The Integration of Grid-enabled Internet GIServices and Geographic Semantic Web Technologies

Tong Zhang[1], Ming-Hsiang Tsou[2]

Department of Geography, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-4493, USA

[1]E-mail: zhangt@rohan.sdsu.edu
[2]E-mail: mtsou@mail.sdsu.edu

Abstract

This paper presents a new framework for Grid-enabled GIService web portals to facilitate the building of high-level intelligent Internet GIServices. The five-tier architecture suggested here can support advanced semantic search and query functions for distributed GIServices by combining Grid computing, Semantic Web, and software agent technologies. The design of the web portal user interface with software agents can help end users combine and integrate computing power with geospatial data and services. Geospatial ontologies are incorporated into the framework by using geographic web ontology language (G-OWL). Intelligent software agent (GeoAgent) techniques are used to automate the procedures for searching, retrieving and processing geospatial data. The proposed Internet GIService architecture will provide a blue print for the establishment of Grid-enabled Internet GIServices and will help the GIS community to identify potential technical challenges to implement intelligent Internet GIServices.

Keywords

grid computing, semantic Web, internet GIServices, software agents

I. INTRODUCTION

Internet Geographic Information Services (GIServices) are information services that utilize both wired and wireless network through Internet to access geographic information, spatial analytical tools and other GIS resources [1]. To establish comprehensive Internet GIServices, many innovative web technologies have been recently adopted by the GIS community, ranging from programming languages, software architectures, to communication protocols. These web technologies, from a technical perspective, can provide different levels of information services and analytical functions, including static web mapping, interactive GIS querying, online data mining tools and distributed GIServices. Many commercial and public domain software packages already utilize these innovative web technologies in their Internet GIServices. However, there are two major hurdles that impede the technological progress of Internet GIServices: (1) the communication speed across the Internet, and (2) the interoperability between different GIS engines and map servers.

Constrained by network communication speeds, the current transfer methods of geo-spatial data and services are too slow to enable complicated web-based spatial analysis procedures. The heterogeneous geospatial data formats also complicate the situation since geospatial data retrieval and conversion require immense computing resources and processing time. Communication speed thus becomes the first major challenge for providing high-level GIServices which usually require intensive geocomputation and cumbersome data exchanges. The second challenge of Internet GIServices is the interoperability between different GIS engines and map servers. It is difficult to share and exchange heterogeneous data and services across the networks. Moreover, the vocabulary and terminology used to describe geo-spatial objects and spatial relations are quite different among GIS users, scientific communities, and Internet GIServices providers. These difficulties intensify the challenge for searching web-based GIServices for on-line geospatial data.

Many international organizations have been formulating interoperability standards (e.g. Open Geospatial Consortium, International Organization of Standardization). However, most of these standards, such as Geography Markup Language (GML) and Web Map Services (WMS), are represented by UML static structures [2]. It is a challenge to represent dynamic Internet GIServices by these UML static structures. Another problem is that many interoperability issues are boiled down to semantic problems. The heterogeneity of GIServices naming, data type, and ontology domain concepts are typical semantic problems [3]. One possible solution for the interoperability issues in distributed network environments is to adopt the Semantic Web technology, which can provide well-defined ontologies for different knowledge domains across the network.

Ideally, multiple geodata and GIS functions from dispersed servers or different platforms should be connected and integrated dynamically to conduct a specific geospatial analysis task. To dynamically construct distributed GIServices, large amount of geospatial data and computational results would be requested and transferred across the network. Therefore, the GIS community needs to establish a comprehensive GIService framework in order to keep all available heterogeneous GIService components synchronized in terms of data flow and operation procedures. Such framework

provides levels of abstractions that makes a design of internet GIS less difficult to dynamically adjust to complicated and changing network environments (dial-up modem, cable modem service, ISDN, T1/T3, wireless channels, etc.). The recent development of Grid technology[4], with its powerful computational and resource management capabilities, provides a possible framework for the deployment of Internet GIServices. In the computer science community, the focus of Grid technology is to resolve low-level Grid computing technique issues (e.g. communication, protocols and resource management) and to build a collaborative network architecture of Grid (e.g. Globus Toolkit). However, the deployment of Internet GIServices needs a high-level application framework rather than the low-level Grid computing architecture.

This paper presents a new framework for Grid-enabled GIServices web portals to facilitate the building of high-level intelligent Internet GIServices. The concept of web portals is to provide an easily accessible interface for scientists and general users to access high-performance computational resources and geospatial data. They can perform complex GIS operations and spatial analysis without knowing the details of configuration. The design of this kind of portals will integrate Grid technology, Semantic Web technology and intelligent software agent (geo-agent) techniques into a high-level service-oriented framework. Users can use the portal to analyze geospatial problems, obtain useful mapping services, request geospatial data, and aggregate multiple GIS analysis results. The intelligence of the web portal will be provided by Semantic Web technologies which organize geospatial ontologies to facilitate the interoperability between users and the machines.

