Geospatial Service Web: towards integrated cyberinfrastructure for GIScience

GONG Jianya*, WU Huayi, ZHANG Tong, GUI Zhipeng, LI Zhenlong, YOU Lan, SHEN Shengyu, ZHENG Jie, GENG Jing, QI Kunlun, YANG Wenjing, LI Zhenqiang and YU Jingmin

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China

(Received 10 March 2012; final version received 27 June 2012)

A geospatial cyberinfrastructure is needed to support advanced GIScience research and education activities. However, the heterogeneous and distributed nature of geospatial resources creates enormous obstacles for building a unified and interoperable geospatial cyberinfrastructure. In this paper, we propose the Geospatial Service Web (GSW) to underpin the development of a future geospatial cyberinfrastructure. The GSW excels over the traditional spatial data infrastructure by providing a highly intelligent geospatial middleware to integrate various geospatial resources through the Internet based on interoperable Web service technologies. The development of the GSW focuses on the establishment of a platform where data, information, and knowledge can be shared and exchanged in an interoperable manner. Theoretically, we describe the conceptual framework and research challenges for GSW, and then introduce our recent research toward building a GSW. A research agenda for building a GSW is also presented in the paper.

Keywords: Geospatial Service Web; geospatial cyberinfrastructure; GIScience; GIServices

1. Introduction

Recent information technology developments will shape the future evolution of Geographic Information Systems (GISys) and Geographical Information Science (GIScience). The emergence of the virtual globe dramatically enhances the expressiveness of digital maps thus improving the user experience given it is highly interactive and explorative three-dimensional global paradigms (1). Diffusion of Web 2.0 technology increases the dissemination of geographical information on the web. Developers can easily implement map mash-ups through simple web mapping application programming interfaces (APIs) (e.g. Google Maps API). Taking advantage of individual efforts, crowdsourcing has been used extensively in large-scale web mapping projects (e.g. OpenStreetMap) (2). Voluntary web mapping efforts and geographic information sharing create many opportunities for geographic research and applications (3).

Over the last few years, location-based social networks or geosocial networking created enormous market potential for the advertising business (e.g. FourSquare.com). The sensor web is another pervasive monitoring technology emerging, aiming to collect, store, manage, and retrieve information from sensors through coordinated and interoperable web interfaces. Earth observation systems heavily rely on the Geosensor web to integrate and consolidate heterogeneous sensor data stream (4). With salient advantages and benefits, cloud computing has been widely embraced in many scientific and engineering domains. Geospatial researchers and professionals can also adopt cloud computing technologies to deliver cost-effective yet highly scalable solutions for geospatial decision-making (5).

Above-mentioned emerging information technologies have laid solid technical foundation to develop, deploy, and manage comprehensive and integrated geospatial tools in cyber space, namely, cyberGISystems. Novel concepts in other disciplines also encourage GIScience researchers and professionals to continuously produce promising research ideas and business models.

Therefore, traditional GISystems are migrating from desktop to network deployment environments, where end users, application/research areas, data sources, model complexity, and visualization methodologies are undergoing significant changes, as summarized below:

1. End users include both general public users and professional users, while previously GISystems are only intended for being used by professional users.
2. Application and research areas are moving from isolated ones into cross-disciplinary applications and collaborative research.
3. Visualization is evolving from simple image-based browsing into Internet-based three-dimensional virtual globe representations that allow interactive and expressive display of digital geospatial data.

*Corresponding author. Email: gongjy@whu.edu.cn

ISSN 1009-5020 print/ISSN 1993-5153 online
© 2012 Wuhan University
http://dx.doi.org/10.1080/10095020.2012.714098
http://www.tandfonline.com
(4) Data to be processed is now beyond static geographic data and include dynamic and continuous spatio-temporal data collected with geographically distributed sensors.

(5) Large-scale simulation and complicated geocomputation has become common for regular geospatial analysis and decision-making tasks. Obviously, we are at a stage where digital geospatial data, including digital maps, remote sensing imagery, and positioning data, are increasing rapidly. Effective spatial data management is of significant importance for geospatial research and application. Since the 1990s, the idea of building a spatial data infrastructure (SDI) to achieve broad spatial data access through the Internet has received a considerable amount of attention worldwide (6). Most countries have launched national SDI projects to facilitate spatial data sharing primarily for government agencies. International efforts such as INPIRE and Global SDI have made tremendous progress. Since the interoperability issue is central to SDI, Open Geospatial Consortium (OGC) standards are widely embraced to support seamless integration and complex queries from multsource heterogeneous geospatial data (7). Spatial data harmonization methodologies have also been developed to address data integration issues (8). Beyond data access services, interoperable geoprocessing services have been developed and deployed in SDI, to extract geospatial information effectively from distributed geospatial data (9).

