A STUDY OF GEOGRAPHIC MODELS INTEGRATION IN DISTRIBUTED COLLABORATIVE DESIGN

Hui. Yang [a, *], Guonian Lu , Yongning Wen , Yi Yang, Ying Hu

[a] Key Laboratory of Virtual Geographic Environment, MOE,Nanjing Normal University,China - whinee@gmail.com

Commission II,SS-15

KEY WORDS: Model, Integration, Distributed, Interoperability, Design, GIS

ABSTRACT:

To better adapt to modern distributed and heterogeneous network environment, Geographic Information Systems (GIS) will be used by more and more geographers, who are from different regions and work for different institutes, to cooperate to solve the complicated geographic problems. Distributed collaborative geographic model integration environment (DCGMIM), an integrated technology, offers an interactive, efficient, and feasible environment through which geographically separated geographers can solve complicated geographic problems and conduct collaborative work. In this paper, four key technologies, geographic model organization, representation, wrapping and integration strategies, are tightly coupled to develop a DCGMIM systems. Using a Web Services-based architecture it is efficient to integrate and share distributed resources as well as modelling procedures built on different platforms. Finally, a DCGMIM prototype system is implemented for collaboration on model building on the groundwater. The experimental results show that the architecture developed in this paper is efficient and feasible.

1. INTRODUCTION

Solving geographic problems has becomemore multidisciplinary and multiobjective in nature. A geographic problem such as the analysis of drainage basin is akind of multidisciplinary research, including geomorphology, sedimentology, biology and so on. In addition, various subjects such as physics, chemistry, dynamics, and hydraulics, should be interacted and negotiated to solve complicated geographic problems.

Geographic models attempt to describe, analyze and simulate the geographic problems in real world. It has become a trend to build different geographic models in different regions for many research institutes. For instance, geographic model A might be built in location A while the model B is designed in location B. All of these pre-existing geographic models are heterogeneous and distributed, so it is necessary for us to research geographic model integration in collaborative design.

The request to integrate geographic models in distributed collaborative design has been the fundamental motivation and original driving force. On the one hand, the growing requirement of spatial analyses needs geographic models should be integrated from variable related regions. On the other hand, various geographic models lead to the high cost of model building.

Traditional geographic modelling approaches are no longer appropriate for modern distributed, heterogeneous network environments, due to their closed architecture, and lack of interoperability, reusability, and flexibility. With the advanced progress of networked technology, geographers and geo-analysis experts now can solve model geographic problems which were beyond the capability of a sole computing power in years ago.

Along with the rapid progress of computer technology and networked technology, the geographic model integration in distributed collaborative design has become not only possible, but also essential. The Web technologies developed in the field of computer network can provide such a platform for sharing and integrating resources, such as heterogeneous geographic models, heterogeneous computing power, database, and knowledge base, across distributed, wide area networks (see Figure 1).However, the diversity and complexity of existing geographic models make the development of a networked, distributed framework become difficult.

This paper mainly focuses on the design and implementation of a distributed collaborative geographic model integration environment (DCGMIM) using Web Services technologies. Several key techniques, such as geographic model organization, geographic model representation, geographic model wrapping, and integration strategies, which are needed to build DCGMIM, are discussed in the following sections. The system aims to efficiently integrate disparate geographic model resources, offer powerful geographic model services (GMS) and allow large-scale distributed users to solve complicated geographic problems and conduct collaborative work.

The remainder of this paper is organized as follows. In Section 2, the related work of model integration is introduced and their advantages and disadvantages are discussed. In Section 3, a unified geographic models organization, representation, wrapping and integration strategies are presented. In Section 4, detailed descriptions of the distributed and collaborative four-layer architecture are discussed, and a prototype case study

* Corresponding author. Address: Nanjing Normal University, School of Geography Science, Xianlin Campus, WenYuan Road, 06307 box, Nanjing 210046, PR China. Tel.: +86 25 85898270; fax: + 86 25 85898270.
E-mail address: whinee@gmail.com (H. Yang).

given in this section is used to illustrate the distributed collaborative geographic integration architecture. Finally, in Section 5, the work is summarised and some future research directions are highlighted.

