TRANSPORTATION RESEARCH RECORD

No. 1516

Planning and Administration

Intelligent Transportation Systems

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1995

Transportation Research Record 1516
ISSN 0361-1981
ISBN 0-309-06207-1
Price: $26.00

Subscriber Category
IA planning and administration

Printed in the United States of America

Sponsorship of Transportation Research Record 1516

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 Passenger Travel Demand Forecasting
Chairman: Eric Ivan Pas, Duke University
Kay W. Axhausen, Moshe E. Ben-Akiva, Chandra R. Bhat, Daniel Brand, William A. Davidson, Christopher R. Fleet, David A. Hensher, Alan Joel Horowitz, Ron Jensen-Fisher, Peter M. Jones, Frank S. Koppelman, David L. Kurth, T. Keith Lawton, Fred L. Mannering, Eric J. Miller, Maren L. Outwater, Ram M. Pendyala, Joseph N. Prashker, Charles L. Purvis, Martin G. Richards, Thomas F. Rossi, Gordon W. Schultz, Peter R. Stopher, Jan van der Waard

Committee on Transportation Supply Analysis
Chairman: Randolph W. Hall, University of Southern California
Hadi Baaj, Michael G. H. Bell, David Bernstein
David E. Boyce, Ennio Cascetta, Michael J. Cassidy, Yupo Chan, Teodor Gabriel Crainic, Carlos F. Daganzo, Robert B. Dial, Terry L. Friesz, Michel Gendreau, Ali E. Haghani, Ronald Hagquist, Rudi Hamerslag, Yasunori Iida, Bruce N. Janson, R. Jayakrishnan, Haris N. Koutsopoulos, Hani S. Mahmassani, Chryssi Malandraki, Fred L. Mannering, Mark R. McCord, Eric J. Miller, Anna Nagurney, Srinivas Peeta, Earl R. Ruiter, Yorgos J. Stephanedes, Konstantinos Zograsos

GROUP 2—DESIGN AND CONSTRUCTION OF TRANSPORTATION FACILITIES
Chairman: Michael G. Katona, U.S. Air Force Armstrong Laboratory

Construction Section
Chairman: Donn E.Hancher, University of Kentucky

Committee on Applications of Emerging Technology
Chairman: Chris T. Hendrickson, Carnegie-Mellon University
Secretary: Thomas A. Fuca, New Jersey Department of Transportation
Robert J. Benke, Aviva Brecher, Darcy Bullock, Carl Haas, Richard C. Hanley, Hozayen Hozayen, C. JotinKhisty, Haris N. Koutsopoulos, Sue McNeil, Malcolm Ray, Justin J. Rennilson, Kumares C. Sinha, Yorgos J. Stephanedes, James A. Wentworth

GROUP 3—OPERATION, SAFETY, AND MAINTENANCE OF TRANSPORTATION FACILITIES
Chairman: Jerome W. Hall, University of New Mexico

Facilities and Operation Section
Chairman: Jack L. Kay, JHK & Associates

Committee on Communications
Chairman: T. Russell Shields, Navigation Technologies Inc.
Walter A. Albers Jr., Kenneth C. Allen, James W. Bourg, E. Ryerson Case, Kan Chen, Dennis L. Foderberg, Robert L. French, Charles J. Glass, L. F. Gomes, Robert L. Gordon, Jan E. Hellaker, Gabriel Heti, Randall Jones, Kevin Kelley, Gerard J. Kerwin, Allan M. Kirson, Joseph A. LoVecchio, Wesley S. C. Lum, Roger D. Madden, Said Majdi, Frank J. Mammano, Patrick F. McGowan, B. F. Mitchell, Corwin D. Moore, Jr., Heinz Sodeikat, Philip J. Tarnoff

Committee on Transportation System Management
Chairman: Wayne Berman, Federal Highway Administration, U.S. Department of Transportation
Secretary: Gary R. Erenrich, County of Fairfax
Guzin Akan, Thomas J. Higgins, Frannie F. Humplick, Rajendra Jain, Harvey R. Joyner, Steven Z. Levine, Jonathan L. Levy, Timothy J. Lomax, Michael D. Meyer, Susan L. Moe, Jouko A. Parviainen, George J. Scheuernstuhl, R. Sivanandan, Susanne Pelly Spitzer, Andrzej B. Tomecki, Cy Ulberg, Douglas W. Wiersig

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

Committee on Intelligent Transportation 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, H. Douglas Robertson, 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

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1994.

Transportation Research Record 1516

Contents

Foreword

The papers in this volume focus on various elements of intelligent transportation systems (ITS), automated highway systems (AHS), transportation modeling, defense conversion, and high-frequency automated traveler information systems (ATIS).

One paper focuses on a review of electronic toll collection (ETC) systems and discusses the basic knowledge necessary for implementation of various ETC systems.

Two papers present results of ongoing studies on advanced traveler information systems in Los Angeles and Sacramento.

Two papers focus on automated highway systems; one discusses comparable systems from four main categories: highway-based systems, vehicle-driven systems, other transportation systems, and non-transportation-related systems. The second paper examines lessons to be learned form the BART system as applied to the growing research on automated highway systems.

Two modeling papers discuss a methodological framework for evaluating operating and pricing policies in an intermodal auto-transit (commuter rail) network and the results of implementing a dynamic traffic assignment model.

Two papers focus on information dissemination; one examines computer-integrated transportation as an integrated network of public and private organizations with a common mission of facilitating travel across all modes of transportation; the second presents a case for using a little-used part of the shortwave spectrum as a means of broadcasting digital traveler information to rural and intercity users.

Primer on Electronic Toll Collection Technologies

LAZAR N. SPASOVIC, WEN ZHANG, ATHANASSIOS K. BLADIKAS, LOUIS J. PIGNATARO, EDIP NIVER, AND STANLEY CISZEWSKI

A comprehensive review of electronic toll collection (ETC) systems and related issues is presented with the primary purpose of providing transportation professionals with the basic knowledge necessary for the selection, evaluation, and implementation of various ETC systems. In addition, a tentative framework is offered for demonstrating, testing, and evaluating the performance of ETC systems. Many agencies are expected to procure ETC systems in the near future, creating a clear need for developing a standardized evaluation procedure. Once developed, this procedure can be used to demonstrate the capabilities of various systems and allow the results to be transferred among agencies.

The purpose of this article is to present a review of current designs, as well as developments in and implementation of electronic toll collection (ETC) technologies. ETC is based on vehicle-roadside communication systems; more precisely, it is an application of electronic signature detection to passenger and commercial vehicle traffic for the purpose of collecting tolls. ETC technology is part of automatic vehicle identification (AVI), a functional area of intelligent transportation systems (ITSs). [For a description of the systems approach to ITS, including a discussion of user services, functions to support these services, and technologies that provide certain functions, see Yablonski (*1*).] ETC allows drivers to pay tolls without stopping and can potentially generate significant economic and environmental benefits.

Because of the proprietary nature of ETC technologies, the limited information available about them comes from trade brochures and presentations at specialized trade conferences. These scant resources are not enough. Transportation professionals need to become more knowledgeable about ETC systems. The primary motivation for writing this article is to synthesize the available information, organize it, and present it to the transportation community, thus framing a "breathing document" that will evolve as the technology develops.

This article has two objectives. The first is to present a comprehensive review of ETC systems and related issues. The second is to offer a tentative framework for demonstrating, testing, and evaluating the hardware and functional performance of ETC systems.

L. N. Spasovic, School of Industrial Management and Institute for Transportation, New Jersey Institute of Technology; W. Zhang, Transportation Program, New Jersey Institute of Technology; A. K. Bladikas, Department of Industrial and Manufacturing Engineering and Institute for Transportation, New Jersey Institute of Technology; L. J. Pignataro, Institute for Transportation, New Jersey Institute of Technology; E. Niver, Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, N.J. 07102. S. Ciszewski, New Jersey Highway Authority, Woodbridge, N.J. 07095.

BASIC STRUCTURE OF AN ETC

An ETC system, shown in Figure 1, consists of a vehicle with an on-board unit, a two-way microwave link, and roadside (or tollgate) equipment. The in-vehicle equipment is a transponder, which is usually a tag, an integrated circuit (IC) card with a card holder, or a combination of the two. It stores the information needed for toll transactions, such as vehicle type, account identification, balance, etc. The roadside equipment consists of (a) transceiver (transmitter and receiver) or reader and decoder, the main functions of which are to verify the functionality of the in-vehicle equipment and to conduct the transaction; (b) a lane controller, which monitors activities occurring in a toll lane; and (c) a primary processing computer system, used to access account information and process the transaction requests. A schematic diagram of a generic system is shown in Figure 2. This paper describes the system in which a connection between the in-vehicle and roadside units is established via mobile telecommunication techniques primarily, a microwave or radio frequency (RF) link.

Transaction Process

A typical ETC transaction, shown in Figure 3, starts with the roadside equipment continuously sending interrogating signals via an antenna to the toll lane. This is the transmission phase, in which digital data received from a computer are used to modulate a microwave (or RF) carrier signal, which is then sent by the transceiver to the in-vehicle equipment via a downlink.

When a moving vehicle enters a communication area, its in-vehicle equipment unit receives the interrogating (wake-up) signal. The unit then sends an answering signal back to the transceiver via an uplink. This signal usually consists of elements of the unit's memory (vehicle identification, account number, balance, etc.). Sophisticated units may also have a security mechanism to prevent unauthorized access.

The transceiver receives the answering signal in the so-called reception phase. The roadside unit first verifies its data integrity by checking an error detection and correction code. If an error is detected, the transceiver repeats the interrogation. This process is repeated several times while the moving vehicle is in the communication range. The successfully completed process is referred to as a "handshake." The signal is demodulated to a digital signal (binary data) and sent to the processing computer. The computer deducts the toll from an account and sends this information back to the vehicle via the transceiver. The in-vehicle equipment then sends a signal back confirming that the information was successfully recorded. The process is concluded by the transceiver's reading of the content

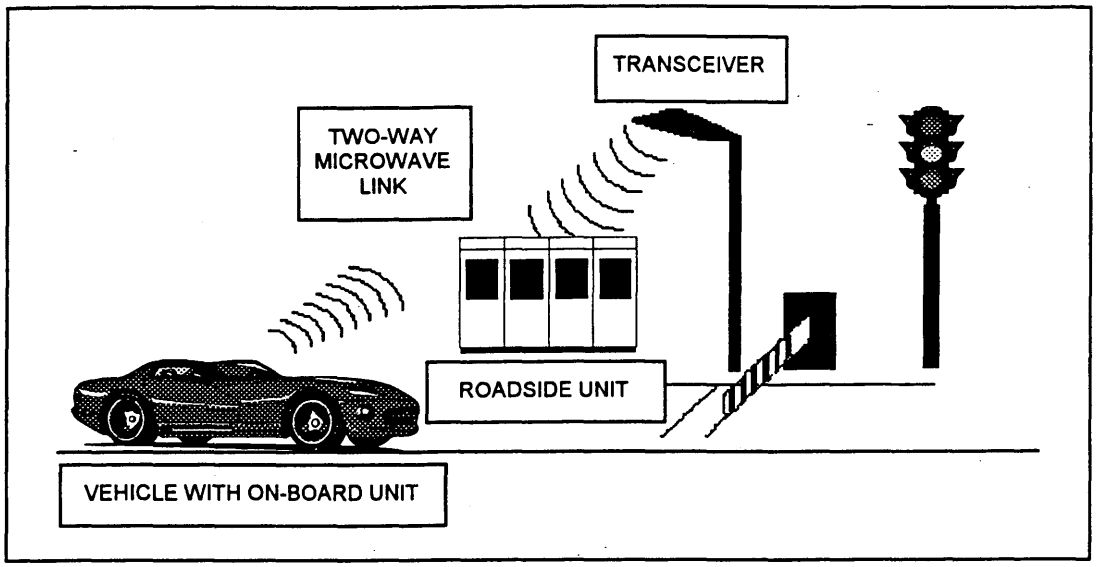

FIGURE 1 ETC system.

of the in-vehicle equipment's memory to verify that the appropriate toll is indeed deducted.

The solid lines in Figure 3 connecting processes represent a normal transaction. The dashed lines represent a transaction in which errors (both in RF communication and insufficient funds) are encountered, and the way the system responds by activating violation enforcement systems (VES).

The above process represents an "open" system operation, in which each time a vehicle passes a toll lane, the roadside unit instructs the transceiver to debit the vehicle's account. The transaction process is more complex when a vehicle enters a "closed" ETC system. At the point of entry, a roadside transceiver reads the memory content of the in-vehicle equipment unit. The computer then verifies the vehicle identification and account number, and whether sufficient funds are available in the account. The transceiver then writes a date, time, location, and lane number "stamp" in the appropriate fields of the unit's memory. When the vehicle exits the system, the transceiver reads the on-board unit's memory and the computer calculates the toll and debits the vehicle's account. This information is then written back to the unit's memory.

Antenna Types and Location

The entire ETC transaction occurs while a vehicle is traveling through a coverage area. The length of the area and the communication distance are determined by the receiver sensitivity, antenna type, location, and transmitted power. The distance is usually no longer than 40 m (137 ft) (*2*).

RF or microwave transmission uses antennas in both the in-vehicle and roadside equipment. In the in-vehicle equipment, the antenna is integrated within the equipment itself and has to be compact. In the roadside equipment, the antenna is installed at the toll plaza. Different manufacturers have different antenna configurations. Some use only one antenna for both triggering (sending an interrogating signal) and communicating, whereas others have sep-

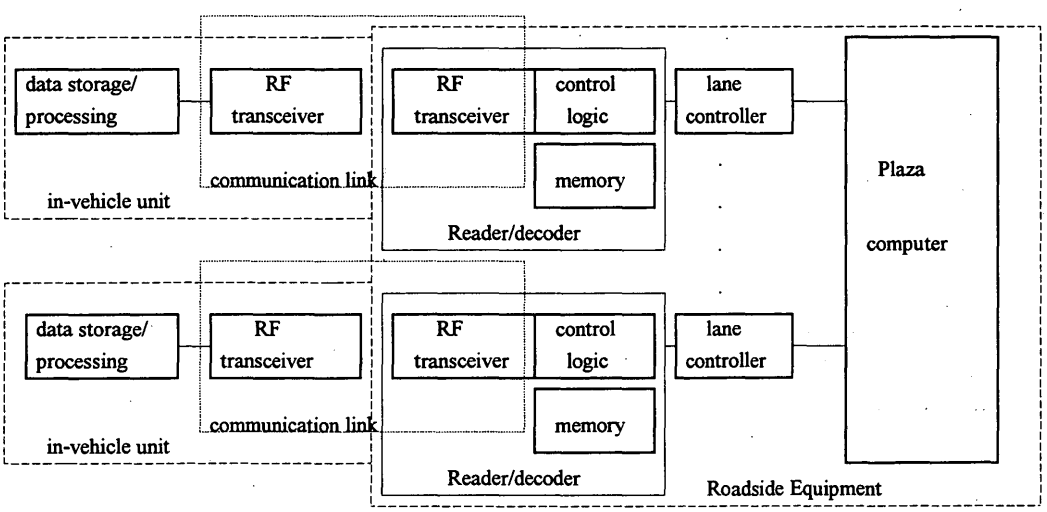

FIGURE 2 Schematic diagram of generic ETC system.

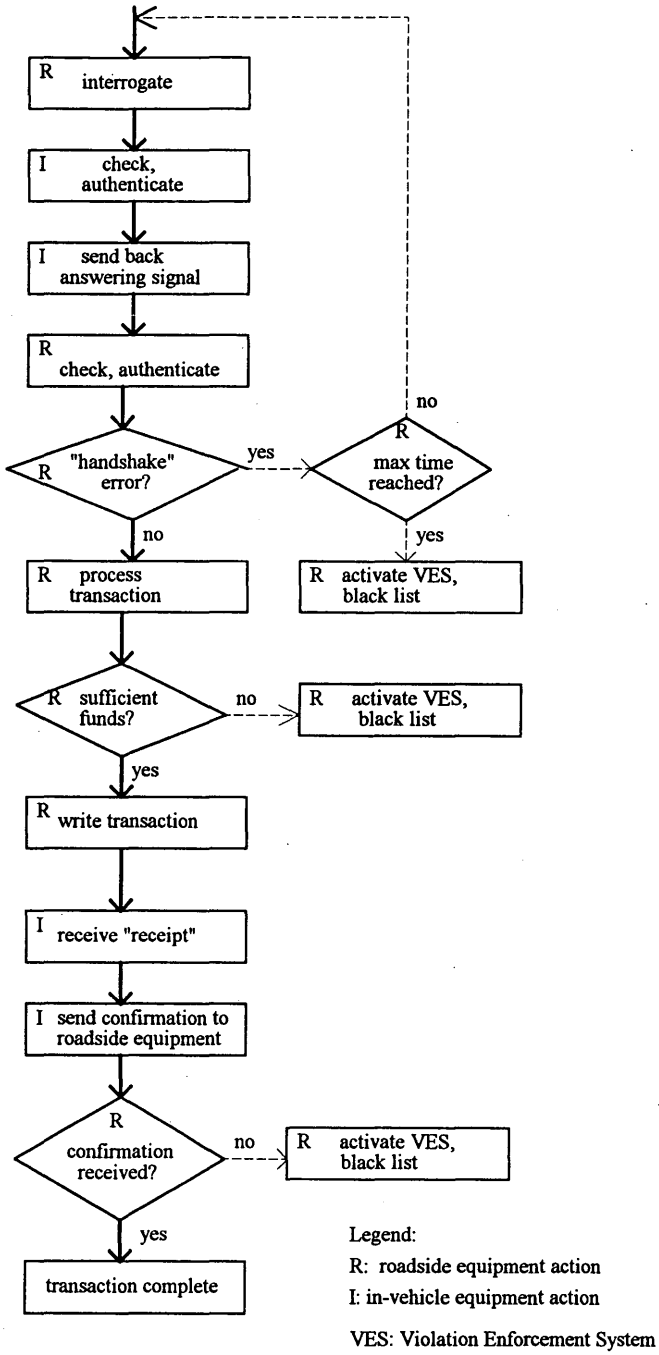

FIGURE 3 ETC transaction process.

arate antennas for each function. Still others have three antennas, one for triggering, one for communicating, and one for collecting position information to adjust the RF power output level or adjust antenna propagation patterns to optimize communication.

In terms of location, there are three kinds of antennas: in-pavement, distributed overhead, and focused beam. When one antenna is used to read multiple tags over multiple lanes, it is called a multiplexing antenna. Antenna selection is closely related to the frequency used, communication distance, and properties of the in-vehicle unit.

TAG AND IC CARD–BASED ETC SYSTEMS

The variations in the design of roadside units are not considered major. In-vehicle equipment, however, may differ widely. It may consist of a tag (a small electronic device attached to the windshield of a car) or an IC card and a card reader, which are usually installed in the dashboard. Based on the design of the in-vehicle device, ETC systems can be classified as tag or IC card–based.

Tag-Based ETC Systems

In a tag-based system, the tag has memory, data processing, and communication capabilities, and it is also an RF device. On receiving a designated signal, it emits a radio signal of its own that is used for detection, identification, and location. A tag can be called a transponder. It can communicate without physical contact with a purpose-specific station over a distance ranging from a few millimeters up to several meters. The basic structure of a tag consists of several functional electronic circuits such as memory, control logic, RF modulation unit, and antenna. A tag can have its own power supply or operate without it.

Classification

According to its communication capability, a tag can be classified as read-only or read and write. A read-only tag can only be read by the roadside reader and is also referred to as a Type I transponder. A read and write tag (Type II transponder) can be read and also be written on by the reader. The information that is written can contain the tollgate identification, account balance, etc.

A tag can be classified as active, semiactive, or passive according to the source of power and its ability to generate its own RF signal or simply reflect the incoming signal. An active tag is always operational. It gets power from an internal battery or from the vehicle via a converter. Some active tags can receive a confirmation signal by the transceiver via the tag's RF unit, thus possessing a read and write capability.

A semiactive tag can also receive and transmit signals. However, it is not active in the absence of a transaction. It is activated only when it receives an interrogating signal from the reader's antenna.

In marked contrast to active and semiactive tags, passive tags have no power supply and thus can only be read after they collect enough power from the reader's RF signal to activate the electronics. They are primarily used in automatic vehicle identification (AVI) and early ETC systems.

Transmission and Modulation Methods

Whether a tag is active, semiactive, or passive, its transmission and modulation methods depend on the vendor. Intellitag's, Texas Instruments', SAAB's, and Amtech's tags, which are semiactive, use a "modulated backscatter" technology (*3*). The tag sends information back to the reader by changing or modulating the amount of RF energy reflected back to the reader antenna from a continuous-wave RF signal beamed from a reader. The RF energy is either allowed to continue traveling past it or is intercepted and "scattered" according to the tag's antenna pattern back to the reader antenna.

The operation of switching the "scattering" on and off can be done with very little cost in hardware and with very little electrical power. As a result, a tag can use so little energy that it can be powered with an internally integrated battery (similar to an electronic watch) or with power derived by rectifying the RF signal intercepted from the transceiver. Their efficiency, read and write capability, and low cost make these tags widely used in AVI and ETC.

Design and Data Storage

The information on the tag could be structured and stored in various ways, requiring different memory capabilities. The data storage capabilities usually differ among vendors. An example of a tag with read and write capabilities is given in Table 1. This particular tag, which happens to be the PREMID 3100, has a serial data package capability of 60 bytes (*2*). This data packet is divided into read and write portions. Table 1, Part *a* shows its read portion (bytes 0 to 23), which stores fixed data.

Table 1*b* shows that the write portion of the data (bytes 24 to 59) stores real-time information such as date, time, account value, and agency data. During each transaction, this portion of the memory must be rewritten to ensure that the toll amount is deducted from the account. In this example, the write portion contains sample data from a 1991 New York State Thruway Authority ETC field test.

Early tag designs called for data allocation to a specific physical portion of a tag's memory location. Current design calls for a "smart transponder," which is a microprocessor-based system with a flexible data structure (*4*).

Critical Tag-Based ETC Design Issues

Currently, there are two problems that plague tag-based ETC systems. The first problem is that products from different vendors use incompatible uplink and downlink hardware. The exceptions to this are the Amtech, Intellitag, and Texas Instruments systems, all of which use the same hardware. The second problem is that tags from different vendors have different data structure and storage capabilities. Either of these issues can lead to the problem of on-board to roadside unit incompatibility.

There are three ways to alleviate this problem. First, regional toll agencies can get together and select a common technology. This was the logic behind establishing interagency groups (IAGs) that

TABLE 1 Premid 3100 System Ample Serial Data Packet (*2*)

a. Read only portion

Information type	Location	Format	Content type
Tag type	Byte 0	W	Fixed ASCII char.
Application ID	Byte 1	T	Fixed ASCII
Group ID	Byte 2-3	01	Fixed code
Agency ID	Byte 4-5	01	Fixed code
Serial No.	Byte 6-11	nnnnn	Range 00001-000041
Checksum	Byte 12	C	Fixed ASCII char.
Vehicle type	Byte 13	1	Fixed
Vehicle profile	Byte 14	1	Fixed
Vehicle axis	Byte 15	4	Fixed
Revenue type	Byte 16	0-9	Fixed to 1
Agency data	Byte 17-21	NYSTA	Fixed ASCII string
Checksum	Byte 22-23	CS	Fixed

b. Write portion (for toll collection)

Information type	Location	Format	Content type
Plaza ID	Byte 26-27	01	01=Spring Valley 02=Tappan Zee
Lane ID	Byte 28-29	01	Fixed
Julian date	Byte 30-34	JUL04	Varies JAN01 to DEC31
Hour	Byte 35-36	hh	Varies 00 to 23
Minute	Byte 37-38	mm	Varies 00 to 59
Second	Byte 39-40	ss	Varies 00 to 59
Txn Number	Byte 41-44	nnnn	Increments with every successful transaction
Vehicle type	Byte 45	0-9	Revenue amount for debiting
Vehicle profile	Byte 46	1	Fixed
Vehicle axis	Byte 47	4	Fixed
Value	Byte 48-52	nnnnn	Reduced every transaction according to vehicle type stored in byte 45
Agency data	Byte 53-57	NYSTA	Fixed
Checksum	Byte 58-59	CS	Fixed

brought together various regional toll agencies to establish a common standard in the region. By ensuring that a selected technology will be used on all facilities, IAGs could increase ETC use and market penetration. The IAGs that are working out well include the following:

- The E-Z Pass, formed by toll agencies in the New York-New Jersey-Pennsylvania region, the operations of which account for some 37 percent of all U.S. toll transactions;
- The New England IAG;
- A California IAG, consisting of the Department of Transportation, Golden Gate Bridge, Highway and Transportation District, the Transportation Corridor Agency, and developers of four private toll roads; and
- The Greater New Orleans Expressway Authority and the Louisiana Department of Transportation and Development (*5*).

Second, different vendor processing software can be used at each tollgate, but this solution could prove impractical and prohibitively costly. Furthermore, a tag's fixed data structure makes it difficult to make modifications to adapt to future applications (e.g., traveler information service) that one of the regional agencies might offer.

The third and final solution is to develop tags that have flexible data structures so that they can operate in a multivendor environment. For example, Intellitag's tag (*6*) has an extended memory consisting of 20 data frames, each frame having 128 bits (a total of 2,560 bits). All but one of the frames can be programmed as fixed or variable according to a toll agency's data storage format. During each transaction, the tag's memory is read or written according to the "frame." This tag could be used with different toll agencies, as well as for other services (e.g., access control, parking management, etc.) (*7*).

IC Card–Based ETC Systems

The IC-based in-vehicle equipment consists of an IC card, a card reader (usually installed on the dashboard), and an RF link unit. An IC card, which incorporates a microprocessor, memory, and other electronic components, provides a means of storing information during each transaction. The RF link unit can be a tag mounted on the windshield or any other form of transponder that can establish a communication link between an IC card reader and the roadside equipment. This link can vary among different vendors from an ordinary frequency-modulated microwave signal to a spread-spectrum modulated signal.

System Description

The IC card consists of a microprocessing chip, memory, and input and output devices that are sealed with hard plastic or ceramic coating into a card. This design incorporates two innovative features. First, although in an ordinary IC (the reader can think of a personal computer as an analogy) the power line is linked directly to the IC's input pin, the IC card has no direct link with a card reader. Instead, an ultrathin printed circuit board with electronic components is "wire-bonded" to its surface. The card uses an "etched coil" to receive power and clock a signal from a card reader. When the card is inserted in the reader, the coil in the reader is coupled with the coil in the card, thus enabling the transfer of power and the clock signal between the reader and the card. Second, the data transfer is realized by using two pairs of "capacitor plates" in the card. The card has a custom IC to transfer data to and from the capacitor plates. The relative positions of a card and a card reader are very important, because the "etched coil" and "capacitor plates" of the two devices must be coupled for power and data transfer.

The reader has a data transfer circuit that uses the same custom IC that is used in the card to transfer data to and from the capacitor plates. This IC provides transistor-transistor-logic level signals that can be connected to a microprocessor serial port or RS232 drivers and receivers, which provide an interface between data terminal equipment and data communication equipment. Using this serial port, a card reader can be connected to any equipment such as a handheld reader, a personal computer, an automatic teller machine (ATM), a simple liquid crystal display (LCD), or a personal assistant system.

The card reader can also take on many different forms. It can be mounted inside a slotted plastic housing to form an insertion-type card reader or it can be mounted into the housing itself to form a surface-type card reader. The flexibility of interface also means that a card reader can be designed to be a freestanding device or it can be integrated into another device similar to the way a disk drive is mounted into a personal computer (*8*).

Special Security Device

Because an IC card contains a microcomputer, it is possible to employ many sophisticated security measures to protect the data integrity. For instance, the AT&T Smart Card has a personal identification number (PIN) protection. A special PIN is assigned for each directory and file access control security features. The security features may include many options, such as:

- One-time proof of user identity;
- User identity required every time a file is accessed;
- Card identity check before each transaction to prevent use of bogus cards; and
- User identity checking only at the IC card holder's request.

The above-mentioned file access control methods can be used in any combination, providing different degrees of data security (*9*). These features are especially desirable if an IC card is used in a multitoll operator region because they can be used to control access to each specific memory location.

IC Card Design Flexibility

An IC card has higher data storage capabilities than that of a typical tag [5 to 10 kilobytes (Kb) memory versus 2 to 3 Kb]. In addition, a change in the system's control function can be easily made by modifying its software at relatively low cost. This gives an IC card-based ETC two advantages over a tag based system. First, it can be modified easily to satisfy new standards. Second, if any portion of the memory is not functioning well, the problem can be corrected simply by redesigning the memory map of the card using the operating system command.

RECENT ETC APPLICATIONS

Because ETC can increase highway efficiency by reducing toll collection time and cost, many toll agencies in the United States and abroad are actively considering, testing, and implementing ETC technologies. Some of them are shown in Table 2, which is our updated 1992 Survey by the International Bridge, Tunnel, and Turnpike Association (*10*). A comprehensive review of the current application of United States and European electronic toll traffic management systems is given by Schuster (*5*).

Cost of ETC Systems

The costs of ETC systems (in 1993 dollars), obtained from the agencies that implemented them (e.g., Dallas North Tollway, Oklahoma Turnpike) are shown in Table 3 (*11*). It is indicated that early applications of ETC systems are very cost-effective (*5*). The Oklahoma Turnpike Authority reported $160,000 in savings per lane from replacing a single manual lane by the agency's PIKEPASS system.

ETC COMMUNICATION DESIGN ISSUES

Even the simplest ETC systems have rather complex communications needs. In-vehicle equipment must be identified and the relevant information must be transmitted between a vehicle and roadside unit. This information needs to be converted from a radio wave to a computer message, which must be processed through a computer system that will act on that message.

To guarantee a reliable communication data rate, operating frequencies and antenna selection must be considered. Reliable communication also depends on interoperability. Therefore a protocol

TABLE 2 Selected ETC Survey Results

Country, Agency and Facility	Lanes	Vendor	AVC	Enforcement
System under operation				
E-470 Public Highway Authority Colorado	4 dedicated 2 mixed	X-cyte	No	Video image
Dallas North Tollway Texas Turnpike Authority	4 dedicated 59 mixed	Amtech	Yes	Violation system
Oklahoma Turnpike Authority	56 dedicated 117 mixed	Amtech	Yes	Violation Surveillance
Crescent City Connection Louisiana Dept. of Transport.	3 dedicated 9 mixed	Amtech	Yes	None
Dartford River Crossing England	2 dedicated 6 mixed	SABB PREMID	No	Pre-classification Blacklisting
BRISA-Auto-Estradas de Portugal Portugal	8 dedicated	MICRO DESIGN	Yes	Video Retrieval System
Autostrade S.p.A Italy (IC card based ETC systems)	1800 dedicated	Autostrade S.p.A	Yes	Post-Payment Card
System under consideration				
E-Z Pass Group - New York - -- New Jersey --Pennsylvania (17 toll authorities)	varies between agencies	Finished the evaluation -- selected the AT&T-MARK IV	Yes	varies between agencies
Tobin Memorial Bridge Massachusetts Port Authority	N/A	Tested Amtech -- not decided	N/A	N/A
All roads North/South Illinois State Highway Authority	N/A	AT/Comm	N/A	N/A
Florida Turnpike Florida DOT	N/A	Finished testing - not decided	N/A	N/A
DRIVE Sweden (IC card based ETC systems)	N/A	Philips IC card	N/A	N/A
CGA/Gemplus HAMLET 2 French (IC card based ETC systems)	N/A	Premid and Schumberger IC card	N/A	N/A

Sources: International Bridge, Tunnel, and Turnpike Association, March 1992 (*10*),

TABLE 3 Equipment, Operating, and Maintenance Costs by Lane Type

Lane Type	Equipment Cost ($/lane)t	Operating and Maintenance Cost ($/lane)
Manual	58,500	141,900
Automatic	58,000	43,300
Manual/Automatic	107,500	111,000
Manual/AVI	72,700	146,100
Automatic/AVI	69,500	47,500
Manual/Automatic/AVI	121,300	115,200
AVI dedicated	15,400	4,200
Express AVI	15,400	4,200

Source: Pietrzyk, M. C. and E. A. Mierzejewski (*11*).

governing data format, error detection and correction method, and memory allotment is needed, just as for a computer communication network.

In an effort to resolve these issues, the Standards and Protocols Committee of ITS America published *Electronic Toll and Traffic Management (ETTM) User Requirements for Future National Interoperability* (Draft Version 2.0) (*12*). The standardization of different system design elements will not be easy, because the current systems are operating in various frequency ranges. Comparisons between the ITS America recommendations and vendors' data are shown in Table 4. For example, setting a certain frequency band for ETC communications (e.g., 5.8 GHz) will severely limit the competitive edge of vendors whose technology is outside the approved band.

European Experience

European promoters of ETC moved swiftly and set up an Europe-wide nonproprietary specification standard for ETC systems. This has been accomplished via a Vehicle Information and Transaction Aid (VITA) program. In late 1990, a VITA I document was developed by French, Italian, and Spanish toll agencies. It was revised in 1993, yielding the VITA II version. The document defined requirements for an open-protocol, two-way vehicle-roadside communication system with a high data rate, low bit-error rate, and low transaction error. The 5.8 GHz European Radio Communication Committee band was chosen for ETC communications. Additional VITA I and VITA II specifications can be found in Yacoubi (*13*).

TABLE 4 Comparison Between Draft 2.0 Recommendations and Vendors' Data

	Draft 2.0 recommendations	Intellitag	Amtech	Mark IV	SAAB Combitech
Frequency location	902-928 MHZ 2.45 GHZ, 5.8 GHZ	902-928 MHZ	850-950 MHZ 2400-2500 MHZ	902-928 MHZ	5.795-5.805 GHZ
Data rates	550 Kbps± 10%	300-600 Kbps	N/A	550 Kbps±10%	250 Kbps
Minimum capture rate	>99.95% or >99.00%	N/A	>99.9999%	N/A	99.99999%
Error rate	<0.00001%	N/A	(1-0.999999) %	N/A	0.0000001
RF characteristics					
Modulation (downlink)	ASK (Manchester)	ASK (Manchester)	ASK	ASK (Manchester)	ASK (Manchester)
Modulation (uplink) Active Backscatter	ASK (Manchester) AM subcarrier or FSK modulated at 600-1200 KHZ±10%	 FSK modulated	 FSK	ASK (Manchester)	 FSK
Power density	<20μw/cm^2 (max)	N/A	100μw/cm^2 at surface	N/A	2.0W EIRP
Sensitivity	<1μw/cm^2 at toll booth when field strength<500 mv/m, tag should not respond	N/A	N/A	N/A	downlink >-43 dbm uplink>-100 dbm
Protocol					
data	HDLC, TDMA, binary	HDLC, TDMA, binary	N/A	TDMA, binary	HDLC, binary
error correcting	CRC-CCITT, Luhn code, ID check	CRC-CCITT, Luhn code, ID check	N/A	CRC-CCITT	N/A
Tags					
memory	PROM>256 bits, read/write>1Kbits	2560 bits total, flexible ROM and RAM allocation	N/A	N/A	N/A

TAG, IC CARD, OR COMBINED ETC SYSTEMS?

Currently, in the United States, tag-based ETC systems are widely used, whereas ETC IC card applications are nonexistent. However, the European experience of the Autostrade's AT&T Smart Card–based ETC system indicates that IC cards have enormous potential in ETC. This potential comes primarily from four unique features:

- The card's interface can be integrated into any other device such as a personal computer, ATM, or computer security device with little additional cost.
- The card can be attached to auxiliary on-board devices such as sensors for monitoring vehicle operation or display devices such as a CRT, LCD, or on-board computer.
- The card can be removed after use and is easy to carry.
- The communication link with other devices is not limited to RF (or microwave) communication.

It is this flexibility of the IC card and the potential of using it for a myriad of services that can make it the dominant medium for all electronic transactions. For example, the card can be used as a credit card, bank card, calling card, transit pass, parking card, and toll card, provided, of course, that these services have been equipped with an appropriate reader.

Europeans have already recognized the potential of IC cards. They have undertaken two major projects with the goal of designing a standardized ETC system with in-vehicle equipment based on an IC card. The Pricing and Monitoring Electronically for Automobiles (PAMELA) project was implemented to design and test a two-way microwave link within the Dedicated Road Infrastructure for Vehicle Safety in Europe (DRIVE 1) program. The Automatic Debiting and Electronic Payment project was implemented within the DRIVE 2 program to design, prototype, and field test a complete ETC system based on a microwave link developed in PAMELA. The system design calls for in-vehicle equipment built around the IC card. Figure 4 shows a schematic diagram of the in-vehicle equipment (*13*).

The French Toll Road Operators Association (USAP) should be given credit for initiating an international call for awarding a contract to build a French standard system for ETC. The USAP developed strict nonproprietary specifications, similar to those listed in VITA. In 1992, two consortia (one was led by CSEE Péage and included GEA and SAAB-SCANIA, and the other was led by CEG-ELEC-CGA and included GEM+), were awarded contracts to build the prototype systems, called télépéage intersociété. The prototype systems are currently undergoing indoor and field hardware and functional testing. It is expected that the selected system will be available in 1995 (*13*).

SECURITY AND LEGAL ISSUES

Clearly, ETC services may yield tremendous benefits to toll agencies and users. These technologies, however, monitor and control vehicle operations and trajectories and may elicit images of "Big Brother" watching. More to the point, those participating in the ETC program need to know whether information gathered by an ETC system may legally be used against them. The resolution of these issues will have an impact on market penetration, because people do not tend to participate in an activity in which they are being monitored.

Furthermore, two issues need to be distinguished: security and privacy. Security deals with protecting the integrity of the data transaction. For this purpose, the systems use various sophisticated techniques to prevent unauthorized users from gaining access to the authority and client information. These techniques vary from sophisticated data encryption techniques to password checking. For example, the actual user (tag) identity is scrambled and carried through the system as an unintelligible text.

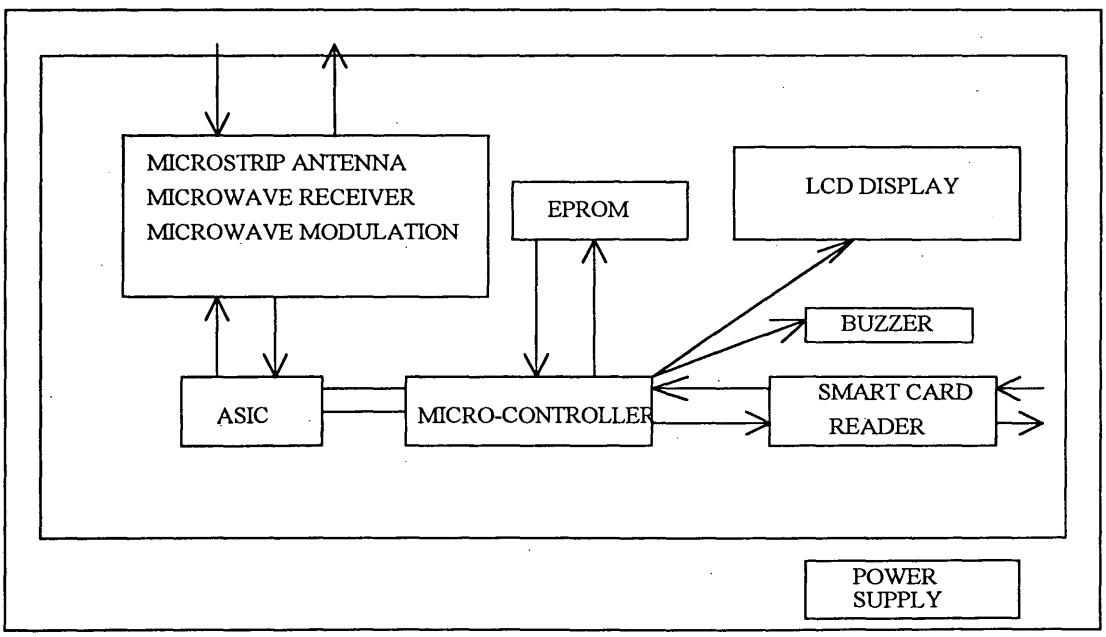

FIGURE 4 Schematic diagram of the in-vehicle equipment (*13*).

Assuming that the transaction has not been compromised, however, the question arises about whether a record of the transaction resides within the toll authority's files and whether it can be unscrambled and used to expose the user to potential legal liability.

In the read-only tag-based ETC system, the central data base stores all information concerning users' identification, account numbers, and account balances. The scrambled data communication records are also stored there. This stored information can be used to trace the movement of a particular vehicle. For example, law enforcement agencies may ask courts to order toll authorities to release data on a particular user's travel route, time, etc. This procedure is similar to the one in which a court asks a telephone company to provide detailed records on all calls, including the time of day for local calls, which usually do not appear on typical phone bills.

Introducing technology that records all transactions on a user device can alleviate this problem. For example, if one uses the read and write technology with a tag or an IC card as electronic money stored in the tag or card memory, the information resides with the user. Once the transaction between the roadside reader (authority) and the user is completed and all information regarding travel is stored on a card or tag and the user has left the system, the authority could purge all information about the transaction from its records. This operating procedure can put the user in control somewhat. However, the courts could subpoena the tags or IC cards from users as well.

TEST PROTOCOLS

With more agencies moving toward procurement of ETC systems, a set of consistent procedures must be developed to evaluate the performance of the systems. ETC system testing should include static, functionality, and long-term tests.

Static or controlled tests are performed in a laboratory to check a system's communication range and speeds, ability to withstand interference, reliability under harsh weather conditions, and electronic component reliability. A detailed description of the test protocols is given by Spasovic et al. (*14*).

Functionality tests are focused on assessing the functionality of tags or IC cards and can be performed simultaneously with static tests. They include tests to verify whether transaction information is correctly recorded, and fraud or cheating tests to identify potential loopholes in the systems.

Long-term tests are designed to test operational reliability of the system in an actual toll collection setting. They require that a large number of vehicles pass through the tollgate.

It is imperative that an ETC system can deliver a high degree of accuracy regardless of the adversity or hostility of the operating environment. The recommendation in ITS America Draft 2.0 (*12*) defines this accuracy in terms of capture rate (percent of successful reading and decoding of a properly mounted tag's data message) and error rate (percent reading error of a tag's identification fields). To satisfy these requirements, an ETC system must have a capture rate larger than 99.95 percent and an error rate smaller than 0.00001 percent (1 in 100,000).

Sample Size Estimation

If p denotes the probability of a successful transaction (neither read nor write errors occur) and q the probability of failure (i.e., $1 - p$), then the probability of k successful transactions during n tests is obtained by the binomial distribution:

$$P(Y=k)=\binom{n}{k}p^k q^{n-k}$$

Using the maximum likelihood estimation method, an estimate for p is:

$$\hat{p}=\frac{Y}{n}$$

and the estimation error is:

$$p\left\{\left|\frac{Y}{n}-p\right|<d\right\}\geq 1-\alpha$$

where

$1 - d$ = measurement accuracy and
$1 - \alpha$ = confidence level.

If it is necessary to have a data capture of 99.95 percent and a confidence level of x, then:

$d = 1 - 99.95$ percent $= 0.0005$ and

$\alpha = 1 - x$.

For the binomial distribution:

$$P\left\{\left|\frac{Y}{n}-p\right|<d\right\}=P\left\{-d\sqrt{\frac{n}{pq}}<\frac{Y-np}{\sqrt{npq}}<d\sqrt{\frac{n}{pq}}\right\}$$

According to the Central Limit Theorem, the above probability can be approximated by the normal distribution as:

$$\Phi\left(d\sqrt{\frac{n}{pq}}\right)-\Phi\left(-d\sqrt{\frac{n}{pq}}\right)=2\Phi\left(d\sqrt{\frac{n}{pq}}\right)-1 \qquad (p+q=1)$$

From

$$\mathrm{F}\left(d\sqrt{\frac{n}{pq}}\right)=1-\frac{\mathrm{a}}{2}$$

the sample size is:

$$n=\left(\frac{z\frac{\alpha}{2}}{d}\right)^2 * p * q$$

The required sample size depends on the measurement accuracy, confidence level, and probability of success. For a given measurement accuracy of 0.0005 and a probability of success, $p = 0.9995$, the required sample size varies with confidence level. For a 0.9 confidence level, the sample size is 1,645; it increases to 7,679 for a 0.95 confidence level.

The sample size is much more sensitive to the measurement accuracy than it is to the confidence level. For example, if the measurement accuracy is reduced from 0.0005 to 0.00001, and leaving $p = 0.9995$, the sample size becomes 13.5 million for a 0.9 level of confidence and 19.2 million for a 0.95 level.

CONCLUSIONS

Several issues relating to ETC technologies have been presented. The potential of these technologies is enormous. In addition to benefits related to toll collection, these technologies are capable of providing added value because they can be used in many other transportation services such as traffic management and traveler information, as well as for nontransportation services such as access control, distribution of benefits, and so forth. The technologies, however, are not mature, and as such, they have not been thoroughly field tested. The proprietary nature of the devices and the large sample sizes required to validate claims of accuracy may force the transportation profession and the general public to wait a relatively long time for the best devices to rise and become industry leaders and reliable off-the-shelf items. Industry, public sector, and academic consortia formed for the testing of existing or future devices and technologies would speed up the evaluation process in realistic operational environments, while at the same time maintaining the confidentiality and safeguarding the proprietary nature of the technologies.

ACKNOWLEDGMENT

This research was partially supported by a grant from FHWA Cooperative Agreement DTFH61-92-X-00024. This support is gratefully acknowledged but implies no endorsement of the conclusions by the FHWA.

REFERENCES

1. Yablonski, A. *IVHS User Services and Functions*. Working Paper. Mitre Corporation, Washington, D.C., 1992.
2. *Premid Introduction*. Saab Automation, Combitech Traffic Systems, Jonkoping, Sweden, 1991.
3. Koelle, A. R. Advances in Practical Implementation of AVI Systems. *Proc., Vehicle Navigation and Information Systems Conference*, Dearborn, Mich., 1991, pp. 969–975.
4. *AT/Comm Brochure*. AT/Comm, Inc., Marblehead, Mass., 1994.
5. Schuster, N. D. *ETTM Technology: Current Successes and Future Potential*. International Bridge, Tunnel, and Turnpike Association, Washington, D.C., 1994.
6. Intellitag Documentation. Intellitag Products, Scottsdale, Arizona, 1994.
7. Brown, R. Can One Tag Fit All? *Intelligent Highway Systems. ENR* Supplement, November 28, 1994, pp. 28–29.
8. Komaneck, M. R. IVHS Applications of Smart Card. *Proc., Vehicle Navigation Information Systems Conference*, Dearborn, Mich., 1991, pp. 977–988.
9. AT&T Brochure. AT&T Smart Cards, Bridgewater, N.J., 1994.
10. *Survey of ETC Systems*. International Bridge, Tunnel, and Turnpike Association, Washington, D.C., 1992.
11. Pietrzyk, M. C., and E. A. Mierzejewski. *NCHRP Synthesis of Highway Practice 194: Electronic Toll and Traffic Management (ETTM) Systems*. TRB, National Research Council, Washington, D.C., 1993.
12. ETTM User Group. *Electronic Toll and Traffic Management (ETTM) User Requirements for Future National Interoperability* (Draft 2.0), Special Group of the Standards and Protocols Committee, ITS America, Washington, D.C., 1994.
13. Yacoubi, S. An Electronic Toll and Traffic Management System. *Microwave Journal*, July 1994, pp. 64–72.
14. Spasovic, L. N., W. Zhang, and E. Niver. *Radio Frequency (RF) Tag–IC Card Based Electronic Toll Collection (ETC) System Test Plan*. Institute for Transportation, New Jersey Institute of Technology, Newark, N.J., 1994.

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

Los Angeles Smart Traveler Information Kiosks: A Preliminary Report

GENEVIEVE GIULIANO AND JACQUELINE M. GOLOB

Advanced Traveler Information Systems (ATISs), a key part of new technology applications in transportation, provide accurate and timely information that helps travelers select routes, times of travel, and travel modes. The potential for new technology applications in transportation is extensive, and substantial investments are being made to further this rapid technological development. Comprehensive, systematic testing and evaluation of new technology elements, therefore, is needed to ensure that the most promising and cost-effective technologies are pursued. To this end, several Field Operational Tests (FOTs) are being conducted throughout the United States. Preliminary results from an ongoing FOT evaluation, the Los Angeles Smart Traveler project, are presented. Usage patterns based on automated data and user surveys are examined for a sample of 41 kiosks throughout Los Angeles County. The data indicate that kiosks located in retail establishments are used more than those located in employment centers, and that that usage is related to the level of pedestrian activity at the site. Satisfaction with the kiosks is very high among users. The authors conclude that kiosks are used to obtain information more often for nonroutine or new trips than for the regular commute trip.

Choices travelers make regarding route, mode, time of departure, etc., are limited by the information available to them. Advanced Traveler Information Systems (ATISs), which seek to improve the accuracy and timeliness of this information, are being developed to improve traffic conditions and make transportation systems more efficient.

Advocates of pre-trip ATISs argue that access to information on ridesharing and transit is limited. Travelers must telephone the local rideshare or transit agency and request information. Because telephone lines are often busy, prospective transit riders must spend long periods of time on hold. Typical ridematching procedures take days or weeks, as match requests must be verified, potential partners identified, and results returned. Some argue that if such information were more accessible, transit use and ridesharing would increase.

THE LOS ANGELES SMART TRAVELER FIELD OPERATIONAL TEST

Smart Traveler is being implemented by the California Advanced Public Transportation Systems (CAPTS) Group, within California's Department of Transportation (Caltrans). Originally designed as a limited test within a confined freeway corridor (I-110), a reorientation and major expansion of the project was ordered after the Northridge earthquake to target affected areas of Los Angeles (Figure 1).

G. Giuliano, School of Urban and Regional Planning, University of Southern California, Los Angeles, Calif. 90089-0042. J. M. Golob, Jacqueline Golob Associates, 15 Whitman Court, Irvine, Calif. 92715.

