Integrating Geographic Information System Technology and Transportation Models

HOWARD J. SIMKOWITZ

The potential of geographic information system (GIS) technology in transportation has been recognized for some time. However, only recently has GIS technology been adapted to meet the requirements of transportation systems. The results presented are based on research being conducted by the Caliper Corporation on integrating GIS technology and transportation models. First, the data requirements of various transportation models are explored. The data requirements include sequential text files for both inputs and outputs, tables containing vectors and matrices, and connected networks of nodes and links with their associated attributes. The data requirements, in turn, make a variety of demands on the structure of the GIS data bases and the way the model interfaces with the GIS. Next, the content and structure of a transportation organization's data bases are discussed and compared with the requirements imposed by the various model types. The traditional GIS formulation is specified, and its limitations for transportation analysis discussed. The concept of a transportation GIS is developed. Research has shown that rather than a mere addition of features to the existing GIS formulation, a transportation GIS requires a hybrid architecture that incorporates important transportation data structures and specialized procedural input, processing, and output modules. This new formulation makes it possible to optimize both the GIS functionality (e.g., thematic mapping, complex spatial manipulations, and rapid spatial and key field queries) as well as the transportation and operations research modeling needs (e.g., integration with compact, highly efficient connected networks and tables containing vectors and two-dimensional matrices to produce and analyze chains, tours, and other model outputs). The introduction of these transportation objects facilitates nearly all transportation applications of GIS and greatly increases ease of use by transportation professionals. The use of modern software concepts for the user interface makes GIS more accessible to nonprogrammers and persons not trained in GIS, and an extensible architecture increases the potential scope and integration of transportation applications. Of critical importance is the introduction of numerous transportation application modules. Other research results being reported include dynamic segmentation of linear feature data bases and the ability to handle extremely large networks and data bases on microcomputers. Operational experience with Caliper Corporation's GIS for transportation, TransCAD, is reviewed. The breadth of applications is startling and illustrates the power of a generic transportation GIS used as a platform for all types of transportation analysis. Experience indicates that important productivity gains for transportation organizations and professionals will result from the adoption of a transportation GIS.

A geographic information system (GIS) is a computerized data base management system for the capture, storage, retrieval, analysis, and display of spatial (i.e., locationally defined) data. The advent of inexpensive, easily accessible GIS technology has placed the entire transportation analysis and modeling process in a new and exciting light. Whereas in the past, transportation data were displayed and analyzed in a tabular format or, at best, by using greatly simplified straight line network abstractions, today it is possible to integrate transportation data bases and models into a GIS system. A transportation GIS requires important extensions of the traditional GIS formulation that was developed for environmental research.

The results presented in this paper are based on research being conducted by the Caliper Corporation on expanding the definition of a GIS to contain necessary transportation extensions. These extensions not only result in a flexible representation of reality but also provide spatial analytical tools that enable the analyst to ask fundamentally new questions about diverse data sets that have important bearing on the solutions to transportation problems. These extensions facilitate the integration of GIS technology and transportation models. Caliper's research has resulted in the development of a transportation GIS, TransCAD. All reported concepts and results have been incorporated into the TransCAD software.

CLASSIFICATION OF TRANSPORTATION MODELS

There are a number of ways of classifying transportation models. For the purpose of this paper, classification will focus on the required data formats for inputs and outputs. Some models require as inputs and outputs one or more attributes from a selected set of records in a GIS data base. Models in this category include the following: accident analysis in which each record is a point on the roadway and the attributes describe the accident's setting and participants; pavement management in which each record is a road segment, the input attributes describe distress, vehicle volumes, and road type, and the output attributes define a condition ranking or a strategy of treatment; and a trip-generation model in which each record is a traffic analysis zone, the input attributes describe the propensity for trip making, and the output attributes describe the estimated trip productions and attractions.

Spatial distribution models require the ability to manipulate vectors and matrices of data. The gravity model, used to estimate trip distribution between a set of origins and destinations, requires vectors containing trip productions and attractions at each traffic analysis zone and matrices that con-

---

Caliper Corporation, 4819 Cumberland Avenue, Chevy Chase, Md. 20815.

tain the minimum cost of traveling between zones as inputs to the model. The model outputs a new matrix of estimated trips between all traffic analysis zone pairs.

