# Spatial Data Model for Feature-based GIS


Wang Feng*
Dept. of System Engineering & Mathematics
National Univ. of Defense Technology
Changsha, Hunan, 410073, P.R.China
fengwang@nudt.edu.cn

Sha Jichang
Dept. of System Engineering & Mathematics
National Univ. of Defense Technology
Changsha, Hunan, 410073, P.R.China
jcsha@nudt.edu.cn

Yang Shuqiang
Dept. of Computer Science
National Univ. of Defense Technology
Changsha, Hunan, 410073, P.R.China
sqyang@nudt.edu.cn

Chen Huowang
Dept. of Computer Science
National Univ. of Defense Technology
Changsha, Hunan, 410073, P.R.China
hwchen@nudt.edu.cn

*State Key Lab. For Novel Software Technology, Nanjing Univ., Nanjing, 210093, P.R.China



## Abstract

In recently advanced GISs, geographic phenomena are represented as features with associated topologic relations and classification attributes. A feature is defined as two-fold: both a real world geographic entity and its digital representation [10,4,8,9,7]. Semantic relations and intrinsic interrelations of the features themselves are generally neglected. In this paper, the intension of a feature' concept is extended. Features, as the fundamental depiction of geographic phenomena, encompass geometric objects, geographic entities and graphic symbols. A feature-based model, TLDM (Three-Leveled Data Model) is described, that represents geographic phenomena from three levels: geometry, geography and graph, among which the interrelations are analyzed in details. Lastly, the development of an object-oriented prototype feature-based GIS, YH-GIS, is described briefly.



**Keywords:** GIS, spatial databases, data model





**Proceedings of the Workshop on Computer Science and Information Technologies CSIT'99**
**Moscow, Russia, 1999**


## 1. Introduction

The traditional approaches to representing spatial features in exiting GISs, such as vector and raster, are primarily based on the geometric components of spatial phenomena. Semantic and intrinsic interrelations of the objects themselves are generally neglected. The deficiency may significantly affect the analytical power of GIS [7].

Problems and errors generated by a GIS are, to a great extent, affected by the design of the data model and the method of data representation [5]. Researches have recognized the need for a high-level, unifying representational scheme for geographical phenomena and a unique path to describe the fundamentals of GIS [2,3]. Since the world is perceived as a set of material entities possessing highly correlated structures, a comprehensive model is needed to embody all types of spatial objects and the relations among them. Therefor, representation of geographic phenomena becomes a fundamental issue in developing a GIS application.

A feature-based approach models the geographical features while the layer-based approach models a map or a set of thematic maps. In the context of a feature-based GIS, geographical features are represented according to geography of locations, plus the complex interrelations of the human and physical characteristics of the locations [8].

One method of implementing the feature-based approach in GIS is to develop the model using object-orientation. The resulting feature model can include human views of geographical reality, spatial relations among geometric objects, and feature-feature relations among non-spatial objects.

In this paper, the intension of a feature's concept is extended. Features, as the fundamental depiction of geographic phenomena, encompass geometric objects, geographic entities and graphic symbols. A feature-based model, TLDM (Three-Leveled Data Model) is described, that represents geographic phenomena from three levels: geometric, geographic and graphic, among which the interrelations are analyzed in details. Lastly, the development of an object-oriented prototype feature-based GIS, YH-GIS, is described briefly.

## 2. People's three viewpoints on Map

Maps are the primary media of geographical information and the elementary objects manipulated in GISs. From the maps sculpted on clay tablets by Mesopotamian 5000 years ago [6] to today's more and more atlas, there has been a long history for humanity to use maps, which play a very important role in our work and lives.

Maps are graphic representations of geographic surfaces on plane, generalized by mathematical rules and represented by visual symbols for different purposes. Maps imply the distributions, states and associativities of diverse natural or social phenomena. Every map contains three aspects: (a) Mathematical Structure. There are rigorous mapping relations between the points of the geographic surface and those of the map plain. Taking advantage of the projections, we can research the spatial relations between geographic phenomena represented by map. (b) Proprietary Symbol System. Map adopts special graphic symbols, which enable spatially locating, to represent geographic phenomena and interrelations of the human and physical characteristics of the locations. (c) Generalization. The generalization of geographic phenomena are determined by the purpose and the resolution (scale) of map, of which some irrelative real phenomena are neglected in the course.