In this data-rich era, GISystem researchers and GIS professionals feel compelled to achieve significant advances in the geospatial hardware, software, and algorithm developments, desperately required to transform enormous geospatial data into usable geospatial information and knowledge. In response to these needs, geospatial cyberinfrastructure (CI) has been an idea gaining momentum in the GIScience community (10–12), primarily inspired by the initial proposition for a cyberinfrastructure by the US National Science Foundation (US National Science Foundation (13, 14). Geospatial CI is concerned with building ubiquitous digital environments that allow GIScience researchers to perform various geospatial analysis and decision-making tasks, supported by highly scalable network-connected geospatial data collection and storage facilities as well as interactive geospatial data computation and visualization tools. Geospatial CI is still largely a long-term vision since it gives rise to many severe research challenges and implementational/operational difficulties. Current geospatial data processing and visualization models were mostly developed in desktop computing environments and are typically run in a serial manner. The capacity for handling large volumes of geospatial data with these models is sorely lacking. Although high-capacity data storage, high-speed networking, and high-performance computing facilities are readily available, special attention must be paid to adapt geospatial tools to distributed computing environments in order to attain as high performance as possible.

Despite impressive strides in SDI and geospatial CI over the last two decades, the ultimate goal of establishing advanced scientific discovery and education environments is still elusive due to numerous technical and institutional challenges. From the pure technical perspective, the following four formidable research challenges can be identified:

(1) Building intelligent mechanisms to enable effective search, discovery, recognition, and composition of distributed and heterogeneous geospatial resources across the Internet.

(2) Defining and establishing a unified framework for geospatial data access, processing, and visualization standards to achieve interoperability in high performance computing and networking environments.

(3) Constructing a geospatial knowledge base capable of accumulating ever-evolving geospatial knowledge to help solve complicated geospatial problems.

(4) Developing efficient yet cost-effective data-intensive computing cyber tools to support scalable geospatial data analytical operations.

To address the above challenges, we propose a new research paradigm, Geospatial Service Web (GSW), to underpin the development of future geospatial cyberinfrastructure. The development of the GSW will focus on the establishment of a platform where data, information, and knowledge can be shared and exchanged in an interoperable manner. Sections 2 and 3 describe the conceptual framework and research challenges for GSW. Section 4 introduces our recent research efforts for building GSW. A research agenda for building GSW is presented in Section 5, followed by a brief conclusion in Section 6.

2. GSW: conceptual framework

Existing SDI or web-based geospatial applications allow users to access, share, and visually explore registered geospatial data. However, high-performance and intelligent cyber tools are needed by users in many complicated geospatial decision-making scenarios. Meanwhile, advances in mainstream information technology, including the widespread of high-speed Internet access, web service architectures, high-performance, and cloud computing, provide ambient technical environments for the geospatial community to transform voluminous geospatial data into information and knowledge efficiently. Clearly, we see that the research frontier has gradually moved from data-oriented SDI to information-oriented SDI, towards knowledge-oriented geospatial CI. This trend indicates a complete network-based transformation.
workflow, that is, ‘Earth observation data–geospatial information–geographic knowledge.’ In order to realize geographic knowledge discovery and management in a geospatial cyberinfrastructure, we need middleware to bridge the data, information, and knowledge gaps. Figure 1 shows the underlying role played by GSW to support effective data–information–knowledge transformation. GSW collects sensor data from one end and deliver geospatial knowledge for domain applications on the other end. GSW is distinct from traditional Internet GIServices in that its data resources are extended from static databases to real-time data collection sensors. Furthermore, the domain applications that GSW bolsters may include real-time automated service orchestration for geospatial decision-making.

GSW can be defined as a geospatial middleware that integrates various geospatial-related resources through the Internet. It provides a systematic framework for timely discovery, retrieval, processing, and integration of geospatial resources. The GSW middleware is responsible for establishing intelligent connections between end users, applications, data, services, and sensor devices that enables real-time geospatial data acquisition, massive data management, high-performance computing, and visualization capabilities. For example, GSW can be used to mine useful information from various geospatial data ranging from structured to unstructured and time-varying data. In the GSW paradigm, we also need to develop a series of domain- or application-oriented cyber tools to solve different geospatial problems (e.g. resources management, public health, urban planning, disaster prevention, and emergency response). More importantly, a GSW will integrate data sharing, visualization, and processing cyber tools and to construct comprehensive tool chains upon request.

In the GSW paradigm, geospatial data, information, knowledge, software, hardware, and collaboration can be abstracted as geospatial resources. Hardware infrastructure, such as computing, storage, and networking facilities are supporting resources for geospatial research and applications. Geospatial data, information, and knowledge are also resources. Data collection instruments and sensors can be regarded as geospatial resources as well. From the web services perspective, all geospatial resources can be packaged as GIServices.

GSW excels over traditional SDI in that it provides a highly intelligent geospatial middleware that integrates various geospatial resources through the Internet based on interoperable Web services technologies. At the core of GSW is web-based geospatial intelligence that enhances and promotes the level of geospatial information sharing and processing services by enabling them with automated reasoning capabilities. In essence, GSW is a virtual network of intelligent geographical information analysis and decision support tools that incorporate latest progress in GIScience research. GSW provides intelligent web-based services including the semantic integration of geospatial information, automated geographic ontology mapping, geospatial semantic reasoning and knowledge discovery, and integration and sharing of geospatial semantics in heterogeneous environments.