The Smart Traveler Kiosks

The kiosks are a multimedia, personal computer-based system for accessing information on bus routes, carpools, and freeway congestion, and for viewing videos on various transportation topics. Designed and developed by IBM, the kiosks operate from a personal computer. Bus route, carpool, and freeway congestion data bases are accessed via modem. Videos are stored on a laser disk within the unit. The computer interface, for input and output, is a touch-screen monitor. Displays are in textual form, except for freeway congestion, which is displayed on a map. A 40-column printer allows users to print and take home bus route directions and carpool match lists. The kiosks are totally self-contained; only the touch-screen is visible to the user.

Kiosk usage is menu-driven. Progress through menu items is determined by user requests and responses. The first choice is language; all menu items are available in English and Spanish. The second choice is type of information (e.g., freeway conditions, bus routes, etc.). The number and extent of menu branches is determined by the complexity of the information requested. For example, only two commands are required to access the map of freeway conditions, whereas 10 or more commands may be required to obtain complete route information for a given trip.

The kiosks are linked and managed by a communications network. The system is depicted in Figure 2. The kiosks are connected to the IBM 3090 computer at the California Health and Welfare Agency Data Center (HWDC) in Sacramento. The HWDC mainframe is the central processing unit of the system. Its role is to obtain information from the three data bases (Caltrans for freeway conditions; Commuter Transportation Services, Inc. for carpool information; and the Los Angeles County Metropolitan Transportation Authority for transit information) in response to requests from kiosks and to process the information in a format that is compatible with the kiosk software.

Location and Distribution of Smart Traveler Kiosks

As noted previously, the incorporation of Smart Traveler for emergency response activities after the Northridge earthquake substantially changed the magnitude of the FOT. The change led to an increase in kiosks from 3 to 77, all within the earthquake impact area. Kiosks were installed during May, June, and July 1994. Sites were selected based on the density of foot traffic; hours of availability of the site to the public; security; and willingness of businesses in the area to take part in the project. Kiosk sites are diverse and include dense employment centers with high-rise offices; retail centers such as shopping malls, grocery stores, and multi-purpose high-volume retailers; transportation centers; hospitals; and public office buildings.

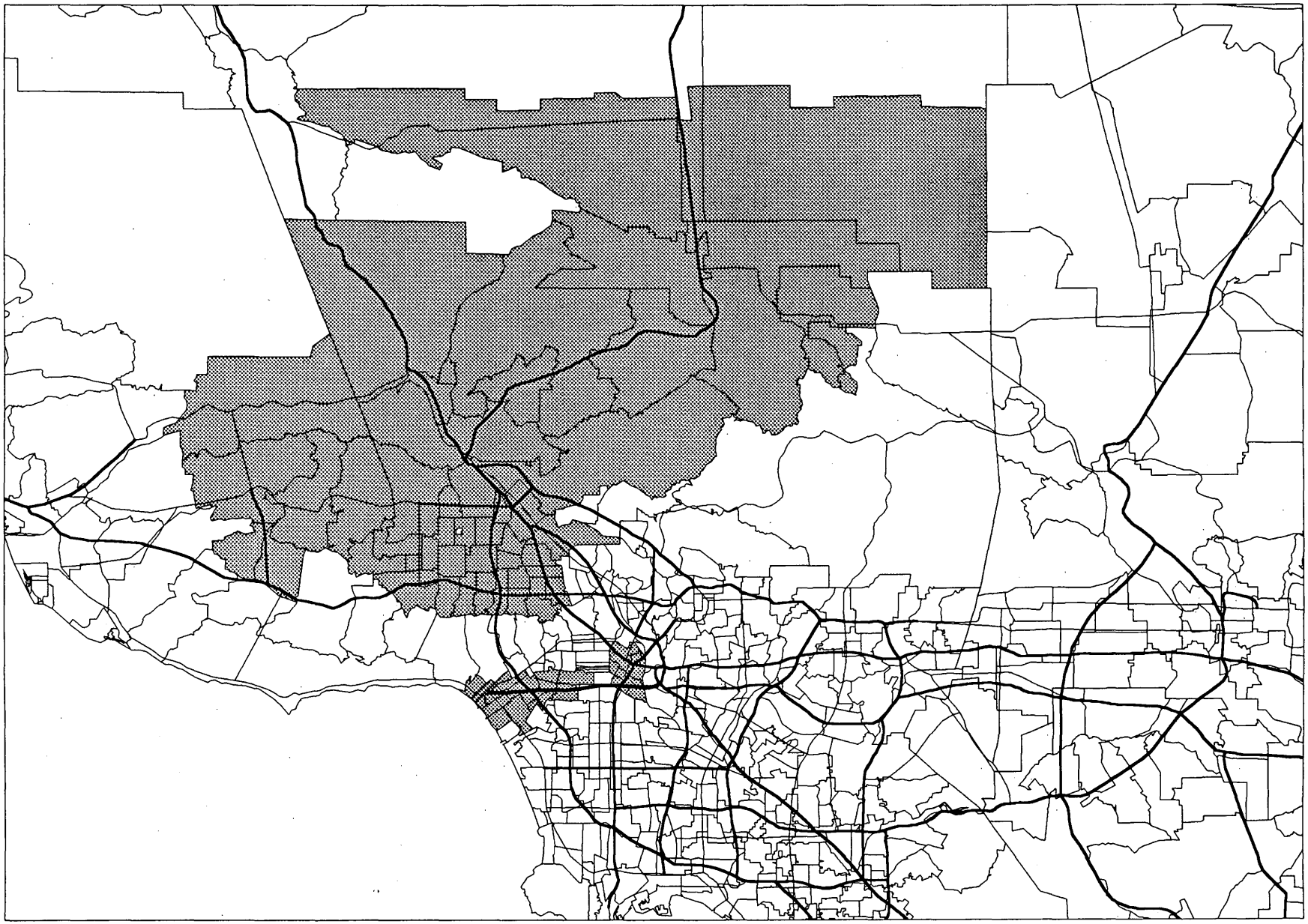

FIGURE 1 Smart Traveler emergency response target area.

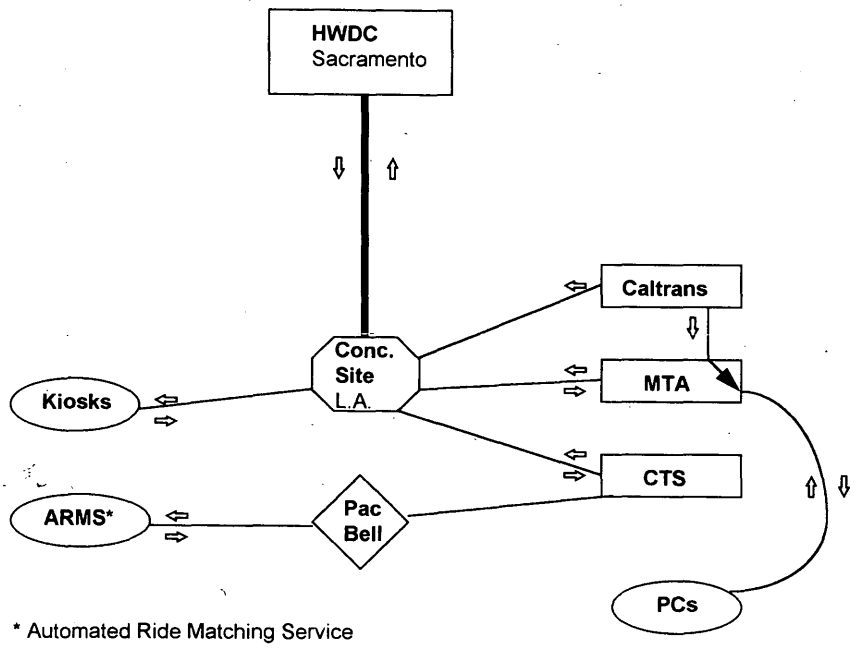

FIGURE 2 Smart Traveler system diagram.

USER RESPONSE

Data

User response is examined using the three data sources described in the following sections.

Automated Kiosk Data

Smart Traveler kiosks display a continuous video, called an "attract loop," when the machine is not being used. The attract loop is interrupted when the screen is touched, and a welcoming screen appears containing icons and text. The user can activate the menu options by touching the screen icons and text. Each touch is logged by task element and time by the kiosk computer software, creating a data log file. The data log file is stored on the kiosk's hard drive. The kiosk project design calls for an automated procedure for periodically polling each kiosk, transferring the files to the HWDC computer, and downloading them to tape or diskette. To date, this process has not been implemented, and the log files must be manually collected and downloaded.

In theory, the log files provide a complete record of use for each kiosk over the life of the project. Every use is tracked, making it possible to determine which menu items are accessed, in what order, and over what period of time. In practice, however, many problems have arisen in the collection and transfer of data files, thus limiting the analysis to a sample of data from 41 kiosks for which site field observations had been completed. The sample data spans August 17, 1994 through November 17, 1994; however, most sites do not have continuous data over the entire period. In addition, only basic measures of use are included. More comprehensive analysis will require additional programming and data file construction.

Site Field Observations of the Kiosks

Field observations were conducted to evaluate the kiosks based on their location. It was postulated that differences in levels of kiosk use can be expected to be a function of (a) the type of site where they are located; (b) the level of activity in the immediate vicinity of the kiosk; (c) the relative quality of the area where the kiosk stands; and (d) other factors, such as the maintenance and operating condition of the kiosk. It was anticipated that data collected at each site would help with the interpretation of the automated data and perhaps also help in clarifying the responses to the kiosk user surveys. The field observations were conducted in May, June, and July 1994, immediately after the installation of the kiosks.

Kiosk User Survey

The kiosk user surveys were conducted to determine user responses and perceptions of the kiosks. Because the survey was financed with funds from the Federal Emergency Management Agency (which must be used within 6 months), it had to be completed by July 17, 1994. To meet the deadline, surveys were distributed to users and observers of kiosks on site with pre-paid envelopes provided for survey returns. The kiosks had only recently been activated, and for many users this was their first opportunity to use Smart Traveler. It was therefore not possible to ask questions about repeated use or whether kiosk use had influenced travel decisions.

The following survey locations were selected based on relatively high estimated foot traffic:

1. An upscale food court serving two downtown high-rise office towers.
2. A high pedestrian traffic area in a large, up-scale suburban shopping mall.

3. A food court in a middle-market urban shopping mall.
4. A fourth location, a downtown plaza with a luxury hotel, multiple eating facilities, and a major anchor store, was added when the first location yielded a low response.

The range of locations selected for the distribution of the surveys helped increase the mix of survey respondents even though the method of distribution was non-random.

Analysis

Patterns of Kiosk Use

As noted previously, the kiosks selected for the preliminary analysis do not all have data for each of the 92 consecutive days. Explanations for missing data include:

1. Some kiosks are not available on weekends, and therefore no transactions were recorded;
2. Kiosks may not have been operating and were awaiting maintenance;
3. There were no users; and
4. The kiosk was used, but the data were not downloaded.

It was therefore necessary to correct for the number of days in which use was registered for each kiosk to calculate the average use per day.

A "use" is defined as each time the attract loop is interrupted and at least one menu item is selected. A "touch" is defined as each time the attract loop is interrupted. The attract loop is interrupted whenever the screen is touched, whether or not the kiosk is actually used. The number of touches is generally about 20 percent greater than the number of uses, indicating considerable touching of the screens without subsequent use. Table 1 gives the mean, standard deviation, and range for number of days, as well as average uses and touches per day for the 41 kiosks. The average number of days of data per kiosk is about 71, and the range is from 7 to the full 93 days.

Kiosk usage varies. Average use is about 20 per day, but ranges from a high of 49 to a low of 4. The kiosk with the highest average daily use is in Union Station in downtown Los Angeles. The remaining four of the top five locations are all in shopping malls. Of the five least-used kiosks, three are in office locations, one is in a grocery store, and one is at the Port of Los Angeles. Figure 3 illustrates daily usage at three locations: Union Station (Kiosk 11), Fox Hills Mall (Kiosk 1), and Warner Center (Kiosk 22). Fox Hills Mall is a shopping center on the west side of Los Angeles. Warner Center is a large suburban employment center about 67 km (40 mi) northwest of central Los Angeles. Not only is there a great difference between the highest and lowest performing kiosks, but each pattern shows a lot of day-to-day variability. A difference between weekends and weekdays also is discernible. Differences in use by day of the week were further analyzed. The single highest day is Saturday, followed by Sunday and Friday; the differences, however, are not statistically significant. Figure 3 also shows that the level of use at each kiosk appears to be fairly consistent over the 3-month period. Whether this will remain true for the whole sample of kiosks over a longer period of time still must be tested.

Kiosk location, as expected, is an important explanatory factor for differences in use. Kiosk locations were categorized as follows: shopping centers, grocery stores, discount stores, office, and other. The "other" category includes transportation facilities, hospitals, libraries, and other hard-to-classify locations. An analysis of variance was conducted using the location categories and a dummy variable for weekday/weekend. All effects are significant (F-statistic = 32.92, sig. = 0.000, N = 284). Location type accounts for most of the explained variation in average usage; time of week is significant both independently and jointly.

These differences are further illustrated in Table 2, which gives average daily usage for each type of location, by weekday and weekend. As expected, usage is higher on weekends than weekdays at the retail locations, and lower on weekends than weekdays at office locations. However, even on weekdays the office locations have less use than any other type of location. Because the "other" category includes such a diverse set of locations, no conclusions should be drawn about the patterns observed for this group.

A dummy variable regression was conducted to determine the relative effect of location and time of week on average daily usage. Results are given in Table 3. All variables except the grocery dummy are significant. The value of the constant is close to the actual sample mean value, and the R^2 is reasonable. Because all the coefficients are effectively in the same units, they can be interpreted directly. The shopping center and discount stores have much higher average usage than the sample as a whole. The equation predicts a use rate of about 35 users per day for these types of locations, which is about 75 percent higher than the sample average. Office locations have significantly lower than average usage by comparison. Time of week has a relatively weaker effect.

The kiosk menus can be accessed in English or Spanish. Spanish accounts for 17 percent of the kiosk daily average. It ranges from a low of 3 percent to a high of 55 percent. The two kiosk locations where Spanish was the menu of choice were both discount stores, ranked 7 and 8, respectively, in terms of average daily usage. Figure 4 shows average daily usage, in English and Spanish, for the 41 kiosks in rank order. The figure shows that the higher usage kiosks generally also have a greater than average proportion of Spanish use.

Explaining Patterns of Kiosk Use

If kiosk use were equally attractive to all passers-by, the level of usage would simply be determined by the level of pedestrian traffic

TABLE 1 Average Kiosk Use: 41 Kiosks

Total Kiosks	Use Days	Touches Per Day	Uses Per Day
Average	71.4	25.7	20.3
Std. deviation	21.4	15.6	12.8
Range	7 - 93	5 - 56	4 - 49

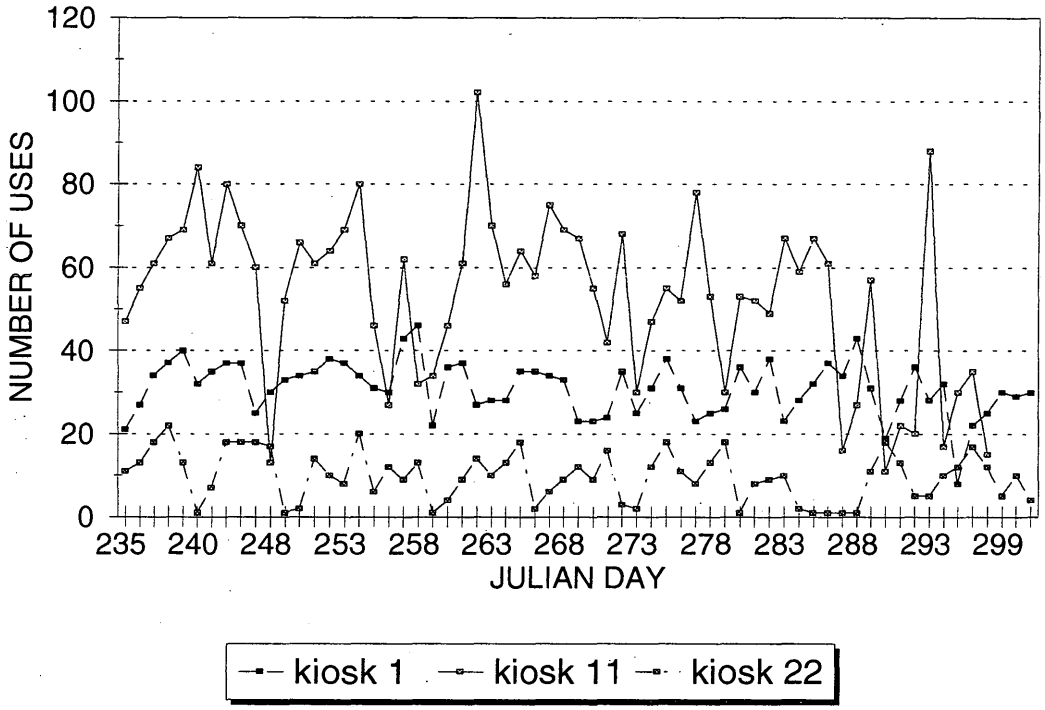

FIGURE 3 Total daily usage—high, average, and low examples.

in the area of the kiosk. Thus, one explanation for higher use at retail locations is that such locations get more pedestrian traffic. This hypothesis cannot be rigorously tested because average traffic per site is not known. However, the field site observations included a subjective assessment of activity levels both within the buildings where the kiosks are located (site area) and for the area within 15 to 20 ft of the kiosk (kiosk area). These subjective assessments are quite consistent with level of usage. Table 4 gives site and area activity ratings for the top five, middle five, and bottom five kiosk locations for which complete field information is available. Differences are found only among the bottom five locations, suggesting that other problems at these sites (for example frequent equipment breakdowns) may be deterring kiosk usage.

Total traffic is also a function of the number of hours per day that the kiosk is available. Retail locations have more hours per day of exposure than office complexes. The information on hours of operation was examined to determine whether it was possible to test directly for an effect. Hours of operation were found to be clearly established at retail sites and were found to have little variability within location categories; this was not the case at other sites. Many office complexes are open at night or on weekends, for example, yet little business activity takes place at these times. The conclusion is that the stated hours of operation are not a good indicator of pedestrian traffic at kiosk sites.

Another explanation for the patterns that were observed may be associated with the purposes for which the kiosks are being used. Traveler information systems are typically aimed at the work trip. The commuter is the stereotypical user (i.e., the motorist checks the freeway map before departure and the prospective transit user or carpooler searches for bus routes or carpool partners). Office locations thus seem appropriate. However, kiosk information may be more relevant and beneficial for non-routine trips (i.e., for tourists, for non-work destinations, or for new work trips). High usage at shopping centers and other retail locations suggests that kiosk use is more of a leisure time activity; travelers are gathering information about possible future trips, rather than for their current trip. Even for commuting, travelers may find it more convenient to learn about alternative modes during their off hours.

TABLE 2 Group Means, Average Daily Usage by Location and Time of Week

	Shopping Center	Grocery Store	Discount Store	Office	Other
Weekend	40.19 (20)	19.87 (14)	38.83 (8)	3.67 (19)	17.84 (18)
Weekday	30.8 (50)	14.27 (35)	29.07 (20)	10.87 (50)	16.98 (50)

( ) = number of observations in each group

TABLE 3 Regression Results, Dependent Variable = Average Daily Usage

Variable	Coeff	T-statistic	Sig. of T
Weekday	-2.77	2.032	.0431
Shopping Center	16.22	9.241	.000
Office	-8.34	4.737	.000
Discount Store	14.59	6.304	.000
Grocery Store	-1.39	.722	.471
Constant	19.24	12.004	.000
N = 284			
R^2 (adj) = .467			

Anecdotal evidence suggests that kiosks may not be a very convenient way to obtain real-time travel information. In a focus group discussion held at one of the office complexes, participants noted that while it would be helpful to access the freeway map to check on traffic conditions before leaving for home, it was not worth a trip to the kiosk to do so. Having access to the map in the office on their PC (another Smart Traveler element) was considered greatly superior.

Smart Traveler Kiosk Users

In this section, kiosk user survey results are discussed. Of the 1,785 surveys that were distributed to kiosk users and observers, 325 were returned, yielding a response rate of 18.2 percent. As noted previously, the kiosk user survey is not based on a random sample of kiosk users or potential kiosk users. In fact, there is no information on the population of kiosk users. The population of potential kiosk users is determined by the kiosk locations. Because few of the kiosks are located in areas that serve large numbers of lower income households, the authors did not expect to find many individuals from lower income households among the kiosk users. In addition, survey responses are generally correlated with education and income, and the surveys were written only in English, thus eliminating non-English speakers from the sample population.

It is therefore useful to compare characteristics of survey respondents with those of the Los Angeles County population. Table 5 gives gender, employment, education, and income level for survey

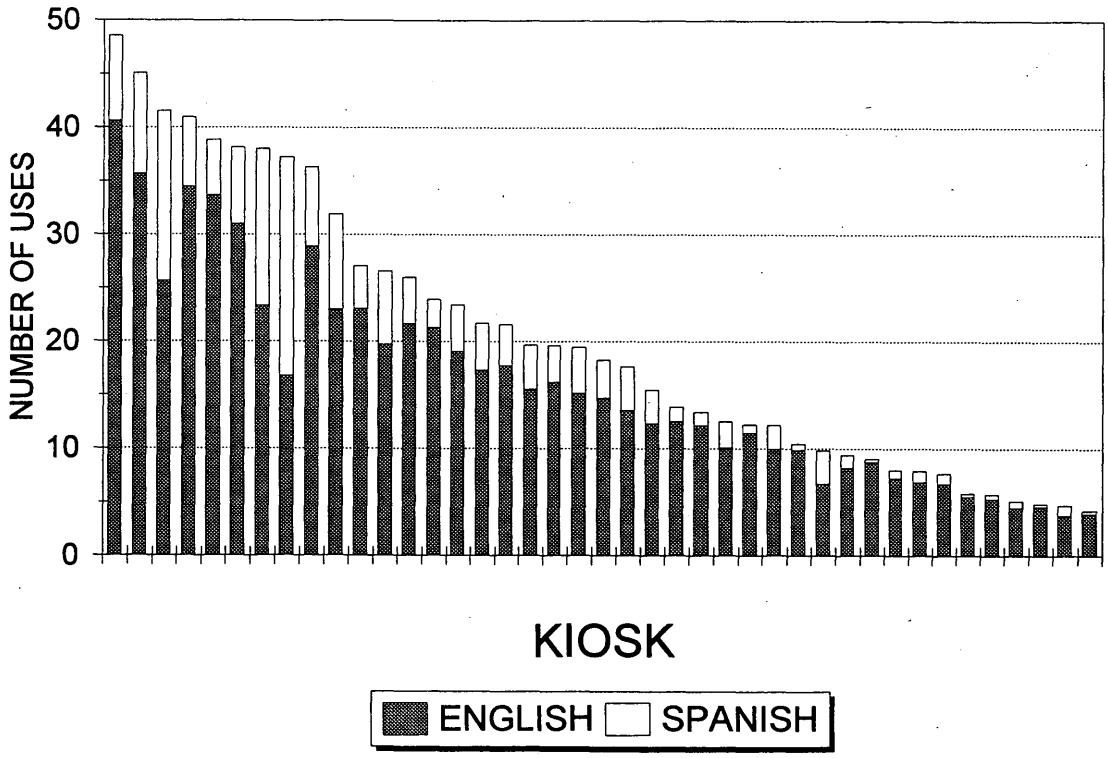

FIGURE 4 Kiosk average daily usage by language.

TABLE 4 Average Daily Usage, Site and Area Activity Level, Selected Kiosks

Category	Mean Use	Site Activity	Area Activity
High Usage			
Union Station	49	high	high
Shopping Mall	45	high	high
Shopping Mall	42	high	high
Shopping Mall	41	high	high
Shopping Mall	39	high	high
Moderate Usage			
Grocery store	20	high	high
Discount Store	19	high	high
Shopping Mall	18	medium	medium
Office building	18	medium	medium
Grocery store	15	medium	medium
Low Usage			
Hospital	6	high	low
Office complex	6	low	low
Office complex	5	medium	medium
Grocery store	5	low	low
Office complex	4	medium	medium

respondents and Los Angeles County population, taken from the 1990 Census. As expected, survey respondents differ from the general population. They are more likely to be employed, are more educated, and have higher household incomes. Although our survey respondents are not representative of the general population, they likely are representative of the potential users at the locations where the surveys were conducted.

Most survey respondents are employed. Of those employed, 83 percent work 40 hours or more per week. Vehicle access and ownership is extensive among those employed; 66 percent report household ownership of two or more vehicles, and 82 percent report having a vehicle available to drive to work. An additional 10 percent report having a vehicle available to drive "sometimes." Only 4 percent report having no household vehicles, and 8 percent do not have a vehicle to drive to work.

Given the level of vehicle access, there is a higher than expected use of public transit for the trip to work, as shown in Table 6, where work trip mode shares for survey respondents and Los Angeles County workers are listed. The drive-alone share is close to the regional average, and the carpool share is slightly lower. Survey respondents also have longer trips to work. Reported mean travel time and distance are 36 min and 33 km (19.8 mi), with medians of 30 min and 25 km (15 mi), respectively. The Los Angeles County 1990 census data give a mean travel time of 26.5 min. An annual survey of Los Angeles metropolitan area commuters reports a mean distance of 27.5 km (16.5 mi) and a travel time of 31 min in its 1994 survey report (*1*).

Further examination of the survey data revealed that the downtown mixed-use plaza survey site generated a large proportion of transit users: 24 of the 64 commuters (37 percent) used bus or rail transit, compared with 9 percent for commuters surveyed at the downtown office site, 7 percent at the urban mall, and none at the suburban mall. The mixed-use plaza is adjacent to the Metro subway line and is one of the most "transit accessible" locations downtown.

Most of the survey respondents had used the kiosk. (Surveyors were instructed to distribute surveys to people near the kiosk). Respondents were asked whether they were aware of a Smart Traveler kiosk at the survey location, and, if so, whether they used the kiosk. Eighty-one percent of the 325 respondents were aware of the kiosk, and of this group 84 percent had used the kiosk. Given that the kiosks are a new technology, it is possible that willingness to use the kiosks and perceptions about the kiosks are related to individual characteristics such as education level or gender. For example, people with college degrees may have had more exposure to computers than high school graduates. Cross-tabulations were conducted of

TABLE 5 Comparison of Los Angeles County and Survey Respondent Characteristics

Characteristic	LA County (%)	Survey (%)
Gender		
Male (%)	49.4	56.8
Female (%)	51.6	43.2
Employment		
Employed (%)	62.0	81.5
Not employed (%)	38.0	18.5
Education		
No high school diploma	30.0	0.3
High school diploma	21.0	19.0
Some college	27.0	41.0
College degree	22.0	40.0
Household Income		
Up to $34,999	50.0	32.0
$35,000 to $49,999	17.0	20.0
$50,000 to $99,999	25.0	38.0
$100,000 or more	8.0	10.0

TABLE 6 Usual Mode of Travel to Work, Percent Shares, Survey and Los Angeles County

Mode of Travel to work	Survey (%)	LA County (%)
Drive Alone	70.3	72.1
Carpool with others	12.6	16.0
Bus or train	13.8	6.6
Vanpool	0.8	n/a
Walk or bike	1.6	4.0
Other	0.8	1.3
N	249	4,002,048

kiosk awareness and use with education level, income level, and gender. Only the results on gender and use were significant. While 83 percent of men and 79 percent of women were aware of the kiosk, 90 percent of men and 75 percent of women who were aware of the kiosks actually used them (chi-square = 11.30, df = 1, n = 262, sig. = 0.001). The greater propensity of men to use the kiosk may indicate a greater interest or willingness to try out the kiosk, or it may reflect joint use. When couples use the kiosk, the man may be more likely to navigate the menus.

Respondents who had used the kiosks also were asked whether they found them easy to use, whether they would use the kiosk again, and whether they would encourage others to use the kiosk. The results are quite positive, as shown in Table 7, and suggest that the kiosks have been well designed for their intended use. Users also were asked whether they would recommend any improvements to the kiosks. Recommendations for improvement were made by 62 percent of those who had used the kiosk. Of those responding to the question, the most frequently mentioned comment was a need to make the kiosk quicker (24 percent). Cross-tabulations were conducted to determine whether perceptions about using the kiosks were associated with individual characteristics; none were found to be significant. Apparently kiosk users are self-selected; those not favorably inclined to using the kiosks do not even try to use them.

Users of Smart Traveler kiosks were then asked about the particular menu items they had used, whether they found the given item easy to use, and whether they found the information obtained useful. Table 8 gives the frequency of menu items requested in rank order. The freeway conditions map is the most commonly requested menu item, followed by MTA bus and train routes. Rideshare or transit videos and carpool information are requested much less frequently. The vast majority of those who used a given menu item found the information useful and stated that they would use the kiosk again to obtain such information.

Because the survey was conducted shortly after the kiosks were installed, no information is available on whether kiosk users acted on the information they received. To get some indication of their willingness to use the information, respondents were asked whether they used the menu just to see how the kiosk works, or whether they actually requested information. Most users (90 percent or more, depending on the item) were experimenting, but the majority also requested and obtained transit (83 percent) or rideshare (67 percent) information. In the case of the freeway conditions map, 71 percent of those who used the map stated that they would use it before starting a trip.

CONCLUSIONS

The survey results show a very positive response to the Smart Traveler kiosks, yet the average usage of the kiosks is quite low. An average of 20 per day is equivalent to fewer than two uses per hour, indicating that the kiosks are idle most of the time. One explanation

TABLE 7 Perceptions of the Kiosks

Question	Easy	Neither	Difficult	N
How easy or difficult did you find the Smart Traveler Kiosk to use?	79%	16%	5%	217
	Yes	No		
Would you use the Smart Traveler Kiosk again?	85%	15%		219
Would you encourage other people to use Smart Traveler Kiosk services?	88%	12%		214

TABLE 8 Menu Items Requested

ITEM	YES	NO	N
Did you request freeway conditions map?	83%	17%	218
Did you request MTA bus and train routes?	56	44	220
Did you request rideshare or MTA bus and train videos?	28	72	217
Did you request the carpool service?	26	74	217

for this is that the kiosks were installed with minimal marketing effort. Although the kiosks are large and have a continuous moving light display across the top denoting them as Smart Traveler kiosks, there is no information at the site that describes what they do. Passers-by must be curious enough to investigate what the kiosks offer. Given this absence of descriptive information, kiosk location becomes more important. For example, the Union Station kiosk is adjacent to the MTA ticket office, whereas the kiosk in Fox Hill Malls is in the food court. It is likely that the function of the kiosk is more obvious in Union Station, where people are traveling, purchasing tickets, etc., than in a shopping mall food court.

A second explanation is that some of the kiosks have maintenance problems and are often out of service. In a location where repeated use might be expected (e.g., office locations), frequent breakdowns would lower kiosk use. Although an analysis of maintenance data has just begun, preliminary indications are that reported maintenance problems are not related to low kiosk usage. The relationship may in fact be the reverse: breakdowns should be a function of usage. This issue will be explored further in subsequent research.

How Kiosks Are Used

Preliminary evidence suggests that kiosks are used either for non-routine trip planning or for trips to be made at some future time. Most of the kiosk usage takes place in non-work environments where people apparently have the time available to explore travel options. Use for future trip planning seems logical. For transit and rideshare information, kiosks have a considerable advantage over telephone inquiries. Users can obtain transit information for more than one trip, and they can obtain printouts of specific trip itineraries. Given the size and complexity of transit services in the Los Angeles area, printouts are likely very helpful. Easily accessible transit information may be particularly valuable to tourists. In the case of rideshare information, the kiosks provide another alternative outside of one's place of work to obtain such information. If the usage patterns observed are supported in further analysis, they imply a kiosk deployment strategy oriented to train stations and other major transportation facilities, large hotel complexes, and shopping malls.

There is also evidence of high levels of use in locations that serve lower income households (e.g., discount stores in the central city area). These locations also show above average use of the Spanish menus suggesting that the kiosks are being used by those who are most likely to be transit dependent. Kiosks may prove an effective means of reaching such groups, and should be considered for social welfare and employment offices as well as major shopping areas.

Finally, the lack of use of the kiosks in office locations suggests that kiosks are not an effective way to provide more "real-time" travel information. On-line services delivering information directly to users at their desks are likely to be the preferred media interface. This is an element of the Smart Traveler Program that is still in development. (The Caltrans freeway traffic conditions map has just become available via the World Wide Web. This service is not part of Smart Traveler).

Further Kiosk Evaluation

The automated kiosk data are a rich source of information. Data for all 77 kiosks are still being processed. Subsequent analysis will examine usage across the different menus, as well as by location and over time. Such analysis should provide a deeper understanding of the extent to which the kiosks are being used for different purposes (e.g., planning of transit trips). It also should allow researchers to investigate use by time of day and test hypotheses about non-routine trip planning. The field site location data for the kiosks will be reviewed in detail to determine whether particular location characteristics (e.g., proximity to other machines, such as ATMs and lottery ticket dispensers) encourage or inhibit kiosk use. Finally, usage data will be supplemented with cost and maintenance data to establish how well the system has performed, and to make an overall assessment of the effectiveness of kiosks in providing traveler information.

ACKNOWLEDGMENT

This research was supported by the California Department of Transportation CAPTS branch and by the Federal Emergency Management Agency. In addition to the authors, the research team includes Randolph Hall, School of Engineering, University of Southern California; and graduate research assistants David Huang, Xiaobo Huang, Amer Qureshi, and Benjamin Wang.

REFERENCE

1. Commuter Transportation Services, Inc. *State of the Commute, 1994.* Commuter Transportation Services, Los Angeles, 1994.

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

Ten Lessons for Automated Highway System Design: A Comparable Systems Analysis

CAREN LEVINE AND DOUG FUNKE

A comparable systems analysis was performed to identify issues to be considered in the development of an automated highway system (AHS). In the analysis, many comparable systems from four main categories: highway-based systems, vehicle-driver systems, other transportation systems, and nontransportation related systems were considered. This article summarizes the results of this comprehensive analysis. Each of the comparable systems studied shares important features with AHS and is therefore relevant to different aspects of AHS. For example, some of the comparable systems are similar to AHS in that they involve the introduction of a new technology, whereas others have introduced new ways of doing old things irrespective of the technologies involved. Some involved large amounts of financing, and some introduced automation to traditionally manual tasks. The comparable systems have been analyzed to identify relevant lessons for the introduction of AHS from their own applicable perspective. The results from the analysis are synthesized within 10 overreaching issues that are relevant for AHS development. Each issue is supported by evidence from several of the comparable systems. This article presents and defends each issue based on the past experiences from comparable systems.

Our society has undergone extensive changes over the past century as a result of new systems and technology applications to public, business, and private activities. These changes have required new ways of thinking, behaving, and performing everyday tasks affecting large portions of the American public. People's fears and reluctance to change have had to be overcome (e.g., acceptance of commercial flight), new public policies at all levels of government have had to be reconciled with private interests across the spectrum of American society [e.g., introduction of high-occupancy vehicle (HOV) lanes], and innovative approaches for introducing new technology within preexisting systems and to old technology users (e.g., the introduction of the automobile) have had to be developed.

The objective of this comparable systems analysis was to draw relevant lessons about designing and managing technologically based change from past experience relevant to AHS. These AHS comparisons and similarities may be associated with specific aspects of AHS, or bear relevance across broad ranges of design, development, and deployment issues.

For each comparable system, an analysis was performed to identify the relevant lessons learned for AHS. This process involved review of literature, interview of subject matter experts, and application of expertise within the study team.

Calspan Advanced Technology Center, Buffalo, N.Y. 14225.

DISCUSSION

The following paragraphs describe 10 of the overreaching issues identified during this comparable systems analysis. Evidence from selected comparable systems supporting each issue is also discussed.

Issue 1. The Public Must Perceive Overall Benefits

For a new technology to successfully replace an existing technology, the new system must offer clear and obvious advantages and benefits over the older system. If these benefits are nonexistent or nonevident, potential users will likely be unwilling to give up the preexisting trusted system for the newer system, especially if the changeover involves significant costs (e.g., money to purchase the new system, time to learn new procedures, license fees). AHS design and deployment should proceed in ways that will make the benefits obvious to all potential users.

Evidence from the comparable systems studied supports this conclusion. Experience has shown that people's willingness to purchase new systems and services has been enhanced when these systems and services offer clear benefits in convenience, safety, or cost effectiveness. Findings from the comparable systems studied provide the following examples.

Toll Roads

When the first toll roads were constructed, many believed that they would be unable to generate enough profits and would eventually fail. There were many local roads (already in existence) that could be utilized for no charge, which allowed drivers to travel between the same origins and destinations as the toll roads. Nonetheless, the toll roads were built as limited-access highways, intended for long-distance travel between cities. The toll roads allowed drivers to travel at faster speeds and more safely than they could travel on local roadways. After implementation, toll road use significantly surpassed expectations. Drivers found that the toll was well worth the very significant time savings and improved safety that the toll road offered (*1*).

Video Cassette Recorders (VCRs)

Similarly, people have been unwilling to pay for new systems and services that did not offer significant advantages over existing systems.

RCA discontinued work on video cassette recorders (VCR) in 1978, to concentrate on the development of the videodisk. The videodisk was released to the public in 1980, almost 5 years after the introduction of the VCR. By this time, hundreds of thousands of VCRs, which could both play and record, had been sold in the United States. The videodisk offered the same service to the public (the ability to watch movies at home), but offered fewer features than the VCR (e.g., the videodisk player could not record television programs). Few people were willing to invest in the new technology, and the videodisk has not become as successful in the marketplace as the VCR (*2*).

Our study of comparable systems provides evidence that for AHS to succeed, it must offer the public some obvious advantages and benefits over the conventional highway system. These benefits should be highlighted in the marketing of AHS.

Issue 2. Safety and Reliability Must Be Clearly Demonstrated

There are many factors that may contribute to the public acceptance of AHS. Some important factors include: cost relative to other transportation modes, convenience and ease of use, ability to match users' origins and destinations, impact on pollution, and obviousness of fail-safe features. The safety and reliability of the new technologies used by AHS must also be clearly demonstrated before public acceptance can be achieved.

Any new technology must be proven to be safe and reliable before the general public is willing to accept and use it. Evidence from the comparable systems studied has shown that even systems that have a reputation for safe operation may face a loss of users if a safety incident does occur. Systems that have a reputation for safety problems have a very difficult time achieving public acceptance.

For these reasons, the safety and reliability of AHS will need to be proven to the public before it will be accepted. The current highway system, although not perfect, is considered by the public to be relatively safe, and experience has shown that people do not completely trust computers and automation. The public is unlikely to believe that AHS is safe until they have observed it for themselves, and the public is also unlikely to accept an automated system that has experienced failures or safety problems. AHS must be demonstrated to operate as expected under all circumstances. If the initial users of AHS experience problems or delays, they will be discouraged from continuing to use the system, and are likely to be reluctant to use the system even after the problems are corrected. This has been the experience of comparable systems which share common attributes with AHS.

It is therefore necessary to demonstrate the safety and reliability of AHS before full implementation. However, public demonstrations that raise safety concerns may do more harm than good, and even prevent the system from ever being accepted. This implies that AHS developers should perform extensive testing before a prototype is demonstrated to the public. The comparable systems analysis has provided many examples to show that demonstrations of safety and reliability can help a new technology to be successful, and that demonstrations that do not accomplish this can have an opposite effect.

Elevators

Elevator-like lifting devices were used by almost all ancient societies. The primary function of these early elevators was to lift heavy objects, although they were occasionally used for transporting people. The modern elevator was developed in the late 1700s, and until 1854, was primarily used to haul freight because of serious concerns about passenger safety. In 1854, Elisha Otis demonstrated his safety elevator at the Crystal Palace Exhibition in New York City. As the public gathered, Otis stood on an elevator platform that was heavily loaded with freight and raised himself 12.2 m (40 ft) in the air. Otis' assistant then cut the rope supporting the platform. The platform jerked downward slightly, then stopped. The safety mechanism successfully engaged and stopped the elevator from falling. The public was shocked by the demonstration, but began to believe that such lifts could be safe for passengers. Soon after these demonstrations, orders began coming in for the "safety elevator," and they began to be installed in buildings worldwide for the transportation of people (*3*).

Supersonic Transport (SST)

Our comparable systems analysis also found examples of detrimental effects from unsuccessful demonstrations. When demonstrations experience problems or result in failure, a negative view of the technology or system can be created. The resulting concerns and negative image can persist even after the problems are corrected.

Advocates for the SST believed that supersonic transportation had no serious drawbacks, because supersonic flight allowed faster air travel than conventional subsonic flight. However, opposition groups identified many problems associated with supersonic flight, including the noise associated with sonic booms. In 1964, during the development phase of the SST, the FAA decided to conduct a series of supersonic tests over Oklahoma City. They wanted to prove to the public that the typical sonic boom was not overly annoying, and that it caused no physical damage on the ground. However, after the tests began, the FAA received over 15,000 complaints about the noise from the booms, and 5,000 claims for damage caused by the tests. The negative publicity resulting from the Oklahoma City tests was never really overcome (*4,5*). Public opposition to SST grew immensely, and the project was eventually canceled.

From the study of these systems, we have concluded that AHS must demonstrate its reliability and safety to the general public before the system can be fully implemented. People must believe that AHS is as safe, or safer, than the traditional highway system. The best way for them to gain this belief is to see the system in operation. However, care must be taken not to provide public demonstrations too early in the design phase, because one unsuccessful demonstration can seriously hinder eventual acceptance by the public.

Issue 3. Secure Long-Term and Continuous Financial Support Deployment

For the long-term success of AHS, it is important to ensure that funding for the project is sufficient and guaranteed. If the funding is not sufficient, it may be difficult to raise the additional funds at a later date. If the funds are not guaranteed, they may be cut at any

time. This issue is especially important if AHS must depend on funds from the federal (or state and local) government.

If funding for AHS is derived from government sources, AHS will have to compete with many other projects. The allocation of funds must be approved by many committees, in both the Senate and the House of Representatives, and possibly even by the President of the United States. Therefore, if sufficient funding is not obtained up front, it may be a long, difficult process to obtain additional funding midway through the project. Also, if complete funding is not guaranteed, AHS may be subject to the priorities of a new administration following elections.

Evidence also supports the benefits of pay-as-you-go financing, instead of borrowing funds for a project. Under a pay-as-you-go approach revenues are raised especially for the project and used to fund the project on an as-needed, pay-as-you-go, basis. The funds may be raised from specific and dedicated taxes, private contributions, or user fees. This approach allows the project to control its own resources, and to ensure that the correct funding level is achieved. Borrowing funds for a project usually makes the project dependent on another agency (i.e., the loaning agency) for its resources and may limit the amount of available funds. There are many examples from the comparable systems analysis that provide support for this issue.

Interstate Highway System

In 1955, a bill was presented in Congress that would raise $25 billion through the sale of bonds for the development of the Interstate highway system. However, this bill was defeated because the Senate was unwilling to borrow funds for this project. The Federal Highway Act of 1956, for the construction of the Interstate highway system, promised something for everyone and established a trust fund through which highway users would pay for the Interstate system (*1,6,7*). The premise of the trust fund was that revenue raised from highway users (through gasoline taxes, user fees, etc.) would be placed in a separate account earmarked for the ongoing maintenance of the highway system. The money from this fund could not be diverted to other projects, nor could the Interstate system obtain funding above the level available in the trust fund account. The trust fund (although controversial) allowed the entire Interstate highway system to be constructed and maintained without borrowing funds, whereas the original approach of leveraged funding met with failure because of political opposition.

Downtown People Movers (DPM) Project

Because of the high level of uncertainty for eventual success, and the complexity involved in implementing (and maintaining) a public transportation system, the federal government is usually the only possible customer for the construction of a new people mover system. Unfortunately, the government is often an unstable customer for such expensive and innovative technologies. Priorities change rapidly in the federal government, and often projects must continuously request funding (e.g., annually or at the start of every project phase). At any time, the funding request may be turned down, leaving the project unfinished and with no other financial support. Local governments are reluctant to contribute to the development of these systems unless the financial risks are underwritten by the federal government (*8*). One specific example is the DPM project. The DPM project was established in 1976 to help six cities design and construct fully automated urban transportation systems. In 1981, the federal government cut off funding to the DPM Project before any of the systems could be built. Only three of the six cities (Miami, Detroit, and Jacksonville) managed to continue their projects without federal funding. Each of these cities had received firm commitments for funding from state and local governments, as well as private sources, early in the project (*9*).

To ensure the success of AHS in the United States, it is necessary to secure the necessary funding before the project has begun. This will prevent AHS from having to participate in constant battles over project financing, and will help the project to be completed according to schedule.

Issue 4. High-Level Support Enhances Success of Innovative, Infrastructure-Intensive Projects

The success of many large-scale projects has been facilitated through the commitment of high-ranking officials from government or industry who were willing to work hard to ensure the success of the project. AHS would benefit from such an individual (or group) to help secure the necessary financing and support and to help maintain enthusiasm for the project during all stages of design and implementation.

The importance of a strong proponent for large projects was evident in many of the systems we studied. Projects without one or more strong supporters have often faced great difficulties at all stages, whereas projects with strong and influential supporters have generally been more successful. Influential supporters for such projects are important throughout all stages of project planning, development, implementation, and during initial stages of operation. Projects with such support are also more readily accepted by the American public.

Automobile Proponents

A number of organizations including lobbying groups for the automobile supply industry, the American Automobile Association, the National Association of Automobile Manufacturers, and the Automobile Club of America, were established to further the rights of motorists and the automobile industry (*10,11*). These groups were willing to spend considerable time and money convincing the government to support the automobile industry, and to provide the necessary funding to improve highways (and other services) for drivers. Although some groups were formed to oppose the automobile, these groups never wielded the power of the automobile proponents. The influence and support from these groups were important factors in the development of the highway system in the United States.

Airplane Travel

In 1932, Franklin D. Roosevelt used an airplane to travel from New York to Chicago to accept the Democratic nomination for the pres-

idency of the United States. His flight gave the public confidence in flying, and helped call attention to the speed, availability, and safety of air transport as a routine mode of transportation (*12*). Similarly, when the restriction barring most American movie stars from flying was deleted from film contracts in 1935, many film stars began to fly regularly, again demonstrating to the public that flying did not pose a significant risk. In fact, some European airlines even published a monthly list of all of the celebrities who had flown on their airline that month (*12,13*). Both of these examples illustrate the influence a celebrity endorsement can have on the success of a system. The public often perceives a system to be safe if someone important or famous is willing to use it.

These experiences provide evidence that AHS will be best served if strong supporters can be found to help maintain support and enthusiasm for the project. Such support may help in obtaining funding for the project, in gaining necessary political approval, and in persuading the public to accept AHS.

Issue 5. Evolutionary Development Is Recommended

An evolutionary approach to the development and implementation of AHS is recommended based on the experience of several large-scale public systems studied during this project. An evolutionary approach will allow the public to gradually learn to use and accept AHS, building from their current experience with Interstate highways. Using this approach, the public will be able to experience AHS on a small scale and to develop confidence in its safety and reliability, before large-scale implementation is introduced.

In addition, the evolutionary approach will make it possible to begin implementation of AHS sooner than if the entire system was to be implemented all at once. A limited capability AHS, requiring fewer changes in highway infrastructure and vehicle technology, could be implemented far earlier than a fully capable system. For example, a mixed-traffic approach to AHS that uses existing highways (in which automatic and manual vehicles share the same lanes) requires less development and preparation time than a system requiring dedicated lanes for automated traffic. As AHS popularity grows, the system may then naturally evolve into a more advanced AHS system.

An evolutionary approach also provides an opportunity to incorporate user feedback into subsequent system designs. User comments and opinions about the early stages of implementation can be collected and applied to later stages. This will ensure that the latter stages of AHS design meet the needs and desires of users, and will help to ensure user acceptance of the final system. Many other technologies have taken an evolutionary approach, and have been quite successful.

High-Occupancy Vehicle (HOV) Lanes

One reason for the success of high occupancy vehicle (HOV) lanes is that they are relatively quick and inexpensive to implement, most often as retrofits into medians or by using existing lanes of congested freeways. They become an evolutionary enhancement to, and an integral part of, the larger highway system. The rules and behaviors required for HOV use differ only slightly from non-HOV lanes, and drivers have been able to adapt easily to their use and to see the immediate improvements in traffic flow resulting from HOV implementation.

Computer Technology

The rapid growth of computer technology has been realized through many small incremental steps. Each improvement has been made in a way that preserves compatibility between the newest technology and the previous generation of technology, facilitating the acceptance and improving the likelihood of success for that improvement. As each new generation of computers has been introduced, it has improved on features of the older systems, better meeting the needs of the users. Computer users have been willing to upgrade their systems to obtain the better technology. This evolutionary approach has made possible many revolutionary changes in the computer industry. Users have been able to slowly become adjusted to computer technology, and to provide feedback to hardware and software developers about what is lacking from the current generation of technology.

Based on the study of these comparable systems, arguments for an evolutionary development of AHS can be made. First, it allows people to become adjusted to the new technology at a slower pace, and prevents them from being totally overwhelmed by the implementation of a large, radically new system. An evolutionary approach also allows the AHS system to be installed in stages, which does not necessitate the system being fully developed before implementation begins. Finally, an evolutionary approach allows the AHS to be changed as necessary to meet the needs and expectations of the users based on initial experience with the system, and allows new technologies to be incorporated as they are developed.

Issue 6. AHS Must Be Designed for Integration Within the Overall Transportation System

Clearly, AHS will be one of many transportation systems available in the United States. It should, therefore, be designed as an integrated part of the overall transportation system. As an integral component of the U.S. transportation system, rather than as an independent competing mode, a large and stable user base will be encouraged.

There may be regions in which geographic or traffic conditions favor AHS, whereas other areas may be less favorable. On the one hand, this will make it possible to select locations for AHS demonstrations, where AHS can provide significant benefits within the larger transportation system. It also will help guide the planning of AHS evolution and system expansion. On the other hand, it will be difficult to gain political support from legislators representing areas with little to gain from AHS.

To meet the goal of integration within the larger transportation system, AHS should be designed to be compatible with the existing highway system. For example, AHS should minimize the number of highway signs and markers that must be redefined. Such incompatibilities would make it difficult for people to adjust to AHS and could cause confusion on manual trips.

