GIS-T DATA MODELS

by Harvey J. Miller and Shih-Lung Shaw

Excerpts from **Geographic Information Systems for Transportation: Principles and Applications**, *Oxford University Press (ISBN 0195123948)*

Geographic information systems for transportation (GIS-T) are interconnected hardware, software, data, people, organizations and institutional arrangements for collecting, storing, analyzing and communicating particular types of information about the earth. The particular types of information are transportation systems and geographic regions that affect or are affected by these systems. Transportation is one of the most important and growing applications of GIS (Fletcher 2000; Wiggins et al 2000).

At first glance, designing data models for GIS-T applications may not appear difficult. In contrast with many other GIS applications, GIS-T has a central object of study, namely, the transportation network in a study area (at least for ground transportation). However, digital representation of these networks is not easy. Transportation network data is complex since it is often multi-modal, exists across many different jurisdictions and has different logical views depending on the particular user or stakeholder. There is often a need to reference events (e.g., accidents, pavement quality) within the network. A network can have varying representations depending on the map scale of interest. There is often a need to represent the relationships between the network and other non-network data. Advanced GIS-T applications require the ability to track conditions or objects over time as well as solve navigation problems.

A large gulf exists between the rich features and attributes of transportation systems in the real world and the data models used for their representation within a computer. First, many GIS software packages only recognize pre-defined, simple geometric entities (i.e., points, lines, and polygons). Most GIS software cannot easily handle data such as origin-destination flows, complex paths, and temporal changes (Goodchild 1998). Many existing GIS have limited and clumsy representations of transportation features such as an underpass/overpass and relationships such as an intermodal transfer between a commuter rail line and a highway (Spear and Lakshmanan 1998). The separation of transportation features into different GIS layers results in a lack of topological information between entities (Shaw 1993).

The Multifaceted Nature of Transportation Data

Table 1 highlights the multifaceted nature of transportation data (Fletcher 1987). Transportation entities have obvious *physical* descriptions but can also have *logical* relationships with other transportation entities. Second, entities exist both in the *real world* and in the database or virtual world. The relationships between the physical and logical realms are often one-to-many, creating database design complexities.

	Logical	**Physical**
Real	*Legal definitions* - Route - State trunk network - County trunk network - Street network - Political boundary	*Actual facilities* - Highways - Roads - Interchanges - Intersections
Virtual	*Data structures* - Networks - Chains - Links - Nodes - Lattices	*Data values* - Lines - Points - Polylines - Polygons - Attributes

Table 1: GIS-T modeling transformations (after Fletcher 1987)

The *real/physical* mode corresponds to transportation facilities as constructed and used in the real world (e.g., physical facilities such as highways, intersections and interchanges). Also in the real world are *real/logical* or legally defined transportation entities such as state and federal routes. The relationship between real/physical entities and real/logical entities are often one-to-many. These one-to-many relationships occur in both directions. For example, two state routes may share the same physical highway. Conversely, a state route can (and often will) traverse several physical streets in an urban area.

Virtual/logical entities correspond to data structures such as nodes, links, networks and polygons. *Virtual/physical* entities correspond to geometric and attribute data associated with the transportation entity. This latter data is often the information displayed graphically by the GIS. A one-to-many relationship occurs when two or more network links correspond to the same graphical line when displaying the network at a given scale (e.g., displaying a two-way street represented logically by two directed arcs as a single cartographic line at small map scales). Also, several cartographic lines can represent one link (e.g., displaying modal-specific flow in a network link).

Transportation Networks in a GIS: The Node-arc Model

A network is a type of graph, a mathematical structure that represents relationships among entities. Rather than relationships, a network represents interaction or movement between point locations. *Nodes* are point locations where flow originates, terminates or relays while arcs are the conduits for flow between nodes. *Arcs* connect nodes; these can represent physical conduits (e.g., a road segment) or logical relationship (e.g., airline service between two cities). Arcs are directed or undirected. If the arc is directed, the node ordering indicates the flow direction. An important difference between a network and a graph is that a network can accommodate weights associated with each arc. Each arc has a *weight* that represents the cost incurred by one unit of flow when traversing the arc.

In the basic "node-arc" representation of a transportation network, we deal exclusively with directed networks (that is, a network consisting of directed arcs) since transportation systems typically have important directional flow properties (e.g., one-way streets, differences in directional travel times depending on the time-of-day).

For a street transportation network, nodes generally correspond to street intersections while arcs correspond to street segments between intersections. Similarly, nodes correspond to interchanges and arcs correspond to highway segments when representing limited access highways. Two directed arcs oriented in opposite directions represent a two-way street or parallel limited access highway segments with opposite directional flow. A generalized cost function represents the unit flow cost for traversing the arc. These typically consist of two major components, namely, any out-of-pocket expenses (e.g., tolls) and the travel time required. The latter component often relates to current flow.

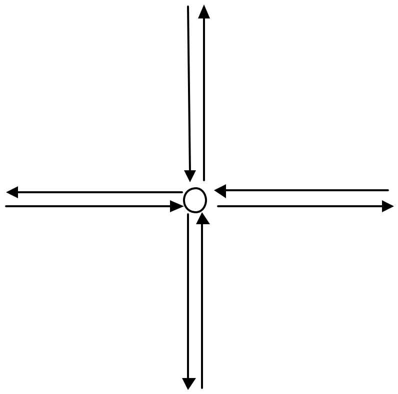


Figure 1a: Single node representation of an intersection

Although nodes correspond to the street network's intersections, these can be represented at varying levels of resolution using the node-arc model. Figures 1a and 1b illustrate two methods for network representation of a street network. In Figure 1a, the intersection is aggregated to a single node. Although parsimonious, this method is simplistic and does not capture a critical intersection property, namely, the varying turn impedances associated with different directions of travel through the intersection. For example, a left turn can require more time than a right turn or traveling straight through the intersection. In addition, turn restrictions may be present (e.g., "no left turn").

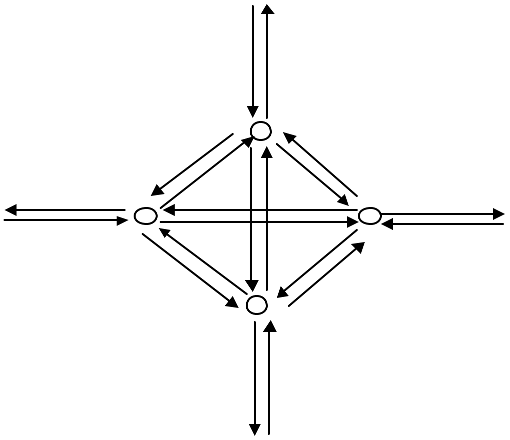


Figure 1b: Expanded representation of an intersection

To capture these features, we can use the expanded representation in Figure1b. This method expands the intersection to four nodes with connecting arcs representing direction-specific travel. Although this method can capture the necessary turn properties of the intersection, a problem is the large increase in the number of nodes and arcs in the network. This not only rapidly expands the data storage requirements for the network but can also degrade the performance of network analytical procedures. For example, the run-time for shortest path routines generally increases as a function of the number of nodes in the network.