The ultimate objective of GSW is to build a new generation of multilevel, multigranularity, multidimensional spatio-temporal data management, visualization, and processing service network. GSW connects to a variety of sensors and partially endows the sensor web with distributed data management and dynamic visualization capabilities. Researchers can use GSW to develop high-precision and high-performance geospatial analysis algorithms and modeling tools in web-based environments. GSW also supports automated coordination and optimal use of distributed geospatial resources based on spatial principles (15). Practically, GSW will establish a web-based intelligent service platform that integrates Earth observation sensor networks to support real-time geospatial information and decision support services.

Figure 2 illustrates the conceptual framework of GSW. It consists of five components: geospatial resources, geospatial services, geospatial applications as well as GSW interoperability and security standards. Two additional protocol layers, resource access and standard service protocols, are also included to facilitate interactions between geospatial resource, service, and application components. As discussed previously, all geospatial resources can be released as services. Additionally, critical geospatial management and analytical operations can be configured as services, such as resource registration services, resource search services, data processing, visualization services, data mining services, and model

![Figure 1. GSW and data–information–knowledge transformation workflow.](image_url)
chaining services. Based on GSW APIs and GeoVisualization/GeoCollaboration environments, domain-oriented applications tools can be developed and deployed to GIScience gateway (16) or geospatial application gateways. All elements in Figure 2 can be combined together to function as a primitive geospatial CI.

Figure 3 further describes the logic relations between geospatial resources and services. GSW includes utilities to register heterogeneous geospatial resources automatically. Once geospatial resources are registered in registration center, users can search for desired resources through resource search services. To meet application requirements, sensor, data, information, and knowledge resources are organized hierarchically and connected by various services. For instance, sensor web services gather and fuse sensor data streams autonomously before they are transferred through the network as data resources. As soon as data resources are available, data processing and modeling chaining services are employed to extract useful spatio-temporal information, which is later further processed through data mining services. Finally knowledge management services effectively promote the use of geospatial knowledge resources. Visualization services are critical in the logic model since they are commonly
used to help understand and manipulate heterogeneous geospatial resources.

According to above description of the conceptual framework and the logic model of GSW, we can see that GSW establishes a complete geospatial knowledge discovery procedure, where interoperable web services and GIService technologies play an important role to utilize geospatial resources as effective as possible. We emphasize knowledge resources and services in GSW since they are necessary to reduce semantic gaps when different geospatial resources are involved.

3. GSW: challenges

The long-term vision of GSW is to realize automated and intelligent geospatial data–information–knowledge transformation based on heterogeneous geospatial resources. Nevertheless, great challenges lie ahead and we try to identify these key research and technical issues for GSW in this section. Theoretically, pressing and fundamental problems include abstraction and management of geospatial resources, interoperability and standardization of geospatial resources and services, and representation of geospatial information and knowledge. These theoretical issues are briefly discussed below.

3.1. Effective management of geospatial resources

In the GSW paradigm, geospatial resources are the underlying foundation upon which services and applications are built (Figure 1). Therefore, effective management of geospatial resources is central to GSW. The primary challenge for resource management is that geospatial resources cover a wide range of types. Resources are offered by different providers through various network protocols. In many time-critical geospatial applications, GSW are required to deliver needed resources on demand. These resources must be instantly identified, retrieved, and aggregated from heterogeneous and geographical distributed providers. Current geographic information discovery and retrieval is largely ontology-based (Brisaboa et al. (17)) and can be extended to accommodate rich types of geospatial resources. In distributed resource-constrained environments, the efficient use of limited geospatial resources is another big concern. In GSW, the coordination of geospatial resources when performing specific geospatial analysis tasks is also challenging.

Key research issues include:

(1) To formalize standard hierarchical geospatial resource categorization and annotation approaches; To define general geospatial resource description models and operation modeling approaches that are compliant with current Earth observation sensor web standards.

(2) To develop automated geospatial resource registration and query service interfaces as well as resource access protocols.

(3) To propose resource-aware parallel and autonomous agent-based management framework that is capable of aggregating highly coordinated geospatial resources.

(4) To design optimal geospatial resource deployment, configuration, and allocation mechanisms given stringent resource constraints.

3.2. Interoperability and standardization

Interoperability has long been a problematic issue for Internet-based GIServices. In addition to web services standard defined by the World Wide Web Consortium (W3C) and other international organizations, a number of geospatial service standards have been released by OGC to enable interoperable geospatial information sharing and processing. In GSW, the major challenge focuses on geospatial resources and complex geoprocessing services because regular Internet GIServices have been well-standardized. In GSW, integration of various geospatial information and knowledge introduce severe semantic ambiguity given the overwhelming amounts of dynamic and inconsistent data, information, and knowledge. Service-oriented architecture provides an ideal framework for geospatial resource integration and advanced application development. Cloud computing holds considerable promise for data-intensive geospatial applications. Emerging cloud computing standards certainly offer good reference for GSW to deliver scalable, on-demand, and cost-effective GIServices.

Key research issues include:

(1) To define abstract resource reference models to facilitate geospatial resource discovery and sharing.

(2) To develop abstract GIService chain model to enable complex GIService composition.