This paper is organized as follows. Section II presents a survey of related Grid computing, Semantic web, software agents (GeoAgents) studies. In section III, we detail the proposed architecture for integrating Grid computing and Semantic web technologies. This is followed by a user scenario and the description of research agenda for GIS community in section IV. We conclude in section V with the summary of our research findings and the vision for future research direction.

II. THE SUPPORTING TECHNIQUES TOWARD THE DEVELOPMENT OF A SPATIAL WEB

The concept of "Spatial Web" is becoming popular in the GIS community [5]. A Spatial Web is a communication network for discovering, accessing, and integrating geospatial data and GIServices across different computing/network environments. In December 2004, the National Center for Geographic Information and Analysis (NCGIA) held a specialists meeting on Spatial Webs at UC Santa Barbara. The specialists meeting foresaw the future of the Spatial Web concept and future applications of intelligent GIServices. During the meeting, Grid computing, Semantic Web, and software agents were suggested as three potential web technologies for the deployment of Spatial Webs [5]. As one way to implement

such a vision, this paper proposes a new Internet GIServices architecture that integrates the three web technologies (Grid Computing, Semantic Web, and software agents). In the following section, we will provide an overview of current progress on these web technologies.

A. Grid computing

Grid technologies have been rapidly evolving since the mid 1990s. They are primarily concerned with the issues of integrating large-scale computational resources and services [6]. The increasing diversity of computational and human resources created a "Grid problem" which requires new mechanisms for sharing resources dynamically [4]. There are two types of Grid technologies: data Grid and computational Grid. A data Grid [7] offers a basic architecture for managing geo-spatial data with coordinated data storage, data access, metadata and security services over dispersed locations. Data Grids can provide various downloadable GIS data and on-line mapping services via its metadata search engine. Computational Grid [8] is primarily concerned about related issues for building the infrastructure, which is supposed to meet the requirements for high-performance computational needs. Both data Grid and computational Grid are emerging technologies to provide effective computing resources sharing and integration. Currently, researchers in the computer science community have been focusing on the low-level Grid infrastructure construction. One famous example is the *Globus* Toolkit [9]. *The Globus* Toolkit can provide a standard Grid framework and a set of coordination functions. Many ongoing Grid projects around the world are based on the *Globus* Toolkit. The newest release of *Globus* Toolkit is version 3.2.1 which is based on OGSA (Open Grid Services Architecture) standards. The OGSA extends the capability of Web Services into the Grid computing framework. With the solid technical foundations laid by the *Globus* Toolkit and other computational/data Grid architectures, many application services could be created to solve problems in many disciplines. These application services have to be connected through Grid middleware to the Grid fabric layer [4][10]. For specific application areas, generic middleware has to be modified to efficiently handle particular tasks or scenarios. For geospatial applications, it is necessary to design and develop geospatial oriented Grid middleware [11]. Wang and his colleagues [11] have started to explore the conceptual and practical issues to implement geospatial Grid-based middleware. They have been conducting research on GIScience Grid portal (GGP), which is based on Grid-based geo-middleware (GGM) [12]. The research community gradually realized that in order to make Grid technologies more widely used by different scientific communities, it is critical to simplify the low-level configurations. A web portal offers a centralized and uniform interface to access the distributed and heterogeneous resources and services [13]. The web portal could be built upon a Grid portal toolkit (e.g. Commodity Grid Kits, GridPort, GPDK, PCT4G and ASCPortal and GridSphere). VisPortal [14] is a web portal example that has implemented

Grid-enabled visualization capabilities and is based on GPDK (Grid portal Development Toolkit) [15]. Since most GIServices require intensive computation as well as a large amount of geospatial data, both computational Grid and data Grid techniques will be required to enable GIServices run smoothly. So far, considerable progresses have been made in addressing software, hardware and system integration issues for building the Grid infrastructure [9] [10] [16]. GIServices, which usually demand intensive computing (e.g. spatial analysis, Geovisualization, simulation), could be greatly enhanced if deployed under the framework built upon the computational Grid and data Grid technologies.

B. Semantic Web

Semantic webs can facilitate web-based data sharing within a global network system. The technology can provide better definitions of web data and services. Thus large scale data sharing and reuse could become possible [17]. A work group in the World Wide Web Consortium (WC3) has defined related standards and languages for the applications of semantic web technologies (www.w3.org). Resource Description Framework (RDF) was designed to organize web information into triple terms to facilitate data retrieval (http://www.w3.org/RDF/). To better handle terms and relations in semantic webs, a Web Ontology Language (OWL) was also proposed to define terminologies used for specific contexts and properties in terms of classes and relations (http://www.w3.org/TR/owl-guide/). Besides the reorganization of online contents, recent breakthroughs have been made in defining Web services with mark-up languages. OWL-S is such a specification to help Web services discovering and usage (http://www.w3.org/Submission/2004/SUBM-OWL-S-20041122/). In addition to all the standards set by W3C and other organizations, semantic web tools have been developed for data parsing, metadata processing, ontology management and RDF/OWL formatting (e.g. Jena, RDF Gateway, Unicorn system et al.). To facilitate data sharing and the integration of various services, a geospatial Semantic Web will need to be developed utilizing a geographic web ontology language (G-OWL). The G-OWL, as a media, will help create a geographic semantic ontology/knowledge database. Also, every operation (data/service register, data query, visualization, online spatial analysis) will strictly follow the ontology guidelines and rules. By formalizing GIServices with G-OWL, semantic inconsistencies can be reduced and thus enhancing the accuracy and efficiency of the GIServices.