In addition, it is important that AHS be standardized throughout the United States and even worldwide. Standardization helps to prevent user confusion, and will support long-distance AHS trips in the future. Designing a standardized system from the early stages will prevent expensive system alterations in the future. Support for this conclusion was found from the study of several comparable systems.

Interstate Highway System

The Interstate highway system has often been implemented in urban areas without high-level system planning. In these cases, the Interstate highway system was perceived as a stand-alone system, not part of a larger interdependent roadway network. The impact of the Interstate on local roadways was not considered during the design stage and was only determined after implementation. This has led to many unforeseen problems of congestion in areas in which local roads were not capable of adequately serving the Interstate highways (*6,7*). In addition, the designers of the Interstate highway system did not anticipate the societal changes that would accompany such a system. They never imagined that the Interstate system could ever become incapable of meeting traffic demands. The highways were designed for long-distance travel, not the commuter traffic that plagues many Interstate highways today. In many areas, it is not possible to expand the highways to meet current traffic demands. Traffic planners in these areas are struggling to find alternate solutions to congestion problems. This lesson emphasizes the need to consider, in advance, the impact of AHS on non-AHS roads and facilities and on the larger transportation system.

Regional Railroads

Each regional railroad in the United States was originally built according to its own specifications for track gauge (the distance between the tracks). The trains were designed to operate only on one specific gauge and thus could not travel beyond the limits of their own system. Over time, the railroads developed methods to permit limited access of trains across a few different gauges. Eventually, all of the railroads standardized their track gauge, requiring the relocation of thousands of miles of railroad tracks to establish an efficient, national railroad system (*14*). Similarly, regional differences in AHS design could require serious rework to permit interoperability in the future.

The study of these comparable systems provides evidence that AHS should be developed using an integrated design approach that considers all aspects of the larger transportation system. To be most effective, AHS should be one integrated part of a larger system. It is also important that AHS developers consider the long-term requirements, and design the system accordingly.

Issue 7. Cost and Time Estimates Must Be Accurately Determined

It is important for the AHS project to maintain a good public image throughout the design, development, and implementation phases. Any negative publicity associated with AHS will simply make the system more difficult for the public to accept and will reduce the number of users on the system. Cost and schedule "bad news" can reduce public acceptance of the system, even when the shortfalls are because of estimation errors, rather than more serious system problems.

The public (and media) often closely monitor the progress of large-scale projects, such as the development of AHS. Projects are expected to be well managed and to be kept under tight control in terms of costs and scheduling.

For this reason, AHS developers must carefully make realistic estimates concerning the amount of time the system will take to implement, and the amount of money it will cost to complete. Neither the financial backers nor the general public is pleased when a project requires sudden increases in financial support when it is halfway through, or when the project takes significantly longer to complete than predicted. This is especially true for projects financed by public funds. Evidence for this issue has been found in the study of the following comparable systems.

Morgantown Personal Rapid Transit (PRT)

The PRT project was undertaken with an $18 million grant, based on a rough estimate by university officials. In addition, The Urban Mass Transportation Administration, the sponsoring agency, set unrealistic deadlines for the project, which led to incomplete analysis during the initial design phases. Because of the short deadline imposed on the project, the design team attempted to design, develop, and construct the system simultaneously, causing many system deficiencies and problems that required redesign work, increased development costs, and schedule slippage. The extensive cost overruns led to a poor public perception of the Morgantown project, and hindered support for other PRT projects.

Denver International Airport (DIA)

During the 1980s, when the idea for the construction of the new DIA was first proposed, the project was widely supported by the public in Denver. Plans for the new airport included revolutionary new technology, and the airport was supposed to be the most advanced and modern airport in the world. The scheduled opening for the airport was established as October 1994. However, because early progress on the airport was far ahead of schedule, and system promoters were trying to improve public perception of the project, the completion date for the airport was moved up to October 1993. Unfortunately, the opening date for the airport was postponed four times, and the airport did not officially open until February 1995. During the delay, public support for the project wavered. The public began to question whether Denver really needed the new airport in the first place, and whether the money spent on the project should have been spent on other, more important, social and economic programs. A more realistic schedule would have avoided at least some of the devastating negative publicity surrounding DIA.

From the study of these comparable systems, the importance of realistic and achievable budgets and schedules is reinforced. The public expects large-scale projects, such as AHS, to be well managed and to be kept under tight control. Negative publicity resulting from budget overruns or schedule slippage may help sour the public about the system. They may be less willing to accept the system or even to support future AHS projects.

Issue 8. Consortiums of Private and Public Agencies Ensure Long-Term Success

A consortium approach to AHS development can help to ensure that the AHS system is successfully implemented. The consortium approach allows the project to benefit from a wide range of expertise and perspectives, and to share the costs involved with system

implementation. Even more importantly, cooperation among the various industries and organizations interested in AHS will facilitate efficient and effective designs that can be supported by products and services developed independently, yet which must operate within a common infrastructure. This recommendation is based on the study of several large-scale systems developed for implementation using a common infrastructure.

It can be noted that the consortium approach differs significantly from the alternative approach of allowing many different companies to bid for the privilege of being the single system designer. This competitive approach frequently involves prototype competition in which prototype systems, designed and built by competing bidders, are judged to see which is the best. This approach has been applied for the development of many large-scale systems. The competitive environment created by this approach encourages the companies involved to invest significant fiscal and sweat equity and are thought to provide the government the best value. However, the systems developed with this approach, although often successful, are usually not oriented for public use and are not dependent on public infrastructure. Also, they do not necessarily promote compatibility across a wide range of independently developed products and services.

In contrast, under the consortium approach there is not a single winner selected on the basis of a parameter-based process. Winners and losers are sorted out in the marketplace. The motivation for investment, participation in the consortium, and diligence to the task comes from increased market share potential that results from design participation. A consortium approach to system development has proven effective in many situations similar to AHS (i.e., large, market-driven systems).

Anytime Teller Machine (ATM) Networks

To compete with Citibank's ATM network, other New York City-area banks joined in a cooperative venture to create an ATM network, which they called The New York Cash Exchange (NYCE). The network approach allowed the participating banks to share start-up and development costs, as well as operating and equipment costs. By pulling together, the member banks were able to develop a more integrated and capable system for providing banking services. This cooperative approach has been very successful. The NYCE system has grown to become a national ATM network. Also, other cooperative banking networks have been developed based on the consortium approach to system development (*15*).

International Air Transport Association (IATA)

In 1944, the airlines and associated government agencies from all nations (except Russia) joined to form the IATA. The IATA established international standards for safety, navigational controls, air maps, and even the international setting of air fares (*12*). This organization helped to make international air travel safe and efficient. These accomplishments would have been very difficult without widespread cooperation.

The consortium approach has proved successful in developing large-scale projects like AHS. The cooperation between various industries and agencies (both public and private) helps ensure that the target system will be well designed, and that the various subsystems will work together effectively. Experience has shown that the use of consortiums can reduce the total cost of system implementation.

Issue 9. Keep the General Public Educated and Informed Throughout Planning, Design, and Development

Our study of comparable systems has shown that it is important for AHS to have public support throughout the development process. Further, the best way to obtain this support is to keep the public well informed about the project, and to provide them with as much information as they require. AHS developers and supporters should make the public aware of the benefits of AHS, and immediately deal with any criticisms or concerns raised. In addition to maintaining support for the program, this will help attract users to the system by allowing them to understand how the system works and the benefits it offers.

It is possible that organized groups may oppose AHS. Our study of comparable systems has led us to conclude that it is necessary to respond immediately to opposition groups, and to address the concerns they may raise. If these groups are able to operate with no response from AHS developers, the AHS project could be hindered, possibly seriously. Public resistance to large-scale projects can be very powerful, and experience on projects similar to AHS has shown that public education can help avoid such resistance.

Ramp Metering

Some locales have rejected ramp metering because of public resistance. Often, perceptions held by some members of the public (e.g., that the metering of ramps will cause unequal access to the highways) has led to increased resistance. When programs of public education are associated with ramp-metering projects, they have been much more successful in avoiding such resistance. This approach has usually facilitated a smooth and orderly implementation of the ramp metering projects.

SST

Throughout the first half of the 1960s, important SST activities and decisions were more or less contained within government agencies. Toward the end of the 1960s, the SST slowly emerged as a matter of public concern. In 1966, the Citizens League Against the Sonic Boom was established, and eventually joined with 13 other organizations to form one unified consortium against SST. The Coalition Against the SST succeeded in lobbying Congress to vote against additional funding for the SST program. In 1971, funding was cut off for the SST, after $623 million had already been spent (*4,5*).

We have concluded, based on our study of comparable systems, that public support for large projects like AHS is critical for their success. Developers cannot expect strong public support unless the public has been involved (or at least well informed) throughout the development process. If AHS does not deal immediately with public questions as they arise, the public may develop serious concerns about the project. Concerned individuals may form organized opposition groups. Such opposition groups can be extremely powerful, especially when the project is dependent on public funding.

In addition to avoiding the potential for public opposition, our research has found that full public disclosure and education is important for avoiding liability problems. According to the definitions of the legal system in the United States, the definition of a defective product and dangerous conditions is based on the perceptions of the

general public. It is necessary to inform and educate the public about AHS operation and limitations to help mitigate legal responsibility.

Issue 10. Marketability Is Influenced by Design and Economic Factors

AHS will be just one of several options for travelers. Its design and pricing approaches will affect its potential market base. Innovative approaches to AHS pricing, and the particular sales approaches used, can increase the potential achievable market. Also, the development of the AHS market can be facilitated by "piggybacking" on other markets (e.g., market AHS to those currently using ETTM systems). In planning for AHS marketing, it will also be important to consider prevailing economic conditions.

Obviously, the more people to whom AHS is marketable, the more potential users there are for the system. Limiting the potential market for AHS could exclude a large number of otherwise potential users and result in poor public perception of AHS. To maximize the potential for AHS success, it is best to open up the system to as many categories of users as possible. This approach of seeking the broadest possible market is recommended on the basis of the study of several comparable systems.

Automobiles

Initially, automobiles were marketed and sold as toys for the rich. With the introduction of the mass-produced Model-T, Henry Ford popularized the idea that owning a car was not a luxury, but was something to which every family could aspire. The Model-T was a basic car, offering few luxuries for the occupants. This concept rev-

TABLE 1 Issues Supported by Comparable Systems Analysis

	Issue	Supporting Comparable Systems Studied
1	The public must perceive overall benefits of AHS. Approach AHS design and deployment in ways that will make these benefits obvious.	• Automated teller machines • Automobiles • Commercial flight • Domestic appliances • Electric streetcars • HOV lanes • Office automation • Ramp Metering • Supersonic transport (SST) • Toll roads and limited access highways • Video cassette recorders
2	The safety and reliability of AHS must be clearly demonstrated before public acceptance can be ensured and successful commercial deployment made possible.	• Automated guide way transit systems • Automobiles • Commercial flight • Denver International Airport • Elevators • Interurbans • Maglev rail systems • Supersonic transport (SST)
3	Secure long-term and continuous financial support for AHS deployment. Funding must be sufficient, specific to the goals of AHS, and continuous. Pay-as-you-go financing is preferable to borrowing.	• Automated guideway transit systems • Interstate highway system • Interurbans • Maglev rail systems • Supersonic transport (SST)
4	The success of innovative, infrastructure-intensive projects is greatly enhanced by high level support from influential persons in government and industry. Without high level support, projects like AHS are likely to suffer failures.	• Automobiles • Automated guideway transit • Commercial flight • Maglev rail systems • Supersonic transport (SST) • Toll roads and limited access highways • Typewriters
5	Evolutionary development of AHS is recommended. This will provide for incremental development and deployment, allow safety and reliability to be demonstrated on a small scale before system level integration is attempted, and provide a gradual approach to achieving public acceptance.	• Automated guideway transit • Automated teller machines • Air traffic control (ATC) system • Automobiles • Chunnel • Commercial flight • Domestic appliances • Elevators • HOV lanes • Personal computers • Railroads • Ramp metering

(continued on next page)

TABLE 1 *(continued)*

	Issue	Supporting Comparable Systems Studied
6	AHS must be designed for integration within the overall transportation system in the United States and worldwide. AHS components should be standardized for all AHS applications, and should be as compatible as possible with existing conventions.	• Automated guideway transit • Domestic appliances • HOV lanes • Interstate highway system • Railroads • Ramp metering • Supersonic transport (SST) • Video cassette recorders
7	Cost and time estimates for developing AHS must be carefully and accurately determined. System design, testing, and implementation must remain within budgetary guidelines and time constraints for the project. Serious budget overruns or schedule slippage can lead to negative publicity and poor public acceptance of the system.	• Automated guideway transit • Chunnel • Denver International Airport • Supersonic transport (SST)
8	Consortiums of private and public agencies can help to ensure the long-term success of AHS research, development, and implementation.	• Automated teller machines • Automobiles • Commercial flight • Railroads
9	It is wise to keep the general public educated and informed throughout the AHS planning, design, and development phases. AHS developers and promoters should build coalitions with opposition groups, and deal forthrightly with public concerns.	• Automated guideway transit • Automobiles • Commercial flight • Interstate highway system • Liability considerations for automobile systems • Ramp metering • Supersonic transport (SST)
10	Do not overlook potential markets for AHS. The wider the potential market-base, the easier it will be to gain widespread acceptance of the new technology. This may also help to keep AHS operation costs low.	• Automated guideway transit • Automated teller machines • Automobiles • Domestic appliances • HOV lanes • Office automation • Typewriters • Video cassette recorders

olutionized the automobile industry and ingrained the automobile into American society. It also ultimately made the automobile industry a major part of the American economy.

Typewriters

The early typewriter was initially marketed to clergy, writers, and scholars, who were unwilling to accept the new invention. In their work, typed letters were considered offensive and raised questions of authenticity. There was, however, a growing need for such a device in the business community. Businessmen were less concerned with the social norms involved with handwritten letters, and desired a method to quickly and neatly record business activities for internal use. The typewriter was ideally suited for this purpose. It took almost 10 years after the introduction of the typewriter for the developers to recognize their market. Once the market for the typewriter was expanded to include business, the previously fledgling typewriter business became one of the fastest growing markets in the country (*16*).

The study of these comparable systems provides evidence that limitations in the market for AHS (e.g., by limiting the systems to passenger cars only) can limit the eventual success of the system. AHS developers need to assess the impact of including various segments of the transportation market within the AHS market served. Widening the market for AHS can lead to increased use of the system, which in turn will help reduce operating costs per person and allow AHS to gain more sufficient market size. This can be an important factor in long-term AHS success.

CONCLUSIONS

Significant lessons for AHS can be learned from the study of past systems that share important features with AHS. For example, insights for AHS may be provided from systems that involve new technology, new ways of doing old things, large infrastructure requirements, long development time requirements, extensive financing needs, and ideas that initially sound radical.

The study of the selected comparable systems has led to the identification of 10 issues that are highly relevant to AHS. These issues have been identified based on the experiences gained from previously implemented systems. Evidence from several of the comparable systems studied supports each issue. Table 1 describes each

issue and summarizes the comparable systems from our larger study that provide supporting evidence for the issue.

ACKNOWLEDGMENTS

This research was performed under the Precursor Systems Analyses of AHS Project, funded by the FHWA. The authors would like to thank everyone who contributed to this work, especially Parsons Brinckerhoff, Dr. Alain Kornhauser and his students from Princeton University, and other Calspan Team members.

REFERENCES

1. Heppenheimer, T. A. The Rise of the Interstates. *American Heritage of Invention and Technology*, Vol. 7, No. 2, 1991, pp. 8–19.
2. Diebold, J. Audiotape, Videotape, and Videocassette Recorders: American Gifts to the Japanese. In *The Innovators: The Discoveries, Inventions, and Breakthroughs of Our Time.* Dutton, New York, 1990, pp. 203–226.
3. Ford, Barbara. *The Elevator*, Walker and Company, New York, 1982.
4. Horwitch, M. *Clipped Wings: The American SST Conflict.* Massachusetts Institute of Technology Press, Cambridge, Mass. 1982.
5. Levy, E. *The People Lobby: The SST Story.* Delacorte Press, New York, 1973.
6. Kelly, Ben. *The Pavers and the Paved.* Donald W. Brown, New York, 1971.
7. Leavitt, Helen. *Super Highway—Super Hoax.* Doubleday & Company, Garden City, N.J., 1970.
8. *Impact of Advanced Group Rapid Transit Technology*, Report OTA-T-106. Office of Technology Assessment, Washington, D.C., 1980.
9. *People Mover Profile*, Report UMTA-MA-06-0081-77-1. Urban Mass Transportation Administration, Washington, D.C., 1977.
10. Finch, Christopher. *Highways to Heaven: The AUTO Biography of America.* Harper Collins Publ., New York, 1992.
11. Rae, John B. *The American Automobile: A Brief History.* University of Chicago Press, Chicago, Ill., 1965.
12. Solberg, Carl. *Conquest of the Skies: A History of Commercial Aviation in America.* Little, Brown, Boston, Mass., 1979.
13. Hudson, Kenneth. *Air Travel: A Social History.* Rowman and Littlefield, Totowa, N.J., 1972.
14. Stover, John F. One Gauge: How Hundreds of Incompatible Railroads Became a National System. *American Heritage of Invention and Technology*, Vol. 8, No. 3, 1993, pp. 54–62.
15. Diebold, J. The Electronic Bankers: The Rise and Spread of ATMs. In *The Innovators: The Discoveries, Inventions, and Breakthroughs of Our Time.* Dutton, New York, 1990, pp. 177–202.
16. Monoco, C. The Difficult Birth of the Typewriter. *American Heritage of Invention and Technology*, Vol. 4, No. 1, 1988, pp. 10–21.

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

Comparable Systems Analysis of San Francisco's Bay Area Rapid Transit System: Lessons for Automated Highway Systems

MARK D. HICKMAN

This study examines the lessons to be learned from the experience of the San Francisco Bay Area Rapid Transit (BART) system, particularly as applied to the growing research on automated highway systems (AHS). By examining the technical and nontechnical issues surrounding the development and implementation of BART in the 1960s and 1970s, the insights gained may be applied to future research and ultimate deployment of AHS. The first section of the report describes the analogy of BART by comparing some of the technical and nontechnical performance factors surrounding AHS and BART. Several pertinent technical and nontechnical issues surrounding BART are described in more detail, emphasizing the decision making that went in to BART's development, testing, and opening for revenue service. Several key issues are pursued in detail. Technically, the issues of safety, reliability, and maintenance were identified and investigated. It appears that sound system engineering principles were not applied in the BART case, and specific recommendations for improving this practice for AHS are described. In addition, the nontechnical issues of political pressure and loss of public confidence are also investigated. In this case, these pressures have severely hindered BART from achieving its full potential. The insights from the BART experience are directed toward improving the planning, design, development, and ultimate deployment of AHS.

The FHWA, as part of a recent broad agency announcement, commissioned a precursor systems analysis of the technical and nontechnical issues surrounding automated highway systems (AHS). Part of that research included an analysis of comparable systems to summarize the lessons learned from similar experiences with new technology. By reviewing these lessons, we may avoid the mistakes of the past and also make better decisions for the development and deployment of AHS. Below, the basic tenets of an AHS are outlined and the choice of San Francisco Bay Area Rapid Transit (BART) system as a comparable system is described. [In this report, BART refers to the rail transit network and its operation. The organization that runs the transit system is the Bay Area Rapid Transit District (BARTD).]

AUTOMATED HIGHWAY SYSTEMS

The basic operational concept for an AHS is that many of the typical driver functions are automated, primarily the control of steering, throttle, and brake, allowing "hands off, feet off" driving. Generally, an AHS involves an equipped vehicle and a reserved right-of-way.

California PATH, University of California, Berkeley, Richmond, Calif. 94804.

There are a number of technical and nontechnical issues associated with automated driving. Technically, there are issues associated with the sophistication, accuracy, and reliability of vehicle control. To this end, many technical needs have been identified, including:

- Proximity sensors/detectors,
- Vehicle lateral control (steering),
- Vehicle longitudinal control (throttle and brake),
- Coordination of maneuvers with other vehicles (lane changing, merging),
- Transition into/out of automated driving from/to manual driving,
- System safety in normal and degraded service conditions,
- System reliability,
- Communications requirements, and
- Response to degraded service conditions (e.g., vehicle breakdowns, accidents).

Perhaps greater than these significant technical issues are questions about the implementation and public acceptance of AHS, including:

- Ability to solve traffic congestion,
- Ability to solve safety problems,
- Environmental impacts (emissions, right-of-way requirements, etc.),
- Coordination of infrastructure provider (public) with vehicle provider (private),
- Public acceptance of the technology and of the loss of manual driving control, and
- Public response to accidents and other major problems.

These technical and nontechnical issues may appear very daunting, but many believe that an AHS is feasible. The U.S. Department of Transportation is sponsoring a national AHS consortium to identify possible solutions. With an investment of $210 million over the next 7 years, this consortium represents a considerable research and development effort. As we enter this intensive effort, it is important to review the lessons learned from other comparable experiences.

Why BART?

There are several factors that suggest that the BART system may be reasonably comparable to an AHS, including:

1. Interaction with the public. BART represents a comparable transportation system, in which the urban public is given a new transportation alternative. Currently, BART carries 260,000 trips per day. As a new transportation alternative, AHS may also face high utilization by the public.
2. Degree of intelligence. BART train operation, control, and supervision are fully automated. Train movements are under centralized control using sophisticated signals and communications. An AHS will also incorporate a high degree of intelligence.
3. Severe safety constraint. There were considerable safety issues associated with full automation of BART: train control in accident or emergency situations, hazards associated with car-borne components, safety of passengers, failures of train and central control systems, and other infrastructure failures. There are also considerable safety implications with automated vehicle control within an AHS.
4. Severe reliability constraint. Under full automation, BART has considerable reliability constraints: maintaining train schedules, interpreting speed commands, maintaining safe distances between trains, sensing the location of trains, and coordinating train movements. As with BART, the complexity and accuracy of vehicle and system-wide control systems will help determine the reliability of an AHS.
5. Many diverse subsystems. Many subsystems are required for BART, including car-borne, wayside, infrastructure, and central control systems. These subsystems manage train propulsion, operation, detection, signaling, and central control. AHS also will involve a comparable set of diverse sub-systems.
6. Failure modes. Failures in the automatic train control and train detection systems have resulted in several well-publicized accidents on BART. Subsystem failures require the removal of trains from service and associated disruptions. Vehicle and infrastructure system failures are also possible for an AHS and may significantly impact public perceptions of this technology.
7. Outage time constraints. Safe and prompt response to service disruptions are necessary on BART. Although fail-safe principles apply, continued operations under degraded service conditions are critical to system performance. For an AHS, operation in periods of system failures and degradation must be critically examined.

Based on these factors, the BART experience is examined as a comparable system to derive insights for AHS development and operation. Section 2 highlights a preliminary investigation of the key technical and nontechnical issues surrounding BART's development and draws some insights for AHS. The third section describes a detailed analysis of several key technical and nontechnical issues, making specific recommendations for AHS based on the BART experience. A fourth section offers a few conclusions.

PRELIMINARY DISCUSSION

A review of the literature regarding the planning, development, and operation of BART was conducted, yielding a set of salient issues that are particularly relevant for the continuing development of AHS. This section summarizes these technical and nontechnical issues from BART.

Technical Issues

Level of Technical Sophistication

In the 1950s, BART represented an opportunity to capitalize on new technologies. In order to lure travelers, planners envisioned a high level of service, with headways of 90 sec between trains and top speeds of 128 kph (80 mph). One such technology to reach these goals was an automatic train control (ATC) system. At that time, there was little opposition to this technology, although it was untested when the choice of technology was made. BART represented an opportunity to bring train control systems into the 20th century using more sophisticated vehicle detection, communication, and train control technologies. Similar choices about the sophistication of vehicle and roadway technologies are pending for AHS.

Level of Technical Verification and Testing

Having chosen advanced train monitoring and control technologies, BARTD and the prime contractor (the team of Parsons-Brinckerhoff, Tudor and Bechtel, or PBTB) developed specifications for these systems. However, contracts for the technical systems were not always awarded to contractors with appropriately tested and proven technologies. For example, the ATC contract was awarded to the lowest bidder (Westinghouse Electric) based on a design that was not previously demonstrated. Also, before revenue service, each car-borne and wayside system was to undergo substantial testing and quality assurance; however, these standards were not rigorously maintained, largely because of project delays. Similar specifications for AHS systems will be developed, and a rigorous program of verification and testing for these AHS systems is needed.

Consideration of Safety and Reliability

In the legislative act creating BARTD, the California Public Utilities Commission (CPUC) was given authority to monitor the safety of BART operations. The CPUC had little experience with transit systems and provided very little oversight during the early years of system development. In general, there were few safety standards included in the original system specifications. Moreover, PBTB and BARTD did not have any safety, reliability, or systems engineers on the project until the early 1970s, as BART moved into prerevenue operation. However, a large number of safety problems surfaced during system testing, including: unintended station runthroughs at 80 kph (50 mph), large gaps between cars and platforms, and a lack of information displays for the train operator. Also, there were considerable reliability problems with the ATC system during the early years of operation. For an AHS, safety and reliability will be critical to successful operation.

Other Shortcomings in Technical Performance

Original expectations of 90-sec headways through San Francisco and Oakland, with peak speeds of 128 kph (80 mph), have not been met, largely because of safety problems with the ATC system, train

and car reliability problems, lower than expected acceleration and deceleration rates, and considerable control delays at track junctions and endpoints. Moreover, BART operations have shown little tolerance for service degradation. First, there was little consideration for a means of conveying information to the train operator in normal or degraded service (e.g., location of train system malfunctions, speed limits, block occupancies). Second, there were severe limitations on degraded system operation, restricting train speeds below 40 kph (25 mph) and requiring significantly longer clearances (headways) between trains. These problems caused significant disruptions to service, especially during the first several years of operation. These observations suggest more concerted examination of AHS operations in degraded service conditions.

Maintenance Requirements

BART's experience reinforces the supposition that higher technology leads to higher maintenance costs. Studies comparing BART with other rail transit systems suggest that the operating personnel and expenditures saved through system automation are less than those required to maintain the system. Initially, many problems attributed to maintenance were, in fact, because of poor workmanship and quality control of car systems. Also, BARTD lacked the know-how within the maintenance staff to deal with car problems, resulting in high dependence on the car supplier (Rohr). The maintenance requirements of infrastructure and vehicles for AHS should likewise be identified and addressed.

Nontechnical Issues

Level of Expectations

At its inception, BART was intended to be a panacea to the problems of urban sprawl, decentralized commercial activity, and traffic congestion. Planners believed that BART would focus development in the urban core areas of Oakland and San Francisco. However, research to date suggests that BART has had little impact on commercial activity in Oakland and San Francisco and has done little to alleviate traffic congestion. There are similarly high expectations for an AHS system that should be critically examined to determine their credibility.

Public and Private Responsibilities in Development

BARTD selected PBTB for both the system design and construction management and awarded a cost plus fees contract. PBTB was answerable directly to the BARTD board of directors, leaving little oversight from BARTD staff to manage PBTB's costs or engineering practices. Moreover, there was little technical experience in rail transit systems among personnel at BARTD, leaving much of the technical oversight with PBTB. PBTB also managed all of the system subcontractors, many of whom were traditional defense contractors with little or no experience in public transit. In this way, the BART project blurred the roles of public agencies and private firms. AHS will likely bring both public and private interests into project development, and responsibilities of each should be clearly defined.

Political Pressure

BART ran over budget and opened for revenue service much later than expected for a variety of reasons: construction problems, contract negotiations and disputes, technical problems in prerevenue testing, and securing of additional project funding. Significant political pressure, however, brought the system into revenue service before the full system was operable and before all technical components had undergone sufficient testing. As a result, significant degradation in service and several accidents marred the first few years of operation. AHS will also come under significant political pressure to begin operation, and this pressure must be dealt with appropriately.

Market Prediction

Actual ridership on BART was much lower than forecast. Figures for 1975 generally show BART daily ridership at 51 percent of the forecast value (133,000 actual versus 260,000 forecast). Even today, ridership levels are lower than originally planned. Some reasons for this shortfall include: lack of rigor in the forecasting methodology, unanticipated growth in automobile ownership and low gasoline costs, poor station access, and public concerns for system reliability and safety. Similarly, caution is necessary in predicting public acceptance and demand for AHS.

Loss of Public Confidence

During the early years of BART operation, there were significant delays and disruptions in service, mostly because of problems with the ATC system and other car-borne and wayside systems. Also, several accidents were attributed to system failures or poor operating procedures, resulting in much negative publicity. Coupled with significant financial problems in the first several years, these factors led to a loss of public confidence in BART. The public perception of BART has only slowly recovered from these initial setbacks. As with other high-technology systems, AHS will also face considerable early scrutiny of system performance, and the handling of initial setbacks may determine the ultimate success of AHS.

DETAILED DISCUSSION OF BART ISSUES

The following section discusses in greater detail the technical and nontechnical issues of greatest interest and similarity with current AHS issues. The first section discusses the technical issues of safety, reliability, and maintenance (Items 2, 3, and 5 from the technical issues list above), while the second section details the nontechnical issues of handling political pressure and the loss of public confidence (Items 3 and 5 from the nontechnical issues list).

Technical Issues

Before going into the specific issues of safety, reliability, and maintenance, it is important to make some general observations about BART's technical development process. During development of the technical systems in the 1960s, the role of BARTD was intention-

ally primarily managerial as opposed to technical. PBTB, as prime contractor, was responsible for system integration and technical oversight. It was not until the system went in to prerevenue testing that many technical responsibilities shifted from PBTB to BARTD. As one can note, many of the pitfalls noted below fall in the gray area of technical responsibility between PBTB and BARTD, often during this critical time just before BART opened.

The delegation of virtually all technical development tasks to PBTB meant that there was little oversight by BARTD staff. This is widely considered to be the most significant error in the development of BART [e.g., see Burck (*1*,p.105); Profet (2,pp.124ff); and Legislative Analyst, State of California (*3*,pp.51ff)]. The primary problems did not stem from poor technical choices; rather, their root cause lay in poor project management and oversight by BARTD. It is noted that up until the late 1960s, only one member of BARTD staff was an engineer, and he had served as a consultant to PBTB in some of their BART work before his appointment at BARTD. Thus, there was little review of PBTB's technical work, either by BARTD or an independent review board, during critical times in the development process.

Recommendation

In the development and procurement of an AHS, a competent and independent technical review team should be retained during each phase of technical development and system testing. Also, the operating organization should hire capable technical personnel from the early stages of system development.

There are several other characteristics of BART that deserve mention. First, through the technical development process, BARTD and PBTB lacked any individuals specifically assigned to systems engineering or integration. Although such systems engineers are common in detailed aerospace technologies, they are rare in transportation. During the development of new technical subsystems, a systems engineering function at BART would have aided in integrating vehicle, wayside, and central systems; anticipating system hazards; and responding to system problems.

Recommendation

In program development, as well as in each field operational test and proposed implementation, a systems engineering function should be incorporated that integrates AHS subsystems for the vehicle, wayside, and infrastructure.

Second, PBTB chose to use functional rather than design specifications for the development of several technical subsystems. These specifications allow characterization of a system in terms of its function, rather than determining detailed design standards, and allow the greatest innovation by the system developer. In the BART experience, examples of liberties in design include the novel train control system, the car design (by an aerospace contractor), and a nonstandard gauge and concrete ties for the track to improve ride stability. However, this type of specification makes it difficult to verify contractual obligations of each system contractor when the system does not perform as desired; BARTD entered litigation separately against Westinghouse and Rohr over the issue of system specifications and the resulting contractual obligations (*4*,p.144; *1*,p.105). In addition, the high degree of innovation in design may also lead to difficulties in system integration.

Recommendation

AHS specifications must carefully balance the need for technical innovation with the need for more specific design criteria to assure a safe and reliable system.

Safety

During the first several years of operation, BART was plagued with safety problems. Many of these problems resulted not from operator error but rather as the result of technical faults. Several safety issues emerged in prerevenue testing as many of the technical bugs were worked out of the system. This period of testing, however, was cut short, resulting in safety problems in revenue service that received high publicity (*4*,p.8).

The first major accident in revenue service occurred only 3 weeks after the system opened in 1972. According to the investigation by the Legislative Analyst, State of California (*3*,pp.25ff), the car-borne ATC equipment misinterpreted a speed command, causing the train to speed past the Fremont station and crash at the end of the line. In January 1975, a nonrevenue train had a fatal collision with a maintenance vehicle; the accident was blamed on the inability of the train detection system to detect maintenance vehicles on the main right-of-way. A third serious accident in 1979 involved a train fire in the Transbay Tube. Further investigation revealed that the material from which the cars were manufactured was not sufficiently flame-retardant.

These incidents raise specific concerns about the treatment of safety by BARTD and PBTB because these problems were primarily because of technical error. It seems that the root causes of these safety problems resulted from several factors in the system development process (*4*,p.85). The following describes these factors and identifies some lessons learned about the treatment of safety for the technical development of AHS.

1. Specification of safety requirements for system components. The system specifications put forth by PBTB for each of the technical systems were primarily functional and not design specifications. In using this approach, specific safety standards for each proposed new technology were basically nonexistent: the technology for critical subsystems (such as the ATC system) lacked widespread industry safety standards (*4*,pp.166–167).

Recommendation

Regardless of the decision for functional or design specifications, safety and reliability requirements for system operation should be explicitly incorporated.

2. Hazard analysis of the system. Because many of the subsystems were developed as new technology, it would have been helpful to have a systems engineering function to determine appropriate ways of integrating these subsystems. One part of this function would be a complete hazard analysis of the various system components and possible modes of failure. Oddly, a hazard analysis was performed on the car-borne and wayside ATC equipment in 1971, and it identified several critical deficiencies in system design, including possible misinterpretation of speed commands on the vehicle (*5*,pp.169–170). Unfortunately, PBTB had not investigated this matter further before the related accident during revenue service in 1972 (*5*,pp.232–233; *3*,pp.27–28).

Recommendation

A critical function of systems engineering for AHS should be a detailed hazard analysis of vehicle, wayside, and infrastructure systems. This hazard analysis must be performed as early in the design process as possible to allow revisions to the system design.

Recommendation

Safety issues should be given highest priority in determining the readiness of an AHS system before start of service.

3. Technical experience at BARTD and the CPUC. The CPUC was given responsibility for assuring safe operation in BARTD's enabling legislation in 1957. However, PBTB controlled technical system specification and development up until the system opened for revenue service. Personnel at both BARTD and the CPUC during the 1960s and early 1970s had little experience with rapid transit systems or their associated technologies (*3*,pp.37–45). Both agents may have been aided by hiring technical personnel much earlier in the technical development process.

Recommendation

A staff of technically competent safety engineers should be hired (or retained) to conduct independent safety analyses for an AHS. This staff should be brought in to the AHS development process as early as possible.

4. Organizational treatment of safety within BARTD. Up until April 1972, a few months before the system opened, safety engineering was only a small organization within the operations department and relied heavily on the technical expertise of the operations and maintenance personnel. To many, this did not allow a fair and independent safety review, because the operating personnel were under considerable political pressure to bring the system into revenue service (*3*,pp.43–45). In May 1972, the safety group was moved under the finance department, creating a new insurance and safety organization that was thus free of the political pressure but nonetheless distant from the technical expertise of operations and maintenance. In 1973, the group was moved up to the departmental level (the insurance and safety department), largely because of political pressure resulting from the revenue service accident and other studies of system safety (*3*,pp.43–45). The technical competence of the safety group was still inadequate, leading BARTD to retain the Lawrence Berkeley Laboratory as safety consultants for several years after beginning revenue service (*6*,pp.11ff). It was not until July 1975 that an independent safety department was formed at BART (*7*,pp.20–21).

Recommendation

A safety engineering function should include staff members at the highest possible level within the AHS development team, who can effectively communicate safety concerns to project management.

5. Capabilities of a safety program. Now that BART has been in operation for over 20 years, the safety department has been given considerable responsibility and broad authority to improve safety within BARTD. The responsibilities of the BART safety program now, as detailed in the following list, may be transferable to an AHS safety organization (*8*,pp.15–16).

- Setting reasonable safety goals and objectives for BARTD.
- Informing BARTD management of safety status, problems, and improvements.
- Participating in the planning and review process for system design, construction, reliability, maintenance, and personnel training.
- Reviewing engineering tests to ensure compliance to safety requirements.
- Monitoring and inspection of system operation.
- Conducting hazard analyses to identify and mitigate safety risks.
- Analyzing operating rules, procedures, and practices to limit exposure to hazardous situations.
- Collecting and reviewing historical information on hazards, system failures, and accidents.
- Investigating system failures, mishaps, and accidents.
- Ensuring operability of hazard detection and warning systems.
- Ensuring compliance with regulatory agencies.
- Organizing and coordinating safety programs within BARTD.
- Conducting scheduled and unscheduled disaster and emergency exercises and drills.

Reliability

Because many of the subsystems in BART relied on new technology, system reliability was a significant issue in system development and early operations. The facts of the BART experience are clear: during this time, major reliability problems emerged. As late as 1975 (3 years after opening for service), an average of 40 percent of BART cars were out of service on a given day because of failed components. Car-borne system failures occurred very frequently in revenue service, seriously degrading train and network-wide performance. Failures in the wayside ATC system also caused considerable delays. In time, however, BART has been able to recover from many of these early reliability problems but not without generating considerable public dismay over system performance.

AHS, because it represents an entirely new technology, has very severe reliability constraints associated with successful deployment. In contrast with BART, however, an implementation of AHS may come under significantly greater pressure to ensure a high level of safety and reliability in early operation. Also, AHS may not be so fortunate to have a long "grace period" to work out the bugs in the system; perhaps today's public is less forgiving and patient. To this end, the following issues in system design and development may serve as learning experiences from BART.

1. Design for "graceful decay." BART was intended to be, and ultimately achieved its goal, a completely automated train operation, even under degraded conditions. However, during the first several years, operating procedures for degraded conditions resulted in significant disruptions in service. Statistics from the first 3 years of operation show that passengers had to be off-loaded for one of every four equipment failures, a measure at least seven times worse than other peer rail transit systems. Moreover, during any single car-borne failure, "fail-safe" procedures were applied; in almost all cases, this implied a full stop of the train, after which the train was limited to a maximum speed of 40 kph (25 mph). Because there are few yards or sidings on the BART system, these trains would continue over a significant portion of the network at this reduced speed. These frequent stops and speed restrictions resulted in serious delays that propagated through the system (*9*,p.IV–14).

Recommendation

Consideration of automated systems should focus on a graceful decay under degraded conditions. System specifications for AHS should focus on design for possible service degradation, including equipment malfunctions in the vehicle, at the wayside, and in the infrastructure.

2. Design for human interaction. As originally designed, the train operator is responsible for operation only in the case of a major service disruption or emergency. However, because the ATC system was not fully operational when BART opened for revenue service and because service disruptions occurred frequently, the operator played a more significant role during the first few years of operation. This role was impeded by a cab design that assumed a much more passive role of the operator: there were no information displays in the cab for the operator to know the intended vehicle speed or information on subsystem failures within that train. As a result, operators used line-of-sight rules for train operation or held trains in a station for extended periods to locate car problems. This was a serious design flaw that led to substantial train delays in early revenue service. It was several years after beginning operation before the cab interfaces were upgraded (*9*,p.II–7 and III–3; *4*,p.157).

Recommendation

An AHS design must be sensitive to the information provided to drivers during automated operation, especially during degraded conditions. Human factors research should emphasize the driver's response to information, especially in degraded service or emergency situations.

3. System specification and development. With some federal financial assistance, PBTB developed a test track to test alternative system configurations. The track ultimately had two purposes: to allow prospective system suppliers to test their products and to assist BARTD and PBTB in developing specifications for each of the required subsystems (*4*,p.141). Many suppliers participated in the testing program. Moreover, PBTB often incorporated the abilities of several products tested on the track in developing the functional specifications for subsystems. This testing program was very successful, considering the lack of research and development on these systems nationally at that time (*4*,p.152).

In deciding on contract awards, however, the testing experience was largely ignored (*4*,p.152). Because the specifications were functional, subsystem design was left to each contractor. Moreover, contract award criteria were independent of whether vendors had successfully demonstrated their product either on the test track or in any other application. As a result, many of the contracts were awarded to suppliers with little experience or no proven product. For example, the contract to supply rail cars was given to a supplier with no experience in rail transit, and the ATC system contract was awarded to Westinghouse Electric in spite of the fact the proposed system had never been tested and no prototype existed.

Recommendation

As much as possible, AHS operational test sites should be flexible to allow various manufacturers to test new technologies. In selecting system suppliers, technical experience, proven technology, and test results should be given considerable weight in the evaluation criteria.

4. Prerevenue system testing and quality assurance. BARTD had no internal quality control organization for the delivered systems (*10*,p.6). As a result, operating and maintenance personnel at BARTD relied heavily on PBTB for early product testing and quality control. Simultaneously, political pressure was mounting on PBTB to bring the system into revenue operation; construction delays had already pushed back the opening for revenue service from 1969 to 1972. For this reason, testing and quality control functions were rushed, leaving considerable doubt regarding the effectiveness of test procedures (*4*,p.149). One report indicates that less than one-half of the rolling stock had been subject to adequate yard departure testing, and none of the cars had undergone complete ATC system tests, before revenue service began (*3*,p.69). This inadequacy of system testing also had significant repercussions for the maintenance function at BARTD.

Recommendation

Sufficient time in the AHS development process must be left for product testing and quality control. This involves allowing ample time and resources for suppliers to debug new technical subsystems and to test and debug the fully integrated AHS on site before beginning operation.

Maintenance

Maintenance was the responsibility of BARTD once the various contractors began delivering each of the subsystems. The maintenance organization within BARTD's operations department was responsible for checking car-borne systems on arrival of the car at the yards. However, the maintenance department relied heavily on PBTB to supervise these testing procedures (*2*,pp.78ff). Once revenue service began, BARTD alone was responsible for approving trains for release into revenue service each day. Because many of the delivered subsystems had not been adequately tested, the maintenance function faced a considerable workload once the system opened. Anywhere from 30 to 60 percent of the cars were in the shop on a given day, and about 25 percent of the cars were brought into the shops three or more times with the same problem (*6*,p.17; *11*,p.III–22).

Several factors that influenced the planning and management of maintenance at BARTD offer insights for AHS.

1. Design for maintenance. In terms of component specifications, PBTB took a novel approach by including reliability, maintainability, and availability (RMA) specifications directly. Despite this approach, a number of contractors did not adequately consider product failures and maintenance requirements in designing their systems. For the cars, critical train control systems were located in very troublesome positions on the car, requiring significant time to repair or replace. The car manufacturer also did not adequately consider some of the environmental hazards of rail operations; for example, several critical components were mounted on the undercarriage, where there is considerable wear and tear in normal operation (*12*,p.331). On the other hand, some components were too accessible. For example, the emergency door release equipment was placed just below a passenger seat and attached only with Velcro brand fasteners. Thus, a passenger might accidentally (or deliberately) open the doors while the train was in motion (*10*,pp.113–114). Such problems required considerable work to modify the location of components.

Recommendation

RMA specifications should be used for any AHS implementation, including explicit mean-time-between-failures and mean-time-to-replace requirements. These requirements should be specified for both vehicle and wayside equipment, ensuring that parts are

easily accessible and that component trouble-shooting requires minimal effort, both on board the vehicle and in the automated lanes.

2. Maintenance information. Initially, BARTD maintenance personnel were very dependent on PBTB and its subcontractors, largely because the system specifications had been developed by PBTB and ultimate product designs were approved most often without adequate oversight by BARTD personnel (*2*,pp.78–80). Another significant problem with BARTD's maintenance efforts in the early years can be attributed to a lack of information about the systems built: significant discrepancies were often noted between car-borne systems as delivered and the blueprints on hand at BARTD. Information was inadequate, placing additional demands on the contractors to assist with maintenance (*2*,pp.27–34). The maintenance effort was also poorly managed within BARTD: there was initially no consistent information reporting format to identify problems on cars as they were brought to the shops, making it difficult to know the type and severity of the problem (*2*,pp.87ff).

Recommendation

AHS system operators should develop substantial maintenance capabilities in house during system development. Because of the large number of diverse subsystems involved with an AHS, capabilities must include a common information system to track component specifications, performance, and failures.

3. Maintenance planning and management. In addition to the information reporting problems noted previously, initially there were inadequate supplies of common parts. This resulted primarily from the management's inexperience with traditional inventory stocking practice (*11*,pp.III–40ff). Also, because of the magnitude of initial system problems, resources were not managed effectively; because of the great need to keep rolling stock on the rails, resources were funneled into crisis management, detracting from detailed trouble-shooting or other preventive maintenance practices (*11*,pp.III–24ff). During one maintenance audit, the ratio of hours spent on unscheduled versus scheduled maintenance was 1.48:1 (*10*,p.4). As a result, problems were not adequately diagnosed, and cars often returned to the shops with the same problem as a previous visit.

Recommendation

The provision and maintenance of in-vehicle components will obviously be the responsibility of equipment suppliers; these suppliers should carefully consider maintenance requirements in designing and developing these systems. Infrastructure providers should also begin planning for maintenance requirements during the development process. In both cases, requirements will include maintenance equipment to identify and repair failures, common information systems, and clearly defined procedures for addressing scheduled and unscheduled maintenance needs.

Nontechnical Issues

The success or failure of large public projects such as BART is typically driven not by the level of technical sophistication but rather by nontechnical issues. The political conditions and overall public perception of the project may have significant ramifications for its success, especially if there is a large investment of public money in a project. The challenge to the project management is to deal with these pressures appropriately. The BART experience suggests that if the these concerns are not handled appropriately, the project faces an uphill battle for political and public acceptance.

In public transit projects, the loss of confidence either in the political realm or among the public at large rarely results in the full project being canceled. In the BART case, although mistakes were made in the development process and in the early years of operation, system operation and ridership have improved significantly. This ability to tolerate short-term problems for long-term benefits results in part from the long-term success of other cities' rail systems (Boston, New York, Chicago, etc.). For AHS, however, no such long-term experience with the technology exists, and the early years of AHS implementation will be critical to the acceptance of this technology. Thus, alleviating the early political and public acceptance issues is important to sustain long-term development of AHS.

Political Pressure

The political stakes in BART during the early 1970s were the culmination of a political process beginning more than 20 years earlier. The genesis of BART was the result of strong political forces in the Bay Area in the 1950s, when BART was sold as the core element of the regional planning program (*13*,pp.3–5; *14*,pp.10–11). Politicians and the business community supported the rail system because they thought it would solve the problems of urban sprawl, decentralized development, and increasing traffic congestion. Thus, the political forces were sold on rather unrealistic expectations of what the rail system might do (*13*,pp.7ff).

The resulting political energy was compounded by the large number of actors involved, including:

- Local officials,
- The BARTD board of directors,
- The regional planning commissions: the Bay Area Rapid Transit Commission (1951–1957) and the Metropolitan Transportation Commission (since 1970),
- The California state legislature,
- The Federal Department of Housing and Urban Development (HUD), and
- The Federal Urban Mass Transportation Administration (UMTA).

The California legislature created BARTD in 1957 and provided some funding for the project through the 1960s, whereas HUD and UMTA provided funding for the BART system development in the late 1960s and early 1970s. Thus, many political interests had a financial or political stake in the success of BART.

As with most public works projects, delays were considerable. The initial starting date was pushed back from 1969 to 1972, and the Transbay Tube was not opened for revenue service until 1974. In the late 1960s, delays were primarily related to the final systems design, procurement, and funding (*14*,pp.21–22). Yet, technical concerns and procurement problems with the ATC system and the cars contributed significant delay in the early 1970s (*3*,p.12). Political pressures mounted to begin revenue service as early as possible.

The high political expectations, the many institutions, and the inevitable project delays all resulted in great political pressure on BARTD and PBTB. Several measures, detailed in the following

list, may have either contributed to or alleviated some of this pressure.

1. Interaction of technical and political forces in the development process. During the 2 years before opening, delays resulted primarily from technical problems with delivered systems. Most researchers believe there was insufficient time to work out these bugs before BART entered revenue service. Unfortunately, the technical personnel on the project (primarily at PBTB) either were not in a position to influence decision making or simply did not speak strongly enough for a longer testing period. It seems that there was inadequate representation of technical concerns in the political process, attributed to the poor management of technical issues at BARTD (*1*,pp.106–107).

Recommendation

Technical personnel should maintain high visibility in AHS decision making throughout the development process. Administrative and management boards should include staff with a high degree of technical competence.

2. Ability to develop the system incrementally. One advantage of the radial nature of the BART system design is that it permitted incremental implementation. BARTD was able to open the Fremont-Oakland line first in September 1972, alleviating at least some of the pressure to bring the system on line. This early operating experience could also be translated to other lines before they opened for revenue service. Political pressure was obviously greatest to open the Transbay Tube between Oakland and San Francisco [e.g., see the emphasis in Legislative Analyst, State of California (*3*,pp.12–13)]; unfortunately, that section was the last to open in September 1974.

Recommendation

As system design will allow, AHS projects should take advantage of incremental implementation. This may imply that an AHS be deployed in a small corridor initially, allowing for system expansion to other corridors in the future. The selection of an initial corridor should be based in part on that corridor's ability to demonstrate significant first user benefits.

Loss of Public Confidence

From the seemingly strong voter support in 1962, public opinion on BART deteriorated. Public relations activities during the development and construction were minimal. Also, in the first years of revenue service, passengers found the stations difficult to get to and encountered frequent delays and disruptions in service. These problems were compounded by the state legislature's discovery of widespread system safety and reliability problems after the system opened [reported in Legislative Analyst, State of California (*3*)]. As a sign of this loss, BART ridership is just now reaching levels initially predicted for 1975.

In hindsight, several factors contributed to the deterioration of public support for BART.

1. Level of public interaction before opening. Following the voters' approval of a bond bill in 1962, contact between BARTD and the public diminished rapidly, because of, in part, the obvious shift in focus toward design and construction and away from consensus building (*15*,p.49). However, communities could hold public hearings at any time after the vote; sadly, few communities took advantage of these, except where there was considerable opposition to development plans [e.g., Berkeley, summarized in Zwerling (*15*,pp.56–65)]. Moreover, responsibility for public relations passed from BARTD to PBTB, despite the fact that the contractor had little expertise in this area (*14*,p.22; *15*,p.43). Moreover, little effort was made by PBTB to solicit public comment during design and construction for fear this would contribute to additional delays and costs (*14*,p.45; *15*,p.49).