People read, understand and use maps from three different viewpoints. First, geometric object-based viewpoint. Map is considered a plane composed of fundamental geometric elements, such as nodes, chains, and polygons, between which are there certain topologic relations, i.e. meeting, containing, disjoint, overlapping. Only location coordinates, no geographic semantic attribute, are directly associated with nodes, chains and the boundaries of polygons. The topologic relations are kept unchanged under transformations such as rotating, translating, scaling. For example, a topologic relation of a chain lying in a polygon is kept true in any transformation of the map, only if the two elements are still existing. Second, geographic entity-based viewpoint. Map is thought as a collection of diverse geographic phenomena, which have explicit classification attributes, such as tower, road, river, park, and city, etc. Last, graphic symbol-based viewpoint. Map is regarded as a set of graphic symbols, which are separated into three kinds: area-, line- and point-pattern, possessing different colors, styles, sizes, widths, brightnesses, etc. The graphic symbols reflect the cognitive stipulations between the cartographers and map-readers to exchange spatial information. For examples, area-pattern symbols with different colors and filling styles stand for vegetation or population density of one field, line-pattern symbols with different colors and widths for rivers, point-pattern symbols with different brightnesses for cities or towns.

## 3. Feature-based concept and TLDM

In a feature-based GIS, features are the fundamental representations of geographic phenomena. The U.S. National Committee for Digital Cartographic Data Standards and the U.S. Geological Survey define a feature as two-fold, both a real world geographic entity and digital/graphic representation [10,4,8,9,7]. According to people's three viewpoints of reading, understanding and using maps, we extend the definition of a feature's concept as three-fold: geographic entity, geometric object and graphic symbol, as shown in figure 1.

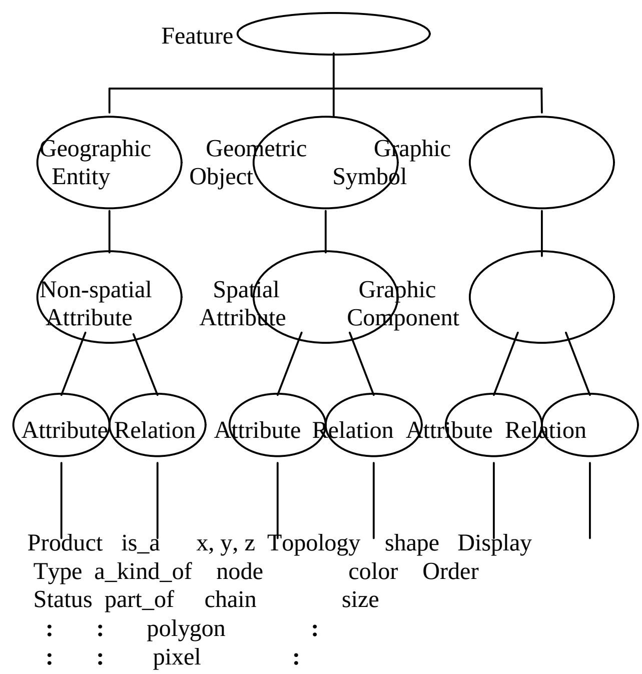


**Figure 1 Representation of a feature**

A geographic entity, which defines a geographic phenomenon, shares common attributes and relations. For example, a *road* object as a class comprises a set of real world road entity instances associated with common attributes and functions [7]. It may have name, type, number of lanes, starting point and terminating point, etc. A geometric object represents the spatial attributes of the feature, such as location and shape. For example, a *shop* can generally be represented as a node, a *river* as a chain, and a *park* as a polygon. The most important usage of a map is its readability. A graphic symbol object is a visual presentation of a geographic phenomenon to the users. For example, a *town* can be displayed as a red, green, or blue area, an icon or a dot, and can be neglected in certain resolutions (scales) of the map.

TLDM (Three-Leveled Data Model) is an object-oriented data model, which represents maps from three levels: geometric, geographic and graphic, according to people's three viewpoints on maps. It separates GIS data into three categories correspondingly, and defines their interrelations, as shown in figure 2.

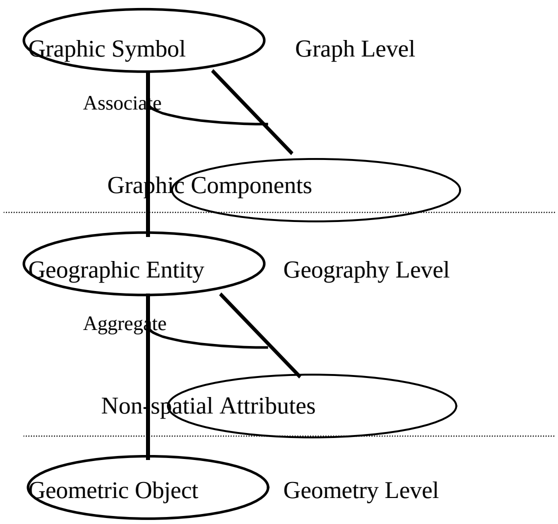


**Figure 2  Three-Leveled Data Model (TLDM) of map**

## 3.1 Geometric Model

Geometry model simulate a map as a set of geometric objects, which possessing locations, shapes, topologic relations and geometric sizing information, but no concrete geographic semantic meaning. Geometric objects are divided into nodes, chains, and polygons, whose interrelations are shown as figure 3.