C. Software agents

Software agent technologies have has been explored since the 1960s when scientists began to develop a type of self-managed tools that are capable of dealing with heterogeneous data and computational components within distributed computing environments. As modern computer science and technology evolved, software agents have gradually become more and more intelligent and can interact with the user's inputs and other agents without human interference [18]. Intelligent software agents may possess capabilities to execute tasks based on their learning experiences. This process takes advantage of prior domain knowledge they have obtained [18]. Intelligent software agents have already been applied to help solve practical problems such as network traffic monitoring, web-based data mining, information retrieval, and collaborative decision-making. According to Li [19], there have only been a few cases where intelligent software agents have been specifically used to address geographic problems. Among these efforts, Tsou and Buttenfield [20] proposed a GeoAgent model to build a distributed geographic information service. GeoAgents perform communication mechanism to help users search and convert required data. Nolan et al. [21] proposed an agent-based architecture for imagery and geospatial computing. The architecture employs an Agent Communication Language, I-XML, to enable collaborations among agents. Thus intelligent software agents have great potential in integrating comprehensive yet heterogeneous Internet GIServices. Intelligent software agents or GeoAgents can also be used to construct geographic semantic webs. These agents are able to perform functions such as GIServices registration/management, communication coordination, metadata/ontology-based management, interpretation of user's intention, among others.

D. Integrated spatial Web

The goal of proposing intelligent Grid-enabled GIServices in this paper is similar to the concept of Spatial Webs. The integration of geographic Semantic Web and GeoAgents with Grid computing architecture would be a promising step toward the construction of an intelligent Spatial Web. Currently, the actual implementation of Spatial Webs has not been realized due to the lack of high-level cyber-infrastructure for GIServices. However, there have been some research projects in investigating semantic geospatial services. Arpinar et al. [22] have introduced several semantic analytical methods in their efforts on geospatial ontology development. Codex [23] is a web portal with an interactive graph visualization interface for geospatial ontologies and knowledge management (http://flatbox.geog.psu.edu/codex) as the development was funded by the GEON and HERO projects. Another effort on interoperability was by Lin and Ludaescher [24], who implemented an ontology-enabled geospatial data integration system based on ArcIMS, a commercial Internet Map Server. In addition to the efforts to facilitate distributed data access, the Open Geospatial Consortium (OCG) is developing a standard specification for an online GIServices integration framework (http://www.intl-interfaces.com/servicemodel/), with which interoperable Internet GIServices will be easier to integrate, discover and access. At George Mason University, researchers are in the process to integrate Grid technologies with OGC web services [25].

To summarize, the future development of spatial webs will provide integrated, intelligent, and collaborative Internet GIServices by combining Grid computing, Semantic Web, and intelligent

software agents. Grid computing architecture will facilitate the integration of end users' computing power and the selection of geospatial data and services. This paper will present a prototype which can combine geospatial ontologies, geospatial Semantic Web technologies, and intelligent software agent techniques. Intelligent software agent techniques will help users to automate the procedures of geospatial data searching, retrieving, and processing. The next section will illustrate such framework for the future development of Internet GIServices.

III. A NEW INTERNET GISERVICE ARCHITECTURE INTEGRATING GRID COMPUTING, SEMANTIC WEB, AND SOFTWARE AGENTS

The following section will introduce an integrated framework for intelligent GIServices. The term "intelligent" indicates that the design of the architecture must consider the dynamic nature of the network environment and the Internet GIServices become "automatic" or "semi-automatic"-style operations rather than the current "manual"-style searching and operations. The integration of Grid computing, Semantic Web, and software agents will be helpful to build a new generation of Internet GIServices with much higher performance and intelligence. To achieve this goal, three critical design challenges of Internet GIServices have to be addressed.

A. The design challenges

Problem 1: How can we Grid-enable Internet GIServices?

Current Grid computing architecture can be regarded as a middleware linking Grid fabrics (computing, storage and network resources) with specific applications. It consists of several layers (Fabrics, Connectivity, Resources, Collective, and Applications) [4] while all kinds of protocols work to manage resources, exchange data, discover services and protect the security of the applications. However, geospatial applications, due to their particularities, cannot directly be implemented on the *Applications* layer. A layer of Geospatial Grid Middleware is necessary to take into account all of the above-mentioned particularities of geospatial applications.

Problem 2: Where can GIS users find appropriate GIServices and data to meet their needs?