Recommendation

AHS project development should include mandatory public forums to discuss system implementation, both before initial project authorization and during the project design and construction. Also, other public information strategies should be implemented, such as local site offices, telephone information lines, and other forums for public information and feedback.

2. Public perception of the ease of use. From the initial system design, access to BART was difficult because of the large interstation spacing. The system needed substantial in-station parking and feeder bus service to provide station access for both drivers and transit-dependent passengers (*13*,pp.33–34). However, parking facilities were and remain inadequate to handle demand. For the feeder bus service, BART was largely unable to coordinate services with local providers such as AC Transit and Muni. Although there were clearly stated policies regarding service coordination between BART and these transit providers, little actually changed once BART opened (*15*,pp.91–104). For example, BART is still competing with AC Transit for passengers traveling across the bay. These problems may have resulted in part because BARTD was not responsible to any regional transportation planning body during development in the 1960s (*14*,p.45).

Recommendation

AHS should be incorporated in a regional transportation planning process (likely to be mandated under current federal legislation) and should be coordinated with other regional transportation system improvements. Specifically, adverse and beneficial impacts of an AHS should be addressed in the context of the entire regional transportation system.

3. Overcoming early problems. Finally, BARTD officials were not candid with the public about early problems on the system, including service delays and disruptions. Statistics compiled in 1979 indicated that equipment failures alone totaled about seven per day; such failures resulted in train off-loads, unscheduled train removals, or schedule delays over 10 min (*9*,p.IV–13). In addition, BARTD had significant financial problems in its first few years of operation, because revenues were unable to cover operating costs as expected (*6*,pp.24ff). Many observers believe the first general manager of BARTD had difficulty admitting publicly the scope of technical and financial problems within the system. As a result, the public (and the media) tended to control the investigation of these problems, rather than BARTD [(*1*,p.164); and e.g., see Legislative Analyst, State of California (*6*)]. Although there were substantial changes in management policies within 2 to 3 years after the system opened, the gradual improvement in public attitudes about BART are because of the considerable patience of the public during that time (*13*,pp.37–38).

Recommendation

As much as is politically feasible, problems with AHS development and implementation should be addressed candidly, both internally within the organization and externally with the public.

CONCLUSIONS

Research and development efforts in the AHS field are growing steadily and will likely continue for many years to come. Hopefully, these efforts can take advantage of the history and experience of BART and of other similar experiences of new technology in transportation. As evidenced through this study, there is wealth of insight into both good and bad practice that emerges from a detailed review of these experiences. As the transportation profession looks to the future, this review can be useful for developing both technically sound and politically and publicly acceptable innovations for the transportation system.

ACKNOWLEDGMENTS

The financial support of the FHWA, through the Precursor Systems Analysis for Automated Highway Systems, is gratefully acknowledged. Dr. Steven Shladover at PATH provided oversight for this research. Other individuals also contributed valuable comments and support, including Stein Weissenberger at PATH, Loren Bonderson and Herb Hall at Delco, and Fred Mangarelli and Amy Cochran at Hughes. However, the author remains solely responsible for the content and accuracy of the material in this report.

REFERENCES

1. Burck, C. G. What We Can Learn from BART's Misadventures. *Fortune*, Vol. XCII, No. 1, July 1975, pp.104ff.
2. Profet, Robert A. *BART Automatic Train Control System Quality and Safety Independent Study*. Presented to the BARTD Board of Directors, 1973.
3. Legislative Analyst, State of California. Investigation of the Operations of the Bay Area Rapid Transit District with Particular Reference to Safety and Contract Administration. 1972.
4. Office of Technology Assessment, U.S. Congress. *Automatic Train Control in Rail Rapid Transit*. Washington, D.C., 1976.
5. Crooks, R. B., H. H. Hunter, R. W. Mowery, and C. B. Shields. *Battelle Limited Safety Analysis of the BART Automatic Train Control System*. Prepared for Parsons-Brinckerhoff-Tudor-Bechtel, Battelle Columbus Laboratories, 1971.
6. Legislative Analyst, State of California. Analysis of BART's Operational Problems and Fiscal Deficits, with Recommendations for Corrective Action. 1974.
7. Legislative Analyst, State of California. Public Utilities Commission Review of Rapid Transit Safety. December 1977.
8. BARTD Safety Department. BART System Safety Program Plan. 1978.
9. Brumberger, Neil. Bay Area Rapid Transit District System Performance Study. 1980.
10. Lancaster, E. J., and J. H. Teske. BART Quality Audit Report. 1973.
11. Cresap, McCormick and Paget, Inc. A Management Audit of the San Francisco Bay Area Rapid Transit District. Prepared for the California Senate Committee on Public Utilities and Corporations, 1974.
12. Strobel, Horst. *Computer Controlled Urban Transportation*, Vol. 10, International Series on Applied Systems Analysis. Wiley-Interscience, New York, 1982.
13. Webber, Melvin M. The BART Experience—What Have We Learned? Monograph No. 26, University of California, Berkeley. Institute for Transportation Studies, Berkeley, Ca., 1976.
14. Office of Technology Assessment, U.S. Congress. *An Assessment of Community Planning for Mass Transit; Volume 8: San Francisco Case Study*. Office of Technology Assessment, Washington, D.C., 1976.
15. Zwerling, Stephen. *Mass Transit and the Politics of Technology: A Study of BART and the San Francisco Bay Area*. Praeger Publ., New York, 1974.

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

Modeling Intermodal Auto-Rail Commuter Networks

MARIA P. BOILE, LAZAR N. SPASOVIC, AND ATHANASSIOS K. BLADIKAS

This study is a methodological framework for evaluating the operating and pricing policies in an intermodal auto-rail network. Commuters departing from their homes access their final destination, a Central Business District (CBD), via auto, rail, and intermodal auto-to-rail modes (e.g., park and ride). If a commuter chooses to begin the trip by auto, there are numerous paths to reach the final destination. Once on the highway, the commuter can switch to rail at stations along the rail route. The commuter may also choose to walk to the rail station closest to the trip's origin. The intermodal network is considered as one system, and traffic flows and travel costs are optimized for the whole system and not for separate auto and transit networks, as is the case in most previous studies. The central part of the framework is an equilibrium demand-supply model that consists of a mode choice and a traffic assignment submodel. The commuters' choice of using auto or rail is represented by a mode choice submodel. The traffic assignment submodel assigns transit (walk-to-rail and drive-to-rail) and pure auto trips to routes in the network, based on the minimization of an individual traveler's generalized (travel time and out-of-pocket) cost. The model yields modal shares, equilibrium flow pattern, and resulting generalized costs of the network assignment. The framework is applied to an intermodal network, modeled closely after a northern New Jersey commuter corridor, and used to analyze policies that include increasing parking capacity on commuter rail lots, increasing tolls on highways, increasing the parking fee at the CBD, decreasing rail fares, and improving rail service quality. These policies are evaluated based on user costs, and transit and highway operators' costs and revenues.

This study is a methodological framework for analyzing the effects of various operating and pricing policies for assigning traffic flows according to the performance of intermodal passenger transportation systems. In an intermodal transportation network, commuters leaving their homes can access their final destination via auto, transit, and intermodal auto-to-transit modes. If a commuter chooses to begin a trip by auto, there are numerous paths to reach the final destination. Once on the highway, the commuter can switch to transit at any station along the transit route. The commuter may also choose to walk to the transit station closest to the trip's origin.

Central to the methodological framework is a combined mode choice-traffic assignment model. This model allocates trips over the auto and transit modes of the network, and assigns pure transit (walk-to-transit), intermodal (auto-to-transit), and auto trips over the network routes, yielding equilibrium flows between demand and supply.

The primary motivation for this work was the necessity to include all transportation modes in a unified, interconnected manner to form a National Intermodal System as dictated by the Intermodal Surface Transportation Efficiency Act of 1991. In addition, the Clean Air Act Amendments of 1990 require employers located in nonattainment areas and having more than 100 employees to increase average vehicle occupancy by 25 percent. The objective is to use the proposed framework to evaluate various policies that can induce a commuter shift to public transit. More precisely, the framework could be used to evaluate the effects of policies on network performance, the cost of capital improvements, and operating cost changes for highway and rail operators. To demonstrate its capabilities, the framework is used to evaluate a variety of policies in an intermodal commuter network.

M. P. Boile, Department of Civil and Environmental Engineering, Lafayette College, Easton, Penn. 18042-1775. L. N. Spasovic and A. K. Bladikas, National Center for Transportation and Industrial Productivity, NJIT, Newark, N.J. 07102.

LITERATURE REVIEW

A review of the literature revealed several papers related to the development of demand-supply network equilibrium models, the formulation of mode and transit station choice models, and the analysis of multimodal networks.

Network equilibrium models that consider travel by private car and one or more public transit modes have been developed by Florian (*1*), Florian and Nguyen (*2*), Abdulaal and LeBlanc (*3*), Dafermos (*4*), LeBlanc and Farhangian (*5*), and Florian and Spiess (*6*). All these formulations consider only pure modes. Once travelers have chosen modes, they are assigned over separate modal networks without the possibility of switching modes during their journey. A recent paper by Fernandez et al. (*7*) is the only published work that explicitly analyzes intermodal trips in a network equilibrium framework. The paper presents three model formulations with auto, metro, and a combined mode (auto-to-metro), and analyzes the resulting equilibrium conditions. The underlying assumption is that the combined mode is considered only at those origins where metro is not available. When metro is available, travelers can choose only between auto and metro. This assumption does not consider that even when there is a metro station at an origin, a traveler may prefer to drive to or be dropped off at a station along the metro route.

Mode and station choice model formulations, Fan et al. (*8*), Miller (*9*), and Ortuzar (*10*), concentrate on estimating travel volumes by mode, transit station, or both. They have not been implemented within a network equilibrium context.

Turnquist et al. (*11*), developed a framework to evaluate improvements in transit operating strategies that meet various service objectives. It was assumed that transit and the competing modes have fixed travel impedance, thus precluding the consideration of competitive interactions among modes.

Manheim (*12*) presented a conceptual framework for analyzing transportation improvements using a simple intermodal network with one origin, one destination and three paths (auto, rail, and auto-to-rail). A software package was used to obtain modal splits and

traffic assignment. It was not clear what method was used to equilibrate demand and supply.

Current Urban Transportation Planning Software

The majority of planning software packages such as QRSII (*13*), MINUTP (*14*), and TRANPLAN (*15*), follow the Urban Transportation Modeling System (UTMS), which consists of four steps: (a) trip generation, (b) distribution, (c) modal split, and (d) traffic assignment. Although, in theory, the UTMS should consider the effect of network performance (*i.e.*, travel times) on demand, this is not how the software actually works.

These software have three shortcomings. First, an inexact methodology is used for assigning flows over the networks. When assignment is completed using the all-or-nothing method, the impact of congestion on travel times is not recognized, because travel times are assumed to be constant. When the minimization of networkwide cost is used (*13*), the assignment is inconsistent with driver behavior, because according to Wardrop's First Principle (*16*), drivers will attempt to switch paths between an origin-destination (O-D) as long as this can decrease their individual travel times. This inconsistency could lead to unrealistic flows, especially in networks with moderate congestion. In rare cases when an "equilibrium solution" (Sheffi (*17*)) is computed, this is accomplished using an inexact heuristic (*14*).

Second, the software do not recognize the interaction between network performance and modal splits. The final resulting equilibrium travel times are not considered in adjusting the initial modal splits. However, because of the previously mentioned flaws with the traffic assignment, this interaction, even if established, will likely produce unrealistic flows.

Third, the UTMS packages ignore intermodal flows by not permitting trips to shift between the auto and transit networks. This is critical because the highway portion of an intermodal trip will affect highway travel times and thus modal splits. This shortcoming makes it difficult, if not impossible, to use the packages for evaluating the effects of transit on highway network performance. An exception in the ability to model intermodal trips is the new version of EMME/2—Release 7 (*18*). A new module, "Matrix Convolutions," allows the enumeration of intermediate zones between an origin and a destination. By having an intermediate zone, for example, serve as a destination zone for the highway network, and as an origin zone for the transit network, it is possible to consider intermodal trips.

In marked contrast to the current software, the model presented here performs an interactive mode choice-traffic assignment over an integrated highway-transit network, wherein different performance (time-volume) functions are used to model travel over each portion of an intermodal trip. Therefore, the model can be used to evaluate effects of transit improvement policies on highway network performance.

METHODOLOGICAL FRAMEWORK

The framework for analyzing the effects of various policies on network flow patterns and associated travel costs in an intermodal network is shown in Figure 1. The framework is designed to answer questions of interest to transportation planners (Is it possible to assign travelers to the network more efficiently? What are the benefits of the improved assignment? At what cost are these benefits achieved?) and to investigate trade-offs between travel time reductions and the cost of capacity increases. The model answers these questions by providing:

- Equilibrium assignment of flows over an intermodal network that minimizes user costs;
- Total network travel cost for each policy;
- Incremental changes in cost;
- Estimation of benefits for each alternative; and
- Rail service and parking capacity additions needed to accommodate rail ridership increase.

Given an intermodal highway-rail transportation network, the framework starts with the collection of network data, costs, and travel demand parameters. The data include capacities of highway and rail links, performance functions that relate traffic volumes, capacities and travel times, time impedances, out-of-pocket costs (e.g., rail fares, tolls, parking fees, auto operating cost), commuter rail data (frequency and capacity of trains and capacities of station parking facilities), and travel volumes for each O-D. In addition,

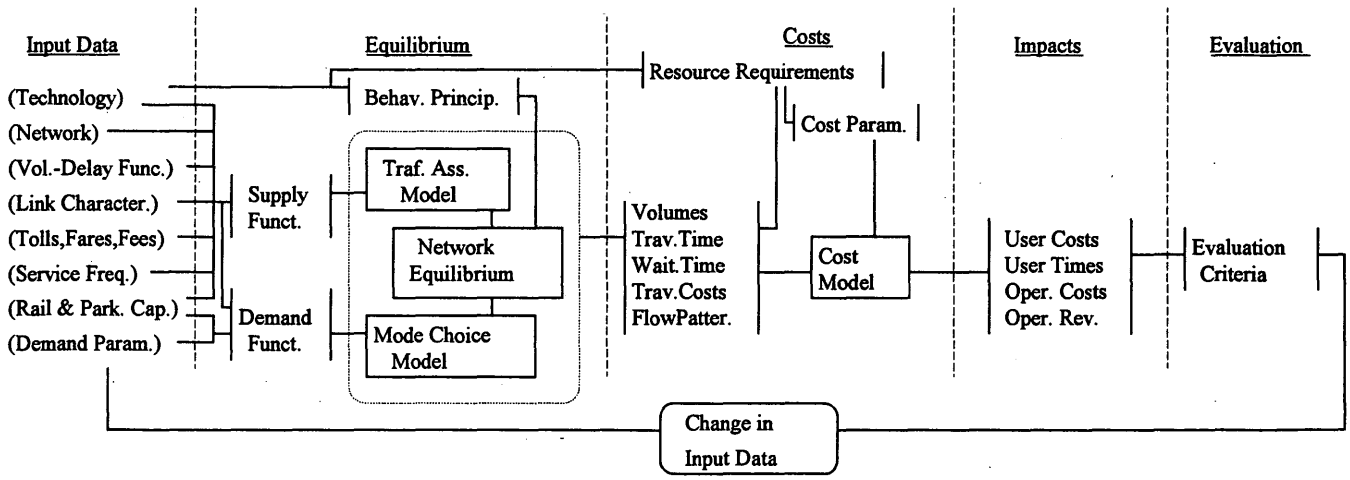

FIGURE 1 Methodological framework.

background traffic volumes that originate, terminate, or both, outside the study network, are also collected.

The data are entered into a combined mode choice-traffic assignment network equilibrium model. The model calculates modal shares, equilibrium flow patterns, and the resulting generalized cost (a sum of out-of-pocket costs, in-vehicle, and out-of-vehicle time costs) of the network assignment, and can perform sensitivity analysis by varying input parameters such as congestion levels, out-of-pocket costs, and demand data.

Several policies can be developed by selecting (and later changing) the input parameters. The cost module estimates the policy costs. Finally, the evaluation module evaluates the effect of policies by trading off user costs and travel times with operators' costs and revenues, and determining the desirability of each policy.

MODELING INTERMODAL DECISIONS

There are basically three distinctive approaches for modeling mode choice and route choice decisions in an intermodal setting as shown in Figure 2. They differ on what choices are modeled within the demand side and what choices are modeled within the supply side of the formulation. Within the demand side, a mode choice could be formulated using disaggregate mode choice models, such as multinomial, or nested logit. This formulation assumes that each mode is chosen with some finite probability and considers the relative attractiveness of one mode over the others. Within the supply side, a choice is performed as a route choice (routing) problem, in which a traveler chooses the mode-specific routes that minimize generalized travel cost.

In the first approach, shown in Part *a* of Figure 2, mode choice is modeled within the supply side of the formulation as a routing problem. As a result, the total demand is distributed among auto, rail, and intermodal routes, to minimize the individual travelers' generalized cost. The summation of the resulting route flows over the auto, rail, and intermodal routes yields modal shares.

The second approach, shown in Part *b* of Figure 2, formulates the commuters' choice of auto or rail transit within the demand side of the model formulation via a binomial logit model, which splits the total demand between auto and transit. Then, within transit, the choice between pure rail (walk-to-rail) and intermodal (auto-to-rail) trips is treated as a least-cost routing problem.

The third approach, shown in Part *c* of Figure 2, performs the choice among auto, rail, and intermodal trips within the demand side of the formulation via a nested logit model. The model splits the total demand between auto and transit in the so-called upper-level decision. Then, the demand for transit is split between rail and intermodal in the lower-level decision. After the modal shares have been determined, the choice of routes within each mode is performed within the supply side as a routing problem.

A detailed discussion of these three approaches, mathematical formulations of the models, derivations of the equilibrium conditions, and selected case studies are presented by Boile (*19*).

THE MODE CHOICE-TRAFFIC ASSIGNMENT MODEL

The second approach has been adopted here for the formulation of the model, which is essentially a mathematical program with a nonlinear objective function and linear constraints. Its objective function follows the formulation of user equilibrium with elastic demand by Beckmann et al. (*20*). Its conceptual statement is:

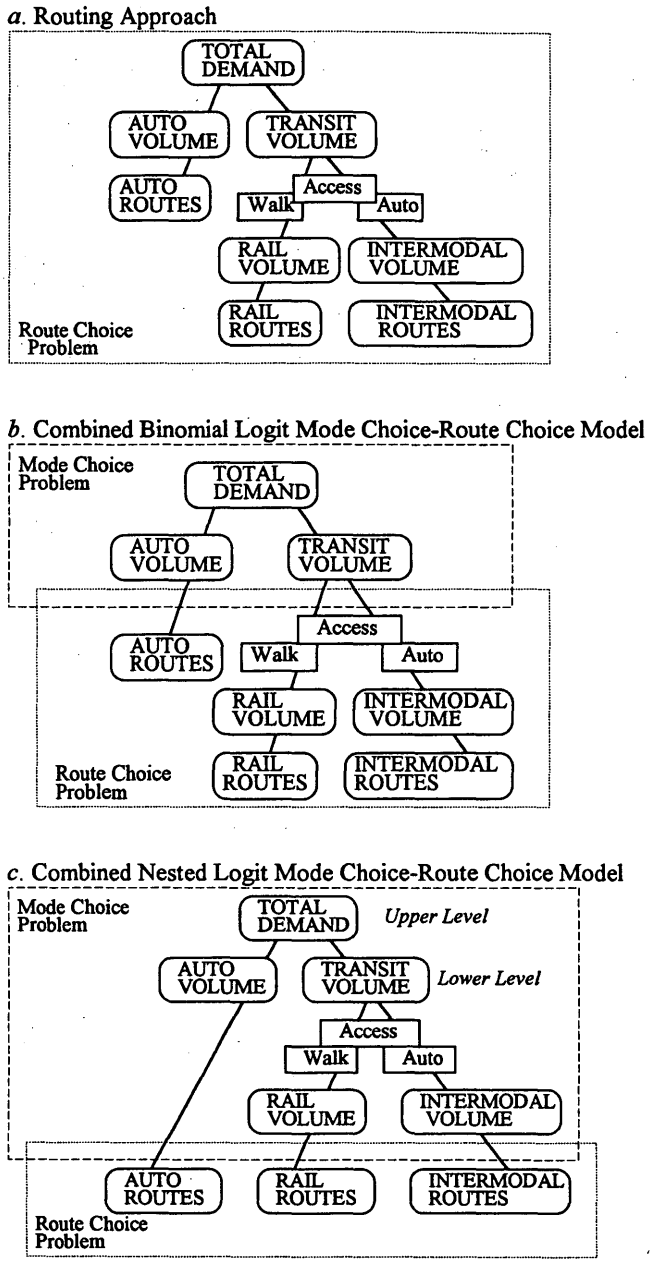

FIGURE 2 Approaches for modeling intermodal decisions.

Minimize:

Total Individual User Cost minus the Integral of the Inverse Demand Function subject to:

- Demand conservation constraints;
- Link flow conservation constraints;
- Rail and transfer link capacity constraints; and
- Non-negativity constraints.

The mathematical formulation is shown in Table 1. Equation 1.1 ensures that all trips between an O-D pair are accounted for. Equa-

TABLE 1 Combined Mode Choice-Traffic Assignment Model

Nodes and Links:	Choice Variables:	Sets:
i,j = origin, destination	f_l = flow on link l	E = all O-D pairs
l, p = link, path.	h_p = flow on path p	L = all links
		$R(L), A(L), T(L), W(L)$ = all rail, highway, transfer, walking links
		Pa, Pr, Pm = all auto, rail, intermodal paths
		$CR(R)$ = highest passenger load rail links

Parameters:

T^{ij} = demand between origin i and destination j

T_t^{ij} = demand for transit between i and j $[T^{ij} * \exp(a_t + b * GC_t^{ij}) \;/\; \sum_{k=auto,transit} \exp(a_k + b * GC_k^{ij})]$,

$D^{-1}{}_{ij}$ = inverted demand function,
δ_{lp} = binary parameter, 1 if link l is in path p, and 0 otherwise,
c_l = generalized cost on link l.
occ = occupancy rate for autos,
$Space_l$ = number of parking spaces at a rail station (link l), $Seats$ = number of train seats per peak hour,
λ = transit vehicle load factor,

$$\text{Minimize } Z = \sum_{l \in L} \int_0^{fl} c_l(\phi)d\phi - \sum_{ij \in E} \int_0^{T^{ij}} D_{ij}^{-1}(\omega)d\omega$$

subject to:

$$T^{ij} = T_a^{ij} + T_t^{ij} \qquad \forall\, i,j \in E \qquad 1.1$$

$$T_a^{ij} = \sum_{p \in P_a^{ij}} hp \qquad \forall\, i,j \in E \qquad 1.2$$

$$T_t^{ij} = \sum_{p \in P_r^{ij}} hp + \sum_{p \in P_m^{ij}} hp \qquad \forall\, i,j \in E \qquad 1.3$$

$$fl = \frac{1}{occ}[\sum_p \delta lp * hp] \qquad \forall\, l \in A,T \qquad 1.4$$

$$fl = \sum_p \delta lp * hp \qquad \forall\, l \in R,W \qquad 1.5$$

$$fl \le Space_l \qquad \forall\, l \in T \qquad 1.6$$

$$fl \le (Seats) * \lambda \qquad \forall\, l \in CR \qquad 1.7$$

$$h_p \ge 0 \qquad \forall p \in Pa, Pr, Pm \qquad 1.8$$

tions 1.2 and 1.3 are the definitional constraints for auto and transit shares of the travel volume. Equations 1.4 and 1.5 equate flows on a link with the sum of the flows on all the paths through that link. Equations 1.6 and 1.7, respectively, ensure that the number of commuters using parking lots, and the number of rail users, do not exceed the parking and rail line capacities. Equation 1.8 ensures that all path flows are non-negative.

Equilibrium Conditions

An equilibrium solution to the model must satisfy two conditions. First, for each O-D pair and mode, namely auto and transit, no traveler has an incentive to unilaterally change routes, because travel cost cannot be minimized further. This is known as Wardrop's First Principle (*16*), and for this model it is expressed as:

$$GC_{pk}^{ij} - C_k^{ij} \begin{cases} = 0, \text{ if } h_p^{ij} > 0 \\ \geq 0, \text{ if } h_p^{ij} = 0 \end{cases} \quad \forall i, j, k \tag{1}$$

where

GC_{pk}^{ij} = generalized cost on path p from origin i to destination j via mode k;
C_k^{ij} = minimum generalized cost from origin i to destination j via mode k; and
h_p^{ij} = flow on path p from origin i to destination j.

This condition indicates that for a given O-D pair i, j and a mode k, a path p is used only if its generalized cost equals the minimum generalized cost. If the flow on this path is zero, its cost must be greater than (or equal to) the cost on the least cost path.

Second, at equilibrium, no traveler has an incentive to unilaterally change modes. This condition is derived from the inverted demand function (given as a logit model):

$$T_t^{ij} = T^{ij} \frac{\exp(a_t + b * GC_t^{ij})}{\exp(a_t + b * GC_t^{ij}) + \exp(a_a + b * GC_a^{ij})} \quad \forall i, j \tag{2}$$

where

T_t^{ij} = number of travelers from i to j by transit;
T^{ij} = total number of travelers from i to j;
T_a^{ij} = number of travelers from i to j by auto, $T_a^{ij} = T^{ij} - T_t^{ij}$;
a_a, a_t = mode-specific parameters;
b = mode-independent parameter; and
GC_k^{ij} = the generalized cost of traveling from i to j via mode k (k = auto, transit).

Then, the inverse of this function:

$$GC_a^{ij} - GC_t^{ij} = \frac{1}{b}\left[\ln \frac{T_t^{ij}}{T_a^{ij}} - (a_a - a_t)\right] \quad \forall i, j \tag{3}$$

becomes the second equilibrium condition that gives the relationship between the generalized costs of the two utilized modes. Details on the equilibrium condition derivations are explained by Boile (*19*).

CASE STUDY

The methodological framework was applied to an intermodal network shown in Figure 3. The network is a realistic representation of a portion of the Raritan Valley Corridor located in Union County, NJ. It consists of five origins: Westfield, Garwood, Cranford, Kenilworth, and Roselle Park (designated O1 through O5); and one destination, which is Newark (D). The network is composed of three major highways, I-78, Route 22, and the Garden State Parkway (GSP), local county roads that run between the major highways, and a NJ Transit commuter rail line. According to 1980 U.S. Bureau of the Census data, there were 540, 130, 620, 220, and 920 peak-hour work trips, respectively from Westfield, Garwood, Cranford, Kenilworth, and Roselle Park, to Newark (*22*).

Assumptions and Input Data

While actual demand and traffic volume data are used in the analysis, the mode choice model coefficients for Equation 3 were estimated to be: $a_t = 0$, $a_a = 0.02$, and $b = 0.33$. Currently, NJ Transit and the regional MPO are attempting to calibrate access choice models in the corridor. Obviously, the mode choice-traffic assignment model can be easily re-run once accurate and origin-specific mode choice coefficients are available.

The highway links are congested and travel times were modeled using the Bureau of Public Roads (*23*) congestion curves. For each link (road) type (*i.e.*, arterial and freeway), the capacities used in the curve were calculated using the 1985 Highway Capacity Manual (*24*). Transfer link travel times were assumed constant and included time to park and wait for the train. It was assumed that travelers are experiencing an average waiting time of one-half the headway (Manheim (*12*)). Walking times were determined by dividing the walking link length by an average walking speed of 4.39 km/hr (2.73 mph). Rail link travel times were estimated from a service schedule. Trains on the route operate in local service (*i.e.*, stop at each station).

Regarding out-of-pocket costs, there is a standard toll of $0.35 on the GSP. Rail fares are $2.15 from O1 and from O2, $1.75 from O3 and $1.50 from O5. There is no train station at O4. Parking fees are $3.00, $1.15, $1.50, and $5.00 for the parking lots at O1, O3, O5, and D, respectively. There is no commuter parking at O2. Vehicle operating cost is calculated by multiplying the distance traveled by $0.40 per vehicle-km ($0.25/vehicle-m). Travel time was translated into cost by multiplying it by $15/hr.

Train seating capacity was estimated at 500 seats (four cars at 125 seats each). Trains run every 20 minutes yielding a line capacity of 1,500 seats per peak hour. The number of parking spaces at O1, O3, and O5 is 759, 373, and 239, respectively. There is no constraint on parking capacity at D.

The Cost Model

A cost model was developed to estimate the cost of capital investments such as parking expansion and train capacity improvements. An annual cost of $805.90 per parking space was obtained by spreading the acquisition cost of $10,000 over a 30-year period and using 7 percent interest (capital recovery factor 0.08059). Assuming that half of this cost was allocated to the peak hour during 265 working days per year, the daily cost per peak hour was $1.52 per space.

The transit operator cost includes maintenance and overhead as well as the more direct cost of operation (wages, power, etc.), and is represented by the all-inclusive hourly operating cost per vehicle c. The total hourly operator cost is obtained by multiplying the active fleet (defined as the ratio of the total round trip time over the headway) with the hourly operating cost per vehicle. The total round trip time is the round trip route length divided by the average speed. The total hourly operator cost is then:

$$C = \frac{2cL}{HV} \tag{4}$$

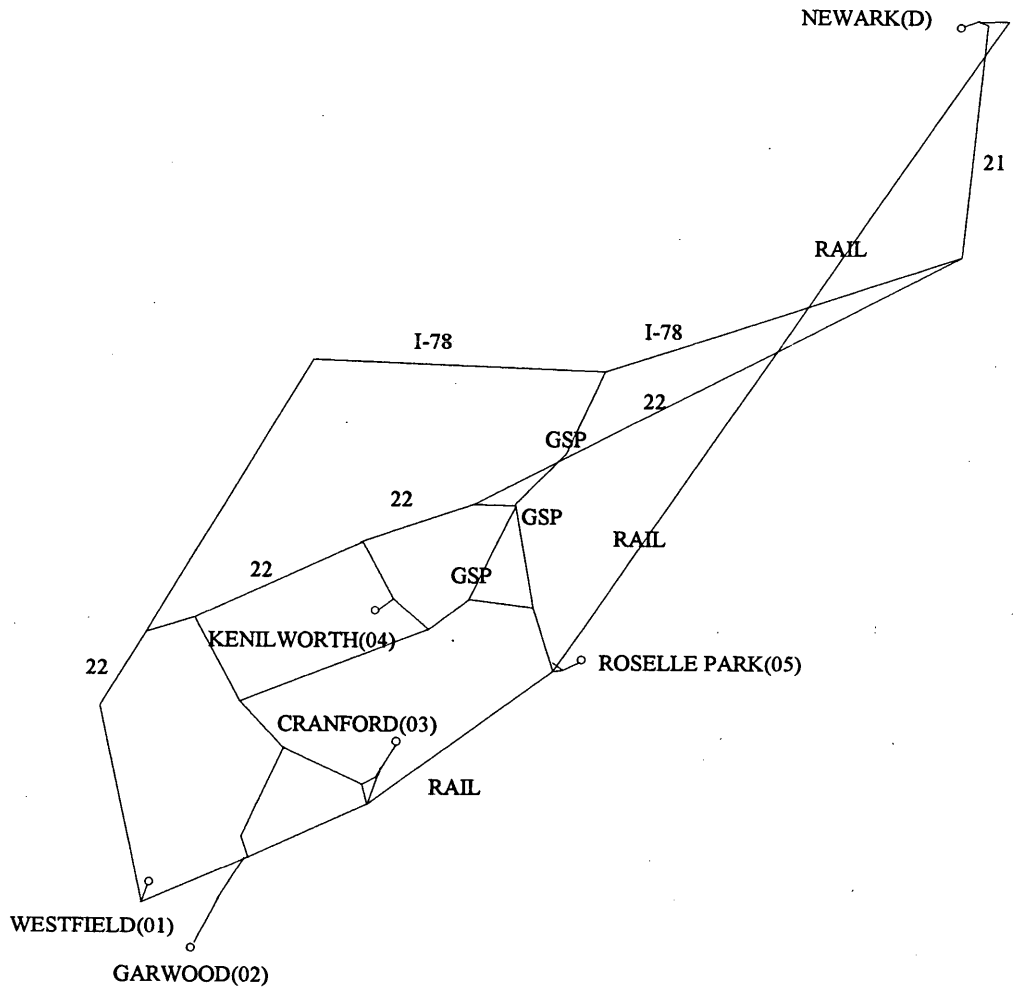

FIGURE 3 Experimental network.

where the variables and their assumed values are:

c = vehicle operating cost	500 [$/vehicle-hr]
L = length of transit route	5.47 [km] (3.4 [mi])
H = route headway	0.333 [hr/vehicle]
V = average transit speed	32.2 [km/hr] (20 [mi/hr])

With the above assumptions, the total hourly operating cost (C) is $510.51/hr.

Policies Analyzed

Six policies were analyzed to determine the best alternative in terms of its impact on the users and transportation system (highway and rail) operators. They are as follows:

Policy 1: Baseline Case

It models the current situation, represents the "Do Nothing" alternative, and serves as a basis for comparison.

Policy 2: Increase Parking Capacity

Currently, some of the parking lots operate at capacity, preventing an increase in intermodal travel. This policy expands parking capacity by adding 60 new spaces to the lot at O3 and 300 new spaces to the lot at O5.

Policy 3: Increase Tolls

Highway tolls were doubled (from $0.35 to $0.70) to induce travelers to shift from auto to rail.

Policy 4: Increase Parking Fees at the CBD

The CBD parking fee was increased 40 percent ($2.00), to make auto trips less attractive and reduce the number of auto commuters.

Policy 5: Decrease Rail Fares

The rail fare was reduced by $1.00 for all trips.

Policy 6: Increase Train Frequency

The train frequency was doubled from three to six trains per hour, halving the average waiting time from 10 to 5 min.

RESULTS

For each policy, the model generated equilibrium modal splits and network flow assignments, and calculated generalized costs. Table 2 presents the flows and costs by mode for each O-D pair and policy.

Evaluation of Policies

The policy effects in terms of modal shares are shown in Table 3. The effects of policies, in terms of user and operator impacts are presented in Table 4. The operator gross revenues are estimated by multiplying the toll (for highways) and fare (for transit) by the number of users. The last column of the table labeled "Net User and Operator Cost" is the sum of all operator (rail, highway, and CBD parking) revenues, which are included in the three preceding columns, minus the user costs (first column). It is assumed that the rail operator is responsible for the rail line and the station parking lots. Therefore, net rail operator revenues are computed by subtracting rail operating costs and parking capital investments from the fare-box and station parking revenues.

The highest reduction in user cost (and time), results from the doubling of rail frequency. User cost decreased by 9.11 percent (and time by 15.44 percent) in comparison with the Baseline. Rail operator's revenues and station parking lot revenues increased by 20.5 percent and 27.7 percent, respectively, and the highway and the CBD parking operators' revenues decreased by 20.3 percent. The rail revenue came primarily from 281 travelers who shifted from auto to rail and intermodal. Of the people who shifted to transit, 200 or 71.17 percent were shifted to pure rail and 81 or 28.83 percent to intermodal trips. The increase in rail frequency doubled the operator costs to $1021.02/peak hr, but the net rail operator revenue increased by $226.3.

The increase in station parking lot capacities resulted in a decrease of user cost by 3.4 percent (and time by 3.6 percent) compared to the Baseline. The $547.4 capital investment in 360 spaces increased station parking and rail revenues by $511.5 and $40.5, respectively. Thus, per $1 invested, $1.008 was generated in revenues for station parking and rail combined. The increase in rail revenue came both from highway and rail travelers who could not have driven to rail in the Baseline because of parking lot capacity constraints. The increase in parking capacities reduced the total number of auto users by 2.8 percent.

The increase in CBD parking lot fees increased the user cost by 3.5 percent and reduced the time by 5.36 percent compared to the Baseline. This policy increased CBD parking revenues by 20.4%. It also increased rail and station parking revenues by 14.3 percent and 19.2 percent, respectively. Note that this policy also reduced highway toll revenues by 14 percent. The increase in rail revenue came from 194 travelers shifting from auto to rail and intermodal. As

TABLE 2 Results of the Mode Choice-Traffic Assignment Model

		Policy											
		1 -- Baseline		2 -- Increase Parking Capacity		3 -- Increase Tolls		4 -- Increase CBD Parking Fee		5 -- Decrease Rail Fare		6 - Double Rail Frequency	
Origin -- Destination	Path Type	flow users/peak hr.	cost $/user	flow users/peak hr.	cost $/user	flow users/peak hr.	cost $/user	flow users/peak hr.	cost $/user	flow users/peak hr.	cost $/user	flow users/peak hr.	cost $/user
	Auto	410	20.1	412	20.1	401	20.4	354	21.7	383	19.9	329	19.5
O1-D	Rail	--	--	--	--	--	--	--	--	--	--	--	--
	Intermod.	130	23.6	128	23.6	139	23.6	186	23.6	157	22.6	211	20.8
	Auto	63	21.7	67	21.3	60	21.9	51	22.8	56	21.3	46	20.5
O2-D	Rail	67	21.4	11	21.4	70	21.4	79	21.4	74	20.4	84	18.6
	Intermod.	--	--	52	20.0	--	--	--	--	--	--	--	--
	Auto	247	22	239	21.6	247	22.2	219	23.1	244	21.6	200	20.8
O3-D	Rail	--	--	--	--	--	--	28	21.2	3	20.2	47	18.5
	Intermod.	373	18.7	381	18.7	373	18.7	373	18.7	373	17.7	373	15.9
	Auto	166	19.4	155	19	163	19.6	149	20.6	157	19	142	18.2
O4-D	Rail	--	--	--	--	--	--	--	--	--	--	--	--
	Intermod.	54	18.6	65	18.6	57	18.6	71	18.6	63	17.6	78	15.8
	Auto	500	20.7	446	20.3	485	20.9	419	21.8	457	20.3	388	19.4
O5-D	Rail	235	21.2	--	--	253	21.2	333	21.2	286	20.2	371	18.4
	Intermod.	185	17.1	474	17.1	182	17.1	168	17.1	177	16.1	161	14.3
	Auto	1,386		1,319		1,356		1,192		1,297		1,105	
Total	Rail	302		11		323		440		363		502	
	Intermod.	742		1,100		751		798		770		823	

TABLE 3 Mode Shares

Policy	O-D Pair	Auto (%)	Rail (%)	Intermodal (%)
1 -- Baseline	O1 - D	75.9	0.0	24.1
	O2 - D	48.2	51.8	0.0
	O3 - D	39.8	0.0	60.2
	O4 - D	75.6	0.0	24.4
	O5 - D	54.4	25.6	20.0
	TOTAL	57.1	12.4	30.5
2 -- Increase in Parking Capacity	O1 - D	76.4	0.0	23.6
	O2 - D	51.4	8.4	40.2
	O3 - D	38.6	0.0	61.4
	O4 - D	70.6	0.0	29.4
	O5 - D	48.4	0.0	51.6
	TOTAL	54.3	0.4	45.3
3 --Increase Tolls	O1 - D	74.2	0.0	25.8
	O2 - D	46.4	53.6	0.0
	O3 - D	39.8	0.0	60.2
	O4 - D	74.2	0.0	25.8
	O5 - D	52.7	27.5	19.8
	TOTAL	56.0	13.1	30.9
4 -- Increase Parking Fees	O1 - D	65.5	0.0	34.5
	O2 - D	38.9	61.1	0.0
	O3 - D	35.3	4.5	60.2
	O4 - D	67.7	0.0	32.3
	O5 - D	45.5	36.2	18.3
	TOTAL	49.0	18.2	32.8
5 --Decrease Rail Fares	O1 - D	71.0	0.0	29.0
	O2 - D	43.2	56.8	0.0
	O3 - D	39.4	0.4	60.2
	O4 - D	71.6	0.0	28.4
	O5 - D	49.7	31.1	19.2
	TOTAL	53.4	15.0	31.6
6 -- Double the Rail Frequency	O1 - D	61.0	0.0	39.0
	O2 - D	35.5	64.5	0.0
	O3 - D	32.1	7.7	60.2
	O4 - D	64.4	0.0	35.6
	O5 - D	42.0	40.5	17.5
	TOTAL	45.5	20.7	33.8

shown in Table 3, the increase in rail and parking revenue came primarily from auto travelers shifting to intermodal (56 travelers) and rail (138 travelers). An increase in the CBD parking lot fee by $2.00 decreased auto's share by 8.1 percent.

The remaining policies can be examined in the same manner to determine the most attractive one, in terms of its impacts on users and operators. The increase in highway tolls increased the highway operator's revenue by 95.7 percent (from $485.1 to $949.5). This increase caused 15 auto and 3 intermodal users to shift to rail, thus increasing the rail operator's revenue by $51.6. User costs increased marginally by 0.7%, while the time savings decreased marginally by 0.9%.

The decrease in rail fare does not have a substantial impact on rail ridership. Only 89 auto travelers diverted to rail and intermodal. This policy also resulted in a large loss of farebox revenue (a 54.7 percent reduction compared to the Baseline).

In terms of their effect on the net user and operator costs, the order of preference of the policies is as follows:

1. Double the rail frequency.
2. Increase station parking capacity.
3. Increase the CBD parking fee.
4. Decrease rail fare.
5. Increase tolls.

Combinations of Policies

Two additional policies were developed. The first increased station parking lot capacities and parking fees at the CBD. The second, in addition to the changes included in the first, decreased parking fees at an underused lot, added more capacity to the O3 and O5 station lots, and doubled the rail frequency. These policies were deemed to have the best potential for further reducing user cost while increasing rail revenues.

The previous discussion of equilibrium flows indicated that user cost could have been reduced further if parking capacity constraints

TABLE 4 User and Operator Impacts

Policy	User Cost ($/peak hour)	User Time (min/ peak hour)	Parking Capital Investm. ($/peak hour)	Station Parking Reven. ($/peak hour)	Rail Operating Cost ($/peak hour)	Rail Fare-Box Reven. ($/peak hour)	Net Rail Operator Revenues ($/peak hour)	Highway Toll Revenues ($/peak hour)	CBD Parking Reven. ($/peak hour)	Net User & Oper. Cost ($/peak hour)
a. 1--Baseline	49,282	81,005	0	1,177.8	510.51	1,787.7	2,455	485.1	6,930	39,412
b. 2--Increase park. capacity	47,610	78,084	547.4	1,689.3	510.51	1,864.1	2,495.5	461.8	6,597.5	38,055
c. 3--Increase tolls	49,638	80,277	0	1,204.8	510.51	1,839.3	2,533.6	949.5	6,782	39,373
d. 4--Increase CBD park. fee	51,017	76,663	0	1,345.7	510.51	2,130.7	2,965.9	417.0	8,340	39,294
e. 5--Decrease rail fare	47,825	78,927	0	1,257.6	510.51	809.7	1,556.8	454.4	6,491	39,323
f. 6--Double rail Frequency	44,793	68,498	0	1,419.3	1021.02	2,283	2,681.3	386.9	5,527	36,198
R* (-) or I* (+) of *b* over *a* (%)	-3.4	-3.6	N/A	+43.4	0	+4.3	+1.65	-4.8	-4.8	-3.44
R (-) or I (+) of*c* over *a* (%)	+0.7	-0.9	0	+2.3	0	+2.9	+3.2	+95.73	-2.14	-0.10
R (-) or I (+) of*d* over *a* (%)	+3.5	-5.36	0	+14.3	0	+19.2	+20.81	-14	+20.4	-0.30
R (-) or I (+) of*e* over *a* (%)	-2.96	-2.6	0	+0.7	0	-54.7	-36.58	-6.3	-6.3	-0.22
R (-) or I (+) of*f* over *a* (%)	-9.11	-15.44	0	+20.5	+100	+27.7	+9.22	-20.3	-20.24	-8.15

R*: Reduction, I*: Increase

were relaxed. The model provides a convenient measure for evaluating the effect of such a capacity increase on user cost in the form of dual prices. In optimization terminology, dual prices indicate a change in the objective function (in this case user costs) resulting from a marginal increase in the availability of scarce resources (parking spaces). A careful review of dual prices indicated that providing an additional parking space at O3 or O5 would decrease user costs by $1.95/traveler and $4.13/traveler, respectively. Based on this observation, the first combination policy added 170 spaces at O5, in addition to the spaces already added under Policy 2. In addition, the parking charge at the CBD was doubled to $10. This policy increased user costs by 3.92 percent while time decreased by 11.67 percent compared to the Baseline. The CBD parking, rail operator and station parking revenues increased by 43 percent, 37.6 percent, and 87.6 percent, respectively, while highway revenues decreased by 28.5 percent. The rail revenue came primarily from 395 travelers who shifted from auto to intermodal. This policy, implemented with a total capital investment of $805.9 for 530 spaces resulted in an increase of $1031.9 and $672.1 in station parking and rail revenues, respectively. Thus, per $1 invested, $2.1 were generated in revenues from station parking and rail.

The results indicated that out of 759 spaces at the O1 parking lot, only 216 spaces were occupied. In addition, a marginal increase in the parking lot capacity at O3 and O5 would further decrease user costs by $1.95/traveler and $2.7/traveler, respectively. This led to the second combination policy that added 60 additional spaces at O3 and 100 at O5. Since there was no point to further increase station parking capacity without increasing rail capacity, this policy also included the doubling of rail frequency. In addition, the parking fee at O1 was decreased from $3 to $1.

This policy decreased user costs by 13.2 percent and reduced time by 28.7 percent compared to the Baseline. Rail and station parking revenues increased by 90.7 percent and 88.4 percent, respectively, while the CBD parking revenues decreased by 30.2 percent. The increase in rail revenue came primarily from 903 travelers shifting from auto to intermodal. The expenditures of $1079.6 for 690 spaces, and $510.51 for more frequent rail operations resulted in increases of $1041.7 and $1621 in station parking and rail revenues respectively. Therefore, the combined revenues for station parking and rail increased by $1.67 per $1 invested.

CONCLUSIONS

The cases analyzed in this study are a small sample of all possible policies that the outlined methodological framework is capable of evaluating. The results can be used to screen policies aimed at affecting travel demand. Policies aimed at improving rail service and making it more accessible are the best in terms of their ability to divert auto drivers to transit. The best deterrent to driving appears to be an increase in the CBD parking fee. Direct user fee charges for rail or highway users appear to be the least able to affect travelers' choices. However, it needs to be recognized that, in general, the final decision on policy selection is likely to depend on social, political, economic and environmental considerations as well.

FUTURE EXTENSIONS

The model presented here can be improved to consider travelers' preferences between pure rail (walk-to-rail) and intermodal (auto-to-rail) trips, as well as between transfer points. Second, the assumption of constant travel time on rail needs to be relaxed, and accelerated regimes such as skip stop and express services could be considered. These regimes would be advantageous for rail, because as auto commuters are shifted to rail, the average travel time on rail can decrease (Morlok (*25*)). Of course, this is true only if there is excess physical rail capacity to accommodate accelerated operations. Therefore, the model can be improved to handle accelerated regimes on rail, and to select optimal operating schedules. Third, the assumption of travelers having perfect information on travel times and making rational decisions was rather strong and needs to be relaxed. In the real world, lack of information is likely to yield traffic assignment with worse total travel costs than those predicted by the model. Finally, each policy was examined for current levels of congestion. The model can be rerun to analyze policies under congestion levels that might be encountered in future years.

ACKNOWLEDGMENT

The research results presented here were partially supported by a grant from the U.S. Department of Transportation, University Transportation Centers Program, through the National Center for Transportation and Industrial Productivity (NCTIP) at NJIT.

REFERENCES

1. Florian, M. A Traffic Equilibrium Model of Travel by Car and Public Transit Modes. *Transportation Science,* Vol. 8, 1977, pp. 166–179.
2. Florian, M., and S. Nguyen. A Combined Distribution Modal Split and Trip Assignment Model. *Transportation Research,* 12-B, 1978, pp. 241–246.
3. Abdulaal, M., and L. LeBlanc. Methods for Combining Modal Split and Equilibrium Assignment Models. *Transportation Science,* Vol. 13, 1979, pp. 292–314.
4. Dafermos, S. C. The General Multimodal Network Equilibrium Problem with Elastic Demand. *Networks,* Vol. 12, 1982, pp. 57–72.
5. LeBlanc, L. J., and K. Farhangian. Efficient Algorithms for Solving Elastic Demand Traffic Assignment Problems and Mode Split-Assignment Problems. *Transportation Science,* Vol. 15, No. 4, 1981, pp. 306–317.
6. Florian, M., and H. Spiess. On Binary Mode Choice/Assignment Models. *Transportation Science,* Vol. 17, No. 1, 1983, pp. 32–47.
7. Fernandez, E., J. DeCea, M. Florian, and E. Cabrera. Network Equilibrium Models with Combined Modes. *Transportation Science,* Vol. 28, No. 3, 1994, pp. 182–192.
8. Fan, K., E. Miller, and D. Badoe. Modeling Rail Access Mode and Station Choice. *Transportation Research Record,* 1413, 1993, pp. 49–59.
9. Miller, E. Central Area Mode Choice and Parking Demand. *Transportation Research Record,* 1413, 1993, pp. 60–69.
10. Ortuzar, J. Nested Logit Models for Mixed-Mode Travel in Urban Corridors. *Transportation Research A,* Vol. 17A, No. 4, 1983, pp. 283–299.
11. Turnquist, M. A., A. H. Meyburg, and S. G. Ritchie. Innovative Transit Service Planning Model that Uses a Microcomputer. *Transportation Research Record,* 854, 1982.
12. Manheim, M. *Fundamentals of Transportation Systems Analysis: Volume 1: Basic Concepts,* MIT Press, 1979.
13. National Cooperative Highway Research Program, Report 187. *Quick-Response Urban Travel Estimation Techniques and Transferable Parameters: User's Guide.* TRB, 1978.
14. Comsis. *MinUTP Technical User Manual,* 92 B-Version, 1992.
15. The Urban Analysis Group. *TRANPLAN: User Manual,* Version 7.0, 1990.
16. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. *Proceedings Institute of Civil Engineers Part II,* 1952.
17. Sheffi, Y. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods,* Prentice-Hall Inc., 1985.
18. INRO Consultants. *Users Manual EMME/2 Release 7,* Montreal, Canada, June 1994.
19. Boile, M. P. *Demand–Supply Equilibrium over Intermodal Transportation Networks.* National Center for Transportation and Industrial Productivity (NCTIP), NJIT, NCTIP-WP94-01, 1994, 89 pages.
20. Beckmann, M. J., C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation,* Yale University Press, New Haven, Conn.
21. *A Manual on User Benefit Analyses of Highway and Bus Transit Improvements,* AASHTO, 1965.
22. *1980 Census of Population.* U.S. Bureau of Census, U.S. Department of Commerce.
23. U.S. Bureau of Public Roads. *Traffic Assignment Manual.* U.S. Department of Commerce, 1964.
24. Transportation Research Board. *Highway Capacity Manual.* Special Report 209, 1985.
25. Morlok, E. K. *Introduction ot Transportation Planning and Engineering.* McGraw Hill, 1978.

