RESEARCH ON A DISTRIBUTED ARCHITECTURE OF MOBILE GIS BASED ON WAP

Wang Fangxiong [a], Jiang Zhiyong [b]

[a] Research Centre of Spatial Information and Digital Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China. wfxwhu@163.com

[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China. houyingzi@sohu.com

Commission VI, WG VI/4

KEY WORDS: Mobile GIS, Architecture, J2EE, WAP, Mobile Internet

ABSTRACT:

Mobile Geographical Information System (Mobile GIS) is a Geographical Information System based on mobile computing and mobile Internet. It is not a conventional GIS modified to operate on a smaller computer, but an extension of Web GIS to mobile Internet including wireless Internet/Intranet and mobile communication network. But architectures of Web GIS are unsuitable for Mobile GIS, because of several bottlenecks such as the low-bandwidth of wireless network, the diversity of mobile devices, limited processing power and screen display limitation of mobile devices and the diversity of mobile system platform. So the research on the architecture of Mobile GIS is hot in the GIS field. WAP-based Mobile GIS can be described as mobile users can perform the almost same functionality as of Web GIS but in a mobile environment at any time, any place and without the limitation of operating system and wired link. In this paper, a distributed architecture for WAP-based Mobile GIS is proposed. Using Java Servlet Engine technology and spatial information Web caching mechanism at the fat server side, the architecture based on J2EE is composed of four logic tiers: presentation tier, WAP service tier, application tier and data service tier. A prototype system based on the architecture has been developed. It can successfully work on WAP phones, which proves that the architecture is feasible.

1. INTRODUCTION

Since 1990, geospatial information technologies and mobile wireless Internet have been rapidly developed. It is easy to see that the integration of geospatial information and mobile Internet is inevitable, which is simultaneity driven by market demands and technologies (Li Deren, et al, 2002). The integrated system is designed to work on mobile intelligent terminals, and brings new dimension - at any time, any place - to access geospatial and attribute information in GIS. It is called Mobile Geographic Information System (Mobile GIS). Mobile GIS offers another new perspective for the use of GIS and further extends the "office" GIS works in mobile environment (Shi Wenzhong and Kwan Kawai, 2002). Mobile GIS was early applied to assist office and collect data in the field (Ness S. T. Lee, 1993; K. North, 1997). It is not a conventional GIS modified to operate on a smaller computer, but a system build using a fundamentally new paradigm (David Maguire, 2001). It extends unlimited information on the Internet and powerful service functions of GIS to mobile devices, and can provide mobile users with geospatial information services. Mobile GIS creates a new channel of business practice, and thousands of potential applications and services can also be developed. Hence, it is exploring a new era of mobile geographic information services.

The rest this paper is arranged as follow: WAP-based Mobile GIS and its superiority to SMS/MMS-based Mobile GIS are briefly introduced in Section 2. Three mainstream architecture solutions for distributed systems are presented in Section 3. A distributed architecture for WAP-based Mobile GIS based on J2EE is proposed in Section 4. And a prototype system implemented based on J2EE is presented in Section 5. Finally, the conclusion is drawn and outlook of WAP-based Mobile GIS are discussed.

2. MOBILE GIS BASED ON WAP

Mobile GIS can be simply divided into two categories, according to currently popular solutions accessing to the Mobile Internet. The one is based on Short Message Service (SMS)/Multimedia Message Service (MMS), the other is based on Wireless Application Protocol (WAP).

SMS/MMS-based Mobile GIS (Ma Lingbin, et al, 2002; Wei Zhongya, et al, 2002; Zhang Jianqun and Liang Juanzhu, 2003) can only be suitable for mobile phones with sample system functions, unfriendly graphical user interface (GUI), poor information presentation and the restricted application fields, because of the limitations of SMS/MMS such as restricted carrying information, time lag, unfriendly interactive mode, and so on.

