantly affect how the system is to be implemented. The five major functions of AHS command and control are described.

Traffic Flow Management

Traffic flow management senses conditions that affect traffic flow, determines changes required in that overall traffic flow, and provides overall guidance to AHS traffic through "traffic flow parameters." It is assumed that this function manages traffic after it has entered an AHS; the regulation of traffic entering and exiting the AHS is assumed to be a function of the entry/exit infrastructure component, which is not addressed in this paper.

Sensing of conditions will occur within a fixed geographic area of operation called a region. The size of a region is not defined, but it is assumed to be a reasonably large traffic management segment (e.g., a 100-mi^2 section of an urban area). It is assumed that a region can be segmented into many smaller local areas called zones so that local problems such as construction can be managed locally.

Sensing may be accomplished with organic sensors or through interaction with other IVHS systems or other AHS regions, or both. Conditions to be sensed could include the following:

- System condition: Malfunctions of the region's AHS, including communications and infrastructure malfunctions, must be detected; malfunctions of individual vehicles should also be detected.
- Environment: The roadway conditions, including wetness, temperature, and wind, will need to be monitored; monitoring could include the region's AHS roadway and connecting roadways (AHS and non-AHS); this could include both existing and projected conditions.
- Traffic conditions: Conditions sensed could include rate of flow in AHS lanes in the region and on connecting roadways (AHS and non-AHS), check-in and check-out plaza congestion, abnormalities, and projected traffic in the near term.
- Roadway impediments: AHS roadway impediments may include those that are planned, such as construction, or those that are unplanned, such as accidents; impediments on connecting roadways should also be sensed.

On the basis of the information received from condition sensing, the "traffic flow parameters" for the region as a whole and for specific zones will be constantly adjusted. Traffic flow parameters may include speed per lane segment or zone, spacing between vehicles and platoons, lane restrictions, passing/merging restrictions, and check-in and check-out restrictions (e.g., this plaza is closed). For some malfunctions, the traffic flow parameters may bring the traffic movement in a zone to a temporary halt, or divert it to non-AHS roadways, or both. For some weather conditions, the traffic flow parameters may lower the speed, increase the spacing between vehicles, and actuate special roadway conditioners.

Connectivity within and among the command and control component layers and with other AHS components will be provided by the communications component. It will consist of a variety of regional, wayside, or on-board (or all of these) communications systems. In this paper, it is assumed that the necessary communications links will be established and that the kinds of communications links will not be influential in defining AHS concepts.

Intervehicle Coordination

Intervehicle coordination provides management among vehicles in a vehicle cluster to ensure that crashes do not occur and that the overall flow of the vehicles is smooth. A vehicle cluster could be as small as one vehicle and all the vehicles within its sensing range or as large as sets of platoons of vehicles on multiple AHS lanes within a zone. The control must (*a*) operate within the constraints of the traffic flow parameters determined by the traffic flow management function (i.e., speed and spacing); (*b*) detect and track the relative positions of the vehicles in the cluster and maintain their status; (*c*) detect and be responsive to the needs of each of the vehicles in the vehicle cluster for merging, exiting, and lane changing; (*d*) determine potential openings in the traffic for merging, exiting, or lane changing, or all of these, and plan the maneuver including development of "control profiles" for each of the vehicles involved in the planned maneuver; and (*e*) transmit those control profiles to the vehicle control functions of the respective vehicles and track the vehicles to ensure that they follow the profiles. In theory, the intervehicle coordination function could be accomplished among autonomous, communicating vehicles if the proper cooperative algorithms were developed.

Incident Management

Incident management is, in effect, an extension of the intervehicle coordination function. It will (*a*) detect impending potential incidents, such as a vehicle not maintaining its position (e.g., not staying in its lane); (*b*) determine and calculate necessary controlling and evasive actions, and transmit these actions as incident management profiles to the vehicle control function of all affected vehicles; and (*c*) notify adjacent zones and the regional control center about the problem and the action being taken.

Vehicle Control

Vehicle control for each vehicle provides precise, millisecond-level (*a*) sensing of the vehicle's longitudinal and lateral position and (*b*) direction to the vehicle's actuators so that the vehicle either maintains its longitudinal and lateral position or tracks its control profile for merging, lane changing, or incident management. Depending on the concept, the on-board sensing of position may be a part of the intervehicle coordination function or may supplement that function. The latter case might provide added safety assurance but could lead to conflicts between the two functions that would need to be resolved.

Vehicle Management

For each vehicle, vehicle management maintains overall status and control awareness. It (*a*) monitors all vehicle status such as temperature and fuel supply; (*b*) maintains driver requests such as desired exit; (*c*) calculates and assesses overall management factors, such as destination versus fuel supply or on-board problem indicators, such as loss of traction versus current speed; (*d*) within the context of the overall traffic flow parameters, communicates with the vehicle control function if immediate actions are needed; and (*e*) interacts with the intervehicle coordination function for merge, lane change, entry/exit, or emergency stop requests/demands. Strategies for handling each of these requests will be handled by the intervehicle coordination function because other vehicles may be affected by the requesting vehicle's action.

Primary Physical Characteristics

The AHS command and control component consists of processors and the software that operates on those processors to perform command and control functions. The AHS command and control is viewed as a single entity so that the necessary integration among the various functions can be more easily envisioned; however, its physical design will consist of processing capabilities (processors with backup and environmental support) at a minimum of three different kinds of locations: on-vehicle, zone roadside, and regional control center. The actual location of the software to perform the functions described above may be influenced by the variations possible in the command and control component, as described below.

On-Board Characteristics

The on-board (i.e., vehicle-mounted) processors will have interfaces with the vehicle's actuators and sensors, the on-board communications equipment, and any other on-board IVHS equipment. Because of the critical nature of the AHS processing, redundant sensing and processing may be necessary to provide the necessary levels of reliability and availability; this, too, may translate into space needs. The algorithms, response times, and reliability of the software must be specified to ensure proper, safe operation while in the system.

Zone Roadside Characteristics

It is assumed that the roadside zone sensor and processing capabilities would provide coverage over a relatively small roadway segment (e.g., a few hundred feet in length). For fully autonomous intervehicle coordination, the "zone" would move with the vehicles. It is assumed that the capability is (*a*) able to operate in all weather conditions; (*b*) very reliable and includes self-diagnostic capabilities, and (*c*) unmanned other than for maintenance. These locations must provide protection against the environment, connectivity to required communications channels, enclosures with adequate space for maintenance access, and space for future growth. Because of the critical nature of the AHS processing, redundant sensing and processing may be necessary to provide the necessary levels of reliability and availability; this, too, may translate into space needs.

The in-zone sensing capability may include sensing of (*a*) vehicles entering the zone; (*b*) vehicle position within the zone; and (*c*) roadway condition within the zone. Communications component instrumentation within the zone may include beacons, receivers, and transmitters for communicating with vehicles in zone; networks for communication among zones; networks for communicating with regional traffic management; and communication links with sensors.

Region Characteristics

It is assumed that regional centers will be needed to provide the broad view and control of AHS traffic. A regional AHS center will be able to interact with other AHS regions and other IVHS systems to provide overall, integrated network control. A regional center probably will be manned. It must have provisions for regional command and control processors, command and control center display capability, and communication links to zones, and other regions for network connectivity.

Command and Control Variables

The basic variations in command and control are (*a*) the extent to which the functions are performed on an automated basis and (*b*) the location at which processing occurs; that is how the "intelligence" is distributed (on board the vehicle, zone roadside, or regional) (*11*). It is assumed that all AHS command and control functions are automated (i.e., performed with software or with software assistance to humans). It is also assumed that the command and control component will have total control at all times, with the exception of extreme malfunctions; thus, only in extreme situations would the driver be allowed to take control while in the system. The primary variations at which command and control component functions can be performed are shown in Table 2.

There are practical constraints that help limit the location at which the functional software is performed (Table 2). For example, vehicle management should be performed on the vehicle because the data sensed are mostly on board, and the use of the information is mostly on board. Similarly, many traffic flow management functions should be performed primarily at the regional level. It appears that the primary command and control variables are the location of the intervehicle coordination and incident management functions and, to a lesser extent, some of the vehicle control function.

The following variables can be characterized as types of strategies for accomplishing the vehicle control functions: roadway centered, vehicle-centered, and combined.

Roadway-Centered Control Strategy

When the intervehicle control function is performed at the zone roadside, this strategy is termed roadway centered; that is, the command and control component located at the fixed zone will be aware of and track the multiple vehicles and their interactions in its area of coverage; it will control and relate the movements of those vehicles in accordance with the traffic flow parameters by giving pre-

TABLE 2 Variable Locations for Command and Control Functions

Command and Control Functional Characteristics	Most Likely Location		
	Regional	Zone	On-Board
Traffic Flow Management	X		
Inter-Vehicle Control			
— Subordinate		X	
— Autonomous			X
— Combined		X — *Both* —	X
Incident Management			
— Subordinate	X — *Both* —	X	
— Autonomous			X
— Combined		X — *Both* —	X
Vehicle Control			
— Subordinate (Combined)		X — *Both* —	X
— Autonomous			X
Vehicle Management			X

cise commands to the vehicle control function of those vehicles. A variation of the roadway-centered control strategy, which is not explored in this paper, is the synchronous control strategy. This strategy, also termed "point following," assumes that moving slots are electronically predefined in the lane, and a vehicle moves into a slot, moves along the lane in its designated slot, and moves out of the slot when it exits the system (12). This strategy implies significant roadside sensing and on-board vehicle sensing; it also implies a reliable, higher-bandwidth communications capability from the zone roadside to the vehicles.

Vehicle-Centered Control Strategy

The intervehicle coordination and vehicle control functions could be performed exclusively among cooperative command and control component processors on board a cluster of vehicles. The roadway infrastructure component role would provide information (probably passive) on lane boundaries and other roadway characteristics. The command and control component located at the zone roadside would provide active information to vehicles on roadway conditions; it might also detect overall traffic flow rates and assess levels of congestion.

This strategy implies a higher level of sensing and computation on board the vehicle and probably significantly higher communications capabilities among vehicles in a cluster. Further investigation is needed of the approach for handling incident management and its relative safety and effectiveness.

Combined Strategy

A combination of the roadway-centered and vehicle-centered control strategies is quite feasible; many variations could be defined. Because these variations are not addressed in this paper, they are all grouped together as one strategy. This is an area in which a focused analysis might prove to be fruitful.

CDFs

The variables identified earlier constitute the characteristics, termed CDFs, that distinguish concepts from each other. Specifically, there were four variables among the vehicle and roadway infrastructure components and one variable in the command and control component, for a total of five CDFs, as follows:

- Vehicle class (size and performance);
- Roadway infrastructure interaction (type of interaction between the roadway and vehicle);
- Power source (on-board vehicle or roadway-provided electric vehicles);
- Lateral control strategy (passive embedded markers, passive physical side lane boundary markers, or active embedded markers);
- Vehicle control strategy (i.e:, vehicle control and intervehicle coordination; three alternatives are roadway-centered, vehicle-centered, or combined approaches)

A major variable within an AHS is the longitudinal spacing between vehicles; however, vehicle spacing is considered a concept variation rather than a CDF. This is because the vehicle spacing of most concepts can and will be varied, and this will be done without any major changes to the vehicle or the roadway. Specifically, the AHS spacing at any given time will be determined by (*a*) the operating tolerances of the command and control component design; (*b*) acceptability to the drivers and passengers; and (*c*) the changing traffic flow parameters calculated by the traffic flow management algorithms. A given AHS command and control component design will have limits on the tolerances of spacing that it can safely allow. Over time, these limits may evolve to closer tolerances as electronic components and sensors are produced that are more responsive; however, this evolution should not cause major redesign of the vehicles or the infrastructure—electronic component replacement is not considered a "fundamental difference."

INITIAL CONCEPTS DEFINITION

Alternative AHS Concepts

The CDFs were used to define an initial set of AHS concepts. This was done by examining the various combination of the factors that could be made without regard to whether the combinations are practical. A total of 147 combinations is possible. Theoretically, each of these combinations could be considered an alternative AHS concept. However, the design or implementation or both, of many of these combinations is either highly unlikely or nearly impossible; for example, a pallet system may not be able to accommodate heavy interstate trucks. The concepts are identified by eliminating these unlikely combinations; the resulting number of concepts is 37.

In Table 3, the various potential combinations and concepts are grouped into eight different categories. Six of the categories are created using the vehicle class and lateral control strategy CDFs; the other two categories are special cases—roadway powered and special pallet. The selection of these categories was somewhat arbitrary; different groupings are certainly possible but would result in the same overall number of combinations and concepts. The assumptions made in assessing whether a given combination of CDFs was unlikely or unfeasible are described as follows.

- Wide truck use: Assumes a freeway type of roadway; no pallets would be used; energy provided from on board the vehicle.
- Passive lateral control strategy: The vehicle control will not be pure roadway centered; it will be primarily vehicle-centered control.
- Active lateral control strategy: The vehicle control strategy will not be purely vehicle centered; it will contain some control from the roadside.
- Roadway-provided electric power: All vehicles will be specially designed and will be narrow; the control strategy will not be purely vehicle-centered; and the lateral control strategy will be active, not passive.

Potential Combined Concepts

The 37 alternatives are simplistic; that is, there is no implied mixing of ideas from one alternative to another. For example, it is assumed that a roadway used by large trucks cannot be a pallet system because a pallet system would be all pallets.

The eventual AHS implementation will probably be a combination of a few of the alternatives. For example, pallets could be in-

TABLE 3 Categories and Number of Alternatives in Each Category

Category	Combinations	Concepts
• Wide trucks or smaller, passive lateral control	36	6
• Wide trucks or smaller, active lateral control	3	2
• Normal passenger vehicle or smaller, passive lateral control	36	8
• Normal passenger vehicle or smaller, active lateral control	18	4
• Narrow vehicle only, passive lateral control	12	8
• Narrow vehicle only, active lateral control	6	4
• Narrow vehicle only, roadway-powered	18	4
• Narrow vehicle only, special pallet	18	1
Total combinations	147	37

termixed with nonpallet traffic, including wide trucks. The AHS concept modeling and simulation capability must be able to accommodate these combinations. A few of the potential combinations of these concepts are defined and discussed as examples.

Separate Lanes for Various Vehicle Classes

In this combined approach, all vehicles would be accommodated in the system, but on different lanes where justified. Wide AHS lanes would be provided to accommodate wide trucks and transit vehicles. Smaller vehicles could also use these lanes; this would be the case where traffic density did not justify the separate lanes.

Transition Pallets

Pallets owned by the roadway operator could be intermixed with the other AHS traffic to allow noninstrumented vehicles to use the AHS roadway. The pallet is viewed as a fully instrumented, four-wheeled chassis upon which a noninstrumented vehicle could park—as on a trailer—and be transported through the AHS system; power would be provided by the pallet or conceivably by the vehicle placed on the pallet. Pallet loading and unloading facilities would be needed at entry and exit points.

Intermixed Roadway-Provided Electric Vehicles

In this approach, a special narrow AHS roadway lane would be constructed to provide partial or full power to electric vehicles as they operate on the roadway or to allow operation by narrow vehicles that have on-board fuel. Because the roadway-powered vehicles would need to be specially designed the number of them in a given AHS area may be low initially. This intermixing would allow the lanes to have higher utilization as the number of roadway-powered vehicles slowly increases.

Evolution from Vehicle-Centered to Combined Vehicle Control

Some AHS instrumentation may be available on some vehicles before the AHS actually becomes operational. For example, AHS instrumentation, if properly designed, could conceivably be used for collision avoidance on noninstrumented highways. Options could include autonomous intelligent cruise control; frontal collision warning and avoidance; lane keeping; and lateral collision warning and avoidance. Assuming that standards and specifications had been established in advance for these collision-avoidance features, then these vehicles could possibly operate on an early AHS roadway with little if any added instrumentation.

CONCLUSIONS

On the basis of the study effort summarized in this paper, the following conclusions are drawn:

1. A process can be developed to identify AHS concepts in a structured manner. The structured process developed and used in this paper seems to allow most, if not all, concepts to be identified.
2. There is at least one approach for structuring an AHS into its major components. An initial definition of each component's functional and physical characteristics can be made, and the major variations from one concept to another can be identified. These major variations are termed the concept definition factors; they can be used to identify a set of AHS concepts.
3. By eliminating the unlikely combinations of the concept definition factors, the number of AHS concepts defined in this paper is 37. Eventual AHS deployments probably will be combinations of two or more of these concepts.
4. The 37 concepts are not necessarily a complete and definitive set of AHS concepts; that complete set will be defined as the AHS program proceeds. However, they do provide an adequate basis for scoping and defining the AHS program in this early planning phase.

ACKNOWLEDGMENTS

The material in this paper is based on work reported in a MITRE Corporation technical report and coalesces many thoughts and ideas that have been discussed in meetings with people interested in, or responsible for, the AHS program. In particular, the comments and thoughts of the following people deserve special mention: Lyle Saxton of FHWA, William Leasure and August Burgett of NHTSA, and Richard Bishop and Milton Heywood of FHWA, who contributed many excellent comments.

REFERENCES

1. Public Law 102–240, December 18, 1991, *Intermodal Surface Transportation Efficiency Act* of 1991 (ISTEA).
2. Rao, B., et al. Investigations into Achievable Capacities and Stream Stability with Coordinated Intelligent Vehicles. Presented at 70th Annual Meeting of the Transportation Research Board, Washington, D.C., 1993.
3. *Our Nation's Highways—Selected Facts and Figures,* FHWA-PL-92-004. FHWA, U.S. Department of Transportation, 1992.
4. *The Status of the Nation's Highways and Bridges: Conditions and Performance*. FHWA, U.S. Department of Transportation, 1992.
5. *Human Factors Design of Automated Highway Systems (AHS).* MITRE Corporation, McLean, Va., Sept. 1992.
6. Stevens, W. *Precursor Systems Analyses for the AHS Program.* Working Paper WP 92W0000382. The MITRE Corporation, McLean, Va., 1992.
7. *Broad Agency Announcement for Precursor Systems Analyses for the Automated Highway System (AHS) Program.* U.S. Department of Transportation, Nov. 1992.
8. Stevens, W. *Automated Highway Systems Concepts Analysis.* Technical Report MTR 93W0000123. The MITRE Corporation, McLean, Va., Aug. 1993.
9. Bender, J. An Overview of Systems Studies of Automated Highway Systems. *IEEE Transactions on Vehicular Technology,* Vol. 40, No. 1, IEEE Vehicular Technology Society, New York, Feb. 1991.
10. Wilson, D. G. et al. *Quadramode Transport: A Class of Controlled Systems.* Industrial Liaison Symposium on Urban Transportation, May 1971.
11. Shladover, S. E. et al. Automatic Vehicle Control Developments in the PATH Program. *IEEE Transactions on Vehicular Technology,* Vol. 40, No. 1, IEEE Vehicular Technology Society, New York, Feb. 1991.
12. Fenton, R. E. et al. Automated Highway Studies at The Ohio State University. *IEEE Transactions on Vehicular Technology,* Vol. 40, No. 1, IEEE Vehicular Technology Society, New York, Feb. 1991.

This paper was prepared for FHWA but does not necessarily express FHWA policy.

Publication of this paper sponsored by Committee on Intelligent Vehicle Highway Systems.

Models of Commuters' Information Use and Route Choice: Initial Results Based on Southern California Commuter Route Choice Survey

MOHAMED A. ABDEL-ATY, KENNETH M. VAUGHN, RYUICHI KITAMURA, PAUL P. JOVANIS, AND FRED L. MANNERING

A statistical analysis of commuters' route choice behavior and the influence of traffic information is presented. The analysis is based on a 1992 computer-aided telephone interview survey of Los Angeles area morning commuters. Cross tabulations were performed on the data to explore interrelationships among variables and provide a basis for subsequent model estimation. Two sets of models were estimated: bivariate probit models of whether individuals follow the same route to work every day and whether they receive traffic information (pretrip or en route) and negative binomial models of the frequency of route changes per month on the basis of pretrip and en route traffic reports. The estimation results underscore the important relationship between the use of traffic information and the propensity to change routes. In addition, important relationships are uncovered relating to the influence that commuters' socioeconomic characteristics and the level of traffic congestion they face has on traffic information use and route change frequency.

The problem of route choice for a commute trip can be defined as choosing the best route through the transportation network, in terms of some criterion or criteria, while facing temporal (i.e., departure and arrival times) and geographic (i.e., origin and destination) constraints. This best route most often is thought of as the one that minimizes travel disutility (e.g., travel time, distance, or generalized travel cost). In reality, the problem of route choice faced by an automobile driver is complex because of the large number of possible alternative routes through the road network and the complex patterns of overlap between the various route alternatives (*1*).

In recent years, an abundance of research has focused on commuters' route choice with an emphasis on how real-time traffic information might affect this choice (*2–6*). In an ongoing Partners for Advanced Transit and Highways (PATH) project at the University of California (UC) Davis entitled ATIS (Advanced Traveler Information Systems) Impact on Travel Demand, a variety of issues about traveler response to information are being investigated [see, for example, previous work (*7–9*)]. These earlier papers focused on development of learning models of drivers' adaptation to traffic advice, particularly when the advice is not always reliable. A second part of the project deals with studying the actual route choices of drivers, with the objective of developing refined route choice models that can include the effect of traveler information. Understanding route choice behavior is essential to improving network assignment methods and to investigating ATIS effectiveness (e.g., how much information drivers have or need or how information affects route choice behavior). This paper is concerned with the second part of the project.

M. A. Abdel-Aty, K. M. Vaughn, R. Kitamura, and P. P. Jovanis, Institute of Transportation Studies, 2028 Academic Surge, University of California at Davis, Davis, Calif. 95616. F.L. Mannering, Department of Civil Engineering, 121 More Hall FX-10, University of Washington, Seattle, Wash. 98915.

To probe into drivers' route choice behavior, a telephone survey of Los Angeles area morning commuters was conducted as part of the project. The survey was designed to investigate how much information drivers have about their routes, their awareness of alternate routes, their awareness of traffic conditions that could affect their route choices, and their use of available traffic information either en route or pretrip. The survey, undertaken in May and June 1992, is differentiated from those of previous studies in that the specific routes taken by individuals were obtained for their morning commute.

This paper describes the survey design and administration. General descriptive statistics are also introduced to show the characteristics, preferences, and perceptions in commuters' route choice behavior. Bivariate probit models of traffic information use and propensity to use alternative routes are also developed. In addition, negative binomial models are used to assess frequency of commuters' route changes on the basis of traffic reports, route characteristics, and individual characteristics. Further details about the survey itself and additional descriptive statistics are contained in a project report (*10*).

SURVEY AND SAMPLE DESCRIPTION

A route choice/traffic information survey was developed to target Los Angeles area morning commuters. A mail-out/mail-back survey instrument was initially designed to gather detailed information on commuters' main and alternate routes, to determine the level of information commuters have about these routes, to measure commuters' attitudes toward, and perceptions of, these routes, and to determine how existing traffic information affects their route choice behavior. The mail survey instrument required several branchings, increasing its level of complexity and potentially jeopardizing the response rate and response accuracy. Therefore, it was decided to perform a computer-aided telephone interview (CATI) survey. A CATI survey allows interviewer–respondent interaction and automatically handles branchings with complete reliability and lower interviewer error. The survey targeted a random sample of adult commuters residing in the area covered by the South Coast Air Quality Management District, which includes most of the contiguously populated areas of Los Angeles, Orange, San Bernardino, and Riverside Counties. The sampling, based on a Mitofsky-Waksberg cluster sampling design (*11*), covered both listed and unlisted numbers. The Mitofsky-Waksberg sampling is known to reduce the number of unproductive dialings and improves efficiency (*12*). In all, 944 commuters were surveyed in May and early June 1992. Summary statistics for the sample are presented in Table 1. The values shown in this table are within expected bounds.

To test the representativeness of the sample, several socioeconomic and commute characteristics were compared and statistically tested with the 1990 Census (*13*), the 1991 California Statewide Travel Survey results (CSTS) (*14*), and the 1990 California Statistical Abstract (*15*). In most cases the null hypothesis that the values from the route choice survey are not different from the corresponding statistical sources was not rejected at the 0.05 level of significance, implying that the sample well represents the population in the study area. Further information on this and other tests can be found elsewhere (*10*).

Traffic Information Use

As Table 1 shows, the survey provided some interesting insight into travelers' use of traffic information and their choice of route. In the survey, traffic information questions were divided into two groups, depending on where the information is received, either before (pretrip) or while (en-route) driving to work. About 36.5 percent of the respondents listen to traffic reports before leaving their homes, and 51.25 percent listen while driving. Close to 27.6 percent of the respondents listen to traffic reports both at home and en route, and 60.1 percent listen to reports either at home or en route, whereas 39.9 percent never listen to reports. These findings are consistent to a great extent with those of Khattak et al. (*2*). Most respondents who receive traffic information perceive traffic reports to be either very accurate or somewhat accurate.

More women (40 percent) listen to traffic reports before leaving home to work than men (33 percent), whereas more men (54.5 percent) listen to reports en route than women (47.7 percent). The hypothesis of no differences between sexes was rejected using Pearson chi-square at a 0.05 level of significance. It was also found that more women change their routes or departure times as a result of listening to traffic reports before leaving their homes, whereas men change their routes more frequently than women as a result of traffic reports they hear while driving to work. However, it is possible that socioeconomic or commute characteristics, or both, associated with gender are the cause of such differences between men and women and not gender itself.

Commuters who use freeways may be more likely to receive traffic information if their freeway traffic conditions are perceived as heavy or very heavy. The relationship was confirmed (using a chi-square test) for pretrip information but not for en-route information. This suggests that commuters try to find out their freeway conditions in advance, possibly because these are the segments of their routes that are exposed most to delays or because they realize that their diversion options, once they get onto a freeway, are very limited, or both.

Route Choice Behavior

Only 15.5 percent of the respondents said they use more than one route to work. Considering the well-developed freeway network in the study area, this may be considered a low percentage. However it indicates that an information system that would make people aware of alternative routes has promising potential. About 50 percent of the respondents had at least one freeway segment in their primary routes (a primary route is the route that the respondent uses most frequently), and 38 percent had at least one freeway segment in their secondary routes (Figure 1); secondary routes tend to have more surface streets than primary routes, possibly as alternatives used to avoid congestion on freeways. The percent of freeway users in the CSTS data is 46.3 percent, which is very close to the results of the present study. Even for an area that is generally considered saturated with freeways, 50 percent of the primary routes involves no freeway at all.

Finally, it is interesting that the most frequent reason for changing routes, cited by 34 percent of respondents, is the traffic that the respondents see on the roads. The need to make stops on the way

TABLE 1 Sample Summary Statistics

• Commute distance on usual route (miles)	12.75
• Travel time on usual route (minutes)	28.14
• Trip duration (including stops)	31.9
• Percent of respondents commuting in single-occupant autos/carpool/public transit	78.8/14.6/4.9
• Percent receiving pre-trip traffic reports	36.5
• Percent receiving en-route traffic reports	51.25
• Percent of respondents with flexible/ somewhat flexible / fixed work starting time	24.4/30.4/45.2
• Percent male/female	51.3/48.7
• No. of household cars	2.31
• No. of years at present address	7.24
• No. of years at present job location	5.52
• Percent own/rent their homes	59/41
• Household income	38,750
• Percent of college graduates	43.8
• Think traffic congestion is a problem or major problem (percent)	61.3
• Think trip time uncertainty is a problem or major problem (percent)	31.9

Note: Values are averages unless noted otherwise.

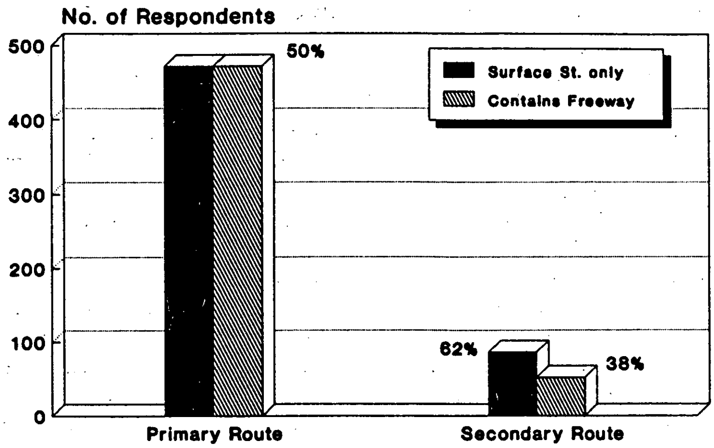


FIGURE 1 Composition of primary and secondary routes by road type.

and traffic reports follow (15.5 and 14 percent, respectively). Additional reasons include the time of day (8 percent) and the day of the week (5.5 percent). If the percent of respondents that base their choice on the traffic they see is added to others who base their choice on traffic reports, then about 50 percent of the commuters depend on real-time information for choosing their routes. This finding reiterates the potential of an ATIS system to alleviate traffic delays.

Individuals with higher incomes tend to report using more than one route to work. The fraction of individuals with alternative routes (percent of multiple route users within each income category) increases from 6.7 percent among those with incomes less than $25,000 to 28 percent among those with incomes more than $100,000. The null hypothesis of independence between income and using alternative routes is rejected. Khattak et al. (*2*) also found that higher-income drivers were more likely to take alternate routes. The same relationship is also found for level of education: highly educated people tend to use alternate routes.

MODELING APPROACH

To assess commuters' propensity to change routes and acquire traffic information, the study focused on the joint decision of whether commuters follow the same route to work every day and whether they receive traffic information (pretrip or en route). The objective is to examine the association between information use and route choice and to verify the results of the cross-tabulation analysis in multivariate modeling contexts. For such a joint decision, the bivariate (two-dimensional) probit formulation is appropriate. Commuters' frequency of route changes on the basis of traffic information is then modeled using negative binomial regression models. The modeling effort reported in this paper represents a preliminary analysis of the interplay of information use and route choice. The variables considered in model development include the attributes of main commute routes, attributes of commuters, and their perception of traffic conditions. Future work could extend the range of variables to include objectively measured traffic characteristics for the respective commuters' main and alternative routes. Figure 2 summarizes the modeling effort presented.

Joint Estimation of Route Switching and Information Choices

There is a need to identify the factors that lead a commuter to use single or multiple routes to work and to receive traffic information. Gaining an understanding of this issue will aid in how traffic conditions and other factors affect the use of traffic information and route switching. In particular, building a model that predicts route-switching behavior as a function of information use will aid in evaluating the potential impacts of ATIS on route choice.

Methodological Approach

The commuters' choice of receiving traffic information and their use of alternate routes are likely to be interrelated. As such, there is a likely correlation of unobserved effects (between information use and route choice) which if not accounted for, would lead to biased model coefficient estimates. An example of such unobserved correlation would be the tendency of a commuter to be "adventurous" and "dynamic." Clearly such a tendency would be difficult to quantify (and therefore likely to show up in model error terms), but adventurous and dynamic commuters would be expected to be much more likely to receive traffic information and to change routes. This would produce a positive correlation in error terms that must be accounted for. An appropriate model for capturing this correlation is the bivariate probit.

The bivariate probit model can be used directly to identify the contributing factors that influence route switching behavior and affect the likelihood of receiving traffic information. In this case, the two choices are (*a*) whether the respondent receives traffic information ($Y_1 = 0$ or 1), and (*b*) whether the respondent uses more than one route to work ($Y_2 = 0$ or 1). These two choices can be represented by the following simultaneous equation system:

$$Y_1^* = \beta X_1 + \epsilon$$

$$Y_1 = \begin{cases} 1 & \text{if } Y_1^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

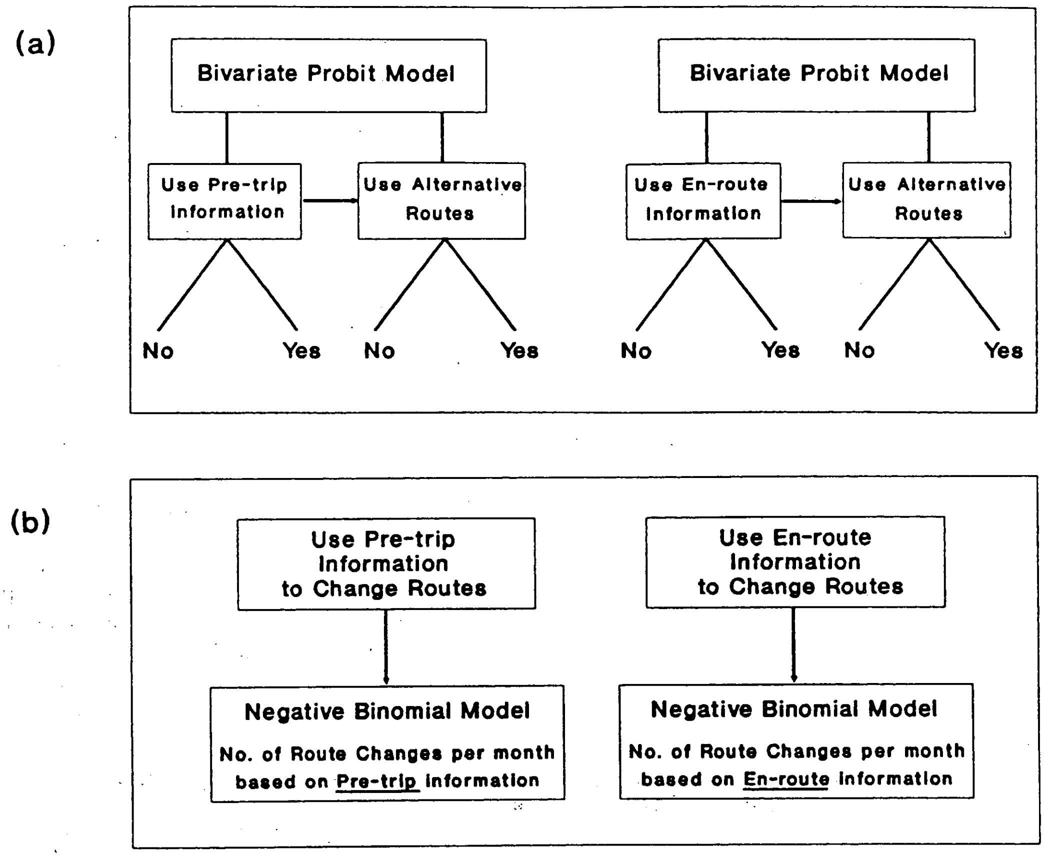


FIGURE 2 **Modeling structure.**

$$Y_2^* = \alpha X_2 + \Theta Y_1 + \xi$$

$$Y_2 = \begin{cases} 1 & \text{if } Y_2^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where

Y_1^* = latent variable indicating the respondent's propensity to listen to traffic information;
Y_1 = observed choice (1 if the respondent listens to information and 0 otherwise);
Y_2^* = latent variable indicating the respondent's propensity to use multiple routes;
Y_2 = observed choice (1 if the respondent is a multiple route user, and 0 if exactly one route is used every day to work);
β, α = coefficient vectors;
Θ = scalar coefficient;
X_1, X_2 = vectors of explanatory variables influencing choice behavior; and
ϵ, ξ = random error terms.

