AN UPDATING METHOD OF GEOSPATIAL DATA BASED ON TRAFFIC FLOW ANALYSIS

Y. H. Zeng [a,c,*], Z.H. Tian [b], J.M. Huang [a,c], S.S. Chen [a,c], K.W. Zhong [a,c]

[a] Gungzhou Institute of Geography, Guangzhou, 510070, P.R. China - silversoft@21cn.com
[b] Zhengzhou University,75 Daxue Road ,Zhengzhou , 40052,P.R. China - iezhtian@zzu.edu.cn
[c] Key Laboratory of Guangdong for Application of Remote Sensing and Geographic Information System,Guangzhou,510070, P.R. China

KEY WORDS: Network GIS , Traffic Flow, Data Updating

ABSTRACT:

The paper proposes a new kind of objected-oriented method of updating geospatial data, that is different from the previous “bottom up” or “top down” updating methods. Through weighted directed graph model and the selection of appropriate network parameters, the proposed method simulates the city traffic network through a light weight network GIS (Geographical Information System). Detailed analysis of the connectivity degree and traffic flow can be used to derive the topology changes in the city road network. Furthermore, such mapping can also be used to directly store and retrieve data from the lower level geospatial database, hence achieving the goal of automatic updates for the road networks. The application of classical graph theory algorithm makes it feasible to quantify and compute the relationship between traffic flow volume and network topology changes. Finally, we may map the changes to the geospatial databases through the interface for the updating operations. This method can also extend to other networks, such as the utility (water, drainage, power, etc.) networks.

1. INTRODUCTION

The data expansion, updates and management are challenging problems for geographic information systems (GIS). It is well known that the building and updating of the database is the bottleneck of the overall GIS development. Traditionally, the resource investment ratio of a GIS is roughly: hardware: software: database building = 1:10:100. Hence the database maintenance in the face of rapidly expanding data is one of the most important areas in GIS. Much development has been focused on the automated operation for database update. The majority of the updating methods can be classified into the following two categories:

(1) Bottom Up method: digitizing the smallest survey cell , including field surveys, GPS, photogrammetry and satellite remote sensing data. The comprehensive data sources make the spatial database complete. However, the global concurrency is often hard to guarantee.
(2) Top Down method: Update of the existing geospatial database is performed according to the end user requirement only. Hence the database may be updated only partially. For example, at certain hotspot or special situations. This method has the advantage of lower cost when compared to the bottom up approach. However, it does not keep the whole database complete and current

In (Shi,W.Z.,2005) proposed the direct and indirect geospatial data capture methods, which are similar to the above approaches.

However, the traditional method lacks the mechanism for automatic data updates, due to its sole reliance on the raw data input. In this paper, we propose an object-oriented method. It borrows a page from the computer software engineering’s OO method. Through introduction of objects and their associated methods, it can be used to describe how data are related. Such ontologies have proven to be a useful tool for data integration across heterogeneous data sets (Lu Feng , 2004).

Traditional maps are based on the real-world geographical model. When this model is extended to include 3-dimensional data, as well as the time-evolution of such data, the massive amount of data often renders such model ineffective, and can sometimes introduce unnecessary complexity into the model. In order to reduce the complexity and difficulty, we introduce a more abstract graph based model: a geographic graph model (GeoGraph) (Catherine Dibble, 2003). The network abstraction allows more targeted study of certain GIS network properties. For example, it has been shown that many GIS networks exhibits small-world or scale-free properties. In (J. Bing, 2004) the authors claimed that the urban traffic network can be modeled as a directed small-world network (W.Aiello,2000).

2. DATA CHANGE RECOGNITION MODEL

The GeoGraph model combines the graph theory with geospatial data. In this paper, we treat the GIS network a directed graph and study the traffic flow on the graph edges.

2.1 Directed Graph

A directed graph G is a finite nonempty set V of objects called vertices together with a set of E of ordered pairs of distinct vertices. The elements in E are called edges or arc. If (x,y) is a directed edge, then we indicate this on the diagram representing G by drawing a directed line segment or curve from x to y. Then x is said to be adjacent to x and y is adjacent from x. Edges (x,y)and (y, x) may both be present in the directed graph. In this paper, in order to simplify the model, we consider only single directed graph where only a single edge exist between two vertices.(B.Bollobás,1998)

2.2 Flow on the Graph

In the directed graph model, we refer the vertex (node) as the road intersection, and the edge connecting any two vertices (x,y) as the street road connecting the two intersections in one direction. The out-degree of a vertex is the number of edges originating from the vertex and the in-degree of the a vertex is the number of edges terminating at the vertex. Such connected vertices form the entire traffic network.