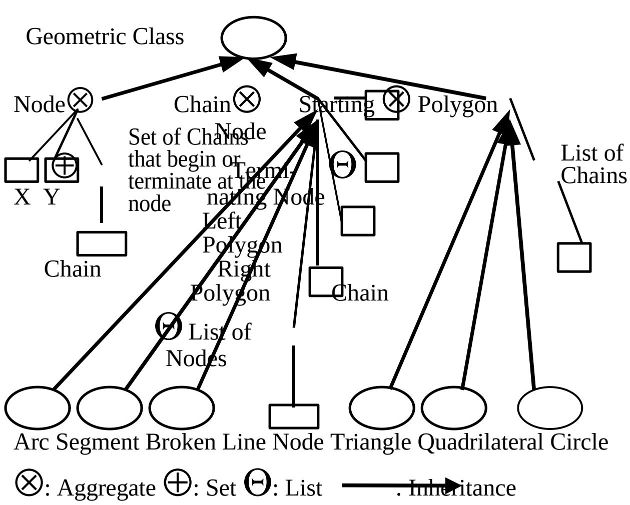


**Figure 3  Geometric objects and their interrelations**

Geometric objects encapsulate all spatial operations, such as determining that whether two chains intersect, or one node lies within another polygon, and computing the common parts of two polygons. Following key phases are often used to describe the topologic relations between geometric objects: disjoint, meet, inside, contain, equal, cross, overlap.

TLDM define a node as an isolated point, or terminals of a chain, a chain as a serial of coordinates with an explicit direction, and a polygon as a 2-dimensional connected field without holes. Concrete geometric classes associated with an application can be defined as their subclasses, as shown in figure 4.

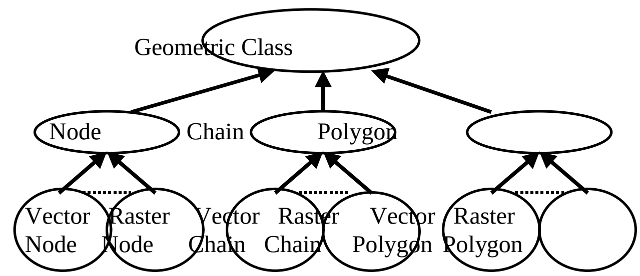


**Figure 4  The implementations of geometric classes**

GISs generally organize data according to thematic layers (maps), among which is there no interrelation. Most applications (for examples, urban planning, and evaluating of flood, etc) involve the spatial operations of data from many thematic layers, such as overlaying. These operations are often performed in the course of spatial queries or spatial analyses, resulting in the defficiencies of the application systems. TLDM establishes only one geometric layer to store all geometric objects for a map or several thematic layers, and each map or thematic layer just preserves its geographic entities' references to geometric object identifiers. The spatial relations between two geographic entities are inferred from the topologic relations of corresponding geometric objects, irrelative to the structures or representations of the geometric components. Therefore, the topologic consistence of the spatial database is assured, and the storage space is reduced, and the efficiency is improved.

Integrating vector and raster to result in a unified spatial data structure has been the dream of GIS researchers for a long time. With object-orientation, TLDM can integrate vector and raster structures in geometric object model, enabling the representations and topologic relations of geometric objects are independent of their implementations. TLDM provides the approach of integration at model level. In one hand, it supports a variety of geometric classes; in another hand, it supports the implementations of vector and various tesselations, such as raster, quadtree, octant tree, and 2-d run length encoding, etc.

## 3.2 Geographic Model

In geometric model, there is no definition or interpretation for the geographic meanings of the geometric objects, in order to preserve their independence, and to manage and manipulate them and their topologic relations in the unified geometric layer. Geography model provides geographic definitions and interpretations for geographic entities. Aggregating geometric object identifiers and non-spatial attributes, a geographic entity is connected with the correspondent geometric objects. A geometric object can be aggregated into different geographic entities. Thus, the geometric layer can be mapped to several maps or thematic maps with the same resolution (scale).

The geographic entities in real world can be separated into points, lines, areas, and complex entities. A complex entity is a set of the former three entities. The mapping relations of geometric objects and geographic entities are shown as figure 5.