Numerous Internet GIServices and data warehouses have been deployed currently. The geospatial applications, more or less, have to deal with the collection of required geospatial data and GIS analytical tools. The new Internet GIServices architecture has to call for a mechanism to facilitate GIServices registration, metadata management, and the communication across heterogeneous network/computing environment.

Problem 3: Can the web portal be intelligent enough to understand user's inputs/requests and respond accordingly?

The profile of Internet GIServices users will likely be broader than traditional GIS users. It is logical to expect that most Internet GIService users will not be very knowledgeable in performing the appropriate geospatial analysis procedures to solve their particular geospatial problems. A user-friendly design of intelligent GIServices is needed to help users launch appropriate GIS procedures. The user's interface should have enough built-in intelligence to "understand" the user's intentions and automatically search for appropriate data and services to meet the needs [26]. Users can manipulate the GIS operations through natural language descriptions and direct manipulations. The GIS engine behind the web portal interface will automatically formalize spatial search criterion and adjust GIS operations to meet the users' demands. Also, a new communication mechanism will be needed to facilitate the interaction between users and the Internet GIServices network, and between geospatial data and GIS operation modules/models. The architecture design will adopt software agent communication mechanism to provide autonomous, adaptive, dynamic, and interoperable GIServices.

B. Architecture description

Although we have seen many successful applications using Grid computing, Web Services, Semantic Web technologies, and software agents in addressing distributed scientific and commercial problems, we cannot directly implemented with these technologies to develop intelligent GIServices due to the particularities of geospatial applications. The major obstacles are huge data volumes, heterogeneous data types, ill-structured geospatial problems, diverse spatial analysis strategies, heavy reliance on visualization, complicated spatio-temporal representation and measurement frameworks (e.g. scale), and security concerns of some geospatial data. All these complexities introduced by geospatial applications requires more considerations in terms of scalability, reliability, resource management, Grid services and data definition, geographical knowledge and rule management and security issues. Given these concerns and based on prior discussions, this paper proposes a five-tier Internet GIServices Architecture (Figure 1). This new architecture can provide a comprehensive framework to integrate Grid-enabled Internet GIServices with geospatial Semantic Web technologies and software agents.

(i) Tier 1: Web portal interface with software agents
The first tier is the web portal user-interface that can help simplify complex configurations of system integration, data conversion and data analysis. The web portal, by means of common Web browsers, can provide visualization tools for data search, spatial analysis, and the display of final GIS operation results. The web portal can also serve as a Visual Problem Solving Environment (PSE) [27] in which users can visually and interactively manipulate active Internet GIServices as analysis modules. Through the web portal, users can implement their personalized analysis strategies. The first tier is a thin client because most intensive analysis procedures and data transfer functions are allocated to the servers within
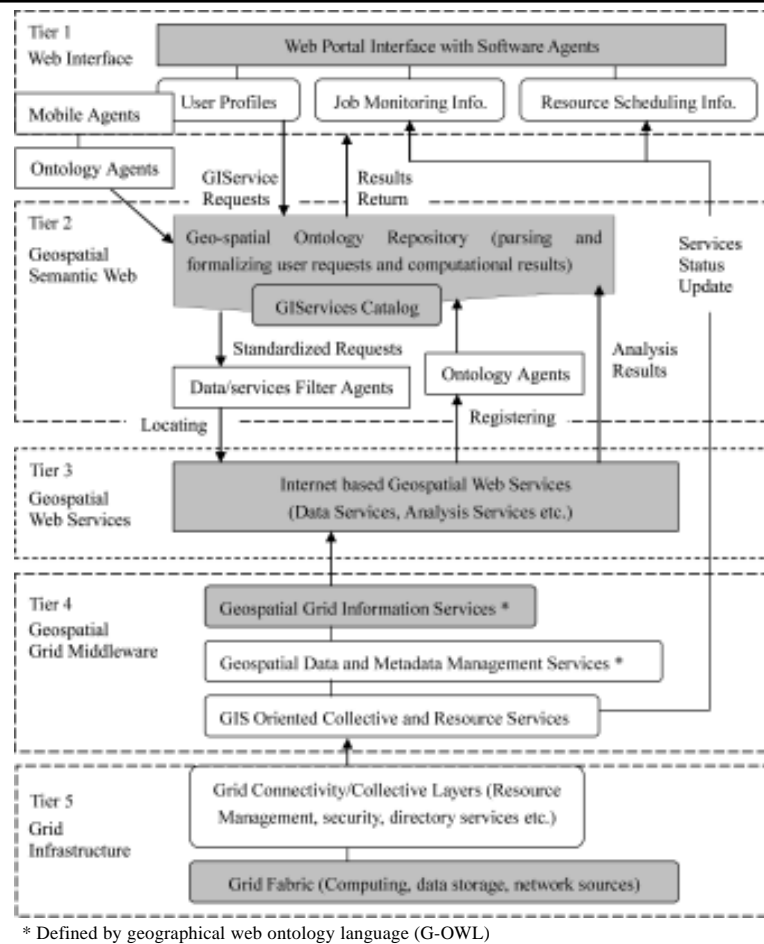
* Defined by geographical web ontology language (G-OWL)

Figure 1. An integrated three-tier internet GIService architecture

the Grid computing platforms. In this tier, a mobile user interface agent will collect information about user profiles, task status, and Grid resource scheduling. Users can monitor current progress of GIS tasks, track the computing resources used, and interact with other analysis processes if needed. Mobile user interface agents will be dynamically downloaded from servers to the user's machine (the client) for the purposes of analyzing user activities and request history, sending and formalizing inputs, and collecting computational results.

