Creating a Municipal Geographic Information System for Transportation: Case Study of Newton, Massachusetts

MARC BAILEY AND SIMON LEWIS

The potential of geographic information systems (GISs) as depositories of urban data has been made clear by a number of authors. As a next step, further consideration should be given to developing the best methods for accessing and manipulating such data for transportation applications. Building simple transportation tools directly within a general GIS package may be an appropriate avenue for smaller agencies without sufficient staff and resources to support multipackage solutions. A case study in the city of Newton, Massachusetts, demonstrated the modification of a GIS to provide an additional range of transportation functionality, including traffic assignment, vehicle routing, reapportionment of traffic zones, and location of centers (such as fire stations) on a network. An important aspect of this case study is that only readily available sources of data were used. Such an approach may be particularly appropriate for simplified "sketch planning" purposes at the local level.

A broad range of transportation modeling software is available to transportation planners at the local level. Relative to the graphics display standards established by new geographic information system (GIS) software, most of the specialized transportation packages appear inherently limited in their general graphical editing and data query capabilities.

Conversely, most GIS packages historically have offered only minimal utilities for transportation analysis. These have included tools for geocoding and rudimentary shortest-path and network allocation capabilities. However, much of the data frequently integrated into a GIS are of considerable use for transportation planning, modeling, and highway inventory. For example, zonal data stored in a GIS may be used for trip generation and distribution; street data may be used to create link files for a transportation model; and highway inventory records may be used to assist street maintenance and emergency response evaluations. Unfortunately, although full-featured GIS software is becoming increasingly accessible, even to municipal-level transportation professionals with limited budgets, most of the available GIS software does not offer specialized transportation tools.

To date, software development in GIS for transportation (GIS-T) has focused heavily on linking existing transportation packages to GIS. Another approach has been to build GIS tools with a strong transportation focus. TransCAD, produced by the Caliper Corporation, is an example of a package developed using such an approach (*1*).

Existing transportation models may access GIS data bases through either "cold links" (data transfer through ASCII files) or "hot links" (interactive use of a common data base). Also, a number of intermediate alternatives are possible—that is, the data links between packages may be more or less automated. Although the potential benefits of such tools are often referred to, a number of problems with linking alternatives have not yet been thoroughly addressed: for example, one street network link in the GIS data base may not translate into one link in the transportation data base.

At the municipal level, a key problem with this linking approach is that the average city traffic engineer or planner must learn a major GIS system as well as a major transportation package and the linking software. The user must learn not only two command sets, each of which may consist of hundreds of directives, but also different command sequences, interface standards, and overall design nuances. This situation may be feasible in a large transportation agency with dedicated staff and resources. However, in the context of technical staff with time and resource constraints, it is important that software tools be as integrated as possible in form and style, even at the possible cost of some marginal functionality.

There are significant advantages to building transportation analysis tools within a GIS. For instance, GIS packages have sophisticated built-in features for display and mapping to facilitate the visualization of model inputs—a loaded transportation network, for example—and the direct comparison of analysis results. Integration of analysis tools within a GIS permits a more direct conversion of geographic data to the specific input data required by analysis models. Such integration is particularly advantageous if the GIS used is what is commonly referred to as a "full-featured" GIS, offering a complete range of facilities for cartographic input, query, and analysis. In addition, in various regions and states, certain GIS products have been adopted as standards for use; if these products can be used, associated benefits are greater still.

If transportation analysis tools are provided within a GIS environment, they may be used in a similar style and form to the more standard GIS tools. The syntax of the commands also may be similar, so that the user may perform all required tasks within a single, familiar environment. Incorporation of transportation functions within GIS may be critical if more sophisticated transportation applications are to be learned and used effectively, particularly at the local level.

The institutional advantages of building transportation applications tools within a comprehensive municipal GIS framework are also clear. Sharing of coverages (digital maps with attribute data) among and across agencies within a common

GIS/Trans, Ltd., 675 Massachusetts Avenue, Cambridge, Mass. 02139.

software environment potentially removes redundancies in data collection and maintenance efforts, while strengthening data quality and analytical consistency. Transportation planning may be significantly improved by direct access to land use and population information from other departments, such as assessing. Other departments, such as planning, might benefit greatly from direct access to transportation department information, such as network modeling results.

The work summarized in this paper concerned the development of a GIS-T framework for transportation analysis. Project technical papers provide more detailed review of this effort, as well as references to the many works in the field on network data and analysis applications. It is not the intent of this paper to describe the particular algorithms and models used in detail or the intricacies of the particular GIS—in principle, it would have been possible to complete the work in several alternative fully functional GIS packages.