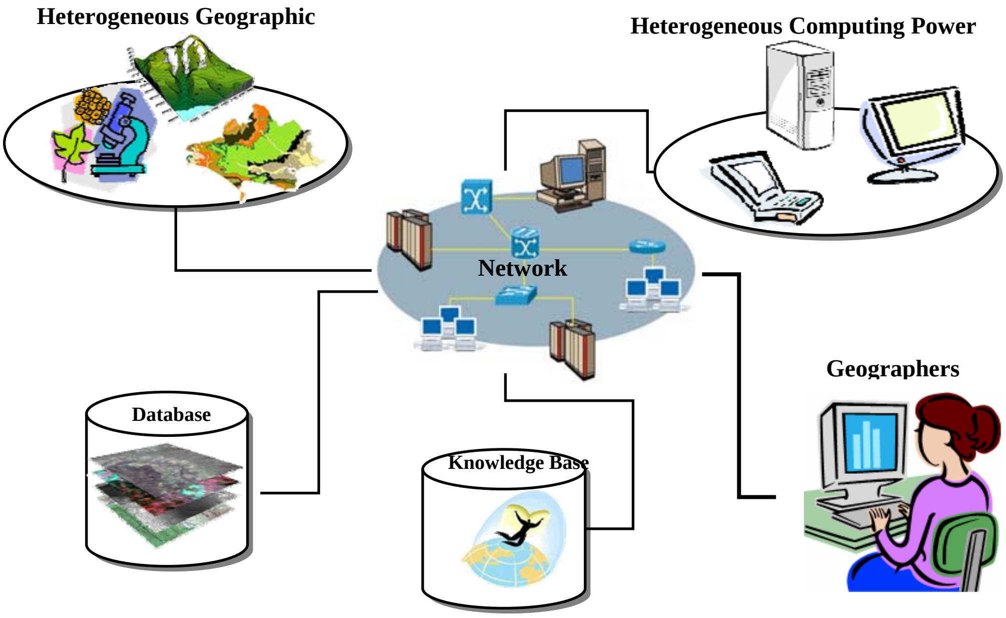

Figure 1. Distributed resources on the Network

2. RELATED WORK

It has become a trend for many research institutes to have a research possibility for model integration. For example, Kottemann, J. E. *et al*. (1992) identify two aspects of an integration specification. Ramirez, R. G. *et al*. address similar but simpler problems in their design for DAMS. The SWAMP system (Abel, D. J. *et al*. 1994) approached the model integration problem with architecture, comprising a graphical user interface, a database, a model driver, and the HSPF solver. Sham, C. H. *et al*. (1995) have developed an application that combines a legacy ground-water model, a custom spatio-temporal nitrogen loading model, and a GIS to study water quality. The GOODES (Djokic, D. *et al*. 1996) approach to model integration in the hydrological domain focuses on the problem of data transfer between models. The work of Bettett, D. A. (1997) also supports the development of spatial systems incorporating multiple users, demonstrated in a hydrological domain. Taylor, K. *et al*. (1999) proposes the FMISDS integration approach that offers a problem-oriented method for designing the integration system.

In other related project, The EU HarmonIT project is the most sophisticated approach for model integration. Great effort is spent on the development of concepts called Open Modelling Interfaces and Environment (OpenMI, 2002), which allow the combination of various models. GeoVISTA Studio, which is developed by the Pennsylvania State University, is a Java-based, visual programming environment that allows for the rapid, programming free development of complex data exploration and knowledge construction applications to support geographic analysis. The European Network for Earth System Modelling (ENES) organized the PRISM project, funded in 2001, and is to share the development, maintenance and support of a comprehensive Earth System Modelling (ESM) software environment.

These ongoing research efforts address ways in which a model integration environment will be able to support geographers and suggest what a modelling tool or software should look like in such an environment. However, they do not provide a distributed and collaborative framework for integrating the multidisciplinary and multiobjective geographic models.

The framework presented in this paper differs in its focus to: create a unified schema that handles the heterogeneous geographic models integration needed in distributed design (i.e. geographic model organization; geographic model representation; geographic model wrapping; geographic model integration strategies); and provide a Web Services-based architecture to facilitate the integration of geographic models and their utilization in an open, distributed and collaborative design environment.

3. GEOGRAPHIC MODEL INTEGRATION IN DISTRIBUTED COLLABORATIVE DESIGN

This paper draws increasing attention in the fields of geographic model integration in collaborative design. It is working on enabling technologies or infrastructure that can assist geographers in the distributed environment. We are intended to help geographers to collaborate in or coordinate solve problems by sharing spatial information and geographic models through formal or informal interactions.

