The Integration of Spatial and Non-spatial Data Model

Daut Daman, Harihodin Selamat, Shafry Rahim
Faculty Science Computer and Information System
University of Technology Malaysia
81310 Skudai, Johor
Tel: 607-5576160 ext 3764
Fax: 607-5565044
{ daut@mel.utm.my, harihodn@itp.utm.my, c-shafry@mel.utm.my}

Abstract – This paper will discuss about the integration of spatial and non-spatial data in a single data model for a Geographical Information System(GIS) database. In recent research findings for GIS data model, geographical data depends on time constraint. We propose the integration that include the time constraint using a double cube data model based on a feature-based model. A unique key is used to access the integrated database.

I. Introduction

The integration between spatial data and non-spatial data has become a major problem in most GIS applications. This is because spatial data is located is a separate database. To access the spatial data normally a unique key is required to identify data item In many research experiments several tasks like storing and managing the spatial and non-spatial data is done in one data model. The results showed that the samples of data can be accessed and integrated into one data model. Some of them are Geographical Information System Entity Relationship (GISER) [7], Tree Level Data Model (TLDM) [3] and GODOT [4]. The integration will allow users to manage the data easily, create the analysis [5,7], manage data continuously [7], generate data model efficiency and control the features changes [7].

In our research, the problem is to increase flexibility of the integration of data model that already exist. The goal is to develop a data model which is capable to store all the spatial and non-spatial data, spatial temporal and temporal data using a unique identifier

II. Related Work

Spatial data has been related with space and shape. In Geographical Information System (GIS), spatial data was defined as a data that represent space and location. Non-spatial data is an attribute data that stored spatial information. Geographical data is a combination of spatial and non-spatial data. In addition a time constarint is added into the data model. The time-change in GIS data is needed in some analysis that involves time constraint.

Currently, there are four methodologies to develop the GIS data model: (i) feature-based [3], (ii) object based [6,7], (iii) field based [6,7] and (iv) directional based[8]. Feature based is a geographical feature oriented method. Sample data that been taken using this method is Tree Level Data Model [3]. The result showed that this method is efficient to integrate data and data presented in the form of map. Field based is a method to define earth surface as one continuous space where at the top level there exists a geographical feature [7]. Object based defines the world as a combination of separated objects (e.g. rivers, mountains, and cities) which all of them have their own coordinates. This method is commonly used by most GIS researches (e.g. GraphDB, GODOT, GeoOOA, OGIS [6]). However, the result showed that this method has a drawback ie. it is unable to support continuous space data.

The integration of spatial and non-spatial data has been a great issue. Several techniques have been suggested for the integration such as the Forward Link, Backward Link and Bi-directional Link [9]. In some integrated GIS data model the use of bi-directional link is better than other two links because its query processing of spatial and spatial data is much easier.

III. An Extended Feature Based Method

The spatial and non-spatial data integration model is focused on Feature-based methododology. Will will first define the feature-based method to represent geographic phenomena. Feature is a set of points, lines or polygons in a spatial database that represent a real-world entity. The term's feature and object are often used interchangeability. The U.S National Committee for the Digital Cartographic Data Standards and the U.S Geological Survey define feature as real world geographic entity and digital/graphic presentation [3]. Latest research [3] suggested that feature is defined as geographic entity, geometric object and graphics symbol.

In this research, feature definition has been extended where time element has been added. With this new addition, there is a parallelism changing between feature, geographic phenomena and human's activities. As an example, we can apply the new model by the changes made by Jabatan Ukur & Pemetaan Malaysia[JUPEM] in producing new release and up-to-date map.

Previous feature-based model is defined as three folds: geographic entity, geometric object, graphic symbol. We now add a time entity to the model so the new model. The new model is now defined as four folds: geographic entity, geometric object, graphic symbol and time entity. Geographic entity is an entity that describes geographic phenomena. It is a non-spatial data attribute and time series data or temporal data. Geometric entity is an element that explains feature form whether it is spatial data or non-spatial data. Vector is represented by coordinates (x,y,z) or longitude and magnitude or other mapping systems (e.g. river is represented by blue color). Time entity is an entity that will record the time for existing and changing feature.

