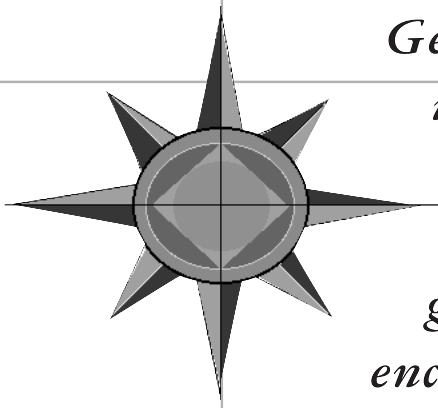
Shashi Shekhar, Mark Coyle, Brajesh Goyal,
Duen-Ren Liu, and Shyamsundar Sarkar

Using object-oriented database technology to model the real world.

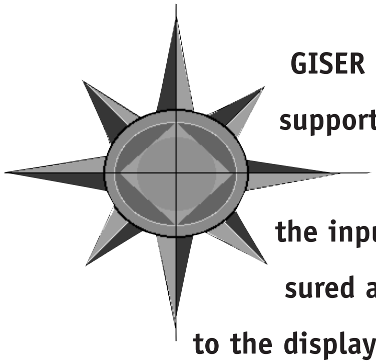
Data Models in Geographic Information Systems



Geographic information systems are used to collect, analyze, and present information describing the physical and logical properties of the geographic world. Geographically referenced data is the spatial data that pertain to a location on the earth's surface.

There are four major functional units in a typical geographic information systems (GIS):

- Data Input Unit. Measurements in GIS are taken by sensors such as cameras and global positioning systems. A manual process is then used for inputting data that cannot easily be processed automatically. The measurements are discretized, for example, by imposing a regular, multidimensional discrete grid over the surface to be measured, allowing points of interest to lie only at the intersection of the grid lines. In addition to the error imposed by this discretization process, measurement errors also reduce the accuracy of attribute values. The data input therefore needs validation. Various integrity constraints including topological constraints also need to be checked. An example of a topological constraint is "Minneapolis should be inside Minnesota."
- Data Model. A conceptual data model is a type of



GISER attempts to support the entire GIS process, from the input of data measured and discretized to the display of this entity and all the data processing that must be performed.

data abstraction that hides the details of data storage [9]. It uses logical concepts, which may be easier for most users to understand. It supports data input, manipulation and result presentation. Many GIS are organized as a collection of themes. Each theme represents the values of a unique attribute of the geographic space. A theme may independently partition, decompose, and fragment the continuous space for a particular value (or value range) of the attribute. The partitions and fragments of space within each theme are often stored within the database and can be treated as entities or objects.

- Data Manipulation Capabilities. Geographic data is queried and analyzed for various operations, including spatial searches and overlays. Operations on primitive vector data-types include geometric operations (e.g., area or boundary, intersection), topological operations (e.g., connectedness) and metric operations (e.g., distance).
- Result Presentation Facilities. A GIS presents results visually (e.g., cartographically) in the form of maps, consisting of graphic images with vector data displayed over raster data; 3D display; animation; and cartographic production. Cartographic maps sometimes highlight semantically interesting information at the expense of locational accuracy. This phenomenon is called map generalization.

Related Work and Contribution

The GIS data models can be categorized into *field-based* models and *object-based* models. Field-based models see the world as a continuous surface (layer) over which features (e.g., elevation) vary. Layer algebra [2] provides a field-based view. It defines a set of operations that can manipulate different layers to produce new layers. The object-based model treats the world as a surface littered with recognizable objects (e.g., cities, mountains, rivers), which exist independent of their locations. GraphDB [7], GODOT [5], Worboy [12], OGIS [3] and GeoOOA [8] are some attempts to model GIS using the object-based approach. OGIS provides a library of spatial types (e.g., point, line, chain) and operations on these types (e.g., intersect, overlap) to facilitate data exchange across different GIS. GeoSAL, Worboy and GODOT propose extensive class hierarchies to model the geometry and topology of spatial objects. GraphDB supports the explicit modeling and querying of graphs. GeoOOA adds a geographic dimension to each object modeling a spatial entity, and it supports a fixed set of geometric types and topological relationships. The GERM model [11] attempts to unify the two approaches and provides a set of concepts as an add-on to the ER model for modeling GIS.

Table 1. Summary of requirements

Functional Unit	Data Modeling Requirement	Examples / Issues
Data Input	Data Consistency and Quality Continuous Space	Constraints, Interpretation Discretization, Interpolation models
Data Modeling	Geometry Topology	Points, curves, polygons, etc. Networks, Partitions
Data Manipulation	Set Operations, Spatial Relationships, Network Analysis	Topological, Direction, Metric
Result Presentation	Visual Representation	Visualization constraints, Maps

These models do not explicitly support the discretization process and interpolation to invert the discretization. This omission leads to several complexities in the existing models, including the dichotomy between field- and object-based approaches.

