A Study of Dynamic Database in Mobile GIS

Ka-wai Kwan and Wen-zhong Shi

Advanced Research Centre for Spatial Information Technology, Department of Land Surveying & Geo-Informatics, Hong Kong Polytechnic University, Kowloon, Hong Kong, Tel: (852) 2766-5974Fax: (852) 2330-2994, {99903830r, lswzshi}@polyu.edu.hk

Abstract

Wireless Internet is one of the fastest growing areas in the information economy, and as a result, location-based services have undergone explosive growth. In mobile computing, considerable research has been focused on the infrastructure and architecture design to try and solve system bottlenecks, such as the development of third generation mobile systems or to resolve cache management of the database. The database design, however, is an essential issue that must be considered. This paper introduces a new conceptual model for a spatial temporal data set for mobile GIS and explores the creation of a dynamic database. The database is updated continuously to reflect the real situation, and captures the movement of a mobile user. The goals are to fine-tune the original database and thus provide the latest and accurate information for mobile users with a limited communication bandwidth and to improve the response time. Issues of spatial temporal data integration in mobile databases, how they work and performance for real time applications and on-line queries are discussed. We assume that the spatial information is in a simplified structure with only one temporal dimension; the GIS are developed based on a relational database.
Keywords: dynamic database, location-based service, mobile GIS, spatial temporal data set, wireless web application

1 Introduction

Currently, workforce trends appear to be increasing service mobility and seizing opportunities presented by the Internet. It is interesting to note that International Data Corporation predicts that the majority of Internet access will be through wireless connection after 2002. As the Internet usage widely increases, the Internet system becomes more complicated and more difficult to ‘manage’. This creates many network problems. In the past, most solutions were to increase resources on the web server such as hardware improvements, for example through expansion of communication bandwidth. However, these are not low-cost methods or ideal solutions. A further important consideration is bandwidth expansion, which is not suitable in a mobile environment.

In general, the major concern of most web sites is the overload on servers. Only 15% of the web sites are effective relative to the system load testing (Mercury Interactive, 2001). The system bottleneck occurs in four main areas: database (27%), network (25%), application server (23%) and web server (20%). The major setback for the system bottleneck originates from the database. As a result, fine-tuning the existing system such as improving the design of existing databases is an important step to enhance the overall performance.

The database is an essential component in Geographic Information Systems (GIS) and without doubt, a poor design is a burden on performance. In this paper, we concentrate on the fine tuning of an existing database. We propose the idea of a spatial temporal data set in wireless GIS applications that can create the dynamic database to support the mobile application. A location-based web service is used to illustrate how it works. The objectives are to minimise the process time and to improve the data accuracy. Thus, these can provide an alternative way to enrich the work in mobile GIS applications. In addition, the assumption is made that only a simplified spatial structure will be used for the temporal dimension.

The remainder of this paper is organised as follows: Section 2 gives a brief introduction to mobile GIS. Section 3 describes the spatial and related attribute data sets and the dynamic database for mobile GIS. Section 4 analyses the performance of the proposed dynamic database. Section 5 concludes the work.

2 Mobile GIS

The intention of a mobile GIS is to bring GIS technology into the field. This presents a relatively new dimension - in any time, any place - to access spatial and attribute information. It offers another new perspective for GIS application and further extends the usage of the “office”.

2.1 Wireless Web GIS Application

Mobile GIS application is the integration of mobile technology into traditional GIS or Internet GIS. The “Thin Client” concept and sharing of information among an unlimited number of users can also be added. The applications can be divided into two main groups: system construction and system utilisation. The former one concentrates on creating the GIS, such as data collection, data checking, data updating among other aspects. The system utilisation aspect concentrates on the use of an existing system, for example, location searching, path finding, and information seeking. In this paper, the utilisation issue is the focus.

A wireless web GIS application extends the use of Internet GIS. Through the web content provider's service offering and by pointing to a web-browser on mobile a device and then to a web site, the mobile users can perform various functions. This kind of application uses all the advantages of Internet, mobile computing and GIS. This also allows users to interact with GIS data and maps on the Web without owning mobile GIS software and data. Mobile GIS applications mainly serve the public and the business to customer (B2C) services.