The transportation and operations research literature is filled with models that require a connected network formulation to solve the algorithms. These include shortest-path routines (useful for truck permitting, hazardous materials routing, and detouring), traffic assignment models, routing and scheduling models (for creating optimal vehicle tours), maximum flow network models, location and allocation models, and network-based signal timing models. Inputs to these models require a selected set of nodes and links with a subset of attributes to be used to determine the generalized cost of traversing each node and link. Outputs include attributes on nodes (e.g., intersection counts and turning movements) and links (e.g., flows, paths, and tours). Network problems can take an unacceptably long time to solve or be limited to small, unrepresentative abstractions if an efficient network formulation is not used.

Although it is possible to fit a generic model into one of these groups, in reality this distinction can become a bit fuzzy, particularly as a modeling system grows in complexity. For example, while a pavement management system requires information about all road segments and their spatial relationship to each other, it is only of secondary concern how one might traverse the roads to get from Point A to Point B. But what if the traffic engineer determines that a detour around the construction site is required? Determining the path of the detour requires a network formulation.

It is obvious that these various model types make different demands on a GIS data base structural design. The traditional GIS formulation is poorly equipped to deal with these demands.

TRANSPORTATION AGENCY DATA BASES

Transportation agencies collect and maintain large amounts of data. If a generalization can be made, it is that these data bases are large, unwieldy, unrelated to each other, and poorly integrated. A data base may contain hundreds of thousands of records, each with hundreds of fields. For example, a typical state accident reporting form contains more than 100 data items describing the vehicles and individuals involved as well as the accident setting. A sign inventory might include information on sign and pole type, date installed, and current condition.

A road segment data base for pavement management might contain entries detailing the original roadway composition plus all repair and reconstruction activities, 10 or more deficiency ratings, and the annual traffic counts. A roadway inventory, possibly maintained by a different office, could include more detailed traffic counts, lane information, and surface type. The segment endpoints might not be the same as those used for pavement management.

The planning staff may have built data bases for predicting traffic generation and origin-destination flows using socioeconomic data anchored to census blocks, tracts, traffic analysis zones, or economic regions. The data bases might include the multitude of data summarized with each decennial census supplemented by agency surveys. The regional planning network is typically an abstraction of the actual highway and transit network. Thus its segmentation is different from both the pavement management and roadway inventory data bases.

It is not unusual for one agency to use a variety of segmenting schemes and referencing systems (e.g., milepoint, reference post, state plane coordinates, longitude and latitude, and Cartesian coordinates), greatly complicating the task of spatially relating data.

USING GIS TO INTEGRATE DATA WITHIN AND ACROSS AGENCIES

A GIS typically can deal with three types of data bases: points, lines, and polygons. The simplest of these is a point data base that might be used for accidents and signs. When dealing with a large area, a point data base might also be used to represent collapsed polygons (e.g., using the city centroid as the spatial reference in a point data base of cities). A segment data base, used for roads, transit routes, and the like, contains segments, nodes at the ends of each segment, and intermediate shape points that give each segment a realistic profile. Shape points are stored differently from nodes. They have coordinates but no attribute data and, therefore, take up little space in the data base. A polygon data base is used for boundary data such as traffic analysis zones, engineering districts, and state outlines.

GIS can relate previously unrelatable transportation data bases into a comprehensive information management system. This can be accomplished in two ways. When an existing data base is converted to a GIS format, a common locational reference (usually longitude and latitude) is added to each record. Thus, a correspondence between longitude and latitude and state plane coordinates, milepoint, street names, and any other geocoding system is determined by the GIS's data base builder. Procedures have been developed that, in most cases, will automatically compute these transformations during the data base building process.

An alternative would be to calculate the common reference each time a data item was used. One such method is termed dynamic segmentation. Changes in attribute data that have been geocoded by milepoint are used to determine a temporary segment partitioning "on the fly." The idea is that the base network is segmented only when one of the attributes being analyzed changes. If not implemented carefully, such a real-time partitioning could greatly impair the efficiency of the GIS software. Procedures have been developed that extract data from route-milepoint files and create true GIS data bases with the minimum necessary segmentation.