THIS ARTICLE PROPOSES A NEW MODEL called the Geographic Information System Entity Relational model (GISER). This model explicitly represents the discretization aspects so as to unify the two approaches. GISER extends the Enhanced Entity Relationship (EER) model [9] to include continuous fields. The continuous fields are associated with discretizations and interpolation models.

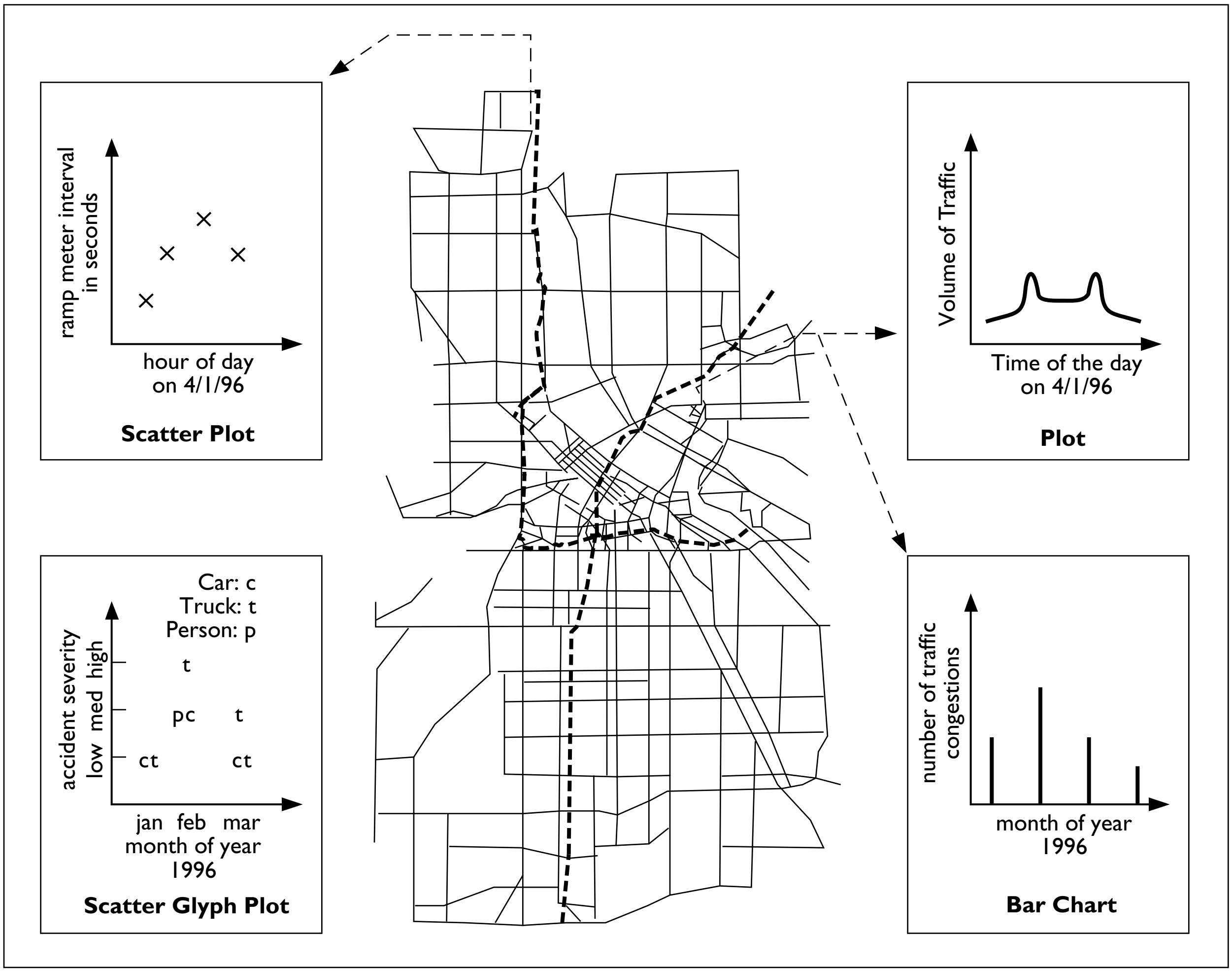


Figure 1. Road map (network) and traffic properties

GISER also uses procedure-valued attributes (e.g., the *interpolation model*) within the EER model. This is not a major departure, as most modern databases support procedure-valued fields. GISER attempts to support the entire GIS process, from the input of data measured and discretized to the display of this entity and all the data processing that must be performed. GISER has distinct identifiable components corresponding to each functional unit.

Data Modeling Requirements

Various requirements of a GIS in relation to each of its functional units will be outlined in an informal manner. Readers are referred to [10] for a more formal discussion. Table 1 summarizes the modeling requirements on which this article will focus.