On the contrary, WAP is a bear-independent international standard protocol that has optimized for mobile devices with limited display and small keyboards of mobile handsets and low bandwidths of wireless networks, and permits applications and services to operate over all existing and foreseeable wireless networks such as GSM, CDMA, PHS, TDMA and WCDMA (WAP Forum, 2002). The WAP specification encompasses a relatively simple and compact version of XML (eXtendable Markup Language) called WML (Wireless Markup Language), which makes it possible to make requests to a mobile service from a mobile terminal and return a map in the form of an embedded bitmap (e.g. WBMP). So WAP-based Mobile GIS has richer information presentation, friendlier GUI, more system functions and more application fields than the former. Moreover, it can work on a wide range of mobile devices with a WAP microbrowser only, from Personal Digital Assistants (PDAs), mobile phones, and in-car computers to other small mobile devices, and keeps the standard browsing style. WAP-based Mobile GIS can be described as mobile users (with a WAP mobile terminal only) can perform the almost same functionality as of Internet GIS but in a mobile environment at any time, any place and without the limitation of operating system and wired link. Altogether, it is because of the advantages of WAP that we believe WAP-based Mobile GIS will play a leading role in our mobile information services markets.

The framework and soul of a distributed system is the system architecture deciding the distribution of system functions and data. WAP-based Mobile GIS, as a new development to Internet GIS, is still in a preliminary stage, whose research issues need be deal with. However, the architecture solution to WAP-based Mobile GIS is the key and core of its whole researches. Although Li Deren, et al (2002) and Li Luqun, et al (2002) used to present an architecture for Mobile GIS respectively, the architectures were conceptual ones and had difficulty playing a substantially directive role to building a Mobile GIS system, let alone WAP-based Mobile GIS.

3. ARCHITECTURE SOLUTIONS

WAP-based Mobile GIS is a “thin client” distributed system via mobile Internet. It must be an open, extendable, stable and cross-platform distributed system, because of the increasing application demands and the diversity of mobile terminals.

Currently, popular system architecture solutions for distributed systems are all based on distributed object technologies. There are three mainstream industry standards: Microsoft's Windows Distributed Network Architecture (Windows DNA)/.NET Architecture, Object Management Group's Common Object Request Broker Architecture (CORBA), Sun's Java 2 Enterprise Edition (J2EE) Architecture. A system based on Windows DNA/.NET Architecture solution can only use Microsoft's platforms from development, deployment to running, including developing platforms and operating system. It can not cross operating system platform especially, which is the fatal weakness of Windows DNA/.NET Architecture. CORBA is too huge and complicated. And its technologies and standards are updated relatively slowly. J2EE is a specification and standard created by Sun and her industry partners, and it evolves quickly. J2EE provides support for the technologies such as Enterprise JavaBeans (EJB), Java Servlets API and Java Server Pagers (JSP), and so on. J2EE solution reduces the cost and complexity of developing a multi-tier distributed system which can be rapidly developed and deployed, and can enhance the portability, security, load balancing and extensibility of a distributed system. There are the following advantages to build distributed GIS based on J2EE (Mao Haifeng, 2004): (1) cross-platform; (2) multi-tier separating to complicated tasks; (3) component reusing and (4) module developing. Thus, J2EE solution can satisfy the system requirements of WAP-based Mobile GIS. Based on J2EE, this paper proposes a distributed architecture for WAP-based Mobile GIS.

4. A DISTRIBUTED ARCHITECTURE

As showed on Figure 1, the J2EE-based distributed architecture is composed of four logic tiers from the client side to the server side: presentation tier, WAP service tier, application tier and data service tier. At the fat server side, Geospatial information Web caching mechanism is used to optimize the performance capability and reducing the access delay of the client. Functions and components of each logic tier and the statement about the distribution of presentation logic, business logic and geospatial data in the architecture are detailedly stated as follow.