Assuming that ϵ and ξ are correlated [$E(\epsilon\xi) \neq 0$], then the two equations should be estimated simultaneously using the full-information maximum likelihood (FIML) or sequentially using the limited-information maximum likelihood (*16,17*). If a limited-information approach is adopted, parameters are estimated in one equation at a time with instrumental variables (*10*) or correction terms (*18*) introduced to account for error correlation. For a linear system, these techniques provide consistent but inefficient estimates of parameters (*16*). However, in a system of two binary-choice equations, as is the case in this study, these approaches may lead to inconsistent estimates [numerical comparisons of alternative estimators are given by Kitamura (*19*)]. The FIML is desirable because it offers consistent and efficient estimates while accounting for possible error correlation across equations. Thus, FIML is adopted in this study.

Distributional assumptions need to be made on the random error terms ϵ and ξ to express response probabilities. A probit model offers a theoretically sound formulation for discrete responses. Adoption of the probit formulation in a situation involving two binary-choice endogenous variables would imply that the joint distribution of ϵ and ξ is given by the bivariate standard normal distribution, with var(ϵ) = var(ξ = 1 for normalization.

For this system of equations (i.e., Equations 1 and 2), the full-information likelihood function for the bivariate probit is developed by first defining sample strata as follows:

S_1: $Y_1 = 1$ $\quad Y_2 = 1$

S_2: $Y_1 = 1$ $\quad Y_2 = 0$

S_3: $Y_1 = 0$ $\quad Y_2 = 1$

S_4: $Y_1 = 0$ $\quad Y_2 = 0$

The likelihood function for the first set of observations, S_1, is derived by considering the joint probability of the event, $Y_1 = 1$ and $Y_2 = 1$:

$$Pr[Y_1 = 1, Y_2 = 1] = Pr[Y_1^* \geq 0, Y_2^* \geq 0]$$
$$= Pr[\epsilon \geq -\beta X_1, \xi \geq -\alpha X_2 - \theta]$$
$$= \int_{-\beta X_1}^{\infty} \int_{-\alpha X_2 - \theta}^{\infty} f(\epsilon, \xi)\, d\epsilon\, d\xi$$

where f is the standard bivariate normal density function:

$$f = \frac{2\pi}{\sqrt{(1 - \rho^2)}} \exp\left[-\frac{(\epsilon^2 - 2\rho\epsilon\xi + \xi^2)}{2(1 - \rho^2)}\right] \quad (3)$$

and ρ is the correlation coefficient between ϵ and ξ.

The likelihood function for this set of observations is

$$L_1 = \prod_{S_1} \int_{-\beta X_1}^{\infty} \int_{-\alpha X_2 - \theta}^{\infty} f(\epsilon, \xi)\, d\epsilon\, d\xi \quad (4)$$

Similarly L_2, L_3, and L_4 can be derived. Therefore, the likelihood function for the entire sample is

$$L = \prod_{S_1} \int_{-\beta X_1}^{\infty} \int_{-\alpha X_2 - \theta}^{\infty} f(\epsilon, \xi) d\epsilon\, d\xi$$
$$\prod_{S_2} \int_{-\beta X_1}^{\infty} \int_{-\infty}^{-\alpha X_2 - \theta} f(\epsilon, \xi)\, d\epsilon\, d\xi$$
$$\prod_{S_3} \int_{-\infty}^{-\beta X_1} \int_{-\alpha X_2}^{\infty} f(\epsilon, \xi)\, d\epsilon\, d\xi$$
$$\prod_{S_4} \int_{-\infty}^{-\beta X_1} \int_{-\infty}^{-\alpha X_2} f(\epsilon, \xi) d\epsilon\, d\xi \quad (5)$$

Parameter vectors β, α, θ, and ρ are estimated so as to maximize L. The statistical significance of the coefficient θ will indicate whether state dependence is present. Also, significant error correlation between ϵ and ξ (ρ) will indicate the presence of unobserved individual factors (heterogeneity) that affect both choices of route and receiving information.

Estimation Results for the Bivariate Probit Models

Two bivariate probit models were developed after investigating several alternative model formulations. The first estimates whether respondents often receive traffic reports before leaving home to work (pretrip) and whether they are multiple-route users. The second estimates whether respondents often receive traffic reports while driving to work (en route) and whether they are multiple-route users (the whole sample is used in estimating these models).

Estimation results for the pretrip information/multiple-route user model are given in Table 2. All variables included are self-explanatory and their coefficients are readily interpretable. The pretrip information model indicates that people who perceive no variation in traffic conditions on their usual commute route are less likely to listen to pretrip traffic reports. Women, long-distance commuters, or respondents who reported uncertainty in travel time as a major problem, or all of these, are more likely to listen to these

TABLE 2 Bivariate Probit Model Estimating Whether Respondents Receive Traffic Reports Before Leaving Home to Work and Whether They Are Multiple-Route Users

	Coefficient	t-statistic
PRE-TRIP INFORMATION MODEL		
Constant	-0.416	-3.79
X_1 No variation in traffic conditions dummy (1 if no variation is perceived, 0 otherwise)	-0.361	-3.68
X_2 Female dummy (1 if female, 0 otherwise)	0.110	1.15
X_3 Uncertainty of travel time dummy (1 if reported that trip time uncertainty is a major problem, 0 otherwise)	0.436	3.23
X_4 Distance from home to work	0.013	3.40
MULTIPLE ROUTE MODEL		
Constant	-2.033	-6.95
X_5 Income dummy (1 if income ≥ $75,000, 0 otherwise)	0.302	2.43
Y_1 Receiving pre-trip information dummy (1 if receive pre-trip information, 0 otherwise)	1.002	2.74
X_6 No. of driving days in the last 2 weeks	0.032	1.26
X_7 Level of education dummy (1 if respondent is a college grad. or completed some college, 0 otherwise)	0.409	2.55
Error-term Correlation	-0.518	-2.38

<u>Summary Statistics</u>
Log Likelihood at zero = -1061.761
Log Likelihood at market share = -790.804
Log Likelihood at convergence = -758.191
Likelihood ratio index = 0.286
Number of observations = 733
Percent correct predicted = 72%

Note: Variables' coefficients are defined for receiving reports and multiple route use

TABLE 3 Bivariate Probit Model Estimating Whether Respondents Receive Traffic Reports While Driving to Work and Whether They are Multiple-Route Users

	Coefficient	t-statistic
EN-ROUTE INFORMATION MODEL		
Constant	-0.303	-2.82
X_1 No variation in traffic conditions dummy (1 if no variation is perceived, 0 otherwise)	-0.244	-2.42
X_2 College graduate dummy (1 if respondent is a college grad, 0 otherwise)	0.195	2.00
X_3 Uncertainty of travel time dummy (1 if reported that trip time uncertainty is a major problem, 0 otherwise)	0.708	4.51
X_4 Distance from home to work	0.026	6.57
MULTIPLE ROUTES MODEL		
Constant	-2.061	-5.92
X_5 Income dummy (1 if income ≥ $75,000, 0 otherwise)	0.306	2.33
Y_1 Receiving en-route information (1 if receive en-route information, 0 otherwise)	0.531	1.66
X_6 No. of driving days in the last 2 weeks	0.035	1.28
X_7 Level of education dummy (1 if respondent is a college grad. or completed some college, 0 otherwise)	0.415	2.44
Error-term Correlation	-0.174	-0.82

Summary Statistics
Log Likelihood at zero = -1061.761
Log Likelihood at market share = -815.902
Log Likelihood at convergence = -762.335
Likelihood ratio index = 0.282
Number of observations = 733
Percent correct predicted = 84.9%

Note: Variables' coefficients are defined for receiving reports and multiple route use

reports. These findings are consistent with those of an earlier analysis of the data.

For the multiple-route choice model, high income (≥ $75,000), high level of education (college graduate or completed some college), and the number of days driving to work in 2 weeks increase the likelihood of using multiple routes. The positive coefficient of receiving pretrip information indicates that commuters who receive this information are more likely to use more than one route to work, whereas the significance of the variable indicates the important effect of Y_1^* on Y_2^*. The significance of the correlation between the two error terms underscores the importance of accounting for cross-equation correlation. The negative sign indicates the presence of unobserved factors that reversely affect the two behavioral aspects. Note that there is no expectation with respect to the sign of the error correlation. For example, a cautious and well-prepared commuter would try to obtain as much information as possible before departure (positive ϵ) but choose to adjust the departure time rather than venture onto unfamiliar alternate route (negative ξ).

Estimation results for the en-route information/multiple-route user model are given in Table 3. The model is similar to the previous model, except that gender is replaced by a college graduate dummy, which significantly increases the likelihood that a respondent receives en-route traffic reports. The positive coefficient of receiving en-route information (Y_1) indicates that commuters who receive en-route information are more likely to use more than one route to work, although this variable is not highly significant (only at the 90 percent level of significance). The correlation between the two error terms is insignificant.

Frequency of Changing Routes On the Basis of Information

To assess commuter frequency in changing routes on the basis of traffic information, an appropriate statistical modeling technique is needed. The Poisson regression was initially attempted, but chi-square tests indicated significant differences between the estimated and observed route switching frequencies. Also, the Poisson distribution was rejected because the mean and variance of the dependent variables are different, indicating substantial overdispersion in the data (number of route changes based on pretrip information: mean, 1.69, variance, 7.34; number of route changes based on en route information: mean, 1.46, variance, 6.61). Such overdispersion suggests a negative binomial model. The negative binomial model is an extension of the Poisson regression model and allows the variance of the process to differ from the mean.

Methodological Approach

This section on methodological approach is drawn from Greene (*20*). The negative binomial model arises from the Poisson model by specifying

$$\ln \lambda_i = \beta X_i + \epsilon \tag{6}$$

where

- λ_i = parameter giving individual i's expected route-changing frequency;
- β = vector of estimable parameters;
- X_i = vector of commuting and socioeconomic characteristics for individual i; and
- ϵ = error term, where $\exp(\epsilon)$ has a gamma distribution with mean 1 and variance α^2. The resulting probability distribution is as follows:

$$Pr[Y = y_i \mid \epsilon] = \exp[-\lambda_i \exp(\epsilon)]\lambda_i^{y_i}/y_i! \tag{7}$$

where y_i is the number of route changes, and all other variables are as previously defined. Integrating ϵ out of this expression produces the unconditional distribution of y_i. The formulation of this distribution is

$$Pr[Y = y_i] = \Gamma(\theta + y_i)/[\Gamma(\theta)y_i!]\, u_i^{\theta}(1 - u_i)^{y_i} \tag{8}$$

where

$P[Y = y_i]$ = probability of commuter i making y_i changes in a specified period of time.
$u_i = \theta/(\theta + \lambda_i)$
$\theta = 1/\alpha$

Compared with the Poisson model, this model has an additional parameter α, such that

$$\mathrm{Var}[y_i] = E[y_i]\{1 + \alpha E[y_i]\} \tag{9}$$

This is a natural form of overdispersion in that the overdispersion rate is

$$\mathrm{Var}[y_i]/E[y_i = 1 + \alpha E[y_i] \tag{10}$$

Such an approach is well suited to modeling frequency of route change because it accounts for the no-change option ($y_i = 0$) as well as all other possible non-negative integer outcomes (*21*). The negative binomial model can be estimated by standard maximum likelihood methods.

Testing for Existence of Selectivity Bias

Before proceeding with the estimation of the negative binomial models, it is important to test for possible selectivity bias. Selectivity bias could be present if the commuters observed to be using traffic information as a basis for changing routes were a self-selected group with route change behavior that systematically differed from those commuters not observed to be using information as a basis for changing routes. Such selectivity creates a problem because frequency data have been collected only on those individuals observed to be using information for changing routes. If their behavior systematically defers from those not observed changing routes, the estimates of β will be biased.

Selectivity bias correction methods in standard regression equations have been derived by other researchers (*22,23*). However, developing corrective techniques for count data (i.e., on the basis of a negative binomial regression) has not been done and is likely to be a difficult task because a closed-form expression for the expected value of the gamma error term (see Equation 6) conditioned on the bivariate probit error terms must be developed [i.e., $E(\epsilon \mid \epsilon\xi)$]. Such a formulation is beyond the scope of this paper. However, a suggestive test of this matter using a standard discrete–continuous selectivity bias correction procedure (*24*) was conducted. In doing so, the bivariate probit model (of whether or not information is used) with a simple independent binary logit model and the negative binomial regression model (of the frequency of route changes) with a standard regression model were approximated.

Formalizing this, the utility, to respondent i, of using traffic information, U_{il}, can be written as

$$U_{il} = \beta_1 X_1 + \epsilon_i \tag{11}$$

where

- β_1 = vector of estimable parameters,
- X_1 = vector of factors influencing information use, and
- ϵ_i = Gumbel distributed error term.

These variables give rise to the binary logit formulation

$$P_{il} = 1/[1 + \exp(-\beta_1 X_1)] \tag{12}$$

where P_{il} is the probability of respondent i using information.

For the regression equation of the frequency of route changes conditioned on the use of information, I, the following is given:

$$E(y_i \mid I) = \beta_2 X_2 + E(\psi_i \mid I) \tag{13}$$

where

- β_2 = vector of estimable parameters,
- X_2 = vector of factors influencing the frequency of route choice (y_i), and
- ψ_i = a normally distributed error term.

Selectivity bias arises because of unobserved effects [i.e., $E(\epsilon_i\psi_i) \neq 0$]. To correct this bias in the estimation of route change frequency, an expression for the conditional expectation of the error term [i.e., $E(\psi_i \mid I)$] is needed. Dubin and McFadden (*23*) have shown this to be

$$E(\psi_i \mid I) = \rho(\sigma_\psi/\sigma_\epsilon)\,\pi_{il} \tag{14}$$

where

- σ_ψ = standard deviation of the normally distributed error term ψ_i,
- σ_ϵ = standard deviation of the logistic error term (from Equation 12),
- ρ = partial correction coefficient for ψ and ϵ,
- $\pi_{il} = [1/P_{il}][P_{il}\log(P_{il}) + (1 - P_{il})\log(1 - P_{il})]$, and
- P_{il} = probability of respondent i choosing to use information I. Thus Equation 13 becomes

$$E(y_i \mid I) = \beta_2 X_2 + \omega\pi_{il} \tag{15}$$

where ω is an estimable coefficient equal to $\rho(\sigma_\psi/\sigma_\epsilon)$. The significance of the coefficient, ω, associated with the selectivity correction term, π_{il}, gives a measure of the importance of selectivity bias in the equation.

The ω coefficient terms in both regression models (number of times per month changing routes on the basis of pretrip information and number of times per month changing routes on the basis of en route information) were statistically insignificant (pretrip information model $\omega = -0.001$, t-statistic $= -0.795$; en route information model $\omega = -0.001$, t-statistic $= -1.280$), suggesting that selectivity bias is not present. It is therefore concluded that estimating the negative binomial models without possible error correlation between the bivariate probit and the negative binomial is not likely to be a significant source of error.

Estimation Results for Negative Binomial Models

Two negative binomial models were developed: the first modeled the number of route changes per month on the basis of listening to pretrip traffic reports, and the second modeled the number of route changes per month on the basis of listening to en-route traffic reports. The estimation results for the first model are illustrated in Table 4. The results show that commuter's perceptions have an important effect on the number of route changes; that is, if respondents perceive substantial variation in traffic conditions from day to day on their primary route, they are likely to make more route changes per month. If it is perceived to be accurate, information will have a positive effect on the number of changes per month; dummy variables representing individual report values (e.g., 1, 2, 3) were attempted but the results showed that the relationship was linear, and therefore, a simple ordering of responses was used.

Turning to the socioeconomic factors, a high level of education (e.g, college graduates) was found to have a positive impact on the number of route changes per month. Also, from the commute characteristics, the log of travel time on the most frequently used route has a positive impact on the number of route changes per month, indicating that longer commutes make travelers more likely to change routes. A possible explanation can be that time-consuming commutes lead to a greater awareness and use of alternate routes, on the basis of pretrip information. However, the log transformation indicates that this effect diminishes with increasing travel time. Finally, the significance of the overdispersion parameter (α) indicates that the negative binomial formulation is preferred to the more restrictive Poisson formulation.

The second model (the frequency of route changes per month on the basis of en-route information) is presented in Table 5. The results show that the carpool dummy has a positive effect on the number of route changes per month on the basis of en-route traffic reports. It appears that once carpoolers are together on the road, en-route information influences their decision to change routes. A perception of substantial traffic variation and bad traffic conditions on the usual route increased the frequency of route changes. Also, the perception that information is accurate has a positive effect (again, dummy variables representing report accuracy and traffic conditions were attempted but a linear relationship was found; therefore, the ordered responses were used). Individuals' perception of reality is important because it ultimately drives their behavior, which indicates that accurate traffic information is vital for commuters who perceive variations or bad traffic conditions on changing routes.

The model also shows that freeway users tend to change routes more frequently on the basis of en-route information, possibly as a means to avoid congestion. The positive coefficient of the log of commute distance depicts that longer distances cause route changes on the basis of en-route information. The use of the log transformation indicates that this effect is nonlinear, with marginal increases in distance playing a stronger role in shorter commutes. Again, the significance of the overdispersion parameter (α) shows that the negative binomial formulation is a preferred specification.

TABLE 4 Negative Binomial Model: Frequency of Route Changes Per Month to Work On the Basis of Pretrip Reports

	ß	t-statistic
β_0 Constant	-2.735	-2.43
X_1 Perceived Variation in traffic conditions dummy (1 if traffic conditions are substantially different from day to day on the usual commute route, 0 otherwise)	0.752	1.50
X_2 Perceived accuracy of traffic reports (1 not at all accurate, 2 not very accurate, 3 somewhat accurate, 4 very accurate, 5 extremely accurate)	0.362	2.42
X_3 College graduate dummy (1 if college graduate, 0 otherwise)	0.354	1.48
X_4 Log driving time on last trip using the usual route	0.507	2.19
α overdispersion parameter	2.065	5.90

Summary Statistics
Log Likelihood at zero = -833.809
Log Likelihood at convergence = -415.779
ρ^2 = 0.501
Number of observations = 238

TABLE 5 Negative Binomial Model: Frequency of Route Changes Per Month to Work On the Basis of En-Route Reports

	ß	t-statistic
$ß_0$ Constant	-2.385	-4.59
X_1 Carpool dummy (1 if mode is carpool, 0 otherwise)	0.473	1.68
X_2 Perceived Variation in traffic conditions dummy (1 if traffic conditions are substantially different from day to day on the usual commute route, 0 otherwise)	0.617	1.61
X_3 Rating of traffic conditions (1 very good, 2 good, 3 OK, 4 bad, 5 very bad)	0.250	2.66
X_4 Perceived accuracy of traffic reports (1 not at all accurate, 2 not very accurate, 3 somewhat accurate, 4 very accurate, 5 extremely accurate)	0.297	2.91
X_5 Freeway user dummy (1 if uses fwy, 0 otherwise)	0.420	1.56
X_6 Log commute distance	0.190	1.40
α overdispersion parameter	2.149	7.97

Summary Statistics
Log Likelihood at zero = -1426.647
Log Likelihood at convergence = -675.750
ρ^2 = 0.526
Number of observations = 443

SUMMARY AND CONCLUSIONS

This paper uses a CATI survey carried out as part of a research project at UC Davis. This survey was designed to gain a basic understanding of drivers' route choice behavior, to collect detailed information about their commute routes, and to explore how commuters use traffic information to decide on what routes to travel to work.

An analysis using general descriptive statistics showed several tendencies in the commuters' route choice decisions. Only 15.5 percent of the respondents reported that they do not always follow the same exact route to work, which indicates a potential benefit from an information system that would make more commuters aware of alternative routes.

The following were cited as the most common reasons for changing from a primary route: the desire to decrease the trip time, receiving traffic reports, and the time the commuters leave their homes. High income and a high level of education were two sociodemographic factors correlated with the use of more than one route. Other factors, such as the commute distance, did not seem to have a significant effect on using alternative routes.

Finally, the statistical exploration of the data also indicated that gender influences the use of traffic information. Women tend to listen to pretrip traffic reports more frequently than men and tend to use freeways less frequently than men.

Bivariate probit models were developed to determine the factors that influence information use and the propensity to use alternative routes. The models showed the significant influence of income, education, frequency of driving to work, and listening to traffic reports on the commuters' route choice. Also, perceived variation in traffic conditions, gender, commute distance, and travel time uncertainty affected the likelihood of listening to traffic information.

Negative binomial models were developed to assess commuters' frequency in changing routes. Two models were developed: the first modeled the number of route changes per month on the basis of pretrip traffic reports and the second modeled the number of route changes per month on the basis of en-route traffic reports. The models showed the significant effect that commuters' perceptions of the accuracy of traffic reports and variation in traffic conditions, travel time, and the level of education had on the frequency of changing routes on the basis of pretrip information. Also, traffic conditions, perceptions of information accuracy and traffic variation, freeway use, commute distance and carpool, were among the variables influencing the frequency of route changes on the basis of en-route traffic information.

The findings of this study suggest at least two important directions for future research. The first direction is methodological in nature. There is a need to develop an FIML procedure for estimating simultaneously the commuter's choice to use information and the frequency of route changes. This task will not be easy because of the complexity of the error term structure, but there are potentially many applications of such a procedure to the analysis of route choice behavior and other ATIS-related concerns.

The second direction relates to the need for information on the specific routes used by travelers. Such information would include highway geometrics, signal timings, and the temporal distribution of traffic. Although it is often tedious and time consuming to process, this information can be gathered and used to explore many detailed relationships that will have a direct impact on ATIS utilization.

ACKNOWLEDGMENTS

The authors wish to acknowledge the comments and suggestions of four anonymous referees. Their recommendations resulted in a substantially improved paper. They also thank Thomas Golob, of the University of California at Irvine, for his significant contribution to the design of the survey. Special thanks to the California Department of Transportation (Caltrans) and PATH for funding this research. The models presented in this paper were estimated using LIMDEP econometric software.

REFERENCES

1. Antonisse R., A. Daly, and M. Ben-Akiva. A Highway Assignment Method Based on Behavioral Models of Car Drivers' Route Choice. In *Transportation Research Record 1220,* TRB, National Research Council, Washington, D.C., 1989.
2. Khattak, A., J. Schofer, and F. Koppelman. Effect of Traffic Reports on Commuters' Route and Departure Time Changes. *Proc., Vehicle Navigation and Information Systems Conference,* Part 2, Society of Automotive Engineers, Warrendale, Pa., Oct. 1991, pp. 669–679.
3. Hatcher, G. and H. Mahmassani. Daily Variability of Route and Trip Scheduling Decisions for the Evening Commute. In *Transportation Research Record 1357,* TRB, National Research Council, Washington, D.C., 1992.
4. Haselkorn, M., J. Spyridakis, and W. Barfield. *Improving Motorist Information Systems: Towards a User-Based Motorist Information System for the Puget Sound Area.* Final Report. Washington State Transportation Center, University of Washington, Seattle, March 1990.
5. Haselkorn, M., J. Spyridakis, and W. Barfield. Surveying Commuters to Obtain Functional Requirements for the Design of a Graphic-Based Traffic Information System. *Proc., Vehicle Navigation and Information Systems Conference,* Part 2, Society of Automotive Engineers, Warrendale, Pa., Oct. 1991, pp. 1041–1044.
6. Mannering, F., S-G. Kim, W. Barfield, and L. Ng. Statistical Analysis of Commuters' Route, Mode, and Departure Time Flexibility. *Transportation Research,* Vol. 2C, No. 1, 1994, pp. 35–47.
7. Vaughn, K., M. Abdel-Aty, R. Kitamura, P. Jovanis, and H. Yang. Experimental Analysis and Modeling of Sequential Route Choice Behavior Under ATIS in a Simplistic Traffic Network. In *Transportation Research Record 1408,* 1993, pp. 75–82.
8. Yang, H., R. Kitamura, P. Jovanis, K. Vaughn, and M. Abdel-Aty. Exploration of Route Choice Behavior with Advanced Traveler Information using Neural Network Concepts. *Transportation* 20, 1993, pp. 199–223.
9. Abdel-Aty, M., K. Vaughn, P. Jovanis, and R. Kitamura. Understanding Traveler Responses to ATIS. *Proc., 26th ISATA Conference,* Aachen, Germany, Sept. 1993.
10. Abdel-Aty, M., K. Vaughn, R. Kitamura, and P. Jovanis. *Survey of Route Choice Behavior: Empirical Results From Southern California and Their Implications for Advanced Traveler Information Systems.* Research Report UCD-ITS-RR-93-12. Institute of Transportation Studies, University of California at Davis, 1993.
11. Groves, R., P. Biemer, L. Lyberg, J. Massey, W. Nicolls, and J. Waksberg. *Telephone Survey Methodology,* John Wiley and Sons, New York, 1988.
12. Frey, J. *Survey Research By Telephone,* Sage Publications, Beverly Hills, 1983.
13. Bureau of the Census. *1990 Census of Population and Housing, California.* U.S. Department of Commerce, Economics and Statistics Administration, Washington D.C., June 1992.
14. *1991 Statewide Travel Survey: Summary of Findings,* California Department of Transportation, Sacramento, Nov. 1992.
15. *California Statistical Abstract 1990,* Department of Finance, Sacramento, California, Oct. 1990.
16. Maddala, G. *Limited-Dependent and Qualitative Variables in Econometrics,* Cambridge University Press, Cambridge, England, 1983.
17. Kitamura, R., R. Pendyala, and K. Goulias. Weighting Methods for Choice Based Panels with Correlated Attrition and Initial Choice. In *Transportation and Traffic Theory* (C.F. Daganzo, ed.), Elsevier Science Publishers, The Netherlands, 1993, pp. 275–294.
18. Heckman, J. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement,* Vol. 5, 1976, pp. 475–492.
19. Kitamura, R. A Review of Dynamic Vehicle Holdings Models and a Proposal for a Vehicle Transaction Model. *Proc., Japanese Society of Civil Engineers,* 440/IV-16, 1992.
20. Greene, W. *LIMDEP: User's Manual and Reference Guide, Version 6.0.* Econometric Software, Inc., Bellport, N.Y., 1992.
21. Mannering, F. Poisson Analysis of Commuter Flexibility in Changing Routes and Departure Times. *Transportation Research,* Vol. 23B, No. 1, 1989, pp. 53–60.
22. Heckman, J. Sample Selection Bias as a Specification Error. *Econometrica,* Vol. 47, No. 1, 1979.
23. Dubin, J., and D. McFadden. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica,* Vol. 52, 1984.
24. Mannering, F. "Selectivity Bias in Models of Discrete and Continuous Choice: An Empirical Analysis. In *Transportation Research Record 1085,* TRB, National Research Council, Washington, D.C., 1986, pp. 58–62.

Publication of this paper sponsored by Committee on Transportation Supply Analysis.

Influence of Traffic Information on Drivers' Route Choice Behavior

AMALIA POLYDOROPOULOU, MOSHE BEN-AKIVA, AND ISAM KAYSI

Commuters' route choice behavior in the presence of traffic information is analyzed. A modeling framework for the acquisition and processing of pretrip and en-route information, drivers' route switching behavior, and the willingness to pay for more useful information is proposed. The estimation of these models is based on revealed preference data obtained from a 1991 survey of commuters to Massachusetts Institute of Technology. Trip characteristics and travelers' perceptions of the relevance and reliability of radio traffic reports were found to be important factors affecting radio traffic information acquisition and its influence on drivers' decisions. The key finding was that en-route diversion is primarily influenced by attitudinal factors and by information acquisition. Moreover, drivers' own observations are important factors affecting route switching. It can be concluded that a reliable and frequently updated traffic information system will stimulate the acquisition of traffic information and affect route diversion.

Increasing attention has been paid in recent years to the use of advanced traveler information systems (ATIS) for alleviating traffic congestion. ATIS can be classified into pretrip information services such as television and radio and telephone and route-planning information services; and en-route information services such as traffic information broadcasting services, telephone information services, in-vehicle navigation, route guidance systems, and variable-message signs. By collecting and transmitting real-time information on traffic conditions and transit schedules, ATIS possess the capability to improve the efficiency of the traffic system.

Figure 1 shows the potential influence of information on travelers' pretrip and en-route decisions. At the beginning of each trip, travelers decide on whether to travel or not to travel; the location of their destination; and their departure time, mode, and route choice. Travelers' decisions are influenced by two types of information: experience-based and real-time information. Experience-based information is acquired by actual traveling; it is subjective, limited, and imperfect because it reflects average values and cannot foresee daily variations in travel times. On the other hand, real-time information provided by ATIS gives travelers the ability to more accurately predict travel times to their destinations.

The information provided by these systems may lead travelers to decide to alter or postpone their trip or to choose a route other than the habitual one. The provision of public transport and parking information may influence their mode, departure time, and destination choices. If travelers acquire dynamic en-route information, they may decide to revise their preselected travel pattern by switching route, destination, or mode. In subsequent trips, travelers' decisions to acquire pretrip and en-route information, and to review their choices, will be based on previous travel experiences and the expected attributes of new travel choices.

This paper investigates the influence of traffic information on drivers' route choice behavior. Specifically, the impact of factors such as drivers' socioeconomic characteristics, travel characteristics, and information characteristics on the following decisions is examined:

- Pretrip traffic information acquisition;
- Pretrip route choice;
- En-route traffic information acquisition; and
- En-route switching decisions.

This paper consists of six sections. The second section discusses the major findings of recent research on route choice and route switching behavior. The third section presents the general framework of route choice behavior as influenced by traffic information. The fourth section describes the data collection method and presents descriptive data analysis results. The fifth section presents model specification and estimation results. The sixth section concludes the paper and offers suggestions for further research.

PREVIOUS RESEARCH

Revealed preference data and stated preference data are the two basic approaches of data collection for modeling users' decisions. The next two subsections describe data collected and models estimated by various researchers on the basis of each of the two approaches. An extensive review of the state of the art of route choice models is provided by Bovy and Stern (*1*).

Revealed Preference Approach

The revealed preference approach analyzes drivers' behavior in real-life situations, on the basis of respondents' reports (usually diaries of actual trips), on previous actions, or by observing traveler behavior in real-life situations (field study approach).

Diary Survey

Khattak et al. (*2*) used revealed preference data to estimate drivers' diversion decisions. The results showed that drivers prefer to stay on the usual route and are more likely to divert after receiving delay information from radio traffic reports than through self-observation. Attitudinal factors included in the models, indicating the drivers' inherent tendency to divert and risk propensity, were found to be significant. Hatcher and Mahmassani (*3*) addressed the day-to-day

A. Polydoropoulou and M. Ben-Akiva, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 02139. I. Kaysi, Department of Civil Engineering, American Institute of Beirut, P.O. Box 11-0236, Beirut, Lebanon.

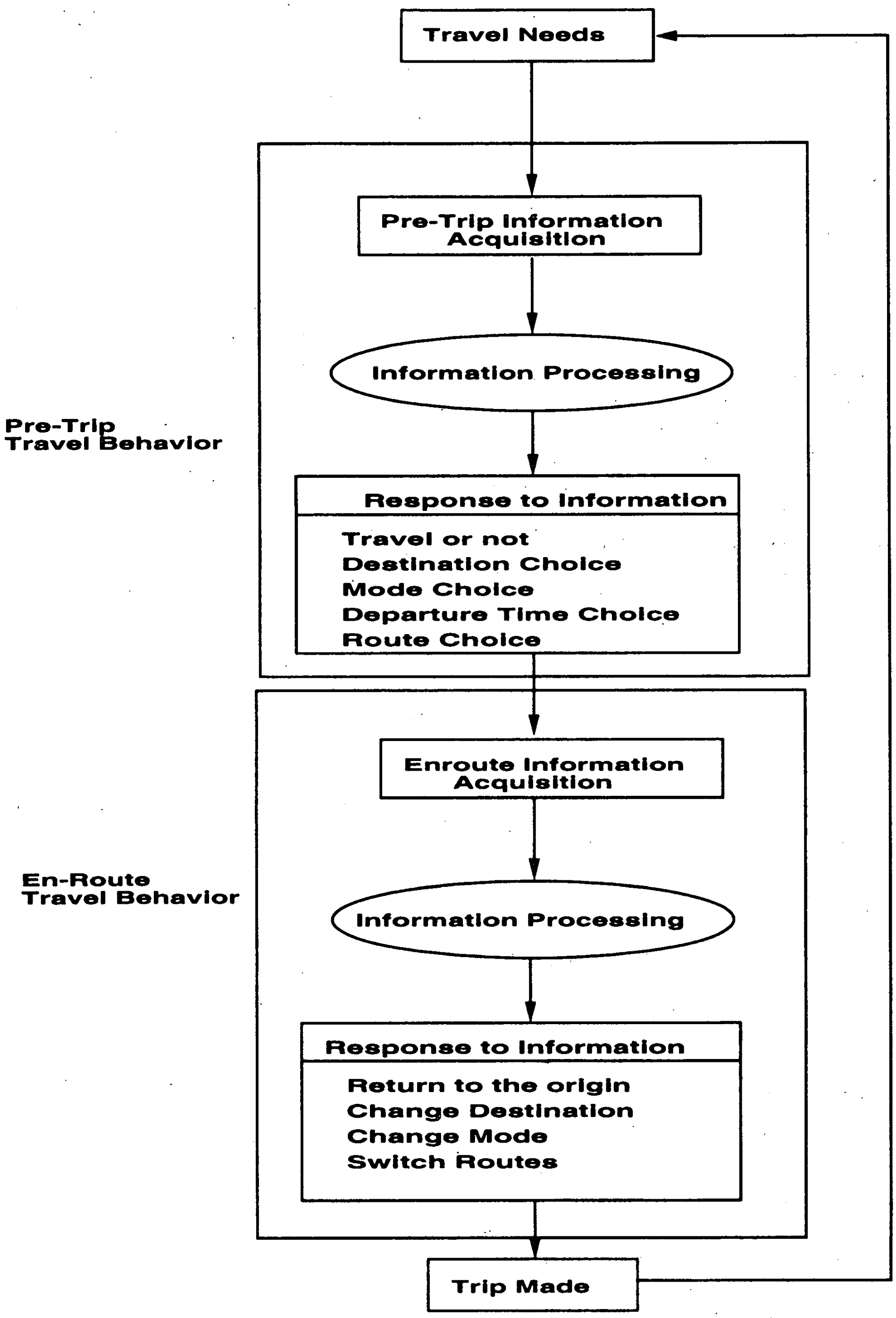


FIGURE 1 Impact of traffic information on travelers' decisions.

variation of individual trip scheduling and route decisions for the evening commute. It was found that trip chaining significantly influences route switching behavior and that commuters tend to change departure times more frequently than they change routes.

Cascetta and Biggiero (*4*) estimated models for departure time and route choice for home-to-work trips. The analysis indicated that travel time spent on secondary roads played an important role in route choice. Safety and comfort variables also were found to be significant.