(3) To establish semantic-enabled dynamic service-oriented integration architecture and to bridge resource and service gaps under this architecture.

(4) To design a complete set of service quality and security standards that ensures robust geospatial service delivery in distributed computing environments.

(5) To collaborate with other domain scientists to investigate cross-discipline application-specific interoperability requirements and accordingly to examine resource and service interoperability issues, particularly in cloud computing environments.

3.3. Multidimensional spatio-temporal representation and reasoning of geospatial information and knowledge

Geospatial knowledge, especially for specific domains, is difficult to represent since there is no widely-accepted standard symbolization for geospatial elements and
phenomena. In the area of Semantic web and Artificial Intelligence, extensive studies have been conducted to define expressive knowledge symbolization and inference systems (18, 19). However, geospatial knowledge has particular spatio-temporal implications such as topological relations, introducing additional complexity for knowledge modeling. Further, it is difficult to establish a unified geospatial knowledge base within distributed network environments where heterogeneous geospatial resources are semantically inconsistent. Geospatial CI usually involves interdisciplinary researchers and knowledge ambiguities must be resolved to the maximum extent. Geographers and GIScience researchers have made significant efforts to investigate geographic knowledge representation in the context of geospatial analysis and modeling (20, 21). Nevertheless, there is no well-established geospatial knowledge representation theory until now.

Key research issues include:

1. To establish a set of consistent vocabulary for geospatial knowledge management.
2. To develop a formal knowledge inference logics and operation standards to facilitate geospatial knowledge reasoning and sharing.
3. To build a flexible semantic framework for geospatial processing and moving representation to accommodate dynamically changing geospatial data.
4. To provide effective extension mechanisms for domain-specific knowledge representation and reasoning.
5. To propose a semantically enabled and efficient geospatial knowledge query, retrieval, and update mechanisms.
6. In addition to theoretical challenges, technical issues arise when building an integrated GSW platform. We discuss technical challenges below.

3.4. An integrated GSW platform

Various geospatial applications present domain-specific requirements for cyberGIS tools, such as spatial clustering analysis (22) and Bayesian geostatistical modeling (23). Of late, most of cyberGIS tools are dedicated for specific algorithms rather than complex geospatial problem solving. Consequently, GIScience gateways or spatial web portals only provide limited geospatial problem solving capabilities. Practically, technical issues must be resolved in order to build a GSW prototype, an open web-based geospatial intelligent service platform that connects to Earth observation sensor networks, and provides real-time GIServices to support multiscale analyses and geospatial problem solving. Based on appropriate interoperability and security standards, the GSW platform provides ‘plug-in-and-play’ standard-compliant high-performance GIServices in a multidimensional integrated environment.

Key research issues include:

1. To investigate resource virtualization techniques and to establish highly reliable resource centers.
2. To develop a complete set of geoprocessing, geocomputation, and geosimulation cyber tools and to deploy them on GSW as on-demand GIServices.
3. To construct automatic/semi-automatic GIService composition models according to specific problem solving scenarios.
4. To design cloud-enabled and high-performance GSW APIs to support the development of scalable and cost-effective GIServices in multitenant environments.
5. To develop interactive GeoCollaboration and GeoVisualization environments and to perform usability tests to enhance user experience.
6. To implement novel GSW applications, particularly time-critical (e.g. emergency response) or large-scale (e.g. global climate change) applications to demonstrate GSW capabilities and effectiveness.

4. Pilot studies

We are currently in the process of developing an initial prototype to demonstrate the core concept of GSW. In the last few years, we have already developed three independent research prototypes for cloud-enabled geoprocessing (i.e. OpenRS), GIService chaining (i.e. GeoChaining), and a large-scale three-dimensional GeoVisualization (i.e. GeoGlobe). Based upon the GSW proposition, these efforts were jointly directed toward establishing a demonstrative GSW platform. These research prototypes are briefly introduced below.

The OpenRS initiative aims to build an open interoperable geoprocessing framework primarily for remote sensing imagery. The goal of OpenRS is to implement highly extensible, scalable, customizable, and configurable remote sensing image processing toolbox that enables the development of advanced geoprocessing functions or applications. The OpenRS framework encompasses basic image processing operations such as image mapping, image enhancement, georeferencing, and classification. We adopted a plug-in architecture to promote the extensibility of OpenRS. The entire OpenRS framework includes four layers: core plug-in management (orsBase), service plug-ins, object plug-ins, and application programs. The orsBase layer is responsible for object creation, registration, and query. The service plug-in layer includes application independent image modules, such as geometric processing (orsGeometry), spatial reference systems (orsSRS), and graphic interface customization (orsGUI). Developers can develop and deploy reusable object plug-ins to build complex application programs, such as a comprehensive image processing workflow program. A set of standard development interfaces is
defined to facilitate the development of various plug-ins that are highly interoperable with each other. We have launched the OpenRS Cloud project, which is supported by OpenRS plug-in architecture. OpenRS Cloud is capable of performing multitask parallel image processing operations. Cloud-enablement services can be implemented through a dedicated distributed programming model and a distributed file system. Web-based graphic interfaces allow for interactive parallel job monitoring.

