IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 1, January 2011
ISSN (Online): 1694-0814
www.IJCSI.org

Triangular Pyramid Framework For Enhanced Object Relational Dynamic Data Model for GIS

Ms. Barkha Bahl[1], Dr. Navin Rajpal [2] and Dr. Vandana Sharma[3],

[1]Research Scholar, GGS IP University,Delhi, India

[2]Professor, GGS IP University,Delhi, India

[3]Dy.Director General, National Informatics Centre,Delhi, India

ABSTRACT

GIS (Geographic Information System) data modeling is a methodology for designing spatial databases. Efficient database design is crucial to efficient GIS implementation. A model is a simplification of the complexity of reality, which makes the reality more understandable and operational.

The Objective of this paper is to present an overview of various data models and to evolve an approach to design a model which can overcome the identified gaps of various models being studied for a GIS application. The proposed data model i.e. Enhanced Object Relational Dynamic Data Model enhances Object Relational Data Model (EORDM) by introducing reusability and dynamicity for the database . Reusability concept allows storage of distinct objects having same spatial representation onto a single storage space and dynamicity allows creation of relations on user's requirement. The conceptual data model for the same has been designed named Triangular Pyramid Framework which has been explained in section IV of this paper.

Keywords: *Dynamic database, Spatial database,reusability,vector database, object relational database.*

1. Introduction

A geographic Information System, GIS, is a system for capturing, storing, querying, analyzing and displaying geospatial data **[12]** .This technology has the capability of spatial searches, overlays, association of spatial data with non spatial to generate new information. These features make this technology different from the CAD systems and conventional database applications. The CAD application becomes GIS only when an image becomes georeferenced. "Every object present on the Earth can be geo-referenced", is the fundamental key of associating any database to GIS.

In the Geographic Information Systems (GIS) community the concept of data modeling is well developed and integrated in the design of databases and related applications. Various data models for GIS applications have been introduced Viz. Multi-Dimentional location referencing system data model**[10],** Multiple representations object data model **[5],[9].** Spatio– temporal data model **[14]** etc .

Multi-Dimensional location referencing system data model allows organizations to implement improved solutions for transportation systems using advanced technologies.

Multiple representations, object data model represents same object more than once at different abstraction levels tailored towards the need of different users or analysis. Another data model i.e.Spatio-temporal model has been developed using object-oriented concepts. Spatial objects have both spatial and temporal dimensions. The geometric attributes and relationship change over time in real world objects. Hence , these aspects have been covered in temporal model .

Conventional GIS Systems represent the real world in a snapshot only ,which is inadequate for analyzing changes and patterns of change over a period of time. Every update creates a view map and no attempt is made to establish any link between the snapshots. These systems are therefore not suitable for many applications where data need to be interpreted in the context of time, such as urban application, environmental monitoring, agricultural applications, forest management, etc. **[15].**

Varying approaches have been used in the design of spatial data models, but no model or abstraction of reality can represent all aspects of reality, as it is impossible to design a generic data model. For example implementation of spatial data model are good for plotting digital environment, but are inefficient for analytic purposes and further for producing graphics. Secondly , it has been observed that in Survey of India, data capture in the digital domain is done in the dgn format with the help of microstation based softwares. The database of each colour separate is generated separately, in which all the cartographic elements, i.e. points, lines, areas and text are present in the same digital file. Unfortunatey in .dgn file interrelationship between the basic cartographic elements do not exist as it does not have any topology built into it [**14**]. Department of Science & Technology, Survey of India, Department of Space and GSI are coordinating to bring out an exchange format which will cater to the exchange of Vector, Raster and GIS data.

Bouille approach includes all identifiable entities and their relationships in deriving the phenomenal structure for phenomenal based design which is extremely complex**[2].** Whereas Mark adopted a philosophy of including only those entities and relationships that are relevant for an application for designing a data model which is known as minimalist approach and has minimum complexity**[11].**

The data model is more robust and flexible if it represents reality more perfectly i.e. all entities and possible relations are incorporated while designing the database. However incorporating essential entities and relations for single application a model will be considered more efficient with respect to storage of space and ease of use. Thus, the selection or design of data model must therefore be based both on the nature of the phenomenon the data represents and the specific manipulation processes which will be required to perform on the data. In this paper , we propose a Triangular framework for enhanced object relational dynamic data model that provides these and other features.