IV. The Cube System

Cube System is one method to manage temporal data or time series data. This method is been implemented in Malaysian Hydrological Information System (MHIS) [2]. Reports, graphs and hydrological analysis have been produced Cube data model. Conceptually, Cube is a 3D object. Each of its cell has coordinates (x,y,z). These coordinates (x,y,z) can be referenced to answer all questions related to hydrological analysis. They are referred to three questions (Where? What? When?). In other words, in each cube has three references (FID, DID, TID).

The Cube system can also be used to manage spatial data and to record any changes of the data. The difference between existing cube and the new model cube is that a unique identifier is used to refer both apatial and non-spatial data. The Cube model has been extended to include reference coordinate: (FID,GID,TID). The role of FID and TID is similar to the existing cube while GID is reference to the spatial data.

V. Double Cube Data Model

Double Cube Data Model (DCDM) is proposed and developed based on feature-based method and the Cube system. Figure 1 shows DCDM data model. The first cube is use to manage spatial data while the second cube is used to manage the non-spatial data. In each cube, there are changes of spatial data process (spatiotemporal data) and changes of non-spatial data (temporal data). The Cube coordinates (FID,TID,GID) are reference for any searching of data by using a unique key FID (a key for feature).

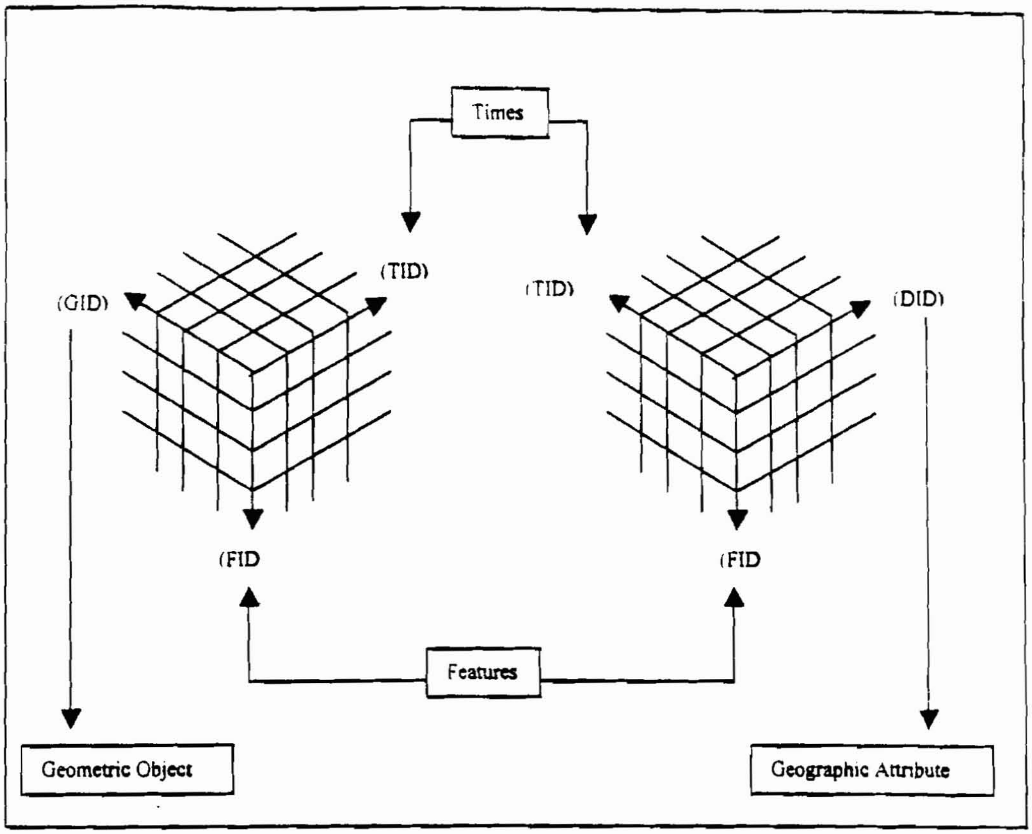

Figure 1. Conceptual Model of DCDM

DCDM can be tested and implemented using relational data base engine. By using this model, the process of accessing data is based on SQL statement. Below is an example of SQL statement:

Example:
How much average of the rainfall in a region A at 1971 at station A.