Field Study

The field study approach analyzes drivers' behavior through field observation of drivers, such as observation of actual diversion behavior in response to information acquisition. Although a number of operational tests of ATIS are being conducted, no data have been available for the estimation and calibration of drivers' behavior models (*5*). However, this approach holds significant potential for providing critical revealed preference data, and efforts are under way to extract useful information from a number of operational tests already in existence or being planned.

Stated Preference Approach

Two different approaches of extracting stated preference data on potential driver behavior in hypothetical situations, namely, by surveys and by simulators are described.

Surveys

Khattak et al. (*6*) used stated preference data to evaluate the drivers' willingness to divert from their usual route to an alternate route. Drivers expressed a higher willingness to divert if expected delays on their usual route increased, if they experienced travel times that were longer than usual, and if congestion was incident induced, as opposed to recurring. Respondents were less willing to divert if the alternate route was unfamiliar or unsafe or had several traffic stops. Socioeconomic characteristics were found significant in predicting willingness to divert. The results of a computer-based survey, conducted by Polak and Jones (*7*) to study the impact of in-home pretrip traffic information on travelers' behavior, showed that even among regular car users, there is a demand for multimodal pretrip information.

Simulators

A review of recent efforts to collect data using travel simulators to study and model travelers' behavior in the presence of information is provided by Koutsopoulos et al. (*8*). Bonsall and Parry (*9*) developed an interactive route choice simulator to investigate drivers' compliance with route guidance advice. Estimations of regression models showed that the acceptance of advice depended on its credibility, which was a function of past experience, local conditions, and psychological factors. Koutsopoulos et al. (*10*) developed a travel simulator and used concepts from fuzzy sets theory, approximate reasoning, and fuzzy control to model route choice process and the drivers' perceptions in the presence of information. It was found that traffic conditions on one path might affect the perceived attractiveness of an alternative path. Preliminary analysis of data obtained from a simulator developed by Adler et al. (*11–13*), showed that en-route diversion behavior is influenced by the familiarity of drivers with the potential alternative routes and their traffic conditions, the information provided, changes in travel speeds, and drivers' risk preference. Vaughn et al. (*14*) and Yang et al. (*15*) also used data from a travel simulator to model sequential route choice behavior. A logit model showed that as the perceived delay on a route increases, the probability that this route is chosen decreases. Furthermore, as the accuracy of the system increases, the probability of following the advised route increases. A second modeling approach using concepts from neural networks revealed that most subjects make route choices on the basis of their most recent experiences. Therefore, the relative accuracy of the information provided plays an important role on the immediate choices. Finally, Chen and Mahmassani (*16*) developed a travel simulator that is connected to a traffic simulation model. This simulator allows multiple users to drive through the network, interact with each other, and influence the systems performance.

FRAMEWORK FOR ROUTE CHOICE BEHAVIOR

This section presents a general framework for route choice behavior, which describes the choice factors and choice structure, as well as the potential impacts of both pretrip and en-route information on drivers' decisions.

Choice Factors

The ultimate choice—that is the route to be taken—is the result of the following factors: (*a*) driver characteristics, (*b*) travel characteristics, and (*c*) information characteristics.

The traveler's route choice depends on socioeconomic characteristics, such as age, gender, income, personality, habits, preference, driving experience, and familiarity with the transportation network. Travel characteristics (such as trip purpose, flexibility in arrival time, availability of alternatives) and traffic conditions associated with the travel alternatives play a significant role in determining the route choice behavior. Constraints imposed by the purpose of the trip also are important factors. Moreover, each traveler has only limited knowledge of all available routes.

Information attributes play an important role in defining the attitudes of travelers toward information (*17*). Information must be reliable and accurate, or else travelers will tend to reject further information acquisition and develop a negative attitude toward information acquisition. Only information provided in a timely manner that is responsive to changing traffic conditions and relevant to the travelers' trips will positively affect travelers' decisions and their attitudes toward information acquisition.

The decision-making process is dynamic because of the feedback from each trip. A learning process is central to the driver's cognition as the information acquired through experience of earlier travel choices is processed before the next decision is made. Moreover, the characteristics of each known alternative route do not have the same importance in a driver's final decision. On the basis of a factor importance hierarchy, the traveler formulates a choice set of sufficiently attractive alternatives (*1*). From this set travelers make their choices, with the chosen route being the one that best satisfies their needs and is consistent with their personal constraints and preferences. Finally, travel choice inertia also plays a role, dictating that certain thresholds be crossed before drivers change their habitual behavior (*2*).

Impact of Pretrip Information on Route Choice

Drivers may adjust their destination, mode, departure time, and route choices on the basis of information acquired before beginning a trip. If they decide not to acquire pretrip information, drivers will rely on their historical perceptions and experiences and, therefore, start their trip following the habitual route.

When information is acquired by drivers, the perceived importance of indicated traffic conditions, combined with the drivers' general attitudes and preferences, influence their pretrip route choice decisions. The derived benefits from acquiring and utilizing information to make pretrip decisions will then serve as feedback to drivers that updates their attitudes and preference toward pretrip information acquisition. Therefore, pretrip decisions for subsequent trips will depend on their perceived benefits on previous trips.

Impact of En-Route Traffic Information on Route Choice

Drivers' attitudes toward traffic information acquisition will play an important role in their en-route travel choices. If drivers do not acquire pretrip traffic information, their route choice will be based on their past experiences. On the other hand, if information is acquired, their route choice might be modified accordingly. En-route information could be acquired either passively (from en-route observations or variable-message signs) or voluntarily. Voluntary en-route information acquisition usually takes place when the level of service on the preselected route is different from what was anticipated; the driver might seek to acquire traffic information for both the actual and alternate routes. After traffic information is acquired, drivers may choose to ignore it (follow preselected route) or to divert. The perceived benefits of drivers' actual decisions are expected to affect the updating of their attitudes toward the acquisition of and response to traffic information as well as their future route choice behavior.

DATA COLLECTION AND ANALYSIS

In this study revealed preference data were obtained from a diary survey of Massachusetts Institute of Technology (MIT) commuters conducted in the spring of 1991 (*18,19*).

Survey Instrument

The survey included the following two parts.

Part 1:

Part 1 contained four groups of questions on the usual commuting trip to MIT, the drivers' socioeconomic characteristics, and the drivers' attitudes and preferences. The questions about the commuting trip were about the usual departure and arrival times, flexibility in arrival time, traffic conditions usually encountered, number of alternative routes used, and the duration and purpose of stops made. The second group of questions sought information regarding the driver's sex, marital status, education, income, profession, number of years at current dwelling unit, and current job.

The third group of questions utilized a five-point scale to indicate the driver's level of agreement with the following factors:

1. Statements that indicate the familiarity of drivers with the network, such as "I am very familiar with at least two significantly different routes to work";
2. Statements that reflect the general attitudes of drivers toward diverting, such as "I like discovering new routes or I often change routes while driving", and
3. Statements that revealed the perceptions of drivers toward the validity and effectiveness of traffic reports, such as "Radio traffic reports are usually reliable" or "I often change my route after listening to radio traffic reports."

Finally, the fourth group of questions indicated the importance of several factors in choosing the route to work. For these statements, a five-point scale was used to indicate how drivers rate the importance of factors such as time of day, commuting time, habit, traffic reports, risk of delay, and weather conditions in choosing their commuting route.

Part 2:

Part 2 consisted of a detailed 1-week diary of morning commute trip information. A questionnaire was included to be completed on each of the 5 days. The questions related to pretrip and en-route traffic information acquisition, their influence on the commuters decisions, the driver's en-route diversion decisions, and the anticipated and observed traffic conditions. The questionnaire also asked about the duration and purpose of stops made.

A related set of questions in this part of the survey used a five-point scale to indicate agreement with day-specific perceptions about the commuting trip, such as "Traffic conditions today were better than usual"; "I am satisfied with my route choice today"; and "Traffic information received today was useful." In the next subsection an overview of the statistics obtained from the conducted survey is presented.

Summary of Results

A total of 1,300 individuals responded to Part 1 of the survey. Of these, 898 completed Part 2 of the survey and reported 3,218 trips. The average range of commuting time was between 29 and 40 min. The average number of routes used was 1.6, in spite of the fact that more than 75 percent of the drivers were very familiar with more than two routes. Among the respondents, 14 percent had no flexibility in arrival time, 17 percent had flexibility up to 15 min, 17 percent had flexibility of 16 to 30 min, 15 percent had flexibility from 31 to 60 min, and 37 percent had flexibility of more than 1 hr.

A total of 22 percent of the commuters usually made stops while commuting to MIT, whereas 78 percent never made stops. The average duration of stops was 2.2 min. Approximately 38 percent of the stops had as a purpose dropping off a passenger, whereas only 5 percent of the stops were made to pick up a passenger. Almost 60 percent of the respondents had a postgraduate degree. Approximately 28 percent were administrative staff, 20 percent support staff, 26 percent faculty, and 8 percent students.

Attitudes and Preferences

Table 1 presents the answers of the statements reflecting the attitudes and preferences of drivers. Category 1 corresponds to the response "strongly disagree," whereas Category 5 corresponds to the response "strongly agree." Although 63 percent of the drivers rarely or never change their planned route while driving, 16 percent often make such a change. From the sample of drivers, 37 percent indicated that they often listen to radio traffic reports, and 27 percent usually follow the recommendations. Only 25 percent of the drivers think that radio traffic reports are reliable, whereas 22 percent consider them as not relevant. Among the drivers who listen to radio traffic reports, 20 percent often change their routes after listening, whereas 50 percent completely ignore traffic reports when they are different from their observations. In the latter case, only 9 percent continue to follow the radio's advice. Among all the surveyed commuters, 10 percent were willing to pay to get more useful information, whereas 68 percent were not.

TABLE 1 Respondents' Attitudes and Preferences

Number	Statement	Strongly disagree 1 %	2 %	3 %	4 %	Strongly agree 5 %	Not relevant %
1	I am very familiar with at least 2 different routes to work	7.3	4.4	5.0	6.4	74.5	2.4
2	I often change my planned route while driving	41.0	21.6	18.8	7.9	8.0	2.5
3	I like discovering new routes	22.4	11.6	25.0	15.6	18.0	7.3
4	I am willing to try new routes to avoid traffic delays	5.2	5.5	13.4	22.4	50.5	3.0
5	I always listen to radio traffic reports	24.4	13.5	21.0	12.1	25.2	3.9
6	I usually follow the recommendations of radio traffic reports	19.1	15.5	21.9	15.8	11.4	16.3
7	Radio traffic reports are usually reliable	8.5	14.7	30.9	17.3	7.9	20.7
8	When traffic reports are different from my own observation I ignore them	5.1	5.4	16.2	18.9	31.3	23.1
9	I often change my route after listening to radio traffic reports	13.6	19.5	25.9	14.0	6.4	20.7
10	I trust my own judgment more than the traffic reports	8.1	13.8	26.5	20.2	16.2	15.2
11	Traffic reports do not provide relevant information	16.9	23.6	23.7	11.7	10.1	14
12	I am willing to pay in order to get more useful traffic information	54.5	13.9	12.5	5.3	3.4	10.4

Route Choice Factors

Table 2 presents the importance of different travel attributes when drivers choose their route to work. Category 1 corresponds to the response "Not important at all," whereas Category 5 corresponds to "Very important." Almost 61 percent of the drivers perceive time of day as a very important factor in choosing their route to work. However, 76 percent of the drivers perceive commuting time as the most important factor in their route choice. Other factors identified as important in route choice process were risk of delay (57 percent), habit (50 percent), radio traffic reports (19 percent), and weather conditions (38 percent).

Commuting time is stated to be the most important route choice factor. At the same time, the small percentage of travelers who attach importance to radio traffic reports can be attributed to the perceived poor reliability and lack of relevance of these reports. Therefore, a route guidance system that provides accurate and relevant instructions would be a more useful information tool for the drivers.

Daily Commute Characteristics

A summary of the results of the second part of the survey, which reflected daily commute characteristics, is reported in Table 3. Pretrip traffic information was acquired in 27 percent of the trips made. A total of 16 percent of the drivers was influenced by this information. En-route traffic information was acquired in 24 percent of the trips made. Of the trips made, 67 percent presented the opportunity for

TABLE 2 Importance of Factors Affecting Route Choice Behavior

Number	Attribute	Not important 1 %	2 %	3 %	4 %	Very important 5 %
1	Time of day	15.7	7.1	11.1	16.3	44.9
2	Commute time	9.2	4.6	10.0	20.6	55.5
3	Habit	12.3	8.6	29.8	26.1	23.3
4	Time spent stopped in traffic	5.3	5.3	10.3	29.0	50.1
5	Number of traffic lights	10.9	11.5	24.2	25.4	27.9
6	Traffic reports	30.7	22.7	28.1	12.4	6.1
7	Risk of delay	8.6	8.5	26.1	30.6	26.2
8	Weather	25.9	16.7	19.3	15.9	22.1

TABLE 3 Traffic Information and Route Choice Behavior

Question		Percentage
Received pre-trip traffic information		
	From radio	22
	From TV	5
	No	73
Pre-trip information influence route choice		
	A lot	4
	Somewhat	12
	Very little	12
	Not at all	68
Received en-route traffic information		
	Yes	24
	No	76
Received traffic information before switching for new route		
	Yes	8
	No	92
Reason for switching		
	Radio Traffic Reports	12
	Own observation	62
	Forced detour	5
	Other	22

switching routes; route switching was actually observed in 6 percent of these trips. Among the trips in which route switching was observed, 8 percent received traffic information before switching. The reasons for switching were split as follows: 12 percent because of radio traffic reports, 62 percent because of drivers' own observations, and the remaining 26 percent for other reasons.

MODEL SPECIFICATION AND ESTIMATION RESULTS

This section presents the estimation results from modeling the acquisition of pretrip traffic information, the influence of this information on drivers' behavior, the acquisition of en-route information, route switching decisions, and the willingness to pay for more useful traffic information.

As discussed in Section 4, 3,218 trips were made by 898 individuals over a 5-day period. The totality of these trips was used to estimate discrete choice models using standard maximum likelihood estimation (MLE) techniques [see Ben-Akiva and Lerman (*20*) for further discussion of discrete choice models]. Because choices over time were observed, the dynamic complications of serial correlation were present. By using the standard statistical packages, obtained parameter estimates would be consistent but not efficient. The Jackknife method (*21*) was used to calculate the correct standard errors of the estimated coefficients. This method gives a nonparametric estimation of the standard errors.

The modeling approach consisted of two stages. In the first stage, it was assumed that the attitudes and preferences of drivers are related to explanatory variables such as socioeconomic characteristics, perceptions about information characteristics, and trip characteristics. To model the above relation, ordered response models were estimated in which the dependent variables are the five-point scale responses that reflect the attitudes and preference of the drivers [see Polydoropolou (22) for more details].

In the second stage, driver behavior was modeled using binary choice models. However, these models incorporated the fitted values of the attitudinal and preference variables obtained from the ordered response models. In that way, any notion of endogeneity appearing in the model was avoided. Note that since revealed preference data were used in the model estimations, the inclusion of reported attitudes as independent variables in the choice models could have rendered endogeneity implications. Finally, to avoid another endogeneity problem, the fitted probabilities of the pretrip and en-route information acquisition were used as explanatory variables, when included in the en-route information acquisition and route switching models, respectively.

The data used in the model estimations may not be representative of the population at large because the sample population was restricted to MIT commuters. However, the estimated models presented are indicative of the underlying travel decision-making process.

Modeling Acquisition of Pretrip Information

Binary logit models were used to model the acquisition of pretrip traffic information. For such models the dependent variable was 1 if a commuter acquired pretrip information and 0 otherwise.

The estimated model parameters are presented in Table 4. If travel time often exceeds its usual range, then drivers are more likely to acquire pretrip information. Moreover, the longer the commuting trip is, the more prone drivers are to acquire pretrip traffic information. On the other hand, if the drivers have more than a 16-min flexibility in arrival time, they are less likely to acquire pretrip traffic information.

Driver perceptions about the reliability and relevance of radio traffic reports also influence pretrip traffic acquisition. When traffic reports are perceived as unreliable or irrelevant, drivers are less likely to acquire pretrip traffic information. This result validates the results of most travel simulator studies that investigate the effects

TABLE 4 Acquisition of Pretrip Traffic Information

	Variable number	Variable name	Coefficient estimate	t statistic
	1	Constant	-1.442	-4.74
Trip Characteristics	2	Travel time often exceeds its usual range	0.654	4.53
	3	Average travel time (max)	0.06	2.65
	4	Flexibility > 16 min.	-0.386	-2.86
Perceptions & attitudes	5	Radio traffic reports are not reliable	-0.231	-3.88
	6	Radio traffic reports are relevant	0.597	6.13

Summary statistics
Number of observations =3218
$\mathcal{L}(\mathbf{0}) = -2230.5$
$\mathcal{L}(\hat{\beta}) = -1799.5$
$-2[\mathcal{L}(\mathbf{0}) - \mathcal{L}(\hat{\beta})] = 862$
$\rho^2 = 0.193$
$\bar{\rho}^2 = 0.191$

of information reliability on travelers' behavior [see, for example, previous work (*9,10,14,15*)]. Furthermore, it indicates that ATIS should provide reliable and relevant information to be accepted by the users.

Modeling Influence of Pretrip Traffic Information

To model the influence of pretrip information, only trips made by drivers who acquired pretrip traffic information were taken into account (728 trips) (see Table 5). Binary logit models were estimated in which the dependent variable was 1 if the driver was influenced by the radio reports and 0 otherwise.

The reliability of traffic reports, as well as the willingness of drivers to try different routes to avoid traffic congestion, are significant factors. The less reliable traffic reports are, the less likely this information is to influence driver decisions. Moreover, the more often drivers have a tendency to change their preselected routes, the more likely they are to be influenced by the acquired information.

If information indicated that traffic conditions were better than usual on the preselected route, drivers were likely to decide to stay on their route. On the other hand, if information indicated that the traffic conditions were worse than usual on the preselected route, drivers were likely to use an alternative route. Therefore, in both cases pretrip information had an influence on the drivers' decision about the route to follow. The results also indicate that information positively influences drivers' behavior only when it is relevant to their trip patterns.

Modeling Acquisition of En-Route Information

The acquisition of en-route traffic information is also a binary logit choice model. The dependent variable is 1 if a commuter acquired en-route information and 0 otherwise (see Table 6).

If travel time often exceeds its usual range drivers are more likely to acquire en-route traffic information. The longer the maximum travel time, the more drivers tend to acquire en-route information. Drivers who consider traffic reports to be relevant are more willing to acquire en-route information. A driver who acquired pretrip information is more likely to acquire en-route information. Finally, if the conditions encountered are worse than expected, drivers are more prone to acquire en-route traffic information.

TABLE 5 Influence of Pretrip Information

	Variable number	Variable name	Coefficient estimate	t statistic
	1	Constant	-1.464	-9.6
Perceptions & Attitudes	2	Radio traffic reports are not reliable	-0.598	-2.5
	3	Often change planned route while driving	0.519	2.6
Pre-trip Information	4	Indicated traffic conditions : none	-1.432	-3.5
	5	Indicated traffic conditions worse than usual	0.607	2.4
	6	Indicated traffic conditions better than usual	0.878	3.1

Summary statistics
Number of observations = 728
$\mathcal{L}(\mathbf{0}) = -504.61$
$\mathcal{L}(\mathbf{c}) = -370.31$
$\mathcal{L}(\hat{\beta}) = -345.16$
$-2[\mathcal{L}(\mathbf{0}) - \mathcal{L}(\hat{\beta})] = 318.90$
$\rho^2 = 0.3160$
$\bar{\rho}^2 = 0.3041$

TABLE 6 Acquisition of En-Route Information

	Variable number	Variable name	Coefficient estimate	t statistic
	1	Constant	-3.598	-16.26
Trip Characteristics	2	Travel Time Often Exceeds its Usual Range	0.412	2.75
	3	Usual Travel Time (max)	0.031	6.77
Perceptions	4	Traffic Reports are Relevant	0.253	1.89
Pre-trip Information	5	Acquired pre-trip information	3.439	2.89
Actual Traffic Conditions	6	Observed Conditions Worse than Usual	0.158	1.90

Summary statistics
Number of observations = 3218
$\mathcal{L}(\mathbf{0}) = -2230.5$
$\mathcal{L}(\hat{\beta}) = -1598.9$
$-2[\mathcal{L}(\mathbf{0}) - \mathcal{L}(\hat{\beta})] = 1263.2$
$\rho^2 = 0.283$
$\bar{\rho}^2 = 0.280$

Modeling En-Route Switching Behavior

A binary logit model was used to model the influence of en-route information. The dependent variable is 1 if drivers switched from the preselected route and 0 otherwise (see Table 7).

The negative sign of the constant shows the tendency of drivers to follow their preselected route. A "switch" is a deviation from their habitual behavior. The importance of risk of delay is a significant factor in the route choice behavior. As the importance drivers attach to it increases they become more likely to divert. When travelers are under time pressure, they try to avoid traffic congestion by switching to alternative routes. Moreover, drivers who often change their preselected routes while driving are more likely to divert from the preselected route. This factor indicates an important attitudinal component in the drivers' diversion decisions: a driver with a risk-taking attitude is more likely to switch to another route than one who prefers following the same route and does not like changes. The above results coincide with those of Khattak et al. (*2*) and Adler et al. (*11*).

The observed traffic conditions are an important factor in the estimation results in Table 7. The results indicate that traffic conditions that are worse than usual encourage travelers to divert. Furthermore, when a driver acquires traffic information about the alternate route and this information indicates traffic conditions that are worse than usual, drivers are less likely to switch to the new route. This result was also obtained by stated preference studies about the expected delays on alternative routes and switching decisions [see, for example, previous work (*6, 13–15*)].

Modeling Willingness to Pay for More Useful Traffic Information

An ordered probit model was used for modeling the subjects' willingness to pay for more useful information. The dependent variable took values from 1 to 5, with 1 indicating a strong disagreement for paying for more useful information and 5 indicating a strong agreement for paying for more useful information (see Table 8).

Age has a significant effect on drivers' willingness to pay. Older drivers are more willing to pay than younger drivers. Income also appears significant because drivers with a yearly income of more than $80,000 are more willing to pay than drivers with less income.

TABLE 7 Route Switching Behavior

	Variable number	Variable name	Coefficient estimate	t statistic
	1	Constant	-2.869	-19.42
Perceptions & Attitudes	2	Risk of Delay Importance	2.236	-2.16
	3	Often Change Planned Route While Driving	2.235	2.16
Actual traffic Conditions	4	Observed Conditions at the Beginning of the Trip Worst than Usual	1.223	7.276
En-Route Information for New Route	5	Indicated Conditions for Alternative Route Worse than Usual	5.611	5.42

Summary statistics
Number of observations = 3218
$\mathcal{L}(\mathbf{0}) = -2230.5$
$\mathcal{L}(\hat{\beta}) = -702.28$
$-2[\mathcal{L}(\mathbf{0}) - \mathcal{L}(\hat{\beta})] = 3056.44$
$\rho^2 = 0.685$
$\bar{\rho}^2 = 0.683$

TABLE 8 Willingness to Pay for More Useful Information

	Variable number	Variable name	Coefficient estimate	t statistic
	1	Constant	-0.649	-3.82
Socioeconomic Characteristics	2	Age > 50	0.128	2.24
	3	Income >80K	0.320	3.29
	4	Faculty	0.153	1.22
	5	Administrative Staff	0.038	0.31
	6	Support Staff	-0.297	-2.19
Trip Characteristics	7	Flexibility > 16 min	0.153	1.62
	8	Travel Time Exceeds Often the Usual Time Range	0.133	1.27
Perceptions & Attitudes	9	Traffic Reports are not Important in my Route Choice	-0.260	-2.61
	10	Traffic Reports are Very Important in my Route Choice	0.155	1.26
	11	Habit Importance	-0.001	-2.56
	12	Like Discovering New Routes	0.279	2.42
	13	Thresh 1	0.494	14.23
	14	Thresh 2	1.115	24.17
	15	Thresh 3	1.616	22.66

Summary statistics
Number of observations = 776
$\mathcal{L}(\mathbf{0}) = -920.19$
$\mathcal{L}(\hat{\beta}) = -858.62$

Profession also plays an important role. With the "base case" used for this variable being the student category, faculty and administration staff are more willing to pay, whereas support staff are less willing to pay than students for more useful information. If travelers do not consider radio traffic reports to be an important factor in their route choice, they are less willing to pay for more useful information. These might be people experiencing relatively short commuting times or having a lot of flexibility when choosing their travel patterns. On the other hand, if traffic reports are considered to be an important factor that influences route choice, then the willingness to pay for more useful information increases. This is a positive indicator for future ATIS implementation issues. ATIS will be successfully implemented if they satisfy travelers' needs by providing reliable and relevant information.

Drivers having more than 16 min of flexibility are more likely to pay for more useful information. The traffic information will give these drivers the opportunity to optimize their departure time choice and avoid traffic congestion. If the travel time often exceeds its usual range, drivers are more willing to pay for more useful information. However, if travelers choose their route by habit, they are less willing to pay for more useful information.

Impact of Driver Perceptions of Information Characteristics on Information Acquisition

The models presented in the fifth section indicate that the acquisition of pretrip and en-route traffic information is influenced by driver perceptions of the reliability and relevance of the provided information. To assess the sensitivity of information acquisition decisions to changes in driver perceptions of the reliability and relevance of traffic information, two improvement levels in these perceptions were assumed. In the first case whereby perceptions improved by one level on the five-level scale, a 4 percent increase in the number of drivers acquiring pretrip traffic information (31 to 35 percent and a 2 percent increase in acquisition of en-route information (from 30 to 32 percent) were forecasted. In the second case, when all drivers were assumed to perceive radio traffic information as reliable and relevant, the forecasted increases in pretrip and en-route information acquisition were 7 percent (from 31 to 38 percent) and 3 percent (from 30 to 33 percent), respectively.

Although the percentage increases were not very significant—possibly a reflection of general attitudes toward information provided by radio traffic reports—the exercise helps visualize the potential applications of the modeling framework in forecasting traveler response to ATIS options, including the acquisition of traffic information characterized by various reliability levels. It is believed that if the modeling framework is utilized with data of traveler behavior obtained from field tests using actual ATIS or from travel simulators, the sensitivity to information characteristics will be more significant.

CONCLUSIONS AND FURTHER RESEARCH

A general framework and model structure for drivers' route choice behavior in the presence of traffic information were formulated in this paper. The following travel decisions were analyzed:

- Acquisition of pretrip traffic information;
- Influence of pretrip traffic information on commuters' decisions;
- Acquisition of en-route traffic information; and
- Drivers' switching behavior.

Model estimation was based on the results of a survey of MIT commuters conducted in 1991. The questions in the survey, related to pretrip and en-route information acquisition, provided the opportunity to evaluate the direct impact of traffic information on users' decisions. In addition, attitudes and perceptions were incorporated into the models using a two-stage process that eliminated inconsistencies or biases from the estimations.

Perceptions about the relevance and reliability of the radio traffic reports were found to be important factors that affect radio traffic information acquisition. Therefore, it can be said that a more reliable and more frequently updated traffic information system than radio would stimulate the acquisition of traffic information. En-route diversion is influenced by attitudinal factors of the drivers and by information acquisition. It was also found that the drivers' own observations are an important factor toward route switching. This finding is important for the efficient implementation of route guidance systems. Only a system that gives accurate and precise directions that correspond to and reflect actual traffic conditions will be successful at gaining driver confidence in following its route choice instructions. Finally, the model showing willingness to pay indicates that drivers with higher incomes and the potential for travel time savings are more willing to pay for more useful information.

REFERENCES

1. Bovy, P. H. L., and E. Stern. *Route Choice: Wayfinding in Transportation Networks.* Kluwers Academic Publishers, Boston, 1990.
2. Khattak, A. J., J. L. Schofer, and F. S. Koppleman. Commuters' Enroute Diversion and Return Decisions: IVHS Design Implications. In *Proc. of International Conference on Travel Behavior,* Quebec City, Quebec, Canada, 1991.
3. Hatcher G., and H. S. Mahmassani. Daily Variability of Route and Trip Scheduling Decisions for the Evening Commute. Presented at 71st Annual Meeting of the Transportation Research Board, Washington, D.C., 1992.
4. Cascetta, E. and L. Biggiero. Analysis and Modeling of Commuters' Departures and Route Choices in Urban Networks. In *2nd International Capri Seminar on Urban Traffic Networks,* 1992.
5. Whitworth, P. *Advanced Traveler Information Systems: Technology, Market Development, User Response and the Government's Role.* M.S. thesis, Massachusetts Institute of Technology, Cambridge. 1993.
6. Khattak, A. J., J. L. Schofer, and F. S. Koppleman. Factors Influencing Commuters' Enroute Diversion Behavior in Response to Delay. In *Transportation Research Record 1318,* TRB, National Research Council, Washington, D.C., 1992.
7. Polak, J., and P. Jones. The Acquisition of Pretrip Information: A Stated Preference Approach. Presented at 71st Annual Meeting of the Transportation Research Board, Washington, D.C., 1992.
8. Koutsopoulos H., A. Polydoropoulou, and M. Ben-Akiva. *Public Acceptance and User Response to ATIS Products and Services: The Use of Travel Simulators to Investigate the Response to Traffic Information.* Volpe National Transportation Systems Center, Cambridge, Mass., 1993.
9. Bonsall, P. W. and T. Parry. Using an Interactive Route-Choice Simulator to Investigate Drivers' Compliance with Route Guidance Advice. In *Transportation Research Record 1306,* TRB, National Research Council, Washington, D.C., 1991.
10. Koutsopoulos, H. N., T. Lotan, and Q. Yang. A Driving Simulator and its Application for Modeling Route Choice in the Presence of Information. Presented at 72nd Annual Meeting of the Transportation Research Board, Washington, D.C., 1993.
11. Alder, J. L., W. W. Recker, and M. G. McNally. *Using Interactive Simulation to Model Driver Behavior under ATIS.* Working Paper. Institute of Transportation Studies, University of California, Irvine, 1992.
12. Adler, J. L., W. W. Recker, and M. G. McNally. *A Conflict Model and Interactive Simulator (FASTCARS) for Predicting Enroute Assessment and Adjustment Behavior in Response to Real-Time Traffic Condition Information.* Working Paper. University of California, Irvine, 1992.
13. Adler, J. L., W. W. Recker, and M. G. McNally. *In-Laboratory Experiments to Analyze Enroute Driver Behavior Under ATIS.* Working Paper. University of California, Irvine, 1993.
14. Vaughn, K. M., M. Abdel-Aty, R. Kitamura, P. P. Jovanis, and H. Yang. Experimental Analysis and Modeling of Sequential Route Choice Behavior Under ATIS in a Simplistic Traffic Network. Presented at 72nd Annual Meeting of the Transportation Research Board, Washington, D.C., 1993.
15. Yang, H., R. Kitamura, P. P. Jovanis, K. V. Vaughn, M. Abdel-Aty, and R. DVG Prasuna. Exploration of Driver Route Choice with Advanced Traveler Information Using Neural Network Concepts. Presented at 72nd Annual Meeting of the Transportation Research Board, Washington, D.C., 1993.
16. Chen, P. S., and H. S. Mahmassani. A Dynamic Interactive Simulator for the Study of Commuter Behavior Under Real-time Traffic Information Supply Strategies. Presented at 72nd Annual Meeting of the Transportation Research Board, Washington, D.C., 1993.
17. Ben-Akiva, M., A. De Palma, and I. Kaysi. Dynamic Network Models and Driver Information Systems. *Transportation Research,* Vol. 25A, No. 5, 1991.
18. Kaysi, I. A. *Framework and Models for the Provision of Real-Time Driver Information.* Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1992.
19. Lotan, T. *Modeling Route Choice Behavior in the Presence of Information Using Concepts from Fuzzy Set Theory and Approximate Reasoning.* Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1992.
20. Ben-Akiva, M. and S. R. Lerman. *Discrete Choice Analysis.* The MIT Press, Cambridge, 1985.
21. Tukey, J. Bias and Confidence in Not Quite Large Samples. *Annals of Mathematical Statistics,* Vol. 29, 1958.
22. Polydoropoulou, A. *Modeling the Influence of Traffic Information on Drivers' Route Choice Behavior.* M.S. Thesis, Massachusetts Institute of Technology, Cambridge, 1993.

Publication of this paper sponsored by Committee on Transportation Supply Analysis.

Effect of Modeling Assumptions on Evolution of Queues in a Single Corridor

CARLOS F. DAGANZO AND WEI-HUA LIN

A qualitative description is presented of queuing patterns under an idealized scenario analogous to the evolution of traffic congestion during the morning peak hour in a long corridor leading to a single destination. The simplicity of the scenario allows the results to be verified independently by hand. Initially the corridor is assumed to consist of a single freeway. Traffic is generated at the freeway's many on-ramps during a short period and then is assumed to subside. Capacity limitations create queues on the ramps and the freeway, whose evolution is then described. A special case with just a few parameters is analyzed in detail. The solution obtained under the assumptions of the hydrodynamic theory of traffic flow (which explicitly recognizes vehicle storage limitations on the freeway) is shown to be drastically different from the solution obtained using "point queue" models, which ignore these limitations. Because the latter models are currently a favored approach in the dynamic traffic assignment literature, results of this study illustrate the need for reevaluating the conditions under which current theories may be applicable.The effect that a slower parallel arterial would have on the system's traffic is also discussed. It was found that a route choice mechanism in which drivers do not anticipate the system's evolution leads to unreasonable traffic patterns—that is, patterns that would not be expected in reality. The anticipation phenomenon must thus be incorporated into any realistic model of dynamic network flows, which unfortunately, increases the difficulty of developing detailed control strategies.

An attempt is made to illustrate the effect that modeling assumptions may have on the evolution of queues in traffic networks. The vehicle for this discussion will be an idealized corridor consisting of a freeway leading to a central business district (CBD) and the parallel set of surface streets. It is assumed that traffic flows without random incidents and that all the vehicles are headed for the CBD. This scenario can be viewed as an idealization of the "morning commute," and its simplicity allows the discussion to be transparent. The qualitative conclusions reached will extend beyond the scenario, and this will be justified as needed. Initially a single freeway is considered and then the results are extended to a corridor.

The following section describes the freeway, its demand, and the traffic flow rules. The core of this note describes the evolution of traffic on the freeway with the hydrodynamic model and with the "point queue" model. The last section introduces the arterial and examines the implications of route choice phenomena.

A SINGLE FREEWAY

A freeway with a series of on-ramps leading to a major destination is considered. Each origin is connected with the freeway by one on-ramp. The origins and on-ramps are consecutively numbered $i = 1, 2, 3, \ldots$, starting from the destination. The index i also refers to the freeway link directly upstream of ramp i. Note that off-ramps are ignored in this network because they have little effect on the scenario considered in this paper.

Department of Civil Engineering and Institute of Transportation Studies, University of California, Berkeley, Calif. 94720-1712.

Demand

It is assumed that the freeway is empty and that the rush commences at time $t = 0$, when A_i vehicles suddenly leave each origin and form a queue at their respective on-ramp. As a result of this model the approximate time d_i, when the queue on ramp i dissipates, will be determined.