The purpose of the GeoGlobe project is to design and implement a multiscale three-dimensional visualization environment that integrates and display multisource raster and vector geospatial data at global scale. It can also function as a public geospatial information management platform that enables scalable sharing of distributed geospatial resources through interoperable service-oriented architecture and OGC standards. The GeoGlobe software suite contains three components: GeoGlobe Server for multisource geospatial data registration and management, GeoGlobe Builder for massive geospatial data organization, and GeoGlobe 3D Viewer for real-time geospatial data visualization, query, and analysis on the client side. GeoGlobe Server contains distributed geospatial data engine to deliver complicated data query and spatial analysis services in GeoGlobe 3D Viewer. High data access performance is ensured as time-varying vector, raster, digital elevation models, and three-dimensional building models are well organized and indexed in GeoGlobe Builder.

The GeoGlobe platform adopts a series of advanced technologies, including global seamless data models, multiresolution wavelet pyramids, progressive peer-to-peer data transfer strategies, multiuser concurrent access, and scheduling mechanisms for efficient data organization, transmission, and visualization of massive disparate geospatial data.

GeoGlobe provides a set of APIs enables third-party developers to customize and extend the functionality and interfaces of GeoGlobe. MapWorld (http://www.tianditu.com, which has been officially released as official online map service by China’s National Administration of Surveying, Mapping and Geoinformation, is developed based on GeoGlobe. GeoGlobe has also been customized to develop a variety of web-based geospatial applications, including public geospatial information service systems, power grid management and emergency response systems, and digital city applications.

GeoChaining is an integrated geospatial web service chain modeling framework with interactive and intuitive drag-and-drop service composition tools. Figure 4 shows the main graphic user interface of GeoChaining and describes a GIService chain built by GeoChaining tools. Users can monitor the execution of GIService chains and visualize the running results on two-dimensional Google Maps or three-dimensional virtual earth environments (e.g. Google Earth or GeoGlobe).

In December 2011, we released the first version of integrated GSW prototype, GeoSquare. The GeoSquare project was initiated by International Collaborative Center for Geo-computation Studies (ICCGS) and sponsored by State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS) at Wuhan University. It provides a web-based platform to encourage geospatial data, information, and knowledge sharing through highly interactive and expressive graphic interfaces. GeoSquare integrates heterogeneous and distributed geospatial data, service, and infrastructure.
resources. It supports popular interoperable web services and OGC standards. Geoprocessing capabilities are incorporated seamlessly with powerful GIService chaining tools. GeoSquare also features high computation performance with built-in scalable parallel programming interfaces. Figure 5 shows that GeoSquare consists of five components: a web portal that hosts registration center and offers supportive functions, a GIService composition tool (GeoChaining), a web-based visualization client (GeoGlobe Viewer), a high-performance geoprocessing service platform (OpenRS), and data/service centers that manage and disseminate geospatial resources. GeoSquare portal (http://www.geosquare.org/) delivers a single entry point from where users can access Data & Service Registration Center, download GeoChaining modeling tool, and find all information related to GeoSquare project such as news, documentation, and development team. We implemented a unified user management mechanism.

Figure 5. GeoSquare architecture.

Figure 6. Data & Service Registration Center.
with the same access control policy and user account configuration as administered across all components in GeoSquare.

The Data & Service Registration Center is shown in Figure 6. In the left panel, registered data and services are organized in tree structures. The right panel includes five tabbed subpanels to illustrate geographic locations where data and services are stored, to show detailed data, geoprocessing service, and Web Map Services (WMSs) information and to deliver keyword-based and map-based query functions. Data and service centers are physical computer servers where geographic data, information, and knowledge are stored.

GeoChaining plays a central role in GeoSquare because it seamlessly connects to a Data & Service Registration Center and incorporates GeoGlobe as built-in visualization environments. All registered geospatial resources can be readily accessed in GeoChaining, in which quality information of resources is also available. With GeoGlobe visualization capabilities, users can directly observe remote geospatial resources (e.g., data, geoprocessing, and WMS service) as well as final execution results of GIService chains.

The GeoSquare project was implemented based on common web technologies such as enterprise web portal (e.g., Liferay portal framework), Rich Internet Applications (e.g. Ext Google Web Toolkit), Java User Interface toolkits (e.g. JFace), and workflow engines (e.g. ActiveBPEL engine). Most of these technologies are built upon open source standards and easy to develop. These technologies combined ensure the delivery of content-rich and cross-platform applications. In addition, the compliance of web services and OGC standards significantly promote GeoSquare interoperability.

GeoSquare has been developed as a GSW prototype that integrates complex service chaining tools (GeoChaining), open high-performance geoprocessing platform (OpenRS), and three-dimensional visualization clients (GeoGlobe) into an integrated environment. GeoSquare is capable of manipulating Internet-based GIServices and other geospatial resources with easy-to-use user interfaces. It demonstrates the feasibility of GSW using current web and GIS technologies. The next version of GeoSquare will focus on the development of advanced geocomputation services and GIService chaining applications.