Query Process:

{
SELECT FID, avg (Rainfall)

FROM [Time] INNER JOIN (Feature INNER JOIN Cube1 ON Feature.FID = Cube1.FID) ON Time.TID = Cube1.TID

WHERE DID= 'St A' and Year = 1971
}

Output:

FID for The Station and average rainfall

Input :

FID For The Station

Query Process:

{
SELECT FID, GID, Color/Symbol , NoSegment, x1, y1, z1, x2, y2, z2,

FROM Cube2 INNER JOIN Geogmetri ON Cube2.GID = Geogmetri.GID

WHERE FID= FID For The Station
}

Output:

Shape of feature A and their location

The representation of spatial data is in the form of a map. The map has a set of points and symbols of color of the object that will represent feature information. A set of data can be produced based on layers upon clients' request. Below is an example of SQL statement to produce a sample set of data:

Input:

Layer Name/ Layer ID

Query Process:

{
SELECT FID, GID, Color/Symbol, NoSegment, x1, y1, z1, x2, y2, z2,

FROM (Feature INNER JOIN (Cube2 INNER JOIN Geometry ON Cube2.GID = Geogmetri.GID) ON Feature.FID = Cube2.FID) INNER JOIN [Time] ON Cube2.TID = Time.TID

WHERE Feature.Layer_Name = Layer Name;
}

Output:

Set of Data to generate one layer

VI. Conclusion

As a conclusion, this paper give the concept modeling an integrated data model for managing a geographical data. We have introduced a time entity for modeling an integrated spatial and non-spatial data that include the changes of the spatial (spatiotemporal) data related to time. The model is encompassing feature based method and cube system.

We have focused and tested our model with a complex query for making spatial analysis (e.g. calculate length of the line, calculate area, searching area). Special queries in hydrological analysis has been implemented in Malaysia Hydrologycal Information system (MHIS).

The proposed data model can also be implemented in other domain of applications such as survey, housing, transportation and site management.

References:

[1] Bellamy, S.P, *"MHIS Developtment, Database/ArcInfo Link-Heigh Level Design"*, Technical Report, Issue V1.R1.M0, p.p. 1-3, Jully 24, 1996 (unpublished)

[2] Claudia, B.M and Pires, F. *" Databases for GIS"*. ACM SIGMOD Record, Vol. 23, No.1, p.p. 107-115. 1994

[3] Feng, W. , Shuqiang, Y. , Huowang, C., Jichang, S., *"Spatial Data Model for Feature-based GIS"* , CSIT'99, Proceedings of 1st International Workshop on Computer Science and Information Technologies, p.p. 107-111, January, 1999

[4] Gunther,O., and Riekert, W. *"The Design of GODOT: An Object-Oriented Geographical Information System"*. IEEE Data Engineering Bulletin 16, p.p. 1-6, March (1993)

[5] Lee, J.Y, Jin, O.K and Ryu, K H, *"Integration with Spatiotemporal Relationship Operators in SQL"* , ACM GIS' 98 11/98 Washington, D.C., USA p.p 113-115 , Nov. 1998

[6] Shekhar, S, Chawla, S., Ravada, S., Fetterer, A., Liu, X. and Liu, C.T., *"Spatial Databases: Accomplishments and Research Needs"* , IEEE Transactions on Knowledge and Data Engineering, p.p. 1-15, Jan.-Feb. 1999.

[7] Shekhar, S. , Coyle, M, Goyal, B, Liu, D. R. and Sarkar, S., *"Data Models in Geographic Information Systems"*, Communications of the ACM , Vol. 40, No. 4, April 1997

[8] Shekhar,S. , Liu, X., Chawla, S., *"An Object Model of Direction and Its Implications"*, Kluwer Academic Publishers; A summary of results was among 5 best papers in Sixth International Symposium on Advances in Geographic Information System, GeoInformatica 3:4, pp. 357-379. 1999

[9] Walid G.A and Samet H. *"Spatial Data Model and Query Processing"* , W. Kim, Ed., Addison-Wesley/ACM Press, p.p. 338-360.1995.