(ii) Tier 2: Geospatial semantic Web
The second tier is the geospatial Semantic Web, which consists of a geospatial ontology repository, GIService catalog, data/ services filter agents and ontology agents. The ontology repository will store basic geographic knowledge written in geographical web ontology language (G-OWL). G-OWL will provide guidelines to define geographic objects and relation concepts. Users' requests are parsed and formatted when evaluated by the ontology repository. Similarly, the analysis results are also reformatted before returning to the client. Specific user's data/services filter agents are responsible for finding appropriate data and analysis GIServices. These agents will search the GIServices catalog, where standard Geospatial Grid Information Services are registered. To find metadata for

the needed services, the agents will also retrieve metadata from the data or service sites to obtain the most updated information. This two-step searching strategy will assist the user in finding the best services. Ontology agents will help the Geospatial Grid Information Services register in the GIServices catalog. Another role the ontology agents will play is in the storage of the user's profile in the ontology repository. Future inquiries by repeated users can be expedited because the data/services agents have already acquired knowledge of the users' profiles.

(iii) Tier 3: Geospatial Web services
In this tier resided the high-level Geospatial Web Services that are defined by Geospatial Grid Information Services specifications. Government agencies, research institutions and private companies as well as individuals can host these Web Services. All of the Geospatial Web Services have to be in compliance with the widely accepted web service standards. Also, a conceptual Geospatial Service model should be designed and implemented with a set of Geospatial Grid Information Services specifications. One important issue is how to wrap around current Web Services with interoperable interfaces so that they will be useable in the new architecture

of Internet GIServices. Internet GIServices utilizes both wired and wireless Internet to access Geospatial data services and analytical services. There have been many Internet GIServices operating across the network. However, as we have discussed in the introduction section, interoperability is still a major concern when connecting these online services. Web Services provide a new communication paradigm to facilitate the composition of distributed Geospatial Services. Thus, current Internet GIServices have to be modified to meet the new requirements of Web GIServices architecture. In other words, Geospatial Web Services are also Web Services, though particularly offer services for geospatial problems.

(iv) Tier 4: Geospatial grid middleware

Tier 4 is the Geospatial Grid Middleware, which will extend the Open Grid Services Architecture (OGSA) with geospatial Web Services specifications. It will consist of interoperable GIService definitions, which are described by G-OWL. The Geospatial Grid Middleware will include three major components: GIS-Oriented Collective and Resource Services; Geospatial Data and Metadata Management Services; and Geospatial Grid Information Services. GIS-Oriented Collective and Resource Services extend the protocols of general Grid collective and resource layers by taking into account the particularities of geospatial applications. Therefore, data accessing, job management, failover, authorization, authentication, service discovery, diagnostic mechanisms have to be modified to reflect the requirements of Geospatial Grid Services. Geospatial Data and Metadata Management Services will be dedicated to handling geospatial data in Grid environment. Geospatial Grid Information Services provide a framework for the standardization of all Internet-base GIServices including both data services and analysis services. Geospatial Data and Metadata Management Services and Geospatial Grid Information Services are all defined in G-OWL and can be built using current OGC-endorsed standards (e.g. OGC Web Services 2, Web Registry Service, Catalog Service Specifications).

(v) Tier 5: Grid infrastructure

Tier five is the lowest tier providing a Grid infrastructure to conduct Grid-enabled GIS analysis. It usually includes the Grid Fabric, Connectivity and Collective Layers [4]. This infrastructure includes all the underlying Grid computing hardware and software. They provide a supportive environment to enable high-performance geospatial analysis. High-speed interconnected networks, processors, and storage space will be the basic fabric elements. Coupled with Grid software (e.g. Globus Toolkit [9]), these hardware can be connected to provide aggregated manageable resources. In addition, our GIS researchers may have other concerns about adding geospatial collection devices into this Grid infrastructure. These devices will include high-resolution remote sensing systems, GPS navigation systems, mobile GIS hand-held devices, etc.