The rest of this paper is organized as follows: Section 1 covers the survey of various GIS data models. Section 2 discusses proposed data model and new features introduced in the same. Section 3 describes Triangular Pyramid Framework for the proposed data model. Section 4 concludes the paper.

Survey of GIS Models

Data models can be vector data model or raster data model. .Vector data represents discrete feature and raster data represents continuous features. This section will discuss various Vector Data Models viz. The Spaghattie Model[4], Topology Model[19], Polyvrt[13]**,** Relational[18], Georelational [16], Object based [1]and Geodatabase [20] etc.

1.1 The Spaghettie Model

In Spaghettie a digital cartographic data file is constructed referred to as a Spaghattie file which is a collection of coordinates, strings heaped together with no inherent structure**[4].** This model is inefficient for most types of spatial analysis, since any spatial relationships must be derived through computation. On the other hand the lack of stored spatial relationships, which are extraneous to the plotting process marks the spaghetti model efficient for reproducing the original graphic image. Thus they are used for simpler forms of computer assisted cartographic production.

1.2 Topological model

In topological model the information allows the spatial definitions of points, lines, and polygon type entities to be stored in a non-redundant manner**[19].** The GBF/DIME (Geographic Base File/Dual Independent Map Encoding) model is built upon the topological concept. The model represents a directed graph,in which an explicit direction is being assigned by recording a from-node and to-node which automatically check for missing segments and other errors in the file. In this model the basic logical entity is a straight line , where street, river, etc is represented as a series of straight line segment which are spatially defined. The main problem with this model is that the individual line segment do not occur in any particular sequence order, so to retrieve all line segments which define the boundary of a polygon, an exhaustive search must be done as many times as there are line segments in the polygon boundary.

1.3 POLYVRT

Peucker and Chrisman (1975) developed POLYVRT**[13].** This model had overcome the very major retrieval and manipulation inefficiencies seen in simpler topologic structures by explicitly and separately storing each type of data entity in a hierarchical data structure. It made distinctions between types of entities both logically and topologically meaningful, so that a chain is denoted as the basic line entity. It facilitates easy search and retrieval and there is partitioned storage. Leads to storage overhead (pointers) and integrity of pointers. It is a multipurpose database model.

1.4 Relational Data model

The power and elegance of the relational model stems from the fact that it uses a single construct, the relation **[18].** Five functional closed operations are defined in relations, namely, union, difference, selection, projection and Cartesian product.

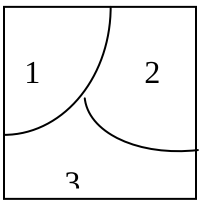
For spatial applications, however, the resulting representation is inadequate.

For example, if layers are represented with plain relation, operations such as overlaying and reclassification cannot be derived from the fundamental relational databases.

In the relational model these operations are hidden in the physical level. As a result important information is lost and the system is tied to some specific implementation. Thus relations are inadequate as the sole modeling construct for geographical applications.

1.5 GeoRelational data model

Data representation for GIS applications includes the spatial and attribute component **[16].** Spatial data describes the location of spatial features, whereas attribute data describes the characteristic of spatial features. The GeoRelational data model stores spatial and attribute data separately in a split system. Spatial data is stored in graphic files and attribute data is in a relational database. A GeoRelational data model uses the feature label or ID to link the two components as shown in figure below. The two components must be synchronized using some ID so that they can be queried, analyzed and displayed in unison.



Graphic Files

Polygon/arc list Arc-coordinate list Left/right list

INFO File		
Polygon-ID	Field 1	. . .
1		
2		
3		

Fig 1 : Georelational database design

1.6 Object Based Data Model

A fundamental requirement for spatial database design is the ability to model spatial properties, i.e., to associate parts of space with an attribute **[1].**

Parts of space are usually represented by points, lines and regions and are known as geometric features. Spatial applications deal with two, orthogonal and generalizations of spatial properties. One is association of the whole of space with an attribute and the other is associations of sets of attribute and geometric feature. The former is modeled with concepts oriented toward objects **[17]**. Object Based Data Model has been used as a means of conceptual structuring of geographic information. In particular it models real-world objects (or entities) with a precise and 'crisp' spatial location and extent.