BACKGROUND

In Massachusetts, the state agencies under the Executive Office of Environmental Affairs provide GIS maps, digital data, and analyses to all levels of government and the private sector. This program, called MassGIS, is extremely successful and recently won several prestigious national awards.

Within this context of intensive state GIS activity, GIS/Trans, Ltd., is working with the city of Newton, Massachusetts, to develop a case study for the integration of transportation applications within a municipal GIS. Some of the lessons learned in Newton may be largely transferable to other cities and towns interested in investigating their capacity for transportation GIS.

Newton is a suburb west of Boston, straddling the metropolitan area's main east-west road and rail routes. The city's 84,000 inhabitants are concentrated around a number of village centers, where congestion and other pressures on traffic management are increasing. At the same time, resources to address these problems are severely restricted.

The city has attempted to be forward thinking in evaluating how to better manage and analyze its planning and engineering data and is currently developing a citywide GIS. This initiative began in part through the running of a semester-long workshop in conjunction with the computer resource laboratory of the Massachusetts Institute of Technology (MIT). In this workshop, MIT students surveyed the different city departments and completed a preliminary information systems analysis. From this work, a view of the potential for GIS was identified and an implementation plan was drafted. GIS hardware and software were acquired, and a parcel-level data base is currently being digitized.

The work discussed here originated in a number of research initiatives at MIT. In further pursuing these directions, GIS/Trans, Ltd., has assisted the city traffic engineer over the past 3 years in preparing to establish various transportation GIS applications for the city. The goal is to improve traffic data management and to allow the execution of transportation analyses that previously were very difficult, if not impossible. A fundamental benefit of the work is the provision of a data set for broad consideration and GIS analysis.

The work has also been completed as a case study representing the range of applications for which GIS-T might be used in a municipal setting. It is used in a 3-day National Highway Institute GIS-T course that will be presented at each state transportation agency over the next 3 years (*2*). The work to date includes activities in three main areas.

Matching TIGER files with Roadway Inventory Files

Some innovative programming was required to match U.S. Census Bureau TIGER records with Massachusetts Department of Public Works (MDPW) road inventory data. The resulting data set provides both a geographic network and street attributes for Newton.

Implementing Link to Transportation Planning Package

The data from the file matching were first entered into a transportation planning package, QRS II. This package was chosen because of its low cost and ease of use. Links to other packages such as TRANPLAN also have been developed and demonstrated. An origin-destination (O-D) matrix was generated by "cordon compression" from the Massachusetts Central Transportation Planning Staff (CTPS) statewide data base, using tools within the Urban Transportation Planning System (UTPS) suite of programs.

Development of Prototype Applications for Routing and Traffic Assignment

Newton was used to test the validity of incorporating transportation tools within a GIS environment (as opposed to merely linking GIS to transportation packages such as QRS II or TRANPLAN). The example applications built within a GIS environment include traffic assignment, routing applications (such as school bus route generation), network location problems (such as fire station location), and traffic zone reapportionment.

This approach differs from the cold and hot links described earlier, as the analytical tools reside as modules within the GIS. Essentially, the analytical tools are composed of the core algorithms and make direct use of the various facilities and programming services provided within the GIS. Thus, although the algorithms are functionally complete if provided with input in appropriate formats, they depend on the GIS's lower-level routines for file input and output, screen display, and other operating system tasks. By improving the various file input and output hooks developed below, other users would be encouraged to develop and incorporate algorithms to handle various analytical tasks of their own.

OPPORTUNITIES FOR LINKS BETWEEN GIS AND APPLICATION TOOLS

There are a number of alternatives for linking GIS and transportation analysis tools. These tools could range from the