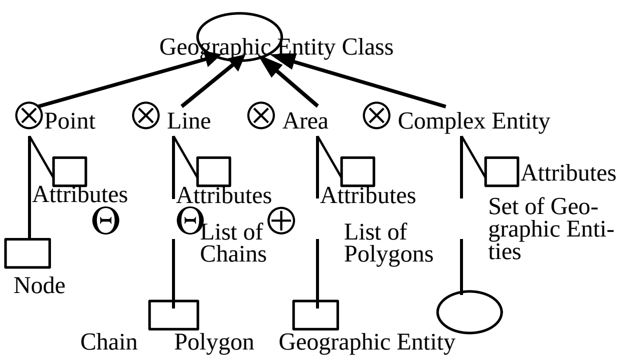


**Figure 5  The relations of geographic entities and geometric objects**

Points, lines, areas, and complex entities, which are called system classes, just abstractly define geographic phonenma in real world. Developers can implement concrete geographic entity classes according to the purpose of application, such as *well, shop, road, river, street, park, lake, university* and *city* etc, by defining them as the subclasses of system ones, as shown in figure 6.

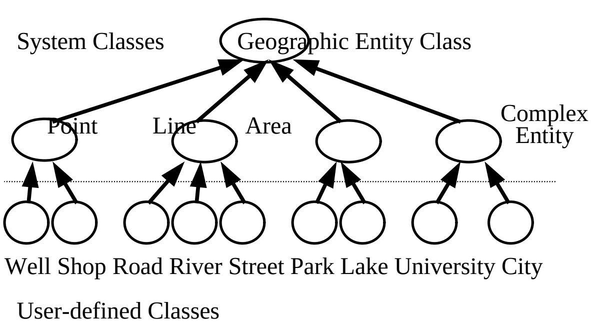


**Figure 6  The hierarchies of geographic entities**

## 3.3 Graphic Model

Map is the main representational form of data outputted by GISs, which is composed of diverse visual graphic symbols. In a conventional map-interpretation problem, each user determines map features through a combination of geographically portrayed symbols and personal cognitive knowledge. TLDM divided map symbols into three classes: dot-, line- and area-pattern. Users can define special symbols for various purposes as their subclasses, as shown in figure 7.

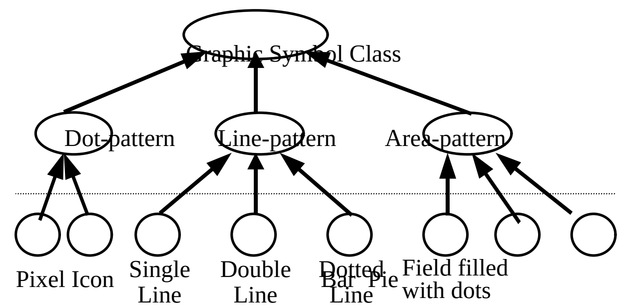


**Figure 7  The hierarchies of graphic symbols**

A geographic entity can be plotted as different graphic symbols in the light of needs. For example, a road may be plotted as a single, double or dotted line, in red, blue, or gray color.

Under a certain resolution (scale) of map, the appearance of a graphic symbol is associated with the geometric objects of the geographic entity defaultly. For example, line entities are usually plotted as line-pattern symbols. In another hand, the color, brightness, style, and size etc of a graphic symbol are associated with the non-spatial attributes of the geographic entity through association functions. For instance, let *TCountyEntity* be a class of geographic entities of administrative counties, *TArea-Pattern* a class of area-pattern graphic symbols with filling colors, then their association function on color can be defined as:

```
void Associate(TCountyEnttity county,
        TArea-Pattern area-pattern)
  { area-pattern.FillingColor =
      (long) county.Population/county.Area(); }
```

That is to say, the population density of the county determines the filling color of the area-pattern symbol.

# 4. The Organization of Map in YH-GIS

YH-GIS is prototype GIS that is realized with MS-Windows 95 and Borland C++ 4.5 on microcomputer. Based on TLDM, it stores spatial and non-spatial data into an object-oriented geographic database system, OOGDB [11,12], supporting the unified approach to integrating vectors and tesselations.

YH-GIS organizes and displays maps from three levels: geometric, geographic and graphic, as shown in figure 8. Adopting spatial index technologies, YH-GIS does its best to store all geometric objects of the same range into a unified geometric layer. Based on the geometric layer, some thematic maps can be established, and different graphic layers would be constructed through associating the graphic objects and the non-spatial attributes of geographic entities. Because the spatial data from geometric layers and non-spatial data from geographic layers are shared highly, the redundancy of the database is greatly reduced, and the inference of spatial relations among geographic entities, the spatial analyses of thematic layers, and the maintenance of integrity and consistence of the database are becoming more easy. With regard to users of YH-GIS, the geometric layer is transparent to them, and they just interact with geographic and display layers directly.