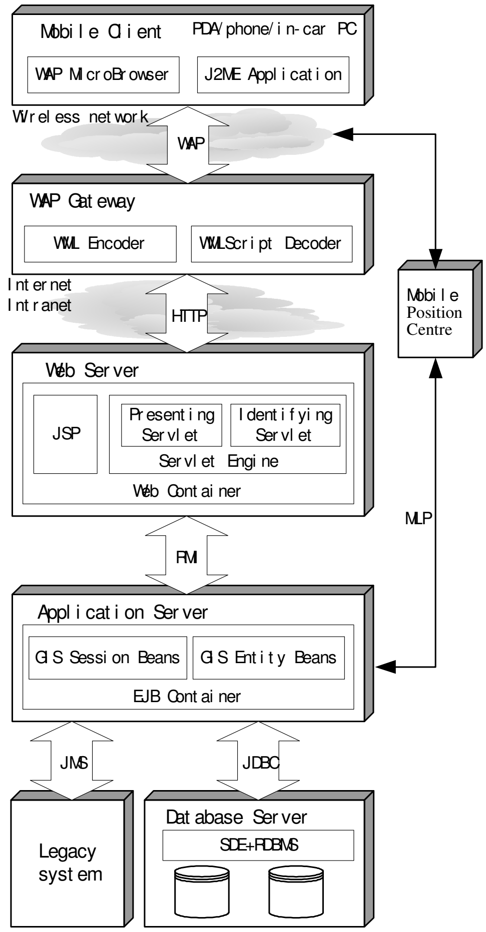

Figure 1: A distributed architecture of WAP-based Mobile GIS

4.1 Presentation Tier

Presentation tier is a carrier of the client of WAP-based Mobile GIS, and is mainly responsible for implementing the presentation logic of GIS data. Generally, the client without local-storage data is a WAP microbrowser that controls the GUI and is analogous to a standard Web browser. The WAP microbrowser needn’t perform any GIS business logic, don’t directly connect to a back-end database server, and don’t store itself state information, so it is a really “Thin Client”. The client also may be a J2ME (Java 2 Micro Edition) application.

4.2 WAP Service Tier

WAP service tier includes a WAP Proxy (often referred to as a WAP Gateway) and a Web Server. The WAP Gateway was required to handle the protocol interworking between the client and the Web Server. As depicted in Figure 1, the WAP Gateway consists of WML encoders and WMLScript decoders. The WAP Gateway can optimize the communication process and may offer mobile service enhancements, such as location, privacy, and presence based services. The WAP Gateway communicates with the client (WAP microbrowser) using the WAP protocols, and it communicates with the Web Server using the standard Internet protocols such as HTTP/HTTPS. Once the WAP Gateway receives WAP requests from the client, it translates the requests to HTTP requests, and then sends them to the Web Server. Once the WAP Gateway receives HTTP responses (web contents) from the Web Server, it translates the web contents to compact encoded binary formats for reducing the size and number of packets traveling over the wireless network to the client for displaying and/or processing.

The Web Server includes a Web Container and Web protocols support, security support, and so on. Web caching mechanism for Geospatial Information The Web Container is responsible for managing a Java Servlet Engine and Java Server Pagers (JSP). An Internet GIS system based on Java often uses Java Applets and/or Servlets to extend the dynamically displaying functionality for the Web browser. This paper uses a Servlet Engine instead of running Applets and/or Servlets inside the Web Server, which has several advantages as follow:

1 The Servlet Engine runs itself inside the Web Server. But applets inside the Web Server must be dynamically downloaded to the client for processing, which would increase the load of wireless and mobile devices (Especially, a WAP phone has not the capability of running applets.).

2 The Servlet Engine can host Servlets and provide them with standard Java Servlet APIs.

3 The Servlet Engine decouples Servlets processing from specific implementation details of the Web Server. This increases the flexibility of the system architecture, as it allows the Web Server to change without impacting on the overall system. The Servlet Engine can also provide various management features that help to shift the load of Servlets processing away from the Web Server.

The client devices of WAP-based Mobile GIS have several types (e.g. PDA and WAP phone) whose displaying and presenting capabilities are different. So there must has a mechanism at the server side, that is, the Web Server may determine the type of the client device and generate corresponding presentation logic for the client. Fortunately, the Servlet Engine can solve the problem effectively. The Servlet Engine may provide two kinds of Servlets: the one is presenting Servlet which is responsible for generating corresponding presentation logic according to the type of the client device for the client; the orther is identifying Servlet which is responsible for determining the type of the client device (through the only ID of the client device accessing mobile networks) and then notify presenting Servlet.