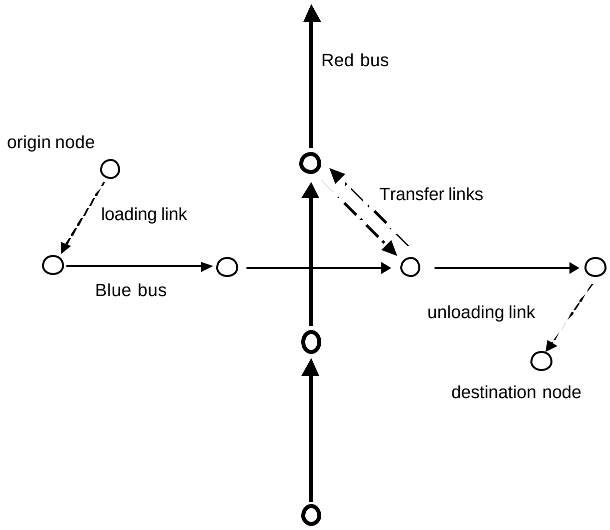


Figure 2: Representing public transit system entrance, egress and transfers

The node-arc representation traditionally partitions a transportation system into separate, modal-specific subnetworks. For example, an urban transportation system may consist of separate networks corresponding to the private transportation network (i.e., street network) and each public transportation network (e.g., bus system, subway system, commuter rail system). Transfer arcs link the separate subnetworks; these arcs represent modal transfers. See Figure 2 for an illustration. Although the traditional node-arc representation is still widely used, emerging transportation network models treat multiple modes in a more sophisticated manner that does not impose an artificial separation.

The Node-arc Model and Relational Databases

The most common logical data model used to support the node-arc representation is the relational model. Figure 3 provides a simple example network and Figure 4 provides the normalized relational structure for this network.

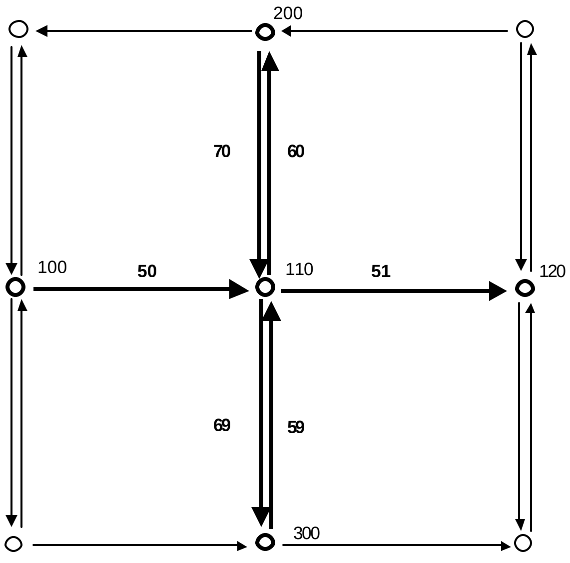


Figure 3: Example network

Arc

arc id	from node	to node	(other attributes)
50	100	110	
51	110	120	
59	300	110	
60	110	200	
69	110	300	
70	200	110	

Node

node id	(attributes)
100	
110	
120	
200	
300	

Turn Table

from arc	to arc	impedance
50	60	-1
50	51	1
50	69	2
70	69	1
70	51	3
59	60	1
59	51	2

Figure 4: Normalized relations for the example network

Other, ancillary relations include turn tables and reference address tables. *Turn tables* are relations for storing data on expanded intersection representations such as the one in Figure 1b. The turn table contains a tuple corresponding to each direction of travel through an intersection. An additional field maintains the travel cost associated with that direction of travel (or perhaps a pointer to a flow cost function). A reserved character (such as a negative number) can indicate a turn restriction. Similar to the expanded intersection representation in the formal node-arc model, the turn table strategy is effective but not efficient. Turn tables require adding twelve tuples to the database for each intersection in the street network. The total can be quite large for a detailed urban street network.

We often want to include information on address locations within the network. This is useful for address matching within the network, i.e., georeferencing entities (such as home addresses, businesses) based on their street address. To maintain address information, we can also add a *Reference Address* relation. This relation includes a “from address” and “to address” fields that provide address range for the given arc, as well as other information such as which side the address range applies and a parity field indicates whether the address numbers on each side are always even or always odd. The street name corresponding to the arc often must be partitioned into the following fields: i) a street prefix (e.g., "North"); ii) street name (e.g., "Oak"); iii) street type (e.g., "Avenue"), and; iv) street suffix (e.g., "East").

Problems with the Node-arc Model

To maximize database integrity, a GIS often requires *planar embedding* of the node-arc network. This means that a node must exist at every arc crossing. Planar embedding ensures that the topology (connectivity) of the resulting spatial data layer is correct, e.g., all polygons are enclosed by arcs. However, the planar network requirement that all arc intersections correspond to network nodes does not match the real-world properties of a transportation network (Goodchild 1998; Spear and Lakshmanan 1998). For example, a limited access highway may pass over or under a surface street. Placing a node at that over- or underpass implies that automobile traffic can turn off or onto the highway. This can cause problems when conducting network routing since the connection may not exist in reality.

A partial resolution of the planar network problem is to relax enforcement of planar topological consistency. Some GIS software allow varying levels of enforced topological consistency when building a spatial data layer (such as not "building" the polygons implied by the node-arc-area topology). However, this can lead to data integrity problems. Another strategy is to use the expanded intersection representation (Figure 1b) and restrict turns at over- or underpasses, although in many respects this is a "work-around" rather than a true extension beyond the planar network. Because of this problem, several emerging GIS-T data models greatly expand on the traditional node-arc model.

Another problem with the traditional network model is the assumption that arcs are homogeneous, i.e., arc characteristics do not change between its end nodes (Goodchild 1998). This is not the case for many transportation applications. An obvious example is pavement management: pavement quality can vary substantially within a given street segment. A less obvious example is traffic flow modeling: the node-arc model implies that network flow levels and related properties (such as travel time) are uniform within the arc. Also, note that this imposes a fixed level of spatial resolution, i.e., we cannot exploit information on lane configurations or street geometry. This information is often useful for advanced applications such as Intelligent Transportation Systems.

A fourth problem is difficulty in supporting one-to-many relationships among transportation entities. As Fletcher (1987) notes, the relationships between real/physical transportation entities (e.g., real world physical entities such as highways, intersections and interchanges) and real/logical transportation entities (e.g., state or federal routes) are often one-to-many. The standard planar network model and its relational counterpart cannot easily accommodate these relationships.

An important type of one-to-many relationship occurs with linear referencing systems (LRS). State Departments of Transportation (DOTs) and metropolitan planning organizations (MPOs) often use LRS (e.g., mileposts) for locating incidents or infrastructure within the network. Since these referencing systems exist at the sub-arc and supra-arc level they cannot be supported effectively using the traditional representation.