Data Input Unit

Consistency and Quality. The measurement and discretization process is prone to errors and inaccuracies. Raw geographic measurement data is often processed and interpreted. It is processed to filter noise or out-of-range data. Then it is interpreted to fit a domain prototype, and from that, outliers are removed. The results of the analysis of this data needs to be interpreted within the error tolerance of the data.

Discretizing Continuous Space. Geographic phenomena occupy continuous space. However, computers and digital databases can only store and manipulate their discrete approximations. The process of discretization raises the issue of interpolation, which estimates the values at various points to generate a cartographic presentation.

Discretization could either be uniform, that is, independent of the spatial entity that it is modeling (e.g., laying a rectangular grid over a continuous two-dimensional domain), or it could be dependent on it (e.g., a thematic layer [11]). A layer is a mapping from a domain of non-overlapping geometric regions to a domain of attribute values. In a simple case, a thematic layer consists of a set of polygons, with each polygon being an area of constant value for an attribute.

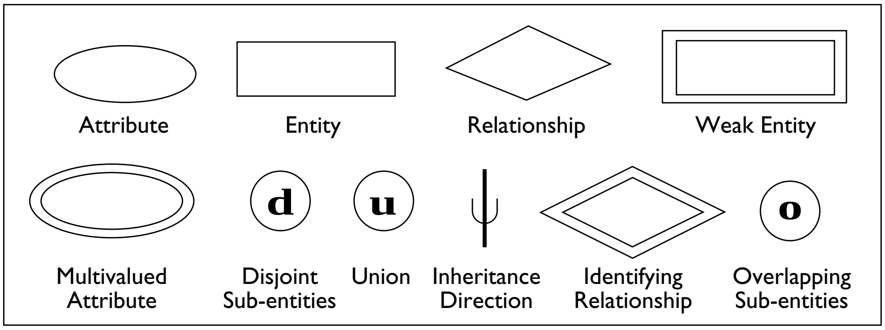


Figure 2. Enhanced entity relationship (EER) model

Data Model

Geometry. Geographical entities have geometric properties that can be modeled by the measurements, properties, and relationships of points, lines, angles, and surfaces. Two types of geometric datatypes are prevalent: *Vector* data and *Raster* data. Vector data include points, lines, and polygons, all of which are representations of the space occupied by real-world entities. Raster data is characterized as an array of points, where each point represents the value of an attribute for a real-world area. Baumann [1] gives a more thorough description of various raster image types.

Topology. Some geographic entities have topological properties that are unaltered by elastic deformations. Examples include the connectedness of a region and the connectivity between road intersections via road segments. Primitives are required for representing networks, graphs, and partitions as high-level entities [7]. Partitions are related to networks in that they associate regions with other regions by relationships such as next or adjacent. It is natural to use a direct construct for networks and partitions, for example, for modeling. Additional discussion of topology in GIS can be found in the Worboy model [12].

Data Manipulation Capabilities

Queries in GIS often involve set operations on the geometric and topological properties of spatial entities. These set operations could be classified into the following groups:

- *Spatial selection* of a subset from an entity set that fulfills a spatial predicate. Some examples are: *Find all cities in Minnesota, Find all cities no more than 500 miles from Minneapolis.*
- *Spatial join* produces a set of pairs of spatial objects from two layers or entities that satisfies a spatial predicate. For example: *For each river, find all cities within 50 miles.*
- *Transformation* synthesizes a set of layers (a set of spatial objects) into a new layer using spatial predicates. Some examples are: *Map generalization, transformation of vector layer to raster representation.*
- *Network analysis* represents a set of queries on spatial networks, such as route evaluation, network overlay, and path optimization. Route evaluation is concerned with aggregating attribute data over route-units. A route-unit represents a collection of arcs with common characteristics (e.g., name). A network overlay enables the integration of disparate network-attribute databases, which join two or more sets of attributes. Path Optimization models several problems, including shortest path analysis and optimum tour routing.

Queries in GIS use spatial relationships within the query predicates. Spatial relationships can be organized into three categories [6]:

- Topological relationships. These include connected, adjacent, inside, and disjoint. These are invariable under topological transformations like translation, scaling, and rotation.
- Direction relationships. These include above, below, or north_of, southwest_of.
- Metric relationships. These include relationships such as the distance between two entities.

Result Presentation Facilities

Results in GIS are visually represented in the form of maps, tables, or plots. Visual representation should be able to present all the required information and should not present any information which is not intended. Geographic data is most often displayed on a map. Information on a map is sometimes distorted to show some useful feature. For example, the road map in Figure 1 shows each road segment as a straight line. In light of such distortions, maps should be used with caution. Different traffic features related to the road map are shown using other types of visual representation including a plot, a bar chart, a scatter plot, and a scatter-glyph-plot.