It should be clear that the behavior of traffic on the freeway would not change if some of the A_i vehicles on ramp i had left the origin after time $t = 0$, but not so late that they would avoid queuing at the ramp. As long as some vehicles are waiting to leave the ramp at all times in $(0, d_i)$, the actual size of the queue is irrelevant to the behavior of the freeway. (Although the focus of this paper is freeway behavior, ramp delay is briefly considered. Thus our traffic generation assumption is not as restrictive as it may seem; it represents any arrival pattern in which the queues at every ramp would build at about the same time and dissipate only once.

Traffic Flow Rules

Traffic on the freeway will be assumed to behave according to the hydrodynamic theory of traffic flow (*1,2*). Further, it will be assumed that the equation of state [the relationship between density k and flow q that holds at every point in space-time, $q(k)$] is triangular. A triangular relationship is defined by three constants (see Figure 1): the free-flow speed, (v_f), the optimum density, k_0 (or alternatively the maximum flow, $q_{max} = v_f k_0$), and the wave speed for $k > k_0$, $\omega > 0$. The expression is

$$q = \begin{cases} v_f k & \text{if } k \leq k_0 \\ v_f k_0 - \omega(k - k_0) & \text{if } k_0 < k < k_0\left(1 + \frac{v_f}{\omega}\right) \end{cases}$$

The parameter ω represents the unique speed at which flow disturbances propagate in the upstream direction within any (moving) queue. The ratio v_f/ω is approximately 6 for most freeways. The expression $k_0(1 + v_f/\omega)$ represents the jam density. Triangular $q(k)$ relations have been proposed by Newell (*3*) for their simplicity of analysis and reasonable realism. Experiments to measure the accuracy of all these assumptions are being planned.

To predict traffic behavior on the freeway with the hydrodynamic theory, the above is not sufficient. Boundary conditions must also be defined at the location of each nonempty ramp to reflect the amount of ramp traffic that is allowed to enter, depending on con-

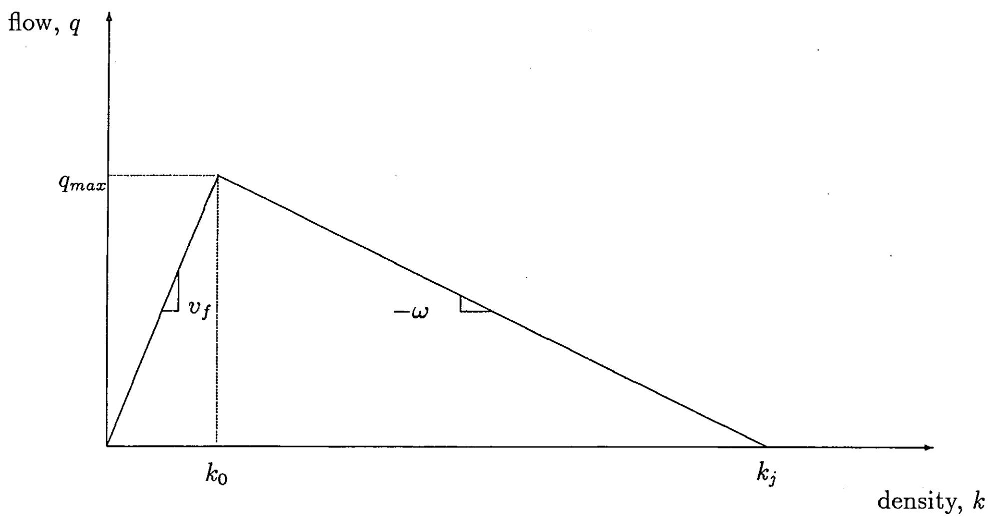


FIGURE 1 Piecewise linear q–k relation.

ditions. If the freeway is congested at the ramp location, the ramp flow might be restricted. This phenomenon can be captured by means of mixing constants that indicate the fraction of ramp vehicles and freeway vehicles that will flow immediately downstream of the junction, assuming that the flow of ramp vehicles does not exceed the ramp capacity. If the ramp flow is restricted by freeway congestion, it will be assumed that α of the vehicles entering the merge comes from the ramp and that $1 - \alpha$ comes from the freeway [more details pertaining to the boundary conditions holding at merge junctions were given previously (4)]. Otherwise the ramp is assumed to flow at its maximum rate.

Although α could depend on the freeway density, it is reasonable to assume that it is constant for these purposes—perhaps close to $1/2m$—where m is the number of freeway lanes (5). If α is constant and the freeway is in a congested steady state during part of the rush hour, its link flow will decrease geometrically by a factor of $(1 - \alpha)$ in the upstream direction. This means that downstream ramps will discharge more flow than their upstream counterparts, essentially giving the commuting advantage to downstream inhabitants.

This qualitative observation is also true if one assumes that the ramps have absolute priority (and that the ramp flow is limited to a fixed amount, q_r). The main difference is that the freeway link flows would now decrease arithmetically instead of geometrically, and the ramp flows would vary more drastically.

The remainder of this paper assumes that the ramps have absolute priority because the same qualitative conclusions are reached, and this simpler model leads to a graphical presentation that involves only a few traffic states and exhibits more contrast; it is better suited for comparison with graphical computer output. The theoretical and numerical results for the more general case are summarized in the appendix of an earlier work (6).

In the following section the evolution of traffic on the freeway is examined. The first part gives a qualitative description of traffic behavior for general conditions about freeway geometry and demand. The second part provides quantitative formulae for a special case that can be described with just a few parameters. The third part explains the implications of the above results for control, considering the impact that a control scheme would have on ramp queuing delay.

TRAFFIC EVOLUTION

Assuming that the priority flow from every ramp is metered at a rate $q_r > 0$, as long as on-ramp traffic is not blocked by downstream freeway congestion, on-ramp queues will discharge at a rate q_r.

Qualitative Description

Immediately after $t = 0$, as vehicles from upstream ramps travel downstream, flow on each freeway link will increase. The hydrodynamic model adopted here predicts that it should do so in steps. [The same results can be obtained with Newell's solution method for the cumulative flow curves at the entrance and exit to every freeway link (3).] The time between steps is the time it takes a vehicle to travel between ramps; and the size of each step is q_r.

Flow will have reached capacity in a freeway section upstream of a ramp in the time that it takes a vehicle to travel q_{max}/q_r freeway links. Because the maximum freeway flow that can get past an active ramp is $q_{max} - q_r$, capacity flows will be precluded from advancing from then on. Our hydrodynamic model predicts that, under these conditions, queues of slow-moving vehicles will form upstream of every ramp and those queues will grow spatially at rate ω. (The result is unusual but true for the scenario considered in this paper. In nonidealized cases with unequal ramp spacing and nonidentical ramp loads, queues on every link will not necessarily grow spatially at the same rate. This, however, does not change the qualitative features of the queuing evolution patterns to be discussed later.) When the queues reach the next upstream ramp, in the time that it takes for the shock wave to travel between ramps (e.g., 5 min), the freeway flow immediately downstream of Ramps 2, 3, . . . will decrease to $q_{max} - q_r$ from q_r. As Ramps 2, 3, . . . continue to emit vehicles at rate q_r, the freeway flow directly upstream of these ramps will be further restricted to $q_{max} - 2q_r$; another set of shock waves will be generated.

Consideration shows that in the time that it takes a wave to travel the length of q_{max}/q_r freeway links (about 30 min) the system will have reached an equilibrium state, as shown in Figure 2, assuming of course that no ramp queues have dissipated. Link flows would be

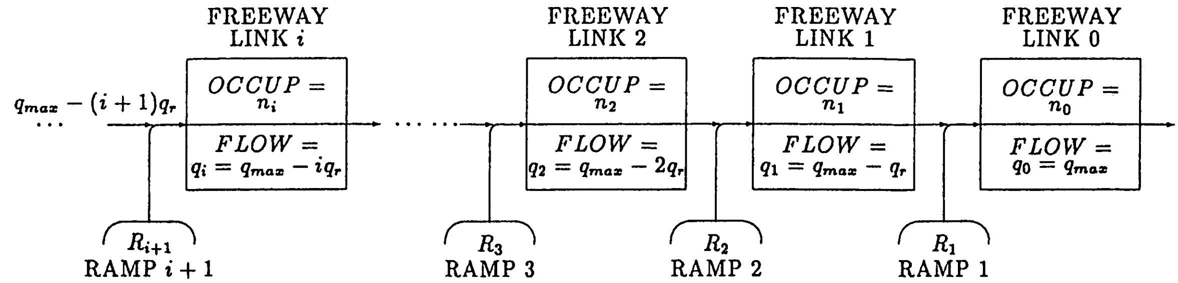


FIGURE 2 Schematic representation of freeway system.

$$q_i = \begin{cases} q_{\max} - iq_r & \text{if } i \leq q_{\max}/q_r \\ 0 & \text{if } i > q_{\max}/q_r \end{cases}$$

This occurs because priority downstream origins block upstream traffic. This stable pattern in which upstream ramps are blocked by the stopped freeway traffic will persist until some of the downstream ramps dissipate their queues.

The stable pattern is characterized by certain link occupancies, n_i, and link travel times, t_i. The freeway link occupancy is a linearly decreasing function of the link flow, given the equation of state (Figure 1). If l_i is the length of Link i, the relationship is

$$n_i = \left(\frac{v_f k_0 - q_i}{\omega} + k_0\right) l_i = \left(\frac{iq_r}{\omega} + k_0\right) l_i$$

The link travel times for a vehicle traversing Link i during the stable pattern is

$$t_i = \frac{n_i}{q_i} = \frac{\left(\frac{iq_r}{\omega} + k_0\right) l_i}{q_{\max} - iq_r}$$

The stable pattern will persist until one of the downstream ramp queues dissipates. Upstream of such a ramp freeway flow will then increase by q_r units and, after a suitable delay, the first ramp to have been blocked will commence to discharge. If no other ramps dissipate their queues during this delay, another stable pattern will then have been reached.

As time progresses other stable patterns may arise. They all have in common that the first $q_{\max}/q_r$ ramps with queues continue emitting vehicles and the rest are either empty or blocked. Flow on Link 0 remains at capacity throughout—during the stable patterns and the transition periods. Note as well that ramps close to the destination clear before those far away. Although the last $q_{\max}/q_r$ ramps are emitting vehicles, the downstream portion of the freeway will still be at capacity. As the queues on these ramps dissipate, freeway speeds will return to normal and downstream freeway flows will decrease.

Quantitative Results

In this section the following quantitative results for an idealized case that can be described with few parameters and is easy to analyze are developed; it is assumed that the freeway ramps are evenly spaced l distance units apart and that $q_{\max}/q_r$ is an integer, m. If t_0 denotes the sum of the times that it takes for a vehicle to travel one link (l/v_f) and for a wave to do the same (l/ω), $t_0 = (l/v_f) + (l/\omega)$, the first stable pattern is reached at time $t_b = mt_0$. This is also the time when ramps such that $i > q_{\max}/q_r$ become blocked:

$$t_b = mt_0 = \frac{q_{\max}}{q_r} t_0 = \frac{q_{\max}}{q_r} l \left(\frac{1}{v_f} + \frac{1}{\omega}\right)$$

The number of vehicles initially discharged by one of these ramps is

$$A_b = q_r t_b = q_{\max} t_0 = l\left(\frac{q_{\max}}{v_f} + \frac{q_{\max}}{\omega}\right) = lk_{\text{jam}}$$

which is also the maximum number of vehicles that fit in a freeway link.

Assuming that $A_i > A_b$ (the usual case if there is a congestion problem during the morning peak), the last $(A_i - A_b)$ vehicles on the ramp will have to wait until $(i - q_{\max}/q_r)$ downstream ramps have discharged and the clearing signal has had the time to reach ramp i. The specific times when ramp i begins to discharge again, t_{ri}, and ends to discharge, t_{di}, will depend on the downstream A_i's, but the formula is not particularly important for these purposes. The simplest case in which all the A_i are equal, $A_i = A$, will suffice to illustrate some issues.

The time-space solution for this problem is diagrammed in Figure 3 for a case with $m = 4$ and $v_f/\omega = 2$ and with only 12 on-ramps. (A value of $v_f/\omega = 6$ close to 6 would be more reasonable, but the diagram would be less clear.) Figure 3 does not show vehicle trajectories; the solid lines represent time-space interfaces between the traffic states prevailing in various regions of the time-space plane. These traffic states, labeled $O, A, B, C, D, C', B', A', O'$, correspond to the points on the $q(k)$ curve also displayed in the figure. Since this $q(k)$ relation exhibits only two wave speeds (v_f and ω), the hydrodynamic solution is easy to construct using the following two well-known facts:

1. Because there is a change in freeway flow at the ramp's location whenever the ramp is emitting vehicles a stationary interface must be located there. The flow on the upstream side of the interface must be q_r flow units less than the downstream flows.
2. Nonstationary interfaces must move with a speed equal to the ratio of the change in flows on both sides of the interface to the change in density. If the scales of representation are chosen appropriately (as has been done) interfaces are parallel to the $q-k$ diagram line that joins the two states.

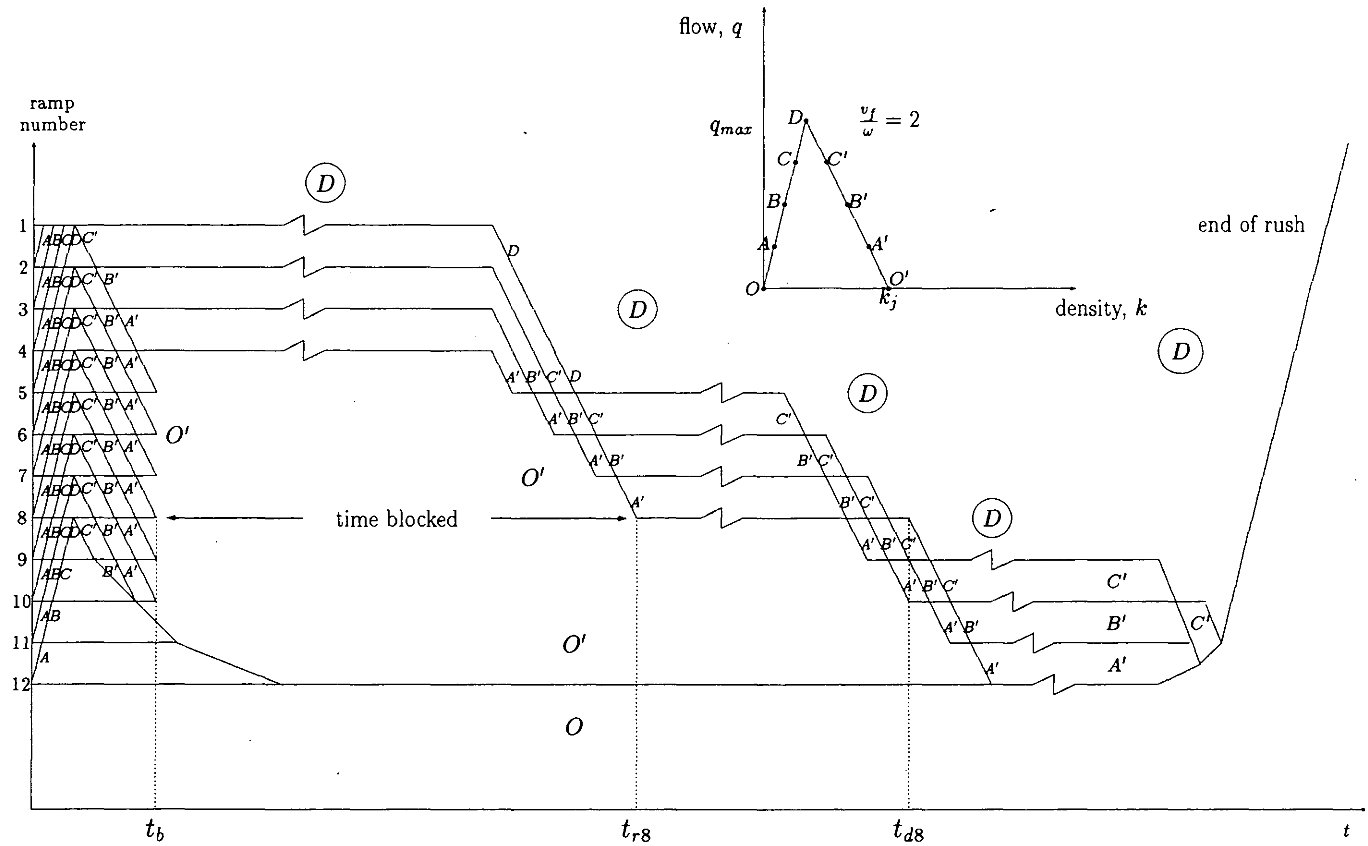


FIGURE 3 Time-space solution to morning traffic problem.

In Figure 3 a horizontal interface corresponds to an active ramp; a blocked ramp is identified by an interruption in a horizontal line at the ramp's position. Figure 3 also displays explicitly the times t_{ri} and t_{di} for ramp $i = 8$. The first and last set of four ramps behave a little differently from the four in between because the former are affected by upstream boundary conditions. The expression for $t_{di} - t_{ri}$, obvious from Figure 3, applies to the middle ramps (excluding the first and last m in general cases):

$$t_{di} - t_{ri} = \frac{A - A_b}{q_r}, \qquad \text{if } m < i < i_{\max} - m \tag{1}$$

It should also be clear from the geometry of Figure 3 that

$$t_{di} = t_b + \left(\frac{A}{q_r} - t_b\right)\left(\frac{i}{m}\right)^{+} + \left(\frac{l}{\omega}\right)\left\{-1 + 2i - \left(\frac{i}{m}\right)^{+} m\right\} \tag{2a}$$

which can be approximated by a smooth function of i:

$$t_{di} \approx \frac{A}{2q_r} + \left(\frac{A}{q_r} - t_b\right)\left(\frac{i}{m}\right) + \frac{il}{\omega} \tag{2b}$$

The first two terms of Equation 2b represent the time that it takes to discharge the queues blocking Ramp i (including Ramp i itself). The third term represents the time for the clearing signal to travel back to Ramp i. Because $t_b = q_{\max} t_0 / q_r$ and $m = q_{\max}/q_r$, it is possible to write

$$t_{di} \approx \frac{A}{2q_r} + \left(\frac{A}{q_{\max}} - t_0 + \frac{l}{\omega}\right) i \qquad \text{if } m < i < i_{\max} - m \tag{3}$$

which increases with i at rate $(A/q_{\max} - t_0 + l/\omega)$.

Because the difference between t_{di} and t_{ri} is constant, t_{ri} also increases at the same rate with i. Figure 4 shows the relationship between i and the times when the ramp actually discharges vehicles, including $i \in (1, m)$.

The times t_{di} represent the times when an additional vehicle could conceivably enter the freeway at Ramp i and travel at the free-flow speed all the way to the destination. As such, it can be viewed as the time of return to normalcy for origin i.

Implications for Control

Except for an initial time, m/v_f, the bottleneck remains at capacity throughout the peak hour. This observation is true independent of q_r and therefore independent of a ramp metering strategy if one assumes (plausibly) that changes in q_r do not change $q_{\max}$. (It is also true if ramps do not have absolute priority.) This indicates that the $i_{\max} A$ vehicles that use the systems are removed at a maximum rate that is independent of q_r. It follows that the total number of vehicle-hours in the system cannot be reduced by ramp metering.

Ramp metering cannot even influence the number of vehicle-hours on the freeway; metering one ramp simply allows vehicles from other ramps to spend time on the freeway. (This, of course, is

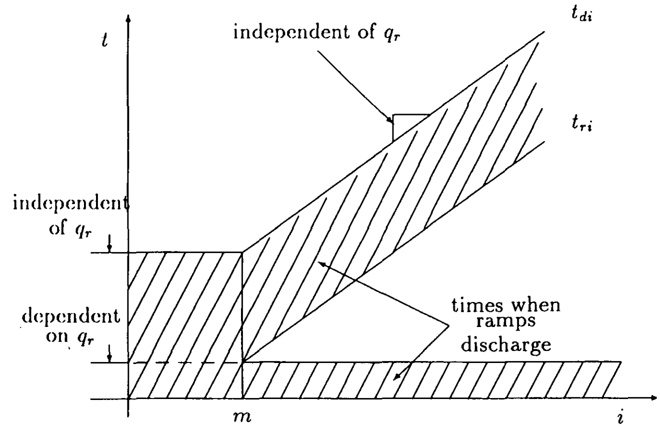


FIGURE 4 Periods of activity for various ramps.

a result of the lack of parallel routes and exiting traffic and does not occur in general.) To illustrate this it is shown that vehicles enter the freeway from all ramps at a combined rate that varies cyclically around a long-term average that is independent of the metering rate.

At any given time, t', the number of ramps discharging (at rate q_r) into the freeway is given by the number of horizontal lines that would be intersected by the line $t = t'$ in Figure 3. During any stable pattern, without moving interfaces, the number is $m = q_{max}/q_r$, but the number is lower during a transition period. The number fluctuates periodically.

For a freeway with $i_{max} \to \infty$, the number of active ramps is on average

$$m\left\{\left(\frac{A}{q_r} - mt_0\right)\Big/\left(\frac{A}{q_r} - mt_0 + \frac{ml}{\omega}\right)\right\}$$

This can be easily seen from the figure if one imagines that there are m "on/off" moving switches that control the ramp flows; a ramp emits flow when it has an on switch. During any stable period there is one on switch on each of the m-active ramps. Whenever a ramp queue dissipates, its switch is turned off and sent to the first upstream ramp that is blocked at the speed of the interface. The trip takes ml/ω time units (see Figure 3). On arrival it is turned on until the queue dissipates; that is, it stays on for a time $(A/q_r) - mt_0$. For a very long freeway, the average number of active ramps is equal to the average number of on switches, which in turn is equal to the product of m and the fraction of time that a switch is on. This is the formula that has just been presented.

The average flow discharged into the freeway before the end of the peak also varies cyclically around an average that is the product of the above formula and q_r:

$$q_{max}\left\{\left(A - q_{max}t_0\right)\Big/\left(A - q_{max}t_0 + q_{max}\frac{l}{\omega}\right)\right\} \quad (4)$$

This quantity is independent of q_r.

TRAFFIC EVOLUTION WITH POINT QUEUES

The term "point queue" is used here to refer to the limiting case of the hydrodynamic model with $k_{jam} \to \infty$ (or $\omega \to 0$). This limiting case exhibits the feature that queues never grow to be so long that they restrict entry into a link.

This limiting model is becoming popular in the dynamic traffic assignment literature (7,8) because it is somewhat tractable and (because it includes transient queuing phenomena) can be quite realistic when congestion is mild. Point queue models are not reasonable, however, for networks so congested that storage is a problem because crowded links may not always be ready to admit vehicles. As the material in this subsection shows, the consequences of the point queue assumption can be drastic.

First it is noted that with the point queue model all the ramp queues will have dissipated by time $\max_i \{A_i/q_r\}$. If the A_i are not too different, this is roughly the same time at which the queues on Ramps 1 through m clear with the hydrodynamic model. The two times coincide for the idealized case with $A_i = A$. With point queues, the freeway does not return to normalcy at that time, however. Large queues will have formed on every freeway link that is away from the upstream system boundary. To determine this, one must note that if the A_i/q_r are large relative to the time it takes to travel m links, which is the case of interest as explained earlier, then queues must begin to grow after time $(m + 1)\, l/v$ on any freeway link that is more than m ramps away from the upstream system boundary. These queues continue to grow until the ramp flows are exhausted. If the A_i are large and the freeway has many ramps, then the queues will be very large.

Unfortunately these queues do not dissipate quickly. Without any ramp flow, all the queues will receive the same number of vehicles as they emit in a unit time (q_{max}) and therefore will remain fixed in size—all the queues, that is, except that the last nonzero upstream queue, which will decline in size at rate q_{max}.

The specific times at which the individual queues dissipate can be found graphically with a construction as in Figure 3 (with $\omega \to 0$), or (more easily) with time versus cumulative count diagrams at every freeway link. This is not done here, however, because the qualitative properties of the result for $A_i = A$ are obvious.

As explained earlier, the m links at the upstream end of the freeway do not develop queues; queues grow only on the remaining links after a delay of $(m + 1)l/v_f$ time units. The build-up will take place at the same rate (q_r) for all links, because links can emit a flow q_{max} but receive the sum of a flow q_{max}, emitted by the upstream freeway link, and the ramp flow q_r. If A is so large that the initial delay $[(m + 1)l/v_f]$ is negligible compared with A/q_r, then the maximum freeway queues will be reached approximately at time A/q_r and will be of size A on each link for the ideal case. (Essentially, the ramp queues will have been transferred to the freeway.) From then on the freeway occupancy will decline.

The last upstream queue will dissipate first as it receives no traffic, but discharges it at rate q_{max}; its dissipation will reduce arrivals to the downstream queue, which will dissipate next, and so on. Interestingly, the result of the point queue assumption is that freeway links return to normalcy from the upstream end to the downstream end. That is, in the reverse order as one would expect for the scenario with no alternative routes and equal ramp traffic.

Figure 5 displays the predictions of the point queue and the hydrodynamic model for a limiting case with $A/q_r, i_{max} \to \infty$. Note that both models predict the same clearing time for the system as a whole: $i_{max}A/q_{max}$. This should not be surprising because both models keep Link 0 saturated throughout the rush.

Constant saturation flow through Link 0 also implies that the number of vehicle-hours in the system is the same as that for the hydrodynamic model. However, the number of freeway vehicle-hours

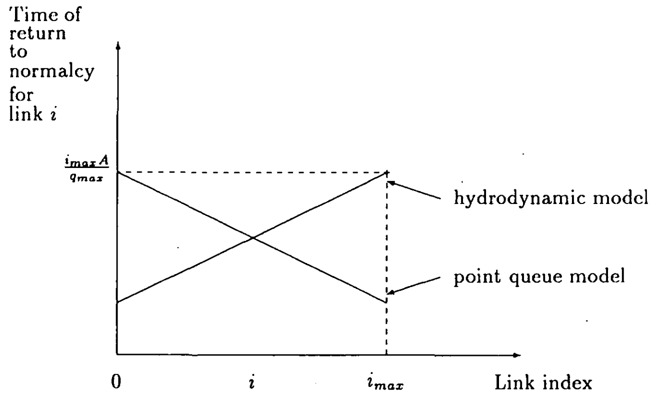


FIGURE 5 Comparison of point queue model and hydrodynamic model.

is overestimated because vehicles are assumed to enter the freeway at a rate $q_r i_{\max}$, which is much larger than that in Equation 4. (In the limiting case with $i_{\max}$ and $A \to \infty$, the fractional overestimation error is unbounded.) This overestimation is a direct consequence of the absence of freeway link-to-link interactions (blocking) in the point queue model.

FREEWAY AND PARALLEL ARTERIAL

The example discussed in this section will show that if drivers choose routes dynamically on the basis of current travel times without anticipating the evolution of traffic, as is done in some dynamic traffic assignment models, they would follow inefficient paths to their destinations—paths that experienced commuters would not take.

With the goal of assessing the reasonableness of a simple model of driver route choice behavior, people are now allowed to travel on a slower one-way arterial road and choose the most convenient ramp. They are not allowed, however, to postpone or cancel their trips, for example, as in using an alternative mode of transportation. It also is assumed that the arterial ends at Ramp 1 and cannot be used to avoid the bottleneck. This could represent a situation in which the bottleneck is a freeway bridge (e.g., the Oakland–San Francisco Bay Bridge).

The arterial/freeway problem is easy to analyze if it is assumed that people evaluate routes continuously as they travel without anticipating future changes in link travel times. That is, the route travel time used in their decision at time t is the sum of the "current" link travel times, defined to be the projected link travel times under current conditions for a vehicle just entering each link. It will be assumed that speed on the arterial is independent of flow; this is reasonable because arterial flows should be modest, and in real life there may be a system of parallel streets ready to handle the flow. As part of this solution, the queue lengths on each ramp as a function of time, $R_i(t)$, are sought.

With this route choice model, a driver on the arterial at time t will compare two active adjoining ramps by considering the projected ramp queuing times, $R_{i+1}(t)/q_r$ and $R_i(t)/q_r$, and the current travel time on the intervening freeway link, $t_i(t) = n_i(t)/q_i(t)$. The travel time difference between the two choices (using Ramp i and Ramp $i + 1$) is then

$$\Delta\text{time} = \frac{R_{i+1}(t) - R_i(t)}{q_r} + t_i(t) \tag{5}$$

Note that an inactive ramp (one with $q_r = 0$) would never be selected.

Because the flow from the first m ramps eventually blocks all flow from the remaining ramps as discussed earlier, it follows that the demand from origins $i = m + 1, m + 2, \ldots$, will eventually reject these ramps, moving on the arterial to one of the active ramps ($i \leq m$). Thus, a pattern of ramp queues will develop in which $R_i(t) = 0$ if $i > m$, and $R_i(t) > 0$, otherwise. The positive $R_i(t)$ will decrease in size with time, but as long as $R_i(t) > 0$ for all $i \leq m$, the freeway will behave as if there were no arterial and the stable pattern of Figure 2 prevailed, as explained earlier:

$$t_i(t) = \frac{\left[\frac{iq_r}{\omega} + k_0\right] l_i}{q_{\max} - iq_r} \qquad 0 \leq i \leq m$$

On substituting this relation for $t_i(t)$ in Equation 5 for the ideal case with $l_i = l$, the following is obtained

$$\Delta\text{time} = \frac{R_{i+1}(t) - R_i(t)}{q_r} + \frac{\left(\frac{iq_r}{\omega} + k_0\right) l}{q_{\max} - iq_r} \qquad 1 \leq i < m \tag{6}$$

which gives the difference in current times between ramps, as a function of the queue sizes, $R_i(t)$. These, of course, are yet to be determined.

If τ denotes the travel time on the arterial between two ramps, the route choice mechanism implies that

$$\Delta\text{time} \leq \tau$$

A pure equality would imply that people are indifferent between the two ramps. Otherwise the downstream ramp is favored, which is possible because drivers cannot backtrack. This inequality yields the following condition:

$$R_{i+1}(t) - R_i(t) \leq q_r \left\{\tau - \frac{l(iq_r/\omega + k_0)}{q_{\max} - iq_r}\right\}$$

or

$$R_{i+1}(t) - R_i(t) \leq q_r \left\{\tau - \frac{l}{v_f}\left[\frac{v_f i/\omega + m}{m - i}\right]\right\} \tag{7}$$

The right side of Equation 7 is positive for $i = 1$ (if $\tau >> l/v_f$ as one would expect) and negative for $i = m - 1$ (since $m\,l/v_f$ should be greater than τ in most cases). It is now argued that Equation 7 should be a pure equality.

Initially, one would expect Equation 7 to be a pure equality for $i \leq m$ because the (many) drivers from upstream blocked origins can choose among all i ($i \leq m$) without traveling the wrong way; that is, if one ramp were favored (with a small R_i), some drivers would travel to it and increase its R_i. As long as there are queues on all $i \leq m$, the $R_i(t)$ will decrease at the same rate and drivers will not have an incentive to jockey among ramps. Thus, one would expect Equation 7 to remain as an equality until one of the queues dissipates. It follows that (as long as they are all positive) the $R_i(t)$ can be easily obtained for any given total number of ramp occupants

$$R(t) = \sum_{i=1}^{m} R_i(t) \tag{8}$$

with the following procedure:

1. Starting with a given R_1, obtain $R_2, R_3, \ldots, R_m$ with Equation 7 interpreted as an equality.
2. Then add $(1/m)(R(t) - \sum_{i=1}^{m} R_i)$ to each of the values.
3. The resulting values satisfy Equations 7 and 8; they are valid if none is negative.

Note that the ramp with the longest queue has an i where the right side of Equation 7 is close to 0. This means that it is located where the freeway travel time roughly matches the arterial travel time. Note as well that the minimum queue length arises at one of the extremes (either $i = 1$ or $i = m$). Consequently, because all the $R_i(t)$ decline at the same rate (q_r), one of the extreme ramps will eventually become idle. Consideration shows that no vehicles would move to the new empty ramp. This is obvious for ramp m because vehicles are not allowed to backtrack. It is also true if Ramp 1 clears first because, after dissipation, the queuing time of Ramp 1 remains fixed at 0 while the remaining queuing times decline. Drivers wouldn't want to move to Ramp 1. It should be clear then that as soon as one of the ramps clears, the combined input flow to the freeway drops. As soon as this happens, freeway traffic must quickly return to an uncongested state with positive flows and rapidly moving vehicles (see Figure 6).

What is disturbing from this outcome is that the flow through the bottleneck will have declined below capacity while ramp queues remain. Vehicles trapped in these queues would have been better served by waiting at Ramp $i = m + 1$ in anticipation of the freeway clearing time. Because they did not, their delay and that of the overall system are increased. It is doubtful that such a phenomenon would be observed in real life. (If it did, closing the arterial would result in less delay.) It seems reasonable to conclude that drivers do anticipate the evolution on the basis of past experience and that providing nonanticipatory information may actually be counterproductive. This suggests that the development of effective dynamic traffic-responsive control strategies may be quite difficult.

Similar conclusions are reached if two-way travel is allowed on the arterial. It is not difficult to see that the two-way system would still evolve in the same way until the dissipation time of the first ramp. After that time, drivers at Ramp m or $m - 1$, or both, may find it attractive to backtrack so as to enter Ramp $m + 1$ without significant queuing or delay. This may or may not happen, depending on the specific values of the parameters τ, l, vf, etc., appearing in Equation 7. But if it does, it would imply that some motorists from upstream origins would travel to a ramp and queue and later decide that a previously bypassed upstream ramp was a better alternative. Commuters are unlikely to behave so naively.

A more realistic model of commuter behavior should be based on past experiences. One would expect experienced commuters always to maintain that the first m ramps are busy. If this is true, it would mean that the bottleneck is never starved for traffic and that the overall system delay is as low as it can be, given that there is a bottleneck and that the arterial ends before the bottleneck. Ramp metering strategies could not reduce the overall system delay.

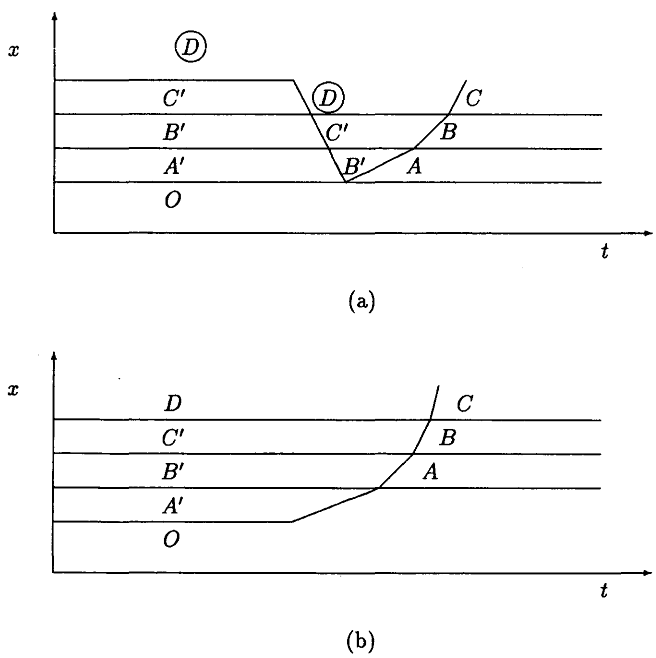


FIGURE 6 Time-space diagram of system's return to normalcy when (*a*) Ramp $i = 1$ clears first and (*b*) Ramp $i = m$ clears first. Figure depicts the case with $m = 4$.