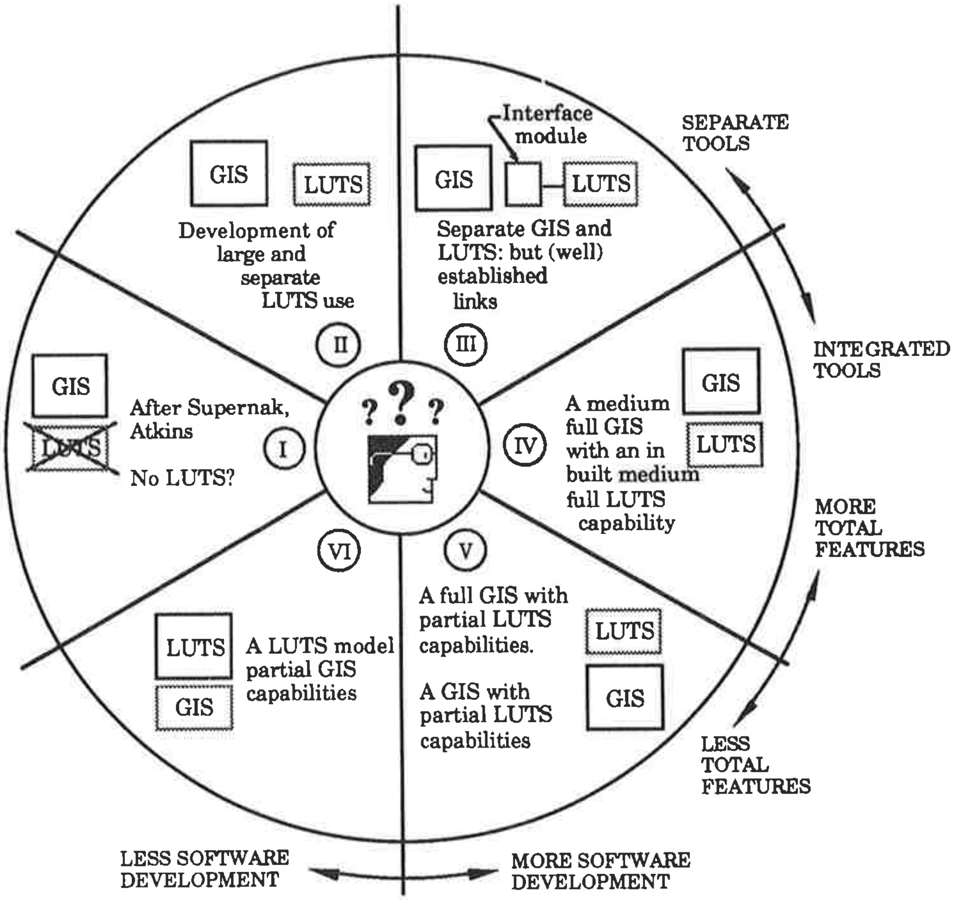

FIGURE 1 Links between GIS and transportation application tools.

traditional transportation models through to routing packages for school buses or snow removal. These alternatives have been described in detail elsewhere (*2–4*), so a full description is not given here.

In brief, Figure 1 summarizes some of these opportunities. The alternatives indicate various trade-offs system designers have made between the integration and separation of various tools. They also show whether the delivered systems have more or less utility and whether there is a greater or lesser degree of software development.

Models I through III assume that GIS toolboxes and transportation planning toolboxes are essentially separated. Model I asserts that there is a decline in use of the classic UTPS family of tools, given the criticisms directed at these models in various "before and after" studies, whereas the use of GIS continues to expand. Model II assumes the separate growth of GIS and transportation modeling tools. This assumption was the case until fairly recently; the GIS and transportation communities maintained their distance with separate professional organizations, conferences, and other activities. However, this situation has changed over the last 2 years. Model III allows for the provision of links between GIS and transportation modeling tools by the use of interface modules. These interface modules may use either hot or cold links for the exchange of data between packages.

Models IV through VI assume a closer linking of GIS and transportation planning code. Model IV is best described as the addition of graphical capabilities to many transportation packages. These additional graphical capabilities have become increasingly more sophisticated, allowing, for example, the maintenance of topological connectivity. Model V is the state of many existing fully functional implementations of GIS. Finally, Model VI represents the "transportation analysis toolbox" approach, in which many tools are disaggregated into component parts within a common GIS platform. The effort summarized in this paper was essentially concerned with an initial review of Model III, which is followed by a more extensive review of Model VI.

BUILDING THE DATA BASE

The first and most essential concern in developing a GIS-T is to generate or secure adequate data. To accommodate the Newton GIS-T model, two distinct and previously independent data bases were joined. The first of these is the TIGER/Line, developed and maintained by the Bureau of the Census as a cartographic information base to assist with the execution of the 1990 national census. TIGER is crucial to the Newton GIS-T because it provides spatial geometry for the various intended analytical tasks and a structure to which data base attributes can be linked. Additionally, TIGER provides certain roadway characteristic data, such as street addresses and intersection longitudes and latitudes. (A number of vendors provide "cleaned" versions of TIGER data, with missing street names and addresses inserted, positional accuracy checked, and so forth.)

The other key data file for the Newton GIS-T is the MDPW Roadway Inventory, a computerized record of geometric and structural data for approximately 33,800 mi of roadway in Massachusetts. Most of the data describe the dimensions of each street, including lane, shoulder, and sidewalk widths, with additional information on surface types and conditions, access control, federal aid, and categorical estimates of traffic flows. The inventory was developed to facilitate effective transportation planning and bookkeeping and to bring into accord the needs and resources of local, state, and federal agencies.