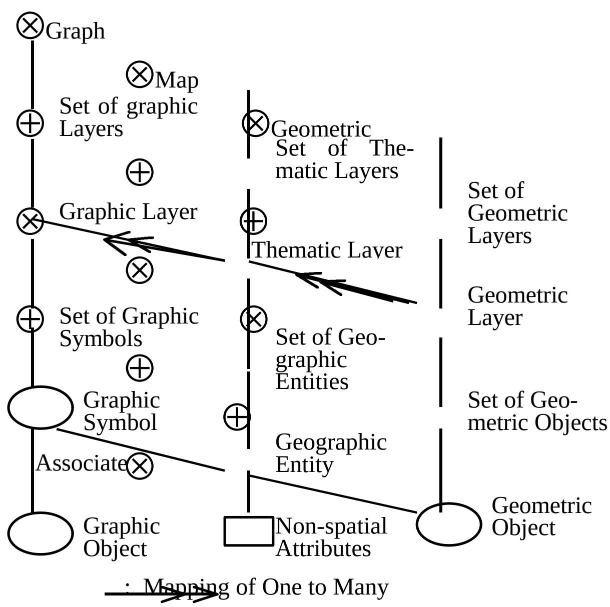


**Figure 8 The organization of maps in YH-GIS**

## 5. Conclusion

One may ask why we need features or to model features in GIS. A feature is essential to represent a geographic phenomenon because it can encompass all aspects of information about the phenomenon [9,7]. In TLDM, the non-spatial attributes are directly attached to the identifiers of geographic entities, and topologic relations to those of geometric objects. Geographic entities are displayed according to their non-spatial attributes and the map's resolution (scale). This can effectively describe the characteristics of the feature. Leveraging object-orientation, associating operations with the feature can implement what the feature does. Digital representation of a feature is not limited to one data structure. A feature can be realized in vector and raster. Geographic information is always so complex that a higher-order abstraction is required for modeling the reality. TLDM provides a means for representing the products of geographic abstraction.

## References


1. Abiteboul, S. and Hull, R., "IFO: A formal semantic database model". *ACM Transactions on Database Systems*, 1987; Vol.12, No.4, pp.525-565.
2. Chrisman, N.R., "Fundamental principles of geographic information systems". *Proceedings of 8th Int. Symp. On Computer Assisted Cartography, Auto-Cart 8*, Mar. 29-April 3, 1987, Baltimore, Maryland, USA. pp.32-41.
3. Goodchild, M.F., "A spatial analytical perspective on geographic information systems". *Int. J. of GIS*, 1987; Vol.1, No.4, pp.327-334.
4. Guptill, S.C., Boyko, K.J., Domaratz, M.A., Fegeas, R.G., Rossmeissl, H.J., and Usery, E.L., "An enhanced digital line graph design", *USGS Circular* 1048. Reston, Virginia: USGS, 1990.
5. Peuquet, D.J., "A conceptual framework and comparison of spatial data models". *Cartographica*, 1984; 21:66-113.
6. Robinson, A.H., Sale, R.D., Morrison, J.L., and Muehrcke, P.C., "Elements of cartography", fifth edition. New York, John Wiley and Sons, 1985.
7. Tang, A.Y., Adams,T.M., and Usery, E.L., "A spatial data model design for feature-based geographical information systems". *Int. J. of GIS*, 1996; Vol.10, No.5, p.643-659.
8. Usery, E.L., "Categogory theory and the structure of features in geographic information systems". *Cartography and GIS*, 1993; Vol.20, No.1, p.5-12.
9. Usery, E.L., "A features-based Geographic information system model". Photogrammetric Engineering & Remoting Sensing, 1996; Vol.62, No.7, p.833-838.
10. USGS, "Spatial data transfer standard", Version 12/90. Department of the Interior, U.S. Geographical Survey, National Mapping Division, December, 1990.
11. Wang Feng, *Object-oriented design principles, methods and implementations of geographical information systems* (in Chinese). Ph.D. thesis, Dept. of Computer Science, National Univ. of Defense Technology, 1996, Changsha, P.R.China.
12. Yang Shuqiang, *Researches and implementations of TLDM and OOGDB* (in Chinese). Ph.D. thesis, Dept. of Computer Science, National Univ. of Defense Technology, 1996, Changsha, P.R.China.