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

Dynamic Traffic Modeling of the I-25/HOV Corridor Southeast of Denver

JUAN ROBLES AND BRUCE N. JANSON

The application of a dynamic traffic assignment model (DYMOD) to the southeast Denver metro area surrounding I-25 and I-225 is described. Hourly volume counts on 20 percent of network links were used to estimate a morning peak-period trip matrix between 110 zones using a three-step procedure developed to estimate origins, destinations, and the origin-destination (O-D) trip matrix. Trip departure times from each zone were estimated using 5-min counts at on-ramps to I-25. Then, 5-min volumes and speeds predicted by DYMOD for I-25 through lanes and on-ramps were compared with observed data from loop detectors. Lane-blocking incidents were modeled and compared with observed traffic volumes and speeds during these incidents. The results show that DYMOD can reproduce and predict network traffic conditions (with or without accidents) as well as generate alternative routes to reduce traffic delays during incidents.

The implementation and testing of a dynamic traffic assignment model (DYMOD) to predict time-varying traffic conditions on a moderate-sized urban network during incidents and congested periods is described. Using nonlinear optimization formulations and solution algorithms (*1–5*), DYMOD performed well in computational tests on small networks and was ready for validation and testing on a suitable freeway-arterial system. An area southeast of Denver including the I-25/HOV corridor presented an excellent test environment for this application because of its (a) density of instrumentation, (b) diversity of highway types, and (c) variations in daily traffic conditions.

The objectives of this project were to:

- Develop computer data bases of system characteristics (both supply and demand) for a network of freeways and arterials southeast of Denver including I-25 and I-225 covering approximately 100 mi^2 (260 km^2).
- Calibrate and validate DYMOD to reproduce time-varying traffic conditions throughout this network based on historical data collected from loop detectors.
- Demonstrate the model's ability to predict volumes, speeds, and delays on alternative routes during incidents such as lane-blocking accidents.

Peak-hour counts for about 20 percent of the network links were collected from city, county, and state traffic engineering departments, and used to estimate a morning peak-period trip matrix between 110 zones covering this area. Volume counts of 5-min collected from loop detectors at the on-ramps to I-25 and I-225 were used to estimate the departure times of these trips from each zone. Average speeds collected in 5-min intervals from the through-lane detectors on I-25 were used to calibrate the model's speed-flow relationships.

With these data, DYMOD was used to predict volumes and speeds during a typical 5:00–10:00 a.m. weekday peak period. On average, predicted flows agreed to within 12 percent of actual 5-min volumes on the I-25 through lanes at detector locations. Three lane-blocking accidents on I-25 were then modeled. The results indicate that DYMOD can successfully model incident conditions to estimate vehicle hours of delay and generate route-diversion planning strategies during lane-blocking accidents.

DYMOD's formulation, math properties, and solution algorithm are described in a companion paper (*6*). The brevity of this paper precludes a discussion of other dynamic traffic modeling approaches. The next two sections describe the development of network supply and demand data bases for the southeast Denver area in preparation for running DYMOD and emphasize the importance of obtaining sufficient network coverage of traffic counts in short time intervals with which to estimate departure times. The results of DYMOD's application to a "base case" with no accidents and to three known accident cases are then presented. The concluding section outlines recommendations for data acquisition and management procedures essential for successful dynamic traffic modeling.

OVERVIEW OF DYMOD

The dynamic user-equilibrium (DUE) version of DYMOD applied in this research is defined as follows:

> Given a network with speed-volume functions to predict travel times, and given a set of zone-to-zone trip tables containing the number of vehicle trips departing from each zone and headed towards each zone in successive time intervals, DYMOD finds the volume of vehicles on each link in each time interval that satisfy DUE conditions. The DUE condition to be satisfied for each pair of zones is that no path can have a lower travel time than any used path between these zones for trips departing in a given time interval.

DUE is formulated in terms of link flows as a bilevel program (*6*). The upper problem solves for DUE flows subject to nonnegativity and conservation of flow constraints. The lower problem updates the node time intervals and ensures temporally continuous trip paths. The solution algorithm solves these two problems successively until suitable convergence is obtained. Link capacity reductions due to accidents or spillback queueing in specific time intervals are made between these problems. DUE as solved here maintains the first-in first-out (FIFO) ordering of trips between all zone pairs.

Department of Civil Engineering, University of Colorado, Denver, Colo. 80217-3364.

NETWORK DEVELOPMENT

The study network shown in Figure 1 contains 110 zones, 1,714 nodes, and 3,417 links, including intersection links. Zone centroids shown as circled dots in Figures 1 and 2 define the origin-destination (O-D) trip-end locations. All legal movements at every intersection are represented by separate links (see Figure 3). An intersection of two two-way streets requires at least 12 through and turn-movement links connecting eight approach and exit nodes. Special lane groups and allowed U-turns require additional links. Thus, in this network, 1,395 nodes and 1,778 links define intersection turn movements, and an additional 702 links are centroid connectors, which are explained later.

Figures 4 and 5 show the detail with which freeway interchanges are represented. The network contains 11 interchanges on I-25 and three on I-225. Figure 6 shows that all but two I-25 interchanges (those at I-225 and University Boulevard) have dual-loop detector arrangements as shown in Figure 7. Three cloverleaf interchanges (Colorado Boulevard, Arapahoe Road, and County Line Road) have full-loop detection at both northbound on-ramps.

Geographic Information System (GIS) software was used to code the network. The GIS platform proved very useful for displaying and analyzing the network links, and for editing and manipulating their attributes. All links are unidirectional, and their geographic representation is such that (a) no two links connect the same two nodes, and (b) each two-way street or highway section is repre-

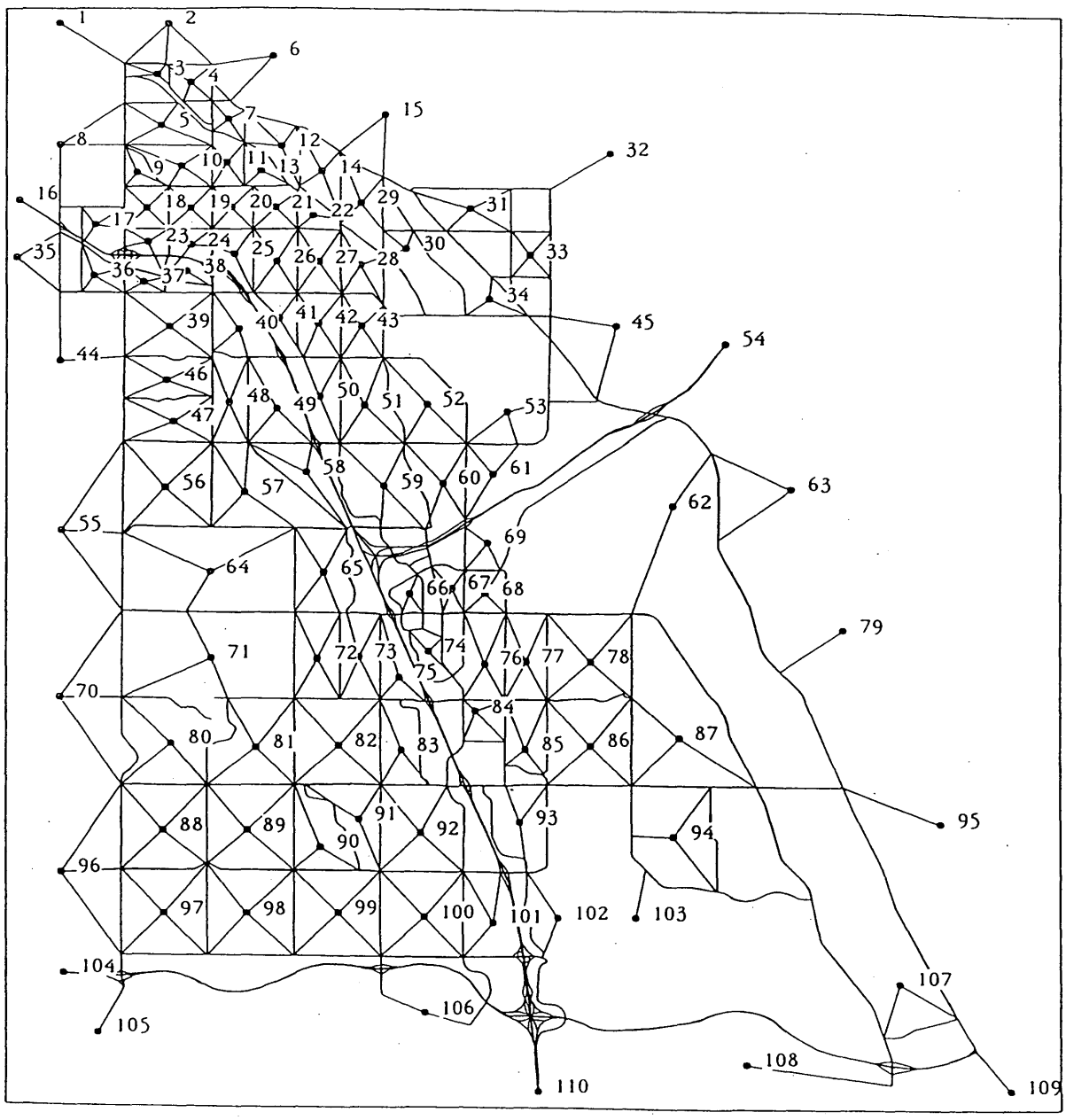

FIGURE 1 Southeast Denver network surrounding I-25.

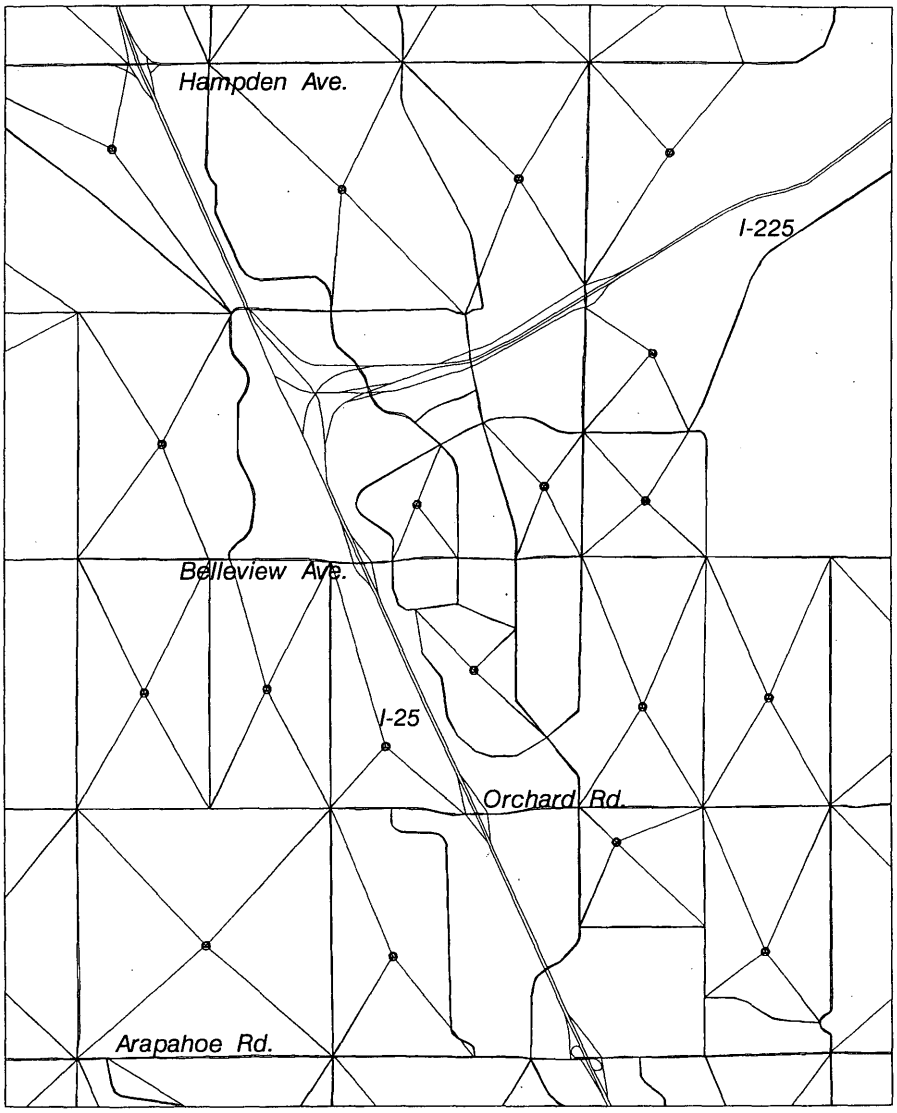

FIGURE 2 I-25 from Arapahoe Road to Hampden Avenue.

sented by two oppositely directed links connecting two distinct node pairs.

The main advantages of representing the network with this level of detail are that (a) delays incurred by vehicles at each through or turn movement at intersections can be estimated more accurately and (b) illegal turn movements cannot occur. This coding is required for the estimation of spillback queueing effects on the capacities of upstream links as explained by Janson and Robles (*6*). The disadvantages of coding a network with too much detail are that (a) larger numbers of nodes and links result in a greater computational burden, and (b) developing the network can become very time-consuming, depending on the size of the study area and the data available.

Perhaps the most tedious task is converting digital line graph (DLG) files into usable form for traffic modeling. This involves dividing every link into two oppositely directed links, and then adding separate turn movement links at each intersection and interchange. No automated GIS procedures or utilities have been developed to perform this task, so a program was created to split the links and make some intersection connections. However, nearly every intersection still required some reconfiguration with the GIS network editing tools.

To incorporate any road or intersection changes that may have occurred since the DLG files were created, survey trips were made to various places throughout the study area to verify (a) the configuration of roads and intersections, (b) number of lanes, (c) turning allowances, and (d) new roads constructed. The entire network building and conversion process required several months of full-time effort by the first author of this paper.

In addition to geographic alignment coordinates, supply attributes of each link stored in the link layer of the data base include:

1. Link identifier,
2. From-node,
3. To-node,
4. Directionality code,
5. Length,
6. Link name,

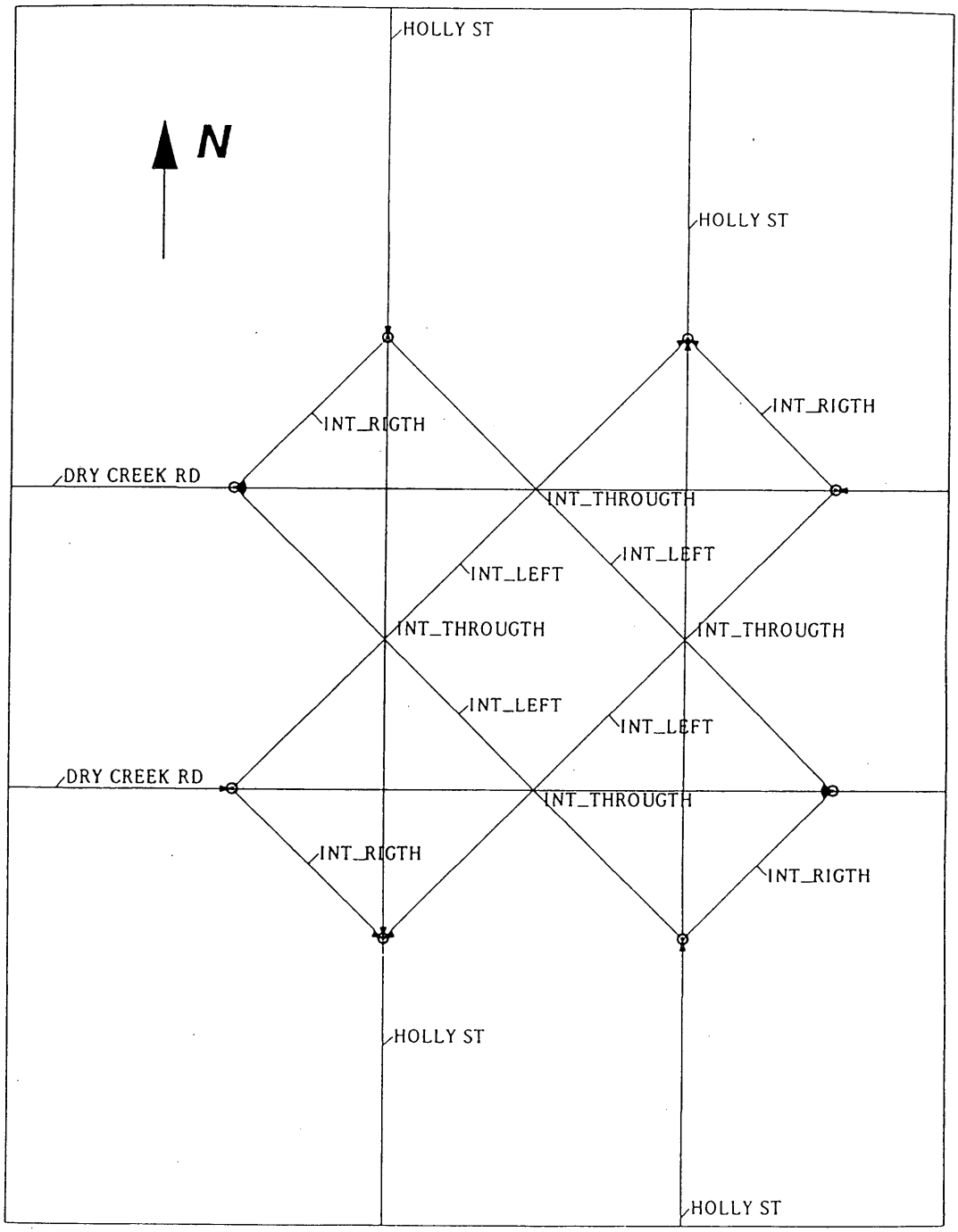

FIGURE 3 Intersection representation.

7. Road class,
8. Number of lanes,
9. Capacity,
10. Speed limit, and
11. Free flow travel time.

Some of the link data (road names, lengths, speed limits, number of lanes, and road classes) were obtained from the data bases of other transportation agencies and matched to the network. Discrepancies found in some of these data were corrected by checking maps and conducting field trips. Fields for importing modeling results (e.g., link volumes, volume to capacity ratios, link travel times and speeds, etc.) were also created. By comparison, the attributes of each node stored in the node layer of the GIS data base were limited to its ID, *X*-coordinate, and *Y*-coordinate.

The next step was to select a sufficient number and coverage of zone centroids and their linkages to the network. A zone centroid was located within every block sub-area surrounded by signalized arterial streets, plus external centroids surrounding the region's boundaries. Zonal areas ranged from less than 0.25 mi^2 (0.65 km^2)

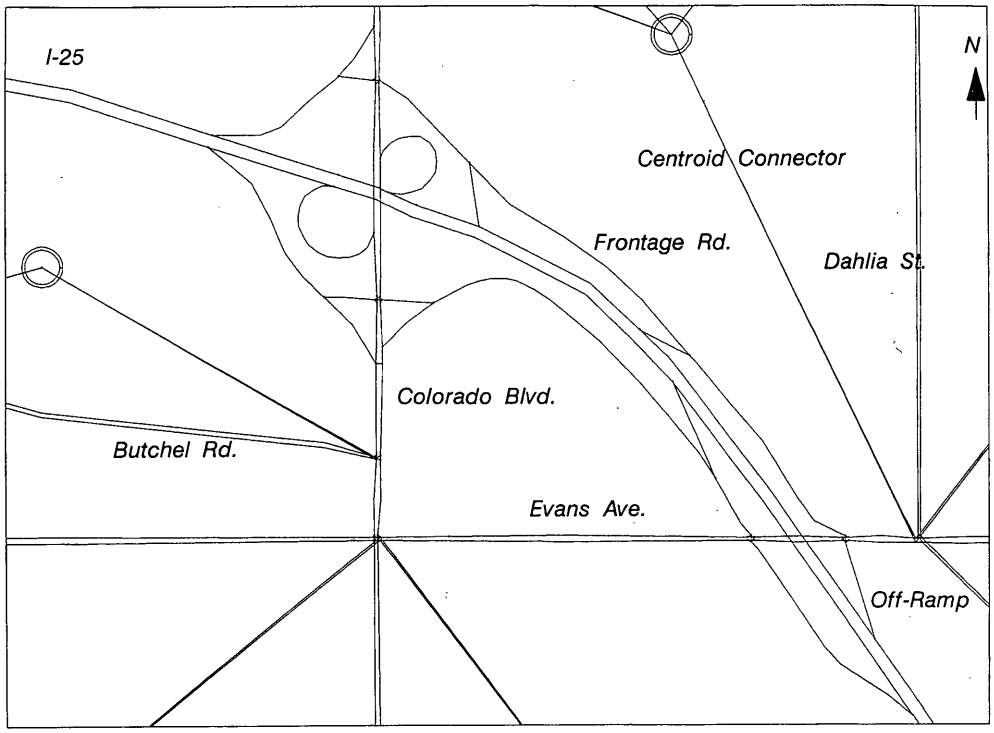

FIGURE 4 Locations of accidents No. 1 and No. 3.

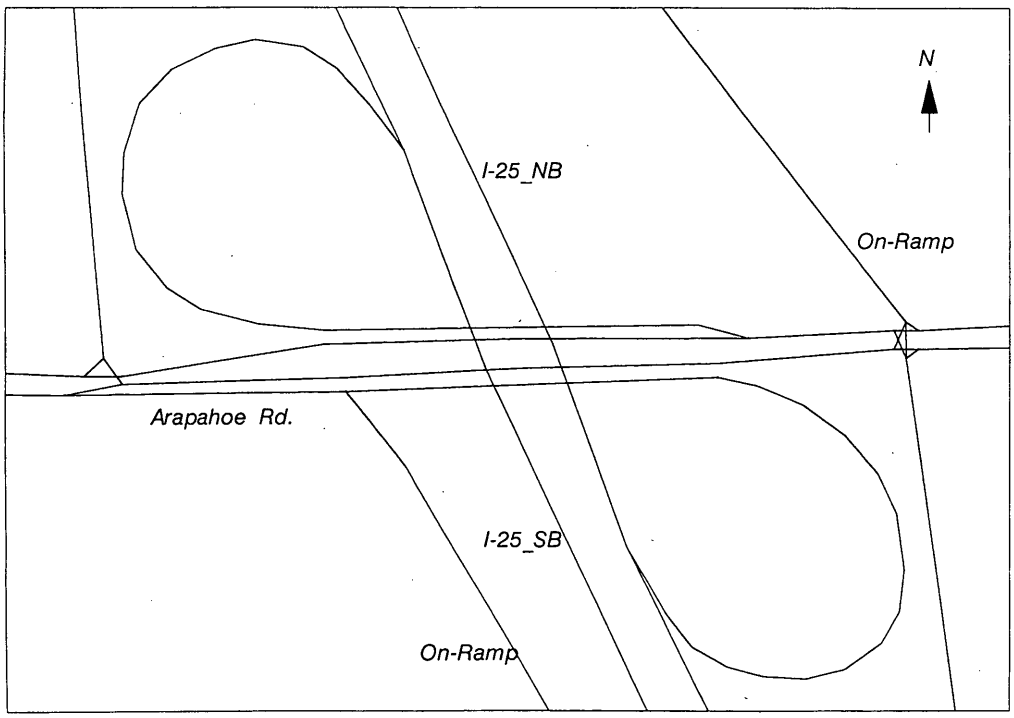

FIGURE 5 Interchange at Arapahoe Road and I-25.

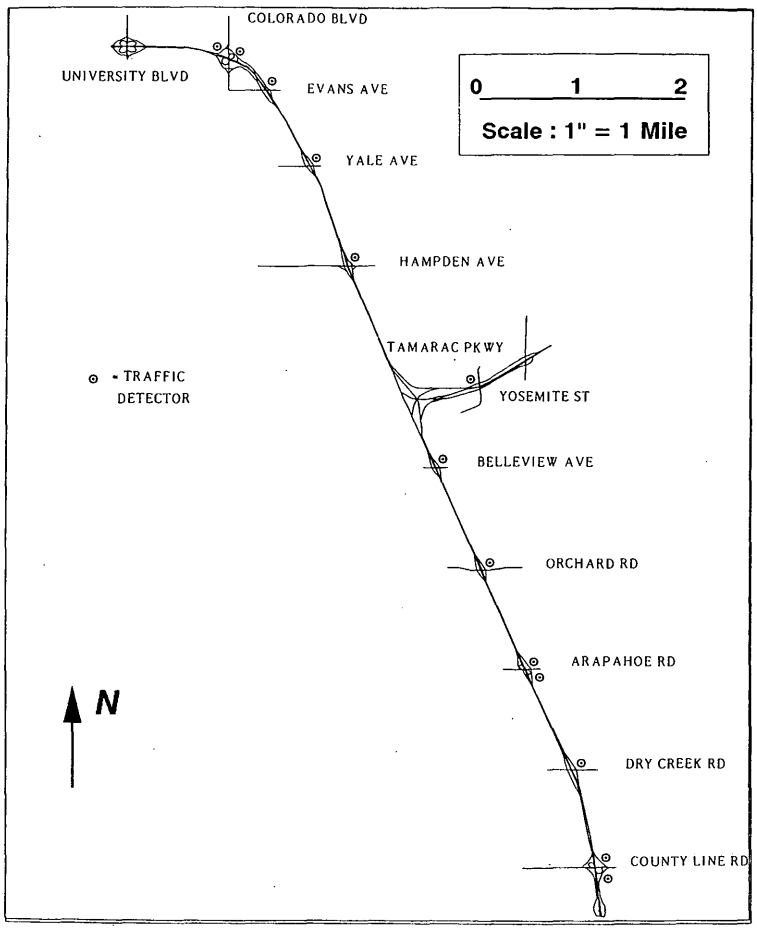

FIGURE 6 I-25 interchanges.

in the northern, denser end of the network to greater than 2 mi^2 (5.2 km^2) in the southern, sparser end of the network. Each zone centroid is connected to the approach or exit nodes of 2 to 4 intersections located on the boundaries of the zone. Access links connect before intersections and egress links connect after intersections.

Trip O-D and Departure Time Estimation

Having defined the traffic analysis zones, the next step is to create a peak-period trip matrix of interzonal trips for the study area. Obtaining an O-D trip matrix by dispatching surveying trip makers is expensive, labor intensive, prone to sampling errors, and not feasible for real-time applications. An alternative is a synthetic technique that uses traffic counts on alternative routes to develop an O-D table.

A conventional zone-to-zone trip distribution matrix represents trips between each O-D pair of zones in a given analysis period. Whether trips depart and arrive within the time period, and when trips actually travel within that period, is unknown once the trip matrix has been compiled from survey data. If the time period of a trip matrix is shortened (e.g., from 1 hr to 5 min), then most trips departing in any given interval will not be completed within that interval. Hence, trip matrices for short time intervals represent "trip departure" matrices of trips departing from each zone to the zone they are headed (or trips arriving at each zone and from the zone they came).

The preceding discussion raises many questions concerning how to estimate O-D trips and their departure times in a combined or sequential manner. Janson and Southworth (*5*) describe a method of using traffic detector data to disaggregate a peak-period trip matrix into the likely departure times of these trips. A prohibitive disadvantage of estimating departure times by O-D pair is that it requires a large coverage of 24-hr detectors reporting data about every 5 min. Moreover, there was no average weekday peak-period trip matrix covering 5:00 to 10:00 a.m. to disaggregate. The Denver Regional Council of Governments (DRCOG) uses either a 24-hr or 3-hr trip matrix for most of its work. Extracting a 5-hr trip matrix for the study area from the regional DRCOG data base with somewhat different zone configurations did not appear to be reliable or up-to-date with current traffic patterns.

It was beyond the scope of this study to collect the necessary data and calibrate the travel demand models with which to estimate trip productions, attractions, and O-D trips. Such a process would also be error-prone due to the large number of external trip ends in this region. Thus, despite the difficulties from traffic counts, O-D estimation was deemed to be the best strategy for developing a 5-hr trip matrix for which approximate departure times could be obtained. In addition to the I-25 detectors, traffic count information was requested from every state, city, or town agency in the area known

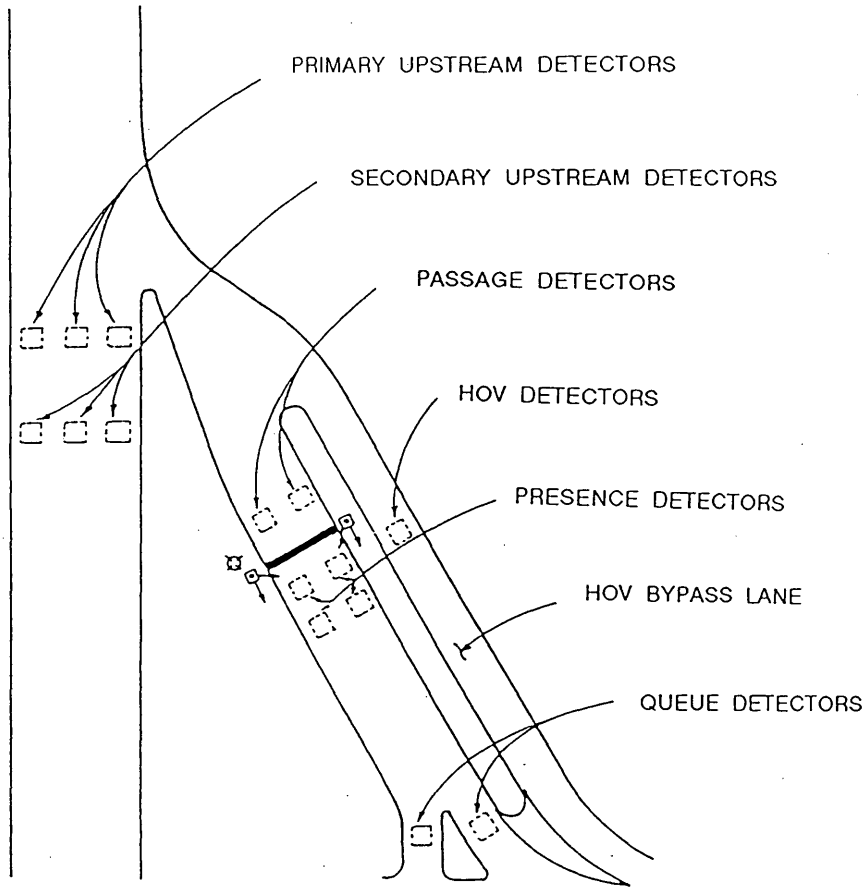

FIGURE 7 Typical metered freeway ramp on I-25.

to have some counts. A mix of 1-hr and 24-hr counts dating back to 1990 was obtained for approximately 20 percent of the network links, including turn movement links. Peak-hour counts from the I-25 detectors were then added to this set. Much reconciliation and judgment were required to pull together a set of usable peak-hour counts from this data.

O-D estimation must reconcile observed counts with link-use probabilities and network flow feasibility. In the study, maximum entropy O-D estimation (with a base-trip matrix and observed link counts) and static user-equilibrium (UE) assignment were repeatedly executed to obtain link-use probabilities, which resulted in an O-D trip matrix that when assigned to the network resulted in similar link-use probabilities. As a base-trip matrix, an approximate pattern of peak-hour O-D trips were obtained from DRCOG, but not one that was compatible with observed counts or up-to-date with recent traffic growth.

The entropy maximizing model for O-D estimation was programmed from traffic counts as described by Van Zuylen and Willumsen (*7*) with the base O-D trip matrix previously mentioned as prior information to improve the reliability of the estimated O-D matrix. Because of relatively sparse count coverage on a fairly large network, the O-D estimation process was broken down into the following steps:

1. Assign a base peak-hour trip matrix to the network with static UE assignment to obtain initial link-use proportions. Also calculate the sum of these base origins and destinations for use in the next step.
2. Use the maximum entropy procedure to separately estimate origins and destinations from traffic counts in proportion to base origins and destinations from the base-trip matrix using the link-use proportions just obtained.
3. If newly estimated origins and destinations are within a small percent change from previously estimated origins and destinations (equal to the base origins and destinations when this step is first executed), then STOP. Otherwise, continue.
4. Distribute these trip ends in a biproportional manner to the base-trip matrix without using a trip deterrence function because of the large proportion of pass-through trips.
5. Use static UE assignment to assign the estimated trip matrix from Step 4 to obtain a new set of link use proportions, and return to Step 2.

Since there is no assurance that the above procedure will converge, an added step before UE assignment in Step 5 that will "force" convergence involves combining the latest trip matrix with the previously estimated matrix using the method of successive averages. This step was not necessary for the application studied here.

The advantage of estimating origins and destinations and then O-D trips in separate steps is that each step requires much less computational burden than (a) generating the entire three-dimensional

matrix of O-D link-use proportions, and (b) estimating the full O-D trip matrix from traffic counts. Also, each matrix of link-use proportions by trip-end zone is much less sparse than a comparable matrix of link-use proportions by O-D pair. The disadvantage of this approach is that it does not use O-D-specific link-use information. Since the Frank-Wolfe assignment algorithm linearly combines successive trip assignments by origin to shortest path trees, it is unclear whether this poses any disadvantage to the outcome of the procedure just explained.

Use of I-25 Loop Detector Data

In defining a study area for dynamic traffic modeling, a key issue is the availability of 24-hr loop detectors from which volumes, speeds, and densities can be obtained for short time intervals (less than 5 min). On I-25 southeast of Denver there are 12 locations at which loop detectors monitor traffic using the northbound through lanes and on-ramps of each interchange for ramp meter operation (see Figures 6 and 7). Volumes and speeds for 65-min intervals of the morning peak-period (5:00–10:00 a.m.) were archived to tape by the Colorado department of transportation for each detector for each day from June 15 to September 15, 1992. These data were used to calibrate the travel time functions of the model. The best-fitting parameters for the Bureau of Public Roads impedance function were found to be 0.7 (instead of 0.15) multiplied by the volume-to-capacity ratio raised to the power of 6 (instead of 4).

These data also were used to factor the peak-hour trip matrix into a 5-hr trip matrix. Because a large percentage of trips in this study area use I-25 for some portion of their journeys, the 5-hr trip matrix was estimated as a scalar multiple of the peak-hour trip matrix that best fit the 5-hr counts on I-25 when assigned in a static UE manner. The best-fitting multiple was found to be 3.2 for this study network. The final 5:00–10:00 a.m. trip matrix represents a total of 222,218 trips, with nonzero trips between each of the (109 × 110) interzonal O-D pairs. Intrazonal trips are not modeled.

The next step was to disaggregate the 5-hr trip matrix into trip departure times. Departure time estimation can be combined with DYMOD as described by Janson and Robles (*4*), but this requires knowledge of desired arrival times and schedule delay penalties of trips by origin zone. Instead, it was assumed that departure times from each origin were distributed similarly to departure times of trips using I-25 (for which there were link crossing times in 5-min intervals). It was assumed that the distribution of departure times from each origin in 5-min intervals was similar to the distribution of I-25 entrance volumes at the nearest on-ramp, but offset by the approximate travel times from each origin to their nearest I-25 on-ramp. This approach is an ad hoc execution of the more formal procedure [defined by Janson and Southworth (*5*)] that worked well for this network.

DISCUSSION OF RESULTS

Figures 8 and 9 show observed and predicted volumes and speeds at three northbound through links of I-25 at detector sites just before the merge points of each northbound on-ramp. In general, predicted speeds declined much earlier than actual speeds, beginning about 6:30 a.m. at each location. The model overestimates speed reductions in most, but not all, cases. This result is satisfying because traditional static models are often criticized for grossly underestimating travel times and delays. Speed comparisons at other interchanges were generally better than the ones shown here.

Figures 10 and 11 show observed and predicted volumes at three off-ramps and three on-ramps at the same interchanges. Detectors were not located at the off-ramps, but observed off-ramp volumes were computed based on the observed volumes on adjacent links. The predicted off-ramp volumes in Figures 10 and 11 generally agree with observed volumes, disregarding the dramatic fluctuations in ramp volumes, which the model is not intended to predict. The off-ramp volumes at Belleview and Evans Avenues are most poorly predicted, but the observed Evans Avenue off-ramp volumes

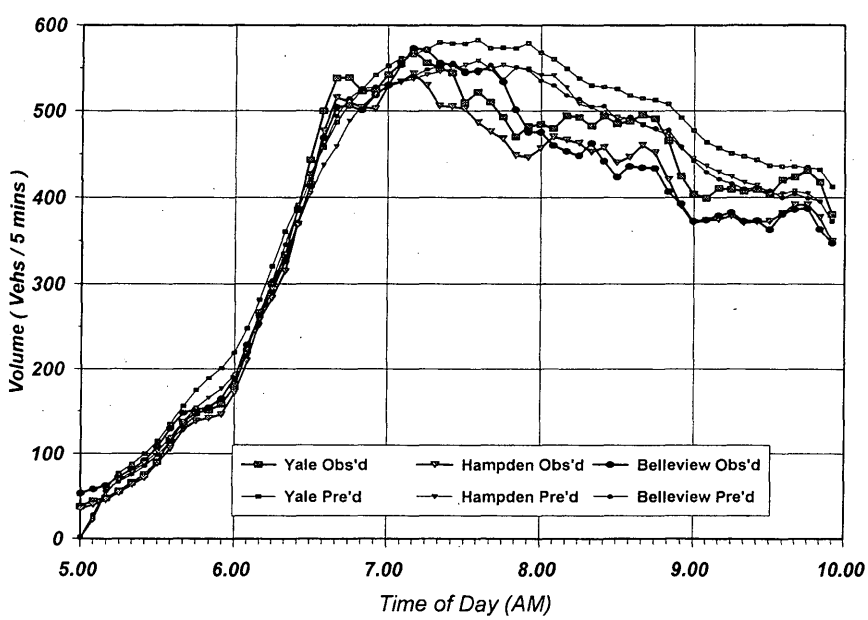

FIGURE 8 Predicted versus observed I-25 volumes.

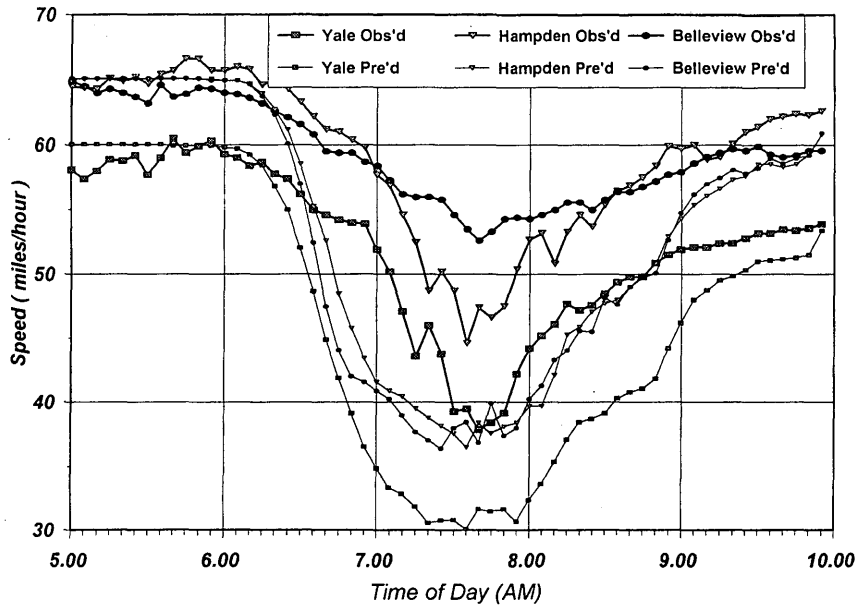

FIGURE 9 Predicted versus observed I-25 speeds.

are very curious. This may be due to some observed data problems. The Colorado southeast cloverleaf on-ramp also shows large disparities and appears to be an underutilized on-ramp compared with the adjacent northwest diamond on-ramp.

Analysis of Lane-Blocking Accidents

A key feature of DYMOD is that it adjusts upstream link capacities for spillback queueing (a) from oversaturated links in specific time intervals or (b) due to accident blockages, weather conditions, construction, or signal timing changes in time intervals when they occur as input to the program. Spillback queueing effects are difficult to predict because of the speed at which they develop and their proportional effect of multiple inflow links to the same intersection or freeway merge section.

Accident Case No. 2 Near Belleview Avenue

Figure 2 shows the section of I-25 along which two accidents occurred on October 20, 1992. According to the Mile High Courtesy Patrol (a motorist assistance service on I-25), the first accident occurred near Hampden Avenue at 7:10 a.m. and was cleared at 7:30 a.m. It occupied the right shoulder and part of the right lane, causing approximately 30 percent of this four-lane section's capac-

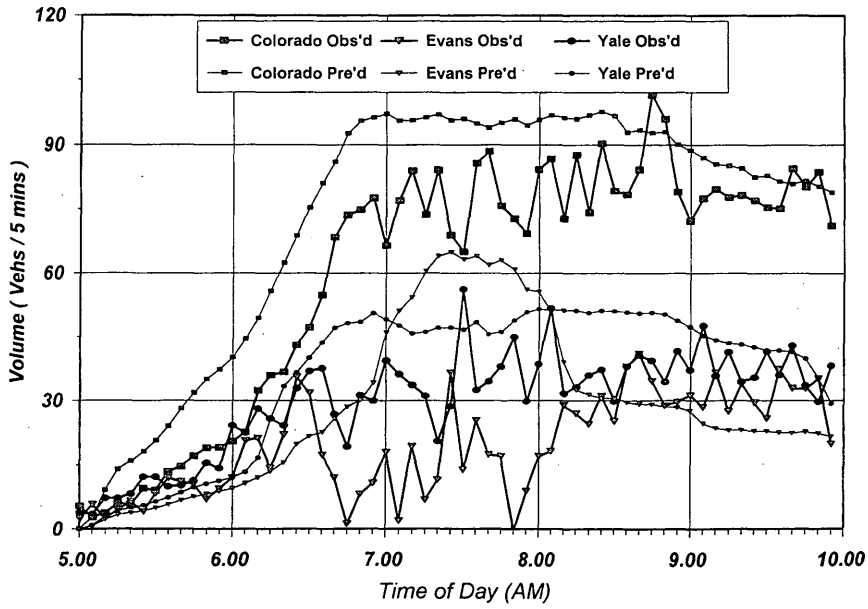

FIGURE 10 Predicted versus observed off-ramp volumes.

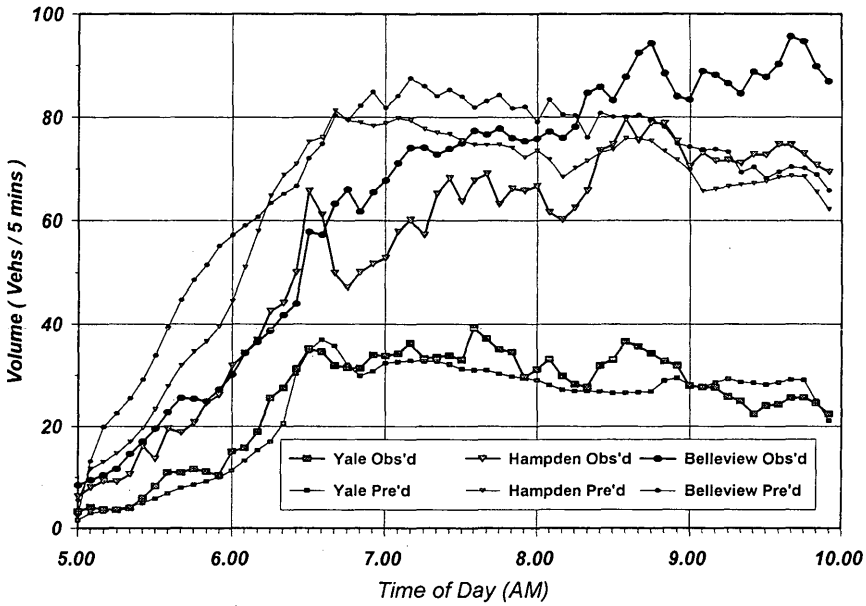

FIGURE 11 Predicted versus observed on-ramp volumes.

ity to be lost for that period of time. The second accident occurred near Belleview Avenue about 7:15 a.m. and was cleared about 8:25 a.m. It was reported in the right lane and part of the adjacent lane, causing about 50 percent of this four-lane section's capacity to be lost for that period of time.

Figure 12 compares predicted and observed I-25 volumes on the day of the second accident. A very close agreement was observed at the Belleview and Orchard detectors, but not at the Hampden detector. The capacity loss and traffic impacts were underestimated due to the first accident at Hampden. Observed volumes rise to the predicted volumes when the first accident is cleared, but drop sharply after the second accident. DYMOD shows high traffic volumes at Hampden during the second accident because DYMOD rerouted traffic onto I-25 via the Belleview on-ramp (which was beyond the accident) and via I-225. These alternate routes could have been used more effectively by many travelers to avoid the accident queue if they had known the location of the accident. Travelers who diverted from I-25 at Orchard apparently did not attempt to reenter I-25 until much farther north, if at all, or they simply waited in the queue on I-25.

Figures 13 and 14 show very good predictions of travel times and speeds on I-25 during the accident. Upstream effects at Dry Creek Road (3 to 4 mi upstream) are shown in Figure 15. DYMOD captured more upstream effects in this case because of less acces-

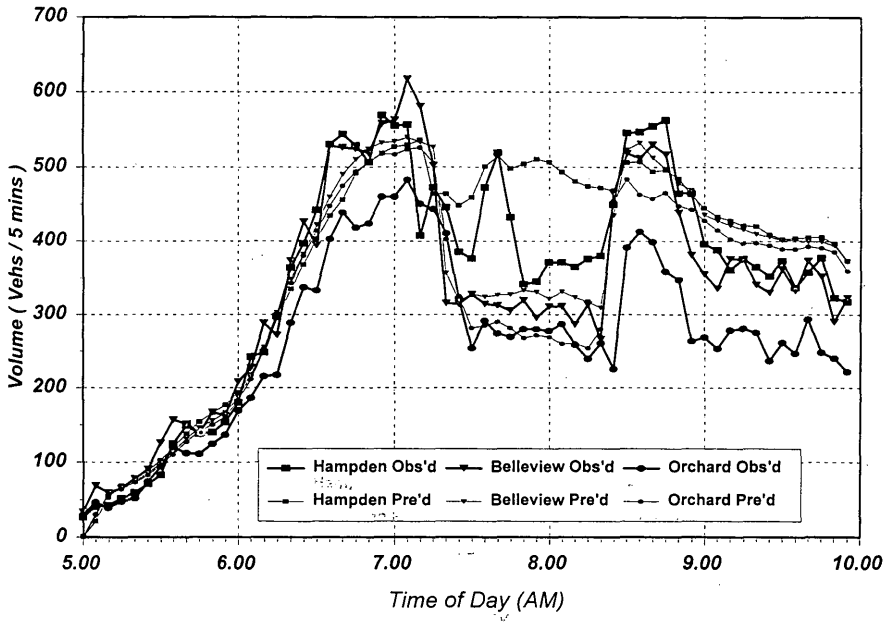

FIGURE 12 Predicted versus observed I-25 volumes (Accident 2).

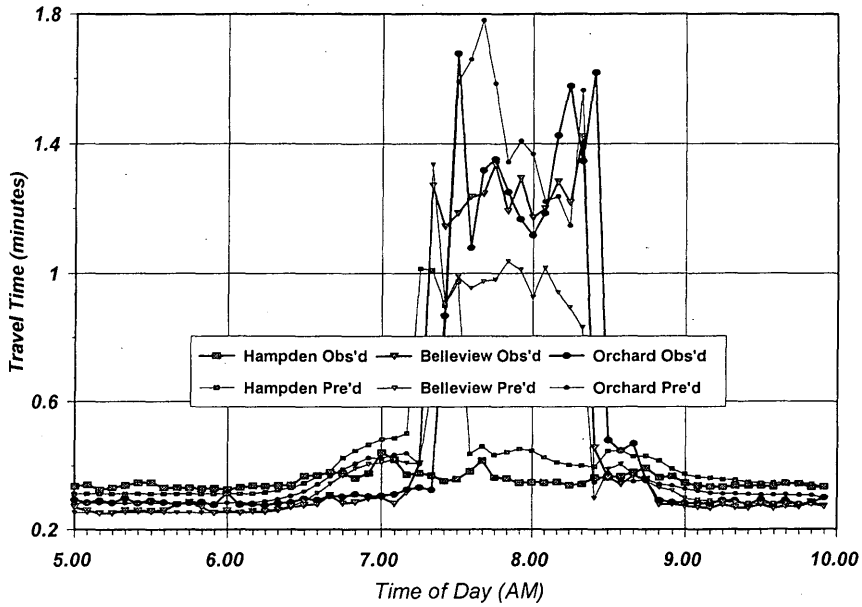

FIGURE 13 Predicted versus observed I-25 times (Accident 2).

sible alternate routes around this site compared with other sections of I-25.