The greatest challenge was the actual matching of TIGER and MDPW records. Having evaluated several approaches to coupling these files, it was ultimately decided to create a single file containing, for each roadway segment defined in the original TIGER file, all the related information from both the original TIGER and MDPW files. These efforts generated a GIS-T data base structure flexible enough to be used in many applications.

In the integration of TIGER and MDPW files, one-to-one mappings based on street name are rare. TIGER street records are segmented according to intersections, whereas MDPW data for each street are linearly referenced and independent of other features. Furthermore, the TIGER and MDPW files have many missing or conflicting address ranges, rendering these virtually useless for matching.

For each street segment, the MDPW inventory provides an odometer distance measured to the hundredth of a mile. Although TIGER does not include a comparable data field, makeshift lengths were generated by using basic geometry from TIGER street segment endpoint longitudes and latitudes. Programs were developed to tie individual MDPW records to the corresponding groups of TIGER records on the basis of street names and these newly calculated link lengths.

The matching program indexed each record in TIGER by adding a field to the data base containing the number of the corresponding record in MDPW. In a first pass, about 86 percent of the records in TIGER were identified for updating with Roadway Inventory data. There are several explanations for the missing links. First, the MDPW inventory does not include entries for private roads, whereas TIGER does. Next, both TIGER and MDPW contain many "unnamed road" entries. Some subset of these records probably describes the same roads, but they cannot be matched using merely segment length (independent of street name). Finally, there are spelling discrepancies between the two files (e.g., Lindberg Ave. versus Lindbergh Ave.). A Soundex function, as well as some manual processing, was used to deal with many of these discrepancies. After the problems that could be identified were addressed, the indexing program was reexecuted and produced a 94 percent match. A separate paper will address the wider issues of coupling networks from different sources and the approach taken in detail in Newton.

The successful creation of an integrated data base for Newton provides some indication of the potential value of such efforts to the modeling field. Most important, it demonstrates that a process for merging such disparate sources as TIGER and state road inventory data may be developed inexpensively. By using the same steps as those taken for Newton, one could establish a similar data base for any city or town in Massachusetts. Furthermore, roadway data foundations analogous to the MDPW inventory might be used to produce similar GIS-T base data for municipalities outside Massachusetts.

FIGURE 2 One-way streets in Newton.

With access to road attribute data in a GIS, a city traffic engineer can respond quickly and efficiently to queries about current road and traffic conditions of the city. For example, Figure 2 shows a map of the city of Newton with one-way links distinguished graphically. Many other graphical displays of road information can be generated, including dimensions, structural condition, and state aid status.

Considerable work is required to maintain any road data base. Geometric and attribute information change constantly. A GIS is currently most useful for graphically tracking the status of attribute data and verifying new attribute data when it is merged into the data base; GISs are less useful for merging geometric data.

APPLICATIONS

Traffic Assignment

A valuable tool in municipal-level transportation planning and management is traffic assignment. A traffic assignment may be used to predict the effect on traffic flows of changes in demand for travel or in the structure of the transportation network. The goal of the tools provided in this study was to meet the traffic engineer's needs for quick first-order assessments, or "sketch planning." For example, what would be the first-order effect on traffic flows across the Newton network of a bridge collapse that made a certain link unpassable? The tool is especially valuable for such "quick-response" situations when there is insufficient time to use a more complex analytical package.

Another potential use for the tool is concurrency management. For example, what additional traffic strain would a new shopping center in Analysis Zone *x* place on the network?

The basic types of input information required for traffic assignment are

- An analysis network, representing major highways and streets, which generally must be abstracted or reduced from a GIS coverage including all streets in the study area. Vehicle flow capacities and traffic volume-delay relationships for each road represented in the analysis network must also be specified.
- A set of internal and external traffic analysis zones (TAZs), defined for the city and its immediate surroundings. The volume of traffic generated is typically determined using the socioeconomic characteristics of these zones.
- An O-D trip table, which describes the demand for travel between all pairs of TAZs in the study area. Such a table is usually developed from a combination of past travel survey data and traffic flow observations. (Tools were also developed to directly generate these from traffic count data in the GIS, although these are not discussed in this paper.)