5. Research agenda

In Section 3, we have pinpointed the unique research and implementation challenges of GSW. To address these challenges, a research plan must be outlined to define future directions in terms of priorities. Based on previous research progress in our laboratory, we attempt to summarize and categorize a list of immediate research topics and development plans in this section.

LIESMARS at Wuhan University will collaborate with ICCGS to conduct innovative research on GSW, with focus on Geocomputation related theories and algorithms, supporting technologies, and domain-oriented geospatial applications. The benefits of these research efforts will be realized through an open collaborative Geocomputation research environment, and developed on GeoSquare. The research agenda of GSW can be categorized at four levels (or themes): theory, technology, system, and application.

5.1. Theme I: theories and algorithms

Topics in this research theme help establish solid theoretical foundation for GSW. Our research plan may not comprehensively cover research challenges discussed in Section 3. Instead, the theoretical focus will be placed on large-scale complex geospatial representation, modeling, and simulation. Being the basis of many geospatial related applications in climatology, urban morphology, atmospheric chemistry, and ecology, the development of mathematical and computation models for dynamic earth surfaces process analysis and simulation certainly attracts our attention. The availability of massive high-resolution earth surface data (remote sensing images) enables the exploration of complex geographical phenomena, but creating daunting challenges to construct a new generation of models and algorithms that can effectively accommodate these data-rich environments. Another interesting topic is the adaptive geospatial representation from a geospatial cognition perspective. This interdisciplinary research area of cognitive geospatial representation brings together many novel ideas from cognitive psychology, GIScience, and artificial intelligence. A promising research venue has been the geospatial statistical modeling of time-varying and scale-dependent geospatial data to reveal intrinsic spatio-temporal properties.

We believe selective topics such as exploratory spatial data analysis, linear and nonlinear spatial regression, fuzzy spatial analysis, and spatial error analysis are of particular theoretical significance for advanced geospatial knowledge discovery in GSW.

5.2. Theme II: supporting technologies

The success of GSW is dependent on several key technologies that effectively and efficiently transform voluminous geospatial data into knowledge. More specially, we will concentrate on the following four technologies: massive spatio-temporal data management, cloud-enabled high-performance Geocomputation and Geovisualization, intelligent GIService chaining, and virtual GeoCollaboration. Based on current data organization technologies we developed for GeoGlobe, we will further explore the data management in GeoSensor web, including data aggregation, query scheduling, data provenance, and security issues. We hope to participate actively in the formulation, implementation, and evaluation of data interoperability standards. Cloud computing provides a potential solution for data intensive Geocomputation research and applications. We are currently building a cloud computing infrastructure to support cloud-enabled geospatial
data processing. More importantly, issues arise when traditional message-passing parallel algorithms are executed in cloud computing environments. Technical challenges involve how to design novel domain decomposition and adaptive load balancing approaches in cloud computing paradigms. Meanwhile, parallel and remote Geovisualization strategies require further research since they are particularly useful for interactive exploration of massive geospatial data. Intelligence is considered to be the essence of GSW and is primarily demonstrated in the on-demand discovery, composition, and optimized utilization of geospatial resources. Future research will be directed toward the formulation of application-oriented GIService chaining approaches that enable user-defined, workflow-based, and integrated service chain construction. Synchronous and asynchronous distributed GeoCollaboration environments enable decision-makers to build tentative virtual organizations and perform geospatial computation together. Major technical issues involve multiuser concurrency control, the design and evaluation of intuitive collaboration tools, and the performance issues related to collaborative geospatial computation and visualization.

5.3. Theme III: open GeoSquare: collaborative Geocomputation research environment

Driven by supporting technologies, we propose to continue our development efforts in GeoSquare and extend it into a network-based collaborative Geocomputation research environment that encourages researchers worldwide to share and disseminate geospatial theories, models, and tools. Open GeoSquare environment serves multiple purposes, ranging from GIScience research, geospatial science and technology education, and real-world applications. A comprehensive set of GIService development, deployment, registration, discovery, composition, and visualization toolbox will be provided in Open GeoSquare in an interactive visual development environment. Multisource and multidimensional geospatial resources at multiple spatial and temporal granularities will be seamlessly integrated and visualized. With the help of intelligent GIService chaining tools, decision-makers can easily create comprehensive geospatial problem solutions and collaborate with different stakeholders. Users are exposed to highly intuitive GeoCollaboration and Geovisualization environments to learn, to develop, and to share geospatial knowledge. Future Open GeoSquare will address needs of different working environments and deliver GSW capabilities through desktops, Internet browsers, and mobile devices over wired and wireless networks.

5.4. Theme IV: domain applications and geospatial knowledge management

Finally, a few typical domain applications will be selected to demonstrate the applicability and effectiveness of our GSW concepts and technologies. These applications usually pose profound challenges when processing a large amount of data incurring intensive Geocomputation. More specially, highly advanced service performance is needed for geospatial applications that require time-critical responses. Tentatively, global change studies, emergency response, and public geospatial information dissemination are candidate applications. Through these demonstrative applications, we will conduct empirical Geocomputation experiments on distributed high-performance computing and cloud computing environments to evaluate the computation efficiency, scalability, security, and reliability for the proposed GSW tools and Open GeoSquare platform. It is important to develop a full set of standard geospatial workbench datasets to facilitate comparative analysis of various GSW algorithms and tools. Also, GSW quality standards are useful for examine how well a GSW solution can meet the needs across different geospatial applications. Geospatial domain knowledge management is of both theoretical and practical significance for real-world geospatial applications. Extensive and comprehensive knowledge management services (e.g. inference, query, retrieval, and update) are needed for semantic-based knowledge integration and intelligent geoprocessing service chaining in the context of highly dynamic GeoCollaboration environments.