Summary of Estimated Accident Delays

The table shown summarizes accident delays estimated by DYMOD compared to travel times estimated by DYMOD without any accidents. The following observations may be made:

- Case 1 caused the least total hours of delay (742 hr), but the most delay per directly affected trip.
- Case 1 was of short duration, but caused a 50 percent reduction in capacity of an already narrow (3-lane) section of I-25, and happened at the peak of rush hour.
- Cases 2 and 3 were of much longer duration, but caused less capacity reduction and occurred mostly on the downside of the peak period.
- Thus, Cases 2 and 3 directly affected more than twice as many trips, and caused nearly twice the total vehicle delay, but caused less delay per directly affected trip than Case 1.

These delay estimates are conservative in that DYMOD diverts trips to alternate routes as accident queues develop. In reality, many

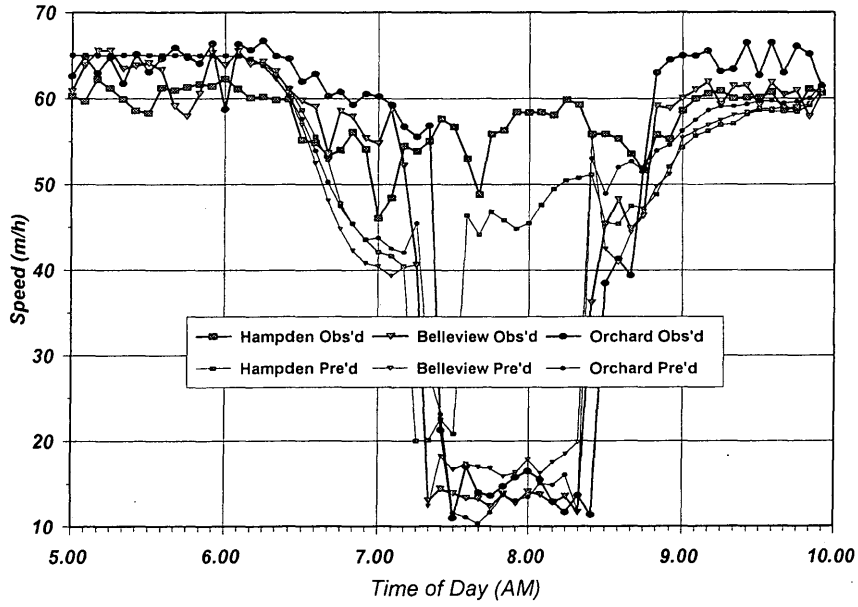

FIGURE 14 Predicted versus observed I-25 speeds (Accident 2).

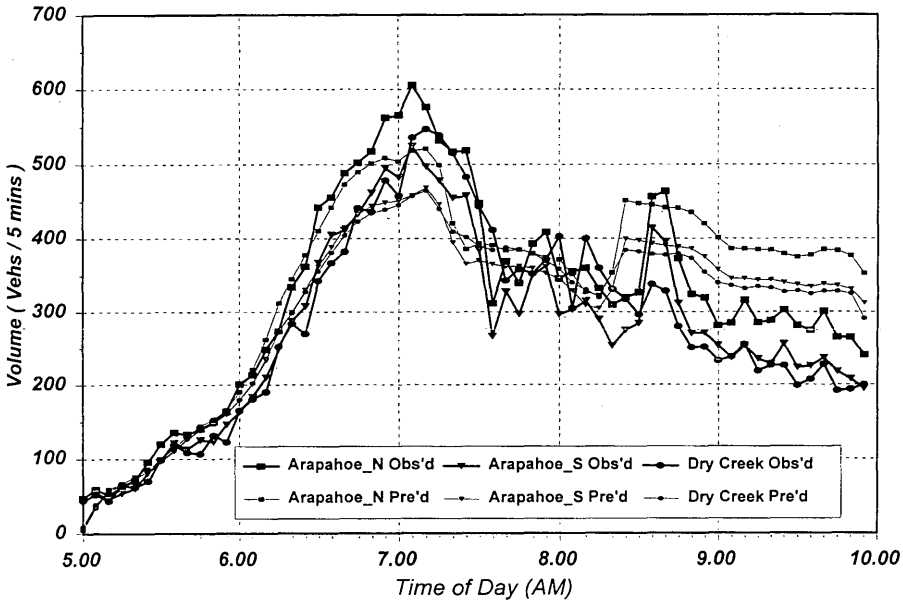

FIGURE 15 Predicted versus observed I-25 upstream volumes (Accident 2).

travelers do not so readily divert from accident queues because of not having good knowledge of alternate route locations and travel times. Estimates of queueing delay assuming less route diversion were approximated for the same accidents in an evaluation study of the Mile High Courtesy Patrol (*8*). Those estimates were approximately 50 percent greater than the ones just mentioned, but still conservative compared with other national reports. Since DYMOD represents route diversions with good alternate route information, the difference in these estimates indicates that significant delays could have been reduced by directing travelers to other routes and adjusting the signal timing along these routes to better handle the diverted flows.

CONCLUSIONS

The results indicate that incident delays can be significantly reduced using travel advisory systems in which further research and development is needed. The results also show that DYMOD can be used "off-line" to develop (a) proactive response plans for accidents at critical network locations, (b) work-zone traffic control and detour routing plans, or (c) traffic impact predictions for a major spectator event or storm. Dynamic traffic models also can be combined with microsimulation models of smaller network sub-areas to make finer traffic control adjustments.

Based on the information in this study, several general recommendations may be made:

- Dynamic traffic modeling yields much closer estimates of traffic conditions than conventional transportation planning models when applied to urban area networks during congested periods.
- The key to successful dynamic traffic modeling is the care with which the supply and demand data bases are developed. Much more detail is needed than is typical of conventional static models.
- Wider regional coverage of traffic detection must be a priority to support the successful development and operation of dynamic traffic modeling and route guidance from a traffic management center. O-D and departure time estimation is operationally the "weak link" in dynamic travel modeling because of such limited count coverage in most urban areas.

TABLE 1 Summary of Estimated Accident Delays

Evaluation Measure	Case #1	Case #2	Case #3
Total Delay (vehicle hours)	742	1426	1248
Number of Directly Affected Trips	3300	6600	7200
Delay per Directly Affected Trip (min)	13.5	13.0	10.4

Directly affected trips are the approximate number of vehicles that would have passed the accident location on I-25 during the accident in the base or "no accident" case.

Eventually, dynamic traffic models will be integrated with traffic control centers that respond directly to real-time conditions through adjustments to arterial signals, ramp meters, and messages sent to travelers. This study examined the implementation and performance of one approach that, for reasonably large networks, can be run concurrently on a high-speed computer with traffic detection input to provide updated travel advisories and traffic management information.

REFERENCES

1. Janson, B. N. Dynamic Traffic Assignment for Urban Road Networks. *Transportation Research,* Vol. 25B, No. 2/3, 1991a, pp. 143–161.
2. Janson, B. N. A Convergent Algorithm for Dynamic Traffic Assignment. *Transportation Research Record,* Vol. 1328, 1991b, pp. 69–80.
3. Janson, B. N. Network Design Effects of Dynamic Traffic Assignment. *Journal of Transportation Engineering,* ASCE, Vol. 120, No. 6, 1995, pp. 1–13.
4. Janson, B. N., and Robles, J. Dynamic Traffic Assignment with Arrival Time Costs. *Transportation and Traffic Theory,* Carlos F. Daganzo, Ed., Elsevier Science Publishers, Amsterdam, 1993, pp. 127–146.
5. Janson, B. N., and Southworth, F. Estimating Departure Times from Traffic Counts Using Dynamic Assignment. *Transportation Research,* Vol. 26B, No. 1, 1992, pp. 3–16.
6. Janson, B. N., and Robles, J. Quasi-Continuous Dynamic Traffic Assignment Model. *Transportation Research Record* (to be published).
7. Van Zuylen, H. J., and Willumsen, L. G. The Most Likely Trip Matrix Estimated from Traffic Counts. *Transportation Research,* Vol. 14B, No. 4, 1980, pp. 281–293.
8. Cuciti, P., and Janson, B. N. Incident Management via a Courtesy Patrol: Evaluation of a Pilot Program in Colorado. *Transportation Research Record* (to be published).

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

Route Choice and Information Use: Initial Results from Simulation Experiments

KENNETH M. VAUGHN, PRASUNA REDDY, MOHAMED A. ABDEL-ATY, RYUICHI KITAMURA, AND PAUL P. JOVANIS

Computer-based simulation experiments were designed and conducted utilizing 100 regular commuters from the Sacramento, Calif. region. Computer-based simulation was used to create a hypothetical traffic network and as a tool for collecting dynamic route choice data. Dynamic route choice data under the influence of incident information, en route guidance, pretrip guidance, and congestion information were collected for a sequence of 20 simulated trial days for each participant. A fractional factorial experimental design was used in the experiment. The experimental treatments included the type of information that subjects received, as well as personal characteristics of the subjects themselves. An analysis of the experimental treatments on subjects' accuracy perceptions, information effects on travel decisions, and consumer demand and information system valuation was performed utilizing analysis of variance and regression techniques with corrections for heteroskedasticity. Study results were compared with previous survey and simulation research. The findings indicate that there are significant differences between the individual characteristics that influence drivers' preferences about how they receive information and the characteristics that influence drivers' travel choices in response to receipt of such information. Results from the experiment indicate that incident and congestion information types were the most important in influencing subjects' route decisions. The findings also indicate the importance of having incident information available to drivers, and that this importance is reinforced by the availability of route guidance information. Descriptive information such as incident locations and congestion levels can provide a strong rationale for compliance with prescriptive information such as route guidance.

Research at the Institute of Transportation Studies, University of California at Davis is being performed to study the impact of Advanced Traveler Information Systems (ATIS) on travel demand. If information systems and services, which can provide accurate up-to-date information on the travel environment, are made available to drivers on board their vehicles, at home, or at high-demand locations such as office complexes or shopping centers, what will be the effect on travelers' behavior? Part of this project has focused on investigating the effect of information on route choice behavior. Our research efforts in this area have employed two state-of-the-art approaches undertaken in tandem.

The first approach used sophisticated computer-aided telephone interview (CATI) surveys of Los Angeles-area morning commuters, with a follow-up mail survey to a large subgroup of the original sample, which was customized for each survey recipient (*1,2*). The goal of this series of surveys was to investigate what routes commuters were using and identify the characteristics of these routes; to determine how much information drivers have about their routes and if they have alternative routes, to investigate what sociodemographic and route specific characteristics influence drivers to use alternative routes; and how the availability and perceived accuracy of existing information affect route choice.

The second approach was to use computer simulation as a data collection tool to investigate drivers' information use and learning with an ATIS. The computer is used to simulate a hypothetical traffic network that creates the decision framework into which subjects are placed. In the first year of this research project, an experiment to collect sequential route choice data under the influence of ATIS was performed using a personal computer (PC)-based simulation (*3–6*). The experiment collected information on drivers' pretrip route choice behavior at three levels of information accuracy, 60 percent, 75 percent, and 90 percent, utilizing a simple binary choice framework in which subjects could choose either a freeway route or an alternative side street. The experimental factors that were controlled in this first experiment included the accuracy level, stops on the side road route, presentation of a justification or rationale for the advised route, feedback of actual and alternative travel times, and identification of route as a freeway. The main findings of this first experiment were that drivers can rapidly identify the accuracy level of information being provided and that they adjust their behavior accordingly. There is also evidence that indicates an accuracy threshold level does exist below which drivers will not follow advice and above which drivers readily follow advice. It was found that female subjects agreed with advice more often than males, that less experienced drivers agreed more often than experienced drivers, and that a "freeway bias" exists with drivers because they were much more willing to follow advice to take a freeway route. The model of route choice behavior had a prediction rate that was 79 percent accurate and also indicated that strong memory effects existed in the updating of the perception of information accuracy. This finding indicates that subjects placed more emphasis on the accumulation of past experiences as opposed to just the latest experiences. Analysis of the experimental treatments revealed that subjects' compliance with advice increased with increasing system accuracy, by providing subjects with feedback and a decision rationale, but that intersection stops on the side street route significantly reduced advice compliance.

As an extension of this previous simulation work, a new PC-based travel simulator has been designed and a new set of experiments were designed and carried out (*7*). Some limitations of our previous experiments included the simplicity of the choice set in the binary framework, the limit to a pretrip information context imposed by this framework, and a limited investigation of sociodemographic and travel characteristics imposed by the use of university students as test subjects. This new set of experiments was

Institute of Transportation Studies, University of California at Davis, Davis, Calif. 95616.

developed to expand the simulation and experimental design complexity to account for these previous limitations.

The main goals in the development of the simulator were to expand the traffic network to improve the realism of the choice set and to allow for investigation of route-, link-, and node-specific characteristics on subjects' decision processes. The main goals of the experimental design were to test hypotheses about the main and interaction effects of four information treatments and three demographic factors, which included incident information, route guidance information, pretrip information, congestion level information, gender, age, and education. To investigate the influence of sociodemographic and travel characteristics, regular commuters and carpoolers were recruited as test subjects.

This article summarizes the initial findings of the data collected during this set of experiments. The network and travel simulator and the experimental design are briefly reviewed. Basic statistical comparisons and an analysis of frequency tables were utilized to investigate the sociodemographic, travel, and information use characteristics of the sample. Analysis of variance (ANOVA) techniques, ordinary least squares (OLS), and generalized least squares (GLS) regression techniques are used to analyze the effects of the experimental treatments on subjects' perceptions of accuracy, on their travel decisions, and on measures of consumer demand and system valuation. The significant findings of this analysis and comparisons to previous findings are summarized in the Conclusions.

DESCRIPTION OF EXPERIMENT

Overview

The driving simulation was conducted on a DOS-based PC that simulated a driving environment and provided different types of traffic information to the subjects. Subjects were randomly recruited from the Sacramento, Calif. area by an independent recruiting firm. One-hundred people were recruited according to a specific criteria that categorized them by commuter type [single occupancy vehicle (SOV)/carpool], gender, age, education level, and driving experience. The experiment consisted of a brief preexperiment survey followed by the use of the driving simulator. All subjects who completed the experiment were paid an incentive of $50.00 in cash.

Participants using the simulator were asked to "drive" from a specific starting point to a specific destination point on the computer screen. There were a series of paths to choose between representing roadways and a freeway. Figure 1 is a sketch of the simulation screen. To help them navigate through the network, the simulator offered four types of traveler information:

- Incident information,
- En route guidance,
- Pretrip guidance, and
- Congestion information.

All participants were preassigned a treatment for their 20-day simulation. Treatments consist of some combination of the four types of information, all types of information, or no information, depending on the experimental design cell to which the subject belonged.

Network and Travel Simulator

The simulation is an interactive PC program. The screen display is composed of three main windows: a network window, an information window, and an instruction window (see Figure 1 for typical screen display). The network window displays a hypothetical traf-

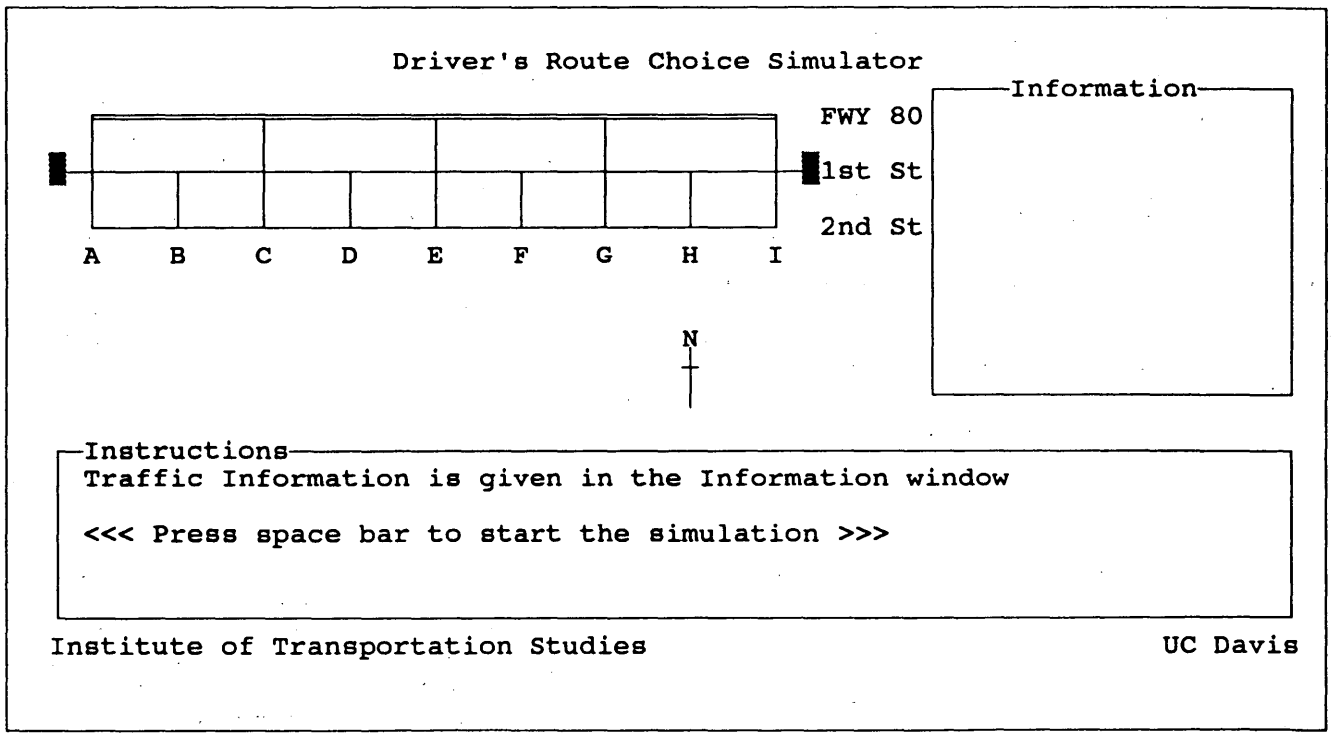

FIGURE 1 Screen display of simulator.

fic network, composed of three primary routes from an origin to a destination. The primary routes are composed of a freeway route and two arterial routes. These primary routes are cross-connected with a series of surface streets creating a network of 34 roadway links and 23 intersections (or potential decision points). The travel environment is generated by a stochastic assignment of travel speeds and stop delays to the network links and nodes. A random incident generator is used to assign an incident of random severity to the network for each travel day. The information and instruction windows are used to simulate an in-vehicle information system. Participants use keyboard arrow keys to represent driver route choices in the network for a sequence of 20 travel days (trials).

The use of the normal distribution is suggested (*8*) to represent speed distributions on roadway links. For this simulation, the network characteristic subroutine created normally distributed link speeds for each link in the network. These link speed assignments were established independent of each other in the current configuration of the simulation program. The program created a scenario in which the three primary routes through the network had similar mean speeds over the sequence of trials, but the variance in travel speed was greater on the freeway than on the arterials, and the surface streets had lower mean speeds but experienced little variation.

The simulation also contained an incident generator that stochastically assigned a traffic incident with a random severity level to one roadway link on each trial day. Links with incidents assigned had their link speed assignments automatically reduced to represent the effects of the incident. The structure of the program also allowed for assignment of incidents to intersection nodes, which would then affect speed assignments on multiple links. This option was not utilized in this series of experiments and all incidents and their effects were confined to a specific link. For a full description of the design and workings of the network and travel simulator see Vaughn et al. (*7*).

Experimental Design

The experimental design selected for this simulation study is a one-fourth fraction of a 2^7 factorial design (*9*). With two levels for each factor, a full factorial design would require a minimum of $2^7 = 128$ runs. A one-fourth fraction of this design can be used requiring a minimum of only 32 runs, and if three-way and higher interactions are assumed to be negligible, then all of the main effects are estimable and 15 of the 21 two-way interactions are also estimable. With three subjects per design block, 96 subjects are required.

Information Treatments

The simulation applied four information treatments and used three blocking factors to make up the seven experimental treatments. The information treatments are labeled A through D and the blocking factors are E through G. All treatments have two levels and are described as:

A. Incident with Description (on/off). A red icon is displayed at the location of a severe incident and a yellow icon is displayed at the location of a moderate incident. In the information window, text is used to describe the location and classification of the incident, for example, "Severe injury accident on First Street between F Street and G Street." The incident information is displayed both pretrip and en route. When the subject begins a trial day, the incident location is initially displayed and remains displayed during the trial.

B. En Route Guidance (on/off). Graphical arrows indicate advised turning movements and text description of advice. At every node, the information system provides turning movement recommendations for the next node in the form of a blinking arrow for a left turn, a right turn, or continuing straight ahead. Also, text information is provided such as "go north on B Street." The turning movement advice is based on the computed minimum path from the current cursor position to the destination. As subjects make decisions and move through the network, the minimum path is recomputed.

C. Pretrip Guidance (on/off). The minimum path is displayed at beginning of the trip. At the start of the trial, the initial calculated minimum path from the origin to the destination is outlined on the network, remains until the subjects make their first decision, and then is turned off as they are under way.

D. Congestion Information (on/off). There are color-coded links for moderate and severely congested links, with green indicating normal congestion, yellow indicating moderate congestion, and red indicating severe congestion. Like incident information, congestion information is displayed both pretrip and en route.

Three blocking factors are:

E. Gender (male/female). If subjects are male, they are in the high level of the gender factor; female subjects are in the low level of the factor.

F. Age (young/old). If subjects are in the young age group (40 years old or less), then they are in the high level of the age factor, whereas subjects in the older age group (greater than 40) are in the low level of the factor.

G. Education (high/low). If subjects are in the high education group (some college or more), they are in the high level of the education factor, whereas subjects with low education (high school graduate or less) are in the low level of the factor.

These three factors are subject characteristics that were used for recruitment and to measure individual effects on driver performance. This experimental design provides estimates of the following effects:

Main effects: A,B,C,D,E,F,G

Two-way interactions: AB,AC,AD,AE,AF,AG,BC,BD,BE,BF, BG,CD,DE,DF,DG

Aliased two-way interactions: CE = FG, CF = EG, CG = EF

SAMPLE COMPARISONS WITH STATE AND LOS ANGELES-AREA SURVEY

Table 1 summarizes and compares several demographic, travel, and information use characteristics of the simulation subjects with the California statewide travel survey (1991), as well as our previous survey of route choice from the Los Angeles area (*1,2*). The Los Angeles-area survey was a CATI survey of commuters' route choice and current information use. The statewide travel survey was a traditional trip diary survey. The characteristics that were found to be significantly different from the state survey were home ownership, commute time, and travel method. The travel method to work was a characteristic that was controlled in the sampling to ensure a significant number of carpoolers; therefore this significant difference from the state mean was to be expected.

TABLE 1 Sample Comparisons

Variable	Simulation (Sacramento)	Survey (Los Angeles)	State Survey
Demographic and travel characteristic comparisons			
Gender(male)	50.5%	51.3%	46.9% (t=.71)
Homeowner	65.3%	59.0%	76.0% (t=2.22)
Household size	2.95	3.35	3.214 (t=1.78)
Income	$42,095	$38,750	$42,021 (t=.04)
SOV commuter	74%	78.8%	91% (t=3.70)
Commute time(min)	28.4	31.9	21.77 (t=4.28)
Commute Distance(mi)	13.8	12.8	-
Comparisons of Information Use			
Listen to pre-trip	43.4%	36.5%	-
Listen to en-route	70.0%	51.3%	-
Listen to both	38.4%	27.6%	-
Rated information extremely accurate or very accurate	40.0%	50.0%	-
Change departure time due to pre-trip information	37.2%	41.1%	-
Change route due to pre-trip information	69.8%	44.2%	-
Change route due to En-route information	60.0%	43.1%	-

The significantly higher average commuting time may also be linked to some self-selection bias, because subjects who commuted shorter distances were less likely to volunteer for the simulation because it required driving a considerable distance to the University of California Davis campus. The state survey data does not have trip distance information for comparison, but it is reasonable to assume that the sample also has inflated travel distances based on the inflated travel time. A comparison between the Sacramento simulation subjects and the Los Angeles survey indicates that although the average travel time was less for the Sacramento subjects, the average commute distance was greater in Sacramento, indicating the significant differences in traffic congestion that exist. Similar findings included that people tended to listen to traffic reports while en route more than before leaving home, and route modification in response to information seemed to dominate over departure time shifts. Both the receipt and use of information tended to be higher in the Sacramento sample. Accuracy perceptions of the information drivers received were found to be very similar between the two studies with 40 percent of Sacramento subjects rating information as either extremely accurate or very accurate, and 50 percent of those in Los Angeles giving a similar ranking. As previously mentioned, females listened to pretrip traffic reports at a higher rate than males, and this was a common finding between the two samples. Another comparable result between the two samples was that of work schedule flexibility.

There were also some common findings between the two studies on the effect of individual characteristics on information acquisition and use. The trend that higher income individuals were less likely to listen to pretrip reports was common between the two studies but not significantly so in the Los Angeles survey. Significant findings that were common in both studies included: females were more likely to listen to pretrip traffic reports, longer distance commuters were more likely to listen to en route traffic reports and also more likely to change routes because of pretrip traffic reports, and subjects with higher education levels and higher incomes were less likely to change departure time because of pretrip traffic reports.

ANALYSIS OF INFORMATION TREATMENTS

At the completion of the simulation experiment, subjects were asked five questions related to their experiences with the simulation. These questions rated subjects' perceptions of the accuracy of the information received in the simulation, the importance of the information in subjects' route selection, subjects' willingness to purchase an information system, and the potential price they might be willing to pay for such a system. The questions have been labeled EQ1 to EQ5 for reference in the analysis. The responses to these questions were treated as the dependent variables for the analysis performed in this section. The wording of these questions was as follows.

EQ1. How accurate was the information you received during this simulation?

EQ2. In general, how important was the information you received in determining which route you followed during the simulation?

EQ3. If a traffic information system similar to what you experienced in this simulation were available on the market today, how likely would you be to purchase such a system?

EQ4. If you were to buy such a system, how much would you be willing to pay?

EQ5. If such a system were only available as a monthly service, how much would you be willing to pay to subscribe to such a service?

In order to investigate the effects of the experimental treatments on the subjects responses, ANOVA models were used. When a background variable (covariate) was strongly related to the dependent variable, an analysis of covariance could increase the precision of comparisons between treatments by reducing the within-group variability in the dependent variable because of the influence of the covariate. The objective then was to test appropriate hypotheses about the treatment effects and to estimate these effects. For hypothesis testing, the model errors were assumed to be normally and independently distributed with zero mean and constant variance σ^2 (*9*). Additional assumptions of the covariate model were that within each group, the dependent variable had a linear relationship with the covariate, and that the slope of the regression for each covariate was the same in each group. The strength of the ANOVA model, and the main reason it was applied in this study, was that it did not require making assumptions about the nature of the statistical relation (except for the covariates), nor did it require that the independent variables be quantitative (*10*). Full models that contained all 25 effects were estimated for each of the five dependent variables (*11*). Reduced form models were obtained by excluding all effects that were not considered at least marginally significant (i.e., at the 90 percent level) and are presented in this paper.

The constant vectors of the ANOVA factor effects model give an indication of the effects of the within-factor levels of the grouping variables on the dependent variable. These factor level constants could be estimated using a regression approach that is equivalent to the ANOVA model (*12*). An underlying assumption of the ANOVA and regression models discussed in this section is that of homoskedasticity, or constant variance of the error term. OLS residuals could be tested for heteroskedasticity by regressing the squared OLS residuals against the OLS predicted values and a constant (*12*). The test statistic was a measure of the fit of the residual regression and was the coefficient of determination (r^2) times the sample size. This test statistic could be compared with the chi-square distribution with 1 degree of freedom to test the null hypothesis of homoskedasticity. An econometric estimation package (*13*) was utilized to carry out tests of heteroskedasticity, with the results presented in Table 2. The tests indicate that three of the five regression models had significant violations of the homoskedastic disturbance assumption and therefore had inefficient estimates. To account for this heteroskedasticity, a heteroskedastic model (HET), which is a special case of the GLS regression model, could be estimated by the method of maximum likelihood using an econometrics computer package (*13*).

Accuracy Perceptions

The accuracy of the information was controlled at the 75 percent level for this experiment. This means that five of the 20 trial days were randomly assigned to have incorrect information displayed for that day. Out of every four trial days, one day was randomly selected to receive incorrect information to obtain a balanced level of accuracy and not have the inaccurate days be grouped in time, which is a possibility with a pure random assignment. On days with inaccurate information, all information was incorrect except for incident location information. The locations of incidents were always correctly displayed. This means that on an inaccurate day, the minimum path displayed was incorrect, the color-coded congestion level was incorrect, and the route guidance provided did not follow the minimum path for that day.

In general, subjects perceived the information to be considerably accurate, with 17 subjects rating the information extremely accurate and 57 subjects rating it as frequently accurate. Still, a considerable number of subjects, 22, rated the information as only moderately accurate and 3 subjects rated the information as frequently inaccurate. This indicated that significant between-subject differences in accuracy perception levels existed.

In the ANOVA model of subjects' perceptions of accuracy, the interaction effects of en route guidance and congestion information were found to be significant ($F = 6.83$). For factors that had significant interaction effects, the significance of the main effects were confounded by the interaction effects, and therefore, the significance of the main effects for en route guidance and congestion information could not be determined independent of the interactions. This meant that even though the F statistics indicated that en route guidance and congestion information were not individually significant, these factors were still contributing to the model via interactions with each other. The main effects of education level and incident information were found to be marginally significant ($F = 3.66$ and 3.24, respectively).

The ANOVA regression model for the accuracy rating of information is presented in Table 3. The results of this regression model suggested that subjects with higher education levels perceived the information to be more accurate than did less educated subjects (note this was a five-point scale with one equal to the highest rating; thus the effect of the negative coefficient was to increase the rating).

TABLE 2 Tests for Heteroskedastic Disturbances

Test Form	Test Statistic[1]				
	EQ1	EQ2	EQ3	EQ4	EQ5
ϵ^2 on $\hat{Y}$	.000	34.85	.759	7.86	12.18
ϵ^2 on $\hat{Y}^2$	.007	35.65	.532	8.47	12.53
ϵ^2 on $\log(\hat{Y}^2)$	.004	31.02	.995	7.13	11.26

[1]Chi-square critical value = 3.84 (1 df, α=0.05)

TABLE 3 HET and OLS Regression Models

Variable	Coefficient(t-statistic)				
	EQ1[a,f]	EQ2[b,g]	EQ3[c,f]	EQ4[d,g]	EQ5[e,g]
Intercept	2.113	1.644	2.299	1.946	2.228
Incident	-0.116(-1.96)	-0.357(-6.00)			0.116(1.69)
En-route	0.018(0.02)	-0.135(-2.72)	0.120(1.16)		-0.051(-0.73)
Pre-trip		-0.062(-1.07)	0.050(0.49)	-0.152(-1.90)	-0.097(-1.43)
Congestion	0.030(0.44)	-0.299(-5.12)	0.155(1.49)	-0.125(-1.56)	
Age		-0.044(-0.86)			0.024(0.37)
Education	-0.130(-1.90)				
Gender			0.060(0.59)	0.282(3.55)	
Incident*pre-trip		0.133(2.29)			
Incident*congestion		0.247(4.23)			
Incident*age		-0.183(-3.61)			0.141(2.15)
En-route*age		-0.278(-4.72)			
Pre-trip*congestion		0.240(4.01)		0.275(3.41)	
En-route*pre-trip			-0.278(-2.69)		0.285(4.18)
Incident*en-route					0.203(3.08)
En-route*congestion	-0.191(-2.79)				
Pre-trip*gender			-0.158(-1.52)		
Income			0.025(0.42)	-0.018(-0.42)	-0.084(-2.20)
Variance equation					
Alpha		0.375(12.43)		0.542(11.17)	0.422(12.08)
R^2	0.135		0.135		
Log Likelihood		-85.18		-139.33	-115.12

[a]EQ1 = dependent variable: accuracy rating of information
[b]EQ2 = dependent variable: importance rating of information
[c]EQ3 = dependent variable: likelihood to purchase a system
[d]EQ4 = dependent variable: information system valuation
[e]EQ5 = dependent variable: information monthly service valuation
[f]OLS model
[g]HET model

Also, incident information tended to increase the subjects' perception of accuracy. En route guidance and congestion information had significant interactions and had to be analyzed jointly. The net effect of having either en route guidance or congestion information or having both resulted in a decreased perception of system accuracy.

Importance of Information on Travel Decisions

Subjects also found the information to be considerably important in making their route choice decisions with 84.5 percent of the subjects indicating that the information was either very important or of some importance. The reduced ANOVA model of the importance rating of information indicated that all of the main effects in the model were paired in a significant interaction effect. The main effects of incident information and congestion information were found to be individually significant. Incident information significantly interacted with pretrip information, congestion information, and age level. Pretrip information also significantly interacted with congestion information, and en route guidance significantly interacted with age level.

The HET regression model for the importance rating of information is presented in Table 3. Because all main effects were involved in interactions, the effects of a treatment must be considered as a net effect of main and interaction effects. As an example of the effects of the treatments on the importance of the information on subjects' decisions, consider an individual in the young age group (less than 40) who had incident and congestion information versus an individual in the same age group who had en route guidance and pretrip information. The regression model indicated that for the individual with incident and congestion information, the mean rating of importance was increased by 9 percent, whereas for the individual with en route guidance and pretrip information, the mean importance rating was lowered by 5.5 percent. The combination of incident, congestion, and en route guidance produced the greatest effect with a 24 percent increase in mean rating. If all information treatments were available, the mean importance rating was increased by only 12 percent. The age level effects indicated that younger subjects found the information to be of more importance in making their route decisions.

The model indicated that the pretrip information treatment had the least effect on increasing the mean importance of information, and that having en route information in conjunction with incident and congestion information significantly increased the importance of the information. These results indicate the importance of having incident information available to drivers and that its importance was reinforced by the availability of route guidance information. Descriptive information such as incident locations and congestion levels can provide a strong rationale for compliance with prescriptive information such as route guidance.

Consumer Demand and System Valuation

The potential demand for such systems, gauged by subjects' willingness to purchase a system, seems to be fairly high, with 51.5 percent of the subjects indicating that they would be likely or very likely to purchase an information system. Thirty-nine percent indicated that they were undecided, and the remainder were either unlikely or very unlikely to purchase.

The value of traffic information services was measured by the dollar value subjects indicated they would be willing to pay for an information system of the type experienced in the simulation. The value of both a fixed-price system and a monthly service type of system were measured. Forty-eight subjects indicated that they were only willing to pay less than $200 for a system, whereas 32 subjects were willing to pay between $200 to $400, and 11 subjects were willing to pay $400 to $600. Eight subjects were willing to pay in excess of $600. The average system value was calculated as $271 based on the assumption that the responses in each group were uniformly distributed within the group. The willingness to pay for an information system as a monthly service had a similar value distribution to that of the fixed price. Forty-one subjects indicated that they were only willing to pay less than $10/month, whereas 37 subjects were willing to pay between $10 and $20/month, and 14 subjects were willing to pay $20 to $30/month. None of the subjects were willing to pay more than $75/month for this type of service. The average monthly service value was calculated as $14.25/month.

The ANOVA regression models for the system demand and valuation variables are presented in Table 3. Income was included as a covariate in all of these models based on consumer theory, even though it was not found to be significant in all models. Inclusion of income did have significant effects on other variables in the model, and this is also an indicator that the variable should be retained in the model.

In the model of system demand, the interaction effects of en route and pretrip information, as well as pretrip information and gender, were found to affect the likelihood of purchasing a system. Congestion information was found to be marginally significant individually and was retained in the model. The regression model is presented in Table 3 and indicated that subjects were more willing to purchase a system that had a combination of pretrip and en route information. Also, a system with only congestion information was more likely to be purchased over a system with either en route or pretrip information individually. For systems that included pretrip information, males were more likely to purchase than females, and for systems without pretrip information, females were more likely to purchase. The income coefficient indicated that those with higher incomes were less likely to purchase a system.

The HET regression model of system valuation is shown in Table 3. The model indicated that the interaction effects of pretrip and congestion information significantly affected the dollar value of the system to the subjects. The gender of the subjects was also shown to be individually significant. This model indicated that subjects placed a higher value on a system with both pretrip and congestion information over systems with either pretrip or congestion information individually. The coefficient on the gender variable indicated that male subjects placed a higher value on the system than did female subjects. Again, the income variable was not significant but indicated that subjects with higher incomes valued the system less.

The HET regression model for service valuation is presented in Table 3. The model indicated that the interaction effects of incident information with both en route guidance and age level, as well as the interaction effects of en route guidance and pretrip information, significantly affected the dollar value of a monthly service to the subjects. Incident information and income were found to be individually significant with incident information increasing the service value, and income decreasing the value. This model indicated that subjects put more value on a service that provided incident information, en route guidance, and pretrip information collectively, than for a system with any other combination of one or two of these treatment types. This finding suggests a certain synergism among the information treatments in which individually, the effects are not significant on modifying behavior, but collectively, the information becomes effective. The model also indicated that younger subjects placed a higher value on the service than older subjects. This is not surprising given that younger subjects were also found to rate the information as more important in their route selection.

CONCLUSIONS

Preexperiment Survey Findings

This article summarizes the initial findings of a simulation experiment carried out at the University of California at Davis to investigate information use and learning with advance traveler information. The simulation experiment was found to be an extremely useful tool in collecting sequential route choice data in the presence of information. The only other data of this type available for analysis comes from other simulation research (*14–18*). To obtain this level of data, a substantial increase in per-subject cost is required. The per-subject costs for this analysis were approximately $100 per completed subject for the 100 subjects. This can be compared with the approximate $20 per-subject required to complete the more traditional CATI survey (*1*). If one considers the trial day in the simulation as the observational unit in lieu of the subject, then this type of data collection becomes very cost effective, on the order of only a few dollars per observation.

The preexperiment survey captured many interesting aspects of commuter's information acquisition and use. It was found that 70 percent of the sample currently listen to en route traffic reports and that 43 percent listen to pretrip reports, and that almost all subjects who listen to traffic reports listen every day or nearly every day. This finding indicates that a large market exists for traffic-related information and is validated by similar findings from the Los Angeles survey. The results also indicate that this information is having an effect on commuters' travel behavior. Of those subjects who listen to pretrip information, 70 percent had changed routes at least once in the last month, and 37.2 percent had changed departure time. Of those subjects who listen to en route information, 60 percent had changed routes at least once in the last month. Although currently available information is having an effect on commuters' behavior, it is a limited effect. The majority of those who had changed routes in the last month had only changed once or twice, and of those who changed departure time, the majority had changed between one and six times. This limited effect of current information may be because of subjects' perceptions of the accuracy they receive. It was found that 60 percent of the sample rated the accuracy of current traffic information as either somewhat accurate or not very accurate. Results from the survey of Los Angeles commuters also found that the more inaccurate information was per-

ceived to be, the less likely respondents were to have changed their routes because of either pretrip or en route information.

Acquisition of pretrip information was found to be significantly dependent on subjects' gender and income. Findings from this study, as well as those of the Los Angeles survey, indicate that females were more likely to listen to pretrip information. Subjects with higher incomes were found to be less likely to listen to pretrip information in both studies. Previous simulation experiments (*3*) involving pretrip information, found that females followed pretrip information more than males. Other researchers have also found gender differences with regard to commute aspects and traffic information use (*19,20*).

Although listening to pretrip information is significantly associated with gender and income, results indicate that the effect of this pretrip information on travel behavior is independent of gender but is dependent on commute distance (for route changing), income, education, and age (for changing departure time). Subjects with longer commute distances were much more likely to have changed their commute route because of pretrip information, whereas older, more educated, higher income subjects were less likely to change their departure time. These effects of commute distance, education, and income had similar significant effects on travel behavior for the respondents in the Los Angeles study. These findings indicate that there are significant differences between the individual characteristics that influence drivers' preferences about how they receive information and the characteristics that influence drivers' travel choices in response to receipt of such information.

Similar results are also found for the acquisition and use of en route information. Commute distance significantly affects whether an individual listens to en route information, with longer distance commuters much more likely to receive en route information. Age and the work schedule flexibility significantly contribute to whether an individual changes routes based on en route information. The finding that older subjects were more likely to change routes because of en route information seems counterintuitive on the surface. However the older age group in this study does not infer "aged," only age 41 or over. Subjects in this age group may be more experienced drivers and have more familiarity with the traffic network and therefore may have a greater awareness of alternatives that are available.

Simulation Experiment Findings

Based on the subjects' rankings, it appears that incident and congestion information types were the most important in influencing subjects' route decisions. These results are supported by the ANOVA, which also reveals some significant interaction effects. In general, subjects indicated that the information provided in the simulation was important in making their route selections. Individually, all of the information types had the effect of increasing the importance rating of the information, collectively, subjects with all but pretrip information rated the importance of the information the highest. The age level of subjects was also significant in determining the importance of information. This finding supports the validity of the use of a simulated environment as age is shown to be significant in affecting the behavior of subjects both in the simulation and in the real travel environment. Another supporting factor is that gender was shown to be insignificant in both environments. In the previous simulation experiments that used university students, it was postulated that age may have significant behavioral effects; the decision to include age effects in this design is validated by these results.

Subjects generally indicated that they found the information in the simulation to be very accurate even though the accuracy was controlled to 75 percent and they actually experienced one incorrect information day out of every five. The distribution of the accuracy rating of the information in the simulation tends to be higher than the distribution for information that subjects currently receive. Subjects' education significantly contributes to the perception of accuracy with more educated subjects perceiving the information as more accurate. Incident information also contributed to increasing the accuracy perception as incident information was always correct. The effects of en route guidance and congestion information tended to decrease accuracy ratings.

The analysis of consumer demand and system valuation found that subjects were more likely to purchase systems that contain combinations of information types over a single type. Another important finding is that although incident information is very important in influencing route choice decisions and the dollar value placed on a system, incident information was not found to significantly influence the decision to purchase a system. This reiterates that the factors that influence the use of information in the decision process are not necessarily the same as the factors that influence the acquisition of that information. Another example of this is in the findings on gender effects. It was found that gender significantly influenced the acquisition of pretrip information with females being more likely to receive pretrip information, but gender had no significant effect on the use of that information in changing routes or departure times. Also, in the simulation, gender was found to significantly contribute to system demand and valuation ratings, but did not contribute to the importance of the information in making route choice decisions.

Future Research

The findings presented in this article are initial findings from a rich data source that will continue to be explored. The analysis performed here can be enhanced by application of more advanced modeling techniques to further separate and identify the characteristics that influence the acquisition and use of information. Qualitative response modeling using logit and probit frameworks is a natural extension of the analysis performed here.

ACKNOWLEDGMENTS

This research has been funded by the California Department of Transportation (Caltrans) and the Partners for Advanced Transit and Highways (PATH). The authors would also like to acknowledge the staff and researchers at the institute who contributed their time during the pretesting of the simulation. The views expressed in this paper are those of the authors and do not necessarily represent the views of the funding agencies.

REFERENCES

1. Abdel-Aty, M. A., K. M. Vaughn, R. Kitamura, and P. P. Jovanis. *Survey of Route Choice Behavior: Empirical Results from Southern California and Their Implications for ATIS*, Research Report UCD-ITS-RR-

93-12. Institute of Transportation Studies, University of California at Davis, Davis, Calif., 1993.
2. Abdel-Aty, M. A., R. Kitamura, P. P. Jovanis, and K. M. Vaughn. *Understanding Commuters' Attitudes, Uncertainties, and Decision-Making and their Implications for Route Choice*, Research Report UCD-ITS-RR-94-5 Institute of Transportation Studies, University of California at Davis, Davis, Calif., 1994.
3. Vaughn, K. M., M. A. Abdel-Aty, R. Kitamura, P. P. Jovanis, H. Yang, N. Kroll, R. B. Post, and B. Oppy. Experimental Analysis and Modeling of Sequential Route Choice Behavior Under ATIS in a Simplistic Traffic Network. In *Transportation Research Record 1408*, TRB, National Research Council, Washington, D.C., 1993.
4. Vaughn, K. M., M. A. Abdel-Aty, R. Kitamura, and P. P. Jovanis. Analysis of Sequential Route Choice Under ATIS. *Proc., Fifth International Conference on Computing in Civil Engineering*, American Society of Civil Engineers, Anaheim, Calif., 1993.
5. Abdel-Aty, M. A., K. M. Vaughn, P. P. Jovanis, and R. Kitamura. Understanding Traveler Responses to ATIS. *Proc., International Symposium on Automotive Technology and Automation (ISATA)*, Aachen, Germany, 1993.
6. Yang, H., R. Kitamura, P. P. Jovanis, K. M. Vaughn, and M. A. Abdel-Aty. Exploration of Route Choice Behavior with Advanced Traveler Information Using Neural Network Concepts. *Transportation*, Vol. 20, 1993, pp. 199–223.
7. Vaughn, K. M., P. Reddy, M. A. Abdel-Aty, F. Mar, R. Kitamura, and P. P. Jovanis. *Investigating Information Use and Learning with ATIS: Simulation Description and Experimental Design*, Research Report UCD-ITS-RR-94-7. Institute of Transportation Studies, University of California at Davis, Davis, Calif., 1994.
8. May, A. D. *Traffic Flow Fundamentals*. Prentice Hall, Englewood Cliffs, N.J., 1990.
9. Montgomery, D. C. *Design and Analysis of Experiments*, 3th ed., John Wiley & Sons, New York, 1991.
10. Neter J., Wasserman W., and Kutner M. H. *Applied Linear Statistical Models*. Irwin, Homewood, Ill., 1990.
11. Vaughn, K. M., P. Reddy, M. A. Abdel-Aty, F. Mar, R. Kitamura, and P. P. Jovanis. *Information Use and Learning with ATIS: Analysis of Network Simulation Experiments*, Research Report, UCD-ITS-RR-94-10. Institute of Transportation Studies, University of California at Davis, Davis, Calif., 1994.
12. Kmenta, J. *Elements of Econometrics*, 2nd ed. Macmillian, New York, 1986.
13. *SHAZAM Econometrics Computer Program, User's Reference Manual version 7.0*. McGraw-Hill, New York, 1993.
14. Chen, Shen-Te, P., 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.
15. Allen, R. W., et al. Laboratory Assessment of Driver Route Diversion in Response to In-Vehicle Navigation and Motorist Information Systems. In *Transportation Research Record 1306*, TRB, National Research Council, Washington, D.C., 1992, pp. 82–91.
16. Polak, J. and Jones, P. The Acquisition of Pre-Trip Information: A Stated Preference Approach. Presented at 71st Annual Meeting of Transportation Research Board, Washington, D.C., 1992.
17. Adler, J. L. and McNally, M. G. In *Laboratory Experiments to Investigate Driver Behavior Under ATIS*. UCI-ITS-WP-93-5, 1993.
18. Iida, Y., T. Akiyama, and T. Uchida. Experimental Analysis of Dynamic Route Choice Behavior. *Transportation Research-B*, Vol. 26B, No.1,, 1992 pp. 17–32.
19. Wenger, J. M., et al. Motorist Behavior and the Design of Motorist Information Systems. In *Transportation Research Record 1281*, TRB, National Research Council, Washington, D.C., 1990.
20. Mannering, F., S. G. Kim, W. Barfield, and L. Ng. Statistical Analysis of Commuters' Route, Mode and Departure Time Flexibility and the Influence of Traffic Information. Presented at 73rd Annual Meeting of Transportation Research Board, Washington, D.C., 1994.

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

Challenges in Defense Conversion: Intelligent Transportation Systems as a Case Study

CHRISTINE N. PETERSON AND DEBBIE S. OLSON

The demise of the cold war sparked an intense debate as to the feasibility of converting formerly defense-dependent, high technology companies to civilian or dual-use endeavors. Recently this topic has been linked to transportation, as many of these companies have identified potential transportation applications of their technology developed under defense sponsorship. Past attempts to diversify and a review of the literature indicate, however, that this transition is neither automatic nor easily accomplished, in part because of nontechnical barriers in the areas of technology transfer, organizational culture, marketing, and finance. Using the emerging domain of Intelligent Transportation Systems (ITS) as a case study, nontechnical barriers to this transition are reviewed. Drawing on a series of lengthy telephone and personal interviews with defense company leaders and the recommendations gleaned from a joint industry-public sector executive panel, this study addresses two questions: what are the specific barriers faced by defense companies entering the ITS and transportation market, and what is the appropriate industry and public sector response to these challenges? In discussing these issues, the conclusion is reached that the defense industry is making progress toward overcoming difficult barriers to entry into the ITS domain; however, continued formation of a public-private partnership with DOT and other transportation officials is needed to provide far-reaching transportation benefits.

The demise of the Cold War signaled for many defense-dependent high technology firms a new sort of conflict: the battle for corporate survival. In the sociopolitical thaw that has followed, one of the most frequently debated topics is the feasibility of conversion or diversification of previously defense budget-dependent U.S. companies. The impetus behind this debate is a recognition that resident within these defense firms is a technological base proven in its ability to defend our nation, but also bearing promise of applications to pressing national needs. Transportation infrastructure revitalization and enhancement represents one such prominent need. Technologies such as Global Positioning Systems (GPS), advanced electronics, communications, and surveillance systems have been identified by defense companies and the public sector alike as potential solutions to mobility and safety problems (*1*). The quest for economic vitality similarly heightens the desire to put the resources embodied in these companies to work on providing significant transportation improvements.

Past experience and recent reviews of the literature indicate, however, that this transition is neither automatic nor easily accomplished. The literature is rife with examples of well-publicized conversion failures in previous periods of defense downsizing (*2,3*). What these examples illustrate, however, is not necessarily the inappropriateness of the technology being applied to new problems, but rather a number of difficult nontechnical barriers faced by these organizations, which are hindering their progress.

The emerging domain of Intelligent Transportation Systems (ITS) represents a case study in which to examine this theory. Jump-started by federal dollars and initiative, the ITS program has grown from a mere $4 million in funding in fiscal year 1990 to a budget of nearly $227 million in fiscal year 1995. Beyond federal involvement, a burgeoning ITS market and industry is beginning to take shape. Some optimistic futurists even project an anticipated market of $200 billion over the next 20 years (*4*).