Linear Referencing Systems

In attempting to maintain information on transportation infrastructure, many government agencies (e.g., Departments of Transportation or DOTs) and planning organizations (e.g., metropolitan planning organizations or MPOs) have developed *linear referencing systems* (LRS) for transportation facilities. LRS support the storage and maintenance of information on events that occur within a transportation network; these can include phenomena such as pavement quality, accidents, functional classes, traffic flow and maintenance districts. As these examples suggest, events can be points or lines (and, in some cases, areas; see below) referenced within the transportation facility.

A LRS typically consists of the following components (Dueker and Butler 1997; Nyerges 1990; Sutton 1997; Vonderohe and Hepworth 1996): i) a transportation network; ii) a *location referencing method* (LRM), and; iii) datum. The transportation network consists of the traditional node-arc topological network. The LRM determines an unknown location within the transportation network using a defined path and an offset distance along that path from some known location. This provides the basis for maintaining event data within the network. The datum is the set of objects with "known" (directly measured) georeferenced locations. The datum ties the LRS to the real world and supports the integration of multiple networks, multiple LRMs for a given network, multiple event databases and cartographic display of the data. Figures 5a, 5b and 5c illustrates three major LRM.

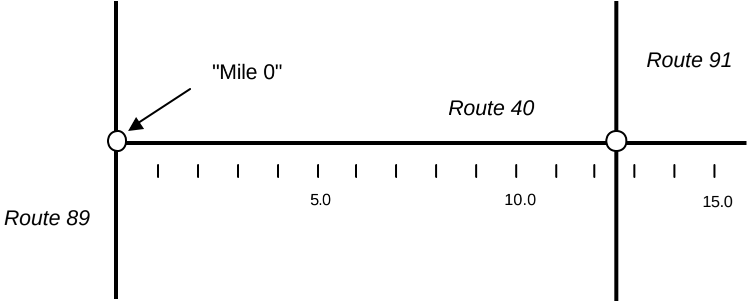


Figure 5a: Linear referencing methods - road name and milepoint

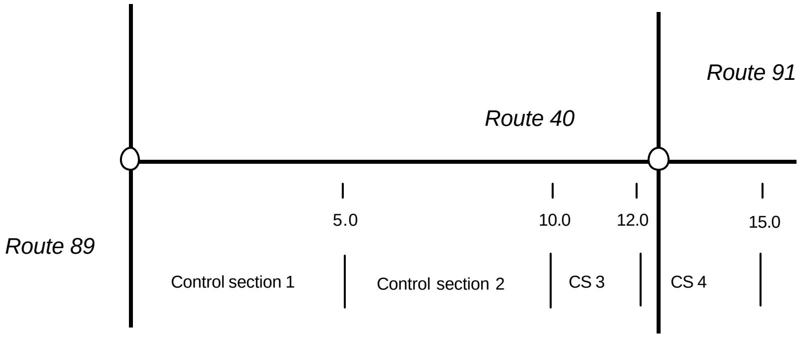


Figure 5b: Linear referencing methods - control section

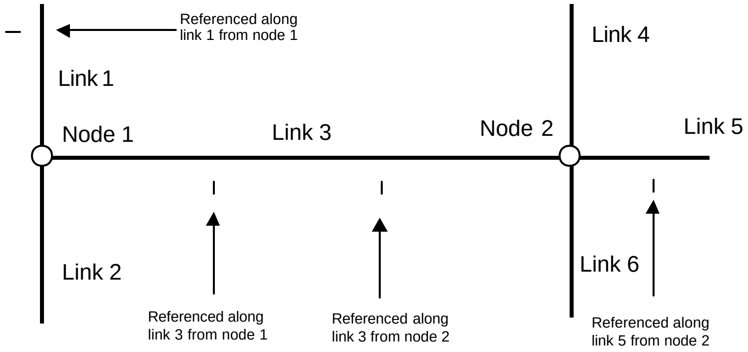


Figure 5c: Linear referencing methods - link and node

Fixed-length and variable-length (dynamic) segmentation

Maintaining data on spatial objects or events referenced using a LRM requires developing a segmentation scheme for the network data model and linking the linearly referenced data to the network model using the segmentation scheme. Two basic types of segmentation schemes are available (Nyerges 1990).

Fixed-length segmentation subdivides each network arc into segments of uniform length and records attribute value for each segment. Each attribute can have its own segment length, although the segment length for that variable is constant across all locations in the network. An advantage of this approach is its simplicity: since we subdivide the arcs, there is a one-to-one correspondence between a segment and an arc (each segment is linked to only one arc). A big disadvantage is that it imposes a fixed level of spatial resolution on the linear data. For example, we cannot determine the spatial distribution of an attribute at a higher level of resolution than designated by the fixed segment length. Also, this approach results in redundant data if an attribute occupies a large contiguous linear region within an arc (i.e., several contiguous segments may record the same attribute value).

Variable-length or *dynamic segmentation* controls the attribute (i.e., holds the attribute value constant) and measures the locations where this attribute exhibits the specified value. In this case, we pre-specify the condition of interest and allow each segment to vary in length to encompass all contiguous locations that exhibit that value. A segment can transcend arcs and therefore have a one-to-many correspondence with arcs. For example, a segment can begin in any location in an arc, traverse several arcs, and then end in another arc. This allows the data model to maintain data referenced using all three major LRMs. Dynamic segmentation can also easily maintain route structures such as bus routes since these are similar in structure to the route milepost LRS. Because of this flexibility, the dynamic segmentation scheme is more popular than fixed-length segmentation (Dueker and Vrana 1992; Nyerges 1990).

Figure 6 illustrates two important components of a dynamic segmentation data structure. A *route attribute table* (RAT) and a *section table* (SEC) to store the relationships between individual routes and network arcs. The RAT has one record for each route and stores an identifier, a "begin post" and an "end post" for each route. The SEC keeps track of the network links that are associated with each route.

Record #	Route#	Route-ID	ROADWAY	BEGIN_POST	END_POST
1	1	2901	87000001	0.000	6.000
2	2	2902	87000002	0.000	5.200
3	3	2903	87000003	0.000	5.400
4	4	2904	87000004	0.000	4.200
5	5	2905	87000005	0.000	7.400
6	6	2906	87000006	0.000	4.100
7	7	2907	87000007	0.000	2.900
8	8	2908	87000008	0.000	1.300
9	9	2909	87000009	0.000	2.400
10	10	2910	87000010	0.000	3.600
11	11	2911	87000011	0.000	2.500
12	12	2912	87000012	0.000	4.984
13	13	2913	87000013	0.000	2.400
14	14	2914	87000014	0.000	5.800
15	15	2915	87000015	0.000	1.013
.	.	.	.	.	.

a. Route attribute table (RAT)

Record #	Route Link#	ArcLink#	F-MEAS	T-MEAS	F-POS	T-POS	Section #	Section-ID
1	1	292	0.000	6.000	0.000	100.000	1	1750
2	2	293	0.000	5.200	0.000	100.000	2	1751
.	.	.	.	.	.	.	.	.

b. Section table (SEC)

Figure 6: Route attribute and section tables in a dynamic segmentation data model

Once a route system is defined under the dynamic segmentation data model, it can be used to associate the locations of event data with the base network GIS map layer. An event is an attribute that is associated with a portion of a route or a single location on a route. Event data are commonly stored in separate data tables known as event tables.