2.2 Limitations of Wireless Web GIS Application

As wireless web GIS application combines the wireless, Internet and GIS technology, it has all the limitations of the above three technologies. In the second generation of mobile systems, the low bandwidth and low reliability are the major obstacles in application development. These obstacles constrain the applications design. Moreover, the small screen and low resolution of mobile devices degrade the visibility and create a non user-friendly interface, and certain sophisticated and enhanced functions cannot be performed through the Internet, such as three-dimensional analysis. Since the amount of data sent over web is only a quarter to one-third of the amount of corresponding raster data, the vector display has a far greater productivity and faster response. Unfortunately, the mobile web browser and plug-ins do not support vector data. Furthermore, there is no common vector format in the Internet world. As a result, the mobile web page contents are mainly composed by text, supplemented with images, but of low resolution. Lastly, the server side performs most processing works. Hence, poor database design has downgraded the overall performance.

2.3 Location-based Service

Industry observers forecast that by 2004 the location-based service will create almost $10 billion per year of new revenue world-wide for wireless operators. The sponsors of the OpenGIS Open Location Services (OpenLS) initiative states: "Spatial connectivity is a primary, universal construct for business planning and modelling, service development and deployment, network provisioning and operation and customer satisfaction. The applications of location-based services are of universal industry service significance and depend upon the availability of relevant spatial information infrastructures in forms useful for small devices."

Location-based services can describe any service or application with GIS capabilities via Internet or wireless networks. Through combining both advanced technologies – i.e. mobile GIS and positioning service, new business opportunities, with many potential new applications and services can be developed. Thus, mobile GIS becomes an essential component to wireless mobility.

3 Dynamic Database

A spatial temporal data set is defined as a collection of spatial objects including their whole lifetime (history). Each member in the spatial temporal data set can be treated as a snapshot of a particular spatial object in a particular moment. The notation is:

$$\mathrm{ST} = \{\mathrm{s_i}(\mathrm{t}) : \mathrm{g}(\mathrm{s_i}(\mathrm{t})) \in \mathrm{A}, i = 1,2,\ldots,n; 0 < \mathrm{t} < \infty\} \tag{1}$$

where si(t) refers to a spatial object 'i' at time t.

Traditionally, a change of attribute can be defined as the difference between the two elements in the same spatial temporal data set that corresponds to the same object. For example, there is only one building in zone A at time t1. In time t2, however, there are three buildings in the same zone.

In a mobile environment, we should focus on the location of the user. At different locations, a mobile user may encounter different objects in different locations. The user also gets a different picture at the same location at different times. The changes of a particular object itself are meaningless unless the 'entire' scene is considered. The spatial temporal change should be described as the environmental change according to the point of view of the mobile user at his current position and time.

In mobile GIS applications, we distinguish between attribute, spatial and temporal changes. Attribute change comprises the inclusion of only the application-related attributes that are retained. Temporal change includes the alteration of the same spatial attributes at a different time, and parts of the object are discarded if they are invalid for the application. Spatial change is the environmental change while the mobile user is at a different position. After applying the new spatial temporal data set in the original database, a new subset is created and the contents will continue to be updated as the location changes. We term this as a "dynamic database".

Dynamic databases consist of the "value-based" relationships where typically, the relationship is specified at a retrieval time and the locations of related records are discovered during retrieval. Therefore, the content and number of records are altered from time and time.

Fig. 1 shows the spatial temporal data set. For experimentation, an assumption is made that a mobile service is requested to find the coffee shop at a particular location. In stage 1, it shows the structure of the original GIS database. Each row corresponds to one spatial object. The attribute constraints are added in stage 2. The relevant fields such as "ID", "Open_time", "Close_time", "District" and "Name" are preserved. "ID" acts as the unique key while "Open_time" and "Close_time" indicate the working hours. "District" is the location identification and "Name" is the label. A spatial constraint is added in stage 3, where all objects within the mobile user's scopes are extracted. For example, the spatial data are kept in the same district of where the user is located. In the final stage, a temporal constraint is implemented and the meaningless information is discarded. For example, even if the buildings are in the area of interest, but do not have the same office hour temporal properties, they can be ignored. As a result, a subset is generated.