The lucrative promise of a vast commercial marketplace and the natural fit between the technology-intensive ITS arena and defense company expertise in these areas, as well as a rich history of systems integration and test and evaluation experience, have led many companies to demonstrate interest. However, the synergy between these talents and ITS national program needs has not been ensured, because of a number of nontechnical barriers earlier alluded to. For example, many defense company leaders report a lack of understanding of their new customers, who differ greatly from the single-point Department of Defense (DOD) buyer they shaped their businesses around. Likewise, whereas performance once superseded cost in the defense environment, in the transportation sector cost is a critical driver of investments and acquisition. In addition, many companies themselves are burdened with cost structures inappropriate for nondefense and commercial markets, as well as risk-averse corporate cultures fostered over decades of a relatively steady defense market. Similarly, this repositioning is taking place against a backdrop of an extremely dynamic public policy environment in which the relationships between industry and government, and the respective roles each must play to achieve the goals of improving the nation's transportation infrastructure, are being reevaluated. Nowhere is this more true than in the domain of ITS, in which a true public-private partnership is deemed necessary to achieve full deployment.

Through SRI's analysis, the barriers faced by defense companies as they position themselves for entry into the ITS marketplace can be grouped into a four-pronged framework of technology transfer, organizational culture, marketing, and finance. The desire to identify and qualify specific barriers in these categories led to the commission of a SRI International study, which filled in the gaps of existing research and used ITS as a case study to examine the possible barriers faced in other conversion efforts, perhaps with lessons for other modes (*5*).

Sponsored by the FHWA and the Volpe National Transportation Systems Center (VNTSC), the first phase of the study involved a

SRI International, 1611 N. Kent Street, Arlington, VA 22209.

combination of survey and field research, as SRI conducted a literature review and then held over 35 interviews with defense company leaders and in-depth focus groups. The interviews included a series of telephone conversations as well as on-site interviews with three representative firms. SRI then convened an executive industry panel for a 1-day joint session with DOT officials on May 25, 1994 at SRI's Washington, D.C. office. The purpose of this meeting was to establish a dialog between key industry representatives, validate barriers uncovered in the telephone and focus group interviews, and examine possible next steps. Indeed, the focus throughout was to address two questions: what are the specific barriers faced by defense companies entering the ITS and transportation marketplace, and what is the appropriate industry and public sector response to these challenges? The recommendations for appropriate actions suggested by the joint industry-public sector executive panel are described in this paper. The conclusion is likewise reached that significant progress is being made in overcoming these barriers, although continued building of a public-private partnership between industry and DOT is needed to provide far-reaching transportation benefits envisioned by both sectors.

CASE STUDY METHODOLOGY

Literature Review

An internally funded SRI assessment of the literature formed the baseline information for the study. Before DOT sponsorship, SRI reviewed the most recent literature on general defense conversion topics as well as more focused materials specific to the transportation and ITS sectors. The amount of information concerning this topic has rapidly increased, as SRI discovered in its compilation of a source list (*6*). Articles related to the study topic that were monitored by SRI roughly doubled in number during January to September of 1993 when compared with the 12-month period of 1991 to 1992. A portion of this increased coverage was due in part to the keen attention given to the federal government's Technology Reinvestment Project (TRP), with its related dual-use focus. As part of the literature review, SRI also examined specific assessments of defense companies and their diversification strategies, which generally covered the relationship of defense conversion to transportation (*7,8*). However, none of the previous research identified nontechnical barriers to entry specifically within the ITS context.

Telephone Interviews

The first phase of the study included telephone interviews of defense company leaders. The approach was qualitative and the questions were largely open-ended, designed to identify the general issues and specific barriers faced by these companies as they prospected for new market opportunities in the ITS domain. The interview sample of 10 companies was drawn from a population of the top 100 defense prime contractors based on fiscal year 1992 award amounts. This interview sample represented companies whose revenues accounted for nearly 15 percent of the total Research, Development, Testing, and Evaluation (RDT&E) budget of the DOD in fiscal year 1992, which totaled $121.4 billion (*9*). To ensure candid responses, SRI agreed to hold confidential the identities of companies and company leaders that participated in the interviews.

In addition to being drawn from the top 100 contractors, the interview sample was selected on the basis of several key criteria. These criteria included geographic location, wherein selected firms had a major operation in at least one of the eight geographic regions where defense cutbacks are likely to have the most impact (*1*), indication of interest in ITS (demonstrated by membership in ITS AMERICA, the primary ITS association, or appearance on bidders or awards lists of ITS-related projects solicited by FHWA), and potential to be responsive to the technical requirements and identified user services of ITS (*10*).

A structured interview guide was developed, and company leaders in each of the 10 companies participated in a telephone interview, conducted December 1993 through February 1994. The 14 study respondents, whose average length of service in their companies was 19 years, were primarily senior-level managers. Three included "ITS Program Manager" as part of their titles. At least one company leader in each organization was asked to provide his or her observations in five areas relating to the ITS and transportation marketplace, including organizational approach, information networks utilized, public-private partnerships, procurement responses, and industry impressions.

Perceived Barriers

The telephone interviews revealed a number of barriers, as well as opportunities for change. Foremost among the perceived barriers and opportunities were the following:

1. Most companies have only recently become involved in ITS and their targets are largely driven by existing core competencies, instead of market pull.
2. An interdisciplinary approach to entry into ITS is being taken, although corporate coordination and organization appears loosely structured.
3. Corporate commitment to ITS is currently tentative, in part because of what is perceived as an undefined and long-term market. Many managers are reluctant to invest in ITS without returns on investment within a few years, according to the respondents.
4. One-way information channels (reports, press coverage) are rated as minimally useful in terms of providing information.
5. Two-way channels (networking, conference attendance) are more productive, although there was a mixed view of the usefulness of some conferences and ITS AMERICA (the primary ITS association) committee participation.
6. Interactions with federal, state, and local transportation agency officials are quite useful and help in understanding the demands of the new ITS domain.
7. Team formation and partnering is necessary for success but should be company initiated.
8. Major differences exist between FHWA and DOD solicitations, and respondents understand little about FHWA's procurement processes.

In addition to these specific barriers, respondents also indicated that their key impressions about the current state of the ITS program centered around the need to sustain ITS program momentum by maintaining a strong federal presence. Similarly, they felt that a technology infusion was needed in federal, state, and local transportation agencies. Respondents thought transportation officials with technical training in advanced technologies would facilitate

communication and enable the best products to be deployed. The respondents also commented on the importance of addressing financial barriers such as cost-share and no-fee provisions in federal solicitations. These requirements are difficult to justify to upper management, which does not see an immediate payoff. Finally, the need for enhanced outreach with a regional orientation was voiced as another important concern.

On-Site Interviews and Focus Groups

Drawing from their responses, SRI shaped the questionnaire for in-depth focus groups and on-site interviews. The objective for these series of interviews was to go beyond the specific, tactical barriers identified in the telephone interviews to a broader understanding of how ITS was fitting into overall corporate diversification strategies and to ascertain whether the identified barriers still held true. This second phase consisted of on-site visits to three companies, representative of the population under study, which included aerospace, electronics, and engineering companies. Each selected company had yearly revenues of $5 to $10 billion and major operations in one or more of the eight states most affected by defense cutbacks (*1*). Lengthy interviews based on an interview guide were held with an average of six company leaders per company, representing technical managers, business development leaders, and corporate officers.

One of the most striking characteristics of the on-site focus groups was the willingness to share their observations on the challenges they face in repositioning themselves, in what is generally a closed and cautious environment. The issues of conversion and their relationship to the transportation marketplace appeared to be of great interest. General characteristics of the three companies studied appear in Figure 1.

Major Findings from Interviews and Focus Groups

From our analysis of the data collected in both the telephone and focus group interviews with defense company leaders, it appears that many intriguingly similar barriers to entry are facing the industry as a whole. However, we are limited in our ability to generalize across the industry because of the relatively small sample size. From the data, however, one fact did become particularly clear: both defense companies and the public sector, to include DOT, have the opportunity to actively reduce barriers within their respective spheres of influence. The challenge appeared to be twofold, with both industry-based and public sector-influenced barriers comprising the complete picture. The barriers most highly emphasized in both the case studies and the interviews are summarized below.

Industry-Based Barriers

1. Market unfamiliarity: defense companies lack knowledge of nondefense government procurement processes and evaluation criteria and have little or no experience in dealing with multiple customers and users.
2. Technology push orientation: many defense companies are committed to high-technology solutions in marketplaces with different requirements.

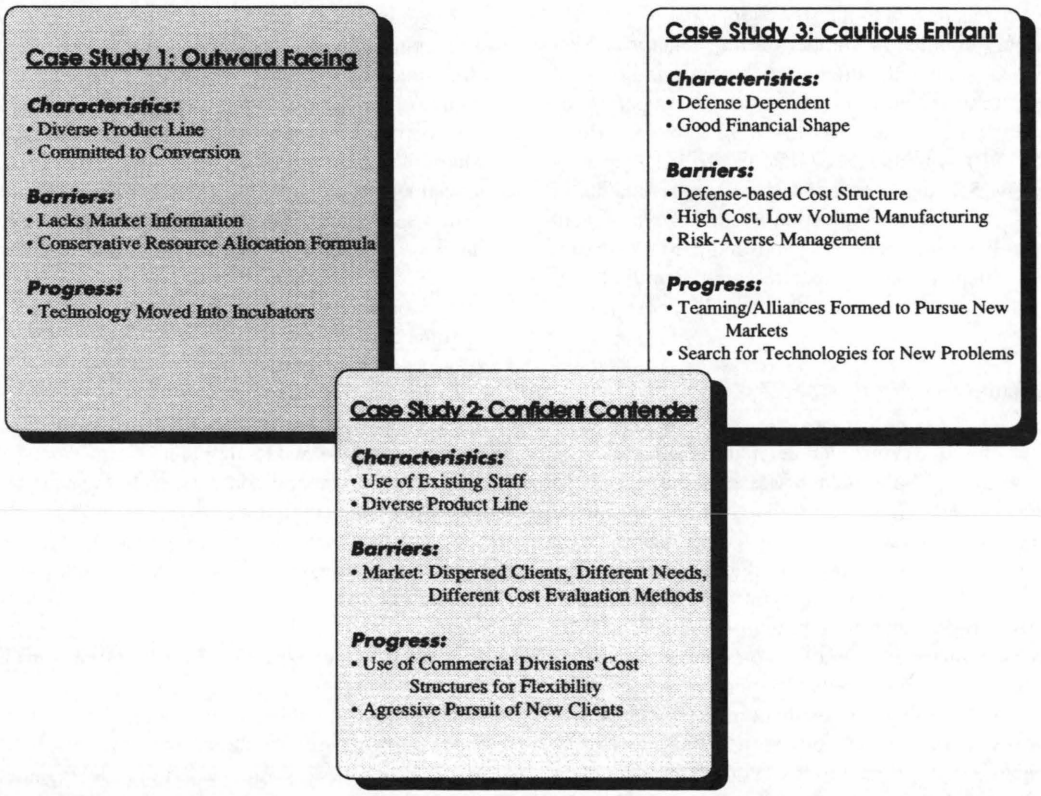

FIGURE 1 Excerpts from on-site interviews.

3. Risk-averse culture: these organizations have a tradition of low risk approaches and caution in the marketplace; they likewise lack internal transportation domain knowledge and expertise.

4. Financial constraints: companies retain defense-configured financial infrastructure with high overheads and little flexibility.

Public Sector-Influenced Barriers

1. Closed communication channels: many respondents indicated that they felt programmatic information on ITS projects and plans were filtered and dissemination modes limited. They likewise saw a shortage of opportunities to interact with customers and users and expressed a desire for enhanced dialog.

2. Technology gap: many company leaders perceived a lack of in-house technological sophistication on the part of the public sector, as well as few standards and protocols that will reduce market risk and enable defense companies to more readily pursue ITS markets.

3. Industry knowledge lacking: many public sector representatives exhibited little understanding of how the defense industry operates, according to the respondents.

4. Cost Requirements: federal solicitation requirements are often incompatible with business practices, particularly cost-share and no-fee provisions in situations such as system integration projects, in which no commercially marketable product will result.

Industry-DOT Executive Panel

A starting point for the panel session involved drawing from the above framework established in the first phase of the study, with its corresponding public sector-influenced and industry-based barriers. For the panel, refinement included distilling the findings into a simplified framework for use as a template in the panel discussion. The four-pronged template of technology transfer, organizational culture, marketing, and finance adequately captured the comments garnered in the interview and focus group process. On May 24, 1994 at SRI's Washington office, a joint industry-DOT panel met to (a) establish a dialog, (b) discuss and validate the barriers identified in the first phase of the study, and (c) explore the next possible actions. The sort of frank exchange envisioned by SRI and the DOT sponsors required identification of appropriate individuals for the discussion and a format for the meeting that would facilitate this openness.

Panel Selection and Panelist Qualifications

The selection of industry candidates was based on similar criteria to that used in the first phase of the study. Drawn from the pool of the top 100 defense prime contractors, the panelists represented a varied spectrum of involvement in the ITS program (as evidenced by ITS AMERICA membership, appearance of FHWA bidder's lists, etc.). Although some industry panelists represented companies that participated in the first phase of the study through the survey interviews and focus groups, several new company leaders represented their organizations on the panel. As a group, the represented companies formed a broad mix in terms of defense dependency (determined as a ratio of defense to nondefense corporate revenues) and business sector, including advanced electronics, aerospace, and information services.

In terms of the industry panelists' personal qualifications, SRI primarily targeted the director and vice presidential levels for the eight industry representatives. Selection was based on inviting panelists who could articulate the broad strategic views of their organizations, and who likewise had a knowledge of corporate culture and financial resources. Selected panelists' professional responsibilities included strategic planning and analysis as well as corporate marketing and technical management.

The seven public sector representatives from DOT mirrored this composition, including program managers from the ITS program at DOT headquarters and the research and development divisions, as well as representatives from the director and associate administrator levels. In keeping with the intermodal nature of ITS, representatives from both FHWA and FTA were present. The balance among public sector and private industry representatives allowed for a sound foundation to begin an exploratory dialog on the salient issues involved in defense conversion, within the context of an emerging ITS market.

Agenda and Proceedings

The agenda included a review of the SRI study and a group dialog on key ITS challenges and barriers for defense companies, using the identified template. This discussion was followed by forming small brainstorming groups to expand the vision of opportunities and identify priority issues. An attempt at action planning closed out the discussion and formed the basis for a series of recommendations.

Underscoring both of these discussions was a sense of urgency on the part of the defense industry. In an environment of intensified downsizing of the defense segments of their companies, company leaders greatly desired the ability to quickly employ the resources embodied in their people and technology before these were lost. This desire to retain and fully utilize these vast resources was woven into many of the comments identified below.

Panel Validation of Barriers

The themes from the panelists' comments are clustered around the template described earlier, which uses the core categories of technology transfer, organizational culture, marketing, and finance.

Technology Transfer

Technology transfer in the context of this panel meeting did not refer to core competencies but rather connoted the movement toward commercialization of technology within the defense company or defense portion of the company. Although this sort of transfer is a difficult process, several defense companies are dealing with the issue. Many of the industry panelists represented defense companies that also had large and profitable commercial divisions. In those companies, technologies developed for defense applications but with commercial possibilities were being transferred to commercial divisions for manufacturing and sale. The sentiment expressed was that only commercial divisions could provide high volume, low cost production competitively.

The panelists also indicated some struggles to apply to their technologies and capabilities to specific markets, whether commercial or nondefense government. This was especially true in the context

of ITS, in which panelists asserted that valuable technologies remained buried within their companies and that requests for new ways to solve a particular problem would help uncover these technologies. In light of this, many panelists called for acceleration of the matchmaking process between existing technology and problems or needs in the ITS arena.

Furthermore, company leaders voiced the belief that a definite advantage exists for DOT to continue its proactive role in enabling this matching process. If defense companies can see how technologies latent within their organizations can apply to ITS or transportation problems, they can write compelling TRP proposals. DOT gains by expanding their development and demonstration effort through leveraging TRP funds. This scenario also applies to tapping into the reservoirs of additional funds under the rubric of defense conversion or dual-use.

The other aspect of this issue was captured as the realization that programmatic understanding of ITS (i.e., familiarity with the ITS user services without in-depth knowledge) was not thorough enough to guarantee application of appropriate technology to all the necessary ITS needs. Comments made throughout the day suggested the efficacy of public sector visits to companies and formalized exchanges on key problems looking for cost-effective technical solutions. The overriding concern about having potential solutions and not knowing enough about the problem to explore application of these solutions was expressed frequently.

In both of these issues, including transfer of technology to other nondefense agencies and leveraging other sources of funds, one mutually agreed-on belief was the importance of having a federal program in ITS. This echoed the views expressed in the telephone interviews. The industry panelists were in agreement that their companies would not be pursuing technology transfer to other government agencies or in commercial marketplaces without clear government leadership and direction.

Organizational Culture

During the panel discussion, management issues and barriers were expanded into a broader aegis of organizational culture. The primary issues became rooted in the nature of relationships: primarily relationships with the new customer, but also including newly emerging public-private alliances.

According to the panelists, the defense culture is not tailored to the proclivities of their newfound dispersed customer base, which includes not only federal officials, but state and local stakeholders as well. Even Congress becomes a customer to satisfy, as one DOT representative pointed out, and the shortcoming in many traditionally defense-oriented companies is a tendency to tell, not show, when describing the benefits of a new technological advance. Many agreed that the power of demonstrated technology extends much further than verbal claims to exciting new products and ideas. This wisdom becomes even more important as these companies approach their ultimate customer and user-base: the driving public.

Other relationships that remain awkward fall under the category of public-private partnerships. In this era, when "industrial policy" is often espoused, many panelists expressed a degree of discomfort with the new relationships emerging between industry and government. Much of this may be because of an unfamiliarity of working partnership models, as well as lacking knowledge of their new ITS customers' needs. Industry panelists also pointed out that government officials responsible for shaping policies which govern these new partnerships lacked in-depth understanding of how companies operate, as was suggested in the telephone and focus group interviews. The need for formalized education and opportunities for exchange was mentioned again within this context.

Marketing

Two issues raised in relationship to technology transfer appeared again in the discussion of current and potential markets for ITS, including the importance of federal momentum and the incompleteness of technology matches with market needs.

According to the panelists, one of the key factors in making a corporate investment decision is lack of market risk, since the company wants to assure themselves that they are not too far in front of the market. In essence, they begin to focus on the shape of a new ITS market and the resulting businesses developed to serve that market. This perception led to the introduction of two new issues during the meeting. The first involved the timeliness of the system architecture effort as well as standards development. Secondly, several defense industry panelists discussed their process for selecting new commercial markets to pursue.

Two efforts underway at DOT and within the committee structure of ITS AMERICA are standards development and the systems architecture projects. Both can contribute in a major way to reducing the market risk to industry. The defense industry, however, has a keen sense of urgency and immediacy as earlier stated. They would like to see all these efforts accelerated to provide market opportunities immediately—not many years from now, as some might predict.

To this end, concentration on standards for interconnectivity was stressed and enthusiastically endorsed by several industry panelists. However, there was not a complete consensus. In a discussion focused on electronic toll collection, one company leader argued for desirability of interconnectivity, particularly for interstate or multijurisdictional travel. Another individual challenged the necessity of accelerating interconnectivity because he saw the lack of standardization as an immediate opportunity for his company to develop the technology to convert a toll card into a usable form for many systems. To this particular panelist, interconnectivity in this case could close off a potential market.

Nevertheless, interconnectivity was brought forth frequently as an important issue to address quickly in order to reduce market risk for defense companies eager to develop ITS products. Likewise, the systems architecture effort was reviewed with similar urgent requests to accelerate the process. One panelist suggested that the critical decisions would be made early on and that if this is the case, quick dissemination of the results would be helpful to industry leaders as they deal with determining what the market will bear.

The second new topic was not only specific to the defense industry but may have broader implications. When an industry such as the defense companies represented by the panel contemplates the ITS market, ultimately they are not concentrating on selling to the government. Rather, they are thinking about creating a new business. This new ITS business is by no means a monolithic enterprise. Developing an ITS industry in information services, vehicle components, or for infrastructure-based systems requires different operating models. How to create these different models and comprehend the barriers and opportunities inherent in each is vital to both a successful ITS program and all forms of resultant industry. The role of the federal government in understanding how the new ITS business

may be shaped is a key to reducing market risk, which when unmitigated, can make for reluctant investors.

Finance

The insights captured under this element differed most from the comments uncovered in the initial telephone interviews. The results of these interviews portrayed financial concerns primarily in tactical terms of government cost share requirements not being consistently applied and inappropriate no-fee situations. In terms of strategic planning, respondents indicated that the market was perceived as too long-term to justify cost sharing, especially to upper management looking for immediate returns on investment. The on-site focus groups further clarified the strategic issues in finance. Two of the companies studied in those sessions defined lack of flexibility in their financial systems as being a problem. One of those companies had embarked on a new accounting system to rectify the situation.

During the panel session, many company leaders confirmed that there was indeed a lack of financial flexibility in the defense divisions of their organizations, including burdensome overheads. However, as the on-site focus groups suggested, several companies are attempting to address those problems. As mentioned under the technology transfer category, technology is simply moved out of defense divisions into commercial divisions experienced in distribution, pricing, and high volume production. According to many industry panelists, inflexibility was no longer a central problem.

What has emerged as a major focus is the question of whether the defense industry will make major investments in future ITS products and processes. As some may assume, the answer to this question is not dependent on the resources available within the company for such an investment. Two industry panelists asserted that if the market exists, their organizations would commit significant funds to develop, demonstrate, and commercialize ITS technology. Rather, the answer rests heavily on the perceived market risks of such a venture. Many panelists similarly believed that DOT has an important and appropriate role to play in reducing the market risk by helping to define the current and future ITS markets.

Another issue that surfaced in the focus group interviews and was further amplified in the panel meeting was the adequacy of financial models. The interviews elicited many comments related to company leaders' attempts to understand new funding mechanisms involved in public-private partnerships. Both government and industry panel members expressed a desire to better utilize those models and to identify best practices. A review of the possible models and verification of those that have been effective was deemed useful to the entire ITS stakeholder community.

Recommendations

With respect to developing appropriate recommendations from the specific suggestions captured in the four-pronged template of technology transfer, organizational culture, market, and finance, the panelists articulated a potential vision for ITS. Underlying this vision was a focus on goals attainable through implementation of specific recommendations. These immediate goals as articulated by the panelists include the following:

1. Financial creativity;
2. Strong consumer economy and awareness;
3. Highly motivated, informed, and excited state and local government officials;
4. First generation architecture;
5. Sufficient standards to accelerate deployment; and
6. Defense technology players involved and integrated into the market.

The following recommendations suggested by the panel represent shared initiatives by both DOT and industry participants which can feed directly into realization of the emerging vision of ITS. However, they also represent only a possible sample of appropriate responses. Further evaluation is needed to determine which sector should take the leadership role in each of these areas.

The recommendations are categorized into three fields: education and outreach, sustaining ITS momentum, and finally, market risk reduction. These three fields can then be overlaid against the four-pronged basic template established to guide the panel discussion.

Regional Outreach and Education

1. Provide access to financial models for state and local governments. One of the shortcomings articulated by the panel participants was a lack of working financial models to assure funding and deployment of important ITS technologies. Recommendations for improvement would include exploration of successful models used in industry in similar emerging high technology areas in order to develop a series of best practices which might be applied to the ITS domain. These best practices would be representative not only of public-private partnerships but also private-private partnerships. After identifying these practices, dissemination to state and local stakeholders in written form and perhaps accompanied by a workshop series with a regional orientation would be recommended.

2. Redefine financial and management roles and further explore privatization. Traditional public roles with respect to financing and operational management of ITS systems might be appropriately transferred to the private sector. This transfer may allow for improved services and may likewise stimulate the market opportunity of ITS technologies. Further exploration of privatization issues was suggested.

3. Arrange for DOT state and local visits to key companies on specific ITS-related technology areas or to investigate potential solutions to specific problems. Many industry panelists encouraged this sort of exchange as a valuable way to uncover a number of technologies suitable for application to needs resident within these public sector stakeholders' purview, as well as a method to continue interaction. Likewise, company leaders reaffirmed an openness and willingness to entertain such mutually educating visits.

4. Continue the process of matching defense companies' technologies to potential markets and cost-effective solutions to problems. Defense companies must maintain their proactive information-gathering processes on the shape and prospects of new markets. They must also continue examining the ways that their technologies can solve problems or provide advantages over existing systems to achieve the goal, as expressed by the panel, of full integration into ITS.

5. Instigate a series of state governors' meetings. In key states affected by the defense downturn, a meeting between corporate CEOs and other company leaders with that state's governor and staff, transportation officials, and others would serve many pur-

poses. Besides bringing visibility to the issue of the applicability of ITS as an economic stimulus with far-reaching social and infrastructure benefits, it would help facilitate dialog and provide a forum in which to exchange ideas. Moreover, it would sustain and perhaps increase momentum on the part of industry by involving upper management. The spinoff effects could include increased corporate investment and commitment.

Sustaining ITS Momentum

6. Leverage other federal funding sources such as TRP and state defense conversion assistance funds. The presence of the federal government in ITS is vital to the defense industry. To upper management in many of the companies represented by the panel, a federal presence provides the impetus for embarking on the difficult process of transferring military technology to a civilian environment. Similarly, if the defense industry understands how its core technologies can be applied to solving difficult transportation problems, they can also leverage the defense conversion funding that became available in fiscal year 1993 and is continuing at a minimum through fiscal year 1995 in the TRP.

Access to additional conversion assistance funds available at the state level may also be possible, representing a very real opportunity to leverage federal ITS investments for the benefit of increased transportation research and development. To do so, however, requires immediate understanding of dual-use opportunities on the part of defense companies, thus closely linking this particular recommendation to the education and outreach activities.

7. Generate excitement via technology demonstrations by the defense industry to state and local officials, Congress, and metropolitan planning organizations, among others. Just as the nation was awed by visual displays of sophisticated weaponry during the Persian Gulf War, the transportation community and its stakeholders can be equally impressed by potential dual-use technologies applicable to ITS. Demonstrations of technology are powerful and should be used with more regularity during the course of conferences and through other mechanisms.

8. Develop and build consumer awareness. Likewise, consumer awareness and excitement can be created by technical demonstrations. A number of other additional avenues in which to build consumer acceptance exist and are currently under study by DOT.

9. Tap into existing transportation associations. Defense companies can further their understanding of new markets such as ITS while at the same time cultivating an appreciation for the technology they have developed by reaching out to associations that represent state and local transportation officials and stakeholders. Associations gain similar benefits by expanding their membership to include these companies and their supplier community. Developing a connection to such organizations as AASHTO and actively participating in special sessions at ITS AMERICA and TRB annual meetings would be possible activities related to this recommendation.

10. Leverage operational tests and results. Operational tests represent a very real opportunity to showcase not only technology but innovative public-private partnerships and financial arrangements. Likewise, operational tests provide for increased market knowledge, state and local government buy-in, and development of consumer acceptance and interest. They also can create test beds in which to highlight defense companies' impressive systems integration and program management skills.

Market Risk Reduction

11. Promote an ITS industry using other emerging industries as models. ITS is certainly not a single, monolithic entity. The emergence of ITS will most likely spawn several new business areas never before in existence. This is a challenge to all companies involved in this evolution, including defense companies, small businesses, and traffic-engineering firms. Similar challenges are faced by their partners: federal, state, and local governments. Over the past decade, however, the creation of new businesses has occurred in other technological domains. Insights and usable models may be available from the computer, electronics, and biotechnology industries. Drawing on experiences from past emerging markets would allow DOT to spur the process of creating an ITS industry.

12. Explore barriers to new ITS businesses. Examination of the new businesses that do emerge as ITS evolves, and the barriers facing this evolution, would be particularly salient in light of the partnership existing between government and industry in this endeavor. Technological barriers as well as institutional and business-related challenges need to be defined and modeled in order to fully understand the risks and opportunities involved. These two topics relate closely to the roles of government and industrial policy, which some industry panelists felt made them particularly timely.

13. Encourage rapid development of appropriate interconnectivity standards. In principle this was generally accepted but careful selection of the topics is still necessary. A few panelists felt that market development could be accelerated greatly based on increased activity in this area.

14. Accelerate systems architecture development and dissemination of results. Many panelists called for the availability of systems architecture results as soon as is feasible. Thorough and timely dissemination of this work was deemed essential, particularly to those in the product development process.

CONCLUSION

The fourteen recommendations under the three fields of education and outreach, sustaining ITS momentum, and market risk reduction can likewise be captured under the four-pronged template that represents the critical components to overcoming barriers to full ITS deployment and integration. Figure 2 represents this overlay. These recommendations can form the building blocks of appropriate industry and public-sector responses in facing these barriers. However, evaluation is needed as to assignment of leadership roles in order to implement these recommendations. In its desire to establish a healthy and functioning public-private partnership, DOT and state and local officials may be required to proactively plan for and implement these recommendations. Defense company leaders have articulated a desire to participate and similarly lead as deemed appropriate.

In conclusion, after assessing the barriers faced by the major defense companies attempting to enter the ITS market (both observed by SRI and articulated by company leaders themselves), this study concludes that some progress is being made toward resolving these challenges in the four areas of technology transfer, organizational culture, marketing, and finance. This appears to be spurred on by many of the companies' own actions. However, several of the recommendations suggested in this report can facilitate and perhaps accelerate this course and provide an opportunity for DOT and other transportation agency officials to nurture the needed

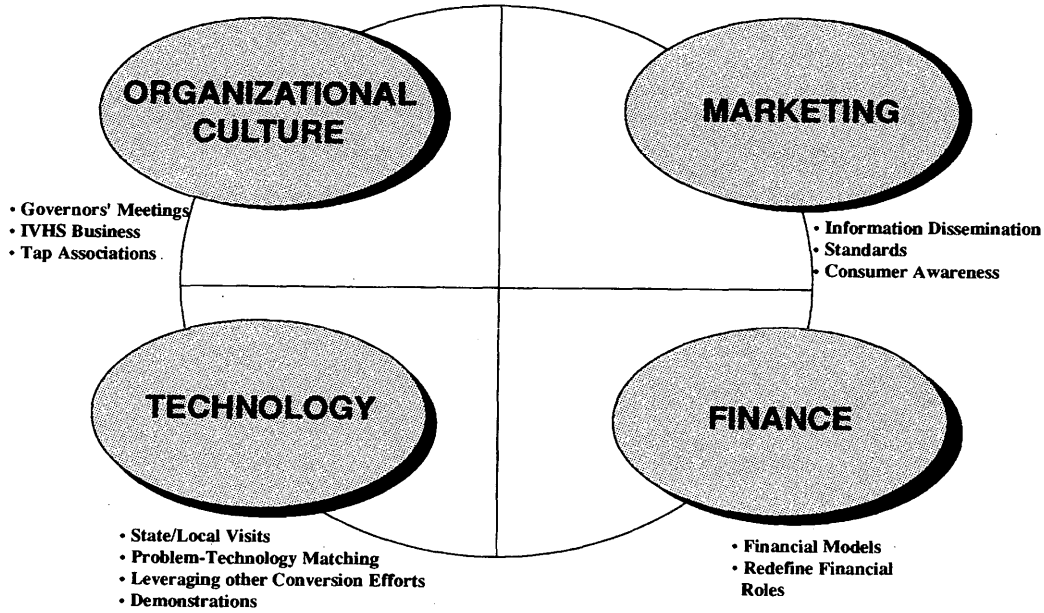

FIGURE 2 Four-pronged template with recommendations overlay.

public-private partnership. Certainly public-sector and industry collaboration in executing selected recommendations will build on the dialog established in the panel meeting and will further enhance this evolving relationship.

Additional analysis may be needed to bring into clear focus the shape and potential of this emerging paradigm for an evolving ITS business which will significantly affect the vitality of the nation's transportation system. However, the opportunity for DOT to be a leader and catalyst in this process is perhaps exactly what President Clinton articulated in his technology plan (*11*):

> American technology must move in a new direction to build economic strength and spur economic growth. The traditional federal role in technology development has been limited to the support of basic science and mission-oriented research in the Defense Department, NASA, and other agencies. This strategy was appropriate for a previous generation but not for today's profound challenges. We cannot rely on the serendipitous application of defense technology to the private sector. We must aim directly at these challenges and focus our efforts on the new opportunities before us, recognizing that the government can play a key role helping private firms develop and profit from innovations.

The promises of an invigorated economy and productive repositioning of national resources lie implicit in the realm of ITS and will require commitment on the part of both public and private stakeholders to ensure the realization of this vision.

ACKNOWLEDGMENTS

This research was undertaken with support from the Office of Traffic Management and ITS at FHWA and the Volpe National Transportation Systems Center. The efforts of Ronald Fresne and Susan Russell of SRI International during the survey research portions of the study were also instrumental in completing this work. Finally, the defense company leaders who participated in the study are gratefully acknowledged for their time and candid insights.

REFERENCES

1. *After the Cold War: Living With Lower Defense Spending.* OTA-ITE-524. Office of Technology Assessment, U.S. Government Printing Office, Feb. 1992, p. 32.
2. Reilly, L. Big Platform Primes Hold Their Own . . . But Smart Arms Have a Bright Future. *Washington Technology*, 24 Sept. 1992, pp. 26–29.
3. Stewart II, W. G. From War to Peace: A History of Past Conversions. *Report of the Defense Conversion Commission*, Jan. 1993.
4. Skinner, L. On The Virtual Road: No More Maps in Laps. *Washington Technology*, 23 June 1994, p. 10.
5. Peterson, C. N., and D. S. Olson. *Institutional Barriers to IVHS: Defense Company Interviews and Case Studies.* Prepared for FHWA, U.S. Department of Transportation, and Volpe National Transportation Systems Center, Cambridge, Mass., April 1994.
6. Peterson, C. N., and D. S. Olson. *Institutional Barriers to IVHS: Source List.* Prepared for FHWA, U.S. Department of Transportation, and Volpe National Transportation Systems Center, Cambridge, Mass., Oct. 1993.
7. Platt, D. J. *Defense Drawdown: Financial Overview and Strategies for the Top 25 Prime Contractors.* Prepared for Defense Conversion Commission, Washington, D.C., 1993.
8. Zoretich, T. J., et al. *The Impact of Reduced Defense Spending on U.S. Defense Contractors.* Prepared for Defense Conversion Commission, DRI/McGraw-Hill, Washington, D.C., Feb. 1993.
9. *100 Companies Receiving the Largest Dollar Volume of Prime Contract Awards: Fiscal Year 1992.* Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports.
10. Peterson, C. N., and D. S. Olson. *Institutional Barriers to IVHS: Sample Contractor List.* Prepared for FHWA, U.S. Department of Transportation, and Volpe National Transportation Systems Center, Cambridge, Mass., Nov. 1993.
11. Clinton, William J., and Al Gore, Jr. *Technology for America's Economic Growth, A New Direction to Build Economic Strength.* 22 Feb. 1993, p. 1.

Publication of this report sponsored by Committee on Applications of Emerging Technology.

Organization for Intelligent Transportation System: Computer-Integrated Transportation

RANDOLPH W. HALL, HONG KAM LO, AND ERIK MINGE

Computer-Integrated Transportation (CIT) is envisioned as an integrated network of public and private transportation organizations, each with unique responsibilities but working toward a common mission of facilitating travel across all modes of transportation. The objective of this study was to evaluate alternative frameworks for CIT from an institutional perspective. This was accomplished through site visits and interview at existing transportation management centers (TMC) along with focus group sessions in which strategies for CIT were presented to TMC managers and staff (only the latter is reported here). The study found that four factors have profound implications of Intelligent Transportation System (ITS) implementation and research: (1) time-frame, (2) linking information to actions, (3) broadcast orientation, and (4) embracing of new technologies.

Advances in information, computation, and communication technologies in the 1970s, 1980s, and 1990s have stimulated remarkable changes in business practices throughout the world. For instance, with the advent of computer-integrated-manufacturing (CIM), it is now possible to track accurately, automate, and control production from the moment raw materials are extracted from the ground until finished products are delivered to customers. Today, intelligent transportation systems (ITS) offer much of the same promise for transportation that CIM held for manufacturing. Encompassing a spectrum of electronic and communication technologies, ITS may one day achieve computer-integrated-transportation (CIT), where both the users and operators of transportation systems can obtain and exchange information effortlessly, to facilitate travel across all modes of transportation.

CIT is envisioned as an integrated network of public and private transportation organizations, each with unique responsibilities, but working toward a common mission of facilitating travel across all modes of transportation. The CIT is designed to achieve effective coordination of the transportation system, while respecting the responsibilities of participating agencies. Within these bounds, the CIT draws on resources (e.g., emergency crews, traffic control, etc.), both internally and externally, as needed to ensure the smooth operation of the transportation system.

The objective of this paper is to evaluate alternative frameworks for achieving CIT from an institutional perspective. To this end, existing transportation management centers (TMCs) both in California and around the country were surveyed in depth to assess existing capabilities. Site visits were conducted at all Caltrans TMCs as well as at three city TMCs (Anaheim, Los Angeles, and San Jose) (*1*). These interviews were followed up by a series of four focus group sessions, in which strategies for CIT were presented to TMC managers and staff for their comment and discussion. Finally, a follow-up survey was administered to each TMC to assess future directions for California TMCs. The initial survey, focus groups, and follow-up surveys are the basis for our evaluation of alternative frameworks for CIT. In-depth results are provided in Hall et al (*2*). Because of length restrictions, this paper focuses on the focus group aspect of the study.

The remainder of the paper is divided into seven major sections. First, the concept of CIT is introduced, along with key organizational issues associated with CIT. Next, a literature review is provided, concentrating on organizational designs for transportation management centers, and issues in CIM. This is followed by summarized results from focus groups with TMC personnel. Finally, survey and focus group findings are interpreted and recommendations are provided on how to implement CIT.

LITERATURE REVIEW

By far the most extensive study on TMC organization is a report by Booz-Allen and Hamilton, *Institutional Impediments to Metro Traffic Management Coordination* (*3*). The study includes a literature review on organizational theory and TMC practices, as well as results of interviews with TMC personnel in six metropolitan areas. The study further provides a list of 30 recommended solutions, most important of which include developing a "vision of evolutionary ATMS [Advanced Traffic Management System] implementation" and developing "work plan guidelines for implementing ATMS" ("a step-by-step 'cookbook' approach for implementing one, or more, ATMS technologies in an area").

Carvell et al. (*4*) provide a case study on improving interagency coordination, based on a traffic signal control project in North Dallas County. The project avoided disagreements by first creating a multi-agency steering committee and then developing guidelines aimed at promoting cooperation. These included funding restrictions (only supporting projects that would benefit multiple cities), procurement coordination (cities used normal procurement procedures, but submitted documentation to the steering committee for approval), and hardware flexibility ("cities were free to use their own controller specification but it had to contain minimum criteria").

Other relevant research includes papers on ITS system architecture. Varaiya (*5*), for instance, describes a layered structure that, to a degree, also defines an organizational structure. However, Varaiya's work is not directed at the institutional issues that arise in

R. W. Hall, Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, Calif. 90089. H. K. Lo and E. Minge, Partners for Advanced Transit and Highways, University of California, 1301 S. 46th Street, Richmond, Calif. 94804.

cross-jurisdictional coordination. Hall (*6*) proposed a framework for defining transportation architectures that spans modes (for both goods and people). He classifies architectures along the dimensions of communication medium (e.g., audio, visual, electronic, mechanical, and verbal/nonverbal), assignment of functions to transportation entities, and degree of coordination. This framework will be used in structuring organizational designs within the proposed work. The National IVHS (Intelligent Vehicle Highway System) System Architecture project has produced a variety of documents defining lines of communication between organizations (*7*) but not addressing specifically organizational responsibilities. As part of this effort, Rockwell International's document on evolutionary deployment (*8*) discusses how user services might be bundled into market packages and deployed over 5-, 10-, and 20-year time scales but does not address organizational responsibilities.

ORGANIZATIONAL ISSUES

This section is divided into three sub-sections: organizational structures, assignment of ITS functions, and internal organization. These concepts served as the framework for focus group discussions and address strategic issues in the design of CIT.

Organizational Structures

The organizational structure defines responsibilities by jurisdiction and defines the patterns of coordination and communication. The goal is to enable both public and private agencies to work effectively with each other. Two fundamental alternatives are "leadership" and "decentralized" structures, as discussed below.

Leadership Structures

Under this vision, certain transportation management centers are designated (or created) to act as leaders among satellite centers. Coordination may occur at any of several levels, ranging from simple exchange of information and responding to requests to active control. Leadership can be defined on a functional basis or on a locational basis, as discussed below.

Functional Leader

For a metropolitan region, the functional leader is responsible for coordinating a function, such as arterial signal control, incident response, traveler information, etc. (Figure 1). The leader TMC is activated when there is a need to pull together resources across jurisdictional lines or to synchronize across jurisdictional lines.

Locational Leader

Within a district, the locational leader is responsible for coordinating all functions (Figure 2). The leader TMC is activated when there is a need to coordinate resources across functions. Locational leaders communicate as equals with each other. For example, a locational leader might be responsible for coordination of all ITS functions (signal control, traveler information, vehicle identification, etc.) within a county. Its satellites could then be TMCs that specialize in particular ITS functions that serve the county.

Decentralized Structures

Under this vision, no TMC is designated as a leader (Figure 3). Coordination is achieved through exchange of information, perhaps accompanied by protocols as to how one jurisdiction should respond to another (in a manner like mutual aid pacts for fire districts). However, no TMC assumes leadership over others.

Information Exchange

The type of information exchanged between agencies and the protocols in place for how an agency responds define the "degree of coordination." The degree of coordination is a spectrum with control at one extreme and isolation at the other. Lo et al. (*1*) identify four steps in this spectrum:

- Coordination via occasional meetings, phone calls, faxes or electronic mail;
- Established data links among TMCs, so that the TMCs can observe each other's real-time traffic patterns and controls;

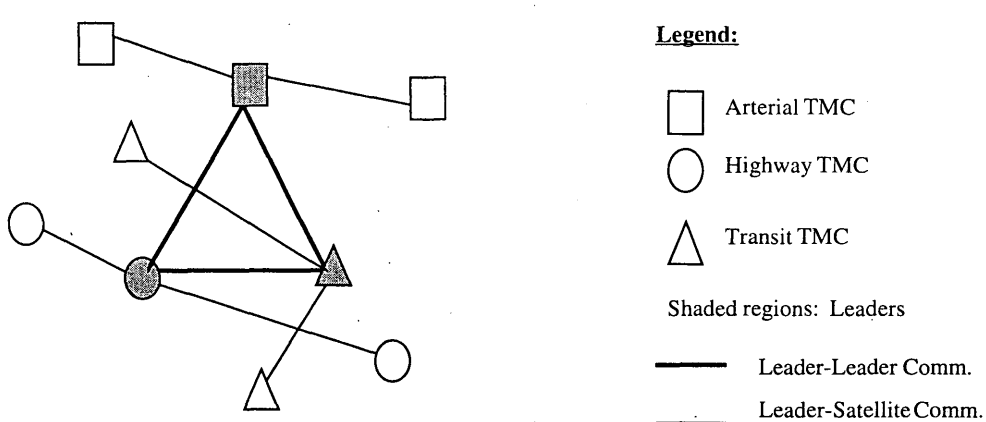

FIGURE 1 Organizational structure for functional leadership

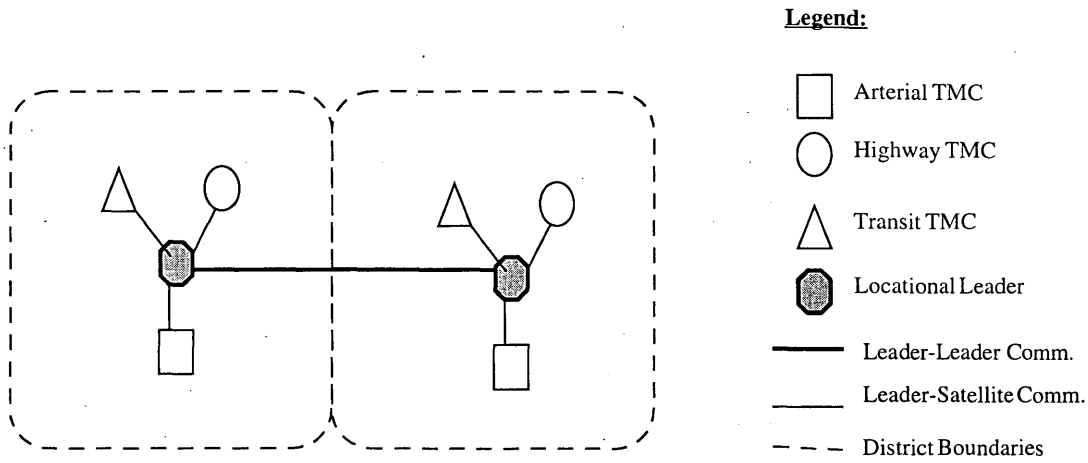

FIGURE 2 Organizational structure for locational leadership.

- TMCs not only observe each other's patterns but respond to patterns of external agencies through prescribed protocols; and
- TMCs not only exchange information, but TMCs are empowered to issue commands to external TMCs under prescribed conditions (such as from a leader to a satellite).

Assignment of ITS Functions

The CIT should be capable of providing a range of ITS functions today and be flexible for expanding to new ITS functions as they are developed. The organizational structure provides a framework for determining where and how these functions are implemented and coordinated. As an ideal, such a framework would remove these decisions from the burden of political wrangling and speed the adoption of ITS.

Managerial Reporting and Control

This category of functions is aimed at long-term improvements in the transportation system. Through daily, weekly, monthly, and annual reporting on system performance (including accident statistics, travel times, on-time performance, patronage, congestion, etc.), the reporting system draws managerial attention to the most urgent problems, and speeds their resolution. Managerial reporting relies on a range of ITS technologies, most importantly surveillance, communication, and management information systems. Most importantly, managerial reporting offers the impetus for continuous improvement in overall system performance by drawing attention to the most critical problems.

Operational Control

The minute-to-minute decisions needed to keep the transportation system up and running, at maximum efficiency, fall in the category of operational control. *Normal Operational Control* pertains to operation in the absence of unusual disturbances or incidents. Mostly these functions do not require human intervention. Examples include adaptive signal control and automated toll collection. *Incident Based Operational Control* pertains to operation in the event of accidents, adverse weather, stalled vehicles, etc. These functions ordinarily demand human intervention, both at the site

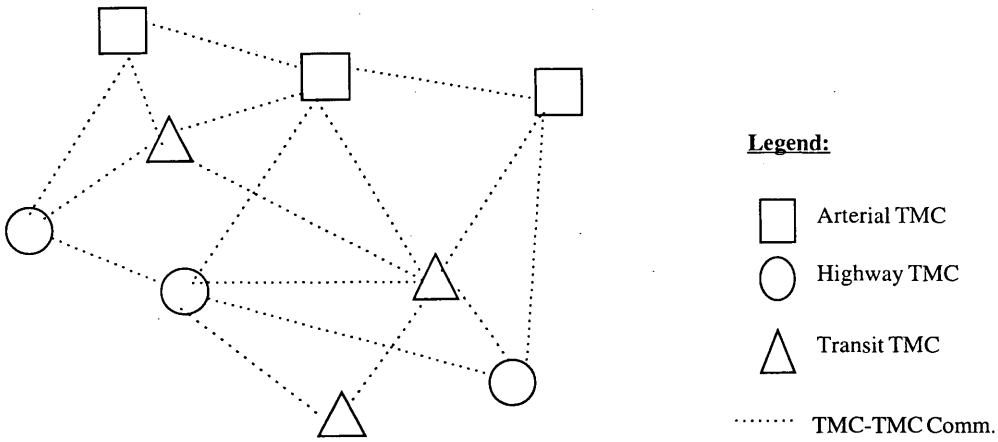

FIGURE 3 Decentralized organizational structure.

by incident response crews and at the TMC, to direct the response, employ changes in signal plans, and send out traveler advisories.

Internal Organization

Within any organization, an ITS function can be implemented in a variety of ways. It might be centralized at a management center, or it might be distributed among field units. A function might be computer-automated, or it might entail extensive human intervention. Finally, a function might be specialized, or it might be an aspect of higher level functions.

Centralized Versus Distributed

Most ITS functions require communication of surveillance information from the field, assessment of the information, and execution of actions. In a centralized system, decision making (whether automated or not) is concentrated at a single location. As a consequence, communication requirements may be large whereas, on the positive side, scale-economies might be exploited. In addition, centralization distances the decision makers from the field (the consequence being a loss in familiarity with actual conditions). In a distributed system, decision making is localized, perhaps in lower level management centers or perhaps in fully automated field units.

Automation

Automation can free humans from the more tedious work, so they can focus on higher level decisions. Already, incident detection algorithms can free up operators' attention, and adaptive signal control can reduce the need for manual overrides. However, during abnormal situations, it may not be appropriate to rely on automatic responses. Computers might then serve as decision support tools rather than decision making tools. TMCs could activate, for instance, incident management procedures only after a human confirmation, or perhaps the human could actively direct the response procedures. The primary issue is then which types of ITS functions demand human involvement, which types would benefit from a combined computer/human approach, and which can be entirely automated.

Specialized Versus Encompassing TMCs

Most existing TMCs are geared toward performing two major functions: incident management and signal control, both of which entail various supporting functions (e.g., dispatching emergency vehicles or traffic surveillance). However, many of the proposed ITS functions do not fall neatly into the existing categories. AVI for automatic toll collection, for instance, conceivably might be integrated into general roadway surveillance and potentially even supplant it. Two distinct avenues for implementation are (1) to incorporate new functions within the functions served by existing TMCs, or (2) to create new TMCs that operate with some degree autonomy.

FOCUS GROUPS

The aim of the focus groups was to facilitate discussion among experts in TMC operations regarding organizational structure, functional assignment, and internal operations, as a way of documenting the strengths and weaknesses of the various alternatives. A total of four meetings were held, two in Northern California (at Caltrans District 4 headquarters) and two in Southern California (at Caltrans District 12 headquarters). In each location, a morning meeting concentrated on managerial issues and an afternoon meeting concentrated on technical issues.

Several weeks before the focus groups, invitees were mailed a copy of the PATH working paper on existing TMCs and a discussion paper on issues in TMC organization. The format of the meetings was discussed in the cover letter, and an agenda was provided. A total of 50 people attended the four meetings, some of whom attended multiple meetings.