Linear events represent the attribute values associated with discontinuous linear segments along a route. *Continuous events* are associated with the entire route; therefore, they only need to be referenced at the locations where the event value changes. *Point events* occur at a specific point location on a route. Examples include accident locations, bus stops, and intersection locations. Since a point event is defined by a single locational measure along a route, its data table structure is similar to that of continuous events. See Figure 7.

Linear events: Curb-side Parking

Route ID	**From Milepost**	**To Milepost**	**Parking Fee**
441	2.53	2.75	$0.75/hr
441	3.64	4.27	$0.50/hr
441	6.82	7.35	$0.75/hr
68	1.25	1.67	$1.00/hr
.	.	.	.

Continuous events: Posted speed limit

Route ID	**To Milepost**	**Posted Speed Limit**
441	2.5	35 mph
441	6.7	45 mph
441	12.6	35 mph
68	4.3	30 mph
.	.	.

Point events: Bus stops

Bus Route ID	**Milepost**	**Shelter**
10	0.8	Y
10	1.4	Y
10	2.1	N
10	3.3	N
10	4.5	Y
.	.	.

Figure 7: Types of events in a dynamic segmentation data model

Enterprise LRS Data Models

The flexibility of linear referencing systems and the dynamic segmentation data models supports their application in a wide variety of transportation domains, including infrastructure management, public transit, freight, intelligent transportation systems, waterway navigation, hydrological analysis, utilities management and seismological sensing (Vonderohe and others 1995). Although flexibility is desirable, less desirable are independent, application-specific LRS and data models among different departments within a DOT or MPO, different political jurisdictions or different government levels. Ideal is a common LRS and data model that can support all relevant applications. This can support data sharing among agencies as well as interoperability among GIS-T applications. In short, a desirable goal is the development of an enterprise data model for

LRS, that is, a data model with common features that support applications across an entire “enterprise” (e.g., agency, corporation, MPO) or across several enterprises.

NCHRP LRS data model

The U.S. National Cooperative Highway Research Program (NCHRP) sponsored the development of the NCHRP 20-27 LRS data model in response to multiple independent efforts at developing LRM and LRS data models in the United States (NCHRP 1997). Vonderohe et. al (1995) present a draft version of the data model while Vonderohe and Hepworth (1996) provide a review that includes some minor revisions and clarifications relative to the initial report. NCHRP (1997) provides a concise summary of the activities that led to the data model as well as the data model itself.

The NCHRP data model supports the following fundamental operations (NCHRP 1997):

i) **Locate**. Establish the location of an unknown point in the field by reference to other objects in the "real world."
ii) **Position**. Translation of a real-world location into a database location.
iii) **Place**. Translation of a database location into a real-world location (i.e., the inverse of "position").
iv) **Transform**. Conversion between various LRMs represented by database locations, between various cartographic representations and between LRMs and cartographic representations.

By supporting these fundamental operations, the data model also supports "higher-level" operations. These include GIS operations such as overlay, connectivity and proximity as well as network analysis operations required in transportation analysis (pathfinding, routing, facility location and allocation of network resources).

Figure 8 provides a conceptual overview of the NCHRP data model. The data model separates the data elements into three layers. These layers are: i) linear referencing system (LRS); ii) business data, and; iii) cartographic representation. The LRS consists of three sublayers that comprise its standard components, namely location referencing methods (LRM), a network and datum. "Business data" refers to the event data referenced through the LRS.

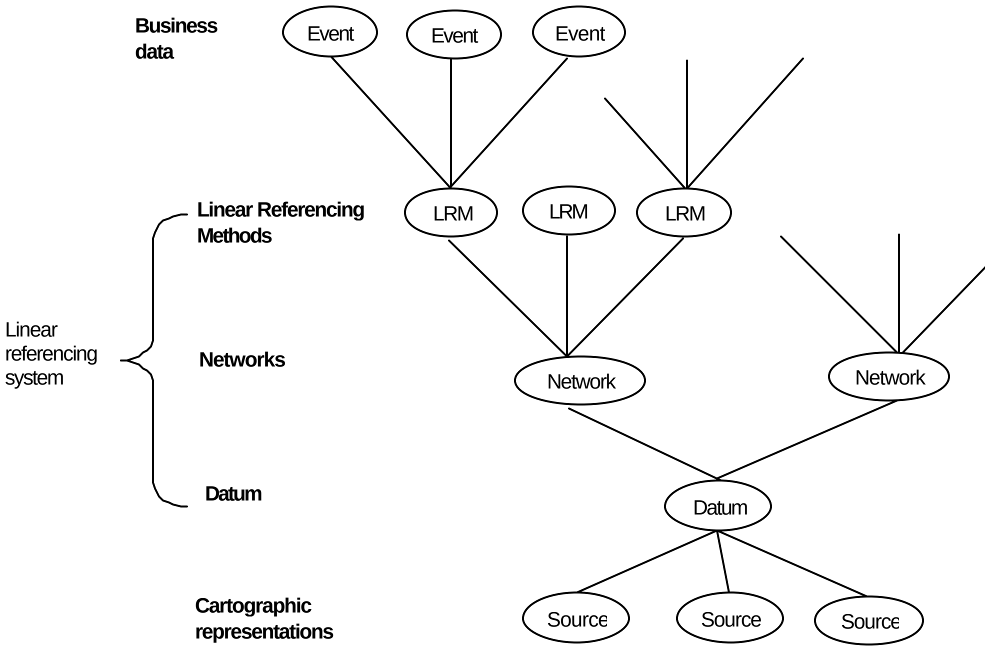


Figure 8: Conceptual overview of the NCHRP enterprise LRS model

Dueker/Butler enterprise data model

Dueker and Butler (1997) provide an enterprise GIS-T data model that extends the enterprise data concept beyond LRS to include other transportation features. Similar to the NCHRP LRS data model, the Dueker/Butler model is based on independence among geographic datum, events and geometry for cartographic display and link-node topology. The model can also accommodate areal transportation features (e.g., airports, railyards) as well as areal events (e.g., "park-and-ride" lots). The Dueker/Butler model is more inclusive (and consequently more complex): in addition to multiple networks, LRS, events and cartographic representation, the model supports other, integrated spatial data (in particular, areal transportation facilities and events).

Figure 9 provides the conceptual data model (in Entity-Relationship format) for the Dueker/Butler enterprise GIS-T data model. The core of the model is the entities `Transportation Feature`, `Jurisdiction` and `Event Point`. A `Transportation Feature` is some identifiable element of the transportation system; this can be a point, line or area. `Jurisdiction` refers to a political or similar entity for designating transportation features and their labels (e.g., zip codes for streets, airport designators such as LAX). An `Event Point` is the location on the transportation feature where some portion of a point, line or area event occurs.