To summarize, research themes I and II are central to GSW and therefore we will concentrate on these two themes in the next 2–3 years. The proposed collaborative research environment will permit implementation of our latest research on GSW theories, algorithms, and supporting technologies. Finally, we plan to develop typical time-critical geospatial decision-making applications to evaluate the applicability of GSW theories and technologies.

6. Conclusion

We present a holistic view of GSW, covering its conceptual and logic paradigms, research challenges, and our initial endeavors towards integrated geospatial cyberinfrastructure to support GIScience research and education activities. By introducing intelligence into existing information infrastructure, the GSW has far-reaching potential to provide a powerful platform of intelligent geospatial analysis and decision support, and to reduce the difficulties of using geospatial information by the general public. GSW is the essential pillar for future geospatial cyberinfrastructure. Therefore, we fully intend to continue our efforts following the proposed research agenda. We also seek international collaboration from GIScience and other disciplines.

Acknowledgements

This work is jointly supported by National Basic Research Program of China (Nos. 2012CB719906 and 2011CB707105) and National Natural Science Foundation of China (Nos. 41023001, 40801153 and 40901190).
Notes on contributors

Gong Jianya is an academician of Chinese Academy of Science and director of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, China. He studied as a PhD candidate at Wuhan Technical University of Surveying and Mapping and Technical University of Denmark during the time period of 1988–1992 and got his PhD in 1992. Up to now, he has experienced several professional careers in different countries. He has been a Changjiang chair professor at LIESMARS, Wuhan Technical University of Surveying and Mapping and Wuhan University since 1996. His research interests include geospatial data structure and data model, geospatial data integration and management, geographical information system software, geospatial data sharing and interoperability, photogrammetry, GIScience, and remote sensing applications. He is the principal investigator (PI) of GeoSquare project.

Huayi Wu is a professor in GIScience in LIESMARS at Wuhan University. He is the team leader of GeoSquare project.

Tong Zhang is an associate professor in LIESMARS at Wuhan University. His current research interests are high-performance GeoComputation, spatial optimization, and geographic information science for transportation (GIS-T). He is the project manager and technical coordinator of the GeoSquare project.

Zhipeng Gui is currently a visiting scholar at the Center of Intelligent Spatial Computing for Water/Energy Science, College of Science, George Mason University. His research interests focus on geospatial service chain modeling (web services composition), QGIS-aware monitoring, and evaluation and optimization of geospatial services. He worked as the principal designer and developer of GeoChaining.

Zhenlong Li is the manager of R&D department of SeaSky Geomatics Inc. His research interests include web service integration, geospatial information mining and visualization, and geospatial cloud computing.

Lan You is a second-year PhD student in LIESMARS at Wuhan University. Her research interest is GIService composition. She has been primarily involved in the development of GeoChaining.

Shengyu Shen is a fourth-year PhD student in LIESMARS at Wuhan University, advised by Professor Wu Huayi. His research interests include quality of geospatial information service and remote sensing image retrieval based on cloud computing.

Jie Zheng is a first-year MS student in LIESMARS at Wuhan University, advised by Professor Jianya Gong. His research interest is open source web GIS. He participated in the design of Registration Center and implemented the component of data registration and query.

Jing Geng is a first-year PhD student in LIESMARS at Wuhan University, advised by Professor Jianya Gong. Her research focuses on GeoComputation.

Kunlun Qi is a first-year PhD student in LIESMARS at Wuhan University, advised by Professors Jianya Gong and Huayi Wu. His research interests include high-performance GeoComputation and GIService.

Wenjing Yang is a first-year MS student in LIESMARS at Wuhan University, advised by Professor Huayi Wu. Her research interest is open source web GIS.

Zhenqiang Li is a senior in the School of Remote Sensing and Information Engineering at Wuhan University. His research interests include GIS and data mining.

Yu Jingmin is a senior in the School of Remote Sensing and Information Engineering at Wuhan University. Her research interests include WebGIS and GIS applications.