CONCLUSION AND SUMMARY

A solution to the hydrodynamic model for a freeway carrying morning commuters to a single destination has been developed. Although one may take issue with the realism of the hydrodynamic model for a microscopic description of traffic (over short distances and small intervals), its behavior on a macroscopic scale would seem satisfactory. It has been extensively studied and is widely used by practicing engineers. (The value of purported improvements such as microscopic car-following simulations can be questioned because the macroscopic implications of all the details have not been examined in sufficient detail. As is well known, for example, from chaos theory, nonlinear systems can be very sensitive to details.) The model certainly is more realistic than the point queue model because it recognizes that as the freeway fills with vehicles the flow from upstream links must be restricted. The hydrodynamic model predicts a different congestion dissipation pattern. It predicts that freeway links reach freeflow speed in the upstream direction, starting from the bottleneck, whereas the point queue model predicts the opposite; this highlights the latter's inability to represent real-life situations where blocking arises.

A by-product of this analysis is a set of expressions that can be used to test numerical implementations of the hydrodynamic model. This is done in Figure 7, which displays the output of a prototype freeway network flow program currently being developed (9). The shading intensity in this figure corresponds to the magnitude of the predicted traffic density (compare with Figure 3). Figure 8 shows the result for the case in which ramps do not have absolute priority (and $\alpha = 1/4$).

The introduction of a parallel arterial revealed that if drivers were to choose routes without anticipating the future evolution of the system, the queuing patterns that would develop would become unrealistic. This observation has important implications for the direction of dynamic traffic network research; it seems to suggest that commuters choose routes on the basis of expectations formed on prior travel days and that anticipation should be built into the models.

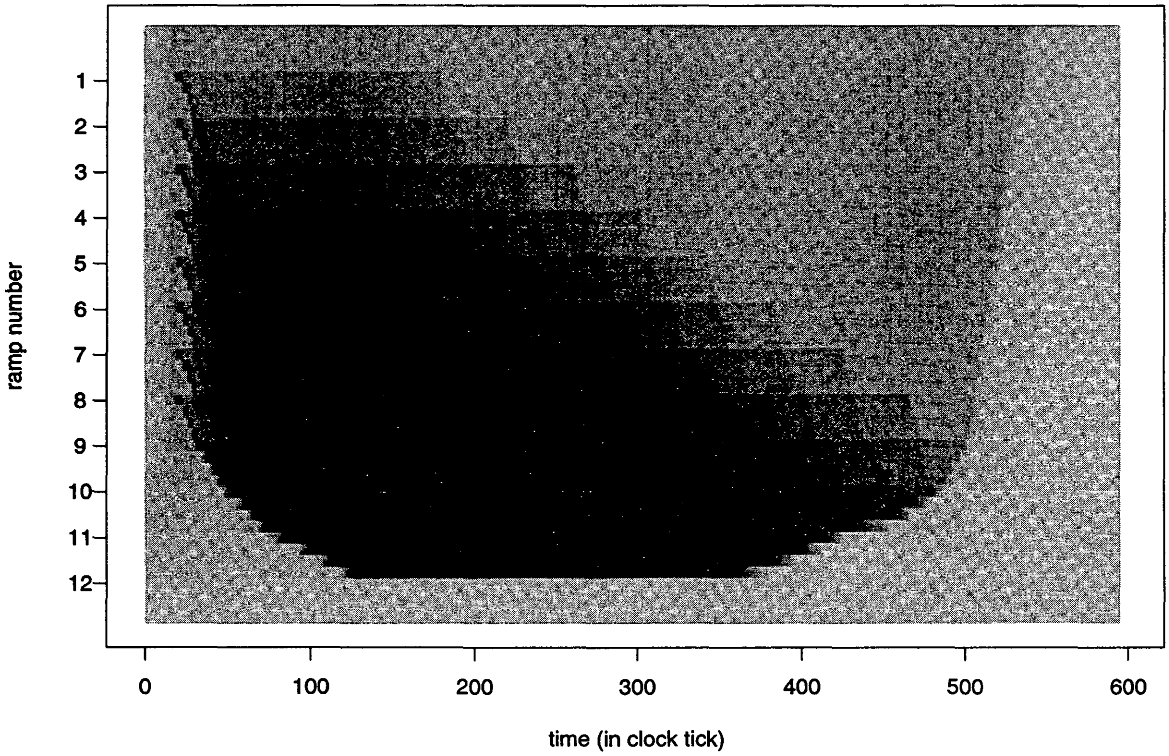


FIGURE 7 **Occupancy intensity plot to morning traffic problem ($\alpha = 1$).**

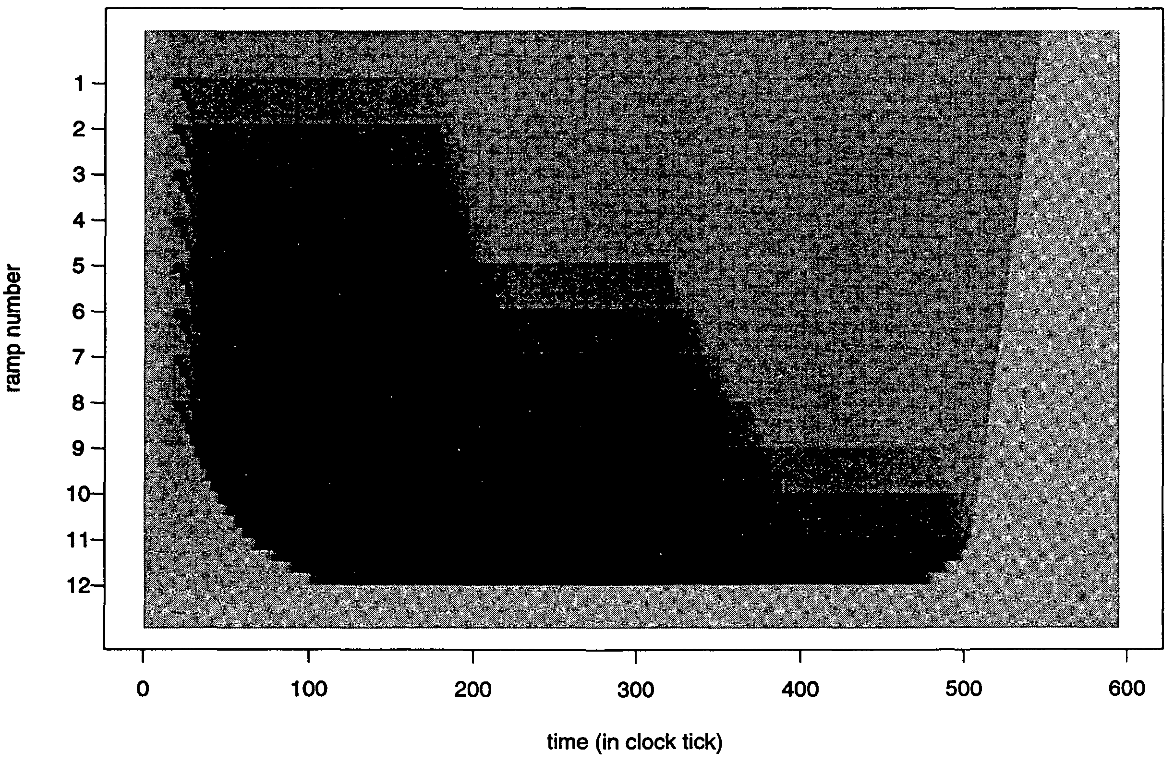


FIGURE 8 **Occupancy intensity plot to morning traffic problem ($\alpha = 1/4$).**

ACKNOWLEDGMENT

The authors appreciate the thoughtful comments of an anonymous reviewer.

REFERENCES

1. Lighthill, M. J. and J. B Whitham. On Kinematic Waves. I Flow Movement in Long Rivers. II A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society,* Vol. A, No. 229, 1955, pp. 281–345.
2. Richards, P. I. Shockwaves on the Highway. *Operations Research* Vol. 4, 1956, pp. 42–51.
3. Newell, G. F. A Simplified Theory of Kinematic Waves: I General Theory; II Queuing at Freeway Bottlenecks; III "Traffic Assignment Problem" for Freeways. *Transportation Research,* Vol. 27B, No. 4, 1993, pp. 281–314.
4. Daganzo, C. F. The Cell Transmission Model: Network Traffic. *Transportation Research,* Vol. 28B, 1994.
5. Yagar, S. Applications of Traffic Flow Theory in Modeling Network Operations. In *Transportation Research Record 567,* TRB, National Research Council, Washington, D.C., 1976, pp. 65–69.
6. Daganzo, C. F., and W. H. Lin. *The Spatial Evolution of Queues During the Morning Commute in a Single Corridor.* PATH Working Paper UCB-ITS-PWP-93-7. Institute of Transportation Studies, University of California at Berkeley, 1993.
7. Bernstein, D., T. L. Priesz, R. L. Tobin, and B. W. Wie. A Variational Control Formulation of the Simultaneous Route and Departure-Time Choice Equilibrium Problem. In *Proc., 12th International Symposium on Transportation and Traffic Theory* (C.F. Daganzo, ed.), Elsevier, New York, 1993, pp. 107–126.
8. Kuwahara, M., and T. Akamatsu. Dynamic Equilibrium Assignment with Queues for a One-to-Many OD Pattern. In *Proc., 12th International Symposium on Transportation and Traffic Theory* (C.F. Daganzo, ed.), Elsevier, New York, 1993, pp. 185–204.
9. Lin, W. H., and C. F. Daganzo. *Technical Description of NETCELL: General Framework and Data Structure.* PATH Technical Report. Institute of Transportation Studies, University of California at Berkeley, 1993.

Publication of this paper sponsored by Committee on Transportation Supply Analysis.

Improved Freeway Incident Detection Using Fuzzy Set Theory

EDMOND CHIN-PING CHANG AND SU-HUA WANG

Freeway incidents often occur unexpectedly and cause undesirable traffic congestion, mobility loss, and environmental pollution even where computerized traffic management systems are installed and in operation. Automatic incident detection, being one of the primary functions of computerized freeway traffic management systems, must be able to detect all freeway incidents as soon as possible with minimum false alarms.

In a study that evaluated the applications of fuzzy set theory to improve existing incident detection algorithms, the potential system performance was compared with that of conventional systems using real-world volume and occupancy data that were collected earlier. The potential benefits and needed improvements in the existing incident detection algorithms to take advantage of the promising fuzzy set methodology are summarized.

Freeway incidents often occur unexpectedly and cause undesirable traffic congestion, mobility loss, and environmental pollution even where computerized traffic management systems are installed and currently in operation. Automatic incident detection (AID) has been used increasingly to improve urban freeway operations and reduce the operational impact of incidents. Being a primary function of computerized freeway traffic management systems (FTMS), the ideal incident detection systems or detection algorithms must be able to detect freeway incidents quickly with minimum false alarms (*1–4*). The operations require the efficient use of available information for reliable incident detection during congested operations and incident conditions.

The commonly used comparative or California-type algorithm requires the continuous evaluation of traffic operational characteristics collected from consecutive detector stations. To develop effective freeway incident management, the algorithms use the principle that an incident will likely increase the occupancy upstream of the incident and decrease occupancy downstream of the incident. Although lane volume and occupancy are the main measures, other algorithms also use measured speeds to distinguish incidents and daily congestion. Most algorithms have been developed to detect freeway incidents through traffic information collected from loop detectors. However, three operational problems often occur when implementing an AID system. These problems include the understanding of the relative operational effectiveness, threshold parameter selection, and better interpretation of the algorithms.

During operations, most conventional incident detection algorithms use a series of decision-making analyses against the predefined thresholds to detect any freeway status changes because these status descriptions often are used with uncertainty measures. Using the "crisp thresholds" cannot reliably distinguish among true and false incidents. In addition, the loss of information also may cause errors, fail to detect incidents, or generate false alarms. Fuzzy logic, which provides approximate reasoning instead of exact reasoning, is an alternative that may improve the reliability of incident detection systems (*5,6*).

Texas Transportation Institute, Texas A&M University System, College Station, Tex. 77843-3135.

STUDY OBJECTIVES

This study examined a feasible software design to improve existing freeway incident detection algorithms through application of fuzzy set theory. The study used real-world traffic volume and occupancy data to improve existing incident detection algorithms. Finally, system performance is measured against conventional systems to enhance existing automatic freeway incident detection algorithms.

This paper investigated the potential application of fuzzy set theory to improve California Incident Detection Algorithm 8. Through the freeway volume and occupancy data collected, three feasible approaches were examined. Other possible enhancements that can be used to improve the existing automatic incident detection algorithms in most computerized FTMs also are summarized.

STUDY BACKGROUND

Most existing incident detection algorithms fall into four categories: pattern recognition approach, statistical analysis approach, catastrophe theory approach, and artificial intelligence approach. Among them, neural network and fuzzy set theory has demonstrated success in representing complicated knowledge and compensating for the difficulty encountered in the conventional decision approach. Experiments, performed initially by the Texas Transportation Institute (TTI), indicated that these two approaches are technologically feasible, especially during insufficient detector information or loss of part of data communication. The purpose of this study is to demonstrate operational performance of fuzzy logic against conventional incident detection algorithms using historical detector data observed from freeway control centers.

The following sections discuss the background of freeway incident detection algorithms, fuzzy set theory, and fuzzy applications. The development tools used in the proposed approaches are described briefly.

Incident Detection Algorithms

Two different incident detection algorithms are commonly used by monitoring data from one or a series of detector stations. The first or the most commonly used incident detection algorithm is the comparative or California-type algorithm (*4*). Ten comparative incident detection algorithms were developed by FHWA. Among these, Algorithm 8 is recommended for use during high-volume conditions. The algorithm can continuously assess freeway incident potential by analyzing volume and occupancy data from paired vehicular detectors at each freeway section.

The second approach uses incident detection triggered by the condition changes as observed at a single detector station (*7,8*). Although not being free from false alarms, the second or point-detection algorithm, such as the McMaster algorithm, was developed. This algorithm does not have to rely on the continuous measures from paired detectors but requires better understanding of freeway operating characteristics at each detector station. As these systems move into daily usage, evaluations of these incident detection elements are essential to improving the operational effectiveness of the traffic management system.

Both the comparative and point-based detection techniques, such as the California and McMaster algorithms, rely on several predefined thresholds to adjust the relative sensitivity for detecting freeway status changes. Because these status descriptors often are associated with uncertainty measures, the use of crisp thresholds cannot clearly distinguish operations among true incidents, congested operations, and impacts from previous incidents. The selection of improper threshold values may result in undesirable detection errors, such as generating high false alarms or failing to detect potential incidents. Fuzzy set theory provides a feasible alternative scheme that may improve the reliability of incident detection systems.

As shown in Figure 1, FHWA Incident Detection Algorithm 8 can be regarded as a series of binary decision trees (*9*). Nine "incident conditions" or "operating states" can be detected by Algorithm 8. The algorithm takes input occupancy data and calculates the following traffic measures: spatial difference in occupancies (OCCDF), relative temporal difference in downstream occupancy (DOCCTD), relative spatial difference in occupancies (OCCRDF), and downstream occupancy (DOCC). The system-operating states can be determined from the decision tree analysis.

Fuzzy Applications

Fuzzy set theory has been applied successfully to many fields, including structural engineering, damage evaluation, manufacturing, medical diagnoses, meteorology, and ramp control (*10–15*). In the crisp system, when the observation is imprecise, noise prone, or near the decision boundaries, the result is easily biased or mistaken. The fuzzy approach allows users to approximate reasoning by specifying boundaries in decision-making. Fuzzy logic provides the approximate reasoning that can improve classical expert system designs using fuzzy techniques by specifying the membership func-

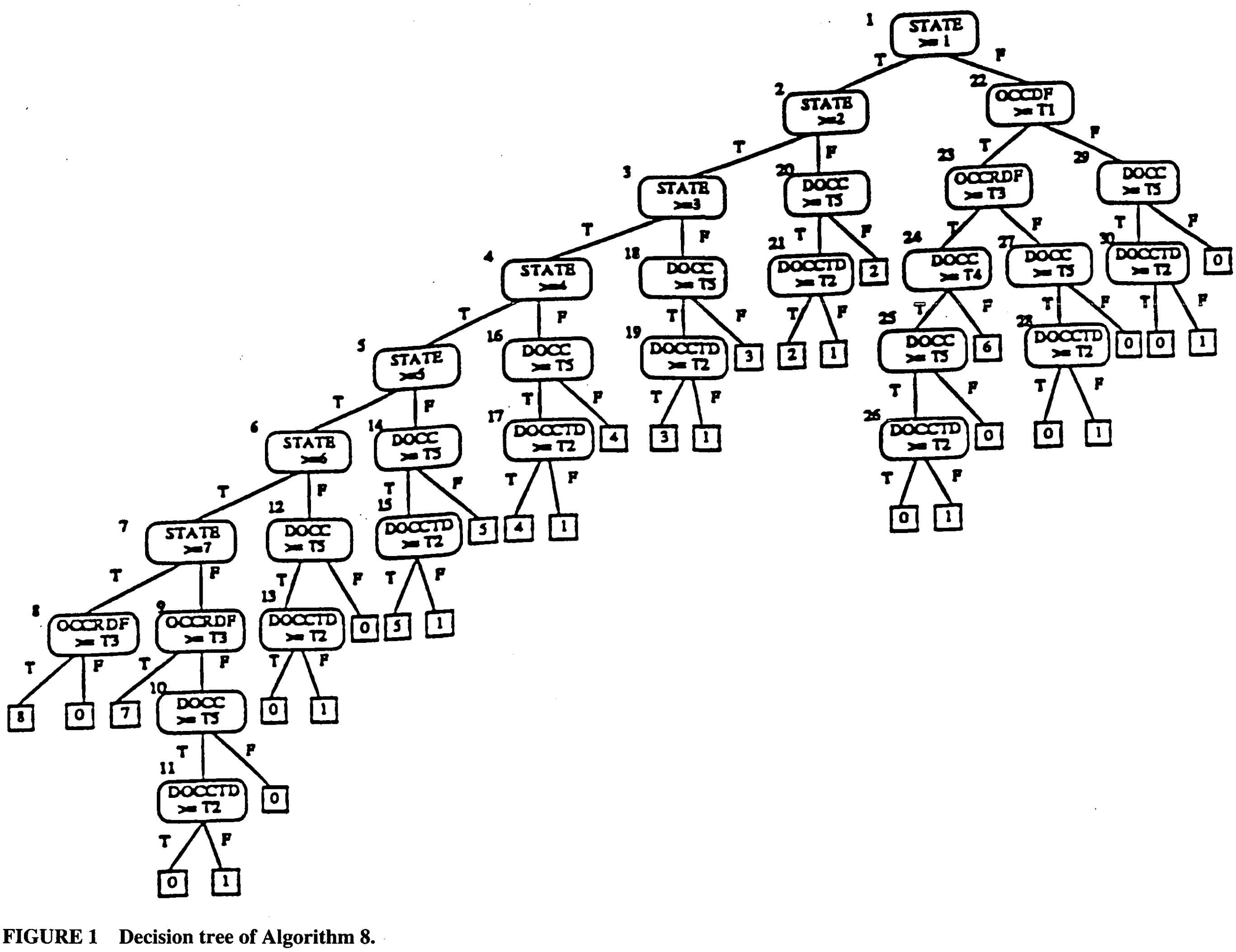


FIGURE 1 Decision tree of Algorithm 8.

tions to fine tune the system by the users. Therefore, when inputs fall near the fuzzy boundaries, fuzzy values show the significance of the inputs that can better tolerate imprecision and provide reliable results.

Fuzzy utilities are programs built with fuzzy logic. Fuzzy approach is an approach to develop "fuzzy expert systems" by combining the fuzzy utilities with existing functions of "crisp expert systems." During inferences, the fuzzy-logic–based fuzzy utilities use the membership functions to represent the likelihood of input values. The range of a membership function falls between 0 and 1, which represent the check boundaries. The membership functions can be modified when the fuzzy expert system is applied to a different location or jurisdiction to screen out the least possible inputs from the preconditions. Once the facts match the preconditions and fuzzy values are greater than the threshold, the postconditions will be propagated into the following stages.

To explore how fuzzy logic can be used to improve California Incident Algorithm 8, two different fuzzy logics are used. The first method uses commercial fuzzy expert system building tools. The other approach embeds the fuzzy utilities into crisp expert system building tools. The following section describes the development tools used to implement these two fuzzy inference systems.

Development Tools

Fuzzy expert system building tools provide built-in fuzzy features and operations so that users need not expend extra effort in developing system operational features. On the other hand, developing fuzzy systems by embedding fuzzy utilities into expert system building tools provides more flexibility in the software design that can be integrated with conventional programs, once the design can be finalized.

In this study, the first approach applies C-Language Integrated Production System (CLIPS) functions to implement fuzzy logic with Algorithm 8. The other two approaches use a fuzzy expert system building tool (FIDE or fuzzy inference development environment) with two different heuristic rules and membership functions, based on realistic freeway traffic measures, to improve the performance of the existing algorithm.

CLIPS System

The CLIPS system is a complete environment for developing expert systems, that is, programs that are specifically intended to model human expertise or knowledge (*16,17*). Designed originally by the NASA Johnson Space Flight Center, CLIPS is designed to allow artificial intelligence research, development, and delivery on conventional computers. CLIPS provides a cohesive tool for handling a wide variety of knowledge with support for three different programming paradigms—rule-based, object-oriented, and procedural paradigms.

The CLIPS system uses a "forward chaining rule language" that is based on the Rete algorithm. The term production system is an acronym of the rule-based programming paradigm. These languages include the fact lists, current state of the "world," if-then knowledge rules, and inference engine. Rule-based programming allows knowledge to be represented as heuristic, or "rule-of-thumb" that specifies a set of actions to be performed for a given situation. This object-oriented programming scheme allows complex systems to be modeled as modular components that can be easily reused to model other systems or create new components.

The procedural programming used in CLIPS 5.1 allows CLIPS to represent a knowledge base in ways similar to those allowed in languages such as C, Pascal, Ada, and LISP. Using CLIPS, one can develop expert system software using either rule-based programming, object-oriented programming, procedural programming, or combinations of the three approaches.

FIDE System

FIDE provides a user-friendly working environment for users to develop fuzzy applications (*18*). At first, the users need to prepare fuzzy sets and fuzzy rules. Several descriptive words can be used to modify the system input. For example, the users may use adjectives such as hot, warm, or cold to represent a variable temperature. These adjectives or labels can be used to associate with a membership function in the fuzzy subset. Finally, fuzzy rules can be constructed on the basis of the system analysis results.

After preparing the needed fuzzy sets and fuzzy rules, FIDE can be used for fuzzy inference (*19*). The input values are fuzzified according to the membership functions of labels. The fuzzified values can be used to refine fuzzy rule evaluation. After the system evaluation is completed, the results are defuzzified. To provide user interface capability, various equations can be further added to provide various degrees of defuzzification in the system interference analysis.

SYSTEM DESIGN APPROACH

Initial experiments, performed by TTI and others, indicated that the advanced techniques can be used to improve incident detection. These techniques include data-smoothing techniques; neural networks and fuzzy set theory are technologically feasible. Many approaches suggested that advanced techniques can improve system operations using historical detector information from preobserved freeway incidents.

The main operational advantage of fuzzy incident detection systems is to eliminate the sharp decision boundaries caused by the predefined crisp thresholds. The systems can also provide approximate reasoning to consider the uncertainty characteristics of incident detection. The proposed development, as discussed, can lead to the possible development scheme for providing the automatic training of decision thresholds.

Study Variables

The California algorithm and its variations detected an incident by determining whether the following three criteria are met:

1. The absolute difference between upstream and downstream occupancy level exceeds an established threshold value;
2. The relative difference between upstream and downstream occupancy levels, with respect to the observation from upstream detector stations, exceeds a second threshold value; and
3. The current downstream occupancy level is significantly different from the occupancy level recorded downstream 2 min before the current system reading.

As shown in Table 1, the analysis results or operating state values will be mapped to nine corresponding operating states to record the intermediate decision-making points and state of the potential incidents. The algorithm uses two basic analyses, including simple features and time series analysis. The simple feature measures site characteristics, such as occupancy and volume on the incident occurrence. The time series features, based on data consistency analysis, detect any temporal discontinuity in occupancy and volume.

Variations of this basic incident detection algorithm were further developed to distinguish incidents from normal bottleneck congestion, previous incident compression shock wave, and random traffic fluctuation by analyzing volume and occupancy from paired detectors.

Development Process

As shown in Figure 2, each fuzzy system can be divided further into three analysis stages, such as fuzzification, fuzzy inference, and defuzzification. In addition to the basic input variables, the current incident condition can be used as an input for the next interval in the fuzzy system.

Fuzzification

The fuzzification part of the fuzzy system is a mapping from the crisp inputs into fuzzy subsets. The fuzzier decides the corresponding degrees of membership functions from the crisp inputs. The resulting fuzzy values are then fed into the fuzzy inference engine.

Fuzzy Inference

The inference compositional rule is mostly adopted in the fuzzy inference (*20*). The fuzzy rule base contains a set of IF-THEN fuzzy rules. The output is obtained from the data input and fuzzy relation. The MAX-MIN operator is used.

Defuzzification

The defuzzification process generates crisp outputs from the fuzzy results. Output membership functions may be discrete or continuous. The weighted average defuzzification is mostly used for discrete membership functions. The commonly used continuous defuzzification strategies are centroid of area and mean of maximum.

Alternative Approaches

The basic idea behind development of fuzzy incident detection systems is to eliminate the sharp boundaries set by the predefined thresholds often needed in the decision-making process. In the fuzzy approaches, the thresholds are defined as fuzzy sets instead of crisp values. The decision tree can then be replaced by fuzzy rules to represent the decision-making process.

Three approaches are proposed for applying fuzzy logic to Incident Detection Algorithm 8. The first approach embeds fuzzy utilities into CLIPS. Approaches 2 and 3 use the fuzzy expert system building tool, FIDE, to develop the systems.

Approach 1

The first fuzzy system approach is developed by embedding fuzzy utilities into the CLIPS system. The original decision tree of Algorithm 8 is used in the inference process along with the MAX-MIN operations. The input variables are compared with fuzzified threshold values in each node of the decision tree. The output is the state with the highest fuzzy value.

Approach 2

The second approach uses the fuzzy expert system building tool, FIDE, to develop the system. Linguistic terms are defined for input variables and operating states. One fuzzy rule is created for each path of the decision tree of Algorithm 8. The weighted average defuzzification is recommended for generating crisp outputs for the discrete membership functions of states.

Approach 3

The third approach uses basically the same structure as that defined in Approach 2. However, simplified fuzzy rules are used to represent the decision-making functions used in Approach 3. Therefore, the

TABLE 1 Operating States of Incident Detection

NO	INCIDENT DETECTION OPERATING STATE	
	STATE	MESSAGE
1	0	INCIDENT-FREE
2	1	COMPRESSION WAVE DOWNSTREAM IN THIS INTERVAL
3	2	COMPRESSION WAVE DOWNSTREAM 2 INTERVALS AGO
4	3	COMPRESSION WAVE DOWNSTREAM 3 INTERVALS AGO
5	4	COMPRESSION WAVE DOWNSTREAM 4 INTERVALS AGO
6	5	COMPRESSION WAVE DOWNSTREAM 5 INTERVALS AGO
7	6	TENTATIVE INCIDENT
8	7	INCIDENT CONFIRMED
9	8	INCIDENT CONTINUING

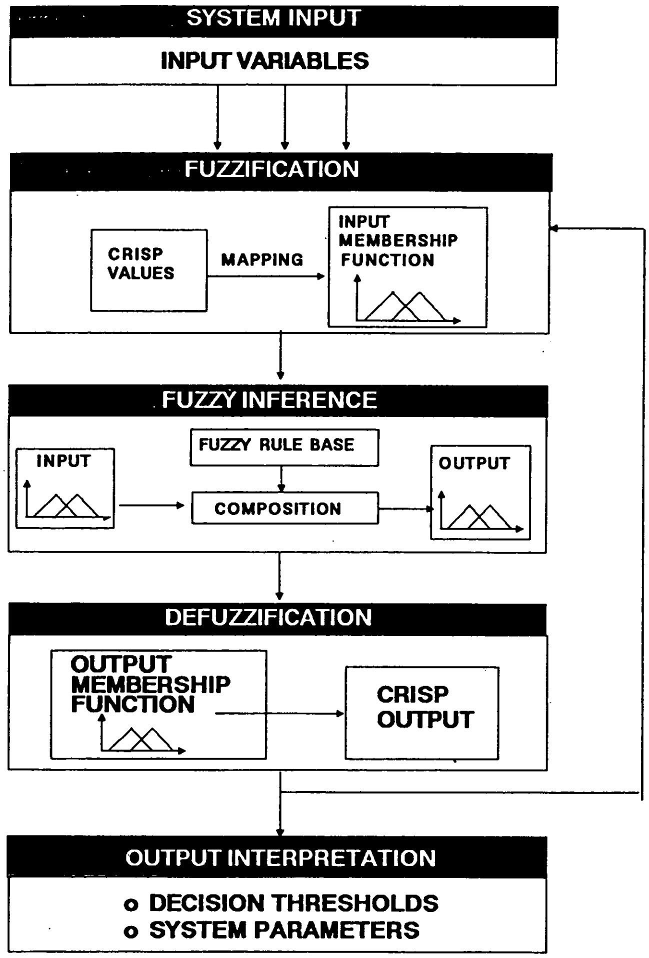


FIGURE 2 System development process.

total number of decision rules can be significantly decreased because one fuzzy rule can represent a group of paths of the decision tree.

Continuous membership functions are defined for all the operating states to reflect the different degrees of incident conditions. The left-most maximum defuzzification is used in this formulation. Figure 3 describes the exact membership functions as used in Approach 3.

SYSTEM EVALUATION

This section describes the results of the proposed approaches, evaluates the operational performance of the fuzzy systems, and summarizes the further development directions.

System Comparisons

Experimental Approaches 1 through 3 were implemented and compared with the original Algorithm 8. As shown in Figure 4, each approach proposes different methods of organizing the main components of the fuzzy systems. Approaches 2 and 3 fuzzify the input values before the fuzzy inference, whereas Approach 1 compares input variables with fuzzified thresholds during the inference process.

Because the fuzzy utilities are not provided in CLIPS, Approach 1 uses more primitive fuzzification and defuzzification processes than Approaches 2 and 3. However, the fuzzy system of Approach 1 has more potential for extending with automatic system learning abilities that can be integrated with conventional system software in the future.

Performance Evaluation

Three types of measures of effectiveness (MOEs) are often used to evaluate the automatic incident detection systems. These measures usually include the following (*21*):

1. Fraction of incidents detected,
2. Fraction of false alarms, and
3. Time to detect.

To illustrate the feasibility of the proposed approaches, several sets of occupancy data are used to test the experimental systems (*9*).

Figure 5 illustrates one example that summarized the comparison of the evaluation results with a known incident occurring at Time 11. All systems detect the incident at Time 13. The bottom half of the graph indicates the results of a set of incident-free data in which all the outputs with state Values 7 and 8 are considered as false alarms.

From this example, it was observed that all the fuzzy systems produced better results than were produced in the original Algorithm 8. The total number of false alarms is obviously reduced and the time to detect remains basically the same as that in Algorithm 8. Applying fuzzy logic to the conventional incident detection algorithm should be a feasible and practical development approach that can improve the accuracy of incident detection.

Approach Evaluation

Because of the development process required, the production rules used in the fuzzy systems can be represented in the form of plain English. As shown in Figure 6, the fuzzy approaches are much easier to understand than the original algorithm.

On the other hand, the binary decision tree of Algorithm 8 takes more time to trace, is difficult to debug, and is hard to understand. The fuzzy systems are also easier to implement initially and provide a tool that allows the user to improve the definition of a specific decision-making process that can be used in the future to maintain the workable decision thresholds.

Further Development

Two challenges still remain for the effective development of an automatic incident detection algorithm based on fuzzy set logics. At first, the performance of both nonfuzzy and fuzzy approaches depends heavily on how to generate and develop suitable thresholds, that is, membership functions in the fuzzy systems. Defining and fine tuning the membership functions in the fuzzy systems are as difficult as determining the right thresholds to be used in the original algorithm. However, if the membership functions can be systematically tuned through an automated procedure, the fuzzy systems can be developed as an effective tool to find the best threshold values (*22*).

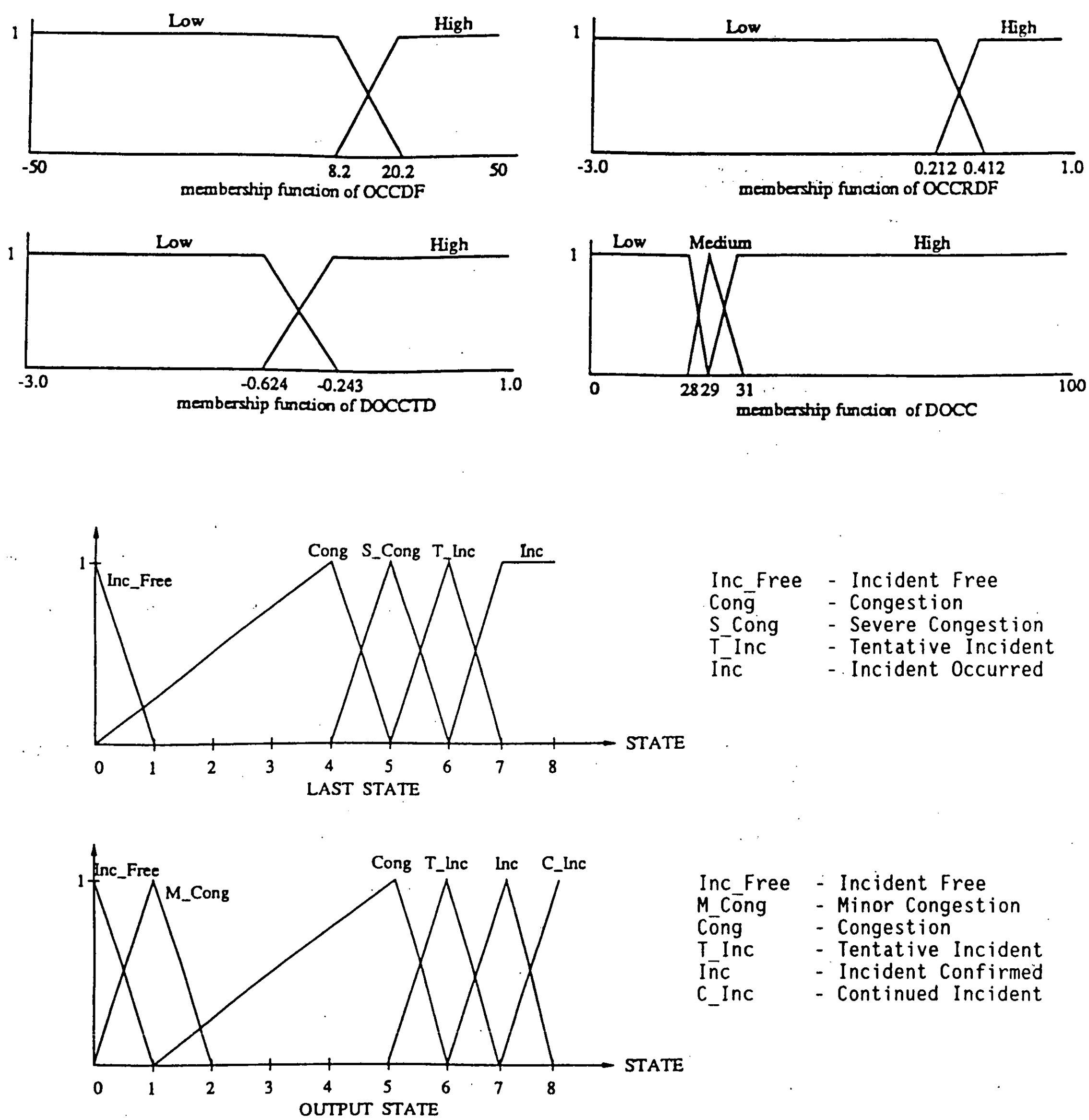


FIGURE 3 Input variable and state function.