3.1 Geographic Model Organization

Different kinds and research purposes of Geographical model make its implied information various. Therefore, when we integrate geographic models in a distributed environment, we should take more specific factors of specific categories into account in order to identify the significant factors to grasp the essence of the law. The research based on Earth System Science (ESS) in this paper develops an organization of geographical models. Using this organization as organizational structure of integrating heterogeneous resources in distributed environment, and distributing subdirectories in accordance with the organization, we can place diversity models in the network environment under a unified organization.

As shown in Figure 2, organizational structure of geographic model is in the form of tree structure, and it is the basis of organizing and managing model resources. According to scientific theory of earth system, the Earth is an open and complex giant system. Internal and external interaction and mutual constraints exists between the Earth layers. Also, in the process of development and evolution, the mechanism of physical, chemical and life processes, and the unity of contradictions between micro and macro phenomenon are intertwined with (Qian, X. S., 1994). Therefore, this paper firstly divides geographical models into earth system synthesized model, water system model, soil system model, atmospheric system model, ecology system model, land system model, ocean system model, snow and ice system model and solid earth system model from a macro perspective. Secondly, according to the relationship of earth layers' interaction, taking the water system model as an example, the paper further divides it into sub-model, water system integrated model, atmospheric water system model, ocean water system model, surface water and groundwater system model. Finally, from a microscopic perspective, in accordance with the physics, chemistry and ecology mechanism, with the groundwater system model as an example, the paper, again, divides it into groundwater system synthesized model, groundwater system physical model, groundwater system chemical model, groundwater system ecological model and groundwater system statistical model.

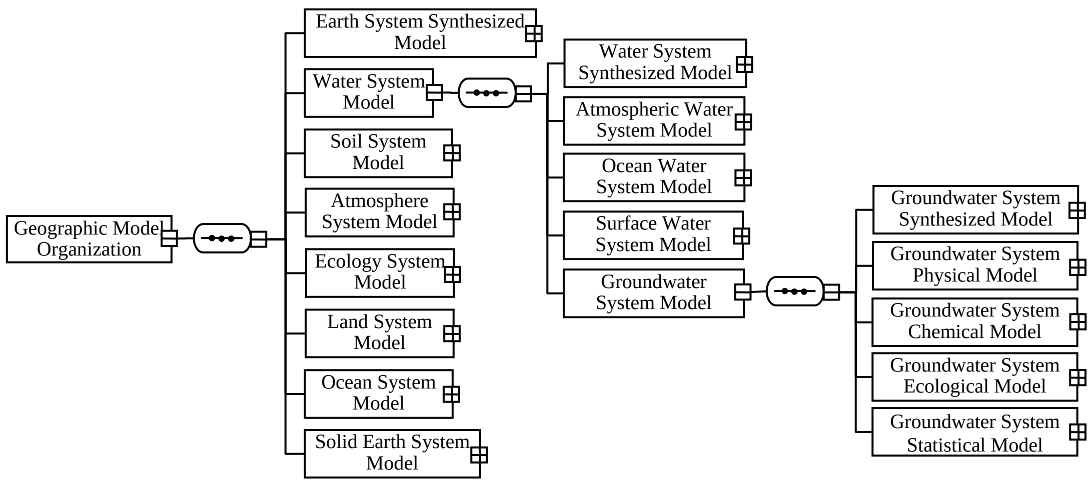

Figure 2. Organization of geographic models

3.2 Geographic Model Representation

The diversity, complexity, and heterogeneous of existing geographic model resources make the development of a networked, distributed framework for geographic model integration become difficult. As a result, we have to face the fact that geographic models often incorporate different syntactic, semantics, formats, and representations. In view of this situation, we put forward geographic model description which uses metadata to provide unified representation to publish, share, and locate model sources.

A solution to the unified represent by metadata for a geographic model is necessary for the integrated architecture. Appropriate and effective metadata standard framework for geographic model plays a key role in the integrated approach. It can help a geographer to locate and understand geographic models. Metadata is usually defined as "data about the content, quality, condition, and other characteristics of data" (FGDC, 1998). Here we extend this definition to include information about geographic models. We contend that geographic model metadata is required to index and categorize the growing number of spatial models available for download from the internet (Raja, R. S. *et al*., 2003).