DATA VISUALIZATION, SPATIAL ANALYSIS, AND QUERIES

At its most basic level, a GIS makes it possible to visualize data quickly in many ways. It is possible to zoom in and out on a map display and show the objects in the data bases color-coded by grouping or highlighted by selection. Accident sites on a highway—near intersections, along curves, clustered near schools—can be color-coded by severity. Traffic analysis zones can be color-coded by trip production, income level, trip attraction, or some other parameter of interest. Traffic volumes can be displayed as single bandwidths and traffic assignments as dual bandwidths. Intermediate shape points (coordinate points that give curvature to road segments) make it possible to display routes according to their actual trajectory and not as straight lines connecting the nodes.

A GIS also makes it possible to ask intelligent questions about how the data are spatially related. How many accidents occurred within 2 mi of a university? How many construction dollars are scheduled for each state engineering district? Which signs are within the boundaries of a construction project? How many people live in a 0.5-mi buffer around a proposed hazardous materials truck route? Which transit routes come closest to the largest number of transit dependents? Which are the deficient bridges within a particular congressional district?

A well-designed GIS should store data in a way that allows quick answers to these types of spatial queries. A unique spatial indexing scheme has been developed to do this. The indexing scheme works in world coordinates (longitude and latitude) and is not dependent on dividing the cartographic data base into small mapsheets stored as physically separate files as is done in traditional environmentally oriented GIS programs.

In addition to rapid spatial querying, a GIS requires rapid attribute querying as well. Which road segments are Interstate 95? Which accidents involved a fatality? Which census tracts have a majority of low-income residents? Which road segments are in the worst condition, in descending order of needed repair?

Spatial and key field queries can be combined in interesting and informative ways. Which fatal accidents occurred within 2 mi of a university? Which road segment contains 505 Main Street (and pinpoint the address location on the map)? Which traffic analysis zones within 1 mi of the central business district have a high concentration of housing?

Not only can a GIS integrate various types of data, it can also integrate data collected at various degrees of resolution. Building lots digitized with great precision can be displayed along with census tracts that were digitized at 1 : 100,000 scale. If there are obvious discrepancies between lot boundaries and census tract boundaries (which follow street centerlines), on-screen editing can be used to correct them.

TRANSPORTATION MODEL DATA REQUIREMENTS

Any particular transportation model requires only a small subset of the agency's data. However, these data may come from diverse data sets. They can, for the first time, be easily related using a GIS. For example, a road reconstruction project could benefit greatly from a thorough understanding of accident experience, vehicle volumes, and turning movements. The best route for a truck carrying hazardous materials depends not just on road classifications, lane widths, and bridge and tunnel restrictions, but also on the number of people living and working near the proposed route. It also should account for hospital and school locations and accident histories by time of day.

From the above discussion, it is clear that transportation models require a variety of data constructs. Depending on these constructs, different demands are made on the GIS. The simplest request by a model is for the GIS to create a sequential file in the format required for input to the model. Complexities of varying degrees arise when trying to get the outputs of the model back into the GIS so that they can be analyzed spatially, working with vectors and matrices that are not part of the traditional GIS formulation, and working efficiently with large, realistic networks.

Network models make important demands on how the data are structured for optimal problem solution. Although a network model might require only a few fields of data as input, the GIS segment data base over which the model is to be solved may have to be recast into a connected link-node formatted network. At each node, the model must have sufficient information to determine quickly which links lead to adjoining nodes.

Because of the enormous amounts of data being processed, GIS operations are of necessity disk-based. That is, the data are read from the disk as needed. By contrast, transportation network models are usually solved by first reading the entire network and the relevant attributes such as length, travel time, and capacity, into random access memory (RAM). Because it is not feasible to read an entire GIS data base into RAM, this is a further argument for a compact network structure different from the one optimized for GIS functionality. A forward star formulation permits adjacent nodes to be quickly identified by the network algorithms. Only the fields required for modeling need to be stored in memory. Shape points, which are important for calculating segment length and for accurately displaying results, are also not needed when solving a network model and need not be brought into memory.

A process that integrates and optimizes the GIS functionality and the modeling functionality has been developed. The GIS data bases can be of enormous size, each with 16 million records of up to 1,000 fields. The network segments are selected by using the GIS querying tools previously described. The network builder turns the GIS segment data base into a compact connected network that contains only the fields required by the model.