References

(1) Butler, D. Virtual Globes: The Web-Wide World. Nature 2006, 439, 776–778.
(2) Hudson-Smith, A.; Batthy, M.; Crooks, A.; Milton, R. Mapping the Masses: Accessing Web2.0 through Crowdsourcing. Social Sci. Comput. Rev. 2009, 27 (4), 524–538.
(3) Goodchild, M.F. Citizens as Sensors: The World of Volunteer Geography. GeoJournal 2007, 69 (4), 211–221.
(4) Van Zyl, T.L.; Simonis, I.; McFerren, G. The Sensor Web: Systems of Sensor Systems. Int. J. Digital Earth 2009, 2 (1), 16–30.
(5) Yang, C.; Goodchild, M.; Huang, Q.; Nebert, D.; Raskin, R.; Xu, Y.; Bambuce M.; Fay, D. Spatial Cloud Computing: How Can The Geospatial Sciences Use And Help Shape Cloud Computing? Int. J. Digital Earth 2011, 4 (4), 305–329.
(6) Scholten, M.; Klamma, R.; Kiehlé, C. Evaluating Performance in Spatial Data Infrastructures for Geoprocessing. IEEE Int. Comput. 2006, 10 (5), 34–41.
(7) Janowicz, K.; Schade, S.; Broring, A.; Kebler, C.; Maue, P.; Stasch, C. Semantic Enablement for Spatial Data Infrastructure. Trans. GIS 2010, 14 (2), 111–129.
(8) Mohammadi, H.; Rajabifard, A.; Williamson, I.P. Development of an Interoperable Tool to Facilitate Spatial Data Integration in the Context of SDI. Int. J. Geogr. Inform. Sci. 2010, 24 (4), 487–505.
(9) Kiehlé, C.; Greve, K.; Heier, C. Requirements for Next Generation Spatial Data Infrastructures-Standardized Web Based Geoprocessing and Web Service Orchestration. Trans. GIS 2007, 11 (6), 819–834.
(10) Zhang, T.; Tsou, M.-H. Developing Grid-enabled Spatial Web Portal for Internet GIServices and Geospatial Cyber-infrastructure. Int. J. Geogr. Inform. Sci. 2009, 23 (5), 605–630.
(11) Yang, C.; Raskin, R.; Goodchild, M.; Gahegan, M. Geospatial Cyberinfrastructure: Past, Present and Future. Comput. Environ. Urban Syst. 2010, 34 (4), 264–277.
(12) Wright, D.J.; Wang, S. The Emergence of Spatial Cyberinfrastructure. Proc. Natl. Acad. Sci. 2011, 108 (14), 5488–5491.
(13) US National Science Foundation. Revolutionizing science and engineering through cyberinfrastructure: Report of the national science foundation blue-ribbon advisory panel on cyberinfrastructure. NSF Panel Reports; 2003. http://www.nsf.gov/pubs/2007/nsf0728/index.jsp (accessed Dec 10, 2011).
(14) US National Science Foundation. Cyberinfrastructure vision for 21st Century Discovery; 2007. http://www.nsf.gov/pubs/2007/nsf0728/index.jsp (accessed Dec 10, 2011).
(15) Yang, C.; Wu, H.; Huang, Q.; Li, Z.; Li, J. Using Spatial Principles to Optimize Distributed Computing for Enabling the Physical Science Discoveries. Proc. Natl. Acad. Sci. 2011, 108 (14), 5488–5491.
(16) Wang, S.; Liu, Y. TerraGrid GIScience Gateway: Bridging Cyberinfrastructure and GIScience. Int. J. Geogr. Inform. Sci. 2009, 23 (5), 631–656.
(17) Brisaboa, N.; Luaces, M.; Places, A.; Seco, D. Exploiting Geographic Reference of Documents in a Geographical Information Retrieval System Using an Ontology-Based Index. GeoInformatica 2010, 14 (3), 307–331.
(18) Brachman, R.J.; Levesque, H.J. *Knowledge Representation and Reasoning*; Morgan Kaufmann: San Francisco, CA, 2004.
(19) Helbig, H. *Knowledge Representation and the Semantics of Natural Language*; Springer: Berlin, 2006.
(20) Miller, H.; Wentz, E.A. Representation and Spatial Analysis in Geographic Information System. *Ann. Assoc. Am. Geogr.* 2003, 93 (3), 574–594.
(21) Goodchild, M.F.; Yuan, M.; Cova, T.J. Towards a General Theory of Geographic Representation in GIS. *Int. J. Geogr. Inform. Sci.* 2007, 21 (3), 239–260.
(22) Wang, S.; Cowles, M.K.; Armstrong, M.P. Grid Computing of Spatial Statistics: Using the TeraGrid for Gi(d) Analysis. *Concurrency Comput.: Practice Exp.* 2008, 20 (14), 1697–1720.
(23) Yan, J.; Cowles, M.K.; Wang, S.; Armstrong, M.P. Parallelizing MCMC for Bayesian Spatiotemporal Geostatistical Models. *Stat. Comput.* 2007, 323–17 (4), 323–335.
(24) Guo, W.; Gong, J.; Jiang, W.; Liu, Y.; She, B. OpenRS-Cloud: A Remote Sensing Image Processing Platform Based on Cloud Computing Environment. *Sci. China: Technol. Sci.* 2010, 53 (Suppl. 1), 221–230.
(25) Gong, J.; Xiang, L.; Chen, J.; Yue, P.; Liu, Y. GeoGlobe: A Virtual Globe for Multi-Source Geospatial Information Integration and Service. In *Advances in Web-based GIS, Mapping Service and Applications*; Li, Dragicevic, Veennendaal, Eds.; CRC Press: London, 2011; Chapter 5; pp. 85–108.