A city-level transportation analysis network and an O-D trip table were extracted from a UTPS-based regional transportation model for metropolitan Boston operated and maintained by the Central Transportation Planning Staff, an adjunct agency of Boston's MPO. The cordon compression technique was used to build the trip table, summarizing trips entering and leaving the city and allocating them to the surrounding ring of external zones. This technique—when combined with the use of GIS technology—can make transportation modeling cost effective and useful for many local agencies that previously did not have the resources to engage in a longer-term modeling effort.

With the data available in appropriate format, four basic steps are executed using the traffic assignment module developed within the GIS:

1. Read in the link data representing the analysis network, converting each two-way street into a pair of one-way analysis links. The travel impedance and capacity for each link are defined based on the coverage link attribute table item values.
2. Read in the O-D trip table data.
3. Perform the traffic assignment. Separate options are provided for "all-or-nothing," user equilibrium, and incremental assignment.
4. Write out the assignment results—final traffic flows and costs in each direction—as new items on the link attribute table.

Programs were developed to produce a graphical display of results for a single assignment (see Figure 3) or a comparison of traffic volumes for two scenarios (see Figure 4). These procedures define a set of line symbols of varying widths and a look-up table to associate each level of traffic flow with a correspondingly wide symbol. When two scenarios are com-

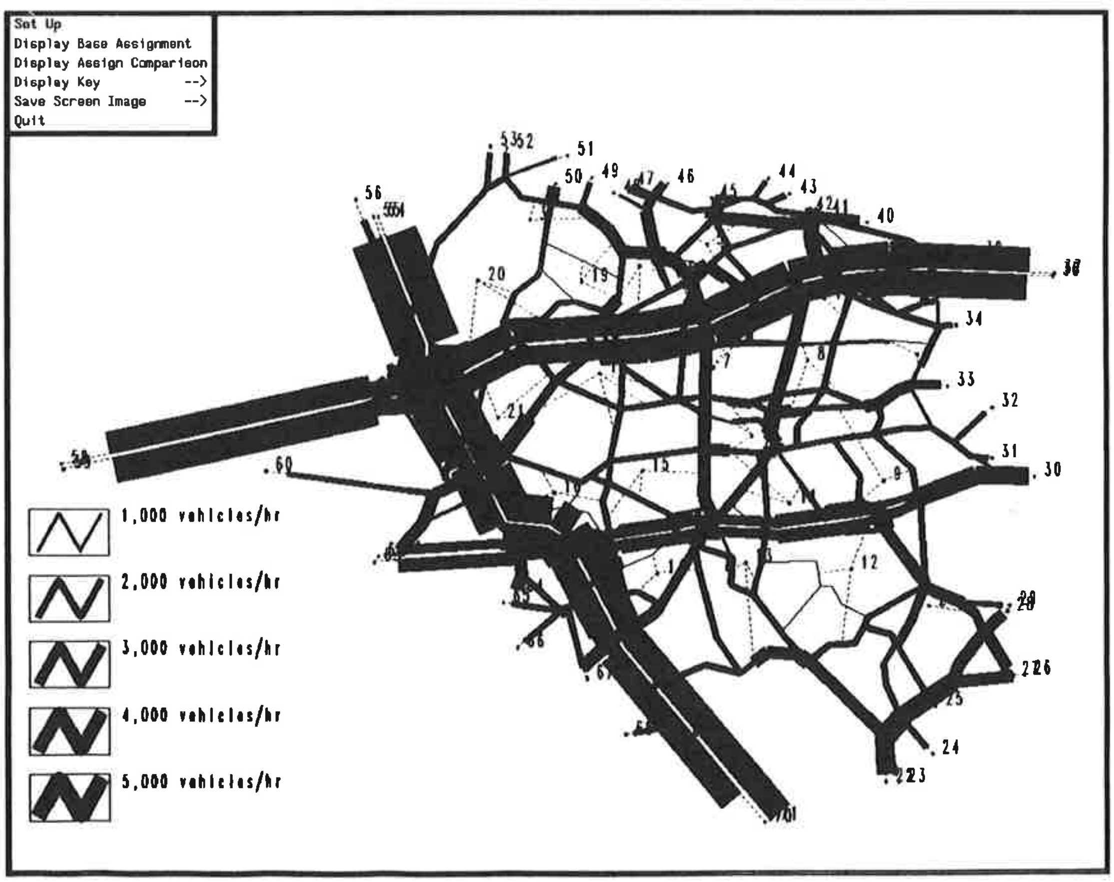

FIGURE 3 Traffic assignment: results for base assignment.