To sum, this paper introduces a five-tier architecture for establishing intelligent high-level Internet GIServices, which can combine Grid computing, Semantic Web, and software agent technologies. The rationale behind this five-tier design is to make the Internet GIServices framework more flexible to support task-oriented geospatial analyses and GIS operations across the Internet. Multi-tier architecture can provide more effective and customizable GIServices to various users who may not be familiar with traditional GIS operations or spatial analysis procedures. The five-tier architecture can also provide a scalable framework, which can be modularized and upgradeable. To operationalize "Spatial Web", two backbones need to be built: Grid-enabled GIServices and Geospatial Semantic Web. Numerous Internet-based Geospatial Web Services will be deployed. They should be implemented with the interface of geospatial Grid middleware and geospatial ontology. These mechanisms are exactly what tier two, three and four are intended to accomplish. Currently, the actual implementation of the proposed framework is still under development as part of the ongoing research effort. The major challenge of the implementation is to combine multiple computing environments and technologies into a seamless framework. To demonstrate the advantages of this new architecture, the following section will provide a hypothetical user scenario to illustrate how this framework can help users to conduct geospatial analysis more effectively and efficiently.

IV. USER SCENARIO AND DISCUSSION

Figure 2 illustrates a hypothetical user scenario for the future Internet GIServices. This scenario is to simulate the occurrence of a serious wildfire in the County of San Diego (similar situation to the winter 2003 wildfire, which destroyed thousands of houses and killed fifteen people). The local governments will need to respond immediately in order to save lives and prevent further *damage*. One of the major tasks for the mayor of San Diego is to decide the locations of emergency shelters for residents who were evacuated from their homes. With the help of Internet GIServices, the mayor can access an Internet GIServices web portal and specify the service area (San Diego) and GIS task (site location analysis). Then the network will return a San Diego county map in which the mayor could continue to narrow down preferences by simply circling a candidate area. After she clarifies her intentions using plain text inputs (finding shelter sites without high wildfire risks, near freeways, easily accessible by emergency management agencies' personnel, etc.), the portal will parse her inputs, refer to the geospatial ontology repository and then locate related GIServices across the network as well as related data and maps. Some basic analytical tools will be wrapped together as Geospatial Web Services and registered in the web portal. When requested, the most suitable analytical tools will be launched as geospatial web services to derive the results. For example, a road network analysis tool will be located somewhere and used to analyze the transportation situation in the specified area. The portal will also search for registered geospatial datasets, which are relevant to mayor's submitted job. Once found, the data and services will be integrated within
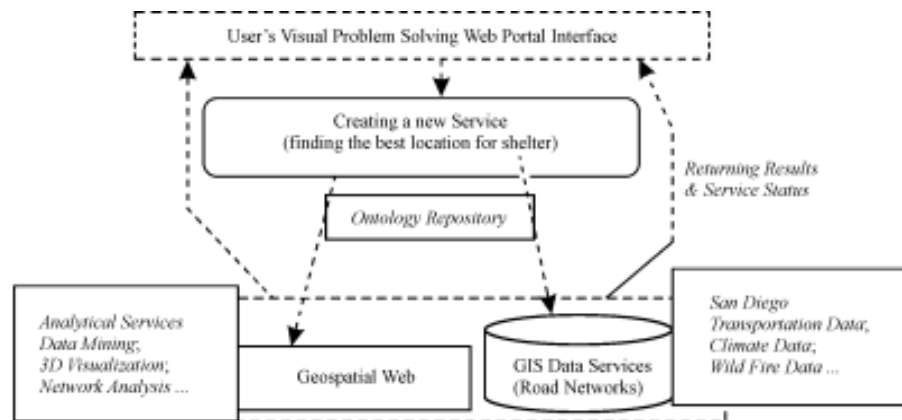
Figure 2. A user scenario for intelligent internet GIServices

the Grid framework and conduct Grid-enabled analyses.

Once finished, all the results will be aggregated and returned to the mayor's machine. The web portal will represent the results with appropriate visualization methods. At the user's end, all council members can visualize the data, intermediate results or even manipulate the Geospatial Web Services that work across the network. The web portal can be a Problem Solving Environment (PSE) [27] in which city council members can build their personalized analysis work-flow with active online Internet GIServices as analysis modules visually and interactively (similar to GeoVISTA Studio [28], which uses Java Bean Component to "design" analysis tasks). Finally, decision makers or participants may discuss overall alternatives and reach a consensus on where the shelters are going to be built. Since the computation and data transfer are Grid-based, the complicated, time-consuming GIS operations (including the network analysis) can be finished almost in real-time with the power of Grid computing. All decisions could be made in a short time, In managing disasters or emergency situations such as wildfire, time-saving technology is certainly critical.

This scenario has not yet been implemented. However, the proposed Internet GIServices architecture will contribute to this future vision by identifying technical issues and providing a blue print for the establishment of Grid-enabled Internet GIServices with Semantic Web and software agents. To operationalize such an integrated architecture, the following is a list of key components to implement the five-tier Internet GIServices.

- GIS-oriented Grid protocols in the Collective and Resource layers. The Grid protocols for data access, job management, failover, authorization, authentication, service discovery, diagnostic mechanisms will be carefully studied and customized for geospatial-specific services. Current Grid platforms (e.g. Globus Toolkits) can provide Application Programming Interfaces (API) and Software Development Kits (SDKs) for GIS researchers to develop geospatial-oriented resource layers.