The GISER Data Model

The Geographical Information System Entity Relationship model is shown in Figure 3. It uses the Enhanced Entity Relationship [9] diagram notation

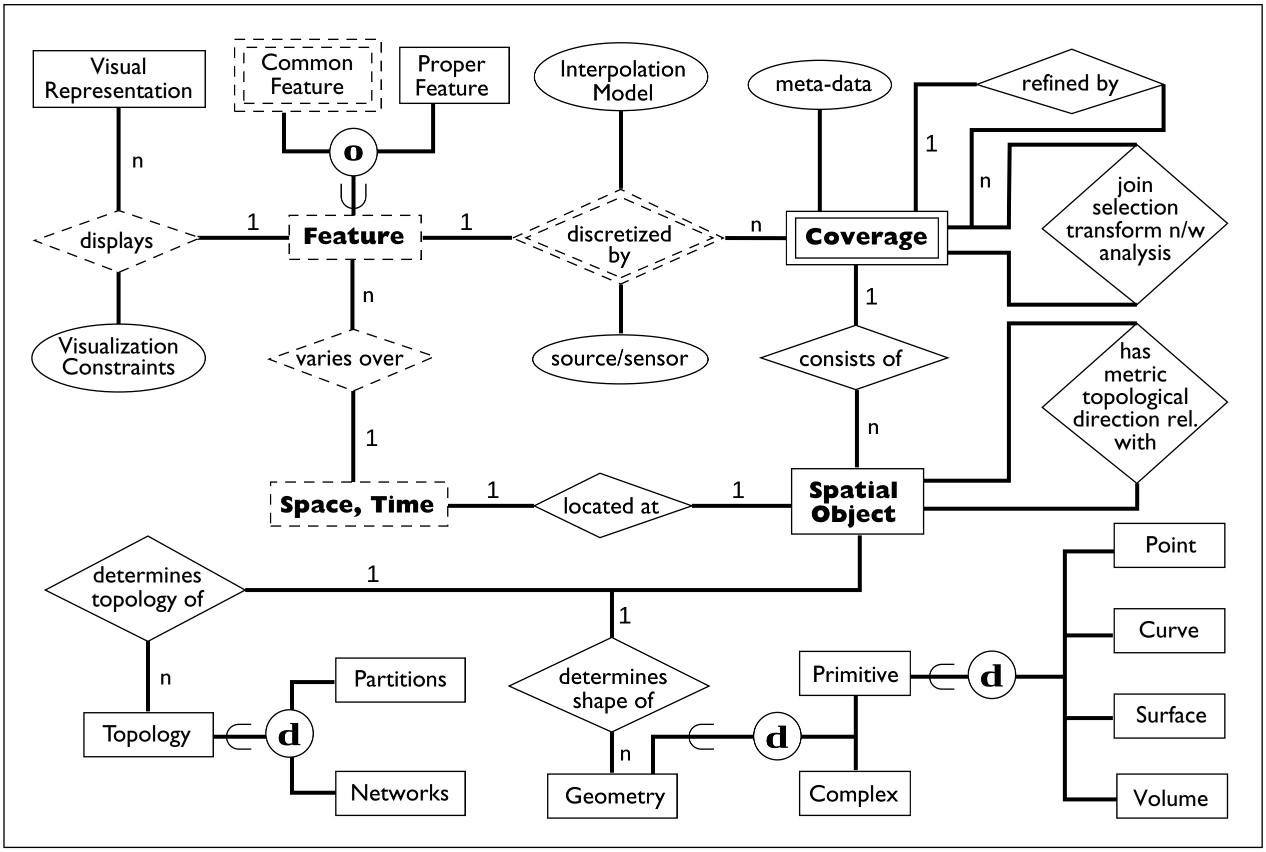


Figure 3. The GISER model

described in Figure 2, along with dashed lines for continuous fields and relationships. The GISER model is based on four major concepts: *Space/Time, Features, Coverages* and *Spatial Objects*. Space/times represents boundless multidimensional extents in which geographic phenomena and events can occur and have relative position and direction. It is a continuous field and may possibly be discretized into realms and calendars if needed. Examples of realms include the surface of the earth and its subsets of interest.

Features represent geographic phenomena such as rivers, vegetation, and cities. GISER models features as continuous fields varying over space and time, thus features as such cannot be stored in a GIS. These must go through the process of discretization in order to compute *coverages,* which are then stored in a GIS. A feature may have multiple coverages based on multiple discretization with varying resolution, accuracy, and sources. A coverage consists of a set of *spatial objects*, which occupy a subset of space and time and have geometric and/or topological properties.

Data Input Unit

Data Consistency and Quality. GISER provides the relationship *refined by* to support the data collection process as well as multiple levels of refinements. GISER states that raw data is refined into processed data, which is refined into interpreted data. Ancillary descriptive data, often called meta-data, is provided to facilitate the interpretation and use of the processed data. Meta-data on each coverage would document the refinement procedure and the source.