The object based data model differs from the georelational data model in two important aspects. First, the object-based data model stores both the spatial and attribute data of spatial features in a single system i.e. an object rather than a split system. Second, the object-based data model allows a spatial feature to be associated with set of properties and methods. The same is represented in Fig 2 .

The shape field stores the spatial data of land-use polygons, other field store attribute data such as landuse-id and category.

Since both spatial data and attribute data is stored in a single system the problem of data synchronization is eliminated that is found in split system (Georelational datamodel).

Objected	Shape	Land use_ ID	Category	Shape_ Length	Shape_ Area
1	Polygon	1	5	14,607.7	5,959,800
2	Polygon	2	8	16,979.3	5,421,216
3	Polygon	3	5	42,654.2	21,021,728

Fig. 2 : Object based data model .

1.7 Geodatabase Data Model

This model is built on arc-objects. It uses the geometries of point, polyline and polygon to represent vector-based spatial features **[20]**. The data structure of geodatabase distinguishes the feature classes and feature datasets. A feature class stores spatial data of the same geometry type and its datasets store feature classes that share the same coordinate system and area extent. In this model, feature classes can be standalone feature classes or members of a feature dataset. The geodatabase constitutes a uniform repository of both spatial and attribute data in a single database system. Objects in the geodatabase can have behavior associated with them. Integration with object-oriented concepts and COM technology allows great level of customization and reuse of the model to create application-specifications, which may (fig. 3) provide the framework for interoperability. The main problem we encountered was the custom domains, but this was not investigated. In addition, although the ArcObjects Library is extensive, more functionality needs to be added to allow high level of customization.

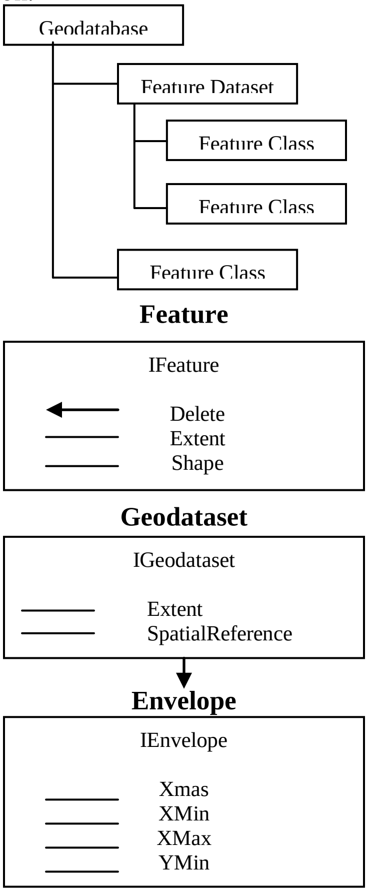


Fig . 3 Geodatabase Model

In a geodatabase , feature classes can be standalone feature classes or members of a feature dataset. A feature object implements the Ifeature interface. A geodataset object supports IGeodataset and an Envelope object supports IEnvelope.

1.8 Bi Level Data Model

There are two separate layers **[3]** in this model as mentioned below:

Higher level data model (Geographic object data model): This level consists of the geographic objects and a set of semantic spatial functions through which the topological relationships among objects can be defined or derived.

Lower level data model (Geometric object data model): This level consists of geometric objects which are actual spatial representations of the geographic objects. It also has a set of functions for retrieval, manipulation, computation of geometric objects. In this model, relationship between geographic and spatial objects is investigated, Spatial object is not PART-OF a geographic object , but is just a representation of the geographic object , similar to the mapping from one object to any other object(s). Where as Triangular pyramid model, the proposed model has three abstraction levels represented using three components – the object component, Geometric component and the location component. The details of the same have been discussed in the next section. In this model “uses” relationship type has been introduced, where the common reference table is used for representing the location component for various maps.