The Servlet Engine is responsible for manage Servlets and provide support for JSP. Since Servlets run inside the Web Server, everyone does not need a GUI. Servlets in the same servlet engine can share resources and chain together. So the Web Server can generate dynamical Web contents by running Servlets (i.e. presenting Servlets) to response for client requests like CGI. And Servlets are more efficient than the CGI approach and the Fast-CGI approach. Through Servlets and/or JSP, the client side may indirectly perform EJB components inside the Application Server to implement GIS business logic, such as spatial analysis, spatial and attribute querying, route planning, geocoding and gazetteer, and so on.

4.3 Application Tier

Application tier is the core of the architecture. It corresponds GIS Application Servers that communicate with the Web Server in WAP service tier through Remote Method Invocation (RMI). An EJB container at an Application Server is the runtime environment of EJB components including GIS Session Beans and GIS Entity Beans, and controls these components to be performed and transferred. At the same time, the container also provides these components with all required services for distributed computing environments. Thus, these EJB components could more efficiently execute in the Application Server. The EJB components can use JDBC (Java Database Connection) technology to access to database servers, and use JMS (Java Message Service) technology to connect to back-end legacy systems. The Application Server has a special Locating Entity Bean to communicate the Mobile Position Centre (MPC), a Server providing geographic location information, with Mobile Location Protocol (MLP) (Location Interoperability Forum, 2002). The mobile position technologies for real-time capturing the location information of mobile users generally are GPS, Cell Of Origin (COO), Time Of Arrival (TOA), Angle Of Arrival (AOA), Enhanced Observed Time Difference (E-OTD) and so on.

4.4 Data Service Tier

Data service tier corresponds Database Servers that are used to manage and store geospatial and attribute data of the whole system. Object oriented database management system (OODBMS) is the most desired database server for a GIS system, but OODBMS is immature and very costly currently, so it is not popular and to be used commonly (Gong Jianya, 2001). At present, the mainstream solution is that large object-relation database systems such as DB2, Oracle, Sybase, SQL Server, and so on are used to manage and store GIS data, at the same time, spatial data engine (SDE) also can be developed to build the communication between data service tier and application tier. SDE is an open standards-based middleware such as ArcSDE, Spatial Ware and Oracle Spatial.

5. J2EE-BASED PROTOTYPE SYSTEM

Using the 100% Pure Java J2EE solution, the above architecture is open and cross-platform, and has the characteristics such as extensibility, stability and reliability, which could satisfy system requirements of WAP-based Mobile GIS. Based on the architecture, a prototype system of WAP-based Mobile GIS has been developed. In the prototype system, the server side uses Apache Tomcat 5 as the web server, BEA WebLogic 8.1 as the application server, Oracle 8i and ESRI ArcSDE as the database server, and it provides a range of basic GIS functions (e.g. spatial and attribute querying, route planning, geocoding, mapping, etc.) to a Mobile GIS user using the Servlet Engine and EJB components to gain and process spatial geometry object in back-end spatial databases. The client side uses WAP phone with a WAP microbrowser and may obtain maps in WBMP format from the server side, and it has the basic map browsing functions such as zooming out, zooming in, moving (up, down, left and right). The running results are showed in Figure 2.