- Geospatial Grid Middleware. This component is particularly important for the implementation of Geospatial Grid Information Services, which function as directory services (similar to generic Grid Information Services [29]). These Geospatial Grid Information Services provide information about existing Internet GIServices to facilitate query, replication and update of these services.

- Geospatial Data and Metadata Management Services. Similarly, Geospatial Data and Metadata Management Services will be based on core Data Grid Services to facilitate the geospatial data and metadata access in a Grid environment. For geospatial data, replica management is of central importance to data transfer because of the large volumes and diverse scopes of geospatial data. Metadata Management Services are concerned about the status of geospatial data currently used, such as connecting, replicating, and storage status.

- G-OWL. Geographic Web Ontology Language is necessary for describing the concepts of geographic objectives and relations. It is the infrastructure required to build the Geospatial Semantic Web (or Spatial Web). Geospatial Data and Metadata Management Services will be defined in G-OWL to make all data and analysis services semantically enabled. Prior geographic ontology research will benefit the design of such language.

- GeoAgent communication mechanisms. Intelligent software agents (GeoAgents) will work as basic communicators for the integrated Grid-enabled Internet GIServices network. To ensure the proper communication between different GeoAgents, communication mechanisms need to be designed for the entire Grid-enabled Internet GIService network (e.g. standardized request and send formats, priority, usage policies, etc.). These communication protocols have to comply with the specifications of Geographic Web Ontology Language.

V. CONCLUSION

This paper introduces an integrated Grid-enabled and

semantic-based Internet GIServices framework which can offer intelligent GIServices with querying, fetching and coordinating capabilities in distributed network environments. The integration of Grid-enabled GIServices, Semantic Web, and software agent technologies was formalized in a five-tier framework designed specifically for the next generation of Internet GIServices. To achieve the full potential of Internet GIServices, the GIS community needs to collaborate with geospatial data providers, web services designers, and the public/private sectors. The integrated five-tier Internet GIServices architecture proposed in this paper can be a pioneer endeavor to build the foundation for a distributed Internet GIService. This paper has investigated the theoretical and technical difficulties in developing such an integrated Grid-enabled and geospatial ontologies- based GIServices network.

Within this new Internet GIServices architecture, we can address the two major problems we discussed at the beginning: communication speed and interoperability. The performance of Internet GIServices will be greatly enhanced by Grid computing. Thus the complicated GIS data and tasks will be distributed across the network more efficiently, avoiding the bottleneck in communication speed on the network. Geographic Semantic Web, which standardizes all the communications between users and distributed GIServices, will make the transfers of information and knowledge transparent on the Internet. The integration of relevant data and analysis services becomes possible in this way. Software agents can handle the communicating tasks in the entire framework (e.g. interpreting of user's inputs, searching and locating GIServices, parsing and formalizing user requests and computational results, etc.). In summary, the integration of Grid computing, Semantic Web, and software agents will provide a new direction for the future development and performance improvement of intelligent Internet GIServices. There are other complementary approaches that can also improve the performance of Internet GIServices. For example, Yang et al. [30] discussed several innovative web techniques to improve the performance of Web-based GIS. These methods include a pyramid technique and harsh index, cluster and multithread, caching and dynamic data request, binary format and compression which work on both the server and client sides to enhance data access efficiency. We believe that all these web techniques can be integrated into our five-tier architecture to achieve better Internet GIServices performance. Hopefully, by integrating these innovative web technologies, the development of Internet GIServices will provide better geospatial services to our end-users and help them accomplish their geospatial analysis tasks more effectively and efficiently.

ACKNOWLEDGEMENT

REFERENCES

[1]   Peng Z. R., Tsou M. H., 2003, Internet GIS: *Distributed Geographic Information Services for the Internet and Wireless Network.* John Wiley & Sons. Inc.

[2]   Probst F., Gibotti F., Pazos A., Esbri M. A., Benigno. M., Gutiérrez M., Kuhn W., 2004, Connecting ISO and OGC Models to the Semantic Web. In *3rd International Conference on Geographic Information Science.*

[3]   Probst F., Lutz M., 2004, Giving Meaning to GI Web Service Descriptions (Extended Abstract). In *7th Conference on Geographic Information Science (AGILE 2004)*, Heraklion, Greece.

[4]   Foster I., Kesselman C., Tuecke S., 2001, The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *International J. Supercomputer Applications*, 15(3).

[5]   Goodchild M. F., 2004, Spatial Webs. http://www.ncgia.ucsb.edu/projects/nga/.

[6]   Baker M., Buyya R., Aforenza D., 2002, Grids and Grid Technologies for Wide-area Distributed Computing, *Software-Practice and Experience*, 32(15): 1437–1466.

[7]   Chervenak A., Foster I., Kesselman C., Salisbury C., Tuecke S., 2001, The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets, *Journal of Network and Computer Applications*, 23, pp.187–200.