Continuous Space Modeling. GISER includes continuous fields named features and space/time as well as the relationship *discretized by*. Each feature represents a mapping from space to a domain of values. Examples of features include elevation, soil type, and water level. Some features are *proper* features, which represent a collection of uniquely named geographic places such as rivers, cities and countries. Each instance of these entities has a unique name, and the entity's geographic location can sometimes change. For example, rivers such as the Mississippi flood and change their courses over time. Features that are not proper features are *common* features, which are identified by their location in space and time and do not have an identification of their own. These entities can be regarded as weak entities that are dependent on and identified by their spatial locations. Examples of common features are land parcels and political boundaries.

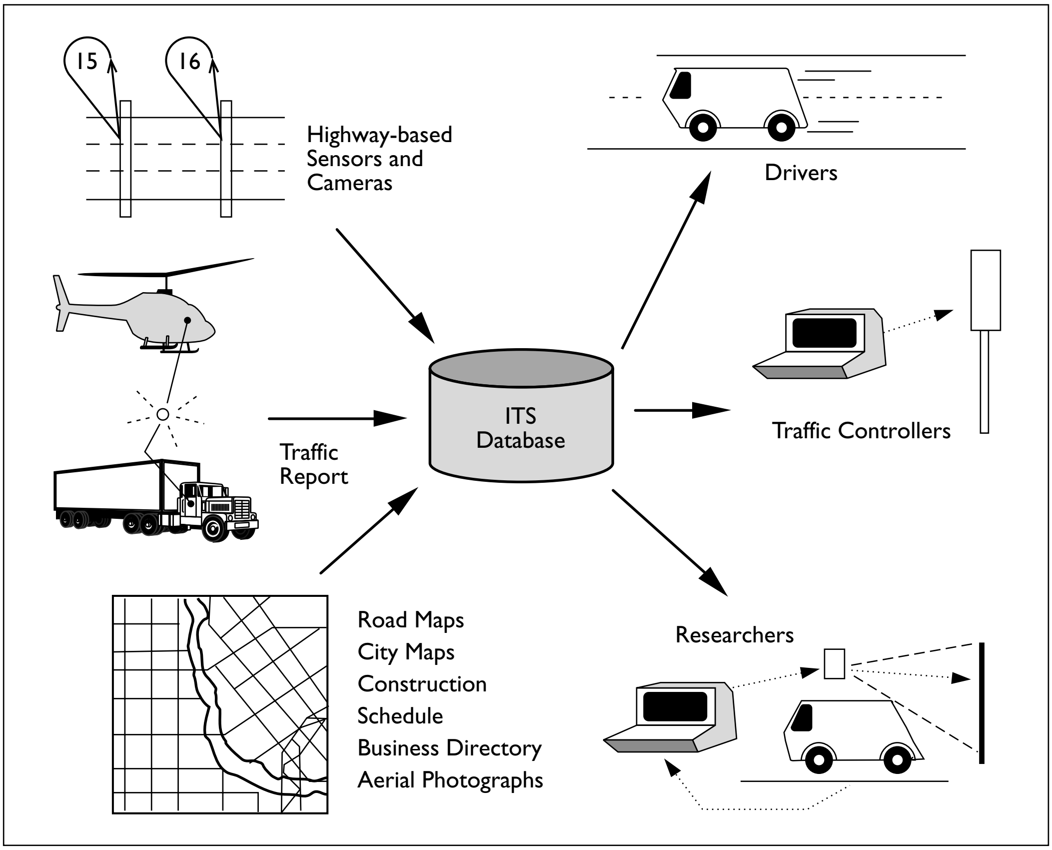


Figure 4. ITS data flow

In order to accurately model these entities, GISER makes explicit the fact that continuous entities are discretized for representation in the database system.

Features are discretized to give coverages. There could be multiple coverages for the same feature depending on the source/sensor. Different interpolation models need to be used for different coverages to retrieve a feature. Specifically, GISER includes the relationship *discretized by* with the attributes interpolation model and source/sensor to model this. A coverage has multiple spatial objects, and is modeled with the relationship *consists of*. A spatial object occupies a subset of space and time, and is modeled using the relationship *located at*. Features occupy a subset of space and time and its attribute takes values defined for a subset of space and time. This is modeled using the relationship *varies over*.

Data Model

Geometry. In the GISER model, geometry is an entity that is related to a spatial object by the relationship *determines shape of*. Additional entities represent the primitives such as points, lines, and polygons as proposed in related models [3, 12].

Topology. Topology is a property belonging to a spatial object and that property remains unaltered even when the object deforms. An example is a road network. The two nodes in the network thus remain connected even if the path between the nodes is changed by road construction. In order to represent the topology, the basic primitives such as networks (i.e., graphs) and partitions are provided. Additional primitives can be added on lines of the Worboy model [12].