2. PROPOSED DATA MODEL – “ENHANCED OBJECT RELATIONAL DYNAMIC DATA MODEL”

An Object relational databases lie in between relational databases and object oriented databases. The approach is eventually that of relational database only.

The new construct added to the core functionality of traditional database includes user defined data types (for creating layers) and the ability to create the same dynamically through database programming interface. The reusability concept of Object oriented paradigm is incorporated using common reference location table. The basic selections, retrievals, manipulations have been included to the capabilities of the function associated with object classes.

For the rest of this section, we first introduce the features & assumptions. Features of our Enhanced Object relational dynamic data model are followed by the four modeling aspects of the proposed data model.

2.1 Features and Assumptions

Below are the features and assumptions, which we have considered for our new data model for GIS applications.

- All the Maps have the same extent. i.e. same (Xmin,Ymin) and (Xmax, Ymax)
- Each layer of the map is treated separately as virtual map sheets.
- All the maps have common reference coordinate table(x, y).
- All the spatial calculations use same scale of measurement.

- Distinct geographic objects having exactly the same spatial representation meaning occupying same physical space on the earth's surface will be represented by distinct geometric objects for their spatial representations.
- The data model is designed so that it is possible to have the dynamic creation of tables for user-defined layers, so that those data may be queried and visualized further.
- It defines the relationship between attribute and spatial data in a manner that is efficient with respect to space and time while executing complex queries.

2.2 Four Modeling Aspects

The purpose of this research paper is to develop a data model called enhanced object relational vector data model that better facilitates the access and storage of both spatial and attribute data than conventional GIS database models. The modeling aspect of this data model includes four concepts. These are ***vector data , dynamicity, reusability, and object-relational concept as*** discussed below:-

1) **Vector data**

   Vector data represents one major category of data managed by GIS **[12]**. It represents each feature as a row in a table and feature shapes are defined by x,y location in the space. Features can be discrete points, lines, and polygons.
   Points: - having a pair of coordinates. Locations such as the address of a customer or the spot of crime are represented as points.
   Lines: - series of coordinate pairs. Streams or roads are represented as lines.
   Polygons: - are defined by borders and are represented by closed regions. They can be defined such as parcel of land, counties, watersheds etc.

2) **Dynamicity**: - As we know, response time and storage space are the major concerns for GIS applications. The proposed model provides a solution for improving the performance of response time and storage space by creating tables dynamically. The dynamic database can be easily created as per the user's requirement **[7]**.
   The design of a dynamic data model increases the availability of spatial data to the users by providing the latest updated information **[6]**.

3) **Reusability**: In the proposed model different object classes could be defined using the "uses" relation type which introduces reusability aspect of object oriented model. The reusability has been incorporated through common reference location table. The details are discussed in section 4.

4) **Object relational concept**: In our model existing relational systems SQL server is enhanced in the direction of object orientation or in terms of front end interfaces by providing an object oriented flavor or outlook to a conventional relational system. Thus, both object and relational views are provided for designing the data model.

2.3 Triangular Pyramid Framework for EORDM

Unlike the data models being discussed, the proposed data model has three levels of abstraction . They are : The Object Component (Highest Level),The Geometric Component (Middle level) The Location Component (Lowest Level).

The model (fig. 4) covers the representation of whole information required by GIS application in three components: Object, Geometric and location. Diagrammatic representation have been given below where the mapping of the model with the database table have been mentioned:

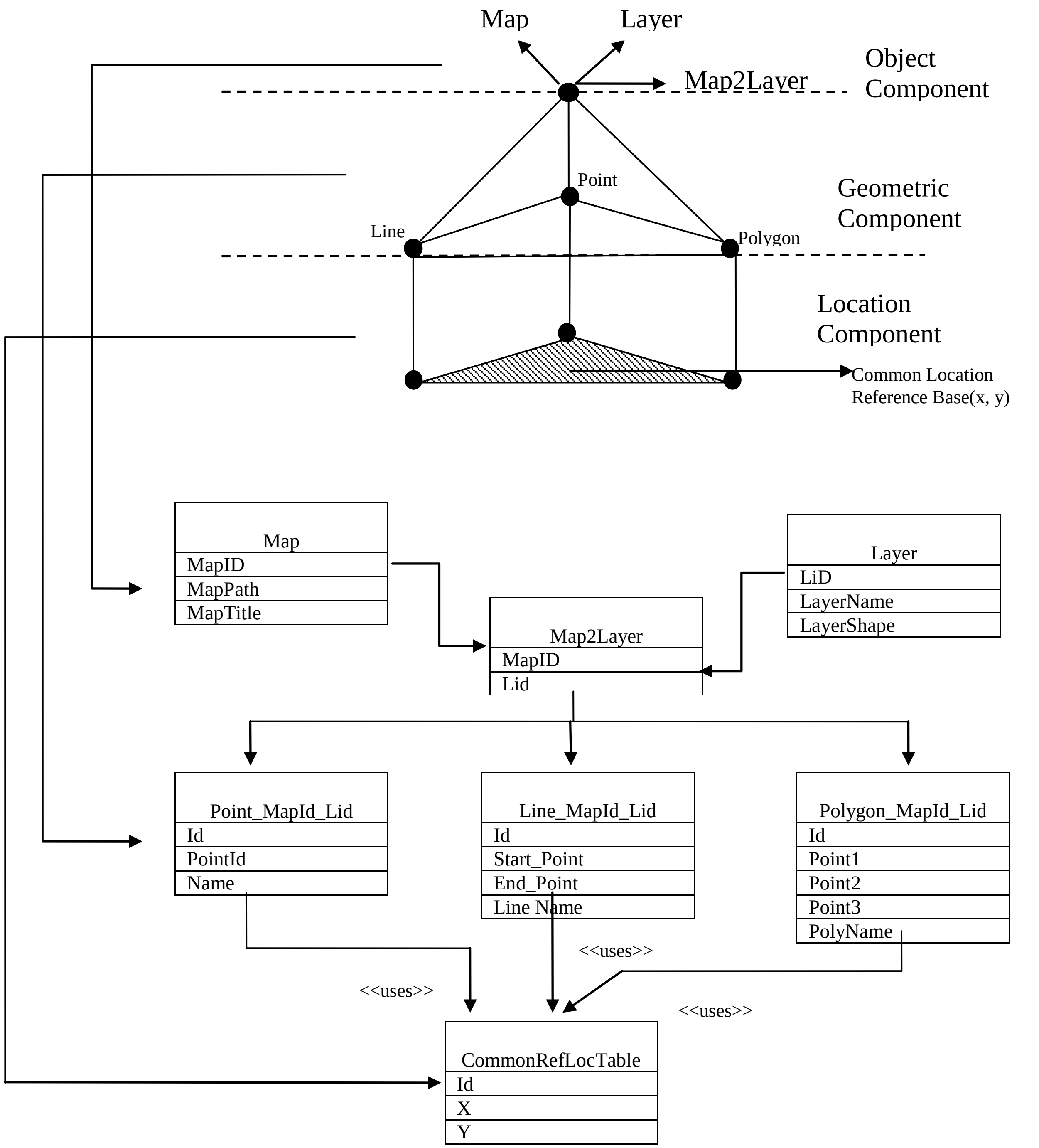


Fig.4 Triangular Pyramid Framework for EORDM

The Object Component (Highest Level)

A map is a symbolic depiction highlighting relationships between elements of that space what we here call them as layers such as rivers, lakes, mountains, districts, forests, temples, Etc. Thus map is a combination of different types of layers. These maps and layers are called objects as they are real life entities having both attribute and behavior.

Thus object level is the highest level of abstraction, and at this level geographic information is represented by layers representing the relative position of spatial objects.

The Geometric Component (Middle Level)

The Geometric component is the middle level as it is the interface between the object and its actual spatial existence on the earth.

Each geographic object in the higher level has its corresponding geometric object. The information at geometric level represents the shape of the geographic object, which is categorized into three: - point, line polygon.

The first is point data where each object is associated with a single location, Example a city, district, school, hospital etc.

The second is line data where the location is described by the string of points, Example: - road, river, railway line etc.

The third is polygon data, where the location of object is represented by a closed string of coordinates. They are thus associated with areas over define space, Example: - mountains, parks, wasteland area etc.