The metadata representing the geographic models is used as the basis of narrating various properties and capabilities, such as input, output and functionality of a geographic model. These metadata formalizes name of properties, numeric constraints on parameters or aliases. Then geographic model sources can be fully understood just by exploring the model metadata. Therefore we define a comprehensive metadata standard framework for geographic models which should be bundled with the web services wrapper (GMS). This standard framework specifies the content of metadata for a set of geographic models. The purpose of the standard framework is to provide a common set of terminology and definitions of concepts related to these metadata. Figure 2 briefly illustrates the metadata standard framework for Geographic model which can be recognized by the geographic model integrated architecture. This framework utilizes components of the metadata standard recommended by FGDC for geospatial data (FGDC, 1998).

The metadata standard framework for geographic model consists of nine major sections and three subordination sections. The main sections are: Identification information, Application Range Information, Spatial Reference Information, Spatio-temporal Information, Model Parameter Information, Modeling Principle and Solving Method Information, Running Condition Information, Model Performance Information and Distribured and Citation Information. The subordination sections are: Management Information, Contact Information and Instance Information.

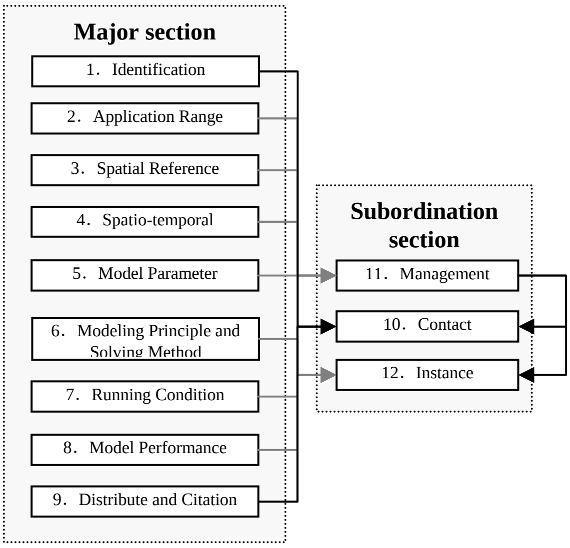

Figure 3. Metadata Standard Framework for Geographic models

3.3 Geographic Model Wrapping

For solution of the heterogeneity of pre-existing geographic models we draw on approaches from the study of database integration for multidatabase (Steth, A. P. *et al*., 1990). The solution requires individual, independent geographic model components to be encapsulated by a wrapper which can interact with each other. It means that existing geographic models are able to cooperate although the cooperation was not designed in advance.

Web services are self-contained and modular applications that can be described, published, located, and invoked over the Web. They use open standards and common infrastructure for their description, discovery and invocation, (e.g., XML over HTTP, WSDL, SOAP, and UDDI). The standard specifications and general introductions to these concepts have been provided in some papers. Certainly for Web Services the importance of XML is paramount; all key Web Services technologies are based on it (Hagel, J. *et al*., 2001). SOAP is the XML-based information which can be used for exchanging structured and typed information between peers in a decentralized, distributed environment (W3C, 2001). WSDL defines an XML grammar for describing network services as collections of communication endpoints capable of exchanging messages (W3C, 2001). UDDI is a "meta services" for locating Web Services by enabling robust queries against rich metadata (UDDI.ORG, 2004). Two of the great advantages of the Web Services architecture are its openness and its modularity (Graham, S. *et al*., 2001).

Overall, Web services present many new potential opportunities to significantly reduce the complexity and costs of spatial analyses for developing and maintaining geographic models. Figure 4 illustrates a general model for Web services which named Service Oriented Architecture (SOA).

The SOA describes three key roles. The services provider implements the Web Service and publishes its description to one or more repositories for potential users to locate. The service requesters are searching for a Web Service that it needs to bind with. Finally, the service broker manages a repository and allows the service requestor to find an adequate service. In actual usage scenarios, multiple service requesters, providers and brokers can interact with one another.