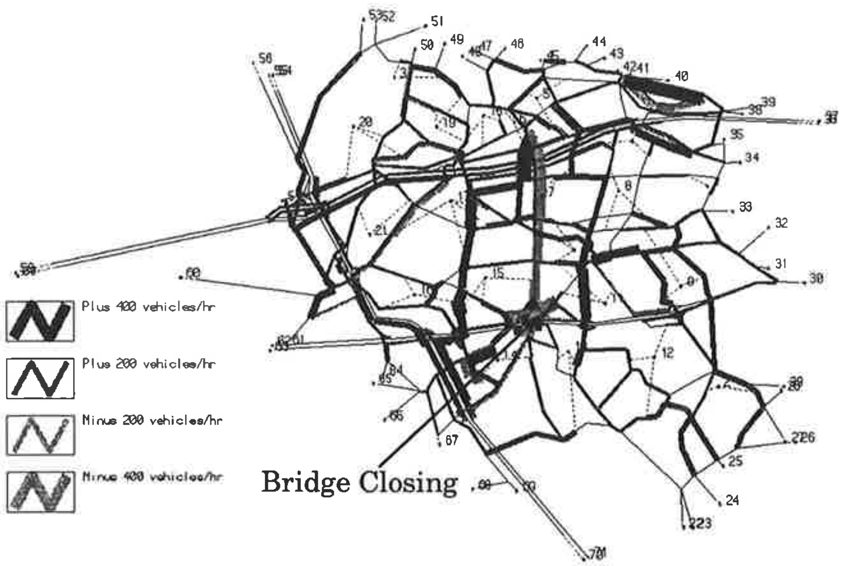

FIGURE 4 Comparison of two assignments—bridge closing.

pared, the display employs red and green line symbols to denote net increases and decreases in volume, respectively.

The ability to modify the transportation network (or trip table) and present modeling results in a GIS environment gives transportation planners the power to generate and thoroughly evaluate many "what if" scenarios within a limited time and resource budget.

Routing Applications

Many city services involve employees or work crews visiting customers, fixed facilities, or other landmarks at specific points. Some component of the total service cost relates to the time spent traveling between successive customers or work sites. Thus, by sending servers along the shortest route to all work sites, the city may realize a reduction in service costs. This is the basis of the well-known traveling salesman problem (TSP). Relatively simple examples might include a single person or crew performing mail collection or servicing street lights.

A related problem considers servicing sets of links rather than points. Typical applications include street sweeping or garbage collection. A current practical application in Newton is to calculate an optimal order in which to visit all streets in the city, so street lights may be inventoried and incorporated into the GIS.

More complicated vehicle routing problems—such as school bus transportation or dial-a-ride service—involve multiple service vehicles, time constraints on pickups and drop-offs, real-time response, or all of the above. The key to solving most, if not all, of these routing problems is to determine the solution to one or more TSPs.

Even for simple versions of the TSP, the difficulty in finding an optimal ordering of the stops is magnified enormously as the number of stops increases. Thus, for problems as small as a few hundred stops, a heuristic procedure is usually applied to find a good solution within a reasonable period of time rather than exhaustively determining the optimal shortest tour.

The GIS module to determine good solutions for TSPs requires the following input information:

- A link coverage defining the transportation network. The link attribute table must include attributes (e.g., distance, travel time) that will be used to specify travel cost.
- A set of points within the network to be visited.

To solve a TSP using this module, the following basic steps are performed:

1. Read in the link coverage, converting each street segment to a pair of one-way links.
2. Read in the list of points to be visited.
3. Build a table of travel costs between each possible pair of visit points. This is done with a shortest-path algorithm using any specified link attribute as the measure of link impedance.
4. Execute an algorithm to find a good ordering of the points in the tour.
5. Write the final order of stops to a data file.

The output results of the TSP module can then be read with a macro to produce a graphical display of the solution route. Figure 5 displays a solution tour to 18 school locations in Newton.

School Bus Applications

Routing school buses in the city of Newton offers an interesting development of base tools to demonstrate some GIS functionality in solving a real transportation problem. A macro-based menu system was developed for student clustering and bus routing. The menu is invoked from within the module used for displaying final bus routes.

Input information was provided by the Newton Schools Department, which maintains a computerized student data

FIGURE 5 TSP: optimized visit of Newton elementary schools.

base for school assignment and bus transportation planning.

The steps toward solving a school bus problem are to

1. Read the street network.
2. Read the school location and bus stop locations (stored as separate point coverages) into the module.
3. Cluster the student bus stops by invoking a specially written C-language program from the menu. The user is prompted to enter the maximum number of students that may be picked up per bus trip.
4. For each cluster of stops, determine the best route by calling the TSP module.
5. Display each route using a separate line color.

