Geospatial Semantic Web for Spatial Data Sharing

Chuanrong Zhang
University of Connecticut, USA

Weidong Li
University of Connecticut, USA

INTRODUCTION

Semantic Web was recently proposed to overcome the semantic heterogeneity problem and provide computers meaningful web content (Berners-Lee et al. 2001). Geospatial Semantic Web is an extension of the current Web, where geospatial information is given well-defined meaning by the ontology so that geospatial contents can be discovered, queried, and consumed automatically by computers. Geospatial Semantic Web aims to add computer-processable meaning (semantics) to the geospatial information on the World Wide Web. Because there are different encodings of geospatial semantics in GIS (Geographic Information System) databases, it is challenging to process requests for geospatial information over the Web. The Geospatial Semantic Web concept was proposed to address the vexing semantic challenges and achieve automation in geospatial web service discovery and execution (Duke et al. 2005).

While Open Geospatial Consortium (OGC) web services provide syntactic ways to encode geospatial information over the Web, they are unable to capture semantics information of geospatial data. Thus, a Web user has difficulty to find an appropriate geospatial data set for a specific task using one of the current search engines, because geospatial data sets encoded using OGC web services lack semantic information and computer programs are unable to understand the meanings expressed by geospatial data contents and requests. However, Geospatial Semantic Web is capable to capture, analyze, and share geospatial information beyond the purely lexical and syntactic level.

This article introduces the spatial data sharing problem and how to make intelligent search and integration of heterogeneous geospatial information by using the Geospatial Semantic Web technologies. We introduce the state-of-art Geospatial Semantic Web technologies such as ontology, ontology-based geospatial web services, ontology-based geospatial web service search engine, and the natural language interface for enabling users to instantly access disparate heterogeneous legacy geospatial data.

BACKGROUND

As GIS has been widely used by a variety of applications, many geographical databases have been developed by different programs and software. However, it is still a big problem to share these geospatial data and use them for the development of GIS applications. Not that spatial data are not available, there is a huge amount of geographical data stored in different places and in different formats; but data reuse for new applications and data sharing are daunting tasks because of the heterogeneity of existing systems in terms of data modeling concepts, data encoding techniques and storage structures, etc. (Devogele et al. 1998).

With the development of open standards, web services emerged for data interoperability over the Web. Within the broader context of web services, OGC web service specifications deal with geographic information on the Internet. OGC web services are evolutionary web standards that enable integration of different online GIS data and location information. With OGC's web service specifications and technologies, users can "wrap" the existing heterogeneous spatial data into a web service and enable many potential clients to use the service (OGC Interoperability Program White Paper 2001). OGC web services can be treated as a "black box" to perform a task by dynamically connecting

DOI: 10.4018/978-1-4666-5888-2.ch735

interoperable service chains for different applications (OGC White Paper 2001).

The major OGC web service specifications include Web Feature Service (WFS) specification, Web Map Service (WMS) Specification, Web Coverage Service (WCS) Specifications, Web Processing Service (WPS) Specification, and Catalogue Service (CS) Specifications (OGC 04-094 2005; OGC 06-042 2006; OGC 09-110r3 2010; OGC 05-007r7 2007; OGC 07-006r1 2007).

Unlike the current proprietary commercial GIS formats, the OGC web services support mapping from a wide variety of sources and enable sharing of geospatial data for online information exchanges. The OGC web services provide public open standards for coding and sharing geospatial data. Thus the databases based on the OGC web services can be easily shared and reused (OGC 2003). In addition, the OGC web services also provide a good solution for reducing the costly conversion processes among different format geospatial databases.

ISSUES

Although OGC web services have undoubtedly improved the sharing and synchronization of geospatial information across diverse resources, there are limitations with the currently implemented OGC web services:

First, although OGC web services facilitate data interoperability at the syntactical level via standard interfaces, they cannot resolve data interoperability problems at the semantic level. However, one of the major problems in spatial data sharing and data interoperability is the semantic heterogeneity of spatial data (Bishr 1998; Fonseca et al. 2002; Levinsohn 2000). The OGC web service descriptions only allow the specification of the syntax of basic service contents such as operation metadata, Feature Type list, and filter capabilities, and they provide no semantic descriptions of the meaning of these contents. Two identical XML descriptions may mean very different things depending on the contexts of their uses. In addition, the Web Feature Service specification of the outputs of each call to the service similarly lacks semantic definitions. All defined search operations return results using the same data structure, regardless of what information is requested. For example, a "*Building*" feature contains a field "*Commercial Building,*" which is used to describe buildings in commercial areas, and a field "*Residential Building,*" which is used to describe houses in residential areas. Even if the types of buildings specified in a "*Building*" file were clearly identified in a "*Type*" field by the interface designer, OGC WFS descriptions provide no uniform way of enabling such interpretations. It is up to the web service client to recognize the values in these fields, which indicate whether it is a commercial or a residential building.

Second, OGC web services only make it possible to search and access geospatial data by keywords in metadata, but cannot allow content-based searching at the semantic level. Because OGC web service descriptions do not support the semantic specification of service contents and operations, they only can allow requesting semi-structured keyword searches based on the metadata. In addition, metadata also have semantic heterogeneity problems. Different metadata creators may use different names for the same feature. For example, by typing keywords "*school*" and "*Storrs, CT*" in a data system implemented using OGC web services in Connecticut, users may get query results of a bunch of feature-level school data such as "*Mansfield Middle School*" and "*E.O Smith High School*" data services for Storrs, Connecticut, if their metadata contain exactly these keywords. However, if they use different names for the same feature, it is unlikely that a software program could read and utilize these data services without human assistance, because the OGC web service descriptions provide no means of including representations of the semantics of the provided services. Therefore, with OGC web services it is difficult to perform intelligent content-based search and users cannot correctly utilize the discovered web services without additional human assistance or programming. Further, metadata contain only limited information to allow users to search. Despite the efforts that the geospatial community has put on providing better tools to manage geospatial metadata, content-based searching at the semantic level remains a challenging problem.

Third, without a formal semantic description of OGC web services, it is difficult to allow users and applications to discover, deploy, compose, and synthesize the OGC web services automatically. The lack of an explicit semantic in the XML-based standard OGC web service descriptions proves to be a major limitation in automatic capability matching. It is unrealistic to expect that advertisements and requests of

Related Content