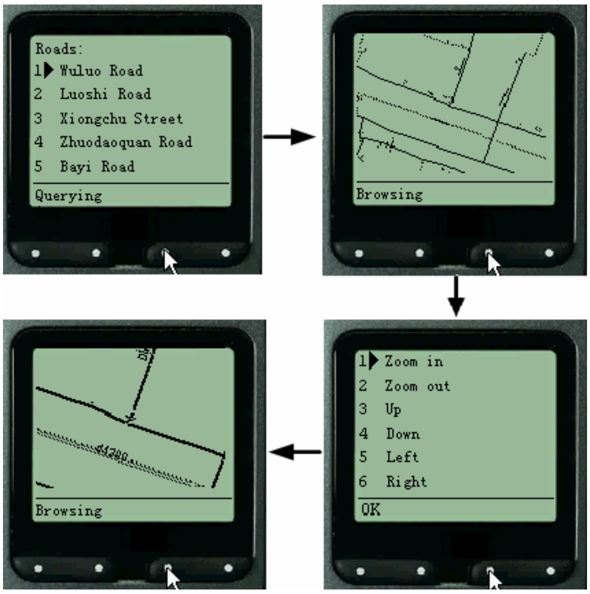

Figure 2: A prototype system of WAP-based Mobile GIS

6. CONCLUSION AND PROSPECT

WAP-based Mobile GIS is a very new study field of Geoinformatics. Above all, the architecture is solved, because it is the key and core of the distributed system. In this paper, based on J2EE, an open distributed architecture is proposed for WAP-based Mobile GIS. And a prototype system is designed and developed based on the architecture which is composed of four logic tiers. The prototype system can successfully work on WAP phones, which proves that the architecture is feasible.

At present, low bandwidths are still the main bottleneck of all mobile applications (Wei Zhongya, Xu Sunxin and Wu Lun, 2003). So researches on the organization of geospatial data at the fat server side and on the presentation of geospatial information at the thin client side should be done farther. Fortunately, the mobile network is towards the development of 3G. In 3G age, when the mobile terminal moves at the same speed as vehicle, the transmission speed is 144kbps, when the mobile terminal moves at the walking speed or un-moves in outdoor, the transmission speed is 384kbps, when the mobile terminal is in the room, the transmission speed is up to 2 Mbps. The 3G bandwidths will satisfy with the requirements of geospatial information wireless transmission. Therefore, it is not hard to understand that WAP-based Mobile GIS will have better development and application perspectives and considerable business value.

REFERENCES

Li Deren, Li Qingquan, Xie Zhiying, Zhu Xinyan, 2002. The Technique Integration of the Spatial Information and Mobile Communication. *Geomatics and Information Science of Wuhan University,* 27(1): 1-6.

Shi Wenzhong, Ka-wai Kwan, 2002. : A Review and Analysis of Mobile GIS Development. *The International Workshop on Mobile and Internet GIS*. Wuhan, 2002.

Ness S. T. Lee, 1993. Single Line Street Network: The Foundation of Mobile GIS. *IEEE-IEE Vehicle Navigation and Information Systems Conference*. Ottawa-VNIS'93, 34-37.

K. North, 1997. Field Information Systems for Managing Your Assets. *Engineering the Benefits of Geographical Information Systems*. IEEE Colloquium, 1997, 6/1-6/7.

David Maguire, 2001. Mobile Geographic Services Come of Age. *Geoinformatics*, March, 6-9.

Ma Linbing, Gong Jianya, Zhang ChunSen, 2002. Research on Application Solution and Key Technology of Mobile GIS. ISPRS, Volume XXXIV, Part2, Commission II: 323-326.

Wei Zhongya, Xu Suning, Wu Lun, 2002. A Model of Mobile Geographic Information Service. ISPRS, Volume XXXIV, Part2, Commission II: 525-529.

Zhang Jianqun, Liang Juanzhu, 2003. Implementation of Mobile Spatial Information Service By MMS. *Geomatics and Information Science of Wuhan University*, 28(1): 115-119.

WAP Forum, 2002. WAP 2.0 Technical White Paper. http://www.wapforum.org/what/WAPWhite_Paper1.pdf (accessed 28 Nov. 2003)

Li Luqun, Li Chengming and Lin Zongjian, 2002. Investigation on the Concept Model of Mobile GIS. ISPRS, Vol. XXXIV, part 4, Commission IV. http://www.isprs.org/commission4/proceedings/pdfpapers/415.pdf (accessed 18 May 2003)

Mao Haifeng, 2004. Distributed GIS for Agriculture Based on J2EE. *Geomatics and Information Science of Wuhan University,* 29(2): 142-143.

Location Interoperability Forum, 2002. Mobile Location Protocol Specification. http://www.openmobilealliance.org/lifdownload/LIF-TS-101-v3.0.0.zip (accessed 18 May 2003)

Gong Jianya, 2001. Concept and Development Trend of Spatial Database Management System. *Science of Surveying and Mapping*, 26 (3): 4-9.

Wei Zhongya, Xu Suning,Wu Lun, 2003. A WAP-based Geographic Information Mobile Service. *Communication Technology Proceedings, ICCT 2003. International Conference on*, Vol2: 1455-1460.