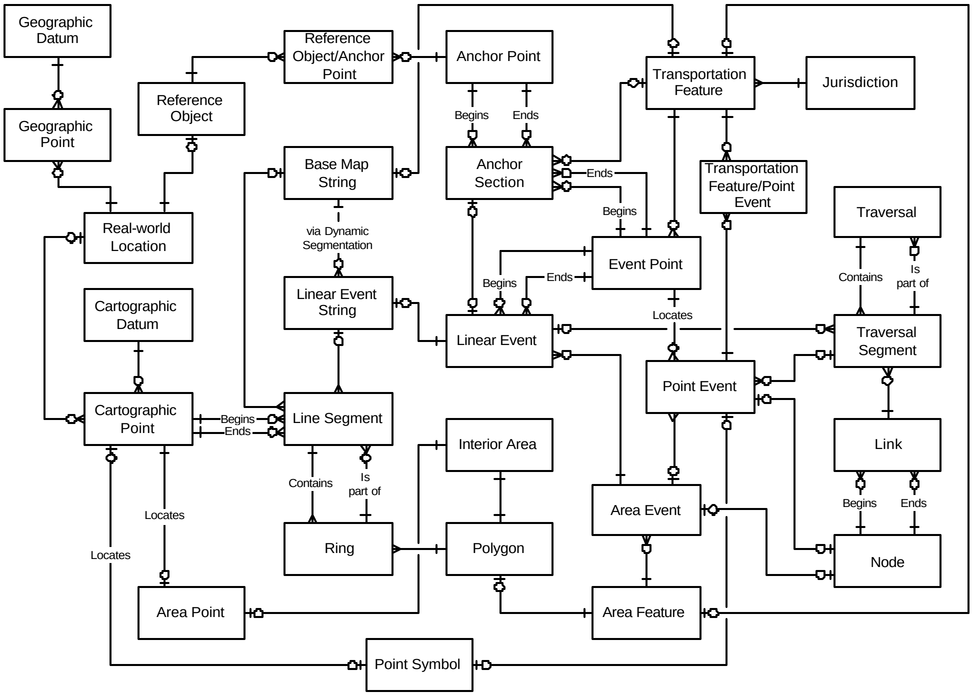


Figure 9: Dueker/Butler enterprise LRS data model

Implementation issues

Since enterprise LRS data models are typically "generic," there are often difficulties in implementing these models in an operational database system. Some of these issues are case-specific and cannot be addressed in a systematic manner: for a review of several implementation case studies, see Bureau of Transportation Statistics (1998).

Several linear referencing "pathologies" can occur due to route definition problems when implementing a LRS. These pathologies include discontinuous routes, dog-leg routes, split roads, cul-de-sacs and ramps (Sutton and Bespalko 1995; Vonderohe and Hepworth 1996). A *discontinuous route* occurs when designated or logical routes stop and start, creating gaps. *Dog-leg routes* occur when designated or logical routes share common sections of a physical transportation facility. A decision must be made with respect to the assignment of events to individual routes along the shared sections unless there is some administrative reason to assign the event to only one of the traversals. The *split road problem* occurs when divided highways have two roadways of unequal length. *Cul-de-sacs* (i.e., a street closed at one end with a circular feature; these are often found in residential areas) can create problems since offset measurements can be arbitrary or non-unique. *Ramps* are often transitions among routes and therefore must be dealt with in a special manner. Some transportation agencies (such as Colorado DOT and Washington DOT in the United States) have developed ramp coding strategies that associate these with particular routes (Vonderohe and Hepworth 1996).

Transportation data models for ITS and related applications

Intelligent transportation systems (ITS) are an attempt to improve the safety, efficiency and capacity of surface transportation systems through the use of information and telecommunications technologies. ITS reflects a shift in thinking about transportation systems. Rather than physically increasing the capacity of transportation infrastructure, ITS intends to use existing capacity more effectively by collecting detailed spatial and temporal data about the transportation system and using this information in transportation system management. This can include techniques such as electronic toll collection, advanced traffic control, information provision to travelers, navigation and vehicle guidance (Branscomb and Keller 1996; Nwagboso 1997).

ITS introduce transportation database requirements beyond the traditional requirements of maintaining arc/node topology, two-dimensional georeferencing and linear referencing of events within transportation features. A fully developed ITS requires a high-integrity, real-time information system that will receive inputs from sensors embedded within transportation facilities and from vehicles equipped with Global Positioning System (GPS) receivers. Information will be updated continuously in a system of databases that maintains a dynamic model of the integrated, multi-modal transportation system. This data will be used to provide route information and navigation to travelers as well as update traffic control devices such as timed traffic lights and variable message signage in real-time (Branscomb and Keller 1996).

ITS data requirements can also go well beyond maintaining the performance of transportation system components. Ancillary information that enhances information provision and route guidance for travelers includes (Golledge 1998):

i) disaggregate data on travelers' characteristics (socio-economic attributes, attitudes/perceptions with respect to decision making and information provision, criteria for route selections);
ii) data on information system characteristics (ITS availability, type of information provided, access costs, reliability), and;
iii) trip/transportation system characteristics (usual trip times, network performance, travel times on each link, accident frequency).

These data must be integrated with the system performance data and accessed often in real-time.

This brief description of ITS implies the following functional requirements for ITS databases. First, ITS requires *navigable data models*, that is, data models that can locate a vehicle within the map reference frame and provide navigation functions based on this position as well as other information about current and anticipated system performance. Second, the diverse ITS databases must be integrated in a manner that is seamless to travelers. *Seamless integration* must occur among the diverse data within an ITS jurisdiction as well as across jurisdictions. Finally, the data must be *interoperable*, that is, easily exchanged and accessed among heterogeneous system components and among different ITS.

Navigable data models

As the name implies, a navigable data model is a digital geographic database of a transportation system that can support vehicle guidance operations at a high level. For ITS, this typically involves the following functions (Dane and Rizos 1998). First, the data model must be able to translate a latitude and longitude into a street address and vice versa. Travelers use address systems for location referencing while the ITS typically tracks a vehicle location using its GPS receiver in a vehicle. Unambiguous translation among locations referenced using both systems must occur if the system is to provide meaningful navigational directions to the traveler.

Second, the data model must support map matching, that is, the ability to "snap" a vehicle's position to the nearest location on a network segment when a reported or estimated location is outside the network. This can occur because of differences in accuracy between the GPS and the digital network database. It can also occur if the ITS uses a "dead-reckoning" system for estimating a vehicle's location between GPS "fixes." Third, the data model must support best route calculation. This refers to assisting the traveler in selecting an "optimal" route based on stated criteria (e.g., travel time, cost, navigation simplicity). This requires a high level of spatial data accuracy including traffic regulations such as turn restrictions. The fourth function is closely related: the data model must support route guidance. This refers to navigation instructions provided to the traveler as the route is executed. This is a challenging task in real time since the ITS must process the traveler's current position, perform address and map matching and then provide navigational cues before the landmark or decision juncture is reached. If a traveler misses a turn, the system should recover gracefully and calculate a new best route and provide directions along this new route.