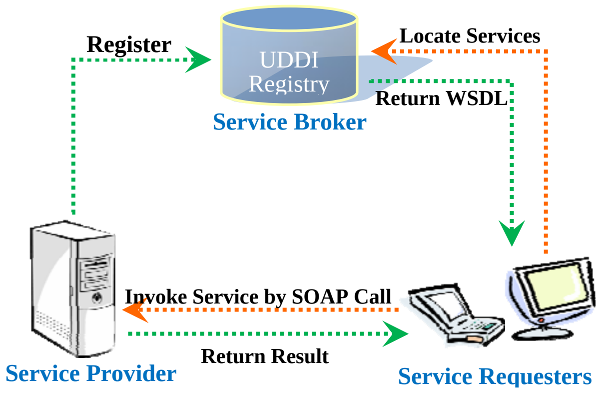

Figure 4. Service Oriented Architecture for Service Providers, Requesters and Broker

Web services perform encapsulated geographic models ranging from simple buffer operation to complex geographic process interactions. Effectively, the Web Services-based approach offers a mechanism which can wrap the pre-existing geographic model as the geographic model services (GMS) at which wrapper linking is performed. The wrapper translates queries on the geographic model services to direct queries on the constituent component in its native style.

One unique feature of Web services is that they can be mixed and matched to construct a new and complete geographic process due to their standardized service interfaces and common communication protocols. They also allow dynamic integration with reduced human interaction through the embedded capability of services discovery and binding. Existing geographic models can, therefore, be integrated and extended to participate as part of a new composite GMS.

Web services are standards-based and suited to build common infrastructure to reduce the barriers of geographic model integration. The simplicity of the overall Web services model helps accelerate its deployment. The dynamic nature of Web services also opens up new opportunities for both geographic and IT support.

3.4 Geographic Model Integration Strategies

The purpose of geographic model integration strategies are to make whole geographic problem solved as soon as possible by allocating a running task to another node with less average executing time. Figure 5 shows the four integration strategies identified this paper.

Sequential: Geographic models are executed sequentially if the execution of one geographic model is followed by the next model. In Figure 5(a) model B is executed after model A has been completed and before model C is stated.

AND: In Figure5 (b) geographic model B and C are executed in parallel. This means that B and C are executed at the same time or in any order. The model B and C can be executed after A has been completed. And model D may start after B and C have been completed.

OR: In Figure 5(c) either model B or C (exclusive OR) is exectuted. If model A is executed, a choice is made between B and C. Model D may start after B or C is completed.
Iteration: sometimes it is necessary to execute a model multiple times. In Figure 5(d) model B is executed one or more times.

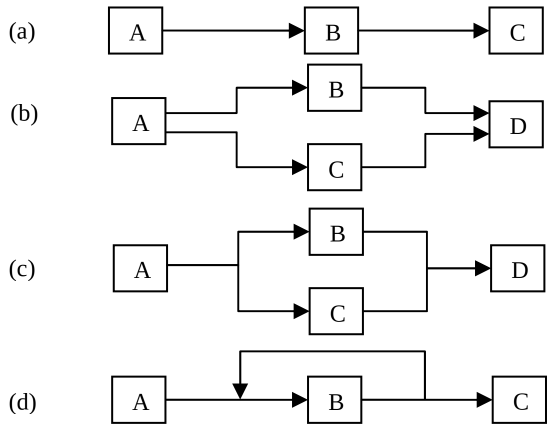

Figure 5. Four Geographic Model Integration Strategies

4. THE ARCHITECTURE OF DISTRIBUTED COLLABORATIVE GEOGRAPHIC MODEL INTEGRATION

In order to effectively apply Web Services technologies to the distributed geographic model integration system, a four-layer system architecture is designed (Figure 5). In the architecture, a series of GMS are built to share and integrate geographic models. Using these GMS, geographic models are wrapped as higher-level services for dealing with application requirements. Within the architecture, users can register and publish their GMS. The Web Services-based approach assists the distributed architecture that increases synergy and user cooperation during the integration between geographic model resources.

Here we describe the distributed geographic model integrated architecture using Web Services-based approach in detail. The architecture is composed of four main layers: Resources Layer, Wrapping Layer, Integration Layer, and Application Layer.

4.1 Layer Description

The Resource Layer aggregates all resources reference to geographic model integration, such as heterogeneous geographic models, computing power, database, and knowledge base. It includes the physical (in this case, geographic models, spatial database and computing power) resources and logical (in this case, knowledge and rule) resources. It provides other layers with its various resources.
The architecture acquires unified representation and wrapping of the geographic models, which includes geographic model metadata and GMS. The metadata is about geographic models and GMS for model wrapping.