Because most commercial fuzzy expert system building tools allow only fixed sets of membership functions and fuzzy rules, it would be difficult to include learning ability in Approaches 2 and 3. Therefore, Approach 1 can be used as the basis for enhancing freeway incident detection systems with learning ability, whereas Approaches 2 and 3 provide insights for understanding fuzzy utilities and validating the test evaluation results. Figure 7 is proposed as one design that can allow for automatic thresholds training in a computerized FTMS. The system learning function may include a meta rule base and a computational unit to automatically learn the membership functions and interface within the fuzzy system.

The second challenge is to devise a performance analysis that can distinguish performance among various fuzzy systems because most MOEs were developed initially to evaluate end states from conventional algorithms. Furthermore, instead of being cut off after each incident decision, the outputs from the previous fuzzy system decision-making state are further aggregated. For instance, during this evaluation, it was found that three different fuzzy systems under evaluation have almost the same detection rates and times to detect, and close false alarm rates. Without a specific evaluation of the intermediate state detection, the currently used MOEs cannot effectively reflect and evaluate the detection of intermediate states. Therefore, a composite evaluation index should be developed for evaluating different fuzzy system designs.

CONCLUSIONS AND RECOMMENDATIONS

AID, based on the real-time detector measurements from the computerized FTMS, are increasingly used to reduce the impacts of unexpected incidents. The successful operations can minimize undesirable congestion and regional mobility loss and provide necessary motorist information. The operations of the overall surveillance,

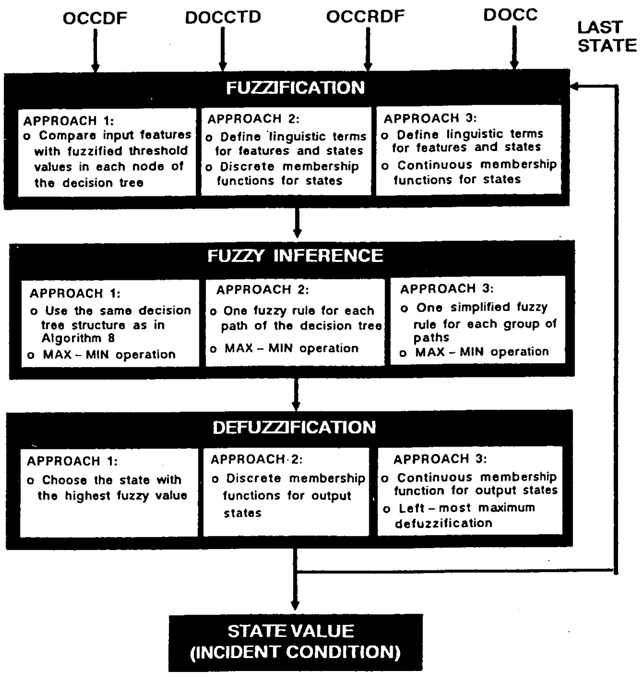


FIGURE 4 Comparison of three approaches.

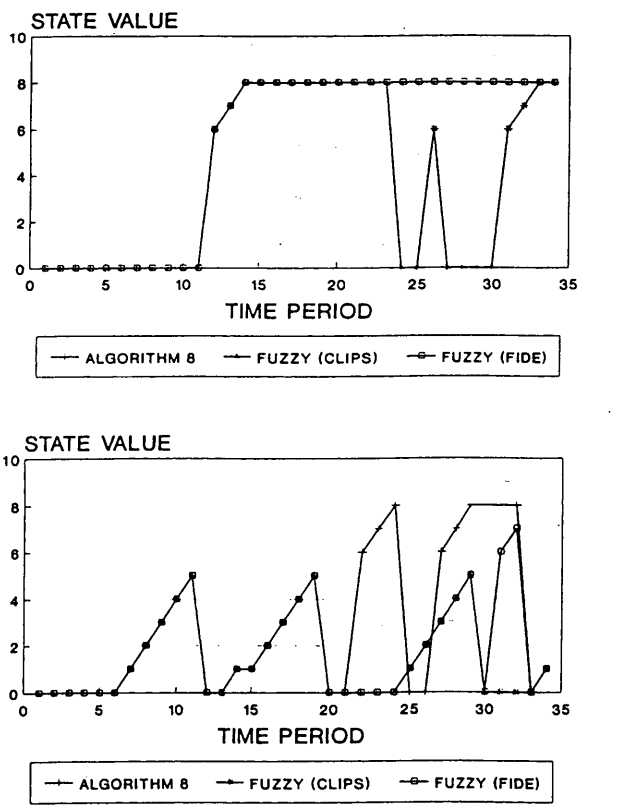


FIGURE 5 Incident test results. Results of fuzzy and nonfuzzy systems: *top*, incidents occurred at Time 11, *bottom,* incident-free data.

```
if last_s is Inc and OCCRDF is High then state is C_Inc;

if last_s is Inc and OCCRDF is Low  then state is Inc_Free;

if last_s is T_Inc and OCCRDF is High then state is Inc;

if last_s is T_Inc and OCCRDF is Low  and DOCC is High and
    DOCCTD is High then state is Inc_Free;

if last_s is T_Inc and OCCRDF is Low  and DOCC is High and
    DOCCTD is Low  then state is M_Cong;

if last_s is T_Inc and OCCRDF is Low  and DOCC is notHigh
    then state is Inc_Free;

if last_s is S_Cong and DOCC is High and DOCCTD is High then
    state is Inc_Free;

if last_s is S_Cong and DOCC is notHigh then state is Inc_Free;

if last_s is Cong and DOCC is High and DOCCTD is High then
    state is Cong;

if last_s is Cong and DOCC is notHigh then state is Cong;

if last_s is Inc_Free and OCCDF is High and OCCRDF is High and
    DOCC is Medium then state is Inc_Free;

if last_s is Inc_Free and OCCDF is High and OCCRDF is High and
    DOCC is Low    then state is T_Inc;

if last_s is Inc_Free and OCCDF is High and OCCRDF is Low  and
    DOCC is notHigh then state is Inc_Free;

if last_s is Inc_Free and OCCDF is Low  and DOCC is notHigh
    then state is Inc_Free;

if last_s is Inc_Free and DOCC is High and DOCCTD is High then
    state is Inc_Free;

if last_s is notInc and DOCC is High
    and DOCCTD is Low then state is M_Cong;
```

FIGURE 6 Example of linguistic rules.

communications, and control system, however, relies on the accurate and effective usage of incident detection algorithms.

Conclusions

This study explores the use of fuzzy logic to improve the operations of California Incident Detection Algorithm 8. Three design approaches were investigated and found to be equal or superior in performance to the conventional systems for effective freeway management. Several issues still remain for the practical development of automatic incident detection techniques, including the following:

1. Evaluation of the operational effectiveness of existing incident detection algorithm(s) in the field;
2. Assistance in developing users' guidelines to provide the suitable operator interface and settings during different conditions; and
3. Recommendations on how to integrate alternative data sources and improve existing automatic incident detection system design.

Recommendations

The fuzzy rules represented by the linguistic terms in fuzzy set systems are much easier to understand and debug than the decision tree commonly used in the nonfuzzy conventional approach. Fuzzy systems are also easier to maintain and adjust to various traffic control environments. Because different operating states can be used to represent degrees of severity of the freeway incidents, fuzzy rules can be grouped to simplify the operating states with similar behavior in

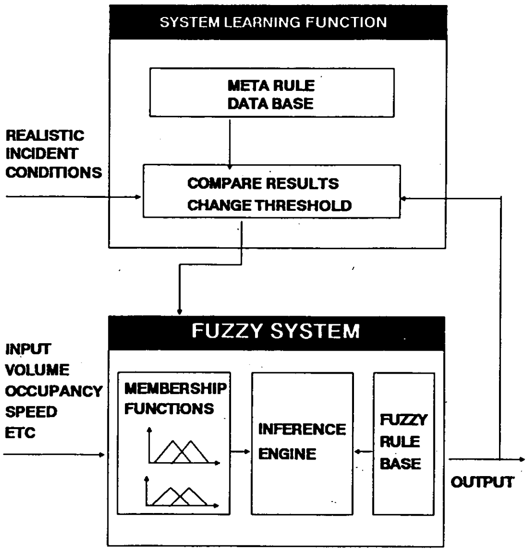


FIGURE 7 Automatic process to train thresholds.

the decision tree analysis. These simplified rules can further improve future operational efficiency and maintenance of fuzzy systems.

However, the performance of incident detection systems depends on how to generate suitable "threshold values" or "membership functions" in the fuzzy systems. Two problems still exist with the uses of fuzzy reasoning, including the lack of methods to determine proper thresholds and lack of automatic learning functions or adaptability in the algorithm (*22*). A fuzzy incident detection system design with automatic learning design can greatly improve the performance of an incident detection system. In addition, the conventional MOEs are not sufficient to evaluate among different fuzzy systems. A composite index is needed to improve the incident detection evaluation during intermediate states. Furthermore, the index can be used to reflect the degree of information usage and to measure the quality of information received for further system improvements.

REFERENCES

1. Chang, E. C., and K. Huarng. Freeway Incident Detection using Advanced Technologies. Presented at 71st Annual Meeting of the Transportation Research Board, Washington D.C., Jan. 1992.
2. Chang, E. C., and K. Huarng. Automatic Freeway Incident Detection Applying Fuzzy Set Concept. *Proc.* 1993 Annual Science, Engineering, and Technology Seminar (SETS), Association of American-Chinese Professionals (AACP), Houston, May 1993.
3. Chang, E. C., and K. Huarng. Fuzzy Set Applications for Freeway Incident Detection. *Proc., Intelligent Vehicles '93 (VNIS '93)*, Institute of Electrical and Electronics Engineers, Society of Automotive Engineers, American Society of Mechanical Engineers, Japan Society for Fuzzy Theory and Systems, Robotics Society of Japan, and Society of Instrument and Control Engineers, Tokyo, July 1993.
4. Payne, H. J., and S. C. Tignor. Freeway Incident-Detection Algorithms Based on Decision Trees with States. In *Transportation Research Record 682,* TRB, National Research Council, Washington D.C., 1978, pp. 30–36.
5. Kaufmann, A., and M. M. Gupta. Introduction to Fuzzy Arithmetic: Theory and Applications. Van Nostrand Reinhold, 1991.
6. Zadeh, L. A. The Calculus of Fuzzy—If/Then Rules. *AI Expert,* March 1992.
7. Courage, K. G., and M. Levin. A Freeway Corridor Surveillance, Information, and Control System. Research Report 488–8. Texas Transportation Institute, Texas A&M University, College Station, 1968.
8. Dudek, C. L., and C. J. Messer. Incident Detection on Urban Freeways. In *Transportation Research Record 495,* TRB, National Research Council, Washington D.C., 1974, pp. 12–24.
9. Payne, H. J., and H. C. Knobel. *Development and Testing of Incident Detection Algorithms, Vol. 3—User Guidelines.* FHWA-RD-76-21, 1978.
10. D. I. Blockley. *The Nature of Structural Design and Safety.* Chichester. 1992.
11. Ishizuka, M., K. S. Fu, and J. P. T. Yao. A Rule-Based Inference with Fuzzy Set for Structural Damage Assessment. In Approximate Reasoning in Decision Analysis (M.M. Gupta and E. Sanchez, eds.), 1982.
12. Mamdani, E. H. Advances in the Linguistic Synthesis of Fuzzy Controllers. Fuzzy Reasoning and Its Applications (E.H. Mamdani and G. R. Gaines, eds.), 1981.
13. Vila, M. A., and M. Delgado. On Medical Diagnosis Using Possibility Measures. *Fuzzy Sets and Systems,* Vol. 10, 1983.
14. Cao, H., and G. Chen. Some Applications of Fuzzy Sets of Meteorological Forecasting. *Fuzzy Sets and Systems,* Vol. 32, 1983.
15. Chen, L. L., and A. D. May. *Freeway Ramp Control Using Fuzzy Set Theory for Inexact Reasoning.* Research Report UCB-ITS-RR-88-10. Institute of Transportation Studies, University of California, Berkeley, May 1988.
16. *CLIPS Reference Manual, Version 5.1 of CLIPS, Vol. I.* Software Technology Branch, Lyndon B. Johnson Space Center, Sept. 1991.
17. *CLIPS User's Guide, Version 5.1 of CLIPS, Vol. II.* Software Technology Branch, Lyndon B. Johnson Space Center, Sept. 1991.
18. *Fide Tutorial,* Aptronix, 1992.
19. *Fide Reference Manual.* Aptronix, 1992.
20. Zimmermann, H.-J. *Fuzzy Set Theory and Its Applications,* 2nd ed. Kluwer Academic, 1991.
21. Stephanedes, Y. J., A. P. Chassiakos, and P. G. Micchalopoulos. Comparative Performance Evaluation of Incident Detection Algorithms. In *Transportation Research Record 1360,* TRB. National Research Council, Washington D.C., pp. 50–57.
22. Takagi and I. Hayashi. Artificial Neural Network-Driven Fuzzy Reasoning. *Proc. International Workshop on Fuzzy System Applications,* Aug. 1988, pp. 217–218.

Publication of this paper sponsored by Committee on Artificial Intelligence.

Applications of Artificial Neural Networks to Intelligent Vehicle-Highway Systems

JIUYI HUA AND ARDESHIR FAGHRI

The potential applications of artificial neural networks (ANNs) to intelligent vehicle-highway systems (IVHS) are evaluated, and the extent of use and the position of ANNs in future IVHS implementations are discussed. The state of the art and the potential implementation needs of IVHS are reviewed, and the characteristics, properties, limitations, and application domains of ANNs are discussed from a technical perspective. On the basis of review and discussion, potential application domains of ANNs to IVHS are evaluated. A technical demand-supply matrix is provided to indicate the most possible potential application domains of ANNs to IVHS. As an application case study of ANNs in the implementation of IVHS, an ANN-based model is established for vehicle travel time estimation. The results and findings associated with the development of the neural network-based travel time estimation model are also reported. It is concluded that (*a*) ANNs can provide most techniques needed by IVHS, and (*b*) for some IVHS implementation domains, ANNs may be superior to conventional techniques. The case study demonstrates the modeling feasibility of ANNs for potential IVHS implementations.

Today's vehicle-highway systems face a variety of problems, including increasing congestion, rising accident rates, alarming environmental pollution, and depletion of energy sources. Building more highways, the traditional solution, is no longer effective because building involves intensive development for which there are limited financial resources. Instead, upgrading existing vehicle highway systems by making them "intelligent" is being considered as one way to curtail or resolve those problems. Such upgrading is made feasible by the computerization of existing facilities and operations with applications of a series of advanced technologies, such as expert systems, neural networks, and fuzzy set theory. Unlike traditional transportation systems, intelligent vehicle-highway systems (IVHS) will possess more power in dealing with dynamic and uncertain problems. It is anticipated that advanced technologies will be a large share of the technical base for future IVHS development.

Inspired by the success of initial efforts, many attempts have been made to apply various advanced technologies to IVHS. Nevertheless, it is also true that each of these advanced technologies has its own proper application domains and limitations. As a branch of artificial intelligence, artificial neural networks (ANNs) are considered useful techniques for a wide variety of IVHS implementations. There is, however, less literature work available for a systematic discussion about the domains, extents, and limitations of ANN applications in IVHS implementations. Being a tool of problem solving, ANNs have been highly recognized for certain tasks in several areas, such as financing, business, and computer science. It is important to the future of IVHS programs to understand how far and to what kind of applications ANNs can be applied to IVHS implementations.

Department of Civil Engineering, University of Delaware, Newark, Del. 19716.

In the past, there have been few reports in the literature about the application of ANNs to IVHS programs (*1,2*). Nevertheless, the successful application of ANNs in various areas strongly implies great potential for the application of ANNs in IVHS. This paper focuses on a comprehensive evaluation of the application of ANNs to IVHS. The technical needs of IVHS and the technical properties of ANNs are analyzed and summarized. On the basis of on an in-depth investigation of the characteristics of IVHS and ANNs, the interface between IVHS and ANNs is drawn in terms of the technical demand-supply relationship. As a result, the technical conjunction of IVHS and ANNs is reclarified. Several matrices that indicate the technical requirements of IVHS, the technical properties, and the potential application domains of ANNs to IVHS are provided. A simplified version of ANN application to a potential IVHS problem is also presented to demonstrate an application of ANNs in a modeling process. Results of the application example also are discussed.

TECHNICAL REQUIREMENTS OF IVHS

IVHS is a collection of advanced applications for information processing and computer technologies. The major benefits provided by IVHS are greater mobility for highway users and a higher level of safety. So far, several implementations of IVHS already have been achieved or are in progress. To illustrate the technical needs of IVHS, the following IVHS technologies have been implemented and are summarized.

- Leit- Und Information System of Berlin (LISB) (*3*),
- Autoscope (*3*),
- Advanced Mobile Traffic Information and Communication System (AMTIC) (*3*), and
- Automatic Vehicle Identification (AVI) System (*4*).

LISB (*3*) is a route guidance information system that encompasses about 500 vehicles and a network of approximately 3000 km of roadways. The network covers 4,500 intersections and about 1,300 signals. Approximately 250 beacons set along the roadside provide information to the vehicles. This is an information processing intensive system.

AUTOSCOPE (*3*), a driver navigation system, was codeveloped by the Minnesota Department of Transportation and the University of Minnesota. AUTOSCOPE uses video cameras and computers to analyze images and extract traffic flow information for surveillance and control purposes. The system includes roadside image sensors and closed-circuit television devices. The computer vision system processes images and generates data identifying congestion and incidents. Artificial intelligence systems are used to aid in the analysis of alternatives for congestion control.

AMTIC (*3*) was developed in Japan. Traffic information is sent to the system's center from a guide terminal and sign posts through a cable network installed along the ground. The sign posts and guide terminal also relay signals from the system's automobiles through the ground network. The automobiles have on-board signal-receiving equipment and advisory display systems.

The AVI system (*4*) identifies vehicles without stopping or even slowing the traffic flow, thereby improving highway mobility. Its most promising application at this time is in toll booth functions. With the success of the AVI system, toll booths are expected to become highly automated by providing nonhuman operation for nonstop toll charge services. A prototype of AVI was set up on the Dallas (Texas) North Tollway in 1988. Another three AVI-equipped highways are also under construction in Orange County, California.

As seen in the aforementioned programs, IVHS makes intensive use of information processing, control, image processing, pattern recognition, optimization, classification, prediction, and knowledge-processing techniques.

IVHS is composed of four elements. Each element is a system that contains a variety of advanced hardware implementations and software technologies. The four elements are as follows:

- Advanced traffic management system (ATMS),
- Advanced driver information system (ADIS),
- Advanced vehicle control system (AVCS), and
- Commercial vehicle operation (CVO).

ATMS consists of several procedures, including detection, classification, prediction, optimization, communication, and control technologies. The information about traffic conditions on the highways is detected and transmitted to a central management processor. After processing the input information, the central management processor sends commands or advisory information back to the vehicles in the traffic flow.

ADIS provides information to individual motorists. The information could be travel related, such as traffic conditions, advice on route selections, or travel costs, or nontravel related, such as the location of a theater, if an extensive ADIS is established. Intensive information processing, massive transportation data, data compression, and knowledge-processing technologies are required because of the wide variety of information communicated to individual drivers.

AVCS enhances the performance of motor vehicles. The implementation of AVCS is expected to bring significant improvements in highway safety, fuel efficiency, and elimination of air and noise pollution. It is also expected to affect some human limitations in driving, such as perception/reaction time, thereby changing the basic models and designs of highway transportation systems. Prediction, classification, detection, and control techniques are used intensively in AVCS.

CVO applies to commercial and emergency vehicles. The implementation of CVO provides faster, more reliable, and more efficient services, such as the collection and delivery of goods through the monitoring of fleet vehicles. Communication, optimization, and control techniques are necessary for the implementation of CVO.

In summary, the major technical requirements involved in IVHS implementation should be in accordance with intelligent techniques; that is, they are capable of performing well even under uncertain or dynamic circumstances, or both, giving heuristic solutions when exact solutions are unnecessary or impossible; and reasonably ignoring noises that exist in the input data. Figure 1 indicates major techniques required for IVHS implementation on a category basis. These techniques could be applied through either hardware implementation or software programming.

IVHS Elements / Technical Requirments	ATMS	ADIS	AVCS	CVO
prediction	●	●		●
classification	●			
control	●		●	●
pattern recognition	●	●	●	
optimization	●			
noise cancellation	●	●	●	
image processing			●	
knowledge processing	●	●		
signal processing			●	
data compression		●		
natural language processing			●	●
speech recognition			●	●

● Useful technique for the corresponding IVHS element

FIGURE 1 Technical requirements of IVHS.

PROPERTIES OF ANNs

ANNs are parallel computing techniques that conceptually mimic human mental neural structure and functions. The number of ANN applications in different fields has been expanding rapidly since the middle 1980s because of their attractive characteristics and capabilities such as learning, abstraction, and generation. Although it is preferable that they be operated on parallel computing devices, ANNs problem-solving abilities are still impressive in their performance of many tasks from a noncomputing speed point of view. Although limited in number, ANN applications have covered a variety of problem domains in transportation engineering (*5*). ANN applicable problem domains include prediction, classification, control, optimization, and modeling. Some of the ANN uses most frequently encountered in transportation engineering are summarized.

Prediction

Prediction is a basic technique of transportation planning. It determines the "appropriate" actions to be taken in the future on the basis of historical data and reasonable logic or inferences. The perceptron types of ANNs can be trained to form appropriate architectures for the prediction procedure. Training data can be obtained in the same way, from either an historical data base or a current survey. The modeling of the input-output relationship is completed by training the ANN. The advantages of applying ANNs to prediction problems are (a) no programming or formulation effort is required; and (b) to some degree the noise existing in the data base can be removed. Neural network models such as ADALINE and backpropagation have been applied to zonal trip production prediction (*5*). The backpropagation model also has been applied to the prediction of short-term traffic volume on a highway segment (*6*).

Control

Control is an important procedure in traffic management and transportation operations such as vehicle driving. Control logic is defined as giving an appropriate output when an input is presented to the controller. The output of a control logic model could be an action to take, the amount of an adjustment, or some other measurements. It is often difficult, however, to formulate the input-output logic model mathematically. Sometimes irregular requirements, such as insensitivity of an output value with a specific range of input values, are necessary. Ordinary mathematical approaches are unlikely to form an arbitrary curve. In most cases, the control logic model could be complicated even more by providing multiple inputs and producing multiple outputs. In contrast, a multilayered nonlinear perceptron ANN is able to approximate any reasonable continuous function to an arbitrary degree of accuracy (*7*). A number of control applications have been reported in the area of transportation engineering. These applications are widely scattered in a variety of domains, including roadway network traffic signal control (*8,9*), vehicle control (*10*), in-vehicle device control (*11*), and air traffic control (*12*).

Optimization

Neural networks in general provide heuristic optimization. A typical optimization application is the traveling salesperson problem. A recurrent network called the Boltzman machine is used to determine the salesperson's optimal route. A simulated annealing technique is incorporated in the architecture of the neural network. The network attempts to discover the global minimum. It allows, however, for an alternative point that is near the true optimal point in the solution space to be the ultimate solution.

Classification

ANNs can perform classification in various fashions. Some of the neural network approaches provide better classification than do conventional approaches. For instance, an application that uses ART1 (adaptive resonance theory), an ANN model designed to solve the stability-plasticity dilemma, to classify dynamic traffic patterns at a network level is much more efficient than an ordinary dynamic programming approach (*8*).

Pattern Recognition

Perhaps the most remarkable property of ANNs is that of pattern recognition. Much success has been reported in this application domain (*13*). The pattern recognition abilities of ANNs are especially useful for computer vision because of their capability for dealing with incomplete, noisy, or distorted patterns. Such properties could be extended to obstacle detection, pedestrian detection, and even traffic incident detection problems.

ANNs possess a wide variety of properties, and although the performance of each ANN property has not yet been fully evaluated, ANNs are most certainly able to provide transportation engineering, including IVHS, with many useful problem-solving approaches. Figure 2 lists some major ANN properties that are frequently needed in transportation engineering.

POTENTIAL APPLICATION OF ANNs TO IVHS

The four IVHS elements—ATMS, ADIS, AVCS, and CVO—are implemented primarily through the use of two advanced techniques: advanced information processing and advanced modeling. Advanced information processing includes numerous techniques such as data acquisition, data arrangement, knowledge processing, classification, and transportation data processing. Modeling techniques are used to build desirable relationships between given inputs and the outputs of models, such as the control logic model and the prediction model. Neural networks provide numerous versions of these two types of techniques as a technical basis for IVHS. In this section, several major technical properties provided by ANNs and their potential applicability to IVHS are discussed. The limitations of ANNs for use in IVHS are discussed at the end of this section.

Data Acquisition

In this study, the term data acquisition does not refer to the capability of hardware to obtain information but instead to the capability of software to correctly and precisely obtain the desired information. For instance, for the purpose of pattern recognition the preferred data acquisition technique should be able to correctly catch patterns that may be incomplete or noisy because one of the major advantages offered by ANNs is recognizing noisy, incomplete, or distorted patterns. Many such ANN applications have proven to be successful (*14–18*). For IVHS implementations, ANNs can be applied to computer vision to acquire information through pattern recognition. They also can provide reliable and feasible input information to other processors such as algorithms for obstacle detection (*19*) or an autonomous driving system (*10*).

Knowledge Processing

ANNs are useful in dealing with abstract information. In real life, including IVHS systems, many pieces of information are abstract, that is, it may not be possible to represent them by a single, clearly defined concept. Sometimes, their definition varies, depending on who is doing the defining. An example of such a case is making a judgment about whether a specific congested traffic condition is recurrent or nonrecurrent. An ANN can store different opinions and summarize them to give a single-valued reasonable "standard." This characteristic of ANNs can be utilized as a decision-making procedure with the varying-input information needed for such IVHS applications as incident detection systems.

Classification

A number of ANN models can be used as classifiers (*20–22*). An interesting ANN model, ART, has been identified as a good tool for distinguishing different attributes of roadway segments on the basis

ANN Applications / ANN paradigms	control logic modeling	pattern recognition	prediction	noise cancellation	knowledge processing	classification	image processing	optimization
(M)ADALINE	●		●	●			●	
Backpropagation	●	●	●		●		●	●
ART	●				●	●	●	
Hopfield Network	●	●				●	●	
Boltzman Machine		●			●		●	●
Neocognitron / Cognitron		●					●	
Counterpropagation						●		
Kohonen Net						●		
Neural GMDH								●
BSB					●	●	●	
LVQ	●						●	●
BAM	●						●	

FIGURE 2 ANN application domains.

of traffic patterns taking place on those roadway segments. The sensitivity of ART in detecting the differences between traffic patterns is adjustable. It is also capable of on-line training, that is, no predetermined classification algorithm is required. The classification is largely based on the analog similarity among the patterns instead of numerically computed error values. Another advantage of using ART is that the haphazard, unseen irregular traffic patterns will not affect the memory before it is stored.

Data Compression

Because intensive information and data on transportation are required for IVHS applications, such as in ADIS and ATMS, the collection of accurate data is extremely important. Also, because IVHS serve dynamic systems and require massive communication with time limits, the efficiency of information transmission becomes a key issue affecting the success of many IVHS applications. Data compression is one way to improve efficiency. ANNs are good tools to compress data for quickening data transportation (*23*). The utilization of data compression by ANNs should be expected soon for ADIS.

Modeling

As mentioned earlier, the neural network (multilayered perceptron with nonlinear transfer function) can approximate any reasonable function. It can be applied to virtually any modeling problem if sufficient data are provided for the training. Especially when the system to be modeled is very complicated, conventional mathematical approaches are often impossible to use for completion of the procedure. In IVHS, many tasks such as control and forecasting need to be completed through models that compute or react reasonably to external stimulus. In the real world, relationships between the external stimulus and the outputs of the model are frequently unknown. Multiple inputs/outputs is another often-seen situation in which the relationships are difficult to be expressed mathematically. For example, travel time in a weaving section of highway can be affected by many factors, such as number of lanes, driver characteristics, and visibility of the highway section. Some of those factors are known, some are unknown. The travel time should be a function of these factors, however, it is difficult to tell how these factors will affect travel time. To deal with such a complex problem using conventional mathematical approaches, in most cases, is establishing empirical equations that require a great deal of effort in calibration and model validation. Similar problems are encountered even more often in control logic modeling. Nonlinearity often prevents building a reasonable model or even the ability to build that model. ANNs can be trained to obtain appropriate architecture that functions desirably. As long as reasonable and sufficient data are available, models can be established. With an ANN approach, it is not necessary to identify how every factor affects the output of the model. The importance of each factor is automatically identified through the training. Therefore, less knowledge about the nature of the problem is required.

Limitations

Although ANNs provide many useful properties that meet IVHS needs, there are several drawbacks to the use of ANNs that should be noted.

- Many ANN paradigms have not yet been well examined;
- In most cases, ANNs can provide only heuristic solutions;
- Some inherent drawbacks, such as determining the proper learning parameter during backpropagation training remain unsolved; and
- Computing speed is highly dependent on hardware implementation.

There is still a long way to go to fully exploit the advantages of ANNs implementation for IVHS. For control IVHS applications, hardware implementations are needed to accompany software development for the full utilization of the ANNs advantages.

In summary, as seen in Figure 1 and Figure 2, the technological requirements of IVHS and the technological contributions made by ANNs meet at many points. Figure 3 indicates the potential application domains of ANNs to IVHS. Many ANN properties can be directly applied to IVHS implementations such as classification, pattern recognition, image processing, knowledge processing, and control logical modeling. Some, such as speech processing (*24*), and natural language processing (*25,26*) are expected to be applicable in cooperation with other supplementary approaches in future applications.

ANN APPLICATION EXAMPLE—ANN TRAVEL TIME ESTIMATION MODEL

As indicated earlier, ANNs could be used for numerous IVHS implementations. In specific circumstances, various IVHS tasks may be completed by various ANN models or their combination with other approaches. In this case study, a limited demonstration is presented to explore the applicability of a specific ANN model, backpropagation (BP) network, in modeling problem. An experiment was undertaken to create an ANN travel time estimation model. The computational details will not be presented here because this is a demonstration version.

As known, travel time estimation is an important process in many decision-making procedures in traffic management. In IVHS, travel time is a basic piece of information needed for ATMS and ADIS. On the other hand, the travel time estimation process is complex because any of a large number of factors could affect motor vehicle travel speed. This case will examine a neural network travel time estimation model for a two-lane roadway segment that is assumed to be an urban roadway with a Class 2 ranking and a 30-mph free flow speed (Figure 4). The model will deal with the following situations: normal condition, blockage in the left lane, and blockage in the right lane.

Problem Statement

Roadway condition is considered the most important factor affecting travel time in an urban network. Many studies have examined travel time estimation for a particular roadway segment under normal conditions. The problem becomes much more complex under abnormal conditions such as a lane blockage. The complexity is the result of several factors:

- The reduction in the number of available lanes,
- Weaving movement,
- Positions of weaving taking place,
- Length of the blockage, and
- Position of the blockage.

It is difficult to determine which and by how much these factors will ultimately affect the average travel time. Empirical mathemat-

ANN Application Domains / IVHS Applications	control logic modeling	pattern recognition	prediction	noise cancellation	temporal pattern recognition	classification	image processing	speech & language
route guidance			●					
vehicle identification		●					●	
neural suspension	●							
obstacle detection		●					●	
driving control	●						●	●
navigation				●				
collision avoidance	●			●				
traffic control		●	●	●	●	●		
incidence detection		●				●	●	
traffic information			●	●				

FIGURE 3 Potential applications of ANN.

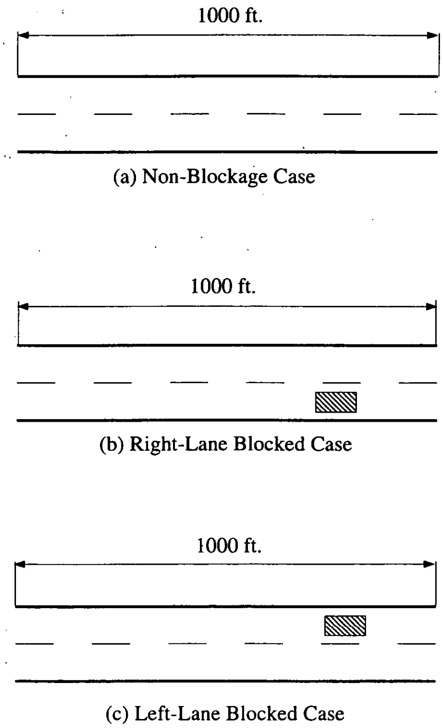


FIGURE 4 Schematic drawing of street segment for simulation of TRAF-NETSIM.

ical modeling has been the primary approach used to express the relation between travel time and the group of influencing factors. However, the method of determining the basic form of the equation, that is, to approximate the curve of travel time versus the influencing factors, is still nondeterministic. Although some mathematical processes such as Fourier transformation are theoretically able to model any curve, it is virtually impossible if the polynomial equation is large.

Establishment of ANN Model

As mentioned earlier, the BP ANN is in wide use to reasonably approximate any curve on a multiple input-output basis. It is also convenient to update simply by feeding new data into the system. For these reasons, the BP is considered to be a flexible neural network paradigm for travel time estimation modeling.

The raw data for training the BP for this case study is obtained from a simulation run on the TRAF-NETSIM software, which was developed for urban traffic network simulation. The simulation was divided into three groups: Group 1, normal roadway conditions; Group 2, right lane blockage; Group 3, left lane blockage. Each group contains 32 cases with different average traffic volumes. The simulation period for each case is 5 min. After running those cases, the simulation results for 5-min traffic volume and average through time for every case was recorded as if for BP training use.