Figure 6 shows the final results of a clustering and routing analysis applied to students of the F.A. Day Junior High School. Each bus route visits up to 20 bus stops before returning to the school.

Traffic Zone Reapportionment Application

In the transportation community, reapportionment is often used almost interchangeably with redistricting. Reapportionment refers to the process of defining and redefining TAZs for transportation analysis. TAZs are areas from which traffic is generated. Various information is used to build a regional O-D matrix, which in turn is used as input data to traffic assignment or routing algorithms.

There are two main criteria for the creation of new TAZs. First, equality of population—the population of each district should be distributed as uniformly as possible. Major deviation from the target population can be justified only by extraordinary local circumstances, such as data problems, racial distribution, county lines, and shapes. The second criterion is compactness, or contiguity of districts.

In the past, redistricting has simply been performed manually. More recently, computerized automation of the redistricting process has been explored. User-specified criteria and conditions are input that allow the program to make judgments and decisions based on the statistics it computes at every stage of the process.

The reapportioning process used in this study has the following characteristics:

1. It is a heuristic approach: it presents one of the many possible solutions on a redistricting problem. Other solutions may be equally practicable, but this system produces a result that is as close to user specifications as possible.
2. It combines the power of GIS and the capabilities of the C programming language in the UNIX environment. The GIS provides access to raw data from census files, a user interface, and graphic display of the area to be redistricted, whereas the C-module performs the actual redistricting.
3. Currently it uses only the two major criteria of redistricting mentioned previously.

The GIS generates three tables: a population table, a correspondence table, and a small adjacency table. The population table contains information on the population at the block level; the correspondence table lists the blocks in each district; and the small adjacency table describes the adjacency relations among the blocks. The module is able to deduce the adjacency relations among the large districts from the correspondence table and the small adjacency table.

The main output file generated by the module is an updated correspondence table, with the new correspondence relations between the districts and the blocks. The summary file calculates the mean population difference of the districts from the target population and provides statistics on the extent of improvement in terms of population distribution after the redistricting process.

Perhaps the most important advantage of such a redistricting system is its automated nature. As a result, there is less human work involved, a great reduction in tedious and mechanical work, and fewer errors. Further description of this application is given in Lin (*5*).

Network Allocation

Optimal siting of emergency response facilities, such as fire stations, is a traditional and typical network decision problem. Partly because of the lack of available network data, many existing approaches have used areal allocation measures without reference to the underlying detailed network.

In a base form, the stated problem may be "to determine the number and location of a given number of centers given a maximum response time." Variants of the problem allow for certain centers to be fixed. Typically, this type of location decision problem is a "brute force problem." A number of simplifications may be made so that a computer-tractable algorithm can be constructed, such as reducing the problem so that only network nodes are considered or assuming that all centers have the same capacity.

The GIS was used to complete the spatial ordering of the data so that a specially prepared network allocation algorithm