The Integration Layer provides three main roles: GMS Provider, GMS Requesters and GMS Broker. Then the GMS Providers can publish GMS, while GMS Requesters can retrieve GMS from GMS Broker. It allows the geographers to publish, manage, query and retrieve services about the geographic models.

Application Layer supplies the collaborative environment with functionality of workflow management. It provides several facilities for remote geographers so that they can link and couple geographic models in four integration strategies.

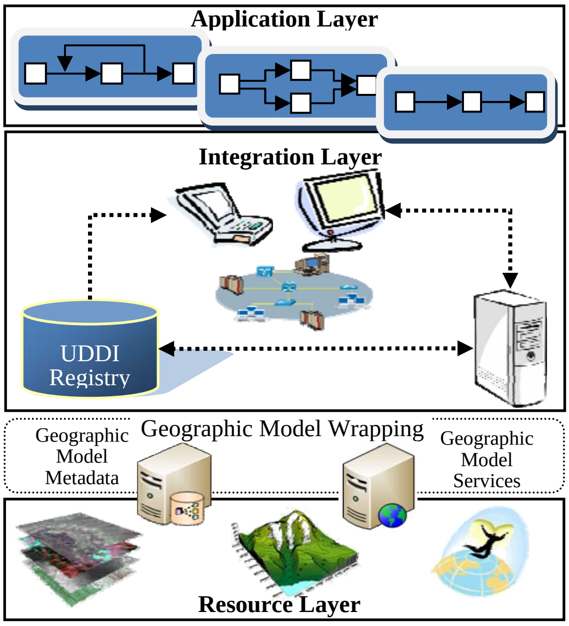

Figure 6. Four-layer architecture of the Web Services-based geographic model integration

Prototype of the Architecture to illustrate the function of DCGMIM system, a prototype was implemented. This system provides a distributed environment for collaborative integrating a geographic model on groundwater. The system is more flexible and scalable. Geographers can extract geographic model metadata, wrap models, and register their wrapper (GMS) in the ESS-based organization. For collaboration, model integration process are designed and shown in the main window. One User can send a request to other users or join group. When collaboratively solving a geographic problem, all users can share their GMS. Any links can couples made by one user will be automatically broadcasted to other collaborators and their displays will be updated as well.

The prototype of an open and flexible architecture and more efficient and interactive collaborative environment is crucial to the distributed integration of geographic models. The results show that the architecture presented in this paper is efficient and feasible and that the DCGMIM can help geographers rapidly manipulate geographic models and solve geographic problems.

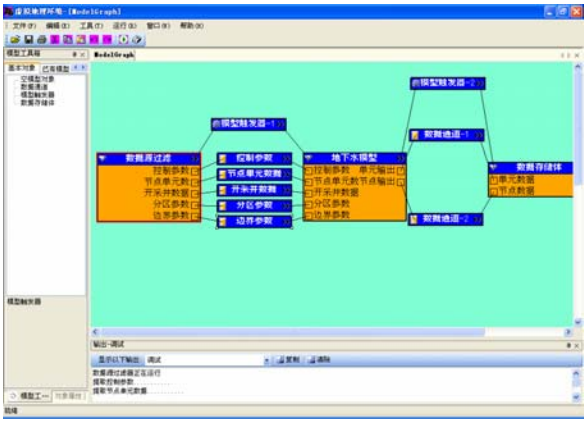

Figure 7. The user interface of the prototype system

5. CONCLUSIONS

Geographic model integration is a hot research of GIS, and helpful to enhance the GIS's ability to solve complex geographic problems. In conclusion, the reader's attention is drawn to the four key parts in a geographic integrating architecture which is based on Web Services. The first is the organization of the geographic model which is tree structure based on the ESS. The second is the geographic model representation, which provides uniform description mechanism for heterogeneity geographic models. The third is the wrapping technology, which provides the GMS to share and integrate heterogeneous geographic models. The last one is the integration strategy of geographic models for connection and communication with each other in the open network environment.