The BP has a two-layered architecture with five hidden processing units. It requires three inputs: 5-min traffic volume, blockage index for the right lane, and blockage index for the left lane. The latter two inputs are binary, that is, set to 0,0 if the roadway is not blocked, 1,0 if the right lane is blocked, and 0,1 if the left lane is blocked. Only one output is given by the BP—the average through time in seconds.

The total of 96 data sets was divided into two groups with a random order. One of the groups was used to train the network and the other was used to test the training effect. In the training, the BP reached the error minimum on the testing data group at Training Iteration 4900. The total squared errors generated by the BP on the training data group and the testing data group were 1,225.08 and 1,449.68 sec ^{2}, respectively.

Figures 5 through 7 indicate the performance of the neural travel time estimation model under three different roadway conditions. The line is drawn by the neural network estimation model, and the dots are the simulation results.

Results and Discussion

The application example shown in this study is simple, but it is detailed enough to show the feasibility of the modeling process. Many factors affecting travel time were not taken into consideration because of the simulation program used. The BP neural network estimated the travel time on a specific roadway segment for all the three scenarios with an average error of 5.5 sec compared with the results obtained by TRAF-NETSIM. Such a result is considered conceptually reasonable because no other models are available with which to compare it. Figures 5 through 7 indicate the performance of the neural network travel time estimation model under three different roadway conditions. The lines were drawn by the neural network estimation model, and the dots are the simulation results. In addition to not having any requirements for addressing explicitly the characteristics of the input-output relation such as polynomial, exponential, and logarithmic, for concepts and operations of the model, the following are possible extended advantages of using ANN modeling.

- More complex traffic and roadway situations can be dealt with,
- Updating the ANN model is simple, and
- The ANN model requires less mathematical effort.

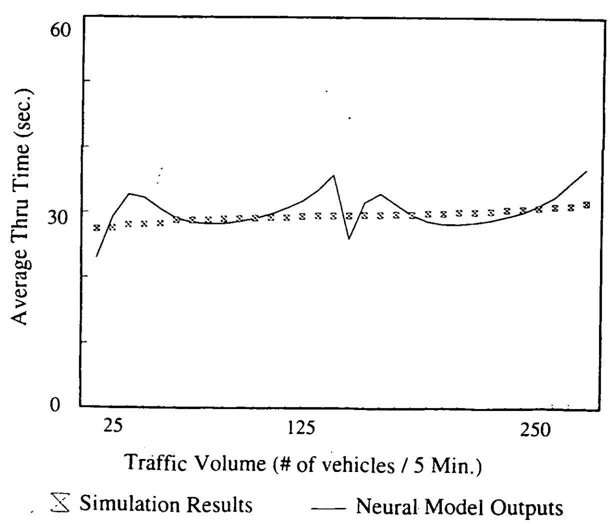


FIGURE 5 Traffic volume versus travel time for nonblockage case.

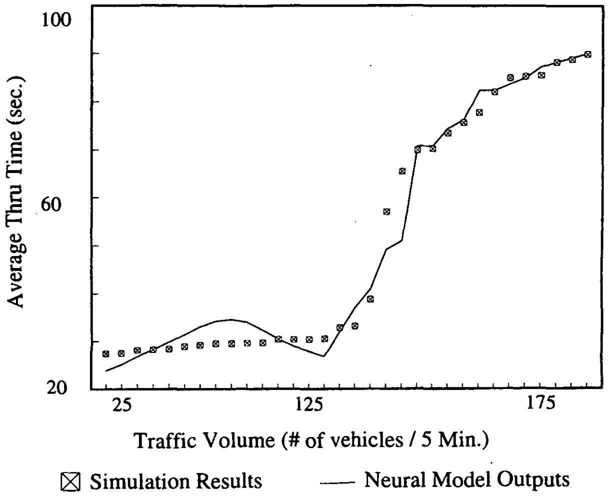


FIGURE 6 Traffic volume versus travel time for right lane-blocked case.

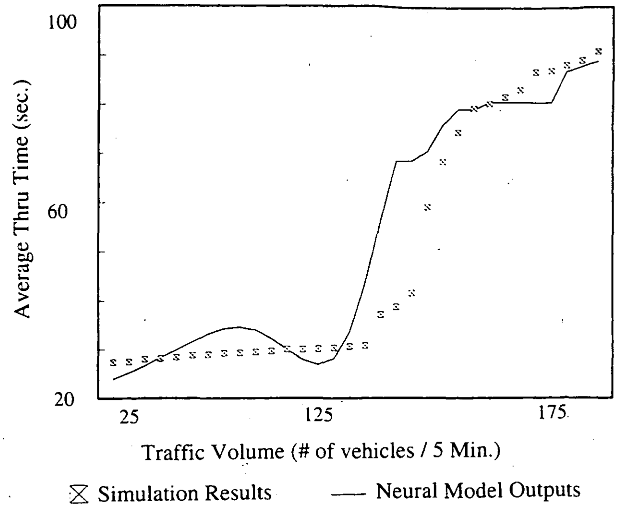


FIGURE 7 Traffic volume versus travel time for left lane-blocked case.

The ANN models provide heuristic solutions instead of exact solutions. Because ANNs are operated on a vector basis and each element of the vector can be treated independently, more complex problems can be dealt with. Furthermore, the relationship between each element in the input vector and the outputs are allowed to be unknown. Therefore, ANNs are considered to be particularly useful for those estimation or prediction problems in which accuracy is not necessary or not guaranteeable.

SUMMARY AND CONCLUSIONS

ANNs provide many properties that meet a number of technological requirements of IVHS, especially in information processing and modeling. The intelligent operations provided by ANNs are, in many aspects, in accordance with the application orientation of IVHS. Some ANN techniques, such as pattern recognition and control logic modeling, are expected to contribute excellent performances when used in IVHS. There is a great potential for ANN techniques to be superior to conventional techniques in such areas as pattern operation and modeling process in certain IVHS applications. The vector-level operations of ANNs—handling multiple input-output—also will better meet the complexity of transportation systems. As seen in Figure 3, numerous direct and indirect applications of ANNs to IVHS should be expected.

In the application example, the BP formulated a model that was able to reasonably estimate the average travel time on a given roadway segment on the basis of two types of information—the traffic volume and lane-blockage index. This model takes into consideration irregular roadway conditions. With the neural network, the complex procedures of formulating traffic weaving movement, car following movement, and tourist lane-blockage detection/reaction behavior could be avoided. The average estimation error was 5.5 sec, whereas the simulation result of travel time at free flow is 27.3 sec for going through the roadway segment (on the basis of results obtained from the Highway Capacity Manual).

Overall, the application potential of ANNs to IVHS is significant. Suitability to IVHS is found in the ANN properties of pattern recognition, classification, data compression, prediction, and control logic modeling. Although superiority to other technologies is not identified in this study, it is possible to expect that ANNs are feasible alternatives to conventional techniques in numerous IVHS application domains, and their use is expected to be extended further in the future.

ACKNOWLEDGMENT

Partial funding for this project was provided by the Delaware Department of Transportation through the Delaware Transportation Center, Department of Civil Engineering, University of Delaware.

REFERENCES

1. Lubin, J. M., E. C. Huber, S. A. Gilbert, and A. L. Kornhauser. Analysis of a Neural Network Lateral Controller for an Automated Road Vehicle. *IVHS—Issues and Technology,* SAE, Warrendale, Pa. 1992, pp. 1–66.
2. Takaba, S., T. Morita, T. Hada, T. Usami, and M. Yamaguchi. Estimation and Measurement of Travel Time By Vehicle Detectors and License Plate Readers. *Society of Automotive Engineers.* Number P-253, Part 1. SAE, Warrendale, Pa. 1991, pp. 257–267.
3. French, R. L. Intelligent Vehicle/Highway System in Action. *ITE Journal,* Vol. 60, No. 11, Nov. 1990, pp. 23–31.
4. Hartje, R. L. Tomorrow's Toll Road. *Civil Engineering,* Feb. 1991, pp. 60–61.
5. Faghri, A., and J. Hua. Evaluation of Artificial Neural Network Applications in Transportation Engineering. In *Transportation Research Record 1358,* TRB, National Research Council, Washington, D.C., 1992, pp. 71–80.
6. Nakatsuji, T., and T. Kaku. Application of Neural Network Models to Traffic Engineering Problems. *Proc., 11th International Symposium on Transportation and Traffic Theory,* Yokohama, Japan, 1990, pp. 291–306.
7. Hertz, J., A. Krogh, and G. R. Palmer. *Introduction to the Theory of Neural Computing.* Addison-Wesley, Redwood City, Calif. 1991.
8. Hua, J., and A. Faghri. Dynamic Traffic Pattern Classification. In *Transportation Research Record 1399,* TRB, National Research Council, Washington, D.C., 1993, pp. 14–19.

9. Nakatsuji, T., and T. Kaku. Development of Self-Organizing Traffic Control System Using Neural Network Models. In *Transportation Research Record 1324,* TRB, National Research Council, Washington, D.C., 1991, pp. 137–145.
10. Lubin, L. M., E. C. Huber, S. A. Gilbert, and A. L. Kornhauser. Analysis of a Neural Network Controller for an Automated Road Vehicle. *IVHS—Issues and Technology,* SAE Special Publications 928. SAE, Warrendale, Pa., 1992, pp. 23–44.
11. Aono, S. Next Step in Automotive Electronic Control. *International Conference on Transportation Electron ProConvergence 88,* IEEE, New York, 1988, pp. 83–89.
12. Haykin, S., W. Stehwien, C. Deng, P. Weber, and R. Mann. Classification of Radar Clutter in An Air Traffic Control Environment. *Proc., IEEE,* Vol. 79, No. 6, June 1991, pp. 742–772.
13. Simpson, P. K. *Artificial Neural Systems.* Pergamon Press Inc., Elmsford, N.Y., 1990.
14. Shea, P., B. Yedidia, and F. Liu. Application of an Artificial Neural Network in the SAICZ Explosive Detection System. *Proc., International Joint Conference on Neural Networks,* IEEE, 1991, pp. 910.
15. Widrow, B., J. Glover, J. Caunitz, C. Williams, and R. Hearn. Adaptive Noise Canceling: Principles and Applications. *Proc., 63rd Conference of IEEE,* 1975, pp. 1692–1716.
16. Widrow, B., P. Mantey, and L. Griffiths. Adaptive Antenna Systems. *Proc., 55th Conference of IEEE,* 1967, pp. 2143–2159.
17. Shimohara, K., Y. Tokunoga, T. Uchiyama, and Y. Kimura. A Neural Network System with an Automatic Generation Mechanism for Distorted Patterns. *Neural Networks Supplement: INNS Abstracts,* 1, 1988. pp 52.
18. Tusgawa, S., K. Kitoh, T. Harada, and K. Kunio. Info-Mobility, A Concept for Advanced Automotive Functions Toward the 21st Century. *IVHS and Vehicle Communications,* SAE Special Publication 877. SAE, Warrendale, Pa., 1991, pp. 67–75.
19. McMahon, D. H., V. K. Narendran, D. Swaroop, J. K. Hedrick, K. S. Chang, and P. E. Devlin. Longitudinal Vehicle Controllers for IVHS: Theory and Experiment. *Proc., American Control Conference,* Vol. 2, American Automatic Control Council, Green Valley, Ariz., 1992, pp. 1753–1757.
20. Wee, W. Generalized Inverse Approach to Adaptive Multiclass Pattern Classification, *IEEE, Transaction on Computers,* C-17, 1968, pp. 1157–1164.
21. Grossberg, S. Adaptive Pattern Classification and Universal Recording: I. Parallel Development and Coding of Neural Detectors. *Biological Cybernetics,* Vol. 23, 1976, pp. 121–134.
22. Hecht-Nielsen, R. Applications of Counterpropagation Networks. *Neural Networks,* Vol. 1, 1988, pp. 131–140.
23. Namphol, A., M. Arozullah, and S. Chin. Higher Order Data Compression with Neural Networks. *International Joint Conference on Neural Networks,* Part 1, IEEE, 1991, pp. 55–59.
24. Lippman, R. Neural Network Classifiers for Speech Recognition. *The Lincoln Laboratory Journal,* Vol. 1, 1988, pp. 107–124.
25. Robinson, A., and F. Fallside. Static and Dynamic Error Propagation with Application to Speech Coding. *Proc., 187th Conference on Neural Information Processing Systems—Neural and Synthetic.* American Institute of Physics, 1988, pp. 632–641.
26. Kohonen, T. The "Neural" Phonetic Typewriter. *IEEE Computer Magazine,* March 1988, pp. 11–22.

Publication of this paper sponsored by Committee on Artificial Intelligence.

Urban Rail Corridor Control Through Machine Learning: An Intelligent Vehicle-Highway System Approach

SNEHAMAY KHASNABIS, TOMASZ ARCISZEWSKI, SYED KHURSHIDUL HODA, AND WOJCIECH ZIARKO

Traffic control along an urban rail corridor with closely spaced stations can be considered a sequence of decision-making stages. A train on an urban rail corridor that connects two terminal points with a number of intermediate stations can follow various regimes of moving and stopping, which identify individual driving scenarios. The execution of these regimes may result in different values of attributes that describe driving scenarios, namely, travel time, energy consumption, passenger comfort, and others. An attempt is made to demonstrate how to develop decision rules for driving scenarios along an urban rail corridor that can optimize travel time, energy consumption, and passenger comfort, using the concept of machine learning. Machine learning is a science that deals with the development and implementation of computational models of learning processes. The concept of knowledge acquisition through inductive learning as an intelligent vehicle-highway system approach is explored to establish some initial decision rules. A computer model, REGIME, was developed for the estimation of values of evaluation criteria, such as travel time, energy consumption, and passenger comfort levels for a hypothetical rail corridor for various driving scenarios. Next, a commercial learning system, ROUGH, was used in conjunction with the examples created through REGIME to develop decision rules. The learning algorithm is based on the theory of rough sets. The feasibility of machine learning in automated knowledge acquisition to develop decision rules for complex engineering problems, such as urban rail corridor control, is demonstrated. Further research is needed to verify the rules developed before these can be applied.

Machine learning so far has had only limited applications to knowledge acquisition in civil engineering (*1*). However, there are some examples of the application of machine learning in the areas of conceptual design (*2,3*). A feasibility study of automated acquisition of knowledge of traffic control along an urban rail corridor was conducted at Wayne State University as a part of a program on intelligent vehicle-highway systems (IVHS). In this paper, the authors explore the concept of machine learning to develop decision rules for optimum control along an urban rail corridor. An earlier version of this paper was published previously (*4*).

PROBLEM STATEMENT

A train on an urban rail corridor connecting two terminal points with a number of intermediate stations can follow various regimes of moving and stopping, which identify individual driving scenarios. The execution of these regimes may result in different values of attributes that describe these scenarios, namely, travel time, energy consumption, passenger comfort, and others.

The concept of preprogrammed driving for urban rail corridors, as proposed in the literature will permit an automated selection of driving scenarios consistent with the distribution of ridership demand along the corridor (*5*). The driving scenarios are likely to change with the time of day as the demand changes. This concept is consistent with IVHS technology that aims at the integration of the vehicle, the facility, and the driver using the state-of-the-art communication, computer, and electronic technology (*6*).

When a large number of train driving scenarios is considered, the evaluation of individual scenarios and selection of the optimal one becomes difficult because of the complexities involved in analyzing these scenarios. An alternative approach is a knowledge-based selection. In this case, the optimal scenario may be selected using a knowledge-based decision support tool. This paper examines such an approach.

OBJECTIVES

The paper is based on a pilot study to explore the concept of knowledge acquisition through inductive learning to establish decision rules for an urban rail corridor. Ideally, one would like to

- Minimize travel time,
- Minimize energy consumption,
- Maximize access, and
- Minimize level of discomfort.

In actuality, it is not possible to minimize travel time and energy consumption at the same time because these two are nearly inversely related entities, as evidenced from empirical studies (*5*). The question of maximizing access has never been satisfactorily resolved in the literature. Passenger discomfort levels are associated with acceleration and deceleration characteristics of the train, and these are difficult entities to quantify. The objective of this paper is to demonstrate how to develop decision rules for driving scenarios along an urban rail corridor that can improve travel time, energy consumption, and comfort levels of passengers.

RAIL TRAFFIC CONTROL

Traffic control along an urban rail corridor with closely spaced stations can be considered a sequence of decision-making stages. In

S. Khasnabis, Department of Civil Engineering, Wayne State University, Detroit, Mich. 48202. T. Arciszewski, Systems Engineering Department, School of Information Technology and Engineering, George Mason University, Fairfax, Va. 22030. S.K. Hoda, H. Carl Walker Engineers, 445 W. Michigan Avenue, Suite 101, Kalamazoo, Mich. 49007. W. Ziarko, Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada.

this study, three evaluation criteria are used: travel time that is based on regimes of motion, energy consumption, and comfort levels.

Regimes of Motion

Typically, the train operator has the option of selecting regimes of motion for individual segments (Figure 1) from four basic regimes, A through D, as discussed (*5*):

- *Regime A:* The interstation spacing is shorter than the critical spacing; critical spacing is the minimum distance between stations needed for the train to attain its maximum speed (Figure 1a).
- *Regime B:* The interstation spacing is longer than the critical spacing. The train maintains a sustained level of maximum speed before deceleration is initiated for the next stop (Figure 1b).
- *Regime C:* The interstation spacing is longer than the critical spacing. However, as an energy conservation measure, the train starts coasting (decelerating at a very slow rate) immediately on reaching its maximum speed and continues to coast until deceleration is initiated as the train approaches the next station (Figure 1c).
- *Regime D:* Regime D represents an intermediate condition between Regimes B and C that allows the train to travel at its maximum speed and to coast between two stops. The train accelerates to its highest speed, travels at the maximum speed for a predetermined period (as described in Regime B), and starts coasting (as described in Regime C) before the deceleration process is started. Within Regime D, infinite combinations are possible, depending on the instant when coasting is initiated. If coasting begins immediately before deceleration, the resulting is Regime B as a limiting condition. If, on the other hand, coasting is initiated immediately upon the attainment of maximum speed, Regime C will result as the other limit (Figure 1d).

Energy Consumption

Studies of Hamburg rail systems by empirical and computer simulation techniques have demonstrated the importance of different driving regimes in the context of energy consumption (*7*). The tradeoff between energy consumption and travel time, developed from time-speed-energy consumption data, was used in this study to develop surrogate measures of energy consumption for varying travel times in the form of an empirical relationship (Figure 2). Although this relationship does not explicitly consider different regimes of motion, lower energy consumption resulting from longer coasting and consequent longer travel times are incorporated in this relationship (*8*).

Four models were developed for estimating energy consumption for the purpose of this study using the data presented in Figure 2. These models included simple, polynomial, logarithmic, and exponential models. The following exponential model was used for the study reported here.

$$Y = 1322.5 * 10^{(-1.0097e-2X} \cdots (A) \qquad R^2 = 0.983$$

where

X = travel time surrogate and
Y = energy consumption surrogate.

Passenger Comfort Levels

Every change in the acceleration/deceleration phase is associated with a level of discomfort for the passenger. The rate of acceleration/deceleration (second derivative of speed with respect to time) is commonly termed "jerk." It was assumed that the level

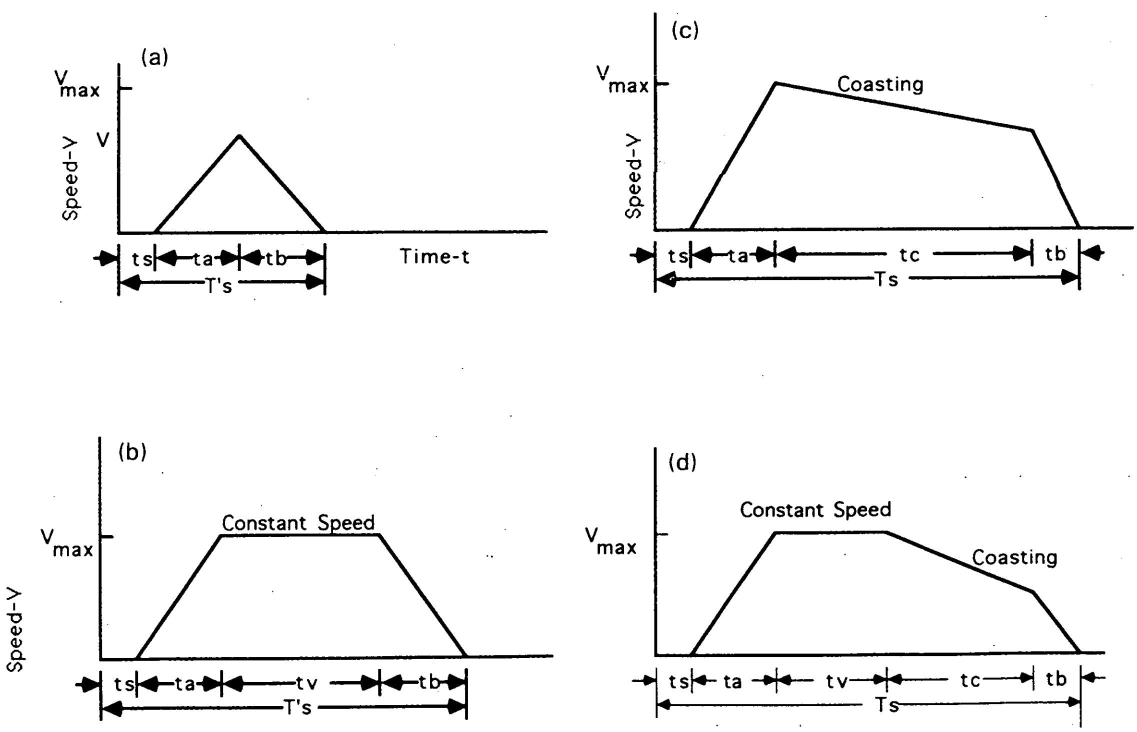


FIGURE 1 Interstation travel regimes (*5*). (*a*) Regime A, (*b*) Regime B, (*c*) Regime C, and (*d*) Regime D.

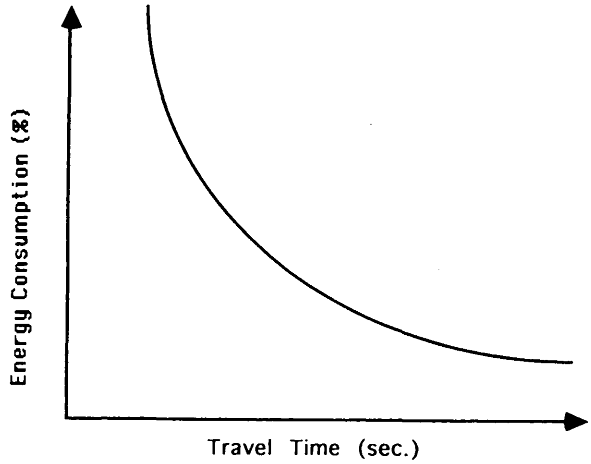


FIGURE 2 Travel time verses energy consumption trade-off (5).

of discomfort experienced by a passenger is measured by the number of jerks during a given pass of the train along the entire corridor. Ideally, not only the frequency of jerks but also their respective magnitudes should be considered. However, magnitude was considered too complex to quantify for the purpose of this study.

A review of acceleration/deceleration characteristics indicates that for a typical interstation travel, a total of two instances of jerks will be experienced during the acceleration phase, two during the deceleration phase, and one during the beginning of the coasting operation. Further, for every skip-stop operation, a total of four instances of jerk can be "saved," resulting from the elimination of deceleration and acceleration operation as the train approaches and leaves the station in question, respectively.

METHODOLOGY

The primary objective of this study was to apply a learning system as an automated knowledge acquisition tool for an urban rail corridor. The following methodology was used.

Machine Learning

Machine learning is the process of generating decision rules representing logical relationships between various combinations of attributes and their values. It is a science that deals with studies and development and implementation of computational models of learning and discovery processes. Learning systems are computer programs that transform input in the form of data (usually examples) into knowledge (usually in the form of decision rules).

A decision rule is a logical relationship between a group of attributes called "independent attributes" and a single attribute called a "dependent attribute." In this case, independent attributes are those that are controlled by the rail traffic operator, whereas dependent attributes can be only indirectly controlled. Therefore, independent attributes affect the values of dependent attributes. For example, an independent attribute "skipping one or two stops" affects the value of the dependent attribute "travel time."

ROUGH as a Machine Learning Tool

Automated knowledge acquisition was conducted using ROUGH, Version 1.1, a commercial inductive learning system developed by Ziarko (*9*). This algorithm utilizes the theory of rough sets proposed by Pawlak (*10*). Its objective is to produce decision rules for the classification of examples into one of the categories of the dependent attribute.

In the theory of rough sets, the determination of decision rules is based on the analysis of individual attributes in the context of a given collection of examples. This analysis includes the determination of the dependency relationship between the dependent attribute and any group of independent attributes, identification of a minimal set of independent attributes that are necessary and sufficient to produce decision rules, and the determination of the relative importance of individual attributes from this group.

ROUGH conducts the learning process in two stages. First, it performs the analysis of dependency factors for individual attributes and identifies a set of "reducts" for a given collection of examples. The term reduct signifies a minimum and sufficient collection of attributes describing a given system. This stage can be considered an analysis and modification of the representation space for given examples. In the second stage, actual learning occurs, and the system uses individual reduct attributes to produce decision rules using these attributes.

Study Approach

A computer program, REGIME, was developed for the estimation of values of evaluation criteria, which included travel time and surrogates of energy consumption and comfort level for a given train driving scenario along an assumed corridor. The necessary algorithms for computing travel times were obtained from Vuchic (*5*). Equation A, derived by the authors of this paper from Hamburg rail data, was used for estimating energy surrogates. The number of jerks experienced during a complete journey for each scenario was computed from first principles.

Although the computation of travel time for Regimes B and C is relatively straightforward, for Regime D it is somewhat complex, because of an infinite number of possibilities when coasting may be initiated. The model provides the user with a range of possible values of the coasting speed for different skip-stop combinations to help the user converge to a definite solution. The model produces the output at the individual station level (microscopic) as well as the corridor level (macroscopic).

All the driving scenarios were then identified by the condition attributes in binary form by assigning yes or no values (Table 1). After identifying the scenarios, the decision attributes, that is, total travel time, energy consumption factor, and passenger comfort, were determined. The decision attributes were identified by high, medium, and low desirability. The condition attributes and decision attributes were then entered in the ROUGH system as inputs to develop decision rules.

Study Area

Evaluation of individual scenarios was produced for a hypothetical urban rail corridor of 50 station spaces (sections) divided into five segments. The two end segments, Segments 1 and 5, consist of four

TABLE 1 Binary Representations of Condition Attributes

Segment	Condition Attributes	Binary Decision	
No. 1 and 5	Constant Speed	Yes	No
	Coasting	Yes	No
	Constant speed and Coasting	Yes	No
	One stop skipped	Yes	No
	Two stops skipped	Yes	No
No. 2 and 4	Constant speed	Yes	No
	Coasting	Yes	No
	Constant speed and Coasting	Yes	No
	One stop skipped	Yes	No
	Two stops skipped	Yes	No
No.3	Constant speed	Yes	No
	Coasting	Yes	No
	Constant speed and Coasting	Yes	No
	One stop skipped	Yes	No
	Two stops skipped	Yes	No

station spaces. Segments 2 and 4 are the two intermediate segments, each consisting of 12 station spaces. The central segment, 3, contains 18 station spaces. Each spacing was assumed to be 2,000 ft for a total corridor length of 100,000 ft. The rail corridor analyzed is thus a symmetrical one, with Segments 1 and 2 being mirror images of Segments 5 and 4, respectively, and Segment 3 being the central portion.

It was assumed that the decisions taken for Segments 1 and 5 would be identical. Similarly, it was assumed that Segments 2 and 4 are described by identical decisions. Therefore, the entire rail corridor is described when decisions for three different segments are known (for Segments 1 through 3). For each of these segments, five binary decisions about train operations are to be made. Thus, the entire problem of train control is represented as a sequence of 15 decision-making stages. At these stages, decision making requires answering binary (yes/no) questions (Table 1).

When all these questions are answered for the three different segments identified, a train-driving scenario is produced in the form of a sequence of 15 answers. In this model, the total number of train-driving scenarios is large, although it is significantly smaller than 2^{15} because some combinations of answers are infeasible.

RESULTS

The results are presented in two sections. First, the output of the software REGIME is presented in both microscopic (station) and macroscopic (corridor) levels. Next, the results of using the macro level output as an input to ROUGH to develop decision rules are presented. Finally, a brief discussion of the rationale of some of the decision rules thus developed is provided.

REGIME Output

The basic input for REGIME is maximum speed, demand, headway, acceleration, deceleration, coasting deceleration, interstation distance, and the total length of the corridor. The input values for the study are shown as follows:

Input	*Value*
Maximum speed	60 mph
Demand	45,000 passengers per hour
Headway	2.00 min
Acceleration	5 ft/sec^2
Deceleration	6 ft/sec^2
Coasting deceleration	1 ft/sec^2
Station waiting time	35 sec
Interstation distance	2,000 ft
Total distance	100,000 ft

Table 2 shows the micro level output of REGIME. Table 2 indicates that when Regime B is used, operating speed improves from 27.08 ft/sec obtained for no-skip operation to 41.41, 50.29, and 56.32 ft/sec for one-stop-skip, two-stop-skip, and three-stop-skip operations, respectively. At Regime C, skipping more than one stop results in a dysfunctional operation as the gradual drop in speed results in 0 speed. This is a consequence of coasting over an extended distance, caused by skipping more than one stop. The speed level attained at Regime D is between corresponding speeds at Regimes B and C for respective skip-stop operations.

Table 3 shows the macro level output of REGIME. The following is an interpretation of the first row in Table 3, a scenario in which the

TABLE 2 Output of REGIME Model (Micro Level)

		Skip Stop Scenario			
Regime	Operating Data	0 Stop	1 Stop	2 Stops	3 Stops
B	Operating Speed (ft/s)	27.08	41.41	50.29	56.32
	Travel Time (sec)	73.86	96.59	119.32	142.04
C	Operating Speed (ft/s)	26.96	37.11	0.00	0.00
	Travel Time (sec)	74.19	107.78	0.00	0.00
	Coasting Speed (ft/s)	79.70	39.39	0.00	0.00
D	Operating Speed (ft/s)	26.97	41.34	50.26	56.31
	Travel Time (sec)	74.16	96.76	119.39	142.06
	Coasting Speed(ft/s)	80.00	82.00	84.00	86.00

TABLE 3 Output of REGIME Model (Macro Level)

Scenario	No. of Jerks	Travel Time (seconds)	Energy Consumption Factor
B1-B2-B2	72	2056.76	129.33
B0-B1-B2	104	2465.83	81.45
D0-D2-D2	110	2298.79	98.37
C0-D2-D2	110	2298.98	98.35
C0-D1-B2	124	2550.49	74.01
C0-B2-D1	117	2478.84	80.26
D0-B2-B2	96	2263.72	102.35
D1-B2-C1	97	2338.22	94.09
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.

train would travel in: (*a*) Regime B, skipping one stop in Segments 1 and 5; (*b*) Regime B and skipping two stops in Segments 2 and 4; and (*c*) Regime B and skipping two stops in Segment 3 will result in 72 jerks, 2,056.76 sec of travel time, and 129.33 surrogate units of energy consumption for the entire corridor consisting of 100,000 ft.

ROUGH Output

REGIME was used to analyze 102 train driving scenarios, and for each scenario, estimates of the values of three evaluation criteria were produced. Next, these cases were used to prepare examples for inductive learning. The preparation of examples required the transformation of the evaluation criteria from interval into nominal attributes high, medium, and low. For instance, it was assumed that short travel time, low energy consumption, and high comfort (small number of jerks) all have high desirability, whereas long travel time, high energy consumption, and low comfort (large number of jerks) all have low desirability.

Automated knowledge acquisition was conducted using ROUGH to learn about driving scenarios for an urban rail corridor. Scenarios were defined by 15 condition attributes (Table 4). Three decision attributes or logical extensions of the three measures of effectiveness are as follows:

1. D1, Desirability for passenger comfort,
2. D2, Desirability for travel time, and
3. D3, Desirability for energy consumption.

Four machine-learning processes were performed using the ROUGH system. In the first process, comfort was considered as a decision attribute, and a total of 46 decision rules were developed (Table 5). Next, travel time was considered as a decision attribute, and a total of 41 decision rules were developed (Table 6). In the third process, energy consumption factor was considered as a decision attribute, resulting in a total of 52 decision rules (Table 7). In the fourth and final learning process, all decision attributes were considered together. A total of 71 decision rules were obtained (Table 8).

It is beyond the scope of this paper to offer exact interpretation of the above rules and, more importantly, to review the rationale behind these rules. Just to provide an example on how to interpret these rules, referring to Table 8, rule 69, the following explanation is offered.

To achieve high desirability in comfort (small number of jerks), medium desirability in travel time (medium travel time), and high desirability in energy consumption (low consumption levels) a driving scenario encompassing the following conditions should be fulfilled:

1. Coasting in Segments 1 and 5;
2. No constant speed and coasting in Segments 1 and 5;
3. Skipping one stop in Segments 1 and 5;
4. Skipping one stop in Segments 2 and 4;
5. Skipping one stop in Segment 3; and
6. No skipping two stops in Segment 3.

A review of these rules indicates that Condition 2 is automatically preempted by Condition 1 and thus is redundant. The essence of the remaining conditions is that the combination of coasting and skipping one stop in each segment will result in high desirability in comfort levels and in energy consumption levels. Further, skipping more than one stop may result in inordinately long travel times (particularly if Regime B is involved). As such, skipping more than one stop is discouraged. These rules thus appear logical and reasonable.

CONCLUSIONS

The objective of this paper is to explore the concept of knowledge acquisition through inductive learning to establish decision rules for an urban rail corridor. The study demonstrates the feasibility of using machine learning in automated knowledge acquisition about complex engineering problems such as urban rail traffic control.