Data Manipulation Capabilities

Spatial relationships are added to the model to accommodate spatial operations in GIS. These relationships include topological relationships, direction relationships and metric relationships, as described previously. These relationships serve as spatial predicates for queries in GIS. Directional relationships involve the location of the objects (examples are north_of, south_of, and northeast_of). Topological relationships involve the regions occupied by the objects (examples are adjacent_to, inside, etc.) Metric relationships could involve the geometry and location of the objects (examples include the distance between two objects, or the area of an object that is occupied). To simplify the figure, only the relationships involving spatial objects are kept. A query in GIS is a set operation on one or more coverages to give another coverage. This set operation could be join, select, transform or analyze network.

Result Presentation Facilities

GISER proposes that visual representation be specified in the database to declaritively specify the essential properties of the map types. Visual representation consists of primitives such as text,

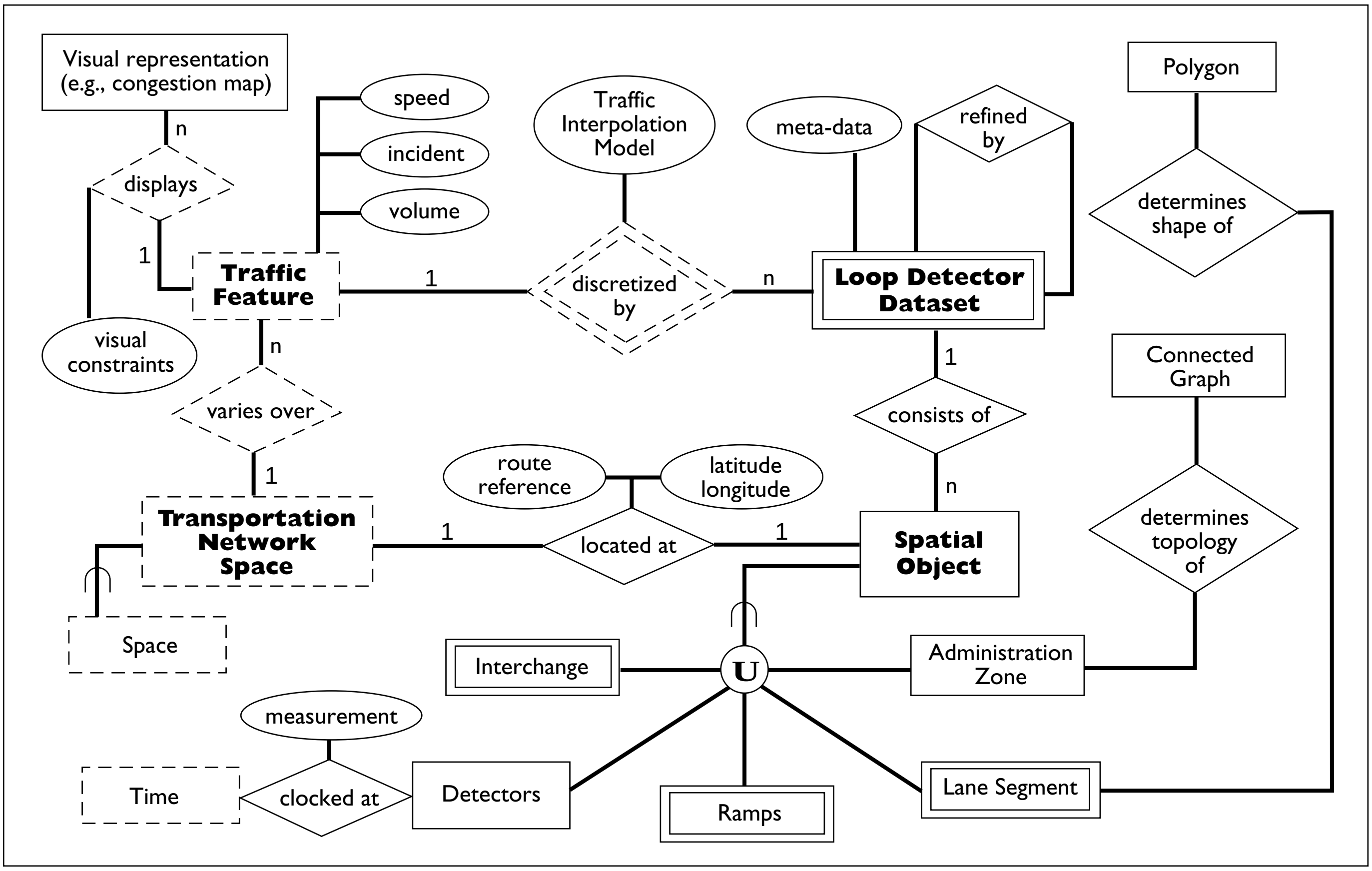


Figure 5. Model of loop detector dataset

icons, graphs and geometries like points, lines, and polygons. These primitives are associated with a location and orientation inside a visual representation. In addition, visual constraints ensure the visual representation does not convey any information which is not present in the geographic data, and that the visual representation should convey all the information requested by the user. Visualization constraints on a road map include: connectivity—that connected road segments should remain connected in the visual representation; and location—that the distortion of the location of various objects should be maintained within acceptable ranges.