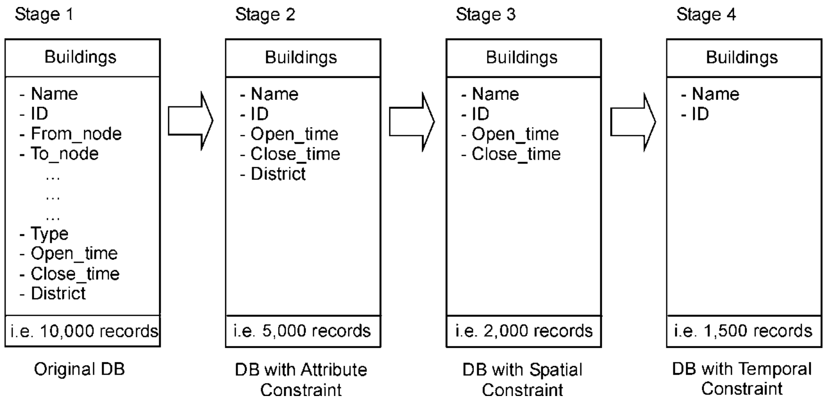
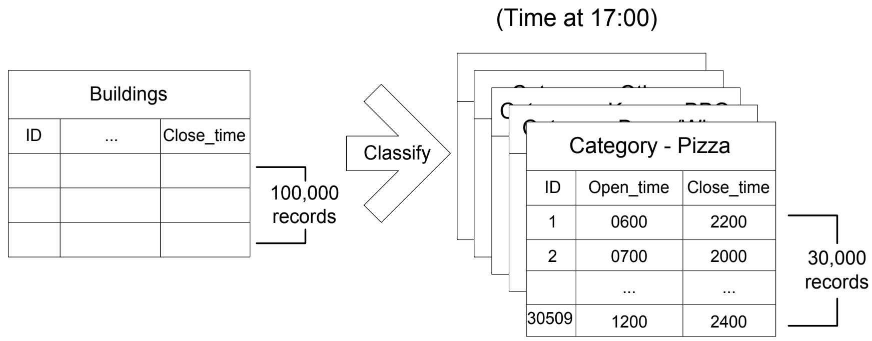


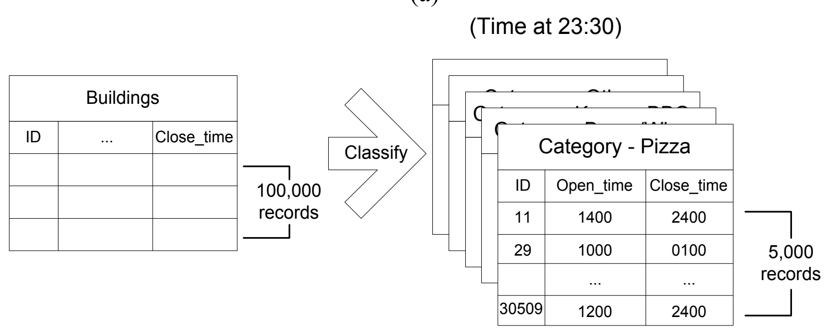
Fig. 1. Spatial-Temporal set concept in spatial data

Physically, the contents in each dynamic database are extracted from the original database and stored separately. However, the sum of records in all dynamic databases may not be equal to the number of records in the master database.

Figs. 2a and 2b illustrate that the number of records varies from time to time. For example, there are 100,000 records in the attribute table and 30,000 are categorised as fast food shops. At time 1700, there are 30,000 records. However at 2330, only 5,000 records are stored because they are still within working hours.



(a)



(b)

Fig. 2(a) Number of records at time 17:00, **(b)** number or records at time 23:30

Fig. 3 shows the attribute value changes in a dynamic database. A particular value, i.e. cost value, is varied from time to time. While there is environmental change, some contents are also changed to reflect the truth. A typical example is the optimum path measurement. Optimum path means the best route from the source to destination path is calculated from all objective factors, such as the traffic condition, the cost to pass, the penalty for the turns, and so on.

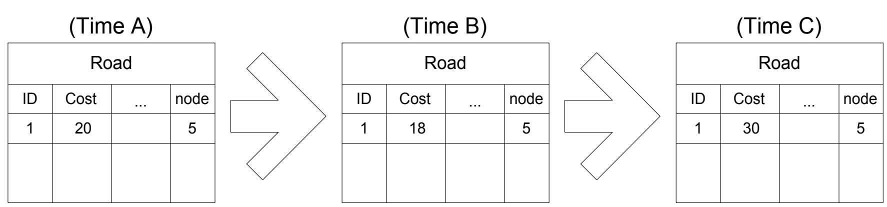


Fig. 3. Change in attribute value

4 Performance Evaluation

The model discussed in this paper has been simplified with the following assumptions:

1. Time is one-dimensional and linearly ordered;
2. Relational databases are used to develop the GIS model;
3. Connection to the server and data retrieval via an active TCP/IP connection using the Hyper Test Transfer Protocol (HTTP) utilised as data transfer protocol; the Mobile user provides an initial position.