A flow f is a non-negative function defined on the edges; the value f(x,y) is the amount of traffic on the edge (x,y). The traffic flow from the source (***s***) to the destination (***t***) satisfies the Kirchhoff's current low: the total current flowing into each intermediate vertex (that is the vertex different from ***s*** and ***t***) is equal to the total current leaving the vertex. For any $x \in V$ we define two neighborhood of x:

$$\Gamma^{+}(x)=\{y \in V:\ xy \in E\} \quad (1)$$

$$\Gamma^{-}(x)=\{y \in V:\ yx \in E\} \quad (2)$$

Then a flow from ***s*** to ***t*** satisfies the following equation:

$$\sum_{y \in \Gamma^{+}(x)} f(x,y) = \sum_{z \in \Gamma^{-}(x)} f(z,x) \quad (3)$$

where + deontes the direction flowing out and - deontes the direction flowing into.

2.3 Maximum Flow and Minimum Cut Theorem

We denote v(f) as the value of "f " or the amount of flow from ***s*** to ***t***; c(x,y) is a non-negative number called the **capacity** of the edge.

Given two subsets *X,Y* of *V,* we write E(X,Y) for the set of directed X-Y edges:

$$E(X,Y)= \{xy \in E:\ x \in X\ \ y \in Y\} \quad (4)$$

If S is a subset of V containing ***s*** but not ***t*** then E **$(S, \bar{S})$** is called a cut separating ***s*** from ***t.*** Here $\bar{S} = V - S$ is the complement set of **S**. If we delete the edges of a cut then no positive-value flow from ***s*** to ***t*** can be defined on the remainder. The capacity of a cut E **$(S, \bar{S})$** is **$c(S, \bar{S})$.** It is easily seen from the definition that the capacity of a cut is at least as large as the value of any flow from ***s*** to ***t***, so the minimum of all cut capacities is at least as large as the maximum of all flow values. That's max-flow min-cut theorem of Ford and Fulkerson:

$$v(f) \leqq \sum c(x,y) \qquad xy \in E \quad (5)$$

$$v(f) = \sum c(x,y) = c(S, \bar{S}) \quad x \in S, y \in \bar{S} \quad (6)$$

Utilizing (5) and (6), Edmond and Karp designed an algorithm that find the maximum flow on G with complexity of $O(m^3)$ (M.H.Alsuwaiyel,1999), where m is vertex size of the network.

2.4 Variation of Flow

Consider a city road networks that goes through changes and repairs. Without losing generality, we view the c(x, y) as a constant value when the road network is working well. If some path is congested or closed, then we model c(xi,yj) as a negative power law function from experience. Hence we have

$$v(f) = C'(x,y) + f(a / T^{n}) = C + f(a / T^{n})$$

$$\text{and} \qquad \Delta v(f) = \Delta f(a / T^{n}) \quad (7)$$

As the time function $f(a / T^{n})$ is negative power law , so the $\Delta f(a / T^{n})$ and $\Delta v(f)$ is also to negative power law .

If we monitor the max-flow continuously and observe that a negative power law phenomenon appears in a time-window, then we can say the congestion has emerged. Through the Depths First Algorithm's Searching, we can easily find the E(x, y) is the congested or closed edges in the graph. We can then display a message regarding the condition of the corresponding road.

3. DATA CAPTURE ANALYSIS

Figure 1 is displays the Guangzhou city's old business area's street networks. We selected the main and large street cross's information as the sample points. All the roads are one-way streets drawed in the graph, then we obtain the dynamic road state information as it displayed. The numbers(0,1,2…29) are the IDs. of the vertices which referred as street crosses. We let them flow in from 0 and flow out on 29. The flow volume data can be acquired from City Traffic Control Centre or recorded from the field sampling points.

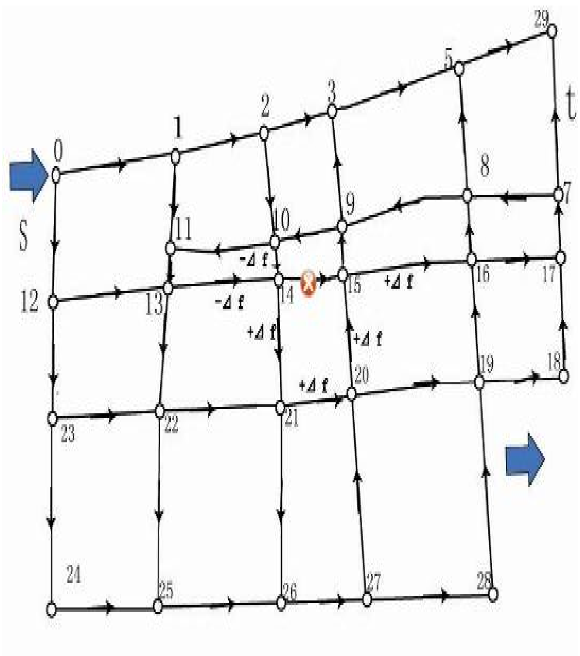

Figure 1: A traffic max-flow network graph in Guangzhou urban center :

From the Fig. 1, we suppose that the edge E(14,15) is congested and the flow is gradually slowing down to 0 as a negative power law, As the equation (1),(2),(3), we have f(13,14)+f(10,14)= f(14,15)+f(14,21). When the f(14,15)= 0, then the f(13,14),f(10,14) will decrease quickly and f(14,21) will increase the flow amount of Δf. We use the Depths First Algorithm's to find the E(14,15) is the true congested edge.