[8]   Foster I., Kesselman C., 1999, Computational Grids. Chapter 2 of *The Grid: Blueprint for a New Computing Infrastructure.* San Francisco: Morgan-Kaufman.

[9]   Foster I., Kesselman C., 1997, Globus: A Metacomputing Infrastructure Toolkit. *International Journal Supercomputer Applications*, 11(2):115–128.

[10] Foster I., Kesselman C., Nick J., Tuecke S., 2002, The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration. *Open Grid Service Infrastructure WG, Global Grid Forum.*

[11] Wang S., Armstrong M. P., Bennett D. A., 2002, Conceptual Basics of Middleware Design to Support Grid Computing of Geographic Information. In *Proceedings of 2nd International Conference on Geographic Information Science*, Boulder, CO, USA.

[12] Wang S., Armstrong M., Cowels M., 2003, Grid-based Geo-Middleware for Geographic Analysis.. Online Materials. http://www.uiowa.edu/%7Egrow/presentations/GGM_Wangetalpdf.pdf.

[13] Jankun-Kelly T. J., Kreylos O., Shalf J., Ma K. B., Hamann B., Joy K., Bethel W., 2003, Deploying Web-based Visual Exploration Tools on the Grid, *IEEE Computer Graphics and Applications*, 23(2): 40–49.

[14] Bethel W., Siegrist C., Shalf J., Shetty P., Jankun-Kelly T. J., Kreylos O., Ma K., 2003, VisPortal: Deploying Grid-enabled Visualization Tools through a Web-portal Interface. In *Third Annual Workshop on Advanced Collaborative Environments*, Seattle, WA, June 2003.

[15] Novotny J., 2000, The Grid Portal Development Kit, *Concurrency Practice and Experience*, pp.1–7.

[16] Tuecke S., Czajkowski K., Foster I., Frey J., Graham S.,

Kesselman S., 2002, Grid Service Specification. *Open Grid Service Infrastructure WG, Global Grid Forum*, Draft 2.

[17] Berners-Lee T., Hendler J., Lassila O., 2001, The Semantic Web, *Scientific American*, 284(5): 34–43.

[18] Nwana H. S., 1996, Software Agents: An Overview. *Knowledge Engineering Review*, 11(3):1–40.

[19] Li Y., 2003, Software Agent Based Intelligent Data Processing System for Environmental Models. Master's thesis, Department of Geography, San Diego State University, San Diego, California.

[20] Tsou M., Buttenfield B., 2004, Integrating Distributed Geographic Information Services: A GeoAgent Approach, Submitted to *International Journal Geographical Information Science*.

[21] Nolan J. J., Simon R., Sood A. K., 2001, An Agent-based Approach to Imagery and Geospatial Computing. In *Proceedings of the fifth International Conference on Autonomous Agents*.

[22] Arpinar I., Sheth A., Ramakrishnan C., Usery E., Azami M., M. Kwan, 2005, Geospatial Ontology Development and Semantic Analytics. Book chapter in *Handbook of Geographic Information Science*, J. Wilson, A. Fotheringham(Eds.). Blackwell Publishing, Oxford (in print).

[23] Pike W. A., Gahegan M., 2003, Constructing semantically scalable cognitive spaces, in: *Spatial Information Theory: Foundations of Geographic Information Science. Conference on Spatial Information Theory COSIT03*, Lecture Notes in *Computer Science* 2825, K. Worboys M, and S. Timpf (Eds.), Springer-Verlag, Berlin: 332–348.

[24] Lin K., Ludaescher B., 2004, GEON: Ontology-Enabled Map Integration. In *Twenty-fourth Annual ESRI International User Conference*, San Diego, CA.

[25] Di L., Chen A., Yang W., Zhao P., 2003, The Integration of Grid Technology with OGC Web Services (OWS) in NWGISS for NASA EOS Data, in *HPDC12 & GGF8*, Seattle, USA.

[26] Cai G., Wang H., MacEachren A. M., Fuhrmann S., 2005, Natural Conversational Interfaces to Geospatial Databases. *Transactions in GIS*, 9(2):199–221.

[27] Rice J. R., Boisvet R. F., 1996, From Scientific Software Libraries to Problem-solving Environment. *IEEE Computational Science & Engineering*, 3(3):44–53.

[28] Takatsuka M., Gahegan M., 2002, GeoVISTA Studio: A Codeless Visual Programming Environment for Geoscientific Data Analysis and Visualization. *Computers & Geosciences*, 28(10): 1131–1144.

[29] Plale B., Dinda P., Laszewski G., 2002, Key Concepts and Services of a Grid Information Service, in *Proceedings of the 15th International Conference on Parallel and Distributed Computing Systems (PDCS 2002)*.

[30] Yang C., Wong D. W., Yang R., Kafatos M., Li Q., 2005, Performance-improving Techniques in Web-based GIS. *International Journal of Geographical Information Sciences*, 19(3):319–342.