A simplified location-based service has demonstrated how the dynamic database works. A mobile user chooses a district, a current location, pass and destination points, and maximum walking distance. Based on the above criteria, the optimum path can be found and displayed on the mobile device. We characterise the performance of this function in the following way. The average response time from the mobile client is measured as the time spent (in seconds) from the moment the query is issued to the moment the results of the query are generated. The hardware settings are summarised in Table 1.

Table 1. Setting of the experiments

Web Server & GIS Server	CPU:	P-III 650
	RAM:	128M
	Web Server:	Personal Web Server
	GIS Server:	TransCAD
Mobile Device	Model:	Compaq iPAQ H3630
	OS:	Windows CE Version 3.0
	Web Browser:	Internet Explorer for CE

4.1 Analysis of the Results

To reflect the real situation in Hong Kong, international cuisine are chosen from Yellow Pages. There are 9 different types of international cuisine, which are listed in Table 2. Moreover, this test has been conducted using the Intranet.

Table 2. Nine types of international cuisine

Type	Beers / Wines	Curry / Satay	Japanese Noodles	Korean BBQ	Pizza	Sashimi / Sushi Bar	Steak House	West. Dessert / Drink	Others
%	40.8	5.5	2.1	1.6	6.8	5.4	4.7	13.5	19.6

The result is depicted in Fig. 4, which is plotted as a two-dimensional graph to illustrate the results of average response time. The line with square nodes corresponds to the time measurement from the dynamic database, while the line with diamond nodes represents the time measurement from the original database.

We observe that the time saved increases as the database size expands. For the number of records, over 60,000, the time required in the original database is three times more than that in the dynamic database. It is noted, however, that this is related to the actual time of the query, since the database volume is varied as time changes. The response time is slightly different to the result shown in Fig. 4. Nevertheless, we can easily discover that the time response on the dynamic database is faster than the original database.

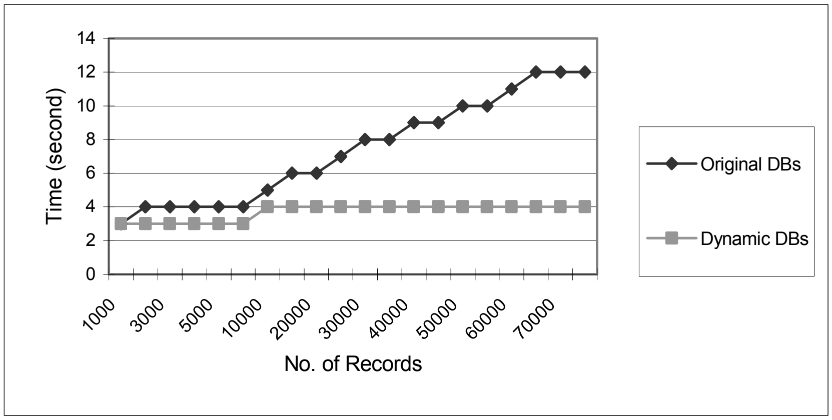


Fig. 4. Average response time

5 Conclusions

This paper described a new conceptual model – a spatial temporal data set for mobile GIS and correspondingly a dynamic database. In traditional GIS, spatial objects are the primary focus whereas in mobile GIS, the focus has been changed to a particular scene of a particular location at a particular time. The relevant objects are only a subset of the overall objects in the GIS and meaningful to the mobile user for a particular situation – involving location and time. The dynamic database is automatically updated to mirror the current world. The application and test were developed using a wireless web GIS environment.

We compared the difference between a traditional database with the proposed dynamic database in terms of response. From the results we can see that the processing time of the conventional database is proportionally increased with the data size in the database while the response time can be dramatically reduced in the proposed dynamic database. This is even more effective for a database with a large data volume. In the dynamic database, the response time is minimised and query accuracy is improved. These two advantages will greatly benefit mobile users in their applications.