The traditional node-arc model can support these capabilities, albeit nominally. Much of the high-resolution positional information provided by in-vehicle GPS units and in-facility sensors is lost when referenced within the traditional network structure. While enhancing the network with dynamic segmentation can improve its ability to support positional information along the "length" of each arc, this cannot fully capture complex road geometry such as ramps and non-planar features such as underpasses and overpasses. Also required to support advanced traffic control and navigational devices is positional information resolved to the individual lanes within each segment. Emerging data models for ITS attempt to represent these complex spatial features in order to support more fully the high-resolution positional information and the detailed route guidance and traffic control features available through ITS.

Lane-based network data models

A straightforward way to enhance the traditional arc-node model for ITS applications is to add information on lanes within each arc. Lane information can be used in ITS to provide turn directions, instructions for avoidance of obstructions or restrictions in particular lanes, monitoring lane-specific flow for analysis and modeling and representing beginning and ending points for lanes for enhanced route guidance. This requires a data model that can represent the appearance/disappearance of lanes, connectivity among parallel lanes and at turns, and movement obstructions and restrictions including forward, lateral and turn barriers or restrictions (Gottsegen, Goodchild and Church 1994).

Several strategies are available for enhancing the node-arc model with lane information (Fohl, et al. 1996; Gottsegen, Goodchild and Church 1994). One possibility is to work directly with the arc-node topology and add a node to "split" the links when the number of lanes changes within an arc. However, as noted previously in this chapter, this approach to capturing one-to-many relationships in the traditional node-arc model results in a rapid proliferation in the number of arcs and nodes, unnecessarily increasing the data storage requirements. Also, updating the database to reflect lane changes (e.g., temporary lane restrictions during construction) would be cumbersome and require rebuilding network topology.

A second strategy is to use dynamic segmentation. In this case, a change in the number of lanes is recording using an offset distance or a percent coverage from the beginning of an arc. These route systems cannot traverse more than one arc since this would alter network topology. Two options exist within this general strategy. The first is to store a route corresponding to each lane. Each route would have an independent beginning and ending offset within the arc. The second option is to store the number of lanes as the phenomenon that occurs along the arc. In other words, we would subdivide the arc into portions that have two lanes, one lane etc, with the number stored as an attribute of the route. While the second method is more parsimonious with respect to data storage, a substantial disadvantage is that it does not retain information on which lanes continue or terminate after a location where the number change.

Another property that must be captured by the data model is connectivity among lanes. Two types of connectivity must be captured, namely, connectivity among lanes at turns and connectivity between parallel lanes within an arc (for capturing lane-changing behavior). With respect to connectivity at turns, one possibility is to record, as an attribute of each lane, the possible movement from each lane at a junction to the intersecting arc as a whole. A second approach is to expand the turntable structure used to represent intersections in the traditional node-arc model. In this case, each record of the turntable represents a lane and possible turns on to the intersecting arc. This requires an even larger increase in data storage than the standard arc-arc turntable structure. Also, since it does not record the lane it cannot handle special use lanes well. We can improve on this by an even greater expansion of the turntable, i.e., store a record corresponding to a movement from each lane in an arc to each lane in the intersecting arc. Although it captures the full topology, it results in an explosion in the database storage requirements.

Movement restrictions between adjacent parallel lanes within an arc can be captured either in the geometry or as an attribute of the arcs. In the former case we would treat each lane with restricted movements as a separate arc. Portions with free travel among lanes can be modeled as a single arc. We could also maintain an attribute associated with each lane record indicating whether "turns" from that lane are possible. This latter method is better at handling changes in lane movement restrictions but is less effective for routing algorithms.

Figure 10 describes the relations for a prototype lane-based navigable data model (Fohl et al. 1996). `Lane` maintains information on the lanes within an arc. Each lane is an independent route structure in a dynamic segmentation-type data schema. The lane is the basic element of the data schema; this is a subtle but important difference with "standard" dynamic segmentation schemas in which a lane would be a linear event referenced within a route structure. Therefore, `lane ID` is the key field for the relation while `street ID` is a foreign key linking `Lane` to a "street" relation maintaining spatial information on the polylines corresponding to each street. `Point Turn Table` reflects fully expanded turntable strategy discussed above, i.e., a record corresponds to a turn from each lane in the origin street to each lane in the destination street. Since lanes are the basic data element, each `lane ID` is unique and the turn table need not maintain street identifier data. The `turn ID` field is the key field for this relation, although this is not required since a key can be derived from the combination of the `lane ID`, `to lane` fields. The `position` and `to position` fields maintain the offset distance along the street where the turn is located. Finally, `Linear Turn Table` maintains movement possibilities along lanes. However, rather than a single turn location as in the `Point Turn Table`, turns now have a `start` position and an `end` position where the turns are allowed. The to offset field maintains the location on the to lane corresponding to the start position of the origin lane, i.e., the locations along the to lane that could theoretically receive flow from the origin lane. The distance field maintains the actual distance required for the lane change.

Three-dimensional data models

A more radical approach to navigable data models for ITS is to abandon the node-arc model entirely. Note that the classical arc-node model breaks down when applied to the following domains:

Lane

lane id	street id	from	to lane
		(start position of lane)	(end position of lane)

Point Turn Table

lane id	turn id	position	to lane	to position	impedance
		(start position in origin lane)	(turn destination lane)	(end position in destination lane)	

Linear Turn Table

lane id	turn id	start	end	to lane	to offset	distance	impedance
		(location within lane at beginning of turn)	(location within lane at end of turn)	(destination lane of turn)	(location on "to lane" corresponding to "start" position of turn	(linear distance required to complete turn)	

Figure 10: Relations for a lane-based navigable data model

i) true distance measurement across sloping or hilly terrain;
ii) representation of 3D structures such as overpasses and on-ramps;
iii) assigning multiple routes over a single arc.

These are fundamental flaws of a two-dimensional geospatial representation. Rather than address these fundamental flaws, the GIS-T community attempted to fix the 2D model by adding additional layers of new topologies, particularly routes and LRS. However, the network pathologies reviewed by Sutton and Bespalko (1995) suggest that these problems are not completely resolved and indeed will be even more problematic when attempting to support ITS applications. Bespalko et al. (1998) argue that we should abandon the two-dimensional paradigm resulting from the origins of GIS in automated cartography. A shift from 2D to 3D GIS can eliminate the "pathologies" that plague implementation of LRS data models. These pathologies can create even greater problems for ITS deployment.