The meetings began with introductions, a 10-minute review of existing TMCs, and a 10-minute presentation on TMC organizational issues (following the format of Section 3). At this point, the thrust shifted to discussion. In the managerial sessions, the discussion was divided into four blocks, centering on the following issues: (1) traffic signal control and coordination, (2) incident management, (3) automated vehicle identification, and (4) managerial reporting and control. Each topic was introduced by presenting a list of opportunities and concerns. The facilitator opened discussion to the group by requesting comments on institutional aspects of the technologies and the strengths and weaknesses of alternative organizational designs. The discussion was directed only to the extent needed to keep on topic and on schedule and to allow all participants opportunity to speak. The format for the technical sessions was similar, with the exception that managerial reporting and control were not included and participants were instructed to focus on technical issues, such as communication protocols and software requirements.

Findings

The findings are categorized according to major themes, which tended to be common among all focus group sessions.

Funding

Perhaps not surprisingly, there was a strong consensus among all four groups that carefully directed funding was essential to implementing ITS. Local agencies, especially small cities, lack both the budget and staff to implement existing technologies, and it seems unrealistic that these agencies would divert already tight funds to advanced technologies. Beyond the basic budget squeeze, there was a strong sentiment that state and federal funding should be used to leverage agencies toward better coordination, by targeting funds toward inter-jurisdictional projects. Comments in this regard were highly consistent with Booz-Allen & Hamilton's conclusion that "the political unit that controls the financial resources has the ability to shape how ATMS is provided."

On the contrary, many felt that funding would be more plentiful if the benefits of ITS were documented carefully. The FHWA Field Operational Test (FOT) program was identified as an important element of this effort. However, evaluation was also viewed as impor-

tant to non-FOT projects, and many felt that ITS systems should routinely generate data for evaluation purposes. As in the Booz-Allen & Hamilton study, rigorous cost-benefit analysis was strongly supported.

Maintainability

There was also a strong sentiment that more care should be given to ensuring the long-term maintenance of systems as they are implemented. Participants expressed strong concern that the systems being implemented today may have diminished effectiveness because they are difficult to maintain or because there is insufficient funding for maintenance. A solution would be to include maintenance cost as an explicit factor in system selection, along with redesigning systems if maintenance costs are prohibitive. In addition, a life-cycle budgeting approach was supported to ensure that future funding is sufficient for adequate maintenance.

Coordination

There was considerable discussion on the viability of leadership-based organizations, with no strong consensus. On one hand, many argued that no agency would yield "control" of their transportation system to another because of liability reasons or a desire to retain "ownership." As stated in the Booz-Allen & Hamilton study, "by agreeing to permit another entity to 'control' their infrastructure, they may believe that others will make decisions and take actions that may not be supported by their own constituents."

On the other hand, several examples were cited of cities that had successfully turned over control of traffic signals to other agencies without encountering major obstacles. Overall, there appeared to be agreement that adoption of leadership type organizations hinged on three critical factors: (1) funding incentives, (2) demonstrated benefits, and (3) coordination from a neutral agency, most likely a Metropolitan Planning Organization. These factors are needed to convince governing bodies to participate in such efforts. On the other hand, Booz-Allen & Hamilton's conclusion that "regional ownership is unlikely" also likely holds true, and that a more realistic scenario would be where each jurisdiction retains ownership but allows regional coordination under tightly prescribed conditions.

There appeared to be few obstacles to decentralized structures, so long as this was interpreted as simple information exchange without control. However, participants were skeptical that information exchange was sufficient to achieve coordination. Success would depend on the procedures enacted to respond to information, which would require careful study.

Overall, participants appeared to be less concerned about the type of organization structure than about ambiguity. For instance, participants saw fewer problems with an outside agency completely taking over operations than with an outside agency that might occasionally assume control (perhaps in response to an incident). Hence, there was a strong consensus that whatever organizational structure is implemented, roles and responsibilities must be defined precisely.

Conflict Resolution

Several participants commented that coordinated systems are difficult to implement because different agencies have different goals and objectives. For instance, some cities are quite supportive of projects to improve throughput on major arterials and to allow their use for diverted freeway traffic, whereas many are absolutely opposed. Support or opposition often can be traced to traffic impacts on residents and the significance of the traffic to the city's tax base; and because city councils respond to differing constituencies, they naturally have different objectives.

As one participant stated, "when the vision is common, the opportunities are there." There was a strong consensus that transportation agencies need to define such a vision and to establish processes for resolving conflicts when they arise.

Information Exchange

No one stated that major technical obstacles stand in the way of coordinating ITS systems. What is most needed is to define the interfaces. As a first step, high priority was given to developing interchange standards for signal plans. This might be followed by interface standards for other elements of the transportation system. In this regard, the TravInfo project was cited as an example for traveler information. There was a consensus, however, that standards should be devised by committees of experts through a consensus process and not legislated or imposed by higher level agencies.

Public Image

Participants were concerned that all projects be sufficiently well conceived to pass the test of public scrutiny. This meant that ITS applications should provide tangible benefits to individuals and should avoid traffic enforcement or other aspects of control. Public image was an especially large concern in AVI systems. There was a consensus that agencies should be totally open with the public about all possible uses for AVI information before deployment. Furthermore, many felt that AVI tags should only be used for limited purposes, such as toll collection. Broadening their use to traffic surveillance, vehicle inspection, etc., especially after the fact, was viewed as risky, because it engenders the "Big Brother" specter.

Management Reporting

There was a strong consensus that the success of ITS hinged on demonstrated cost-effectiveness. To this end, it was suggested that the State of California establish a "mobility index" that would be a common yardstick used in all regions to measure the performance of the transportation system on a daily, weekly, monthly, and annual basis. To this end, it was viewed as essential that new systems have built-in capabilities for archiving data so that these statistics could be generated automatically and that the State should develop standards for how these data are reported. This may result in a uniform management information system that enables access to a broad range of transportation statistics including delays, traffic volumes, transit usage, and accidents. This information could then be used for an array of purposes including staffing, safety improvements, and transportation planning.

Comments on Technologies

Participants also provided specific suggestions on the implications of ITS for signal systems, incident management and AVI.

Signal Systems

Considerable enthusiasm was expressed for using ITS to coordinate signals between adjacent jurisdictions and between arterials and ramp meters. The preferred approach was first to establish uniform protocols for exchanging information on signal plans and to coordinate responses through multi-jurisdiction signal committees. This was viewed as especially important in areas where Caltrans manages signals at diamond interchanges and along major arterials that pass through multiple jurisdictions and/or parallel major freeways. At the same time, participants recognized that these steps might be insufficient in the future with the deployment of adaptive signal systems.

Incident Management

Participants saw considerable opportunity for ITS in incident management including (1) closed-circuit-television (CCTV) for incident verification, (2) remote command centers to coordinate incident response, and (3) improved communication to, from, and at the scene through wireless technologies. The remote command center might be especially effective in dispatching hazardous materials crews, specialized maintenance equipment, or the coroner to the accident scene.

Automated Vehicle Identification (AVI)

Suggested applications of AVI included toll collection, air quality enforcement, mayday/safety devices, commercial vehicle inspections, and fleet management. Many participants were skeptical about the use of AVI for law enforcement purposes, whereas others felt this was viable.

The focus groups concluded with a discussion of ITS opportunities and obstacles. Opportunities included (1) faster and better information for travelers and TMC operators, which would enable better choices, (2) multi-agency and multi-modal coordination, and (3) creation of additional resources. Obstacles included (1) individual agencies that may not strive for the common good, (2) resistance to change and "turf" battles, (3) inability to maintain systems, (4) liability and privacy concerns, and (5) inability to fund deployment and operation.

DISCUSSION OF RESULTS

We have presented the vision of CIT, which is an integrated network of public and private organizations working toward a common mission of facilitating travel across all modes of transportation. ITSs serve as an enabling force for CIT, providing the technological capabilities for its fulfillment. Just as important, however, is how CIT fits within the institutional environment of state and local agencies. We have worked toward understanding how CIT and ITS can be implemented within the organizational framework of these agencies.

Throughout the study, we encountered considerable enthusiasm for ITS. Participants were nearly unanimous in their belief that ITS has improved their working relationships with other agencies and that one of the most important benefits of ITS was improved coordination. Although participants generally felt that a decentralized organizational structure was more practical within the current institutional environment, they felt that leadership-based structures might be viable in the future, provided that the benefits could be demonstrated.

There was also a strong commitment among state agencies and some local agencies toward implementing an array of ITS services. In most cases, participants felt that the ITS services should be assigned along the lines of current agency responsibilities. With respect to internal organization, the ultimate answer would depend on costs analyses and technical feasibility and not so much on institutional considerations.

Despite the enthusiasm for ITS, significant obstacles lie ahead. Participants were worried about the ability to fund and maintain ITS. They were worried that parochial interests might stand in the way of improved coordination. They were worried that some agencies were not sufficiently supportive of innovation and change. They were also concerned that the benefits of ITS might not be documented, which could stand in the way of future deployments.

Although the participants' focus was on the deployment of ITS, their comments also have relevance for the ITS research program. In some cases, this is not reflected so much in specific comments as in participants' priorities and attitudes, as expressed in focus group sessions. These are summarized below.

Time Frame

Nearly unanimously, participants were focused on short-range applications of ITS, mostly in the time frame of two years or less. For instance, participants showed considerable enthusiasm for, and detailed knowledge about, signal control systems and incident response strategies. On the other hand, medium-range applications such as AVI evoked much less discussion and interest.

Linking Information to Actions

Transportation management requires the coordinated effort of multiple agencies and multiple divisions within agencies. Unfortunately, it appears that some of these organizations suffer because the information is not being collected by the organization that is empowered to act on the information. This is most apparent in incident response strategies, where the focal points of transportation information, TMCs, have limited power in responding to incidents.

Broadcast Orientation

TMCs as they exist today disseminate information via broadcast technologies (changeable message signs, radio stations, etc.) and collect information in aggregate (mostly via loop detectors). ITS presents the opportunity for targeting information collection and dissemination to individual vehicles, drivers or travelers. AVI is one aspect of this opportunity. Other aspects include safety devices, in-vehicle signage and, eventually, automated highways. Evolution from a broadcast orientation to a "narrow-cast" orientation likely

will require significant changes in the function and organization of TMCs.

Embracing of New Technologies

Although all participants were enthusiastic toward ITS within the context of their current functions, there was some hesitation toward expanding their functions. Lack of funding was an obvious concern. Just as important, perhaps, was that some agencies are nervous that they will be perceived as invading someone else's "turf." Although this type of caution has helped create a cooperative spirit among agencies, it has also created an obstacle to innovations in how transportation is organized that ultimately may affect ITS implementation.

Organizational Boundaries and Performance Measures

There is a significant opportunity to use ITS systems for calculating transportation performance measures on an ongoing basis. These measures could be the basis for a performance-based incentive system, which could reward TMC personnel for their ability to respond to incidents and otherwise manage the transportation system. The measures further could be the basis for continuous improvement systems aimed at ongoing reduction in delay, accidents, and other impacts. To make such a program effective, further consideration is needed on organizational boundaries so that overlapping responsibilities across agencies and departments do not dilute TMC personnel's ability to manage the transportation system.

To conclude, the above four factors have profound implications for ITS implementation and research. Each demands careful deliberation at a strategic level and, possibly, changes in how transportation agencies are organized and how they relate to each other. From the research perspective, the greatest risk is that innovative ITS concepts may have no home for the following reasons: they are difficult to implement within existing organizations; there is no long-term plan for their incorporation and coordination; there is no one within operating agencies to advocate the concept; and agencies are not empowered to act on the information that they generate.

It should be pointed out that the organizational issues facing ITS are not unusual. Wilson (*9*) describes the importance governmental agencies place on autonomy. He states that by finding a unique functional niche, organizations avoid external competitors. In the process, however, they tend to avoid taking on new tasks that deviate from their traditional core responsibilities. As an example, he cites the Army's decision to develop a large helicopter fleet out of deference to an agreement with the Air Force that forbade purchase of fixed wing aircraft rather than to technical considerations favoring helicopters over alternatives. In ITS there are similar risks: that technological choice will be driven more by long-standing organizational functions than by what is best for the system.

To overcome these potential barriers, we believe that is it essential to reconsider the lines drawn between organizations, both within and between agencies, to determine whether they still make sense in the ITS environment. Further, it is essential to strengthen the dialogue between the researcher and practitioner communities, so that the two groups work together in developing a plan and a vision for computer-integrated transportation.

ACKNOWLEDGMENTS

The authors thank the California Department of Transportation for funding the study and participating in focus group sessions. Special thanks go to Jim Pursell, who was instrumental in coordinating the research.

REFERENCES

1. Lo, H. K., R. W. Hall, and J. R. Windover. *California Transportation Management Centers Part 1. Assessment of Existing Capabilities*, PATH Working Paper 93-17, 1993.
2. Hall, R. W., H. K. Lo, and E. Minge. *Organization for IVHS: Computer Integrated Transportation*, PATH Research Report 94-24, 1994.
3. Booz-Allen & Hamilton. *Institutional Impediments to Metro Traffic Management Coordination*, for Volpe National Transportation Center, Contract DTRS 57-89-D-00087, 1993.
4. Carvell, J. D., E. J. Seymour, R. L. Peterson and G. Human. Interagency Communication: A Vital Element in the Dallas Multijurisdictional Environment. *ITE Journal*, April 1994, pp. 12–16,
5. Varaiya, P. Smart Cars on Smart Roads: Problems of Control. *IEEE Transactions on Automated Control*, Vol. 38, 1993, pp. 195–297.
6. Hall, R. W. The Architecture of Transportation Systems. In *Progress in Material Handling Research*, Braun-Brumfield, 1992, pp. 77–96.
7. USDOT and ITS America. *ITS Architecture Development Program, Phase I Summary Report*, Washington D.C., 1994.
8. Ice, R., T. Smith, T. Horan, M. Al-Kadri, and E. Seymour. *IVHS Architecture, Evolutionary Deployment Strategy*, Rockwell International for USDOT (contract DTFH61-93-C-00212), 1994.
9. Wilson, J. Q. *Bureaucracy: What Government Agencies Do and Why They Do It*. Basic Books, 1989.

Publication of this paper sponsored by Committee on Transportation System Management.

The Case For High-Frequency ATIS

JOHN P. KILJAN

In this paper a case is made for using a little-used part of the shortwave radio spectrum as a cost-effective means of broadcasting digital traveler information to rural and intercity users. Although the subject is technical, the material is provided in an easy-to-digest format. The target audience is state and federal transportation officials who believe that providing real-time traffic information is as much a part of transportation as building and maintaining highways. The approach uses low-cost off-the-shelf technology for both transmitting and receiving, and promises state-wide coverage for a total investment comparable to a simple highway advisory radio installation, about $10,000. The paper also describes how state departments of transportation or state police organizations can set up wide-area low-cost broadcasting facilities using a simple personal computer work station as an interface. The preliminary research needed and a strategy for follow-up operational tests are also described. The proposal describes the benefits and discusses the downside risks of developing this cost-effective medium for broadcasting traveler information.

This paper is written for state and federal transportation officials who believe that providing real-time rural traffic information is as much a part of transportation as building and maintaining our nation's highways and who are interested in broadcasting traveler information over wide areas at low cost.

BACKGROUND

What are the best radio frequencies for providing advanced traveler information in support of the intelligent vehicle-highway systems (IVHSs)? It is a question asked by many who plan for IVHS projects and as yet, there is no clear answer. This paper offers one possibility optimized for rural and intercity traffic using a bearer that currently has no commercial advocates and uses low-cost off-the-shelf technology.

Many frequency bands have been proposed for one-way and two-way advanced traveler information systems (ATIS) services. They include very low data-rate AM broadcast band subcarriers that have wide coverage, abandoned VHF-low frequencies below 50 MHz to FM and television broadcast subcarriers, the newly allocated 220 MHz channels through more conventional (though well-occupied) commercial bands, including piggybacked cellular and pager channels, and microwave beacons to direct satellite broadcasts to infrared roadside beacons with the capacity of gigabits per second.

The available bit rates on each of these, the coverage areas, the cost and complexity of providing receivers and transmitters, and the infrastructure support needed vary widely with each proposal. There are almost too many choices, yet no one frequency band is perfect for all uses and each choice represents some kind of trade-off in performance versus cost.

Colorado Department of Transportation, 4201 East Arkansas Avenue, Denver, Colo., 80222.

Many proposals (though not all) suppose that the cost of establishing a broadcasting infrastructure (dedicated transmitters, broadcast towers, power feeds, data links to favorable sites, satellite uplinks, and the like) will limit general digital information broadcasts to modifications of existing services that can carry IVHS ATIS information on subcarriers or on a time-shared basis. Indeed, the cost of building and maintaining a new broadcast infrastructure—or even expanding it to new coverage areas—can be very high. The conventional wisdom of using what is already out there makes a lot of sense for a new service that must be supported, at least in part, by public agencies.

The problem is that these services still do not provide low-cost universal coverage. The difficulty is greater for rural areas of the country, where low population densities mean that there are few existing broadcast or paging facilities that can be adapted to provide ATIS services.

In this paper, one more possible ATIS bearer is proposed—one that uses dedicated channels at the extreme low end of the high-frequency bands (HF) or the upper ends of the medium-wave (MW) bands. The suggested system goes against conventional wisdom. Unlike proposals that try to maximize the use of existing commercial broadcasting infrastructure, HF ATIS requires establishing a new infrastructure with the hope of doing it at a very low cost.

Although referred to as HF in this paper to avoid confusion with the medium-wave AM broadcast band (.550 to 1.700 MHz), the proposed frequencies (2.00 to 2.10 MHz) are more properly a part of the MW bands. However, the service should work almost as well up to 3.50 MHz, passing the "official" 3.00 MHz dividing line between MW and HF (Figure 1).

THE VISION

The vision is simply a low-cost nationwide digital traveler information service targeted for rural and intercity travelers that is run as a partnership between the pubic and private sectors on a user-fee basis. The partnership would make the best use of a state's transportation and law enforcement agencies' ability to collect and store real-time traveler information and the private sector's ability to provide value-added services and market these services on a user-fee basis.

LOW COST AND RAPID IMPLEMENTATION

Much of the advantage of using the 2 MHz band is the low cost needed to set up an ATIS service—a fact that should particularly appeal to rural states that want to distribute traveler information over a wide area without investing a lot in labor and equipment. A complete low-power HF broadcast station should only incur an equipment cost of $5000. That is roughly comparable to the cost of

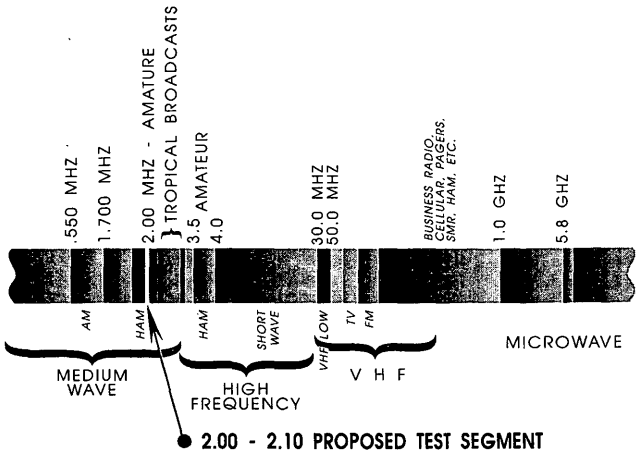

FIGURE 1 Simplified electro-magnetic spectrum.

a conventional highway advisory radio transmitter and would need no special installation site other than proximity to a state's traffic operations center or a link to a remote site.

Although there are currently no receivers being produced for use in vehicles for this kind of service, the cost to manufacture them should be very low. (There are receivers designed for receiving HF digital data in the maritime services and in the amateur services. The former are designed for high reliability in a hostile environment and the latter for low cost. Both are designed for use on many frequencies and many bands. For prototyping and operational tests either could readily be modified.) The circuitry needed to capture and demodulate the signals is only a little more complicated than a conventional AM transistor radio, which contains about $5 worth of parts. Even with the additional circuitry needed to provide digital detection and a standard RS-232 output port, the overall cost of the receiver in production could be as little as $50. It would fit into a package the size of a modestly sized internal or external computer modem. Indeed, the receiver would operate much like a modem, and for evaluation purposes a simple terminal program such as Procomm could be used to decode the messages if they were in a simple ASCII format. Receivers could be designed to provide audio tones compatible with existing computer modems or, for higher data rates and greater user simplicity, they could be combined into a single package that would provide decoded data over a standard RS-232 serial port. Figure 2 sketches some possible operational receiver configurations.

Despite the simplicity of broadcasting plain ASCII text, there is much to be gained by using standard message codes for traffic events and locations. Coded messages can be more readily sorted and stored for selected retrieval—a useful feature to prevent information overload for an expanded system. Much work is underway in this area. The protocols for coding digital traveler information are referred to as ITIS or ITIIS (International Traveler Information Interchange Standard). In North America, both ENTERPRISE (ITIS) and SAE (ITIIS) have technical working groups dedicated to developing an open standard for traffic messages, data storage formats, location coding, and dictionaries needed to support the coding of digital traveler information. Nothing would prevent dual broadcasting of both compact coded messages and plain ASCII text, each adapted to the sophistication level of the user's software.

Because of the low capital costs to be recovered by a completed system, the eventual user fees needed to support the service could be in the order of $10 a month if the service receives even a modest amount of support from its users. The receivers and the associated palm-top or desk-top computers needed to decode the messages could be built in as OEM equipment. However, such an architecture would also lend itself to third-party vendors or to add-on units to be added to rental car fleets in states such as Colorado, where rented cars are often used for intercity and recreational travel. The interface software that will be needed for desk-top or hand-held computers offers rich possibilities for the private sector to create and sell user-friendly programs.

Perhaps the best feature of such a system is that it could be easily and rapidly implemented, which would make it easy for any state or provincial transportation agency that can collect and process wide-area traveler information to provide ATIS services without large capital investments. The frequencies suggested, 2.000 to 2.100 MHz, are set aside for the federal government, which could authorize the state departments of transportation (DOTs) to operate the transmitters for them and set up the public-private partnerships needed to support the system on a user-fee basis.

WHY 2.0 MHz?

The suggested HF frequencies were found simply by listening for a few hours over the course of a year. It is educational for anyone to take a tour of the complete electromagnetic spectrum. If you have some time you can do that easily enough with one of the new high-tech broad-band general coverage receivers. But those of us who are not electrical or electronic engineering types can take a very quick tour with our imaginations.

Imagine the air around us being filled with low-energy photons that oscillate at frequencies from as low as 10 KHz, used to communicate to submarines around the world, through the frequencies of AM radio stations, through the suggested frequencies proposed in this paper, through world-wide shortwave frequencies, through short-range FM, television, and commercial business radio frequencies, through point-to-point microwave services' frequencies, and through the thousands of large "footprints" made by satellites in low-, medium-, and high-earth orbits. The cacophony of conversations that can be heard in a cocktail reception in a large hotel ballroom in a good-sized Washington, D.C. convention pales in comparison to what can be picked up by only a few feet of wire connected to a radio receiver, listening to the world talk to itself each day.

Listening to this raucous uproar, it is hard to believe that only 90 years ago it was all silent. Not a word nor conversation broke the silence that had lasted billions of years. Moreover, it has been a busy 90 years. The use of this limited resource has changed dramatically over the decades. Frequencies desperately needed in war were abandoned in peacetime, and receivers that could not fit into a large room now comfortably fit into the palm of the hand. Data has replaced much of the voice traffic. The spectrum from 2.0 to 3.5 MHz, which for decades reached out to so many maritime listeners, providing mobile services and tropical broadcasting, is now mostly unused in North America.

Where did all the users go? It seems that most have moved on to higher VHF channels for short-range communications or are relying on the global network of satellites for worldwide coverage. Although the international allocations have not changed in decades, there are now only a few dozen North American stations to be heard between 2.0 and 3.5 MHz in a typical evening of listening. The pro-

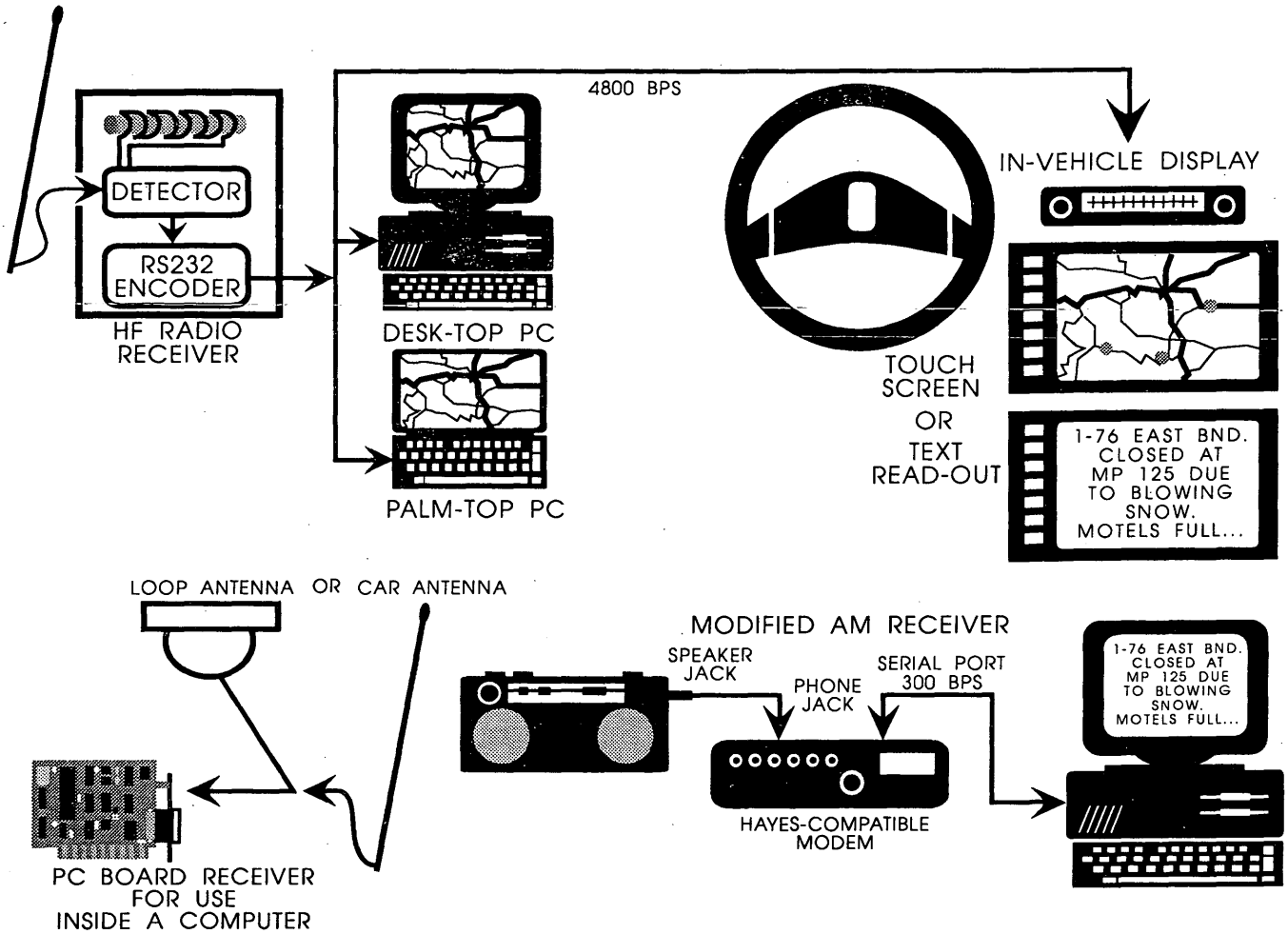

FIGURE 2 Possible receiver configurations.

posed test segment, from 2.00 to 2.10 MHz, is particularly quiet. This listener has regularly heard only three stations on the 20 or so available frequencies. The continuing abandonment of these frequencies by their former users means that there will little displacement, if any, of existing users. Even if the segment from 2.00 to 2.10 MHz is found to be unacceptable for administrative or technical reasons, there are a number of other potential segments between 2.0 and 3.5 MHz that could support a number of HF ITIS channels. Refer to the International Frequency Allocation Tables, which outline a number of segments that may be suitable.

The odd propagation characteristics, which are a disadvantage for other services using this part of the band, may be a plus for digital traveler information broadcasting of the type required by IVHS. Propagation in this frequency range is very different during the day and the night. During the day, ground wave and sky wave propagation is nearly ideal, with good coverage from 50 to 100 miles using a simple antenna. At night, the ground wave remains unchanged, but sky wave propagation is increased to carry signals much farther. The effect is familiar to anyone listening to commercial AM broadcast stations on what used to be called "clear channels." These signals can be heard 1000 miles or more away from the transmitter. Extreme nighttime coverage is undesirable, and the effect can be much reduced by configuring the transmitting antenna to direct the signals more or less vertically. This can be done with a basic dipole or end-fed wire antenna close to the ground. See Figures 3 and 4.

During the summer, the frequencies around 2 MHz are subject to static crashes from nearby thunderstorms. In the early days of digital communications using 5-bit teletype codes, this type of atmospheric noise and signal fading would cause "hits" that introduced errors into the copied text. Modern digital coding methods developed for maritime and other services using checksums, packets, and other forward error correction techniques have eliminated hits. Only correct text is received and interference simply reduces the overall data transfer rate.

Why not higher in frequency? Above about 3.5 MHz the HF spectrum is much better occupied and it is much more difficult to prevent signals from propagating far beyond their intended reception area. If the service is to expand nationwide, then higher frequencies must be avoided to prevent interference with stations in adjacent states. See Figure 5 for a representation of what typical low-power coverage areas might look like in the western United States.

Why not lower in frequency? This is certainly possible, but below 2.00 MHz there is much greater occupancy. The amateur 160-meter band from 1.8 to 2.0 MHz is well occupied in the winter months, and there are also other services in this segment. The segment from

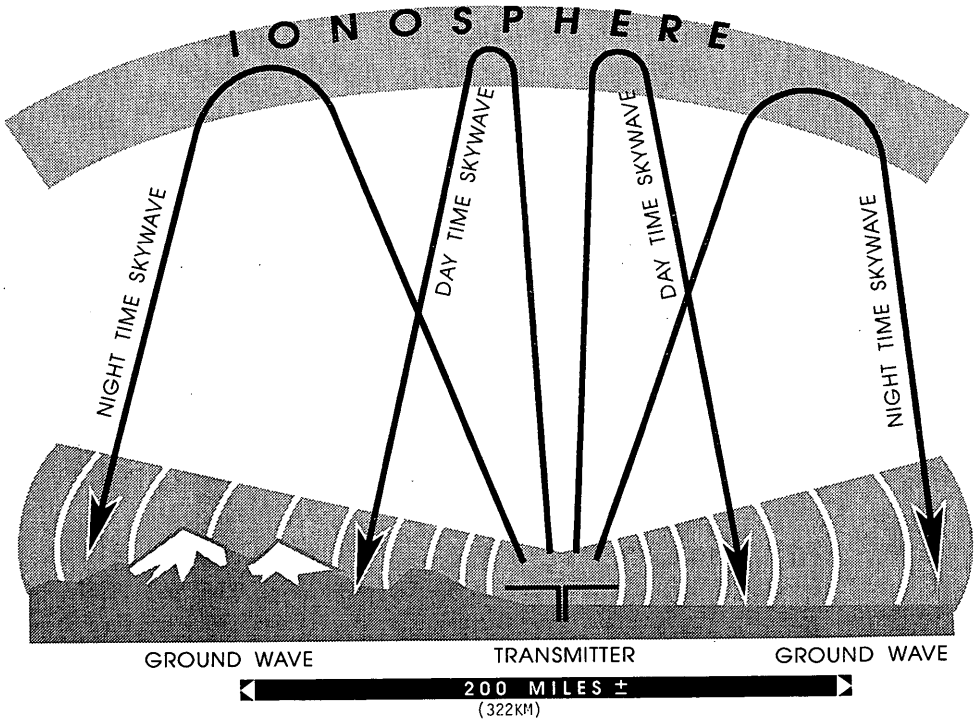

FIGURE 3 Skywave and groundwave propagation.

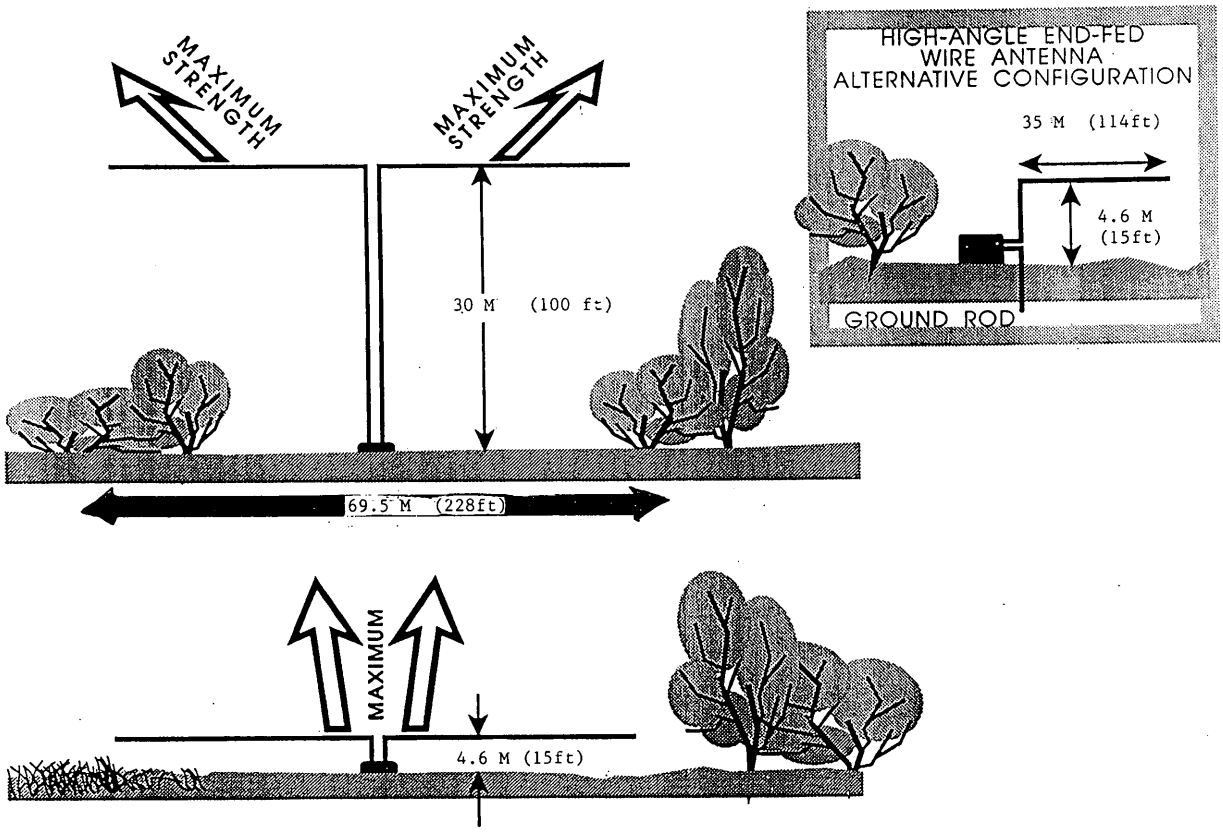

FIGURE 4 Radiation pattern versus antenna height.

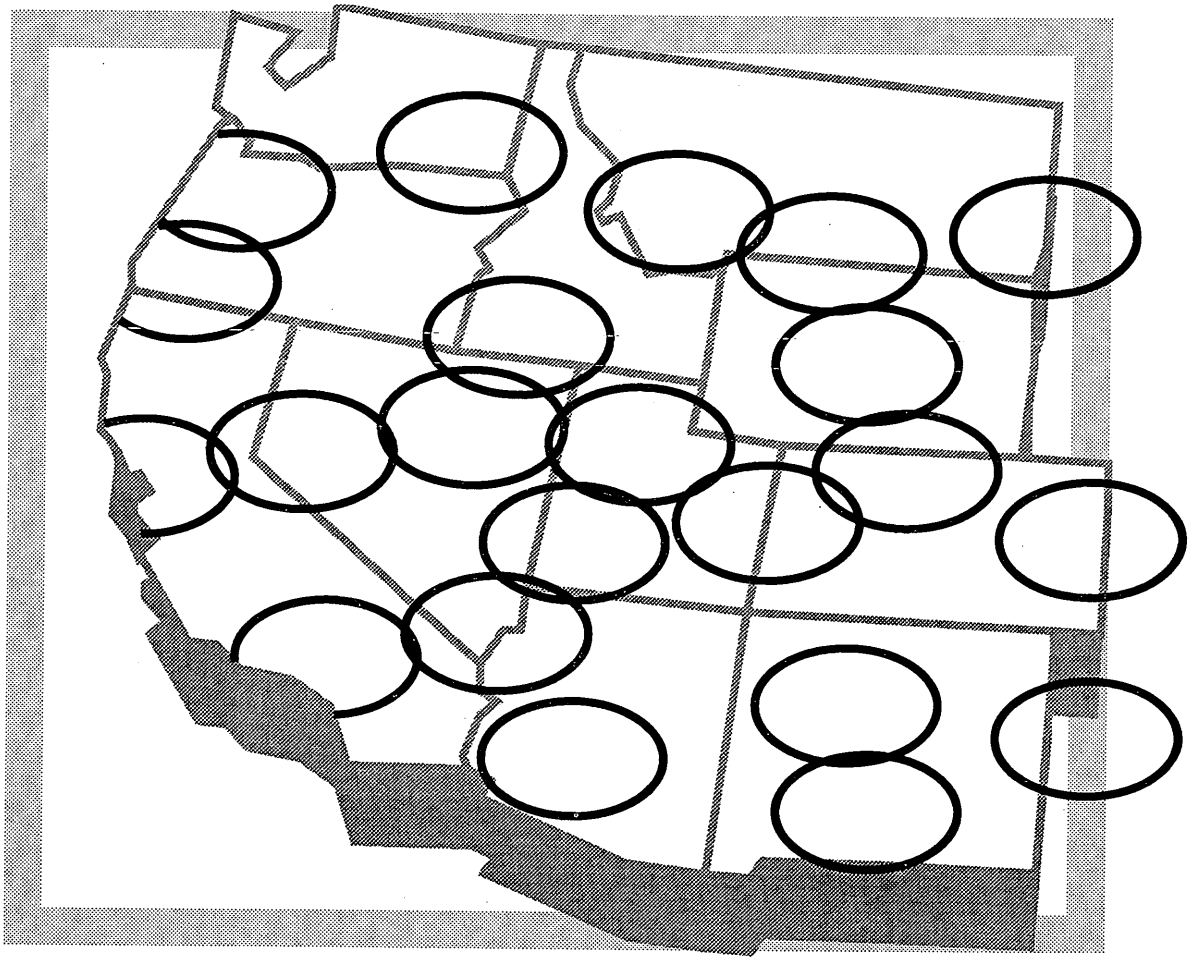

FIGURE 5 Possible western state coverage.

1.7 to 1.8 MHz could be used, but this segment is still being used by some wireless telephones, and the transmitting antennas are inefficient at these frequencies unless they are very long. Moreover, the powerful signals from adjacent AM broadcast stations mix and cause interference in some parts of the segment from 1.7 to 2.0 MHz.

Mountains (or even large foothills) create special problems for broadcasting travel information that long-wavelength carriers such as HF and AM radio address especially well. Unlike the line-of-sight propagation of VHF and FM transmissions, the ground waves from the lower HF and AM frequencies tend to wrap around large objects (such as mountains) and provide good coverage even in deep canyons and valleys where no FM coverage is available. At night, signals also arrive from an ionized part of the atmosphere that is more than 100 miles over our heads, allowing receivers to work anywhere except inside a metal building or a tunnel.

Why not in the AM band itself using subcarriers? AM subcarrier broadcasting may offer the best competition to dedicated digital frequencies at 2.0 MHz. The use of commercial AM broadcast stations from .550 to 1.700 MHz offers many of the same advantages, including wide-area coverage, as dedicated frequencies at 2.0 MHz, and it would certainly take advantage of the large number of available high-power AM stations that can be heard across the country. However, there are three apparent drawbacks: reduced bit rates, technical complexity, and tariffs.

It is uncertain what bit rates can be achieved with AM subcarriers in a field environment. Bit error rate testing scheduled for 1994 and 1995 should define the usable rate better, but expectations are that it will be very low—on the order of 100 bits per second or even less. (ENTERPRISE, a cooperative IVHS research, development, and implementation group among 11 state, provincial, and federal agencies, is planning field testing of AM subcarrier ATIS capabilities in 1994 and 1995.) With the overhead needed for error correction, throughputs may be as low as 25 bits per second. This compares to the 1200 to 4800 bits per second that may be available with a dedicated channel at 2.0 MHz. The low bit rates mean that compact codes must be used from the start with no possibility of transmitting plain ASCII text.

AM subcarriers are also technically complex. Unlike FM subcarrier encoders, which are simple to install and require no license, to install an AM subcarrier a technician must modify the station's carefully controlled master oscillator when it is off the air. RBDS FM subcarrier use is expanding in North America, and RBDS can be used for more than traveler information. One ENTERPRISE state, Minnesota, is currently testing ATIS broadcasting over an RBDS FM subcarrier in Minneapolis. Other organizations are plan-

ning to use FM subcarrier systems to distribute GPS differential correction codes. The Federal Communications Commission currently does not allow the modifications needed for an AM subcarrier, and special permits will be required for trials. The optimum encoding technique has still not been determined, but some of those that are being proposed would preclude the use of stereo systems that are now approved for AM stations. Fortunately, the technical complexity does not appear to raise the cost of the receiver significantly. The parts needed to modify a conventional AM receiver are not expensive.

Unfortunately, this is only true if used in production. For field trials and operational tests, test receivers are still large and expensive. Moreover, the investment needed to reduce AM subcarrier receivers to a single $20 integrated circuit is very large. This is not a problem if manufacturers decide there is a market for these receivers, but it can prevent low-level implementation if there is not. For FM subcarriers such as the North American RBDS and the European RDS standards, this investment has already been made and receivers are readily and inexpensively available.

Station owners need to make money to cover their operating expenses. Subcarrier capabilities, if used to any great extent, will be thought of as real estate to be rented out to users. Although these costs should be minimal to begin with, if competing uses develop for these capabilities, the costs may rise.

ADVANTAGES

Let us quickly summarize the advantages of a dedicated digital broadcast system at 2.0 MHz:

- Rapid implementation;
- Inexpensive transmission;
- Inexpensive reception;
- Use of readily available equipment;
- Easy prototyping for field trials;
- Easy management by modest state DOT or state police traffic management centers (Note: see Figure 6 for conceptual layouts of how a modest traffic information center could broadcast information on a variety of carriers using a single interface);
- Easy commercialization that may be a good introduction to public-private partnership for state DOTs;
- An open architecture that provides maximum potential for competition and service expansion;
- "Recycling" of old frequencies and no competition with the needs of existing services;
- Effective use of proven forward error correcting (SITOR/AMTOR) or packet digital technologies; and
- Useful information service during initial operational tests.

NEGATIVES

HF is sometimes dismissed as a potential IVHS ATIS carrier for various reasons. "It is unreliable, subject to fading and requires frequent frequency changes." Or HF "requires a number of frequency allocations to adapt to the rapidly changing propagation conditions. . . . network control and management [is] very difficult. . . . HF is also very susceptible to both atmospheric noise due to lightning and to artificial noise" (*1*). Because of the need for fairly large antennas when transmitting, HF is awkward to use for two-way mobile transmission. All of these statements are true for some HF systems, some of the time, for some communication needs, and if Texas had a need to broadcast traveler information to Brazil, it would find HF to be a very limiting media. However, none of these

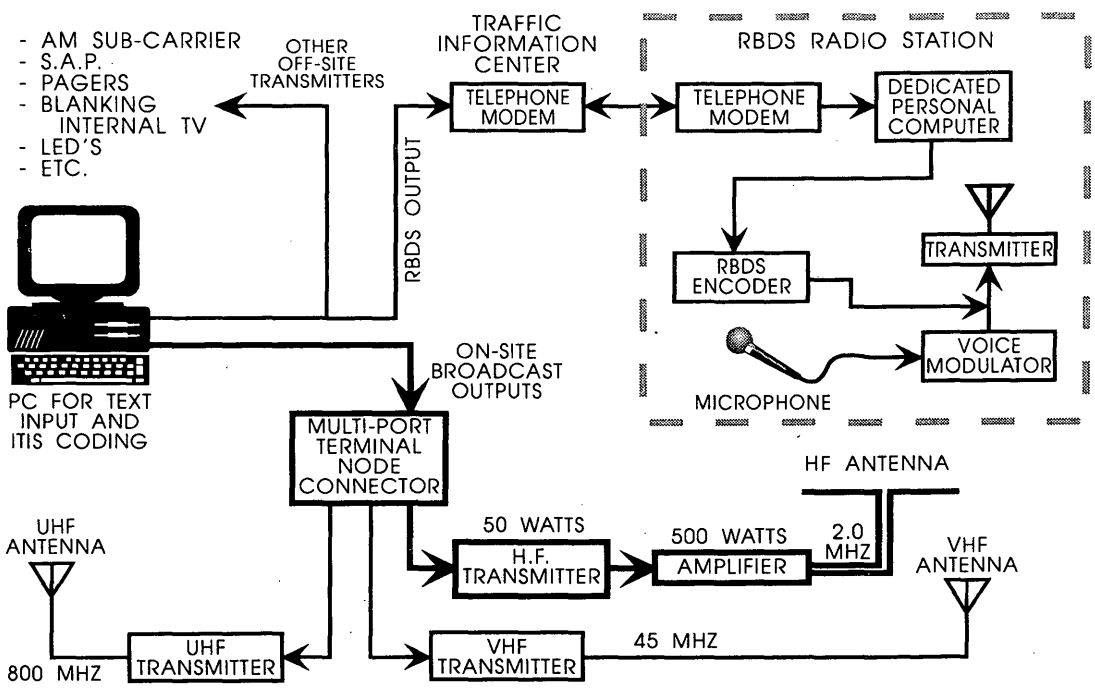

FIGURE 6 Conceptual traffic information center R.F. outputs.

concerns are valid for the types of short-range digital coverage described in this concept paper.

The remaining users in the lower part of the HF spectrum have shown that advances in digital radio broadcasting deal very well with these conditions. That is not to say that there are no unresolved issues: this writer is not aware of any situation in which a number of stations have occupied only a handful of frequencies. The potential for adjacent station interference is still largely unknown. HF antenna configurations are usually optimized for maximum propagation using a low angle of radiation. For ATIS broadcasting, where a shorter range and reduced interference is desired, antennas can be configured to radiate with a high angle of radiation. Not only would this reduce adjacent station interference, but because of the minimal support structures needed, antennas would also be much cheaper to install. Figure 4 shows two possible antenna configurations that provide a high-angle of radiation.

The other major unknowns are those listed in the proposed scopes of work: optimum antenna configuration, data rates, data formats, power levels, and the like.

FIRST STEPS

To reach implementation, the initial impetus must be provided by the federal and state DOTs. Public incubation is needed, because the private sector cannot be expected to front-end the investment needed for an open-architecture system that would provide no particular vendor any advantage over others. Beyond the early stages, transmitting facilities can be supported jointly by regional transportation agencies and the private-sector companies that operate in partnership with them. Conceivably, nonprofit corporations (set up independently, by state DOTs, or by university foundations) could assist in operating multistate systems.

Theories and concepts are just that. The first step to implementation must be proof of concept testing with trials in states with a variety of geographic features (mountains, plains, and Eastern U.S. intercity routes). If the trials do not work, the effort should be abandoned before performing follow-on operational tests, which would actually collect and disseminate useful information. Both concept testing and operational tests should be funded by the U.S. DOT and state and provincial transportation agencies as research and development projects.

WHAT'S NEEDED NOW

State IVHS researchers tend to think in terms of contracts, cooperative agreements, and scopes of work as means of getting programs started. To proceed further, a scope of work is needed to tell those who will potentially be providing the financial support what they would be getting for their investment. Here is a brief outline that is intended to be expanded into a full scope of work for the first two phases:

Phase I—Technical Proof of Concept

- Determine modulation methods and forward error correction methods to be tested;
- Determine hardware needs, location, and antenna configurations to be tested;
- Determine trial frequencies and obtain FCC test approval;
- Field test data rates and bit error rates, including the following:
  - –Coverage zones (day and night),
  - –Year-round testing,
  - –Radiation pattern testing,
  - –Interference testing from adjacent stations, and
  - –Power level requirements;
- Decide optimum configuration for operational tests;
- Make "go/no-go" recommendation for Phase II; and
- Complete detailed scope of work for Phase II.

Phase II—Full Operational Tests

- Make a full operational test description:
  - –Develop a standard architecture and data format;
  - –Use several hundred receivers;
  - –Create a prototype of encoding and decoding software;
  - –Distribute to a wide variety of users;
- Identify and develop potential private sector partners, including the following:
  - –Equipment manufacturers,
  - –Software developers, and
  - –Commercial traffic information providers;
- Complete overall system architecture;
- Establish initial institutional arrangements for a commercial service; and
- Begin informational presentations and outreach program to potential partners across North America.

INTERESTED?

Despite the potential benefits, HF ATIS cannot be implemented by a single state DOT. It must have the interest and support of many agencies responsible for traveler information and public safety. It must also have support from the private sector companies who are best able to find sponsors, set up private investment and manage user fees. Those who would want to support an HF ATIS service and would benefit from its establishment are scattered across North America. At least initially, the Colorado Department of Transportation's IVHS/New Technologies Group is willing to act as a focal point for those interested in this concept. No commitment is required. The Colorado DOT is just trying to put together a working list of those who are interested in the concept and would like to be kept informed about developments.

Contact
John P. Kiljan
Colorado Department of Transportation
4201 E. Arkansas Avenue
Denver, Colorado 80222
U.S.A.
303-757-9508/Voice
303-757-9974/Fax

REFERENCE

1. Weitzen, J. A., J. D. Larsen, and R. S. Mawrey. The Use of Meteor Scatter and Extended Range VHF for Vehicle Tracking and Communications. Presented at the Annual Meeting of IVHS America, Washington, D.C., 1993.

Publication of this paper sponsored by Committee on Communications.