TABLE 4 Description of Condition Attributes

Condition Attributes	Notation
Constant Speed in the 1st and 5th Segments	S11
Coasting in the 1st and 5th segments	S12
Constant speed and Coasting in the 1st and 5th segments	S13
Skipping One stop in the 1st and 5th segments	S14
Skipping Two stops in the 1st and 5th segments	S15
Constant speed in the 2nd and 4th segments	S21
Coasting in the 2nd and 4th segments	S22
Constant Speed and Coasting in the 2nd and 4th segments	S23
Skipping One stop in the 2nd and 4th segments	S24
Skipping Two stops in the 2nd and 4th segments	S25
Constant speed in the 3rd segment	S31
Coasting in the 3rd segment	S32
Constant speed and Coasting in the 3rd segment	S33
Skipping One stop in the 3rd segment	S34
Skipping Two stops in the 3rd segment	S35

TABLE 5 Rules by Decision Attribute D1 (Passenger Comfort) as Output of ROUGH

Rule	Condition Attribute															Decision Attribute
No.	S11	S12	S13	S14	S15	S21	S22	S23	S24	S25	S31	S32	S33	S34	S35	D1
1		N		N			N		N	N		N		N	N	L
2		N		N			N	N	N	N		N		Y	N	M
3		N		N			N	Y	N	N		N		Y	N	L
4		N		N			N		N	N		N			Y	M
5		N		N			N		N	Y		N		N		M
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
43		Y		Y					N	N				Y	N	M
44		Y		Y					Y	N				Y	N	H
45		Y		Y						Y				Y	N	H
46		Y		Y											Y	H

TABLE 6 Rules by Decision Attribute D2 (Travel Times) as Output of ROUGH

Rule	Condition Attribute															Decision Attribute
No.	S11	S12	S13	S14	S15	S21	S22	S23	S24	S25	S31	S32	S33	S34	S35	D2
1		N		N			N		N	N		N			N	L
2		N		N			N		N	N		N			Y	M
3		N		N			N		N	Y		N		N	N	M
4		N		N			N		N	Y		N		N	Y	H
5		N		N			N		N	Y		N		Y		H
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
38		Y		Y					N	N				Y	N	L
39		Y		Y					Y	N				Y	N	M
40		Y		Y						Y				Y	N	H
41		Y		Y											Y	H

TABLE 7 Rules by Decision Attribute D3 (Energy Consumption) as Output of ROUGH

Rule	Condition Attribute															Decision Attribute
No.	S11	S12	S13	S14	S15	S21	S22	S23	S24	S25	S31	S32	S33	S34	S35	D3
1		N	N	N					N	N		N				H
2		N	N	N				N	N	Y		N				L
3		N	N	N				Y	N	Y		N		N		H
4		N	N	N				Y	N	Y		N		Y		M
5		N	N	N					Y			N	N			M
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
49		Y	N	Y				Y				N	Y	Y	N	M
50		Y	N	Y				Y				Y		Y	N	M
51		Y	N	Y				Y							Y	L
52		Y	Y													L

TABLE 8 Rules by Decision Attributes D1, D2, and D3 as Output of ROUGH

Rule No.	Condition Attribute															Decision Attributes		
	S11	S12	S13	S14	S15	S21	S22	S23	S24	S25	S31	S32	S33	S34	S35	D1	D2	D3
1		N	N	N			N	N	N	N				N	N	L	L	H
2		N	N	N			N	N	N	N				N	Y	M	M	H
3		N	N	N			N	N	N	N				Y		M	L	H
4		N	N	N			N	N	Y	N				N		H	H	M
5		N	N	N			N	N	Y	N			N	Y		H	M	M
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.			.	.	.	.			.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
68		Y	N	Y					N	Y		Y		Y	N	H	H	M
69		Y	N	Y					Y					Y	N	H	M	H
70		Y	N	Y											Y	H	H	L
71		Y	Y													H	H	L

The rules developed are based on three separate evaluation criteria: passenger comfort, travel time, and energy consumption. Additionally, a set of rules was developed with all of the three attributes combined. No effort was made in this study to explain these rules, to validate them, or to assess how they can be applied in actual train control. The large number of decision rules and their interaction reflects the complexity of the rail corridor control. To gain further insights into this problem, an automated rule verification method is recommended on the basis of the performance of the learning system, measured by various empirical error rates.

Machine learning in rail traffic control is a new complex and interdisciplinary research, and more work is needed to determine the feasibility of machine learning in rail corridor control, to develop better understanding of the problem, and to prepare a program that would lead from research to practical application of results. The technique of machine learning appears to complement the emerging IVHS area.

ACKNOWLEDGMENT

The study was conducted as a part of the IVHS program at the Department of Civil Engineering, Wayne State University. This study was funded by the U.S. Department of Transportation, through the IVHS program at the University of Michigan, Ann Arbor. The authors are thankful to the sponsoring agencies for their support. The authors would also like to acknowledge the cooperation of REDUCT Systems, Inc., of Regina, Saskatchewan, Canada, which provided the inductive system ROUGH.

REFERENCES

1. Arciszewski, T., and W. Ziarko. Machine Learning in Knowledge Acquisition. In *Knowledge Acquisition in Civil Engineering* (T. Arciszewski and L. Rossman, eds.), ASCE, New York, 1992.
2. Reich, Y., and S. J. Fenves. Automated Acquisition of Design Knowledge by Concept Formation. In *Knowledge Acquisition in Civil Engineering,* (T. Arciszewski and L. Rossman, eds.), ASCE, New York, 1992.
3. Maher, M. L., et al. Automatically Learning Preliminary Design Knowledge from Design Examples. *Journal of Microcomputers in Civil Engineering,* Vol. 7, Special Issue: *4th National Conference on Computing in Civil and Building Engineering,* 1992.
4. Khasnabis, S., T. Arciszewski, S. K. Hoda, W. Ziarko. Automated Knowledge Acquisition for Control of an Urban Rail Corridor. *Proc. 3rd International Conference on the Application of Artificial Intelligence to Civil and Structural Engineering (Civil-Comp 93),* Edinburgh, Scotland, Civil Comp Press, Edinburgh, Aug. 1993.
5. Vuchic, V. R., Urban Transportation. In *System and Technology,* Prentice-Hall, Inc., 1980.
6. Strategic Plan for Intelligent Vehicle Highway Systems in the United States. IVHS America, May 1992.
7. Mies, A. Underground Railway Automation of the Hamburger Hochbahn AG. In *UITP Revue,* No. 1 1974, pp. 45–48, UITP, Brussels.
8. Mies, A. Fahrdynamische Grunduberlegungen fur die Automatisierung von Schnellbahnen. *Verkehr und Technik.* Erich Schmidt Verlang, Bielefeld, West Germany, Aug. 1969, pp. 224–227.
9. Ziarko, W. Data Analysis and Case-Based Expert System Development and Tool. ROUGH, *Proc., of Case-Based Reasoning Workshop,* Pensacola, Fla. 1989
10. Pawlak, Z. Rough Sets, *International Journal of Computer and Information Science,* Vol. 5, No. 11, pp. 341–356, 1982.

Publication of this paper sponsored by Committee on Artificial Intelligence.

Short-Term Traffic Flow Prediction: Neural Network Approach

Brian L. Smith and Michael J. Demetsky

Much of the current activity in the area of intelligent vehicle-highway systems (IVHS) focuses on one simple objective: to collect more data. Clearly, improvements in sensor technology and communication systems will allow transportation agencies to more closely monitor the condition of the surface transportation system. However, monitoring alone cannot improve the safety or efficiency of the system. It is imperative that surveillance data be used to manage the system in a proactive rather than a reactive manner. Proactive traffic management will require the ability to predict traffic conditions. Previous predictive modeling approaches can be grouped into three categories: (*a*) historical, data-based algorithms; (*b*) time-series models; and (*c*) simulations. A relatively new mathematical model, the neural network, offers an attractive alternative because neural networks can model undefined, complex nonlinear surfaces. In a comparison of a backpropagation neural network model with the more traditional approaches of an historical, data-based algorithm and a time-series model, the backpropagation model was clearly superior, although all three models did an adequate job of predicting future traffic volumes. The backpropagation model was more responsive to dynamic conditions than the historical, data-based algorithm, and it did not experience the lag and overprediction characteristics of the time-series model. Given these advantages and the backpropagation model's ability to run in a parallel computing environment, it appears that such neural network prediction models hold considerable potential for use in real-time IVHS applications.

An emerging group of technologies and systems known as intelligent vehicle-highway systems (IVHS) have the potential to serve as powerful tools in combating transportation safety and congestion problems by improving the manner in which the nation's extensive existing surface transportation system operates. The backbone of IVHS is the "smart highway"—advanced traffic management systems (ATMS). ATMS collect, utilize, and disseminate real-time data on the status of the surface transportation system. ATMS rely on extensive traffic surveillance systems, thereby providing all other IVHS components with accurate, real-time information. Furthermore, ATMS provide traffic control in both time and space through techniques such as the optimization of traffic signal timing and ramp metering. These techniques have proven benefits, such as freer traffic flows, shorter journey times, and fuel savings (*1*).

The challenge of effectively using real-time data extends to the area of advanced traveler information systems (ATIS), the basic premise of which is to provide travelers with accurate and timely information to allow them to make sound decisions. Clearly, there is a well-defined link between ATMS and ATIS in that both rely on accurate real-time data that describe the status of the transportation network. In addition, it is possible that ATIS will serve as an additional control measure for ATMS by encouraging individual route selection, which would spread demand across all available capacity.

B. L. Smith, Virginia Transportation Research Council, Box 3817 University Station, Charlottesville, Va. 22901. M.J. Demetsky, Department of Civil Engineering, University of Virginia, Charlottesville, Va. 22901.

Although many of the physical components of ATMS and ATIS are still some years away from wide-scale deployment, the preliminary development of software support systems is feasible and should receive immediate attention. At this time, most research has focused on specific applications, such as incident detection and ramp-metering algorithms; very little consideration has been given to developing more general support systems, such as real-time, short-term prediction of traffic conditions. The development of such software support systems will enhance the performance of current systems and serve as a critical step in developing ATMS and ATIS.

REAL-TIME INFORMATION

Real-time data primarily will consist of vehicle counts, vehicle locations, and vehicle speeds. Clearly, vehicle counts alone cannot help a traveler make a routing decision or a traffic manager set a series of signal timings. It is critical that the raw data be processed to derive true information that will support intelligent decision making.

A particularly important function in transforming raw data into information is the prediction of traffic conditions. The current focus on real-time applications is likely to result in reactive control of the transportation system. There is certain to be some lag between the collection of real-time data and the implementation of a control strategy. Therefore, the system will operate under control strategies that are based on past conditions. To control the system in a proactive manner, ATMS must have some sort of predictive capability: "The ability to make and continuously update predictions of traffic flows and link times for several minutes into the future using real-time data is a major requirement for providing dynamic traffic control" (*2*, p. x). Traffic prediction is also an important function for ATIS: "the rationale behind using predictive information (for route guidance) is that travelers' decisions are affected by future traffic conditions expected to be in effect when they reach downstream sections of the network" (*3*). In fact, traveler information services are hampered by the lack of a capability to predict future traffic conditions. For example, changeable message signs are rarely used to provide travel time information because they are inaccurate. Clearly, the success of IVHS is dependent on the development of a traffic prediction capability. Consequently, "special attention should be given to the ability to make short-term traffic predictions with real-time sensor data" (*2*, p. viii).

PREDICTION OF TRAFFIC CONDITIONS: PREVIOUS EFFORTS

"The short-term forecasting of traffic conditions has had an active but somewhat unsatisfying research history" (*1*). Most efforts have focused on traffic prediction for surface signal control systems, such as the Urban Traffic Control System (UTCS). There have been a limited number of freeway traffic prediction applications. The approaches used for traffic prediction are largely dictated by the fact that traffic conditions are time dependent and often follow fairly well-defined patterns. Previous traffic prediction efforts can be classified as historical, data-based algorithms; time-series models; or simulations.

Historical, Data-Based Algorithms

The basic premise behind historical, data-based algorithms is that traffic patterns are cyclical. In other words, a knowledge of typical traffic conditions on Tuesday at 5:30 p.m.will allow one to predict the conditions on any particular Tuesday at 5:30 p.m. AUTOGUIDE, an ATIS demonstration project in London, utilizes the simplest historical, data-based algorithm possible. AUTOGUIDE simply uses a historical traffic data base to predict travel times on the basis of time of day (*4*). Such an algorithm is attractive in that it requires no real-time data.

The UTCS traffic control system utilizes predictions of traffic conditions in an attempt to control signals in a proactive manner. In general, UTCS relies on historical data as support for predictions. A weakness of this method is that UTCS requires an extensive set of historical data; consequently, it is difficult to install the system in a new setting (*6*). An enhancement to the prediction capabilities of the second-generation UTCS (UTCS-2) is that the system uses "current traffic measures to correct for the traffic deviation from the average historical pattern" (*5*, p. 28). Finally, it is interesting to note that the third-generation UTCS (UTCS-3) does not utilize historical data; it predicts conditions on the basis of current traffic measurements only. Although the predictive models of both UTCS-2 and UTCS-3 have serious problems with time lag, UTCS-3 is incapable of performing at a level comparable to UTCS-2 (*5*).

LISB, which is a European traveler information experiment, uses a simple methodology to predict future traffic conditions. LISB uses both historical data and real-time data. A projection ratio of the "historical travel time on a specific link to the current travel time as reported by equipped vehicles" is used to predict travel times on the link for future intervals. A major weakness of this methodology is that it implicitly assumes that the projection ratio will remain constant (*3*, p. 4.)

Time-Series Analysis Techniques

In a traffic management system, detectors are used to measure the system's condition at time t, $x(t)$. These measurements can easily be stored for use in predicting the system's condition at time $t + D$, where D is the prediction interval. As such, the prediction problem boils down to forecasting $x(t + D)$, given $x(t)$, $x(t - D)$, $x(t - 2D)$, and so on. This representation of the prediction problem describes a time series. There have been a number of techniques developed in the field of statistics to model time series. Transportation researchers have applied many of these time series analysis techniques to traffic prediction.

The Box and Jenkins technique is a widely used approach to specifying a variety of time-series models (*7*). It has been shown to yield accurate forecasting results in a number of application areas. The most developed Box and Jenkins technique is the autoregressive integrated moving average (ARIMA) method. ARIMA models require very little computational time for execution, which makes them useful for applications in real-time traffic management. However, ARIMA models have not shown great promise in traffic applications. For example, in attempts to apply ARIMA models to UTCS, it has been found that they "resulted in unsatisfactory goodness of fit and high errors; in certain cases they have not been more accurate than a simple moving average" (*6*, p. 1).

Simulation Models

Simulation models provide predictive capability because they demonstrate how the system is likely to react to varying conditions and control strategies. Given the importance of predictive capabilities in ATMS, it is natural to consider the application of simulation in a real-time environment: "An effective on-line simulation model would enable the ATMS control center to project promptly future traffic patterns considering any previously implemented strategies in a real-time operating environment" (*8*, p. 13). Unfortunately, at this time, the real-time application of traffic simulation is not feasible because existing model/algorithm constructs cannot support real-time applications (*9*). A need exists for new approaches to the simulation of transportation systems.

An exciting development that may support real-time simulation is parallel computing. Parallel computing, or processing, is defined as "an efficient form of information processing which emphasizes the exploitation of concurrent events" (*8*, p. 14). In other words, a parallel computer has multiple processors that work simultaneously (in parallel). Of course, this parallelization allows for tremendous increases in the speed of execution. However, the programming of a parallel computer is extremely challenging because of the need to synchronize certain procedures. A recent research effort attempted to develop an architecture for a parallel traffic simulation application. Although it shows promising results, the effort is still in preliminary stages (*8*). The wide-scale deployment of parallel traffic simulation appears to be far from realization.

Assessment

Although a number of approaches to the prediction problem have been described in this section, the fact remains that very few traffic control systems include any proven forecasting capability. There is, thus, a need to develop efficient and accurate real-time traffic prediction models. To be effective, such a model must be able to recognize patterns, use historical or time-series data or both, and represent complex, nonlinear relationships. The next section will introduce neural networks, which have shown considerable promise in these areas.

NEURAL NETWORKS

Over the past several years, both in research and in practical applications, neural networks have proven to be a very powerful method of mathematical modeling. In particular, neural networks are well suited for pattern recognition, offer efficient execution, and model nonlinear relationships effectively. Clearly, neural networks are well worth exploring as a tool for the short-term prediction of traffic.

Neural networks may be defined as "an information processing technology inspired by studies of the brain and nervous system" (*10*, p. 30). This inspiration obviously led to the use of the word neural. However, neural networks in no way attempt to produce biological clones; rather, they are simply models with a rigorous mathematical basis (*11*). Although neural networks are typically associated with the field of artificial intelligence, they function as a sophisticated form of regression. The use of neural networks has been proven successful in a number of applications for the following reasons (*12*):

1. Neural networks can perform highly nonlinear mappings between input and output spaces;
2. The parallel structure of neural networks lends them to implementation on parallel computers, which offers the potential for extremely fast processing; and
3. The neural networks approach is nonparametric; therefore, one need not make any assumptions about the functional form of the underlying distribution of the data.

These characteristics have attracted the attention of researchers from a number of disciplines to problems such as classification, forecasting, process control, and signal processing (*10*).

Neural Networks Basics

To gain a fuller understanding of the underlying mechanics of neural networks, it is instructive to consider the following definition: "a neural network is a computing system made up of a number of simple, highly interconnected processing elements" (*13*, p. 71). The basic structure of a neural network is illustrated in Figure 1. A description of the elements follows.

- *Processing element:* The processing element is the basic building block of a neural network. Processing elements on the input layer simply pass the input value to the adjoining connection weights. Processing elements on the hidden and output layers sum their inputs and compute an output according to a transfer function.
- *Connection weight:* Connection weight serves to join processing elements within the neural network. The connections are of varying strength, which *weight* the value that the connection "transports" between processing elements. In effect, the connection weights may be compared with coefficients in a regression model.
- *Layers*: Layers are sets of processing elements in which all processing elements in adjacent sets are connected. A neural network generally has an input layer, a hidden layer (in which all connection weights are internal to the network), and an output layer.
- *Bias:* The bias is a constant input to each processing element. The input is defined solely by the connection weight between the bias input (which outputs a constant value of 1.0) and the processing element.
- *Transfer function:* The transfer function is an operator, usually nonlinear, that is applied to the summed inputs of a processing element to produce the output value.

In a basic feed-forward neural network, raw input data are presented to processing elements in the input layer. The input values are then weighted and passed to the hidden layer through the connections. Processing elements in the hidden layer sum and process their inputs and then pass the output to the output layer. Processing elements in the output layer sum and process their weighted inputs to produce the network output. The following equation represents this process in a functional form:

$$Y = \Phi[W_2\Phi(W_1 X + \Theta_1) + \Theta_2]$$

where

Φ = transfer function,
W_1 = array of connection weights for layer *I*,
X = input values, and
Θ_1 = array of bias values for layer *I*.

The description presents a neural network as a graphical mathematical modeling technique. In a neural network, the fundamental variables are the set of connection weights. The definition of the connection weights, much like the definition of coefficients in a re-

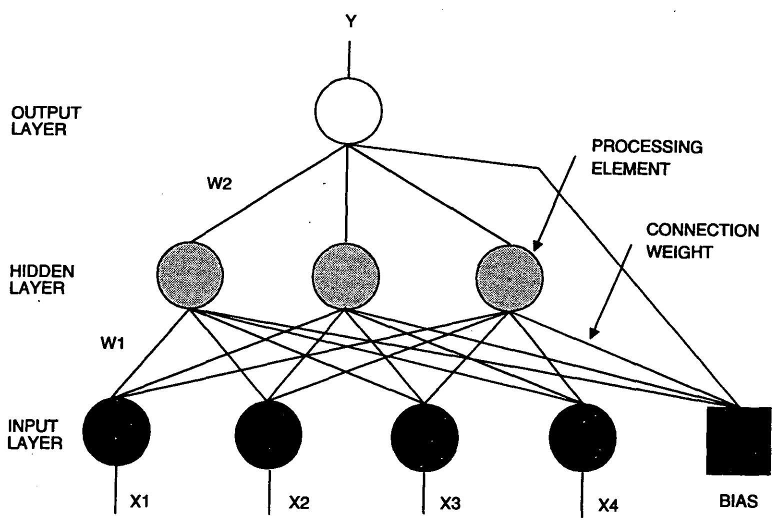


FIGURE 1 Neural network.

gression model, allows the model to be "fitted" to the data. In regression, the method of least squares is utilized to define coefficient values. In the field of neural networks, the process of defining connection weights is generally referred to as "learning." The learning method utilized by a network defines the neural network paradigm. The backpropagation paradigm was chosen for this application. The strengths of this paradigm have led to the conclusion that "almost universally backpropagation has become the standard network paradigm for modeling, forecasting, and classification" (*10,* p. 29). A complete description of this paradigm, can be found in a previous publication (*14*).

APPLICATION OF NEURAL NETWORKS TO TRAFFIC FLOW PREDICTION

The characteristics of neural networks make them excellent candidates for application to the traffic flow prediction problem. In this section, recent studies examining neural network traffic flow prediction models are described.

Gilmore et al. (*15*) applied a backpropagation neural network to predict congestion on surface streets. On the basis of current and past volumes on the surface system, the network predicts traffic flow over the next half hour, in 5-min intervals. The development of this network is based strictly on data obtained from a simulation model (*15*). Although the effort illustrates the potential of neural networks, it is difficult to generalize the results. In effect, a mathematical model (the neural network) was developed to predict the behavior of another mathematical model (the simulation model).

A similar study illustrates the potential of neural networks for the prediction of freeway traffic volume. Zhang et al. describe a backpropagation network to emulate a macroscopic traffic flow model. They chose such an approach on the basis of the fact that "traffic flow on freeways is a complex process that is often described by a set of highly non-linear dynamic equations" (*12,* p. 2). After training and testing the network on data developed by the macroscopic model, it was concluded that the neural network captured the traffic dynamics of the macroscopic model. Clearly, this is an encouraging conclusion. However, again one will note that a mathematical model (the neural network) was developed to predict the behavior of another mathematical model (the macroscopic model).

Another important research effort exploring the applicability of neural networks to the traffic flow prediction problem was conducted at the University of Leeds. A short-term traffic forecasting model was developed for a surface system using a backpropagation neural network. The model simply relies on current network flow levels as well as flow levels 5 and 10 min in the past. Data from a SCOOT traffic control system in England were utilized. Although the neural network model performed well, it was outperformed by a traditional Box and Jenkins time-series model (*16*). Although this result may seem disappointing, this effort is encouraging because it describes a viable neural network application in a real-world situation.

Clearly, these efforts illustrate the potential of neural networks. However, the need remains to demonstrate the effectiveness of a neural network prediction model using data from an actual freeway facility. Data available from a traffic management system often leave much to be desired, particularly when compared with simulation data. The challenge of maintaining loop detectors, noise in communication systems, and other system problems results in data streams that often look much different from those available from a simulation model. Clearly, it is important to examine the effectiveness of a neural network prediction model using data collected in an operational traffic management system.

CASE STUDY: FREEWAY VOLUME PREDICTION

The purpose of the case study was to develop short-term volume prediction capability at a site on the Capital Beltway. The site selected for this study is on the inner loop of the Beltway near the Telegraph Road interchange in Alexandria, Virginia. At this location, the Beltway is a four-lane freeway, carrying a high volume of local and interstate traffic. In addition, the section is affected by one the region's most notorious bottlenecks, the Woodrow Wilson Bridge.

The Northern Virginia Traffic Management System (TMS) monitors this site with a video camera and full loop detector stations in each of the four lanes. The stations provide the following data continuously to the TMS:

- Volume (vehicles/hour),
- Average speed (miles/hour), and
- Average occupancy (percent).

In addition, Virginia Department of Transportation operates an automatic weather monitoring system (SCAN) to collect pertinent weather data. The SCAN station in Rosslyn, Virginia, roughly 8 mi from the freeway site, is utilized to access the following data:

- Air temperature and
- Pavement condition (wet/dry).

To develop predictive models, a traffic and weather data base was created. The data are stored in 15-min intervals from June 3, 1993, through August 11, 1993, resulting in 3,000 records. This set of data was divided into training and test sets for model development. The training set consisted of 2,550 records, and the test set consisted of 450. Each set of data consisted of roughly a uniform distribution of volume levels. Finally, a third set of data was collected after August 11 to serve as a validation data base.

Models Developed

Three models were developed to predict the link volume at the Telegraph Road site on the Beltway. A 15-min prediction interval was utilized. These models were used to compare traditional approaches to short-term predictions of traffic conditions with a neural network model. A brief description of each model follows.

Historical Average

This model is a simple historical, data-based algorithm. The model developed in this case study simply used the historical average volume, which was calculated using the training data set, as the basis for predicting future volume. In other words, to predict volume on Monday, September 10, at 3:00 p.m., the historical average volume on Mondays at 3:00 p.m. was used.

ARIMA

ARIMA models are among the most powerful and advanced statistical time-series techniques. On the basis of an analysis of autocorrelations and partial autocorrelations of the volume time-series, an ARIMA (2,1,0) model was selected for this application. Such a model describes a second-order autoregressive process that is integrated with no moving average. Model coefficients were based on an analysis of the training data set.

Backpropagation Model

The backpropagation neural network was developed using the following variables as inputs: volume (t), volume (t − 15 min), historical volume (t), historical volume (t + 15 min), average speed (t), and wet pavement (t) (a binary variable). It was trained using a learning rate of 0.3 and a momentum of 0.4. The network architecture consisted of one hidden layer of 10 processing elements.

Performance Analysis

To compare the models, the third, independent validation data set was used. This data set was gathered on two consecutive days (a Monday and Tuesday) in September 1993. In general, all three of the models did an excellent job of predicting future volumes in the short term. In fact, a comparison of error measures in Table 1 reveals that the historical average and backpropagation models displayed comparable error measures, whereas the ARIMA model was less accurate. Figure 2 illustrates graphically the performance of the neural network model on the validation data.

Table 2 displays the average estimate percentage of error for all cases and for cases in peak conditions (defined as any period in which the volume exceeds 3,000 vehicles per hour). Interestingly, the historical data model outperforms the neural network when considering all periods, whereas the neural network model demonstrates better accuracy during peak periods. This indicates that the historical data model can be expected to consistently produce estimates within 5 to 6 percent error levels. On the other hand, one would expect better performance from a neural network during peak periods. This expectation is likely because of the neural network's capability to accurately model the complex characteristics of traffic flow during peak conditions. Clearly, performance during peak periods is of the utmost importance to traffic management and traveler information applications. Therefore, a peak period of the validation data in detail will be examined. The period considered is the p.m. peak, from noon to 7:00 p.m. on Monday.

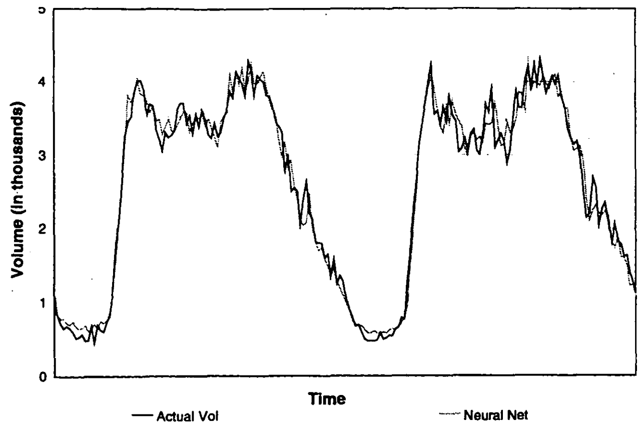


FIGURE 2 Backpropagation mode performance: validation data set.

Peak Period Analysis

Figure 3 illustrates the performance of the historical data model. The model predicts consistently low values for this period. For whatever reason, higher-than-"normal" volumes occurred on this Monday, volumes that the historical model had no capability to predict. This illustrates the significant weakness of such a model; it cannot react to external or abnormal factors that may affect the volume level.

TABLE 1 Error of Prediction Models

Model	Root Mean Square Error	Average Absolute Error
Historical Average	2730	146
ARIMA	3490	195
Backpropagation	2620	144

TABLE 2 Average Percent Forecast Error

Model	Average % Forecast Error All Cases	Average % Forecast Error Volume > 3,000 veh/hr
Historical Average	6.4%	5.0%
ARIMA	9.0%	10.8%
Backpropagation	7.5%	4.3%

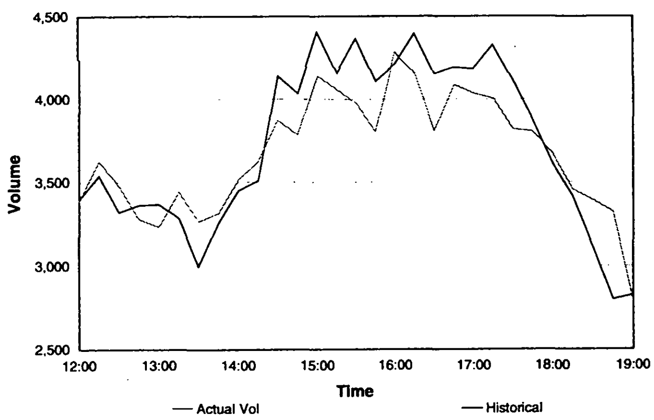


FIGURE 3 Historical data model: p.m. peak.

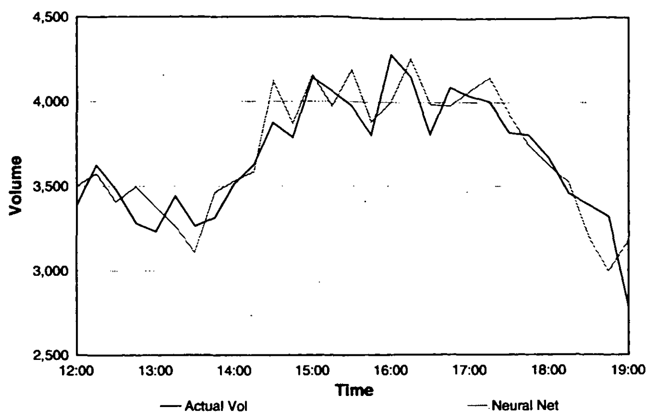


FIGURE 5 Backpropagation model: p.m. peak.

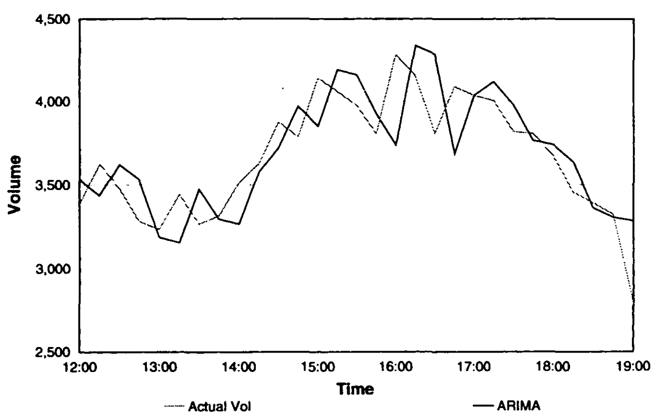


FIGURE 4 ARIMA model: p.m. peak.

Figure 4 illustrates the performance of the ARIMA model. It is clear that the predictions of the ARIMA model tend to lag roughly one interval (15 min) In addition, the ARIMA model tended to overpredict values. This is evident in that volume peaks for the ARIMA model are consistently more extreme than those of the actual volume. The lagging and overpredicting are not surprising given the fact that the ARIMA model uses only time-series data.

Finally, Figure 5 displays the performance of the backpropagation model during the p.m. peak period. This model does an excellent job of predicting volume levels without the lag or overprediction problems of the ARIMA model. This example shows that although all three models have roughly comparable overall error, the backpropagation model clearly does the best job of modeling the underlying relationship between the state of the system and future traffic volume during peak conditions.

CONCLUSION

IVHS technology allows for vastly improved data collection and data communication capability. However, a very real risk is that the world will become data rich and information poor. Thus, a critical effort in the development of IVHS is to create real-time decision support software that will rely on advanced technology, such as expert systems and models. A critical element of such support software that has been identified in this paper is a short-term traffic condition prediction model.

This paper has demonstrated the potential of neural networks to accurately predict short-term traffic conditions in real time. A neural network developed with data from an operational traffic management system performed comparably to traditional prediction approaches when tested with an independent set of validation data. The neural network model, however, outperformed other models during peak conditions, demonstrating its ability to model complex traffic characteristics. On the basis of these promising results, research is continuing to further refine neural network models for ultimate implementation in traffic management systems.

REFERENCES

1. Davis, G. A., and N. L. Nihan. Nonparametric Regression and Short-Term Freeway Traffic Forecasting. *Journal of Transportation Engineering,* 1991, pp. 178–188.
2. Cheslow, M., S. G. Hatcher, and V. M. Patel. *An Initial Evaluation of Alternative Intelligent Vehicle Highway Systems Architectures.* MITRE Report 92W0000063. MITRE Corp., McLean, Va., 1992.
3. Kaysi, I., M. Ben-Akiva, and H. Koutsopoulos. An Integrated Approach to Vehicle Routing and Congestion Prediction for Real-Time Driver Guidance. In *Transportation Research Record* 1408, TRB, National Research Council, Washington, D.C., 1993.
4. Jeffery, D. J., K. Russam, and D. I. Robertson. Electronic Route Guidance by AUTOGUIDE: The Research Background. *Traffic Engineering and Control,* 1987, pp. 525–529.
5. Stephanedes, Y. J., P. G. Michalopoulos, and R. A. Plum. Improved Estimation of Traffic Flow for Real-Time Control. In *Transportation Research Record 795,* TRB, National Research Council, Washington, D.C., 1981, pp. 28–39.
6. Okutani, I., and Y. J. Stephanedes. Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory. *Transportation Research Part B,* 1984, pp. 1–11.
7. Nihan, N. L., and K. O. Holmesland. Use of the Box and Jenkins Time Series Technique in Traffic Forecasting. *Transportation,* 1980, pp. 125–143.
8. Junchaya, T., G. Chang, and A. Santiago. Advanced Traffic Management System: Real-Time Network Traffic Simulation Methodology with a Massive Parallel Computing Architecture. In *Transportation Research Record 1358;* TRB, National Research Council, Washington, D.C., 1992. pp. 13–21.
9. Rathi, A. K., and A. J. Santiageo, *Simulation and Intelligent Vehicle Highway Systems.* FHWA Paper CONF-9209176-1, 1992.
10. Klimasauskas, C. C. Applying Neural Networks—Parts 1–6. *PC AL,* 1991.

11. Nelson and Illingworth. *A Practical Guide to Neural Nets.* Addison-Wesley, Inc., Reading, Mass., 1991.
12. Zhang, H., S. G. Ritchie, and Z. Lo, Macroscopic Modeling of Freeway Traffic Using an Artificial Neural Network. Presented at 72nd Annual Meeting of the Transportation Research Board, Washington, D.C., Jan. 1993.
13. Faghri, A., and J. Hua. Evaluation of Artificial Neural Network Applications in Transportation Engineering. In *Transportation Research Record 1358;* TRB, National Research Council, Washington. D.C., 1992, pp. 71–80.
14. Rumelhart, D. E., and J. L. McClelland. *Parallel Distributed Processing,* Vol. 1. The MIT Press, Cambridge, Mass, 1986.
15. Gilmore, J. F., K. J. Elibiary, and M. Abe. Traffic Management Applications of Neuro Networks Systems. *Working Notes,* AAAI-93 Workshop on AI in Intelligent Vehicle Highway Systems, 1993, pp. 85–95.
16. Dougherty, M. S., H. R. Kirby, and R. D. Boyle. The Use of Neural Networks to Recognize and Predict Traffic Congestion. *Traffic Engineering and Control,* 1993, pp. 311–314.

Publication of this paper sponsored by Committee on Artificial Intelligence.