Bespalko, Ganter and Van Meter (1996) discuss the characteristics of a 3D, object-oriented GIS-T data model suitable for ITS. This model would represent navigable areas as thin, articulated surfaces that are coaxial and branching. For example, several side-by-side "ribbons" would represent multiple lanes that at some point could be joined by additional "ribbons" or could branch off. Ribbons could run above or below other ribbons to represent 3D structures such as on- and off-ramps. This expands substantially on the positional accuracy allowed by the node-arc-route structure of a dynamic segmentation-enhanced network.

By capturing the third spatial dimension, a GIS can support effectively advanced ITS applications such as route guidance, vision enhancement and automated piloting. Guidance applications provide suggested routes to travelers; this requires clear communication of travel directions at the appropriate decision junctures en-route. A 3D object-oriented data model can distinguish between overpasses, underpasses and intersections, thereby providing guidance through complex intersections or ramp structures. Vision enhancement technology estimates vehicle speed, direction and proximity information using a "heads-up" display. Incorporating 3D information allows more accurate predictions of potential vehicle conflicts in real-time. Long-term advanced ITS planning calls for automated piloting of vehicles in dedicated highway lanes. Automatic piloting requires accurate information on in-road sensors and the ability to react to changing conditions (including terrain) in real-time (Bespalko, Ganter and Van Meter 1996).

Distributed and interoperable transportation databases

Distributed databases and interoperability are two related concepts. A distributed database is a data collection that belongs logically to the same system but is distributed over different sites linked by a communications network. Advantages of distributed databases include (Elmasri and Navathe 1994):

i) **Distributed nature of some database applications**. Often, a database can be partitioned into entities relevant only to local users and entities relevant to global users or everyone interacting with the database. Data accessed by local users can be kept locally while data used by global users can be kept at a central site.

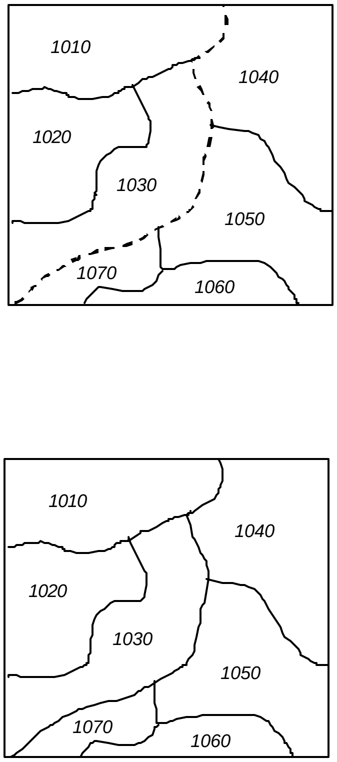
ii) **Data sharing with local control**. A closely related concept to the distributed nature of some database applications is controlled data sharing. This refers to allowing controlled access by local users to local data.

iii) **Reliability and availability**. "Reliability" refers to the probability that a system is up at a particular moment while "availability" refers to the probability that a system is continuously available over some time period. With a distributed database, a site can fail while the rest of the system is still operating. If data are carefully replicated (at the physical level), then the reliability and availability of the entire database increases.

iv) **Improved performance**. Data distributed over several sites means that each site has a smaller database and receives a smaller number of transactions. This improves the system performance, especially as perceived by users.

A distributed database involves the following design principles: i) partitioning; ii) allocation; iii) replication, and; iv) transparency. Partitioning refers to separating the database into fragments suitable for distribution. We can partition any database in two ways, namely horizontal or vertical. A horizontal fragment is selected tuples or records of the database with the records selected based on one or more attribute values meeting a stated criterion. A vertical fragment is selected attributes of the database. It is critical that each vertical fragment retains the primary key for the records so that the original database can be reconstructed (Elmasri and Navathe 1994). Mixed partitioning applies both methods simultaneously.

In digital geographic databases, horizontal fragmentation translates into partitioning according to geographic space. We divide the geographic database into fragments corresponding to different "regions" or geographic domains. Each fragment contains the entire cartographic structure and attributes for that region. Vertical fragmentation in digital geographic databases involves separating the database into thematic layers in a manner similar to the classic "layer" model of GIS database organization. Each thematic layer can serve as a fragment to be distributed; the fragment will contain the cartographic structure for that attribute as well as the recorded attributes. One fragment will also consist of the datum and registration for the geographic database. Figure 11 illustrates these strategies (after Cova and Goodchild 1994).



Parcels (stored at Site A)

polygon id	owner	land use
1010	Lowry	100
1020	Siebeneck	200
1030	LaBonty	300

Parcels (stored at Site B)

polygon id	owner	land use
1040	Zajkowksi	300
1050	Granberg	400
1060	Hofmann	100
1070	Wu	200

Parcels (stored at Site A)

polygon id	owner
1010	Lowry
1020	Siebeneck
1030	LaBonty
1040	Zajkowski
1050	Granberg
1060	Hofmann
1070	Wu

Land use (stored at Site B)

polygon id	land use
1010	100
1020	200
1030	300
1040	300
1050	400
1060	100
1070	200

Figure 11: Distributing a geographic database between two sites

Interoperability is a broader concept. While "distributed database" usually refers to dispersing an integrated and homogeneous database system (i.e., a database with the same conceptual and logical structure) among different nodes in a computer network, "interoperability" refers to integrating independent and heterogeneous database systems. Interoperability is much more complex since it spans the data modeling levels from semantics to data structures as well as encompasses hardware, software and network protocol compatibility (although the boundary between these concepts can be fuzzy in a given application).

Distributed and interoperable databases are critical for ITS since data and system components are likely to be distributed over many sites both within a jurisdiction and among different jurisdictions. However, these concepts can be more broadly applied across all GIS-T domains as distributed network computing and enterprise systems are applied increasingly to many types of information processing activities. In particular, support for integrated highway information systems (IHIS) requires a high degree of data sharing among agencies and interoperability among disparate and heterogeneous Data base management system (DBMS).

ITS deployments are increasingly shifting from centralized to distributed and interoperable configurations. This reflects the widening scope of ITS in terms of geographic coverage and functionality. In the past, ITS configurations have mirrored the "integrate and centralize" principle in non-distributed database systems. Transportation operations centers (TOCs) served as collectors and disseminators of information in computerized traffic control systems. The TOC maintains a high degree of control over the database. "Users" include transportation agencies, local governments, transit authorities, private businesses, researchers and citizens. Users interacted with the database for diverse purposes such as data viewing (e.g., drivers using in-vehicle navigation systems), transferring data to local sites for manipulation and integration with local data, integrating local data into the TOC database for availability to others and modifying the TOC database. This requires users to relinquish rights and control over data and receive access permissions from the TOC database administrator. This impedes participation and removes quality control over local data from the users who are most familiar. Also, large centralized databases can not meet the performance levels necessary for handling dynamic transportation data particularly for ITS operating at a regional or national level (Cova and Goodchild 1994).

A distributed database facilitates data sharing among users; this is essential in an ITS environment where dynamic data from diverse sources must be integrated in real-time. Local autonomy facilitates more effective quality control for local data and increases the likelihood of continuous operation and data availability and reliability. The smaller databases and smaller number of transaction requests at distributed sites facilitates real-time data acquisition and visualization. However, a distributed database potentially introduces problems with respect to data accuracy and integrity. This is a concern since ITS demand very high data accuracy and integrity, particularly for navigational functions (Cova and Goodchild 1994).