Figure 1 shows a few examples of the visual representation of a road map and of the traffic properties associated with the road map. The representation plot assumes both of the keys are continuous and ordered (e.g., time of the day, volume of traffic at the road segment). The bar chart assumes that one attribute is discrete and that the other is continuous (e.g., month of the year, amount of traffic congestion at the intersection). The scatter plot is a visual representation of a discrete feature and an ordered discrete feature (e.g., ramp meter interval, hour of the day). The scatter-glyph plot is a visual representation of two discrete features (e.g., accident severity, month of the year).

An Example GIS Application

Intelligent transportation systems (ITS) are being developed to improve the safety and efficiency of highway travel. Major data sources and users in ITS are shown in Figure 4. Three sources of data are depicted. First, sensors on the highways produce measurements of traffic flow at regular intervals. The road and highway maps represent a second source of data. The traffic reports represent a third source of information to an ITS. Data sent to drivers represents one type of result presentation in an ITS. Vehicles contain graphic display devices that display road maps and current congestions. Traffic controllers at a traffic management center (TMC) may use the data. Also, researchers will often use the data from an ITS for driving simulation to study traffic management and human safety issues.

The GISER Model of the Application

Figure 5 shows a fragment of the GISER data model for a traffic measurement data set. It includes continuous fields (e.g., traffic, visual representation), entities (e.g., loop detector dataset) and relationships

Future work includes the evaluation of the GISER model in different GIS application domains.

(e.g., discretized by, displays). It is based on four major concepts: *transportation network space, traffic, loop-detector dataset* and *spatial objects*. Transportation network space is a subset of geographic space consisting of freeways, highways, and roads. Traffic is the phenomenon of interest in an ITS application, which represents the movements of vehicles and pedestrians over the transportation network space and is characterized by attributes such as speed and volume. The volume of traffic on a road segment refers to the number of vehicles moving across a cross-section. The loop-detector dataset is a set of measurements of the traffic phenomena. The loop detector is a magnetic sensor embedded in the road pavement that measures the volume of traffic. The detectors are linked to the central TMC to collect the traffic measurements over the entire network. The loop detector dataset consists of the traffic measurements and the detector locations. Detector locations are classified into area types such as ramps, interchanges and lane segments. Several sensor groupings are of interest, including stations and administration zones. Stations group the detectors on different lanes of a road segment at a milepost. Administrative zones group detectors belonging to common administrative units such as cities or counties.

Data Input Unit. Loop detectors provide continuous analog measurements. A local analog-to-digital converter discretizes the analog readings and accumulates them for a specified time period (e.g., 30 seconds). Errors can occur due to bad sensors, out-of-calibration sensors, and communication hazards. This erroneous data needs to be filtered out at multiple levels. At one level, erroneous data may be filtered using range checks, parity checks, and similar methods. At another level, it may be removed by using traffic-flow models and statistical methods. Traffic feature could either be modeled as a proper feature or as a common feature: one instance of a proper feature is aggregate traffic on freeway 35W; one instance of a common feature is aggregate traffic on 35W at mile 5. Measurements from magnetic-loop detectors and cameras represent different coverages. Even the transportation network space is discretized by lanes across its width and by road segments across its length. Traffic measurements at adjacent detectors can also be interpolated over space and time using the traffic flow theory [4].

Data Model. The highways are often represented in more than one format, since different users of the data have different requirements. One class of users requires that the lane segments, interchanges, and ramps be represented as polygons. Other users work with the topological relationships of the highway network to perform network computations.

Data Manipulation Capabilities. These are the useful aggregate queries for the TMC application. The metric, directional and topological relationships allow queries such as the following:

- Spatial Selection. Show all the incidents on freeways within 5 miles north of downtown Minneapolis.
- Spatial Join: Show the intersections of railway lines and roads in Minneapolis.
- Spatial Transformation: Show traffic speeds on all freeway segments. Relative to speed limit, transform 80%–100% to green, 60%–80% to yellow, blocked to red, and the remainder to orange.
- Network Analysis. Find the shortest path between two places. Find the shortest path that covers all the road segments to plan removal of snow from all streets.

Result Presentation Facilities. The primary visual representation for this application is a road map. A road map can visually show different road types (i.e., one-way, two-way, or major roads), by using a different color for each road type. A chosen path can be given a different color from other roads. Multiple attributes of the road (e.g., speed limit, volume) can be shown using different visual attributes of the lines (e.g., color, thickness, shape). As an example, on the freeway map, show road segment as green if traffic is moving at 80%–100% of posted speed, as yellow if at 60%–80%, as red if blocked and all others as orange.