References

Aberdeen Group, Inc. (2000) The Benefits of Mobilising Enterprise Applications with Handheld Devices, August 2000

Chan Y, Si A, Leong HV (1998) Cache Management for Mobile Databases: Design and Evaluation. Data Engineering, 1998, Proceedings, 14th International Conference, 1998, pp 54-63

Coates P (2000) Mapping and the Mobile Future, GI News, June 2000

CONNECT Computer Corp. (2000) What is Thin Client Computing? [online] Available from: <http://www.connectcomputer.com/thin-client.asp> [Accessed 13 December 2001]

Darwen H, Date CJ (2000) Temporal Database Systems. Advanced Database Technology and Design, 2000, pp 137-187

Dent BD (1997) The Map Design Process and the Elements of Map Composition. In: Cartography Thematic Map Design, pp 236-267

Dunham MH, Kumar V (1998) Location Dependent data and its Management in Mobile Database [online]. Available from: <http://ieeexplore.ieee.org/iel4/5718/15304/00707433.pdf> [Accessed 15 December 2001]

Francica J (2000) Location-Based Services Wireless Meets GIS. GEOWorld, December 2000

Hanttula D (2001) The Future of the Pocket PC. SemperAptus.com, January, 2001

Hecht L (2000) Stand up for Standards. GEOEurope, September, 2000

Iyengar SS, Pangburn BE (1998) Real-Time Computing: Meeting the Client's Demand for Data Access via the Internet. In: IEEE Workshop on Application-Specific Software Engineering Technology, ASSET-98 Proceedings. 1998, pp 14–19

Koeppel I (2000) What are Location Services? – From a GIS Perspective [online]. Available from: <http://www.jlocationservices.com/company/esri/What%20are%20%Location%20Services.html> [Accessed 8 July 2001]

Laurini R (2001) Real Time Spatio-Temporal Database. Transactions in GIS 5(2)

Lauzac SW, Chrysanthis PK (2001) Programming views for mobile database clients [online]. Available from: <http://ieeexplore.ieee.org/iel4/5718/15304/00707432.pdf> [Accessed 17 December 2001]

Lowe JW (2001) Which Where Will Win? Dueling Wireless Location Solution. Geospatial Solutions, March 2001

Meshulam Y (2001) says report HK Economics Times, 10 October 2001.

Mirghafori N, Fontaine A (1995) A Design for File Access in a Mobile Environment. In: Proceedings Workshop on Mobile Computing Systems and Applications, 1994. 1995, pp 57-62

Open Location Services Initiative (2001) [online]. Available from: <Open Location Services: Vision and Objectives, Corporate Roles and Work Program Summary> [Accessed 16 December 2001]

Palm Inc (2000) Wireless Enterprise Application for Mobile Information Management [online]. Available from: <http://www.palm.com/pdfs/wireless_ent_apps.pdf> [Accessed 15 December 2001]

Peuquet DJ (1999) Time in GIS and Geographical Database. Geographical Information Systems, Vol. 1, 2nd Ed. pp 91-103

Potmesi M (1997) Map Alive: Viewing Geospatial Information on WWW. In: Proceedings of 6th International World Wide Web Conference, 1997.
OGC Releases a Request for Technology for a Major Web Services Initiative. Open GIS Consortium – Press Releases, March, 2001
Stead SD (1996) Temporal Dynamics and Geographic Information Systems. Spatial and Temporal Reasoning in Geographic Information Systems, Chapter 15, pp 214-219
Subhasree N (2001) GIS Data Dissemination: A New Approach through WEB Technology [online]. Available from: http://www.gisdevelopment.net/technology/gis/techgi0028.htm [Accessed 16 December 2001]
Varshney U, Vetter R (2000) Emerging Mobile and Wireless Networks. Communications of the ACM 43(6)
Wachowicz M, Healey RG (1994) Towards Temporality in GIS. In: Innovations in GIS. pp 105-115
Waters NM (1999) Transportation GIS: GIS-T. Geographical Information Systems, Vol. 2, Second Edition. pp 827-844
Wilson JD (2000) Mobile Technology Takes GIS to the Field. GEOWorld, June 2000
Zhang ZQ, Griffith DA (2000) Integrating GIS components and spatial statistical analysis in DBMSs. International Journal of Geographical Information Science, Vol. 14 Number 6, September 2000, pp 543
YP food (2001) [online] Available from: http://ypfood.com/en/html/search/ypfood_basic_eng.asp [Accessed 15 December 2001]