Given the diverse and multifaceted nature of ITS applications, it is likely that data sharing will involve multiple users with diverse semantic, schematic and syntactic data models and implementations. ITS data sharing can range from simple communication of location references that must be reconciled within a local database to exchanging and integration of entire databases. The "players" involved in this exchange can include a multitude of public and private sector organizations over a wide geographic area, even up to national scales. A common language for ITS can be difficult to achieve due to the diversity of data semantics even within a single organization. Schema interoperability is even more problematic since proprietary interests arise with the use of private databases (and, in some cases, "public" databases). Some schema interoperability is possible through data transfer standards (such as SDTS) and architectures such as OGIS. However, data transfer standards and interoperable architectures do not address many of the institutional and organizational barriers to data sharing that can plague full ITS deployment (Ganter, Goodwin and Xiong 1995; Goodwin 1994).

Conclusion

This article provided a brief review of data models specifically oriented to GIS-T applications. We reviewed the foundations of network representations and discussed the node-arc model and its implementation using the relational data model. We discussed linear referencing systems and the dynamic segmentation data model, a major extension of the node-arc model that allows one-to-many relationships to be maintained. We also discussed navigable data models and distributed/interoperability issues in ITS.

References

Bespalko, S. J., Ganter, J. H. and Van Meter, M. D. (1996) “Geospatial data for ITS,” in L. M. Branscomb and J. H. Keller (eds.) *Converging Infrastructures: Intelligent Transportation and the National Information Infrastructure*, Cambridge, MA: MIT Press, 209-226

Branscomb, L. M. and Keller, J. H. (eds.) *Converging Infrastructures: Intelligent Transportation and the National Information Infrastructure,* Cambridge, MA: MIT Press, 1-19

Bureau of Transportation Statistics (1998) *Resource Guide on the Implementation of Linear Referencing Systems in Geographic Information Systems*, Bureau of Transportation Statistics, U.S. Department of Transportation, CD-ROM number BTS-CD-22

Cova, T. J. and Goodchild, M. F. (1994) “Spatially distributed navigable databases for intelligent vehicle highway systems,” *GIS/LIS '94 Proceedings,* 191-200

Dane, C. and Rizos, C. (1998) *Positioning Systems in Intelligent Transportation Systems,* Boston: Artech House

Dueker, K. J. and Butler, J. A. (1997) “GIS-T enterprise data model with suggested implementation choices,” *Discussion paper, Center for Urban Studies, Portland State University.*

Dueker, K. J. and Vrana, R. (1992) “Dynamic segmentation revisited: A milepoint linear data model,” *URISA Journal,* 4, 94-105

Elmasri, R. and Navathe, S. B. (1994) *Fundamentals of Database Systems*, 2ed. Redwood City, CA: Benjamin/Cummings

Fletcher, D. (1987) “Modelling GIS transportation networks,” *Proceedings, 25th annual meeting of the Urban and Regional Information Systems Association,* 84-92

Fletcher, D. (2000) “Geographic information systems for transportation: A look forward,” in *Transportation in the New Millenium: State of the Art and Future Directions*, Washington, DC: Transportation Research Board, CD-ROM

Fohl, P., Curtin, K. M., Goodchild, M. F.and Church, R. L. (1996) “A non-planar, lane-based navigable data model for ITS Proceedings, *7th International Symposium on Spatial Data Handling*, Delft, The Netherlands, 7B.17-7B.29

Ganter, J. H., Goodwin, C. W. H. and Xiong, D. (1995) "The open geodata interoperability specifications (OGIS) as a technology for geospatial transportation computing, *GIS-T Proceedings*, 535-553

Golledge, R. G. (1998) "The relationship between geographic information systems and disaggregate behavioral travel modeling," *Geographical Systems,* 5, 9-17

Goodchild, M. F. (1998) "Geographic information systems and disaggregate transportation modeling," *Geographical Systems*, 5, 19-44

Goodwin, C. W. H. (1994) "Towards a common language for IVHS databases, *GIS/LIS '94*, 351-360

Gottsegen, J., Goodchild, M. and Church, R. (1994) "A conceptual navigable database model for intelligent vehicle highway systems," *GIS/LIS '94* 371-380

NCHRP (1997) "A generic data model for linear referencing systems," *National Cooperative Highway Research Program, Research Results #218*, Transportation Research Board, National Research Council, Washington, DC

Nwagboso, C. (1997) "Introduction to intelligent transportation systems," in C. O. Nwagboso (eds.) *Advanced Vehicle and Infrastructure Systems: Computer Application, Control and Automation,* Chichester, U.K.: John Wiley and Sons, 3-32

Nyerges, T. L. (1990) "Locational referencing and highway segmentation in a geographic information systems," *ITE Journal* 60(3) 27-31

Shaw, S.-L. 1993 GIS for urban travel demand analysis: Requirements and alternatives *Computers, Environment and Urban Systems* 17 15-29

Spear, B. D. and Laksmanan, T. R. (1998) "The role of GIS in transportation planning analysis,"*Geographical Systems* 5 45-58

Sutton, J. (1997) "Data attribution and network representation issues in GIS and transportation," *Transportation Planning and Technology,* 21, 25-44

Sutton, J. and Bespalko, S. (1995) "Network pathologies phase 1 report," *Sandia National Laboratories Transportation Systems Analysis GIS Project*, *Document No. AH-2266*

Vonderohe, A. and Hepworth, T. (1996) "A methodology for design of a linear referencing system for surface transportation," *Research report*, *Sandia National Laboratory, Project AT-4567*

Vonderohe, A. and others (1995) "On the results of a workshop on generic data model for linear referencing systems," *GIS-T '95 Proceedings*

Wiggins, L. Dueker, K. Ferreira, J. Merry, C. Peng, Z. and Spear, B. (2000) "Application challenges for geographic information science: Implications for research, education and policy for transportation planning and management," *URISA Journal,* 12, 52-59

About the authors

Harvey J. Miller is Professor in the Department of Geography at the University of Utah. His research focuses on the use of GIS and other information technology in the analysis of transportation and telecommunication systems. Harvey is a board member of the University Consortium for Information Science and Councilor-at-Large for the North American Regional Science Council. He is currently North American Editor of the *International Journal of Geographical Information Science*. He is also co-editor of a forthcoming book on "Geographic Data Mining and Knowledge Discovery" by Taylor and Francis.

Shih-Lung Shaw is Associate Professor in the Department of Geography at the University of Tennessee - Knoxville. His research interests include geographic information systems, urban transportation planning and modeling, GIS for transportation (GIS-T), and spatial analysis. Shih-Lung currently serves as the Chair of Transportation Geography Specialty Group (2000-2002) of the Association of American Geographers. He also serves on the editorial board of *Journal of Transport Geography*.