Other ancillary visual representations like plots and bar charts can also be used, as shown in Figure 1. The plot for the time of the day and the traffic volume is used for identifying rush hours. The bar chart for the amount of traffic congestion and month of the year is used to analyze seasonal variation. The scatter plot for the ramp-meter interval and the hour of the day can be used to correlate ramp-meter rates with traffic volumes for validation of ramp meter rates. The scatter-glyph plot for accident severity and month of the year can be used to analyze seasonal variation in type and frequency of accidents so as to improve safety.

Conclusion

The GISER data model integrates the field-based and object-based models of geographic data by using the discretizes relationship between feature fields and coverage entity. This leads to a simple data model for geographic data like the loop-detector dataset in transportation. Several data management aspects of data input, modeling, query and result presentation can be supported by the simple integrated model.

GISER implementation using current relational and object databases raises the issues of implementation of continuous fields (i.e., features). A possible approach is to consider continuous fields to be derived data, which is not physically stored in the database. A procedure to derive continuous fields from various discretizations is assumed. In the context of GIS, each coverage has an interpolation model for the discretized feature. Given perfect interpolation models, each coverage will lead to the identical continuous field feature. However, given approximate interpolation models, each coverage would yield different estimates of the feature, and these estimates will differ from each other within the margin of error. The meta-data associated with coverage will allow the interpolation of estimated feature values in the context of the approximation errors.

Future work includes the evaluation of the GISER model in different GIS application domains. We would also like to expand the data model to improve the modeling of dynamic and temporal features such as a waterfront. A particular manifestation of the dynamic nature of geographic data occurs in modeling the boundaries of lakes, oceans, and other natural phenomena. The geometric extent of oceans is dynamic, differing a great deal between low tide and high tide, and therefore presents an interesting challenge. C

ACKNOWLEDGMENTS

The comments from A. Elhaddi of the National Weather Service, Christiane McCarthy, Siva Ravada, and Andrew Fetterer have greatly improved the readability and technical accuracy of this article.

REFERENCES

1. Baumann, P. Management of multidimensional discrete data. *Very Large Databases Journal 3*, 4 (Oct. 1994), 401–444.
2. Delis, V., Hadzilacos, T., and Tryfona, N. *An Introduction to Layer Algebra*. Tech. Rep. CTI-94.01.2, Computer Technology Institute, University of Patras, Greece, 1994.
3. Draft Base Document-OGIS Project Document 94-025R1. The Open Geodata Interoperability Specification, October 1994.
4. Gerlough, D.L., and Huber, M.J. *Traffic Flow Theory: A Monograph. Special Report 165*. Transportation Research Board National Research Council, Washington D.C., 1974.
5. Gunther, O., and Riekert, W. The design of GODOT: An object-oriented geographical information system. *IEEE Data Engineering Bulletin 16*, 3 (1993).
6. Guting, R. An introduction to spatial database systems. *Very Large Databases Journal 3*, 4 (1994).
7. Guting, R.H. GraphDB: Modeling and querying graphs in databases. In *Proceedings of the Int. Conference on Very Large Data Bases* (1994).
8. Kosters, G., Pagel, B.U., and Six, H.W. Object-oriented requirements engineering for GIS applications. In *Proceedings of the ACM International Conference on Geographical Information Systems* (1995).
9. Navathe, S.B. Evolution of data modeling for databases. *Commun. ACM 35*, 9 (Sept. 1992), 112–123.
10. Roman, G. Formal specification of geographic data processing requirements. *IEEE Trans. Knowledge and Data Engineering 2*, 4 (Dec. 1990).
11. Tryfona, N., and Hadzilacos, T. Geographic applications development: Models and tools for the conceptual level. In *Proceedings of the ACM International Conference on Geographical Information System* (1995).
12. Worboy, M.F. Object-oriented approaches to geo-referenced information. *Int. J. Geographical Info. Syst. 8*, 4 (1994).

SHASHI SHEKHAR (shekhar@cs.umn.edu) is an associate professor in the Department of Computer Science at the University of Minnesota.

MARK COYLE (mcoyle@us.oracle.com) is a research scientist at Oracle Corporation.

BRAJESH GOYAL (brajesh@cs.umn.edu) is a graduate student at the University of Minnesota.

DUEN-REN LIU (dliu@iim.nctu.edu.tw) is an assistant professor with the Institute of Information Management at the National Chaio Tung University in Taiwan.

SHYAMSUNDAR SARKAR (sarkar@informix.com) is a senior member of the research staff at Informix Corporation.

This work was supported by the Federal Highway Authority (FHWA), the Intelligent Transportation Institute (University of Minnesota), Computing Devices International (Minneapolis), and the Minnesota Department of Transportation. Portions of this work was done when S. Sarkar was at the Unisys RDBMS group in Roseville, Minnesota.