Distinct geographic objects may have same physical representation, for an instance Roads and rivers both are distinct geographic objects but are represented by same geometric object i.e. line.

Thus In geometric level, the internal information represents the geometry of the object keeping this information hidden from the object level.

The Location Component (Lowest Level)

The lowest level of the proposed data model is Location Component which represents the actual screen coordinate value of the geometric object in middle level.
In the proposed model, we are considering that the maps have same extent i.e. they have same (Xmin, Ymin) and (Xmax, Ymax) screen coordinates. If we have different maps having same map extent on the screen and corresponding to them layers and their screen coordinates are stored. It may happen that many of the layers on different maps have same value of (x,y) coordinate on the screen.

For an example consider maps of four different cities viz: .India, South America, North America, Asia. With an assumption that they have same extent on the screen, layers for these maps are drawn .It may happen that for the screen coordinates (100,200), India and South America may have a same point .Corresponding to (275,145) we have points in South America, North America, Asia. So if we store (100,200) twice and (275,145) thrice ,then it will be a redundant effort of storing the data in the database table., which leads to both wastage of space and time, hence a common reference coordinate table has been developed, which contains all the coordinates that have information corresponding to them in any of the maps.

The main concern in GIS applications is the response time and in turn the response time is dependent on how efficiently underlying database model is designed.
The proposed data model is designed keeping the same in mind. This leads to development of a database model which is dynamic and at the same time it uses the concept of reusability.

3 DDL FOR TRIANGULAR PYRAMID FRAMEWORK. The above mentioned model has been realized using following create command(DDL).

DDL for the proposed model

Table creation :

```
CREATE TABLE MAP
(
      Mapid : int <<PK>>
      Mapname : varchar(50)
       Mappath : varchar(200)
);

CREATE TABLE LAYER
(
 Layerid : int<<PK>>
  Layername : varchar(50)
  LayerShape : varchar(50)
);

Create Table Map2Layer
(
   Id : int
   MId : Int <<PK>><<FK>> References MAPID(Map)
    LID : Int <<PK>><<FK>> References Layerid(Layer)
);
CREATE Table Commonreference Table
(
     LocID : Int <<PK>>
            X: Int <<UNIQUE KEY>>
            Y: Int <<UNIQUE KEY>>
);
```

DDL for the EORDM.

Fig (4) represents a schema used to model GIS applications consisting of Roads , national highways , parks, waste land areas, schools and hospital . Each of these are real world entities/objects being represented by different layers. Layers are functionally related map features. Features types (point, line or polygon) and feature themes (What the feature represents) are the two most common considerations for organizing data layers. In our schema Roads and national highways are both lines (feature type) , but they have to be stored in separate layers , because of different attributes(feature theme).

SCHOOLS (*Name: String* , *P: Point*)
HOSPITALS (*Name: String* , *P: Point*)

ROADS (*RoadNo: integer* , *Length: Real* , *L: Line*)
NATIONALHIGHWAYS (*Name: String* , *Length: Real* , *L: Line*)
PARKS (*Area: Integer* , *P: Polygon*)
WASTELAND (*Area: Integer* , *P: Polygon*)

Fig (4) : GIS Schema

Select *Map2layer.Lid*, *Layer.layerName* , *Map.Mapname*
From *Map2layer,layer.Map*
Where *Map2layer.Mid* = *Map.MapID* and *Map2layer.LID* = *Layer.LID*
And *Mapname* = *'Delhi'*

Sample queries for the proposed model

Query 1 : All the layers of Delhi Map

Query 2: **Select total wasteland area of UP Map**

Select Sum(*UP_Wastelands.Area*) as *TotalArea*
From *UP_Wastelands*

Query 3: Select No of Hospitals in Maharashtra

Select count(*) as *NoOfHospitals*
From *Maharashtra_Hospitals*

Query 4: Display (X,Y) coordinates of Delhi Schools

Select *CRT.X* , *CRT.Y* , *CRT.Name*
From *CommonRefreneceTable* as *CRT* ,*Delhi_Schools* as DS
Where *DS.PointId* = *CRT.XY_Id*

CONCLUSION

In this paper, a new data model named Triangular Pyramid framework for enhanced object relational dynamic vector data model has been introduced for representing the complete information being required for representing the data for GIS based application.