As was seen earlier, in addition to point, segment, and polygon data, some transportation models require data in table format. Vectors are used to store trip productions and attractions, supply and demand, and the like. Two-dimensional matrices are used to store minimum travel times, origin-destination flows, and other centroid-to-centroid data. A full set of table manipulation functions has been developed, including those that provide a direct link between tables and the GIS data bases.

The tables are also available for spatial analysis. For example, a distribution table of origin-destination flows between centroids can be represented on the map display as straight line bands whose widths are proportional to the flow.

Research has resulted in the introduction of additional transportation objects to exploit fully the network model formulations in the GIS framework. Results from the models are stored as paths or tours connecting nodes. These transportation objects are new entities, not commonly found in a GIS. They are available for display, manipulation, and analysis using the GIS tools. Thus, optimal routes follow their cartographic paths by making use of the shape points stored with each segment. The flows that result from an assignment model are displayed as dual bandwidths on the map display.

Even tabular data can be displayed on the map as proportional bandwidths connecting centroids.

A macro language has been created for extracting the information required by transportation models from the GIS data bases and for putting model solutions back into the GIS. The macro language permits the implementation of generic models over which the user exerts considerable control. For example, when solving a shortest-path, traveling-salesman, or assignment model, the user is able to select the variables to be included in the generalized cost function as well as their weights.

GIS AS AN INTEGRATOR OF MODELS

Just as a GIS can integrate formerly diverse data bases, a well-designed GIS can serve as a platform for integrating formerly unrelated models. Because the data bases are accessible through the GIS, if the outputs from one model running on the GIS platform can be inserted into one or more of the GIS data bases, these newly created data can serve as inputs to a different model running on the same platform. In turn, outputs from this second model can be used as inputs to a third model or as improved inputs to the first model.

An example can be found in the pavement management process. The traditional pavement management formulation does not require a connected network for its solution. Deficiencies, road type, and the amount of traffic serve as inputs to the pavement management process. Outputs may be a priority ranking of segments, a list of recommended treatments, and, ultimately, the establishment of projects. At this point, a GIS serving as a platform for analysis can extend the pavement management process. Different projects can be scheduled to have minimum impact on traffic flow over the entire network by testing their impacts, singly and in combination, through network assignment models. For a particular project, segments having the highest volume/capacity ratios can be identified as good candidates for signed detours. A shortest-path model can be used to route traffic around these hot spots.

Combining a traffic assignment model and an intersection optimization model illustrates how models can be linked to improve input data assumptions. Initial phasings and turning movement penalties are assumed for the assignment model. The turning movements output by the assignment model serve as inputs to a signal-timing model which then outputs new phasing along with new turn penalties. The updated turn penalties can then be fed into the traffic assignment model, resulting in changes in the predicted assignments. This iterative process could be continued until a satisfactory convergence is reached.

MODELING ENVIRONMENT

Research results indicate that a differentiation between the GIS functionality and the modeling environment is essential for optimal performance. A macro language has been developed that tightly integrates the GIS and the modeling aspects. Each transportation model is a separate executable computer program. An associated command file containing the macros is used to provide communication between the model and the GIS platform. This structure makes it possible for users to add models of all types to the GIS platform. Because the models are stand-alone computer programs, they can be written in any programming language and can use the full memory of the machine. In fact, a DOS Extender compiler has been used so that models running on 386-based PCs can manipulate RAM-based networks containing hundreds of thousands of links and nodes.

TransCAD contains a large number of models that take advantage of this structure. These models include shortest-path procedures, traffic assignment models, routing and scheduling, general network solvers, and allocation models. In addition, many other models have been interfaced with the software. These include pavement management, accident diagramming, highway capacity, and signal timing. A number of researchers are currently interfacing their models with TransCAD. The structure is being continually expanded to include additional transportation objects as required by the new models.

CONCLUSIONS

The traditional GIS formulation's strengths are in mapping display and polygon processing, but a transportation GIS requires new data structures, data objects, interfaces, and procedures to fulfill its potential. Table and network-based models, in particular, place significant demands on the GIS architecture. Research into data base design, transportation objects, and user interface has resulted in the development of TransCAD, a GIS that is fundamentally different from the traditional environmentally oriented GIS. TransCAD is a transportation GIS that provides all the tools that the transportation analyst needs, and at the same time supports complex transportation and operations research models and algorithms in a comprehensive and cohesive structure.