This paper introduces the four-layer architecture with four key technologies used by geographers from multiple domains to share and integrate geographic model sources via Web in a new perspective view. Various heterogeneous geographic model resources will be presented by metadata and wrapped by GMS. The GMS, managed by the ESS-based organization, are helpful to the uniform management for the distributed and local geographic model services. Through the Integration Layer, they can register, locate and invoke geographic model services in execution environment. Finally, Application Layer supplies the collaborative environment with several facilities for remote geographers so that they can link and couple geographic models in four integration strategies.

We can safely conclude that the work on distributed geographic model integration is not finished yet and that many challenges tasks lie ahead. Some questions should be better studied in future work: parsing pre-existing geographic models with the methods of automatically code analysis and artificial assistant analysis, disassembling the basic geographic model as different kinds of granularities, and putting it to the geographic model repository, algorithms library and method repository. As far as this is concerned, it is well possible that development of geographic model integration in distributed collaborative design is ahead of schedule.

ACKNOWLEDGEMENTS

This research has been supported by Major Program of National Natural Science Foundation of China, National Natural Science Foundation of China (No. 40671147), and Doctoral Fund of Ministry of Education of China (No.20060319004). The authors would like to express special thanks to Professor Guonian Lu who has carefully reviewed the manuscript and provided many valuable comments and suggestions that help improve the quality of the paper. We also thank Yongjun Wang and all other VGE group members for their fruitful discussions.

REFERENCES

Abel, D. J., Kilby, P. J., and Cameron, M. A., 1994. A federated systems approach to design of spatial decision support systems. *In Proceeding of the Sixth International Symposium on Spatial Data Handing*, pp.*46-59.*

Bettett, D. A., 1997a. A framework for the integration of geographical information systems and modelbase management. *International Journal of Geographic Information Systems*, 11(4), pp. 337-357.

Djokic, D., Coates, A., and Ball, J. E., 1995. Generic data exchange: Integrating models and data providers. *In NCGIA Third International Conference/Workshop on Integrating GIS and Environmental Modeling*,

FGDC, 1998. Federal Geographic Data Committee. Content Standard for Digital Geospatial Metadata. http://www.fgdc.gov/metadata/csdgm

Gahegan, M., Takatsuka, M., Wheeler, M., and Hardisty, F., GeoVISTA Studio: a geocomputational workbench, *Computers, Environment and Urban Systems*, 26, pp. 267-292.

Graham, S., and Simeonov, S., 2001. *Building Web services with Java*, SAMS,

Hagel, J., and Brown, J. S., 2001. *Your next IT strategy*. Harvard Business Review.

Kottemann, J. E., and Dolk, D. R., 1992. Model Integration and model languages: A process perspective. *Information Systems Research*, 3, pp.1-16.

Qian, X. S., 1994. *On Geographic Science*. Hangzhou Zhejiang Education Press, pp. 1-325.

Raja, R. S., and David, A. B., 2003. Agent-based modeling environment for spatial decision support. *Geographic information science*, 17(2), pp. 157-180.

Ramirez, R. G., and Ching, C., and Louis, R. D. S., 1993. Independence and mapping in model-based decision support systems. *Decision Support System*, 10(3), pp. 341-358,

Sham, C. H., Brawley, J. W., and Moritz, M. A., 1995. Quantifying septic nitrogen loadings to receiving waters: Waquoit Bay, Massachusetts. *International Journal of Geographical Information Systems*, 9(4), pp. 463-473.

Sheth, A. P., and Larson, J. A., 1990. Federated database systems for managing distributed, heterogeneous, and autonomous databases. *Computing Surveys*. 22(3), pp.183-236.

Taylor, K., Walker, G., and Abel, D., 1999a. A Framework for model integration in spatial decision support systems. *International Journal of Geographical Information Science*, 13(6), pp. 533-555.

UDDI.ORG: UDDI version 3.0.2 Specification, UDDI Spec Technical Committee. http://uddi.org/pubs/uddi_v3.htm

W3C: Web services Description Language (WSDL) version 1.1, W3C standard. http://www.w3.org/TR/wsdl

W3C: SOAP Version 1.2 Part 0-2, W3C working draft. http://www.w3.org/TR/soap12-part0

OpenMI, 2002. Open Modelling Interfaces, http://www.openmi.org (accessed 24 April. 2008)

PRISM, 2001. Progrom for Integrated Earth System Modelling, http://www.prism.enes.org (accessed 24 April. 2008)