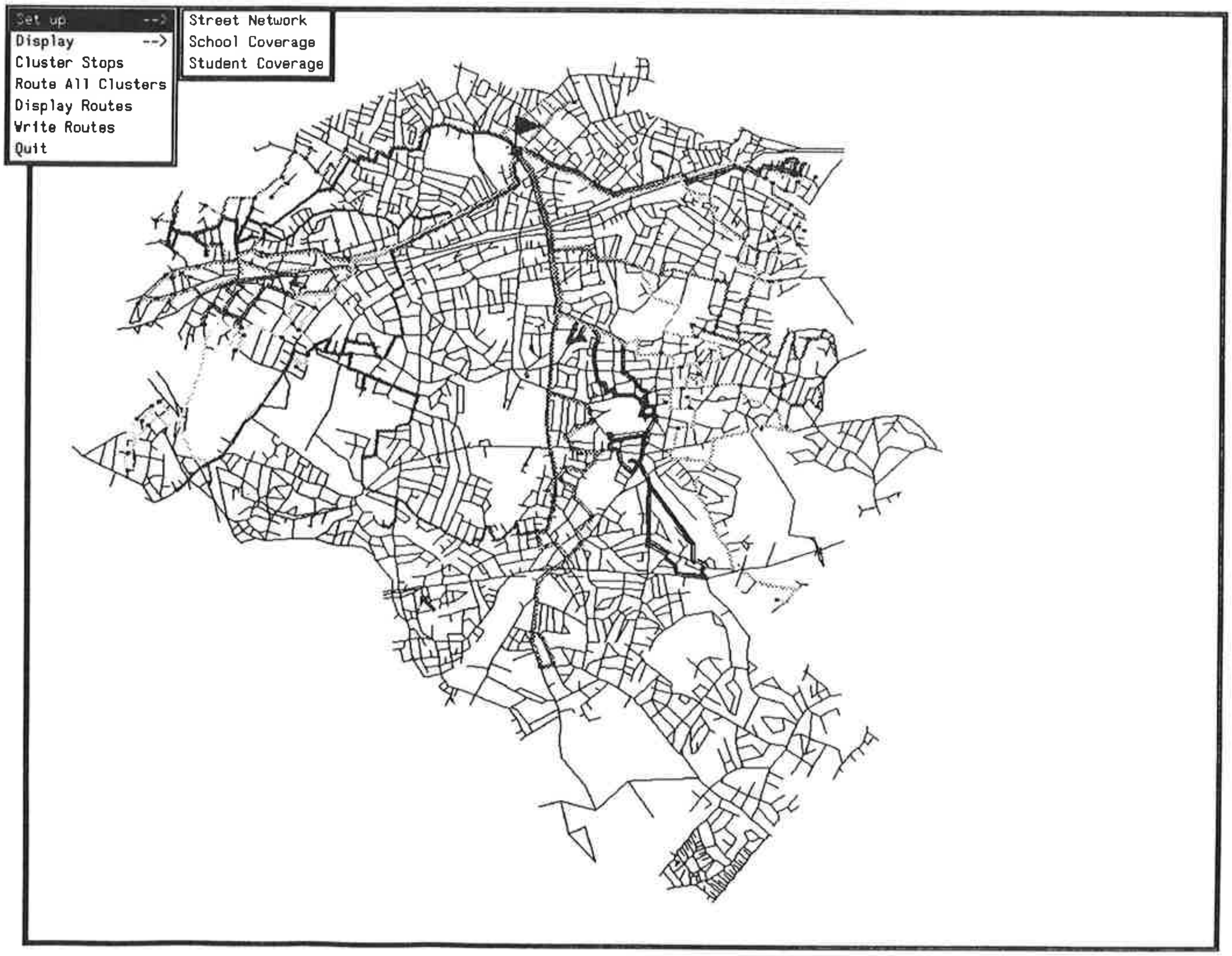

FIGURE 6 Results for school bus routing.

FIGURE 7 Network allocation.

could be run within the GIS. Once the algorithm executed successfully, the GIS was used to organize and display the network data. See Figure 7 for results of network allocation. Further description of this activity is given in Yuan (*6*).

CONCLUSIONS

The assemblage of a GIS-T toolbox approach to planning and analysis has been reviewed; the work described practically outlines the incorporation of one set of transportation applications within GIS. Some focus was made on integrating network data from multiple sources because, although they are technically less challenging to some reviewers, such issues have been major stumbling blocks to GIS-T implementation, both technically and practically. Second, the generation of a useful set of core transportation tools within a GIS is described. The purpose of this paper was to propose and demonstrate the framework, not to describe its instantiation in detail. This is a topic of continuing detailed research.

The simple transportation applications described for the city of Newton are a modest beginning. Greater sophistication will soon be included in the clustering and routing programs in the TSP applications, the interface will be further simplified, and many other applications will be added. The purpose of the initiative thus far has been to explore the general feasibility and value of developing transportation tools (and, in particular, simple tools at the local level) within a GIS package, rather than merely to link them to a GIS. Naturally, there will still be many applications for which more sophisticated and dedicated tools will be required, as was demonstrated in the first part of this work.

The research done here establishes a foundation for future GIS-T work, tailored to the local transportation engineer in particular and to others who may require simple, cost-effective tools for sketch-planning or quick-response applications.

REFERENCES

1. H. J. Simkowitz. Geographic Information Systems: An Important Technology for Transportation Planning and Operations. In *Transportation Research Record 1286*, TRB, National Research Council, Washington, D.C., 1991.
2. GIS/Trans, Ltd. *Introduction to Geographic Information Systems for Transportation. Participant Workbook*, Version 1.05. FHWA, U.S. Department of Transportation, March 1992.
3. S. Lewis. Use of Geographical Information Systems in Transportation Modeling. *ITE Journal*, March 1990.
4. S. Lewis, P. Cook, and M. Minc. Comprehensive Transportation Models: Past, Present, and Future. *Transportation Quarterly*, Vol. 44, No. 2, April 1990; pp. 249–265.
5. C. Lin. Redistricting with GIS: An Automated Approach. Presented at the New England URISA Conference, White River Junction, Vt., March, 1992.
6. W. Yuan. Fire Station Allocation Within a GIS Framework. Presented at the AASHTO GIS-T Symposium, Portland, Oreg., March 1992.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.