Then, we can transform the updating road message to the database.

4. GEOSPATIAL DATABASE UPDATE

4.1 Updating Client

In the GIS business market, there are many kinds of geo-database to used and developed , if we were to develop a updating client for each geo-database, it would be difficult and also unnecessary. Therefore, in this paper, we design a general software middleware agent analogous to a SDE(Spatial Data Engine). Its task is to manage the geospatial data's lifecycles.
The middleware agent can search the flow of any road segments. If the flows on the segments changing with certain predefined pattern (such as negative power law), we can conclude that the whole road is repairing or changing. Hence, we can display the road change message and update the database automatically.

4.2 Geo-database's Structure

Considering the Open-GIS's architectures as a trend for future development, we select the open source and the traditional business database : PostgreSQL. It's similar to the famous MySQL database. But the frontend has a PostGIS extention model for the support for geographic objects. PostGIS "spatially enables" the PostgreSQL server, allowing it to be used as a

Function	Return Type	Description
area(*object*)	double precision	area
Center(*object*)	point	center
diameter(circle)	double precision	diameter of circle
Height(box)	double precision	vertical size of box
isclosed(path)	boolean	a closed path?
Isopen(path)	boolean	an open path?
Length(*object*)	double precision	length
npoints(path)	int	number of points
npoints(polygon)	int	number of points
Pclose(path)	path	convert path to closed
popen(path)	path	convert path to open
width(box)	double precision	horizontal size of box

Table1:PostgreSQL/PostGIS 8.2 Geometric Functions Tables:

backend spatial database for GIS. Much like ESRI's SDE or Oracle's Spatial extension. When we manipulate the geospatial objects, the following geometric functions can be called from library.

In the above function list, we can operate the center objects(vertices), length(edge) and area(graphs) to analogue the network GIS's topology and features(attributes).

Finally, through network GIS model structure, and the object-oriented-relational database, we can achieve the automated database updates. Such approach can effectively use the network's dynamic and mobile ability to mitigate the conservative nature of geospatial data.

5. CONCLUSIONS

In this paper, we proposed a dynamic network data model for the GIS geospatial database. The automatic update of the database are achieved by quickly capture the traffic flow change and through network search algorithm, finding the specific network edge that causes the flow changes. This approach can be conveniently applied to many GIS applications.

REFERENCES

A.Hosseini naveh,et.al.,2006.Studying the effect of traffic elements by GIS.*Map World Forum,Hyderabad, India*.

B.Bollobás,1998.*Modern Graph Theory.* Springer-Verlag New York,pp.68-72.

Catherine Dibble, Philip G. Feldman, 2003.The GeoGraph 3D Computational Laboratory. *the 8th International Command and Control Research and Technology Symposium.*

Chen Tao,Chen Sen-fa, 2006.An Algorithm for Dynamic Core of Transport Network with Applications.*Journal of Transportation Systems Engineering and Information Technology* Vol.6 No.2.*Computing, Orlando.*

Frauke Heinzle, Karl-Heinrich Anders, and Monika Sester, 2006.Pattern Recognition in Road Networks on the Exampleof Circular Road Detection. Raubal et al. (Eds.): *GIScience*, LNCS 4197, Springer-Verlag Berlin Heidelberg, pp. 153 –167.

I. Budak Arpinar, et.al, 2004.Geospatial Ontology Development and Semantic Analytics. *Handbook of Geographic Information Science.*Eds: J. P. Wilson and A. S. Fotheringham, Blackwell Publishing .

Jiang B. ,C. Claramunt, 2004.Topological Analysis of Urban Street Networks, *Envirenment and Planning B,* pp.151-161.

Lu Feng, Shen Paiwei, Zhang Mingbo, 2004. Feature-based Object Oriented Geographical Network Modeling. *Geo-Information Science.*Vol.6,No.3,pp.72-78

M.H.Alsuwaiyel,1999.*Algorithms Techniques and Analysis.* World Scientific Publishing Co,pp.260-269

Mark T. Elmore Thomas E.Potok and Frederick T. Sheldon, 2003. Dynamic Data Fusion Using An Ontology-Based Software Agent System. *Proceedings of the IIIS Agent Based*

Shi,W.Z.,2005.*Principle of modeling uncertainties in spatial data and analysis.* Science Press, Beijing

W.Aiello,F.Chung,L.Lu,2000.Random graph model for massive graphs, *Proceedings of the Thirty-Second Annual ACM symposium on Theory of Computing*,pp.171-180.

Zhang Linguang, 2006.Implement and Optimization Research of GIS Network Analysis Algorithms for Large Amounts of Data, *Graudate Dissertation, Graduate University of Chinese Academy of Sciences, Beijing*.