Here , we have suggested three layer components namely object , geographic and geometric. The details of the same have been explained in the privious section. Besides, we also have specified DDL for implementing the proposed model . With this proposal , we have showen that the layer/component architecture of EORDM is more suitable for modeling GIS application . In the dynamic database less storage space is required as the instances of the databases are created at the time of requirement only.Secondly , the storage space and response time is reduced since the common reference table is being introduced to handle data efficiently.

REFERENCES

[1] Banerjee, J.,1987, "Data Model issues for object-oriented applications", ACM Transactions on office information system 5(1) pages 445-456, 1987.

[2] Bouille, F. 1978. "Structuring Cartographic data and Spatial Processing with the hyper graph-based data structures" Proceedings of the First International Symposium on Topological Data Structures for Geographic Information Systems, Harvard University, Cambridge.

[3] Choi Amelia , Lub.w.s, 1991 " A Bi-level object oriented data model for GIS Application ", IEEE,0730-157/90/0000/02385/0,pp 238-244

[4] Dangermond .J.1982," A Classification of software components commonly used in geographic information systems", proceedings,US, Australia workshop on the design and implementation of computer-Based Geographic Information System, Honolulu, Hawai.,PP 70-91,

[5] Egenhofer , J.M.,Clementini, E., Di Felice P. 1994, "Evaluating inconsistency among multiple representations.", Sixth International Symposium on Spatial Data Handling ,Edinburgh, Scotland, pp. 901-920.

[6] Huang, B. 2003 . "An object Model with Paramteric Polymorphism for Dynamic

Segmentation." International Journal of Geographic Information Sciences 17:343-60

[7] Ka-Wai Kwan and Wen-Zhing Shi,2002, "A Study of Dynamic Database in Mobile GIS", Symposium on Geospatial Theory, Processing and Apllications.Pg 1 -10

[8] Koncz,N.A. and T.M. Adams 2002."A data Model for Multidimensional Transportation Applications.", International Journal of Geographic Information Sciences 16:551-70

[9] Kilpelainen , T. ,2000, Maintenance of multiple representations database for topographic data. The Cartographic Journal, Vol. 37, No.2 , pp. 101-107.

[10] Newell . R.G and Sancha T.L,1990 "The difference between CAD and GIS ",Computer – Aided Design , Volume 22, Issue 3, April 1990, Pages 131-135.

[11] Mark, D.M. 1975." Computer analysis of topography: a comparision of terrain storage methods", Geografiska Annaler, vol . 579, pp. 179-188.

[12] Oliver Kersting and J¨urgen D¨ollner, Nov2002, " Interactive 3d visualization of vector data in gis.", In Proceedings of the 10th ACM International Symposium on Advances in Geographic Information Systems (ACMGIS 2002), ISBN 1-58113-591-2,107-112

[13] Peucker, T. and N. Chrisman, 1975." Cartographic Data Structures", The American Cartographer, Vol. 2, pp 55-69.

[14] Perumal . A. "GIS Applications for resources management-National Initiatives", Natural Resources Data Management System, Department of Science and

[15] Raza, A., 2001, "Object Oriented temporal GIS for urban applications". PhD Thesis, ITC Publication Number 79.

[16] Robinson, A.H.,J.L.Morrison, P.C.Muehrcke, A.J. Kimerling, and S.C.Guptill. 1995. " Elements of Cartography", 6th ed. New York: Wiley.

[17] Shi, W. , B. Yang and Q.Li.2003 "An object Oriented Data Model for complex objects in Three Dimensional Geographic Systems". International Journal of Geographic Information Sciences 17:411-30

[18] Smith, Karen and Zdonik Stanley, 1987, " Intermedia: A case study of the differences between Relational and object-oriented Database Sytems.", In OOPS SLA 37 Proceedings, Pages 452-465,

[19] US Dept. of commerce, Bureau of the census, 1970. the DIME geocoding system, in report no. 4, census use study.

[20] Zeiler, M. 1999. "Modeling our world:" The ESRI Guide to Geo database Design. Red lands, CA: ESRI Press.